

Introduction to Fractal Geometry and Cantor Sets

(Chinese version first published on 2024.2.18)

O. Anqiao

Introduction

In the end of my previous article about Mathematical Analysis - real analysis, I mentioned a very special set called the "Cantor" set. It plays a crucial role in the field of mathematical analysis, so I have decided to write a separate article about it and fractal geometry.

Due to the limitations of equipment conditions during the journey, almost all the derivations in this article are based on some impressions and hand-written deductions. Please point out if any lack of rigor exists. Thank you.

The Cantor set, introduced by the German mathematician *Georg Ferdinand Ludwig Philipp Cantor* in the late 19th century (although discovered earlier by Henry John Stephen Smith), is infinite and uncountable. Due to its peculiar properties, it is often used in real analysis to construct various counterexamples. The Cantor function, defined on the interval [0,1], is uniformly continuous and monotonically increasing. It maps sparse sets to continuous intervals, possessing special properties worthy of further investigation.

I would like to extend the properties of Cantor sets and the Cantor function to discuss them in terms of the Hausdorff dimension in this article, not limited to Euclidean spaces, as indicated by the title. This approach aims to explore their properties from the perspective of fractal geometry. We cannot discuss the fractal nature of Cantor sets in Euclidean spaces, as will be explained later.

About Fractal Geometry

Fractal geometry is a mathematical theory used to study complex and irregular geometric shapes and structures. Unlike the regular geometric figures we typically discuss, such as

rectangles, circles, and triangles, which are regular and relatively easy to study, many things in nature exhibit more irregular and complex forms. Hence, the concept of fractals (from the Latin word "frāctus") was introduced in mathematics to describe situations where local patterns resemble the whole in some form.

Real Numbers and Cantor Set

Strictly speaking, Cantor's initial description of the Cantor set was quite abstract. What we commonly refer to as the "Cantor ternary set" is just one of the most representative methods of constructing Cantor sets. The Cantor set is a compact subset of the real number line, and it has measure zero, yet its cardinality equals that of the set of real numbers.

Cantor's earliest description of the Cantor set can be traced back to his paper *On a Property of the Class of all Real Algebraic Numbers* in 1874. In this paper, he first introduced the concept of "uncountable sets" and described a method for constructing such a set, later known as the Cantor set. His description involved constructing this set through a diagonal argument. He initially considered the existence of a set of real numbers that can be represented by a finite number of real numbers. Then, by constructing a new real number such that it differs from every real number in the original set, he obtained a new uncountable set.

"Infinity"

In the paper "On a Property of the Class of all Real Algebraic Numbers" which spans only three pages of PDF, Cantor asserts that real algebraic numbers are real numbers satisfying some form of non-trivial equation. Specifically, considering a real number ω , where n, a_0, a_1, \ldots, a_n are all integers, we have the equation

$$a_0\omega^n+a_1\omega^{n-1}+\ldots+a_n=0$$

We then call ω a real algebraic number, in other words, a root of a polynomial with integer

coefficients.

If its elements can be put into one-to-one correspondence with the elements of the set of natural numbers, we can then say a set is countable. However, despite the conditions defining real algebraic numbers above, intuitively, one might think that the set of real algebraic numbers is smaller than the set of real numbers, making it a countably infinite set.

But the focus here is not on the "overall properties of all classes of real algebraic numbers," but rather on Cantor's concept of extending finite sets to infinite sets, making this paper regarded as the cornerstone of set theory.

Let's take another look at what infinity means. We know that the cardinality of the set of natural numbers is infinite, generally denoted as \aleph_0 (aleph null). Intuitively, the set of real numbers should be denser than the set of natural numbers, thus larger, right? According to the viewpoint of the continuum hypothesis, we consider the cardinality of the set of real numbers as the power set of the cardinality of the set of natural numbers, 2^{\aleph_0} , and there exists no set A such that $\aleph_0 < |A| < 2^{\aleph_0}$, meaning there's no set larger than the cardinality of the natural numbers but smaller than that of the real numbers. In the generalized continuum hypothesis, this is generalized to $2^{\aleph_{\alpha}} = \aleph_{\alpha+1}$, giving rise to the so-called super infinite cardinals $\aleph_0, \aleph_1, \aleph_2, \ldots$.

The continuum hypothesis is assumed to be independent of ZFC (Zermelo-Fraenkel Set Theory), meaning that under ZF plus the Axiom of Choice, neither the truth nor the falsehood of the continuum hypothesis can be proven.

Moreover, this \aleph_1 is referred to as the "continuum cardinal," and for convenience, we define $\mathfrak{c} = \aleph_1 = 2^{\aleph_0}$ hereafter.

Cantor Ternary Set

As mentioned above, the Cantor ternary set is the most representative way of constructing Cantor sets, probably because it is more intuitively understandable in geometric terms. Therefore, many people sometimes refer to the "Cantor set" as the Cantor ternary set. Let's consider the closed interval I = [0, 1]. We divide it into three equal subintervals by removing the middle third interval, denoted as $I_1 = \left(\frac{1}{3}, \frac{2}{3}\right)$. Thus, the remaining line segment consists of the union of the two intervals that have been separated:

$$C_1=I-I_1=\left[0,rac{1}{3}
ight]~\cup~\left[rac{2}{3},1
ight]$$

Similarly, let's remove one-third of the remaining intervals separately, namely removing $I_{1,1} = \left(\frac{1}{9}, \frac{2}{9}\right)$ and $I_{1,2} = \left(\frac{7}{9}, \frac{8}{9}\right)$, resulting in: $C_2 = I - I_1 - I_{1,1} - I_{1,2} = \left[0, \frac{1}{9}\right] \cup \left[\frac{2}{9}, \frac{1}{3}\right] \cup \left[\frac{2}{3}, \frac{7}{9}\right] \cup \left[\frac{8}{9}, 1\right]$

This process continues iteratively. At the *n*th iteration of the trisection, we remove $\frac{1}{3^n}$ of the original interval length, resulting in 2^{n-1} removed intervals, and obtaining the union of 2^n intervals:

$$C_n = igcup_{i=1}^{2^n} I_{n,2^n}$$

As n approaches infinity, it defines the Cantor ternary set:

$$C = \lim_{n o \infty} C_n = igcup_{n=1}^\infty C_n$$

This process might remind one of the quote "一尺之棰,日取其半,万世不竭。" It can be understood in a similar fashion, though here we take thirds.

Another common method to describe the construction of the Cantor set is using recursion (as found in some Wiki descriptions). Starting with $C_0 = [0, 1]$, for $n \in \mathbb{N}$, the recursive step at iteration n is:

$$egin{aligned} C_n &:= rac{C_{n-1}}{3} \cup \left(rac{2}{3} + rac{C_{n-1}}{3}
ight) \ &= rac{1}{3} \left(C_{n-1} \cup (2+C_{n-1})
ight) \end{aligned}$$

Thus, the Cantor set under this definition is represented as:

$$C:=\lim_{n o\infty}C_n=igcap_{n=1}^\infty C_n$$

To represent the endpoints of the Cantor set, we can consider ternary notation. For example, the portion removed in the first step can be represented in ternary decimals as (0.1, 0.2), and when represented as 1 = 0.2, we no longer need to use 1. Thus, the Cantor set can be defined simply as:

$$C = \left\{ \sum_{n=1}^\infty rac{a_n}{3^n} : orall n \in \mathbb{N}, a_n \in \{0,2\}
ight\}$$

The Cantor set is a special compact totally disconnected subset of the real interval [0, 1], containing infinitely many real numbers and being uncountable, meaning it cannot be mapped one-to-one with the set of natural numbers. Hence, its cardinality is c.

The figure illustrates the step-by-step construction process of a Cantor set of order four (from top to bottom):



Properties

Enumerate some properties of the Cantor set.

Closed Set

The Cantor set is a closed set. It can be regarded as the intersection of infinitely many closed intervals, where each closed interval is the union of a finite number of closed intervals. Therefore, it is also a closed interval.

By the Heine-Borel theorem, if a set is closed and bounded, then it is compact (note that these two properties are equivalent and interchangeable). Since the Cantor set is closed, it is obviously compact.

Perfect Set

Sorry, I'm not sure about its formal Chinese name (many conflicting Chinese articles), so I'll use the name from Wiki for now.

Consider a topological space (X, τ) , where $A \subseteq X$, and A' is the derived set of A, satisfying A = A'. This means that every point in A' is a limit point of A, and all limit points of A are in A, then we call A a perfect set.

A satisfies the following properties in the above definition:

- A is a closed set in (X, τ) .
- A has no isolated points, meaning there is no point a such that there exists a neighborhood U where, except for a itself, there are no points belonging to A in this neighborhood.
- A contains all its limit points.
- A itself is dense.

The Cantor set is a perfect set. As we have already shown that the Cantor set is closed, we only need to prove that it has no isolated points.

Proof

When we consider any point $a \in C$, we have $\forall n \in \mathbb{N}$ such that $a \in C_n$. Therefore, considering the closed interval I_n in C_n , the endpoint a_n of I_n satisfies $a_n \neq a$. As we mentioned before, during the *n*-th iteration of the trisection operation, we remove $\frac{1}{3^n}$ of the original interval length. Hence, we have:

$$|a_n-a|<\frac{1}{3^n}$$

We construct this sequence a_n converging to a point a in C, thus proving that a is a limit point in C. Consequently, any point a in C is not isolated, demonstrating that the Cantor set is a perfect set.

Nowhere dense set

Consider a topological space (X, τ) , where $A \subseteq X$. If the closure \overline{A} of A equals X, we call A a dense set in X. However, if the closure \overline{A} of a set A contains no interior points, meaning its interior is empty, then we call it a sparse set, also known as nowhere dense set.

The Cantor set is a sparse set on \mathbb{R} , meaning that for any point in the Cantor set, any neighborhood of it contains other points not belonging to the Cantor set. This implies that the Cantor set has no interior points.

Proof

Consider the closed interval $(a, b) \subseteq [0, 1]$, where a < b. Let $n \in \mathbb{N}$ satisfy $\frac{1}{3^n} < b - a$. Since each interval in C_n is non-overlapping, there is no interval of length b - a completely contained in C_n .

As the Cantor set C can be seen as the union of C_n , $\bigcup_{n=1}^{\infty} C_n$, there is also no interval of length b - a in C, implying that there are no open intervals in C. Thus, the Cantor set

is nowhere dense.

Lebesgue Measure

We generally discuss Lebesgue measure on \mathbb{R} , while the Cantor set is a set of measure zero. When constructing C_n , we know that it consists of 2^n disjoint closed intervals, each with length $\frac{1}{3^n}$. Considering the Lebesgue measure m and the Lebesgue outer measure m^* , we have:

$$m^*(C) \leq m^*(C_n) \leq rac{2^n}{3^n}$$

According to the definition of the Cantor set, we only need to take the limit of $\frac{2^n}{3^n}$ as $n \to \infty$, which equals 0. Since we know the Lebesgue measure is non-negative, we can conclude that $m(C) = m^*(C) = 0$.

Cantor Function



Some properties of the Cantor function (proofs to be provided later):

• The Cantor function is surjective; every real number in [0,1] corresponds to a

unique real number in [0, 1].

- The Cantor function is continuous but not absolutely continuous; it is monotonically increasing on [0, 1].
- The Cantor function maps rational numbers in the interval [0, 1] to a dense set; its image set in [0, 1] is dense.
- The Cantor function is almost everywhere differentiable, and the set of differentiable points has a cardinality of c.

Another, more intuitive definition of the Cantor function can be given in terms of the ternary expansion $\{a_n\}$ of x:

$$c(x) = egin{cases} & \sum_{n=1}^\infty rac{a_n}{2^n}, x \in C \ & \sup_{y \leq x}, y \in Cc(y), x \in [0,1] ackslash C \end{cases}$$

Since the construction of the Cantor set can be described using ternary expansions, where a_n represents the *n*th digit of the ternary expansion of *x*, this representation directly reflects the construction process of the Cantor set. Additionally, since *c* is continuous and non-decreasing, taking its supremum in $[0, 1] \setminus C$ suffices.

Topology and Fractal Geometry

Now that we have a preliminary understanding of the Cantor set, we can further explore fractal geometry in topology.

Self-similarity

Self-similarity refers to a phenomenon where a part of an object exhibits a similar structure to the whole or another part of the object. This means that regardless of whether we enlarge or reduce the object, its local structures will repeat in a certain way, thereby maintaining overall similarity. This similarity can be either exact or approximate. Many natural phenomena exhibit self-similarity, a one highly representative example is the Romanesco broccoli. The concept of fractals defined as objects possessing self-similarityc.

The figure depicts the Mandelbrot set, which is a collection of points forming a fractal on the complex plane. Consider a complex number c, and first define a complex quadratic polynomial $f_c(z) =$ $z^2 + c$. We iterate $f_c(z)$ starting from $z_0 = 0$, and the choice of parameter c may cause |z| to tend to infinity after an infinite number of iterations. However, if |z| remains within a finite range at all times, we say that c belongs to the Mandelbrot set.



Here, without delving into its algebraic properties, let's observe the image. The Mandelbrot set exhibits self-similarity at various scales. No matter how much you magnify the image, you can see similar structures repeating, which is a typical characteristic of fractals.

Hausdorff Dimension

The Hausdorff measure is an extension of the Lebesgue measure in metric spaces, providing a more generalized measure applicable to subsets of any metric space, including those nonstandard sets with fractal structures. However, in this article, we focus on the concept of Hausdorff dimension, which describes the dimension of sets under the Hausdorff measure. As for various properties of the Hausdorff measure, we'll delve into them later in a specialized short discourse on topology.

In Euclidean space, the dimension of a line is 1, a plane is 2, and a solid is 3, which is basic common knowledge. However, the "dimension" we often use in Euclidean space usually applies only to regular sets. For many fractal sets, such as the Koch curve, Mandelbrot set,

etc., we cannot describe their geometric features using integer dimensions, hence the need for a dimension to describe them. The Hausdorff dimension is a way to describe the dimension of a topological space, which is why sometimes when people refer to "fractals," they mean geometric bodies with fractional Hausdorff dimensions.

We will introduce the definition of the Hausdorff dimension through the definition of the Hausdorff outer measure. First, consider a metric space (X, d), and let $A \subseteq X, \delta > 0$. $\mathcal{D}(U_i)$ denotes the maximum distance between points in U_i , defining:

$$H^s_\delta(A) = \inf\left\{\sum_{i=1}^\infty \mathcal{D}(U_i): A \subseteq igcup_{i=1}^\infty U_i, \ \mathcal{D}(U_i) < \delta
ight\}$$

It meaning that H takes the infimum over all possible covering ways, and then we define the s-dimensional Hausdorff outer measure of this subset A

$$H^s(A):=\lim_{\delta o 0^+}H^s_\delta(A)$$

The Hausdorff dimension H_H of a set A is defined as the smallest value α for which the above outer measure exists with a finite nonzero value.

$$D_H(A) = \inf \left\{ lpha : \mathcal{H}^lpha(A) = 0
ight\}$$

 D_H is the dimension of the set, and our above definition ensures that the outer measure of set A in this dimension is nonzero, thus this dimension is considered as the minimal dimension describing the geometric structure of the set. You see, in essence, doesn't this theoretically solve the problem of the "Lebesgue zero measure" of the Cantor set?

Of course, this is only theoretical; we need a more intuitive approach and method to calculate the Hausdorff dimension, looking at it from a different perspective. If we can cover set A with N sets scaled by ϵ times the original set, then its Hausdorff dimension is

$$D_H(A) = rac{\log N}{\log rac{1}{\epsilon}}$$

Isn't this more straightforward?

We consider a bounded Lebesgue null-measure set X on the real number line. Then $D_H(X) \in [0,1]$ possesses some obvious properties:

- If X is an interval, then $D_H = 1$.
- If X is a single point or a countable set, then $D_H = 0$.
- If $X \subset Y$, then $D_H(X) \leq D_H(Y)$.

Some type of Cantor Set

There are many kinds of Cantor sets, and classical Cantor sets are precisely self-similar. This means that every subset of the Cantor set is "homeomorphic" to the entire Cantor set itself (defined later), straightforwardly implying their equivalence in the topological sense.

Simple Cantor Sets

From an intuitive geometric perspective, we know that the Cantor ternary set is constructed by infinitely dividing into three equal subintervals. If we have a Cantor ternary set C, we can divide it into two parts (left and right), denoted as C_L and C_R , which have equal lengths on the real number line, each being one third of C. Therefore, we can consider covering C with two sets scaled down to $\frac{1}{3}$ of the original length, i.e., N = 2, $\epsilon = \frac{1}{3}$, and then calculate the Hausdorff dimension of C as follows:

$$D_H(C) = rac{\log 2}{\log 3} = \log_3 2 pprox 0.630929754$$

The Cantor ternary set can be further generalized to multifractal sets. For example, if we choose to remove five-thirds of the interval, i.e., the lengths at both ends are $\frac{1}{5}$ of the original, we can still take N = 2, $\epsilon = \frac{1}{5}$. If this new set is called \hat{C} , then we have:

$$D_H(\hat{C}) = rac{\log 2}{\log 5} = \log_5 2 pprox 0.430676558$$

Of course, such sets are very simple, and we can continue this process indefinitely. Now, we

can move on to discussing more complex Cantor sets in higher dimensions.

Sierpiński carpet

We can extend the construction of the Cantor set on [0, 1] to two-dimensional, threedimensional, or even higher dimensions. The Sierpiński carpet (Polish: Dywan Sierpińskiego) is a self-similar set, a type of fractal. In the two-dimensional case, we divide a square into nine equal squares and remove the middle one, leaving eight, and repeat the process.



We use iteration here. In the first step, we can define 1.

For $n \in \mathbb{N}$, we can easily define the area of the nth iteration as follows:

$$S_n = \left(rac{8}{9}
ight)^n = S_{n-1} - rac{1}{9} \cdot \left(rac{8}{9}
ight)^n$$

It's worth mentioning that this is based on the premise of an initial side length of 1. If the side length is a, we can express it directly as follows:

$$S_n = \left(rac{8}{9}
ight)^n \cdot a^2 = S_{n-1} - rac{1}{9} \cdot \left(rac{8}{9}
ight)^n \cdot a^2$$

When n approaches infinity, it is defined as the Sierpinski carpet:

$$S = \lim_{n o \infty} S_n$$

Many aspects of the Sierpiński carpet bear a striking resemblance to the Cantor set, where each small square is a fraction of $\frac{1}{3}$ of the original, allowing us to describe it using base-3 notation. We consider the Sierpiński carpet as a subset in the two-dimensional plane, denoted

as $S_0 = [0,1] imes [0,1].$ Thus, for large n, we have:

$$S = \left\{ (x,y) \in S_0 \; \middle| \; x = \sum_{n=1}^\infty rac{a_n}{3^n}, \; y = \sum_{n=1}^\infty rac{b_n}{3^n}, (a_n,b_n)
eq (1,1), \; a_n, b_n \in \{0,1,2\}
ight\}$$

Clearly, as mentioned earlier in this section, the Sierpinski carpet is also a precise selfsimilar example. After one iteration, each small square is $\frac{1}{3}$ of the original, resulting in eight squares. Therefore, the Hausdorff dimension of the Sierpinski carpet is:

$$D_H(S) = rac{\log 8}{\log 3} = \log_3 8 pprox 1.892789261$$

Expanding the Sierpinski carpet to three dimensions yields the Menger sponge, dividing the cube into 27 parts and removing the central 7 parts. Similarly, its Hausdorff dimension is:

$$D_H(M) = rac{\log 20}{\log 3} = \log_3 20 pprox 2.7268330274$$

However, strictly speaking, this definition of the Sierpinski carpe is not a two-dimensional extension of the Cantor set. To be so, the square should be divided into nine parts and then removing the three squares in the middle column after dividing each into third represented in base-3 notation as:

$$\hat{C} = \left\{ (x,y) \in S_0 \; \middle| \; x = \sum_{n=1}^{\infty} \frac{a_n}{3^n}, \; y = \sum_{n=1}^{\infty} \frac{b_n}{3^n}, \; a_n, b_n \in \{0,2\} \right\}$$

It is obvious that the Hausdorff dimension of this object is:

$$D_H(\hat{C}) = rac{\log 4}{\log 3} = \log_3 4 pprox 1.261859507$$

Koch Curve

The Koch curve is also a classic fractal, initially an equilateral triangle. Each iteration divides the sides into three equal parts and constructs an equilateral triangle outward with the middle segment as the base, then removes the middle part as the base.



The example above depicts a Koch curve iterated four times, forming an overall snowflake-like shape (thus sometimes referred to as Koch snowflake). The blue box represents a local segment iterated four times.

It is evident that when the initial triangle has a side length of a, the perimeter after the nth iteration is:

$$L_n = 3a \left(rac{4}{3}
ight)^n$$

As n approaches infinity, the perimeter also tends to infinity:

$$L=\lim_{n o\infty}L_n=3a\left(rac{4}{3}
ight)^\infty=\infty$$

Hence the notion of the "infinite length of the British coastline" should actually be interpreted as the coastline's uncertainty, depending on the measurement precision. In reality, your precision cannot be infinitely high, can it?

The Koch curve is divided into four equal parts in one iteration, each of which is $\frac{1}{3}$ of the original, so the Hausdorff dimension of the Koch curve is obviously:

$$D_H(K) = rac{\log 4}{\log 3} = \log_3 4 pprox 1.261859507$$

Suppose we want to construct the initial equilateral triangle of the Koch curve with side length $K_0 = 1$, then its initial area is $A_0 = \frac{\sqrt{3}}{4}$, and the number of triangles added in the (n+1)th iteration is $3 \cdot 4^n$. Thus, we can obtain the area formula:

$$A_{n+1} = A_n + rac{3\cdot 4^n}{9^{n+1}}A_0$$

Each iteration adds triangles four times the original quantity, and the area is one-ninth of the triangles from the previous iteration, so we have $4 \cdot \frac{1}{9} = \frac{4}{9}$, but we can calculate the area of the Koch snowflake using geometric series:

$$egin{aligned} A &= \lim_{n o \infty} A_n = A_0 \lim_{n o \infty} \left[1 + rac{1}{3} \sum_{i=0}^{n-1} \left(rac{4}{9}
ight)^i
ight] \ &= rac{\sqrt{3}}{4} \left(1 + rac{1}{3} \cdot rac{1}{1 - rac{4}{9}}
ight) \ &= rac{\sqrt{3}}{4} \cdot rac{8}{5} \ &= rac{2\sqrt{3}}{5} \end{aligned}$$

As we just derived, for an equilateral triangle with initial area A_0 , the area of the Koch curve is $\frac{8}{5}A_0$.

Homeomorphism

A topological space constructed from the Cantor set is called the Cantor space. A topological space that is homeomorphic to the Cantor set is defined as the Cantor space.

Consider two topological spaces $(X, \tau_X), (Y, \tau_Y)$, and a mapping $f : X \to Y$. If both f and its inverse f^{-1} are continuous, and f is a bijection, then we say (X, τ_X) and (Y, τ_Y) are homeomorphic. Homeomorphisms preserve the topological properties of spaces, meaning two homeomorphic spaces are topologically equivalent; they share the same topological properties, although their geometric structures and metrics may differ.

Cantor sets share some common topological properties; they are all disconnected, perfect, have no interior points, and are compact sets, and they are metrizable. A topological space is the Cantor space if and only if it satisfies these properties. According to Brouwer's theorem, the Cantor set is, in the sense of homeomorphism, the unique perfect, compact, zerodimensional, metrizable space.

References

- 1. Cantor, G.(1874). On a Property of the Class of all Real Algebraic Numbers.
- 2. Jason Baker, Kyle Henke, Michael Sanchez. *Lebesgue Measure and The Cantor* Set
- 3. ZIQIAN (ALEXA) JIN. CANTOR SETS IN TOPOLOGY, ANALYSIS, AND FINANCIAL MARKETS
- 4. Falconer, K. J. (1986). The Geometry of Fractal Sets
- 5. Mark Pollicott. Cantor sets, Hausdorff dimension and their applications
- 6. Whyburn, Gordon (1958). Topological chcracterization of the Sierpinski curve

Wiki:

- https://en.wikipedia.org/wiki/Cantor_set
- https://en.wikipedia.org/wiki/Hausdorff_space

MAKIROR gzanan@gmail.com 16:24 17/02/2024