CANTOR SETS
Ferdinand Chovanec

Abstract: The Cantor set is an interesting example of an uncountable set of measure zero and has many interesting properties and consequences in the fields of set theory, topology, and fractal theory. The principal aim of this paper is to introduce a generator of finite subsets of the basic Cantor (ternary) set and its generalization to the Cantor \( n \)-ary set. We compute the fractal dimension of these Cantor sets.

Keywords: fractal, Cantor set, fractal dimension

1. INTRODUCTION

Two years ago we have gained the grant project of Ministry of Defence of the Slovak Republic AGA-01-2008 „Statistical analysis of the influence of the semiconductor system interface nanoroughness on its optical properties“. In this project we have used fractal approach to the study of the surface of solid materials (see [5], [6], [7]). In this study we used the multifractal singularity spectrum function \( f(\alpha) \) to describe the development of the surface fractal properties. We compared experimental \( f(\alpha) \) curves with theoretical singularity spectra, obtained by computer simulation of surface structure based on Cantor numbers properties. This required the use of the Cantor set with the cardinality more than \( 10^5 \). Moreover, we applied different types of Cantor sets.

In the following paper we introduce a generator of the basic Cantor (ternary) set and the generalized Cantor \( (n\text{-ary}) \) set. We compute the Hausdorff-Besicovitch fractal dimension of the Cantor sets.

2. THE CANTOR SETS

The Cantor (ternary) set was first published in 1883 by German mathematician Georg Cantor [1]. The Cantor set plays a very important role in many branches of mathematics, above all in set theory, chaotic dynamical systems and fractal theory.

2.1 The Cantor ternary set

The basic Cantor (ternary) set is a subset of the interval \([0,1]\) and has many definitions and many different constructions. Although Cantor originally provided a purely abstract definition, the most accessible is the „middle-thirds“ or ternary set construction. Begin with the closed real interval \( I_0 = [0,1] \) and divide it into three equal subintervals. Remove the central open interval \( \left(\frac{1}{3}, \frac{2}{3}\right) \) such that

\[ I_1 = [0,1] - \left(\frac{1}{3}, \frac{2}{3}\right) = \left[0, \frac{1}{3}\right] \cup \left[\frac{2}{3}, 1\right]. \]

Next, subdivide each of these two remaining intervals into three equal subintervals and from each remove the central third and continue in the previous manner.

\[ I_2 = \left[\left[0, \frac{1}{3}\right] - \left(\frac{1}{9}, \frac{2}{9}\right)\right] \cup \left[\left[\frac{2}{3}, \frac{7}{9}\right] - \left(\frac{5}{9}, \frac{8}{9}\right)\right] = \left[\left[0, \frac{1}{9}\right] \cup \left[\frac{2}{9}, \frac{7}{9}\right] \cup \left[\frac{8}{9}, 1\right]\right]. \]

In this way we obtain a sequence of closed intervals – one in the zero step, two after the first step, four after the second step, eight after the third step, etc. \((2^k\text{ intervals of length } \left(\frac{1}{3}\right)^k) \) after the \( k \)th step. This process is visible in the Figure 1.

\[ \begin{array}{cccc}
  k = 0 & 1 & 2 & 3 \\
  I_0 & 0 & 1 & \frac{1}{3} \\
  I_1 & 0 & 1 & \frac{1}{3} \\
  I_2 & 0 & 1 & \frac{1}{3} \\
  I_3 & 0 & 1 & \frac{1}{3} \\
\end{array} \]

Fig. 1 Initial steps of the construction of the Cantor ternary set

Finally, we define the Cantor ternary set \( \mathbb{C}(3) \) as follows:

\[ \mathbb{C}(3) = \bigcap_{k=0}^{\infty} I_k. \]

This construction does not provide a sufficient view of elements of the Cantor set. For these reasons we propose the construction of Cantor numbers, where subsets \( \mathbb{C}_k(3) \) of the Cantor set \( \mathbb{C}(3) \) are endpoints of the closed intervals creating \( I_k \), \( k = 0, 1, 2, \ldots \) (see Fig. 1).

Denote by the symbol \( |A| \) the cardinality of a set \( A \). Then we have

\[ \mathbb{C}_0(3) = \{0, 1\} \text{ and } |\mathbb{C}_0(3)| = 2 = 2^0. \]
It is very well visible that the set \( C(3) \) is easily programmable for every nonnegative integer \( k \).

If \(|C_k(3)| = x\) then \( k = \left\lfloor \frac{\log x}{\log 3} \right\rfloor \). For example, if we need 1000 numbers of the Cantor set \( C(3) \), then it suffices to take the set \( C_{9}(3) \).

We should like to emphasize that this method does not allow to construct any number of the Cantor set, only endpoints of the closed intervals remaining after removing the middle thirds. There are numbers in the Cantor ternary set which are not interval endpoints. One example of such number is \( 0.132 \).

There is a natural question, how can we recognize elements (numbers) of the Cantor ternary set. It allows the triadic expansion of its numbers.

Let \( x \in [0, 1] \). Then its expansion with respect to base 3 (3-adic expansion) is given by the following expression

\[
x = a_1 3^{-1} + a_2 3^{-2} + a_3 3^{-3} + \cdots + a_n 3^{-n} + \cdots
\]

where \( a_n \in \{0,1,2\} \) for every \( n = 1,2,3,\ldots \). Then we write

\[
x = 0. a_1 a_2 a_3 \ldots a_n \ldots \in (3).
\]

For example, \( 0.5 = 0.1111 \ldots \in [3] (= 0. \bar{1} \in [3]) \), because

\[
1.3^{-1} + 1.3^{-2} + \cdots + 1.3^{-n} + \cdots = \sum_{n=1}^{\infty} \frac{1}{3^n} = \frac{1}{3} \sum_{n=1}^{\infty} \left( \frac{1}{3} \right)^{n-1} = \frac{1}{3} \cdot \frac{1 - \frac{1}{3}}{1 - \frac{1}{3}} = \frac{1}{3} \cdot \frac{2}{3} = \frac{1}{3}
\]

There are some numbers that have a terminating expansion and simultaneously an infinite expansion. Let us take \( x = \frac{2}{3} \). Then \( x = 0.1_{13} \) and on the other hand we have

\[
2.3^{-2} + 2.3^{-n} + \cdots = \sum_{n=2}^{\infty} \left( \frac{1}{3} \right)^{n-1} = \frac{2}{3} \sum_{n=2}^{\infty} \left( \frac{1}{3} \right)^{n-2} = \frac{2}{3} \left( \frac{1}{3} \right)^2 \frac{1}{1 - \frac{1}{3}} = \frac{1}{3}
\]

In ternary notation we have the similar equivalence that \( 0.1_{13} \) equals \( 0.0222 \ldots \in [3] (= 0. \bar{0} \bar{2} \bar{2} \bar{2} \bar{2} \bar{2} \bar{2} \bar{2} \bar{2} \ldots \in [3] \). All numbers strictly between \( \frac{1}{3} \) and \( \frac{2}{3} \) = \( 0.2_{13} \) must have a digit „1“ somewhere in the middle of the digit sequence. Therefore, these numbers are not in the Cantor ternary set.

In the general we can characterize any number of the Cantor ternary set in the following way.

**Theorem 1** [10] The Cantor ternary set is the set of numbers in \([0, 1]\) for which there is a triadic expansion that does not contain the digit „1“.

We are able easily to verify that \( \frac{1}{4} = 0.02020202\ldots \) (= 0.02) and hence \( \frac{1}{4} \in \mathbb{C}(3) \).

From the mathematical point of view, the Cantor set has many interesting properties:

- The Cantor set is compact (i.e. closed and bounded).
- The Cantor set does not contain any open set.
- The Cantor set is perfect (and hence uncountable).
- The Cantor set has length zero.

We refer to readers the book [11] for detailed proofs of the above mentioned properties.
2.2 The Cantor quintuple set

Motivated by the ternary Cantor set $C(3)$, we construct the Cantor quintuple set $C(5)$. We begin with the closed real interval $J_0 = [0,1]$ again and divide it into five equal subintervals. Remove the open intervals $\left(\frac{1}{5}, \frac{2}{5}\right)$ and $\left(\frac{2}{5}, \frac{3}{5}\right)$ such that

$$J_1 = [0,1] - \left(\frac{1}{5}, \frac{2}{5}\right) \cup \left(\frac{3}{5}, \frac{4}{5}\right) = \left[0, \frac{1}{5}\right] \cup \left[\frac{2}{5}, \frac{3}{5}\right] \cup \left[\frac{4}{5}, 1\right].$$

We subdivide each of these three remaining intervals into five equal subintervals and from each remove the second and fourth open subinterval, and continue in the previous manner. In this way we obtain a sequence of closed intervals one in the zero step, three after the first step, nine after the second step, etc. ($3^k$ intervals of length $\left(\frac{1}{5}\right)^k$ after the $k$th step). This process is visible in the Figure 2.

![Fig. 2 Initial steps of the construction of the Cantor ternary set](image)

We define the Cantor quintuple set $C(5)$ by the formula

$$C(5) = \bigcap_{k=0}^{\infty} J_k.$$

Now we construct a generator of numbers of the set $C(5)$. Let $C_k(5)$ be sets of endpoints of closed intervals creating $J_k$, $k = 0,1,2,\ldots$ (see Fig. 2). Then

$$C_0(5) = \{0,1\} \quad \text{and} \quad |C_0(5)| = 2 = 2.3^0.$$

$$C_1(5) = \left\{0, \frac{1}{5}, \frac{2}{5}, \frac{3}{5}, \frac{4}{5}, \frac{5}{5}\right\} = \bigcup_{j=0}^{5} \left\{\frac{j}{5}\right\},$$

$$|C_1(5)| = 6 = 3 \cdot 2.3^0 = 2.3^1.$$

$$C_2(5) = \left\{0, \frac{1}{5^2}, \frac{2}{5^2}, \frac{3}{5^2}, \frac{4}{5^2}, \frac{5}{5^2}\right\} \cup \bigcup_{j=0}^{5} \left\{\frac{j+2(5^2)}{5^2}\right\},$$

$$|C_2(5)| = 18 = 3 \cdot |C_1(5)| = 3 \cdot 2.3^1 = 2.3^2.$$

$$C_3(5) = \bigcup_{j=0}^{5} \bigcup_{i_1=0}^{2} \bigcup_{i_2=0}^{2} \left\{\frac{j+2(5^2i_2+5i_1)}{5^3}\right\},$$

$$|C_3(5)| = 54 = 3 \cdot |C_2(5)| = 3 \cdot 2.3^2 = 2.3^3.$$

We get a sequence $\left(C_k(5)\right)_{k=0}^{\infty}$ of finite subsets of the Cantor quintuple set such that

$$C_0(5) \subset C_1(5) \subset \ldots \subset C_k(5) \subset C_{k+1}(5) \subset \ldots$$

If $|C_k(5)| = x$ then $k = \frac{\log x}{\log 3}$. To have 1000 numbers of the Cantor set $C(5)$ it suffices to take the set $C_9(5)$.

There is a natural question: How can we characterize numbers of the Cantor quintuple set? Expected answer is – by means of their 5-adic expansion.

It is not difficult to verify that $\frac{1}{5} = 0.1_5 = 0.041_5$, $\frac{3}{5} = 0.3_5 = 0.241_5$ and for that reason all numbers strictly between $\frac{1}{5} = 0.041_5$ and $\frac{2}{5} = 0.21_5$, as well as between $\frac{3}{5} = 0.241_5$ and $\frac{4}{5} = 0.41_5$, must have digits „1“and „3“ somewhere in the middle of the digit sequence. Therefore, these numbers are not in the Cantor ternary set. For these reasons, the following assertion is true.

**Theorem 2** The Cantor quintuple set is the set of numbers in $[0,1]$ for which there is a 5-adic expansion that does not contain the digits „1“and „3“.

2.3 The Cantor $n$-ary set

In this section we generalize the construction of the Cantor ternary and quintuple set.

Let $n = 2m+1, m = 1, 2, 3, \ldots$ . We start with the closed real interval $K_0 = [0,1]$ and divide it into $n$ equal subintervals. Remove the open intervals $\left(\frac{1}{n}, \frac{2}{n}\right), \left(\frac{3}{n}, \frac{4}{n}\right), \ldots, \left(\frac{n-2}{n}, \frac{n-1}{n}\right)$ such that
We subdivide each of these \((m+1)\)-remaining intervals into \(n\) equal subintervals and from each remove the 2nd, 4th, \(\ldots\), \((2m)\)th open subinterval, and continue in the previous manner. In this way we obtain a sequence of closed intervals \(K_k\) – one in the zero step, \(m+1\) after the first step, \((m+1)^2\) after the second step, etc. \((m+1)^k\) intervals of length \(\left(\frac{1}{n}\right)^k\) after the \(k\)th step.

The Cantor \(n\)-ary set is defined by the formula

\[
\mathcal{C}(n) = \bigcap_{k=0}^{\infty} K_k.
\]

Now we construct a sequence \(\left(\mathcal{C}_k(n)\right)_{k=0}^{\infty}\) of numbers of the the Cantor \(n\)-ary set such that

\[
\mathcal{C}_0(n) = \{0, 1\} \text{ and } |\mathcal{C}_0(n)| = 2 = 2 \cdot (m+1)^0.
\]

\[
\mathcal{C}_1(n) = \left\{0, \frac{2}{n}, \ldots, \frac{n-1}{n}, n\right\} = \bigcup_{j=0}^{n} \left\{\frac{j}{n}\right\},
\]

\[
|\mathcal{C}_1(n)| = n + 1 = 2m + 2 = (m+1)|\mathcal{C}_0(n)| = 2(m+1)^1.
\]

\[
\mathcal{C}_k(n) = \bigcup_{j=0}^{n} \bigcup_{i_k=0}^{m-1} \bigcup_{i_{k-1}=0}^{m-1} \left\{\frac{j + 2(n^{k-1}i_k - 1 + \cdots + ni_1)}{n^k}\right\}.
\]

\[
|\mathcal{C}_k(n)| = |\mathcal{C}_k(2m+1)| = (m+1)|\mathcal{C}_{k-1}(n)| = (m+1)2(m+1)^{k-1} = 2(m+1)^k.
\]

If \(|\mathcal{C}_k(n)| = x\) then \(k = \frac{\ln x}{\ln(m+1)}\).

The sum of the lengths of the removed intervals is equal to 1, because

\[
m \left(\frac{1}{n}\right) + m(m+1)\left(\frac{1}{n}\right)^2 + m(m+1)^2\left(\frac{1}{n}\right)^3 + \cdots + m(m+1)^{k-1}\left(\frac{1}{n}\right)^k + \cdots = \sum_{k=0}^{\infty} m(m+1)^{k-1}\left(\frac{1}{n}\right)^k = \frac{m}{n} \frac{1}{1 - \frac{1}{n}} = \frac{m}{n} = \frac{n}{n-m} = 1.
\]

Videlicet, the Lebesque measure of the Cantor set \(\mathcal{C}(n)\) is zero for every \(n = 2m + 1, m = 1, 2, \ldots\).

### 3. Cantor Sets as Fractals

The Cantor set is the prototype of a fractal. A fractal is an object which appears self-similar under varying degrees of magnification. One of the typical features of fractals is their fractal dimension. The fractal dimension is essentially a measure of self-similarity (it is sometimes referred to as the similarity dimension). The fractal dimension is greater than the topological dimension. There are many specific definitions of fractal dimension. The basic type of fractal dimension is the Hausdorff-Besicovitch dimension, which is based on the definition of the Hausdorff measure [2]. One version of the Hausdorff-Besicovitch dimension is given by the formula

\[
D = \frac{\log N}{\log r},
\]

where \(N\) is the number of self-similar pieces and \(r\) is the contraction factor.

We note that there are several different ways of computing the fractal dimension (see [8], [9]).

Now we compute the fractal dimension of the Cantor sets. Let us assume the Cantor ternary set \(\mathcal{C}(3)\). We have \(2^k\) (self-similar) intervals of length \(\left(\frac{1}{3}\right)^k\) after the \(k\)th step, so \(N = 2^k\) and \(r = \left(\frac{1}{3}\right)^k\). Then

\[
D(\mathcal{C}(3)) = \frac{\log 2^k}{\log 3^k} = \frac{k \log 2}{k \log 3} = \frac{\log 2}{\log 3} \approx 0.631.
\]

The fractal dimension of \(\mathcal{C}(3)\) is the same in every step.

In the case of the Cantor quintuple set \(\mathcal{C}(5)\) we have \(3^k\) (self-similar) intervals of length \(\left(\frac{1}{5}\right)^k\) after the \(k\)th step, therefore, \(N = 3^k\) and \(r = \left(\frac{1}{5}\right)^k\). Then

\[
D(\mathcal{C}(5)) = \frac{\log 3^k}{\log 5^k} = \frac{\log 3}{\log 5} \approx 0.683.
\]

In the case of the Cantor set \(\mathcal{C}(n)\), \(n = 2m + 1, m = 1, 2, 3, \ldots\), we have \((m+1)^k\) (self-similar) intervals of the length \(\left(\frac{1}{n}\right)^k\) after the \(k\)th step, so \(N = (m+1)^k\) and \(r = \left(\frac{1}{n}\right)^k\). Then

\[
D(\mathcal{C}(n)) = \frac{\log(m+1)^k}{\log(2m+1)^k} = \frac{\log(m+1)}{\log(2m+1)} < 1.
\]

The Cantor set \(\mathcal{C}(n)\) is an object with fractal dimensionality less than one, between a point (topological dimensionality of zero) and a line (topological dimensionality one), for every \(n = 2m + 1, m = 1, 2, 3, \ldots\).
**Theorem 4** Let \( \{C(2m+1)\}_{m=1}^{\infty} \) be a sequence of Cantor sets. Then a sequence of their dimensions \( \{D(C(2m+1))\}_{m=1}^{\infty} \) is increasing and, moreover,
\[
\lim_{m \to \infty} D(C(2m+1)) = 1.
\]

**Proof** We define three real functions
\[
f: [1, \infty) \to \mathbb{R}, \quad f(x) = \log(x+1),
g: [1, \infty) \to \mathbb{R}, \quad g(x) = \log(2x+1),
h: [1, \infty) \to \mathbb{R}, \quad h(x) = \frac{\log(x+1)}{\log(2x+1)}.
\]

To prove that the function \( h \) is increasing on the interval \([1, \infty)\), we compute its first derivation.
\[
h'(x) = \left( \frac{\log(x+1)}{\log(2x+1)} \right)' = \frac{(\log e)(\log(2x+1)) - (\log(x+1))\log e}{x+1} - \frac{2x+1}{\log(2x+1)}
\]
\[
= \frac{(\log e)[\log(2x+1)^{2x+1} - \log(x+1)^{x+1}]}{(\log(x+1))^{2}(x+1)(2x+1)}
\]
\[
= \frac{(\log e)\left[ \frac{\log(2x+1^{2x+1})}{(x+1)^{x+1}} \right]}{(\log(x+1))^{2}(x+1)(2x+1)}.
\]

Note that
\[
\frac{\log e}{(\log(x+1))^{2}(x+1)(2x+1)} > 0
\]
for every \( x \in [1, \infty) \).

We have
\[
\log \frac{(2x+1)^{2x+1}}{(x+1)^{x+1}} = \log \frac{(2x+1)^{x+1}(2x+1)^x}{(x+1)^{x+1}} = \log \left( \frac{2x+1}{x+1} \right)^{x+1}(2x+1)^x
\]
\[
= \log \left( \frac{2x+1}{x+1} \right)^{x+1}(2x+1)^x
\]
\[
= \log \left(1 + \frac{x}{x+1} \right)^{x+1} + \log(2x+1)^x
\]
\[
= (x+1) \log \left(1 + \frac{x}{x+1} \right) + x \log(2x+1) > 0
\]
for every \( x \in [1, \infty) \).

We proved that \( h'(x) > 0 \) for every \( x \in [1, \infty) \), so the function \( h(x) = \frac{\log(x+1)}{\log(2x+1)} \) is increasing on \([1, \infty)\) and hence, the sequence \( \{\frac{\log(m+1)}{\log(2m+1)}\}_{m=1}^{\infty} \) is increasing too.

Let us calculate
\[
\lim_{x \to \infty} \frac{\log(x+1)}{\log(2x+1)} = \lim_{x \to \infty} \frac{(\log(x+1))'}{(\log(2x+1))'} = \lim_{x \to \infty} \frac{\log e}{x+1} = \lim_{x \to \infty} \frac{2x+1}{2(x+1)} = \lim_{x \to \infty} \frac{2}{2} = 1.
\]

Hence,
\[
\lim_{m \to \infty} D(C(2m+1)) = \lim_{m \to \infty} \frac{\log(m+1)}{\log(2m+1)} = 1.
\]

\[\blacksquare\]

5. CONCLUSION

The Cantor set has many interesting properties and consequences in the fields of set theory, topology, and fractal theory. An application of fractal theory to the theory of algebraic structures was presented at the Tenth International Conference on Fuzzy Sets Theory and Applications in Liptovský Ján (February 1–5, 2010) [3]. A fractal difference poset (a fractal D-poset, in short) was defined as a special pasting of MV-algebras [4]. In this sense, the Cantor fractal D-poset is the „0-1-pasting“ of MV-algebras.

This work has been supported in part by the Slovak Research and Development Agency under the contract No. APVV-0071-06, Grant VEGA 2/0032/09 and AGA-01-2008 Ministry of Defence of the Slovak Republic Grant Agency.

References


Ferdinand Chovanec
Department of Informatics
Armed Forces Academy of Gen. M. R. Štefánik
Demänová 393
031 06 Liptovský Mikuláš 6
E-mail: ferdinand.chovanec@aos.sk