**Real Analysis Exchange**
Vol. 23(2), 1997–1998, pp. 767–771

Eduardo J. Dubuc, Departamento de Matemáticas, F.C.E.y N. UBA, (1428) Buenos Aires, Argentina. e-mail: `edubuc@dm.uba.ar`

# A NOTE ON CANTOR SETS

**Abstract**

The Cantor set is constructed by the iterate deletion of a middle interval equidistant from the end points. It is well known that the sums of points in the set cover completely the real line. It was an open problem to know if this property was still true for the sets obtained when the deleted interval is not any more equidistant from the end points. In this note we answer this question positively. We give a simple proof that reflects the geometric nature of the problem, and that is a variation on an old idea that goes back to Steinhaus[2].

## 1 Introduction

Let us consider the classical Cantor ternary set C. It is obtained by first removing from [0, 1] the middle third $(1/3, 2/3)$, then removing the middle thirds $(1/9, 2/9)$ and $(7/9, 8/9)$ of the remaining intervals, and so on. That is, every time we leave the closed 1/3 left and right parts of each interval. Although C has measure zero, in 1917 Steinhaus [2] proved that $C+C = [0, 2]$. He did so by means of a beautiful and powerful geometric idea. In what follows we shall describe this idea, and show how it can be modified so as to prove a suitable statement for generalized non central cantor sets.

Consider a non symmetric version of the classical Cantor set. We start by first removing from $[0, 1]$ the interval $(1/3, 1/2)$, then removing the intervals $(1/9, 1/6)$ and $(4/6, 3/4)$ of the remaining intervals, and so on. That is, every time we leave the closed 1/3 left part and the closed 1/2 right part of each interval. Note that this subdivision of $[0, 1]$ that generates C produces a corresponding subdivision of $[0, 1] \times [0, 1]$ that generates $C \times C$. Let $\mathbf{C}_k$ be the collection of rectangles corresponding to step $k$. Given a set $S \subset [0, 1]^2$, we shall indicate by $sum(S)$ the subset of the interval $[0, 2]$ formed by those numbers which are the sums of the two coordinates of points in $S$. The geometric idea of Steinhaus is to show that for every $k$ $sum(\mathbf{C}_k) = [0, 2]$. He does this

Key Words: **Cantor ternary set**
Mathematical Reviews subject classification: Primary: 26A03, Secondary: 54F50
Received by the editors June 22, 1997

for the classical case where the removed interval is the middle $(1/3, 2/3)$, and thus the rectangles are squares. In our case, the rectangles become unboundedly distorted, and the argument fails. However, we can still make the things work by means of a simple but fundamental variation on the idea of Steinhaus. Consider the collection $\mathbf{G}_1$ of squares, two of step 1 and one of step 2 shown in Figure 1.

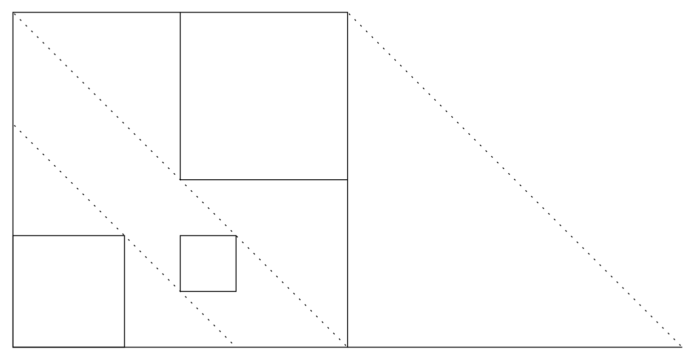


Figure 1.

It is clear that $sum(\mathbf{G}_1) = [0, 2]$. Now, each square in $\mathbf{G}_1$ contains three squares (from steps 2, 3 and 4) forming a similar configuration with the same proportions. It follows then that we also have $sum(\mathbf{G}_2) = [0, 2]$. We utilize only squares, avoiding thus any distortion. In this way we show that $sum(\mathbf{G}_k) = [0, 2]$ for every $k$. From this it readily follows that $sum(\mathbf{C}_k) = [0, 2]$ for every $k$. Since the $\mathbf{C}_k$ form a decreasing chain of subsets whose intersection is $C \times C$, it is immediate then that $sum(C \times C) = C + C = [0, 2]$.

This argument leads to the following generalization: Let $a$, $b$ be any two real numbers, $0 < a < b$, $a + b < 1$ (in the case above we have $a = 1/3$ and $b = 1/2$). We start by first removing from $[0, 1]$ the interval $(a, 1 - b)$, then removing the intervals $(aa, a - ab)$ and $(1 - b + ab, 1 - bb)$ of the remaining intervals, and so on. That is, every time we leave the closed $a$-percent left part and the closed $b$-percent right part of each interval. Then:

$$\underbrace{C + C + \cdots + C}_{n-\text{times}} = [0, n] \text{ if } (1 - (a + b))/ab \leq (n - 1).$$

Thus, the sum of sufficiently many copies of the Cantor set covers a whole interval of the line.

We shall prove now this statement. Our proof has been generalized recently to the case of Cantor sets obtained by iterate deletion, but with a different proportion of the intervals cut at each step [1]. We start setting some notation and conventions. We consider Cantor sets defined in the interval $[0,1] \subset \mathbb{R}$.

**Definition 1.** Let $a$, $b \in \mathbb{R}$, $0 < a < b$, $a + b < 1$.

a) Consider two symbols $\mathsf{a}$, $\mathsf{b}$ (the symbols are always different, even if the real numbers are the same). Given an interval $K = [v, w]$, of length $|K| = t = w - v$, we define $[\mathsf{a}]K = [v, v + a/t]$, and $[\mathsf{b}]K = [w - b/t, w]$. Given any word $w$ in the alphabet $\{\mathsf{a},\mathsf{b}\}$, we define inductively $[\omega]K = [e]([\omega']K)$ (for $\omega = e\omega'$, $e = \mathsf{a}$ or $e = \mathsf{b}$). Given two words $\alpha$ and $\beta$, of the same length, we say that they are *similar* if they have the same number of $\mathsf{a}$'s (thus, also the same number of $\mathsf{b}$'s). Clearly, for similar words we have $|[\alpha]K| = |[\beta]K|$.

b) Let $\mathbf{I}_s$ be the union of the intervals $[\omega][0,1]$, for all the words of length $s$. Thus, $\mathbf{I}_0 = [0,1]$, $\mathbf{I}_1 = [\mathsf{a}][0,1] \bigcup [\mathsf{b}][0,1] = [0,a] \bigcup [1-b,1]$ , etc. (the intervals in $\mathbf{I}_s$ are the intervals of step $s$ in the construction of the Cantor set). Clearly $\mathbf{I}_s \supset \mathbf{I}_{s+1}$. The intersection $C_{ab} = \bigcap_s \mathbf{I}_s$ is the Cantor set of ratios $a$ and $b$.

c) Given any set $A \subset \mathbb{R}$, and a natural number $n$, we denote by $(n)A$ the sum of $n$ copies of $A$. That is, $(n)A = A + A + \cdots + A = \{x \mid x = a_1 + a_2 + \cdots + a_n, a_i \in A\}$. Clearly $(n)C_{ab} = \bigcap_s (n)\mathbf{I}_s$.

**Definition 2.** a) Let $I = [0,1]^n \subset \mathbb{R}^n$ be the unit hypercube. By an *hypercube* we shall understand a set $H$ of the form $H = \mathbf{p} + tI$, with $\mathbf{p} \in \mathbb{R}^n$, and $t \in \mathbb{R}$, $t > 0$. The set $H$, which has sides of length $t$, is determined by the two diagonally opposed vertices $\mathbf{v} = \mathbf{p}$, and $\mathbf{w} = \mathbf{p} + (t, t, \ldots, t)$. Vice versa, any pair of points $\mathbf{v} = (v_1, \ldots, v_n)$, $\mathbf{w} = (w_1, \ldots, w_n)$, $v_i < w_i$, such that $w_i - v_i = t$ (for some constant value $t$), determine an hypercube $H = \mathbf{p} + tI$, $\mathbf{p} = \mathbf{v}$. We shall denote $H = [\mathbf{v}, \mathbf{w}] = [v_1, w_1] \times \cdots \times [v_n, w_n]$.

b) Given two finite collections of hypercubes, $\mathbf{G} = (G_1, \ldots G_m)$, $\mathbf{H} = (H_1, \ldots, H_k)$, we say that $\mathbf{G}$ is *contained* in $\mathbf{H}$, and write $\mathbf{G} \leq \mathbf{H}$, if for all $G_i$ there exists $H_j$ such that $G_i \subset H_j$.

c) By a *construction on hypercubes* we mean a rule that given an hypercube $H$ as input it assigns a finite collection of hypercubes $cH = \mathbf{H} = (H_1, \ldots, H_k)$ as output. We extend the construction to collections of hypercubes by defining $c\mathbf{H} = c(H_1, \ldots, H_k) = (cH_1, ..., cH_k)$.

d) Given an hypercube $H = [\mathbf{v}, \mathbf{w}]$, we indicate by $sum(H)$ the interval of length $nt$, $sum(H) = [\sum_i v_i, \sum_i w_i] = \{x \mid x = \sum_i x_i,\ x_i \in [v_i, w_i]\}$, and we extend this to collections of hypercubes by defining $sum(\mathbf{H}) = \bigcup_j sum(H_j)$. Notice that this will not be in general an interval.

**Proposition 3.** *Let $c$ be a construction on hypercubes such that: $cH \leq H$, and $sum(H) = sum(cH)$ for any $H$. Then:*

1) *Given any collection of hypercubes $\mathbf{H} = (H_1, \dots, H_k)$, we have $c\mathbf{H} \leq \mathbf{H}$, and $sum(\mathbf{H}) = sum(c\mathbf{H})$.*

2) *Given any hypercube $H$, let $\mathbf{G}_s$ be the sequence of collections of hypercubes determined by iterating $c$. That is, $\mathbf{G}_0 = (H)$, $\mathbf{G}_{s+1} = c\mathbf{G}_s$. Then, for all $s$ we have $(H) \geq \mathbf{G}_s \geq \mathbf{G}_{s+1}$, and $sum(H) = sum(\mathbf{G}_s)$.*

PROOF. The first part is clear. The second follows immediately by induction. $\square$

We say that a construction $c$ is *linear* if given any hypercube $H = \mathbf{p} + tI$, we have $cH = \mathbf{p} + t(cI)$. A linear construction is completely determined by defining it at the unit hypercube $I$.

**Definition 4. (the construction $c_{ab}$).** Let $0 < a < b$, $a + b < 1$. $c_{ab}$ is the linear construction defined at the unit hypercube as follows (see the figure 1 above for the case $n = 2$):

$c_{ab}(I) = (H_0, ..., H_n)$, $H_i = [\mathbf{v}_i, \mathbf{w}_i]$, $i = 0, 1, \dots, n$, where:

$$\begin{aligned}
v_0 =& (0, 0, \dots, 0).\\
v_i =& (1-b, \dots, \underbrace{1-b}_{i^{th} coor.}, a-ab, \dots, a-ab), \qquad i = 1, \dots, n-1.\\
v_n =& (1-b, 1-b, \dots, 1-b).\\
w_0 =& (a, a, \dots, a).\\
w_i =& (1-b+ab, \dots, \underbrace{1-b+ab}_{i^{th} coor.}, a, \dots, a), \qquad i = 1, \dots, n-1.\\
w_n =& (1, 1, \dots, 1).
\end{aligned}$$

Thus, $H_0 = \mathbf{v}_0 + aI$, $H_n = \mathbf{v}_n + bI$, and $H_i = \mathbf{v}_i + abI$, for $i = 1, \dots, n-1$.

**Proposition 5.** *Given any hypercube $H$,*

1) *$c_{ab}H \leq H$.*

2) *$sum(H) = sum(c_{ab}H)$ if and only if $(1 - (a + b))/ab \leq (n - 1)$.*

PROOF. It is enough to show the proposition for $H = I$. The fact that 1) holds follows directly from the definition, and $sum(I) = sum(c_{ab}I)$ will hold if and only if

$$\sum_j (\mathbf{v}_{i+1})_j \leq \sum_j (\mathbf{w}_i)_j, \text{ for } i = 0, \dots, n-1.$$

But,

$$\sum_j (\mathbf{w}_i)j = i(1 - b + ab) + (n - i)a, \text{ and}$$
$$\sum_j (\mathbf{v}_{i+1})_j = (i + 1)(1 - b) + (n - i - 1)(a - ab).$$

A simple calculation now shows that all of these conditions reduce to $(1-(a+b))/ab \leq (n-1)$. □

Given a word $\omega$ of length $s$ in the alphabet $\{\mathsf{a}, \mathsf{b}\}$, the interval $[\omega][0, 1]$ is an interval of step $s$ in the construction of the Cantor set $C_{ab}$. If $\alpha_1$, $\alpha_2, \dots, \alpha_n$ are $n$ similar words of length $s$, we say that the hypercube $H = [\alpha_1][0, 1] \times \cdots \times [\alpha_n][0, 1]$ is a *Cantor hypercube* of step $s$. Clearly, $c_{ab}(H)$ consists of two Cantor hypercubes of step $s+1$, and $n-1$ Cantor hypercubes of step $s+2$. Thus, if we iterate the construction $c_{ab}$ starting at the unit hypercube $I$, we obtain a sequence $\mathbf{G}_s$ of collections of hypercubes which are Cantor hypercubes of steps $i$, with $s \leq i \leq 2s$.

We now put together Propositions 3 and 5 and have:

**Theorem 6.** *Let $C_{ab}$ be the Cantor set in the interval $[0, 1]$ of ratios $a$ and $b$, $0 < a < b$, $a + b < 1$. If $(1 - (a + b))/ab \leq (n - 1)$, then $(n)C_{ab} = [0, n]$.*

PROOF. Let $\mathbf{F}_s$ be the collection of all Cantor hypercubes of steps $i$, where $s \leq i \leq 2s$. All the sides of an hypercube in $\mathbf{F}_s$ are contained in some interval of $\mathbf{I}_s$. It follows that $sum(\mathbf{F}_s) \subset (n)\mathbf{I}_s$. On the other hand, if $\mathbf{G}_s$ is the sequence defined by iterating the construction $c_{ab}$ starting at the unit hypercube $I$, $\mathbf{G}_s \leq \mathbf{F}_s$, thus $sum(\mathbf{G}_s) \subset sum(\mathbf{F}_s)$. By propositions 1 and 2, $sum(\mathbf{G}_s) = [0, n]$. It follows that $[0, n] \subset (n)\mathbf{I}_s$, for all $s$. Thus (see definition 1, c) $(n)C_{ab} = \bigcap_s (n)\mathbf{I}_s = [0, n]$. □

## References

[1] Carlos A. Cabrelli, Kathryn E. Hare and Ursula M. Molter, *Sums of Cantor sets*, Ergodic theory and Dynamical Systems, (to appear, 1997).

[2] Hugo. Steinhaus, *Nowa vlasnosc mnogosci G. Cantora*, Wektor (1917).