The Koch snowflake curve is tube-null

Viktor Harangi

April 15, 2010

Abstract

We show that the Koch curve is tube-null, that is, it can be covered by strips of arbitrarily small total width. In fact, we prove the following stronger result: the Koch curve can be decomposed into three sets such that each can be projected to a line in such a way that the image has Hausdorff dimension less than 1. The proof contains geometric, combinatorial, algebraic and probabilistic arguments.

2010 Mathematics Subject Classification: 28A12, 28A80 (primary).

DEPARTMENT OF ANALYSIS
EÖTVÖS LORÁND UNIVERSITY
PÁZMÁNY PÉTER SÉTÁNY 1/C, H-1117 BUDAPEST, HUNGARY
Email address: `bizkit@cs.elte.hu`

1 Introduction

In $\mathbb{R}^n$ an infinite tube is the closed r-neighbourhood of l for some positive real r and some straight line l. The *tube-measure* of a set $E \subset \mathbb{R}^n$ is defined as

$$\mu(E) = \inf\left\{\sum_i \gamma_{n-1} r_i^{n-1} : \bigcup_i T_i \supset E\right\},$$

where T_i is a tube with cross-sectional radius r_i, and γ_{n-1} denotes the volume of the unit ball of $\mathbb{R}^{n-1}$. The set E is called *tube-null* if $\mu(E) = 0$.

Csörnyei and Wisewell showed that the only μ-measuruble sets are the tube-null sets and their complements [2]. Tube-null sets come up in Fourier analysis: Carbery, Soria and Vargas proved that every tube-null set is a "set of divergence" for the localisation problem [1]. From this point of view, it could be useful to see non-trivial examples for tube-null sets. In many cases, it is hard to tell whether a set is tube-null or not (even for *simple* sets).

In this note we prove that the Koch curve is tube-null. In the plane tubes are infinite strips and tube-nullity simply means the existence of a covering with strips of arbitrarily small total width. Actually, we will prove more than that. For some $s < 1$ we will show that K can be covered by strips such that the sum of the s-powers of the widths is arbitrarily small, and we will get such coverings by using strips in only three directions. This will give a decomposition of the Koch curve into three sets, each of which can be projected to a line in such a way that the image has Hausdorff dimension less than 1.

Theorem 1.1. *The Koch curve K is tube-null, that is, it can be covered by strips of arbitrarily small total width.*

Moreover, there exists a decomposition $K = K_0 \cup K_1 \cup K_2$ and projections π_0, π_1, π_2 such that the Hausdorff dimension of $\pi_i(K_i)$ is less than 1 for $i = 0, 1, 2$.

We mention that in a conference talk T.C. O'Neil proved that a certain variant of the Koch curve (which uses only right angles) is tube-null [4]. He also asked whether this holds for the Koch curve.

2 The proof

Let $A_0 A_1 A_2$ be an equilateral triangle with side length $2/\sqrt{3}$ so that each height of the triangle is 1. This is our level 0 triangle. Let e_i be the line that is parallel to $A_{i+1}A_{i+2}$ and goes through A_i (indices are cyclic). The strip bounded by the lines $A_{i+1}A_{i+2}$ and e_i is the level 0 strip in direction i. For some positive integer n we decompose this strip into 3^n strips with equal width 3^{-n}. These strips will be called the level n strips in direction i. The boundary lines of these strips (in all three directions) determine a triangle grid. The triangles in this grid are called level n triangles.

Figure 1: The covering numbers corresponding to level 2 strips

Let us consider the Koch curve K connecting A_1 with A_2 and contained in the triangle $A_0 A_1 A_2$. It is a self-similar set: it is the union of 4^n pieces, each similar to K. Each of these level n pieces is contained in one of the level n triangle of the grid and connects two vertices of that triangle.

Our goal is to find a collection of level n strips such that they cover K and they have a small total width. For a level n strip we define its *covering number* as the number of level n pieces covered by the strip (see Figure 1). The idea is to use strips with large covering number. The next lemma shows that each piece is covered by at least one strip with a large covering number.

Lemma 2.1. *For each level n piece (at least) one of the three level n strips through this piece contains at least $2^{n/3}$ level n pieces.*

Proof. For an arbitrary level n piece take all three level n strips covering this piece. It is sufficient to prove that the product of the covering numbers of these strips is at least 2^n. We will prove this by induction on n. It clearly holds for $n = 0$. For arbitrary $n \geq 1$, a level n piece can be viewed as a level $(n-1)$ piece in one of the four level 1 pieces. Due to the reflection symmetry of K the level n strip in direction 0 covers at least twice as many level n pieces in the whole curve as it covers in any of the level 1 pieces. For the other two directions, we simply use the fact that the strips cover at least as many pieces in K as in a level 1 piece. It follows that the product is at least the double of the product corresponding to the same piece when it is considered as a level $(n-1)$ piece of a level 1 piece, which completes the proof. □

Now take all level n strips that contain at least $2^{n/3}$ pieces. The lemma yields that these strips cover K. Our goal is to prove that the number of such strips is *very small* (compared to 3^n). Since the width of a level n strip is 3^{-n}, this would imply that the total width is also *very small*.

For a given strip we distinguish two different ways it can cover a piece. A piece connects two points lying on the border lines of the strip. If these endpoints lie on the same border line then we say that it is a *border piece*. If, on the other hand, its endpoints are on different border lines, then it is a *crossing piece*. Note that a piece can have different types when covered by different strips. In fact, for each level n piece out of the three level n strips covering the piece, two cover it as a crossing piece and one covers it as a border piece.
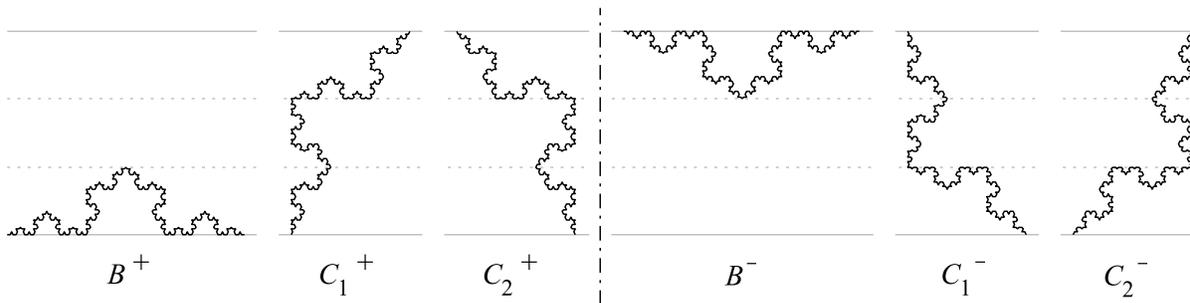
Figure 2: The different types of pieces covered by a strip

To every strip we associate a two dimensional vector called the *covering vector*, the first and second coordinate of which denotes the number of border and the number of crossing pieces in the strip, respectively. Clearly, the covering number of a strip is simply the sum

of the coordinates of the covering vector. First we show that the covering vector of a strip determines the covering vectors corresponding to the three offspring strips. (By *offspring strips* of a level n strip we mean the three level $n + 1$ strips contained in the strip.)

Proposition 2.2. *A covering vector* (v_1, v_2) *yields the following three vectors on the next level:*

$$(2v_1, 2v_1 + v_2); \quad (0, v_2); \quad (v_2, v_2).$$

In other words, to get a next level covering vector we simply right-multiply with one of the three 2×2 matrices below:

$$A = \begin{pmatrix} 2 & 2 \\ 0 & 1 \end{pmatrix}; \quad B = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}; \quad C = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}.$$

Proof. Take an arbitrary strip and the pieces covered by the strip. Theoretically, there are 6 possible types of these pieces (two types of border pieces labelled with B^+ and B^- in Figure 2 and four types of crossing pieces labelled with C_1^+, C_2^+, C_1^- and C_2^-.) However, the truth is that each strip has an *orientation* and depending on this orientation either all the pieces (covered by the strip) are of types B^+, C_1^+, C_2^+ or all of them are of types B^-, C_1^-, C_2^-. This can be proved by induction on the level of the strip: using Figure 2 the reader can easily check that the middle offspring strip always changes orientation while the other two offspring strips have the same orientation as the original strip. Now the statement of the proposition is immediate. $\square$

Now we fix a direction (0, 1 or 2) and take the level 0 strip in this direction. The covering vector $\mathbf{v}$ associated to this strip is either $(1, 0)$ or $(0, 1)$ depending on the direction. The covering vectors of level n strips in the fixed direction can be obtained in the following way. We take the product of n matrices, each matrix being A, B or C and right-multiply $\mathbf{v}$ with this product matrix. If we do this for all possible 3^n products, then we get the covering vectors of all 3^n level n strips in the fixed direction.

So we need to compute such matrix products. It is not that complicated due to the following relations between A,B and C:

$$BA = B; \quad BB = B; \quad BC = C; \quad CC = C. \tag{1}$$

So there are a lot of cancellations in such a product: a matrix B cancels all the subsequent A's and B's until a C comes which cancels B. (For example, $BAABAC = BC = C$.) Also, if there are more than one successive C's, then we can write only one C instead. After all possible cancellations have been done we get a product of the following form:

$$(C)A^{k_1}CA^{k_2}C\cdots CA^{k_r}(B \text{ or } C).$$

By induction, we get that

$$A^k = \begin{pmatrix} 2^k & 2^{k+1} - 2 \\ 0 & 1 \end{pmatrix}, \text{ so } CA^k = \begin{pmatrix} 0 & 0 \\ 2^k & 2^{k+1} - 1 \end{pmatrix}.$$

Now it is easy to see that the sum of the elements in the product matrix is at most

$$L \cdot 2^{(k_1+1)+(k_2+1)+\cdots+(k_r+1)} \leq 2^{c_0 + reduced\_length},$$

4

where L, c_0 are absolute constants and *reduced_length* denotes the length of the product after the cancellations.

The covering number of a strip is the sum of the elements in the covering vector which is bounded above by the sum of the elements in the corresponding product matrix that has been shown to be at most $2^{c_0+reduced\_length}$. So we have proved that

$$covering\_number \le 2^{c_0+reduced\_length}. \tag{2}$$

Now we forget for a moment that A, B, C denote matrices. We just take a random sequence of letters A, B, C, choosing every letter independently and with uniform distribution. We do all the cancellations implied by the relations in (1). The reduced length of the sequence is defined as the number of letters that *survive* cancellation. The next lemma claims that the reduced length of a random sequence of length n is less than $n/3 - c_0$ with high probability.

Lemma 2.3. *There exists a constant $a < 1$ such that*

$P(\text{the reduced length of a random sequence of length } n \text{ is at least } n/3 - c_0) < a^n.$

Before proving this lemma, we first show how it can be used to complete the proof of Theorem 1.1.

Proof of Theorem 1.1. Let $\mathcal{S}_i^n$ be the set of level n strips in direction i with covering number at least $2^{n/3}$, and put $\mathcal{S}^n = \mathcal{S}_0^n \cup \mathcal{S}_1^n \cup \mathcal{S}_2^n$. On one hand, Lemma 2.1 yields that $\mathcal{S}^n$ is a covering of K. On the other hand, (2) and Lemma 2.3 entail that a random level n strip is in $\mathcal{S}^n$ with probability less than a^n for some constant $a < 1$. Thus $|\mathcal{S}^n| < 3a^n 3^n$. It follows that K is tube-null for $\mathcal{S}^n$ has total width at most $3a^n$.

To obtain the decomposition claimed in the theorem we define the set K_i^n as the set of those points in K which are covered by at least one strip in $\mathcal{S}_i^n$. Since $\mathcal{S}^n$ is a covering of K, $K = K_0^n \cup K_1^n \cup K_2^n$. Set

$$K_i := \{x : x \in K_i^n \text{ for infinitely many values of } n\}.$$

Clearly, $K = K_0 \cup K_1 \cup K_2$. By definition, K_i is covered by $\mathcal{S}_i^m \cup \mathcal{S}_i^{m+1} \cup \ldots$ for any positive integer m. Let π_i be the projection in direction i. Then $\pi_i(K_i)$ is covered by $\pi_i(\cup \mathcal{S}_i^m) \cup \pi_i(\cup \mathcal{S}_i^{m+1}) \cup \ldots$ where $\pi_i(\cup \mathcal{S}_i^n)$ is the union of at most $(3a)^n$ segments of length 3^{-n}. It easily follows that $\pi_i(K_i)$ has Hausdorff dimension at most $s = \log_3(3a) < 1$. $\square$

Proof of Lemma 2.3. First we give a heuristic proof. A typical sequence contains about $n/3$ of each letter. About half of the A's survive (depending on whether the first preceding non-A letter is B or C), basically no B's survive and about one third of the C's survive (depending on whether the next letter is A or not). Thus the reduced length of a typical sequence is about $n/3(1/2 + 0 + 1/3) = 5n/18$. In the sequel we make these heuristics precise.

First we compute the expected value of the reduced length of a random sequence of length n. Consider the letter in position k. We will determine the probability that this letter survives cancellation. Clearly, the sum of these probabilities is the expected value in question. However, for these probabilities to be well defined we need to agree on which letter is cancelled in case of two successive B's or C's. When we have two successive

5

B's, let the first B survive and the second one be cancelled. On the other hand, for two successive C's let the first be cancelled and the second survive. (In other words, B's have a *forward-mouth* and they *eat* A's and other B's, while C's have a *backward-mouth* eating B's and other C's.) Now it is a well-defined question whether a letter survives or not. Let the random sequence be $M_1 M_2 \cdots M_n$.

Case $M_k = A$: with probability $1/3^{k-1}$ it holds that for each $i \leq k-1$ $M_i = A$ when M_k survives. If it is not so, then there is an index $i < k$ for which $M_i \neq A$ but $M_{i+1} = M_{i+2} = \cdots = M_k = A$. If $M_i = B$, then M_i cancels all the subsequent A's so it cancels M_k. If $M_i = C$, then M_k survives. The probability of this is clearly $(1 - 1/3^{k-1})/2$. Consequently:

$$P(M_k \text{ survives}|M_k = A) = \frac{1}{2} + \frac{1}{2 \cdot 3^{k-1}}.$$

Case $M_k = B$: it survives only if M_i equals A or B for each $i \geq k+1$ (and even in this case it might be cancelled due to a preceding B):

$$P(M_k \text{ survives}|M_k = B) \leq \frac{1}{3}\left(\frac{2}{3}\right)^{n-k}.$$

Case $M_k = C$: if $M_{k+1} = A$, then M_k survives; if $M_{k+1} = C$, then M_k is cancelled. If $M_{k+1} = B$, then M_k survives if and only if M_{k+1} survives which holds if and only if M_i equals A or B for each $i \geq k+2$. Thus

$$P(M_k \text{ survives}|M_k = C) = \frac{1}{3} + \frac{1}{3}\left(\frac{2}{3}\right)^{n-k-1} \quad (1 \leq k \leq n-1),$$

$$P(M_n \text{ survives}|M_n = C) = 1.$$

It follows that

$$P(M_k \text{ survives}) \leq \frac{5}{18} + \frac{1}{2 \cdot 3^k} + \frac{1}{9}\left(\frac{2}{3}\right)^{n-k} + \frac{1}{9}\left(\frac{2}{3}\right)^{n-k-1} \quad (1 \leq k \leq n-1).$$

When we add up these terms, the sum of the geometric progressions will be bounded so there exists an absolute constant c_1 such that

$$E_n := E(\text{reduced length of a random sequence of length } n) \leq \frac{5}{18}n + c_1.$$

Let $0 < \varepsilon < 1/36$ and let us fix n_0 in such a way that $E_{n_0} < (1/3 - 2\varepsilon)n_0$. Now let $n = kn_0$ for some positive integer k. We take a random sequence of length n and split it up into subsequences of length n_0. Let X_j be the random variable defined as the reduced length of the j-th subsequence ($j = 1, 2, \ldots, k$), and let X be the reduced length of the whole sequence. Clearly, $X \leq X_1 + \cdots + X_k$. The X_j's are independent random variables with $E(X_j) = E_{n_0}$ and $X_j \in (0, n_0]$. We know that under these conditions the sum $X_1 + \cdots + X_k$ is highly concentrated around its expectation which is $kE_{n_0} < (1/3 - 2\varepsilon)n$. For example, we can use Hoeffding's inequality [3] (since $X_1, \ldots, X_k$ are independent and bounded). For sufficiently large k it holds that $c_0 < \varepsilon n$, thus

$$P\left(X \geq \frac{n}{3} - c_0\right) \leq P\left(X > \left(\frac{1}{3} - \varepsilon\right)n\right) \leq P\left(\sum_{j=1}^{k}(X_j - E_{n_0}) > \varepsilon n_0 k\right) <$$

$$\exp\left(-\frac{2\varepsilon^2 n_0^2 k^2}{kn_0^2}\right) = a^n$$

for some constant $a < 1$. This already proves the lemma for n's that are sufficiently large multiples of n_0. However, with a larger $a < 1$ the lemma clearly holds for arbitrary positive integer n. □

References

[1] A. Carbery, F. Soria, A. Vargas, *Localisation and weighted inequalities for spherical Fourier means*, Journal d'Analyse Mathématique **103** (2007) no. 1, 133–156.

[2] M. Csörnyei, L. Wisewell, *Tube-Measurability*, Real Anal. Exchange **33** (2007) no. 1, 247–252.

[3] W. Hoeffding, *Probability inequalities for sums of bounded random variables*, Journal of the American Statistical Association **58** (1963) no. 301, 13–30.

[4] T.C. O'Neil, *Some geometric problems concerning sets and measures in Euclidean space*, Workshop on Harmonic Analysis and Geometric Measure Theory in memory of Laura Wisewell, University of Glasgow, 2008 December.