TOPOLOGY AND SEPARATION OF SELF-SIMILAR FRACTALS IN THE PLANE

CHRISTOPH BANDT AND HUI RAO

Abstract. Even though the open set condition (OSC) is generally accepted as the right condition to control overlaps of self-similar sets, it seems not clear how it relates to the actual size of the overlap. For connected self-similar sets in the plane, we prove that finite overlap implies OSC. On the other hand, there are Cantor sets with arbitrary small dimension which do not fulfil OSC.

MSC classification: Primary 28A80, Secondary 37F20
short title: self-similar fractals in the plane

Christoph Bandt
Institute for Mathematics and Informatics
Arndt University
17487 Greifswald, Germany
bandt@uni-greifswald.de

Hui Rao
Department of Mathematics
Tsinghua University
P.O. Box 100084, Beijing, China
HRao@math.tsinghua.edu.cn

1. Introduction

Given contracting similarity maps $f_1, ..., f_m$ on $\mathbb{R}^d$, the corresponding self-similar set is the unique compact set $A \neq \emptyset$ which satisfies the set equation

$$A = f_1(A) \cup ... \cup f_m(A) \ .$$

A consists of similar copies $A_i = f_i(A)$ of itself, each A_i consists of smaller copies $A_{ij} = f_i(f_j(A))$, and so on. For any integer n, we can consider the set S^n of words $\mathbf{i} = i_1...i_n$ from the alphabet $S = \{1, ..., m\}$. Writing $f_\mathbf{i} = f_{i_1} \cdots f_{i_n}$ and $A_\mathbf{i} = f_\mathbf{i}(A)$, we have $A = \bigcup\{A_\mathbf{i} \mid \mathbf{i} \in S^n\}$. When n tends to infinity, this induces a continuous map $\pi : S^\infty \to A$ from the set S^∞ of sequences $\mathbf{s} = s_1s_2s_3...$ onto the self-similar set, the so-called address map. See [10, 7, 6, 3] for more details.

When the contraction factors r_i of f_i are small, the pieces A_i are disjoint, π is a homeomorphism and A a Cantor set. For large r_i, however, π identifies many addresses, and the overlaps $A_i \cup A_j$ are usually too large to analyse A mathematically. In between there is the "just-touching case" [6] where overlaps are nonempty but sufficiently thin. It is defined by four equivalent conditions:

1. Moran's open set condition (OSC) [12]: there exists a nonempty open set $V \subset \mathbb{R}^n$ with $\bigcup_{i=1}^m f_i(V) \subseteq V$ and $f_i(V) \cap f_j(V) = \emptyset$ for $i \neq j$.
2. Positivity of α-dimensional Hausdorff measure: $\mu^\alpha(A) > 0$, where α denotes the similarity dimension given by $\sum r_i^\alpha = 1$ [12, 15].
3. The finite clustering property [15]: there exists an integer N such that for every piece $A_\mathbf{i}$ of A, with diameter ε, say, there are at most N incomparable pieces $A_\mathbf{j}$ of diameter $\geq \varepsilon$ with distance $< \varepsilon$ from $A_\mathbf{i}$. We call $A_\mathbf{j}$ and $A_\mathbf{k}$ incomparable if $\mathbf{j}$ is not a prefix of $\mathbf{k}$, and $\mathbf{k}$ not a prefix of $\mathbf{j}$.
4. The neighbor map condition [4]: the identity map id is not an accumulation point of the set of neighbor maps of A. A neighbor map has the form $h = f_\mathbf{i}^{-1} f_\mathbf{j}$ where $\mathbf{i}, \mathbf{j} \in S^* = \bigcup_{n\geq 1} S^n$ and $i_1 \neq j_1$. Convergence of similarity maps on $\mathbb{R}^d$ is given by the norm $\|g\| = \sup_{|x|\leq 1} |g(x)|$.

With so many equivalent formulations, OSC has become generally accepted as the adequate separation condition for self-similar fractals. But does it really say that each overlap $A_i \cap A_j$ is small? At first glance, yes. If OSC holds, the overlap is contained in $\overline{V_i} \cap \overline{V_j}$, so it has no interior points, and $\mu^\alpha(A_i \cap A_j) = 0$. We do not know, however, whether the Hausdorff dimension of overlaps must be smaller than α, except for the case of finite type [13].
And what about the converse? The finite clustering condition implies that the cardinality of $\pi^{-1}(x)$ for $x \in A$ is uniformly bounded by some number N. We do not know whether this property is equivalent to OSC. Here we consider a more special case.

Problem. *Will OSC be true when all overlaps are finite sets?*

We start with some evidence for a negative answer. A sequence $s_1s_2...$ is called recurrent if for each $K \geq 1$ there is a $n \geq 1$ with $s_1...s_K = s_{n+1}...s_{n+K}$. We show that any identification of a recurrent address will destroy OSC. However, it is not clear whether such an identification implies other identifications of addresses.

Theorem 1. *a) In a self-similar set A, if one point $a \in A_{s_1}$ with recurrent address* $\mathbf{s}$ *belongs to a piece A_{t_1} with $t_1 \neq s_1$, then OSC cannot hold.*
b) There are self-similar Cantor sets A in $\mathbb{R}$ or $\mathbb{R}^2$, with arbitrary small Hausdorff dimension and with $A_{\mathbf{i}} \neq A_{\mathbf{j}}$ for $\mathbf{i}, \mathbf{j} \in S^*$ with $\mathbf{i} \neq \mathbf{j}$, which do not fulfil the OSC.

Our main result is an affirmative answer to the problem for $d = 2$.

Theorem 2. *Let A be a connected self-similar set in the plane. If $A_i \cap A_j$ is a finite set for $i \neq j$, then OSC holds.*

In section 3 we deal with the general case and in section 4 with the case that A is homeomorphic to an interval. The proof uses plane topology at some key places. We expect that Theorem 2 is not true in higher dimensions.

2. Recurrent addresses as obstacles to OSC

Proof of Theorem 1a. According to the finite clustering property, a self-similar set A cannot fulfil the OSC if for every $N \in \mathbb{N}$ there is a piece $A_{\mathbf{i}}$ which intersects at least N other pieces $A_{\mathbf{j}}$, and no piece is a subpiece of another one, with diam $A_{\mathbf{j}} \geq$ diam $A_{\mathbf{i}}$.

Let N be given, and let $a \in A_{s_1} \cap A_{t_1}, s_1 \neq t_1$ be a point with recurrent address $\mathbf{s} = s_1 s_2 s_3 ...$ and a second address $\mathbf{t}$. There is an initial word $\mathbf{i} = i_1 ... i_n = s_1 ... s_n$ of $\mathbf{s}$ which has N suffixes which coincide with prefixes of $\mathbf{s}$ (see below). In other words, there are $\ell_1 < \ell_2 < ... < \ell_N < n$ such that $i_{\ell_k+1} ... i_n = s_1 ... s_{n-\ell_k}$ for $k = 1, ..., N$. Now we define

$$\mathbf{j}_k = i_1 ... i_{\ell_k} t_1 ... t_{n_k} \qquad \text{for } k = 1, ..., N$$

where n_k is chosen as large as possible so that still diam $A_{\mathbf{j}_k} \geq$ diam $A_{\mathbf{i}}$. (If the factors r_i are very different, then to guarantee $n_k \geq 1$ exists we have to choose $n - \ell_N$ so large that $r_{s_1} \cdot ... \cdot r_{s_{n-\ell_N}} < r_{t_1}$.)

Now $f_{i_1 ... i_{\ell_k}}(a)$ is in $A_{\mathbf{i}} \cap A_{\mathbf{j}_k}$. So $A_{\mathbf{i}}$ intersects N pieces $A_{\mathbf{j}_k}$ of at least the same size, and they are incomparable: for $k < k'$ the $(\ell_k + 1)$-st coordinate of $\mathbf{j}_k$ is t_1 and the $(\ell_k + 1)$-st coordinate of $\mathbf{j}_{k'}$ is $i_{\ell_k+1} = s_1$. OSC does not hold. □

Examples of recurrent sequences. $\mathbf{s}$ is a recurrent sequence if arbitrarily long prefixes $s_1 ... s_K$ will occur inside the sequence. An example is the Cantor sequence

$$\mathbf{s} = 2121112121111111112121112 12... \qquad (1)$$

obtained as limit of the words $s^{(n)}$ where $s^{(0)} = 2$ and $s^{(n+1)} = s^{(n)} 1^{3^n} s^{(n)}$ for $n \geq 0$. Another example is given by taking $s^{(0)} = 2$ and $s^{(n+1)} = s^{(n)} 1^n s^{(n)}$:

$$\mathbf{s} = 2121121211121211212...$$

A third example is the prominent Fibonacci sequence generated by the substitution $1 \mapsto 2, 2 \mapsto 21$:

$$\mathbf{s} = 21221212212212...$$

If $\mathbf{s}$ is recurrent, then for each $N \geq 1$ there is an index k_N such that the word $\mathbf{i}(N) = s_1 ... s_{k_N}$ has N different suffixes which coincide with prefixes of $\mathbf{s}$. The k_N are constructed by induction: let $k_1 = 1$, and let k_2 be the smallest number for which

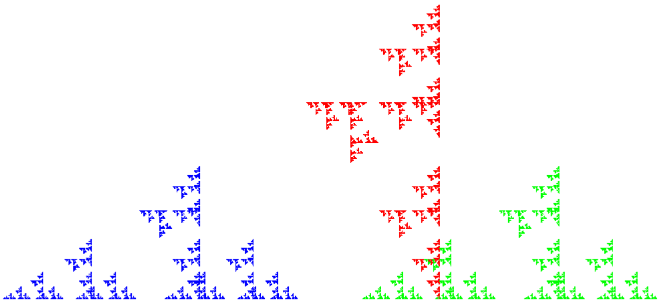

FIGURE 1. Self-similar Cantor set without OSC.

$s_{k_2} = s_1$. Let k_3 denote the end point of the first repetition of the word $s_1 ... s_{k_2}$ inside $\mathbf{s}$, etc. Our example sequences all end with k_4.

Proof of Theorem 1b. We take similitudes in the complex plane with equal factor r, that is, $f_j(z) = a_j rz + b_j, j \in S$, with $a_j, b_j \in \mathbb{C}$ and $|a_j| = 1$. Then the address $\mathbf{t} = t_0 t_1 t_2 ... \in S^\infty$ is mapped to the point

$$\pi(\mathbf{t}) = b_{t_0} + \sum_{k=1}^{\infty} r^k b_{t_k} \prod_{\ell=0}^{k-1} a_{t_\ell} \ . \tag{2}$$

Here we start with t_0 since this gives a power series in r. We also write $\mathbf{s} = s_0 s_1 ...$ To prove the equation, start with $f_{t_0 t_1}(z) = b_{t_0} + r a_{t_0} b_{t_1} + r^2 a_{t_0} a_{t_1} z$ and continue by induction. Note that $r^n z$ tends to zero for $n \to \infty$. We apply the formula to some very simple mappings:

$$f_1(z) = rz \ , \quad f_2(z) = rz + 1 \ , \quad f_3(z) = \omega rz + c$$

where c, ω are complex numbers, $|\omega| = 1$. Suppose we want to identify the points corresponding to the Cantor sequence (1) $\mathbf{s} = 212111212...$ and to $3\overline{1} = 3111...$ Formula (2) with $a_1 = a_2 = 1, a_3 = \omega$ and $b_1 = 0, b_2 = 1, b_3 = c$ gives $\pi(\mathbf{s}) = 1 + r^2 + r^6 + r^8 + ... = \sum_{s_k=2} r^k$ and $\pi(3\overline{1}) = c$ since $\pi(\overline{1}) = 0$. Thus the condition $\pi(3\overline{1}) = \pi(\mathbf{s})$ holds if and only if

$$c = 1 + r^2 + r^6 + r^8 + ... = \sum_{s_k=2} r^k = \prod_{k=0}^{\infty} (1 + r^{2 \cdot 3^k}) \ .$$

Given $r \in]0,1[$, we obtain the corresponding c, and ω can be chosen on the unit circle. For the self-similar set A in Figure 1, we took $\omega = i, r = 0.45$ and got $c \approx 1.2125$. (For $r < \frac{1}{3}$ which we assume below the picture would be hardly visible.)

Let us verify the properties of Theorem 1b for the above example. The Hausdorff dimension of A is not larger than the similarity dimension α determined by the equation $\sum r_j^\alpha = 3r^\alpha = 1$ [7, 10]. In particular, A is a Cantor set whenever $r < 1/3$ since then $\alpha < 1$ and any compact connected set has Hausdorff dimension at least 1 ([8], 2.10.12). Moreover, taking r small enough we get $\alpha = \ln 3 / |\ln r|$ as small as we want. For $\omega = 1$ we have $A \subset \mathbb{R}$. All this holds for arbitrary small r.

Finally, for $\omega = 1$ we show that $A_{\mathbf{i}} = A_{\mathbf{j}}$ with $\mathbf{i} \neq \mathbf{j}$ can hold only for countably many r. It suffices to show this for fixed words $\mathbf{i}, \mathbf{j}$ of the same length n. Now $A_{\mathbf{i}} = A_{\mathbf{j}}$ means $f_{\mathbf{i}} = f_{\mathbf{j}}$ and $f_{\mathbf{i}}(z) = r^n z + \sum_{k=0}^{n-1} b_{n-k} r^k$ where $b_k \in \{0, 1, c\}$. Thus $f_{\mathbf{i}} = f_{\mathbf{j}}$ leads to an equation of the form $p(r) + q(r)c(r) = 0$ where p, q are polynomials with coefficients -1,0,1 and $c(r)$ is the above power series. However, an analytic function has only finitely many zeros in $[0, 1]$. □

Remark 3. *It is easy to check that the set of recurrent addresses in S^∞ has full measure with respect to every product measure $\{p_1, ..., p_m\}^\infty$ with $p_i > 0$ and $\sum p_j = 1$. From this viewpoint, almost every "random identification of two points" will destroy the OSC. This throws some light on results by several authors* [11, 14] *on families of self-similar sets on the line similar to the above example although it does not prove these results.*

3. Proof of Theorem 2 when A is not a Jordan curve

Let $B_r(x)$ denote the ball around x with radius r. Take a number $R \geq 1$ with

$$\bigcup_{i=1}^{m} f_i(B_R(0)) \subseteq B_R(0) \quad \text{and define} \quad \|g\| = \sup_{|x| \leq R} |g(x)| \quad \text{for} \quad g : \mathbb{R}^d \mapsto \mathbb{R}^d.$$

Lemma 4. *If g and f are similitudes and $f(B_R(0)) \subseteq B_R(0)$, then*

$$\|f^{-1} g f - id\| < c_f \|g - id\|,$$

where c_f is a constant depending only on f.

Proof. Let $f^{-1}(x) = Gx + b$ where G is linear, and $c_f = \|G\|$. For $x \in B_R(0)$, we have

$$\begin{aligned} |f^{-1} g f(x) - x| &= |f^{-1} g f(x) - f^{-1} f(x)| \\ = \quad |G(g(f(x)) - f(x))| &\leq \|G\| \cdot |g(f(x)) - f(x)| \quad \leq \quad \|G\| \cdot \|g - id\|. \end{aligned}$$ □

Lemma 5. *Let A be a self-similar set which is not a singleton. For any integer $M > 0$ there exists k_0 and $\mathbf{j}_1, ..., \mathbf{j}_M \in S^{k_0}$ such that the $A_{\mathbf{j}_k}$ are all disjoint.*

Proof. Two maps, say f_1 and f_2, have different fixed points x_1, x_2. So all $f_1^k(x_2)$ with $k = 1, 2, ..., M$ are different. Take $\mathbf{j}_k = 1^k 2^{(k_0 - k)}$ for sufficiently large k_0. □

A simple Jordan curve is the image set of a homeomorphism from $[0, 1]$ to $\mathbb{R}^2$.

Lemma 6. *Let A be a connected self-similar set which is not a simple Jordan curve. Then there exist four points $a, b, c, q \in A$ such that*

(i) there exist Jordan curves $\widehat{qa}, \widehat{qb}, \widehat{qc} \in A$;

(ii) $\widehat{qa}, \widehat{qb}, \widehat{qc}$ intersect each other only at the point q;

(iii) $|qa| = |qb| = |qc| = r'$;

(iv) $\widehat{qa} \setminus \{a\}, \widehat{qb} \setminus \{b\}$ and $\widehat{qc} \setminus \{c\}$ are contained in the interior D° of the closed disk D with center q and radius r'.

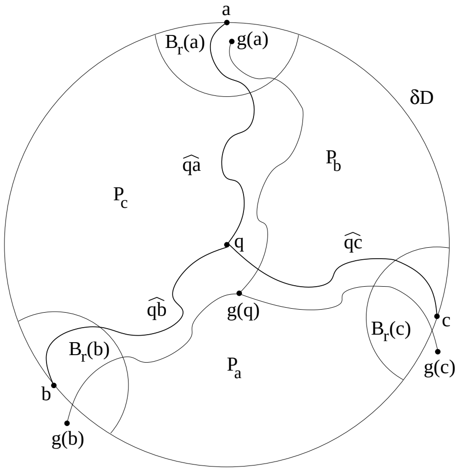

FIGURE 2. Proof of Lemma 7. Drawing by M. Mesing

Proof. A connected self-similar set is arcwise connected [9, 6, 5]. If A is not a Jordan curve, there is a point q and a neighborhood U of q such that $A \cap U \setminus \{q\}$ has at least three components which have q in their closure. Taking a', b', c' in those components, we get disjoint arcs $\widehat{qa'}, \widehat{qb'}, \widehat{qc'}$. Let r' be smaller than the distances of a', b', c' to q, and let D be the closed disk with center q and radius r'. Starting in q, let a, b, c denote the first points where the arcs hit the boundary ∂D. $\square$

We denote by K the union of the three Jordan curves, that is, $K = \widehat{qa} \cup \widehat{qb} \cup \widehat{qc}$.

Lemma 7. *(Perturbation lemma) There is a constant* $\delta > 0$ *such that* $K \cap g(K) \neq \emptyset$ *for every similitude* g *with* $\|g - id\| < \delta$.

Proof. In Lemma 6, $D \setminus K$ is divided into three parts P_a, P_b and P_c, where the closure of P_x does not contain x. Choose r so small that $B_{2r}(a)$ does not intersect P_a, $B_{2r}(b)$ does not intersect P_b and $B_{2r}(c)$ does not intersect P_c.

Let σ be the distance between $K \setminus (B_r(a) \cup B_r(b) \cup B_r(c))$ and the circle ∂D. Clearly $\sigma > 0$. We set

$$\delta = \min\{r'/2,\, r,\, \sigma/2\}$$

and show that $\|g - id\| < \delta$ implies $g(K) \cap K \neq \emptyset$.

Since $\|g - id\| < r'$, the point $g(q)$ is in D°. If $g(q)$ belongs to K, our assertion is true. So we assume, without loss of generality, that $g(q)$ belongs to P_a.

Since $\|g - id\| \leq r$, the point $g(a)$ is in $B_r(a)$, and hence not in P_a. Now we show that $g(\widehat{qa})$ *does not intersect* $\widetilde{bc}$, the arc from b to c on the circle ∂D. The part $g(\widehat{qa} \cap B_r(a))$ is still in $B_{2r}(a)$, so it will not intersect $\widetilde{bc}$, and $g(\widehat{qa} \setminus B_r(a))$ is contained in the disk with center q and radius $r' - \sigma/2$, and will not intersect ∂D at all. Now $g(\widehat{qa})$ must intersect $\widehat{qb} \cup \widehat{qc}$, and $g(K)$ must intersect K. $\square$

Proof of Theorem 2 when A is not a Jordan curve. Let

$$M = \max\{\mathrm{card}\,(A_i \cap A_j)|\; i \neq j\} + 1,$$

k_0 the constant in Lemma 5, and $\mathbf{j}_1, \ldots, \mathbf{j}_M \in S^{k_0}$ the words with disjoint $A_{\mathbf{j}_k}$. Let

$$c = \max\{c_{f_{\mathbf{j}}} |\ \mathbf{j} \in S^{k_0}\}, \tag{3}$$

where c_f is the constant of Lemma 4.

Suppose the f_i do not satisfy OSC. By the neighbor map condition, there are $\mathbf{i}, \mathbf{i}' \in S^*$ with $i_1 \neq i_1'$ and

$$\|f_{\mathbf{i}'}^{-1} f_{\mathbf{i}} - id\| < \delta / c$$

where δ is the constant in Lemma 7 and c is from (3). Then by Lemma 4, we have

$$\|f_{\mathbf{i}'\mathbf{j}}^{-1} f_{\mathbf{i}\mathbf{j}} - id\| = \|f_{\mathbf{j}}^{-1} f_{\mathbf{i}'}^{-1} f_{\mathbf{i}} f_{\mathbf{j}} - id\| < c \|f_{\mathbf{i}'}^{-1} f_{\mathbf{i}} - id\| < \delta$$

for each $\mathbf{j} \in S^{k_0}$. Hence $f_{\mathbf{i}'\mathbf{j}}(A) \cap f_{\mathbf{i}\mathbf{j}}(A) \neq \emptyset$ by Lemma 7. Let $p_{\mathbf{j}}$ be a point in the intersection. Then $p_{\mathbf{j}_1}, \ldots, p_{\mathbf{j}_M}$ are all different, and they belong to the set $f_{i_1'}(A) \cap f_{i_1}(A)$ which contradicts the definition of M. So the f_i satisfy OSC. ☐

4. Finite type and self-similar Jordan curves

Definition 8. *A self-similar set A is of* finite type *if there are only finitely many neighbor maps $h = f_{\mathbf{i}}^{-1} f_{\mathbf{j}}$ with $A \cap h(A) \neq \emptyset$ (or equivalently $A_{\mathbf{i}} \cap A_{\mathbf{j}} \neq \emptyset$), and with similarity factor $r_h \in (r_*, 1/r_*)$ where $r_* = \min\{r_1, ..., r_m\}$.*

The meaning of the last condition is that we accept only neighbors $h(A)$ which fit the size of A, otherwise we take their pieces or supersets as $h(A)$. If all r_i are equal, we take only neighbors $h(A)$ which have the same size as A, and the maps h are isometries.

Compared with the finite type concept in Ngai and Wang [13], this definition is a bit more restrictive, but simpler and in our opinion more natural. The following was proved for equal factors in [2], section 4.

Theorem 9. *A self-similar set of finite type fulfils OSC if $f_{\mathbf{i}} \neq f_{\mathbf{j}}$ for $\mathbf{i} \neq \mathbf{j}$.*

Proof. Let $H_0 = \{h = f_{\mathbf{i}}^{-1} f_{\mathbf{j}} \,|\, i_1 \neq j_1,\ A \cap h(A) \neq \emptyset,\ r_h \in (r_*, \frac{1}{r_*})\}$, let $\tilde{H} = \{f_i^{-1} h f_j,\ f_i^{-1} h,\ h f_j,\ f_i^{-1} f_j \mid h \in H_0 \cup \{id\},\ i, j \in S\}$ denote the "immediate successors" of maps of $H_0 \cup \{id\}$, and let

$$H_1 = \{\tilde{h} \in \tilde{H} \,|\, \tilde{h}(A) \cap A = \emptyset\} .$$

Since for $h \in H_1$ the compact sets A and $h(A)$ are disjoint, their distance $d_h = d(A, h(A)) = \inf\{|x - y| \,|x \in A, y \in h(A)\}$ is positive. Moreover, $\|h - id\| \geq \max\{|h(x) - x| \,|x \in A\} \geq d_h$. If A is of finite type, H_0 and hence H_1 are finite. Thus $\delta_0 = \min\{\|h - id\| \mid h \in H_0\} > 0$ since $f_{\mathbf{i}} \neq f_{\mathbf{j}}$ implies $id \notin H_0$, and $\delta_1 = \min\{d_h \,|\, h \in H_1\} > 0$.

Any neighbor map $g \notin H_0 \cup H_1$ has the form $f_{\mathbf{i}}^{-1} h f_{\mathbf{j}}, f_{\mathbf{i}}^{-1} h$, or $h f_{\mathbf{j}}$ with suitable $h \in H_0 \cup H_1$ and words $\mathbf{i}, \mathbf{j} \in S^*$. Then $d_g \geq d_h$ since $d(B', C) \geq d(B, C)$ for $B \supseteq B'$:

$$d(f_{\mathbf{i}}^{-1} h(A), A) \geq d(f_{\mathbf{i}}^{-1} h(A), f_{\mathbf{i}}^{-1}(A)) = \frac{d_h}{r_{\mathbf{i}}}\ ,\ d(h f_{\mathbf{j}}(A), A) = d(h(A_{\mathbf{j}}), A) \geq d_h$$

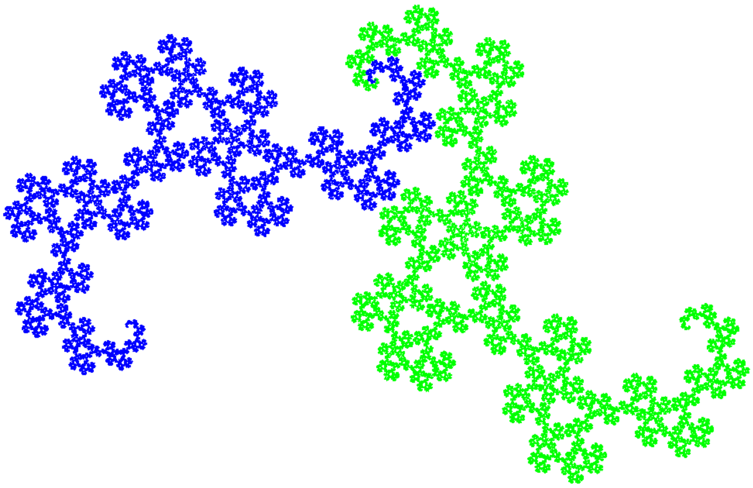

FIGURE 3. Self-similar Jordan arc of type (2) with two pieces.

(for $f_{\mathbf{i}}^{-1} h f_{\mathbf{j}}$ we combine both estimates). This implies $\|g - id\| \geq \min\{\delta_0, \delta_1\}$. This is true for *all* neighbor maps g. So id cannot be an accumulation point, OSC holds. □

Now we consider self-similar sets homeomorphic to [0,1], and write $J = f_1(J) \cup ... \cup f_m(J)$ instead of A. If the pieces have finite intersection, each intersection is at most one point, and we can assume that $J_i \cap J_{i+1} = \{c_i\}$ for $i = 1, ..., m-1$. Furthermore, let $c_0 \in J_1$ and $c_m \in J_m$ be the two endpoints of J (those points x for which $J \setminus \{x\}$ is connected). Concerning the addresses of c_0, c_m four cases are possible (cf. [9, 5]):

(1) $f_1(c_0) = c_0$ and $f_m(c_m) = c_m$, i.e. $c_0 = \pi(\overline{1}), c_m = \pi(\overline{m})$.
(2) $f_1(c_0) = c_0$ and $f_m(c_0) = c_m$, i.e. $c_0 = \pi(\overline{1})$ and $c_m = \pi(m\overline{1})$.
(3) $f_1(c_m) = c_0$ and $f_m(c_m) = c_m$, i.e. $c_0 = \pi(1\overline{m}), c_m = \pi(\overline{m})$.
(4) $f_1(c_m) = c_0$ and $f_m(c_0) = c_m$, i.e. $c_0 = \pi(\overline{1m}), c_m = \pi(\overline{m1})$.

Theorem 10. *A self-similar Jordan curve in the plane is of finite type unless*
(i) it has endpoint type (1), *and*
(ii) there exists an $i \in \{1, 2, ..., m-1\}$ *such that* $f_i^{-1}(c_i) \neq f_{i+1}^{-1}(c_i)$, *and*
(iii) $\frac{\log r_m}{\log r_1}$ *is irrational for the contraction factors* r_1, r_m *of* f_1, f_m.

Proof. We need only check neighbor maps $h = f_{\mathbf{i}}^{-1} f_{\mathbf{j}}$ for pieces $J_{\mathbf{i}} \subseteq J_i$ and $J_{\mathbf{j}} \subseteq J_{i+1}$ of approximately the same size which intersect in the point c_i ($i = 1, ..., m-1$). We show that neighbor maps at c_i repeat periodically when we go to smaller pieces.

If $f_i^{-1}(c_i) = f_{i+1}^{-1}(c_i)$, or if we have endpoint type (2)(cf. Figure 3), (3) or (4), then both addresses of c_i are eventually periodic with the same periodic part:

$$c_i = \pi(iu\overline{w}) = \pi((i+1)v\overline{w})$$

where u, v can be 1 or m or the empty word (in which case $f_u = id$), and $w \in \{1, m, 1m, m1\}$ where $\overline{w}$ is the address of an endpoint of J. This endpoint which we call 0 is the fixed point of f_w. What is more important, it is also the fixed point of all neighbor maps

$$h = f_{\mathbf{i}}^{-1} f_{\mathbf{j}} \qquad \text{with} \quad \mathbf{i} = iuw^n \quad \text{and} \quad \mathbf{j} = (i+1)vw^{n'} .$$

since our assumption was $c_i = f_i f_u(0) = f_{i+1} f_v(0)$. In other words, when (i) or (ii) is not true then *the neighbor maps are rotations around one endpoint of J, composed with a stretching and/or a reflection.* The condition $r_* < r_h < \frac{1}{r_*}$ for $r_h = \frac{r_{\mathbf{j}}}{r_{\mathbf{i}}} = \frac{r_{i+1} r_v}{r_i r_u} \cdot r_w^{n'-n}$ says that only finitely many differences $n' - n$ are possible.

Now let us first assume that all similitudes f_i are orientation-preserving, $f_i(z) = a_i z + b_i$. We choose the origin of our coordinate system at the fixed point 0 of f_w so that $f_w(z) = az$ for some $a \in \mathbb{C}$ with $|a| = r_w < 1$. Then

$$h = f_{\mathbf{i}}^{-1} f_{\mathbf{j}} = f_w^{-n} f_u^{-1} f_i^{-1} f_{i+1} f_v f_w^{n'} = a^{n'-n} f_u^{-1} f_i^{-1} f_{i+1} f_v \tag{4}$$

because orientation-preserving similitudes in the plane with common fixed point commute. So the neighbor map h depends only on $n' - n$, not on n or n' separately. Consequently, the number of neighbor maps at c_i is finite. (Actually, for case (4) with $w = 1m$, we still have to consider $\mathbf{i} = iuw^n 1$ and/or $\mathbf{j} = (i+1)vw^{n'} 1$ which increases the number of neighbor maps at most by a factor 4.) We have verified finite type if either (i) or (ii) is not fulfilled.

What will change if we admit orientation-reversing similitudes, $f_i(z) = a_i \overline{z} + b_i$? If $f_w(z) = a\overline{z}$ we can work with $w^2 = ww$ instead of w in the above calculations. Since $f_{ww}(z) = f_w(f_w(z)) = |a|^2 z$, equation (4) again holds. The vertex c_1 in Figure 4 shows such a case (Type (1) but $c_1 = f_1(c_4) = f_2(c_4)$).

Commutativity could only fail if $f_w(z) = az$ with $a \notin \mathbb{R}$ and at the same time $f_u^{-1} f_i^{-1} f_{i+1} f_v$ is orientation-reversing. In that case one of the maps $f_i f_u$, $f_{i+1} f_v$ would preserve and the other one reverse orientation. However, this case is not possible for a Jordan curve: by assumption, the fixed point 0 of f_w is an endpoint of J (namely c_0 for $w = 1$ or $w = 1m$, and c_m for $w = m$ or $w = m1$). If z_0 denotes the other endpoint of J, the curve consists of Jordan arcs connecting points $z_0, az_0, a^2 z_0, \ldots$ In other words, the curve approaches 0 as a fractal spiral. The mappings $f_i f_u$ and $f_{i+1} f_v$ map this spiral to two spirals with center c_i which represent J_i and J_{i+1}. If these spirals have different orientation, they have plenty of intersection points, contradicting the Jordan curve structure of J.

To finish the proof, we assume both (i) and (ii), and show finite type if (iii) is not true - that is, $r_1^k = r_m^{k'}$ for positive integers k, k'. By (i) and (ii), c_i has addresses $i\overline{m}$ and $(i+1)\overline{1}$ (or $i\overline{1}$, $(i+1)\overline{m}$). Let us take the origin of our coordinate system at c_i. In the orientation-preserving case $f_1(z) - c_0 = a_1(z - c_0)$, $f_m(z) - c_m = a_m(z - c_m)$ the sets J_i and J_{i+1} form two fractal spirals approaching $c_i = 0$, and these spirals are mapped into themselves by multiplication with a_m and a_1, respectively. (Seen from the center c_m, the fractal spiral J connects $c_0, z_1, z_2, \ldots$ where $z_k - c_m = a_m^k(c_0 - c_m)$. Now $f_i(z) = \alpha z + \beta$ maps J to J_i with c_m to 0, c_0 to c_{i-1} and hence z_k to $y_k = a_m^k c_{i-1}$. Similar for a_1 and f_{i+1}. If f_i, f_{i+1} are orientation-reversing the factor is $\overline{a_m}, \overline{a_1}$.) By Lemma 12 below, $a_m = a_1^t$ with $t = k/k'$. Using the maps f_1^k and $f_m^{k'}$ instead of

f_1, f_m we can use the above argument to show finite type at c_i (instead of factor 4 for $w = 1m$ we have factor kk'). The case that f_1 or f_m is orientation reversing leads to real factors as above. The proof is complete. □

Remark 11. *In the proof of Theorem 2 below, we shall see that we have infinite type if (i), (ii) and (iii) hold. We should mention that one-dimensional Cantor sets of this kind where studied by Ayer and Strichartz, and their infinite type had the consequence that there is no point with maximum density with respect to Hausdorff measure (*[1]*, Theorem 3.2).*
We should also mention that Theorem 10 does not hold in dimension 3.

Lemma 12. *If* $J_1, J_2 \subset \mathbb{C}$ *are Jordan curves with* $J_1 \cap J_2 = \{0\}$ *and* a_1, a_2 *complex numbers with* $a_1 J_1 \subset J_1$ *and* $a_2 J_2 \subset J_2$ *then there is* $t > 0$ *with* $a_2 = a_1^t$.

Proof. Let $a_1 = r \cdot e^{i\alpha_1}$, $a_2 = r^t \cdot e^{it\alpha_2}$. We must show $\alpha_1 = \alpha_2$, so let us assume $\alpha_2 - \alpha_1 = \varepsilon \in (0, 2\pi)$. There are points $z_1, z_2 \neq 0$ on J_1, J_2 with $|z_1| = |z_2|$. We take z_1, z_2 as endpoints of J_1, J_2 (forgetting points with larger modulus), and we can assume that no other points of $J_1 \cup J_2$ has modulus $|z_2|$. Express the curve J_1 between z_1 and $a_1 z_1$ in parametric form $\varphi(s) = r(s)e^{i\alpha(s)}$, $0 \le s \le 1$ where $\alpha(s)$ is continuous (does not jump from 2π to 0). Let $\beta = \max\{\alpha(s) - \alpha(0) \,|\, 0 \le s \le 1\}$ and choose n so large that $nt\varepsilon > 4\pi + \beta$.

Let $z_2' = z_2 a_2^n$ and let z_1' be the first point of J_1 (starting from z_1) which has modulus equal $|z_2'| = |z_2| \cdot r^{nt}$. Then z_1' lies between $z_1 a_1^k$ and $z_1 a_1^{k+1}$ where k is the integer part of nt. Now we parametrize the two Jordan curves $[z_i, z_i'] \subset J_i$ between the circles $|z| = |z_2|$ and $|z| = |z_2'|$ by $\varphi_i(s) = r_i(s)e^{i\gamma_i(s)}$, $0 \le s \le 1$ where $\gamma_i(s)$ is continuous and $\gamma_2(0) \in (-2\pi, 0]$, $\gamma_1(0) \in [0, 2\pi)$. Then $\gamma_2(0) < \gamma_1(0)$ but

$$\gamma_2(1) = \gamma_2(0) + nt\alpha_2 \ge \gamma_2(0) + nt\varepsilon + k\alpha_1 > \gamma_2(0) + 4\pi + \beta + k\alpha_1 > \gamma_1(0) + \beta + k\alpha_1 \ge \gamma_1(1).$$

This proves that J_1 and J_2 have an intersection point with modulus $\ge |z_2'|$ which contradicts the assumption. □

Proof of Theorem 2 for Jordan curves. We show that each self-similar Jordan curve J fulfils OSC, studying neighbor maps at each point c_i. By Theorem 10, we can assume (i),(ii) and (iii) because otherwise J is of finite type and hence OSC by Theorem 9. Also if J is contained in a line, OSC is obvious.

As in the end of the proof of Theorem 10, c_i has addresses $i\overline{m}$ and $(i+1)\overline{1}$. We assume that f_1, f_m preserve orientation so that J_i and J_{i+1} form spirals around c_i which remain invariant under similitudes with factors a_m, a_1 and center c_i. Lemma 12 says $a_m = a_1^t$ but this time t is irrational.

We show that the neighbor maps $h = f_{\mathbf{i}}^{-1} f_{\mathbf{j}}$ with $i_1 = i$ and $j_1 = i+1$ cannot approach id. We first study arbitrary small pieces $J_{\mathbf{i}}, J_{\mathbf{j}}$ intersecting at c_i, that is, $\mathbf{i} = im^n$, $\mathbf{j} = (i+1)1^{n'}$. Using $r_m = r_1^t$, we see that we have infinite type here:

$$r_h = \frac{r_{i+1}}{r_i} \cdot r_1^{n' - nt} \, .$$

Since for irrational t the set $\{n' - nt \,|\, n, n' \in \mathbb{N}\}$ is dense in $\mathbb{R}$, the factors r_h include a dense set in $(0, \infty)$.

We apply $f_{\mathbf{i}}^{-1}$ in each case so that each $J_{\mathbf{i}}$ is mapped onto J, and each $J_{\mathbf{j}}$ is transformed into some $h(J)$. The intersection point is $c_m = f_{\mathbf{i}}^{-1}(c_i) = h(c_0)$ for every n and n'. Let $c_m = 0$ be our origin now. While J is kept fixed, the neighbors $h(J)$ are obtained from $h_0(J) = f_i^{-1} f_{i+1}(J)$ by multiplication with $a_1^{n'-nt}$. We define $U = \bigcup\{h(J) \,|\, h = f_{im^n}^{-1} f_{(i+1)1^{n'}}\}$ as union of all these neighbors. $h(J) \cap J = \{0\}$ for each h implies $U \cap J = \{0\}$.

Now let us consider the trajectories of the flow $s \mapsto z_0 a_1^s$ within the unit circle C. For each $z_\alpha = e^{i\alpha} \in C$ we have the spiral

$$S_\alpha = \{z_\alpha a_1^s \,|\, s > 0\}$$

Since J is compact and connected, and the spirals do not form self-similar sets (consider curvature), the set of all z_α for which $S_\alpha \cap J \neq \emptyset$ will be an interval $[z_\beta, z_\gamma]$ on C, or C itself.

The same is true for the neighbor set $h_0(J)$. Since it is a scaled and rotated copy of J, it intersects as many spirals S_α as J. More precisely, the set of all z_α for which $S_\alpha \cap J \neq \emptyset$ will be an interval $[z_{\beta'}, z_{\gamma'}]$ of the same length as $[z_\beta, z_\gamma]$, or C.

The other neighbors $h(J)$ determine exactly the same interval, since they are obtained from $h_0(J)$ by multiplication with $a_1^{n'-nt}$ which leaves each S_α invariant. Thus also U determines the same interval. However, U is a dense union of Jordan curves and will occupy a dense set on each S_α which it intersects. From this fact it will follow that $[z_\beta, z_\gamma]$ and $[z_{\beta'}, z_{\gamma'}]$ are proper intervals which can intersect only in their endpoints. That is, *there are at most two spirals S_α which intersect both J and U.*

To prove this, we assume the contrary: there are two spirals S_1, S_2 which, together with all spirals between them, belong to the interior of both $[z_\beta, z_\gamma]$ and $[z_{\beta'}, z_{\gamma'}]$. Thus J intersects S_1, S_2 in z_1, z_2 and joins them with a Jordan arc $\widehat{z_1, z_2}$, and $h_0(J)$ intersects S_1, S_2 in y_1, y_2, say, and contains a Jordan arc $\widehat{y_1, y_2}$ between them. This arc as well as all its multiples $a_1^s \cdot \widehat{y_1, y_2}$, where s is taken from a dense set of positive numbers, must not intersect J. This is only possible if $\{y_1, y_2\} = b \cdot \{z_1, z_2\}$ for some $b = a_1^u \in \mathbb{C}$. It follows that $\widehat{y_1, y_2} = b \cdot \widehat{z_1, z_2}$, and that the arcs J and $h_0(J)$ continue to intersect the spirals in a parallel manner, because as soon as one of the arcs would turn back, a multiplication of $h_0(J)$ by a_1^s with s very near to 0 would result in an intersection point. On the other hand, since J contains a spiral point, it contains a dense set of spiral points, and can never intersect the spirals in successive order, it must turn back which is a contradiction.

We proved that J intersects S_α for z_α in a proper interval $[z_\beta, z_\gamma]$, and U can only intersect the two boundary spirals S_β and S_γ. These two spirals will be used to separate J and U although they may contain points of both sets.

Now it is easy to give a uniform estimate $\|h - id\| \geq \eta > 0$ for all these neighbor maps h, just using a point $z \in J$ which lies on the spiral $S_{(\beta+\gamma)/2}$ which has distance $> \eta$ from U. However, since we have no finite type, we must study also neighbor maps between *disjoint* pieces of J_i and J_{i+1}. To get the uniform estimate for all these, we need a more general argument.

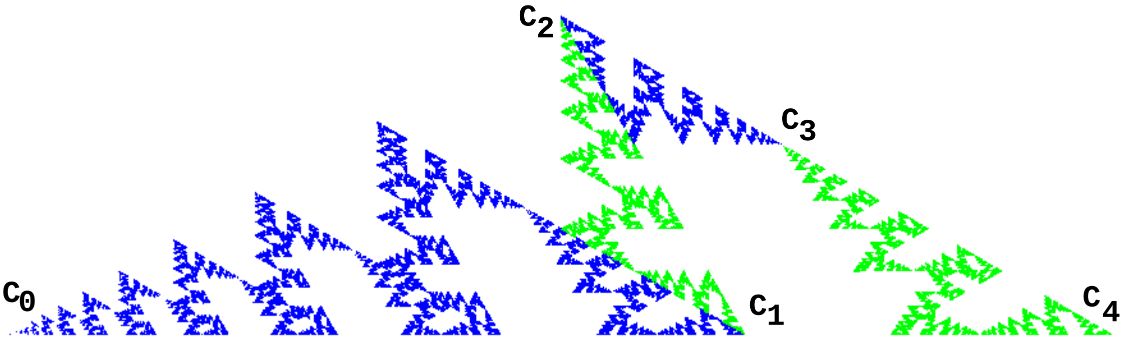

FIGURE 4. A self-similar arc with four pieces. At c_1 and c_3 we have finite type, at c_2 infinite type.

We consider one of the spirals, say S_0, on all its length from 0 to ∞. Let $\delta > 0$ be taken so that

$$B_{2\delta}(c_0) \cap J \subseteq J_1 \quad \text{and} \quad B_{2\delta}(c_m) \cap J \subseteq J_m. \tag{5}$$

A topological argument below says there is $\varepsilon > 0$ such that *for any isometry g of the plane there exists a point $x_g \in J$ with*

$$d(x_g, g(S_0 \setminus B_\delta(0))) = \inf\{d(x_g, y) \,|\, y \in g(S_0 \setminus B_\delta(0))\} \geq \varepsilon\,. \tag{6}$$

This implies that a map h which transforms J onto some $h(J)$ on the other side of a spiral isometric to S_0, outside the δ-neighborhood of its center, must fulfil $\|h - id\| \geq \varepsilon$.

Let us set $\eta = \min\{\delta, r_*\varepsilon\}$ and let us consider the neighbor map $h = f_{\mathbf{i}}^{-1} f_{\mathbf{j}}$ with intersecting $J \cap h(J) = \{c_m\}$ as above: $\mathbf{i} = im^n$, $\mathbf{j} = (i+1)1^{n'}$. By (5) J_1 does not intersect $B_\delta(c_m)$. By (6) there is $x^* \in J_1$ with $d(x^*, h(J) \setminus B_\delta(c_m)) \geq r_*\varepsilon$ (note that $a_1^{-1} J_1$ is isometric with J). By (5) $d(x^*, B_\delta(c_m)) \geq \delta$. Thus $\|h - id\| \geq \eta$.

Now we take disjoint pieces $J_{\mathbf{i}} \subset J_i$ and $J_{\mathbf{j}} \subset J_{i+1}$. If $J_{\mathbf{i}} = J_{im^n u}$ with $u_1 \neq m$ and we standardize by this new $f_{\mathbf{i}}^{-1}$ the separating spiral will be $f_u^{-1}(S_\beta)$ which is isometric to S_0, and the distance of its center to J is larger than $2\delta/r_u$. Again we get $\|h - id\| \geq \eta$.

We get all neighbor maps of J when we add the inverses of those considered (interchanging $\mathbf{i}$ and $\mathbf{j}$) and finitely many others. So id is not approached by neighbor maps, OSC holds.

At the end, we give the topological argument for (6). In the space $\mathcal{F}$ of closed subsets of the closed ball $\overline{B_R(0)}$ with Hausdorff metric [6, 10] we consider the subspace $\mathcal{F}_0$ of all isometric copies of compact subsets of $S_0 \setminus B_\delta(0)$ or of $\mathbb{R}$ which have approximately the same diameter as J.

$$\mathcal{F}_0 = \{g(F) \subseteq \overline{B_R(0)} \,|\, g \text{ isometry, } F \subset \mathbb{R} \text{ or } F \subset S_0 \setminus B_\delta(0),\ \frac{\operatorname{diam} J}{\operatorname{diam} F} \in (r_*.\frac{1}{r_*})\}$$

It is well-known that $\mathcal{F}$ is compact. Moreover, $\mathcal{F}_0$ is a closed subset and so is also compact (Take a sequence $g_n(F_n) \to G$. If $F_n \to \infty$ then G is a subset of a line. If infinitely many F are within a bounded part of S_0 there is a subsequence for which both F_n and g_n converge.) Thus $J \in \mathcal{F}$ has positive Hausdorff distance ε from $\mathcal{F}_0$. (J cannot be in a spiral by a simple curvature argument, and J was assumed not to be in a line, so $J \notin \mathcal{F}_0$.) Take $F \in \mathcal{F}_0$. Since $d_H(J, G) \geq \varepsilon$ for every closed $G \subset F$, there is $x^* \in J$ with $d(x^*, F) \geq \varepsilon$. □

Remark 13. *The infinite type case appears at vertex* c_2 *in Figure 4. Here we have orientation-reversing maps* f_2, f_4*. However,* f_1 *is a homothety so by taking* f_4^2 *we have orientation-preserving maps with real factor so that spirals in the above proof are just rays. Although* $J_2 \cap J_3 = \{c_2\}$*, there is a limit map* g *of neighbor maps at* c_2 *such that* J *and* $g(J)$ *do intersect in an infinite set. Since* $c_1 = f_1(c_4) = f_2(c_4)$ *and* $c_3 = f_3(c_0) = f_4(c_0)$ *we have finite type at these points.*

Acknowledgement. Most of this work was done when Hui Rao visited the University of Greifswald, supported by Deutsche Forschungsgemeinschaft. We express our gratitude also to both referees who gave a number of helpful suggestions.

References

[1] E. Ayer and R.S. Strichartz, Exact Hausdorff measure and intervals of maximum density for Cantor sets, *Trans. Amer. Math. Soc.* **351** (1999), 3725-3741

[2] C. Bandt, Self-similar measures, *Ergodic Theory, Analysis and Efficient Simulation of Dynamical Systems* (ed. B. Fiedler), Springer 2001, 31-46

[3] C. Bandt, N.V. Hung and H. Rao, On the open set condition for self-similar fractals, *Proc. Amer. Math. Soc.* **134** (2005), 1369–1374.

[4] C. Bandt and S. Graf, Self-similar sets VII. A characterization of self-similar fractals with positive Hausdorff measure. *Proc. Amer. Math. Soc.* **114** (1992), 995–1001.

[5] C. Bandt and K. Keller, Self-similar sets 2. A simple approach to the topological structure of fractals, Math. Nachrichten 154 (1991), 27-39

[6] M.F.Barnsley, Fractals Everywhere, 2nd ed., Academic Press 1993

[7] K.J. Falconer, *Fractal Geometry,* Wiley 1990

[8] H. Federer, *Geometric Measure Theory,* Springer 1969

[9] M. Hata, *On the structure of self-similar sets*, Japan J. Appl. Math. **2** (1985), 281–414.

[10] J.E. Hutchinson, Fractals and self-similarity, *Indiana Univ. Math. J.* **30** (1981), 713–747.

[11] M. Keane, M. Smorodinsky and B. Solomyak, On the morphology of γ-expansions with deleted digits, *Trans. Amer. Math. Soc.* **347** (1995), 955–966.

[12] P.A.P. Moran, Additive functions of intervals and Hausdorff measure. *Proc. Cambridge Philos. Soc.* **42**, (1946) 15–23.

[13] S.M. Ngai and Y. Wang, Hausdorff dimension of overlapping self-similar sets. *J. London Math. Soc.* **63** (2001), 655–672.

[14] M. Pollicott and K. Simon, The Hausdorff dimension of λ-expansions with deleted digits, *Trans. Amer. Math. Soc.* **347** (1995), 967–983.

[15] A. Schief, Separation properties for self-similar sets. *Proc. Amer. Math. Soc.* **122** (1994), 111–115.