A CHARACTERIZATION OF SELF-SIMILAR FRACTALS WITH POSITIVE HAUSDORFF MEASURE

CHRISTOPH BANDT AND SIEGFRIED GRAF

Abstract. For self-similar sets with nonoverlapping pieces, Hausdorff dimension and measure are easily determined. We express “absence of overlap” in terms of discontinuous action of a family of similitudes, thus improving the usual “open set condition”.

1. Definitions and result

Among mathematical fractals, self-similar sets with nonoverlapping pieces seem to be the most tractable. This note will help to clarify what “overlap” means. Let \( f_1, \ldots, f_m \) be contracting similitudes on Euclidean \( \mathbb{R}^d: |f_i(x) - f_i(y)| = r_i \cdot |x - y|, \) where \( 0 < r_i < 1 \). Then there is a unique compact set \( A \neq \emptyset \) with

\[
A = f_1(A) \cup \cdots \cup f_m(A).
\]

\( A \) is called the self-similar set with respect to \( f_1, \ldots, f_m \) [5, 3, 4]. The similarity dimension of \( A \) is the number \( \alpha \) for which

\[
1 = r_1^\alpha + \cdots + r_m^\alpha.
\]

Since the \( \alpha \)-dimensional Hausdorff measure \([3, 4]\) clearly fulfills \( \mu^\alpha(A) \leq (\text{diam } A)^\alpha < \infty \), the Hausdorff dimension \( \text{dim } A \) is \( \leq \alpha \). Let us say \( A \) has positive measure if \( \mu^\alpha(A) > 0 \). This implies \( \text{dim } A = \alpha \). If the pieces \( A_i = f_i(A) \) are pairwise disjoint, it is not hard to show that \( A \) has positive measure. On the other hand, if the \( A_i \) overlap so strongly that for instance \( f_1f_2(A) = f_2f_1(A) \), we can represent \( A \) as a self-similar set with respect to \( m^2 - 1 \) mappings \( f_if_j \) and thus find an \( \alpha' < \alpha \) with \( \text{dim } A \leq \alpha' \).

Moran [7] and Hutchinson [6] gave a criterion that guarantees that there is not too much overlap. The open set condition (OSC) says that there is an open set \( V \neq \emptyset \) with \( f_i(V) \subseteq V \) and \( f_i(V) \cap f_j(V) = \emptyset \) for \( i, j \in \{1, \ldots, m\}, i \neq j \). Let us say that the strong OSC holds if there is such a \( V \) with \( V \cap A \neq \emptyset \).

The OSC implies that \( A \) has positive measure [7, 6, 4] but it is not known whether the converse is true. It is also not clear whether the OSC implies the strong OSC. Moreover, there is no method to check the OSC when the \( f_i \) or
even $A$ are given (except for simple examples like the Sierpinski gasket). In fact, the set $V$ can be quite exotic, and there need not exist a convex or even simply connected $V$ [1]. We were not able to prove the OSC for all self-similar sets with two pieces that intersect in a single point. That is why we looked for other, more handy conditions.

Our basic idea was that the OSC means that a certain group of isometries is acting discontinuously and that the set $V$ is a fundamental domain of that group. The situation is a bit more intricate, as we shall see. Let us introduce some notation. Let $S = \{1, \ldots, m\}$, and let $s = s_1 s_2 \cdots s_p$ and $t = t_1 \cdots t_q$ be two words from $S^* = \bigcup \{S^n \mid n = 0, 1, 2, \ldots\}$. The length of $s$ is $|s| = p$. We write $s \subseteq t$ if $s$ is an initial word of $t$, that is, $p \leq q$ and $s_k = t_k$ for $k = 1, \ldots, p$. $s$ and $t$ are incomparable if neither $s \subseteq t$ nor $t \subseteq s$; in other words, $s_k \neq t_k$ for some $k \leq \min\{p, q\}$. Let $r_s = r_{s_1} \cdot r_{s_2} \cdots \cdots r_{s_p}$, $f_s = f_{s_1} \cdot f_{s_2} \cdot \cdots \cdots f_{s_p}$ and $A_s = f_s(A)$.

The map $g = f_t \cdot f_s^{-1}$ maps $A_s$ onto $A_t$. If $s$ and $t$ are incomparable, we want to express the fact that "$A_s$ and $A_t$ do not overlap" by saying that "$g$ is far from id". However, $g$ may be near to id just due to the small size of $A_s$ and $A_t$. For that reason we "renormalize $g" and take $h = f_s^{-1} \cdot f_t$, which maps $A$ onto $h(A)$. There is a commutative diagram $g \cdot f_s = f_s \cdot h$. We consider

$$F = \{f_s^{-1} \cdot f_t \mid s, t \in S^*, \ s, t \text{ incomparable}\}$$

as a subset of the topological group $G$ of all similitudes on $\mathbb{R}^d$. On $G$, the topology of uniform convergence on bounded sets agrees with the topology of pointwise convergence on an arbitrary subset of $\mathbb{R}^d$ that is not contained in a hyperplane. Thus given any $d + 1$ points $x_0, \ldots, x_d$ in $\mathbb{R}^d$ in general position, the sets

$$W_\epsilon = \{h \in G \mid |h(x_j) - x_j| < \epsilon, \ j = 0, \ldots, d\}, \quad \epsilon > 0$$

form a neighbourhood base of id in $G$. (Note that $h$ is determined by the values $h(x_j)$.)

**Theorem.** Let $f_1, \ldots, f_m$ be contracting similitudes, $A$ the corresponding self-similar set, and $F$ the associated family of similitudes. $A$ has positive measure iff $id \notin \text{cl}(F)$.

### 2. Remarks and examples

(1) For every $c$ between 0 and 1, the family of all similitudes with factor $c < \leq c$ or $\geq 1/c$ is closed in $G$. Thus if $F_c$ denotes all similitudes from $F$ with factor in $[c, 1/c]$, id belongs to $\text{cl}(F)$ iff id belongs to $\text{cl}(F_c)$. In particular, if $r_1 = \cdots = r_m = r$, it suffices to study $F_1$, the set of all isometries in $F$, instead of $F$. In this case, $f_s^{-1} f_t \in F_1$ iff $|s| = |t|$. (2) The Sierpinski gasket with vertices $y_1 = (0, 0)$, $y_2 = (1, 0)$, and $y_3 = (0, 1)$ is generated by $f_i(x) = (x + y_i)/2$, $i = 1, 2, 3$. We have

$$f_{t_1} \cdots f_{t_q}(x) = 2^{-q} x + 2^{-q} y_{t_q} + 2^{-q-1} y_{t_q-1} + \cdots + 2^{-1} y_{t_1},$$

$$f_{s_p}^{-1} \cdots f_{s_1}^{-1}(x) = 2^p x - 2^p y_{s_1} - 2^{p-1} y_{s_2} - \cdots - y_{s_p}, \quad \text{and for } p = q$$

$$f_{s_p}^{-1} f_{t_1}(x) = x + 2^{-p}(y_{t_1} - y_{s_1}) + 2(y_{t_p} - y_{s_p})$$
is a translation by a vector $z = (n, n')$, $n, n' \in \mathbb{Z}$, with $(n, n') \neq (0, 0)$ for $s \neq t$. Thus $\text{id} \notin \text{cl}(F_1)$. Moreover, each $(n, n')$ will appear for some $s, t$, so that $F_1 \cup \{\text{id}\}$ is a discrete group of translations.

(3) A more general situation is treated in [1, Theorem 1]: Let $g$ be an expansive similitude on $\mathbb{R}^d$ that transforms the lattice $L$ of integer linear combinations of certain basis vectors $b_1, \ldots, b_d$ into itself, let $y_1, \ldots, y_m$ be members of $m$ different cosets of $g(L)$ in $L$, and let $f_i(x) = g^{-1}(x) + y_i$, $i = 1, \ldots, m$. The calculation above gives for $p = |s| = |t|

$$f_s^{-1}f_t(x) = x + gp(y_i) - y_s + gp^{-1}(y_i) - \cdots - g(y_p) - y_s,$$

which is a translation by a vector $z \in L$, $z \neq 0$ for $s \neq t$. Thus $\text{id} \notin \text{cl}(F_1)$, $A$ has positive measure. More generally,

**Corollary.** If $r_1 = \cdots = r_m = r$, $f_s \neq f_t$ for $s \neq t$, and the group of isometries generated by the $f_s^{-1}f_t$, $|s| = |t|$, is discrete, then $\mu^d(A) > 0$.

(4) Our corollary can be seen as a generalization of the results in [1] on fractal tilings of $\mathbb{R}^n$. (Strictly speaking, we do only prove that $A$ has positive area. To show that $a$ has interior points one has to use the methods of [1].) Figure 1 shows a “terdragon” that is not obtained by [1, Theorem 2]. The mappings in the complex plane are $f_1(z) = -\eta z$, $f_2(z) = -\eta z + 9 - i\sqrt{3}$, $f_3(z) = \eta z - 1 - 3i\sqrt{3}$ with $\eta = \frac{1}{2} + \frac{i}{2}\sqrt{3}$. It is curious that the self-similar set generated by $f_1$ and $f_2$ is just the well-known von Koch curve.

(5) Unfortunately, $\text{id} \notin \text{cl}(F)$ does not always mean that the group generated by the isometries of $F$ is discrete. For an example in $\mathbb{R}$ let $f_1(x) = x/4$, $f_2(x) = (x + 3)/4$, $f_3(x) = (x + c)/4$. Then $A$ is a Cantor set with $\text{dim}A$
\[ h(x) = x + \sum \{ 4^k a_k \mid k = 0, \ldots, n \} \]
with \( n \in \mathbb{N}, a_k \in \{0, \pm 3, \pm c, \pm (3 - c)\}, a_n \neq 0. \]

For \( 1 \leq c \leq 2 \), the OSC will hold with \( V = ]0, 1[ \), and it is also easily seen that the translation vectors of \( h \) are smallest for \( n = 0 \). However, if \( c \) is irrational, \( F_1 \) will not be a group and the group of isometries generated by \( F_1 \) will not be discrete.

It is easy to determine rational \( c \) for which there are \( s \neq t \) with \( f_s = f_t \), by solving equations like \( 4c - 3 = 0 \). Obviously, \( h = \text{id} \) is only possible for rational \( c \). One can construct irrational \( c \) for which the translations \( h \) come arbitrary close to \( \text{id} \), for example \( c = 3/4 + 3/4^2 + 3/4^4 + 3/4^8 + \cdots \). We do not know \( \dim A \) for irrational \( c \), cf. remark (7).

We cannot decide whether our condition \( \text{id} \notin \text{cl}(F) \) implies the OSC, even for \( \alpha = d \). Here we compare the conditions:

**Proposition 1.** Let \( f_1, \ldots, f_m \) be contracting similitudes on \( \mathbb{R}^d \).

(i) The OSC holds iff there is \( x \in \mathbb{R}^d \) and \( \varepsilon > 0 \) such that for all incomparable \( s, t \in S^* \)
\[ |f_s^{-1}f_t(x) - x| \geq \varepsilon. \]

(ii) The strong OSC holds iff there is such an \( x \) in \( A \).

(iii) \( \text{id} \notin \text{cl}(F) \) holds iff there are \( x_0, \ldots, x_d \) in general position in \( \mathbb{R}^d \) and \( \varepsilon > 0 \) such that for all incomparable \( s, t \in S^* \) there is \( j \in \{0, \ldots, d\} \) with \( |f_s^{-1}f_t(x_j) - x_j| \geq \varepsilon. \)

**Proof.** (i) Suppose the OSC holds for \( V \). Then \( f_i(V) \cap f_j(V) = \emptyset \) for incomparable \( s, t \). Take an arbitrary \( x \in V \). There is an \( \varepsilon > 0 \) with \( U_\varepsilon(x) \subseteq V \). Since \( f_s(U_\varepsilon(x)) = U_\varepsilon(f_s(x)) \), the inequality holds. Conversely, if the inequality is true for some \( x, \varepsilon \) and all incomparable \( s, t \) then \( f_i(U_\delta(x)) \cap f_j(U_\delta(x)) = \emptyset \) for \( \delta = \varepsilon/3 \) and incomparable \( s, t \). The set \( V = \bigcup \{ f_s(U_\delta(x)) \mid s \in S^* \} \) is open and \( f_i(V) \subseteq V \). Moreover, \( f_i(V) \cap f_j(V) \neq \emptyset \) for \( i \neq j \) would imply \( f_if_j(U_\delta(x)) \bigcap f_jf_i(U_\delta(x)) \neq \emptyset \) for some \( s, t \in S^* \), which is not possible since the words \( is \) and \( jt \) are incomparable. Thus (i) is proved and (ii) too. (iii) is obvious: the condition just means \( F \cap W_\varepsilon = \emptyset \) for some \( \varepsilon > 0 \).

As noted in §1, assertion (iii) is independent of the choice of the points, provided the \( x_j \) are in general position. In particular, the \( x_j \) can always be chosen in \( A \). (If \( A \) is contained in an affine plane \( P \) of dimension \( k < d \), then \( f_i(P) = P \) for \( i = 1, \ldots, m \), and \( k + 1 \) points in \( P \) will suffice.)

It should be noted that in (i) one cannot define \( V = \{ x \} \) there exists \( \varepsilon > 0 \) with \( |f_s^{-1}f_t(x) - x| \geq \varepsilon \) for all incomparable \( s, t \). In general there is no largest \( V \). If we replace \( f_3 \) in Figure 1 by a similitude with the same factor and fixed point, but with opposite angle of rotation, one obtains a mirror-image of Figure 1. The interior of each dragon can be used as open set for the von Koch curve, but not their union.

(7) It would also be interesting to have a characterization of those self-similar sets for which similarity and Hausdorff dimension agree. Consider the shift space \( S^\infty = \{ i_1 i_2 \cdots \mid i_k \in \{1, \ldots, m\} \} \) and the product measure \( \nu \) on \( S^\infty \) that assigns the cylinder \( C_s = \{ i_1 i_2 \cdots i_1 \cdots i_q = s \} \) the measure \( \nu(C_s) = r_s^q \), for each \( s = s_1 \cdots s_q \) in \( S^* \). There is the projection \( p: S^\infty \to A \), \( p(i_1 i_2 \cdots) = \cap \{ A_{i_1 \cdots i_q} \mid q = 1, 2, \ldots \} \), and the measure \( \mu(B) = \nu(p^{-1}(B)) \) could be called...
the natural measure on $A$ \cite{5, 3}. We think that whenever $\mu(A_i \cap A_j) = 0$ for $i \neq j$, $\mu$ is a Hausdorff measure with respect to some $h(t) = t^\alpha \cdot \varphi(t)$, where $\varphi$ is a function like $|\log t|$ that for $t \to 0$ increases more slowly than any $t^{-\varepsilon}$. However, we are only able to prove a converse

**Proposition 2.** If $\dim A = \alpha$, then $\mu(A_i \cap A_j) = 0$ for $i \neq j$.

*Proof.* Assume $\mu(A_i \cap A_j) > 0$ for $i \neq j$. Let $D = A_i \cap A_j$, and let $B_k = \{t \in S^\infty | t_1 = k, p(t) \in D\}$ for $k = 1, \ldots, m$. Then $\nu(B_k) > 0$ for at least one $k$, say $k = i$. Clearly $p(S^\infty \setminus B_i) = A$, and the open set $S^\infty \setminus B_i$ is the disjoint union of cylinders $C_{s_n}, n = 1, 2, \ldots, s_n \in S^*$. Thus $\{r_{s_n}^a | n = 1, 2, \ldots\} = \nu(S^\infty) - \nu(B_i) < 1$. There is $\beta < \alpha$ with $\sum r_{s_n}^\beta = 1$. Now one can see that $\mu^\beta(A) < (\text{diag } A)^\beta$ since one can cover $A$ by the $A_{s_n}, n = 1, 2, \ldots$, and also by the smaller sets $A_{s_{n_1} \ldots s_{n_k}}, n_1, \ldots, n_k \in \{1, 2, \ldots\}$, for any fixed $k$.

3. **Proof of the theorem**

First we assume $\text{id} \notin \text{cl}(F)$ and prove $\mu^\alpha(A) > 0$. We can assume that $A$ is not contained in a hyperplane of $\mathbb{R}^d$. Choose $x_0, \ldots, x_d$ in $A$ in general position. There is an $\varepsilon > 0$ with $W_\varepsilon \cap F = \emptyset$. That is, for incomparable $s$ and $t$ there exists $j \in \{0, \ldots, d\}$ with $|f_s^{-1}(x_j) - x_j| > \varepsilon$, hence $|f_t(x_j) - f_s(x_j)| \geq \varepsilon$. Let $a = \text{diam } A$. Then $\text{diam } A_\varepsilon = r_\varepsilon a$.

Now we show that there is a constant $K$ such that for any set $U$, the following family has at most $K$ elements:

$$\mathcal{C} = \{s = s_1 \cdots s_p \in S^* | r_\varepsilon a < \text{diam } U \leq r_1, \ldots, r_{s_p-1}, a, A_s \cap U \neq \emptyset\}.$$  

If $s, t \in C$ are different, they are incomparable and  

$$|f_s(x_j) - f_t(x_j)| \geq r_\varepsilon a \geq \frac{r_\varepsilon}{a} \cdot \text{diam } U$$  

for some $j = j(s, t) \in \{0, \ldots, d\}$ and $r = \min\{r_1, \ldots, r_m\}$. Now suppose $\mathcal{C}'$ is a subfamily of $\mathcal{C}$ such that $j(s, t)$ is the same $j$ for all $s, t \in \mathcal{C}'$. Then for $g = \frac{r_\varepsilon}{a}$, the balls of radius $g \cdot \text{diam } U$ around the points $f_s(x_j), s \in \mathcal{C}'$, are pointwise disjoint, and their centres $f_s(x_j) \in A_s$ are contained in a fixed ball of radius $2 \cdot \text{diam } U$ with centre in $U$. Using Lebesgue measure we see that the number of the balls is $\text{card } \mathcal{C} < (2g)^d / g^d := N$ (cf. \cite[p. 173]{3}). Ramsey's theorem \cite{2} says that for any $N$ and $d$ there is an integer $K = K(N, d)$ such that whenever the edges of a complete graph with $K$ vertices are coloured with $d + 1$ colours, there will be a complete subgraph of one colour with $N$ vertices. Interpreting the words $s \in \mathcal{C}$ as vertices and $j(s, t) \in \{0, \ldots, d\}$ as colour of the edge $(s, t)$ we obtain $\text{card } \mathcal{C} \leq K$.

The rest is standard \cite{3, 4, 6}. For the natural measure $\mu$ on $A$ (cf. remark (7)) and any set $U$,

$$\mu(U) \leq \sum_{s \in \mathcal{C}} \mu(A_s) = \sum_{s \in \mathcal{C}} r_{s}^{\alpha} \leq K \cdot (\text{diam } U/a)^\alpha.$$  

Thus any covering $U_1, U_2, \ldots$ of $A$ fulfils 

$$\sum (\text{diam } U_j)^\alpha \geq (a^\alpha/K) \cdot \sum \mu(U_j) \geq a^\alpha/K$$  

since $\mu(A) = 1$. Consequently, $\mu^\alpha(A) > a^\alpha/K$.

To show the other implication of the theorem, we note some interesting properties of the Hausdorff measure $\mu^\alpha$ on $A$. 

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use
Proposition 3. Let \( f_1, \ldots, f_m \) be contracting similitudes and \( \alpha \) the similarity dimension of the associated self-similar set.

(i) For measurable \( B \subseteq A \), \( \mu^\alpha(B) \) coincides with the outer measure

\[
\mu^*(B) = \inf \left\{ \sum_{i \in I} (\text{diam } U_i)^\alpha \mid \text{open, } \bigcup_{i \in I} U_i \supseteq B \right\}.
\]

(ii) \( \mu^\alpha(f_s(A) \cap f_t(A)) = 0 \) for incomparable \( s, t \in S^* \).

Proof. For (i) we first consider \( B = A \). By the definition of Hausdorff measure [3], \( \mu^* \leq \mu^\alpha \), and it suffices to show that for each open covering \( \{U_i\} \) of \( A \) and each \( \delta > 0 \) there is another covering \( \{V_j\} \) with \( \text{diam } V_j < \delta \) for each \( j \) and \( \sum (\text{diam } V_j)^\alpha = \sum (\text{diam } U_i)^\alpha \). To this end, we choose \( n \) such that \( r_x < \delta \cdot \text{diam}(\bigcup U_i) \) for each \( s \in S^n \), and we replace each \( U_i \) with the family \( \{f_s(U_i) \mid s \in S^n\} \). The resulting sets cover each \( f_s(A) \), so they cover \( A \), and the sum remains unchanged by the definition of similarity dimension.

For measurable \( B \subseteq A \) we now have

\[
\mu^\alpha(A) = \mu^\alpha(B) + \mu^\alpha(A \setminus B) \geq \mu^*(B) + \mu^*(A \setminus B) = \mu^*(A) = \mu^\alpha(A),
\]

which implies \( \mu^\alpha(B) = \mu^*(B) \).

From

\[
\mu^\alpha \left( \bigcup_{i \in S} f_i(A) \right) = \mu^\alpha(A) = \sum_{i \in S} r_i^\alpha \cdot \mu(A) = \sum_{i \in S} \mu^\alpha(f_i(A))
\]

we conclude that \( \mu^\alpha(f_i(A) \cap f_j(A)) = 0 \) for \( i \neq j \). The same argument works for sums over \( s \in S^n \).

Let us finish the proof of the theorem. Assume that \( \text{id} \in \text{cl}(F) \) and \( \mu^\alpha(A) > 0 \). Take \( \eta \in ]1, 3/2[ \) and an open covering \( U_1, \ldots, U_n \) of \( A \) with \( \sum (\text{diam } U_i)^\alpha < \eta \cdot \mu^\alpha(A) \). Let \( U = U_1 \cup \cdots \cup U_n \) and \( \delta = \inf \{|a - x|/a \in A, x \notin U\} \). Since \( \text{id} \in \text{cl}(F) \), there are incomparable \( s, t \in S^* \) with

\[
(r_s^{-1} r_t)^\alpha \geq 2 - \eta
\]

and \( \sup \{|x - f_s^{-1} f_t(x)| \mid x \in A\} < \delta \). Thus \( f_s^{-1} f_t(A) \subseteq U \) and

\[
f_t(A) \cup f_s(A) \subseteq f_s(U).
\]

By Proposition 3(i), (ii) this implies

\[
\mu^\alpha(f_t(A)) + \mu^\alpha(f_s(A)) = \mu^\alpha(f_t(A) \cup f_s(A)) \leq \sum_{i=1}^{n} (\text{diam } f_s(U_i))^\alpha
\]

\[
= r_s^\alpha \cdot \sum (\text{diam } U_i)^\alpha < r_s^\alpha \cdot \eta \cdot \mu^\alpha(A).
\]

Consequently, \( r_s^\alpha \leq (\eta - 1) r_s^\alpha \). Together with (\(*)\) we get \( 2 - \eta \leq \eta - 1 \), which contradicts the choice of \( \eta \).

References


FACHBEREICH MATHEMATIK, ERNST-MORITZ-ARNDT-UNIVERSITÄT, O-2200 GRETZWALD, GERMANY

FAKULTÄT FÜR MATHEMATIK AND INFORMATIK, UNIVERSITÄT PASSAU, W-8390 PASSAU, GERMANY