FUNDAMENTA
MATHEMATICAE
141 (1992)

Topological spaces admitting a unique fractal structure

by

Christoph Bandt (Greifswald)
and **Teklehaimanot Retta** (Addis Ababa)

Abstract. Each homeomorphism from the n-dimensional Sierpiński gasket into itself is a similarity map with respect to the usual metrization. Moreover, the topology of this space determines a kind of Haar measure and a canonical metric. We study spaces with similar properties. It turns out that in many cases, "fractal structure" is not a metric but a topological phenomenon.

1. Introduction. The Sierpiński gasket. Fractals are commonly defined in metric terms (Hausdorff dimension, similarity maps between a space and its pieces etc.). This note will show, however, that often the fractal structure is completely determined by the underlying topology.

The typical example is the Sierpiński gasket (Fig. 1). It can be defined in $\mathbb{R}^n$ as a self-similar set $A = f_1(A) \cup \ldots \cup f_{n+1}(A)$ with respect to the mappings $f_i(x) = \frac{1}{2}(x + e_i)$, $i = 1, \ldots, n+1$, where the e_i are vertices of a regular n-simplex C^n [8, 13]. For convenience we work with barycentric coordinates, i.e. we take $\mathbb{R}^n$ as the hyperplane $\{x = (x_1, \ldots, x_{n+1}) \mid \sum x_i = 1\}$ and the e_i as coordinate unit vectors in $\mathbb{R}^{n+1}$, so that $f_i(x) = \frac{1}{2}(x_1, \ldots, x_{i-1}, 1+x_i, x_{i+1}, \ldots, x_{n+1})$. Since a self-similar set contains the points which may be approached by repeated application of the f_i [8], we easily get the following *alternative definition* for the Sierpiński gasket.

Write the coordinates of $x \in C^n$ as binary numbers $x_j = 0.s_{1j}s_{2j}\ldots$, $s_{ij} \in \{0, 1\}$. Then x is in A iff $\sum x_i = 1 = 0.111\ldots$ holds "digitwise":

$$A = \{x \in \mathbb{R}^{n+1} \mid \text{for each } m > 0, \text{ there is an } i_m \text{ with } s_{mi_m} = 1 \text{ and } s_{mj} = 0 \text{ for } j \neq i_m\}.$$

So each point of A is given by a sequence $i_1 i_2 \ldots$ There is a slight ambiguity in using binary numbers since $0.0111\ldots = 0.1$. Consequently, two sequences $ijjj\ldots$ and $jiii\ldots$ describe the same point. These points are called *critical*

Research of the second author supported by Deutscher Akademischer Austauschdienst.

points (cf. Fig. 1).

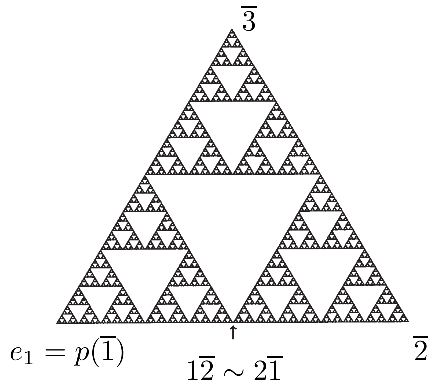

Fig. 1

For a word $w = i_1 \ldots i_k \in \{1, \ldots, n+1\}^k$, the map $f_w = f_{i_1} \ldots f_{i_k}$ is a similarity, and $A_w = f_w(A)$ is the set of all points with prescribed i_m, $m = 1, \ldots, k$. The partition $A = A_1 \cup \ldots \cup A_{n+1}$ and the resulting partitions of A into A_w, where w runs through all words of a fixed length $|w| = k$, will be called the *fractal structure* of A. Note that two pieces A_i, A_j have exactly one critical point in common and so the fractal structure is determined by the critical points.

Proposition 1.1. *The topology of A determines the fractal structure.*

The proof consists in characterizing the critical points by certain separation properties. As a consequence, we have a kind of Haar measure determined by the topology: The measure μ with $\mu(A_w) = (n+1)^{-k}$ for each word of length $|w| = k$ is the only Borel probability measure which assigns equal values to sets of the same partition. Another corollary is

Proposition 1.2. *Each homeomorphism from A into A is a similarity map with respect to the Euclidean metric.*

A metric d on a space X is called an *interior metric* if for each x, y there is a $z \neq x, y$ with $d(x,y) = d(x,z) + d(z,y)$. Since we deal with compact spaces, this implies existence of a path of length $d(x,y)$ between x and y [18]. The Euclidean metric d_{e} on A induces an interior metric $d_{\mathrm{i}}(x,y) = \inf\{d_{\mathrm{e}}\text{-lengths of paths within } A \text{ between } x \text{ and } y\}$.

Proposition 1.3. *The metric d_{i} is the unique* (*up to a constant factor*) *interior metric on A which transforms each homeomorphism into a similarity.*

The measure μ is the only Borel probability measure assigning equal values to d_{i}-isometric sets. Thus in the case of the Sierpiński gasket, the bare topology determines a canonical metrization as well as a "Haar" measure!

Numerical constants like Hausdorff dimension [8] or average distance [11, 3] become "topological invariants". Our aim is to find many spaces which share these extraordinary properties.

In the present paper, we concentrate on the fractal structure and homeomorphism group. The question of a "canonical metrization" seems more complicated [5]. Penrose [17] proved a remarkable analogue of Proposition 1.3 for fractal dendrites with two pieces, but he had to assume that the fractal structure is given.

Standard spaces like the interval, manifolds, the pseudoarc and solenoids have a lot of homeomorphisms. On the other hand, there are many examples of "rigid" topological spaces which do not admit any homeomorphisms onto their subspaces [9, 12]. We are interested in having a few homeomorphisms, but not too many—just enough to have the assertion of Proposition 1.2 satisfied. That is why we restrict our attention to a—sufficiently large—class of recursively defined spaces which have been considered by Thurston [19], Kigami [14] and others [1, 2, 10, 16, 17] in connection with Julia sets and self-similar sets.

2. Invariant factors and their cutpoints. We start with two examples. Fig. 2 shows a self-similar set with respect to $f_i(x) = \frac{1}{2}(3e_i - x)$, $i = 1, 2, 3$. Clearly, there is a homeomorphism h from Fig. 2 *onto* itself with $h(e_i) = a_i$. In contrast, Fig. 3 shows the analoguous construction for five homotheties with similarity factor $-(3-\sqrt{5})/2$, for which the above statements are true.

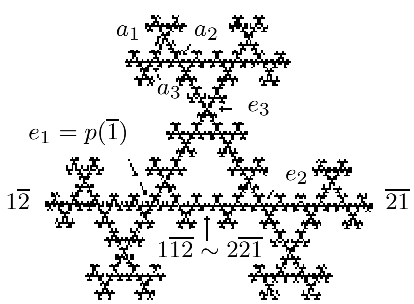

Fig. 2

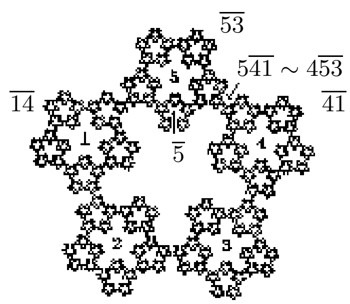

Fig. 3

Let us forget all metric features and describe such spaces topologically, using the terminology developed in [1]. Let $S = \{1, \ldots, m\}$, $C = S^\infty$ the space of sequences $s = s_1 s_2 \ldots$ with the product topology, $S^* = \bigcup_{k=0}^\infty S^k$ the set of words from S and $S^{<n} = \bigcup_{k<n} S^k$ the set of words of length smaller than n. For $w \in S^*$, the length is denoted by $|w|$, and the concatenation with $s \in S^* \cup C$ is written ws. Further, $\overline{w}$ is the periodic sequence with period w, and the initial word of length k of s is denoted by $s_{|k} = s_1 \ldots s_k$.

An equivalence relation $\sim$ on C is called *invariant*, and the quotient space $A = C/\!\sim$ is an *invariant factor* if $\sim$ is closed in $C \times C$ and for all $s, t \in C$ and $i \in S$

$$s \sim t \quad \text{if and only if} \quad is \sim it\,.$$

The set $M = \{s \in C \mid \text{there is a sequence } t \sim s \text{ with } t_1 \neq s_1\}$ completely determines the invariant factor A. We call M the *generator* of $\sim$ and $Q = M/\!\sim$ the set of *critical points.* The latter term comes from Julia sets [19, 17, 2]. We shall consider invariant factors $A = C/\!\sim$ for which M is finite and contains no periodic sequence, or, equivalently, the projection $p : C \to A$ is *finite-to-one* ([1], Th. 8). If there do not exist $s \in M$ and $w \in S^*$ such that ws also belongs to M, the relation $\sim$ and the factor A are called *simple.*

For an invariant factor $A = p(C)$ and $w \in S^*$, we define $A_w = p(C_w)$, where $C_w = \{ws \mid s \in C\}$. There is a unique homeomorphism $f_w : A \to A$ with $f_w p(s) = p(ws)$ [1]. The sequence of coverings $\{A_i \mid i \in S\}$, $\{A_{ij} \mid ij \in S^2\}, \ldots, \{A_w \mid w \in S^n\}, \ldots$ of A will be called the *fractal structure* of A.

For the Sierpiński gasket, the points of Q are $\{i\overline{j}, j\overline{i}\}$ with $i \neq j$. It seems more suggestive to describe the critical points by generating rules $i\overline{j} \sim j\overline{i}$. The rules for Fig. 1 are $i\overline{ij} \sim j\overline{ji}$, $i \neq j$. In Fig. 2, $2\overline{31} \sim 3\overline{24}$, $3\overline{42} \sim 4\overline{35}$, $4\overline{53} \sim 5\overline{41}$, $5\overline{14} \sim 1\overline{52}$, $1\overline{25} \sim 2\overline{13}$. (To check this, note that the fixed point of f_w is $x_w = p(\overline{w})$, and $f_v(x_w) = p(v\overline{w})$.)

We have defined what was called "finitely ramified fractals" by Mandelbrot and many physicists. Related mathematical papers are [14, 16, 17, 10, 1, 2]. Finite-to-one invariant factors have dimension ≤ 1. They can be approximated by undirected graphs G_n, $n = 1, 2, \ldots$ The vertex set of G_n is S^n, the edge set is $S^{<n} \times Q$. An edge can be written as vq, where v is a word with $|v| < n$, and q a critical point, i.e. an equivalence class in M. The endpoints of this edge will be the words $(vs)_{|n}$ with $s \in q$. Clearly, the edge vq represents the point $f_v(q) \in A$, and the adjacent vertices w correspond to the pieces A_w which contain this point.

The graphs can have multiple edges. If some equivalence class q contains more than 2 elements, the G_n are hypergraphs. In such cases, which we neglect in the present paper, each hyperedge can be replaced by a star of edges with a central point, so that we get graphs again. Anyway, it seems most appropriate to consider G_n as a T_0-space, where the vertices are open points, (hyper-) edges are closed points and the open hull of an edge consists of this edge together with all adjacent vertices (the dual of the natural topology). For $k < n$, the projection $p_{nk} : G_n \to G_k$ with $p_{nk}(w) = w_{|k}$, $p_{nk}(vq) = v_{|k}$ for $|v| \leq k$ and $p_{nk}(vq) = vq$ otherwise is continuous. A is essentially the inverse limit of (G_n, p_{nk}) [4].

For our purposes, G_1 and G_2 will be sufficient, however. Already G_1

gives important information on the topology of the factor [1, 10, 17]. A is connected iff G_1 is connected, and A is a dendrite iff G_1 is a tree. We shall assume throughout that A is *connected*, which implies that A is locally connected and has dimension 1.

A point x in A is a *cutpoint* if $A \setminus \{x\}$ is not connected. x is a *local cutpoint* if x is a cutpoint of a connected neighbourhood of x. Obviously all critical points q and their images $f_v(q)$ are local cutpoints. In order to characterize the critical points by separation properties, we must exclude other cutpoints like the $e_i = p(\bar{i})$ in Fig. 1. A connected graph G is said to be 2-*connected* if it has no cutpoints, i.e. if $G \setminus \{u\}$ is connected for each vertex u in G. Note that G_2 is 2-connected in Fig. 1 but not in Fig. 2.

PROPOSITION 2.1. *If $A = C/\sim$ is a finite-to-one invariant factor and G_2 is* 2-*connected*, *then A has no global cutpoints*, *and no local cutpoints other than the f_v-images of the critical points* ($v \in S^*$).

Proof. We show by induction that G_n is 2-connected for $n > 2$. Assume G_{n-1} is 2-connected and $G_n \setminus \{u\} = H_1 \cup H_2$ for some vertex $u = u_1 \dots u_n$, where no edge connects the disjoint graphs H_1 and H_2. Since G_1 is connected, any copy of G_1 in G_n is contained in either H_1 or H_2. Thus the projections of H_1 and H_2 are graphs in G_{n-1} which have no common vertex, with the possible exception of $u' = u_1 \dots u_{n-1}$ which is the projection of u. Now u' is not a cutpoint of G_{n-1}, so one of the sets H_i must be contained in the subgraph of G_n which corresponds to u' and is isomorphic to G_1. Consider the larger subgraph G of G_n which contains u and is isomorphic to G_2. Both H_1 and H_2 intersect G, contradicting our assumption.

Since G_n is 2-connected, it cannot be disconnected by deletion of an edge vq. Thus the points $f_v(q)$ are not global cutpoints. Finally, suppose $a = p(s)$ is a local cutpoint in A, but not of the form $f_v(q)$. Then the A_w with $w = s_{|k}$ form a neighbourhood base of a, and $A_w \setminus \{a\}$ is disconnected for some k. This implies that $A_{s_{|k}} \setminus A_{s_{|n}}$ is disconnected for some $n > k$, which contradicts the 2-connectedness of G_{n-k}. ■

Note that if G_2 is not 2-connected, it may still happen that G_3 is 2-connected. For an example, take $m = 5$ and any generating rules of the form $122\ldots \sim 211\ldots$, $144\ldots \sim 334\ldots$, $255\ldots \sim 311\ldots$, $332\ldots \sim 522\ldots$, $355\ldots \sim 411\ldots$ and $455\ldots \sim 544\ldots$ Even if all G_n have cutpoints, A need not have global cutpoints: let $m = 3$ and $123\ldots \sim 313\ldots$, $122\ldots \sim 211\ldots$ and $233\ldots \sim 322\ldots$

3. Edge-balanced graphs. Here is a combinatorial concept which will be very helpful for our proofs. A connected graph G with m vertices and c edges is said to be *edge-balanced* if for each k with $1 < k < m$, the graph cannot be divided into k components by deleting $(k-1)c/(m-1)$ or less edges. This property seems to be interesting enough to justify a brief discussion.

A cycle with m vertices and edges is edge-balanced for each m. If one edge is added to the cycle, the resulting graph is edge-balanced only for $m = 4$. However, a pentagon with two chords or a hexagon with two chords without common endpoint is again edge-balanced. On the other hand, no vertex in an edge-balanced graph has degree 1.

Remark 3.1. *An edge-balanced graph is* 2*-connected.*

Proof. Suppose a vertex u in G is a cutpoint. Then there are m_1 vertices connected with each other and with u by c_1 edges, and $m_2 = m - m_1 - 1$ remaining vertices connected with each other and with u by $c_2 = c - c_1$ edges. If G is edge-balanced, then

$$c_1 > \frac{m_1}{m-1}c \quad \text{and} \quad c_2 > \frac{m_2}{m-1}c\,,$$

which implies $c > c$. ∎

Remark 3.2. *Let G be obtained from the complete graph K_m with m vertices by the removal of r edges, where $0 \le r \le m/2 - 1$. Then G is edge-balanced.*

Proof. Suppose we can divide G into $k < m$ components by deleting not more than $(k-1)c/(m-1)$ edges. Then since $(m-1)/2 \le c/(m-1)$, we can separate any singleton from a component with not more than $m/2$ vertices, obtaining $k' = k+1$ components by deleting not more than $(k'-1)c/(m-1)$ edges. Thus we can assume from the beginning that $k-1$ of the components are singletons. The number of deleted edges is at least $m-1+\ldots+m-k+1\ -r$. By assumption,

$$\frac{1}{2}(k-1)(2m-k) - r \le (k-1)\left(\frac{m}{2} - \frac{r}{m-1}\right),$$

which implies $(k-1)(m-1) \le 2r$. Even the smallest value, for $k = 2$, contradicts our choice of r. ∎

We can remove even more edges provided they have no common vertex.

4. Separation properties of critical points. In this section we show that for many factors, the fractal structure is fully determined by the topology of A. It will be sufficient to show that the family $\{A_i \mid i \in S\}$ is determined by the topology: since each A_i is homeomorphic to A, we can then determine the A_{ij} and, by induction, the A_w. Next, since the A_i are the closures of components of $A \setminus Q$, it suffices to describe Q in terms of the topology of A. We shall characterize Q by separation properties.

A finite set F in a connected space X is said to *cut* X *into* k *pieces* if $X \setminus F$ has k components. For the "n-dimensional" Sierpiński gasket, Q is the only subset with $n+1$ points which cuts A into more than n pieces. Here is a more general statement. Let us say that a set V of words *contains predecessors* if V contains the empty word, and $v_1 \ldots v_n \in V$ implies $v_1 \ldots v_{n'} \in V$ for $n' < n$.

THEOREM 4.1. *Let A be a simple finite-to-one invariant factor such that G_1 is edge-balanced and G_2 is* 2*-connected. Then a finite subset F which cuts A into k pieces must satisfy*

$$\operatorname{card} F \geq \frac{k-1}{m-1} \operatorname{card} Q .$$

Equality holds iff F has the form $F = \bigcup \{f_v(Q) \mid v \in V\}$, where V contains predecessors.

Proof. Let $c = \operatorname{card} Q$ and $d = \operatorname{card} F$. We can assume that all points of F have the form $f_v(q)$ with $v \in S^*$ and $q \in Q$, since by Proposition 2.1, other points are not local cutpoints. Since A is simple, this representation is unique. Let $F_v = F \cap f_v(Q)$, and let $V = \{v \mid F_v \neq \emptyset\}$. For $v \in V$ let $d_v = \operatorname{card} F_v$ and k_v the number of components of $A_v \setminus F_v$. Since $1 \leq k_v \leq m$ and G_1 is edge-balanced, $d_v \geq (k_v - 1)c/(m-1)$. Summing over v, we get

$$(*) \qquad \sum_{v \in V} (k_v - 1) \leq \frac{d}{c}(m-1) .$$

Equality holds iff $k_v = m$ for all $v \in V$.

Let k denote the number of components of $A \setminus F$. We show

$$(**) \qquad k - 1 \leq \sum_{v \in V} (k_v - 1) ,$$

using induction on $\operatorname{card} V$. Let $V = V' \cup \{w\}$, let $(A \setminus F) \cup F_w$ have k' components, and assume $k' - 1 \leq \sum_{v \in V'} (k_v - 1)$. If we subtract F_w, we have only to regard that component B which contains the interior of A_w. Now $A_w \setminus F_w$ has k_w pieces, and $B \setminus F_w$ cannot have more. Thus $k \leq k' + k_w - 1$, which implies $(**)$.

Combining $(*)$ with $(**)$, we get

$$\frac{k-1}{m-1} \leq \frac{d}{c} ,$$

which we wanted to prove. Equality in $(*)$ was true iff $F_v = f_v(Q)$ for each v in V. Under this condition, let us discuss when $(**)$ turns into equality. In our induction, assume the words are ordered by increasing length, and w is the last word of length $|w|$. If the predecessor of w is contained in V', then $B \subset A_w$ and $k = k' + k_w - 1$. Otherwise, by the 2-connectedness of the graphs G_n, there is a path in $A \setminus A_w$ with endpoints in two different components of $A_w \setminus F_w$. Either the path is contained in $A \setminus \bigcup\{F_v \mid v \in V'\}$, or part of this path connects one component of $A_w \setminus F_w$ with a component of some $A_v \setminus F_v$ with $v \in V'$. In both cases, $k < k' + k_w - 1$. The theorem is proved. ∎

THEOREM 4.2. *Let A be a simple finite-to-one invariant factor of $\{1, \ldots, m\}^\infty$ such that G_1 is edge-balanced and G_2 is 2-connected. Then the fractal structure of A is determined by the topology.*

Proof. Theorem 4.1 says that Q is the only set with c points which cuts A into m pieces. ∎

To show that the theorem applies to the Sierpiński gasket, it remains to check that G_2 is 2-connected. This graph consists of $m = n+1$ copies of the complete graph K_m, where any two copies are joined by an edge. This graph can only have a cutpoint if all edges from one copy to the others start in the same vertex. For the generating rules $i\overline{j} \sim j\overline{i}$, $i \neq j$, this is not the case. More generally, we have

PROPOSITION 4.3. *Define an invariant factor A by the rules $is^{ij} \sim js^{ji}$, $1 \leq i, j \leq m$, $i \neq j$. If none of these sequences is a tail sequence of another one, and if for each i, not all s^{ij} have the same initial letter, the assumptions of Theorem 4.1 are satisfied.* ∎

Obviously there is a continuum of different sets of rules, even if we require that for each i, all s^{ij} have different initial letters. From Remark 5.7 it will follow that only $m!$ of these factors can be mutually homeomorphic, so that the number of non-homeomorphic factors of this type has at least the cardinality of the continuum.

5. The structure of the homeomorphism group

THEOREM 5.1. *Let A be a simple finite-to-one invariant factor such that G_1 is edge-balanced and G_2 is 2-connected. Then each subset of A homeomorphic to A is of the form A_w.*

Proof. It is enough to show $h(A) = A$ for each homeomorphism $h : A \to A$ for which $h(A)$ is not contained in some A_i. First assume $h(A)$ intersects the interiors of $A_1, \ldots, A_k$ with $2 \leq k < m$, let Q_1 denote the set of critical points which are contained in two sets A_i with $i \leq k$, and $Q_2 = Q \setminus Q_1$. Deletion of the edges corresponding to Q_2 divides G_1 into $m - k + 1$ components. Thus $\operatorname{card} Q_2 > (m-k)c/(m-1)$. On the other hand, deletion of Q_1 from $h(A)$ divides $h(A)$ into k components, so that by Remark 3.2, $\operatorname{card} Q_1 > (k-1)c/(m-1)$. Adding we get $\operatorname{card} Q > c$.

The contradiction shows that $h(A)$ must intersect the interiors of *all* pieces A_i, so that Q divides $h(A)$ into m pieces. Thus $h(Q) = Q$ by our characterization of Q, and there is a permutation π of $\{1,\ldots,m\}$ such that $h(A_i)\subseteq A_{\pi(i)}$. Now we repeat the above argument to conclude that $h(A_i)$ contains $f_{\pi(i)}(Q)$. Proceeding by induction, we show that $h(A)$ contains $f_w(Q)$ for all $w\in S^*$. Hence $h(A)=A$. ■

COROLLARY 5.2. *Each homeomorphism h from A into A can be written as $h=f_w g$, where $w\in S^*$ and g is an onto homeomorphism.*

Proof. If $h(A)=A_w$, define $g=f_w^{-1}h$. ■

Let A be a finite-to-one invariant factor. A homeomorphism h from A onto A is said to *preserve the fractal structure* if for each $n\geq 1$ and each $w\in S^n$, there is a word $v\in S^n$ with $h(A_w)=A_v$. Theorem 5.1 may be rephrased as follows.

COROLLARY 5.3. *Each homeomorphism h from A onto A preserves the fractal structure.* ■

We add some remarks concerning the structure of the group of all homeomorphisms h preserving the fractal structure of a finite-to-one invariant factor A. Since h permutes the pieces A_i as well as their intersection points, the critical points of A, it induces a graph isomorphism $h_1 : G_1\to G_1$. Similarly, h permutes the A_{ij} and induces an isomorphism h_2 of G_2. By induction it is easy to show

Remark 5.4. *A homeomorphism h of a finite-to-one invariant factor A preserves the fractal structure iff it is the inverse limit of a sequence of graph isomorphisms $h_n : G_n\to G_n$ with $h_{n-1}p_{n,n-1}=p_{n,n-1}h_n$, $n=1,2,\ldots$* ■

Let H_λ denote the group of all graph isomorphisms of G_1 which can be extended to such a compatible sequence. For each $h_\lambda^1\in H_\lambda$ we fix one extension $h_\lambda=(h_\lambda^n)_{n\in\mathbb{N}}$. The map h_λ^1 describes a permutation of the A_i. If $h_\lambda^1=\mathrm{id}$, the A_{ij}, $j\in S$, can still be interchanged, but those A_{ij} which are incident to some A_k, $k\neq i$, have to be fixed. For $w\in S^n$, $n=1,2,\ldots$, let I_w denote the set of all $i\in S$ such that wi is connected in G_{n+1} to some vertex vj with $v\neq w$. Let H_w be the stabilizer of I_w in H_λ. For each $h_w^1\in H_w$ we again fix one extension $h_w=(h_w^n)_{n\in\mathbb{N}}$. Then each homeomorphism g on A preserving the fractal structure is given by a family $(h_w\in H_w)_{w\in S^*}$.

Remark 5.5. *For any finite-to-one invariant factor*, *there is a one-to-one correspondence between $\prod_{w\in S^*}H_w$ and the group of all structure-preserving homeomorphisms on A which assigns to $(h_w)_{w\in S^*}$ the following compatible sequence of graph isomorphisms $g_n : G_n\to G_n$*:

$$g_n(i_1\ldots i_n)=h_\lambda^n(i_1\ldots i_n)h_{i_1}^{n-1}(i_2\ldots i_n)\ldots h_{i_1\ldots i_{n-1}}^1(i_n)\,.$$

We omit the straightforward proof and discuss those cases where the group of structure-preserving homeomorphisms is finite. Obviously, $I_w\subset$

I_{uw} and hence $H_{uw} \subset H_w$ for arbitrary words $u, w \in S^*$. Consequently, we have

COROLLARY 5.6. *If there is a k such that $H_w = \{\mathrm{id}\}$ for all $w \in S^k$, then H_w is also trivial for all words of length greater than n, and the group of structure-preserving homeomorphisms is finite.* ∎

Remark 5.7. *Let A be a simple finite-to-one invariant factor. If either G_1 is an odd cycle and G_2 is 2-connected (as in Fig. 3), or $G_1 = K_m$, $m \geq 3$, and in G_2 only m vertices have degree $m-1$ (as for the Sierpiński gasket), then H_w is trivial for $|w| \geq 1$, and each homeomorphism from A onto A is fully determined by the corresponding permutation of the pieces A_i of A.* ∎

For the Sierpiński gasket and Fig. 3, each permutation of the pieces is realized by an isometry. This proves Proposition 1.2. In general, however, the above assumptions on G_1, G_2 do not guarantee that non-trivial homeomorphisms from A onto A exist. An arbitrary choice of generating rules rather leads to a rigid factor. Let us also mention an example with infinite automorphism group. Let $m = 4$, and let A be generated by the rules $1\overline{2} \sim 2\overline{1}$, $2\overline{3} \sim 3\overline{2}$, $3\overline{4} \sim 4\overline{3}$ and $4\overline{1} \sim 1\overline{4}$. Then H_w has two elements if w ends with ii or $i(i+2 \bmod 4)$ for some i, and is trivial otherwise.

6. The minimal fractal structure. In Theorem 4.2, we had to assume that m is given in order to make the fractal structure unique. The ordinary Sierpiński gasket ($n = 2$), for instance, has $m = 3$ pieces, but it can also be divided into 5 pieces by $Q \cup f_1(Q)$. This flaw can be overcome by requiring that either m is *minimal*, or G_1 is edge-balanced.

THEOREM 6.1. *Let A be a simple finite-to-one invariant factor such that G_1 is edge-balanced and G_2 is 2-connected. Then $A = A_1 \cup \ldots \cup A_m$ is the only covering of A by m or fewer sets homeomorphic to A. If $A = B_1 \cup \ldots \cup B_r$ is another covering of A by homeomorphic copies of A, where no B_i is contained in another B_j, then the corresponding graph G_1 is not edge-balanced.*

Proof. Both assertions are obvious from Theorem 5.1. For the second one, reverse the roles of m and r. ∎

It is easy to show that the B_i must be the closures of components of $A \setminus F$, where $F = \bigcup_{v \in V} f_v(Q)$ and V contains predecessors (cf. Theorem 4.1). This is a stronger form of minimality of the covering of A by the A_i.

It seems unclear whether the uniqueness of a minimal fractal structure holds under more general conditions. In Fig. 2, for instance, it can be shown that the closure of each component of $A \setminus \{e_1\}$ is homeomorphic to A. Nevertheless, this covering of A by two homeomorphic copies does not lead to an invariant factor: with two pieces and one critical point, such a factor must be a dendrite.

It remains to show Proposition 1.3. If d is a metric such that each homeomorphism from A onto A is an isometry, then $d(e_i, e_j) = t > 0$ for all $i \neq j$. Now suppose f_i is a similarity with factor r_i, and d is an interior metric. Since each path from e_i to e_j passes through a critical point, each of these paths is longer than $r_i t + r_j t$, except for the shortest path through the point corresponding to $i\overline{j} \sim j\overline{i}$, which has exactly this length. Thus $t = r_i t + r_j t$ for $i \neq j$, which implies $r_i = 1/2$ for all i. The side length of A_w is $2^{-|w|}t$, and this uniquely determines an interior metric.

References

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

[2] —, —, *Symbolic dynamics for angle-doubling on the circle I. The topology of locally connected Julia sets*, in: U. Krengel, K. Richter and V. Warstat (eds.), Ergodic Theory and Related Topics III, Lecture Notes in Math. 1514, Springer, 1992, 1–23.

[3] C. Bandt and T. Kuschel, *Self-similar sets* 8. *Average interior distance in some fractals*, Proc. Measure Theory Conference, Oberwolfach 1990, Rend. Circ. Mat. Palermo (2) Suppl. 28 (1992), 307–317.

[4] C. Bandt and T. Retta, *Self-similar sets as inverse limits of finite topological spaces*, in: C. Bandt, J. Flachsmeyer and H. Haase (eds.), Topology, Measures and Fractals, Akademie-Verlag, Berlin 1992, 41–46.

[5] C. Bandt and J. Stahnke, *Self-similar sets.* 6. *Interior distance on deterministic fractals*, preprint, Greifswald 1990.

[6] M. T. Barlow and E. A. Perkins, *Brownian motion on the Sierpiński gasket*, Probab. Theory Related Fields 79 (1988), 543–623.

[7] P. F. Duvall, J. W. Emert and L. S. Husch, *Iterated function systems, compact semigroups and topological contractions*, preprint, University of North Carolina, 1990.

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

[9] J. de Groot, *Groups represented as homeomorphism groups I*, Math. Ann. 138 (1959), 80–102.

[10] M. Hata, *On the structure of self-similar sets*, Japan J. Appl. Math. 2 (2) (1985), 381–414.

[11] A. Hinz and A. Schief, *The average distance on the Sierpiński gasket*, Probab. Theory Related Fields 87 (1990), 129–138.

[12] V. Kannan and M. Rajagopalan, *Constructions and applications of rigid spaces*, *I*, Adv. in Math. 29 (1978), 89–130; *II*, Amer. J. Math. 100 (1978), 1139–1172.

[13] J. Kigami, *A harmonic calculus on the Sierpiński spaces*, Japan J. Appl. Math. 6 (2) (1989), 259–290.

[14] —, *Harmonic calculus on p.c.f. self-similar sets*, Trans. Amer. Math. Soc., to appear.

[15] K. Kuratowski, *Topology*, Vol. 2, Polish Scientific Publishers and Academic Press, 1968.

[16] T. Lindstrøm, *Brownian motion on nested fractals*, Mem. Amer. Math. Soc. 83 (1990).

[17] C. Penrose, *On quotients of the shift associate with dendritic Julia sets of quadratic polynomials*, Ph.D. thesis, University of Warwick, 1990.

[18] W. Rinow, *Die innere Geometrie der metrischen Räume*, Grundlehren Math. Wiss. 105, Springer, 1961.

[19] W. P. Thurston, *On the combinatorics and dynamics of iterated rational maps*, preprint, Princeton 1985.

FACHBEREICH MATHEMATIK
ERNST-MORITZ-ARNDT-UNIVERSITÄT
D-O-2200 GREIFSWALD, GERMANY

MATHEMATICS DEPARTMENT
ADDIS ABABA UNIVERSITY
P.O. BOX 1176, ADDIS ABABA, ETHIOPIA

Received 11 November 1991;
in revised form 17 March 1992

Added in proof (October 1992). Proposition 1.2 and a more general statement were proved in: W. Dębski and J. Mioduszewski, *Simple plane images of the Sierpiński triangular curve are nowhere dense,* Colloq. Math. 59 (1990), 125–140.