J. DIFFERENTIAL GEOMETRY
28 (1988) 203-221

ON THE SHAPE OF CANTOR SETS

DARYL COOPER & THEA PIGNATARO

There has been much interest recently in sets which have a self-similar, fractal nature. As a prototype of a more general problem, we study a class of Cantor sets on the line from the point of view of bi-Lipschitz geometry, i.e., quasi-isometry. These investigations reveal a surprising general principle: A quasi-isometry between such objects is essentially the same thing as a map which is linear on the level of measure theory, i.e., has constant Radon-Nikodým derivative with respect to Hausdorff measure. This principle provides new invariants which enable us to classify generic Cantor sets of the type we consider.

Motivated by a question of Dennis Sullivan, we begin the classification of certain types of Cantor sets on the line. These Cantor sets each arise as the maximal invariant set of a map to the line defined on a neighborhood of the Cantor set. The middle third Cantor set arises in this way: Let $J_1 = [0, 1/3]$, $J_2 = [2/3, 1]$ and $J = [0, 1]$, and define $\tau\colon J_1 \cup J_2 \to J$ by

$$\tau(x) = \begin{cases} 3x & \text{if } x \in J_1, \\ 3x - 2 & \text{if } x \in J_2; \end{cases}$$

then $C = \tau^{-\infty}(J)$ is the middle third Cantor set. The ultimate goal is to understand the structure of C when τ is a locally expanding map. In this paper we consider the case when τ is defined on a finite number of disjoint closed intervals $J_1, \cdots, J_q$ contained in an interval J and τ maps each J_i affinely onto J. The topological and C^1 classifications of such sets are trivial, however the classification up to quasi-isometry (= order preserving bi-Lipschitz homeomorphism) gives a rich theory. The basic result is that a quasi-isometry can always be replaced by a quasi-isometry, possibly no longer surjective, with constant Radon-Nikodým derivative. The measure involved is the Hausdorff measure on the Cantor set, which is positive in its Hausdorff dimension. This result alone gives a powerful new invariant; previously the only known invariant was the Hausdorff dimension. These Cantor sets are homogeneous in the sense that any two points x, y have arbitrarily small affinely isomorphic neighborhoods in the Cantor set, although the isomorphism need not map x to y.

Received May 28, 1986 and, in revised form, September 22, 1987.

A similarity point can be described as a point, x, of the Cantor set which has a neighborhood in the Cantor set which is homeomorphic to a proper subset by an affine map fixing x. Measure preserving quasi-isometries preserve these points, together with a numerical invariant assigned to each such point. These numbers are related to the invariants of any dynamical system producing the Cantor set.

The classification is completely understood in the generic case in which the ratios of the masses (using Hausdorff measure) of $J_2, \cdots, J_q$ to the mass of J_1 are algebraically independent over $\mathbb{Q}$. In this case any affinely generated Cantor set of the type we are considering which is quasi-isometric to a generic example can only be generated by affine maps defined on intervals $J_1', \cdots, J_r'$ which are obtained from the intervals $J_1, \cdots, J_q$ by two geometric operations called splitting and sliding. In this case, the gap invariant of §2 is a complete invariant. In the nongeneric case, splitting and sliding still produce quasi-isometric Cantor sets, but there are quasi-isometric Cantor sets not related by these operations. For some of these examples, the gap invariant is not a complete invariant, the extra information comes from the self-similarity invariant. The middle third set is a nongeneric example for which the classification is worked out in §4. A gap remains in our knowledge of what the classification is for some cases.

1. Definitions

The Cantor sets we study have a self-similarity structure arising from some dynamical system which produces the set. This structure is not unique, but it is convenient to utilize some specific decomposition of the Cantor set determined by a choice of dynamical system τ. This decomposition is called a *clone structure* and is defined below. At various times it is useful to be able to consider alternative clone structures, just as it is useful to be able to change coordinate systems in a manifold.

Let $J_1, J_2, \cdots, J_q$ be disjoint closed intervals contained in a closed interval J (which we will always assume is the unit interval $[0,1]$) and ordered from left to right. Thus J_i is to the left of J_{i+1}. We further assume that J_1 contains the left-hand endpoint of J, and that J_q contains the right-hand endpoint of J. This ensures that the endpoints of J will be in the Cantor set. Let $\tau_i\colon J_i \to J$ be the unique order preserving affine map of J_i onto J. Set $\tau = \bigcup \tau_i\colon \bigcup J_i \to J$; then the Cantor set $C(\tau)$, or just C, is $\bigcap_{n=1}^{\infty} \tau^{-n}(J)$. The components of $\tau^{-n}(J)$ are called the level-n *clones* of C. Thus $J_1, \cdots, J_q$ are the level-1 clones. We also think of a clone as the subset of C contained in the clone. If A is a level-n clone, then $\tau^n|A$ maps A affinely onto J, and so the subset of the Cantor set contained in A is just a linearly scaled down copy of C. $\tau^n|A$ is a *level-n clone map* and may be written uniquely as an n-fold composition $\tau_{i_1} \circ \tau_{i_2} \circ \cdots \circ \tau_{i_n}$, $1 \leq i_j \leq q$, of level-1 clone maps. We say that C is *affinely generated* by the maps $\tau_1, \cdots, \tau_q$. Notice that C is uniquely determined by the collection of intervals $J_1, \cdots, J_q$. A *gap*, G, is a component of $J - C$, and G is at *level-n* if G is a component of $[J - \tau^{-n}(J)] - [J - \tau^{1-n}(J)]$. Thus there are $q-1$ level-1 gaps, which we label from left to right, $G_1, \cdots, G_{q-1}$. Suppose G is a level-n gap, then there is a unique level-$(n-1)$ clone map sending G to a level-1 gap G_i for some $1 \leq i \leq q-1$. The *name* of G is the integer i. Suppose A is a level-m clone, and $B \subset A$ is a level-$(m+n)$ clone, then B has *level-n relative* to A. Similarly a gap $G \subset A$ at level-$(m+n)$ has level-n relative to A. By a *clopen* we mean a subset of C which is both open and closed as a subset of C. Note that every clopen is a union of finitely many disjoint clones. A *co-gap* is a particular type of clopen, it is the interval between two gaps in C. Alternatively we think of it as the subset of C contained in that interval. The following is obvious.

Lemma (1.1). *Let $A \subset J$ be an interval, and let B be the smallest clone containing A. Then there is a gap G at level-1 relative to B such that $A \cap G \neq \emptyset$.*

The Hausdorff dimension of C is that $d \in (0,1)$ satisfying the equation $\sum_{j=1}^{q} |J_j|^d = |J|^d$ and the associated Hausdorff measure μ_C gives, for every clone A, $\mu_C(A) = |A|^d$. We will assume that J is the unit interval $[0,1]$ and that μ_C has been normalized to be a probability measure on C. (See [2] for more information about the Hausdorff measure on C.) Let x be a point in J; the *measure* (or *mass*) *coordinate* of x is $\mu_C([0,x])$. If x and y are two points of J which have the same measure coordinate, then there are no points z with $x < z < y$ and $z \in C$. Thus if $x, y \in C$ have the same measure coordinate, then they are endpoints of the same gap. We are interested in the classification of Cantor sets up to order preserving bi-Lipschitz maps. This means we are considering maps $\phi \colon \mathbb{R} \to \mathbb{R}$ for which there is a (bi-Lipschitz) constant $K > 0$ such that $\forall x,\ y \in \mathbb{R}$, $K^{-1} < |\phi(x) - \phi(y)|/|x-y| < K$. We call such a map a *quasi-isometry.* If C_1 and C_2 are quasi-isometric Cantor sets we write $C_1 \simeq C_2$.

2. General results

The main result of this section is that the classification of affinely generated Cantor sets is the same for the relations of quasi-isometric bijections and measure linear quasi-isometric injections (measure linear means constant Radon-

Nikodým derivative). This result provides two new invariants of affinely generated Cantor sets. The first is the *gap invariant* which is the countable subset of [0,1] consisting of the mass coordinates of the gaps. The second invariant comprises real numbers assigned to each of the countable set of *similarity points* defined below. These numbers are scale factors so that if a neighborhood of a similarity point is contracted by its scale factor, the neighborhood and its contraction appear identical. It is convenient, and in keeping with the philosophy of this paper, to measure the size of objects using Hausdorff measure instead of Euclidean length. Thus the scale factor measures the enlargement of mass, not distance, at a similarity point. In Sullivan's work [3] on the classification of the dynamical systems producing these Cantor sets, an important set of invariants are the eigenvalues of periodic points of the dynamical system. These eigenvalues are closely related to the scale factors. However, the notions are distinct because there are always many different dynamical systems generating any given affinely generated Cantor set.

There are two simple ways of producing affinely generated Cantor sets quasi-isometric to a given one. These methods are the operations of *splitting* and *sliding*. Let C be a Cantor set determined by a set of disjoint closed intervals $J_1, \cdots, J_q$ contained in an interval J, and $\tau_i\colon J_i \to J$ the associated affine maps. Let C', $J_1', \cdots, J_r' \subset J'$, and $\tau_i'\colon J_i' \to J'$ be another such set. If $q = r$, $|J| = |J'|$ and $|J_i| = |J_i'|$ for each i, then we say that the intervals $\{J_i'\}$ are obtained from the intervals $\{J_i\}$ by *sliding* (see Figure 1). Note that according to the conventions of §1, the sequence of $J_1, \cdots, J_q$ is from left to right on J and this is also the sequence of $J_1', \cdots, J_q'$ on J'. Thus the $\{J_i'\}$ differ from the $\{J_i\}$ only by their relative positions on J' and J respectively. There is a natural quasi-isometry $\phi\colon J \to J'$ defined as follows. Let $G_1, \cdots, G_{q-1}$ be the level-1 gaps of C, and $G_1', \cdots, G_{q-1}'$ the level-1 gaps of C'. ϕ is defined so that it preserves the name and level of gaps, i.e., $\phi(G_i) = G_i'$ for $1 \leq i \leq q-1$ and if G is the level-n gap $(\tau_{i_1} \circ \cdots \circ \tau_{i_n})^{-1} G_i$, then $\phi(G) = (\tau_{i_1}' \circ \cdots \circ \tau_{i_n}')^{-1} G_i'$. Since τ_j and τ_j' have the same derivative, it follows that $|\phi(G)|/|G| = |G_i'|/|G_i|$ so ϕ has bi-Lipschitz constant $\max_{1 \leq i \leq q-1}\{|G_i'|/|G_i|, |G_i|/|G_i'|\}$. Note that ϕ is measure preserving because it preserves the length, and hence mass, of clones. ϕ is called the *slide map*. Another operation called *splitting* also produces quasi-isometric Cantor sets. In this case $r = 2q-1$ and

$$J_1', \cdots, J_r' = J_1, \cdots, J_{i-1}, \tau_i^{-1}(J_1), \cdots, \tau_i^{-1}(J_q), J_{i+1}, \cdots, J_q$$

for some fixed $1 \leq i \leq q$ (see Figure 2). Thus the level-1 clones of C' consist of the level-2 clones of C contained in J_i, together with the level-1 clones of C other than J_i. We say that J_i has been *split*. More generally, let $J_1', \cdots, J_r'$

$(r = q + n(q-1))$ be any collection of disjoint clones of C whose union contains C. Then we say that these clones are obtained by splitting from those defining C. It is an elementary fact that such a collection of clones can always be obtained from $\{J_i\}$ by a sequence of single splits, a single split being the replacement of a level-p clone by the level-$(p+1)$ clones it contains. The Cantor set C' determined by these intervals is quasi-isometric to C by the identity map. What has happened is that a new dynamical system generating the same Cantor set has been defined. Alternatively, we have changed the clone structure of C. Finite sequences of splitting and sliding generate an equivalence relation on affinely generated Cantor sets. In §3 we show that for generic affine Cantor sets this is the same equivalence relation as quasi-isometry.

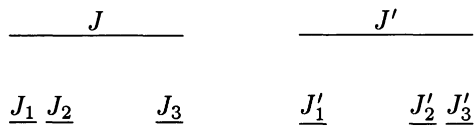

FIGURE 1. Sliding; $q = 3$.

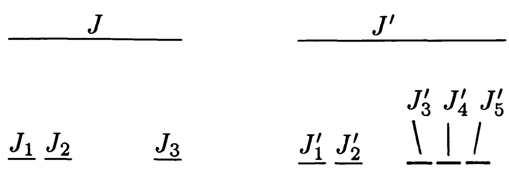

FIGURE 2. Splitting; $q = 3$, $i = 3$.

Example. This is an example of a quasi-isometry ϕ of the unit interval which maps the middle third Cantor set onto itself. The map ϕ is given by the formula:

$$\phi(x) = \begin{cases} x/3, & 0 \le x \le 2/3, \\ x - 4/9, & 2/3 \le x \le 7/9, \\ 3x - 2, & 7/9 \le x \le 1. \end{cases}$$

A more instructive description of ϕ is obtained by labelling the level-2 clones of C from left to right as J_1, J_2, J_3, and J_4, each of which has length $1/9$. Then ϕ maps $J_1 \cup J_2$ linearly onto J_1, maps J_3 linearly onto J_2 and maps J_4 linearly onto $J_3 \cup J_4$.

Definition. Suppose (X, μ_X) and (Y, μ_Y) are measure spaces. A map $\phi\colon X \to Y$ is *measure linear* if there is a constant $\lambda > 0$ such that for all μ_X-measurable sets $B \subset X$, ϕB is μ_Y-measureable and $\mu_Y(\phi B) = \lambda \mu_X(B)$.

Theorem (2.1). *Suppose that C and C' are affinely generated Cantor sets and that ϕ is a quasi-isometry of the line with $\phi(C) \subseteq C'$. Then there is a clone $A \subset C$ such that $\phi|A$ is measure linear with respect to the Hausdorff measures μ_C and $\mu_{C'}$ on C and C' respectively.*

Proof. If K is the bi-Lipschitz constant for ϕ and d is the Hausdorff dimension of C, then for every μ_C-measurable $B \subset C$, $K^{-d} \leq \mu_{C'}(\phi B)/\mu_C(B) \leq K^d$. The *mass ratio* of B is defined to be $\mathrm{MR}(B) \equiv \mu_{C'}(\phi B)/\mu_C(B)$. Fix $\varepsilon > 0$ and choose a clone A of C with $\mathrm{MR}(A) + \varepsilon > \sup\{\mathrm{MR}(B)\colon B \text{ a clone of } C\} \equiv M$. We show that for ε small enough, $\phi|A$ is measure linear.

Claim. There is a finite set $S \subset \mathbb{R}$ such that for every clone D of C and every clone $E \subset D$ at level-1 relative to D, $\rho(E, D) \equiv \mathrm{MR}(E)/\mathrm{MR}(D) \in S$.

We remark that S depends only on C, C' and K, not on ϕ. Let F' be the smallest clone of C' containing $\phi(D)$ and $\sigma\colon F' \to C'$ the corresponding clone map. Because σ is linear, and therefore measure linear,

$$\rho(E, D) = \frac{\mu_{C'}(\sigma \circ \phi E)}{\mu_C(E)} \Big/ \frac{\mu_{C'}(\sigma \circ \phi D)}{\mu_C(D)}.$$

Let $H_1, \cdots, H_{q-1}$ be the gaps in D at relative level-1, and set $H_i' = \sigma \circ \phi(H_i)$. Then H_i' is a gap of C'. By Lemma (1.1), $\sigma \circ \phi(D)$ contains a level-1 gap of C' and so $|\sigma \circ \phi(D)|$ is bounded below. Now $|H_i| \cdot |D|^{-1}$ is bounded below and ϕ is a K-quasi-isometry; therefore $|\sigma \circ \phi H_i| \cdot |\sigma \circ \phi D|^{-1}$ is bounded below. It follows that $|H_i'|$ is bounded below and, since there are only finitely many gaps in C' with size larger than a given lower bound, there are only finitely many possibilities for H_i'. Similarly there are only finitely many possibilities for the image under $\sigma \circ \phi$ of the left- and right-hand end points of D. This is because adjacent to D in C are gaps G_L, G_R on the left and right of D, and $|G_L| \cdot |D|^{-1}$, $|G_R| \cdot |D|^{-1}$ are bounded below (though not above) independently of the choice of D. Then, arguing as above shows that there are only finitely many possibilities for $\sigma \circ \phi(G_L)$ and $\sigma \circ \phi(G_R)$. Thus there are only finitely many possibilities for the image under $\sigma \circ \phi$ of D and its relative level-1 gaps, and thus there are only finitely many possibilities for the image under $\sigma \circ \phi$ of relative level-1 clones of D. Thus there are only finitely many possibilities for $\rho(E, D)$ as claimed.

If $S = \{1\}$, then ϕ is measure linear, otherwise S contains a number larger than 1. This is because

$$1 = \sum_{\substack{\text{level } -p \text{ clones} \\ B \text{ in } C}} \mu_{C'}(\phi B)/\mu_{C'}(\phi C) = \sum \rho(B, C)\mu_C(B)/\mu_C(C),$$

and so if $\rho(B, C) \leq 1$ for all clones B, then $\rho(B, C) = 1$ for all clones. Define $R = \min\{r \in S\colon r > 1\}$; in a sense this is the minimum amount of nonlinearity which can occur at the measure level for ϕ. If $\phi|A$ is not measure linear there is a clone $B \subset A$ with $\rho(B, A) \neq 1$. Choose such a B of the least depth, i.e. to have level-p with p as small as possible. Then there must be a level-p clone $B' \subset A$ with $\rho(B', A) > 1$. Let B'' be the unique level-$(p-1)$ clone containing B', thus $B'' \subset A$, and $\rho(B'', A) = 1$ by choice of B'. Now $\rho(B', A) = \rho(B', B'')\rho(B'', A)$ so $\rho(B', A) = \rho(B', B'') > 1$; hence $\rho(B', B'') \geq R$ and so $\rho(B', A) \geq R$. This implies that $\mathrm{MR}(B') \geq R \cdot \mathrm{MR}(A)$. Now choose ε small enough that $R(M - \varepsilon) > M$. Then $\mathrm{MR}(B') > M$ which contradicts the definition of M, thus $\phi|A$ is measure linear. q.e.d.

As an example of this result, a slide map is measure preserving, thus measure linear everywhere. The theorem can be rephrased by noting that the clone, A, is just a linearly scaled copy of the entire Cantor set C, so that composing $\phi|A$ with a linear map from C onto A gives a measure linear quasi-isometry of C into, but not necessarily onto, C':

Corollary. *If C and C' are quasi-isometric affinely generated Cantor sets, then there is a quasi-isometry $\phi\colon C \to C'$ which is measure linear but not necessarily surjective. The image of ϕ is a co-gap.*

Remark. This result may be strengthened slightly as follows. The image of ϕ is a co-gap, thus a union of clones of C'. Let B be the left-most clone in $\phi(C)$. Then $\phi^{-1}|B$ composed with the affine map of C' onto B is a measure linear quasi-isometry of C' into C with $\phi^{-1}(0) = 0$ (recall, 0 is the left most point of both C and C').

Definition. The *gap-invariant*, $G(C)$, of a Cantor set, C, is the countable subset of $[0,1]$: $G(C) = \{\mu_C([0,x])\colon x \text{ lies in a gap of } C\}$.

The remark after the corollary implies that if $C \simeq C'$ are affinely generated Cantor sets, then there is a linear map $\theta\colon \mathbb{R} \to \mathbb{R}$, $\theta(z) = \alpha z$ for some fixed $\alpha \in (0,1]$, and $\theta(G(C)) = [0,\alpha] \cap G(C')$. In §3 for the generic case it is shown that $\alpha = 1$ so that $G(C)$ is actually an invariant of quasi-isometry. In general, however, we can only say it is invariant in the above sense.

Proposition (2.2). *Let B be a co-gap of an affinely generated Cantor set C. Then $B \simeq C$. Equivalently, if $m \in \mathbb{N}$, and let mC denote the union of m affinely isomorphic copies of C contained in pairwise disjoint intervals, then $C \simeq mC$.*

Proof. The idea is that $C = \infty C \simeq nC$. Let $J_1, \cdots, J_q$ be the level-1 clones of C, and define

$$A = (J_1 \cup \cdots \cup J_{q-1}) \cup \bigcup_{n=1}^{\infty} \tau_q^{-n}(J_1 \cup \cdots \cup J_{q-1}).$$

Then A is a union of disjoint clones, with $(q-1)$ level-n clones for each $n \geq 1$. Label these clones from left to right as $A_1, A_2, \cdots$; thus $A_i = J_i$ for

$1 \leq i \leq q-1$, and

$$A_{[(q-1)n+i]} = \tau_q^{-n}(J_i) \quad \text{for } n \geq 1,\ 1 \leq i \leq q-1.$$

The co-gap B is a union of finitely many disjoint clones of C, call these clones $B_1, \cdots, B_m$. Define $A'_i = B_i$ for $1 \leq i \leq m-1$, and now decompose B_m as above. Formally, let $\theta\colon B_m \to C$ be the clone map, and define $A'_{i+m-1} = \theta^{-1}(A_i)$ for $i \geq 1$. Define $\phi\colon B \to C$ by $\phi|A'_i$ maps A'_i affinely onto A_i. Extend ϕ over the remaining gaps of B affinely. We show that ϕ is the required quasi-isometry. Let J be the smallest interval containing C and set $D = \overline{J - J_q}$, then D is a fundamental domain for the action of τ_q^{-1} on J. Now $\theta^{-1} \circ \tau_q^{-1} \circ \theta$ acts on B_m and has fundamental domain $\theta^{-1}(D)$ on which ϕ is piecewise linear with finitely many changes of derivative. From the definition of ϕ, we obtain $\phi^{-1} \circ \tau_q^{-1} \circ \phi|B_m = \theta^{-1} \circ \tau_q^{-1} \circ \theta$ and so ϕ takes on this same finite collection of derivatives on each fundamental domain, thus $\phi|B_m$ is a quasi-isometry. But on $B - B_m$, ϕ is piecewise linear with only finitely many changes of derivative, therefore ϕ is a quasi-isometry.

Corollary. *If C and C' are affinely generated Cantor sets and $C \simeq C'$, then there is a quasi-isometry ψ mapping C onto C' which is piecewise measure linear. This means ψ is measure linear on each of a countable set of intervals whose union contains C.*

Proof. By the corollary to (2.1) there is a measure linear quasi-isometry $\theta\colon C' \to B$, with B a co-gap of C. The map $\phi\colon B \to C$ constructed in (2.2) is piecewise linear, so $\phi \circ \theta\colon C' \to C$ is piecewise measure linear.

Definition. A (*local*) *similarity* of a Cantor set C is an orientation preserving homeomorphism σ defined on an interval $U \subset J$ with $C \cap \operatorname{Int}(U) \neq \varnothing$ and such that

(i) $\sigma(U \cap C) \subset C$.

(ii) σ has a unique fixed point x, and $x \in C$.

(iii) σ is measure linear.

(iv) Either σ or σ^{-1} (if σ is expanding) is BD.

A map σ is BD, or has *bounded distortion*, if the family $\{\sigma^n\}_{n\in\mathbb{N}}$ is uniformly quasi-linear, i.e. there is a constant, K, and affine maps $\{L_n\}_{n\in\mathbb{N}}$ of $\mathbb{R}^1$ such that each of the maps $\sigma^n \circ L_n$ is a K-quasi-isometry. It is well known that C^2 maps satisfying the Renyi condition are BD (see [2] for example). A (*local*) *similarity point* of C is a fixed point of a local similarity. A point x in C is *eventually periodic for τ* if there are integers m, $p > 0$ with $\tau^{m+p}(x) = \tau^m(x)$. The following result establishes a link between the dynamics of τ and the geometry of C.

Theorem (2.3). *Suppose C is an affinely generated Cantor set generated by τ. Then a point x in C is a local similarity point if and only if x is an eventually periodic point of τ.*

Proof. Let σ be a local similarity with fixed point x and choose a clone $A \subset \mathrm{domain}(\sigma)$ with $x \in A$. If $d\sigma/d\mu_C > 1$, replace σ by σ^{-1}, so that we may assume σ is contracting. Thus $\sigma A \subset A$. Set $B_r = \sigma^r A$ and define A_r to be the smallest clone containing B_r. Let L be the gap adjacent to A on the left and R the gap adjacent to A on the right. Since σ is BD, $|\sigma^r L| \cdot |A_r|^{-1}$ and $|\sigma^r R| \cdot |A_r|^{-1}$ are bounded below independently of r. Then, as in the proof of (2.1), there are only finitely many possibilities for the way B_r is contained in A_r. More formally, let $\psi_r \colon A_r \to C$ be the clone map taking A_r onto C. Then there are only finitely many possibilities for $\psi_r(B_r)$. Thus, for some $r, s > 0$, $\psi_r(B_r) = \psi_{r+s}(B_{r+s})$. Define $\lambda = \psi_{r+s}^{-1} \circ \psi_r$; then $\lambda B_r = B_{r+s}$. Now λ and σ^s are both measure linear and both map B_r onto B_{r+s}, therefore $\lambda|(B_r \cap C) = \sigma^s|(B_r \cap C)$. Let λ_1 be the clone map $\psi_r(A_{r+s}) \to C$; then $\lambda = \psi_r^{-1} \circ \lambda_1^{-1} \circ \psi_r$. Let $m = \mathrm{level}(A_r)$ and $p = \mathrm{level}(A_{r+s}) - m$; then $\psi_r = \tau^m|A_r$ and $\psi_{r+s} = \tau^{m+p}|A_{r+s}$. Now $x = \sigma^s x = \lambda x$ and so

$$\tau^{m+p}x = \tau^{m+p} \circ \lambda x = \tau^{m+p} \circ \psi_{r+s}^{-1} \circ \psi_r x = \tau^m x.$$

Hence x is an eventually periodic point of τ with period p. The converse is obvious. q.e.d.

It follows from the proof of (2.3) that:

Corollary. *If σ is a local similarity of a Cantor set C with fixed point x, then there are a clone A containing x, and clone maps λ and ψ, and an integer $n > 0$ with $\sigma^n|(A \cap C) = \psi^{-1} \circ \lambda^{-1} \circ \psi|(A \cap C)$.*

This corollary says that given a local similarity map, then, modulo taking a suitable iterate and restricting to a smaller domain, it is essentially linear, i.e., linear on C.

Definition. The *local similarity group* at a point x of C is the group of germs of local similarity maps fixing x; the maps are considered to be defined only on a subset of C. This group is written SG_x. Because a similarity map is measure linear, and two germs represent the same element of SG_x if they agree on C, the map $\lambda\colon \mathrm{SG}_x \to \mathbb{R}$ given by $\lambda(\sigma) = d\sigma/d\mu_C$ for $\sigma \in \mathrm{SG}_x$ is injective. Thus we may regard SG_x as a subgroup of $\mathbb{R}^*$, the group of real numbers under multiplication.

Proposition (2.4). *If C is an affinely generated Cantor set, then SG_x is a discrete subgroup of $\mathbb{R}^*$.*

Proof. The subgroup L of SG_x consisting of germs of affine maps of $\mathbb{R}$ is nontrivial by the corollary to (2.3). It is easy to check that L is discrete and therefore cyclic. It follows from the corollary to (2.3) that if $\sigma \in \mathrm{SG}_x$ then

$\sigma^n \in L$ for some $n \geq 1$. Therefore if SG_x is not discrete, there is $\sigma' \in \mathrm{SG}_x$ which is infinitely divisible, i.e., $\infty = \sup\{n \in \mathbb{N};\, \sigma' = \sigma_1^n$ for some $\sigma_1 \in \mathrm{SG}_x\}$. Let A be a clone containing x with $A \subset \mathrm{domain}(\sigma')$, and let $\psi\colon A \to C$ be the corresponding clone map. Then $y = \psi x$ is a similarity point, and ψ induces an isomorphism $\psi_*\colon \mathrm{SG}_x \to \mathrm{SG}_y$. Therefore it suffices to show SG_y has no infinitely divisible elements. Let $\sigma = \psi_*(\sigma') = \psi \circ \sigma' \circ \psi^{-1}$; then σ is infinitely divisible. By replacing σ by σ^{-1} if needed, we may assume $d\sigma/d\mu_c < 1$. Given $\sigma_1 \in \mathrm{SG}_y$, then $\sigma_1 = (\sigma^{-k} \circ \sigma_1 \circ \sigma^k)|\,\mathrm{domain}(\sigma_1)$ because SG_y is abelian. This shows, by choosing k large, that σ_1 can be extended measure linearly over $\mathrm{domain}(\sigma) = C$. Thus every element of SG_y induces a linear map $G(C) \to G(C)$ fixing the measure coordinate of y. A level-p clone of C has mass $m_{i_1} m_{i_2} \cdots m_{i_p}$ for some choice of $i_1, \cdots, i_p \in \{1, \cdots, q\}$ where $m_1, \cdots, m_q$ are the masses of the level-1 clones of C. It follows that $G(C) \subset \Lambda$ where $\Lambda = \mathbb{Q}(m_1, \cdots, m_q)$. Let $m = \mu_C([0, y])$. Since $\sigma \in \mathrm{SG}_y$, there must be a gap with measure coordinate $m(1-\alpha)$ where $\alpha = d\sigma/d\mu_C$. If $\sigma = \sigma_1^n$, $n \geq 1$, then because σ_1 is defined on all of C, there must be a gap with measure coordinate $m(1-\alpha^{1/n})$. Hence $m(1-\alpha^{1/n}) \in \Lambda$, and so $\alpha^{1/n} \in \Lambda(m) \equiv K$. Thus α is an infinitely divisible element of the multiplicative group of K^* of nonzero elements of $\Lambda(m)$. Now K is a finitely generated field extension of Q, and it is well known that the only infinitely divisible elements of such a K^* are roots of unity. But $0 < \alpha < 1$, so α is not a root of unity, therefore α is not infinitely divisible—a contradiction.

Definition. If x is a local similarity point of C, the *local scale factor* at x is

$$S(x) = \max\{d\sigma/d\mu_C < 1 : \sigma \in \mathrm{SG}_x\}.$$

Corollary 1. *Suppose C and C' are affinely generated Cantor sets, and $\phi\colon C \to C'$ is a measure linear quasi-isometry, but not necessarily surjective. If x is a local similarity point of C, then ϕx is a local similarity point of C' and $S(x) = S(\phi x)$.*

Proof. Suppose $\sigma \in \mathrm{SG}_x$; then $\sigma' = \phi \circ \sigma \circ \phi^{-1}$ is BD because σ is BD and ϕ is a quasi-isometry. σ' is measure linear because σ and ϕ are measure linear, in fact $d\sigma'/d\mu_{C'} = d\sigma/d\mu_C$. The map $\phi_*\colon \mathrm{SG}_x \to \mathrm{SG}_{\phi x}$ given by $\phi_*(\sigma) = \sigma'$ is an isomorphism with inverse $(\phi^{-1})_*$. Thus $S(x) = S(\phi x)$.

Definition. The *spectrum* of C is

$$\mathrm{spec}(C) = \{S(x)\colon x \text{ is a similarity point of } C\}.$$

Corollary 2. *Suppose C and C' are affinely generated Cantor sets, and $C \simeq C'$. Then* $\mathrm{spec}(C) = \mathrm{spec}(C')$.

3. Classification in the generic case

Let C^* denote the set of all affinely generated Cantor sets of unit diameter contained in $[0,1]$. C^* may be made into a metric space using Gromov's definition [1]; thus if $C_1, C_2 \in C^*$, distance$(C_1, C_2) = \inf\{\varepsilon > 0 | \exists$ an ε-isometry ρ: $C_1 \to C_2\}$ (recall ρ is an *ε-isometry* if ρ is a homeomorphism and $\forall x, y \in C_1$, $|\,\text{distance}(x,y) - \text{distance}(\rho x, \rho y)| < \varepsilon$).

Let $\mathscr{C}$ be the set of quasi-isometry classes of C^*, and give $\mathscr{C}$ the quotient topology. A collection of q disjoint intervals $J_1, \cdots, J_q$ contained in $[0,1]$ determines an element of C^*; since sliding produces quasi-isometric Cantor sets, the q-tuple $(|J_1|, \cdots, |J_q|)$ determines an element of $\mathscr{C}$. Let

$$M_q = \left\{ (l_1, \cdots, l_q) \middle| \sum_{i=1}^{q} l_i < 1,\ l_i > 0 \right\}.$$

Then there is a continuous map $\pi_q \colon M_q \to \mathscr{C}$ which maps a q-tuple to the quasi-isometry class determined by intervals $J_1, \cdots, J_q$ with $|J_i| = l_i$. M_q is topologized as a subspace of $\mathbb{R}^q$. Let $\sim$ be the equivalence relation on $\bigcup_{q=2}^{\infty} M_q$ generated by splitting, and let $\mathscr{M} = (\bigcup_{q=2}^{\infty} M_q)/\sim$; $\mathscr{M}$ is given the weak topology. Then π_q induces a continuous map $\pi \colon \mathscr{M} \to \mathscr{C}$. We show below that $\pi | M_q$ is injective on a subset of M_q of full Lebesgue measure. In §4 we show that π is not injective.

Definition. Let C be an affinely generated Cantor set with level-1 clones $J_1, \cdots, J_q$ of lengths $l_1, \cdots, l_q$, and let $m_i = \mu_C(J_i)$. The *mass ratios* are $\lambda_i = m_i m_1^{-1}$. An affinely generated Cantor set C' is *generic* if $C' \simeq C$, where C is as above and $\{\lambda_2, \cdots, \lambda_q\}$ are algebraically independent over $\mathbb{Q}$. We remark that this definition can be broadened, it is only necessary to require that certain specific polynomials do not vanish for $\{\lambda_2, \cdots, \lambda_q\}$.

Theorem (3.1). *If $\mathbf{x}, \mathbf{y} \in \mathscr{M}$, and $\pi(\mathbf{x})$ is generic, then $\pi\mathbf{x} = \pi\mathbf{y} \Leftrightarrow \mathbf{x} = \mathbf{y} \Leftrightarrow G(\pi\mathbf{x}) = G(\pi\mathbf{y})$, and $\pi\mathbf{x}$ and $\pi\mathbf{y}$ have the same Hausdorff dimension.*

The first equivalence may be restated in a slightly stronger form:

Theorem (3.2). *Suppose that C and C' are affinely generated Cantor sets and $C \simeq C'$, and also that the level-1 clones $J_1, \cdots, J_q$ of C satisfy the condition above on algebraically independent mass ratios. If $J'_1, \cdots, J'_r$ are the level-1 clones of C', then $J'_1, \cdots, J'_r$ are obtained from $J_1, \cdots, J_q$ by splitting and sliding. In particular, $r \geq q$.*

We now introduce a standing hypothesis: (†) The clone structure on C has level-1 clones $J_1, \cdots, J_q$ with $\{\lambda_2, \cdots, \lambda_q\}$ algebraically independent over $\mathbb{Q}$.

Lemma 1. *Suppose C satisfies (†), A is a co-gap of C, B is a clone of C, and $\mu_C(A) = \mu_C(B)$. Then A is a clone of C.*

Proof. For p large enough both A and B are a union of level-p clones. The mass of a level-p clone D is $\mu_C(D) = m_{i_1} m_{i_2} \cdots m_{i_p}$ for some choice of $i_j \in \{1, \cdots, q\}$. μ_C is a probability measure so $\sum_{i=1}^{q} m_i = 1$ and hence $m_1^{-1} = 1 + \lambda_2 + \cdots + \lambda_q$. Thus

$$\mu_c(D) = \lambda_2^{n_2}\lambda_3^{n_3}\cdots\lambda_q^{n_q}(1+\lambda_2+\cdots+\lambda_q)^{-p}$$

for suitable integers $n_i \geq 0$.

The equation $\mu_C(A) = \mu_C(B)$ becomes, on multiplying by m_1^{-p},

$$\sum_{\substack{D\subset A\\ D \text{ a level}-p \text{ clone}}} \lambda_2^{n_2(D)}\lambda_3^{n_3(D)}\cdots\lambda_q^{n_q(D)} = \sum_{\substack{D\subset B\\ D \text{ a level}-p \text{ clone}}} \lambda_2^{n_2(D)}\lambda_3^{n_3(D)}\cdots\lambda_q^{n_q(D)}. \tag{1}$$

The algebraic independence of the λ_i's over $\mathbb{Q}$ implies that there must be the same number of terms on both sides of (1). Thus A and B contain the same number of level-p clones. If $r = \text{level}(B)$, then B contains q^{p-r} level-p clones. Thus A is contained in the union of at most two adjacent level-r clones, and is contained in a single level-r clone if and only if A is a level-r clone. So suppose that A meets two adjacent level-r clones J' and J''. Let n' and n'' be the number of level-p clones in $J' \cap A$ and $J'' \cap A$ respectively, so $n' + n'' = q^{p-r}$ and $n', n'' > 0$. Set $m' = \mu_C(J')$ and $m'' = \mu_C(J'')$. Then

$$\mu_C(A) = m'l' + m''l'',$$

where $l' = \mu_C(A\cap J')/\mu_C(J')$ and $l'' = \mu_C(A\cap J'')/\mu_C(J'')$. Now l' =(the n' rightmost level-$(p-r)$ clones in C) and $l'' = \mu_C$(the n'' leftmost level-$(p-r)$ clones in C) because there are clone maps taking $J' \to C$ and $J'' \to C$. Now $n' + n'' = q^{p-r}$, therefore $l' + l'' = 1$ and so equation (1) can be rewritten as

$$(m_1^{-r}\mu_C(J'))(m_1^{r-p}l') + (m_1^{-r}\mu_C(J''))(m_1^{r-p}l'') = (m_1^{-r}\mu_C(B))(m_1^{r-p}).$$

Each bracketed term is a polynomial in the λ_i's. Define

$$\begin{aligned}
p_1 &= m_1^{-r}\mu_C(J') = \lambda_2^{c_2}\cdots\lambda_q^{c_q} \quad \text{for some } c_i \geq 0,\\
p_2 &= m_1^{-r}\mu_C(J'') = \lambda_2^{d_2}\cdots\lambda_q^{d_q} \quad \text{for some } d_i \geq 0,\\
p_3 &= m_1^{-r}\mu_C(B) = \lambda_2^{e_2}\cdots\lambda_q^{e_q} \quad \text{for some } e_i \geq 0,\\
p_4 &= m_1^{r-p} = (1+\lambda_2+\cdots+\lambda_q)^{p-r},\\
p_5 &= m_1^{r-p}l''.
\end{aligned}$$

Using $m_1^{r-p}(l' + l'') = p_4$ to rewrite the above equation we obtain

$$p_1(p_4 - p_5) + p_2p_5 = p_3p_4 \quad \text{or} \quad p_4(p_1 - p_3) = p_5(p_1 - p_2).$$

Now p_4 and $(p_1 - p_2)$ have no common factor, therefore p_4 divides p_5. Evaluating p_4 and p_5 at $\lambda_2 = \lambda_3 = \cdots = \lambda_q = 1$ gives $p_4(1, \cdots, 1) = q^{p-r}$

and $p_5(1,\cdots,1)=n''$. Thus q^{p-r} divides n'', but $0<n''<q^{p-r}$, which is a contradiction.

Lemma 2. *Suppose C satisfies (†) and G is a gap in C. Let A_L be any clone adjacent on the left to G, and A_R any clone adjacent on the right to G. Then $\mu_C(A_L)/\mu_C(A_R)$ determines the name of G. In other words, if G' is another gap of C, A'_L and A'_R are clones adjacent to G', and $\mu_C(A'_L)/\mu_C(A'_R)=\mu_C(A_L)/\mu_C(A_R)$, then G and G' have the same name.*

Remark. Nothing is assumed about the level of A_L and A_R; in particular they can be at different levels.

Proof. Let $i\in\{1,\cdots,q-1\}$ be the name of G, and suppose initially that G, A_L and A_R are all at the same level, p say. Then $\mu_C(A_R)/\mu_C(A_L)=\lambda_{i+1}/\lambda_i$ and the algebraic independence of the λ_i's implies that if $\lambda_{j+1}/\lambda_j=\lambda_{i+1}/\lambda_i$, then $i=j$. Proceeding with the general case,

$$\mu_C(A_R)/\mu_C(A_L)=\lambda_{i+1}m_1^r/\lambda_i m_q^s,$$

where $r,s\geq 0$ are defined by $\text{level}(A_L)=s+\text{level}(G)$ and $\text{level}(A_R)=r+\text{level}(G)$. Thus

$$\mu_C(A_R)/\mu_C(A_L)=\lambda_{i+1}\lambda_i^{-1}\lambda_q^{-s}(1+\lambda_2+\cdots+\lambda_q)^{s-r},$$

and again it is clear that if this equals $\lambda_{j+1}\lambda_j^{-1}\lambda_q^{-s'}(1+\lambda_2+\cdots+\lambda_q)^{s'-r'}$ then $i=j$. q.e.d.

Lemma 3. *Suppose the same hypotheses as in Lemma 2. Let A be the smallest clone containing G, and A' the smallest clone containing G'. Let $\phi\colon A\to C$ and $\phi'\colon A'\to C$ be the corresponding clone maps and set $\psi=(\phi')^{-1}\circ\phi$. Then $\psi(G)=G'$, $\psi(A_R)=A'_R$ and $\psi(A_L)=A'_L$.*

Proof. G has level-1 relative to A, G' has level-1 relative to A' and, by Lemma 2, G and G' have the same name, therefore $\psi(G)=G'$. From the proof of Lemma 2,

$$\frac{\mu_C(A_R)}{\mu_C(A_L)}=\frac{\mu_C(A'_R)}{\mu_C(A'_L)}=\lambda_{i+1}\lambda_i^{-1}\lambda_q^{-s}(1+\lambda_2+\cdots+\lambda_q)^{s-r},$$

where

$$\begin{aligned}
\text{level}(A_L)&=\text{level}(G)+s=\text{level}(A)+s+1,\\
\text{level}(A'_L)&=\text{level}(G')+s=\text{level}(A')+s+1,\\
\text{level}(A_R)&=\text{level}(G)+r=\text{level}(A)+r+1,\\
\text{level}(A'_R)&=\text{level}(G')+r=\text{level}(A')+r+1.
\end{aligned}$$

This implies $\psi(A_L)=A'_L$ and $\psi(A_R)=A'_R$ because a clone adjacent on a specified side to a specified gap is completely determined by its level. But the levels of A_L and A_R relative to A are the same as the levels of A'_L and A'_R relative to A', by the above.

Lemma 4. *Suppose C satisfies (†). Let U be a co-gap in C, $\sigma\colon U \to U$ a local similarity, and A the smallest clone containing U. Then there is a linear map $\psi\colon A \to A$ such that $\psi|U \cap C = \sigma|U \cap C$. Furthermore, ψ^{-1} is the conjugate of one clone map by another clone map.*

Proof. By the corollary to (2.3), there are an $n > 0$ and a clone $A' \subset U$, containing the fixed point x of σ, with $\sigma^n|A' \cap C$ linear. Define $\alpha = d\sigma/d\mu_C$. Then the Radon-Nikodým derivative of σ^n is $\alpha^n = \mu_C(B)/\mu_C(A')$ where $B = \sigma^n A'$ is a clone. Set $p = \text{level}(B) - \text{level}(A')$; then $\mu_C(B)/\mu_C(A') = \mu_C(B')$ for some level-p clone B'. Let $\Lambda = \mathbb{Q}(\lambda_2, \cdots, \lambda_q)$; then

$$\alpha^n = \mu_C(B') = \lambda_2^{n_2}\lambda_3^{n_3}\cdots\lambda_q^{n_q}(1 + \lambda_2 + \cdots + \lambda_q)^{-p}$$

for some $n_2, \cdots, n_q \geq 0$. But $\sigma A'$ is a co-gap and $\mu_C(\sigma A') = \alpha\mu_C(A') \in \Lambda$, hence

$$\alpha = \sqrt[n]{[\lambda_2^{n_2}\cdots\lambda_q^{n_q}(1 + \lambda_2 + \cdots + \lambda_q)^{-p}]} \in \Lambda.$$

This implies $n|n_i$ for each $2 \leq i \leq q$ and also $n|p$, since the λ_i's are algebraically independent over $\mathbb{Q}$. Write $n'_i = n_i n^{-1}$ and $p' = pn^{-1}$. Then

$$\mu_C(\sigma A') = \lambda_2^{n'_2}\cdots\lambda_q^{n'_q}(1 + \lambda_2 + \cdots + \lambda_q)^{-p'}\mu_C(A'),$$

and so $\sigma A'$ has the same mass as a clone, for example the clone $(\tau_1^{n'_1}\cdots\tau_q^{n'_q})^{-1}(A')$ where $n'_1 = p' - \sum_{i=2}^{q} n'_i$. By Lemma 1, $\sigma A'$ is a clone. Hence the restriction of σ to $(A' \cap C)$ is affine. Let $\psi\colon A' \to \sigma A'$ be the affine orientation preserving surjection. It remains to show that the affine extension of ψ over A maps C into itself. Because $A' \not\supset U$, there is a clone, E say, with $E \subset U - A'$ and E may be chosen to be adjacent to a gap G say, with G adjacent to A'. Either E is on the left or right of A'; assume without loss that E is on the left of A'. Now $F = \sigma E$ is a co-gap and has mass $\mu_C(F) = \alpha\mu_C(E)$ which is the same mass as a certain clone, therefore by Lemma 1, F is a clone. Furthermore $\mu_C(E)/\mu_C(A') = \mu_C(\sigma E)/\mu_C(\sigma A')$ because σ is measure linear. So by Lemma 3, there is an affine map $\psi''\colon A'' \to \psi''(A'')$ with $\psi''(A'' \cap C) \subset C$, where A'' is the smallest clone containing G. Since now A'' contains A' and E, ψ'' is an affine extension of ψ so that $\text{level}(A'') \leq \text{level}(A') - 1$. Repeating this extension process a finite number of times gives $\text{level}(A'') \leq \text{level}(A)$ and therefore the required extension of ψ over A. ψA is a clone because it has the correct mass to be a clone. Let $\psi_1\colon A \to C$ be the clone map corresponding to A, and $\psi_2\colon \psi_1(\psi A) \to C$ the clone map corresponding to $\psi_1(\psi A)$. Then $\psi = \psi_1^{-1} \circ \psi_2^{-1} \circ \psi_1$. q.e.d.

We are now ready to prove Theorem (3.2): the idea is to conjugate the dynamical system producing C' over to C by using the quasi-isometry, and then linearize this dynamical system. Let $\phi\colon C' \to C$ be the given quasi-isometry, by the corollary to (2.1) there is a measure linear quasi-isometry, which we again call ϕ, $\phi\colon C' \to U$ where U is a co-gap of C. We may assume that the smallest clone containing U is C, by composing ϕ with a clone map of C if needed. Let τ_i', $1 \le i \le r$, be the level-1 clone maps of C, and $\psi_i' = (\tau_i')^{-1}$, and let $J_i' = \psi_i'([0,1])$ be the level-1 clones of C'. Set $\psi_i'' = \phi \circ \psi_i' \circ \phi^{-1}$ and $J_i'' = \phi(J_i')$, then $\psi_i''\colon U \to J_i''$ is BD and measure linear. By Lemma 4, there are affine maps $\psi_i'''\colon J \to \psi_i'''(J) \equiv J_i'''$ where $J = [0,1]$ is the level-0 clone of C, and $\psi_i'''|C \cap U = \psi_i''|C \cap U$, and $(\psi_i''')^{-1}$ is the conjugate of one clone map of C by another clone map of C. Hence J_i''' is a clone of C. Suppose $U \neq J$. Let x be the fixed point of ψ_1''; then x is the leftmost point of J_1''. Because U is a clopen, x is isolated from the left in C. If x is not the leftmost point of J, then x is an interior point of J_1''', and is the fixed point of ψ_1''', so x cannot be isolated from the left in C, a contradiction. Thus the leftmost point of U and of J are the same, similarly for the rightmost points, therefore $U = J$, $J_i'' = J_i'''$, and ϕ is measure linear and surjective. This implies $G(C) = G(C')$. Now $J_1'', \cdots, J_r''$ are disjoint clones of C, which contain C, and so are obtained by splitting the level-1 clones of C. Since ϕ is measure preserving, $\mu_{C'}(J_i') = \mu_C(J_i'')$ so $|J_i''| = |J_i'|$, and so the J_i' are obtained by sliding the J_i''. This proves (3.2) and part of (3.1). It only remains to prove that if $G(C) = G(C')$, C and C' have the same Hausdorff dimension, and C satisfies (†), then $C \simeq C'$. We would like to thank Dennis Sullivan and Curt McMullen for help with the following argument.

Let ψ be a clone map of C'. Then ψ induces an affine map $\psi_*^{-1}\colon G(C') \to G(C')$ with derivative $\alpha^{-1} = d\psi^{-1}/d\mu_{C'}$. Since $G(C) = G(C')$, $\psi_*^{-1}\colon G(C) \to G(C)$. Now $\overline{\psi_*^{-1}[0,1]} = [a_1, a_2]$ where a_1, $a_2 \in G(C)$, so there is a co-gap A of C with $\{\mu_C[0,x]\colon x \in A\} = \overline{\psi_*^{-1}[0,1]}$ and we will show that A is a clone of C. Let U be the multiplicative group generated by $m_1, \lambda_2, \cdots, \lambda_q$. Then $G(C) \subset \mathbb{Z}[U] \equiv \Lambda$, and $\alpha^{-1} = a_2 - a_1 \equiv p_1 \in \Lambda$. But U is the group of units of Λ, and $\mu_C(B) \in U$ if B is any clone of C. Let B be a clone of C such that

$$X = \{\mu_C[0,x]\colon x \in B\} \subset \overline{\psi_*^{-1}[0,1]}.$$

Then $X = [b_1, b_2]$ for some b_1, $b_2 \in G(C)$. Write $\overline{\psi_* X} = [c_1, c_2]$; then $c_2 - c_1 \equiv p_2 \in \Lambda$ and $b_2 - b_1 \equiv p_3 \in U$, so $\alpha = p_2 p_3^{-1}$. Thus $p_1^{-1} = p_2 p_3^{-1}$, so $p_3 = p_1 p_2 \in U$, which implies $p_1 \in U$. For r large enough, A is a union of level-r clones of C, thus

$$p_1 = \sum_{\substack{\text{level}-r \\ \text{clones } D \subset A}} m_1^r \lambda_2^{n_2(D)} \cdots \lambda_q^{n_q(D)} \in U.$$

Observe that $r \geq n_2(D) + \cdots + n_q(D)$ for each $D \subset A$, so that for some n_1', $n_2, \cdots, n_q$, $p_1 = m_1^{n_1'} \lambda_2^{n_2} \cdots \lambda_q^{n_q}$ where $n_1 \equiv n_1' - (n_2 + \cdots + n_q) \geq 0$. Thus $p_1 = m_1^{n_1} m_2^{n_2} \cdots m_q^{n_q}$, and there is a clone E of C with $\mu_C(E) = p_1$. Hence A is a clone by Lemma 1. Let $J_1', \cdots, J_r'$ be the level-1 clones of C'. Then the above shows that there are clones $J_1'', \cdots, J_r''$ of C such that J_i' and J_i'' have the same mass coordinates, and therefore the same length. Now $J_1'', \cdots, J_r''$ are obtained by splitting the level-1 clones of C, and $J_1', \cdots, J_r'$ by sliding $J_1'', \cdots, J_r''$. Since splitting and sliding produce quasi-isometric Cantor sets, $C \simeq C'$. This completes the proof of (3.1) and (3.2).

Remark. The proof shows that a measure linear map $\phi\colon C' \to C$, which is a quasi-isometry onto its image, has image a clone of C, and is the slide map for some splitting of the clones of C, when C is generic.

4. A nongeneric example

In this section we calculate which affinely generated Cantor sets are quasi-isometric to an affinely generated Cantor set determined by two level-1 clones of equal mass, hence equal length. This includes Cantor's middle third set, and in fact there is precisely one example for each Hausdorff dimension $d \in (0, 1)$. The classification is different from the generic case; there are additional quasi-isometries arising from the symmetry. In the language of §3, $\pi\colon \mathscr{M} \to \mathscr{C}$ is not injective.

Theorem (4.1). *Let C be an affinely generated Cantor set determined by two level-1 clones each of mass $1/2$. Let C' be another affinely generated Cantor set with the same Hausdorff dimension as C. The following statements are equivalent:*

(1) $C \simeq C'$.

(2) $G(C) = G(C')$.

(3) *Every level-1 clone of C' has mass an integral power of $1/2$, and $1/2 \in G(C')$.*

Example. Fix a Hausdorff dimension d, and recall the relationship between mass and length of clones given in §1. It follows that an affinely generated Cantor set is determined up to quasi-isometry by the Hausdorff dimension, and the masses of the level-1 clones taken in order from left to right on the line. Suppressing the Hausdorff dimension and denoting such a Cantor set by the n-tuple of level-1 clone masses, the above theorem gives:

(i) $\left(\frac{1}{2}, \frac{1}{2}\right) \simeq \left(\frac{1}{8}, \frac{1}{4}, \frac{1}{8}, \frac{1}{2}\right)$,
(ii) $\left(\frac{1}{2}, \frac{1}{2}\right) \not\simeq \left(\frac{1}{4}, \frac{1}{2}, \frac{1}{4}\right)$ because $\frac{1}{2} \notin G\left(\frac{1}{4}, \frac{1}{2}, \frac{1}{4}\right)$.

The Cantor sets in (i) are not equivalent under splitting and sliding. To prove this, observe that sliding may be ignored using this notation. Given a q-tuple $\boldsymbol{\alpha} = (2^{-n_1}, \cdots, 2^{-n_q})$ determining an affinely generated Cantor set C, let $m = \min\{2^{-n_1}, \cdots, 2^{-n_q}\}$ and define $M(\boldsymbol{\alpha})$ to be 0 if every maximal consecutive sequence $2^{-n_i}, 2^{-n_{i+1}}, 2^{-n_{i+2}}, \cdots, 2^{-n_{i+j}}$, each term of which is equal to m, has an even number of terms in it, i.e. $j \equiv 1 \bmod 2$. Otherwise define $M(\boldsymbol{\alpha}) = 1$. If an r-tuple $\boldsymbol{\beta}$ is obtained by splitting $\boldsymbol{\alpha}$, then it follows by a simple combinatorial argument that $M(\boldsymbol{\alpha}) = M(\boldsymbol{\beta})$. Now $M(\frac{1}{2}, \frac{1}{2}) = 0$ and $M(\frac{1}{8}, \frac{1}{4}, \frac{1}{8}, \frac{1}{2}) = 1$, therefore $(\frac{1}{2}, \frac{1}{2})$ and $(\frac{1}{8}, \frac{1}{4}, \frac{1}{8}, \frac{1}{2})$ are not equivalent under splitting.

The level-p clones in a Cantor set are obtained by splitting once each level-$(p-1)$ clone. However, since the clone structure is not unique, this produces a rather arbitrary decomposition. A geometrically more natural procedure is to split the largest clones. This gives rise to an operation called *rolling*, which for simplicity we will describe only for those Cantor sets all of whose clones have mass a power of $1/2$.

Definition. Let C be an affinely generated Cantor set with level-1 clones $J_1, \cdots, J_q$, and suppose that $\mu_C(J_i) = 2^{-n_i}$ for some $n_i \in \mathbb{N}$. The *roll-1 clones* of C are $J_1, \cdots, J_q$. The *roll-n clones* of C are obtained from the roll-$(n-1)$ clones by splitting once those roll-$(n-1)$ clones with mass 2^{1-n}. Let $M = \max\{n_1, \cdots, n_q\} - 1$; then the mass of every roll-n clone is 2^{-p} for some $p \in \mathbb{N}$ depending on the clone, with $n \leq p < n + M$. This is proved by induction on n. A gap, G, of C is a *roll-n gap* if it first appears at roll-n, i.e., G is a complementary component of the roll-n clones, but not of the roll-$(n-1)$ clones. A roll-n gap G is contained in a unique roll-$(n-1)$ clone, A, of mass 2^{1-n}, and G has level-1 relative to A. If d is the Hausdorff dimension of C and $\alpha^d = 2$, then $|A| = \alpha^{1-n}$. Thus $L \leq |G|\alpha^{n-1} \leq 1$ where $L = \min\{|G| \colon G$ is a level-1 gap of $C\}$. Thus every roll-n gap has length within a bound multiple, independent of the gap or n, of α^{-n}.

Proof of (4.1). (3) $\Rightarrow$ (1) Define $\phi \colon C' \to C$ to be the unique measure preserving continuous map. Extend ϕ affinely across the gaps of C'; we will show that ϕ is a quasi-isometry. Let G' be a roll-n gap of C' and set $G = \phi(G')$. Then the mass coordinate of G' and of G is $p2^{-(n+M)}$ for some $p \in \mathbb{N}$. This is because the mass of a roll-n clone is an integral multiple of $2^{-(n+M)}$. It follows that G is a roll-k gap of C for some $k < n + M$. Hence $|G'| \cdot |G|^{-1} < \alpha^{-n}/L\alpha^{-(n+M)} = \alpha^M L^{-1}$ where L is the length of the unique level-1 gap of C. Thus ϕ^{-1} is Lipschitz. To see that ϕ is Lipschitz, let G be a roll-k gap of C and $G' = \phi^{-1}(G)$. The mass coordinate of G and of G' is $p2^{-k}$ for some $p \in \mathbb{N}$, because the mass of every roll-k (=level-k) clone of G is 2^{-k}. At roll-$(k-M)$ of C', every clone has mass 2^{-r} for some $k-M \leq r < k$. At roll-$(k-M+aM)$ of C', $a \in \mathbb{N}$, every roll-$(k-M)$ clone of C' has gaps dividing it into 2^a co-gaps of equal mass, although there may be additional gaps as well. This means that at roll-$(k-M+M^2)$ every roll-$(k-M)$ clone of C' has gaps dividing it into 2^M co-gaps of equal mass. Therefore there is a gap G'' with mass coordinate $p2^{-k}$, and G'' has appeared by roll-$(k-M+M^2)$. Thus $G' = G''$, and

$$|G'| \cdot |G|^{-1} \geq \alpha^{M-k-M^2} L' / \alpha^{-k} = L' \alpha^{M-M^2},$$

where $L' = \min\{|H| \colon H \text{ is a level-1 gap of } C'\}$. Hence ϕ is Lipschitz, and ϕ is a quasi-isometry, proving (3) $\Rightarrow$ (1).

Next we prove (1) $\Rightarrow$ (3). If $C \simeq C'$, then by the corollary to (2.1) there is a measure linear quasi-isometry $\phi \colon C' \to C$, not necessarily surjective. If x is a local similarity point of C' then by Corollary 1 to (2.4), ϕx is a local similarity point of C' and $S(x) = S(\phi x)$. By the corollary to (2.3), there is a clone map of C with Radon-Nikodým derivative $[S(x)]^n$ for some $n \in \mathbb{N}$. The level-1 clone maps of C both have Radon-Nikodým derivative 2, therefore every clone map of C has derivative an integral power of 2. Thus $S(x) = 2^r$ for some $r \in \mathbb{Q}$. It follows that every clone map of C' has Radon-Nikodým derivative a rational power of 2. Let the level-1 clone maps of C' be $\tau_i' \colon J_i' \to [0,1]$ for $1 \leq i \leq q$. Then $(d\tau_i'/d\mu_{C'})^{-1} = \mu_{C'}(J_i') = 2^{-r_i}$ for some $r_i \in \mathbb{Q}$ and $\sum_{i=1}^q 2^{-r_i} = 1$. It follows that $r_i \in \mathbb{N}$.

Now $\phi(C') = A$ is a co-gap of C, and is therefore a union of level-p clones of C for p large enough. So $\mu_C(A) = m2^{-p}$ where m is the number of level-p clones in A. ϕ is measure linear, therefore there are gaps in C' dividing C' into m co-gaps of mass m^{-1}. Since every level-1 clone of C' has mass an integral power of 1/2, $G(C') \subset \mathbb{Z}[1/2]$, so $m^{-1} \in \mathbb{Z}[1/2]$ and $m = 2^a$ for some $a \in \mathbb{N}$. Thus $1/2 \in G(C')$, proving (1) $\Rightarrow$ (3).

To show (3) $\Rightarrow$ (2) is trivial.

Finally to show (2) $\Rightarrow$ (3), observe that $G(C) = [0,1] \cap \mathbb{Z}[1/2]$. Let A' be a clone of C'. Then $\mu_{C'}(A') \in \mathbb{Z}[1/2]$, and $\mu_{C'}(A') = q2^{-p}$ for some $p, q \in \mathbb{N}$, with q odd. Let x be the leftmost point of A', $y \in A'$ the point with $\mu_{C'}([x,y]) = 2^{-p}$, and $\sigma \colon A' \to C'$ the clone map. Then $\mu_{C'}(\sigma[x,y]) = q^{-1}$. Since $\sigma([x,y])$ is a co-gap of C', $q^{-1} \in \mathbb{Z}[1/2]$ and $q = 1$. Hence the mass of every clone of C' is an integral power of 1/2.

Remark. In the generic case, and for the examples of this section, the gap invariant is a complete invariant of quasi-isometry classes of affinely generated Cantor sets. However, it is not yet known if the gap invariant is a quasi-isometry invariant for all affinely generated Cantor sets. It is not a complete invariant, for example writing $(6 \times \frac{1}{6})$ for $(\frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6})$, $C = (6 \times \frac{1}{6}) \not\simeq (12 \times \frac{1}{12}) = C'$, although both have gap invariant $[0,1] \cap \mathbb{Z}[1/6]$. This is because if x is the leftmost point of C' then $S(x) = 12^{-q}$ for some $q \in \mathbb{Q}$, but if y is a local similarity point of C then $S(y) = 6^{-r}$ for some $r \in \mathbb{Q}$. Now $12^{-q} = 6^{-r}$ is impossible for r, q nonzero rationals. So by Corollary 2 of (2.4), $C \not\simeq C'$.

Bibliography

[1] M. Gromov, *Groups of polynomial growth and expanding maps*, Inst. Hautes Études Sci. Publ. Math. **53** (1981) 53–73.

[2] T. Pignataro, *Hausdorff dimension, spectral theory and applications to the quantization of geodesic flows on surfaces of constant negative curvature*, Ph. D. Thesis, Princeton University, Princeton, NJ, 1984.

[3] D. Sullivan, *Limit sets and the expanding properties of iterated transformations*, to appear.

University of California, San Diego
Institut des Hautes Études Scientifiques