Nonlinearity **9** (1996) 403–412. Printed in the UK

# Cantor sets in the line: scaling functions and the smoothness of the shift-map

Feliks Przytycki† and Folkert Tangerman‡

† Institute of Mathematics at the Polish Academy of Sciences, ul. Śniadeckich 8, 00-950 Warsaw, Poland

‡ Department of Applied Mathematics and Statistics, SUNY Stony Brook, NY 11794-3600, USA

Received 30 March 1994
Recommended by P Cvitanović

**Abstract.** Consider $d$ disjoint closed subintervals of the unit interval and consider an orientation preserving expanding map which maps each of these subintervals to the whole unit interval. The set of points where all iterates of this expanding map are defined is a Cantor set. Associated with the construction of this Cantor set is the scaling function which records the infinitely deep geometry of this Cantor set. This scaling function is an invariant of $C^1$ conjugation. Dennis Sullivan posed the inverse problem: given a scaling function, determine the maximal possible smoothness of any expanding map which produces it. We solve this problem in the case of finite smoothness and in the real-analytic case.

AMS classification scheme numbers: 58F03, 58F08, 58F15

## Preliminaries

Consider the space $\Sigma_d = \{1, \ldots, d\}^{\mathbb{N}}$, with its standard shift-map $\sigma$

$$\sigma(\alpha_1\alpha_2\ldots) = (\alpha_2\ldots).$$

Denote by $\sigma_i^{-1}$ the $d$ right-inverse of $\sigma$:

$$\sigma_i^{-1}(\alpha_1\alpha_2\ldots) = (\mathrm{i}\alpha_1\alpha_2\ldots)$$

Our convention will be to use no separating commas in strings of symbols.

$\Sigma_d$ with the product topology is a Cantor set. Consider an embedding $h$ of the space $\Sigma_d = \{1, \ldots, d\}^{\mathbb{N}}$ into $\mathbb{R}$ with the standard order:

$$h(\alpha) > h(\beta) \qquad \text{iff } \alpha_m > \beta_m$$

where $m$ is the first integer for which $\alpha_m \neq \beta_m$. The image of $h$ is also a Cantor set. Denote by $f$ the induced shift-map on the image of $h$ and by $f_i^{-1}$ the $d$ right-inverses of $f$. Let $r > 1$. We say that $h$ is $C^r$ if each of the right-inverses $f_i^{-1}$ have $C^r$ extensions to $\mathbb{R}$ which are contractions. We say then that the Cantor set is $C^r$.

Every $C^{1+\epsilon}$ Cantor set has a scaling function, defined below, and there is a simple characterization of those functions which are scaling functions for some $C^{1+\epsilon}$ Cantor set. In this paper we describe those scaling functions which actually have up to $C^{k+\epsilon}$ realizations. Here $k$ is any integer greater than or equal to 1 and $0 < \epsilon \leqslant 1$. We follow the convention that $\epsilon = 1$ means a Lipschitz condition.

0951-7715/96/020403+10$19.50 

The theory for $r = 1 + \epsilon$ is essentially due to Feigenbaum [F] and Sullivan [S] who introduced the scaling function. It is defined in the following manner. Given an embedding $h$, then the shift-map allows a canonical definition of the image of $h$ as an intersection of nested collections of intervals. More precisely, define for any finite sequence $(j_1 \ldots j_n)$ $I_{j_1 \ldots j_n}$ as the convex hull of $h(\{\alpha : \alpha_1 = j_n, \ \ldots, \ \alpha_n = j_1\})$. Note the order in which the indices occur. Then for any $j_0$, $I_{j_0 j_1 \ldots j_n} \subset I_{j_1 \ldots j_n}$ and the shift-map maps $I_{j_1 \ldots j_n}$ to $I_{j_1 \ldots j_{n-1}}$. For the empty string, $I$ denotes the image of $h$. The sets thus constructed are not intervals, but actually small pieces of the image of $h$. It is however convenient to think of them as intervals.

For any subset $J$ in the reals denote by $\langle J \rangle$ its convex hull and by $|J|$ the length of its convex hull. We will in the remainder always assume that $\langle I \rangle$ is the unit interval [0,1].

Denote the set of finite strings $j_1, \ \ldots j_n$ of length $n$ by $\Sigma^{\text{dual}}_{d,n}$. The scaling function (ratio geometry) at level $n$ is a function $S^n$:

$$S : \Sigma^{\text{dual}}_{d,n} \to (0,1)^{2d-1}$$

defined in the following manner. For each $j_1 \ldots j_n$ $S(j_1 \ldots j_n)$ records the geometrical location of the $d$ intervals $\{\langle I_{j_0 j_1 \ldots j_n} \rangle\}_{j_0 = 1 \ldots d}$ in $\langle I_{j_1 \ldots j_n} \rangle$ by the ratios of lengths of these $d$ intervals (first $d$ coordinates) and $d - 1$ gaps (last $d - 1$ coordinates) to the length of $\langle I_{j_1, \ \ldots j_n} \rangle$. In particular for $j_0 = 1, \ \ldots d$ the $j_0$th coordinate of $S$ is given by the following formula:

$$S(j_1 \ldots j_n)_{j_0} = \frac{|I_{j_0 \ldots j_n}|}{|I_{j_1 \ldots j_n}|}.$$

The sum of all ratios of lengths equals one. Therefore $S$ actually takes values in the $(2d - 2)$-dimensional simplex $\mathrm{Simp}_{2d-2}$ of $(0,1)^{2d-1}$ where the sum of the coordinates equals 1. Moreover lengths of intervals are determined by the scaling functions at all levels:

$$|I_{j_1 \ldots j_n}| = \prod_k S(j_{k+1} \ldots j_n)_{j_k}$$

Consider two finite sequences $j = j_1 \ldots j_n$ and $j' = j'_1 \ldots j'_m$. There is a canonical identification between $I_j$ and $I_{j'}$ defined as follows. Let $j \cap j'$ be the longest string which agrees with both the beginning of $j$ and the beginning of $j'$. Then suitable iterates of the shift-map map $I_j$ to $I_{j \cap j'}$ respectively $I_{j'}$ to $I_{j \cap j'}$ (see diagram):

$$\begin{tikzcd} & I_{j' \cap j} & \\ I_{j'} \arrow[ur] & & I_j \arrow[ul] \end{tikzcd}$$

The fundamental observation is that if the embedding is $C^{1+\epsilon}$ then the identification map is close to being linear in the following precise sense. Define the nonlinearity of a diffeomorphism $f$ on an interval as

$$\log \sup_{x,y,x \neq y} \frac{Df(x)}{Df(y)}.$$

Then the nonlinearity of the identification map can be estimated from above in terms of the length of the intermediary interval $I_{j \cap j'}$. But then if $j \cap j'$ is long (i.e. $|I_{j \cap j'}|$ small), the subdivision of $I_j$ is close to that of $I_{j'}$. One concludes that there exists a uniform $\gamma$ such that $0 < \gamma < 1$

$$|S(j_1 \ldots j_n) - S(j'_1 \ldots j'_m)| \leqslant \gamma^{\sharp(j \cap j')}.$$

Here $\sharp(j \cap j')$ denotes the length of $j \cap j'$. Therefore for any infinite sequence $j = (j_1 j_2 \ldots)$ the scaling function $S$:

$$S(j) = \lim_{n \to \infty} S(j_1 \ldots j_n)$$

is well defined and has a Hölder modulus of continuity:

$$|S(j) - S(j')| \leqslant \gamma^{\sharp(j \cap j')}.$$

This scaling function is canonically defined on the dual Cantor set $\Sigma_d^{\mathrm{dual}}$, whose elements are infinite sequences $(j_1 j_2 \ldots)$. Each such sequence should be thought of as a prescribed sequence of inverse branches of the shift-map.

Say that a map is $C^{1+}$ if it is $C^{1+\epsilon}$ for some $\epsilon$.

**Theorem.** [S] *Every $C^{1+}$ embedding has a Hölder continuous scaling function. The scaling function is a $C^1$ invariant. Moreover it is a $C^{k+\epsilon}$ complete invariant, namely two $C^{k+\epsilon}$ embeddings ($k$ positive integer and $0 < \epsilon \leqslant 1$) with the same scaling function are $C^{k+\epsilon}$ isomorphic. Every Hölder continuous function on the dual Cantor set with values in* $\mathrm{Simp}_{2d-2}$ *is the scaling function of a $C^{1+}$ embedding.*

Here the Hölder continuity of the scaling function is defined with respect to a metric on $\Sigma_d^{\mathrm{dual}}$:

$$\rho_\delta(j, j') = \exp(-\delta \, \sharp(j \cap j')).$$

In the theorem $\delta$ (the metric on $\Sigma_d^{\mathrm{dual}}$) is not specified so we cannot specify $\epsilon$.

The problem which remained was to understand which functions occur as scaling functions for $C^{1+\epsilon}$ (concrete $\epsilon$) and higher smoothness. Here we give necessary and sufficient conditions for a function $S$ to arise as a scaling function for a $C^{k+\epsilon}$ ($k$ positive integer and $0 < \epsilon \leqslant 1$) embedding. The main observation is that, given an embedding, we should be able to extend the identification map between $I_j$ and $I_{j'}$ to their convex hulls $\langle I_j \rangle$ and $\langle I_{j'} \rangle$ to be $C^{k+\epsilon}$ close to affine provided $j \cap j'$ is long. Here close to affine is measured after affinely rescaling $\langle I_j \rangle$ and $\langle I_{j'} \rangle$ to the unit interval. We refer to the process of changing the map by rescaling domain and range to the unit interval as renormalization.

## 1. $C^{1+\epsilon}$ theory

We will first characterize those functions which are scaling functions of $C^{1+\epsilon}$ Cantor sets. This is a special case of the main theorem. We state it separately because of its simpler form. Given a function $S : \Sigma_d^{\mathrm{dual}} \to \mathrm{Simp}_{2d-2}$. We replace an arbitrary metric $\rho_\delta$ on $\Sigma_d^{\mathrm{dual}}$ with a metric $\rho_S$ so that for an embedding with $S$ as scaling function there exists $K$ so that for every $j, j'$:

$$\frac{1}{K} \leqslant \frac{|I_{j \cap j'}|}{\rho_S(j, j')} \leqslant K.$$

This metric is defined as:

$$\rho_S(j, j') = \sup_w \prod_{t=1}^{n=\sharp(j \cap j')} S(j_{t+1} j_{t+2} \ldots j_n w)_{j_t}.$$

(3) holds by (1) because any infinite tail $w$ changes the product by a uniformly bounded factor (by (2)).

**Theorem 1.** *Fix* $0 < \epsilon \leqslant 1$. *The following are equivalent:*

***1.*** *There exists a* $C^{1+\epsilon}$ *embedding with scaling function* $S$.

***2.*** $S$ *is* $C^{\epsilon}$ *on* $(\Sigma_d^{\text{dual}}, \rho_S)$. *(Here* $C^1$ *means Lipschitz).*

**Proof.** That **1** ⇒ **2** follows when one observes that a stronger form of (2) holds:

$$|S(j_1 \ldots j_n) - S(j_1' \ldots j_n')| \leqslant K\,|I_{j \cap j'}|^{\epsilon}$$

This inequality carries over to the scaling function. Next apply (3).

That **2** ⇒ **1**, i.e. the construction of a $C^{1+\epsilon}$ Cantor set will be done in the proof of the Main Theorem. □

**Example 1.** For every $0 < \epsilon_1 < \epsilon_2 \leqslant 1$ there exists $S$ admitting a $C^{1+\epsilon_1}$ embedding but not $C^{1+\epsilon_2}$. We find it as follows: For an arbitrary $0 < \nu < \frac{\epsilon_2 - \epsilon_1}{2}$ we can easily find a function $S$ to $\text{Simp}_{2d-2}$ which is $C^{\epsilon_1+\nu}$ but not $C^{\epsilon_2-\nu}$ on $\Sigma_{d,n}^{\text{dual}}$ with a standard metric $\rho_\delta$, $\delta > \log d$. We can find in fact $S$ so that for every $j \in \Sigma_{d,n}^{\text{dual}}$, $i = 1, \ldots d$ $|-\log S(j)_i/\delta - 1| < \nu/\epsilon_2$. This is chosen so that $S$ is $C^{\epsilon_1}$ but not $C^{\epsilon_2}$ with respect to the metric $\rho_S$.

## 2. $C^{k+\epsilon}$ theory.

We now turn to the more intricate case of higher smoothness.

Let $A_1$ and $A_2$ be two subsets of the unit interval [0, 1] such that both sets contain the endpoints of [0, 1] and both have equal cardinality. Denote the $k$th derivative operator by $D^k$ and denote by $D^k(A_1, A_2)$ the space of $C^k$ diffeomorphisms on [0, 1] which map $A_1$ to $A_2$. For every constant $M > 0$ consider the space of $C^k$-diffeomorphisms:

$$D^k_{\text{var}}(M)(A_1, A_2) = \{\phi \in D^k(A_1, A_2) : \sup |D^k\phi(x) - D^k\phi(y)| \langle M\}$$

**Lemma.** *Assume that* $A_1$ *and* $A_2$ *consist of* $2d$ *points. Assume that* $k < 2d$. *Then for each* $f, g$ *in* $D^k_{\text{var}}(M)(A_1, A_2)$ *we have for all integers* $0 \leqslant t \leqslant k$*:*

$$\sup |D^t f - D^t g| \leqslant 2M$$

**Proof.** Consider two such maps $f$ and $g$. Their difference vanishes on $A_1$. Since $2d \rangle k$, there exists (mean value theorem) for each $t$ a point $x_t$ in [0, 1] for which:

$$D^t f(x_t) - D^t g(x_t) = 0$$

The lemma follows by induction and integration. □

Given a function $S$ as above and a point $j$ in $\Sigma_d^{\text{dual}}$. Consider $S(j)$. It encodes a partition of [0, 1] in $2d - 1$ intervals. Denote by $A(j)$ the $2d$ end points of these intervals. Consider any $j_0 = 1, \ldots d$ and consider the point $j_0 j$ in $\Sigma_d^{\text{dual}}$. Then $S(j_0 j)$ specifies how the $j_0$th interval in $j$ is subdivided. Consider two points $j$ and $j'$ in $\Sigma_d^{\text{dual}}$ Every element in $D^k(A(j), A(j'))$ maps the $j_0$th interval in the domain to the $j_0$th interval in the range, which we again can renormalize. This defines a map (restrict to $j_0$th interval and renormalize):

$$R_{j_0} : D^k(A(j), A(j')) \to D^k(\{0, 1\}, \{0, 1\})$$

**Main Theorem (**$C^{k+\epsilon}$ **case).** *Suppose* $k < 2d$. *Suppose that we are given a function* $S$ *as above. The following are equivalent.*

***1.*** *There exists a* $C^{k+\epsilon}$ *embedding with scaling function S.*

***2.*** *There exists a constant $C$ so that for all $j$ and $j'$ in $\Sigma_d^{\mathrm{dual}}$ and all $j_0 = 1, \ldots d$:*

$$D^k_{\mathrm{var}}(C(j_0 j,\ j_0 j'))(A(j_0 j), A(j_0 j')) \cap R_{j_0}(D^k_{\mathrm{var}}(C(j,\ j'))(A(j),\ A(j')) \neq \emptyset$$

*where for all $j,\ j' \in \Sigma_d^{\mathrm{dual}}$,*

$$C(j,\ j') = C\,\rho_S(j,\ j')^{k+\epsilon-1}.$$

**Discussion of statement of theorem.** The statement of the theorem may appear obscure. We briefly discuss in an informal manner how the scaling function records smoothness beyond $C^1$.

**(1)** Consider two strings $j$ and $j'$ and the identification map between $I_j$ and $I_{j'}$. The scalings $S(j)$ and $S(j')$ record how $2\,d$ specific points in $I_j$ map to $2\,d$ specific points in $I_{j'}$. Consider the renormalized identification map , and assume that we know that the variation of the $k$th derivative of this identification map is small. Consider any $k + 1$ of the $2\,d$ specific points. Since we know where these points map, we can compute a value of the $k$th derivative (just as the standard mean value theorem computes a value of the first derivative given 2 points and their values). Because the variation of the $k$th derivative is small, we obtain combinatorial relations between any two choices of $k + 1$ points. Condition **2** of the theorem captures this idea. It does not attempt to give an algebraic description of these combinatorial relations†.

**(2)** In fact we do not need all $2d$ points which appear in the definition of the ratio geometry to be involved in the definition of $D^k_{\mathrm{var}}$, $k+1$ would be enough (see lemma). If we define these spaces in this way then condition **2** seems to be vacuous for the $C^{1+\epsilon}$ case, because $k = 1$. In particular it gives the impression we do not need the geometry at all. However, condition **2** in theorem 1 (hence property (4) in the proof of theorem 1) is still hidden in condition **2** in the Main Theorem. Without condition **2** in theorem 1, a map $[0, 1] \to [0, 1]$ in $D^k_{\mathrm{var}}(A(j_0 j), A(j_0 j'))$, even a linear one, after renormalizing by $R^{-1}_{j_0}$ may not be extendible to a map belonging to the second $D$ in **2** of the Main Theorem.

**(3)** The condition of the Main Theorem seems to imply that high smoothness is not discussed when $d$ is small. We can however replace $d$ by any positive power $d^n$ in the following manner. $\Sigma_d$ is canonically homeomorphic to $\Sigma_{d^n}$, by the homeomorphism which groups the digits of a point in $\Sigma_d$ in groups of $n$ digits. This homeomorphism conjugates the $n$th iterate of the shift-map on $\Sigma_d$ to the shift-map on $\Sigma_{d^n}$.

**Proof of Main Theorem.** We first show that **1** implies **2**. Assume that we are given a $C^{k+\epsilon}$ embedding $h$. Denote the induced shift-map on the image by $f$. We may assume that its $d$ right-inverses extend as $C^{k+\epsilon}$ contractions to the unit interval, the convex hull of the image of $h$. Denote by $f_{j'|j}$ the identification between $\langle I_j \rangle$ and $\langle I_{j'} \rangle$ for finite strings $j$, $j'$ and denote by $F_{j'|j}$ the renormalized identification defined on the unit interval. Then $f_{j'|j}$, respectively $F_{j'|j}$, factors as a composition:

$$f_{j'|j} = f_{j'|j'\cap j} \circ f_{j'\cap j|j}$$

$$F_{j'|j} = F_{j'|j'\cap j} \circ F_{j'\cap j|j}.$$

Since $f_{j'|j'\cap j} : \langle I_{j'\cap j} \rangle \to \langle I_{j'} \rangle$ is a composition of $C^{k+\epsilon}$ contractions the derivatives of $f_{j'|j'\cap j}$ are controlled by the first derivative.

† Added in revision. In the analogous case of expanding mappings of the circle algebraic conditions are worked out up to smoothness 3 by A Pinto and D Sullivan in [PS].

More precisely, by a standard computation which we leave to the reader, there exists a constant $C$ so that for all $j$ and $j'$, all $1 \leqslant \tau \leqslant k + \epsilon$

$$|f_{j'|j'\cap j}|_\tau \leqslant C\,|f_{j'|j'\cap j}|_1.$$

Here $|.|_\tau$ denotes the sup-norm of the $\tau$th derivative for $\tau$ integer and the $\alpha$- Hölder norm of the $n$th derivative if $\tau = n + \alpha$, $0 < \alpha \leqslant 1$.

But then:

$$|F_{j'|j'\cap j}|_\tau = \frac{|I_{j'\cap j}|^\tau}{|I_{j'}|}\,|f_{j'|j'\cap j}|_\tau$$

$$\leqslant |I_{j'\cap j}|^{\tau-1}\,C.$$

The last inequality follows because:

$$\frac{|I_{j'}|}{|I_{j'\cap j}|} = Df_{j'|j'\cap j}(x)$$

for some point $x \in I_{j'\cap j}$ and the bounded nonlinearity of the maps.

Now let $j$ and $j'$ be two distinct points in $\Sigma_d^{\mathrm{dual}}$. Denote by $j_n$, respectively $j'_n$, the beginning strings of length $n$. Then for $n$ large enough $j \cap j' = j_n \cap j'_n$ and the sequence of maps $\{F_{j'_n|j'\cap j}\}$ is $C^{k+\epsilon}$-equicontinuous (for $\epsilon > 0$). Since moreover:

$$F_{j'_{n+m}|j'\cap j} = F_{j'_{n+m}|j'_n} \circ F_{j'_n|j'\cap j}$$

this sequence of maps is in fact $C^{k+\epsilon}$ convergent. Denote by $F_{j'|j'\cap j}$ the limit map. By the same argument $F_{j|j'\cap j}$ is defined. Therefore the limiting map:

$$F_{j'|j} = F_{j'|j'\cap j} \circ F^{-1}_{j|j'\cap j}$$

is well-defined and $C^{k+\epsilon}$ (until here $k > 1, \epsilon = 0$ has been allowed). Since $\rho_S(j, j')$ is uniformly comparable to $|I_{j'\cap j}|$ we obtain this limiting map $F_{j'|j}$ in $D^k_{\mathrm{var}}(C'\rho_S(j, j')^{k+\epsilon-1})$ for some uniform constant $C'$. Since moreover:

$$R_{j_0}\,F_{j'|j} = F_{j_0 j'|j_0 j}$$

we automatically have an element in the intersection. **2** now follows.

We next show that **2** implies **1**. Since $S$ is given, we first construct an embedding of the Cantor set with $S$ as scaling function. We then show that this embedding is $C^{k+\epsilon}$.

Fix an arbitrary infinite word $w$. Construct a Cantor set $C$ in the unit interval [0, 1] by consecutively subdividing any interval $\langle I_j\rangle$ according to $S(jw)$. We obtain an embedding with scaling function $S$. Denote the induced shift-map on the image by $f_0$. It is defined on a Cantor set $C$. In order to show that this shift-map has a $C^{k+\epsilon}$ extension, we verify the assumptions to Whitney's Extension Theorem [Stein]. We will construct functions $f_1, \ldots f_k$ on $C$ so that for all $x$, $y$ in $C$ and $l = 0, \ldots k$ (Whitney conditions):

$$f_l(y) = \sum_{t=l}^{t=k} \frac{1}{(t-l)!}\,f_t(x)(y-x)^{t-l} + O(|y-x|^{k-l+\epsilon}).$$

These functions $f_1, \ldots f_k$ play the role of the first $k$ derivatives of $f_0$.

The interval $\langle I\rangle = [0, 1]$ is subdivided in $d$ intervals $\langle I_i\rangle$, $i = 1, \ldots, d$. On each of the intervals, $f_0$ maps $I_i = C \cap \langle I_i\rangle$ to $I = C \cap \langle I\rangle$ by $f_0$. Now fix $i = 1, \ldots d$. We will work on each $\langle I_i\rangle$ separately. For each $t = 1, \ldots, k$ define $f_t$ on $I_i$ as:

$$f_t = \lim_{n\to\infty}\{\,D^t\phi_{j_1\ldots j_n, j}\,\}_{(j_1\ldots j_n)}$$

Here $\phi_{j_1\dots j_n,i} : \langle I_{j_1\dots j_n i}\rangle \to \langle I_{j_1\dots j_n}\rangle$ is any map whose renormalization is in

$$D_k^{\mathrm{var}}(C(j_1\dots j_n i w,\ j_1\dots j_n w)(A(j_1\dots j_n i w),\ A(j_1\dots j_n w)).$$

We need to see that $f_t$ is in fact well-defined on the Cantor set. We first verify that $f_t$ is defined point wise on the Cantor set. Consider a string $j_1\dots j_n$ and an element $j_0$. For $x \in I_{j_0 j_1\dots j_n i}$, consider $\phi_{j_0 j_1\dots j_n i}(x)$ and $\phi_{j_1\dots j_n i}(x)$ and their $t$th derivatives. Then by assumption **2** and the lemma:

$$\begin{aligned}
&|D^t\phi_{j_0 j_1\dots j_n i}(x) - D^t\phi_{j_1\dots j_n i}(x)| \\
&\leqslant \frac{|I_{j_0 j_1\dots j_n}|}{|I_{j_0 j_1\dots j_n i}|^t}\, C\, \rho_S(j_0 j_1\dots j_n i w, j_0 j_1\dots j_n w)^{k+\epsilon-1} \\
&\leqslant C\,|I(j_0 j_1\dots j_n i)|^{k+\epsilon-t}.
\end{aligned}$$

Therefore we obtain the convergence on the Cantor set in fact exponentially fast.

We need to check that the Whitney conditions hold on the Cantor set. Let $x$ and $y$ be distinct points in the Cantor set in $\langle I_i\rangle$. Consider the first time that they wind up in different intervals in the subdivision:

$$x \in \langle I_{j_0 j_1\dots j_n i}\rangle,\quad y \in \langle I_{j_0' j_1\dots j_n i}\rangle \qquad j_0 \neq j_0'.$$

Then again by **2**:

$$|\phi_{j_1\dots j_n i}(y) - \phi_{j_1\dots j_n i}(x) - \sum_{t=0}^{t=k}\frac{1}{t!}D^t\phi_{j_1\dots j_n i}(x)(y-x)^t| \leqslant C\,|x-y|^{k+\epsilon}$$

(and similarly for the higher derivatives) where $C$ is a uniform constant. Since $|D^t\phi_{j_1\dots j_n i}(x) - f_t(x)| \leqslant C\,|x-y|^{k+\epsilon-t}$, we can take limits and obtain the Whitney conditions for the family $f_0, f_1, \dots, f_k$. Consequently there exists a $C^{k+\epsilon}$ extension of $f$ to each $\langle I_i\rangle$ and we have produced a $C^{k+\epsilon}$ embedding of $\Sigma_d$ with scaling function $S$. □

**Corollary†.** *Assume that $h_1$ and $h_2$ are $C^1$-equivalent $C^{k+\epsilon}$ embeddings: $h_2 \circ h_1^{-1}$ is $C^1$. Then $h_2 \circ h_1^{-1}$ is $C^{k+\epsilon}$.*

**Proof.** To show that the conjugacy $h_2 \circ h_1^{-1}$ has a $C^{k+\epsilon}$ extension, it suffices to construct its higher derivatives on the Cantor set and apply the Whitney Extension Theorem. This can be achieved using the same manner as that employed in the second half of the proof of the Main Theorem. Both embeddings have the same scaling function $S$ so , as the embeddings are $C^{k+\epsilon}$, the ratio geometries on finite levels are close to one another in the sense of condition **2** of the Main Theorem. □

**Remark.** The preceding is not totally satisfactory; for example we do not understand how to extract $C^k$-smoothness ($k$ integer!) from the 'finite condition' on scaling function. This is because in the previous scheme everything which needs to be controlled is dominated by geometric series. More refined finite smoothness categories like $C^{1+zygmund}$ can however be treated in much the same way.

Finally, for every $k$, $d$, $\epsilon$ with $k < 2d - 1$, $0 \leqslant \epsilon < 1$ and $k + \epsilon > 1$, we construct an example of a scaling function with a $C^{k+\epsilon}$ realization and none of higher degree of smoothness.

† Added in revision. This is a part of Sullivan's theorem formulated in the preliminary section. A detailed proof along Sullivan's line appeared recently in [BF] and will appear also in [PS].

**Example 2.** Let $J_i = [\frac{2i-2}{2d-1}, \frac{2i-1}{2d-1}]$, $i = 1, \ldots, d$.

For $\epsilon > 0$ define $f : \cup_i J_i \to J = [0, 1]$ as

$$f(x) = A((2d-1)x + x^{k+\epsilon}) \qquad x \in J_1$$

while $f$ is affine onto $J$ on each $J_i$, $i \geqslant 2$. Here the the constant $A$ is chosen so that $f(J_1) = J$.

Of course the resulting Cantor set is $C^{k+\epsilon}$. We will show that its scaling function on the dual Cantor set has no $C^{k+\epsilon_1}$ realization for all $\epsilon_1 > \epsilon$, by explicitly checking that condition **2** of the Main Theorem does not hold for $k + \epsilon_1$.

Let $w$ be any element in $\Sigma_d^{\mathrm{dual}}$ which does not contain the symbol 1. Denote by $1_n$ the string of length $n$ consisting of 1s only:

$$1_n = 11\ldots1.$$

Consider the infinite strings $j = 1_n w$ and $j' = 1_n 1 w = 1_{n+1} w$. Consider the subdivision $A(j)$, respectively $A(j')$, of the unit interval dictated by $S(j)$ and $S(j')$. Let $\Phi_n$ be any map in $D^k_{\mathrm{var}}(A(j), A(j'))$ for which its renormalized restriction $R_1\Phi$ is in fact in $D^k_{\mathrm{var}}(A(1j), A(1j')$. We will bound the variation of the $k$th derivative of $\Phi_n$ from below and conclude that condition **2.** of the Main Theorem is not satisfied with $k + \epsilon_1$.

We denote by $A(j)_m$ the $m$th point from the left in $A(j)$. Because $k < 2d - 1$ there exists $x \in [A(j)_2, A(j)_{2d}]$ such that:

$$D^k F_{j'|j}(x) = D^k \Phi_n(x).$$

Recall that $F_{j'|j}$ is the renormalization of $f_{j'|j}$ for the map $f$ defined above. See the notation of the proof of the Main Theorem. Similarly there exists $y \in [A(1j)_1, \ldots, A(1j)_{2d-1}]$ so that

$$D^k F_{1j'|1j}(y) = D^k(R_1\Phi_n)(y).$$

We have that:

$$D^k F_{j'|j}(x) = B\ (2d-1)^{-n(k-1+\epsilon)}\, x^{\epsilon}$$

(note that $(2d-1)^{-n} \sim \rho_S(j', j)$). The map $F_{1j'|1j}$ is just the renormalization of the restriction of the limit map $F_{j'|j}$ to the left most interval $[A(j)_1, A(j)_2]$ in the unit interval. Let $y'$ be the point in the interval $[A(j)_1, A(j)_2]$, corresponding to $y$ after rescaling the unit interval back to $[A(j)_1, A(j)_2]$. Then we have that:

$$D^k(F_{j'|j})(y') = B\ (2d-1)^{-n(k-1+\epsilon)}\,(y')^{\epsilon}$$

where $B$ is a computable constant.

But $|x - y'| > const\,(2d-1)^{-2}$. Consequently:

$$D^k\Phi_n(x) - D^k\Phi_n(y') = const\ (2d-1)^{-n(k-1+\epsilon)}\,(x^{\epsilon} - (y')^{\epsilon}$$

and is comparable to

$$\rho_s(j', j)^{k-1+\epsilon}$$

i.e. the variation of $D^k\Phi_n$ is at least on the order of: $\rho_s(j', j)^{k-1+\epsilon}$.

Since

$$\lim_{n\to\infty} \frac{\rho_s(j', j)^{k-1+\epsilon_1}}{\rho_s(j', j)^{k-1+\epsilon}} = 0$$

condition **2** of the theorem cannot be satisfied for

$$C(j', j) = C\,\rho_S(j', j)^{k-1+\epsilon_1}.$$

For $\epsilon = 0$ we consider, say, $f(x) = A((2d-1)x + x/\log x)$.

## 3. Real-analytic theory

Here we shall characterize scaling functions which admit real-analytic embeddings. This characterization will be of an 'infinite' type. In example 3 we will present a scaling function which has a smooth realization but no real-analytic one.

Given a scaling function $S$ on $\Sigma_d^{\mathrm{dual}}$ we define a scaling function $\hat{S}$ with values in embeddings of $\Sigma_d$ in $[0,1]$ rather than in $\mathrm{Simp}_{2d-2}$. For each $j = j_1 \ldots \in \Sigma_d^{\mathrm{dual}}$ we define an embedding by induction as follows.

Suppose for $i_n \ldots i_0$ an interval $J_{i_n \ldots i_0} \in [0,1]$ is already defined (for the empty string we set $J = [0,1]$). Then for every $i_{n+1} = 1, \ldots, d$ we define $J_{i_{n+1} i_n \ldots i_0} \subset J_{i_n \ldots i_0}$ as an interval which after an affine rescaling of $J_{i_n \ldots i_0}$ to $[0,1]$ becomes the $(2i_{n+1} - 1)$th interval of the partition of $[0,1]$, determined by $S(i_n i_1 \ldots i_0 j)$. Then $\hat{S}(j)$ is the embedding of $\Sigma_d : (i_0 i_1 \ldots) \mapsto \bigcap_{n=0}^{\infty} J_{i_n \ldots i_0}$. We denote the image Cantor set i.e. $\hat{S}(j)(\Sigma_d)$ by $\mathrm{Can}(j)$. (Observe that this is the same construction as in section 2, Proof of Main Theorem, with $j = w$.)

**Remark.** For every $j$, $\hat{S}(j)$ has the same scaling function $S$ on $\Sigma_d^{\mathrm{dual}}$.

**Main Theorem (real-analytic case).** *Suppose we are given a scaling function $S$. The following are equivalent:*

***1.*** *There exists a $C^\omega$ (real-analytic) embedding with scaling function $S$.*

***2.*** *For every $j, j' \in \Sigma_d^{\mathrm{dual}}$ there exists a real-analytic orientation preserving diffeomorphism $\Phi_{j'|j} : [0,1] \to [0,1]$ mapping* $\mathrm{Can}(j)$ *onto* $\mathrm{Can}(j')$.

**Proof. 1** implies **2** because given $j, j' \in \Sigma_d^{\mathrm{dual}}$ we can take $\Phi_{j|j'} = \lim_{n\to\infty} F_{j'_n|j_n}$ as in the proof of the Main Theorem. $F_{j'_n|j_n}$ are defined on a fixed complex neighbourhood of $[0,1]$ and the sequence is even exponentially convergent because $F_{j_n|j_{n+1}}$ and $F_{j'_{n+1}|j'_n}$ converge exponentially to the identity.

**2** implies **1** in a trivial way. Just take as the Cantor set we are looking for $\mathrm{Can}(j)$ for an arbitrary $j$. The extensions of the shift are renormalizations of analytic maps $\Phi_{j|i_0 j}$ to maps from $J_{i_0}$, $i_0 = 1, \ldots, d$ to $[0,1]$. □

Now we will show the existence of a scaling function which admits a $C^\infty$ embedding but no $C^\omega$

**Example 3.** Take $f$ as in example 2 except that now $A((2d-1)x + x^{r+1})$ is replaced by

$$A((2d-1)x + \exp(-1/x)).$$

Suppose that **2** holds, i.e. there exists a real-analytic $\Phi_{w|1w} : [0,1] \to [0,1]$ mapping $\mathrm{Can}(1w)$ onto $\mathrm{Can}(w)$. From the fact $\Phi_{w|1w}$ coincides with $F_{w|1w}$ on $\mathrm{Can}(1w)$ we conclude that their all derivatives coincide on sequences of points converging to 0. But $F_{w|1w}$ is defined with the use of the same formula (5) as $f$ (up to a renormalization) so all the $r$th derivatives , $r > 1$ , of $F_{w|1w}$, hence $\Phi$, are 0 at 0. This implies $\Phi$ is not real-analytic.

**Remark** (added in revision). One can simplify section 2 by writing an 'infinite' condition as above, instead of **2** in the Main Theorem. For example:

There exists $j \in \Sigma_d^{\mathrm{dual}}$ such that for every $i_0 = 1, \ldots, d$ $\Phi_{j|i_0 j}$ mapping $\mathrm{Can}(i_0 j)$ onto $\mathrm{Can}(j)$ is $C^{k+\epsilon}$.

This includes $k > 1, \epsilon = 0$. One obtains in particular a corollary for $\epsilon = 0, k \geqslant 2$ ! Indeed assuming $h_1$ and $h_2$ are $C^k$ embeddings with the same scaling function $S$, they have the same scaling function $\hat{S}$. So $F^2_{\emptyset|j} \circ F^1_{j|\emptyset}$ is a $C^k$-conjugacy. ($F$ denotes the renormalized $f$ as in section 2, the superscript is 1 or 2 depending as one considers the embedding $h_1$ or $h_2$, $j$ is an arbitrary infinite string, $\emptyset$ is the length 0 string.)

## Acknowledgments

FP thanks the hospitality of the Institute of Mathematical Sciences at SUNY Stony Brook and the Department of Mathematics at Yale University. This paper was written during his stay at these institutions in 1991/92. He acknowledges also support by Polish KBN Grants 210469101 'Iteracje i Fraktale' and 210909101 '. . . Uklady Dynamiczne'.

## References

[BF] Bedford T and Fisher A 1994 Ratio geometry, rigidity and the scenery process for hyperbolic Cantor sets *SUNY Stony Brook preprint 9*

[F] Feigenbaum M 1978 Quantitative universality for a class of non-linear transformations *J. Stat. Phys.* **19** 25–52

[PS] Pinto A and Sullivan D 1994 Asymptotic geometry applied to expanding dynamical systems *Preprint* IHES

[PUbook] Przytycki F and Urbański M 1996 Fractals in the complex plane—ergodic theory methods, in press

[Stein] Stein E 1970 *Singular Integrals and Differentiability Properties of Functions* (Princeton, NJ: Princeton University Press)

[S] Sullivan D 1988 Differentiable structures on fractal-like sets, determined by intrinsic scaling functions on dual Cantor sets *The Mathematical Heritage of Hermann Weyl, AMS Proc. Symp. Pure. Math.* vol 48