Ark. Mat., **40** (2002), 335–362

Level sets of harmonic functions on the Sierpiński gasket

Anders Öberg(1), Robert S. Strichartz(2) and Andrew Q. Yingst(3)

Abstract. We give a detailed description of nonconstant harmonic functions and their level sets on the Sierpiński gasket. We introduce a parameter, called *eccentricity*, which classifies these functions up to affine transformations $h \mapsto ah+b$. We describe three (presumably) distinct measures that describe how the eccentricities are distributed in the limit as we subdivide the gasket into smaller copies (cells) and restrict the harmonic function to the small cells. One measure simply counts the number of small cells with eccentricity in a specified range. One counts the contribution to the total energy coming from those cells. And one counts just those cells that intersect a fixed generic level set. The last measure yields a formula for the box dimension of a generic level set. All three measures are defined by invariance equations with respect to the same iterated function system, but with different weights. We also give a construction for a rectifiable curve containing a given level set. We exhibit examples where the curve has infinite winding number with respect to some points.

1. Introduction

Interesting geometric objects may be described as solutions of equations. This insight, which already appears in the concept of "locus" in Greek geometry, and which became the central idea in Descartes' geometry, culminated in a precise form in the implicit function theorem.

When we consider the realm of fractals, no such analogous result presents itself. Indeed, it is not at all clear what should be considered the interesting geometric subsets of a given fractal. On the other hand, at least for certain fractals, we have an idea of some natural classes of functions that are analogous to smooth functions on

(1) Part of this research was carried out during a visit to the Mathematics Department of Cornell University.

(2) Research supported in part by the National Science Foundation, grant DMS–9970337.

(3) Research supported by the National Science Foundation through the Research Experiences for Undergraduates (REU) program at Cornell.

Euclidean space or manifolds. So we can begin to explore this problem by trying to understand the properties of the level sets of these functions. If the implicit function theorem is any guide, not all the level sets should be equally nice, but perhaps a "generic" level set will be a candidate for a class of "interesting geometric subsets".

In this work we deal with only the simplest case: nonconstant harmonic functions on the Sierpiński gasket, as described in Kigami [K1]. These functions are perhaps more closely analogous to linear functions on an interval rather than harmonic functions on a multi-dimensional domain. They satisfy a simple local interpolation identity which lends itself both to numerical computations and theoretical work (see [DSV] for some examples). The space of harmonic functions is three-dimensional, but by composing on the right with an isometry of the Sierpiński gasket, and on the left with an affine mapping, which does not change the level sets, we can reduce to a one-parameter family of harmonic functions. We refer to the parameter as *eccentricity*.

We give next a description of these harmonic functions. The Sierpiński gasket SG is the invariant set for the iterated function system in the plane given by

$$F_i x = \tfrac{1}{2}(x-p_i)+p_i, \quad i=1,2,3, \tag{1.1}$$

where p_i are the vertices of an equilateral triangle. We regard SG as the limit of the graphs Γ_m with vertices V_m and edge relations $x\sim_m y$ defined inductively as follows. Let Γ_0 be the complete graph on $V_0=\{p_1,p_2,p_3\}$. Then $V_m=\bigcup_i F_i V_{m-1}$ with $x\sim_m y$ if and only if there exists i such that $x=F_i x'$, $y=F_i y'$ and $x'\sim_{m-1} y'$. Note that $V_{m-1}\subseteq V_m$. We regard V_0 as the boundary of SG and each of the graphs Γ_m, so that $V_m\setminus V_0$ consists of all nonboundary vertices in Γ_m. Note that every such vertex has exactly four neighbors in V_m (since it belongs to $V_{m'}$ for every $m'\geq m$, it also has exactly four neighbors in $V_{m'}$, but the set of neighbors changes with m').

A function on V_m is said to be *harmonic* if for every nonboundary vertex its value is the average of its values at the four neighbors,

$$h(x)=\frac{1}{4}\sum_{y\sim_m x} h(y). \tag{1.2}$$

This is a standard definition from graph theory (assuming equal weights on all edges). It is necessary to exclude boundary vertices in order to have nonconstant harmonic functions. It is not hard to show that every harmonic function on V_{m-1} (no condition imposed when $m=1$) has a unique harmonic extension to V_m, given by the following local harmonic extension algorithm:

$$h(x)=\tfrac{2}{5}h(y)+\tfrac{2}{5}h(z)+\tfrac{1}{5}h(w), \tag{1.3}$$

where $x \in V_m \setminus V_{m-1}$, where y and z are the two neighbors of x in V_m that belong to V_{m-1}, and w is the third vertex of the triangle in V_{m-1} that contains y and z. In fact it is trivial to show from (1.3) that h satisfies (1.2) at the new vertices $V_m \setminus V_{m-1}$, but then one needs to show that condition (1.2) continues to hold for the old vertices $V_{m-1} \setminus V_0$ on level m, using the assumption that it holds on level $m-1$ together with (1.3).

A continuous function h on SG is said to be *harmonic* if its restriction to V_m is harmonic for all m. If we assign any boundary values $h(p_j)$, then applying the extension algorithm (1.3) inductively we can determine the values of h on V_m, and since $\bigcup_m V_m$ is dense in SG, this determines h (it is not hard to see that (1.3) always produces continuous functions in the limit). Thus the space of harmonic functions is three-dimensional, and contains the constants. For any nonconstant harmonic function we may relabel the vertices so that $h(p_1) \le h(p_2) \le h(p_3)$ with $h(p_1) < h(p_3)$. This amounts to composing h on the right with an isometry of SG. We define the eccentricity $e(h)$ by

$$e(h) = \frac{h(p_2)-h(p_1)}{h(p_3)-h(p_1)}, \quad 0 \le e(h) \le 1. \tag{1.4}$$

By composing on the left with an affine function (which leaves the eccentricity unchanged) we may assume without loss of generality that $h(p_1)=0$, $h(p_2)=e$, $h(p_3)=1$. Also note that $e(1-h)=1-e(h)$, so that harmonic functions with eccentricities e and $1-e$ are essentially equivalent. Thus we could impose the restriction $0 \le e \le \frac{1}{2}$, if desired.

The theory of harmonic functions on SG is part of a more general theory that includes a Laplacian Δ, so that h is harmonic if and only if $\Delta h=0$. For the most part we will not refer to this more general theory here (see the books [Ba] and [K4] and the references there, as well as the expository article [S1]). We will need the concept of *energy*, which for harmonic functions is easily explained. Let

$$\mathcal{E}_m(u,u) = \left(\frac{5}{3}\right)^m \sum_{y \in V_m} \sum_{x \sim_m y} (u(x)-u(y))^2. \tag{1.5}$$

For harmonic functions, $\mathcal{E}_m(h,h)$ is independent of m, and is called the energy. The mysterious renormalization factor $\left(\frac{5}{3}\right)^m$ in (1.5) is explained by this condition. Kusuoka has shown that there is a measure ν_h associated with the energy so that the energy is the measure of the whole SG, but this measure is singular with respect to the normalized Hausdorff measure on SG (see [Ku] or [BST]). In fact one can already see in (1.5) how the energy is made up of a sum of contributions from the various small triangles of level m in SG.

The first problem we consider is the description of the self-similarity of the harmonic functions, or more precisely, the whole family of harmonic functions. For a fixed harmonic function h, we may restrict to each of the level one cells $F_i(SG)$ and rescale to obtain three harmonic functions $h\circ F_i$. But these are different harmonic functions with different eccentricities. In fact there are three explicit functions ψ_i such that

$$e(h\circ F_i)=\psi_{i'}(e(h)). \tag{1.6}$$

The mapping $i\mapsto i'$ is a permutation that depends on the orientation of vertices for which $h(p_1)\le h(p_2)\le h(p_3)$. More generally, on level m we may decompose SG into 3^m cells $F_w(SG)$, where $w=(w_1,\dots,w_m)$ denotes a word of length $|w|=m$, each $w_j=1,2,3$, and $F_w=F_{w_1}\circ F_{w_2}\circ\dots\circ F_{w_m}$. The restriction of h to $F_w(SG)$ rescaled is just $h\circ F_w$, and we have

$$e(h\circ F_w)=\psi_{w'}(e(h)), \tag{1.7}$$

where $\psi_{w'}=\psi_{w'_1}\circ\dots\circ\psi_{w'_m}$. Thus a harmonic function of eccentricity e is composed of 3^m harmonic functions with eccentricities $\{\psi_w(e):|w|=m\}$ when viewed at level m. The self-similarity of harmonic functions is governed by the dynamics of the iterated function system $\{\psi_i\}_{i=1}^3$.

We will show that there is a unique non-atomic probability measure μ on $[0,1]$ invariant under $\{\psi_i\}_{i=1}^3$ with equal probability weights:

$$\mu=\frac{1}{3}\sum_{i=1}^{3}\mu\circ\psi_i^{-1}, \tag{1.8}$$

and this measure is attractive in the Wasserstein (or Monge–Kantorovich) metric

$$d_W(\mu,\nu)=\sup\left\{\left|\int_0^1 f\,d\mu-\int_0^1 f\,d\nu\right|:\|f\|_{\mathrm{Lip}}\le 1\right\} \tag{1.9}$$

on probability measures. (See [R] for a general reference on probability metrics.) This means that if we look at the distribution of the eccentricities at level m (the discrete measure $3^{-m}\sum_{|w|=m}\delta(\psi_w(e))$), we obtain the invariant measure μ in the limit as $m\to\infty$ in the metric (1.9). This is true no matter what the initial eccentricity of h. From the point of view of the statistics of the eccentricities of the restrictions to small cells, all harmonic functions are the same.

There are two ways to prove the existence and uniqueness of the solution to (1.8). The first refers to the theory of products of random matrices. This is presented in Section 2. This approach has the advantage that it yields additional information relating the eccentricity in a specific cell to the location of the cell, almost independent of the particular harmonic function. The second approach, given in Section 5, is based on an average contractivity in the geometric sense of the double iterate of the iterated function system. The advantage of this approach is that it gives a geometric rate of convergence.

In Sections 3 and 4 we introduce two other measures describing the distribution of eccentricities. The measure μ_E counts the amount of energy contained in cells with specified eccentricity. The measure μ_L describes the distribution of eccentricities encountered among the cells that meet a generic level set. The invariant equation for μ_L involves a parameter 2^α, and the value α gives the box dimension of a generic level set. The existence and uniqueness of these measures (and the parameter 2^α) is also established in Section 5.

In Section 4 we describe the self-similarity of the entire family of level sets. This leads to an invariant measure ν on the two-dimensional family of level sets. In this case we do not have uniqueness for solutions of the invariant equation for ν, since there do exist non-generic level sets. However, the generic situation leads to the conclusion that ν has a product structure, namely $\nu=\mu_L\times dt$ where dt is Lebesgue measure on the parameter t that indexes the particular level set.

In Section 6 we show how to construct a rectifiable curve γ in the plane such that a given level set is the intersection of SG with γ. We give an example to show that γ may have infinite spiral points.

In Section 7 we discuss briefly another example, the hexagasket. This is a fractal in the class of post-critically finite fractals for which Kigami's theory of harmonic functions applies [K4]. We see that the situation there is quite different, owing to the existence of vertices that are not in the boundary and not junction points. In particular this leads to harmonic extension matrices that are not invertible. There exist nonconstant harmonic functions that are nevertheless constant on open sets. Also, the analogs of the measures μ, μ_E and μ_L are discrete. A reasonable conjecture is that the results of this paper may be extended to harmonic structures on most post-critically finite fractals having the property that all harmonic extension matrices are invertible.

We note that in [DSV] it is shown that certain level sets of eigenfunctions of the Laplacian on SG contain line segments. This is presumably nongeneric behavior, but it indicates the difficulty in extending this work to more general functions.

Acknowledgments. We are grateful to A. Teplyaev for helpful discussions.

2. Products of random matrices

The space of harmonic functions on SG is three-dimensional. If we use the boundary values $(h(p_1), h(p_2), h(p_3))$ as coordinates, then the local extension algorithm (1.3) can be written in matrix form

$$h \circ F_i = M_i h \tag{2.1}$$

with

$$M_1 = \frac{1}{5}\begin{pmatrix} 5 & 0 & 0 \\ 2 & 2 & 1 \\ 2 & 1 & 2 \end{pmatrix}, \quad M_2 = \frac{1}{5}\begin{pmatrix} 2 & 2 & 1 \\ 0 & 5 & 0 \\ 1 & 2 & 2 \end{pmatrix}, \quad M_3 = \frac{1}{5}\begin{pmatrix} 2 & 1 & 2 \\ 1 & 2 & 2 \\ 0 & 0 & 5 \end{pmatrix}. \tag{2.2}$$

The iterated form is

$$h \circ F_w = M_{w_m} \dots M_{w_1} h \quad (\text{denoted } M_w h) \tag{2.3}$$

which fits exactly the theory of products of random matrices, with the matrices drawn from the three-element set (2.2). It is desirable to factor out the constant functions to get a two-dimensional vector space and corresponding 2×2 matrices. The choice of coordinate representations is not canonical, but we can obtain an orthonormal basis with respect to the energy (degenerate) inner product by choosing $h_1 = (0, \sqrt{2}, \sqrt{2})$ and $h_2 = (0, \sqrt{2/3}, -\sqrt{2/3})$. Then any harmonic function may be written $a_1 h_1 + a_2 h_2 + c$, and (a_1, a_2) will be our coordinates for harmonic functions mod constants, denoted informally $\tilde{h}$. Then

$$\widetilde{h \circ F_i} = \widetilde{M}_i \tilde{h} \tag{2.4}$$

with

$$\widetilde{M}_1 = \begin{pmatrix} \frac{3}{5} & 0 \\ 0 & \frac{1}{5} \end{pmatrix}, \quad \widetilde{M}_2 = \begin{pmatrix} \frac{3}{10} & \frac{\sqrt{3}}{10} \\ \frac{\sqrt{3}}{10} & \frac{1}{2} \end{pmatrix}, \quad \widetilde{M}_3 = \begin{pmatrix} \frac{3}{10} & -\frac{\sqrt{3}}{10} \\ -\frac{\sqrt{3}}{10} & \frac{1}{2} \end{pmatrix} \tag{2.5}$$

and the iterated form is

$$\widetilde{h \circ F_w} = \widetilde{M}_{w_m} \dots \widetilde{M}_{w_1} \tilde{h} \quad (\text{denoted } \widetilde{M}_w \tilde{h}). \tag{2.6}$$

It is natural to introduce polar coordinates in this two-dimensional space and to delete the origin, which represents the constant functions. Thus each nonconstant harmonic function is associated with a unique angle θ (mod 2π). The eccentricity

e is a continuous function of θ which can be given piecewise by explicit rational functions of $\sin\theta$ and $\cos\theta$. For example

$$e=\frac{\sqrt{3}\,\cos\theta-\sin\theta}{\sqrt{3}\,\cos\theta+\sin\theta}\quad\text{on }\left[0,\frac{\pi}{3}\right] \tag{2.7}$$

with similar expressions on five other intervals. Of course θ carries slightly more information, since it also determines the orientation. Also, if θ corresponds to e then $\theta+\pi$ corresponds to $1-e$. Since the results from the theory of products of random matrices concern the projective circle $P(\mathbf{R}^2)$ (θ and $\theta+\pi$ are identified since they determine the same line through the origin), they will only translate to statements about eccentricities with the identification of e and $1-e$. We will use θ to represent a variable in $P(\mathbf{R}^2)$, so in what follows it should be understood that every function $f(\theta)$ satisfies $f(\theta)=f(\theta+\pi)$.

From the theory of products of random matrices we immediately obtain a great deal of information about the angles $\theta(\widetilde{M}_w\tilde{h})$. Note that we are in the fortunate position of dealing with 2×2 matrices where the theory is more complete. It is clear by inspection that the matrices (2.4) generate a noncompact semigroup with no invariant finite union of lines, so Theorem 4.1 of [Bo] applies. The first conclusion is the existence of a unique invariant probability measure μ on $P(\mathbf{R}^2)$ associated with the independent choice of $\widetilde{M}_i$ with equal $\frac{1}{3}$ probabilities. This may be written

$$\int_0^{2\pi} f(t)\,d\mu(t)=\frac{1}{3}\sum_{i=1}^{3}\int_0^{2\pi} f\left(\theta\left(\widetilde{M}_i\begin{pmatrix}\cos t\\ \sin t\end{pmatrix}\right)\right)\,d\mu(t) \tag{2.8}$$

for any function f. The measure μ is nonatomic (continuous). If we start with any initial nonzero vector v we may obtain μ as

$$\int_0^{2\pi} f(t)\,d\mu(t)=\lim_{m\to\infty}\frac{1}{3^m}\sum_{|w|=m} f(\theta(\widetilde{M}_w v)). \tag{2.9}$$

For large m, the direction $\theta(\widetilde{M}_w v)$ effectively depends only on w: there exists a function $\theta(w)$ such that for every v and ε,

$$\lim_{m\to\infty}\frac{1}{3^m}\#\{w:|w|=m\text{ and }|\theta(\widetilde{M}_w v)-\theta(w)|\ge\varepsilon\}=0. \tag{2.10}$$

Note that for any w, the function $\theta(\widetilde{M}_w v)$ maps onto $P(\mathbf{R}^2)$, but for most choices of v the values will be close to each other, and we may choose $\theta(w)$ to be $\theta(\widetilde{M}_w v)$ for one of these typical choices of v.

We may formulate a general principle, *geography is destiny*, to describe this phenomenon. The behavior of a harmonic function in a neighborhood of a generic point x is essentially determined, except for an affine transformation $h \mapsto ah+b$, by the location of the point, not the particular harmonic function. This seems paradoxical, because the space of harmonic functions is three-dimensional. Of course, for each point x there will be exceptional harmonic functions, but they will constitute a set of measure zero with respect to Lebesgue measure in three-space. There will also be nongeneric points x, including in particular all vertex points. The computational implication is that for large m and most choices of words w of length m, the values of $\theta(\widetilde{M}_w(\cos t, \sin t))$ will cluster around some value θ_0 (and $\theta_0+\pi$) for all t except for a few narrow regions. (Similarly, the eccentricities $e(h\circ F_w)$ will cluster around some e_0 and $1-e_0$ for all initial eccentricities e except for a few narrow regions.) Unfortunately, the convergence of this process is slow, so that m has to be taken to be quite large for this to be apparent.

The "geography is destiny" principle seems likely to hold more generally for functions in the domain of the Laplacian and perhaps even functions of finite energy, using the methods of Kusuoka [Ku] and Teplyaev [T].

3. Distribution of eccentricities

Suppose h is a harmonic function with boundary values $h(p_1)=0$, $h(p_2)=e$ and $h(p_3)=1$. Then $h\circ F_1$ has boundary values

$$h(F_1(p_1)) = 0 \le h(F_1(p_2)) = \tfrac{1}{5}(2e+1) \le h(F_1(p_3)) = \tfrac{1}{5}(e+2),$$

so $e(h\circ F_1)=(2e+1)/(e+2)$. Similarly $h\circ F_3$ has boundary values

$$h(F_3(p_1)) = \tfrac{1}{5}(e+2) \le h(F_3(p_2)) = \tfrac{1}{5}(2e+2) \le h(F_3(p_3)) = 1,$$

so $e(h\circ F_3)=e/(3-e)$. For $h\circ F_2$, the order of the boundary values

$$h(F_2(p_1)) = \tfrac{1}{5}(2e+1), \quad h(F_2(p_2)) = e, \quad h(F_3(p_2)) = \tfrac{1}{5}(2e+2)$$

depends on the size of e. When $0\le e\le \frac{1}{3}$, we have $h(F_2(p_2))\le h(F_2(p_1))\le h(F_2(p_3))$, hence $e(h\circ F_2)=(1-3e)/(2-3e)$. When $\frac{1}{3}\le e\le \frac{2}{3}$, we have $h(F_2(p_1))\le h(F_2(p_2))\le h(F_2(p_3))$, so $e(h\circ F_2)=3e-1$. When $\frac{2}{3}\le e\le 1$, we have $h(F_2(p_1))\le h(F_2(p_3))\le h(F_2(p_2))$, so $e(h\circ F_2)=1/(3e-1)$.

We may summarize these computations in the form (1.6) with

$$
(3.1)\qquad \begin{cases} \psi_1(x)=\dfrac{2x+1}{x+2}, \\ \psi_2(x)=\begin{cases} \dfrac{1-3x}{2-3x}, & \text{if } 0\le x\le \dfrac{1}{3}, \\ 3x-1, & \text{if } \dfrac{1}{3}\le x\le \dfrac{2}{3}, \\ \dfrac{1}{3x-1}, & \text{if } \dfrac{2}{3}\le x\le 1, \end{cases} \\ \psi_3(x)=\dfrac{x}{3-x}. \end{cases}
$$

Figure 3.1 below shows the simultaneous graphs of ψ_1, ψ_2, ψ_3. Note that while ψ_1 and ψ_3 are one-to-one and contractive, ψ_2 is two-to-one and expansive. If we identify x and $1-x$ we obtain an iterated function system on $0\le x\le \frac{1}{2}$ given by

$$
(3.2)\qquad \begin{cases} \tilde\psi_1(x)=\dfrac{1-x}{x+2}, \\ \tilde\psi_2(x)=\begin{cases} \dfrac{1-3x}{2-3x}, & \text{if } 0\le x\le \dfrac{1}{3}, \\ 3x-1, & \text{if } \dfrac{1}{3}\le x\le \dfrac{1}{2}, \end{cases} \\ \tilde\psi_3(x)=\dfrac{x}{3-x}. \end{cases}
$$

It is clear that questions about invariant measures for the iterated function system $\{\psi_j\}_{j=1}^3$ have equivalent formulations for the iterated function system $\{\tilde\psi_j\}_{j=1}^3$ provided the weights respect the $e\mapsto 1-e$ symmetry, which is true in all cases we consider. Therefore, we will work with whichever formulation is most convenient for each question.

According to the results of the last section, there is a unique probability measure satisfying (1.8), which is just the image under $\theta\mapsto e$ of the invariant measure satisfying (2.8) (by slight abuse of notation we will use the same letter μ to denote both measures). In Figure 3.2 we show an approximation of this measure obtained after 10 iterations of the iterated function system starting from a random initial value. By a standard argument based on uniqueness, the measure is either singular or absolutely continuous, but not a mixture of both.

The measure μ gives equal consideration to all cells of level m. It is also natural to ask how much energy is contained in the cells with eccentricity in a given range. This will be described by a different measure, which will be denoted μ_E.

The total energy of our choice of h is $E(h)=2(e^2-e+1)$, and this splits into a sum of $\frac{5}{3}E(h\circ F_i)$. We can easily compute the values, and express the result in the form

$$\tfrac{5}{3}E(h\circ F_i)=p_i(e)E(h) \tag{3.3}$$

with

$$\begin{cases} p_1(e)=\dfrac{1}{5}\dfrac{e^2+e+1}{e^2-e+1}, \\ p_2(e)=\dfrac{1}{5}\dfrac{3e^2-3e+1}{e^2-e+1}, \\ p_3(e)=\dfrac{1}{5}\dfrac{e^2-3e+3}{e^2-e+1}. \end{cases} \tag{3.4}$$

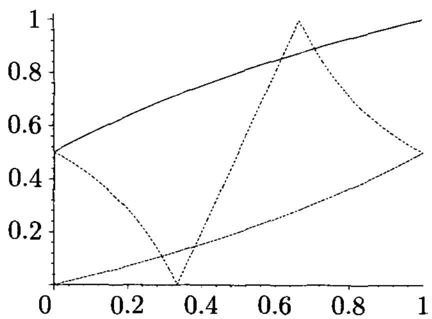
Note that we have

$$p_1(e)+p_2(e)+p_3(e)=1. \tag{3.5}$$

Figure 3.1 also shows the simultaneous graphs of $p_i(e)$. The advantage of (3.3) is that it remains valid for any harmonic function (provided that we permute the indices $i\mapsto i'$ as in (1.6) for different orientations of the vertices). Thus the measure μ_E satisfies the invariance

$$\mu_E=\sum_{i=1}^{3}(p_i\mu_E)\circ\psi_i^{-1} \tag{3.6}$$

with respect to the variable probability weights (3.5). On the right-hand side of Figure 3.2 we show an approximation of the measure μ_E analogous to the approximation to μ. The appearance of the approximation of μ_E is considerably more singular than that of μ displayed on the left of Figure 3.2. Not only do the measures μ and μ_E appear to be singular with respect to Lebesgue measure, but they also appear to be mutually singular.



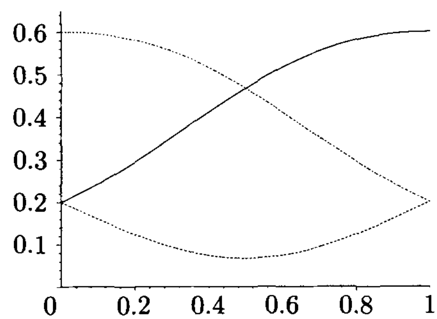


Figure 3.1. On the left: The graphs of ψ_1 (solid line), ψ_2 (dashed line) and ψ_3 (dotted line). On the right: The graphs of p_1 (solid line), p_2 (dashed line) and p_3 (dotted line).

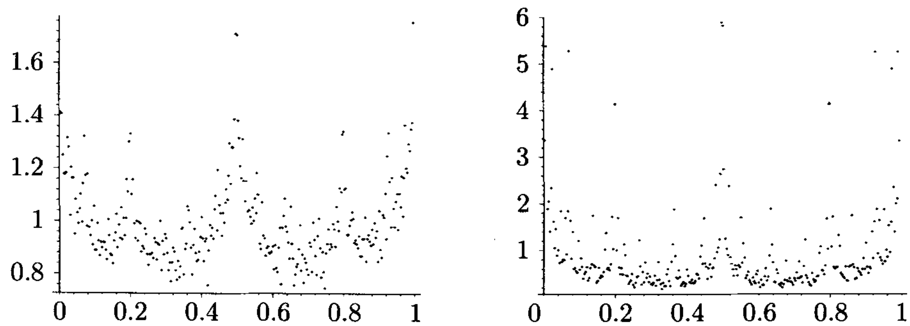


Figure 3.2. Histogram approximations to μ (left) and μ_E (right) with 300 equally spaced bins in $[0,1]$, after 10 iterations of the iterated function system from a random initial value.

4. Level sets

Let $L(e,t)$ denote the level set $\{x{:}h(x)=t\}$ for the harmonic function h of eccentricity e as in the previous section, and $0\le t\le 1$. If h is any nonconstant harmonic function with $h(p_1)\le h(p_2)\le h(p_3)$, then

$$\{x:h(x)=s\}=L\left(\frac{h(p_2)-h(p_1)}{h(p_3)-h(p_1)},\frac{s-h(p_1)}{h(p_3)-h(p_1)}\right) \tag{4.1}$$

so these level sets and their rotations and reflections contain all level sets of nonconstant harmonic functions.

At level m, the level set $L(e,t)$ intersects a number $N(m,e,t)$ of cells $F_w(SG)$ with $|w|=m$, and each intersection is of the form $F_wRL(e',t')$ for some (e',t') and a rotation or reflection R. By abuse of notation we will omit R since it plays an inessential role in our analysis. The number $N(m,e,t)$ will be of the order of $2^{\alpha m}$, where α is the box dimension of $L(e,t)$. More precisely

$$\alpha=\lim_{m\to\infty}\frac{\log N(m,e,t)}{m\log 2} \tag{4.2}$$

is the box dimension, if the limit exists (otherwise the lim sup and lim inf give the upper α^* and lower α_* box dimensions). We are also interested in the distribution of the values (e',t') that occur, and the measure they approximate on the square,

$$\nu(A)=\lim_{m\to\infty}\frac{\#\{(e',t')\in A \text{ at level } m\}}{N(m,e,t)} \quad \text{for } A\subseteq[0,1]\times[0,1]. \tag{4.3}$$

We will see that for a generic choice of (e,t) the measure ν is a product of a measure μ_L in e and the uniform Lebesgue measure in t, with μ_L being independent of the

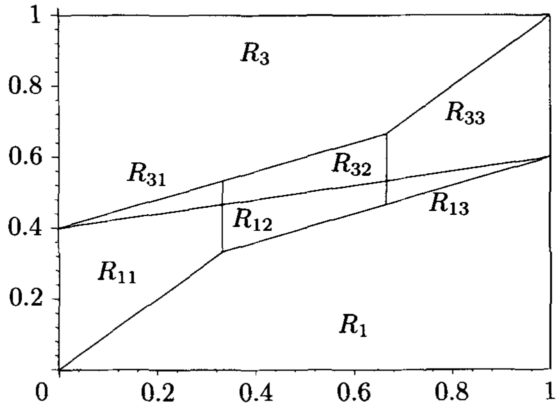


Figure 4.1. The subdivision of the unit square into 8 regions.

particular level set. There are exceptional level sets, such as $L(e,0)$ and $L(e,1)$ which consist of single points, and $L\left(\frac{1}{2},\frac{1}{2}\right)$ which consists of a countable set of points with a single limit point p_2.

We now describe in detail what happens when $m=1$. We divide the square up into 8 regions R_1, R_3, R_{11}, R_{12}, R_{13}, R_{31}, R_{32} and R_{33}. See Figure 4.1. The union $R_1 \cup R_{11} \cup R_{12} \cup R_{13}$ is just the region defined by $t \leq \frac{1}{5}(e+2)$ of all (e,t) such that $L(e,t)$ intersects $F_1(SG)$, and there we have

$$L(e,t) \cap F_1(SG) = F_1(L(\varphi_1(e,t))) \tag{4.4}$$

for

$$\varphi_1(e,t) = \left(\frac{2e+1}{e+2}, \frac{5t}{e+2}\right). \tag{4.5}$$

Similarly, $R_3 \cup R_{31} \cup R_{32} \cup R_{33}$ is defined by $t \geq \frac{1}{5}(e+2)$ and represents the (e,t) values for which $L(e,t)$ intersects $F_3(SG)$, with

$$L(e,t) \cap F_3(SG) = F_3(L(\varphi_3(e,t))) \tag{4.6}$$

for

$$\varphi_3(e,t) = \left(\frac{e}{3-e}, \frac{5t-e-2}{3-e}\right). \tag{4.7}$$

The intersection of $L(e,t)$ with $F_2(SG)$ occurs in the regions $R_{11} \cup R_{12} \cup R_{13} \cup R_{31} \cup R_{32} \cup R_{33}$. The region $R_{11} \cup R_{31}$ is defined by $0 \leq e \leq \frac{1}{3}$ and $e \leq t \leq \frac{1}{5}(2e+2)$, and there

$$L(e,t) \cap F_2(SG) = F_2(L(\varphi_{21}(e,t))) \tag{4.8}$$

for

$$\varphi_{21}(e,t)=\left(\frac{1-3e}{2-3e},\frac{5t-5e}{2-3e}\right). \tag{4.9}$$

The region $R_{13}\cup R_{33}$ is defined by $\frac{2}{3}\leq e\leq 1$ and $\frac{1}{5}(2e+1)\leq t\leq e$, and there

$$L(e,t)\cap F_2(SG)=F_2(L(\varphi_{23}(e,t))) \tag{4.10}$$

for

$$\varphi_{23}(e,t)=\left(\frac{1}{3e-1},\frac{5t-2e-1}{3e-1}\right). \tag{4.11}$$

Finally, the region $R_{12}\cup R_{32}$ is defined by $\frac{1}{3}\leq e\leq\frac{2}{3}$ and $\frac{1}{5}(2e+1)\leq t\leq\frac{1}{5}(2e+2)$, and there

$$L(e,t)\cap F_2(SG)=F_2(L(\varphi_{22}(e,t))) \tag{4.12}$$

for

$$\varphi_{22}(e,t)=(3e-1,5t-2e-1). \tag{4.13}$$

Apart from trivial point intersections,

$$N(1,e,t)=\begin{cases}1 & \text{in } R_1\cup R_3,\\ 2 & \text{in } R_{11}\cup R_{12}\cup R_{13}\cup R_{31}\cup R_{32}\cup R_{33}.\end{cases} \tag{4.14}$$

We can summarize the above by the identities

$$L(e,t)=F_j(L(\varphi_j(e,t))) \qquad \text{on } R_j,\ j=1,3, \tag{4.15}$$

$$L(e,t)=F_j(L(\varphi_j(e,t)))\cup F_2(L(\varphi_{2k}(e,t))) \quad \text{on } R_{jk},\ j=1,3,\ k=1,2,3. \tag{4.16}$$

We therefore have a description of the self-similarity of the family of sets $L(e,t)$.

We can write the decomposition at level m as

$$L(e,t)=\bigcup_{w\in A_m}F_w(L(e(w),t(w))), \tag{4.17}$$

where A_m is a set of words of length m (in our notation we suppress the dependence on the initial (e,t) values). Let B_m denote the set of points $(e(w),t(w))$ in the square for $w \in A_m$. Both sets have cardinality $N(m,e,t)$. The identities (4.15) and (4.16) yield an inductive description of these sets:

$$\begin{aligned}&\text{if } w\in A_m \text{ and } (e(w),t(w))\in R_j \ (j=1,3) \text{ then } wj\in A_{m+1}\\&\text{with } (e(wj),t(wj))=\varphi_j(e(w),t(w));\end{aligned} \tag{4.18}$$

$$\begin{aligned}&\text{if } w\in A_m \text{ and } (e(w),t(w))\in R_{jk} \ (j=1,3) \text{ then both}\\&wj \text{ and } w2 \text{ are in } A_{m+1} \text{ with } (e(wj),t(wj))=\varphi_j(e(w),t(w))\\&\text{and } (e(w2),t(w2))=\varphi_{2k}(e(w),t(w)).\end{aligned} \tag{4.19}$$

The distribution of points in the set B_m is described by the measure

$$\frac{1}{N(m,e,t)}\sum_{A_m}\delta(e(w),t(w)). \tag{4.20}$$

In the weak limit as $m\to\infty$ we obtain the invariant probability measure ν which satisfies the identity

$$\nu=\frac{1}{2^\alpha}(\nu\circ\varphi_1^{-1}+\nu\circ\varphi_3^{-1}+\nu\circ\varphi_{21}^{-1}+\nu\circ\varphi_{22}^{-1}+\nu\circ\varphi_{23}^{-1}). \tag{4.21}$$

Since the functions $\varphi_1,\dots,\varphi_{23}$ all have the form where the first variable is independent of t and the second is linear in t, we may assume that ν has the product form $\mu_L(e)\times dt$. If we substitute a measure with this product form in the right-hand side of (4.21), we obtain a measure with this product form in the left-hand side.

Substituting the product form into (4.21) and simplifying, we find the invariant identity for μ_L to be

$$\begin{aligned}\int_0^1 f(e)\,d\mu_L(e)=\frac{1}{2^\alpha}\bigg(&\int_0^1\frac{e+2}{5}f\Big(\frac{2e+1}{e+2}\Big)\,d\mu_L(e)+\int_0^1\frac{3-e}{5}f\Big(\frac{e}{3-e}\Big)\,d\mu_L(e)\\&+\int_0^{1/3}\frac{2-3e}{5}f\Big(\frac{1-3e}{2-3e}\Big)\,d\mu_L(e)+\frac{1}{5}\int_{1/3}^{2/3}f(3e-1)\,d\mu_L(e)\\&+\int_{2/3}^1\frac{3e-1}{5}f\Big(\frac{1}{3e-1}\Big)\,d\mu_L(e)\bigg).\end{aligned} \tag{4.22}$$

Here the weights do not satisfy the probability condition, and the value of 2^α is an unknown. We can obtain a formula for 2^α in terms of the measure by substituting a constant function into (4.22). After some simplification, and using the $e\mapsto 1-e$ symmetry, we find

$$2^\alpha=\frac{6}{5}+2\int_0^{1/3}\frac{1-3e}{5}\,d\mu_L(e). \tag{4.23}$$

This gives the obvious bounds $\frac{6}{5}<2^\alpha<\frac{7}{5}$. If we used uniform measure in place of μ_L this would yield a value of $\frac{19}{15}$ for the right-hand side of (4.23), which would give a value of around $\log\frac{19}{15}/\log 2\approx 0.3410369$ for α.

In Section 5 we will prove (using the estimate $\frac{6}{5}<2^\alpha$) existence and uniqueness for solutions of (4.22), with a geometric rate of convergence. This implies that we have a geometric rate of convergence of (4.20) to a solution $\nu=\mu_L(e)\times dt$ starting from any generic level set $L(e,t)$.

Figure 4.2 shows an approximation to μ_L. Starting from random (e,t) values, we found 9358 cells at level 35 meeting $L(e,t)$. The figure shows the distribution of e-values among these cells. The number of cells gives a rough estimate of 0.3769 for α.

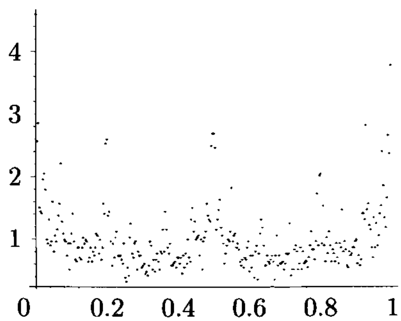


Figure 4.2. A histogram approximation to μ_L, with 300 equally spaced bins in $[0,1]$, starting from random initial values of (e,t). After 35 iterations there are 9358 cells meeting $L(e,t)$, and the e-values only are displayed.

Since there are no exceptional eccentricities, it is reasonable to conjecture that for every e there are at most a countable number of t values for which $L(e,t)$ fails to be generic.

5. Uniqueness and convergence

In this section we present some general results on invariant measures for an iterated function system with weights, and then we show how these results apply to the invariant measures described in Sections 3 and 4. In the following we assume that we have an iterated function system with continuous maps $S_1,\dots,S_m$, $m\geq 2$, with corresponding nonnegative weight functions $q_1,\dots,q_m$, on a compact metric space (X,d). If the weight functions $q_j=p_j$ satisfy $\sum_{j=1}^m p_j(x)=1$ for all $x\in X$, then we say that the iterated function system is probabilistic.

Definition 5.1. The associated *Perron–Frobenius operator* is defined by

$$Lf(x)=\sum_{j=1}^{m} q_j(x)f(S_j(x)). \tag{5.1}$$

We will assume that L acts on Lip, the space of Lipschitz functions equipped with the norm $\|\cdot\|_{\mathrm{Lip}}$ defined by

$$\|f\|_{\mathrm{Lip}}=\sup_{x\neq y}\frac{|f(x)-f(y)|}{|x-y|}+\|f\|_\infty. \tag{5.2}$$

The operator L has an adjoint operator L^*, which we restrict to the probability measures on X,

$$L^*\nu(A)=\sum_{j=1}^{m}\int_{S_j^{-1}(A)} q_j(x)\,d\nu(x), \tag{5.3}$$

where ν is any probability measure and A is any Borel subset of X.

Definition 5.2. The *Wasserstein metric* for probability measures μ and ν is defined as

$$d_W(\mu,\nu)=\sup_{\|f\|_{\mathrm{Lip}}\leq 1}\left|\int_X f\,d\mu-\int_X f\,d\nu\right|. \tag{5.4}$$

Definition 5.3. A probabilistic iterated function system is *average contractive* in the *arithmetic* sense if

$$\sup_{x\neq y}\sum_{j=1}^{m} p_j(x)\,\frac{d(S_j(x),S_j(y))}{d(x,y)}<1. \tag{5.5}$$

A probabilistic iterated function system is *average contractive* in the *geometric* sense if

$$\sup_{x\neq y}\prod_{j=1}^{m} d(S_j(x),S_j(y))^{p_j(x)}<1. \tag{5.6}$$

Lemma 5.4. (See [Ö]) *For an iterated function system with constant probability weights* $p_1, \dots, p_m$ *we have*

$$d_H((L^*)^n\nu, (L^*)^n\mu) \le E\,\mathrm{diam}(S_{w_n}\circ\dots\circ S_{w_1}X), \tag{5.7}$$

where E *denotes the mean taken with respect to the choices of* $w_1, \dots, w_n$*, according to the probabilities* $\{p_j\}_{j=1}^m$*. In particular, if* $E\,\mathrm{diam}(S_{w_n}\circ\dots\circ S_{w_1}X)\to 0$*, then there is a unique probability measure* μ *satisfying* $L^*\mu=\mu$*, i.e., a unique invariant measure for the iterated function system.*

Proof. For probability measures μ and ν, and $f\in\mathrm{Lip}$ with $\|f\|_{\mathrm{Lip}}\le 1$, we have

$$\begin{aligned}\left|\int_X f\,d(L^*)^n\mu-\int_X f\,d(L^*)^n\nu\right| &= \left|\sum_{|w|=n} p_w \int_X\int_X (f(S_wx)-f(S_wy))\,d\mu(x)\,d\nu(y)\right| \\ &\le \sum_{|w|=n} p_w \int_X\int_X d(S_wx, S_wy)\,d\mu(x)\,d\nu(y) \\ &\le E\,\mathrm{diam}(S_wX).\end{aligned}$$

Taking the supremum over f we obtain (5.7). □

We now specialize to the case where (X,d) is an interval and the maps are differentiable. We write, e.g., the arithmetic average contractivity condition as $\sup_{0\le x\le 1}\sum_{j=1}^m p_j(x)|S_j'(x)|<1$. It is an elementary observation to make that if a probabilistic iterated function system is average contractive in the geometric sense, then there exists $0<q\le 1$, such that $\sup_{0\le x\le 1}\sum_{j=1}^m p_j(x)|S_j'(x)|^q<1$. The case when $q=1$ covers the case when the probabilistic iterated function system is also average contractive in the arithmetic sense.

Lemma 5.4 makes it possible to derive an upper estimate for the rate of convergence if we assume only the geometric average contraction condition, but with constant probability weights. If we try to apply this directly to the iterated function system given by (3.2) with weights $\frac{1}{3}, \frac{1}{3}, \frac{1}{3}$ we discover that (5.6) fails. However, if we pass to the double iteration of the iterated function system then (5.6) holds. In contrast, the arithmetic average contractivity condition fails no matter how many times we iterate the iterated function system.

Lemma 5.5. *Suppose that an iterated function system is average contractive in the geometric sense with respect to the constant probability weights. Then* $d_H((L^*)^n\nu, (L^*)^n\mu)\le Ca^n$*, where* $C>0$ *and* $0<a<1$*.*

Proof. We know that there is a $0<q\le 1$ such that $\sup_{0\le x\le 1}\sum_{j=1}^m p_j|S_j'(x)|^q\le \varrho^q$ for some $\varrho<1$. This means that, for any ε satisfying $0<\varepsilon<1$,

$$\begin{aligned}E\,\mathrm{diam}(S_{w_n}\circ\dots\circ S_{w_1}X) &= E(\mathrm{diam}(S_{w_n}\circ\dots\circ S_{w_1}X) : \|(S_{w_n}\circ\dots\circ S_{w_1})'\|_\infty > \varrho^{\varepsilon n}) \\ &\quad + E(\mathrm{diam}(S_{w_n}\circ\dots\circ S_{w_1}X) : \|(S_{w_n}\circ\dots\circ S_{w_1})'\|_\infty \le \varrho^{\varepsilon n}).\end{aligned}$$

Using Chebyshev's inequality we obtain

$$E\operatorname{diam}(S_{w_n}\circ\ldots\circ S_{w_1}X)\le \varrho^{(1-\varepsilon)qn}\operatorname{diam}X+\varrho^{\varepsilon n}\operatorname{diam}X. \quad \square$$

Remark. The optimal choice of ε is to take $\varepsilon=(1-\varepsilon)q$.

Theorem 5.6. *For the iterated function system* (3.2) *with constant probability weights* $\frac{1}{3},\frac{1}{3},\frac{1}{3}$, *the measure* μ *defined by* (1.8) *is uniquely determined by this equation. For any probability measures* ν, $d_H((L^*)^n\nu,\mu)$ *converges to* 0 *at a uniform geometric rate.*

Proof. We consider the double iterate of the iterated function system (3.2) consisting of the nine mappings $\tilde{\psi}_w$ with $|w|=2$ with all probabilities equal to $\frac{1}{9}$. To apply Lemma 5.5 we need to verify the average contractivity in the geometric sense, namely

$$\sup_{0\le x\le 1}\prod_{|w|=2}|\tilde{\psi}'_w(x)|<1.$$

This follows immediately from the straightforward calculation $\prod_{|w|=2}\|\tilde{\psi}'_w\|_\infty^{1/9}=0.8874\ldots$. $\square$

We will now treat the functional equation (4.22), where both the measure μ_L and 2^α are unknown. 2^α can be regarded as an eigenvalue of L and L^* and we will prove that it is in fact the spectral radius and the only eigenvalue with a positive measure solution of (4.22). In the process we will also prove uniqueness of μ_L and we will use the techniques for accomplishing this goal also to provide a proof of uniqueness of μ_E, the invariant measure produced by the iterated function system (3.2) with the variable probability weight functions given by (3.4). Since μ_E is defined using strictly positive probability functions (summing pointwise to one), we could have referred to Barnsley et al. [BDEG], since with these weights the iterated function system is average contractive even in the arithmetic sense; it is very easy to see that it is average contractive in the geometric sense, which is enough, since

$$\sup_{0\le x\le 1}\prod_{j=1}^{3}|\tilde{\psi}'_j(x)|^{p_j(x)}\le\left(\left(\frac{12}{25}\right)^7\left(\frac{3}{4}\right)^3 3^3\right)^{1/15}\approx 0.83499<1.$$

But we will use a different method in order to obtain rates of convergence.

Definition 5.7. The Perron–Frobenius operator L is said to be *quasi-compact* on Lip, if outside a radius strictly smaller than the spectral radius the spectrum consists of discrete eigenvalues with finite multiplicities.

It is well known (see [H] or [KL]) that L is quasi-compact on Lip if

$$\|L^n f\|_{\mathrm{Lip}} \le Ca^n \|f\|_{\mathrm{Lip}} + R_n \|f\|_\infty, \tag{5.8}$$

where $C>0$, $0<a<1$ and $R_n>0$ depend on $n\ge 1$ only.

A sufficient condition for (5.8) to be true is that

$$\sup_{0\le x\le 1} \sum_{j=1}^{m} q_j(x)|S_j'(x)| < \lambda = \text{ the spectral radius,} \tag{5.9}$$

see e.g. [H]. If this is true for L and, in addition (see [H]), we have for every $x\in[0,1]$ and any nontrivial nonnegative continuous function f

$$L^n f(x) > 0 \quad \text{for some } n\ge 1, \tag{5.10}$$

then, since the multiplicity of the spectral radius, which is an eigenvalue, is one, we get

$$\left\| \frac{1}{\lambda^n} L^n f(x) - h(x) \int_0^1 f\, d\mu \right\|_{\mathrm{Lip}} \le Ca^n \|f\|_{\mathrm{Lip}}, \tag{5.11}$$

where, as before, λ is the spectral radius, $C>0$ and $0<a<1$. Furthermore, we have $Lh=\lambda h$ and $h>0$ (a consequence of (5.10)). This implies that μ is the unique invariant measure, in the sense that $L^*\mu=\lambda\mu$.

Theorem 5.8. *In the functional equation* (4.22) *we have a unique* 2^α, *hence a unique* α. *Also,* μ_L *is uniquely defined by* (4.22), *and we have for all Lipschitz functions* f

$$\left\| \frac{1}{2^{\alpha n}} L^n f(x) - h(x) \int_0^1 f\, d\mu_L \right\|_{\mathrm{Lip}} \le Ca^n \|f\|_{\mathrm{Lip}}, \tag{5.12}$$

where L *is defined by using the maps in* (4.22), $h>0$ *and* $Lh=2^\alpha h$.

Proof. The measure μ_L defined by (4.22) is an invariant measure for the iterated function system (3.1) in the sense of

$$2^\alpha \mu_L(A) = \sum_{j=1}^{3} \int_{\psi_j^{-1}(A)} q_j(e)\, d\mu_L(e),$$

or equivalently $L^*\mu_L = 2^\alpha \mu_L$, where

$$\begin{cases} q_1(x) = \dfrac{x+2}{5}, \\ q_2(x) = \begin{cases} \dfrac{2-3x}{5}, & \text{if } 0 \le x \le \dfrac{1}{3}, \\ \dfrac{1}{5}, & \text{if } \dfrac{1}{3} \le x \le \dfrac{2}{3}, \\ \dfrac{3x-1}{5}, & \text{if } \dfrac{2}{3} \le x \le 1, \end{cases} \\ q_3(x) = \dfrac{3-x}{5}. \end{cases}$$

We then have

$$\sup_{0 \le x \le 1} \sum_{j=1}^{3} q_j(x) |\psi_j'(x)| \le \frac{303}{280} < \frac{6}{5} < 2^\alpha. \tag{5.13}$$

The condition (5.10) is true since we have strictly positive weight functions and the two maps ψ_1 and ψ_3 are strictly contractive and nonoverlapping with images covering the whole interval.

Uniqueness of 2^α (and α) follows now from standard Perron–Frobenius theory. □

Theorem 5.9. *Consider the iterated function system in* (3.2) *with the weights given by* (3.4). *Then for all Lipschitz functions* f

$$\left\| L^n f(x) - \int_0^{1/2} f \, d\mu_E \right\|_{\mathrm{Lip}} \le C a^n \|f\|_{\mathrm{Lip}}, \tag{5.14}$$

where as usual $C>0$ *and* $0<a<1$.

Proof. It is easy to check that $\sup_{0 \le x \le 1} \sum_{j=1}^{3} p_j(x) |\tilde{\psi}_j'(x)| < 1$. (5.10) follows from the same argument as in the proof of Theorem 5.8, but now with reference to the maps $\tilde{\psi}_1$ and $\tilde{\psi}_3$. □

Remarks. (1) It is easy to see that we can turn the statement of Theorem 5.9 into an upper estimate of the rate of convergence in the d_H-distance of L^*-iterates of any probability measure to μ_E using a similar type of argument as in the proof of Lemma 5.4.

(2) The value 2^α is given by $\lim_{n\to\infty} \|L^n 1\|_\infty^{1/n} = \lim_{n\to\infty} \|\sum_{|w|=n} q_w\|_\infty^{1/n}$, with reference to the same weights q_j as in the proof of Theorem 5.8. The iterates q_w, $|w|=n$, are defined multiplicatively (according to the definition of iterates of L)

as $q_{w_n}(\psi_{w_{n-1}}\circ\ldots\circ\psi_{w_1}(x))\ldots q_{w_1}(x)$. In the present case with the iterated function system (3.1) and the weights of Theorem 5.8 this product simplifies due to cancellation.

We have used $2^\alpha=\lim_{n\to\infty}\|\sum_{|w|=n}q_w\|_\infty^{1/n}$ to obtain the numerical approximations $2^\alpha\approx1.28121$ and $\alpha\approx0.35751$ as follows. We computed the function $\sum_{|w|=n}q_w$ for $n=1,2,3,4,5$ and observed that it was decreasing on $[0,\frac{1}{2}]$ and symmetric about $\frac{1}{2}$. Thus the supremum is the value at $x=0$. Assuming that this continues to hold for all n, we may compute

$$\left\|\sum_{|w|=n}q_w\right\|_\infty^{1/n}=\left(\sum_{|w|=n}q_w(0)\right)^{1/n}. \tag{5.15}$$

This simplification makes it feasible to carry the computation out to $n=10$. We then fitted a straight line (least mean square) to the data $\left(n,\log\sum_{|w|=n}q_w(0)\right)$ for $3\leq n\leq10$ and took the slope of the line as our approximation for 2^α.

6. A string of pearls

In this section we describe the level sets $L(e,t)$ in SG as the restriction of a rectifiable curve $\gamma(e,t)$ in the plane to SG. Such a curve is not unique, of course; in particular, what it does in the complement of SG is rather arbitrary. But the construction of the curve will specify an order on the level set. The level set will appear as a Cantor set in γ, and so we can imagine γ as consisting of some "string" $\gamma\backslash L$ on which the "pearls", namely L, are strung.

Our approach is based on the standard embedding of SG in the plane. The theory of harmonic functions on SG is based only on the topology of SG, not on its geometry. There is a more natural metric on SG, called the effective resistance metric [K3], but it is not even known whether or not SG embeds isometrically for this metric in any Euclidean space. There are also other natural embeddings of SG in the plane which are harmonic (in fact energy minimizing) [K2]. For a harmonic embedding the level sets are nothing other than the restriction to SG of straight lines in the plane. However, it is not clear how to pull back the straight lines from the harmonic embedding to the standard embedding, since we do not know how to extend the identification of SG in the two embeddings to a mapping of the complements of SG in the two embeddings. Nevertheless, it is not hard to see that the ordering on the level set induced by the ordering of the straight line in the harmonic embedding is the same as the ordering that is induced by our string of pearls γ.

We now give an inductive description of γ for a harmonic function in standard position and $L(e,t)$. The key observation we use is Theorem 2 in [DSV] which says that the restriction of h to two of the edges (p_1,p_3) and (p_1,p_2) if $e\geq\frac{1}{2}$ (or (p_2,p_3) if $e\leq\frac{1}{2}$) is monotone increasing, while on the third edge ((p_1,p_2) if $e<\frac{1}{2}$ or (p_2,p_3) if $e>\frac{1}{2}$) it has a single extremum. In all cases $L(e,t)$ intersects the edge (p_1,p_3) in a single point, and that point will be the initial point of γ. Except in the trivial cases $t=0$ or 1, $L(e,t)$ will intersect the other two edges in one or three points, with exactly one of the edges meeting $L(e,t)$ in one point (there is a degenerate case when we consider $L(e,t)$ meeting an edge in a point of multiplicity 2). That point is the terminal point of γ.

If $L(e,t)$ lies entirely in $F_1(SG)$ or $F_3(SG)$, we pass to the next level and iterate the process. This happens in the first case if $t\leq e\leq\frac{1}{3}$, or $\frac{1}{3}\leq e$ and $t\leq\frac{1}{5}(2e+1)$. This happens in the second case if $\frac{2}{3}\leq e\leq t$, or $\frac{2}{3}\geq e$ and $t\geq\frac{1}{5}(2e+2)$.

If $L(e,t)$ lies in $F_2(SG)$ and one of the others, say $F_1(SG)$ (it cannot lie in all three except in the degenerate case $t=\frac{1}{5}(e+2)$) we consider two separate cases. Note that $t\leq\frac{1}{5}(e+2)$.

Case I: $e\leq\frac{1}{3}$ *and* $e\leq t\leq\frac{1}{5}(2e+1)$. Then $L(e,t)$ intersects the edge (p_2,p_3) exactly once, and this point lies in $F_2(SG)$ and is the terminal point of γ. But $L(e,t)$ intersects the edge (p_1,p_2) twice, once at a point q_1 in $F_1(SG)$ and once at a point q_2 in $F_2(SG)$ (the degenerate case $t=\frac{1}{5}(2e+1)$ has $q_1=q_2$). We place a curve (string) joining q_1 to q_2 and going outside SG in the middle of γ. On one side we have the $F_1(SG)$ part of γ joining the initial point of γ to q_1, and on the other side we have the $F_2(SG)$ part of γ joining q_2 to the terminal point of γ. We iterate to determine these two parts of γ.

Case II: $\frac{1}{5}(2e+1)\leq t\leq\frac{1}{5}(e+2)$. In this case $L(e,t)$ intersects both the segments $(F_1(p_3),F_1(p_2))$ and $(F_2(p_1),F_2(p_3))$ exactly once, in points q_1 and q_2, respectively (when $t=\frac{1}{5}(2e+1)$ we have $q_1=q_2$). We place the line segment (string) joining q_1 and q_2 (passing through the largest inverted triangle in the complement of SG) in the middle of γ. On one side we have the $F_1(SG)$ part of γ joining the initial point of γ to q_1, and on the other side we have the $F_2(SG)$ part of γ joining q_2 to the terminal point of γ. We iterate to determine these two parts of γ.

Note that when we iterate the orientation of the parts of γ in $F_i(SG)$ may be switched, but that does not matter. It is easy to see that γ is rectifiable, since at each level m we add at most $N(m,e,t)$ pieces of string, each of length bounded by a constant multiple of 2^{-m}, and the sum

$$\sum_{m=1}^{\infty}\frac{1}{2^m}N(m,e,t)$$

is convergent because $\alpha<1$.

We now give an example to show that γ may have points around which it spirals infinitely often. In particular this means that γ cannot be chosen to be a C^1 curve. In fact, in our example an initial segment of γ will be a self-similar spiral.

The idea is to find a fixed point (e_0,t_0) for the mapping $\varphi_3\circ\varphi_{23}\circ\varphi_1\circ\varphi_1\circ\varphi_{21}$. We will verify that $e_0\leq\frac{1}{3}$ so that φ_{21} is the correct mapping for the intersection of $L(e_0,t_0)$ with $F_2(SG)$, and this changes the orientation (permuting p_1 and p_2). The two mappings φ_1 that follow (now corresponding to F_2) preserve orientation. We will verify that the e value of $\varphi_1(\varphi_1(\varphi_{21}(e_0,t_0)))$ lies in $\frac{2}{3}\leq e\leq 1$, so that φ_{23} is the correct mapping for the intersection of $L(e_0,t_0)$ with $F_2\circ F_2\circ F_2\circ F_1(SG)$. This again changes the orientation (permuting p_2 and p_3), so the net result of the two changes is a rotation R through $-\frac{2}{3}\pi$. The final mapping φ_3 (corresponding now to F_1) preserves orientation. Thus we will have

$$L(e_0,t_0)|_{F_w(SG)}=RF_wL(e_0,t_0) \tag{6.1}$$

for $w=(2,2,2,1,1)$ and we may iterate this indefinitely. Let γ' denote the initial segment of γ that contains $L(e_0,t_0)\cap F_1(SG)$. This is followed by a string segment γ'' connecting q_1 to q_2 (Case I). Then q_2 is the initial point in $L(e_0,t_0)\cap F_w(SG)$, so the next segment of γ is $RF_w\gamma'$, and so on. Thus we have the infinite self-similar spiral γ', γ'', $RF_w\gamma'$, $RF_w\gamma''$, $R^2F_{w^2}\gamma'$, ..., as an initial segment of γ.

To find the fixed point (e_0,t_0) we first solve for e_0. We obtain the quadratic equation $(14-27x)/(61-135x)=x$ and we choose the solution $e_0=\frac{1}{135}(44-\sqrt{46}\,)=0.2756864\ldots$ which satisfies $e_0\leq\frac{1}{3}$. We then find that the e value after applying $\varphi_1\circ\varphi_1\circ\varphi_{21}$ is $(21+\sqrt{46}\,)/(26+\sqrt{46}\,)=0.8474788\ldots$ which is above $\frac{2}{3}$ as required.

If we let h be the harmonic function with eccentricity e_0, then $h|_{F_w(SG)}$ assumes all values between the minimum $h(F_wp_2)$ and the maximum $h(F_wp_1)$. The t-component of $\varphi_3\circ\varphi_{23}\circ\varphi_1\circ\varphi_1\circ\varphi_{21}$ is just the affine map of this interval to $[0,1]$. Clearly this has a unique fixed point t_0.

A related example involves finding a solution of

$$\varphi_{22}(\varphi_1(\varphi_{21}(e_0,t_0)))=(1-e_0,t_0). \tag{6.2}$$

Then

$$L(e_0,t_0)|_{F_w(SG)}=R^{-1}F_wL(e_0,t_0) \tag{6.3}$$

for $w=(2,2,1)$, but now the ordering on γ is reversed. Thus we must iterate a second time before we come back to the initial order. In this case

$$(e_0,t_0)=\left(\frac{\sqrt{22}-2}{9},\frac{89\sqrt{22}-16}{9(132-\sqrt{22}\,)}\right). \tag{6.4}$$

Figure 6.1 illustrates the construction of γ for this example, including the first four steps in the algorithm, a "complete" approximation after nine levels (beyond that it is impossible to see any changes) and a highlighting of the self-similar region.

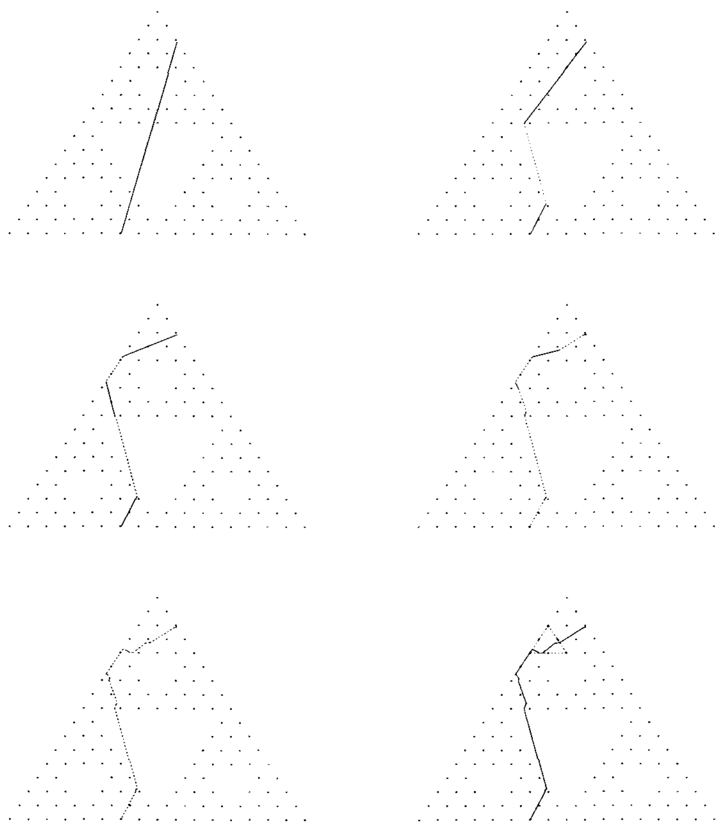


Figure 6.1. An example of the construction of γ for the parameter values (6.4). The first four steps of the algorithm are shown, with dotted lines for the string segments (always taken as straight line segments). The SG is outlined in dots in the background. The "completed" approximation at level nine is shown alone, and with the level three triangle highlighting the region where it is self-similar.

7. The hexagasket

To illustrate the kind of behavior we may expect for general post-critically finite fractals having vertices that are neither boundary points nor junction points, we describe briefly the case of the hexagasket (or fractal star of David) shown in Figure 7.1(a). To simplify the discussion we will take the boundary to consist of just three points, every other vertex of a regular hexagon H. The iterated function system consists of six similarities with contraction ratio $\frac{1}{3}$, so that each $F_i(H)$ contains one of the vertices of H, but the ones associated to the nonboundary vertices must include a rotation through angle $\frac{1}{3}\pi$ so that the intersections $F_i(H) \cap F_{i+1}(H)$ occur at images of the three boundary vertices. See [S2] for the details of the construction of a harmonic structure on the hexagasket analogous to the theory on SG sketched in the introduction.

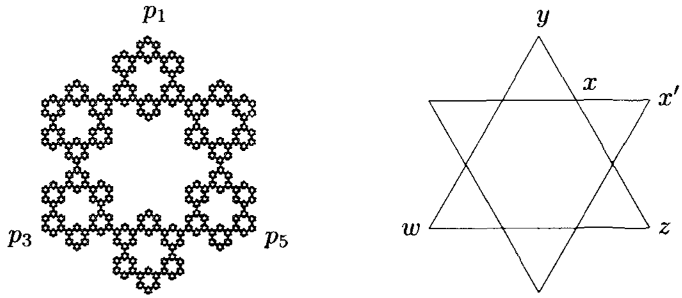


Figure 7.1. (a) The hexagasket, with boundary points p_1, p_3 and p_5. (b) The relative placement of the points x, x', y, z and w.

We define V_m as before, but we note that there are two distinct types of nonboundary vertices: the *junction points*, which belong to two distinct cells of level m and have four neighbors in Γ_m, and the *nonjunction points*, which belong to only one cell of level m and have two neighbors in Γ_m. The condition for h to be a harmonic function is (1.2) for junction points, and the analogous formula with $\frac{1}{4}$ replaced by $\frac{1}{2}$ for nonjunction points. The harmonic extension algorithm (1.3) also must be replaced by two formulas,

$$h(x) = \tfrac{4}{7}h(y) + \tfrac{2}{7}h(z) + \tfrac{1}{7}h(w) \tag{7.1}$$

if x is a junction vertex in $V_m \setminus V_{m-1}$, and y, z and w are three junction vertices in V_{m-1} closest to x, and

$$h(x') = \tfrac{3}{7}h(y) + \tfrac{3}{7}h(z) + \tfrac{1}{7}h(w) \tag{7.2}$$

if x' is a nonjunction vertex in $V_m \setminus V_{m-1}$ (here y and z are equidistant to x'). See Figure 7.1(b).

Writing the harmonic extension algorithm in the form (2.1), we find that the six matrices M_i are of two types,

$$M_1 = \frac{1}{7}\begin{pmatrix} 7 & 0 & 0 \\ 4 & 2 & 1 \\ 4 & 1 & 2 \end{pmatrix} \quad \text{and} \quad M_2 = \frac{1}{7}\begin{pmatrix} 3 & 1 & 3 \\ 4 & 1 & 2 \\ 2 & 1 & 4 \end{pmatrix} \tag{7.3}$$

and their cyclic permutations, with the odd indices associated to junction points and the even indices associated to nonjunction points. Note that the second type of matrix has rank two, so it is not invertible. In particular, this means that the theory of products of random matrices is no longer applicable.

We may define eccentricity exactly as before. The images of the matrices M_{2k} correspond to harmonic functions with eccentricity $\frac{1}{2}$, or to constant functions. We obtain a constant function just in the case that the initial eccentricity is 0 or 1. Thus if we define mappings ψ_j as before we have $\psi_2(x)=\psi_4(x)=\psi_6(x)=\frac{1}{2}$ except for $x=0,1$. Furthermore a simple computation shows that the remaining ψ's are identical to the functions in (3.1), with the indices suitably changed.

A nonconstant harmonic function may be constant on an entire cell, and this happens with initial eccentricity e if and only if there exists a word w with $\psi_w(e)=0$ or $\psi_w(e)=1$. There is a countable set E of rational numbers e for which this holds. Presumably E is dense. It seems likely that $\frac{1}{2}$ does not belong to E. The rest of this discussion is predicated on this assumption.

Let A denote the set of numbers $\psi_w\left(\frac{1}{2}\right)$ as w varies over all words. We are assuming that A does not contain 0 or 1. If we start with any initial eccentricity and look at the harmonic function restricted to cells of level m, we will find more than half of them with eccentricity $\frac{1}{2}$, and at least a set proportion with eccentricity equal to each number in A (for fixed m only a finite number will occur, because the proportions tend to zero). In fact, only the cells associated with odd words (all w_j odd) will have eccentricity that depends on the initial eccentricity. Since there are only 3^m out of 6^m such cells, we see a very strong form of the geography is destiny principle: aside from the presumably nongeneric possibility that the restriction to a cell is constant, the eccentricity of the restriction of a harmonic function to a cell of level m is exactly determined by the location of the cell, not the particular harmonic function, with the exception of 3^m out of the 6^m cells.

The analog of the measure μ is therefore discrete, not continuous, with atoms at each of the points in A. In fact it is easy to see that

$$\int_0^1 f\,d\mu = \frac{1}{2}\left(f\left(\frac{1}{2}\right) + \sum_{w \text{ odd}} \frac{1}{6^{|w|}} f\left(\psi_w\left(\frac{1}{2}\right)\right)\right). \tag{7.4}$$

This is not quite an explicit formula because each $a \in A$ may be represented as $\psi_w(\frac{1}{2})$ with w odd in many ways. Similarly, the measure μ_E is discrete with atoms in A, but with different weights. A simple computation shows that the energy of a harmonic function in a cell $F_w H$ splits at the next level as follows: $\frac{2}{7}$ of the energy goes into the three even cells $F_{wk}H$ (k even) and the remaining $\frac{5}{7}$ splits according to weights $p_k(e)$ (k odd) given by the analog of (3.4) with the factor $\frac{1}{5}$ replaced by $\frac{1}{7}$, and the indices relabeled. Thus

$$\int_0^1 f\,d\mu_E = \frac{2}{7}\left(f\left(\frac{1}{2}\right) + \sum_{w \text{ odd}} p_w\left(\frac{1}{2}\right) f\left(\psi_w\left(\frac{1}{2}\right)\right)\right), \tag{7.5}$$

where p_w is defined as in Remark (2) following Theorem 5.9.

We may parameterize level sets by $L(e,t)$ as before. Nongeneric level sets, for $e \in E$ and the appropriate choice of t, will contain entire cells and so have dimension equal to the dimension of the hexagasket ($\log 6/\log 3$). Generic level sets will avoid this difficulty (under our assumption $\frac{1}{2} \notin E$), and can be described in terms of level sets with $e \in A$ (we can safely throw away the contributions from odd words). Thus the measure ν analogous to (4.3) will be the product of a discrete measure μ_L with atoms in A and uniform measure in t. We omit the details.

References

[Ba] Barlow, M. T., *Diffusion on Fractals,* Lecture Notes in Math. **1690**, Springer-Verlag, Berlin–Heidelberg, 1998.

[BDEG] Barnsley, M. F., Demko, S. G., Elton, J. H. and Geronimo, J. S., Invariant measures for Markov processes arising from place-dependent probabilities, *Ann. Inst. H. Poincaré Probab. Statist.* **24** (1988), 367–394.

[BST] Ben-Bassat, O., Strichartz, R. and Teplyaev, A., What is not in the domain of the Laplacian on a Sierpiński gasket type fractal, *J. Funct. Anal.* **166** (1999), 197–217.

[Bo] Bougerol, P., Limit theorems for products of random matrices, in *Products of Random Matrices with Applications to Schrödinger Operators* (Bougerol, P. and Lacroix, J., authors), pp. 1–180, Birkhäuser, Basel–Boston, 1985.

[DSV] Dalrymple, K., Strichartz, R. S. and Vinson, J. P., Fractal differential equations on the Sierpiński gasket, *J. Fourier Anal. Appl.* **5** (1999), 203–284.

[H] Hennion, H., Sur un théorème spectral et son application aux noyaux lipschitziens, *Proc. Amer. Math. Soc.* **118** (1993), 627–634.

[KL] Keller, G. and Liverani, C., Stability of the spectrum for transfer operators, *Ann. Scuola Norm. Sup. Pisa Cl. Sci.* **28** (1999), 141–152.

[K1] Kigami, J., A harmonic calculus on the Sierpiński spaces, *Japan J. Appl. Math.* **8** (1989), 259–290.

[K2] KIGAMI, J., Harmonic metric and Dirichlet form on the Sierpiński gasket, in *Asymptotic Problems in Probability Theory: Stochastic Models and Diffusions on Fractals (Sanda/Kyoto, 1990)* (Elworthy, K. D. and Ikeda, N., eds.), pp. 201–218, Pitman Research Notes in Math. **283**, Longman, Harlow, 1993.

[K3] KIGAMI, J., Effective resistances for harmonic structures on p.c.f. self-similar sets, *Math. Proc. Cambridge Philos. Soc.* **115** (1994), 291–303.

[K4] KIGAMI, J., *Analysis on Fractals*, Cambridge Univ. Press, Cambridge, 2001.

[Ku] KUSUOKA, S., Dirichlet forms on fractals and products of random matrices, *Publ. Res. Inst. Math. Sci.* **25** (1989), 659–680.

[Ö] ÖBERG, A., Approximation of invariant measures for iterated function systems, *Ph.D. thesis, Umeå University,* 1998.

[R] RACHEV, S. T., *Probability Metrics and the Stability of Stochastic Models*, Wiley, Chichester, 1991.

[S1] STRICHARTZ, R., Analysis on fractals, *Notices Amer. Math. Soc.* **46** (1999), 1199–1208.

[S2] STRICHARTZ, R., Taylor approximations on Sierpiński gasket type fractals, *J. Funct. Anal.* **174** (2000), 76–127.

[T] TEPLYAEV, A., Gradients on fractals, *J. Funct. Anal.* **174** (2000), 128–154.

Received June 28, 2001

Anders Öberg
Department of Mathematics
Uppsala University
P. O. Box 480
SE-751 06 Uppsala
Sweden
email: anders@math.uu.se

Robert S. Strichartz
Mathematics Department
Malott Hall
Cornell University
Ithaca, NY 14853
U.S.A.
email: str@math.cornell.edu

Andrew Q. Yingst
Mathematics Department
University of North Texas
Denton, TX 76203
U.S.A.
email: andyq@unt.edu