TRANSITION DENSITY ESTIMATES FOR DIFFUSION PROCESSES ON POST CRITICALLY FINITE SELF-SIMILAR FRACTALS

B. M. HAMBLY *and* T. KUMAGAI

[Received 16 May 1997—Revised 18 December 1997]

1. *Introduction*

The recent development of analysis on fractal spaces is physically motivated by the study of diffusion in disordered media. The natural questions that arise concern the existence and uniqueness of a suitable Laplace operator, and the behaviour of the associated heat semigroup, on a space which is fractal. The classes of fractals for which these questions were first answered were classes of exactly self-similar fractals, with strong spatial symmetry, such as nested or affine nested fractals (see, for example, [**2**, **8**]). The existence of a Laplacian and estimates on the heat kernel were obtained by considering the associated diffusion process and using the symmetry of the space. The uniqueness of the Laplacian for nested and affine nested fractals has recently been solved through consideration of their Dirichlet forms [**23**].

In [**15**] the framework of post critically finite (which we abbreviate to p.c.f.) self-similar sets was introduced in order to capture the notion of exactly self-similar finitely ramified fractals as used in the physics literature. Finitely ramified fractals have the property that the intersection of any connected subset of the fractal with the rest of the set should occur only at a finite number of points. This makes these structures much easier to analyse than infinitely ramified sets such as the Sierpinski carpet [**5**]. The p.c.f. self-similar sets do not have spatial symmetry in general and have provided a mathematical test bed for analysis on fractals. In this paper we will obtain uniform short time estimates on the heat kernel associated with a natural Laplacian on the fractal.

A Laplacian can be constructed on a p.c.f. self-similar fractal as a limit of discrete Laplacians on graph approximations to the fractal based on the ramification points. We construct these operators via their Dirichlet forms, which can be set on any L^2-space with a full measure, ν. In [**23**], criteria are given which give a partial answer to the existence and uniqueness of Laplace type operators on p.c.f. self-similar sets.

In a series of papers [**16**, **17**, **18**, **20**, **25**, **7**] some of the interesting spectral properties of p.c.f. sets have been elucidated. It has been shown that there can exist localized eigenfunctions if there is a high degree of symmetry in the set and corresponding Laplace operator. This corresponds to oscillation in the leading order term for the asymptotics of the eigenvalue counting function. Such behaviour occurs for nested fractals [**7**] and p.c.f. sets with strong harmonic

1991 *Mathematics Subject Classification*: 60J60, 60J25, 28A80, 31C25.

Proc. London Math. Soc. (3) 78 (1999) 431–458.

structure [**25**]. However, for general p.c.f. sets with an arbitrary Bernoulli measure the oscillation only occurs under a certain condition [**18**].

From [**17**], [**18**] and [**10**] it has become clear that the analytic properties of these sets are best discussed in terms of an effective resistance metric. If we write d_s^ν for the spectral exponent which describes the asymptotic scaling of the eigenvalue counting function for the Laplacian defined with respect to the measure ν, then the spectral dimension d_s is defined to be the number which maximizes d_s^ν over all measures ν. The maximizing measure is equivalent to the Hausdorff measure in the effective resistance metric and the spectral dimension $d_s = 2S/(S+1)$, where S is the Hausdorff dimension of the set in the effective resistance metric [**16**]. We will be concerned here with the behaviour of the heat kernel associated with the Laplacian with respect to the maximizing measure.

The heat kernel associated with the Laplace operator on fractals has been considered probabilistically for some sub-classes of p.c.f. sets, such as nested and affine nested fractals [**19**, **8**], as it is the transition density of the diffusion process generated by the operator. We will extend this work to general p.c.f. self-similar sets by obtaining best possible heat kernel estimates in terms of the effective resistance metric.

By applying the results to some examples we show that uniform Aronson type estimates do not hold in general on fractals. This is a situation in which we have a Poincaré inequality and a doubling condition for the measure, but Aronson type estimates do not hold. For the Laplace–Beltrami operator on a Riemannian manifold, it was shown independently in [**24**, **9**] that a Poincaré inequality and a doubling condition are equivalent to a parabolic Harnack inequality. Using this result one can obtain upper and lower bounds on the heat kernel of Aronson type. In order to get sharp off-diagonal estimates we need to investigate the relationship between the short paths on the fractal and the effective resistance metric. The results show that in general, given a time and two points in the set, the heat kernel estimate is not influenced by structure in the fractal below a certain scale, determined by the relationship between the time and effective resistance between the points.

The plan of the paper is as follows. In § 2 we will introduce p.c.f. self-similar sets, their Dirichlet forms and the effective resistance metric $R(x,y)$. This will be followed by some preliminary results in § 3, which will be used to obtain the estimates. In particular, we introduce the shortest path counting function, $N_m(x,y)$, for the effective resistance metric and give a version of the Einstein relation for p.c.f. self-similar sets; expressing the random walk dimension in the effective resistance metric as $d_w = S+1$. In §§ 4 and 5 we will deduce the following upper and lower uniform short time estimates on the heat kernel for the p.c.f. self-similar set K.

THEOREM 1.1. *There exist constants* $c_{1.1}, c_{1.2}, c_{1.3}, c_{1.4} > 0$ *such that for all* $x, y \in K$ *and* $0 < t < 1$, *if*

$$e^{-m-1} \leqslant R(x,y) \leqslant e^{-m}$$

and

$$k(m,t) := \inf\{j\colon N_{m+j}(x,y)e^{-(S+1)(m+j)} < t\},$$

then

$$p_t(x,y) \leqslant c_{1.1} t^{-d_s/2} \exp(-c_{1.2} N_{m+k(m,t)}(x,y)),$$

and

$$p_t(x, y) \geqslant c_{1.3} t^{-d_s/2} \exp(-c_{1.4} N_{m+k(m,t)}(x, y)).$$

Note that this theorem contains the diagonal estimate as $N_{m+k(m,t)}(x, x) = 0$ for all $m \in \mathbb{N}$, $0 < t < 1$ and $x \in K$. We can write this in terms of the effective resistance and a chemical exponent for this metric, $d_k^c(x, y) = \log N_{[\log R(x,y)]+k}(x, y)/k$ as follows.

COROLLARY 1.2. *There exist constants* $c_{1.5}, c_{1.6}, c_{1.7}, c_{1.8}$ *such that, for* $x, y \in K$ *and* $0 < t < 1$, *with* m *and* k *as above,*

$$p_t(x, y) \leqslant c_{1.5} t^{-d_s/2} \exp\left(-c_{1.6} \left(\frac{R(x, y)^{S+1}}{t} \right)^{d^c_{k(m,t)}(x,y)/(S+1-d^c_{k(m,t)}(x,y))} \right),$$

and

$$p_t(x, y) \geqslant c_{1.7} t^{-d_s/2} \exp\left(-c_{1.8} \left(\frac{R(x, y)^{S+1}}{t} \right)^{d^c_{k(m,t)}(x,y)/(S+1-d^c_{k(m,t)}(x,y))} \right).$$

In the final section we will show how Theorem 1.1 can be used in several examples. The first will be the affine nested fractals of [**8**], where Aronson type estimates can be obtained. We then consider a one-parameter family of diffusions on the Sierpinski gasket which are not spatially symmetric and have a multifractal Bernoulli measure for invariant measure. The heat kernel estimates in this case are not of Aronson type. We will also discuss the non-unique diffusion on the Vicsek set and show that each diffusion can be controlled by estimates of the same functional form but different constants. The *abc*-gaskets of [**12**] present a family of p.c.f. self-similar sets with no spatial symmetry. We can obtain heat kernel estimates in this case. We will conclude with two conjectures concerning the chemical exponent.

2. *P.C.F. self-similar fractals and their Dirichlet forms*

In this section we briefly introduce post critically finite (p.c.f. for short) self-similar sets and summarize known results about their Dirichlet forms. This class was introduced in [**15**] and a detailed discussion can be found in [**2**]. At the end of the definitions we will describe the Sierpinski gasket (see Fig. 1) as a p.c.f. fractal. First, we introduce the one-sided shift space and give some notation.

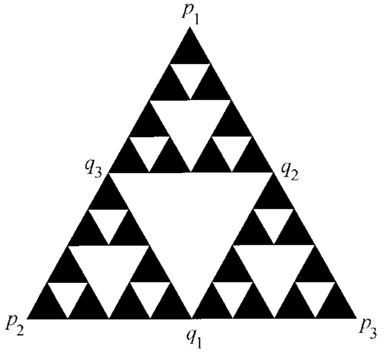

FIG. 1. The Sierpinski gasket.

NOTATION. (1) Let $S = \{1, 2, \dots, N\}$. The one-sided shift space Σ is defined by $\Sigma = S^{\mathbb{N}}$. Also, let $W_n = S^n$.

(2) For $w \in \Sigma$, we denote the ith element in the sequence by w_i and write $w = w_1 w_2 w_3 \dots$.

(3) If $w \in W_n$, we define $|w| = n$.

(4) We denote $\dot{i} = iiii \dots$ for $i \in S$.

(5) Let σ: $\Sigma \to \Sigma$ be the shift map, that is, $\sigma w = w_2 w_3 \dots$ if $w = w_1 w_2 \dots$. Define $\tilde{\sigma}_s$: $\Sigma \to \Sigma$ as $\tilde{\sigma}_s w = sw$ for $s \in S$.

We now introduce the notion of self-similar structures and p.c.f. self-similar sets.

DEFINITION 2.1. Let K be a compact metrizable space and for each $s \in S$, let F_s: $K \to K$ be a continuous injection. Then, $(K, S, \{F_s\}_{s \in S})$ is said to be a *self-similar structure* on K if there exists a continuous surjection π: $\Sigma \to K$ such that $\pi \circ \tilde{\sigma}_s = F_s \circ \pi$ for every $s \in S$.

For $w \in W_n$, we denote $F_w = F_{w_1} \circ F_{w_2} \circ \dots \circ F_{w_n}$ and $K_w = F_w(K)$. In particular, $K_s = F_s(K)$ for $s \in S$. Note that Hutchinson's self-similar set, as defined in [**14**], is a self similar structure in the sense of Definition 2.1 by taking

$$\pi(w) = \bigcap_{n \geqslant 1} F_{w_1} \circ F_{w_2} \circ \dots \circ F_{w_n}(K).$$

DEFINITION 2.2. Let $\mathscr{L} = (K, S, \{F_s\}_{s \in S})$ be a self-similar structure on K. Then the *critical set* of $\mathscr{L}$ is defined by

$$C(\mathscr{L}) = \pi^{-1}\Bigg(\bigcup_{s,t \in S, s \neq t} (K_s \cap K_t)\Bigg),$$

and the *post critical set* of $\mathscr{L}$ is defined by

$$P(\mathscr{L}) = \bigcup_{n \geqslant 1} \sigma^n(C(\mathscr{L})).$$

We say that $\mathscr{L}$ is *post critically finite* (*p.c.f.*) if $P = P(\mathscr{L})$ is a finite set.

In the following, we only consider a connected p.c.f. self-similar set $(K, S, \{F_s\}_{s \in S})$.

NOTATION. (1) For $m \geqslant 0$, let

$$P^{(m)} = \bigcup_{w \in W_m} wP, \quad V_m = \pi(P^{(m)}), \quad V_* = \bigcup_{m \geqslant 0} V_m \quad \text{and} \quad \mathring{V}_m = V_m - V_0.$$

Moreover, $B_w = F_w(\pi(P))$ for $w \in W_m$ for any $m \geqslant 0$.

(2) For the finite sets V, V', we define

$$l(V) = \{f \mid f\colon V \to \mathbb{R}\},$$

$$L(V, V') = \{A \mid A\colon l(V) \to l(V') \text{ and } A \text{ is linear}\},$$

$$L(V) = L(V, V).$$

DEFINITION 2.3. A pair $(D, r) \in L(V_0) \times l(S)$ is called a *quasi-harmonic structure* on K if it satisfies the following:

(1) $r_s > 0$ for each $s \in S$,

(2) $D = {}^tD$,

(3) D is irreducible,

(4) $D_{pp} < 0$, $\sum_{q \in V_0} D_{pq} = 0$ for each $p \in V_0$,

(5) $D_{pq} \geqslant 0$ if $p \neq q$.

From the quasi-harmonic structure (D, r), we have a difference operator H_m on V_m.

DEFINITION 2.4. A *difference operator* $H_m \in L(V_m)$ is defined by

$$H_m = \sum_{w \in W_m} r_w^{-1}\, {}^tR_w D R_w,$$

where R_w: $l(V_m) \to l(V_0)$ is defined by $R_w(u) = u \circ F_w$ and $r_w = r_{w_1} \dots r_{w_m}$.

We decompose H_m into

$$H_m f = \begin{pmatrix} T_m & {}^tJ_m \\ J_m & X_m \end{pmatrix} \cdot \begin{pmatrix} f|_{V_0} \\ f|_{\mathring{V}_m} \end{pmatrix},$$

where $T_m \in L(V_0)$, $J_m \in L(V_0, \mathring{V}_m)$, $X_m \in L(\mathring{V}_m)$. We will write $T = T_1$, $J = J_1$, and $X = X_1$.

We now give the notion of the harmonic structure, under which harmonic functions on V_m (with respect to H_m) automatically become harmonic functions on V_{m-1} (with respect to H_{m-1}).

DEFINITION 2.5. A quasi-harmonic structure (D, r) is called a *harmonic structure* if there exists $\lambda > 0$ such that

$$T - {}^tJX^{-1}J = \lambda^{-1}D. \tag{2.1}$$

Furthermore, a harmonic structure (D, r) is said to be *regular* if $r_s < \lambda$ for each $s \in S$.

Let $\rho_i = \lambda / r_i$ $(1 \leqslant i \leqslant N)$, which is greater than 1 if the harmonic structure is regular. Note that regularity of the harmonic structure is a condition for the reproducing kernel of the following quadratic form $(\mathscr{E}, \mathscr{F})$ to be bounded continuous. Throughout this paper, we treat a regular harmonic structure (D, r).

The uninitiated reader is advised to think of the above definitions in the context of the 2-dimensional Sierpinski gasket, Fig. 1. This can be considered as a p.c.f. self-similar set by setting

$$S = \{1, 2, 3\},$$

$$\pi(C) = \{q_1, q_2, q_3\}, \qquad \pi^{-1}(q_1) = \{2\dot{3}, 3\dot{2}\},$$

$$\pi^{-1}(q_2) = \{1\dot{3}, 3\dot{1}\}, \qquad \pi^{-1}(q_3) = \{1\dot{2}, 2\dot{1}\},$$

$$\pi(P) = \{p_1, p_2, p_3\}, \quad \pi^{-1}(p_i) = \{\dot{i}\} \text{ for } i = 1, 2, 3.$$

The set V_0 is then the vertices of the largest triangle and V_m denotes the set of vertices of all triangles of size 2^{-m}.

The natural Laplace operator on the Sierpinski gasket can be constructed from a quasi-harmonic structure given by

$$D = \begin{pmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{pmatrix}, \quad r = (1,1,1).$$

The matrix D is the generator of a continuous time random walk on the complete graph formed by the vertices V_0 with unit conductors on each edge. The equation (2.1) is satisfied with $\lambda = \frac{5}{3}$ and there exists a unique regular harmonic structure. Thus the difference operator H_m is the generator of the continuous time nearest neighbour random walk on the graph formed from V_m with edges which are images of the edges of V_0 and of conductivity $\left(\frac{5}{3}\right)^n$. The fact that the structure is harmonic means that the random walk on V_m is *decimation invariant* in that, if it is stopped when it hits V_0, it has the same probabilistic structure as the random walk on V_0. Another way to view the existence of a harmonic structure is that it ensures that the networks (V_0, H_0) and (V_m, H_m) are electrically equivalent.

DEFINITION 2.6. We say that $f \in C(K) = \{f\colon f \text{ is a continuous function on } K\}$ is *m-harmonic* if and only if $H_n f|_{V_n \setminus V_m} = 0$ for all $n \geqslant m$. A 0-harmonic function is called a *harmonic* function.

PROPOSITION 2.7 [**15**, Theorem 4.12]. *Let (D,r) be a harmonic structure. For any $\rho \in l(V_m)$, there exists a unique m-harmonic function f with $f|_{V_m} = \rho$.*

DEFINITION 2.8. For $f \in l(V_*)$, we define $P_m f$ to be a continuous function on K satisfying the following:

(1) $P_m f|_{V_m} = f|_{V_m}$,

(2) $H_n(P_m f)|_{V_n \setminus V_m} = 0$ for all $n > m$.

By Proposition 2.7, if (D,r) is a harmonic structure, $P_m f$ exists uniquely for all $f \in l(V_*)$ and $m \geqslant 1$.

For $u, v \in l(V_m)$, define

$$\mathscr{E}_m(u,v) = -\lambda^{m}\, {}^t u H_m v.$$

By Corollary 6.14 in [**15**], we see that $\mathscr{E}_m(u,u) \leqslant \mathscr{E}_{m+1}(u,u)$ for $u \in l(V_{m+1})$ (equality holds if and only if $P_m u|_{V_{m+1}} = u$). Using this fact, let

$$\mathscr{F} = \{f \in l(V_*)\colon \lim_{m\to\infty} \mathscr{E}_m(f,f) < \infty\}, \quad \mathscr{E}(f,g) = \lim_{m\to\infty} \mathscr{E}_m(f,g) \text{ for all } f,g \in \mathscr{F}.$$

In order to embed $\mathscr{F}$ into some $\mathbb{L}^2$-space, we define the following effective resistance between $p, q \in V_*$:

$$R(p,q)^{-1} = \inf\{\mathscr{E}(f,f)\colon f \in l(V_*), f(p) = 1, f(q) = 0\}.$$

PROPOSITION 2.9 [**16, 17**]. (1) *The function $R(\cdot,\cdot)$ is a metric on V_*. It can be extended to a metric on K (which will be denoted by the same symbol R), and it induces the same topology as the original one on K.*

(2) *For* $p, q \in V_*$ *with* $p \neq q$,

$$R(p,q) = \sup\{|f(p) - f(q)|^2 / \mathscr{E}(f,f) \colon f \in \mathscr{F}, f(p) \neq f(q)\}.$$

We remark that the regularity condition for (D, r) is needed for $R(\cdot,\cdot)$ to be a metric on K. By Proposition 2.9(2), we know that $|f(p) - f(q)|^2 \leqslant R(p,q)\mathscr{E}(f,f)$ for all $f \in \mathscr{F}$ and $p, q \in V_*$, so that $f \in \mathscr{F}$ can be extended uniquely to a continuous function on K (we thus consider $\mathscr{F} \subset C(E)$). Now, let μ be a Bernoulli measure on K such that $\mu(F_i(K)) = \mu_i > 0$ (note that $\sum_{i=1}^N \mu_i = 1$). Then, $\mathscr{F} \subset C(E) \subset \mathbb{L}^2(K,\mu)$ and we have the following theorem [**15**, **16**, **18**].

THEOREM 2.10. *The form* $(\mathscr{E}, \mathscr{F})$ *is a local regular Dirichlet form on* $\mathbb{L}^2(K,\mu)$ *which has the following property:*

$$|f(p) - f(q)|^2 \leqslant R(p,q)\mathscr{E}(f,f) \quad \text{for all } f \in \mathscr{F} \text{ and all } p, q \in K, \tag{2.2}$$

$$\mathscr{E}(f,g) = \sum_{i=1}^N \rho_i \mathscr{E}(f \circ F_i, g \circ F_i) \quad \text{for all } f, g \in \mathscr{F}. \tag{2.3}$$

Further, if we set $\mathscr{E}_\beta(\cdot,\cdot) = \mathscr{E}(\cdot,\cdot) + \beta(\cdot,\cdot)_{\mathbb{L}^2(K,\mu)}$ *for* $\beta > 0$, *then,* $\mathscr{E}_\beta$ *admits a positive symmetric continuous reproducing kernel* $g_\beta^K(\cdot,\cdot)$.

If we let Δ_μ be the generator for $(\mathscr{E}, \mathscr{F})$ on $\mathbb{L}^2(K,\mu)$, then, by the above theorem, we know that $-\Delta_\mu$ has a compact resolvent and hence the spectrum of $-\Delta_\mu$ consists of eigenvalues. We will call the operator associated with any of the harmonic structures a Laplacian on the fractal. Our interest here is to obtain estimates on the heat kernel for the corresponding diffusion process. For this purpose, we need a 'natural' measure, which is suggested by the following theorem from [**18**].

THEOREM 2.11. (1) *Let* $n^\mu(x) = \sharp\{\lambda |\ \lambda$ *is an eigenvalue of* $-\Delta_\mu \leqslant x\}$. *Then, for the unique positive number* $d_s(\mu)$ *satisfying* $\sum_{i=1}^N (\mu_i/\rho_i)^{d_s(\mu)/2} = 1$, *the following inequality holds:*

$$0 < \liminf_{x \to \infty} n^\mu(x)/x^{d_s(\mu)/2} \leqslant \limsup_{x \to \infty} n^\mu(x)/x^{d_s(\mu)/2} < \infty.$$

(2) *Let* S *be the unique constant which satisfies* $\sum_{i=1}^N \rho_i^{-S} = 1$. *Then,*

$$\max\{d_s(\mu) \colon \mu \text{ is a Bernoulli measure on } K\} = \frac{2S}{S+1} (\equiv d_s), \tag{2.4}$$

where the maximum is attained only at the Bernoulli measure μ *satisfying*

$$\mu_i = \rho_i^{-S} \quad \text{for } 1 \leqslant i \leqslant N.$$

In the following we define $\mu_i = \rho_i^{-S}$ for $1 \leqslant i \leqslant N$ and consider this measure unless otherwise stated. We note that, using Lemma 3.4, this measure is equivalent to the Hausdorff measure with respect to the effective resistance metric $R(\cdot,\cdot)$.

The harmonic structure for the Sierpinski gasket is regular and we have a regular local Dirichlet form corresponding to the Brownian motion on the

Sierpinski gasket. The Hausdorff dimension of this set is $\log 3/\log 2$, while the Hausdorff dimension in the resistance metric is $S = \log 3/\log(5/3)$. Thus the spectral dimension is $2\log 3/\log 5$.

3. *Preliminary results*

In this section we establish some of the results that we will use in order to prove our heat kernel estimates. To begin we set up a new sequence of graph approximations to the p.c.f. self-similar set. This sequence will have the property that the resistance of any edge in the nth graph approximation will be within constants of e^{-n}. We will call this approximation conductivity coordinates.

Let Λ_n be defined by

$$\Lambda_n = \left\{ w = w_1 \dots w_k \in \bigcup_{i \geqslant 0} S^i \colon \rho_1 \dots \rho_{k-1} \leqslant e^n < \rho_1 \dots \rho_k \right\}.$$

For typographical reasons we will sometimes write $\Lambda(n)$ instead of Λ_n. The approximation that we use for the fractal will be denoted by V_{Λ_n} and is the graph obtained from the vertex set

$$V_{\Lambda_n} = \pi\left(\bigcup_{w \in \Lambda_n} wP \right),$$

where we include an edge whenever $D_{pq} > 0$. We can think of this as taking a sequence of cross-sections of the tree associated with the shift space, determined by the conductivity of the corresponding cells.

The conductance of the Λ_n-cell at w when $w = w_1 \dots w_k \in V_{\Lambda_n}$ is given by $\rho_w = \prod_{i=1}^k \rho_{w_i}$. From the approximation we have, by definition, that there exist constants $c_{3.1}, c_{3.2}, c_{3.3}, c_{3.4} > 0$ such that for $w \in \Lambda_n$,

$$\begin{aligned} c_{3.1} e^n &\leqslant \rho_w \leqslant c_{3.2} e^n, \\ c_{3.3} e^{-Sn} &\leqslant \mu_w \leqslant c_{3.4} e^{-Sn}. \end{aligned} \tag{3.1}$$

We need some further properties of the Dirichlet form $\mathscr{E}$, which are essentially corollaries of earlier results. We have a Poincaré inequality and a decomposition of the Dirichlet form. For $u \in C(F)$ we write $\overline{u} = \int_F u \, d\mu$.

LEMMA 3.1. *There exists $c_{3.5} > 0$ such that for $f \in \mathscr{F}$ and $n \geqslant 0$,*

$$\mathscr{E}(f,f) \geqslant c_{3.5} \| f - \overline{f} \|_2^2, \tag{3.2}$$

$$\mathscr{E}(f,f) = \sum_{w \in \Lambda_n} \rho_w \mathscr{E}(f \circ F_w, f \circ F_w). \tag{3.3}$$

Proof. Let $g = f - \bar{f}$. Then from Theorem 2.10, for $x, y \in F$,

$$(g(x) - g(y))^2 = (f(x) - f(y))^2 \leqslant R(x,y)\mathscr{E}(f,f).$$

So, as the harmonic structure is regular, $R(x,y) \leqslant c_1$ and

$$\mathcal{E}(f,f) = \iint \mathcal{E}(f,f)\,\mu(dx)\,\mu(dy) \geqslant \frac{1}{c_1}\iint (g(x)-g(y))^2\,\mu(dx)\,\mu(dy)$$
$$= \frac{2}{c_1}\int g(x)^2\,\mu(dx).$$

Equation (3.3) follows from the repeated application of (2.3).

LEMMA 3.2. *There exist constants $c_{3.6}$, $c_{3.7}$ such that if $x,y \in V_{\Lambda_n}$ and are neighbours, then*

$$c_{3.6}\,e^{-n} \leqslant R(x,y) \leqslant c_{3.7}\,e^{-n}.$$

Proof. This follows from the construction of the Dirichlet form. By definition we have for all f with $f(x)=0$, $f(y)=1$ with $x,y \in \partial K_w$, $w \in \Lambda_n$,

$$c_{3.1}\,e^n \leqslant \rho_w \leqslant \mathcal{E}_n(f,f) \leqslant \mathcal{E}(f,f).$$

Thus there is a constant such that

$$(\inf\{\mathcal{E}(f,f)\colon f(x)=0, f(y)=1\})^{-1} \leqslant c_1 e^{-n}.$$

For the lower bound we take the Λ_n-harmonic function f_1 which is 1 at y and 0 at all other points of V_{Λ_n}. Then

$$\inf\{\mathcal{E}(f,f)\colon f(x)=0, f(y)=1\} \leqslant \mathcal{E}(f_1,f_1) = \mathcal{E}_{\Lambda_n}(f_1,f_1) \leqslant c_2\,e^n,$$

where $\mathcal{E}_{\Lambda_n}(f,f) = \sum_{w\in\Lambda_n} \rho_w (f(x)-f(y))^2$, and we have the lower bound.

We will also need to define a means of determining the behaviour of the shortest paths on the fractal in the effective resistance metric. For $x,y \in K$ define the set of paths between them on V_{Λ_m} to be

$$\Pi_m(x,y) = \{\pi = (p_i, p_{i+1})_{i=1}^{|\pi|}\colon p_1 = x,\ p_{|\pi|+1} = y,\ p_k \in V_{\Lambda_m}\ (2 \leqslant k \leqslant |\pi|),$$
$$p_k, p_{k+1} \text{ are in the same } \Lambda_m\text{-complex}\ (1 \leqslant k \leqslant |\pi|)\},$$

where $|\pi|$ is the cardinality of the steps in the path π. The number of steps in the shortest path is defined to be

$$N_m(x,y) = \inf\{|\pi|\colon \pi \in \Pi_m(x,y)\}.$$

This shortest path counting function will determine the off-diagonal behaviour of the heat kernel. As we are using the sequence of approximations determined by the resistance there may not be geodesics in the effective resistance metric. Indeed for the ordinary Sierpinski gasket we see that there is no point $x \in \mathrm{Int}(K)$ for which $R(0,x)+R(x,1)=R(0,1)$. This means that we must determine the behaviour of the short paths in the effective resistance metric.

We define a chemical exponent in the effective resistance metric for the fractal by setting $d_k^c(x,y) = k^{-1}\log N_{m+k}(x,y)$, when $e^{-m-1} \leqslant R(x,y) \leqslant e^{-m}$. Note that we have the following upper bound on the length of the shortest path.

LEMMA 3.3. *There exist constants $c_{3.8}$ such that for all $x,y \in K$ and $m > 0$,*

$$N_m(x,y) \leqslant c_{3.8}\,e^{(S+1)m/2}.$$

Proof. First we prove the theorem when $x, y \in V_0$. For $z_0 \in V_0$, let

$$f(y) = \min\left(1, \frac{N_m(z_0, y)}{\max_{z \in V_0} N_m(z_0, z)}\right).$$

Then observe that

$$1 \leqslant \mathscr{E}_0(f, f) \leqslant \mathscr{E}_{\Lambda_m}(f, f) \leqslant c_1 \sum_{x, y \in V_{\Lambda(m)}} (f(x) - f(y))^2 e^m.$$

The increments of f are bounded above by $(\max_{z \in V_0}\{N_m(z_0, z)\})^{-1}$ and, as there are of the order of e^{Sm} vertices in V_{Λ_m}, we have

$$1 \leqslant c_2 \frac{e^m e^{Sm}}{\max_{z \in V_0}\{N_m(z_0, z)\}^2}.$$

Rearranging gives the result for V_0. For $y \in K$, take a sequence $\{y_i\}_{i=0}^M$ such that $y_i \in V_{\Lambda_i}$, y_i and y_{i+1} are in the same Λ_i-complex $(0 \leqslant i \leqslant M-1)$, and $y_{M+1} = y$. Using the above result and self-similarity, we find that $N_m(y_i, y_{i+1}) \leqslant c_3 e^{(S+1)(m-i)/2}$ (for all i with $0 \leqslant i \leqslant M-1$). Thus,

$$N_m(z_0, y) \leqslant \sum_{i=0}^{M-1} N_m(y_i, y_{i+1}) \leqslant c_4 e^{(S+1)m/2}.$$

Using the triangle inequality, we see that the result follows.

Thus the chemical exponent $d_k^c(x, y)$ satisfies

$$d_k^c(x, y) \leqslant c_{3.9} \frac{(S+1)(\log R(x, y) + k)}{2k}, \quad \text{for all } x, y \in K.$$

Now, we will obtain estimates of the mean hitting times for the diffusion. In order to do this, we introduce some more notation and prove some lemmas. For $x \in K$, let $B_r(x) = \{y \in K\colon R(x, y) < r\}$. Also, for $x \in K$ and $l \geqslant 0$, let

$D^0_{\Lambda_l}(x) = \{C\colon C$ is a Λ_l-complex which contains $x\}$,

$D^1_{\Lambda_l}(x) = D^0_{\Lambda_l}(x) \cup \{C\colon C$ is a Λ_l-complex which is connected to $D^0_{\Lambda_l}(x)\}$,

and $\partial D^i_{\Lambda_l}(x) = \mathrm{Cl}(K \setminus D^i_{\Lambda_l}(x)) \cap D^i_{\Lambda_l}(x)$ for $i = 0, 1$.

LEMMA 3.4. *There exists* $k_0 \in \mathbb{N}$ *such that*

$$D^1_{\Lambda(l+k_0)}(x) \subset B_r(x) \subset D^1_{\Lambda(l-k_0)}(x)$$

for all $x \in K$, $r \geqslant 0$ *such that* $e^{-(l+1)} \leqslant r \leqslant e^{-l}$.

Proof. Using Lemma 3.2, we can show that the first inclusion holds, by taking k_0' to be such that

$$\max_{w \in \Lambda(l+k_0')} \max_{x, y \in K_w} R(x, y) < \tfrac{1}{2}\rho_w^{-1}.$$

For the second inequality, it is enough to prove that there exists $c_1 > 0$ such that for all $k \geqslant 0$ and for all $y \notin D^1_{\Lambda_k}(x)$, $c_1 e^{-k} \leqslant R(x, y)$. This is because, by choosing k_0'' so that $c_1 e^{-(l-k_0'')} \geqslant r$ and taking $k = l - k_0''$, we obtain the desired inclusion.

Now, take $f \in l(K)$ so that $f(x) = 1$ for $x \in \partial D^0_{\Lambda_k}$, $f(x) = 0$ for $x \in \partial D^1_{\Lambda_k}$, and f is harmonic inside V_{Λ_k}. Then,

$$\mathscr{E}(f, f) = \sum_{\substack{w' \in \Lambda_k, \\ K_{w'} \subset D^1_{\Lambda(k)} \setminus D^0_{\Lambda(k)}}} \rho_{w'} \mathscr{E}_0(f \circ \Psi_{w'}, f \circ \Psi_{w'}) \leqslant c_2 e^k,$$

where c_2 depends only on the shape of the fractal we treat. Thus, for $y \notin D^1_{\Lambda_k}(x)$, we have $R(x, y) \geqslant \mathscr{E}(f, f)^{-1} \geqslant c_2^{-1} e^{-k}$, and by taking $k_0 = \max\{k'_0, k''_0\}$ we complete the proof.

For the diffusion $\{X_t\}_{t \geqslant 0}$ associated with the Dirichlet form $(\mathscr{E}, \mathscr{F})$, define the hitting time for a set $A \subset K$ by $T_A = \inf\{t \geqslant 0\colon X_t \in A\}$.

LEMMA 3.5. (1) *There exist* $c_{3.10}$, $c_{3.11} > 0$ *such that for all* $r \geqslant 0$ *and for all* $z \in V_{\Lambda_r}$,

$$c_{3.10} e^{-r(1+S)} \leqslant E^z(T_{\partial D^0_{\Lambda(r)}(z)}) \leqslant c_{3.11} e^{-r(1+S)}.$$

(2) *There exist* $c_{3.12} > 0$ *such that for all* $0 \leqslant r \leqslant m$ *and for all* $z \in V_{\Lambda_m}$,

$$E^z(T_{\partial D^0_{\Lambda(r)}(z)}) \leqslant c_{3.12} e^{-r(1+S)}.$$

Proof. First, we take $n \gg m \geqslant r$ and consider the continuous-time Markov chain X_n on V_{Λ_n} associated with $\mathscr{E}_{\Lambda_n}$. By construction of the Markov chain, the mean holding time for X_n at each vertex is a constant multiplied by $e^{-(S+1)n}$. Denote by $E^z_{X_n}$ the mean with respect to the probability measure for X_n starting at z. Then, by the same argument as in the proof of Theorem 2.5 in [**1**], $E^z_{X_n}(T_{\partial D^0_{\Lambda(r)}(z)})$ is estimated from above and below by constants multiplied by

$$\frac{\gamma(z, n)}{\gamma(z, r)} E^z_{X_n}[\inf\{t > 0\colon X_n(t) \in V_{\Lambda_r}, \exists s < t, X_n(s) \neq X_n(0)\}],$$

where $\gamma(z, i) = \{$sum of conductances on Λ_i-bonds containing $z\}$. Now, the first term can be estimated from above and below by $e^{(n-r)}$. Using the ergodic theorem, one can estimate the second term by $e^{-(S+1)n}\left(\int_{V_{\Lambda(n)}} 1_{V_{\Lambda(r)}}(x)\, d\mu_n(x)\right)^{-1}$, where μ_n is an invariant measure for X_n. This value is estimated from above and below by $e^{-(S+1)n} e^{S(n-r)}$ so that we have

$$c_1 e^{-r(1+S)} \leqslant E^z_{X_n}(T_{\partial D^0_{\Lambda(r)}(z)}) \leqslant c_2 e^{-r(1+S)} \tag{3.4}$$

for all $n \gg r$.

We next calculate $E^z_{X_n} T_{\partial D^0_{\Lambda(r)}(z)}$ by conditioning on the exit from a series of domains about the point z. First, condition on the exit place on $\partial D^0_{\Lambda_{r+1}}(z)$, to get

$$E^z_{X_n} T_{\partial D^0_{\Lambda(r)}(z)} = E^z_{X_n} T_{\partial D^0_{\Lambda(r+1)}(z)} + \sum_{x \in \partial D^0_{\Lambda(r+1)}(z)} P^z_{X_n}\left(X_{T_{\partial D^0_{\Lambda(r+1)}(z)}} = x\right) E^x_{X_n}(T_{\partial D^0_{\Lambda(r)}(z)}).$$

Thus

$$E^z_{X_n} T_{\partial D^0_{\Lambda(r)}(z)} \leqslant E^z_{X_n} T_{\partial D^0_{\Lambda(r+1)}(z)} + \max_{x \in \partial D^0_{\Lambda(r+1)}(z)} E^x_{X_n}(T_{\partial D^0_{\Lambda(r)}(z)}).$$

Repeating this procedure gives

$$E^z_{X_n} T_{\partial D^0_{\Lambda(r)}(z)} \leqslant \sum_{i=r}^{\infty} \max_{x_i \in \partial D^0_{\Lambda(i+1)}(z)} E^{x_i}_{X_n}(T_{\partial D^0_{\Lambda(i)}(z)}).$$

From (3.4) and an easy Markov chain calculation we have an upper bound

$$E^{x_i}_{X_n}(T_{\partial D^0_{\Lambda(i)}(z)}) \leqslant c_3 e^{-i(1+S)}.$$

Summing these terms, we have

$$E^{z}_{X_n}(T_{\partial D^0_{\Lambda(r)}(z)}) \leqslant c_4 e^{-r(1+S)} \tag{3.5}$$

for all $z \in V_{\Lambda_m}$, $n \gg m \geqslant r$.

Using Theorem 3.7 below, we have the results from (3.4) and (3.5).

We need estimates for the tail of the crossing time distribution. Define for $i \geqslant 1$,

$$T^{(0)}_{\Lambda_r} = \inf\{t \geqslant 0\colon X_t \in V_{\Lambda_r}\},$$

$$T^{(i)}_{\Lambda_r} = \inf\{t \geqslant T^{(i-1)}_{\Lambda_r}\colon X_t \in V_{\Lambda_r} \setminus \{X_{T^{(i-1)}_{\Lambda(r)}}\}\},$$

$$W^{(i)}_{\Lambda_r} = T^{(i)}_{\Lambda_r} - T^{(i-1)}_{\Lambda_r}.$$

LEMMA 3.6. (1) *There exist* $c_{3.13} > 0$, $0 < c_{3.14} < 1$ *such that for all* $0 \leqslant r \leqslant m$,

$$P^z(W^{(i)}_{\Lambda_r} \leqslant t) \leqslant c_{3.13} e^{(1+S)r} t + c_{3.14} \quad \textit{for all } 0 < t < 1,\ z \in V_{\Lambda_r}.$$

(2) *There exist* $c_{3.15}$, $c_{3.16} > 0$ *such that for all* $0 \leqslant r \leqslant m$,

$$P^z(W^{(i)}_{\Lambda_r} \leqslant t) \leqslant c_{3.15} \exp\{-c_{3.16} \overline{N}_{r+k}(z)\}, \quad \textit{for all } 0 < t < 1,\ z \in K, \tag{3.6}$$

where

$$\overline{N}_{r,j}(z) = \inf_{\substack{x_1, x_2 \in V_{\Lambda(r)} \cap D^1_{\Lambda(r)}(z) \\ x_1 \neq x_2}} N_{r+j}(x_1, x_2), \quad k = \inf\left\{ j\colon \frac{\overline{N}_{r+j}(z)}{e^{(r+j)(S+1)}} \leqslant t \right\}.$$

REMARK. Note that from Lemma 3.3, $k < \infty$ for each $r \geqslant 0$, $0 < t \leqslant 1$, $z \in K$.

Proof. As in the last lemma, we first take $n \gg m \geqslant r$ and consider the continuous-time Markov chain X_n on V_{Λ_n} associated with $\mathcal{E}_{\Lambda_n}$. The random variables $W^{(i)}_{\Lambda_r}$ $(i \in \mathbb{N})$ are i.i.d. so that it is enough to show the result for $i = 1$. Then, following [**3**], we have

$$T_{\partial D^0_{\Lambda(r)}(z)} \leqslant t + 1_{(T_{\partial D^0_{\Lambda(r)}(z)} > t)} \cdot (T_{\partial D^0_{\Lambda(r)}(z)} - t).$$

Thus, by the Markov property,

$$E^z_{X_n}(T_{\partial D^0_{\Lambda(r)}(z)}) \leqslant t + E^z_{X_n}(1_{(T_{\partial D^0_{\Lambda(r)}(z)} > t)} E^{X_n(t)}_{X_n}(T_{\partial D^0_{\Lambda(r)}(z)})). \tag{3.7}$$

On the other hand, by Lemma 3.5(1), $c_{3.10} e^{-(1+S)r} \leqslant E^z_{X_n}(T_{\partial D^0_{\Lambda(r)}(z)})$. Also, using Lemma 3.5(2), one can easily see that $E^x_{X_n}(T_{\partial D^0_{\Lambda(r)}(z)}) \leqslant c_1 e^{-(1+S)r}$ for $x \in D^0_{\Lambda_r}(z)$. Substituting these results into (3.7) and calculating, we have

$$P^z_{X_n}(W^{(1)}_{\Lambda_r} \leqslant t) \leqslant \frac{1}{c_1} e^{(1+S)r} t + 1 - \frac{c_{3.10}}{c_1}$$

(as $z \in V_{\Lambda_r}$, $T_{\partial D^0_{\Lambda(r)}(z)} = W^{(1)}_{\Lambda_r}$) so that the $P^z_{X_n}$ version of (1) is proved.

Next, we prove that for all $z \in V_{\Lambda_m}$ $(m \geqslant r)$,

$$P^z_{X_n}(W^{(i)}_{\Lambda_r} \leqslant t) \leqslant c_{3.15} \exp(-c_{3.16}\overline{N}_{r,k}(z)), \quad \text{for } t_n < t < 1, \tag{3.8}$$

where $t_n \to 0$ as $n \to \infty$. For $\hat{c} > 0$, let

$$k_{\hat{c}} = \inf\left\{ j\colon \frac{\overline{N}_{r,j}(z)}{e^{(r+j)(S+1)}} \leqslant \hat{c}t \right\}, \tag{3.9}$$

and $k(\hat{c}) = k_{\hat{c}}$. Then, provided $k_{\hat{c}} < n$, that is, $t > t_n = \overline{N}_n(z)e^{-n(S+1)}\hat{c}^{-1}$, we have

$$\begin{aligned} P^z_{X_n}(W^{(i)}_{\Lambda_r} \leqslant t) &\leqslant P^z_{X_n}\left(\sum_{i=1}^{\overline{N}_{r,k(\hat{c})}(z)} W^{(i)}_{\Lambda_r+k_{\hat{c}}} \leqslant t \right) \\ &\leqslant \exp(c_2(\overline{N}_{r,k_{\hat{c}}}(z)e^{(1+S)(r+k_{\hat{c}})}t)^{1/2} - c_3\overline{N}_{r,k_{\hat{c}}}(z)) \\ &= \exp(-c_4\overline{N}_{r,k_{\hat{c}}}(z)), \end{aligned}$$

where we use Lemma 1.1 of [**3**] and Lemma 3.6 in the second inequality, and (3.9) in the last equality (we choose $\hat{c}$ so that $c_2 < \hat{c}c_3$). On the other hand, note that there exists $c_5 > 0$ such that

$$N_{l+l'}(x,y) \leqslant c_5 e^{Sl'} N_l(x,y) \quad \text{for all } x,y \in K,\ l,l' \geqslant 0. \tag{3.10}$$

This is because, when we add vertices and construct $V_{\Lambda_{l+l'}}$ from V_{Λ_l}, then there are at most $c_5 e^{Sl'}$ $\Lambda_{l+l'}$-cells inside each Λ_l-cell. Using this fact and by the definition of k and $k_{\hat{c}}$, we easily see that there exists $c_6 > 0$ such that $N_{r+k_{\hat{c}}} \geqslant c_6 N_{r+k}$ so that (3.8) is proved. Using the following Theorem 3.7, we can prove the lemma.

We now prove that $X_n \to X$ in law.

THEOREM 3.7. *We have*

$$E^{x_n}_{X_n}[f(w(\cdot))] \to E^x_X[f(w(\cdot))] \quad \text{as } n \to \infty$$

for any $f \in C_b(D([0,\infty) : K) \to \mathbb{R})$ and any sequence $x_n \in V_{\Lambda_n}$ with $x_n \to x \in K$ as $n \to \infty$. Here the expectations are taken over $w \in D([0,\infty) : K)$.

Proof. The idea of the proof is the same as that of Theorem 6.1 in [**11**], so we just sketch the proof. We need tightness and convergence of the finite-dimensional distributions. The latter is deduced from the general theory as $\mathscr{E}_{\Lambda_n}$ is a monotone increasing sequence of quadratic forms which converges to $\mathscr{E}$. To prove the former, we first show that there exist $c_1, c_2 > 0$ such that for all $x \in K$, $m, j \in \mathbb{N}$,

$$c_1 e^j \leqslant \overline{N}_{m,j}(x) \leqslant c_2 e^{(S+1)(m+j)/2}. \tag{3.11}$$

The upper bound is proved in Lemma 3.3. Let x_m and y_m be elements of $D^1_{\Lambda_m}(x) \cap V_{\Lambda_m}$ such that $\overline{N}_{m,j}(x) = N_{m+j}(x_m, y_m)$. The lower bound uses the observation that

$$R(x_m, y_m) \leqslant \sum_{i=1}^{\overline{N}_{m,j}(x)} R(z_i, z_{i+1}) \leqslant \overline{N}_{m,j}(x) c_{3.7} e^{-(m+j)},$$

where $\{z_i\}$ is a minimum Λ_{m+j}-walk from x_m to y_m. As $R(x_m, y_m) \geqslant c_{3.6}\, e^{-m}$, we have (3.11). Using this and (3.8), with the definition of k, we have for each $t > t_n$,

$$P^x_{X_n}(W^{(i)}_{\Lambda_m} \leqslant t) \leqslant c_{3.15} \exp(-c_{3.16}\overline{N}_{m+k}(x)) \leqslant c_{3.15}\, e^{-c_3 e^k}.$$

As $1 \leqslant \overline{N}_{m,k}(x) \leqslant te^{(m+k)(S+1)}$, we have $e^k \geqslant e^{-m} t^{-1/(S+1)}$. Thus,

$$\limsup_{n \to \infty} \sup_{x_n \in V_{\Lambda(n)}} P^{x_n}_{X_n}(W^{(i)}_{\Lambda_m} \leqslant t) \leqslant c_{3.15}\, e^{-c_3 e^{-m} t^{-1/(S+1)}}.$$

The right-hand side converges to 0 as $t \to 0$ and the tightness of $P^{x_n}_{X_n}$ is proved.

This result also completes the proof of Lemma 3.5 and Lemma 3.6. Note that this argument shows that the result of Lemma 3.5(2) holds for all $z \in K$.

THEOREM 3.8. *There exist $c_{3.17}$, $c_{3.18} > 0$ such that for all $z \in K$,*

$$c_{3.17}\, r^{1+S} \leqslant E^z(T_{\partial B_r(z)}) \leqslant c_{3.18} r^{1+S}.$$

Proof. First, by Lemma 3.4,

$$E^z(T_{\partial D^1_{\Lambda(l+k(0))}(z)}) \leqslant E^z(T_{\partial B_r(z)}) \leqslant E^z(T_{\partial D^1_{\Lambda(l-k(0))}(z)})$$

if $e^{-(l+1)} \leqslant r < e^{-l}$. By Lemma 3.5(1), the left-hand side can be estimated from below by $c_1 e^{-l(1+S)} \geqslant c_1 r^{1+S}$. By Lemma 3.5(1), (2) and simple Markov chain arguments, the right-hand side can be estimated from above by $c_2 e^{-l(1+S)} \leqslant c_3 r^{1+S}$.

We remark that S is the Hausdorff dimension of K with respect to the resistance metric and it coincides with the box dimension (due to Theorem 3.2 of [**16**] and Lemma 7.3 of [**17**], as we have Lemma 3.2). We call

$$d_w = \lim_{r \to 0} \frac{\log E^x(T_{\partial B_r(x)})}{\log r}$$

the *random walk dimension*, if the limit exists. Thus, Theorem 3.8 states that $d_w = 1 + S$. This relationship is called the Einstein relation in the physics literature. It states that the walk dimension should be given by the sum of the resistance dimension and the 'fractal' dimension, which reduces to $1 + d_f$ where d_f denotes the Hausdorff dimension of K in the resistance metric. The Einstein relation occurs in a variety of physical models (see, for example [**13**]). We also remark that, with the formula (2.4), we have the relation $d_s/2 = d_f/d_w$ which is a generalization of the formula given in [**5**, **8**, **19**].

4. *Transition density estimates: upper bounds*

Let P_t be the semigroup of positive operators associated with the Dirichlet form $(\mathscr{E}, \mathscr{F})$ on $L^2(K, \mu)$. As $(\mathscr{E}, \mathscr{F})$ is local and regular, there exists a Feller diffusion $(\{X_t\}_{t \geqslant 0}, \{P^x\}_{x \in K})$ with semigroup P_t, on K. As in [**8**, Lemma 2.9] the existence of a reproducing kernel ensures that the transition function has a density $p_t(x, y)$ with respect to μ which satisfies the Chapman–Kolmogorov equations.

We will obtain upper bounds on $p_t(x, y)$, beginning with the on-diagonal upper bound, where we follow closely the argument of [**21**, **6**].

LEMMA 4.1. (1) *There is a constant* $c_{4.1}$ *such that*

$$\|P_t\|_{1\to\infty} = \sup_{x,y\in K} p_t(x,y) \leqslant c_{4.1} t^{-d_s/2}, \quad \textit{for } 0<t<1. \tag{4.1}$$

(2) *The transition density* $p_t(x,y)$ *is jointly continuous in* $(t,x,y) \in (0,1]\times K\times K$.

Proof. (1) For $w\in\Lambda_n$ write $f_w = f\circ F_w$ and

$$\overline{f}_w = \int_K f_w(x)\,\mu(dx).$$

Note that for $v\in C(K)$, $\overline{v} = \int v\,d\mu = \sum_{w\in\Lambda_n} \overline{v}_w\,\mu_w$.

Let $u_0 \in \mathscr{D}(\Delta)$ with $u_0 \geqslant 0$ and $\|u_0\|_1 = 1$. Set $u_t(x) = (P_t u_0)(x)$ and $g(t) = \|u_t\|_2^2$. We remark that g is continuous and decreasing. As the semigroup is conservative, $\|u_t\|_1 = 1$, and using Lemma 3.1 and (3.1) we have

$$\begin{aligned}
\frac{d}{dt}\,g(t) &= -2\mathscr{E}(u_t,u_t) \\
&= -2\sum_{w\in\Lambda_n} \rho_w\,\mathscr{E}(u_t\circ F_w, u_t\circ F_w) \quad \text{(by (3.3))} \\
&\leqslant -2c_1 e^n \sum_w \int (u_{t,w} - \overline{u}_{t,w})^2\,d\mu \\
&\leqslant -2c_2 e^n e^{Sn} \int u_t^2\,d\mu + 2c_1 e^n \sum_w (\overline{u}_{t,w})^2 \\
&\leqslant -2c_2 e^{(S+1)n} \|u_t\|_2^2 + 2c_3 e^{(2S+1)n}.
\end{aligned} \tag{4.2}$$

Thus $g'(t) \leqslant -c_2 e^{(S+1)n}(g(t) - c_4 e^{Sn})$, for all $n\geqslant 0$. Therefore

$$-\frac{d}{dt}\log(g(t) - c_4 e^{Sn}) \geqslant c_2 e^{(S+1)n}, \quad \text{if } g(t) > c_4 e^{Sn}. \tag{4.3}$$

Let $s_n = \inf\{t\geqslant 0\colon g(t)\leqslant c_4 e^{Sn}\}$ for $n\in\mathbb{N}$. Thus (4.3) holds for $0<t<s_n$. Integrating (4.3) from s_{n+2} to s_{n+1} we obtain

$$\begin{aligned}
c_2 e^{(S+1)n}(s_{n+1} - s_{n+2}) &\leqslant -\log(g(s_{n+1}) - c_4 e^{Sn}) + \log(g(s_{n+2}) - c_4 e^{Sn}) \\
&= \log(e^{S(n+2)} - c_4 e^{Sn})/(e^{S(n+1)} - c_4 e^{Sn}) \leqslant c_5.
\end{aligned}$$

Thus $s_{n+1} - s_{n+2} \leqslant c_6 e^{-(S+1)n}$, and iterating this we have

$$s_n \leqslant c_6 \sum_{k=n-1}^{\infty} e^{-(S+1)k} \leqslant c_7 e^{-(S+1)n}.$$

This implies that $g(c_7/e^{(S+1)n}) \leqslant g(s_n) = c_4 e^{Sn}$. It follows that there exists $c_6<\infty$ such that if $e^{-(S+1)n} \leqslant t < e^{-(S+1)(n-1)}$ then

$$g(t) \leqslant c_8 e^{Sn} = c_9 t^{-S/(S+1)}.$$

Using the fact that $\|P_t\|_{1\to\infty} = \|P_t\|_{1\to 2}^2$, we deduce that $p_t(x,y) \leqslant c_{4.1} t^{-S/(S+1)}$ for all $x,y\in K$.

For (2), the joint continuity will follow from the upper bound on the heat kernel, as in [**8**, Lemma 4.6].

We next prove an off-diagonal upper bound for $p_t(x,y)$.

THEOREM 4.2. *There exist constants* $c_{4.2}, c_{4.3} > 0$ *such that for* x, y *in* K, $t \in (0,1)$, *with*

$$e^{-m-1} \leqslant R(x,y) \leqslant e^{-m}$$

and

$$k = \inf\left\{ j\colon \frac{N_{m+j}(x,y)}{e^{(m+j)(S+1)}} \leqslant t \right\},$$

then

$$p_t(x,y) \leqslant c_{4.2} t^{-d_s/2} \exp\left(-c_{4.3} \left(\frac{R(x,y)^{S+1}}{t} \right)^{d_k^c(x,y)/(S+1-d_k^c(x,y))} \right). \tag{4.4}$$

Proof. Fix $x \neq y$ and t as above and let $\varepsilon > 0$ be sufficiently small. Let $\nu_x = \mu|_{B_\varepsilon(x)}$, $\nu_y = \mu|_{B_\varepsilon(y)}$,

$$A_i(x) = \{ z \in K\colon N_{m+k}(x,z) \geqslant i^{-1} N_{m+k}(x,y) \},$$

and $C_i(x) = K \setminus A_i(x)$. Then

$$\begin{aligned} P^{\nu_x}(X_t \in B_\varepsilon(y)) &= P^{\nu_x}(X_t \in B_\varepsilon(y), X_{t/2} \in A_2(x)) + P^{\nu_x}(X_t \in B_\varepsilon(y), X_{t/2} \in C_2(x)) \\ &\equiv I_1 + I_2. \end{aligned}$$

Choose ε small enough such that $B_\varepsilon(x) \subset C_6(x)$. By the same argument as in the proof of Lemma 3.6(2),

$$\begin{aligned} P^{\nu_x}(X_{t/2} \in A_2(x)) &\leqslant \max_{z \in B_\varepsilon(x)} P^z(X_{t/2} \in A_2(x)) \mu(B_\varepsilon(x)) \\ &\leqslant \max_{z \in B_\varepsilon(x)} P^z\big(T_{\partial A_2(x)} \leqslant \tfrac{1}{2} t\big) \mu(B_\varepsilon(x)) \\ &\leqslant \max_{z \in B_\varepsilon(x)} P^z\left(\sum_{i=1}^{\frac{1}{3} N_{m+k(\hat{c})}(x,y)} W^{(i)}_{\Lambda(m+k_{\hat{c}})} \leqslant \tfrac{1}{2} t \right) \mu(B_\varepsilon(x)) \\ &\leqslant \exp\{ c_1 (N_{m+k_{\hat{c}}}(x,y) e^{(1+S)(m+k_{\hat{c}})} t)^{1/2} - c_2 N_{m+k_{\hat{c}}}(x,y) \} \mu(B_\varepsilon(x)) \\ &= \exp(-c_3 N_{m+k_{\hat{c}}}(x,y)) \mu(B_\varepsilon(x)) \\ &\leqslant \exp(-c_4 N_{m+k}(x,y)) \mu(B_\varepsilon(x)), \end{aligned}$$

where $k_{\hat{c}}$ is defined in the same way as in (3.9).

On the other hand, if $q(z) \equiv P(X_t \in B_\varepsilon(y) |\ X_{t/2} = z)$, then by Lemma 4.1,

$$q(z) = \int_{B_\varepsilon(y)} p_{t/2}(z,w)\, \mu(dw) \leqslant c_5 t^{-d_s/2} \mu(B_\varepsilon(y)).$$

Thus

$$\begin{aligned} I_1 &= E^{\nu_x}(q(X_{t/2})\colon X_{t/2} \in A_2(x)) \\ &\leqslant c_6\, \mu(B_\varepsilon(x))\, \mu(B_\varepsilon(y))\, t^{-d_s/2} \exp(-c_7 N_{m+k}(x,y)). \end{aligned}$$

For I_2, by the symmetry of $p_t(x,y)$,

$$P^{\nu_x}(X_t \in B_\varepsilon(y), X_{t/2} \in C_2(x)) = P^{\nu_y}(X_t \in B_\varepsilon(x), X_{t/2} \in C_2(x))$$

which is bounded in exactly the same way as I_1.

Adding the bounds for I_1 and I_2, we have

$$P^{\nu_x}(X_t \in B_\varepsilon(y)) \leqslant 2c_6\,\mu(B_\varepsilon(x))\,\mu(B_\varepsilon(y))\,t^{-d_s/2}\exp(-c_7 N_{m+k}(x,y)). \tag{4.5}$$

We remark that, from the definitions, it is easy to check that there are constants c_8 and c_9 such that

$$c_8 N_{m+k}(x,y) \leqslant \left(\frac{R(x,y)^{S+1}}{t}\right)^{d_k^c(x,y)/(S+1-d_k^c(x,y))} \leqslant c_9 N_{m+k}(x,y). \tag{4.6}$$

Dividing both sides of (4.5) by $\mu(B_\varepsilon(x))$, $\mu(B_\varepsilon(y))$ and using the continuity of $p_t(x,y)$ in (x,y) proves the theorem.

5. *Lower bounds*

In this section we use techniques developed in [**5**, **8**] to obtain lower bounds on $p_t(x,y)$ which will be identical, apart from the constants, to the upper bound.

LEMMA 5.1. *There exists a constant $c_{5.1} > 0$ such that*

$$p_t(x,x) \geqslant c_{5.1} t^{-d_s/2} \quad \textit{for all } x \in F,\ 0 < t < 1. \tag{5.1}$$

Proof. Note that from (3.6), we have

$$P^x(X_t \notin D^1_{\Lambda_r}(x)) \leqslant P^x(W^1_{\Lambda_r} \leqslant t) \leqslant c_1 \exp\{-c_2 \overline{N}_{r,k-1}(x)\}, \tag{5.2}$$

for all $x \in K$, $r \geqslant 0$, $0 < t < 1$. Let $a > 0$ satisfy $c_1 \exp(-c_2 a) \leqslant \frac{1}{2}$. And take $1 > t \geqslant ae^{-r(S+1)}$ (we choose r large enough so that $ae^{-r(S+1)} < 1$). Then, $\overline{N}_{r,k-1}(x) \geqslant te^{(r+k-1)(S+1)} \geqslant ae^{(k-1)(S+1)} \geqslant a$ so that the right-hand side of (5.2) is less than $\frac{1}{2}$. Thus, $P^x(X_t \in D^1_{\Lambda_r}(x)) \geqslant \frac{1}{2}$. On the other hand, by (3.1), $\mu(D^1_{\Lambda_r}(x)) \leqslant c_2 e^{-rS} \leqslant c_3 t^{S/(S+1)}$. Now, using the Cauchy–Schwarz inequality, we have

$$\begin{aligned}\tfrac{1}{4} \leqslant P^x(X_t \in D^1_{\Lambda_r}(x))^2 &= \left(\int_{D^1_{\Lambda(r)}(x)} p_t(x,y)\,\mu(dy)\right)^2 \\ &\leqslant \mu(D^1_{\Lambda_r}(x)) \int_{D^1_{\Lambda(r)}(x)} p_t(x,y)^2 \mu(dy) \\ &\leqslant c_3 t^{S/(S+1)} p_{2t}(x,x).\end{aligned}$$

Hence we deduce that $p_t(x,x) \geqslant c_4 t^{-S/(S+1)}$.

We need to extend this 'on-diagonal lower bound' to a 'near-diagonal lower bound', which we do via an estimate on the Hölder continuity of the heat kernel.

LEMMA 5.2. *There exists a constant $c_{5.2} > 0$ such that*

$$|p_t(x,y) - p_t(x',y)| \leqslant c_{5.2} R(x,x')^{1/2} t^{-(2S+1)/2(S+1)}, \quad \textit{for all } x,x',y \in K,\ 0 < t < 1. \tag{5.3}$$

In particular, $p_t(\cdot,\cdot)$ is uniformly continuous on $K \times K$ for each $t > 0$.

Proof. By (2.2), we have

$$|p_t(x,y) - p_t(x',y)|^2 \leqslant R(x,x')\mathscr{E}(p_t(\cdot,y), p_t(\cdot,y)). \tag{5.4}$$

As in [**8**, Lemma 6.4], we have, writing $u(x) = p_{t/2}(x,y)$,

$$\begin{aligned}\mathscr{E}(P_{t/2}u, P_{t/2}u) &\leqslant c_1\left(\tfrac{1}{2}t\right)^{-1}\|u\|_2^2 \\ &\leqslant 2c_1 t^{-1}p_t(y,y) \leqslant c_2 t^{-1-S/(S+1)},\end{aligned}$$

so that (5.3) holds.

LEMMA 5.3. *There exist $c_{5.3}, c_{5.4} > 0$ such that for all $x, y \in K$, $0 < t < 1$,*

$$p_t(x,y) \geqslant c_{5.3}t^{-d_s/2} \quad \textit{whenever } R(x,y) \leqslant c_{5.4}t^{1/(S+1)}. \tag{5.5}$$

Proof. If $R(x,y) \leqslant c_{5.4}t^{1/(S+1)}$ then by Lemmas 5.1 and 5.2,

$$\begin{aligned}p_t(x,y) &\geqslant p_t(x,x) - |p_t(x,y) - p_t(x,x)| \\ &\geqslant t^{-d_s/2}(c_{5.1} - c_{5.2}R(x,y)^{1/2}t^{-1/2(S+1)}) \\ &\geqslant \tfrac{1}{2}c_1 t^{-d_s/2},\end{aligned}$$

where $c_{5.4}$ is chosen such that $c_{5.1} - c_{5.4}^{1/2}c_{5.2} > c_1$.

We can now use a standard chaining argument to obtain general lower bounds on p_t from Lemma 5.3.

THEOREM 5.4. *There exist constants $c_{5.5}, c_{5.6} > 0$ such that for x, y in K, $t \in (0,1)$, with*

$$e^{-m-1} \leqslant R(x,y) \leqslant e^{-m}$$

and

$$k = \inf\left\{j \colon \frac{N_{m+j}(x,y)}{e^{(m+j)(S+1)}} \leqslant t\right\},$$

then

$$p_t(x,y) \geqslant c_{5.5}t^{-d_s/2}\exp\left(-c_{5.6}\left(\frac{R(x,y)^{S+1}}{t}\right)^{d_k^c(x,y)/(S+1-d_k^c(x,y))}\right). \tag{5.6}$$

Proof. Fix x, y and t. Using (5.5) we see that the bound is satisfied if $R(x,y)^{S+1}/t < c_{5.4}^{S+1}$. Thus we assume that $D = R(x,y)^{S+1}/t > c_{5.4}^{S+1}$. As in (3.9), for $\hat{c} > 0$, let

$$k' = k_{\hat{c}} = \inf\left\{j \colon \frac{N_{m+j}(x,y)}{e^{(m+j)(S+1)}} \leqslant \hat{c}t\right\}. \tag{5.7}$$

By our choice of k' there is a $c_1 > 0$ such that

$$c_1 N_{m+k'}(x,y)^{-1}e^{(S+1)k'} \leqslant D \leqslant \hat{c}N_{m+k'}(x,y)^{-1}e^{(S+1)k'}. \tag{5.8}$$

Thus we have

$$\frac{R(x,y)}{e^{k'}} \leqslant \hat{c}^{1/(S+1)}\left(\frac{t}{N_{m+k'}(x,y)}\right)^{1/(S+1)}.$$

Now choose a minimum $\Lambda_{m+k'}$-walk $\pi \in \Pi_{m+k'}(x,y)$ with $\pi = \{x_i\}_{i=0}^{|\pi|}$ where $x_0 = x$, $x_{|\pi|} = y$ and $|\pi| = N_{m+k'}(x,y)$. Then

$$R(x_i, x_{i+1}) \leqslant c_2 e^{-m-k'} \leqslant c_2 e \frac{R(x,y)}{e^{k'}} \leqslant c_2 e \hat{c}^{1/(S+1)} \left(\frac{t}{N_{m+k'}(x,y)} \right)^{1/(S+1)}.$$

For $\varepsilon = e^{-m-k'-1}$, write $G_i = B_\varepsilon(x_i)$. If $z_i \in G_i$ and $z_{i+1} \in G_{i+1}$, we have

$$R(z_i, z_{i+1}) \leqslant 2\varepsilon + R(x_i, x_{i+1}) \leqslant (2 + c_2 e) \hat{c}^{1/(S+1)} \left(\frac{t}{N_{m+k'}(x,y)} \right)^{1/(S+1)}.$$

Choose $\hat{c}$ small enough so that $(2 + c_2 e)\hat{c}^{1/(S+1)} < c_{5.4}^{S+1}$. We can then apply the chaining argument with $N = N_{m+k'}(x,y)$, $s = t/N$,

$$\begin{aligned} p_t(x,y) &\geqslant \int_{G_1} \cdots \int_{G_{N-1}} p_s(x,x_1) \ldots p_s(x_{N-1},y)\, \mu(dx_1) \ldots \mu(dx_{N-1}) \\ &\geqslant \left(\prod_{i=1}^{N-1} \mu(G_i) \right) (c_3 s^{-S/(S+1)})^N \\ &\geqslant c_4 s^{-S/(S+1)} \exp(-c_5 N). \end{aligned}$$

In the last inequality, we use the fact that $\mu(G_i) s^{-S/(S+1)}$ is bounded from above and below, which comes from Lemma 3.4 and the definition of k'. As in (3.10), $N_{m+k'} \geqslant c_6 N_{m+k}$ and using (4.6) completes the proof.

Theorem 1.1 and Corollary 1.2 then follow as a consequence of (4.4), (5.6) and (4.6) with appropriate identification of constants.

6. *Examples*

In this section, we discuss some examples for which more exact estimates can be obtained.

1. *Affine nested fractals*

The first examples are affine nested fractals. This is a subclass of p.c.f. self-similar sets which consists of fractals with strong symmetries. Specifically, a connected p.c.f. self-similar fractal $\mathscr{L} = (K, S, \{F_i\}_{i \in S})$ is called an affine nested fractal if the following hold:

(A1) F_i: $\mathbb{R}^D \to \mathbb{R}^D$ is a contraction such that

$$|F_i(x) - F_i(y)| = \alpha_i^{-1} |x - y| \quad \text{for all } x, y \in \mathbb{R}^D,$$

for some $\alpha_i > 1$ and $\{F_i\}_{i \in S}$ satisfies the open set condition;

(A2) if $x, y \in V_0$, then reflection in the hyperplane $H_{xy} = \{z \in \mathbb{R}^D\colon \|z - x\| = \|z - y\|\}$ maps V_n to itself.

For the heat kernel $p_t(x,y)$ of the Dirichlet form constructed by a regular harmonic structure with the measure mentioned in Theorem 2.11(2), we have the following Aronson type estimates [**8**].

THEOREM 6.1. *There exist constants* $c_{6.1}$, $c_{6.2}$, $c_{6.3}$, $c_{6.4} > 0$ *and* $0 < \gamma' < S + 1$ *such that*

$$\Psi(c_{6.1} R(x,y), c_{6.2} t) \leqslant p_t(x,y) \leqslant \Psi(c_{6.3} R(x,y), c_{6.4} t),$$

for all $0 < t < 1$, $x, y \in E$ *where*

$$\Psi(z,t) = t^{-d_s/2} \exp\left(-\left(\frac{z^{S+1}}{t}\right)^{\gamma'/(S+1-\gamma')} \right).$$

The symmetry condition (A2) is essential for such an estimate. Under this symmetry, the exact chemical exponent γ' can be found as the solution to an optimization problem.

2. *Sierpinski gasket*

Another example is provided by the operator on the Sierpinski gasket mentioned in [**20**, § 6]. In § 2, it was shown how to view the 2-dimensional Sierpinski gasket as a p.c.f. self-similar fractal. Instead of constructing the Brownian motion, we let

$$D = \begin{pmatrix} -2p & p & p \\ p & -1 & 1-p \\ p & 1-p & -1 \end{pmatrix}, \quad r = (s^{-1}, 1, 1).$$

Straightforward calculations show that the equation (2.1) is equivalent to the following:

$$(p-1)(p-3)s^2 - 2(p-1)^2 s + p(p-2) = 0, \quad \lambda = \frac{2+s-p}{s(2-p)}.$$

As this equation has a unique positive solution s for each $0 < p < 1$, we know that there exists a unique harmonic structure for each $0 < p < 1$. It is easy to check that this harmonic structure is regular. Thus we have a regular local Dirichlet form for each $0 < p < 1$. Note that when $p = \frac{1}{2}$, it is the Brownian motion on the Sierpinski gasket.

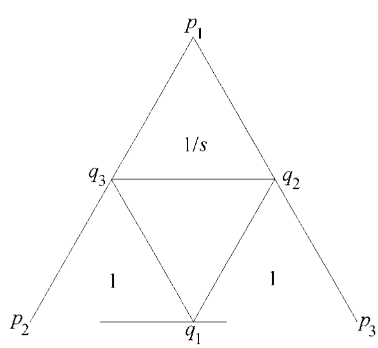

FIG. 2. The Sierpinski gasket with different resistors.

In order to estimate the heat kernel of the corresponding Dirichlet form, we prepare a lemma.

LEMMA 6.2. *For* $p_1, p_2, p_3 \in V_0$, *labelled as above, there exist* $c_{6.5}, c_{6.6}, c_{6.7}, c_{6.8} > 0$ *and*

$$d_c^h = \frac{\log 2}{\log \rho_2}, \quad d_c^v = \{s\colon \rho_1^{-s} + \rho_2^{-s} = 1\},$$

such that

$$c_{6.5}e^{d_c^v n} \leqslant N_n(p_1, p_i) \leqslant c_{6.6}e^{d_c^v n}, \quad \textit{for } i = 2, 3,$$

and for $0 < p < \frac{1}{2}$,

$$c_{6.7}e^{d_c^h n} \leqslant N_n(p_2, p_3) \leqslant c_{6.8}e^{d_c^h n},$$

while for $\frac{1}{2} \leqslant p < 1$,

$$c_{6.5}e^{d_c^v n} \leqslant N_n(p_2, p_3) \leqslant c_{6.6}e^{d_c^v n},$$

for all $n \geqslant 0$.

Proof. By symmetry $N_n(p_1, p_2) = N_n(p_1, p_3)$. For each case we can prove that

$$c_1 N_l N_m \leqslant N_{l+m} \leqslant c_2 N_l N_m \quad \text{for all } n, m \geqslant 0.$$

Thus, by an argument using super-(sub-)additive sequences (see, for example, [**4**, Theorem 5.1]), we see that

$$c_3 \alpha_{ij}^l \leqslant N_l(p_i, p_j) \leqslant c_4 \alpha_{ij}^l \quad \text{for all } l \geqslant 0, \tag{6.1}$$

for some $\alpha_{ij} > 0$.

In order to compute the values for the scale factors we observe that on the Sierpinski gasket the shortest paths consist of straight lines. The chemical exponent is then the box dimension of the shortest path in the resistance metric. To see this, note that the box dimension d_B, when it exists, is given by

$$d_B = \lim_{\varepsilon \to 0} \frac{\log N_\varepsilon^R}{-\log \varepsilon},$$

where N_ε^R is the minimum number of sets of radius ε in the resistance metric required to cover the path. For $\varepsilon = e^{-k}$ this is $N_k(p_i, p_j)$. Hence, by (6.1), we see that the limit exists for the sequence (and it is easy to show d_B exists), giving

$$d_B = \lim_{k \to \infty} \frac{\log N_k(x, y)}{k} = \lim_{k \to \infty} d_k^c(x, y).$$

We thus see that the scaling in the length of the shortest path will be determined by the box dimension.

Next, we calculate the exact values. We first consider the horizontal direction $\overline{p_2 p_3}$ for $p < \frac{1}{2}$. In this case, we can easily calculate the dimension of the straight line path as the resistance of each triangle on a given level is the same. Thus $N_k(p_2, p_3) = 2^l$ where l is the number of maps applied in order that the resistance of a piece is of order e^{-k}. Hence we see that $d_c^h = \lim_{k \to \infty} \log N_k(p_2, p_3)/k = \log 2/\log \rho_2$. For $p \geqslant \frac{1}{2}$, as it is possible to make a horizontal step via two diagonal ones, this becomes the shortest path, and we see that the horizontal estimate coincides with the diagonal one. For the diagonal walk each step in the path on V_{Λ_k} is of resistance e^{-k} and the box dimension of the straight line is required. By comparing with the calculation of the dimension for self-similar fractals with different scale factors, which was remarked on after Theorem 3.8, we obtain the result.

Let $d_c = d_c^v \wedge d_c^h$. Also, for $x \in K$, $\varepsilon > 0$, set

$$L(x) = \{y \in K \colon \exists l \text{ horizontal line, } x, y \in l, \ l \subset K\},$$

$$L_\varepsilon(x) = \{y \in K \colon R(L(x), y) < \varepsilon\},$$

where a horizontal line is a line in $\mathbb{R}^2$ which is parallel to $\overline{p_2 p_3}$.

THEOREM 6.3. *Set*

$$\Psi(z,t:d) = t^{-d_s/2}\exp\left(-\left(\frac{z^{S+1}}{t}\right)^{d/(S+1-d)}\right).$$

Then, there exist constants $c_{6.9}, c_{6.10}, c_{6.11}, c_{6.12}, c_{6.9.\varepsilon}, c_{6.11.\varepsilon} > 0$ *(*$c_{6.9.\varepsilon}$ *and* $c_{6.11.\varepsilon}$ *depend on* ε*) such that the following hold.*

(1) *For the case* $\frac{1}{2} \leqslant p < 1$,

$$\Psi(c_{6.9}R(x,y), c_{6.10}t : d_c) \leqslant p_t(x,y) \leqslant \Psi(c_{6.11}R(x,y), c_{6.12}t : d_c),$$

for all $0 < t < 1$, $x, y \in K$.

(2) *For the case* $0 < p < \frac{1}{2}$, *we have for all* $x \in K$, $y \in K \setminus L_\varepsilon(x)$, *that there exists a function* $f(x,y,\varepsilon) > 0$ *such that*

$$\Psi(c_{6.9.\varepsilon}R(x,y), c_{6.10}t : d_c^h) \leqslant p_t(x,y) \leqslant \Psi(c_{6.11.\varepsilon}R(x,y), c_{6.12}t : d_c^h), \tag{6.2}$$

if $f(x,y,\varepsilon) < t < 1$, *while*

$$\Psi(c_{6.9.\varepsilon}R(x,y), c_{6.10}t : d_c^v) \leqslant p_t(x,y) \leqslant \Psi(c_{6.11.\varepsilon}R(x,y), c_{6.12}t : d_c^v), \tag{6.3}$$

for all $0 < t < f(x,y,\varepsilon)$. *For* $x \in K$, $y \in L(x)$, (6.2) *with* $c_{6.9}$ *and* $c_{6.11}$ *instead of* $c_{6.9.\varepsilon}$ *and* $c_{6.11.\varepsilon}$ *holds for all* $0 < t < 1$.

Proof. Observe that if $p \geqslant \frac{1}{2}$, then $s \geqslant 1$ and $\rho_1 \geqslant \rho_2 = \rho_3$ so that $N_n(p_1, p_j) \leqslant N_n(p_2, p_3)$ $(j = 2, 3)$ and $d_c^v \leqslant d_c^h$. On the other hand, when $p < \frac{1}{2}$, $d_c^v > d_c^h$. Thus, we see that $d_c = d_c^v$ in Case (1), while $d_c = d_c^h$ in (2).

In order to obtain the result we apply Theorem 1.1. We will compute the shortest path in the effective resistance metric as above.

For Case (1) the shortest paths between all points involve diagonal steps so that the result is easily deduced from Theorem 1.1 and Lemma 6.2.

For Case (2), choose m so that $e^{-m-1} \leqslant R(x,y) \leqslant e^{-m}$ and decompose the shortest path $\pi = \{x_i\}_{i=1}^{|\pi|}$ between x and y in the following way. For each $m+1 \leqslant j \leqslant m+k$ $(k = k(m,t))$, let $x_{i_j}^1$ denote the first point in the path when it hits a vertex of V_{Λ_j}, and let $x_{i_j}^2$ denote the last such vertex. We regard our path π as consisting of the pairs $(x_{i_j}^1, x_{i_{j-1}}^1)$, $(x_{i_{j-1}}^2, x_{i_j}^2)$, which each consist of a path containing a certain number of horizontal and diagonal steps on V_{Λ_j}. Thus, using Lemma 6.2, we can write

$$N_{m+k}(x,y) = \sum_{l=1}^{2}\sum_{j=m+2}^{m+k} N_{m+k}(x_{i_j}^l, x_{i_{j-1}}^l) + N_{m+k}(x_{i_{m+1}}^1, x_{i_{m+1}}^2)$$

$$\sim \sum_{j=m+1}^{m+k} c_j^h(x,y)\, e^{d_c^h(m+k-j)} + c_j^v(x,y)\, e^{d_c^v(m+k-j)},$$

where the coefficients $c_j^{h(v)}(x,y)$ give the number of horizontal (diagonal) steps of the j size between x and y. (Here $f \sim g$ means that g/f is bounded from above and below by some positive constant which is independent of the choice of x, y, k.) By the construction of the path we see that these are bounded above by 2 for $m+2 \leqslant j \leqslant m+k$. For $j = m+1$, we see from Lemma 3.4 that it is also bounded from above and below by some uniform constant.

For $y \in L(x)$, we have $c_j^v(x,y) = 0$ for all j so that

$$c_1 e^{d_c^h k} \leqslant N_{m+k}(x,y) \leqslant c_2 e^{d_c^h k}, \quad \text{for all } k \geqslant 0.$$

We now fix a constant $M > 0$. By examining the coefficients $c_j^h(x,y)$, $c_j^v(x,y)$, we see that for each $x, y \in K$ with $y \notin L_\varepsilon(x)$ there exists a $k_\varepsilon^* < \infty$ (independent of x and y) such that

$$c_3 e^{d_c^h k} \leqslant N_{m+k}(x,y) \leqslant M e^{d_c^h k}, \quad \text{for all } k \leqslant k_\varepsilon^*,$$

$$M' e^{d_c^v k} \leqslant N_{m+k}(x,y) \leqslant c_4 e^{d_c^v k}, \quad \text{for all } k > k_\varepsilon^*,$$

where $M' = M \exp(k_\varepsilon^*(d_c^h - d_c^v))$ (note that M' depends on ε). By the definition of $k(m,t)$ if $k(m,t) \leqslant k_\varepsilon^*$, then for all $x \in K$, $y \notin L_\varepsilon(x)$,

$$N_{m+k}(x,y) \sim e^{d_c^h k} \sim \left(\frac{R(x,y)^{S+1}}{t}\right)^{d_c^h/(S+1-d_c^h)}.$$

However if $k(m,t) \geqslant k_\varepsilon^*$, then we get

$$N_{m+k}(x,y) \sim e^{d_c^v k} \sim \left(\frac{R(x,y)^{S+1}}{t}\right)^{d_c^v/(S+1-d_c^v)}.$$

(Here $f \sim g$ has the same meaning as above, but this time the constants depend on ε.) Using the fact that $k(m,t) \nearrow \infty$ as $t \downarrow 0$, we have the existence of the function $f(x,y,\varepsilon) = \{t\colon k(m,t) = k_\varepsilon^*\} \wedge 1$ with the desired properties.

This example suggests that without strong symmetry in the operator on the fractal the Aronson type uniform estimates do not hold. By rearranging this result and using the fact that for each ε, the function $f(x,y,\varepsilon) > 0$ for $y \notin L_\varepsilon(x)$, we have the following result.

COROLLARY 6.4. *For all $x, y \in K$, the following hold.*

(1) *For the case $\frac{1}{2} \leqslant p < 1$,*

$$\lim_{t \to 0} \frac{\log(-\log(t^{d_s/2} p_t(x,y)))}{\log(R(x,y)^{S+1}/t)} = d_c^v/(S+1-d_c^v).$$

(2) *For the case $0 < p < \frac{1}{2}$,*

$$\lim_{t \to 0} \frac{\log(-\log(t^{d_s/2} p_t(x,y)))}{\log(R(x,y)^{S+1}/t)} = \begin{cases} d_c^h/(S+1-d_c^h) & \text{if } y \in L(x), \\ d_c^v/(S+1-d_c^v) & \text{if } y \notin L(x). \end{cases}$$

3. *Vicsek sets*

We consider briefly the Vicsek set as discussed in [**22**]. It was shown that if the operator is not invariant under the complete symmetry group, there can exist a family of non-unique fixed points for this particular fractal. We show that if we think of this set as a p.c.f. self-similar set, then this non-uniqueness can be extended to symmetric resistances on the set. If we write the Vicsek set as a p.c.f. self-similar set with $S = \{1, \ldots, 5\}$, where the index 5 refers to the central square, we have

$$D = \begin{pmatrix} -(2+\beta) & 1 & \beta & 1 \\ 1 & -(2+1/\beta) & 1 & 1/\beta \\ \beta & 1 & -(2+\beta) & 1 \\ 1 & 1/\beta & 1 & -(2+1/\beta) \end{pmatrix},$$

$$r = (1/s, 1/s, 1/s, 1/s, 1).$$

For each $s > 0$, this D is a fixed point for all $\beta \in \mathbb{R}_+$ with $\lambda = 1 + 2s$ (note that $s = 1$ is the case considered in [**22**]). Using the results derived here it is easy to see that, for each s, the heat kernel estimates are of the same form for each diffusion on the Vicsek set defined by a fixed point from this one-parameter family of fixed points.

From the symmetry of the fractal and the resistances, and the essentially tree-like structure of this set, we see that for the vertices of the unit square p_i, for $i = 1, 2, 3, 4$, the number of steps in the shortest path $N(p_i, p_j)$, with $i, j = 1, 2, 3, 4$, will be the same. By sub and super additivity arguments there will be a unique exponent for all paths. As the values of r are given and these determine the behaviour of the shortest path, we see by calculation that for $s = 1$ the effective resistance metric is equivalent to the Euclidean. Thus $d_c(x, y) = 1$ for all x, y. The result for nested fractals ($s = 1$), derived in [**19**] holds for each β.

Further, our results show that for each s, we have $S = \log(1 + 4s)/\log(1 + 2s)$, and for each pair of points the chemical exponent $d_c = d_c(x, y)$ will be given by

$$d_c = \{d\colon (1 + 2s^{-d})(1 + 2s)^{-d} = 1\}.$$

Thus for the Vicsek set F with harmonic structure (D, r) as above, we find that, for each $\beta > 0$, there exist constants $c_1, c_2, c_3, c_4 > 0$ such that

$$\Psi(c_1 R(x, y), c_2 t) \leqslant p_t(x, y) \leqslant \Psi(c_3 R(x, y), c_4 t),$$

for all $0 < t < 1$, $x, y \in F$ where

$$\Psi(z, t) = t^{-d_s/2} \exp\left(-\left(\frac{z^{S+1}}{t} \right)^{d_c/(S+1-d_c)} \right).$$

4. *abc-gaskets*

Finally, we consider fractals which are not symmetric. These are the *abc*-gaskets of [**12**]. It was shown that for these fractals there may not be a fixed point corresponding to the 'natural' choice of resistance. The sets are constructed by setting a number $a + 1$ of triangles on the bottom side, $b + 1$ on the left side and $c + 1$ on the right side. Thus there are $a + b + c$ triangles in total; we always assume that $a \leqslant b \leqslant c$. The Sierpinski gasket is the case $a = b = c = 1$.

We write this as a p.c.f. set by setting

$$S = \{1, \ldots, a + b + c\},$$

$$\pi(C) = \{q_1, \ldots, q_{a+b+c+3}\},$$

$$\pi(P) = \{p_1, p_2, p_3\}, \quad \pi^{-1}(p_i) = \{\dot{i}\} \text{ for } i = 1, 2, 3.$$

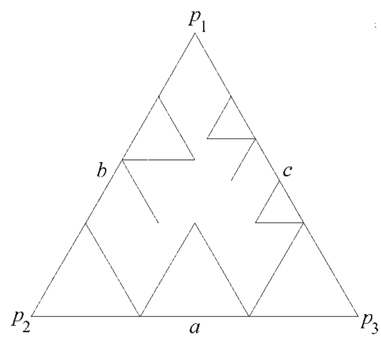

FIG. 3. An *abc*-gasket with $a = 2$, $b = 3$, $c = 4$.

There are several possible ways to construct a natural Laplace operator on this set. In the paper [**12**], the resistance vector r is assumed to be a vector of unit resistances and then the fixed point equation can be solved to find a matrix D, with positive elements, under the condition that $1/a < 1/b + 1/c$. In this case the scale factor $\lambda = (2bac + ab + ac + cb)/(ac + ab + cb)$ and hence, as each $r_i = 1$, we have $\rho_i = \lambda$ for each i. To find the dimension of the straight line paths across the triangle in the resistance metric is then straightforward. Let

$$d_c^{23} = \frac{\log(a+1)}{\log \lambda}, \quad d_c^{12} = \frac{\log(b+1)}{\log \lambda}, \quad d_c^{13} = \frac{\log(c+1)}{\log \lambda},$$

so that $d_c^{23} \leqslant d_c^{12} \leqslant d_c^{13}$.

For all pairs of points in K we can either move directly along a horizontal line between them, or move along a horizontal line and a +60 diagonal. This leads to the following result, proved by calculating the behaviour of the shortest path as in the Sierpinski gasket example. As the proof is essentially the same, we omit it.

THEOREM 6.5. *For the abc gasket there exist constants and a function* $f(x,y,\varepsilon) > 0$ *such that the following holds for all* $x \in K$, $y \in K \setminus L_\varepsilon(x)$:

$$\Psi(c_{6.13.\varepsilon} R(x,y), c_{6.14} t : d_c^{23}) \leqslant p_t(x,y) \leqslant \Psi(c_{6.15.\varepsilon} R(x,y), c_{6.16} t : d_c^{23}), \tag{6.4}$$

for all $f(x,y,\varepsilon) < t < 1$, *while*

$$\Psi(c_{6.17.\varepsilon} R(x,y), c_{6.18} t : d_c^{12}) \leqslant p_t(x,y) \leqslant \Psi(c_{6.19.\varepsilon} R(x,y), c_{6.20} t : d_c^{12}), \tag{6.5}$$

for all $0 < t < f(x,y,\varepsilon)$.

For $x \in K$, $y \in L(x)$, (6.4) *with* $c_{6.13}, c_{6.15}$ *instead of* $c_{6.13.\varepsilon}, c_{6.15.\varepsilon}$ *holds for all* $0 < t < 1$.

COROLLARY 6.6. *For all* $x, y \in K$, *the following holds:*

$$\lim_{t \to 0} \frac{\log(-\log(t^{d_s/2} p_t(x,y)))}{\log(R(x,y)^{S+1}/t)} = \begin{cases} d_c^{23}/(S+1-d_c^{23}) & \text{if } y \in L(x), \\ d_c^{12}/(S+1-d_c^{12}) & \text{if } y \notin L(x). \end{cases}$$

It is interesting to note that for each $x \in K$, the estimates are determined by the middle exponent d_c^{12} for almost all $y \in K$.

The other natural approach to this particular p.c.f. set is to consider the matrix D to consist of unit conductors and then scale this by a resistance factor which depends on the size of the mapping on the triangle. Thus we will need three factors to correspond to the three different sizes of triangle. Our operator is defined by

$$D = \begin{pmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{pmatrix}, \quad r = (r_1, r_2, r_3).$$

Then there exist a λ and a vector r which makes D a fixed point of the conductivity map, and the values of the conductivities ρ_i are given by

$$\rho_1 = \frac{-2a + 2ab - 3 + 4b + 2B}{3(2b-1)},$$

$$\rho_2 = \frac{(-2a + 2ab - 3 + 4b + 2B)(2b+1)}{3(2ab - 2a - 1 + 2B)},$$

$$\rho_3 = \frac{(-2a + 2ab - 3 + 4b + 2B)(2b+1)(c-1)}{3(1 + 2a - b - 4ba + 2ab^2 + 2(b-1)B)},$$

where $B = \sqrt{a^2 - 2ba^2 - ab + b^2a^2 + 4ab^2 + b^2}$. Note that by the choice of r we obtain a set of ρ_i with $1 \leqslant \rho_1 \leqslant \rho_2 \leqslant \rho_3$ for all a, b, c for this matrix D, unlike the previous case of the operator from [**12**].

To calculate the heat kernel estimates we return to our basic result and determine the shortest paths in the resistance metric. For this we need three different exponents corresponding to the three possible directions in this fractal. By calculating the dimension of the straight paths between points in the resistance metric, we obtain the following where the upper indices refer to the direction:

$$d_c^{23} = \frac{\log(a+1)}{\log \rho_1}, \quad d_c^{12} = \{s\colon \rho_1^{-s} + b\rho_2^{-s} = 1\},$$

$$d_c^{13} = \{s\colon \rho_1^{-s} + \rho_2^{-s} + (c-1)\rho_3^{-s} = 1\}.$$

Again there will be a particular direction which has the shortest path. We write d_c^1 for the smallest element of the set $\{d_c^{12}, d_c^{13}, d_c^{23}\}$, and d_c^2 for the next smallest. In this case, the result is the following. Here, for the pair $i, j \in \{1,2,3\}$ such that $d_c^{ij} = d_c^1$, we denote by $L^1(x)$ a line which is parallel to $\overline{p_i p_j}$ and contains the point $x \in K$ ($L^1_\varepsilon(x)$ is defined in the same way as before).

THEOREM 6.7. *For the abc gasket with harmonic structure (D, r) there exist constants and a function $f(x, y, \varepsilon) > 0$ such that the following holds for all $x \in K$, $y \in K \setminus L^1_\varepsilon(x)$:*

$$\Psi(c_{6.21.\varepsilon} R(x,y), c_{6.22}t : d_c^1) \leqslant p_t(x,y) \leqslant \Psi(c_{6.23.\varepsilon} R(x,y), c_{6.24}t : d_c^1), \qquad (6.6)$$

for all $f(x, y, \varepsilon) < t < 1$, while

$$\Psi(c_{6.25.\varepsilon} R(x,y), c_{6.26}t : d_c^2) \leqslant p_t(x,y) \leqslant \Psi(c_{6.27.\varepsilon} R(x,y), c_{6.28}t : d_c^2), \qquad (6.7)$$

for all $0 < t < f(x, y, \varepsilon)$.

For $x \in K$, $y \in L^1(x)$, (6.6) *with $c_{6.21}, c_{6.23}$ instead of $c_{6.21.\varepsilon}, c_{6.23.\varepsilon}$ holds for all* $0 < t < 1$.

COROLLARY 6.8. *For all $x, y \in K$, the following holds:*

$$\lim_{t\to 0} \frac{\log(-\log(t^{d_s/2}p_t(x,y)))}{\log(R(x,y)^{S+1}/t)} = \begin{cases} d_c^1/(S+1-d_c^1) & \text{if } y \in L^1(x), \\ d_c^2/(S+1-d_c^2) & \text{if } y \notin L^1(x). \end{cases}$$

Finally, we conjecture that Corollaries 6.4, 6.6, 6.8 could be generalized as follows.

CONJECTURE 6.9. There exists a function $d_c(x,y) = \lim_{k\to\infty} d_k^c(x,y)$ so that the following holds for all $x, y \in K$:

$$\lim_{t\to 0} \frac{\log(-\log t^{d_s/2}p_t(x,y))}{\log(R(x,y)^{S+1}/t)} = d_c(x,y)/(S+1-d_c(x,y)).$$

We further conjecture that the chemical exponent can be expressed in terms of the box counting dimension of the path in the effective resistance metric. Let $\Pi(x,y)$ be the set of all paths from x to y and $d_B^R(\pi)$ be the box dimension of path $\pi \in \Pi$, in the effective resistance metric.

CONJECTURE 6.10. The limiting chemical exponent $d_c(x,y)$ for the shortest path between x, y exists, and can be expressed in terms of the box counting dimension of the path in the resistance metric as

$$d_c(x,y) = \inf_{\pi\in\Pi(x,y)} d_B^R(\pi) = \lim_{k\to\infty} d_k^c(x,y),$$

for all $x, y \in K$.

Acknowledgment. This work was done while the first author was visiting Nagoya University with the aid of EPSRC grant GR/L07840.

References

1. M. T. BARLOW, 'Random walks, electrical resistance, and nested fractals', *Asymptotic problems in probability theory: stochastic models and diffusion on fractals* (eds K. D. Elworthy and N. Ikeda, Pitman, Montreal, 1993), pp. 131–157.
2. M. T. BARLOW, 'Diffusions on fractals', *Lectures in probability theory and statistics. École d'été de probabilités de St Flour XXV—1995* (M. T. Barlow and D. Nualart), Lecture Notes in Mathematics 1690 (Springer, Berlin, 1998).
3. M. T. BARLOW and R. F. BASS, 'Construction of Brownian motion on the Sierpinski carpet', *Ann. Inst. H. Poincaré* 25 (1989) 101–102.
4. M. T. BARLOW and R. F. BASS, 'On the resistance of the Sierpiński carpet', *Proc. Roy. Soc. London Ser. A* 431 (1990) 345–360.
5. M. T. BARLOW and R. F. BASS, 'Transition densities for Brownian motion on the Sierpinski carpet', *Probab. Theory Related Fields* 91 (1992) 307–330.
6. M. T. BARLOW and B. M. HAMBLY, 'Transition density estimates for Brownian motion on scale irregular Sierpinski gaskets', *Ann. Inst. H. Poincaré* 33 (1997) 531–557.
7. M. T. BARLOW and J. KIGAMI, 'Localized eigenfunctions of the Laplacian on p.c.f. self-similar sets', *J. London Math. Soc.* (2) 56 (1997) 320–332.
8. P. J. FITZSIMMONS, B. M. HAMBLY and T. KUMAGAI, 'Transition density estimates for Brownian motion on affine nested fractals', *Commun. Math. Phys.* 165 (1994) 595–620.
9. A. A. GRIGORYAN, 'The heat equation on noncompact Riemannian manifolds', *Math. Sb.* 182 (1991) 55–87.
10. B. M. HAMBLY, 'Brownian motion on a random recursive Sierpinski gasket', *Ann. Probab.* 25 (1997) 1059–1103.
11. B. M. HAMBLY and T. KUMAGAI, 'Heat kernel estimates and homogenization for asymptotically lower dimensional processes on some nested fractals', *Potential Anal.* 8 (1998) 359–397.

12. K. Hattori, T. Hattori and H. Watanabe, 'Gaussian field theories on general networks and the spectral dimensions', *Progr. Theoret. Phys. Suppl.* 92 (1987) 108–143.
13. S. Havlin and D. Ben-Avraham, 'Diffusion in disordered media', *Adv. in Phys.* 36 (1987) 695–798.
14. J. Hutchinson, 'Fractals and self-similarity', *Indiana Univ. Math. J.* 30 (1981) 713–747.
15. J. Kigami, 'Harmonic calculus on P.C.F. self-similar sets', *Trans. Amer. Math. Soc.* 335 (1993) 721–755.
16. J. Kigami, 'Effective resistances for harmonic structures on p.c.f. self-similar sets', *Math. Proc. Cambridge Philos. Soc.* 115 (1994) 291–303.
17. J. Kigami, 'Harmonic calculus on limits of networks and its application to dendrites', *J. Funct. Anal.* 128 (1995) 48–86.
18. J. Kigami and M. L. Lapidus, 'Weyl's problem for the spectral distribution of Laplacians on P.C.F. self-similar fractals', *Commun. Math. Phys.* 158 (1993) 93–125.
19. T. Kumagai, 'Estimates of transition densities for Brownian motion on nested fractals', *Probab. Theory Related Fields* 96 (1993) 205–224.
20. T. Kumagai, 'Regularity, closedness and spectral dimensions of the Dirichlet forms on P.C.F. self-similar sets', *J. Math. Kyoto Univ.* 33 (1993) 765–786.
21. S. Kusuoka and X. Y. Zhou, 'Dirichlet forms on fractals: Poincare constants', *Probab. Theory Related Fields* 93 (1992) 169–196.
22. V. Metz, 'How many diffusions exist on the Vicsek snowflake?', *Acta. Appl. Math.* 32 (1993) 227–241.
23. C. Sabot, 'Existence and uniqueness of diffusions on finitely ramified self-similar fractals', *Ann. Sci. École Norm. Sup.* 30 (1997) 605–673.
24. L. Saloff-Coste, 'A note on Poincaré, Sobolev and Harnack inequality', *Internat. Math. Res. Notices* 2 (1992) 27–38.
25. T. Shima, 'On eigenvalue problems for Laplacians on P.C.F. self-similar sets', *Japan. J. Indust. Appl. Math.* 13 (1996) 1–23.

Department of Mathematics and Statistics
University of Edinburgh
The King's Buildings
Mayfield Road
*Edinburgh EH*9 3*JZ*

E-mail: bmh@maths.ed.ac.uk

Graduate School of Informatics
Kyoto University
Sakyoku
Kyoto 606-8501
Japan