Measurable Riemannian geometry on the Sierpinski gasket: the Kusuoka measure and the Gaussian heat kernel estimate

Jun Kigami
Graduate School of Informatics
Kyoto University
Kyoto 606-8501, JAPAN
e-mail:kigami@i.kyoto-u.ac.jp

Abstract

We study the standard Dirichlet form and its energy measure, called the Kusuoka measure, on the Sierpinski gasket as a prototype of "measurable Riemannian geometry". The shortest path metric on the harmonic Sierpinski gasket is shown to be the geodesic distance associated with the "measurable Riemannian structure". The Kusuoka measure is shown to have the volume doubling property with respect to the Euclidean distance and also to the geodesic distance. Li-Yau type Gaussian off-diagonal heat kernel estimate is established for the heat kernel associated with the Kusuoka measure.

1 Introduction

The main purpose of this paper is to present a prototype of a theory which could be called "measurable Riemannian geometry" on metric spaces. Our prototype is the standard Dirichlet form on the Sierpinski gasket, which is a typical example of self-similar sets. (See Figure 1.) The Brownian motion on the Sierpinski gasket has been constructed as a scaling limit of random walks independently by Kusuoka[10] and Goldstein[3]. Then analytical counterpart of the Brownian motion has been developed through the works [6], [11] and [7]. In particular, they have identified the Dirichlet form $(\mathcal{E}, \mathcal{F})$, where $\mathcal{E}$ is the bilinear form whose domain is $\mathcal{F}$, associated with the Brownian motion. This form $(\mathcal{E}, \mathcal{F})$ is called the standard Dirichlet form on the Sierpinski gasket. See Section 2. Moreover, in [11], Kusuoka has revealed a structure of the Dirichlet form $(\mathcal{E}, \mathcal{F})$ which is analogous to the Riemannian geometry. Precisely, he has shown the existence of a measure ν, a non-negative symmetric matrix Z and a operator $\widetilde{\nabla}$ such that, for any $u, v \in \mathcal{F}$,

$$\mathcal{E}(u,v) = \int_K (\widetilde{\nabla}u, Z\widetilde{\nabla}v)d\nu,$$

where Z and $\widetilde{\nabla}u, \widetilde{\nabla}v$ are ν-measurable (matrix or vector-valued) functions defined on ν-a.e. $x \in K$. See Section 2 for the brief review of those results. The measure ν is now called the Kusuoka measure. Obviously ν, Z and $\widetilde{\nabla}$ correspond

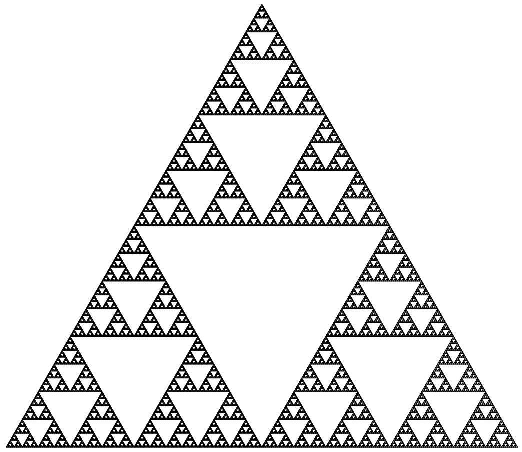

Figure 1: the Sierpinski gasket

to the Riemannian volume, the Riemannian metric and the gradient operator respectively, but they are only defined in measurable sense. This is the reason why $(\nu, Z, \widetilde{\nabla})$ is sometimes called the "measurable Riemannian structure" on the Sierpinski gasket. Later in [8], it has been shown that $\widetilde{\nabla}$ can be replaced by the natural gradient operator for differentiable functions if we use the harmonic functions as a coordinate. The Sierpinski gasket with harmonic functions as a coordinate is called the harmonic Sierpinski gasket, or the Sierpinski gasket with the harmonic coordinate. See Figure 2. We will give the details on the harmonic Sierpinski gasket in Section 4.

There is still one missing piece in the above mentioned measurable analogy of the Riemannian geometry. Namely, we have not found the counterpart of the geodesic distance. In Section 5, we will propose that the proper distance in this case is the shortest path distance d_* on the harmonic Sierpinski gasket, i.e. the minimum of lengths of rectifiable curves between points. We will show the existence of the shortest path (i.e. geodesic) for any two points. Moreover, if $\gamma : [0,1] \to K$ is the geodesic between x and y, then we will see that

$$d_*(x,y) = \int_0^1 (\frac{d\gamma}{dt}, Z\frac{d\gamma}{dt})dt.$$

in Section 5. This justifies that d_* is the geodesic distance associated the measurable Riemannian structure $(\nu, Z, \widetilde{\nabla})$.

Secondly, we will show that the Kusuoka measure (i.e. the Riemannian volume) has the volume doubling property with respect to the Euclidean distance in Section 3 and then to the shortest path metric in Section 6. This fact will be the key to show the Li-Yau type Gaussian estimate of the heat kernel, which will be shown in Section 6, associated with the measurable Riemannian structure on the Sierpinski gasket.

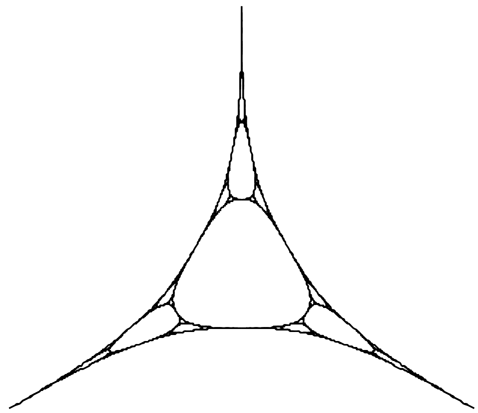

Figure 2: the harmonic Sierpinski gasket

Let $p(t,x,y)$ be the canonical heat kernel on a complete Riemannian manifold with non-negative Ricci curvature. In [13], Li and Yau have shown

$$\frac{c_1}{V(\sqrt{t},x)}\exp\Big(-c_2\frac{d(x,y)^2}{t}\Big)\le p(t,x,y) \\ \le\frac{c_3}{V(\sqrt{t},x)}\exp\Big(-c_4\frac{d(x,y)^2}{t}\Big), \quad (1.1)$$

where d is the geodesic distance and $V(r,x)$ is the Riemannian volume of a ball $\{y|d(x,y)<r\}$. (1.1) is called the Li-Yau type Gaussian estimate. Such an estimate has been shown to be equivalent to the Poincarè inequality and the volume doubling property by Grigor'yan[4] and Saloff-Coste[15].

Let $p_\nu(t,x,y)$ be the heat kernel on the Sierpinski gasket associated with the measurable Riemannian structure or, more precisely, the Dirichlet form $(\mathcal{E},\mathcal{F})$ on $L^2(K,\nu)$. In Section 6, $p_\nu(t,x,y)$ will be shown to satisfy the Li-Yau type Gaussian estimate (1.1), where we replace d and $V(r,x)$ by the harmonic shortest path metric d_* and $\nu(\{y|d_*(x,y)<r\})$ respectively.

Note that the diffusion process corresponding to this heat kernel is the time change of the Brownian motion on the Sierpinski gasket, which corresponds to the Dirichlet form $(\mathcal{E},\mathcal{F})$ on $L^2(K,\mu_*)$, where μ_* is the normalized Hausdorff measure on the original Sierpinski gasket. (The Kusuoka measure ν and the normalized Hausdorff measure μ_* are mutually singular. See Kusuoka[11].) In [1], Barlow and Perkins have shown that the heat kernel $p_{\mu_*}(t,x,y)$ associated with the Brownian motion on the Sierpinski gasket satisfy the following estimate:

$$
\begin{aligned}
c_1 t^{-d_S/2} \exp\left(-c_2 \left(\frac{|x-y|^{d_w}}{t} \right)^{\frac{1}{d_w-1}} \right) &\le p_{\mu_*}(t,x,y) \\
&\le c_3 t^{-d_S/2} \exp\left(-c_4 \left(\frac{|x-y|^{d_w}}{t} \right)^{\frac{1}{d_w-1}} \right) \qquad (1.2)
\end{aligned}
$$

for any $t \in (0,1]$ and any x, y in the Sierpinski gasket, where $d_S = \log 9/\log 5$ is called the spectral dimension and $d_w = \log 5/\log 2$ is called the walk dimension. The estimate (1.2) is called the sub-Gaussian estimate since $d_w > 2$ and hence the heat diffuses slower than the Gaussian case. In comparison with (1.2), our result (1.1) tells that the diffusion process regain the Gaussian scaling between time and space under the Kusuoka measure ν and the harmonic shortest path metric d_* which are more natural than the normalized Hausdorff measure μ_* and the Euclidean distance from the viewpoint of the measurable Riemannian structure.

Finally we will mention several papers related to measurable Riemannian structure. In [14], Metz and Sturm have obtained a weak Gaussian upper estimate of the heat kernel on the Sierpinski gasket with respect to the Kusuoka measure. See Section 6 for details. In [17], Teplyaev has studied gradient operator on post critically finite self-similar set and obtained some continuity results of gradient of certain functions. Recently, he has introduced the notion of fractals with finitely ramified cell structure and constructed the measurable Riemannian structure on them in [16]. It is not known that the results in this paper can be extended to this class or not.

2 Measurable Riemannian structure on the Sierpinski gasket

In this section, we first define the Sierpinski gasket and then introduce analytical objects which can be thought of as measurable analogies with those in the Riemannian geometry.

Let $N \ge 2$. Choose $\{p_1, \ldots, p_N\} \subset \mathbb{R}^{N-1}$ so that $|p_i - p_j| = 1$ for any i and j with $i \ne j$. Note that $\{p_1, \ldots, p_N\}$ is the set of vertices of a regular N-simplex.

Definition 2.1. Define $F_i : \mathbb{R}^{N-1} \to \mathbb{R}^{N-1}$ by $F_i(x) = (x - p_i)/2 + p_i$ for $i = 1, \ldots, N$. Let K be the self-similar set associated with the family of contractions $\{F_i\}_{i=1,\ldots,N}$, that is, K is the unique non-empty compact set which satisfies

$$K = \bigcup_{i=1}^{N} F_i(K).$$

We call K the N-Sierpinski gasket.

The 2-Sierpinski gasket is the closed interval $[p_1, p_2]$. (We assume that $p_1 < p_2$.) See Figure 1 for the 3-Sierpinski gasket.

The followings are standard notations.

Definition 2.2. (1) Let $S = \{1, \ldots, N\}$. Define $\Sigma = S^{\mathbb{N}} = \{i_1 i_2 \ldots | i_n \in S$ for any $n \in \mathbb{N}\}$ Also define $W_m = S^m = \{w_1 \ldots w_m | w_1, \ldots, w_m \in S\}$ for any $m = 0, 1, \ldots$, where $W_0 = \{\emptyset\}$, and set $W_* = \cup_{m \geq 0} W_m$. For $w \in W_*$, we define $|w| = m$ if $w \in W_m$.
(2) For $w = w_1 \ldots w_m \in W_*$, define $F_w = F_{w_1} \circ \ldots \circ F_{w_m}$ and $K_w = F_w(K)$. Also we set $\Sigma_w = \{i_1 i_2 \ldots | i_1 i_2 \ldots \in \Sigma, i_1 \ldots i_m = w_1 \ldots w_m\}$.
(3) Let $V_0 = \{p_1, \ldots, p_N\}$. Set $V_m = \cup_{w \in W_m} F_w(V_0)$ for any $m \geq 0$ and $V_* = \cup_{m \geq 0} V_m$.

Note that $V_m \subseteq V_{m+1}$ for any $m \geq 0$ and that the closure of V_* under the Euclidean metric is the Sierpinski gasket. The map π in the following proposition determines the structure of the Sierpinski gasket as a quotient of Σ.

Proposition 2.3. *There exists a continuous surjective map $\pi : \Sigma \to K$ such that $\{\pi(\omega)\} = \cap_{m \geq 0} K_{\omega_1 \ldots \omega_m}$ for any $\omega = \omega_1 \omega_2 \ldots \in \Sigma$. Define $\sigma_i : \Sigma \to \Sigma$ for $i \in S$ by $\sigma_i(i_1 i_2 \ldots) = i i_1 i_2 \ldots$. Then $\pi \circ \sigma_i = F_i \circ \pi$ for any $i \in S$. Moreover,*

$$\#(\pi^{-1}(x)) = \begin{cases} 2 & \text{if } x \in V_* \backslash V_0, \\ 1 & \text{otherwise.} \end{cases}.$$

Now we define a quadratic form $(\mathcal{E}, \mathcal{F})$ on the Sierpinski gasket, which corresponds to the Riemannian energy from the viewpoint of "measurable Riemannian geometry".

Definition 2.4. Let $C(K)$ be the collection of real-valued continuous functions on K. For $u, v \in C(K)$, define

$$\mathcal{E}_m(u, v) = \Big(\frac{N+2}{N}\Big)^m \sum_{p, q \in V_m . p \sim q} (u(p) - u(q))(v(p) - v(q)),$$

where $p \sim q$ if and only if $p, q \in F_w(V_0)$ for some $w \in W_m$.

If is known that $\mathcal{E}_m$ is non-negative definite symmetric quadratic form on $C(K)$ and that $\mathcal{E}_m(u, u) \leq \mathcal{E}_{m+1}(u, u)$ for any $u \in C(K)$ and $m \geq 0$.

Definition 2.5. Let $\mathcal{F} = \{u | u \in C(K), \lim_{m \to \infty} \mathcal{E}_m(u, u) < +\infty\}$. For $u, v \in \mathcal{E}$, define $\mathcal{E}(u, v) = \lim_{m \to \infty} \mathcal{E}_m(u, v)$.

Theorem 2.6. *If μ is a Borel regular probability measure on K which satisfies that $\mu(O) > 0$ for any non-empty open subset of K and that $\mu(F) = 0$ for any finite set F, then $(\mathcal{E}, \mathcal{F})$ is a local regular Dirichlet form on $L^2(K, \mu)$ and the corresponding diffusion process possesses a jointly continuous heat kernel $p_\mu(t, x, y)$.*

The above theorem has been proven in [9, Theorem 3.4.6] except the existence of a jointly continuous heat kernel $p_\mu(t,x,y)$, which we will prove in Section 6.

Remark. The above results, the construction of $(\mathcal{E},\mathcal{F})$ and Theorem 2.6, are true for more general class of self-similar sets called post critically self-similar sets. See [9] for details.

Next we introduce a counterpart of "Riemannian volume", which is called the Kusuoka measure, defined by Kusuoka [11]. It turns out to be the energy measure of the quadratic form $(\mathcal{E},\mathcal{F})$ as well. Set $e_i = \begin{pmatrix}\delta_{1i}\\ \vdots \\ \delta_{Ni}\end{pmatrix} \in \mathbb{R}^N$ for $i=1,\dots,N$, where δ_{kl} is Kronecker's delta defined by

$$\delta_{kl} = \begin{cases} 1 & \text{if } k=l,\\ 0 & \text{otherwise.}\end{cases}$$

Let $(\cdot,\cdot)$ be the standard inner product of $\mathbb{R}^N$: $(x,y) = \sum_{i=1}^N x_i y_i$ for $x = \sum_{i=1} x_i e_i$ and $y = \sum_{i=1}^N y_i e_i$. We define $|x| = \sqrt{(x,x)}$ for any $x \in \mathbb{R}^N$.

Definition 2.7. Define

$$M_0 = \{x | x = \sum_{i=1}^N x_i e_i \in \mathbb{R}^N, \sum_{i=1}^N x_i = 0\}.$$

Let $P : \mathbb{R}^N \to M_0$ be the orthogonal projection. Define $p_i = Pe_i/\sqrt{2}$ for $i=1,\dots,N$. Note that $|p_i - p_j| = 1$ if $i \neq j$ and $\{p_1,\dots,p_N\}$ is the collection of vertices of a regular N-simplex. For any i, we choose $\{f_{i,1},\dots,f_{i,N-2}\}$ so that $(p_i/|p_i|, f_{i,1},\dots,f_{i,N-2})$ is an orthonormal base of M_0. Then define a linear map $T_i : M_0 \to M_0$ by

$$T_i p_i = \frac{N}{N+2} p_i \qquad \text{and} \qquad T_i f_{i,j} = \frac{1}{N+2} f_{i,j}$$

for any $j = 1,\dots,N-2$.

Remark. Restricting the inner product $(\cdot,\cdot)$ on $\mathbb{R}^N$, we have a natural inner product on M_0, which is also denoted by $(\cdot,\cdot)$.

Hereafter, we regard M_0 as $\mathbb{R}^{N-1}$ in a natural manner and identify $p_1,\dots,p_N$ in Definition 2.7 with those in Definition 2.1.

Lemma 2.8. *Let $L(M_0)$ be the collection of linear operators form M_0 to itself. Let $(a_1,\dots,a_{N-1})$ be an orthonormal base of M_0. For any $X \in L(M_0)$, define the Hilbert-Schmidt norm $||X||_{\mathrm{HS}}$ by*

$$||X||_{\mathrm{HS}} = \sqrt{\sum_{i=1}^N |Xa_i|^2}.$$

Then $||\cdot||_{\mathrm{HS}}$ is a norm on $L(M_0)$. If $(x_{ij})_{1\leq i,j\leq N-1}$ is the matrix representation of X with respect to the orthonormal base $(a_1,\ldots,a_{N-1})$, then $||X||_{\mathrm{HS}} = \sqrt{\sum_{i,j}(x_{ij})^2}$. In particular, $||X||_{\mathrm{HS}} = ||{}^tX||_{\mathrm{HS}}$. Moreover, $||X||_{\mathrm{HS}}$ is independent of the choice of a orthonormal base $(a_1,\ldots,a_{N-1})$.

For a liner operator $X : M_0 \to M_0$, the adjoint operator of X with respect to the inner product $(\cdot,\cdot)$ is denoted by ${}^tX$.

Lemma 2.9. *Let $I : M_0 \to M_0$ be the identity. Then*

$$\sum_{i=1}^{N} T_i{}^t(T_i) = \frac{N}{N+2}I.$$

The next proposition is due to Kusuoka [11, (1.6)]

Proposition 2.10. *For any $w = w_1\ldots w_m \in W_*$, define $T_w = T_{w_1}\cdots T_{w_m}$. Then, there exists a unique Borel regular probability measure ν on Σ such that*

$$\nu(\Sigma_w) = \frac{1}{N-1}\Big(\frac{N+2}{N}\Big)^{|w|}(||T_w||_{\mathrm{HS}})^2$$

for any $w \in W_*$. Moreover ν is non-atomic.

Define $\pi_*\nu(A) = \nu(\pi^{-1}(A))$ for any Borel set $A \subseteq K$. Then by Proposition 2.3, $\pi_*\nu$ is a Borel regular probability measure on K and $\pi_*\nu(K_w) = \nu(\Sigma_w)$ for any $w \in W_*$. Moreover, since $\nu(V_*) = 0$, $(K,\pi_*\nu)$ may be identified with (Σ,ν) from measurable point of view. So, we abuse the notation and use ν to denote $\pi_*\nu$. This ν on K is called the Kusuoka measure on the Sierpinski gasket, which corresponds to "Riemannian volume".

Next we introduce "Riemannian metric" $Z(\cdot)$. Kusuoka has shown the following results in [11, (1.7)].

Proposition 2.11. *For $w \in W_m$, define $Z_m(w) = T_w{}^t(T_w)/||T_w||^2_{\mathrm{HS}}$. Then $Z(\omega) = \lim_{m\to\infty} Z_m(\omega_1\ldots\omega_m)$ exists for ν-a.e. $\omega = \omega_1\omega_2\ldots \in \Sigma$. Moreover,* $\mathrm{rank}Z(\omega) = 1$ *and $Z(\omega)$ is the orthogonal projection onto its image for ν-a.e. $\omega \in \Sigma$.*

Let $Z_*(x) = Z(\pi^{-1}(x))$. Since $\nu(V_*) = 0$, the above proposition implies that $Z_*(x)$ is well-defined, that $\mathrm{rank}Z(x) = 1$ and that $Z(x)$ is the orthogonal projection onto its image for ν-a.e. $x \in K$. This $Z_*(\cdot)$ plays the role of "Riemannian metric" on the Sierpinski gasket. For ease of notation, we use $Z(\cdot)$ instead of $Z_*(\cdot)$.

Remark. Let $\pi^{-1}(x) = \{\omega,\tau\}$ for $x \in V_*$. Then both $Z(\omega)$ and $Z(\tau)$ exist and $Z(\omega) = Z(\tau)$. (This fact can be easily shown by the discussion in the proof of Lemma 3.5.) Hence $Z(x)$ is well-defined on V_*.

Now we have "Riemannian energy" $\mathcal{E}(\cdot,\cdot)$, "Riemannian volume" ν and "Riemannian metric" $Z(\cdot)$. The following theorem obtained by Kusuoka [11, (5.1)] gives the legitimacy of such analogy with the Riemannian geometry.

Theorem 2.12. *There exists* $\widetilde{\nabla} : \mathcal{F} \to \{Y | Y : K \to M_0, Y$ *is* ν*-measurable*$\}$ *such that*

$$\mathcal{E}(u, v) = \int_K (\widetilde{\nabla} u, Z \widetilde{\nabla} v) d\nu$$

for any $u, v \in \mathcal{F}$.

The definition of $\widetilde{\nabla}$ will be given in Definition 4.11 by using the language of the harmonic Sierpinski gasket.

Remark. In terms of the theory of Dirichlet forms, $(\widetilde{\nabla} u, Z \widetilde{\nabla} u) d\nu$ is the energy measure of u associated with the quadratic form $(\mathcal{E}, \mathcal{F})$, that is, if $d\mu_u = (\widetilde{\nabla} u, Z \widetilde{\nabla} u) d\nu$, then μ_u is the unique Radon measure on K which satisfies

$$\int_K f d\mu_u = 2\mathcal{E}(uf, u) - \mathcal{E}(u^2, f)$$

for any $f \in \mathcal{F}$. See [2] for the general theory of Dirichlet forms. Comparing with Theorem 2.6, we notice that μ_u is thoroughly determined by $(\mathcal{E}, \mathcal{F})$ and it does not depends on the measure μ. Note that the measure μ has no role in the definition of $(\mathcal{E}, \mathcal{F})$.

Remark. Kusuoka has given the definition of the Kusuoka measure and the "Riemannian metric" $Z(\cdot)$ and the associated results in the last half of this section not only for the Sierpinski gasket but also for a class of finitely ramified self-similar sets. See [11, 12] for details.

3 Volume doubling property of the Kusuoka measure

In this section, we will show that the Kusuoka measure has the volume doubling property with respect to the Euclidean distance.

Definition 3.1. Let (X, d) be a metric space. Define $B_r(x, d) = \{y | y \in X, d(x, y) < r\}$ for any $x \in X$ and any $r > 0$. A Borel measure μ on X is said to have the volume doubling property with respect to the distance d if and only if there exists $c > 0$ such that $\mu(B_{2r}(x, d)) \leq c\mu(B_r(x, d))$ for any $r > 0$ and any $x \in X$.

Theorem 3.2. *The Kusuoka measure on the* N*-Sierpinski gasket has the volume doubling property with respect to the Euclidean distance.*

The rest of this section is devoted to proving the above theorem.

Lemma 3.3. (1) *Let* $\mathbf{a} = (a_1, \ldots, a_N)$ *be a base of* M_0. *For any* $X \in L(M_0)$, *define*

$$||X||_{\mathrm{b}, \mathbf{a}} = \sum_{i=1}^{N} |Xa_i|.$$

Then $||\cdot||_{\mathrm{b},\mathbf{a}}$ *is a norm on* $L(M_0)$.
(2) *Let* D *be a subset of* M_0. *For any* $X \in L(M_0)$, *define*

$$||X||_{\mathrm{d,D}} = \mathrm{diam}(X(D)) = \sup_{x,y\in D} |Xx - Xy|.$$

If D *has an interior point with respect to the topology of* M_0, *then* $||X||_{\mathrm{d,D}}$ *is a norm on* $L(M_0)$.

Lemma 3.4. *For* $x \in M_0$, *define* $B_r(x) = \{y | y \in M_0, |x-y| < r\}$ *for* $r > 0$. *Let* D_1 *and* D_2 *be subsets of* M_0. *If there exist* $x_1, x_2 \in M_0$ *and positive real numbers* r *and* R *such that* $D_1 \subseteq B_R(x_1)$ *and* $B_r(x_2) \subseteq D_2$, *then for any* $X \in L(M_0)$,

$$r||X||_{\mathrm{d,D_1}} \leq R||X||_{\mathrm{d,D_2}}.$$

Lemma 3.5. *There exists* $c > 0$ *such that*

$$\nu(K_{wi(j)^n}) \leq c\nu(K_{wj(i)^n})$$

for any i, j *with* $1 \leq i < j \leq N$, *any* $w \in W_*$ *and any* $n \geq 0$.

Proof. We may assume that $i = 1$ and $j = 2$ without loss of generality. Let $\mathbf{f}_i = (p_i, f_{i,1}, \ldots, f_{i,N-2})$. We write $||\cdot||_i = ||\cdot||_{\mathrm{b},\mathbf{f}_i}$. Then

$$||T_{wi(j)^n}||_j = \Big(\frac{N}{N+2}\Big)^n |T_w T_i p_j| + \Big(\frac{1}{N+2}\Big)^n \sum_{k=1}^{N-2} |T_w T_i f_{j,k}|.$$

Note that $T_1 p_2 = -T_2 p_1$. Let $x_* = T_1 p_2$. If $g(t) = (\alpha + \beta t)(\gamma + \delta t)^{-1}$ for $\alpha, \beta, \gamma, \delta > 0$, then $\sup_{t\in[0,1]} g(t) = \max\{g(0), g(1)\}$. Hence we see that

$$\frac{||T_{w1(2)^n}||_2}{||T_{w2(1)^n}||_1} = \frac{|T_w x_*| + N^{-n}\sum_{k=1}^{N-2} |T_w T_1 f_{2,k}|}{|T_w x_*| + N^{-n}\sum_{k=1}^{N-2} |T_w T_2 f_{1,k}|} \leq \max\Big\{1, \frac{||T_{w1}||_2}{||T_{w2}||_1}\Big\}. \tag{3.1}$$

Now let J be the convex hull of $\{p_1, \ldots, p_N\}$. (Note that J is an N-simplex.) Since $T_1 J$ and $T_2 J$ is congruent (i.e. $T_1 J = f(T_2 J)$ for some isometry f of M_0), Lemma 3.4 implies that

$$\frac{||T_{w1}||_{\mathrm{d,J}}}{||T_{w2}||_{\mathrm{d,J}}} = \frac{||T_w||_{\mathrm{d,T_1J}}}{||T_w||_{\mathrm{d,T_2J}}} \tag{3.2}$$

is uniformly bounded with respect to $w \in W_*$. Since $||\cdot||_{\mathrm{HS}}$, $||\cdot||_{\mathrm{b},\cdot}$ and $||\cdot||_{\mathrm{d},\cdot}$ are all equivalent, (3.1) and (3.2) implies that

$$\frac{||T_{w1(2)^n}||_{\mathrm{HS}}}{||T_{w2(1)^n}||_{\mathrm{HS}}}$$

is uniformly bounded with respect to $w \in W_*$ and $n \geq 0$. By the definition of the Kusuoka measure, this immediately deduce the lemma. □

For the moment, we briefly review the notions and results in [5], which are needed to complete the proof of Theorem 3.2. The notion of a gauge function gives the sizes of K_w's and the scale associated with a gauge function is the collection of K_w's whose sizes are about the same.

Definition 3.6. (1) A function $g : W_* \to (0, 1]$ is called a gauge function on Σ if and only if it satisfies the following two conditions (G1) and (G2):
(G1) $g(wi) \leq g(w)$ for any $i \in S$ and any $w \in W_*$.
(G2) $\max\{g(w)|w \in W_m\} \to 0$ as $m \to \infty$.
(2) Let $g : W_* \to (0, 1]$ be a gauge function. Define $\Lambda_s(g)$ for $s \in (0, 1]$ by

$$\Lambda_s(g) = \{w|w = w_1 \ldots w_m \in W_*, g(w_1 \ldots w_{m-1}) > s \geq g(w)\},$$

where $g(w_1 \ldots w_{m-1})$ is regarded as 2 for $w = \emptyset$. Then $\{\Lambda_s(g)\}_{0<s\leq 1}$ is called a scale induced by the gauge function g.
(3) Let $\mathcal{S} = \{\Lambda_s(g)\}_{0<s\leq 1}$ be the scale induced by a gauge function g. Define

$$K_s(x) = \bigcup_{w\in\Lambda_s(g), x\in K_w} K_w \quad \text{and} \quad U_s(x) = \bigcup_{w\in\Lambda_s(g), K_w\cap K_s(x)\neq\emptyset} K_w.$$

$\mathcal{S}$ is called locally finite if

$$\sup_{(w,s)\in W_*\times(0,1]} \#\{v|v \in \Lambda_s(g), K_v \cap K_w \neq \emptyset\} < +\infty.$$

(4) A distance $d(\cdot, \cdot)$ on K is said to be adapted to the scale $\{\Lambda_s(g)\}$ induced by a gauge function g if and only if there exist positive constants c_1 and c_2 such that

$$B_{c_1 r}(x, d) \subseteq U_r(x) \subseteq B_{c_2 r}(x, d)$$

for any $x \in K$ and any $r \in (0, 1]$.

Remark. In [5], the definition of "adaptedness" is a little more general than the version in here.

The set $U_s(x)$ is a kind of "ball of radius s" associated with the scale. It can be regarded as a real ball if there exists some distance which is adapted to the scale.

Definition 3.7. (1) A function $f : W_* \to (0, \infty)$ is called elliptic if and only if it satisfies the following two conditions (EL1) and (EL2);
(EL1) There exists $c_1 > 0$ such that $f(wi) \geq c_1 f(w)$ for any $w \in W_*$ and $i \in S$.
(EL2) There exist $k \in \mathbb{N}$ and $c_2 \in (0, 1)$ such that $f(wv) \leq c_2 f(w)$ for any $w \in W_*$ and any $v \in W_k$.
The scale induces by a gauge function g is called elliptic if and only if g is elliptic. Also a Borel regular probability measure μ on K is called elliptic if and only if the map $w \to \mu(K_w)$ is elliptic.
(2) Let $\mathcal{S}$ be the scale induced by a gauge function g. A function of $f : W_* \to (0, \infty)$ is said to be gentle with respect to $\mathcal{S}$ if and only if there exists $c_G > 0$ such that $f(w) \leq c_G f(v)$ for any $s \in (0, 1]$ and any $w, v \in \Lambda_s(g)$ with $K_w \cap K_v \neq \emptyset$. A Borel regular probability measure μ on K is said to be gentle with respect to $\mathcal{S}$ if and only if the map $w \to \mu(K_w)$ is gentle with respect to $\mathcal{S}$.

The following theorem is what we need to prove Theorem 3.2. See [5, Section 1.3] for the proof.

Theorem 3.8. *Let μ be a Borel regular probability measure on K. Also let $\mathcal{S}$ be the scale induced by an elliptic gauge function g. Assume that a distance d is adapted to the scale $\mathcal{S}$. Then μ has the volume doubling property with respect to d if and only if the following three conditions* (EL), (LF) *and* (GE) *are satisfied;*
(EL) *μ is elliptic.*
(LF) *The scale $\mathcal{S}$ is locally finite.*
(GE) *μ is gentle with respect to the scale $\mathcal{S}$.*

Proof of Theorem 3.2. Let $\mathcal{S}_0$ be the scale induced by the gauge function $g(w) = (1/2)^{|w|}$, which is obviously elliptic. Then the Euclidean metric is adapted to the scale $\mathcal{S}_0$. By Theorem 3.8, it is sufficient to prove that $\mathcal{S}_0$ is locally finite, ν is elliptic and ν is gentle with respect to $\mathcal{S}_0$.

Since the N-Sierpinski gasket is post critically finite, $\mathcal{S}_0$ is locally finite. Next we show that ν is elliptic. Let J be the convex hull of $\{p_1, \ldots, p_N\}$. Then, by Lemma 3.4,

$$\frac{||T_w||_{\mathrm{d,J}}}{||T_{wi}||_{\mathrm{d,J}}} = \frac{||T_w||_{\mathrm{d,J}}}{||T_w||_{\mathrm{d,T_i(J)}}}$$

is uniformly bounded with respect to $i = 1, \ldots, N$ and $w \in W_*$. Since $||\cdot||_{\mathrm{d},\cdot}$ and $||\cdot||_{\mathrm{HS}}$ are equivalent, there exists $c_1 > 0$ such that

$$||T_{wi}||_{\mathrm{HS}} \geq c_1 ||T_w||_{\mathrm{HS}} \tag{3.3}$$

for any $i = 1, \ldots, N$ and any $w \in W_*$. This shows (EL1). Let $||X||$ be the operator norm of $X \in L(M_0)$, i.e. $||X|| = \sup_{x \in X, x \neq 0} |Xx|/|x|$. Note that $||T_i|| = N/(N+2)$. Hence if $w, v \in W_*$, then $||T_{wv}|| \leq ||T_w||(N/(N+2))^{|v|}$. Since $||\cdot||$ and $||\cdot||_{\mathrm{HS}}$ are equivalent, there exists $c > 0$ such that

$$|||T_{wv}||_{\mathrm{HS}} \leq c\left(\frac{N}{N+2}\right)^{|v|} ||T_w||_{\mathrm{HS}}. \tag{3.4}$$

This implies (EL2). Thus ν is elliptic.

Finally we prove that ν is gentle with respect to $\mathcal{S}_0$. Note that $\Lambda_s = W_m$ for some m. Hence if $w(1), w(2) \in \Lambda_s, w(1) \neq w(2)$ and $K_{w(1)} \cap K_{w(2)} \neq \emptyset$, then there exist $i, j \in \{1, \ldots, N\}$, $n \geq 0$ and $w \in W_*$ such that $w(1) = wi(j)^n$ and $w(2) = wj(i)^n$. By Lemma 3.5, we obtain that $\nu(K_{w(1)}) \leq c\nu(K_{w(2)})$, where c is independent of $w(1)$ and $w(2)$. Therefore, ν is gentle with respect to $\mathcal{S}_0$. □

4 Harmonic Sierpinski gasket

In this section, we introduce the harmonic Sierpinski gasket, which is the Sierpinski gasket parametrized by harmonic functions. The harmonic Sierpinski gasket turns out to be a geometrical realization of the measurable Riemannian structure given in Section 2.

First we give the definition of harmonic functions. See [9] for the proofs of the following propositions 4.1 and 4.3.

Proposition 4.1. *Let $\rho : V_0 \to \mathbb{R}$. Then there exists a unique $\psi : K \to V_0$ such that $\psi|_{V_0} = \rho$ and*

$$\mathcal{E}(\psi, \psi) = \min\{\mathcal{E}(u, u) | u \in \mathcal{F}, u|_{V_0} = \rho\}.$$

Definition 4.2. We denote ψ in the above proposition by ψ_ρ and call it the harmonic function with boundary value ρ. In particular, define $\psi_p = \psi_{\chi_p}$ for $p \in V_0$, where

$$\chi_p(x) = \begin{cases} 1 & \text{if } x = p, \\ 0 & \text{otherwise.} \end{cases}$$

Proposition 4.3. *For any $\rho : V_0 \to \mathbb{R}$, $\psi_\rho = \sum_{p \in V_0} \rho(p)\psi_p$. In particular, the collection of harmonic functions is N-dimensional vector space.*

The harmonic functions $\{\psi_p\}_{p \in V_0}$ can be a system of coordinates of the Sierpinski gasket. See [8] for the proof of the next proposition.

Proposition 4.4. *Define $\Phi : K \to M_0$ by*

$$\Phi(x) = \frac{1}{\sqrt{2}}\left(\begin{pmatrix} \psi_{p_1}(x) \\ \vdots \\ \psi_{p_N}(x) \end{pmatrix} - \frac{1}{N}\begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \right).$$

If $K_H = \Phi(K)$, then Φ is a homeomorphism between K and K_H. Moreover, define $H_i : M_0 \to M_0$ by

$$H_i(x) = T_i(x - p_i) + p_i$$

for $i = 1, \ldots, N$. Then $K_H = \cup_{i=1}^N H_i(K_H)$ and $\Phi \circ F_i = H_i \circ \Phi$ for any $i = 1, \ldots, N$.

Definition 4.5. K_H is called the harmonic N-Sierpinski gasket.

By Proposition4.4, K_H is the self-similar set associated with the collection of contractions on M_0, $\{H_1, \ldots, H_N\}$.

Now we define the gradient operator ∇ on the Sierpinski gasket as the usual gradient on the harmonic Sierpinski gasket through Φ. Note that we can naturally define the gradient (i.e. ∇) of a smooth function on an open subset of M_0. More precisely, if we fix an orthonormal base of M_0 and regard M_0 as $\mathbb{R}^{N-1}$, then $\nabla u = {}^t(\partial u/\partial x_1, \ldots, \partial u/\partial x_{N-1})$. The following proposition has been shown in [8].

Proposition 4.6. *Let U be an open subset of M_0 which contains K_H. If $v_1, v_2 \in C^1(U)$ and $v_1|_{K_H} = v_2|_{K_H}$, then $(\nabla v_1)|_{K_H} = (\nabla v_2)|_{K_H}$.*

By the above proposition, the gradient (i.e. ∇) of a "smooth function" on K_H is well-defined by the restriction of the gradient on an open subset of M_0. Through Φ, we define C^1-functions on K and the gradient (i.e. ∇) of them.

Definition 4.7. Define

$$C^1(K) = \{u | u = (v|_{K_H}) \circ \Phi, \text{where } v \text{ is a } C^1\text{-class function} \\ \text{on some open subset of } M_0 \text{ containing } K_H.\}$$

Also, for $u \in C^1(K)$, define $\nabla u = (\nabla v|_{K_H}) \circ \Phi$, where v is the same as in the above definition of $C^1(K)$.

In fact, the gradient operator ∇ has been shown to be the natural one from the viewpoint of the "Riemannian structure" as well. The main part of the following theorem has been obtained in [8].

Theorem 4.8. *$C^1(K)$ is a dense subset of $\mathcal{F}$ under the norm $||u|| = \sqrt{\mathcal{E}(u,u)} + ||u||_\infty$. Moreover, $\widetilde{\nabla} u = Z\nabla u$ for any $u \in C^1(K)$ and*

$$\mathcal{E}(u,v) = \int_K (\nabla u, Z\nabla v) d\nu \tag{4.1}$$

for any $u, v \in C^1(K)$.

To give a proof of the above theorem, we translate the definition of the Kusuoka gradient $\widetilde{\nabla}$ into the language of the harmonic Sierpinski gasket.

Definition 4.9. (1) Define $U_0 = \{p_1, \ldots, p_N\}$ and $B_w = H_w(U_0)$ for any $w = w_1 \ldots w_m \in W_*$, where $H_w = H_{w_1} \cdots H_{w_m}$. Also, for any $m \geq 0$, set $U_m = \cup_{w \in W_m} B_w$.
(2) Let $u : K \to \mathbb{R}$. For any $w \in W_*$, $\nabla_w u$ is the gradient of the affine function on M_0 whose values at the points of B_w coincide with those of $u \circ \Phi^{-1}$. Moreover, define $\nabla_m u : K \to \mathbb{R}^{N-1}$ by $(\nabla_m u)|_{K_w} = \nabla_w u$.

Note that $U_0 = V_0$ by the convention after Definition 2.7. We also see $U_m = \Phi(V_m)$. Strictly speaking, $\nabla_m u$ should be thought of as a ν-measurable function since the values on V_m may not be well-defined. The following proposition has been obtained by Kusuoka in [11].

Proposition 4.10. *For any $u \in \mathcal{F}$, $Z\nabla_m u$ converges as $m \to \infty$ for ν-a.e. $x \in K$.*

Now, the definition of the $\widetilde{\nabla}$ appearing in Theorem 2.12 can be given in the following way.

Definition 4.11. For $u \in \mathcal{F}$, the ν-a.e. limit of $Z\nabla_m u$ as $m \to \infty$ is denoted by $\widetilde{\nabla} u$.

Proof of Theorem 4.8. The fact that $C^1(K)$ is dense subset of $\mathcal{F}$ and (4.1) have been shown in [8]. We need to prove $\widetilde{\nabla} u = Z\nabla u$. It is easy to see that $(\nabla u, Z\nabla u)d\nu$ is the energy measure of u with respect to the Dirichlet form as well. Hence we have $(\nabla u, Z\nabla u) = (\widetilde{\nabla} u, Z\widetilde{\nabla} u)$ for ν-a.e. $x \in K$. Recall that rank$Z = 1$ and Z is the orthogonal projection onto its image. Note that $\widetilde{\nabla}u$ is in the image of Z by definition. Combining these facts, we obtain $Z\nabla u = \pm\widetilde{\nabla}u$, where $\pm$ depends on $x \in K$. It follows that $Z\nabla\psi = \widetilde{\nabla}\psi$ for any harmonic function ψ. Choose a harmonic function ψ so that $\psi \neq 0$. Let $U = \{x | x \in K, Z\nabla u \neq \widetilde{\nabla}u\}$. Then on U, we have $Z\nabla(u+\psi) = Z\nabla u + Z\nabla\psi = -\widetilde{\nabla}u + Z\nabla\psi$ on U. On the other hand, $Z\nabla(u+\psi) = \pm\widetilde{\nabla}(u+\psi)$ on U. If the plus sign holds, then $Z\nabla u = 0$, which contradicts to the definition of U. Hence $Z\nabla(u+\psi) = -\widetilde{\nabla}(u+\psi)$ on U. This implies $\psi = -\psi$ and this is also a contradiction. Therefore $\nu(U) = 0$ and $Z\nabla u = \widetilde{\nabla}u$ for ν-a.e. $x \in K$. □

5 Geodesic distance on the Sierpinski gasket

In the previous sections, we have introduced "Riemannian metric" Z, "Riemannian volume" ν, the "Riemannian energy" $(\mathcal{E}, \mathcal{F})$ and the gradient operator ∇ on the Sierpinski gasket and have obtained, in Theorem 4.8, the relation between these objects which is analogous to that in the Riemannian geometry. There have still been one missing element, namely, the geodesic distance associated with "Riemannian structure". In this section, we will introduce the harmonic shortest path metric on the Sierpinski gasket, which is the counterpart of the geodesic distance.

The next theorem shows that there exists a geodesic distance on K_H.

Theorem 5.1. *Define*

$$h_*(p,q) = \inf\{\ell(\gamma) | \gamma \text{ is a rectifiable curve in } K_H \text{ between } p \text{ and } q\}$$

for $p, q \in K_H$, where $\ell(\gamma)$ is the length of the curve. Then for any $p, q \in K_H$, there exists a continuous curve $\gamma_* : [0,1] \to K_H$ such that γ_* is C^1 on $(0,1)$, $\gamma_*(0) = p, \gamma_*(1) = q$, $Z(\Phi^{-1}(\gamma_*(t)))$ exists and $\frac{d\gamma_*}{dt} \in \operatorname{Im} Z(\Phi^{-1}(\gamma_*(t)))$ for any $t \in (0,1)$ and

$$h_*(\gamma_*(a), \gamma_*(b)) = \int_a^b \left(\frac{d\gamma_*}{dt}, Z(\Phi^{-1}(\gamma_*(t)) \frac{d\gamma_*}{dt} \right) dt \tag{5.1}$$

for any $a, b \in [0,1]$ with $a \leq b$. In particular, the infimum in the definition of $h_*(p,q)$ can be replaced by the minimum, which is attained by γ_*. γ_* is called a geodesic between p and q.

(5.1) connects $h_*(\cdot,\cdot)$ with the measurable Riemannian structure.

If $N = 2$, then $K_H = K$ is an interval and the above theorem is trivial. We will give a proof of this theorem for $N = 3$. One may show the case of $N > 3$ with the same ideas and more complicated arguments.

We need several steps to show Theorem 5.1. The following result is a well-known fact in the convex geometry.

Theorem 5.2. *Let C and D be compact subsets of $\mathbb{R}^2$ with $C \subseteq D$. Assume that C is convex and that ∂D is a rectifiable Jordan curve. Then $\ell(\partial C) \leq \ell(\partial D)$.*

Remark. In the above theorem, since C is a compact convex set, its boundary ∂C is a rectifiable Jordan curve.

Definition 5.3. (1) Define $J_w = H_w(J)$, where J is the convex hull of U_0, i.e. the regular triangle with vertices U_0. Also define $K_{H,w} = H_w(K_H)$.
(2) Let $p, q \in B_w$ with $p \neq q$. We define $\widehat{pq}$ by $\widehat{pq} = \Phi(\overline{p_* q_*})$, where $p_* = \Phi^{-1}(p)$ and $q_* = \Phi^{-1}(q)$ and $\overline{xy}$ is the line segment between x and y.

We now study the curve $\widehat{p_1p_2}$, which can be regarded as a graph of a function defined on $\overline{p_1p_2}$. To see this, we introduce a new coordinate system on M_0 which identify $\overline{p_1p_2}$ as $[0,1]$. Set $\mathbf{e}_1 = p_2 - p_1$ and $\mathbf{e}_2 = (p_3 - (p_1+p_2)/2))/(\sqrt{3}/2)$. Note that $(\mathbf{e}_1, \mathbf{e}_2)$ is an orthonormal base of $\mathbb{R}^2$. Associating $(s,t) \in \mathbb{R}^2$ to the point $p_1 + s\mathbf{e}_1 + t\mathbf{e}_2 \in M_0$, we have a new coordinate (s,t) of M_0. More precisely, $((p-p_1, e_1), (p-p_1, e_2)) \in \mathbb{R}^2$ is the coordinate of $p \in M_0$. If we identify $\overline{p_1p_2}$ as $[0,1]$ in the natural manner, $\widehat{p_1p_2}$ may be regarded as a graph of a function defined on $[0,1]$ as follows.

Theorem 5.4. (1) *$Z(x)$ exists for any $x \in \overline{p_1p_2}$.*
(2) *Define $\theta : \widehat{p_1p_2} \to [0,1]$ by $\theta(p) = (p - p_1, e_1)$. Then θ is bijective. Denote the inverse of θ by γ and define $\eta(t) = (\gamma(t) - p_1, e_2)$. Then $\eta : [0,1] \to [0,+\infty)$ is concave, C^1 but not C^2. Moreover, $\frac{d\gamma}{dt} = \begin{pmatrix} 1 \\ \frac{d\eta}{dt} \end{pmatrix} \in \operatorname{Im} Z(\Phi^{-1}(\gamma(t)))$ for any $t \in [0,1]$ and*

$$\ell(\widehat{p_1p_2}) = \int_0^1 \left(\frac{d\gamma}{dt}, Z(\Phi^{-1}(\gamma(t))) \frac{d\gamma}{dt} \right) dt. \tag{5.2}$$

Remark. The results in Theorem 5.4 have been announced by Teplyaev in [18]. Since a preprint with a proof, which was cited in [18], is not available at this moment, we will give our version of proof for the sake of readers.

Proof. In the new coordinates,

$$T_1 = \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{10} \\ \frac{\sqrt{3}}{10} & \frac{3}{10} \end{pmatrix} \quad \text{and} \quad T_2 = \begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{10} \\ -\frac{\sqrt{3}}{10} & \frac{3}{10} \end{pmatrix}.$$

Let

$$\mathcal{A} = \{f | f \in C^1([0,1]), f(0) = f(1) = 0, \\ f'(0) = \frac{1}{\sqrt{3}}, f'(1) = -\frac{1}{\sqrt{3}}, f'(t_1) \geq f'(t_2) \text{ if } t_1 \leq t_2.\}$$

Define $d(f,g) = \sup_{0 \leq t \leq 1} |f(t) - g(t)| + \sup_{0 \leq t \leq 1} |f'(t) - g'(t)|$ for any $f, g \in \mathcal{A}$. Then $(\mathcal{A}, d)$ is a complete metric space. Now, for any $f \in \mathcal{A}$,

$$H_1 \begin{pmatrix} t \\ f(t) \end{pmatrix} = \begin{pmatrix} \frac{1}{2}t + \frac{\sqrt{3}}{10} f(t) \\ \frac{\sqrt{3}}{10} t + \frac{3}{10} f(t) \end{pmatrix} \quad \text{and} \quad H_2 \begin{pmatrix} t \\ f(t) \end{pmatrix} = \begin{pmatrix} \frac{1}{2}t - \frac{\sqrt{3}}{10} f(t) + \frac{1}{2} \\ -\frac{\sqrt{3}}{10} t + \frac{3}{10} f(t) + \frac{\sqrt{3}}{10} \end{pmatrix}.$$

Set $h(t) = \frac{1}{2}t + \frac{\sqrt{3}}{10}f(t)$. Then $h(0) = 0, h(1) = \frac{1}{2}$ and $h'(t) = \frac{1}{2} + \frac{\sqrt{3}}{10}f'(t) \geq \frac{2}{5}$. Therefore, by the inverse function theorem, the inverse of h exists and belongs to $C^1([0, \frac{1}{2}])$. Define $F(s) = \frac{\sqrt{3}}{10}h^{-1}(s) + \frac{3}{10}f(h^{-1}(s))$. Then $F \in C^1([0, \frac{1}{2}])$. In the same manner, we may define $F(s)$ for $s \in [\frac{1}{2}, 1]$ by $H_2 \begin{pmatrix} t \\ f(t) \end{pmatrix} = \begin{pmatrix} s \\ F(s) \end{pmatrix}$. It follows that

$$F'(s) = \sqrt{3}\Big(1 - \frac{4}{5 + \sqrt{3}f'(h^{-1}(s))}\Big) \tag{5.3}$$

for $s \in [0, \frac{1}{2}]$. Using (5.3) and its counterpart on $s \in [\frac{1}{2}, 1]$, we easily see that $F \in \mathcal{A}$. Thus if $\mathcal{G}(f) = F$, then $\mathcal{G}$ is a well-defined map from $\mathcal{A}$ to itself. Moreover, it follows that $d(\mathcal{G}(f_1), \mathcal{G}(f_2)) \leq \frac{3}{4}d(f_1, f_2)$. By the contraction principle, there exists a unique fixed point of $\mathcal{G}$, which is denoted by η. Let $G_\eta = \{(t, \eta(t)) | t \in [0, 1]\}$. Then $G_\eta = H_1(G_\eta) \cup H_2(G_\eta)$ and hence $G_\eta = \widehat{p_1p_2}$. Obviously, $\eta \in C^1([0, 1])$ and η is concave. Assume that $\eta \in C^2([0, 1]) \cap \mathcal{A}$. Then

$$\eta''(h_1(t)) = \frac{120\eta''(t)}{\big(5 + \sqrt{3}\eta'(t)\big)^3} \quad \text{and} \quad \eta''(h_2(t)) = \frac{120\eta''(t)}{\big(5 - \sqrt{3}\eta'(t)\big)^3} \tag{5.4}$$

for any $t \in [0, 1]$, where $h_1(t) = \frac{1}{2}t + \frac{\sqrt{3}}{10}\eta(t)$ and $h_2(t) = \frac{1}{2}t - \frac{\sqrt{3}}{10}\eta(t) - \frac{1}{2}$. Letting $t = 0$ and 1, we see that $\eta''(0) = \eta''(1) = \eta''(1/2) = 0$. Also by (5.4), if $\eta''(t) = 0$, then $\eta''(h_1(t)) = \eta''(h_2(t)) = 0$. Set $h_w = h_{w_1} \circ \cdots \circ h_{w_m}$ for $w = w_1 \ldots w_m \in W_*(\{1, 2\})$, where $W_*(\{1, 2\}) = \cup_{k \geq 0}\{1, 2\}^k$. Then it follows that $\eta''(h_w(0)) = 0$ for any $w \in W_*(\{1, 2\})$. Note that $T_w 0 = \begin{pmatrix} h_w(0) \\ \eta(h_w(0)) \end{pmatrix}$ and that $\{T_w 0\}_{w \in W_*(\{1,2\})}$ is dense in $\widehat{p_1p_2}$. Hence $\{h_w(0)\}_{w \in W_*(\{1,2\})}$ is dense in $[0, 1]$. Therefore, $\eta''(t) = 0$ for any $t \in [0, 1]$. Since $\eta(0) = \eta(1) = 0$, we have $\eta(t) = 0$ for any $t \in [0, 1]$. This contradiction implies that $\eta \notin C^2([0, 1])$.

Next, let $\omega = \omega_1\omega_2 \ldots \in \{1, 2\}^{\mathbb{N}}$. Set $\omega(m) = \omega_1 \ldots \omega_m \in \{1, 2\}^m$ and $T_{\omega(m)} = \begin{pmatrix} a_m & b_m \\ c_m & d_m \end{pmatrix}$. Then,

$$H_{\omega(m)} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a_m & b_m \\ c_m & d_m \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} h_{\omega(m)}(0) \\ \eta(h_{\omega(m)}(0)) \end{pmatrix}$$

and $h_{\omega(m)}(t) = a_m t + b_m\eta(t) + h_{\omega(m)}(0)$. Since $h_1'(t)$ and $h_2'(t)$ are no less than $\frac{2}{5}$ for any $t \in [0, 1]$, we see that $h_{\omega(m)}'(t) = a_m + b_m\eta'(t) \geq (\frac{2}{5})^m$. In particular, $h_{\omega(m)}(\frac{1}{2}) = a_m > 0$ because $\eta'(\frac{1}{2}) = 0$. Moreover, it follows that $a_m + b_m s > 0$ for any $s \in [-\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}]$. Hence

$$|b_m/a_m| \leq \sqrt{3} \tag{5.5}$$

for any m. On the other hand, set $B_m = b_m/a_m, C_m = c_m/a_m$ and $D_m = d_m/a_m$. Then $\eta'(h_{\omega(m)}(t)) = \frac{C_m + D_m\eta'(t)}{1 + B_m\eta'(t)}$. Define $\tau(\omega) = \lim_{m\to\infty} h_{\omega(m)}(0)$.

Note that $\tau(\omega) = \lim_{m\to\infty} h_{\omega(m)}(t)$ for any $t \in [0,1]$. Since η' is continuous, it follows that

$$\lim_{m\to\infty} \frac{C_m + D_m\eta'(t)}{1 + B_m\eta'(t)} = \eta'(\tau(\omega))$$

for any $t \in [0,1]$. In particular, $\lim_{m\to\infty} C_m = \eta'(\tau(w))$. Hence,

$$\frac{C_m + D_m\eta'(t)}{1 + B_m\eta'(t)} - C_m = \eta'(t)\frac{D_m - C_mB_m}{1 + B_m\eta'(t)} \to 0$$

as $m \to \infty$. Choosing t so that $\eta'(t) = (2\sqrt{3})^{-1}$ and using (5.5), we have

$$\lim_{m\to\infty} D_m - C_mB_m = 0. \tag{5.6}$$

Now, let $Z_m = Z_m(\omega(m))$. Then

$$Z_m = \frac{1}{1 + B_m^2 + C_m^2 + D_m^2}\begin{pmatrix} 1 + B_m^2 & C_m + B_mD_m \\ C_m + B_mD_m & C_m^2 + D_m^2 \end{pmatrix}.$$

By (5.6),

$$\begin{aligned}\frac{1}{1 + \eta'(\tau(\omega))^2}\begin{pmatrix} 1 & \eta'(\tau(\omega)) \\ \eta'(\tau(\omega)) & \eta'(\tau(\omega))^2 \end{pmatrix} &= \lim_{m\to\infty} \frac{1}{1 + C_m^2}\begin{pmatrix} 1 & C_m \\ C_m & C_m^2 \end{pmatrix} \\ &= \lim_{m\to\infty} \frac{1 + B_m^2}{1 + B_m^2 + C_m^2 + D_m^2}\begin{pmatrix} 1 & C_m \\ C_m & C_m^2 \end{pmatrix} = \lim_{m\to\infty} Z_m\end{aligned}$$

Thus $Z(\Phi^{-1}(\gamma(\tau(\omega))))$ exists and $\begin{pmatrix} 1 \\ \eta'(\tau(\omega)) \end{pmatrix} \in \operatorname{Im} Z(\Phi^{-1}(\gamma(\tau(\omega))))$. Since there exists $\omega \in \{1,2\}^{\mathbb{N}}$ such that $t = \tau(\omega)$ for any $t \in [0,1]$, we have shown (1) and $\gamma'(t) \in \operatorname{Im} Z(\Phi^{-1}(\gamma(t)))$ for any $t \in [0,1]$. Now, (5.2) is obvious. $\square$

Lemma 5.5. *Let $p, q \in B_w$ with $p \neq q$. Then $\overline{pq} \cup \widehat{pq}$ is a closed curve. Moreover, if D_{pq} is the bounded domain whose boundary is $\overline{pq} \cup \widehat{pq}$, then D_{pq} is convex. In particular, $\widehat{pq}$ is rectifiable.*

Proof. Note that η in Theorem 5.4 is convex. Hence $D_{p_1p_2}$ is convex. Since $D_{p,q}$ is an affine image of D_{p_1,p_2}, we immediately verify the lemma. $\square$

Lemma 5.6. *Let $p, q \in B_w$ with $p \neq q$. Then the following minimum is attained by $\widehat{pq}$.*

$$\min\{\ell(\gamma) | \gamma \textit{ is a rectifiable curve between } p \textit{ and } q \textit{ and is contained in } K_{H,w}.\}$$

Moreover,

$$\frac{2}{5}||T_w||_{\mathrm{d,J}} \leq \ell(\widehat{pq}) \leq 2||T_w||_{\mathrm{d,J}} \tag{5.7}$$

Proof. Let γ be a rectifiable curve between p and q which is contained in $K_{H,w}$. We may assume that γ does not have self-intersection without loss of generality. Let D be the bounded domain whose boundary is $\overline{pq} \cup \gamma$. Then $D_{pq} \subseteq D$. Hence by Theorem 5.2, it follows that $\ell(\overline{pq} \cup \widehat{pq}) \leq \ell(\overline{pq} \cup \gamma)$. Therefore $\ell(\widehat{pq}) \leq \ell(\gamma)$.

Note that J is a regular triangle with vertices $U_0 = \{p_1, p_2, p_3\}$. Suppose that $p = H_w(p_1)$ and $q = H_w(p_2)$. Let $x = H_w(p_3)$. Then by the triangle inequality,

$$||T_w||_{\mathrm{d,J}} = \mathrm{diam}(J_w) \leq \ell(\overline{px}) + \ell(\overline{xq}) \leq 2||T_w||_{\mathrm{d,J}}. \tag{5.8}$$

Since $D_{pq} \subset J_w$, Theorem 5.2 implies that $\ell(\widehat{pq}) \leq \ell(\overline{px}) + \ell(\overline{xq})$. Hence by (5.8), $\ell(\widehat{pq}) \leq 2||T_w||_{\mathrm{d,J}}$. Let y' be the midpoint of $\Phi^{-1}(p)$ and $\Phi^{-1}(q)$ and let $y = \Phi(y)$. Then $y \in \widehat{pq}$ and the line xy passes the center of the triangle J. If z is the intersection of the lines xy and pq, then $|zy| = 2|zx|/5$. Hence $\ell(\overline{py}) + \ell(\overline{yq}) = 2(\ell(\overline{px}) + \ell(\overline{xq}))/5$. On the other hand, the triangle pyq is contained in D_{pq}. Again by Theorem 5.2, $\ell(\overline{py}) + \ell(\overline{yq}) \leq \ell(\widehat{pq})$. Using (5.8), we obtain (5.7). $\square$

Proof of Theorem 5.1. First assume that $p, q \in U_m$ for some $m \geq 0$. Define

$$\begin{aligned} P_m(p,q) = \{(q_i)_{i=0}^k | q_0 = p, q_k = q, &\text{there exists } w \in W_m \text{ such that} \\ &p_i, p_{i+1} \in B_w \text{ for any } i = 0, \ldots, k-1\}. \end{aligned}$$

Also set $\ell((q_i)_{i=0}^k) = \sum_{i=0}^{k-1} \ell(\widehat{q_i q_{i+1}})$. Since $\{(q_i)_{i=0}^k \in P_m(p,q) | q_i \neq q_j \text{ if } i \neq j\}$ is a finite set, there exists $(x_i)_{i=0}^n \in P_m(p,q)$ that attains the minimum of $\ell((q_i)_{i=0}^k)$. We use γ_* to denote the rectifiable curve $\widehat{x_0 x_1} \cup \ldots \cup \widehat{x_{n-1} x_n}$.

Now let γ be a rectifiable curve in K_H between p and q. If $\gamma \cap U_m = \{y_0, \ldots, y_l\}$, then $(y_i)_{i=0}^l \in P_m(p,q)$. Using Lemma 5.6, we see that $\ell(\gamma) \geq \ell((y_i)_{i=0}^l)$. Hence $\ell(\gamma) \geq \ell(\gamma_*)$. Therefore, γ_* attains the minimum of $\ell(\gamma)$ among rectifiable curves in K_H between p and q. By a suitable parametrization, we verify that $\gamma_* : [0, h_*(p,q)] \to K_H$ and $h_*(\gamma_*(s), \gamma_*(t)) = |s - t|$ for any $s, t \in [0, h_*(p,q)]$. Thus there exists a geodesic between p and q if $p, q \in U_m$.

Next assume that $p \in U_0$ and $q \notin \cup_{m \geq 1} U_m$. Then there exists $\omega = \omega_1 \omega_2 \ldots \in \Sigma$ such that $q = \cap_{m \geq 1} K_{H,\omega_1 \ldots \omega_m}$. Making use of induction, we obtain $q_1, q_2, q_3, \ldots$ satisfying that $q_m \in B_{\omega_1 \ldots \omega_m}$, $\gamma_{m+1}|_{[0, h_*(p,q_m)]} = \gamma_m$ for any $m \geq 0$, where γ_k is a geodesic between p and q_k. Set $T = \lim_{m \to \infty} h_*(p, q_m)$. By Lemma 5.6, $h_*(q_m, q_{m+1}) \leq 2||T_w||_{\mathrm{d,J}}$ for some $w \in W_{m+1}$. Since the maximum eigenvalue of T_i is $3/5$, we see that $||T_w||_{\mathrm{d,J}} \leq c(3/5)^m$, where c is independent of $w \in W_*$. Thus we see that $h_*(q_m, q_{m+1}) \leq 2c(3/5)^m$. Hence T is finite. Define $\gamma_* : [0, T]$ by letting $\gamma_*|_{[0, h_*(p,q_m)]} = \gamma_m$ and $\gamma_*(T) = q$. Then γ_* is the geodesic between p and q, and $T = h_*(p,q)$. By the similar arguments, we can construct a geodesic between p and q for any $p, q \in K_H$.

By the above construction, a geodesic between p and q is a combination of $\{\widehat{p_i p_{i+1}}\}$ where $\{p_i\} \in \cup_{m \geq 0} U_m$ and $p_i, p_{i+1} \in B_w$ for some w. Therefore, by Theorem 5.4, we have (5.1). $\square$

Definition 5.7. Define $d_*(x, y) = h_*(\Phi(x), \Phi(y))$ for any $x, y \in K$. $d_*(\cdot, \cdot)$ is called the harmonic shortest path metric on the N-Sierpinski gasket.

Proposition 5.8. *Define $l(w) = ||T_w||_{\mathrm{HS}}$ for any $w \in W_*$. Then l is a gauge function and is elliptic.*

Proof. By Lemma 2.9, we have $\sum_{i=1}^{N} ||T_{wi}||_{\mathrm{HS}}^2 = \frac{N}{N+2}||T_w||_{\mathrm{HS}}^2$. Hence we have (G1). (G2) is immediately deduced from (3.4). Also (3.3) and(3.4) imply for l being elliptic. $\square$

Definition 5.9. We use $\mathcal{S}_*$ to denote the scale induced by the gauge function $l(\cdot)$.

By the last proposition, $\mathcal{S}_*$ is elliptic. Next, we define a (pseudo)distance associated with the scale S_*.

Definition 5.10. For $x, y \in K$, a sequence of words $(w(1), \ldots, w(m))$ is called a chain between x and y if and only if $x \in K_{w(1)}$, $K_{w(j)} \cap K_{w(j+1)} \neq \emptyset$ for any $j = 1, \ldots, m-1$ and $y \in K_{w(m)}$. Define

$$D_{S_*}(x, y) = \inf\{\sum_{j=1}^{m} ||T_{w(j)}||_{\mathrm{HS}} | (w(1), \ldots, w(m)) \text{ is a chain between } x \text{ and } y\}$$

for any $x, y \in K$.

Obviously, D_{S_*} is a pseudo-distance, i.e. $D_{S_*}(x, y) \geq 0$, $D_{S_*}(x, x) = 0$, $D_{S_*}(x, y) = D_{S_*}(y, x)$ and $D_{S_*}(x, z) \leq D_{S_*}(x, y) + D_{S_*}(y, z)$.

The next theorem shows that D_{S_*} and d_* are equivalent.

Theorem 5.11. *There exist positive constants c_1 and c_2 such that*

$$c_1 D_{\mathcal{S}_*}(x, y) \leq d_*(x, y) \leq c_2 D_{\mathcal{S}_*}(x, y) \tag{5.9}$$

for any $x, y \in K$. In particular, $D_{\mathcal{S}_*}(\cdot, \cdot)$ is a distance on K Moreover, $d_*(\cdot, \cdot)$ and $D_{\mathcal{S}_*}(\cdot, \cdot)$ are adapted to the scale $\mathcal{S}_*$.

We write $D_*(\cdot, \cdot) = D_{\mathcal{S}_*}(\cdot, \cdot)$.

Proof. Since $|| \cdot ||_{\mathrm{HS}}$ and $|| \cdot ||_{\mathrm{d,J}}$ is equivalent, (5.7) implies that there exist positive constants c_3 and c_4 such that

$$c_1 ||T_w||_{\mathrm{HS}} \leq \ell(\widehat{pq}) \leq c_2 ||T_w||_{\mathrm{HS}}, \tag{5.10}$$

for any $w \in W_*$ and any $p, q \in B_w$ with $p \neq q$. Let $x, y \in K$. Set $p = \Phi(x)$ and $q = \Phi(y)$. Let γ_* be the geodesic between p and q. First we assume that $p, q \in U_m$ for some $m \geq 0$. Then there exist $w(1), \ldots, w(k) \in W_m$ and $q_0, \ldots, q_k \in U_m$ such that $q_0 = p, q_k = q$, $q_i, q_{i+1} \in B_{w(i+1)}$ for $i = 0, \ldots, k-1$ and $\gamma_* = \widehat{q_0 q_1} \cup \ldots \cup \widehat{q_{k-1} q_k}$. Since $(w(i))_{i=1}^k$ is a chain between x and y, (5.10) implies that $c_1 D_*(x, y) \leq c_1 \sum_{i=1}^k ||T_w||_{\mathrm{HS}} \leq \sum_{i=1}^k \ell(\widehat{q_{i-1} q_i}) = h_*(p, q) = d_*(x, y)$. Next let $(w(i))_{i=1}^k$ be a chain between x and y. We may assume that $K_{w(i)} \cap K_{w(i+1)}$

is a finite set for any i. Then $K_{H,w(i)} \cap K_{H,w(i+1)} = B_{w(i)} \cap B_{w(i+1)}$ for any i. Choose $q_i \in B_{w(i)} \cap B_{w(i+1)}$ for any $i = 1, \ldots, k-1$. Also set $q_0 = p$ and $q_k = q$. Then $h_*(p,q) \leq \sum_{i=1}^{k} \ell(\widehat{q_{i-1}q_i}) \leq c_2 \sum_{i=1}^{k} ||T_{w(i)}||_{\mathrm{HS}}$. This immediately implies that $d_*(x,y) \leq c_2 D_*(x,y)$. Therefore we obtain (5.9) for any $x, y \in V_*$.

For general x and y, we choose ω and τ which satisfy $\pi(\omega) = x$ and $\pi(\tau) = y$. Then there exist $\{x_m\}_{m\geq 1}$ and $\{y_m\}_{m\geq 0}$ such that $x_m \in F_{\omega_1\ldots\omega_m}(V_0)$, and $y_m \in F_{\tau_1\ldots\tau_m}(V_0)$. Using the same arguments as in the proof of Theorem 5.1, we obtain that $d_*(x_m, x_{m+1}) \leq c(3/5)^m$ and $d_*(y_m, y_{m+1}) \leq c(3/5)^m$. Therefore, $d_*(x_m, y_m) \to d(x,y)$ as $m \to \infty$. Also it follows that $D_*(x, x_m) \leq ||T_{\omega_1\ldots\omega_m}||_{\mathrm{HS}}$ and $D_*(y, y_m) \leq ||T_{\tau_1\ldots\tau_m}||_{\mathrm{HS}}$. Hence $D_*(x_m, y_m) \to D_*(x,y)$ as $m \to \infty$. Thus (5.9) holds for any $x, y \in K$.

Now we show that d_* and D_* are adapted to $\mathcal{S}_*$. Let $y \notin U_s(x)$. If γ_* is a geodesic between $p = \Phi(x)$ and $q = \Phi(y)$, then there exist $w \in \Lambda_s$ and $q_1, q_2 \in B_w$ with $q_1 \neq q_2$ such that $\gamma_* \cap K_{H,w} = \widehat{q_1q_2}$. Hence $d_*(x,y) = h_*(p,q) \geq \ell(\widehat{q_1q_2}) \geq c_1||T_w||_{\mathrm{HS}}$. Therefore there exists $c_3 > 0$ such that $d_*(x,y) \geq c_3 s$. This immediately implies that $B_{c_3 s}(x, d_*) \subseteq U_s(x)$. On the other hand, if $y \in U_s(x)$, then there exists $c_4 > 0$ such that $D_*(x,y) \leq c_4 s$. Therefore, $U_s(x) \subseteq B_{c_4 s}(x, D_*)$. By (5.9), it follows that d_* and D_* are adapted to $\mathcal{S}_*$. $\square$

Remark. The harmonic shortest path metric d_* is not equivalent to the so called "harmonic metric" d_H on K introduced in [8], which is defined by $d_H(x,y) = |\Phi(x) - \Phi(y)|$. In fact, let $x_n = (F_1)^n(p_2)$ and let $y_n = (F_1)^n(p_3)$. Then $d_H(x_n, y_n) = |(H_1)^n(p_2 - p_3)|$. Since $p_2 - p_3$ is orthogonal to p_1, we have $d_H(x_n, y_n) = 5^{-n}|p_2 - p_3|$. On the other hand, $||T_{(1)^n}||_{\mathrm{HS}} = \sqrt{(3/5)^{2n} + 5^{-2n}}$. Hence there exist positive constants α_1 and α_2 such that, for any n, $\alpha_1(3/5)^n \leq d_*(x_n, y_n) \leq \alpha_2(3/5)^n$.

6 Gaussian heat kernel estimate

In this section, we will show the Li-Yau type Gaussian estimate of the heat kernel $p_\nu(t,x,y)$ using the geodesic distance d_*. As we mentioned in the introduction, such an estimate holds for the heat kernels on a class of the Riemannian manifolds.

First we establish the continuity of the heat kernel. The next theorem is the repetition of Theorem 2.6. As we mentioned after Theorem 2.6, we only need to prove the existence of jointly continuous hear kernel.

Theorem 6.1. *If μ is a Borel regular probability measure on K which satisfies that $\mu(O) > 0$ for any non-empty open subset of K and that $\mu(F) = 0$ for any finite set F, then $(\mathcal{E}, \mathcal{F})$ is a local regular Dirichlet form on $L^2(K,\mu)$ and the corresponding diffusion process possesses a jointly continuous heat kernel $p_\mu(t,x,y)$.*

Proof. Let $g(x,y)$ be Green's function associated with $(\mathcal{E}, \mathcal{F}_0)$, where $\mathcal{F}_0 = \{u | u \in \mathcal{F}, u|_{V_0} = 0\}$. (See [9, Section 3.8] for the existence and the continuity of Green's function.) Since the non-negative self-adjoint operator H_D associated with the quadratic form $(\mathcal{E}, \mathcal{F}_0)$ on $L^2(K, \mu)$ has compact resolvent, there exist a complete orthonormal system of $L^2(K, \mu)$, $\{\varphi_n^D\}_{n \geq 1}$, and $\{\lambda_n^D\}_{n \geq 1}$ such that $\varphi_n^D \in \mathcal{F}_0$, $H_D \varphi_n^D = \lambda_n^D \varphi_n^D$ and $0 < \lambda_n^D \leq \lambda_{n+1}^D$ for any $n \geq 1$. Let $g^x(y) = g(x, y)$ for any $x, y \in K$. Then $g^x \in \mathcal{F}_0$ and $\mathcal{E}(g^x, u) = u(x)$ for any $u \in \mathcal{F}_0$. Note that Green's function is the integral kernel of H_D^{-1}. Hence

$$\int_K g^x(y) \varphi_n^D(y) \mu(dy) = \frac{\varphi_n^D(x)}{\lambda_n^D}$$

Therefore $g^x = \sum_{n \geq 1} \frac{\varphi_n^D(x)}{\lambda_n^D} \varphi_n$, where the infinite sum converges in $L^2(K, \mu)$. This implies

$$g(x, x) = \mathcal{E}(g^x, g^x) = \sum_{n \geq 1} \frac{\left(\varphi_n^D(x)\right)^2}{\lambda_n^D}.$$

Integrating this, we obtain

$$\sum_{n \geq 1} \frac{1}{\lambda_n^D} < +\infty. \tag{6.1}$$

Since the non-negative self-adjoint operator H associated with the quadratic form $(\mathcal{E}, \mathcal{F})$ on $L^2(K, \mu)$ has compact resolvent, there exist a complete orthonormal system of $L^2(K, \mu)$, $\{\varphi_n\}_{n \geq 1}$, and $\{\lambda_n\}_{n \geq 1}$ such that $\varphi_n \in \mathcal{F}$, $H \varphi_n = \lambda_n \varphi_n$ and $0 \leq \lambda_n \leq \lambda_{n+1}$ for any n. Also, we see that $\varphi_1 \equiv 1$, $\lambda_1 = 0$ and $\lambda_n > 0$ for any $n \geq 2$. By [9, Theorem 4.1.7], it follows that $\lambda_n^D - N \leq \lambda_n \leq \lambda_n^D$. This along with (6.1) shows that

$$\sum_{n \geq 2} \frac{1}{\lambda_n} < +\infty \tag{6.2}$$

Now, since $(\mathcal{E}, \mathcal{F})$ is a resistance form,

$$|\varphi_n(x) - \varphi_n(y)|^2 \leq R(x, y) \mathcal{E}(\varphi_n, \varphi_n) = R(x, y) \lambda_n$$

for any $x, y \in K$. Since $\int_K \varphi_n d\mu = 0$ for $n \geq 2$, φ_n is continuous and K is arcwise connected, there exists $p \in K$ such that $\varphi_n(p) = 0$. Hence, $|\varphi_n(x)|^2 \leq R(x, p) \lambda_n$ for any $x \in K$ and any $n \geq 2$. Therefore,

$$||\varphi_n||_\infty^2 \leq C \lambda_n \tag{6.3}$$

for any $n \geq 2$, where $C = \sup_{x, y \in K} R(x, y) < +\infty$. On the other hand, for any $T > 0$, there exists $c_T > 0$ such that $e^{-\lambda_n t} \leq c_T / (\lambda_n)^2$ for any $t \in [T, \infty)$ and any $n \geq 2$. Combining this and (6.3), we see that $|e^{-\lambda_n t} \varphi_n(x) \varphi_n(y)| \leq C c_T / \lambda_n$ for any $x, y \in K$ and any $n \geq 2$. By (6.2), the sum $\sum_{n \geq 1} e^{-\lambda_n t} \varphi_n(x) \varphi_n(y)$ converges uniformly on $(t, x, y) \in [T, \infty) \times K \times K$. Hence if

$$p_\mu(t, x, y) = \sum_{n \geq 1} e^{-\lambda_n t} \varphi_n(x) \varphi_n(y),$$

then $p_\mu(t, x, y)$ is continuous on $(0, \infty) \times K \times K$. It is easy to see that $p_\mu(t, x, y)$ is the heat kernel. □

Remark. The essence of the above proof is that $(\mathcal{E}, \mathcal{F})$ is the resistance form and that the associated self-adjoint operator has compact resolvent. Therefore, Theorem 6.1 is true for Dirichlet forms derived from a regular harmonic structure on a post critical finite self-similar set. See [9] on the post critically finite self-similar sets. Moreover, with proper (technical) assumptions, it is even true for a resistance form on a general set.

The second step to the heat kernel estimate is the volume doubling property. Recall that the Kusuoka measure ν has the volume doubling property with respect to the Euclidean distance, which is, unfortunately, not the right one to describe the behavior of the heat kernel.

Theorem 6.2. *The Kusuoka measure ν has the volume doubling property with respect to the harmonic shortest path metric d_*.*

Proof. Let $w(1), w(2) \in W_m$ with $w(1) \neq w(2)$ and $K_{w(1)} \cap K_{w(2)} \neq \emptyset$. Then, there exist $n \geq 0$, $i, j \in S$ and $w \in W_*$ such that $w(1) = wi(j)^n$ and $w(2) = wj(i)^n$. Hence by the arguments in the proof of Lemma 3.5, $||T_{w(1)}||_{\mathrm{HS}} \leq c||T_{w(2)}||_{\mathrm{HS}}$, where c is a constance which is independent of $w(1), w(2)$ and m. This implies that $\mathcal{S}_*$ is gentle with respect to $\mathcal{S}_0$. Using [5, Theorem 1.4.3], we obtain that ν is gentle with respect to $\mathcal{S}_*$. Note that d_* is adapted to $\mathcal{S}_*$ by Theorem 5.11. As ν is elliptic, $\mathcal{S}_*$ is locally finite and ν is gentle with respect to $\mathcal{S}_*$, ν has the volume doubling property with respect to d_* by Theorem 3.8. $\square$

Finally we are ready to show the Li-Yau type Gaussian estimate of the heat kernel $p_\nu(t, x, y)$.

Theorem 6.3. *Let $p_\nu(t, x, y)$ be the heat kernel associated with the Dirichlet form $(\mathcal{E}, \mathcal{F})$ on $L^2(K, \nu)$. Then, there exist positive constants c_1, c_2, c_3 and c_4 such that*

$$\frac{c_1}{\nu(B_{\sqrt{t}}(x, d_*))} \exp\Big(- c_2 \frac{d_*(x,y)^2}{t}\Big) \leq p_\nu(t, x, y) \leq \frac{c_3}{\nu(B_{\sqrt{t}}(x, d_*))} \exp\Big(- c_4 \frac{d_*(x,y)^2}{t}\Big)$$

for any $t \in (0, 1]$ and any $x \in K$.

Proof. Since we are in the recurrent case, [5, Proposition 3.3.1] ensures all the prerequisites of [5, Theorem 3.2.3]. By Theorem 6.2, we have the condition (b) in the statement of [5, Theorem 3.2.3], which immediately deduces the desired heat kernel estimate. $\square$

Remark. In [14], Metz and Strum have obtained the following weak upper Gaussian estimate of the averaged heat kernel:

$$\frac{1}{\nu(A)\nu(B)} \int_A \int_B p_\nu(t, x, y)\nu(dx)\nu(dy) \leq \frac{1}{\sqrt{\nu(A)\nu(B)}} \exp\Big(- c_2 \frac{\inf_{x \in A, y \in B} |\Phi(x) - \Phi(y)|^2}{t}\Big)$$

for any compact sets A and B with $\nu(A), \nu(B) > 0$. Note that they have used the Euclidean distance on the harmonic Sierpinski gasket called the harmonic metric, which is not equivalent to the harmonic shortest path distance d_*. Recall the remark at the end of Section 5.

References

[1] M. T. Barlow and E. A. Perkins, *Brownian motion on the Sierpinski gasket*, Probab. Theory Related Fields **79** (1988), 542–624.

[2] M. Fukushima, Y. Oshima, and M. Takeda, *Dirichlet Forms and Symmetric Markov Processes*, de Gruyter Studies in Math. vol. 19, de Gruyter, Berlin, 1994.

[3] S. Goldstein, *Random walks and diffusions on fractals*, Percolation theory and ergodic theory of infinite particle systems (H. Kesten, ed.), IMA Math. Appl., vol. 8, Springer, 1987, pp. 121–129.

[4] A. Grigor'yan, *The heat equation on noncompact Riemannian manifolds. (in Russian)*, Mat. Sb. **182** (1991), 55–87, English translation in Math. USSR-Sb. 72(1992), 47–77.

[5] J. Kigami, *Volume doubling measures and heat kernel estimates on self-similar sets*, to appear in Memoirs of the American Mathematical Society.

[6] ______, *A harmonic calculus on the Sierpinski spaces*, Japan J. Appl. Math. **6** (1989), 259–290.

[7] ______, *Harmonic calculus on p.c.f. self-similar sets*, Trans. Amer. Math. Soc. **335** (1993), 721–755.

[8] ______, *Harmonic metric and Dirichlet form on the Sierpinski gasket*, Asymptotic problems in probability theory: stochastic models and diffusions on fractals (K. D. Elworthy and N. Ikeda, eds.), Pitman Research Notes in Math., vol. 283, Longman, 1993, pp. 201–218.

[9] ______, *Analysis on Fractals*, Cambridge Tracts in Math. vol. 143, Cambridge University Press, 2001.

[10] S. Kusuoka, *A diffusion process on a fractal*, Probabilistic Methods on Mathematical Physics, Proc. of Taniguchi International Symp. (Katata & Kyoto, 1985) (Tokyo) (K. Ito and N. Ikeda, eds.), Kinokuniya, 1987, pp. 251–274.

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

[12] ______, *Lecture on diffusion process on nested fractals*, Lecture Notes in Math., vol. 1567, pp. 39–98, Springer, 1993.

[13] P. Li and S.-T. Yau, *On the parabolic kernel of the Schrödinger operator*, Acta Math. **156** (1986), 153 –201.

[14] V. Metz and K. Sturm, *Gaussian and non-Gaussian estimates for heat kernels on the Sierpinski gasket*, Dirichlet forms and stochastic processes (Beijing, 1993) (Berlin), de Gruyter, 1995, pp. 283–289.

[15] L. Saloff-Coste, *A note on Poincaré, Sobolev, and Harnack inequalities*, Internat. Math. Res. Notices (1992), 27–38.

[16] A. Teplyaev, *Harmonic coordinates on fractals with finitely ramified cell structure*, to appear in Canadian Journal of Mathematics.

[17] ———, *Gradients on fractals*, J. Functional Analysis **174** (2000), 128–154.

[18] ———, *Energy and Laplacians on the Sierpinski gasket*, Fractal Geometry and Applications: A Jubilee of Benoit Mandelbrot (M. Lapidus and M. van Frankenhuijsen, eds.), Proc. Symposia in Pure Math., vol. 72, Part 1, Amer. Math. Soc., 2004, pp. 131–154.