Heat Kernels on Metric Measure Spaces

Alexander Grigor'yan, Jiaxin Hu and Ka-Sing Lau

Contents

A. Grigor'yan
Department of Mathematics, University of Bielefeld, 33501 Bielefeld,
Germany

J. Hu (✉)
Department of Mathematical Sciences, Tsinghua University,
Beijing 100084, China
e-mail: hujiaxin@mail.tsinghua.edu.cn

K.-S. Lau
Department of Mathematics, The Chinese University of Hong Kong,
Shatin, N.T., Hong Kong

1 What is the Heat Kernel

In this section we shall discuss the notion of the heat kernel on a metric measure space (M, d, μ). Loosely speaking, a heat kernel $p_t(x, y)$ is a family of measurable functions in $x, y \in M$ for each $t > 0$ that is *symmetric*, *Markovian* and satisfies the *semigroup* property and the *approximation of identity* property. It turns out that the heat kernel coincides with the *integral kernel* of the *heat semigroup* associated with the *Dirichlet form* in $L^2(M, \mu)$.

Let us start with some basic examples of the heat kernels.

1.1 Examples of Heat Kernels

1.1.1 Heat Kernel in Euclidean Spaces

The classical Gauss-Weierstrass heat kernel is the following function

$$p_t(x, y) = \frac{1}{(4\pi t)^{n/2}} \exp\left(-\frac{|x - y|^2}{4t}\right), \tag{1.1}$$

where $x, y \in \mathbb{R}^n$ and $t > 0$. This function is a fundamental solution of the heat equation

$$\frac{\partial u}{\partial t} = \Delta u,$$

where $\Delta = \sum_{i=1}^{n} \frac{\partial^2}{\partial x_i^2}$ is the Laplace operator. Moreover, if f is a continuous bounded function on $\mathbb{R}^n$, then the function

$$u(t, x) = \int_{\mathbb{R}^n} p_t(x, y) f(y) \, dy$$

solves the Cauchy problem

$$\begin{cases} \frac{\partial u}{\partial t} = \Delta u, \\ u\,(0, x) = f\,(x)\,. \end{cases}$$

This can be also written in the form

$$u\,(t, \cdot) = \exp(-t\mathcal{L})\,f,$$

where $\mathcal{L}$ here is a self-adjoint extension of $-\Delta$ in $L^2(\mathbb{R}^n)$ and $\exp(-t\mathcal{L})$ is understood in the sense of the functional calculus of self-adjoint operators. That means that $p_t\,(x, y)$ is the integral kernel of the operator $\exp(-t\mathcal{L})$.

The function $p_t(x, y)$ has also a probabilistic meaning: it is the transition density of Brownian motion $\{X_t\}_{t \geq 0}$ in $\mathbb{R}^n$ (Fig. 1). The graph of $p_t(x, 0)$ as a function of x is shown here:

The term $\frac{|x-y|^2}{t}$ determines the *space/time scaling*: if $|x - y|^2 \leq Ct$, then $p_t\,(x, y)$ is comparable with $p_t\,(x, x)$, that is, the probability density in the $C\sqrt{t}$-neighborhood of x is nearly constant.

1.1.2 Heat Kernels on Riemannian Manifolds

Let (M, g) be a connected Riemannian manifold, and Δ be the Laplace-Beltrami operator on M. Then the heat kernel $p_t\,(x, y)$ can be defined as the integral kernel of the heat semigroup $\{\exp(-t\mathcal{L})\}_{t \geq 0}$, where $\mathcal{L}$ is the Dirichlet Laplace operator, that is, the minimal self-adjoint extension of $-\Delta$ in $L^2\,(M, \mu)$, and μ is the Riemannian volume. Alternatively, $p_t\,(x, y)$ is the minimal positive fundamental solution of the heat equation

$$\frac{\partial u}{\partial t} = \Delta u.$$

The function $p_t\,(x, y)$ can be used to define Brownian motion $\{X_t\}_{t \geq 0}$ on M. Namely, $\{X_t\}_{t \geq 0}$ is a *diffusion process* (that is, a Markov process with continuous trajectories), such that

$$\mathbb{P}_x\,(X_t \in A) = \int_A p_t\,(x, y)\,d\mu\,(y)$$

for any Borel set $A \subset M$ (Fig. 2).

Let $d\,(x, y)$ be the geodesic distance on M. It turns out that the *Gaussian* type space/time scaling $\frac{d^2(x,y)}{t}$ appears in heat kernel estimates on general Riemannian manifolds:

1. (*Varadhan*) For an arbitrary Riemannian manifold,

$$\log p_t\,(x, y) \sim -\frac{d^2\,(x, y)}{4t} \quad \text{as } t \to 0.$$
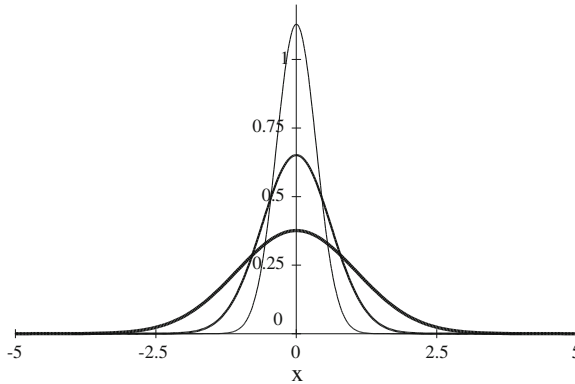
Fig. 1 The Gauss-Weierstrass heat kernel at different values of t
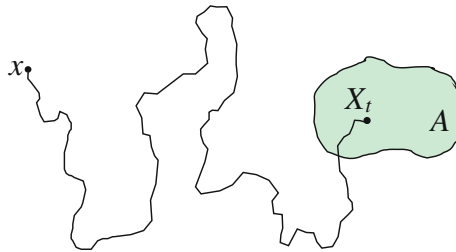
Fig. 2 The Brownian motion X_t hits a set A

2. (*Davies*) For an arbitrary manifold M, for any two measurable sets $A, B \subset M$

$$\int_A \int_B p_t(x, y) \, d\mu(x) \, d\mu(y) \leq \sqrt{\mu(A)\mu(B)} \exp\left(-\frac{d^2(A, B)}{4t}\right).$$

Technically, all these results depend upon the property of the geodesic distance: $|\nabla d| \leq 1$.

It is natural to ask the following question:

Are there settings where the space/time scaling is different from Gaussian?

1.1.3 Heat Kernels of Fractional Powers of Laplacian

Easy examples can be constructed using another operator instead of the Laplacian. As above, let $\mathcal{L}$ be the Dirichlet Laplace operator on a Riemannian manifold M, and consider the evolution equation

$$\frac{\partial u}{\partial t} + \mathcal{L}^{\beta/2} u = 0,$$

where $\beta \in (0, 2)$. The operator $\mathcal{L}^{\beta/2}$ is understood in the sense of the functional calculus in $L^2(M, \mu)$. Let $p_t(x, y)$ be now the heat kernel of $\mathcal{L}^{\beta/2}$, that is, the integral kernel of $\exp\left(-t\mathcal{L}^{\beta/2}\right)$.

The condition $\beta < 2$ leads to the fact that the semigroup $\exp\left(-t\mathcal{L}^{\beta/2}\right)$ is *Markovian*, which, in particular, means that $p_t(x, y) > 0$ (if $\beta > 2$ then $p_t(x, y)$ may be signed). Using the techniques of subordinators, one obtains the following estimate for the heat kernel of $\mathcal{L}^{\beta/2}$ in $\mathbb{R}^n$:

$$p_t(x, y) \asymp \frac{C}{t^{n/\beta}} \left(1 + \frac{|x - y|}{t^{1/\beta}}\right)^{-(n+\beta)} \asymp \frac{C}{t^{n/\beta}} \left(1 + \frac{|x - y|^\beta}{t}\right)^{-\frac{n+\beta}{\beta}}. \quad (1.2)$$

(the symbol $\asymp$ means that both $\leq$ and $\geq$ are valid but with different values of the constant C).

The heat kernel of $\sqrt{\mathcal{L}} = (-\Delta)^{1/2}$ in $\mathbb{R}^n$ (that is, the case $\beta = 1$) is known explicitly:

$$p_t(x, y) = \frac{c_n}{t^n} \left(1 + \frac{|x - y|^2}{t^2}\right)^{-\frac{n+1}{2}} = \frac{c_n t}{\left(t^2 + |x - y|^2\right)^{\frac{n+1}{2}}},$$

where $c_n = \Gamma\left(\frac{n+1}{2}\right)/\pi^{(n+1)/2}$. This function coincides with the Poisson kernel in the half-space $\mathbb{R}^{n+1}_+$ and with the density of the Cauchy distribution in $\mathbb{R}^n$ with the parameter t.

As we have seen, the space/time scaling is given by the term $\frac{d^\beta(x, y)}{t}$ where $\beta < 2$. The heat kernel of the operator $\mathcal{L}^{\beta/2}$ is the transition density of a *symmetric stable process of index* β that belongs to the family of Lévy processes. The trajectories of this process are discontinuous, thus allowing jumps. The heat kernel $p_t(x, y)$ of such process is nearly constant in some $Ct^{1/\beta}$-neighborhood of y. If t is large, then

$$t^{1/\beta} \gg t^{1/2},$$

that is, this neighborhood is much larger than that for the diffusion process, which is not surprising because of the presence of jumps. The space/time scaling with $\beta < 2$ is called *super-Gaussian*.

1.1.4 Heat Kernels on Fractal Spaces

A rich family of heat kernels for diffusion processes has come from Analysis on *fractals*. Loosely speaking, fractals are subsets of $\mathbb{R}^n$ with certain self-similarity
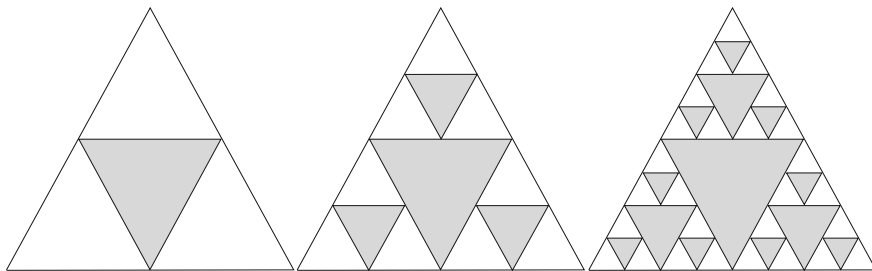
Fig. 3 Construction of the Sierpinski gasket
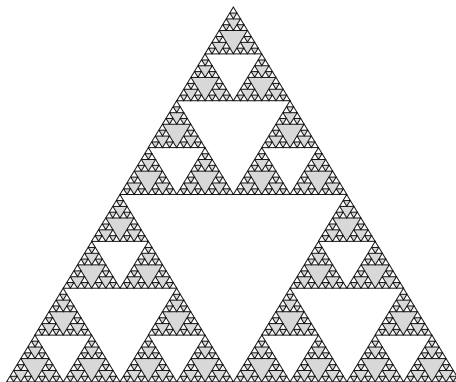
Fig. 4 The unbounded SG is obtained from SG by merging the latter (at the *left lower corner* of the diagram) with two shifted copies and then by repeating this procedure at larger scales

properties. One of the best understood fractals is *the Sierpinski gasket* (SG). The construction of the Sierpinski gasket is similar to the Cantor set: one starts with a triangle as a closed subset of $\mathbb{R}^2$, then eliminates the open middle triangle (shaded on the diagram), then repeats this procedure for the remaining triangles, and so on (Fig. 3).

Hence, SG is a compact connected subset of $\mathbb{R}^2$. The *unbounded* SG is obtained from SG by merging the latter (at the left lower corner of the next diagram) with two shifted copies and then by repeating this procedure at larger scales (Fig. 4).

Barlow and Perkins [BP88], Goldstein [Gol87] and Kusuoka [Kus87] have independently constructed by different methods a natural diffusion process on SG that has the same self-similarity as SG. Barlow and Perkins considered random walks on the graph approximations of SG and showed that, with an appropriate scaling, the random walks converge to a diffusion process. Moreover, they proved that this process has a transition density $p_t(x, y)$ with respect to a proper Hausdorff measure μ of SG, and that p_t satisfies the following elegant estimate:

$$p_t(x, y) \asymp \frac{C}{t^{\alpha/\beta}} \exp\left(-c\left(\frac{d^\beta(x, y)}{t}\right)^{\frac{1}{\beta-1}}\right), \tag{1.3}$$

where $d(x, y) = |x - y|$ and

$$\alpha = \dim_H SG = \frac{\log 3}{\log 2}, \quad \beta = \frac{\log 5}{\log 2} > 2.$$

Similar estimates were proved by Barlow and Bass for other families of fractals, including Sierpinski carpets, and the parameters α and β in (1.3) are determined by the intrinsic properties of the fractal. In all cases, α is the Hausdorff dimension (which is also called the *fractal dimension*). The parameter β, that is called the *walk dimension*, is larger than 2 in all interesting examples.

The heat kernel $p_t(x, y)$, satisfying (1.3) is nearly constant in some $Ct^{1/\beta}$-neighborhood of y. If t is large, then

$$t^{1/\beta} \ll t^{1/2},$$

that is, this neighborhood is much smaller than that for the diffusion process, which is due to the presence of numerous holes-obstacles that the Brownian particle must bypass. The space/time scaling with $\beta > 2$ is called *sub-Gaussian*.

1.1.5 Summary of Examples

Observe now that in all the above examples, the heat kernel estimates can be unified as follows:

$$p_t(x, y) \asymp \frac{C}{t^{\alpha/\beta}} \Phi\left(c\frac{d(x, y)}{t^{1/\beta}}\right), \tag{1.4}$$

where α, β are positive parameters and $\Phi(s)$ is a positive decreasing function on $[0, +\infty)$. For example, the Gauss-Weierstrass function (1.1) satisfies (1.4) with $\alpha = n, \beta = 2$ and

$$\Phi(s) = \exp\left(-s^2\right)$$

(Gaussian estimate).

The heat kernel (1.2) of the symmetric stable process in $\mathbb{R}^n$ satisfies (1.4) with $\alpha = n, 0 < \beta < 2$, and

$$\Phi(s) = (1 + s)^{-(\alpha+\beta)}$$

(super-Gaussian estimate).

The heat kernel (1.3) of diffusions on fractals satisfies (1.4) with $\beta > 2$ and

$$\Phi(s) = \exp\left(-s^{\frac{\beta}{\beta-1}}\right)$$

(sub-Gaussian estimate).

There are at least two questions related to the estimates of the type (1.4):

1. What values of the parameters α, β and what functions Φ can actually occur in the estimate (1.4)?
2. How to obtain estimates of the type (1.4)?

To give these questions a precise meaning, we must define what is a heat kernel.

1.2 Abstract Heat Kernels

Let (M, d) be a locally compact, separable metric space and let μ be a Radon measure on M with full support. The triple (M, d, μ) will be called a *metric measure space*.

Definition 1.1 (*heat kernel*) A family $\{p_t\}_{t>0}$ of measurable functions $p_t(x, y)$ on $M \times M$ is called a *heat kernel* if the following conditions are satisfied, for μ-almost all $x, y \in M$ and all $s, t > 0$:

(i) Positivity: $p_t(x, y) \geq 0$.
(ii) The total mass inequality:

$$\int_M p_t(x, y)d\mu(y) \leq 1.$$

(iii) Symmetry: $p_t(x, y) = p_t(y, x)$.
(iv) The semigroup property:

$$p_{s+t}(x, y) = \int_M p_s(x, z)p_t(z, y)d\mu(z).$$

(v) Approximation of identity: for any $f \in L^2 := L^2(M, \mu)$,

$$\int_M p_t(x, y)f(y)d\mu(y) \xrightarrow{L^2} f(x) \text{ as } t \to 0+.$$

If in addition we have, for all $t > 0$ and almost all $x \in M$,

$$\int_M p_t(x, y)d\mu(y) = 1,$$

then the heat kernel p_t is called *stochastically complete* (or *conservative*).

1.3 Heat Semigroups

Any heat kernel gives rise to the family of operators $\{P_t\}_{t \geq 0}$ where $P_0 = \mathrm{id}$ and P_t for $t > 0$ is defined by

$$P_t f(x) = \int_M p_t(x, y) f(y) d\mu(y),$$

where f is a measurable function on M. It follows from (i) –(ii) that the operator P_t is *Markovian*, that is, $f \geq 0$ implies $P_t f \geq 0$ and $f \leq 1$ implies $P_t f \leq 1$. It follows that P_t is a bounded operator in L^2 and, moreover, is a contraction, that is, $\|P_t f\|_2 \leq \|f\|_2$.

The symmetry property (iii) implies that the operator P_t is *symmetric* and, hence, self-adjoint. The semigroup property (iv) implies that $P_t P_s = P_{t+s}$, that is, the family $\{P_t\}_{t \geq 0}$ is a *semigroup* of operators. It follows from (v) that

$$s\text{-}\lim_{t \to 0} P_t = \mathrm{id} = P_0$$

where s-lim stands for the *strong* limit. Hence, $\{P_t\}_{t \geq 0}$ is a strongly continuous, symmetric, Markovian semigroup in L^2. In short, we call that $\{P_t\}$ is a *heat semigroup*.

Conversely, if $\{P_t\}$ is a heat semigroup and if it has an integral kernel $p_t(x, y)$, then the latter is a heat kernel in the sense of the above Definition.

Given a heat semigroup P_t in L^2, define the *infinitesimal generator* $\mathcal{L}$ of the semigroup by

$$\mathcal{L}f := \lim_{t \to 0} \frac{f - P_t f}{t},$$

where the limit is understood in the L^2-norm. The *domain* $\mathrm{dom}(\mathcal{L})$ of the generator $\mathcal{L}$ is the space of functions $f \in L^2$ for which the above limit exists. By the Hille–Yosida theorem, $\mathrm{dom}(\mathcal{L})$ is dense in L^2. Furthermore, $\mathcal{L}$ is a self-adjoint, positive definite operator, which immediately follows from the fact that the semigroup $\{P_t\}$ is self-adjoint and contractive. Moreover, P_t can be recovered from $\mathcal{L}$ as follows

$$P_t = \exp(-t\mathcal{L}),$$

where the right hand side is understood in the sense of spectral theory.

Heat kernels and heat semigroups arise naturally from Markov processes. Let $\left(\{X_t\}_{t \geq 0}, \{\mathbb{P}_x\}_{x \in M}\right)$ be a Markov process on M, that is reversible with respect to measure μ. Assume that it has the transition density $p_t(x, y)$, that is, a function such

that, for all $x \in M$, $t > 0$, and all Borel sets $A \subset M$,

$$\mathbb{P}_x \left(X_t \in A \right) = \int_M p_t \left(x, y \right) d\mu \left(y \right).$$

Then $p_t \left(x, y \right)$ is a heat kernel in the sense of the above Definition.

1.4 Dirichlet Forms

Given a heat semigroup $\{P_t\}$ on a metric measure space (M, d, μ), for any $t > 0$, we define a bilinear form $\mathcal{E}_t$ on L^2 by

$$\mathcal{E}_t \left(u, v \right) := \left(\frac{u - P_t u}{t}, v \right) = \frac{1}{t} \left(\left(u, v \right) - \left(P_t u, v \right) \right),$$

where $(\cdot, \cdot)$ is the inner product in L^2. Since P_t is symmetric, the form $\mathcal{E}_t$ is also symmetric. Since P_t is a contraction, it follows that

$$\mathcal{E}_t \left(u \right) := \mathcal{E}_t \left(u, u \right) = \frac{1}{t} \left(\left(u, u \right) - \left(P_t u, u \right) \right) \geq 0,$$

that is, $\mathcal{E}_t$ is a positive definite form.

In terms of the spectral resolution $\{E_\lambda\}$ of the generator $\mathcal{L}$, $\mathcal{E}_t$ can be expressed as follows

$$\mathcal{E}_t \left(u \right) = \frac{1}{t} \left(\left(u, u \right) - \left(P_t u, u \right) \right) = \frac{1}{t} \left(\int_0^\infty d\|E_\lambda u\|_2^2 - \int_0^\infty e^{-t\lambda} d\|E_\lambda u\|_2^2 \right)$$

$$= \int_0^\infty \frac{1 - e^{-t\lambda}}{t} d\|E_\lambda u\|_2^2,$$

which implies that $\mathcal{E}_t \left(u \right)$ is decreasing in t, since the function $t \mapsto \frac{1-e^{-t\lambda}}{t}$ is decreasing. Define for any $u \in L^2$

$$\mathcal{E} \left(u \right) = \lim_{t \downarrow 0} \mathcal{E}_t \left(u \right)$$

where the limit (finite or infinite) exists by the monotonicity, so that $\mathcal{E} \left(u \right) \geq \mathcal{E}_t \left(u \right)$. Since $\frac{1-e^{-t\lambda}}{t} \to \lambda$ as $t \to 0$, we have

$$\mathcal{E}\,(u) = \int\limits_0^\infty \lambda d\,\|E_\lambda u\|_2^2.$$

Set

$$\mathcal{F} := \{u \in L^2 : \mathcal{E}\,(u) < \infty\} = \mathrm{dom}\left(\mathcal{L}^{1/2}\right) \supset \mathrm{dom}\,(\mathcal{L})$$

and define a bilinear form $\mathcal{E}\,(u, v)$ on $\mathcal{F}$ by the polarization identity

$$\mathcal{E}\,(u, v) := \frac{1}{4}\left(\mathcal{E}\,(u + v) - \mathcal{E}\,(u - v)\right),$$

which is equivalent to

$$\mathcal{E}\,(u, v) = \lim_{t \to 0} \mathcal{E}_t\,(u, v)\,.$$

Note that $\mathcal{F}$ contains $\mathrm{dom}(\mathcal{L})$. Indeed, if $u \in \mathrm{dom}(\mathcal{L})$, then we have for all $v \in L^2$

$$\lim_{t \to 0} \mathcal{E}_t\,(u, v) = \left(\lim_{t \to 0}\frac{u - P_t u}{t}, v\right) = (\mathcal{L}u, v)\,.$$

Setting $v = u$ we obtain $u \in \mathcal{F}$. Then choosing any $v \in \mathcal{F}$ we obtain the identity

$$\mathcal{E}(u, v) = (\mathcal{L}u, v)\ \text{for all}\ u \in \mathrm{dom}(\mathcal{L})\ \text{and}\ v \in \mathcal{F}.$$

The space $\mathcal{F}$ is naturally endowed with the inner product

$$[u, v] := (u, v) + \mathcal{E}\,(u, v)\,.$$

It is possible to show that the form $\mathcal{E}$ is *closed*, that is, the space $\mathcal{F}$ is *Hilbert*. Furthermore, $\mathrm{dom}\,(\mathcal{L})$ is dense in $\mathcal{F}$.

The fact that P_t is Markovian implies that the form $\mathcal{E}$ is also *Markovian*, that is

$$u \in \mathcal{F} \Rightarrow \widetilde{u} := \min(u_+, 1) \in \mathcal{F}\ \text{ and }\ \mathcal{E}\,(\widetilde{u}) \le \mathcal{E}\,(u)\,.$$

Indeed, let us first show that for any $u \in L^2$

$$\mathcal{E}_t\,(u_+) \le \mathcal{E}_t\,(u)\,.$$

We have

$$\mathcal{E}_t\,(u) = \mathcal{E}_t\,(u_+ - u_-) = \mathcal{E}_t\,(u_+) + \mathcal{E}_t\,(u_-) - 2\mathcal{E}_t\,(u_+, u_-) \ge \mathcal{E}_t\,(u_+)$$

because $\mathcal{E}_t\,(u_-) \ge 0$ and

$$\mathcal{E}_t\,(u_+, u_-) = \frac{1}{t}\,(u_+, u_-) - \frac{1}{t}\,(P_t u_+, u_-) \le 0.$$

Assuming $u \in \mathcal{F}$ and letting $t \to 0$, we obtain

$$\mathcal{E}\,(u_+) = \lim_{t \to 0} \mathcal{E}_t\,(u_+) \le \lim_{t \to 0} \mathcal{E}_t\,(u) = \mathcal{E}\,(u) < \infty$$

whence $\mathcal{E}\,(u_+) \le \mathcal{E}\,(u)$ and, hence, $u_+ \in \mathcal{F}$.

Similarly one proves that $\tilde{u} = \min(u_+, 1)$ belongs to $\mathcal{F}$ and $\mathcal{E}\,(\tilde{u}) \le \mathcal{E}\,(u_+)$.

Conclusion *Hence, $(\mathcal{E}, \mathcal{F})$ is a Dirichlet form, that is, a bilinear, symmetric, positive definite, closed, densely defined form in L^2 with Markovian property.*

If the heat semigroup is defined by means of a heat kernel p_t, then $\mathcal{E}_t$ can be equivalently defined by

$$\mathcal{E}_t\,(u) = \frac{1}{2t} \int_M \int_M (u(x) - u(y))^2\, p_t(x, y) d\mu(y) d\mu(x) \qquad (1.5)$$

$$+ \frac{1}{t} \int_M (1 - P_t 1(x))\, u^2(x) d\mu(x).$$

Indeed, we have

$$u(x) - P_t u(x) = u\,(x)\, P_t 1\,(x) - P_t u\,(x) + (1 - P_t 1(x))\, u\,(x)$$

$$= \int_M (u(x) - u(y))\, p_t(x, y) d\mu(y) + (1 - P_t 1(x))\, u\,(x),$$

whence

$$\mathcal{E}_t\,(u) = \frac{1}{t} \int_M \int_M (u(x) - u(y))\, u(x) p_t(x, y) d\mu(y) d\mu(x)$$

$$+ \frac{1}{t} \int_M (1 - P_t 1(x))\, u^2(x) d\mu(x).$$

Interchanging the variables x and y in the first integral and using the symmetry of the heat kernel, we obtain also

$$\mathcal{E}_t\,(u) = \frac{1}{t} \int_M \int_M (u(y) - u(x))\, u(y) p_t(x, y) d\mu(y) d\mu(x)$$

$$+ \frac{1}{t} \int_M (1 - P_t 1(x))\, u^2(x) d\mu(x),$$

and (1.5) follows by adding up the two previous lines.

Since $P_t 1 \leq 1$, the second term in the right hand side of (1.5) is non-negative. If the heat kernel is stochastically complete, that is, $P_t 1 = 1$, then that term vanishes and we obtain

$$\mathcal{E}_t(u) = \frac{1}{2t} \int\limits_M \int\limits_M (u(x) - u(y))^2 \, p_t(x, y) d\mu(y) d\mu(x). \qquad (1.6)$$

Definition 1.2 The form $(\mathcal{E}, \mathcal{F})$ is called *local* if $\mathcal{E}(u, v) = 0$ whenever the functions $u, v \in \mathcal{F}$ have compact disjoint supports. The form $(\mathcal{E}, \mathcal{F})$ is called *strongly local* if $\mathcal{E}(u, v) = 0$ whenever the functions $u, v \in \mathcal{F}$ have compact supports and $u \equiv \mathrm{const}$ in an open neighborhood of supp v.

For example, if $p_t(x, y)$ is the heat kernel of the Laplace-Beltrami operator on a complete Riemannian manifold, then the associated Dirichlet form is given by

$$\mathcal{E}(u, v) = \int\limits_M \langle \nabla u, \nabla v \rangle d\mu, \qquad (1.7)$$

and $\mathcal{F}$ is the Sobolev space $W_2^1(M)$. Note that this Dirichlet form is strongly local because $u = \mathrm{const}$ on supp v implies $\nabla u = 0$ on supp v and, hence, $\mathcal{E}(u, v) = 0$.

If $p_t(x, y)$ is the heat kernel of the symmetric stable process of index β in $\mathbb{R}^n$, that is, $\mathcal{L} = (-\Delta)^{\beta/2}$, then

$$\mathcal{E}(u, v) = c_{n,\beta} \int\limits_{\mathbb{R}^n} \int\limits_{\mathbb{R}^n} \frac{(u(x) - u(y))(v(x) - v(y))}{|x - y|^{n+\beta}} dx dy,$$

and $\mathcal{F}$ is the Besov space $B_{2,2}^{\beta/2}(\mathbb{R}^n) = \{u \in L^2 : \mathcal{E}(u, u) < \infty\}$. This form is clearly non-local.

Denote by $C_0(M)$ the space of continuous functions on M with compact supports, endowed with sup-norm.

Definition 1.3 The form $(\mathcal{E}, \mathcal{F})$ is called *regular* if $\mathcal{F} \cap C_0(M)$ is dense both in $\mathcal{F}$ (with $[\cdot, \cdot]$-norm) and in $C_0(M)$ (with sup-norm).

All the Dirichlet forms in the above examples are regular.

Assume that we are given a Dirichlet form $(\mathcal{E}, \mathcal{F})$ in $L^2(M, \mu)$. Then one can define the generator $\mathcal{L}$ of $(\mathcal{E}, \mathcal{F})$ by the identity

$$(\mathcal{L}u, v) = \mathcal{E}(u, v) \quad \text{for all } u \in \mathrm{dom}(\mathcal{L}), \ v \in \mathcal{F}, \qquad (1.8)$$

where $\mathrm{dom}(\mathcal{L}) \subset \mathcal{F}$ must satisfy one of the following two equivalent requirements:

1. $\mathrm{dom}(\mathcal{L})$ is a maximal possible subspace of $\mathcal{F}$ such that (1.8) holds
2. $\mathcal{L}$ is a densely defined self-adjoint operator.

Clearly, $\mathcal{L}$ is positive definite so that $\mathrm{spec}\mathcal{L} \subset [0, +\infty)$. Hence, the family of operators $P_t = e^{-t\mathcal{L}}$, $t \geq 0$, forms a strongly continuous, symmetric, contraction semigroup in L^2. Moreover, using the Markovian property of the Dirichlet form $(\mathcal{E}, \mathcal{F})$, it is possible to prove that $\{P_t\}$ is Markovian, that is, $\{P_t\}$ is a heat semigroup. The question whether and when P_t has the heat kernel requires a further investigation.

1.5 More Examples of Heat Kernels

Let us give some examples of stochastically complete heat kernels that do not satisfy (1.4).

Example 1.4 (*A frozen heat kernel*) Let M be a countable set and let $\{x_k\}_{k=1}^{\infty}$ be the sequence of all distinct points from M. Let $\{\mu_k\}_{k=1}^{\infty}$ be a sequence of positive reals and define measure μ on M by $\mu(\{x_k\}) = \mu_k$. Define a function $p_t(x, y)$ on $M \times M$ by

$$p_t(x, y) = \begin{cases} \frac{1}{\mu_k}, & x = y = x_k \\ 0, & \text{otherwise.} \end{cases}$$

It is easy to check that $p_t(x, y)$ is a heat kernel. For example, let us check the approximation of identity: for any function $f \in L^2(M, \mu)$, we have

$$P_t f(x) = \int_M p_t(x, y) f(y) \, d\mu(y) = p_t(x, x) f(x) \mu(\{x\}) = f(x).$$

This identity also implies the stochastic completeness. The Dirichlet form is

$$\mathcal{E}(f) = \lim_{t \to 0} \left(\frac{f - P_t f}{t}, f \right) = 0.$$

The Markov process associated with the frozen heat kernel is very simple: $X_t = X_0$ for all $t \geq 0$ so that it is a frozen diffusion.

Example 1.5 (*The heat kernel in $\mathbb{H}^3$*) The heat kernel of the Laplace-Beltrami operator on the 3-dimensional hyperbolic space $\mathbb{H}^3$ is given by the formula

$$p_t(x, y) = \frac{1}{(4\pi t)^{3/2}} \frac{r}{\sinh r} \exp\left(-\frac{r^2}{4t} - t \right),$$

where $r = d(x, y)$ is the geodesic distance between x, y. The Dirichlet form is given by (1.7).

Example 1.6 (*The Mehler heat kernel*) Let $M = \mathbb{R}$, measure μ be defined by

$$d\mu = e^{x^2} dx,$$

and let $\mathcal{L}$ be given by

$$\mathcal{L} = -e^{-x^2} \frac{d}{dx} \left(e^{x^2} \frac{d}{dx} \right) = -\frac{d^2}{dx^2} - 2x \frac{d}{dx}.$$

Then the heat kernel of $\mathcal{L}$ is given by the formula

$$p_t(x, y) = \frac{1}{(2\pi \sinh 2t)^{1/2}} \exp \left(\frac{2xye^{-2t} - x^2 - y^2}{1 - e^{-4t}} - t \right).$$

The associated Dirichlet form is also given by (1.7).

Similarly, for the measure

$$d\mu = e^{-x^2} dx$$

and for the operator

$$\mathcal{L} = -e^{x^2} \frac{d}{dx} \left(e^{-x^2} \frac{d}{dx} \right) = -\frac{d^2}{dx^2} + 2x \frac{d}{dx},$$

we have

$$p_t(x, y) = \frac{1}{(2\pi \sinh 2t)^{1/2}} \exp \left(\frac{2xye^{-2t} - \left(x^2 + y^2 \right) e^{-4t}}{1 - e^{-4t}} + t \right).$$

2 Necessary Conditions for Heat Kernel Bounds

In this Chapter we assume that $p_t(x, y)$ is a heat kernel on a metric measure space (M, d, μ) that satisfies certain upper and/or lower estimates, and state the consequences of these estimates. The reader may consult [GK08, GHL03] or [GHL09] for the proofs.

Fix two positive parameters α and β, and let $\Phi : [0, +\infty) \to [0, +\infty)$ be a monotone decreasing function. We will consider the bounds of the heat kernel via the function $\frac{1}{t^{\alpha/\beta}} \Phi \left(\frac{d(x,y)}{t^{1/\beta}} \right)$.

2.1 Identifying Φ in the Non-local Case

Theorem 2.1 (Grigor'yan and Kumagai [GK08]) *Let $p_t(x, y)$ be a heat kernel on* (M, d, μ).

(a) *If the heat kernel satisfies the estimate*

$$p_t (x, y) \leq \frac{1}{t^{\alpha/\beta}} \Phi \left(\frac{d (x, y)}{t^{1/\beta}} \right),$$

for all $t > 0$ and almost all $x, y \in M$, then either the associated Dirichlet form $(\mathcal{E}, \mathcal{F})$ is local or

$$\Phi (s) \geq c (1 + s)^{-(\alpha+\beta)}$$

for all $s > 0$ and some $c > 0$.
(b) If the heat kernel satisfies the estimate

$$p_t (x, y) \geq \frac{1}{t^{\alpha/\beta}} \Phi \left(\frac{d (x, y)}{t^{1/\beta}} \right),$$

then we have

$$\Phi (s) \leq C (1 + s)^{-(\alpha+\beta)}$$

for all $s > 0$ and some $C > 0$.
(c) Consequently, if the heat kernel satisfies the estimate

$$p_t (x, y) \asymp \frac{C}{t^{\alpha/\beta}} \Phi \left(c \frac{d (x, y)}{t^{1/\beta}} \right),$$

then either the Dirichlet form $\mathcal{E}$ is local or

$$\Phi (s) \asymp (1 + s)^{-(\alpha+\beta)} .$$

2.2 Volume of Balls

Denote by $B (x, r)$ a metric ball in (M, d), that is

$$B(x, r) := \{y \in M : d(x, y) < r\} .$$

Theorem 2.2 (Grigor'yan et al. [GHL03]) *Let p_t be a heat kernel on (M, d, μ). Assume that it is stochastically complete and that it satisfies the two-sided estimate*

$$p_t (x, y) \asymp \frac{C}{t^{\alpha/\beta}} \Phi \left(c \frac{d (x, y)}{t^{1/\beta}} \right). \tag{2.1}$$

Then, for all $x \in M$ and $r > 0$,

$$\mu(B(x, r)) \asymp r^{\alpha},$$

that is, μ is α-regular.

Consequently, $\dim_H (M, d) = \alpha$ and $\mu \asymp H^\alpha$ on all Borel subsets of M, where H^α is the α-dimensional Hausdorff measure in M.

In particular, the parameter α is the invariant of the metric space (M, d), and measure μ is determined (up to a factor $\asymp 1$) by the metric space (M, d).

2.3 Besov Spaces

Fix $\alpha > 0$, $\sigma > 0$. We introduce the following seminorms on $L^2 = L^2 (M, \mu)$:

$$N_{2,\infty}^{\alpha,\sigma} (u) = \sup_{0 < r \leq 1} \frac{1}{r^{\alpha+2\sigma}} \iint\limits_{\{x,y \in M : d(x,y) < r\}} |u(x) - u(y)|^2 \, d\mu(x) d\mu(y), \qquad (2.2)$$

and

$$N_{2,2}^{\alpha,\sigma}(u) = \int\limits_0^\infty \frac{dr}{r} \frac{1}{r^{\alpha+2\sigma}} \iint\limits_{\{x,y \in M : d(x,y) < r\}} |u(x) - u(y)|^2 \, d\mu(x) d\mu(y). \qquad (2.3)$$

Define the space $\Lambda_{2,\infty}^{\alpha,\sigma}$ by

$$\Lambda_{2,\infty}^{\alpha,\sigma} = \left\{ u \in L^2 : N_{2,\infty}^{\alpha,\sigma}(u) < \infty \right\},$$

and the norm by

$$\|u\|_{\Lambda_{2,\infty}^{\alpha,\sigma}}^2 = \|u\|_2^2 + N_{2,\infty}^{\alpha,\sigma}(u).$$

Similarly, one defines the space $\Lambda_{2,2}^{\alpha,\sigma}$. More generally one can define $\Lambda_{p,q}^{\alpha,\sigma}$ for $p \in [1, +\infty)$ and $q \in [1, +\infty]$.

In the case of $\mathbb{R}^n$, we have the following relations

$$\begin{aligned}
\Lambda_{p,q}^{n,\sigma} \left(\mathbb{R}^n\right) &= B_{p,q}^\sigma \left(\mathbb{R}^n\right), \ 0 < \sigma < 1, \\
\Lambda_{2,\infty}^{n,1} \left(\mathbb{R}^n\right) &= W_p^1 \left(\mathbb{R}^n\right), \\
\Lambda_{2,2}^{n,1} \left(\mathbb{R}^n\right) &= \{0\}, \\
\Lambda_{p,q}^{n,\sigma} \left(\mathbb{R}^n\right) &= \{0\}, \ \sigma > 1.
\end{aligned}$$

where $B_{p,q}^\sigma$ is the Besov space and W_p^1 is the Sobolev space. The spaces $\Lambda_{p,q}^{\alpha,\sigma}$ will also be called Besov spaces.

Theorem 2.3 (Jonsson [Jon96], Pietruska-Pałuba [Pie00], Grigor'yan et al. [GHL03]) *Let p_t be a heat kernel on (M, d, μ). Assume that it is stochastically*

complete and that it satisfies the following estimate: for all $t > 0$ *and almost all* $x, y \in M$,

$$\frac{1}{t^{\alpha/\beta}} \Phi_1 \left(\frac{d(x, y)}{t^{1/\beta}} \right) \leq p_t(x, y) \leq \frac{1}{t^{\alpha/\beta}} \Phi_2 \left(\frac{d(x, y)}{t^{1/\beta}} \right), \qquad (2.4)$$

where α, β *be positive constants, and* Φ_1, Φ_2 *are monotone decreasing functions from* $[0, +\infty)$ *to* $[0, +\infty)$ *such that* $\Phi_1(s) > 0$ *for some* $s > 0$ *and*

$$\int_0^\infty s^{\alpha+\beta} \Phi_2(s) \frac{ds}{s} < \infty. \qquad (2.5)$$

Then, for any $u \in L^2$,

$$\mathcal{E}(u) \asymp N_{2,\infty}^{\alpha,\beta/2}(u),$$

and, consequently, $\mathcal{F} = \Lambda_{2,\infty}^{\alpha,\beta/2}$.

By Theorem 2.1, the upper bound in (2.4) implies that either $(\mathcal{E}, \mathcal{F})$ is local or

$$\Phi_2(s) \geq c (1 + s)^{-(\alpha+\beta)}.$$

Since the latter contradicts condition (2.5), the form $(\mathcal{E}, \mathcal{F})$ must be local. For non-local forms the statement is not true. For example, for the operator $(-\Delta)^{\beta/2}$ in $\mathbb{R}^n$, we have $\mathcal{F} = B_{2,2}^{\beta/2} = \Lambda_{2,2}^{n,\beta/2}$ that is strictly smaller than $B_{2,\infty}^{\beta/2} = \Lambda_{2,\infty}^{n,\beta/2}$. This case will be covered by the following theorem.

Theorem 2.4 (Stós [Sto00]) *Let* p_t *be a stochastically complete heat kernel on* (M, d, μ) *satisfying estimate* (2.4) *with functions*

$$\Phi_1(s) \asymp \Phi_2(s) \asymp (1 + s)^{-(\alpha+\beta)}.$$

Then, for any $u \in L^2$,

$$\mathcal{E}(u) \asymp N_{2,2}^{\alpha,\beta/2}(u).$$

Consequently, we have $\mathcal{F} = \Lambda_{2,2}^{\alpha,\beta/2}$.

2.4 Subordinated Semigroups

Let $\mathcal{L}$ be the generator of a heat semigroup $\{P_t\}$. Then, for any $\delta \in (0, 1)$, the operator $\mathcal{L}^\delta$ is also a generator of a heat semigroup, that is, the semigroup $\{e^{-t\mathcal{L}^\sigma}\}$

is a heat semigroup. Furthermore, there is the following relation between the two semigroups

$$e^{-t\mathcal{L}^\delta} = \int_0^\infty e^{-s\mathcal{L}} \eta_t(s)\,ds,$$

where $\eta_t(s)$ is a *subordinator* whose Laplace transform is given by

$$e^{-t\lambda^\delta} = \int_0^\infty e^{-s\lambda} \eta_t(s)\,ds, \ \lambda > 0.$$

Using the known estimates for $\eta_t(s)$, one can obtain the following result.

Theorem 2.5 *Let a heat kernel p_t satisfy the estimate (2.4) where $\Phi_1(s) > 0$ for some $s > 0$ and*

$$\int_0^\infty s^{\alpha+\beta'} \Phi_2(s) \frac{ds}{s} < \infty,$$

where $\beta' = \delta\beta$, $0 < \delta < 1$. Then the heat kernel $q_t(x, y)$ of operator $\mathcal{L}^\delta$ satisfies the estimate

$$q_t(x, y) \asymp \frac{1}{t^{\alpha/\beta'}} \left(1 + \frac{d(x, y)}{t^{1/\beta'}}\right)^{-(\alpha+\beta')} \asymp \min\left(t^{-\alpha/\beta'}, \frac{t}{d(x, y)^{\alpha+\beta'}}\right),$$

for all $t > 0$ and almost all $x, y \in M$.

2.5 The Walk Dimension

It follows from definition that the Besov seminorm

$$N_{2,\infty}^{\alpha,\sigma}(u) = \sup_{0 < r \le 1} \frac{1}{r^{\alpha+2\sigma}} \iint\limits_{\{x,y \in M : d(x,y) < r\}} |u(x) - u(y)|^2\,d\mu(x)d\mu(y)$$

increases when σ increases, which implies that the space

$$\Lambda_{2,\infty}^{\alpha,\sigma} := \left\{u \in L^2 : N_{2,\infty}^{\alpha,\sigma}(u) < \infty\right\}$$

shrinks. For a certain value of σ, this space may become trivial. For example, as was already mentioned, $\Lambda_{2,\infty}^{n,\sigma}(\mathbb{R}^n) = \{0\}$ for $\sigma > 1$, while $\Lambda_{2,\infty}^{n,\sigma}(\mathbb{R}^n)$ is non-trivial for $\sigma \leq 1$.

Definition 2.6 Fix $\alpha > 0$ and set

$$\beta^* := \sup\left\{\beta > 0 : \Lambda_{2,\infty}^{\alpha,\beta/2} \text{ is dense in } L^2(M,\mu)\right\}. \tag{2.6}$$

The number $\beta^* \in [0, +\infty]$ is called the *critical exponent* of the family $\left\{\Lambda_{2,\infty}^{\alpha,\beta/2}\right\}_{\beta>0}$ of Besov spaces.

Note that the value of β^* is an intrinsic property of the space (M, d, μ), which is defined independently of any heat kernel. For example, for $\mathbb{R}^n$ with $\alpha = n$ we have $\beta^* = 2$.

Theorem 2.7 (Jonsson [Jon96], Pietruska-Pałuba [Pie00], Grigor'yan et al. [GHL03]) *Let p_t be a heat kernel on a metric measure space (M, d, μ). If the heat kernel is stochastically complete and satisfies* (2.4), *where $\Phi_1(s) > 0$ for some $s > 0$ and*

$$\int_0^\infty s^{\alpha+\beta+\varepsilon} \Phi_2(s) \frac{ds}{s} < \infty \tag{2.7}$$

for some $\varepsilon > 0$, then $\beta = \beta^*$.

By Theorem 2.1, condition (2.7) implies that the Dirichlet form $(\mathcal{E}, \mathcal{F})$ is local. For non-local forms the statement is not true: for example, in $\mathbb{R}^n$ for symmetric stable processes we have $\beta < 2 = \beta^*$.

Theorem 2.8 *Under the hypotheses of Theorem 2.7, the values of the parameters α and β are the invariants of the metric space (M, d) alone. Moreover, we have*

$$\mu \asymp H^\alpha \quad \text{and} \quad \mathcal{E} \asymp N_{2,\infty}^{\alpha,\beta/2}.$$

Consequently, both the measure μ and the energy form $\mathcal{E}$ are determined (up to a factor $\asymp 1$) by the metric space (M, d) alone.

Example 2.9 Consider in $\mathbb{R}^n$ the Gauss-Weierstrass heat kernel

$$p_t(x, y) = \frac{1}{(4\pi t)^{n/2}} \exp\left(-\frac{|x-y|^2}{4t}\right)$$

and its generator $\mathcal{L} = -\Delta$ in $L^2(\mathbb{R}^n)$ with the Lebesgue measure. Then $\alpha = n$, $\beta = 2$, and

$$\mathcal{E}(u) = \int\limits_{\mathbb{R}^n} |\nabla u|^2 \, dx.$$

Consider now another elliptic operator in $\mathbb{R}^n$:

$$\mathcal{L} = -\frac{1}{m(x)} \sum_{i,j=1}^n \frac{\partial}{\partial x_i} \left(a_{ij}(x) \frac{\partial}{\partial x_j} \right),$$

where $m(x)$ and $a_{ij}(x)$ are continuous functions, $m(x) > 0$ and the matrix $\left(a_{ij}(x) \right)$ is positive definite. The operator $\mathcal{L}$ is symmetric with respect to measure

$$d\mu = m(x) \, dx,$$

and its Dirichlet form is

$$\mathcal{E}(u) = \int\limits_{\mathbb{R}^n} a_{ij}(x) \frac{\partial u}{\partial x_i} \frac{\partial u}{\partial x_j} dx.$$

Let $d(x,y) = |x-y|$ and assume that the heat kernel $p_t(x,y)$ of $\mathcal{L}$ satisfies the conditions of Theorem 2.7. Then we conclude by Corollary 2.8 that α and β must be the same as in the Gauss-Weierstrass heat kernel, that is, $\alpha = n$ and $\beta = 2$; moreover, measure μ must be comparable to the Lebesgue measure, which implies that $m \asymp 1$, and the energy form must admit the estimate

$$\mathcal{E}(u) \asymp \int\limits_{\mathbb{R}^n} |\nabla u|^2 \, dx,$$

which implies that the matrix $\left(a_{ij}(x) \right)$ is uniformly elliptic. Hence, the operator $\mathcal{L}$ is uniformly elliptic.

By Aronson's theorem [Aro67, Aro68] the heat kernel for uniformly elliptic operators satisfies the estimate

$$p_t(x,y) \asymp \frac{C}{t^{n/2}} \exp\left(-c\frac{|x-y|^2}{t} \right).$$

What we have proved here implies the converse to Aronson's theorem: if the Aronson estimate holds for the operator $\mathcal{L}$, then $\mathcal{L}$ is uniformly elliptic.

The next theorem handles the non-local case.

Theorem 2.10 *Let p_t be a heat kernel on a metric measure space (M, d, μ). If the heat kernel satisfies the lower bound*

$$p_t(x, y) \geq \frac{1}{t^{\alpha/\beta}} \Phi_1\left(\frac{d(x, y)}{t^{1/\beta}}\right),$$

where $\Phi_1(s) > 0$ *for some* $s > 0$, *then* $\beta \leq \beta^*$.

Proof In the proof of Theorem 2.3 one shows that the lower bound of the heat kernel implies $\mathcal{F} \subset \Lambda_{2,\infty}^{\alpha,\beta/2}$ (and the opposite inclusion follows from the upper bound and the stochastic completeness). Since $\mathcal{F}$ is dense in L^2, it follows that $\beta \leq \beta^*$. ∎

As a conclusion of this part, we briefly explain the walk dimension from three different points of view. As we have seen, there is a parameter appears in three different places:

- A parameter β in heat kernel bounds (2.4).
- A parameter θ in Markov processes: for a process X_t, one may have (cf. [Bar98, formula (1.1)])

$$\mathbb{E}_x\left(|X_t - x|^2\right) \asymp t^{2/\theta}.$$

Then θ is a parameter that measures how fast the process X_t goes away from the starting point x in time t. Alternatively, one may have that, for any ball $B(x, r) \subset M$,

$$\mathbb{E}_x\left(\tau_{B(x,r)}\right) \asymp r^\theta,$$

where $\tau_{B(x,r)}$ is the *first exit time* of X_t from the ball

$$\tau_B = \inf\{t > 0 : X_t \notin B(x, r)\}.$$

- A parameter σ in function spaces $N_{2,\infty}^{\alpha,\sigma}$ or $N_{2,2}^{\alpha,\sigma}$. By (2.2) or by (2.3), it is not hard to see that σ measures how much smooth of the functions in the space $N_{2,\infty}^{\alpha,\sigma}$ or $N_{2,2}^{\alpha,\sigma}$.

In general the three parameters $\beta, \theta, 2\sigma$ may be different. However, it turns out that, under some certain conditions, all these parameters are the same:

$$\beta = \theta = 2\sigma. \tag{2.8}$$

For examples, by Theorems 2.3 and 2.4, we see that $\sigma = \frac{\beta}{2}$, whilst by Theorems 3.8 and 4.3 below, we will see that $\beta = \theta$.

2.6 Inequalities for the Walk Dimension

Definition 2.11 We say that a metric space (M, d) satisfies the *chain condition* if there exists a (large) constant C such that for any two points $x, y \in M$ and for any positive integer n there exists a sequence $\{x_i\}_{i=0}^{n}$ of points in M such that $x_0 = x$, $x_n = y$, and

$$d(x_i, x_{i+1}) \leq C \frac{d(x, y)}{n}, \quad \text{for all } i = 0, 1, \ldots, n - 1. \tag{2.9}$$

The sequence $\{x_i\}_{i=0}^{n}$ is referred to as a *chain* connecting x and y.

Theorem 2.12 (Grigor'yan et al. [GHL03]) *Let (M, d, μ) be a metric measure space. Assume that*

$$\mu(B(x, r)) \asymp r^{\alpha} \tag{2.10}$$

for all $x \in M$ and $0 < r \leq 1$. Then

$$\beta^* \geq 2.$$

If in addition (M, d) satisfies the chain condition, then

$$\beta^* \leq \alpha + 1.$$

Observe that the chain condition is essential for the inequality $\beta^* \leq \alpha + 1$ to be true. Indeed, assume for a moment that the claim of Theorem 2.12 holds without the chain condition, and consider a new metric d' on M given by $d' = d^{1/\gamma}$ where $\gamma > 1$. Let us mark by a dash all notions related to the space (M, d', μ) as opposed to those of (M, d, μ). It is easy to see that $\alpha' = \alpha\gamma$ and $\beta^{*'} = \beta^*\gamma$. Hence, if Theorem 2.12 could apply to the space (M, d', μ) it would yield $\beta^*\gamma \leq \alpha\gamma + 1$ which implies $\beta^* \leq \alpha$ because γ may be taken arbitrarily large. However, there are spaces with $\beta^* > \alpha$, for example on SG.

Clearly, the metric d' does not satisfy the chain condition; indeed the inequality (2.9) implies

$$d'(x_i, x_{i+1}) \leq C \frac{d'(x, y)}{n^{1/\gamma}},$$

which is not good enough. Note that if in the inequality (2.9) we replace n by $n^{1/\gamma}$, then the proof of Theorem 2.12 will give that $\beta^* \leq \alpha + \gamma$ instead of $\beta^* \leq \alpha + 1$.

Theorem 2.13 (Grigor'yan et al. [GHL03]) *Let p_t be a stochastically complete heat kernel on (M, d, μ) such that*

$$p_t(x, y) \asymp \frac{C}{t^{\alpha/\beta}} \Phi\left(c \frac{d(x, y)}{t^{1/\beta}}\right).$$

(a) If for some $\varepsilon > 0$

$$\int_0^\infty s^{\alpha+\beta+\varepsilon} \Phi(s) \frac{ds}{s} < \infty, \qquad (2.11)$$

then $\beta \geq 2$.

(b) If (M, d) satisfies the chain condition, then $\beta \leq \alpha + 1$.

Proof By Theorem 2.2 μ is α-regular so that Theorem 2.12 applies.

(*a*) By Theorem 2.12, $\beta^* \geq 2$, and by Theorem 2.12, $\beta = \beta^*$, whence $\beta \geq 2$.

(*b*) By Theorem 2.12, $\beta^* \leq \alpha + 1$, and by Theorem 2.10, $\beta \leq \beta^*$, whence $\beta \leq \alpha + 1$. ∎

Note that the condition (2.11) can occur only for a local Dirichlet form $\mathcal{E}$. If both (2.11) and the chain condition are satisfied, then we obtain

$$2 \leq \beta \leq \alpha + 1. \qquad (2.12)$$

This inequality was stated by Barlow [Bar98] without proof.

The set of couples (α, β) satisfying (2.12) is shown on the diagram (Fig. 5):

Barlow [Bar04] proved that any couple of α, β satisfying (2.12) can be realized for the heat kernel estimate

$$p_t(x, y) \asymp \frac{C}{t^{\alpha/\beta}} \exp\left(-c\left(\frac{d^\beta(x, y)}{t}\right)^{\frac{1}{\beta-1}}\right) \qquad (2.13)$$

For a non-local form, we can only claim that

$$0 < \beta \leq \alpha + 1$$

(under the chain condition). In fact, any couple α, β in the range

$$0 < \beta < \alpha + 1$$

can be realized for the estimate

$$p_t(x, y) \asymp \frac{1}{t^{\alpha/\beta'}} \left(1 + \frac{d(x, y)}{t^{1/\beta'}}\right)^{-(\alpha+\beta')}.$$

Indeed, if $\mathcal{L}$ is the generator of a diffusion with parameters α and β satisfying (2.13), then the operator $\mathcal{L}^\delta$, $\delta \in (0, 1)$, generates a jump process with the walk dimension

Fig. 5 The set $2 \leq \beta \leq \alpha + 1$

$\beta' = \delta\beta$ and the same α (cf. Theorem 2.5). Clearly, β' can take any value from $(0, \alpha + 1)$.

It is not known whether the walk dimension for a non-local form can be equal to $\alpha + 1$.

2.7 Identifying Φ in the Local Case

Theorem 2.14 (Grigor'yan and Kumagai [GK08]) *Assume that the metric space* (M, d) *satisfies the chain condition and all metric balls are precompact. Let* p_t *be a stochastically complete heat kernel in* (M, d, μ). *Assume that the associated Dirichlet form* $(\mathcal{E}, \mathcal{F})$ *is regular, and the following estimate holds with some* $\alpha, \beta > 0$ *and* $\Phi : [0, +\infty) \to [0, +\infty)$:

$$p_t(x, y) \asymp \frac{C}{t^{\alpha/\beta}} \Phi\left(c \frac{d(x, y)}{t^{1/\beta}}\right).$$

Then the following dichotomy holds:

- *either the Dirichlet form* $\mathcal{E}$ *is local,* $2 \leq \beta \leq \alpha + 1$, *and* $\Phi(s) \asymp C \exp(-c s^{\frac{\beta}{\beta-1}})$.
- *or the Dirichlet form* $\mathcal{E}$ *is non-local,* $\beta \leq \alpha + 1$, *and* $\Phi(s) \asymp (1 + s)^{-(\alpha+\beta)}$.

3 Sub-Gaussian Upper Bounds

3.1 Ultracontractive Semigroups

Let (M, d, μ) be a metric measure space and $(\mathcal{E}, \mathcal{F})$ be a Dirichlet form in $L^2(M, \mu)$, and let $\{P_t\}$ be the associated heat semigroup, $P_t = e^{-t\mathcal{L}}$ where $\mathcal{L}$ is the generator of $(\mathcal{E}, \mathcal{F})$. The question to be discussed here is whether P_t possesses the heat kernel, that is, a function $p_t(x, y)$ that is non-negative, jointly measurable in (x, y), and satisfies the identity

$$P_t f(x) = \int_M p_t(x, y) f(y) \, d\mu(y)$$

for all $f \in L^2$, $t > 0$, and almost all $x \in M$. Usually the conditions that ensure the existence of the heat kernel give at the same token some upper bounds.

Given two parameters $p, q \in [0, +\infty]$, define the $L^p \to L^q$ norm of P_t by

$$\|P_t\|_{L^p \to L^q} = \sup_{f \in L^p \cap L^2 \setminus \{0\}} \frac{\|P_t f\|_q}{\|f\|_p}.$$

In fact, the Markovian property allows to extend P_t to an operator in L^p so that the range $L^p \cap L^2$ of f can be replaced by L^p. Also, it follows from the Markovian property that $\|P_t\|_{L^p \to L^p} \leq 1$ for any p.

Definition 3.1 The semigroup $\{P_t\}$ is said to be $L^p \to L^q$ *ultracontractive* if there exists a positive decreasing function γ on $(0, +\infty)$, called the *rate* function, such that, for each $t > 0$

$$\|P_t\|_{L^p \to L^q} \leq \gamma(t).$$

By the symmetry of P_t, if P_t is $L^p \to L^q$ ultracontractive, then P_t is also $L^{q^*} \to L^{p^*}$ ultracontractive with the same rate function, where p^* and q^* are the Hölder conjugates to p and q, respectively. In particular, P_t is $L^1 \to L^2$ ultracontractive if and only if it is $L^2 \to L^\infty$ ultracontractive.

Theorem 3.2 (a) *The heat semigroup $\{P_t\}$ is $L^1 \to L^2$ ultracontractive with a rate function γ, if and only if $\{P_t\}$ has the heat kernel p_t satisfying the estimate*

$$\operatorname*{esup}_{x, y \in M} p_t(x, y) \leq \gamma(t/2)^2 \quad \text{for all } t > 0.$$

(b) *The heat semigroup $\{P_t\}$ is $L^1 \to L^\infty$ ultracontractive with a rate function γ, if and only if $\{P_t\}$ has the heat kernel p_t satisfying the estimate*

$$\operatorname*{esup}_{x, y \in M} p_t(x, y) \leq \gamma(t) \quad \text{for all } t > 0.$$

This result is "well-known" and can be found in many sources. However, there are hardly complete proofs of the measurability of the function $p_t(x, y)$ in (x, y), which is necessary for many applications, for example, to use Fubini. Normally the existence of the heat kernel is proved in some specific setting where $p_t(x, y)$ is continuous in (x, y), or one just proves the existence of a family of functions $p_{t,x} \in L^2$ so that

$$P_t f(x) = (p_{t,x}, f) = \int_M p_{t,x}(y) f(y) \, d\mu(y)$$

for all $t > 0$ and almost all x. However, if one defines $p_t(x, y) = p_{t,x}(y)$, then this function does not have to be jointly measurable. The proof of the existence of a jointly measurable version can be found in [GH10]. Most of the material of this section can also be found there.

3.2 Restriction of the Dirichlet Form

Let Ω be an open subset of M. Define the function space $\mathcal{F}(\Omega)$ by

$$\mathcal{F}(\Omega) = \overline{\{f \in \mathcal{F} : \operatorname{supp} f \subset \Omega\}}^{\mathcal{F}}.$$

Clearly, $\mathcal{F}(\Omega)$ is a closed subspace of $\mathcal{F}$ and a subspace of $L^2(\Omega)$.

Theorem 3.3 *If $(\mathcal{E}, \mathcal{F})$ is a regular Dirichlet form in $L^2(M)$, then $(\mathcal{E}, \mathcal{F}(\Omega))$ is a regular Dirichlet form in $L^2(\Omega)$. If $(\mathcal{E}, \mathcal{F})$ is (strongly) local then so is $(\mathcal{E}, \mathcal{F}(\Omega))$.*

The regularity is used, in particular, to ensure that $\mathcal{F}(\Omega)$ is dense in $L^2(\Omega)$. From now on let us assume that $(\mathcal{E}, \mathcal{F})$ is a regular Dirichlet form. Other consequences of this assumptions are as follows (cf. [FOT11]):

1. The existence of cutoff functions: for any compact set K and any open set $U \supset K$, there is a function $\varphi \in \mathcal{F} \cap C_0(U)$ such that $0 \leq \varphi \leq 1$ and $\varphi \equiv 1$ in an open neighborhood of K.
2. The existence of a Hunt process $\left(\{X_t\}_{t \geq 0}, \{\mathbb{P}_x\}_{x \in M}\right)$ associated with $(\mathcal{E}, \mathcal{F})$.

Hence, for any open subset $\Omega \subset M$, we have the Dirichlet form $(\mathcal{E}, \mathcal{F}(\Omega))$ that is called a *restriction* of $(\mathcal{E}, \mathcal{F})$ to Ω.

Example 3.4 Consider in $\mathbb{R}^n$ the canonical Dirichlet form

$$\mathcal{E}(u) = \int_{\mathbb{R}^n} |\nabla u|^2 \, dx$$

in $\mathcal{F} = W_2^1(\mathbb{R}^n)$. Then $\mathcal{F}(\Omega) = \overline{C_0^1(\Omega)}^{W_2^1} =: H_0^1(\Omega)$.

Using the restricted form $(\mathcal{E}, \mathcal{F}(\Omega))$ corresponds to imposing the Dirichlet boundary conditions on $\partial\Omega$ (or on Ω^c), so that the form $(\mathcal{E}, \mathcal{F}(\Omega))$ could be called the Dirichlet form with the Dirichlet boundary condition.

Denote by $\mathcal{L}_\Omega$ the generator of $(\mathcal{E}, \mathcal{F}(\Omega))$ and set

$$\lambda_{\min}(\Omega) := \inf \operatorname{spec}\mathcal{L}_\Omega = \inf_{u\in\mathcal{F}(\Omega)\setminus\{0\}} \frac{\mathcal{E}(u)}{\|u\|_2^2}. \tag{3.1}$$

Clearly, $\lambda_{\min}(\Omega) \geq 0$ and $\lambda_{\min}(\Omega)$ is decreasing when Ω expands.

Example 3.5 If $(\mathcal{E}, \mathcal{F})$ is the canonical Dirichlet form in $\mathbb{R}^n$ and Ω is the bounded domain in $\mathbb{R}^n$, then the operator $\mathcal{L}_\Omega$ has the discrete spectrum $\lambda_1(\Omega) \leq \lambda_2(\Omega) \leq \lambda_3(\Omega) \leq \ldots$ that coincides with the eigenvalues of the Dirichlet problem

$$\begin{cases} \Delta u + \lambda u = 0, \\ u|_{\partial\Omega} = 0, \end{cases}$$

so that $\lambda_1(\Omega) = \lambda_{\min}(\Omega)$.

3.3 Faber-Krahn and Nash Inequalities

Continuing the above example, we have by a theorem of Faber-Krahn

$$\lambda_1(\Omega) \geq \lambda_1(\Omega^*),$$

where Ω^* is the ball of the same volume as Ω. If r is the radius of Ω^*, then we have

$$\lambda_1(\Omega^*) = \frac{c'}{r^2} = \frac{c}{|\Omega^*|^{2/n}} = \frac{c}{|\Omega|^{2/n}},$$

whence

$$\lambda_1(\Omega) \geq c_n |\Omega|^{-2/n}.$$

It turns out that this inequality, that we call *the Faber-Krahn inequality*, is intimately related to the existence of the heat kernel and its upper bound.

Theorem 3.6 *Let $(\mathcal{E}, \mathcal{F})$ be a regular Dirichlet form in $L^2(M, \mu)$. Fix some constant $\nu > 0$. Then the following conditions are equivalent:*

(i) (The Faber-Krahn inequality) There is a constant $a > 0$ such that, for all nonempty open sets $\Omega \subset M$,

$$\lambda_{\min}(\Omega) \geq a\mu(\Omega)^{-\nu}. \tag{3.2}$$

*(ii) (*The Nash inequality*) There exists a constant $b > 0$ such that*

$$\mathcal{E}(u) \geq b\|u\|_2^{2+2\nu}\|u\|_1^{-2\nu}, \tag{3.3}$$

for any function $u \in \mathcal{F} \setminus \{0\}$.

*(iii) (*On-diagonal estimate of the heat kernel*) The heat kernel exists and satisfies the upper bound*

$$\operatorname*{esup}_{x,y \in M} p_t(x, y) \leq ct^{-1/\nu} \tag{3.4}$$

for some constant c and for all $t > 0$.

The relation between the parameters a, b, c is as follows:

$$a \asymp b \asymp c^{-\nu}$$

where the ratio of any two of these parameters is bounded by constants depending only on ν.

In $\mathbb{R}^n$, we see that $\nu = 2/n$.

The implication (ii) $\Rightarrow$ (iii) was proved by Nash [Nas58], and (iii) $\Rightarrow$ (ii) by Carlen-Kusuoka-Stroock [CKS87], and (i) $\Leftrightarrow$ (iii) by

Grigor'yan [Gri94] and Carron [Car96].

Proof of (i) $\Rightarrow$ (ii) $\Rightarrow$ (iii) Observe first that (ii) $\Rightarrow$ (i) is trivial: choosing in (3.3) a function $u \in \mathcal{F}(\Omega) \setminus \{0\}$ and applying the Cauchy-Schwarz inequality

$$\|u\|_1 \leq \mu(\Omega)^{1/2}\|u\|_2,$$

we obtain

$$\mathcal{E}(u) \geq b\mu(\Omega)^{-\nu}\|u\|_2^2,$$

whence (3.2) follow by the variational principle (3.1).

The opposite inequality (i) $\Rightarrow$ (ii) is a bit more involved, and we prove it for functions $0 \leq u \in \mathcal{F} \cap C_0(M)$ (a general $u \in \mathcal{F}$ requires some approximation argument). By the Markovian property, we have $(u - t)_+ \in \mathcal{F} \cap C_0(M)$ for any $t > 0$ and

$$\mathcal{E}(u) \geq \mathcal{E}\big((u - t)_+\big). \tag{3.5}$$

For any $s > 0$, consider the set

$$U_s := \{x \in M : u(x) > s\},$$

which is clearly open and precompact. If $t > s$, then $(u - t)_+$ is supported in U_s, and whence, $(u - t)_+ \in \mathcal{F}(U_s)$. It follows from (3.1)

$$\mathcal{E}\left((u-t)_+\right) \geq \lambda_{\min}\left(U_s\right) \int_{U_s} (u-t)_+^2 \, d\mu. \tag{3.6}$$

For simplicity, set $A = \|u\|_1$ and $B = \|u\|_2^2$. Since $u \geq 0$, we have

$$(u-t)_+^2 \geq u^2 - 2tu,$$

which implies that

$$\int_{U_s} (u-t)_+^2 \, d\mu = \int_M (u-t)_+^2 \, d\mu \geq B - 2tA. \tag{3.7}$$

On the other hand, we have

$$\mu(U_s) \leq \frac{1}{s} \int_{U_s} u \, d\mu \leq \frac{A}{s},$$

which together with the Faber-Krahn inequality implies

$$\lambda_{\min}\left(U_s\right) \geq a\mu\left(U_s\right)^{-\nu} \geq a\left(\frac{s}{A}\right)^{\nu}. \tag{3.8}$$

Combining (3.5)–(3.8), we obtain

$$\mathcal{E}\left(u\right) \geq \lambda_{\min}\left(U_s\right) \int_{U_s} (u-t)_+^2 \, d\mu \geq a\left(\frac{s}{A}\right)^{\nu} (B - 2tA).$$

Letting $t \to s+$ and then choosing $s = \frac{B}{4A}$, we obtain

$$\mathcal{E}\left(u\right) \geq a\left(\frac{s}{A}\right)^{\nu} (B - 2sA) = a\left(\frac{B}{4A^2}\right)^{\nu} \frac{B}{2} = \frac{a}{4^{\nu}2} B^{\nu+1} A^{-2\nu},$$

which is exactly (3.3).

To prove (ii) $\Rightarrow$ (iii), choose $f \in L^2 \cap L^1$, and consider $u = P_t f$. Since $u = e^{-t\mathcal{L}} f$ and $\frac{d}{dt} u = -\mathcal{L}u$, we have

$$\frac{d}{dt} \|u\|_2^2 = \frac{d}{dt} (u, u) = -2\left(\mathcal{L}u, u\right) = -2\mathcal{E}\left(u, u\right)$$
$$\leq -2b\|u\|_2^{2+2\nu} \|u\|_1^{-2\nu} \leq -2b\|u\|_2^{2+2\nu} \|f\|_1^{-2\nu},$$

since $\|u\|_1 \leq \|f\|_1$. Solving this differential inequality, we obtain

$$\| P_t f \|_2^2 \leq c t^{-1/v} \| f \|_1^2 ,$$

that is, the semigroup P_t is $L^1 \to L^2$ ultracontractive with the rate function $\gamma (t) = \sqrt{c t^{-1/v}}$. By Theorem 3.2 we conclude that the heat kernel exists and satisfies (3.4). ∎

Let M be a Riemannian manifold with the geodesic distance d and the Riemannian volume μ. Let $(\mathcal{E}, \mathcal{F})$ be the canonical Dirichlet form on M. The heat kernel on manifolds always exists and is a smooth function. In this case the estimate (3.4) is equivalent to the on-diagonal upper bound

$$\sup_{x \in M} p_t (x, x) \leq c t^{-1/\nu}.$$

It is known (but non-trivial) that the on-diagonal estimate implies the Gaussian upper bound

$$p_t (x, y) \leq C t^{-1/\nu} \exp \left(- \frac{d^2 (x, y)}{(4 + \varepsilon) t} \right),$$

for all $t > 0$ and $x, y \in M$, which is due to the specific property of the geodesic distance function that $|\nabla d| \leq 1$.

In the context of abstract metric measure space, the distance function does not have to satisfy this property, and typically it does not (say, on fractals). Consequently, one needs some additional conditions that would relate the distance function to the Dirichlet form and imply the off-diagonal bounds.

3.4 Off-diagonal Upper Bounds

From now on, let $(\mathcal{E}, \mathcal{F})$ be a regular *local* Dirichlet form, so that the associated Hunt process $\left(\{X_t\}_{t \geq 0} , \{\mathbb{P}_x\}_{x \in M} \right)$ is a diffusion. Recall that it is related to the heat semigroup $\{P_t\}$ of $(\mathcal{E}, \mathcal{F})$ by means of the identity

$$\mathbb{E}_x (f (X_t)) = P_t f (x)$$

for all $f \in \mathcal{B}_b (M)$, $t > 0$ and almost all $x \in M$ (Fig. 6).

Fix two parameters $\alpha > 0$ and $\beta > 1$ and introduce some conditions.

(V_α) (Volume regularity) For all $x \in M$ and $r > 0$,

$$\mu (B (x, r)) \asymp r^\alpha.$$

(FK) (The Faber-Krahn inequality) For any open set $\Omega \subset M$,

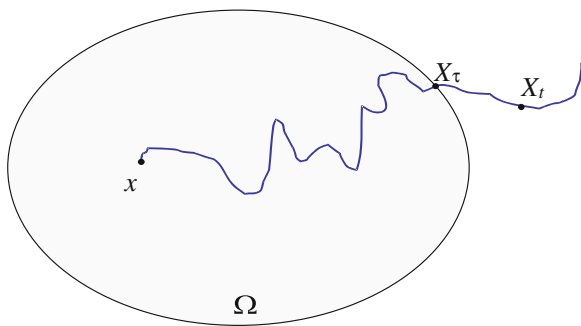$$\lambda_{\min} (\Omega) \geq c \mu (\Omega)^{-\beta/\alpha}.$$

Fig. 6 First exit time τ

For any open set $\Omega \subset M$, define the *first exist time* from Ω by

$$\tau_\Omega = \inf\{t > 0 : X_t \notin \Omega\}.$$

A set $N \subset M$ is called *properly exceptional*, if it is a Borel set of measure 0 that is almost never hit by the process X_t starting outside N. In the next conditions N denotes some properly exceptional set.

(E_β) (*An estimate for the mean exit time from balls*) For all $x \in M \setminus N$ and $r > 0$

$$\mathbb{E}_x\left[\tau_{B(x,r)}\right] \asymp r^\beta$$

(the parameter β is called the walk dimension of the process).
(P_β) (*The exit probability estimate*) There exist constants $\varepsilon \in (0, 1)$, $\delta > 0$ such that, for all $x \in M \setminus N$ and $r > 0$,

$$\mathbb{P}_x\left(\tau_{B(x,r)} \leq \delta r^\beta\right) \leq \varepsilon.$$

$(E\Omega)$ (*An isoperimetric estimate for the mean exit time*) For any open subset $\Omega \subset M$,
$$\sup_{x \in \Omega \setminus N} \mathbb{E}_x(\tau_\Omega) \leq C \mu(\Omega)^{\beta/\alpha}.$$

If both (V_α) and (E_β) are satisfied, then we obtain for any ball $B \subset M$

$$\sup_{x \in B \setminus N} \mathbb{E}_x(\tau_B) \asymp r^\beta \asymp \mu(B)^{\beta/\alpha}.$$

It follows that the balls are in some sense optimal sets for the condition $(E\Omega)$.

Example 3.7 If X_t is Brownian motion in $\mathbb{R}^n$, then it is known that

$$\mathbb{E}_x \tau_{B(x,r)} = c_n r^2,$$

so that (E_β) holds with $\beta = 2$. This can also be rewritten in the form

$$\mathbb{E}_x \tau_B = c_n \, |B|^{2/n},$$

where $B = B(x, r)$.

It is also known that for any open set $\Omega \subset \mathbb{R}^n$ with finite volume and for any $x \in \Omega$,

$$\mathbb{E}_x (\tau_\Omega) \leq \mathbb{E}_x (\tau_{B(x,r)}),$$

provided that ball $B(x, r)$ has the same volume as Ω; that is, for a fixed value of $|\Omega|$, the mean exist time is maximal when Ω is a ball and x is the center. It follows that

$$\mathbb{E}_x (\tau_\Omega) \leq c_n \, |\Omega|^{2/n}$$

so that $(E\Omega)$ is satisfied with $\beta = 2$ and $\alpha = n$.

Finally, introduce notation for the following estimates of the heat kernel:

$(U E_{loc})$ (*Sub-Gaussian upper estimate*) The heat kernel exists and satisfies the estimate

$$p_t(x, y) \leq \frac{C}{t^{\alpha/\beta}} \exp\left(-c \left(\frac{d^\beta(x, y)}{t} \right)^{\frac{1}{\beta-1}} \right)$$

for all $t > 0$ and almost all $x, y \in M$.

$(\Phi U E)$ (Φ-*upper estimate*) The heat kernel exists and satisfies the estimate

$$p_t(x, y) \leq \frac{1}{t^{\alpha/\beta}} \Phi\left(\frac{d(x, y)}{t^{1/\beta}} \right)$$

for all $t > 0$ and almost all $x, y \in M$, where Φ is a decreasing positive function on $[0, +\infty)$ such that

$$\int_0^\infty s^\alpha \Phi(s) \, \frac{ds}{s} < \infty.$$

(DUE) (*On-diagonal upper estimate*) The heat kernel exists and satisfies the estimate

$$p_t(x, y) \leq \frac{C}{t^{\alpha/\beta}}$$

for all $t > 0$ and almost all $x, y \in M$.

$(T_{\exp})$ (*The exponential tail estimate*) The heat kernel p_t exists and satisfies the estimate

$$\int\limits_{B(x,r)^c} p_t(x, y) \, d\mu(y) \leq C \exp\left(-c\left(\frac{r}{t^{1/\beta}}\right)^{\frac{\beta}{\beta-1}}\right), \tag{3.9}$$

for some constants $C, c > 0$, all $t > 0, r > 0$ and μ-almost all $x \in M$.

Note that it is easy to show that (3.9) is equivalent to the following inequality: for any ball $B = B(x_0, r)$ and $t > 0$,

$$P_t \mathbf{1}_{B^c}(x) \leq C \exp\left(-c\left(\frac{r}{t^{1/\beta}}\right)^{\frac{\beta}{\beta-1}}\right) \text{ for } \mu\text{-almost all } x \in \frac{1}{4}B$$

(see [GH08, Remark 3.3]).

(T_β) (*The tail estimate*) There exist $0 < \varepsilon < \frac{1}{2}$ and $C > 0$ such that, for all $t > 0$ and all balls $B = B(x_0, r)$ with $r \geq Ct^{1/\beta}$,

$$P_t \mathbf{1}_{B^c}(x) \leq \varepsilon \quad \text{for } \mu\text{-almost all } x \in \frac{1}{4}B.$$

(S_β) (*The survival estimate*) There exist $0 < \varepsilon < 1$ and $C > 0$ such that, for all $t > 0$ and all balls $B = B(x_0, r)$ with $r \geq Ct^{1/\beta}$,

$$1 - P_t^B \mathbf{1}_B(x) \leq \varepsilon \quad \text{for } \mu\text{-almost all } x \in \frac{1}{4}B.$$

Clearly, we have

$$(UE_{loc}) \Rightarrow (\Phi UE) \Rightarrow (DUE).$$

Theorem 3.8 (Grigor'yan and Hu [GH10]) *Let (M, d, μ) be a metric measure space and let (V_α) hold. Let $(\mathcal{E}, \mathcal{F})$ be a regular, local, conservative Dirichlet form in $L^2(M, \mu)$. Then, the following equivalences are true:*

$$\begin{aligned}
(UE_{loc}) &\Leftrightarrow (FK) + (E_\beta) \Leftrightarrow (E\Omega) + (E_\beta) \\
&\Leftrightarrow (FK) + (P_\beta) \Leftrightarrow (E\Omega) + (P_\beta) \\
&\Leftrightarrow (DUE) + (E_\beta) \Leftrightarrow (DUE) + (P_\beta), \\
&\Leftrightarrow (\Phi UE) \\
&\Leftrightarrow (FK) + (S_\beta) \Leftrightarrow (FK) + (T_\beta) \\
&\Leftrightarrow (DUE) + (S_\beta) \Leftrightarrow (DUE) + (T_\beta) \\
&\Leftrightarrow (DUE) + (T_{\exp}).
\end{aligned}$$

Let us emphasize the equivalence

$$(UE_{loc}) \Leftrightarrow (E\Omega) + \left(E_\beta\right)$$

where the right hand side means the following: the mean exit time from all sets Ω satisfies the isoperimetric inequality, and this inequality is optimal for balls (up to a constant multiple). Note that the latter condition relates the properties of the diffusion (and, hence, of the Dirichlet form) to the distance function.

Conjecture 3.9 *Under the hypotheses of Theorem 3.8,*

$$(UE_{loc}) \Leftrightarrow (FK) + \left\{\lambda_{\min}(B_r) \asymp r^{-\beta}\right\}$$

Indeed, the Faber-Krahn inequality (*FK*) can be regarded as an isoperimetric inequality for $\lambda_{\min}(\Omega)$, and the condition

$$\lambda_{\min}(B_r) \asymp r^{-\beta}$$

means that (*FK*) is optimal for balls (up to a constant multiple).

Theorem 3.8 is an oversimplified version of a result of [GH10], where instead of (V_α) one uses the volume doubling condition, and other hypotheses must be appropriately changed.

The following lemma is used in the proof of Theorem 3.8.

Lemma 3.10 *For any open set $\Omega \subset M$*

$$\lambda_{\min}(\Omega) \geq \frac{1}{\mathrm{esup}_{x \in \Omega} \mathbb{E}_x(\tau_\Omega)}.$$

Proof Let G_Ω be the Green operator in Ω, that is,

$$G_\Omega = \mathcal{L}_\Omega^{-1} = \int_0^\infty e^{-t\mathcal{L}_\Omega} dt.$$

We claim that

$$\mathbb{E}_x(\tau_\Omega) = G_\Omega 1(x)$$

for almost all $x \in \Omega$. We have

$$G_\Omega 1(x) = \int_0^\infty e^{-t\mathcal{L}_\Omega} 1_\Omega(x)\, dt = \int_0^\infty \mathbb{E}_x\left(1_\Omega\left(X_t^\Omega\right)\right)$$

$$= \int_0^\infty \mathbb{E}_x\left(1_{\{t<\tau_\Omega\}}\right) dt = \mathbb{E}_x \int_0^\infty \left(1_{\{t<\tau_\Omega\}}\right) dt = \mathbb{E}_x(\tau_\Omega).$$

Setting

$$m = \operatorname*{esup}_{x \in \Omega} \mathbb{E}_x (\tau_\Omega) ,$$

we obtain that $G_\Omega 1 \leq m$, so that $m^{-1} G_\Omega$ is a Markovian operator. Therefore, $\|m^{-1} G_\Omega\|_{L^2 \to L^2} \leq 1$ whence $\operatorname{spec} G_\Omega \in [0, m]$. It follows that $\operatorname{spec} \mathcal{L}_\Omega \subset [m^{-1}, \infty)$ and $\lambda_{\min} (\Omega) \geq m^{-1}$. ∎

A new analytical approach is developed in [GH10] to prove Theorem 3.8, which is different from the Davies-Gaffney approach [Dav92] . The difficult part in proving Theorem 3.8 is to deduce (UE_{loc}) from various conditions.

Sketch of proof for Theorem 3.8 We sketch the main steps.

- By a direct integration, we have

$$(\Phi UE) \Rightarrow (T_\beta) .$$

Indeed, for any $x \in \frac{1}{4} B$, we see that $B(x, \frac{1}{2} r) \subset B$. Thus, setting $r_k = 2^k (r/2)$ and using condition (ΦUE) and the monotonicity of Φ, we obtain that

$$
\int_{M \setminus B} p_t(x, y) d\mu(y) \leq \int_{M \setminus B(x, r/2)} p_t(x, y) d\mu(y) \qquad (3.10)
$$

$$
= \sum_{k=0}^{\infty} \int_{B(x, r_{k+1}) \setminus B(x, r_k)} p_t (x, y) d\mu(y)
$$

$$
\leq \sum_{k=0}^{\infty} \int_{B(x, r_{k+1}) \setminus B(x, r_k)} C t^{-\alpha/\beta} \Phi \left(\frac{r_k}{t^{1/\beta}} \right) d\mu(y)
$$

$$
\leq \sum_{k=0}^{\infty} C r_{k+1}^{\alpha} t^{-\alpha/\beta} \Phi \left(\frac{r_k}{t^{1/\beta}} \right)
$$

$$
= C' \sum_{k=0}^{\infty} \left(\frac{2^{k-1} r}{t^{1/\beta}} \right)^{\alpha} \Phi \left(\frac{2^{k-1} r}{t^{1/\beta}} \right)
$$

$$
\leq C' \int_{\frac{1}{4} r / t^{1/\beta}}^{\infty} s^{\alpha} \Phi(s) \frac{ds}{s} .
$$

The integral (3.10) converges, and its value can be made arbitrarily small provided that r^β / t is large enough. Hence, condition (T_β) follows.
- The following implications hold:

$$(E\Omega) \overset{\text{L. 3.10}}{\Rightarrow} (FK) \overset{\text{T. 3.6}}{\Rightarrow} (DUE) .$$

In particular, we see that the heat kernel exists under any of the hypotheses of Theorem 3.8.

- We can also show that

$$\left(E_\beta\right) \Rightarrow \left(P_\beta\right) \Longrightarrow \left(T_\beta\right)$$

(the implication $\left(E_\beta\right) \Rightarrow \left(P_\beta\right)$ was pointed out in [Bar98]).

- By a bootstrapping technique, we obtain (hard!) the implication

$$\left(T_\beta\right) \Longrightarrow \left(T_{\exp}\right)$$

(see also [GH08]). Hence, any set of the hypothesis of Theorem 3.8 imply both (DUE) and $\left(T_{\exp}\right)$.

- Finally, it is easy to check the implication

$$(DUE) + \left(T_{\exp}\right) \Rightarrow (U\,E_{loc})\,. \tag{3.11}$$

Indeed, using the semigroup identity, we have that, for all $t > 0$, almost all $x, y \in M$, and $r := \frac{1}{2}d\,(x, y)$,

$$p_t\,(x, y) = \int_M p_{\frac{t}{2}}\,(x, z)\, p_{\frac{t}{2}}\,(z, y)\, d\mu(z) \tag{3.12}$$

$$\leq \left(\int_{B(x,r)^c} + \int_{B(y,r)^c}\right) p_{\frac{t}{2}}\,(x, z)\, p_{\frac{t}{2}}\,(z, y)\, d\mu(z)$$

$$\leq \operatorname*{esup}_{z \in M} p_{\frac{t}{2}}\,(z, y) \int_{B(x,r)^c} p_{\frac{t}{2}}\,(x, z)\, d\mu(z)$$

$$+ \operatorname*{esup}_{z \in M} p_{\frac{t}{2}}\,(x, z) \int_{B(y,r)^c} p_{\frac{t}{2}}\,(y, z)\, d\mu(z).$$

On the other hand, by condition (DUE),

$$\operatorname{esup} p_t \leq C t^{-\alpha/\beta},$$

whilst by condition $\left(T_\beta\right)$,

$$\int_{B(x,r)^c} p_{\frac{t}{2}}\,(x, z)\, d\mu(z) \leq C \exp\left(-c\left(\frac{d^\beta\,(x, y)}{t}\right)^{\frac{1}{\beta-1}}\right).$$

Therefore, it follows from (3.12) that, for almost all $x, y \in M$,

$$p_t\,(x,\,y) \le \frac{C}{t^{\alpha/\beta}} \exp\left(-c\left(\frac{d^{\beta}\,(x,\,y)}{t}\right)^{\frac{1}{\beta-1}}\right),$$

proving the implication (3.11).   ∎

Recently, Andres and Barlow [AB] gave a new equivalence condition for (UE_{loc}). Consider the following functional inequality.

(CSA_{β}) (*The cutoff Sobolev annulus inequality*) There exists a constant $C > 0$ such that, for all two concentric balls $B(x, R)$, $B(x, R + r)$, there exists a cutoff function φ satisfying

$$\int\limits_U f^2 d\mu_{\langle\varphi\rangle} \le \frac{1}{8}\int\limits_U \varphi^2 d\mu_{\langle f\rangle} + Cr^{-\beta}\int\limits_U f^2 d\mu$$

for any $f \in \mathcal{F}$, where $U = B(x, R + r) \setminus B(x, R)$ is the annulus and $\mu_{\langle\varphi\rangle}$ is the energy measure associated with φ:

$$\int\limits_M u d\mu_{\langle\varphi\rangle} = 2\mathcal{E}(u\varphi, \varphi) - \mathcal{E}(\varphi^2, u) \text{ for any } u \in \mathcal{F} \cap C_0(M).$$

We remark here that constant C is universal that is independent of two concentric balls $B(x, R)$, $B(x, R + r)$ and function f, whilst the cutoff function φ may depend on the balls but is independent of function f. The coefficient $\frac{1}{8}$ is not essential and is chosen for technical reasons.

Theorem 3.11 (Andres, Barlow [AB]) *Let (M, d, μ) be an unbounded metric measure space and let (V_{α}) hold. Let $(\mathcal{E}, \mathcal{F})$ be a regular, local Dirichlet form in $L^2(M, \mu)$. Then, the following equivalence is true:*

$$(UE_{loc}) \Leftrightarrow (FK) + \left(CSA_{\beta}\right).$$

We mention that here the Dirichlet form is not required to be conservative as in Theorem 3.8.

The key point in proving Theorem 3.11 is to derive the "Davies-Gaffney" bound [Dav92], and then use the technique developed in [Gri92, CG98] to show a mean value inequality for weak solutions of the heat equation. It is quite surprising that the Davies-Gaffney method still works when the walk dimension β may be greater than 2.

4 Two-Sided Sub-Gaussian Bounds

4.1 Using Elliptic Harnack Inequality

Now we would like to extend the results of Theorems 3.8, 3.11, and obtain also the lower estimates and the Hölder continuity of the heat kernel. As before, (M, d, μ) is a metric measure space, and assume in addition that all metric balls are precompact. Let $(\mathcal{E}, \mathcal{F})$ is a local regular conservative Dirichlet form in $L^2(M, \mu)$.

Definition 4.1 We say that a function $u \in \mathcal{F}$ is *harmonic* in an open set $\Omega \subset M$ if

$$\mathcal{E}(u, v) = 0 \text{ for all } v \in \mathcal{F}(\Omega).$$

For example, if $M = \mathbb{R}^n$ and $(\mathcal{E}, \mathcal{F})$ is the canonical Dirichlet form in $\mathbb{R}^n$, then a function $u \in W_2^1(\mathbb{R}^n)$ is harmonic in an open set $\Omega \subset \mathbb{R}^n$ if

$$\int_{\mathbb{R}^n} \langle \nabla u, \nabla v \rangle dx = 0$$

for all $v \in H_0^1(\Omega)$ or for $v \in C_0^\infty(\Omega)$. This of course implies that $\Delta u = 0$ in a weak sense in Ω and, hence, u is harmonic in Ω in the classical sense. However, unlike the classical definition, we a priori require $u \in W_2^1(\mathbb{R}^n)$.

Definition 4.2 (*Elliptic Harnack inequality* (H)) We say that M satisfies the *elliptic Harnack inequality* (H) if there exist constants $C > 1$ and $\delta \in (0, 1)$ such that for any ball $B(x, r)$ and for any function $u \in \mathcal{F}$ that is non-negative and harmonic in $B(x, r)$,

$$\underset{B(x, \delta r)}{\text{esup }} u \leq C \underset{B(x, \delta r)}{\text{einf }} u.$$

We remark that constants C and δ are independent of ball $B(x, r)$ and function u. We introduce the near-diagonal lower estimate of heat kernel.

(NLE) (*Near-diagonal lower estimate*) The heat kernel $p_t(x, y)$ exists, and satisfies

$$p_t(x, y) \geq \frac{c}{t^{\alpha/\beta}}$$

for all $t > 0$ and $\mu \times \mu$-almost all $x, y \in M$ such that $d(x, y) \leq \delta t^{1/\beta}$, where $\delta > 0$ is a sufficiently small constant.

Denote by (UE_{strong}) a modification of condition (UE_{loc}) that is obtained by adding the Hölder continuity of $p_t(x, y)$ and by restricting inequality in (UE_{loc}) to all $x, y \in M$. In a similar way, we can define *condition* (NLE_{strong}).

Theorem 4.3 (Grigor'yan, Telcs [GT12, Theorem 7.4]) *Let (M, d, μ) be a metric measure space and let (V_α) hold. Let $(\mathcal{E}, \mathcal{F})$ be a regular, strongly local Dirichlet form in $L^2(M, \mu)$. Then, the following equivalences are true:*

$$(H) + \left(E_\beta\right) \Leftrightarrow (UE_{loc}) + (NLE)$$
$$\Leftrightarrow \left(UE_{strong}\right) + \left(NLE_{strong}\right).$$

This theorem is proved in [GT12] for a more general setting of volume doubling instead of (V_α).

Observe that the following implications hold [GT12, Lemma 7.3]:

$$(H) \Rightarrow (M, d) \text{ is connected,}$$
$$\left(E_\beta\right) \Rightarrow (\mathcal{E}, \mathcal{F}) \text{ is conservative,}$$
$$\left(E_\beta\right) \Rightarrow diam\,(M) = \infty.$$

Proof Sketch of proof for Theorem 4.3 First one shows that

$$(V_\alpha) + \left(E_\beta\right) + (H) \Rightarrow (FK),$$

which is quite involved and uses, in particular, Lemma 3.10. Once having $(V_\alpha) + \left(E_\beta\right) + (FK)$, we obtain (UE_{loc}) by Theorem 3.8.

Using the elliptic Harnack inequality, one obtains in a standard way the oscillating inequality for harmonic functions and then for functions of the form $u = G_\Omega f$ (that solves the equation $\mathcal{L}_\Omega u = f$) in terms of $\|f\|_\infty$.

If now $u = P_t^\Omega f$ then u satisfies the equation

$$\frac{d}{dt}u = -\mathcal{L}_\Omega u,$$

and whence

$$u = -G_\Omega\left(\frac{d}{dt}u\right).$$

Knowing an upper bound for u, which follows from the upper bound of the heat kernel, one obtains also an upper bound for $\frac{d}{dt}u$ in terms of u. Applying the oscillation inequality one obtains the Hölder continuity of u and, hence, of the heat kernel.

Let us prove the on-diagonal lower bound

$$p_t\,(x, x) \geq ct^{-\alpha/\beta}.$$

Note that (UE_{loc}) and (V_α) imply that

$$\int\limits_{B(x,r)} p_t\left(x, y\right) d\mu\left(y\right) \geq \frac{1}{2}$$

provided $r \geq K t^{1/\beta}$ (cf. [GHL03, formula (3.8)]). Choosing $r = K t^{1/\beta}$, we obtain

$$p_{2t}\left(x, x\right) = \int\limits_{M} p_t^2\left(x, y\right) d\mu\left(y\right)$$

$$\geq \frac{1}{\mu\left(B\left(x, r\right)\right)}\left(\int\limits_{B(x,r)} p_t\left(x, y\right) d\mu\left(y\right)\right)^2$$

$$\geq \frac{c}{r^{\alpha}} = \frac{c'}{t^{\alpha/\beta}}.$$

Then (NLE) follows from the upper estimate for

$$\left| p_t\left(x, x\right) - p_t\left(x, y\right)\right|$$

when y close to x, which follows from the oscillation inequality.   ∎

We next characterize $(UE_{loc}) + (NLE)$ by using the estimates of the capacity and of the Green function.

Definition 4.4 *(capacity)* Let Ω be an open set in M and $A \Subset \Omega$ be a Borel set. Define the *capacity* cap(A, Ω) by

$$\text{cap}(A, \Omega) := \inf\left\{\mathcal{E}\left(\varphi\right) : \varphi \text{ is a cutoff function of } (A, \Omega)\right\}. \qquad (4.1)$$

It follows from the definition that the capacity cap(A, Ω) is increasing in A, and decreasing in Ω, namely, if $A_1 \subset A_2$, $\Omega_1 \supset \Omega_2$, then cap$(A_1, \Omega_1) \leq \text{cap}(A_2, \Omega_2)$. Using the latter property, let us extend the definition of capacity when $A \subset \Omega$ as follows:

$$\text{cap}(A, \Omega) = \lim_{n\to\infty} \text{cap}(A \cap \Omega_n, \Omega) \qquad (4.2)$$

where $\{\Omega_n\}$ is any increasing sequence of precompact open subsets of Ω exhausting Ω (in particular, $A \cap \Omega_n \Subset \Omega$).

Note that by the monotonicity property of the capacity, the limit in the right hand side of (4.2) exists (finite or infinite) and is independent of the choice of the exhausting sequence $\{\Omega_n\}$.

Next, define the *resistance* res (A, Ω) by

$$\text{res}\left(A, \Omega\right) = \frac{1}{\text{cap}(A, \Omega)}. \qquad (4.3)$$

We introduce the notions of the Green operator and the Green function.

Definition 4.5 For an open $\Omega \subset M$, a linear operator $G^\Omega : L^2(\Omega) \to \mathcal{F}(\Omega)$ is called a *Green operator* if, for any $\varphi \in \mathcal{F}(\Omega)$ and any $f \in L^2(\Omega)$,

$$\mathcal{E}(G^\Omega f, \varphi) = (f, \varphi). \tag{4.4}$$

If G^Ω admits an integral kernel g^Ω, that is,

$$G^\Omega f(x) = \int_\Omega g^\Omega(x, y) f(y) d\mu(y) \text{ for any } f \in L^2(\Omega), \tag{4.5}$$

then g^Ω is called a *Green function*.

It is known (cf. [GH00, Lemma 5.1]) that if $(\mathcal{E}, \mathcal{F})$ is regular and if $\Omega \subset M$ is open such that $\lambda_{\min}(\Omega) > 0$, then the Green operator G^Ω exists, and in fact, $G^\Omega = (-\mathcal{L}^\Omega)^{-1}$, the inverse of $-\mathcal{L}^\Omega$, where $\mathcal{L}^\Omega$ is the generator of $(\mathcal{E}, \mathcal{F}(\Omega))$. However, the issue of the Green function g^Ω is much more involved, and is one of the key topics in [GH00].

For an open set $\Omega \subset M$, function E^Ω is defined by

$$E^\Omega(x) := G^\Omega \mathbf{1}(x) \ (x \in M), \tag{4.6}$$

namely, the function E^Ω is a unique weak solution of the following Poisson-type equation

$$-\mathcal{L}^\Omega E^\Omega = 1, \tag{4.7}$$

provided that $\lambda_{\min}(\Omega) > 0$.

It is known that

$$E^\Omega(x) = \mathbb{E}_x(\tau_\Omega) \text{ for } \mu\text{-a.a. } x \in M. \tag{4.8}$$

Clearly, if the Green function g^Ω exists, then

$$E^\Omega(x) = G^\Omega \mathbf{1}(x) = \int_\Omega g^\Omega(x, y) d\mu(y) \tag{4.9}$$

for μ-almost all $x \in M$.

We introduce the following hypothesis.

(R_β) (*Resistance condition* (R_β)) We say that the *resistance condition* (R_β) is satisfied if, there exist constants $K, C > 1$ such that, for any ball B of radius $r > 0$,

$$C^{-1} \frac{r^\beta}{\mu(B)} \leq \text{res}(B, KB) \leq C \frac{r^\beta}{\mu(B)}, \tag{4.10}$$

where constants K and C are independent of the ball B. Equivalently, (4.10) can be written in the form

$$\mathrm{res}\,(B,\,KB) \asymp \frac{r^\beta}{\mu\,(B)}.$$

$\left(E'_\beta\right)$ (*Condition* $\left(E'_\beta\right)$) We say that *condition* $\left(E'_\beta\right)$ holds if, there exist two constants $C > 1$ and $\delta_1 \in (0, 1)$ such that, for any ball B of radius $r > 0$,

$$\operatorname*{esup}_B E^B \leq Cr^\beta,$$

$$\operatorname*{einf}_{\delta_1 B} E^B \geq C^{-1} r^\beta.$$

$\left(G_\beta\right)$ (*Condition* $\left(G_\beta\right)$) We say that *condition* $\left(G_\beta\right)$ holds if, there exist constants $K > 1$ and $\dot{C} > 0$ such that, for any ball $B := B\,(x_0,\,R)$, the Green kernel g^B exists and is jointly continuous off the diagonal, and satisfies

$$g^B\,(x_0,\,y) \leq C \int_{K^{-1}d(x_0,y)}^{R} \frac{s^\beta ds}{s\,V\,(x,\,s)} \quad \text{for all } y \in B \setminus \{x_0\},$$

$$g^B\,(x_0,\,y) \geq C^{-1} \int_{K^{-1}d(x_0,y)}^{R} \frac{s^\beta ds}{s\,V\,(x,\,s)} \quad \text{for all } y \in K^{-1}B \setminus \{x_0\},$$

where $V(x,\,r) = \mu(B(x,\,r))$ as before.

Theorem 4.6 (Grigor'yan and Hu) [GH00, Theorem 3.14]) *Let* $(M,\,d,\,\mu)$ *be a metric measure space and let* (V_α) *hold. Let* $(\mathcal{E},\,\mathcal{F})$ *be a regular, strongly local Dirichlet form in* $L^2(M,\,\mu)$. *Then, the following equivalences are true:*

$$(H) + \left(E'_\beta\right) \Leftrightarrow \left(G_\beta\right) \Leftrightarrow (H) + \left(R_\beta\right)$$
$$\Leftrightarrow (UE_{loc}) + (NLE)$$
$$\Leftrightarrow \left(UE_{strong}\right) + \left(NLE_{strong}\right).$$

We mention that condition (V_α) can be replaced by conditions (VD) and (RVD), the latter refers to the reverse doubling condition (cf. [GH00]).

Sketch of proof for Theorem 4.6 The proofs of Theorem 4.6 consists of two parts.

- *Part One*. Firstly, the following implications hold:

$$\boxed{\left(UE_{strong}\right) + \left(NLE_{strong}\right)}$$
$$\Updownarrow$$
$$\boxed{\left(UE_{loc}\right) + \left(NLE\right)}$$
$$\Updownarrow \qquad\qquad \Uparrow$$
$$\boxed{(H) + \left(E_\beta\right)} \Rightarrow \boxed{(H) + \left(E'_\beta\right)}$$

In fact, by Theorem 4.3, we only need to show that

$$\left(E_\beta\right) \Rightarrow \left(E'_\beta\right), \tag{4.11}$$

$$(H) + \left(E'_\beta\right) \Longrightarrow (UE_{loc}) + (NLE). \tag{4.12}$$

The implication (4.11) can be proved directly by using the probability argument, see [GH00, Theorem 3.14]. And the implication (4.12) can be done by showing the following

$$(H) + \left(E'_\beta\right) \Rightarrow (FK) \ \ ([GT12, \text{formula (3.17) and T.3.11}])$$
$$\left(E'_\beta\right) \Rightarrow \left(S_\beta\right) \ \ (\text{by [GHL00, formula (6.34)]})$$
$$(FK) + \left(S_\beta\right) \Rightarrow (UE_{loc}) \ \ (\text{by Theorem 3.8})$$
$$(H) + \left(E'_\beta\right) \Rightarrow (NLE) \ \ (\text{by [GT12, Section 5.4]}).$$

- *Part Two*. Secondly, we need to show that

$$(H) + \left(E'_\beta\right) \Leftrightarrow \left(G_\beta\right) \Leftrightarrow (H) + \left(R_\beta\right).$$

This is the hard part. The cycle implications are obtained in [GH00, Section 8] as follows:

$$(H) + \left(R_\beta\right) \Longrightarrow \left(G_\beta\right) \Longrightarrow (H) + \left(E'_\beta\right) \Longrightarrow (H) + \left(R_\beta\right).$$

One of the most challenging results (cf. [GH00, Lemma 5.7]) is to obtain an *annulus Harnack inequality* for the Green function, without assuming any specific properties of the metric d, unlike previously known similar results in [Bar05], [GT02] where the geodesic property of the distance function was used. ∎

4.2 Matching Upper and Lower Bounds

The purpose of this subsection is to improve both (UE_{loc}) and (NLE) in order to obtain matching upper and lower bounds for the heat kernel. The reason why (UE_{loc}) and (NLE) do not match, in particular, why (NLE) contains no information about lower bound of $p_t(x, y)$ for distant x, y is the lack of *chaining properties* of the distance function, that is an ability to connect any two points $x, y \in M$ by a chain of balls of controllable radii so that the number of balls in this chain is also under control.

For example, the chain condition considered above is one of such properties. If (M, d) satisfies the chain condition, then as we have already mentioned, (NLE) implies the full sun-Gaussian lower estimate by the chain argument and the semigroup property (see for example [GHL03, Corollary 3.5]).

Here we consider a setting with weaker chaining properties. For any $\varepsilon > 0$, we introduce a modified distance $d_\varepsilon(x, y)$ by

$$d_\varepsilon(x, y) = \inf_{\{x_i\} \text{ is } \varepsilon\text{-chain}} \sum_{i=1}^{N} d(x_i, x_{i-1}), \qquad (4.13)$$

where an ε-*chain* is a sequence $\{x_i\}_{i=0}^N$ of points in M such that

$$x_0 = x, \quad x_N = y, \quad \text{and } d(x_i, x_{i-1}) < \varepsilon \text{ for all } i = 1, 2, ..., N.$$

Clearly, $d_\varepsilon(x, y)$ is decreases as ε increases and $d_\varepsilon(x, y) = d(x, y)$ if $\varepsilon > d(x, y)$. As $\varepsilon \downarrow 0$, $d_\varepsilon(x, y)$ increases and can go to ∞ or even become equal to ∞. It is easy to see that $d_\varepsilon(x, y)$ satisfies all properties of a distance function except for finiteness, so that it is a distance function with possible value $+\infty$.

It is easy to show that

$$d_\varepsilon(x, y) \asymp \varepsilon N_\varepsilon(x, y),$$

where $N_\varepsilon(x, y)$ is the smallest number of balls in a chain of balls of radius ε connecting x and y (Fig. 7):

N_ε can be regarded as the graph distance on a graph approximation of M by an ε-net.

If d is geodesic, then the points $\{x_i\}$ of an ε-chain can be chosen on the shortest geodesic, whence $d_\varepsilon(x, y) = d(x, y)$ for any $\varepsilon > 0$. If the distance function d satisfies the chain condition, then one can choose in (4.13) an ε-chain so that $d(x_i, x_{i+1}) \le C \frac{d(x, y)}{N}$, whence $d_\varepsilon(x, y) \le Cd(x, y)$. In general, $d_\varepsilon(x, y)$ may go to ∞ as $\varepsilon \to 0$, and the rate of growth of $d_\varepsilon(x, y)$ as $\varepsilon \to 0$ can be regarded as a quantitative description of the chaining properties of d.

We need the following hypothesis

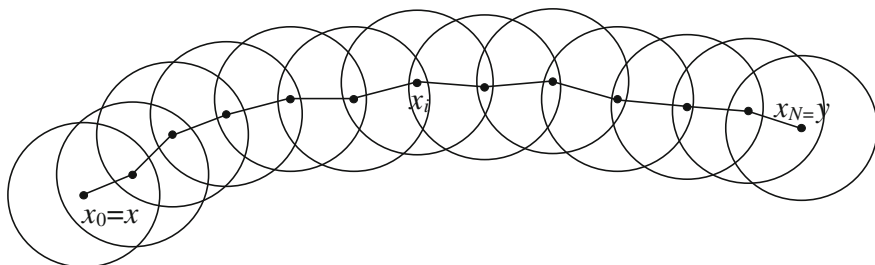C_β (*Chaining property*) For all $x, y \in M$,

Fig. 7 Chain of balls connecting x and y

$$\varepsilon^{\beta-1} d_\varepsilon (x, y) \to 0 \ \text{ as } \varepsilon \to 0,$$

or equivalently,

$$\varepsilon^\beta N_\varepsilon (x, y) \to 0 \ \text{ as } \varepsilon \to 0.$$

For $x \neq y$ we have $\varepsilon^{\beta-1} d_\varepsilon (x, y) \to \infty$ as $\varepsilon \to \infty$, which implies under (C_β) that for any $t > 0$, there is $\varepsilon = \varepsilon (t, x, y)$ satisfying the identity

$$\varepsilon^{\beta-1} d_\varepsilon (x, y) = t \tag{4.14}$$

(always take the maximal possible value of ε). If $x = y$, then set $\varepsilon (t, x, x) = \infty$.

Theorem 4.7 (Grigor'yan, Telcs [GT12, Section 6]) *Assume that all the hypothesis of Theorem 4.6 hold. If $(E_\beta) + (H)$ and (C_β) are satisfied, then*

$$p_t (x, y) \asymp \frac{C}{t^{\alpha/\beta}} \exp\left(-c \left(\frac{d_\varepsilon^\beta (x, y)}{t} \right)^{\frac{1}{\beta-1}} \right) \tag{4.15}$$

$$\asymp \frac{C}{t^{\alpha/\beta}} \exp\left(-c N_\varepsilon (x, y) \right), \tag{4.16}$$

where $\varepsilon = \varepsilon (t, x, y)$.

Since $d_\varepsilon (x, y) \geq d (x, y)$, the upper bound in (4.15) is an improvement of (UE_{loc}); similarly the lower bound in (4.15) is an improvement of (NLE). The proof of the upper bound in (4.15) follows the same line as the proof of (UE_{loc}) with careful tracing all places where the distance $d (x, y)$ is used and making sure that it can be replaced by $d_\varepsilon (x, y)$. The proof of the lower bound in (4.16) uses (NLE) and the semigroup identity along the chain with N_ε balls connecting x and y. Finally, observe that (4.15) and (4.16) are equivalent, that is

$$N_\varepsilon \asymp \left(\frac{d_\varepsilon^\beta (x, y)}{t} \right)^{\frac{1}{\beta-1}},$$

which follows by substituting here $N_\varepsilon \asymp d_\varepsilon / \varepsilon$ and $t = \varepsilon^{\beta-1} d_\varepsilon (x, y)$.

By Theorem 4.6, the same conclusion in Theorem 4.7 is true if $(E_\beta) + (H)$ is instead replaced by the one of conditions $(H) + \left(E_\beta' \right)$, (G_β) and $(H) + (R_\beta)$.

Example 4.8 A good example to illustrate Theorem 4.7 is the class of post critically finite (p.c.f.) fractals. For connected p.c.f. fractals with regular harmonic structure, the heat kernel estimate (4.16) was proved by Hambly and Kumagai [HK99], see also [KS05, Theorem 5.2]. In this setting $d(x, y)$ is the resistance metric of the fractal M and μ is the Hausdorff measure of M of dimension $\alpha := \dim_H M$. Hambly and Kumagai proved that (V_α) and (E_β) are satisfied with $\beta = \alpha + 1$. The condition (C_β) follows from their estimate

$$N_\varepsilon (x, y) \le C \left(\frac{d(x, y)}{\varepsilon} \right)^{\beta/2},$$

because

$$\varepsilon^\beta N_\varepsilon (x, y) \le C d(x, y)^{\beta/2} \varepsilon^{\beta/2} \to 0 \text{ as } \varepsilon \to 0.$$

The Harnack inequality (H) on p.c.f. fractals was proved by Kigami [Kig01, Proposition 3.2.7, p.78]. Hence, Theorem 4.7 applies and gives the estimates (4.15) and (4.16).

The estimate (4.16) means that the diffusion process goes from x to y in time t in the following way. The process firstly "computes" the value $\varepsilon(t, x, y)$, secondly "detects" a shortest chain of ε-balls connecting x and y, and then goes along that chain (Fig. 8).

This phenomenon was first observed by Hambly and Kumagai on p.c.f. fractals, but it seems to be generic. Hence, to obtain matching upper and lower bounds, one needs in addition to the usual hypotheses also the following information, encoded in the function $N_\varepsilon (x, y)$: the graph distance between x and y on any ε-net approximation of M.

Example 4.9 (*Computation of ε*) Assume that the following bound is known for all $x, y \in M$ and $\varepsilon > 0$

$$N_\varepsilon (x, y) \le C \left(\frac{d(x, y)}{\varepsilon} \right)^\gamma,$$

where $0 < \gamma < \beta$, so that (C_β) is satisfied (since $N_\varepsilon \ge d(x, y) / \varepsilon$, one must have $\gamma \ge 1$). Since by (4.14) we have $\varepsilon^\beta N_\varepsilon \asymp t$, it follows that
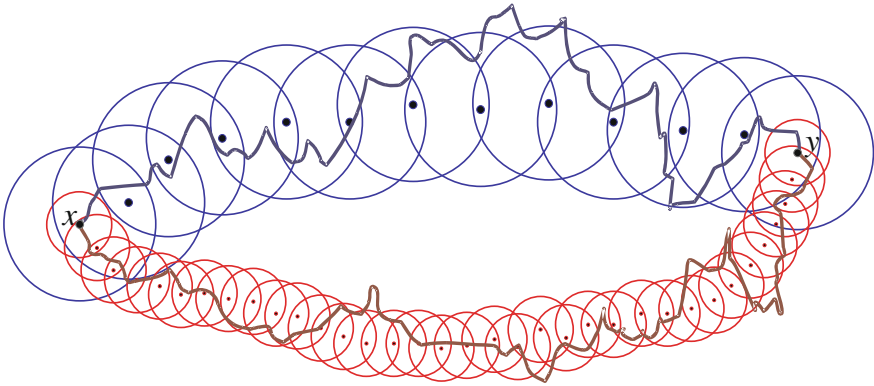
Fig. 8 Two shortest chains of ε-ball for two distinct values of ε provide different routes for the diffusion from x to y for two distinct values of t

$$\varepsilon^{\beta} \left(\frac{d(x, y)}{\varepsilon} \right)^{\gamma} \geq ct,$$

whence

$$\varepsilon \geq c \left(\frac{t}{d(x, y)^{\gamma}} \right)^{\frac{1}{\beta - \gamma}}.$$

Consequently, we obtain

$$N_{\varepsilon}(x, y) \leq Cd(x, y)^{\gamma} \varepsilon^{-\gamma} \leq Cd(x, y)^{\gamma} \left(\frac{d(x, y)^{\gamma}}{t} \right)^{\frac{\gamma}{\beta - \gamma}} = C \left(\frac{d(x, y)^{\beta}}{t} \right)^{\frac{\gamma}{\beta - \gamma}},$$

and so

$$p_t(x, y) \geq \frac{c}{t^{\alpha/\beta}} \exp \left(- \left(\frac{d(x, y)^{\beta}}{ct} \right)^{\frac{\gamma}{\beta - \gamma}} \right).$$

Similarly, the lower estimate of N_{ε}

$$N_{\varepsilon}(x, y) \geq c \left(\frac{d(x, y)}{\varepsilon} \right)^{\gamma}$$

implies an upper bound for the heat kernel

$$p_t(x, y) \leq \frac{C}{t^{\alpha/\beta}} \exp \left(- \left(\frac{d(x, y)^{\beta}}{Ct} \right)^{\frac{\gamma}{\beta - \gamma}} \right).$$

Remark 4.10 Assume that (V_α) holds and all balls in M of radius $\geq r_0$ are connected, for some $r_0 > 0$. We claim that (C_β) holds with any $\beta > \alpha$. The α-regularity of measure μ implies, by the classical ball covering argument, that any ball B_r of radius r can be covered by at most $C\left(\frac{r}{\varepsilon}\right)^\alpha$ balls of radii $\varepsilon \in (0, r)$. Consequently, if B_r is connected then any two points $x, y \in B_r$ can be connected by a chain of ε-balls containing at most $C\left(\frac{r}{\varepsilon}\right)^\alpha$ balls, so that

$$N_\varepsilon (x, y) \leq C \left(\frac{r}{\varepsilon}\right)^\alpha.$$

Since any two points $x, y \in M$ are contained in a connected ball B_r (say, with $r = r_0 + d(x, y)$), we obtain

$$\varepsilon^\beta N_\varepsilon (x, y) \leq C\varepsilon^{\beta-\alpha} r^\alpha \to 0$$

as $\varepsilon \to 0$, which was claimed.

4.3 Further Results

We discuss here some consequences and extensions of the above results. For this, we introduce two-sided estimates of the heat kernel.

(ULE_{loc}) (*Upper and lower estimates*) The heat kernel $p_t(x, y)$ exists and satisfies

$$p_t (x, y) \asymp \frac{C}{t^{\alpha/\beta}} \exp\left(-c\left(\frac{d^\beta(x, y)}{t}\right)^{\frac{1}{\beta-1}}\right). \tag{4.17}$$

Theorem 4.11 *Let (M, d, μ) be a metric measure space, and let $(\mathcal{E}, \mathcal{F})$ be a regular, conservative Dirichlet form in $L^2(M, \mu)$. If (M, d) satisfies the chain condition, then the following equivalences take place:*

$$(V_\alpha) + \begin{cases} \left(E_\beta\right) + (H) \\ \left(E'_\beta\right) + (H) \\ \left(R_\beta\right) + (H) \\ \left(G_\beta\right) \end{cases} + (locality) \Longleftrightarrow (ULE_{loc}),$$

where condition (locality) means that $(\mathcal{E}, \mathcal{F})$ is local.

Remark 4.12 Observe that if $(\mathcal{E}, \mathcal{F})$ is regular, conservative and local, then $(\mathcal{E}, \mathcal{F})$ is strongly local; this is easily seen by using the Beuling-Deny decomposition [FOT11, Theorem 3.2.1, p. 120] and by noting that both killing and jump measures disappear.

Remark 4.13 Observe also that $(V_\alpha) + (NLE) +$ (chain condition) implies that the off-diagonal lower estimate

$$p_t(x, y) \geq \frac{C'}{t^{\alpha/\beta}} \exp\left(-c'\left(\frac{d^\beta(x, y)}{t}\right)^{\frac{1}{\beta-1}}\right) \tag{4.18}$$

for μ-almost all $x, y \in M$ and all $t > 0$, see for example [GHL08, Proposition 3.1] or [Bar98], [GHL03, Corollary 3.5].

Sketch of proof for Theorem 4.11 (1) "$\Rightarrow$".
Let us show the implication

$$(V_\alpha) + (E_\beta) + (H) + (\text{locality}) \Rightarrow (ULE_{loc}). \tag{4.19}$$

Indeed, by Remark 4.12, we have that $(\mathcal{E}, \mathcal{F})$ is strongly local. Now, using Theorem 4.3, we obtain $(UE_{loc}) + (NLE)$. Using Remark 4.13, we see that (4.18) holds, showing that (ULE_{loc}) is true.

Similarly, using Theorem 4.6, we obtain the other three implications "$\Rightarrow$".
(2) "$\Leftarrow$".
Let us show the opposite implication

$$(ULE_{loc}) \Rightarrow (V_\alpha) + (E_\beta) + (H) + (\text{locality}). \tag{4.20}$$

Indeed, note that

$$(ULE_{loc}) \Rightarrow (V_\alpha) \text{ (by Theorem 2.2)}$$
$$(UE_{loc}) \Rightarrow (\text{locality}) \text{ (by Theorem 2.14)}$$
$$(UE_{loc}) + (NLE) \Rightarrow (E_\beta) + (H) \text{ (by Theorem 4.3)}$$

showing that the implication (4.20) holds.

Similarly, all the other three implications "$\Longleftarrow$" also hold.   ∎

Remark 4.14 The implication (4.19) can also be proved by using Theorem 4.7 and the fact that $d_\varepsilon \asymp d$.

Conjecture 4.15 *The condition* (E_β) *above may be replaced by*

$$\lambda_{\min}(B(x, r)) \asymp r^{-\beta}. \tag{λ_β}$$

In fact, (E_β) in all statements can be replaced by the resistance condition:

$$\text{res}(B_r, B_{2r}) \asymp r^{\beta-\alpha} \tag{R_β}$$

where $B_r = B(x, r)$. In the strongly recurrent case $\alpha < \beta$, it alone implies the elliptic Harnack inequality (H) so that two sided heat kernel estimates are equivalent

to $(V_\alpha) + (R_\beta)$ as was proved by Barlow, Coulhon, Kumagai [BCK05] (in a setting of graphs) and was discussed in M. Barlow's lectures.

An interesting (and obviously hard) question is the characterization of the elliptic Harnack inequality (H) in more geometric terms—so far nothing is known, not even a conjecture.

One can consider also a *parabolic* Harnack inequality (*PHI*), which uses caloric functions instead of harmonic functions. Then in a general setting and assuming the volume doubling condition (*VD*) (instead of (V_α)), the following holds (cf. [BGK12]):

$$(PHI) \Leftrightarrow (UE_{loc}) + (NLE) .$$

On the other hand, (*PHI*) is equivalent to

$$\text{Poincaré inequality} + \text{cutoff Sobolev inequality},$$

see [BBK06].

Conjecture 4.16 *The cutoff Sobolev inequality here can be replaced by* (λ_β) *and/or* (R_β).

5 Upper Bounds for Jump Processes

We have investigated above the heat kernel for the *local* Dirichlet form. In this section we shall study the *non-local* Dirichlet form and present the equivalence conditions for upper bounds of the associated heat kernel. As an interesting example, we discuss the heat kernel estimates for effective metric spaces.

A non-local Dirichlet form will give arise to a jump process, that is, the trajectories of this process are discontinuous, as we have already seen for a symmetric stable process of index β (Lévy process). And the heat kernel decays at a polynomial rate (cf. 1.2), instead of an exponential rate as for a local Dirichlet form.

Jump process have found various applications in science. For instance, a Lévy flight is a jump process and can be used to describe animal foraging patterns, the distribution of human travel and some aspects of earthquake behavior (cf. [BBW08]).

5.1 Upper Bounds for Non-local Dirichlet Forms

The techniques for obtaining heat kernel bounds for non-local Dirichlet forms has been developed by a number of authors, see for example [BBCK09, BGK09, BL02, CK03, CK08] and the references therein. The basic approach to obtaining heat kernel upper estimates used in these papers consists of the two steps. The first step is to obtain the heat kernel upper bounds for a *truncated* Dirichlet form, that is, in the

case when the jump density $J(x, y)$ has a bounded range. In this case one uses the Davies method as it was presented in the seminal work [CKS87] and where the cut-off functions of form $(\lambda - d(x_0, x))_+$ were used (where λ is a positive constant). This method can be used as long as the cut-off functions belong to the domain of the Dirichlet form, which is the case only when $\beta < 2$ (hence, if $\beta \geq 2$ then this method does not work).

The second step is to obtain heat kernel estimates for the original Dirichlet form by comparing the heat semigroup of the truncated Dirichlet form with the original heat semigroup. We remark that while the first step was done by purely analytic means, the second step in the above-mentioned papers used a probabilistic argument.

Here we describe an alternative new approach of [GHL00] for obtaining upper bounds.

Recall that by a theorem of Beurling and Deny, any regular conservative Dirichlet form admits a decomposition

$$\mathcal{E}(u, v) = \mathcal{E}^{(L)}(u, v) + \mathcal{E}^{(J)}(u, v), \tag{5.1}$$

where $\mathcal{E}^{(L)}$ is a *local part* and

$$\mathcal{E}^{(J)}(u, v) = \iint_{M \times M \setminus \text{diag}} (u(x) - u(y))(v(x) - v(y)) \, dj(x, y) \tag{5.2}$$

is a *jump part* with a jump measure j defined on $M \times M \setminus \text{diag}$. In our setting the jump measure j will have a density with respect to $\mu \times \mu$, which will be denoted by $J(x, y)$, and so the jump part $\mathcal{E}^{(J)}$ becomes

$$\mathcal{E}^{(J)}(u, v) = \iint_{M \times M} (u(x) - u(y))(v(x) - v(y)) \, J(x, y) d\mu(y) d\mu(x). \tag{5.3}$$

We introduce the following hypothesis.

$(\mathbf{V}_\leq)$ (*Upper α- regularity*) For all $x \in M$ and all $r > 0$,

$$V(x, r) \leq Cr^\alpha.$$

(**UE**) (*Upper estimate of non-local type*) The heat kernel p_t exists and satisfies the off-diagonal upper estimate

$$p_t(x, y) \leq \frac{C}{t^{\alpha/\beta}} \left(1 + \frac{d(x, y)}{t^{1/\beta}}\right)^{-(\alpha+\beta)}$$

for all $t > 0$ and μ-almost all $x, y \in M$.

$(\mathbf{J}_\leq)$ (*Upper bound of jump density*) The jump density exists and admits the estimate

$$J(x, y) \leq Cd(x, y)^{-(\alpha+\beta)},$$

for μ-almost all $x, y \in M$.

($\mathbf{T}_{strong}$) (*Strong tail estimate*) There exist constants $c > 0$ and $\beta > 0$ such that, for all balls $B = B(x_0, r)$ and for all $t > 0$,

$$P_t \mathbf{1}_{B^c}(x) \leq \frac{ct}{r^\beta} \text{ for } \mu \text{ -almost all } x \in \frac{1}{4}B.$$

Clearly, we have that $(T_{\exp}) \Rightarrow (T_{strong}) \Rightarrow (T_\beta)$.

We now state the main technical result of [GHL00].

Theorem 5.1 (Grigor'yan et al. [GHL00]) *Let (M, d, μ) be a metric measure space with precompact balls, and let $(\mathcal{E}, \mathcal{F})$ be a regular conservative Dirichlet form in $L^2(M, \mu)$ with jump density J. Then the following implication holds:*

$$\left(V_\leq\right) + (DUE) + \left(J_\leq\right) + \left(S_\beta\right) \Rightarrow (UE). \tag{5.4}$$

We remark that by [GHL03, Theorem 3.2], if $(\mathcal{E}, \mathcal{F})$ is conservative then

$$\left(V_\leq\right) + (UE) \Rightarrow (V_\alpha).$$

Hence, the hypotheses of Theorem 5.1 imply that μ is α-regular.

Sketch of proof for Theorem 5.1 We sketch the ideas of the proof.

- *Step* 1. We decompose $\mathcal{E}(u)$ into two parts:

$$\mathcal{E}(u) = \mathcal{E}^{(\rho)}(u) + \int_M \int_{M \setminus B(x, \rho)} (u(x) - u(y))^2 J(x, y)d\mu(y)d\mu(x),$$

where $\rho \in (0, \infty)$ is any fixed number. Then the form $\left(\mathcal{E}^{(\rho)}, \mathcal{F}\right)$ can be extended to a regular Dirichlet form $\left(\mathcal{E}^{(\rho)}, \mathcal{F}^{(\rho)}\right)$. Indeed, since using condition $(J_\leq)$,

$$\operatorname*{esup}_{x \in M} \int_{B(x, \rho)^c} J(x, y)d\mu(y) < \infty,$$

the form $\left(\mathcal{E}^{(\rho)}, \mathcal{F}\right)$ is closable, and its closure $\left(\mathcal{E}^{(\rho)}, \mathcal{F}^{(\rho)}\right)$ in L^2 is a regular Dirichlet form in L^2. Note that $\left(\mathcal{E}^{(\rho)}, \mathcal{F}^{(\rho)}\right)$ is ρ-*local* (non-local): $\mathcal{E}^{(\rho)}(f, g) = 0$ for any two functions $f, g \in \mathcal{F}^{(\rho)}$ with compact supports such that

$$\text{dist}(\text{supp } f, \text{supp } g) > \rho.$$

- *Step* 2. We need to obtain upper estimates of the heat kernel $q_t(x, y)$ of the *truncated* Dirichlet form $\left(\mathcal{E}^{(\rho)}, \mathcal{F}^{(\rho)}\right)$. Indeed, conditions (DUE), $\left(J_\leq\right)$, (S_β) and $\left(V_\leq\right)$

imply the following estimate of $q_t(x, y)$:

$$q_t(x, y) \leq \frac{C}{t^{\alpha/\beta}} \exp\left(4\rho^{-\beta}t\right) \exp\left(-c\left(\frac{d(x, y)}{\rho} \wedge \frac{\rho}{t^{1/\beta}}\right)\right) \qquad (5.5)$$

for all $t > 0$ and μ-almost all $x, y \in M$, where constants $C, c > 0$ depend on the constants in the hypotheses but are independent of ρ. This can be done with a certain amount of effort, by using the bootstrapping technique where the comparison inequality [GHL10, Corollary 4.8, Remark 4.10] for heat semigroups play an important rôle.

• *Step* 3. Next we apply the following useful inequality between two heat kernels:

$$p_t(x, y) \leq q_t(x, y) + 2t \underset{x \in M, y \in B(x,\rho)^c}{\text{esup}} J(x, y) \qquad (5.6)$$

for all $t > 0$ and almost all $x, y \in M$; this inequality follows from the parabolic maximum principle alone. Therefore, by choosing an appropriate ρ, it follows from (5.5), (5.6) that, for any real $n \geq 0$,

$$p_t(x, y) \leq \frac{c(n)}{t^{\alpha/\beta}} \left(1 + \frac{d(x, y)}{t^{1/\beta}}\right)^{-\frac{(\alpha+\beta)n}{n+\alpha+\beta}} \qquad (5.7)$$

for almost all $x, y \in M$ and all $t > 0$.

Note that (5.7) is nearly close to our desired estimate (UE). However, one can not just obtain (UE) by directly taking the limit as $n \to \infty$, since we do not know whether the coefficient $c(n)$ is bounded uniformly in n. We need the second iteration.

• *Step* 4. Finally, we will obtain (UE) by a self-improvement of (5.7). Indeed, one can use (5.7) to obtain

$$\int_{B(x,r)^c} p_t(x, y)d\mu(y) \leq C(n) \left(rt^{-1/\beta}\right)^{-\theta},$$

where $\theta = \frac{n\beta - \alpha(\alpha+\beta)}{n+\alpha+\beta} \in (0, \beta)$ (note that this estimate is sharper than condition (S_β)), and then repeating the above procedure, we arrive at (UE).   ∎

Now we can state some equivalences for (UE).

Theorem 5.2 (Grigor'yan et al. [GHL00]) *Let (M, d, μ) be a metric measure space with precompact balls, and let $(\mathcal{E}, \mathcal{F})$ be a regular conservative Dirichlet form in $L^2(M, \mu)$ with jump density J. If $(V_\leq)$ holds, then the following equivalences are true:*

$$(UE) \Leftrightarrow (UE\Phi) + (J_\le) \tag{5.8}$$
$$\Leftrightarrow (DUE) + (J_\le) + \left(T_\beta\right)$$
$$\Leftrightarrow (DUE) + (J_\le) + \left(S_\beta\right)$$
$$\Leftrightarrow (DUE) + (J_\le) + \left(T_{strong}\right).$$

Proof Observe that the implication $(UE) \Rightarrow (J_\le)$ holds by [BGK09, p. 150], and $(UE) \Rightarrow (UE\Phi)$ is trivial by taking $\Phi(s) = (1+s)^{-(\alpha+\beta)}$. The implication $(UE\Phi) \Rightarrow (DUE)$ is obvious. The implication $(UE\Phi) \Rightarrow \left(T_\beta\right)$ was proved in (3.10) (see also [GHL03, formula (3.6), p. 2072]). Since $(\mathcal{E}, \mathcal{F})$ is conservative, the equivalence $\left(T_\beta\right) \Leftrightarrow \left(S_\beta\right)$ holds by [GH08, Theorem 3.1, p. 96]. By Theorem 5.1 we have

$$(DUE) + (J_\le) + \left(S_\beta\right) \Rightarrow (UE),$$

which closes the cycle of implications, thus proving the first three equivalences.

Finally, the implication $(UE) \Rightarrow \left(T_{strong}\right)$ is true by using (3.10), and hence

$$(UE) \Rightarrow (DUE) + (J_\le) + \left(T_{strong}\right)$$
$$\Rightarrow (DUE) + (J_\le) + \left(T_\beta\right) \Rightarrow (UE),$$

which finishes the proof.   ∎

Remark 5.3 The upper estimate (UE) is best possible for non-local forms in the following sense: if the heat kernel p_t satisfies the estimate

$$p_t(x, y) \le \frac{1}{t^{\alpha/\beta}} \Phi\left(\frac{d(x, y)}{t^{1/\beta}}\right)$$

for all $t > 0$ and μ-almost all $x, y \in M$, where Φ is a continuous decreasing function on $[0, +\infty)$, then necessarily

$$\Phi(s) \ge c(1+s)^{-(\alpha+\beta)}$$

for some $c > 0$ (see Theorem 2.14).

Remark 5.4 Under the standing assumptions of Theorem 5.2, the following equivalence is true
$$(UE_{loc}) \Leftrightarrow (DUE) + (\text{"locality"}) + \left(S_\beta\right).$$

Indeed, since (UE_{loc}) is stronger than (UE), it implies (DUE) and $\left(S_\beta\right)$ by Theorem 5.2. Next, $(UE_{loc}) \Rightarrow (\text{"locality"})$ by Theorem 2.14 above. The opposite implication

$$(DUE) + (\text{"locality"}) + \left(S_\beta\right) \Rightarrow (UE_{loc})$$

was stated in Theorem 3.8.

In order to state some consequence of Theorem 5.2, we need the following Proposition.

Define first the following condition:

$(J_\geq)$ (*Lower bound of jump density*) There exist constants C, α, $\beta > 0$ such that, for μ-almost all $x \neq y$,

$$J(x, y) \geq C^{-1}d(x, y)^{-(\alpha+\beta)}.$$

Proposition 5.5 *Let* (M, d, μ) *be a metric measure space, and let* $(\mathcal{E}, \mathcal{F})$ *be a regular Dirichlet form in* $L^2(M, \mu)$ *with jump density* J. *Then*

$$(V_\alpha) + (J_\geq) \Rightarrow (DUE). \tag{5.9}$$

Proof As was proved in [HuK06, Theorem 3.1], under (V_α) the following inequality holds for all non-zero functions $u \in L^1 \cap L^2$:

$$\int_M \int_M \frac{(u(x) - u(y))^2}{d(x, y)^{\alpha+\beta}} d\mu(x)\, d\mu(y) \geq c||u||_2^{2(1+\beta/\alpha)} \|u\|_1^{-2\beta/\alpha},$$

where c is a positive constant. Using (5.1), (5.3) and $(J_\geq)$ we obtain

$$\begin{aligned}
\mathcal{E}(u) &= \mathcal{E}^{(L)}(u) + \mathcal{E}^{(J)}(u) \\
&\geq C \int_M \int_M \frac{(u(x) - u(y))^2}{d(x, y)^{\alpha+\beta}} d\mu(x)\, d\mu(y) \\
&\geq c||u||_2^{2(1+\beta/\alpha)} \|u\|_1^{-2\beta/\alpha}
\end{aligned}$$

for all $u \in \mathcal{F} \cap L^1$. Hence, (DUE) follows by Theorem 3.6. ∎

We obtain the following consequence of Theorem 5.2.

Theorem 5.6 (Grigor'yan et al. [GHL00]) *Let* (M, d, μ) *be a metric measure space with precompact balls, and let* $(\mathcal{E}, \mathcal{F})$ *be a regular conservative Dirichlet form in* $L^2(M, \mu)$ *with jump density* J. *If* (V_α) *holds and* $J(x, y) \asymp d(x, y)^{-(\alpha+\beta)}$, *then*

$$(UE) \Leftrightarrow (S_\beta). \tag{5.10}$$

Proof Let us show that $(S_\beta) \Rightarrow (UE)$. Indeed, (DUE) holds by Proposition 5.5. Hence, (UE) is satisfied by Theorem 5.2. The opposite implication $(UE) \Rightarrow (S_\beta)$ holds also by Theorem 5.2. ∎

Therefore, if (V_α) holds and $J(x, y) \asymp d(x, y)^{-(\alpha+\beta)}$, then in order to obtain off-diagonal upper bounds of heat kernels, one needs only to verify the survival condition (S_β). In the sequel, we will show that the survival condition (S_β) holds for a class of measure spaces with effective resistance metrics.

5.2 Upper Bounds Using Effective Resistance

We will show how Theorem 5.2 can be applied for a certain class of metric measure spaces with effective resistance.

Let $(\mathcal{E}, \mathcal{F})$ be a regular Dirichlet form in $L^2(M, \mu)$ as before. Recall that the *effective resistance* $R(A, B)$ between two disjoint non-empty closed subsets A and B of M is defined by

$$R(A, B)^{-1} = \inf \{\mathcal{E}(u) : u \in \mathcal{F} \cap C_0, \ u|_A = 1 \text{ and } u|_B = 0\}. \quad (5.1)$$

It follows from (5.1) that, for any fixed A, $R(A, B)$ is a non-increasing function of B. Denote by

$$R(x, B) := R(\{x\}, B) \text{ and } R(x, y) := R(\{x\}, \{y\}).$$

In general, it may happen that $R(x, y) = \infty$ for some points $x, y \in M$. Below we will exclude this case.

Fix a parameter $\gamma > 0$, and introduce conditions (R_1) and (R_2).

$(\mathbf{R_1})$: For all $u \in \mathcal{F} \cap C_0(M)$ and all $x, y \in M$, the following inequality holds:

$$|u(x) - u(y)|^2 \leq C d(x, y)^\gamma \mathcal{E}(u).$$

$(\mathbf{R_2})$: For all $x \in M$ and $r > 0$,

$$R(x, B(x, r)^c) \geq C^{-1} r^\gamma.$$

Theorem 5.7 (Grigor'yan et al. [GHL00]) *Let $(\mathcal{E}, \mathcal{F})$ be a regular Dirichlet form in $L^2(M, \mu)$. Then*

$$(V_\alpha) + (R_1) + (R_2) \Rightarrow (S_\beta) + (DUE),$$

where $\beta = \alpha + \gamma$. Consequently, under the standing conditions $(V_\alpha) + (R_1) + (R_2)$, we have that
$$(UE) \Leftrightarrow (J_\leq). \quad (5.2)$$

Sketch of proof for Theorem 5.7 The proof consists of the following five steps.

- *Step* 1. For any ball $B := B(x_0, r)$, using conditions (R_1) and (R_2), we can obtain the two-sided estimate of the Green functions $g_B(x, y)$:

$$\sup_{x, y \in B} g_B(x, y) \leq C r^\gamma, \quad (5.3)$$

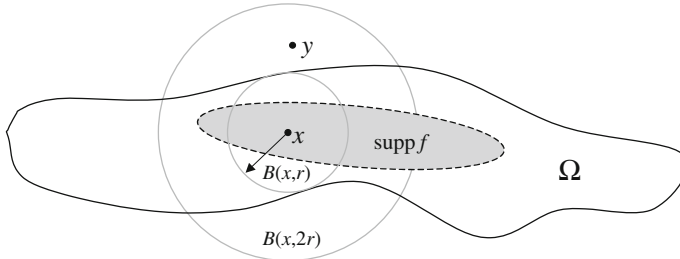$$\inf_{y \in B(x_0, \eta r)} g_B(x_0, y) \geq C^{-1} r^\gamma, \quad (5.4)$$

Fig. 9 Points x, y

where $C > 0$ and $\eta \in (0, 1)$.

- *Step* 2. Therefore, under condition (V_α), it follows from (5.3), (5.4) that condition $\left(E'_\beta\right)$ holds:

$$\operatorname*{esup}_{B} E^B \leq Cr^{\alpha+\gamma}, \tag{5.5}$$

$$\operatorname*{einf}_{\delta_1 B} E^B \geq C^{-1} r^{\alpha+\gamma}, \tag{5.6}$$

where E^B is the weak solution of the Poisson-type equation (4.7) as before, and $C > 0$ and $\delta_1 \in (0, 1/2)$.

- *Step* 3. To show condition $\left(S_\beta\right)$, observe that, for all $t > 0$ and μ-almost all $x \in B$,

$$P_t^B 1_B(x) \geq \frac{E^B(x) - t}{\left\| E^B \right\|_\infty}, \tag{5.7}$$

which follows by using the parabolic maximum principle, nothing else. Hence, using (5.5), (5.6),

$$\begin{aligned} P_t^B 1_B(x) &\geq \frac{E^B(x) - t}{\left\| E^B \right\|_\infty} \\ &\geq c - c_1 t r^{-\beta} \\ &\geq \frac{c}{2}, \end{aligned}$$

for all $t > 0$ and μ-almost all $x \in B(x_0, \delta_1 r)$, provided that $tr^{-\beta}$ is small enough, thus proving $\left(S_\beta\right)$.

- *Step* 4. We show that $(R_1) \Rightarrow (DUE)$. Consider a function $f \in \mathcal{F} \cap C_0(\Omega)$ normalized so that $\sup |f| = 1$, and let $x \in \Omega$ be a point such that $|f(x)| = 1$. Let r be the largest radius such that $B(x, r) \subset \Omega$. Then the ball $B(x, 2r)$ is not covered by Ω so that there exists a point $y \in B(x, 2r) \setminus \Omega$ (note that M is unbounded by condition (V_α)). In particular, $y \notin \operatorname{supp} f$ (see Fig. 9). Noting that

$\mathcal{E}^{(J)}(f) \leq \mathcal{E}(f)$ and by the α-regularity of μ

$$r \leq C\left[\mu\left(B(x, r)\right)\right]^{1/\alpha} \leq C\left[\mu\left(\Omega\right)\right]^{1/\alpha},$$

we obtain from (R_1) that

$$
\begin{aligned}
1 &= |f(y) - f(x)|^2 \\
&\leq Cd(y, x)^{\beta-\alpha}\mathcal{E}^{(J)}(f) \\
&\leq C(2r)^{\beta-\alpha}\mathcal{E}(f) \leq C2^{\beta-\alpha}\left[\mu\left(\Omega\right)\right]^{\beta/\alpha-1}\mathcal{E}(f).
\end{aligned}
$$

Since $\|f\|_2^2 \leq \mu(\Omega)$, it follows that

$$\frac{\mathcal{E}(f)}{\|f\|_2^2} \geq c\left[\mu\left(\Omega\right)\right]^{-\beta/\alpha},$$

for some $c > 0$, thus proving the Faber-Krahn inequality. Hence, condition (*DUE*) follows by using Theorem 3.6.

- *Step* 5. Finally, with a certain amount of effort [GHL00, Proposition 6.5, Lemma 6.4], one can show that

$$(R_1) + (R_2) \Rightarrow \text{conservativeness of } (\mathcal{E}, \mathcal{F}).$$

Therefore, the equivalence (5.2) follows directly by using Theorem 5.2.

References

[AB]   Andres, S., Barlow, M.: Energy inequalities for cutoff fucntions and some applications. J. Reine Angew. Math. (preprint)

[Aro67]   Aronson, D.: Bounds for the fundamental solution of a parabolic equation. Bull. Amer. Math. Soc. **73**, 890–896 (1967)

[Aro68]   Aronson, D.: Non-negative solutions of linear parabolic equations. Ann. Scuola Norm. Sup. Pisa. Cl. Sci. **22**, 607–694 (1968) (Addendum **25**:221–228, 1971)

[Bar98]   Barlow, M.: Diffusions on fractals. Lecture Notes of Mathematics, vol. 1690, pp. 1–121. Springer, Berlin (1998)

[Bar04]   Barlow, M.: Which values of the volume growth and escape time exponents are possible for graphs? Rev. Mat. Iberoamericana **20**, 1–31 (2004)

[Bar05]   Barlow, M.: Some remarks on the elliptic harnack inequality. Bull. London Math. Soc. **37**, 200–208 (2005)

[BBCK09] Barlow, M., Bass, R., Chen, Z.-Q., Kassmann, M.: Non-local Dirichlet forms and symmetric jump processes. Trans. Amer. Math. Soc. **361**, 1963–1999 (2009)

[BBK06] Barlow, M., Bass, R., Kumagai, T.: Stability of parabolic Harnack inequalities on metric measure spaces. J. Math. Soc. Japan **58**, 485–519 (2006)

[BCK05] Barlow, M., Coulhon, T., Kumagai, T.: Characterization of sub-Gaussian heat kernel estimates on strongly recurrent graphs. Comm. Pure Appl. Math. **58**, 1642–1677 (2005)

[BGK09] Barlow, M., Grigor'yan, A., Kumagai, T.: Heat kernel upper bounds for jump processes and the first exit time. J. Reine Angew. Math. **626**, 135–157 (2009)

[BGK12] Barlow, M., Grigor'yan, A., Kumagai, T.: On the equivalence of parabolic harnack inequalities and heat kernel estimates. J. Math. Soc. Japan **64**, 1091–1146 (2012)

[BP88]   Barlow, M., Perkins, E.: Brownian motion on the Sierpínski gasket. Probab. Theory Relat. Fields **79**, 543–623 (1988)

[BBW08] Barthelemy, P., Bertolotti, J., Wiersma, D.: A Lévy flight for light. Nature **453**, 495–498 (2008)

[BL02]   Bass, R., Levin, D.: Transition probabilities for symmetric jump processes. Trans. Amer. Math. Soc. **354**, 2933–2953 (2002)

[CKS87] Carlen, E., Kusuoka, S., Stroock, D.: Upper bounds for symmetric Markov transition functions. Ann. Inst. H. Poincaré-Probab. Statist. **23**, 245–287 (1987)

[Car96] Carron, G.: Inégalités isopérimétriques de Faber-Krahn et conséquences. In: Actes de la table ronde de géométrie différentielle (Luminy, 1992) Collection SMF Séminaires et Congrès, **1** (1996), pp. 205–232

[CK03]   Chen, Z., Kumagai, T.: Heat kernel estimates for stable-like processes on d-sets. Stoch. Process. Appl. **108**, 27–62 (2003)

[CK08]   Chen, Z., Kumagai, T.: Heat kernel estimates for jump processes of mixed types on metric measure spaces. Probab. Theory Relat. Fields **140**, 277–317 (2008)

[CG98]   Coulhon, T., Grigor'yan, A.: Random walks on graphs with regular volume growth. Geom. Funct. Anal. **8**, 656–701 (1998)

[Dav92] Davies, E.: Heat kernel bounds, conservation of probability and the Feller property. J. d'Analyse Math. **58**, 99–119 (1992)

[FOT11] Fukushima, M., Oshima, Y. Takeda, M.: Dirichlet forms and symmetric Markov processes. In: de Gruyter, W. (ed.) Studies in Mathematics, vol. 19, 2nd edn. Springer, Berlin (2011)

[Gol87] Goldstein, S.: Random walks and diffusion on fractals. In: Kesten, H. (ed.) Percolation Theory and Ergodic Theory of Infinite Particle Systems, vol. 8, pp. 121–129, IMA Math. Appl. Springer, New York (1987)

[Gri92] Grigor'yan, A.: The heat equation on noncompact riemannian manifolds. Mat. Sb. **182**(1), 55–87 (1991) (translation in Math. USSR-Sb. **72**(1) (1992), 47–77)

[Gri94] Grigor'yan, A.: Heat kernel upper bounds on a complete non-compact manifold. Rev. Mat. Iberoamericana **10**, 395–452 (1994)

[GH08]   Grigor'yan, A., Hu, J.: Off-diagonal upper estimates for the heat kernel of the dirichlet forms on metric spaces. Invent. Math. **174**, 81–126 (2008)

[GH10]   Grigor'yan, A., Hu, J.: Upper bounds of heat kernels on doubling spaces. Moscow Math. J. **14**, 505–563 (2014)

[GH00]   Grigor'yan, A., Hu, J.: Heat kernels and green functions on metric measure spaces. Canad. J. Math. **66**, 641–699 (2014)

[GHL03] Grigor'yan, A., Hu, J., Lau, K.-S.: Heat kernels on metric-measure spaces and an application to semilinear elliptic equations. Trans. Amer. Math. Soc. **355**, 2065–2095 (2003)

[GHL08] Grigor'yan, A., Hu, J., Lau, K.-S.: Obtaining upper bounds of heat kernels from lower bounds. Comm. Pure Appl. Math. **61**, 639–660 (2008)

[GHL09] Grigor'yan, A., Hu, J., Lau, K.-S.: Heat kernels on metric spaces with doubling measure. In: Bandt, C., Mörters, P., Zähle, M. (eds.) Fractal Geometry and Stochastics IV, Progress in Probability, vol. 61, pp. 3–44. Birkhäuser, Basel (2009)

[GHL10] Grigor'yan, A., Hu, J., Lau, K.-S.: Comparison inequalities for heat semigroups and heat kernels on metric measure spaces. J. Funct. Anal. **259**, 2613–2641 (2010)

[GHL00] Grigor'yan, A., Hu, J., Lau, K.-S.: Estimates of heat kernels for non-local regular dirichlet forms. Trans. AMS (to appear)

[GK08]   Grigor'yan, A., Kumagai, T.: On the dichotomy in the heat kernel two sided estimates. In: P. E. et al. (ed.) Analysis on Graphs and its Applications, Procedings of the Symposia in Pure Mathematics, American Mathematical Society, vol. 77, pp. 199–210 (2008)

[GT02]   Grigor'yan, A., Telcs, A.: Harnack inequalities and sub-Gaussian estimates for random walks. Math. Ann. **324**, 521–556 (2002)

[GT12]   Grigor'yan, A., Telcs, A.: Two-sided estimates of heat kernels on metric measure spaces. Ann. Probab. **40**, 1212–1284 (2012)

[HK99]   Hambly, B., Kumagai, T.: Transition density estimates for diffusion processes on post critically finite self-similar fractals. Proc. London Math. Soc. **79**, 431–458 (1999)

[HuK06]   Hu, J., Kumagai, T.: Nash-type inequalities and heat kernels for non-local Dirichlet forms. Kyushu J. Math. **60**, 245–265 (2006)

[Jon96]   Jonsson, A.: Brownian motion on fractals and function spaces. Math. Zeit. **222**, 495–504 (1996)

[Kig01]   Kigami, J.: Analysis on fractals. Cambridge University Press, Cambridge (2001)

[KS05]   Kumagai, T., Sturm, K.: Construction of diffusion processes on fractals, d-sets, and general metric measure spaces. J. Math. Kyoto Univ. **45**, 307–327 (2005)

[Kus87]   Kusuoka, S.: A diffusion process on a fractal. In: Symposium on Probabilistic Methods in Mathematical Physics, pp. 251–274. Katata, Academic Press, Amsterdam, Taniguchi (1987)

[Nas58]   Nash, J.: Continuity of solutions of parabolic and elliptic equations. Amer. J. Math. **80**, 931–954 (1958)

[Pie00]   Pietruska-PaŁuba, K.: On function spaces related to fractional diffusion on d-sets. Stoch. Stoch. Rep. **70**, 153–164 (2000)

[Sto00]   Stós, A.: Symmetric α-stable processes on d-sets. Bull. Pol. Acad. Sci. Math. **48**, 237–245 (2000)