

# Heat kernels on metric spaces with doubling measure

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

**Abstract.** In this survey we discuss heat kernel estimates of self-similar type on metric spaces with doubling measures. We characterize the tail functions from heat kernel estimates in both non-local and local cases. In the local case we also specify the domain of the energy form as a certain Besov space, and identify the walk dimension in terms of the critical Besov exponent. The techniques used include self-improvement of heat kernel upper bound and the maximum principle for weak solutions. All proofs are completely analytic.

**Mathematics Subject Classification (2000).** Primary: 47D07, Secondary: 28A80, 46E35.

**Keywords.** Doubling measure, heat kernel, maximum principle, heat equation.

## Contents

1. Introduction	1
2. What is the heat kernel	2
2.1. The notion of a heat kernel	2
2.2. Heat semigroup and Dirichlet forms	3
2.3. Examples	6
3. Auxiliary material on metric measure spaces	7
3.1. Besov spaces	7
3.2. Doubling condition and reverse doubling condition	8
4. Consequences of heat kernel estimates	9
4.1. Consequences of lower bound	10
4.2. Consequences of upper bound	12
4.3. Walk dimension	15
4.4. Consequences of two-sided estimates (non-local case)	16
5. A maximum principle and its applications	16
5.1. Weak differentiation	16
5.2. Maximum principle for weak solutions	17
5.3. Some applications of the maximum principle	19
6. Upper bounds in the local case	21
6.1. Exponential tail	21
6.2. Consequences of two-sided estimates (local case)	25
References	29

## 1. Introduction

The heat kernel is an important tool in modern analysis, which appears to be useful for applications in mathematical physics, geometry, probability, fractal analysis, graphs, function spaces and in other fields. There has been a vast literature devoted to various aspects of heat kernels (see, for example, a collection [29]). It is not feasible to give a full-scale panorama of this subject here. In

---

AG was supported by SFB 701 of the German Research Council (DFG) and the Grants from the Department of Mathematics and IMS in CUHK. JH was supported by NSFC (Grant No. 10631040) and the HKRGC Grant in CUHK. KSL was supported by the HKRGC Grant in CUHK.

this article, we consider heat kernels on abstract metric measure spaces and focus on the following questions:

- Assuming that heat kernel satisfies certain estimates of self-similar type, what are the consequences for the underlying metric measure structure?
- Developing of the self-improvement techniques for heat kernel upper bounds of subgaussian types.

Useful auxiliary tools that we develop here include the family of Besov functions spaces and the maximum principle for weak solution for abstract heat equation.

Some of these questions have been discussed in various settings, for example, in [1, 4, 10, 12, 14, 18, 32, 33, 35, 36, 37, 38, 39] for the Euclidean spaces or Riemannian manifolds, in [5, 7, 25] for torus or infinite graphs, in [9, 27, 41] for metric spaces, in [2, 3, 6, 26] for certain classes of fractals. The contents of this paper are based on the work [20], [21], [22] and [24]. Similar questions were discussed in the survey [19] when the underlying measure is Ahlfors-regular, while the main emphasis in the present survey is on the case of doubling measures.

## 2. What is the heat kernel

We give the definition of a heat kernel on a metric measure space, followed by some well-known examples on Riemannian manifolds and on a certain class of fractals.

### 2.1. The notion of a heat kernel

Let  $(M, d)$  be a locally compact separable metric space and let  $\mu$  be a Radon measure on  $M$  with full support. The triple  $(M, d, \mu)$  is termed a *metric measure space*. In the sequel, the norm in the real Banach space  $L^p := L^p(M, \mu)$  is defined as usual by

$$\|f\|_p := \left( \int_M |f(x)|^p d\mu(x) \right)^{1/p}, \quad 1 \leq p < \infty,$$

and

$$\|f\|_\infty := \operatorname{esup}_{x \in M} |f(x)|,$$

where  $\operatorname{esup}$  is the essential supremum. The inner product of  $f, g \in L^2$  is denoted by  $(f, g)$ .

**Definition 2.1.** A family  $\{p_t\}_{t>0}$  of functions  $p_t(x, y)$  on  $M \times M$  is called a *heat kernel* if for any  $t > 0$  it satisfies the following five conditions:

1. *Measurability:* the  $p_t(\cdot, \cdot)$  is  $\mu \times \mu$  measurable in  $M \times M$ .
2. *Markovian property:*  $p_t(x, y) \geq 0$  for  $\mu$ -almost all  $x, y \in M$ , and

$$\int_M p_t(x, y) d\mu(y) \leq 1, \quad (2.1)$$

for  $\mu$ -almost all  $x \in M$ .

3. *Symmetry:*  $p_t(x, y) = p_t(y, x)$  for  $\mu$ -almost all  $x, y \in M$ .
4. *Semigroup property:* for any  $s > 0$  and for  $\mu$ -almost all  $x, y \in M$ ,

$$p_{t+s}(x, y) = \int_M p_t(x, z) p_s(z, y) d\mu(z). \quad (2.2)$$

5. *Approximation of identity:* for any  $f \in L^2$ ,

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

We say that a heat kernel  $p_t$  is *stochastically complete* if equality takes place in (2.1), that is, for any  $t > 0$ ,

$$\int_M p_t(x, y) d\mu(y) = 1 \quad \text{for } \mu\text{-almost all } x \in M.$$

Typically a heat kernel is associated with a Markov process  $(\{X_t\}_{t \geq 0}, \{\mathbb{P}_x\}_{x \in M})$  on  $M$ , so that  $p_t(x, y)$  is the transition density of  $X_t$ , that is,

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

for any Borel set  $A \subset M$  (see Fig. 1).

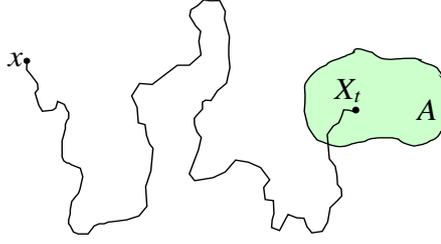


FIGURE 1. Markov process  $X_t$  hits the set  $A$

Here are some examples of heat kernels.

**Example 2.2.** The best known example of a heat kernel is the *Gauss-Weierstrass function* in  $\mathbb{R}^n$ :

$$p_t(x, y) = \frac{1}{(4\pi t)^{n/2}} \exp\left(-\frac{|x-y|^2}{4t}\right). \quad (2.3)$$

It satisfies all the conditions of Definition 2.1 provided  $\mu$  is the Lebesgue measure. This heat kernel is the transition density of the canonical Brownian motion in  $\mathbb{R}^n$ .

**Example 2.3.** The following function in  $\mathbb{R}^n$

$$p_t(x, y) = \frac{C_n}{t^n} \left(1 + \frac{|x-y|^2}{t^2}\right)^{-\frac{n+1}{2}} \quad (2.4)$$

(where  $C_n = \Gamma(\frac{n+1}{2}) / \pi^{(n+1)/2}$ ) is known on the one hand as *the Poisson kernel*, and on the other hand as *the Cauchy distribution*. It is not difficult to verify that it also satisfies Definition 2.1 (also with respect to the Lebesgue measure) and, hence, is a heat kernel. The associated Markov process is the symmetric stable process of index 1.

More examples will be mentioned in the next section.

## 2.2. Heat semigroup and Dirichlet forms

The heat kernel is an integral kernel of the heat semigroup in  $L^2$ . The heat semigroup corresponds uniquely to the Dirichlet form in  $L^2$ .

A *Dirichlet form*  $(\mathcal{E}, \mathcal{F})$  in  $L^2$  is a bilinear form  $\mathcal{E} : \mathcal{F} \times \mathcal{F} \rightarrow \mathbb{R}$  defined on a dense subspace  $\mathcal{F}$  of  $L^2$ , which satisfies in addition the following properties:

- *Positivity*:  $\mathcal{E}(f) := \mathcal{E}(f, f) \geq 0$  for any  $f \in \mathcal{F}$ .
- *Closedness*: the space  $\mathcal{F}$  is a Hilbert space with respect to the following inner product:

$$\mathcal{E}(f, g) + (f, g).$$

- *The Markov property*: if  $f \in \mathcal{F}$  then the function

$$g := \min\{1, \max\{f, 0\}\}$$

also belongs to  $\mathcal{F}$  and  $\mathcal{E}(g) \leq \mathcal{E}(f)$ . Here we have used the shorthand  $\mathcal{E}(f) := \mathcal{E}(f, f)$ .

Any Dirichlet form has the *generator*  $\mathcal{L}$ , which is a non-positive definite self-adjoint operator on  $L^2$  with domain  $\mathcal{D} \subset \mathcal{F}$  such that

$$\mathcal{E}(f, g) = (-\mathcal{L}f, g)$$

for all  $f \in \mathcal{D}$  and  $g \in \mathcal{F}$ . The generator determines the *heat semigroup*  $\{P_t\}_{t \geq 0}$  defined by  $P_t = e^{t\mathcal{L}}$ . The heat semigroup satisfies the following properties:

- $\{P_t\}_{t \geq 0}$  is *contractive* in  $L^2$ , that is  $\|P_t f\|_2 \leq \|f\|_2$  for all  $f \in L^2$  and  $t > 0$ .
- $\{P_t\}_{t \geq 0}$  is *strongly continuous*, that is, for every  $f \in L^2$ ,

$$P_t f \xrightarrow{L^2} f \text{ as } t \rightarrow 0+.$$

- $\{P_t\}_{t \geq 0}$  is *symmetric*, that is,

$$(P_t f, g) = (f, P_t g) \quad \text{for all } f, g \in L^2.$$

- $\{P_t\}_{t \geq 0}$  is *Markovian*, that is, for any  $t > 0$ ,

$$\text{if } f \geq 0 \text{ then } P_t f \geq 0, \text{ and if } f \leq 1 \text{ then } P_t f \leq 1.$$

Here and below the inequalities between  $L^2$ -functions are understood  $\mu$ -almost everywhere in  $M$ .

The form  $(\mathcal{E}, \mathcal{F})$  can be recovered from the heat semigroup as follows. For any  $t > 0$ , define a quadratic form  $\mathcal{E}_t$  on  $L^2$  as follows

$$\mathcal{E}_t(f) := \frac{1}{t} (f - P_t f, f). \quad (2.5)$$

It is easy to show that  $\mathcal{E}_t(f)$  is non-negative and is increasing as  $t$  is decreasing. In particular, it has the limit as  $t \rightarrow 0$ . It turns out that the limit is finite if and only if  $f \in \mathcal{F}$ , and, moreover,

$$\lim_{t \rightarrow 0^+} \mathcal{E}_t(f) = \mathcal{E}(f)$$

(cf. [9]). Extend  $\mathcal{E}_t$  to a bilinear form as follows

$$\mathcal{E}_t(f, g) := \frac{1}{t} (f - P_t f, g).$$

Then, for all  $f, g \in \mathcal{F}$ ,

$$\lim_{t \rightarrow 0^+} \mathcal{E}_t(f, g) = \mathcal{E}(f, g).$$

The Markovian property of the heat semigroup implies that the operator  $P_t$  preserves the inequalities between functions, which allows to use monotone limits to extend  $P_t$  from  $L^2$  to  $L^\infty$  and, in fact, to any  $L^q$ ,  $1 \leq q \leq \infty$ . Moreover, the extended operator  $P_t$  is a contraction on any  $L^q$  (cf. [15, p.33]).

Recall some more terminology from the theory of the Dirichlet form (cf. [15]). The form  $(\mathcal{E}, \mathcal{F})$  is called *conservative* if  $P_t \mathbf{1} = 1$  for every  $t > 0$ . The form  $(\mathcal{E}, \mathcal{F})$  is called *local* if  $\mathcal{E}(f, g) = 0$  for any couple  $f, g \in \mathcal{F}$  with disjoint compact supports. The form  $(\mathcal{E}, \mathcal{F})$  is called *strongly local* if  $\mathcal{E}(f, g) = 0$  for any couple  $f, g \in \mathcal{F}$  with compact supports, such that  $f \equiv \text{const}$  in an open neighborhood of  $\text{supp } g$ .

The form  $(\mathcal{E}, \mathcal{F})$  is called *regular* if  $\mathcal{F} \cap C_0(M)$  is dense both in  $\mathcal{F}$  and in  $C_0(M)$ , where  $C_0(M)$  is the space of all continuous functions with compact support in  $M$ , endowed with sup-norm. For a non-empty open  $\Omega \subset M$ , let  $\mathcal{F}(\Omega)$  be the closure of  $\mathcal{F} \cap C_0(\Omega)$  in the norm of  $\mathcal{F}$ . It is known that if  $(\mathcal{E}, \mathcal{F})$  is regular, then  $(\mathcal{E}, \mathcal{F}(\Omega))$  is also a regular Dirichlet form in  $L^2(\Omega, \mu)$ .

Assume that heat semigroup  $\{P_t\}$  of a Dirichlet form  $(\mathcal{E}, \mathcal{F})$  in  $L^2$  admits an integral kernel  $p_t$ , that is, for all  $t > 0$  and  $x \in M$ , the function  $p_t(x, \cdot)$  belongs to  $L^2$ , and the following identity holds:

$$P_t f(x) = \int_M p_t(x, y) f(y) d\mu(y), \quad (2.6)$$

for all  $f \in L^2$  and  $\mu$ -a.a.  $x \in M$ . Then the function  $p_t$  is indeed a heat kernel, as we will show below. For this reason, we also call  $p_t$  the heat kernel of the Dirichlet form  $(\mathcal{E}, \mathcal{F})$  or of the heat semigroup  $\{P_t\}$ .

Observe that if the heat kernel  $p_t$  of  $(\mathcal{E}, \mathcal{F})$  exists, then by (2.5) and (2.6),

$$\mathcal{E}_t(f) = \frac{1}{2t} \int_M \int_M (f(y) - f(x))^2 p_t(x, y) d\mu(y) d\mu(x) \quad (2.7)$$

$$+ \frac{1}{t} \int_M (1 - P_t \mathbf{1}(x)) f(x)^2 d\mu(x) \quad (2.8)$$

for any  $t > 0$  and  $f \in L^2$ .

**Proposition 2.4.** ([21]) *If  $p_t$  is the integral kernel of the heat semigroup  $\{P_t\}$ , then  $p_t$  is a heat kernel.*

*Proof.* We will verify that  $p_t$  satisfies all the conditions in Definition 2.1. Let  $t > 0$  be fixed unless otherwise is stated.

- (1) Setting  $p_{t,x} = p_t(x, \cdot)$ , we see from (2.6) that, for any  $f \in L^2$ ,

$$P_t f(x) = (p_{t,x}, f) \text{ for } \mu\text{-almost all } x \in M,$$

whence it follows that the function  $x \mapsto (p_{t,x}, f)$  is  $\mu$ -measurable in  $x$ . Let  $\{\varphi_k\}_{k=1}^{\infty}$  be an orthonormal basis of  $L^2$ . Using the identity

$$p_t(x, y) = p_{t,x}(y) = \sum_{k=1}^{\infty} (p_{t,x}, \varphi_k) \varphi_k(y),$$

we conclude that  $p_t(x, y)$  is jointly measurable in  $x, y \in M$ , because so are the functions  $(p_{t,x}, \varphi_k) \varphi_k(y)$ .

(2) By the Markovian property of  $P_t$ , for any non-negative function  $f \in L^2$ , there is a null set  $\mathcal{N}_f \subset M$  such that

$$P_t f(x) \geq 0 \quad \text{for all } x \in M \setminus \mathcal{N}_f.$$

Let  $S$  be a countable family of non-negative functions, which is dense in the cone of all non-negative functions in  $L^2$ , and set

$$\mathcal{N} = \bigcup_{f \in S} \mathcal{N}_f$$

so that  $\mathcal{N}$  is a null set. Then  $P_t f(x) \geq 0$  for all  $x \in M \setminus \mathcal{N}$  and for all  $f \in S$ . If  $f$  is any other non-negative function in  $L^2$ , then  $f$  is an  $L^2$ -limit of a sequence  $\{f_k\} \subset S$ , whence, for any  $x \in M \setminus \mathcal{N}$ ,

$$(p_{t,x}, f) = \lim_{k \rightarrow \infty} (p_{t,x}, f_k) = \lim_{k \rightarrow \infty} P_t f_k(x) \geq 0.$$

Therefore, for any  $x \in M \setminus \mathcal{N}$ , we have that  $p_{t,x} \geq 0$   $\mu$ -a.e. in  $M$ , which proves that  $p_t(x, y) \geq 0$  for  $\mu$ -a.a.  $x, y \in M$ .

Let  $K \subset M$  be compact. Then the indicator function  $\mathbf{1}_K$  belongs to  $L^2$  and is bounded by 1, whence

$$\int_K p_t(x, y) d\mu(y) = P_t \mathbf{1}_K(x) \leq 1$$

for  $\mu$ -a.a.  $x \in M$ . Choosing an increasing sequence of compact sets  $\{K_n\}_{n=1}^{\infty}$  that exhausts  $M$ , we obtain that

$$\int_M p_t(x, y) d\mu(y) = \lim_{n \rightarrow \infty} \int_{K_n} p_t(x, y) d\mu \leq 1$$

for  $\mu$ -a.a.  $x \in M$ .

Consequently, for any compact set  $K \subset M$ , we obtain by Fubini's theorem

$$\begin{aligned} \int_{K \times M} p_t(x, y) d\mu(y) d\mu(x) &= \int_K \left( \int_M p_t(x, y) d\mu(y) \right) d\mu(x) \\ &\leq \int_K d\mu(x) = \mu(K) < \infty, \end{aligned}$$

which implies that  $p_t(x, y) \in L^1_{loc}(M \times M)$ .

(3) For all  $f, g \in L^2$ , we have, again by Fubini's theorem,

$$(P_t f, g) = \int_{M \times M} p_t(x, y) f(y) g(x) d\mu(y) d\mu(x). \quad (2.9)$$

On the other hand, by the symmetry of  $P_t$ ,

$$\begin{aligned} (P_t f, g) &= (f, P_t g) = \int_M P_t g(y) f(y) d\mu(y) \\ &= \int_{M \times M} p_t(y, x) f(y) g(x) d\mu(y) d\mu(x). \end{aligned} \quad (2.10)$$

Comparing (2.9) and (2.10), we obtain  $p_t(x, y) = p_t(y, x)$  for  $\mu$ -almost all  $x, y \in M$ .

(4) Using the semigroup identity  $P_{t+s} = P_t(P_s)$  and Fubini's theorem, we obtain that, for any  $f \in L^2$  and for  $\mu$ -a.a.  $x \in M$ ,

$$\begin{aligned} P_{t+s} f(x) &= P_t(P_s f)(x) \\ &= \int_M p_t(x, z) \left( \int_M p_s(z, y) f(y) d\mu(y) \right) d\mu(z) \\ &= \int_M \left( \int_M p_t(x, z) p_s(z, y) d\mu(z) \right) f(y) d\mu(y), \end{aligned}$$

whence, for any  $g \in L^2$ ,

$$(P_{t+s}f, g) = \int_{M \times M} \left( \int_M p_t(x, z) p_s(z, y) d\mu(z) \right) f(y) g(x) d\mu(y) d\mu(x).$$

Comparing with

$$(P_{t+s}f, g) = \int_{M \times M} p_{t+s}(x, y) f(y) g(x) d\mu(y) d\mu(x),$$

we obtain (2.2).

(5) Finally, the approximation of identity property follows immediately from (2.6) and  $P_t f \xrightarrow{L^2} f$  as  $t \rightarrow 0$ .  $\square$

**Corollary 2.5.** *If  $p_t$  and  $q_t$  are two integral kernels of a heat semigroup  $\{P_t\}$ , then, for any  $t > 0$ ,*

$$p_t(x, y) = q_t(x, y) \text{ for } \mu\text{-a.a. } x, y \in M. \quad (2.11)$$

*Proof.* Similarly to (2.9), we have

$$(P_t f, g) = \int_{M \times M} q_t(x, y) f(y) g(x) d\mu(y) d\mu(x).$$

Comparing with (2.9), we obtain (2.11).  $\square$

**Remark 2.6.** Of course, not every heat semigroup possesses the heat kernel. The existence results for the heat kernel and related to this on-diagonal upper bounds can be found in [4, Theorem 2.1], [2, Propositions 4.13, 4.14], [8], [9], [11], [13], [14, Lemma 2.1.2], [17], [21], [23], [31], [40], [41], [42], [43].

### 2.3. Examples

**Example 2.7.** Let  $M$  be a connected Riemannian manifold,  $d$  be the geodesic distance, and  $\mu$  be the Riemannian measure. The Laplace-Beltrami operator  $\Delta$  on  $M$  can be made into a self-adjoint operator in  $L^2(M, \mu)$  by appropriately defining its domain. Then  $\Delta$  generates the heat semigroup  $P_t = e^{t\Delta}$ , which is associated with the local Dirichlet form  $(\mathcal{E}, \mathcal{F})$  where

$$\mathcal{E}(f) = \int_M |\nabla f|^2 d\mu, \quad \mathcal{F} = W_0^{1,2}(M).$$

The corresponding Markov process is a Brownian motion on  $M$ .

It is known that  $\{P_t\}$  has always a smooth integral kernel  $p_t(x, y)$ , which is called the heat kernel of  $M$ . Although the explicit expression of  $p_t(x, y)$  can not be given in general, there are many important classes of manifolds where  $p_t(x, y)$  admits certain upper and/or lower bounds. For example, as it was proved in [32], if  $M$  is geodesically complete and its Ricci curvature is non-negative then

$$p_t(x, y) \asymp \frac{1}{V(x, \sqrt{t})} \exp\left(-c \frac{d(x, y)^2}{t}\right), \quad x, y \in M, t > 0, \quad (2.12)$$

where  $V(x, r) = \mu(B(x, r))$  is the volume of the geodesic ball

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

The sign  $\asymp$  in (2.12) means that the ratio of the two sides is bounded from above and below by positive constants. Besides,  $c$  is a positive constant, whose value may vary in the upper and lower bounds. We will follow this convention also in the rest of the paper. The letters  $C, C', c, c'$  will always refer to positive constants, whose value is unimportant and may change at each occurrence.

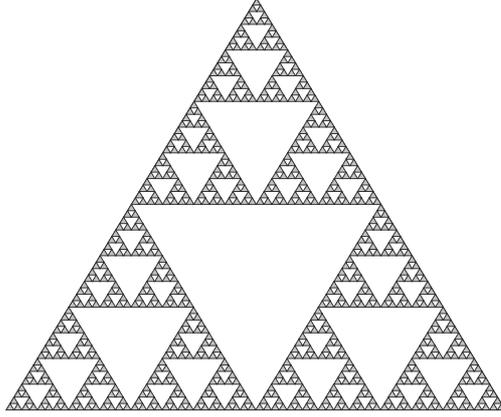
**Example 2.8.** If  $\Delta$  is a self-adjoint Laplace operator as above then the operator  $\mathcal{L} = -(-\Delta)^{\beta/2}$  (where  $0 < \beta < 2$ ) generates on  $M$  a Markov process with jumps. In particular, if  $M = \mathbb{R}^n$  then this is the symmetric stable process of index  $\beta$ , and the corresponding heat kernel admits the following estimate

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

A particular case  $\beta = 1$  was already mentioned above.

**Example 2.9.** Let  $M$  be the Sierpinski gasket<sup>1</sup> in  $\mathbb{R}^n$  (see Fig. 2).

<sup>1</sup>For the background of fractal sets including the notion of the Sierpinski gasket, see [2].

FIGURE 2. Sierpinski gasket in  $\mathbb{R}^2$ 

It is known that the Hausdorff dimension of  $M$  is equal to  $\alpha := \log(n+1)/\log 2$ . Let  $\mu$  be the  $\alpha$ -dimensional Hausdorff measure on  $M$ , which clearly possesses the same self-similarity properties as the set  $M$  itself. It is possible to construct also a self-similar local Dirichlet form on  $M$  which possesses a continuous heat kernel, that is the transition density of a natural Brownian motion on  $M$ ; moreover, the heat kernel admits the following estimate

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

where  $\beta = \log(n+3)/\log 2$  is the *walk dimension* (see [6], [16], [30]). Similar results hold also for a large family of fractal sets, including p.c.f. fractals and the Sierpinski carpet in  $\mathbb{R}^n$  (see [30] and [3]), but with different values of  $\alpha$  and  $\beta$ .

### 3. Auxiliary material on metric measure spaces

Fix a metric measure space  $(M, d, \mu)$  and define its volume function  $V(x, r)$  by

$$V(x, r) := \mu(B(x, r))$$

where  $x \in M$  and  $r > 0$ .

#### 3.1. Besov spaces

Here we now introduce function spaces  $W^{\beta/2, p}$  on  $M$ . Choose parameters  $1 \leq p < \infty$ ,  $\beta > 0$  and define the functional  $E_{\beta, p}(u)$  for all functions  $u \in L^p$  as follows:

$$E_{\beta, p}(u) = \sup_{0 < r \leq 1} r^{-p\beta/2} \int_M \left[ \frac{1}{V(x, r)} \int_{B(x, r)} |u(y) - u(x)|^p d\mu(y) \right] d\mu(x). \quad (3.1)$$

For simplicity, if  $p = 2$ , we denote by

$$E_\beta(u) := E_{\beta, 2}(u).$$

The *Besov* space  $W^{\beta/2, p}$  is defined by

$$W^{\beta/2, p} := \{u \in L^p : E_{\beta, p}(u) < \infty\} \quad (3.2)$$

with the norm

$$\|u\|_{W^{\beta/2, p}} := \|u\|_p + E_{\beta, p}(u)^{1/p}.$$

For Ahlfors regular<sup>2</sup> measures  $\mu$ , the Besov space  $W^{\beta/2, p}$  was introduced in [28, 34, 22] although using different notation.

It is not difficult to verify that for any  $1 \leq p < \infty$  and  $\beta > 0$ , the space  $W^{\beta/2, p}$  is a Banach space. Note that the space  $W^{\beta/2, p}$  decreases as  $\beta$  increases; it may happen that this space becomes trivial for large enough  $\beta$ . For example,  $W^{\beta/2, 2}(\mathbb{R}^n) = \{0\}$  for  $\beta > 2$ .

<sup>2</sup>A measure  $\mu$  on a metric space  $(M, d)$  is said to be *Ahlfors-regular* if there exist  $\alpha, c > 0$  such that  $V(x, r) \asymp r^\alpha$  for all balls  $B(x, r)$  in  $M$  with  $r \in (0, 1)$ .

Define the *critical Besov exponent*  $\beta^*$  by

$$\beta^* := \sup \left\{ \beta > 0 : W^{\beta/2,2} \text{ is dense in } L^2(M, \mu) \right\}. \quad (3.3)$$

**Lemma 3.1.** *We have  $\beta^* \geq 2$ .*

*Proof.* It suffices to show that  $W^{1,2}$  is dense in  $L^2 = L^2(M, \mu)$ . Let  $u$  be a Lipschitz function with a bounded support  $A$  and let  $A_r$  be the closed  $r$ -neighborhood of  $A$ . If  $L$  is the Lipschitz constant of  $u$ , then

$$\begin{aligned} E_2(u) &= \sup_{0 < r \leq 1} r^{-2} \int_{A_r} \frac{1}{V(x, r)} \int_{B(x, r)} |u(y) - u(x)|^2 d\mu(y) d\mu(x) \\ &\leq \sup_{0 < r \leq 1} r^{-2} \int_{A_r} L^2 r^2 d\mu(x) \\ &\leq L^2 \mu(A_1). \end{aligned}$$

It follows that  $E_2(u) < \infty$  and, hence,  $u \in W^{1,2}$ . We are left to show that the class  $Lip$  of all Lipschitz functions with bounded supports is dense in  $L^2$ . Indeed, let now  $A$  be any bounded closed subset of  $M$ . For any positive integer  $n$ , consider the function on  $M$

$$f_n(x) = (1 - nd(x, A))_+,$$

which is Lipschitz and is supported in  $A_{1/n}$ . Clearly,  $f_n \rightarrow 1_A$  in  $L^2$  as  $n \rightarrow \infty$ , whence it follows that  $1_A \in \overline{Lip}$ , where the bar means closure in  $L^2$ . Since the linear combinations of the indicator functions of bounded closed sets form a dense subset in  $L^2$ , it follows that  $\overline{Lip} = L^2$ , which was to be proved.  $\square$

### 3.2. Doubling condition and reverse doubling condition

The measure  $\mu$  on  $M$  is said to be *doubling* if there is a constant  $C_D \geq 1$  such that

$$V(x, 2r) \leq C_D V(x, r) \quad (3.4)$$

for all  $x \in M$  and  $r > 0$ .

**Proposition 3.2.** *If (3.4) holds on  $M$ , then there exists  $\alpha > 0$  depending only on the doubling constant  $C_D$  such that*

$$\frac{V(x, R)}{V(y, r)} \leq C_D \left( \frac{d(x, y) + R}{r} \right)^\alpha \quad \text{for all } x, y \in M \text{ and } 0 < r \leq R. \quad (\text{VD})$$

Hence, the inequality of Proposition 3.2 can be used as an alternative definition of the doubling property of  $\mu$  and will be referred to as (VD) (volume doubling). The advantage of this definition is that it introduces a parameter  $\alpha$  that will be frequently used.

*Proof.* If  $x = y$ , then  $R \leq 2^n r$  where

$$n = \left\lceil \log_2 \frac{R}{r} \right\rceil \leq \log_2 \frac{R}{r} + 1,$$

whence, it follows from (3.4) that

$$\frac{V(x, R)}{V(x, r)} \leq \frac{V(x, 2^n r)}{V(x, r)} \leq (C_D)^n \leq (C_D)^{\log_2 \frac{R}{r} + 1} = C_D \left( \frac{R}{r} \right)^{\log_2 C_D}. \quad (3.5)$$

If  $x \neq y$ , then  $B(x, R) \subset B(y, R + r_0)$  where  $r_0 = d(x, y)$ . By (3.5),

$$\frac{V(x, R)}{V(y, r)} \leq \frac{V(y, R + r_0)}{V(y, r)} \leq C_D \left( \frac{R + r_0}{r} \right)^{\log_2 C_D},$$

which finishes the proof.  $\square$

The measure  $\mu$  is said to be *reverse doubling* if there exist positive constants  $\alpha'$  and  $c$  such that

$$\frac{V(x, R)}{V(x, r)} \geq c \left( \frac{R}{r} \right)^{\alpha'} \quad \text{for all } x \in M \text{ and } 0 < r \leq R. \quad (\text{RVD})$$

**Proposition 3.3.** *If  $(M, d)$  is connected and  $\mu$  satisfies (3.4), then there exist positive constants  $\alpha'$  and  $c$  such that (RVD) holds, provided  $B(x, R)^c$  is non-empty.*

*Proof.* The condition  $\overline{B(x, R)^c} \neq \emptyset$  implies that

$$B(x, \rho') \setminus B(x, \rho) \neq \emptyset \quad (3.6)$$

for all  $0 < \rho < R$  and  $\rho' > \rho$ . Indeed, otherwise  $M$  splits into disjoint union of two open sets:  $B(x, \rho)$  and  $\overline{B(x, \rho)^c}$ . Since  $M$  is connected, the set  $\overline{B(x, \rho)^c}$  must be empty, which contradicts the non-emptiness of  $B(x, R)^c$ .

If  $0 < \rho \leq R/2$ , then we have by (3.6)

$$B\left(x, \frac{5}{3}\rho\right) \setminus B\left(x, \frac{4}{3}\rho\right) \neq \emptyset.$$

Let  $y$  be a point in this annulus. It follows from (VD) that

$$V(x, \rho) \leq CV(y, \rho/3)$$

for some constant  $C > 0$ , whence

$$V(x, 2\rho) \geq V(x, \rho) + V(y, \rho/3) \geq (1 + \varepsilon)V(x, \rho), \quad (3.7)$$

where  $\varepsilon = C^{-1}$ .

For any  $0 < r \leq R$ , we have that  $2^n r \leq R$  where

$$n := \left\lceil \log_2 \frac{R}{r} \right\rceil \geq \log_2 \frac{R}{r} - 1.$$

For any  $0 \leq k \leq n - 1$ , we have  $2^k r \leq R/2$ , and whence by (3.7),

$$V(x, 2^{k+1}r) \geq (1 + \varepsilon)V(x, 2^k r).$$

Iterating this inequality, we obtain

$$\begin{aligned} \frac{V(x, R)}{V(x, r)} &\geq \frac{V(x, 2^n r)}{V(x, r)} \geq (1 + \varepsilon)^n \\ &\geq (1 + \varepsilon)^{\log_2 \frac{R}{r} - 1} = (1 + \varepsilon)^{-1} \left(\frac{R}{r}\right)^{\log_2(1 + \varepsilon)}, \end{aligned}$$

thus proving (RVD).  $\square$

**Remark 3.4.** As one can see from the argument after (3.7), measure  $\mu$  is reverse doubling whenever the following inequality holds

$$V(x, Cr) \geq (1 + \varepsilon)V(x, r) \quad (3.8)$$

for some  $C > 1$ ,  $\varepsilon > 0$  and all  $x \in M$ ,  $r > 0$ .

**Corollary 3.5.** *Assume that  $(M, d)$  is connected and  $\mu$  satisfies (VD). Then*

$$\mu(M) = \infty \Leftrightarrow \text{diam}(M) = \infty \Leftrightarrow (\text{RVD}).$$

*Proof.* If  $\mu(M) = \infty$ , then  $\text{diam}(M) = \infty$ ; indeed, otherwise  $M$  would be a ball of a finite radius and its measure would be finite by (VD). If  $\text{diam}(M) = \infty$ , then  $B^c(x, R) \neq \emptyset$  for any ball  $B(x, R)$ , and (RVD) holds by Proposition 3.3. Finally, (RVD) implies  $\mu(M) = \infty$  by letting  $R \rightarrow \infty$  in (RVD).  $\square$

## 4. Consequences of heat kernel estimates

We give here some consequences of the heat kernel estimates

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

for all  $t > 0$  and  $\mu$ -almost all  $x, y \in M$ . Functions  $\Phi_1(s)$  and  $\Phi_2(s)$  are always assumed to be non-negative and monotone decreasing on  $[0, +\infty)$ , the constant  $\beta$  is positive.

We prove that

- The lower estimate of the heat kernel implies that
  - the measure  $\mu$  is doubling;
  - the space  $\mathcal{F}$  embeds in  $W^{\beta/2, 2}$ ;
  - the lower tail function  $\Phi_1(s)$  is controlled from above by a negative power of  $s$ .
- The upper estimate of the heat kernel implies that

- the space  $W^{\beta/2,2}$  embeds in  $\mathcal{F}$ ;
- if the Dirichlet form is non-local then the *upper tail function*  $\Phi_2(s)$  is controlled from below by a negative power of  $s$  (for large  $s$ ).

#### 4.1. Consequences of lower bound

Let  $p_t$  be a heat kernel on a metric measure space  $(M, d, \mu)$ . Consider the lower estimate of  $p_t$  of the form:

$$p_t(x, y) \geq \frac{1}{V(x, t^{1/\beta})} \Phi_1\left(\frac{d(x, y)}{t^{1/\beta}}\right) \quad (4.1)$$

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

**Lemma 4.1.** *Assume that the heat kernel  $p_t$  satisfies the lower bound (4.1). If  $\Phi_1(s_0) > 0$  for some  $s_0 > 1$ , then  $\mu$  is doubling.*

*Proof.* Fix  $r, t > 0$  and consider the following integral

$$\int_{B(x,r)} p_t(x, y) d\mu(y) := \int_M p_t(x, y) \mathbf{1}_{B(x,r)}(y) d\mu(y).$$

The right-hand side is understood as follows: the function

$$F(x, y) := p_t(x, y) \mathbf{1}_{B(x,r)}(y)$$

is measurable jointly in  $x, y$  so that, by Fubini's theorem, the integral

$$\int_M p_t(x, y) \mathbf{1}_{B(x,r)}(y) d\mu(y)$$

is well-defined for  $\mu$ -almost all  $x \in M$  and is a measurable function of  $x$ . Choose any pointwise version of  $p_t(x, y)$  as a function of  $x, y$ . By Fubini's theorem, there is a subset  $M_0 \in M$  of a full measure such that, for any  $x \in M_0$ , the function  $p_t(x, y)$  is measurable in  $y$  and the inequalities (4.1) and (2.1) hold for  $\mu$ -a.a.  $y \in M$ . It follows that, for all  $x \in M_0$ ,

$$\int_{B(x,r)} p_t(x, y) d\mu(y) \leq 1 \quad (4.2)$$

whence

$$\frac{1}{V(x, r)} \geq \inf_{y \in B(x,r)} p_t(x, y).$$

On the other hand, we have by (4.1)

$$\inf_{y \in B(x,r)} p_t(x, y) \geq \frac{1}{V(x, t^{1/\beta})} \Phi_1\left(\frac{r}{t^{1/\beta}}\right),$$

which together with the previous estimate gives

$$\frac{V(x, r)}{V(x, t^{1/\beta})} \leq \frac{1}{\Phi_1(r/t^{1/\beta})}.$$

Setting here  $t = (r/s_0)^\beta$  we obtain

$$\frac{V(x, r)}{V(x, r/s_0)} \leq \frac{1}{\Phi_1(s_0)}. \quad (4.3)$$

Since  $s_0 > 1$  and  $\Phi_1(s_0) > 0$ , (4.3) implies that measure  $\mu$  is doubling.  $\square$

**Lemma 4.2.** *Assume that the heat kernel  $p_t$  satisfies the lower bound (4.1) with  $\Phi_1(s_0) > 0$  for some  $s_0 \geq 1$ . Then, there is a constant  $c > 0$  such that for all  $u \in L^2(M)$ ,*

$$\mathcal{E}(u) \geq c E_\beta(u). \quad (4.4)$$

*Consequently, the space  $\mathcal{F}$  embeds into  $W^{\beta/2,2}$ .*

*Proof.* Let  $t, r > 0$ . It follows from (2.7) and the lower bound (4.1) that

$$\begin{aligned} \mathcal{E}(u) &\geq \mathcal{E}_t(u) \geq \frac{1}{2t} \int_M \int_{B(x,r)} (u(y) - u(x))^2 p_t(x, y) d\mu(y) d\mu(x) \\ &\geq \frac{1}{2t} \Phi_1\left(\frac{r}{t^{1/\beta}}\right) \int_M \left( \frac{1}{V(x, t^{1/\beta})} \int_{B(x,r)} (u(y) - u(x))^2 d\mu(y) \right) d\mu(x), \end{aligned}$$

where we have used the monotonicity of  $\Phi_1$ . Choosing  $t = (r/s_0)^\beta$  and noticing that  $V(x, r/s_0) \leq V(x, r)$  by  $s_0 \geq 1$ , we obtain

$$\mathcal{E}(u) \geq \frac{s_0^\beta}{2r^\beta} \Phi_1(s_0) \int_M \left( \frac{1}{V(x, r)} \int_{B(x, r)} (u(y) - u(x))^2 d\mu(y) \right) d\mu(x),$$

whence, by taking supremum in  $r$ ,

$$\mathcal{E}(u) \geq \frac{1}{2} s_0^\beta \Phi_1(s_0) E_\beta(u),$$

thus proving (4.4).  $\square$

Finally, we give another consequences of the lower bound (4.1) of the heat kernel.

**Lemma 4.3.** *Assume that the heat kernel  $p_t$  satisfies the lower bound (4.1). If  $\mu$  satisfies the reverse doubling property (RVD), then there is  $c > 0$  such that*

$$\Phi_1(s) \leq c(1+s)^{-(\alpha'+\beta)} \quad \text{for all } s > 0, \quad (4.5)$$

where  $\alpha'$  is the same as in (RVD).

*Proof.* Following [24], let  $u \in L^2$  be a non-constant function. Choose a ball  $B(x_0, R)$  where  $u$  is non-constant and let  $a > b$  be two real values such that the sets

$$A = \{x \in B(x_0, R) : u(x) \geq a\} \quad \text{and} \quad B = \{x \in B(x_0, R) : u(x) \leq b\}$$

have positive measures (see Fig. 3).

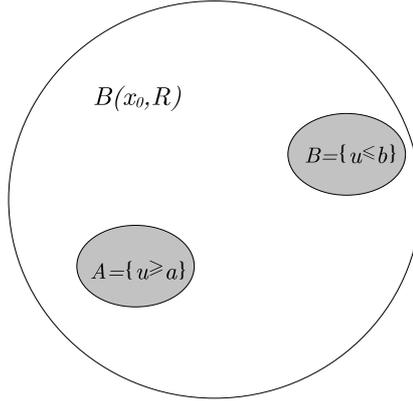


FIGURE 3. Sets  $A$  and  $B$

It follows from (2.7) that

$$\begin{aligned} \mathcal{E}(u) &\geq \frac{1}{2t} \int_M \int_M (u(y) - u(x))^2 p_t(x, y) d\mu(y) d\mu(x) \\ &\geq \frac{1}{2t} \int_A \int_B (a - b)^2 \frac{1}{V(x, t^{1/\beta})} \Phi_1\left(\frac{2R}{t^{1/\beta}}\right) d\mu(y) d\mu(x) \\ &= \frac{(a - b)^2}{2t} \Phi_1\left(\frac{2R}{t^{1/\beta}}\right) \mu(B) \int_A \frac{1}{V(x, t^{1/\beta})} d\mu(x). \end{aligned}$$

For  $x \in A$ , we have that  $B(x, R) \subset B(x_0, 3R)$ , and hence, for small enough  $t > 0$ ,

$$\frac{1}{V(x, t^{1/\beta})} = \frac{1}{V(x, R)} \cdot \frac{V(x, R)}{V(x, t^{1/\beta})} \geq \frac{1}{V(x_0, 3R)} \cdot c \left(\frac{R}{t^{1/\beta}}\right)^{\alpha'},$$

where we have used the reverse doubling property (RVD). Therefore, for small  $t > 0$ ,

$$\mathcal{E}(u) \geq \frac{c'(a - b)^2}{V(x_0, 3R)R^\beta} \mu(A)\mu(B) \left(\frac{2R}{t^{1/\beta}}\right)^{\alpha'+\beta} \Phi_1\left(\frac{2R}{t^{1/\beta}}\right).$$

If (4.5) fails, then there exists a sequence  $\{s_k\}$  with  $s_k \rightarrow \infty$  as  $k \rightarrow \infty$  such that

$$s_k^{\alpha'+\beta} \Phi_1(s_k) \rightarrow \infty \quad \text{as } k \rightarrow \infty.$$

Choose  $t_k$  such that  $s_k = 2R/t_k^{1/\beta}$ . Then

$$\left(\frac{2R}{t_k^{1/\beta}}\right)^{\alpha'+\beta} \Phi_1\left(\frac{2R}{t_k^{1/\beta}}\right) = s_k^{\alpha'+\beta} \Phi_1(s_k) \rightarrow \infty$$

as  $k \rightarrow \infty$ , and hence  $\mathcal{E}(u) = \infty$ . Hence, we see that  $\mathcal{F}$  consists of only constants. Since  $\mathcal{F}$  is dense in  $L^2$ , it follows that  $L^2$  also consists of constants only. Hence, there is a point  $x \in M$  with a positive mass, that is,  $\mu(\{x\}) > 0$ . Then (2.1) implies that, for all  $t > 0$ ,

$$p_t(x, x) \leq \frac{1}{\mu(\{x\})}. \quad (4.6)$$

However, by (RVD), we have  $V(x, r) \rightarrow 0$  as  $r \rightarrow 0$ , which together with (4.1) implies that  $p_t(x, x) \rightarrow \infty$  as  $t \rightarrow 0$ , thus contradicting (4.6).  $\square$

**Remark 4.4.** *The last argument in the above proof can be stated as follows. If (RVD) is satisfied and (4.1) holds with a function  $\Phi_1$  such that  $\Phi_1(0) > 0$ , then  $\mu(\{x\}) = 0$  for all  $x \in M$ . This simple observation will also be used below.*

#### 4.2. Consequences of upper bound

Consider the upper estimate of  $p_t$  of the form:

$$p_t(x, y) \leq \frac{1}{V(x, t^{1/\beta})} \Phi_2\left(\frac{d(x, y)}{t^{1/\beta}}\right) \quad (4.7)$$

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

**Lemma 4.5.** *Assume that  $\mu$  satisfies both (VD) and (RVD), and that the heat kernel  $p_t$  is stochastically complete and satisfies the upper bound (4.7) with*

$$\int_0^\infty s^{\alpha+\beta-1} \Phi_2(s) ds < \infty, \quad (4.8)$$

where  $\alpha$  is the same as in (VD). Then, there is a constant  $c > 0$  such that for all  $u \in L^2(M)$ ,

$$\mathcal{E}(u) \leq c E_\beta(u). \quad (4.9)$$

Consequently, the space  $W^{\beta/2, 2}$  embeds into  $\mathcal{F}$ .

*Proof.* Fix  $t \in (0, 1)$  and let  $n$  be the smallest negative integer such that  $2^{n+1} \geq t^{1/\beta}$ . Since  $p_t$  is stochastically complete, we have that for any  $t > 0$ ,

$$\mathcal{E}_t(u) = \frac{1}{2t} \int_M \int_M (u(x) - u(y))^2 p_t(x, y) d\mu(y) d\mu(x) = A_0(t) + A_1(t) + A_2(t) \quad (4.10)$$

where

$$A_0(t) : = \frac{1}{2t} \int_M \int_{B(x, 1)^c} (u(x) - u(y))^2 p_t(x, y) d\mu(y) d\mu(x), \quad (4.11)$$

$$A_1(t) : = \frac{1}{2t} \int_M \int_{B(x, 1) \setminus B(x, 2^n)} (u(x) - u(y))^2 p_t(x, y) d\mu(y) d\mu(x), \quad (4.12)$$

$$A_2(t) : = \frac{1}{2t} \int_M \int_{B(x, 2^n)} (u(x) - u(y))^2 p_t(x, y) d\mu(y) d\mu(x). \quad (4.13)$$

Observing that by (VD)

$$\frac{V(x, 2^{k+1})}{V(x, t^{1/\beta})} \leq C \left(\frac{2^{k+1}}{t^{1/\beta}}\right)^\alpha \quad \text{for all } k \geq n, \quad (4.14)$$

and using (4.7), we obtain

$$\begin{aligned}
\int_{B(x,1)^c} p_t(x,y) d\mu(y) &\leq \sum_{k=0}^{\infty} \int_{B(x,2^{k+1}) \setminus B(x,2^k)} \frac{1}{V(x,t^{1/\beta})} \Phi_2\left(\frac{2^k}{t^{1/\beta}}\right) d\mu(y) \\
&\leq C \sum_{k=0}^{\infty} \frac{V(x,2^{k+1})}{V(x,t^{1/\beta})} \Phi_2\left(\frac{2^k}{t^{1/\beta}}\right) \\
&\leq C' \sum_{k=0}^{\infty} \left(\frac{2^k}{t^{1/\beta}}\right)^\alpha \Phi_2\left(\frac{2^k}{t^{1/\beta}}\right) \\
&\leq C' \int_{\frac{1}{2}t^{-1/\beta}}^{\infty} s^{\alpha-1} \Phi_2(s) ds
\end{aligned} \tag{4.15}$$

$$\leq ct \int_{\frac{1}{2}t^{-1/\beta}}^{\infty} s^{\alpha+\beta-1} \Phi_2(s) ds. \tag{4.16}$$

Applying the elementary inequality  $(a-b)^2 \leq 2(a^2 + b^2)$ , we obtain from (4.11)

$$\begin{aligned}
A_0(t) &\leq \frac{1}{t} \int_M \int_{B(x,1)^c} (u(x)^2 + u(y)^2) p_t(x,y) d\mu(y) d\mu(x) \\
&= \frac{2}{t} \int_M u(x)^2 \left( \int_{B(x,1)^c} p_t(x,y) d\mu(y) \right) d\mu(x) \\
&\leq 2c \|u\|_2^2 \int_{\frac{1}{2}t^{-1/\beta}}^{\infty} s^{\alpha+\beta-1} \Phi_2(s) ds \\
&= o(1) \|u\|_2^2 \text{ as } t \rightarrow 0,
\end{aligned} \tag{4.17}$$

where we have used (4.8). It follows that

$$\lim_{t \rightarrow 0^+} A_0(t) = 0. \tag{4.18}$$

By (4.7) and (4.14), we obtain that, for  $0 > k \geq n$ ,

$$\begin{aligned}
&\int_{B(x,2^{k+1}) \setminus B(x,2^k)} (u(x) - u(y))^2 p_t(x,y) d\mu(y) \\
&\leq \frac{1}{V(x,t^{1/\beta})} \Phi_2\left(\frac{2^k}{t^{1/\beta}}\right) \int_{B(x,2^{k+1}) \setminus B(x,2^k)} (u(x) - u(y))^2 d\mu(y) \\
&\leq c \left(\frac{2^{k+1}}{t^{1/\beta}}\right)^\alpha \Phi_2\left(\frac{2^k}{t^{1/\beta}}\right) \frac{1}{V(x,2^{k+1})} \int_{B(x,2^{k+1})} (u(x) - u(y))^2 d\mu(y).
\end{aligned}$$

By the definition (3.1) of  $E_\beta$ , for all  $k < 0$ ,

$$\int_M \frac{1}{V(x,2^{k+1})} \int_{B(x,2^{k+1})} (u(x) - u(y))^2 d\mu(y) d\mu(x) \leq (2^{k+1})^\beta E_\beta(u). \tag{4.19}$$

Therefore, we obtain

$$\begin{aligned}
A_1(t) &= \frac{1}{2t} \sum_{n \leq k < 0} \int_M \int_{B(x,2^{k+1}) \setminus B(x,2^k)} (u(x) - u(y))^2 p_t(x,y) d\mu(y) d\mu(x) \\
&\leq \frac{1}{2t} \sum_{n \leq k < 0} c \left(\frac{2^{k+1}}{t^{1/\beta}}\right)^\alpha \Phi_2\left(\frac{2^k}{t^{1/\beta}}\right) \int_M \frac{1}{V(x,2^{k+1})} \int_{B(x,2^{k+1})} (u(x) - u(y))^2 d\mu(y) d\mu(x)
\end{aligned} \tag{4.20}$$

$$\begin{aligned}
&\leq c \sum_{n \leq k < 0} \left(\frac{2^{k+1}}{t^{1/\beta}}\right)^{\alpha+\beta} \Phi_2\left(\frac{2^k}{t^{1/\beta}}\right) E_\beta(u) \\
&\leq c E_\beta(u) \int_0^\infty s^{\alpha+\beta-1} \Phi_2(s) ds,
\end{aligned} \tag{4.21}$$

where the latter integral converges due to (4.8).

For  $k < n$ , we have  $2^{k+1} < t^{1/\beta}$  whence by (RVD)

$$\frac{V(x, 2^{k+1})}{V(x, t^{1/\beta})} \leq c \left( \frac{2^{k+1}}{t^{1/\beta}} \right)^{\alpha'} . \quad (4.22)$$

Similarly to the estimate of  $A_1$ , we obtain

$$\begin{aligned} A_2(t) &= \frac{1}{2t} \sum_{k < n} \int_M \int_{B(x, 2^{k+1}) \setminus B(x, 2^k)} (u(x) - u(y))^2 p_t(x, y) d\mu(y) d\mu(x) \\ &\leq c \sum_{k < n} \left( \frac{2^{k+1}}{t^{1/\beta}} \right)^{\alpha' + \beta} \Phi_2 \left( \frac{2^k}{t^{1/\beta}} \right) E_\beta(u) \\ &\leq c E_\beta(u) \int_0^2 s^{\alpha' + \beta - 1} \Phi_2(s) ds, \end{aligned} \quad (4.23)$$

where the latter integral converges at 0 due to  $\alpha' + \beta > 0$ . It follows from (4.10), (4.18), (4.21) and (4.23) that

$$\mathcal{E}(u) = \lim_{t \rightarrow 0^+} \mathcal{E}_t(u) = \lim_{t \rightarrow 0^+} (A_0(t) + A_1(t) + A_2(t)) \leq C E_\beta(u),$$

which finishes the proof.  $\square$

**Lemma 4.6.** *Assume that  $\mu$  satisfies (VD) and that the heat kernel  $p_t$  satisfies the upper bound (4.7). Then, either  $(\mathcal{E}, \mathcal{F})$  is local, or there is  $c > 0$  such that*

$$\Phi_2(s) \geq c(1+s)^{-(\alpha+\beta)} \quad \text{for all } s > 0. \quad (4.24)$$

*Proof.* Let  $u, v \in \mathcal{F}$  be functions with disjoint compact supports  $A = \text{supp } u$  and  $B = \text{supp } v$  (see Fig. 4).

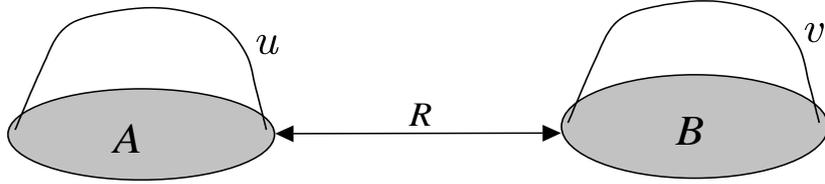


FIGURE 4. Functions  $u$  and  $v$

Noticing that  $(u, v) = 0$ , we obtain, for any  $t > 0$ ,

$$\begin{aligned} \mathcal{E}_t(u, v) &= \frac{1}{t} (u, v - P_t v) \\ &= -\frac{1}{t} (u, P_t v) \\ &= -\frac{1}{t} \int_A u(x) \left( \int_B v(y) p_t(x, y) d\mu(y) \right) d\mu(x). \end{aligned}$$

Setting  $R = d(A, B) > 0$  and using (4.7), we obtain

$$|\mathcal{E}_t(u, v)| \leq \frac{1}{t} \Phi_2 \left( \frac{R}{t^{1/\beta}} \right) \|v\|_1 \int_A \frac{|u(x)|}{V(x, t^{1/\beta})} d\mu(x). \quad (4.25)$$

Choose any fixed point  $x_0 \in A$  and let  $\text{diam}(A) = r$ . Then, using (VD), we see that, for all  $x \in A$  and small  $t > 0$ ,

$$\begin{aligned} \frac{1}{V(x, t^{1/\beta})} &= \frac{1}{V(x_0, r)} \frac{V(x_0, r)}{V(x, t^{1/\beta})} \\ &\leq \frac{c}{V(x_0, r)} \left( \frac{d(x_0, x) + r}{t^{1/\beta}} \right)^\alpha \leq \frac{c}{V(x_0, r)} \left( \frac{2r}{t^{1/\beta}} \right)^\alpha. \end{aligned}$$

Therefore, by (4.25),

$$\begin{aligned} |\mathcal{E}_t(u, v)| &\leq \frac{1}{t} \Phi_2 \left( \frac{R}{t^{1/\beta}} \right) \|v\|_1 \frac{c}{V(x_0, r)} \left( \frac{2r}{t^{1/\beta}} \right)^\alpha \|u\|_1 \\ &= \frac{c(2r)^\alpha}{V(x_0, r) R^{\alpha+\beta}} \|u\|_1 \|v\|_1 \left( \frac{R}{t^{1/\beta}} \right)^{\alpha+\beta} \Phi_2 \left( \frac{R}{t^{1/\beta}} \right). \end{aligned}$$

If (4.24) fails, then there exists a sequence  $\{s_k\}$  such that  $s_k \rightarrow \infty$  as  $k \rightarrow \infty$ , and

$$s_k^{\alpha+\beta} \Phi_2(s_k) \rightarrow 0.$$

Letting  $t_k > 0$  such that  $s_k = R/t_k^{1/\beta}$ , we obtain that

$$|\mathcal{E}_{t_k}(u, v)| \rightarrow 0 \quad \text{as } k \rightarrow \infty,$$

showing that  $\mathcal{E}(u, v) = 0$ . Hence, the  $(\mathcal{E}, \mathcal{F})$  is local, which was to be proved.  $\square$

### 4.3. Walk dimension

Here we obtain certain consequence of a two-sided estimate

$$\frac{1}{V(x, t^{1/\beta})} \Phi_1 \left( \frac{d(x, y)}{t^{1/\beta}} \right) \leq p_t(x, y) \leq \frac{1}{V(x, t^{1/\beta})} \Phi_2 \left( \frac{d(x, y)}{t^{1/\beta}} \right). \quad (4.26)$$

The parameter  $\beta$  from (4.26) is called the *walk dimension* of the associated Markov process.

**Theorem 4.7.** *Assume that  $\mu$  satisfies both (VD) and (RVD). Let a heat kernel  $p_t(x, y)$  be stochastically complete and satisfy (4.26) where  $\Phi_1(s_0) > 0$  for some  $s_0 \geq 1$  and*

$$\int_0^\infty s^{\alpha+\beta+\varepsilon} \Phi_2(s) \frac{ds}{s} < \infty \quad (4.27)$$

for some  $\varepsilon > 0$ . Then  $\beta = \beta^*$  where  $\beta^*$  is the critical Besov exponent defined in (3.3).

**Remark 4.8.** *Assuming in addition that  $\Phi_1(s_0) > 0$  for some  $s_0 > 1$  allows to drop (VD) from the hypothesis, thanks to Lemma 4.1. If one assumes on top of that, that the metric space  $(M, d)$  is connected and has infinite diameter then (RVD) follows from (VD) by Corollary 3.5. Hence, in this case (RVD) can be dropped from the assumptions as well.*

*Proof.* By Lemma 4.2, we have the inclusion  $\mathcal{F} \subset W^{\beta/2, 2}$ . Since  $\mathcal{F}$  is always dense in  $L^2$ , we conclude that  $W^{\beta/2, 2}$  is dense in  $L^2$ , whence  $\beta^* \geq \beta$ .

To prove the opposite inequality, it suffices to verify that, for any  $\beta' > \beta$ , the space  $W^{\beta'/2, 2}$  is not dense in  $L^2$ . We can assume that  $\beta' - \beta$  is sufficiently small so that the condition (4.27) holds with  $\varepsilon = \beta' - \beta$ .

Let us show that  $u \in W^{\beta'/2, 2}$  implies  $\mathcal{E}(u) = 0$ . We use again the decomposition

$$\mathcal{E}_t(u) = A_0(t) + A_1(t) + A_2(t)$$

where  $A_i(t)$  are defined in (4.11)–(4.13). As in the proof of Lemma 4.5, we have

$$\lim_{t \rightarrow 0} A_0(t) = 0.$$

Let us estimate  $A_1(t)$  similarly to the proof of Lemma 4.5 (and using the same notation), but use  $E_{\beta'}$  instead of  $E_\beta$ . Indeed, using instead of (4.19) the inequality

$$\int_M \frac{1}{V(x, 2^{k+1})} \int_{B(x, 2^{k+1})} (u(x) - u(y))^2 d\mu(y) d\mu(x) \leq (2^{k+1})^{\beta'} E_{\beta'}(u), \quad (4.28)$$

we obtain from (4.20) that

$$\begin{aligned} A_1(t) &\leq ct^{\frac{\beta'}{\beta}-1} \sum_{n \leq k < 0} \left( \frac{2^{k+1}}{t^{1/\beta}} \right)^{\alpha+\beta'} \Phi_2 \left( \frac{2^k}{t^{1/\beta}} \right) E_{\beta'}(u) \\ &\leq ct^{\frac{\beta'}{\beta}-1} E_{\beta'}(u) \int_0^\infty s^{\alpha+\beta'-1} \Phi_2(s) ds \end{aligned} \quad (4.29)$$

where the integral converges due to (4.27). In the same way, one obtains

$$A_2(t) \leq ct^{\frac{\beta'}{\beta}-1} E_{\beta'}(u) \int_0^2 s^{\alpha'+\beta'-1} \Phi_2(s) ds.$$

Putting together all the estimates, we obtain

$$\mathcal{E}_t(u) \leq A_0(t) + Ct^{\frac{\beta'}{\beta}-1}E_{\beta'}(u) \rightarrow 0 \quad \text{as } t \rightarrow 0,$$

whence

$$\mathcal{E}(u) = \lim_{t \rightarrow 0} \mathcal{E}_t(u) = 0.$$

Since  $\mathcal{E}_t(u) \leq \mathcal{E}(u)$ , this implies back that  $\mathcal{E}_t(u) \equiv 0$  for all  $t > 0$ .

On the other hand, it follows from (2.7) and the lower bound in (4.26) that

$$\begin{aligned} \mathcal{E}_t(u) &\geq \frac{1}{2t} \int \int_{\{d(x,y) \leq s_0 t^{1/\beta}\}} (u(y) - u(x))^2 p_t(x,y) d\mu(y) d\mu(x) \\ &\geq \frac{\Phi_1(s_0)}{2t} \int \int_{\{d(x,y) \leq s_0 t^{1/\beta}\}} \frac{(u(x) - u(y))^2}{V(x, t^{1/\beta})} d\mu(y) d\mu(x), \end{aligned}$$

which yields  $u(x) = u(y)$  for  $\mu$ -almost all  $x, y$  such that  $d(x, y) \leq s_0 t^{1/\beta}$ . Since  $t$  is arbitrary, we conclude that  $u$  is a constant function.

Hence, we have shown that the space  $W^{\beta'/2, 2}$  consists of constants. However, it follows from Remark 4.4 that the constant functions are not dense in  $L^2$ , which finishes the proof.  $\square$

#### 4.4. Consequences of two-sided estimates (non-local case)

Lemmas 4.3 and 4.6 of the previous subsections imply immediately the following.

**Theorem 4.9.** *Assume that the metric measure space  $(M, d, \mu)$  satisfies (VD) and (RVD). Let  $\{p_t\}$  be a heat kernel on  $M$  such that, for all  $t > 0$  and almost all  $x, y \in M$ ,*

$$\frac{C'_1}{V(x, t^{1/\beta})} \Phi\left(C_1 \frac{d(x, y)}{t^{1/\beta}}\right) \leq p_t(x, y) \leq \frac{C'_2}{V(x, t^{1/\beta})} \Phi\left(C_2 \frac{d(x, y)}{t^{1/\beta}}\right) \quad (4.30)$$

where  $C_1, C'_1, C_2, C'_2$  are positive constants. Then either the associated Dirichlet form  $\mathcal{E}$  is local or

$$c_1(1+s)^{-(\alpha+\beta)} \leq \Phi(s) \leq c_2(1+s)^{-(\alpha'+\beta)} \quad (4.31)$$

for all  $s > 0$  and some  $c_1, c_2 > 0$ , where  $\alpha$  and  $\alpha'$  are the exponents from (VD) and (RVD), respectively.

## 5. A maximum principle and its applications

### 5.1. Weak differentiation

Let  $\mathcal{H}$  be a Hilbert spaces over  $\mathbb{R}$  and  $I$  be an interval in  $\mathbb{R}$ . We say that a function  $u : I \rightarrow \mathcal{H}$  is *weakly differentiable* at  $t \in I$  if for any  $\varphi \in \mathcal{H}$ , the function  $(u(\cdot), \varphi)$  is differentiable at  $t$  (where the outer brackets stand for the inner product in  $\mathcal{H}$ ), that is, the limit

$$\lim_{\varepsilon \rightarrow 0} \left( \frac{u(t+\varepsilon) - u(t)}{\varepsilon}, \varphi \right)$$

exists. By the principle of uniform boundedness, in this case there is  $w \in \mathcal{H}$  such that

$$\lim_{\varepsilon \rightarrow 0} \left( \frac{u(t+\varepsilon) - u(t)}{\varepsilon}, \varphi \right) = (w, \varphi)$$

for all  $\varphi \in \mathcal{H}$ . We refer to the vector  $w$  as the *weak derivative* of the function  $u$  at  $t$  and write  $w = u'(t)$ . Of course, we have the weak convergence

$$\frac{u(t+\varepsilon) - u(t)}{\varepsilon} \rightharpoonup u'(t) \quad \text{as } \varepsilon \rightarrow 0.$$

In the next statement, we collect the necessary elementary properties of weak differentiation.

**Lemma 5.1.** (i) *If  $u$  is weakly differentiable at  $t$  then  $u$  is strongly (that is, in the norm of  $\mathcal{H}$ ) continuous at  $t$ .*

(ii) (The product rule) *If functions  $u : I \rightarrow \mathcal{H}$  and  $v : I \rightarrow \mathcal{H}$  are weakly differentiable at  $t$ , then the inner product  $(u, v)$  is also differentiable at  $t$  and*

$$(u, v)' = (u', v) + (u, v').$$

(iii) (The chain rule) Let  $(M, \mu)$  be a measure space and set  $\mathcal{H} = L^2(M, \mu)$ . Let  $u : I \rightarrow L^2(M, \mu)$  be weakly differentiable at  $t \in I$ . Let  $\Phi$  be a smooth real-valued function on  $\mathbb{R}$  such that

$$\Phi(0) = 0, \quad \sup_{\mathbb{R}} |\Phi'| < \infty, \quad \sup_{\mathbb{R}} |\Phi''| < \infty. \quad (5.1)$$

Then the function  $\Phi(u) : I \rightarrow L^2(M, \mu)$  is also weakly differentiable at  $t$  and

$$\Phi(u)' = \Phi'(u) u'.$$

*Proof.* To shorten the notation, we write  $u_t$  for  $u(t)$ .

(i) It suffices to verify that, for any sequence  $\{\varepsilon_k\} \rightarrow 0$ , we have

$$\|u_{t+\varepsilon_k} - u_t\| \rightarrow 0 \text{ as } k \rightarrow \infty. \quad (5.2)$$

The sequence  $\frac{u_{t+\varepsilon_k} - u_t}{\varepsilon_k}$  converges weakly, whence it follows that it is weakly bounded and, hence, also strongly bounded. The latter clearly implies (5.2).

(ii) Let  $\{\varepsilon_k\}$  be as above. We have the identity

$$\begin{aligned} \frac{(u_{t+\varepsilon_k}, v_{t+\varepsilon_k}) - (u_t, v_t)}{\varepsilon_k} &= \left( \frac{u_{t+\varepsilon_k} - u_t}{\varepsilon_k}, v_t \right) + \left( u_t, \frac{v_{t+\varepsilon_k} - v_t}{\varepsilon_k} \right) \\ &\quad + \left( \frac{u_{t+\varepsilon_k} - u_t}{\varepsilon_k}, v_{t+\varepsilon_k} - v_t \right). \end{aligned}$$

By the definition of the weak derivative, the first two terms in the right hand side converge to  $(u_t', v_t)$  and  $(u_t, v_t')$  respectively. By part (i), the sequence  $\frac{u_{t+\varepsilon_k} - u_t}{\varepsilon_k}$  is bounded in norm, whereas  $\|v_{t+\varepsilon_k} - v_t\| \rightarrow 0$  as  $k \rightarrow \infty$ ; hence, the third term goes to 0, and we obtain the desired.

(iii) By (5.1) the function  $\Phi$  admits the estimate  $|\Phi(r)| \leq C|r|$  for all  $r \in \mathbb{R}$ , which implies that the function  $\Phi(u_t)$  belongs to  $L^2(M, \mu)$  for any  $t \in I$ . By the mean value theorem, for any  $r, s \in \mathbb{R}$ , there exists  $\xi_{r,s} \in (0, 1)$  such that

$$\Phi(r+s) - \Phi(r) = \Phi'(r + \xi_{r,s}(r-s))s.$$

We have then

$$\frac{\Phi(u_{t+\varepsilon_k}) - \Phi(u_t)}{\varepsilon_k} = \Phi'(u_t + \xi(u_{t+\varepsilon_k} - u_t)) \frac{u_{t+\varepsilon_k} - u_t}{\varepsilon_k}$$

where we write for simplicity  $\xi = \xi_{u_t, u_{t+\varepsilon_k} - u_t}$ , which can also be rewritten as

$$\begin{aligned} \frac{\Phi(u_{t+\varepsilon_k}) - \Phi(u_t)}{\varepsilon_k} &= \left( \Phi'(u_t + \xi(u_{t+\varepsilon_k} - u_t)) - \Phi'(u_t) \right) \frac{u_{t+\varepsilon_k} - u_t}{\varepsilon_k} \\ &\quad + \Phi'(u_t) \frac{u_{t+\varepsilon_k} - u_t}{\varepsilon_k}. \end{aligned} \quad (5.3)$$

Since

$$\|(\Phi'(u_t + \xi(u_{t+\varepsilon_k} - u_t)) - \Phi'(u_t))\| \leq \sup |\Phi''| \|u_{t+\varepsilon_k} - u_t\|$$

(where  $\|\cdot\|$  is the  $L^2$ -norm) and the term  $\frac{u_{t+\varepsilon_k} - u_t}{\varepsilon_k}$  is bounded in norm, the first term in the right hand side of (5.3) tends to 0 strongly as  $k \rightarrow \infty$ . We are left to verify that the second term goes weakly to  $\Phi'(u_t) u_t'$ . Indeed, for any  $\varphi \in L^2(M, \mu)$ , we have that

$$\begin{aligned} \left( \Phi'(u_t) \frac{u_{t+\varepsilon_k} - u_t}{\varepsilon_k}, \varphi \right) &= \left( \frac{u_{t+\varepsilon_k} - u_t}{\varepsilon_k}, \Phi'(u_t) \varphi \right) \rightarrow (u_t', \Phi'(u_t) \varphi) \\ &= (\Phi'(u_t) u_t', \varphi), \end{aligned}$$

where we have used the fact that  $\Phi'(u_t)$  is bounded and, hence,  $\Phi'(u_t) \varphi \in L^2(M, \mu)$ .  $\square$

## 5.2. Maximum principle for weak solutions

As before, let  $(\mathcal{E}, \mathcal{F})$  be a regular Dirichlet form in  $L^2(M, \mu)$ . Consider a path  $u : I \rightarrow \mathcal{F}$ . We say that  $u$  is a *weak subsolution* of the heat equation in an open set  $\Omega \subset M$  if  $u$  is weakly differentiable in the space  $L^2(\Omega)$  at any  $t \in I$  and, for any non-negative  $\varphi \in \mathcal{F}(\Omega)$ ,

$$(u', \varphi) + \mathcal{E}(u, \varphi) \leq 0. \quad (5.4)$$

Similarly one defines the notions of weak supersolution and weak solution.

A similar definition was introduced in [20] but with the difference that the time derivative  $u'$  was understood in the sense of the norm convergence in  $L^2(\Omega)$ . Let us refer to the solutions defined in [20] as *semi-weak* solutions. Clearly, any semi-weak solution is also a weak solution.

It is easy to see that, for any  $f \in L^2(M)$ , the function  $P_t f$  is a weak solution in  $(0, \infty) \times \Omega$  for any open  $\Omega \subset M$  (cf. [20, Example 4.10]).

**Proposition 5.2 (parabolic maximum principle).** *Let  $u$  be a weak subsolution of the heat equation in  $(0, T) \times \Omega$ , where  $T \in (0, +\infty]$  and  $\Omega$  is an open subset of  $M$ . Assume in addition that  $u$  satisfies the following boundary and initial conditions:*

- $u_+(t, \cdot) \in \mathcal{F}(\Omega)$  for any  $t \in (0, T)$ ;
- $u_+(t, \cdot) \xrightarrow{L^2(\Omega)} 0$  as  $t \rightarrow 0$ .

*Then  $u(t, x) \leq 0$  for any  $t \in (0, T)$  and  $\mu$ -almost all  $x \in \Omega$ .*

**Remark 5.3.** For semi-weak solutions the maximum principle was proved in [20].

**Remark 5.4.** It was shown in [20, Lemma 4.4] that the condition  $u_+ \in \mathcal{F}(\Omega)$  is equivalent to the following:  $u \in \mathcal{F}$  and  $u \leq v$  for some  $v \in \mathcal{F}(\Omega)$ . We will use this result to verify the boundary condition of the parabolic maximum principle.

*Proof.* Let  $\Phi$  be a smooth function on  $\mathbb{R}$  that satisfies the following conditions for some constant  $C$ :

- (i)  $\Phi(r) = 0$  for all  $r \leq 0$ ;
- (ii)  $0 < \Phi'(r) \leq C$  for all  $r > 0$ .
- (iii)  $|\Phi''(r)| \leq C$  for all  $r > 0$ .

Then  $\Phi(u) = \Phi(u_+) \in \mathcal{F}(\Omega)$  so that we can set  $\varphi = \Phi(u)$  in (5.4) and obtain

$$(u', \Phi(u)) + \mathcal{E}(u, \Phi(u)) \leq 0.$$

Since  $\Phi$  is increasing and Lipschitz, we conclude by [20, Lemma 4.3] that  $\mathcal{E}(u, \Phi(u)) \geq 0$ , whence it follows that

$$(u', \Phi(u)) \leq 0. \tag{5.5}$$

Since  $\Phi$  satisfies the conditions (5.1), we conclude by Lemma 5.1, that the function  $t \mapsto \Phi(u)$  is weakly differentiable in the space  $L^2(\Omega)$  and

$$\Phi(u)' = \Phi'(u) u',$$

and  $(u, \Phi(u))$  is differentiable in  $t$  and

$$\begin{aligned} (u, \Phi(u))' &= (u', \Phi(u)) + (u, \Phi(u)') \\ &= (u', \Phi(u)) + (u, \Phi'(u) u') \\ &= (u', \Phi(u)) + (u', \Phi'(u) u). \end{aligned}$$

Set  $\Psi(r) = \Phi'(r) r$  so that

$$(u, \Phi(u))' = (u', \Phi(u)) + (u', \Psi(u)).$$

Assume for a moment that function  $\Psi$  also satisfies the above properties (i)-(iii). Applying (5.5) to  $\Psi$ , we obtain from the previous line

$$(u, \Phi(u))' \leq 0,$$

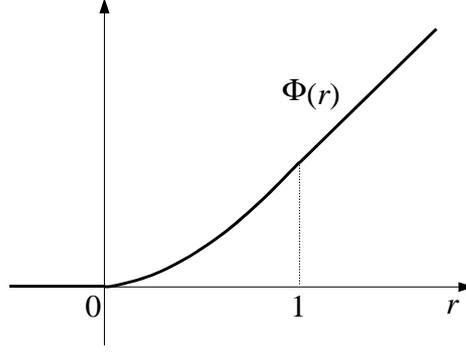
that is, the function  $t \mapsto (u, \Phi(u))$  is decreasing in  $t$ . By the properties (i)-(ii), we have  $\Phi(r) \leq Cr$  for  $r > 0$ , which implies

$$(u, \Phi(u)) = (u_+, \Phi(u_+)) \leq C \|u_+\|^2 \rightarrow 0 \text{ as } t \rightarrow 0.$$

Hence, the function  $t \mapsto (u_+, \Phi(u_+))$  is non-negative, decreasing and goes to 0 as  $t \rightarrow 0$ , which implies that this function is identical 0. It follows that  $u_+ = 0$ , which was to be proved.

We are left to specify the choice of  $\Phi$  so that the function  $\Psi(r) = \Phi'(r) r$  is also in the class (i)-(iii). Let us construct  $\Phi$  from its derivative  $\Phi'$  that can be chosen to satisfy the following:

- $\Phi'(r) = 0$  for  $r \leq 0$ ;
- $\Phi'(r) = 1$  for  $r \geq 1$ ;
- $\Phi''(r) > 0$  for  $r \in (0, 1)$ .

FIGURE 5. Function  $\Phi(r)$ 

(see Fig. 5).

Clearly,  $\Phi$  satisfies (i)-(iii). It follows from the identity

$$\Psi'(r) = \Phi''(r)r + \Phi'(r)$$

that  $\Psi'(r)$  is bounded and  $\Psi'(r) > 0$  for  $r > 0$ . Finally, we have

$$\Psi''(r) = \Phi'''(r)r + 2\Phi''(r)$$

whence it follows that  $\Psi''(r) = 0$  for large enough  $r$  and, hence,  $\Psi''$  is bounded. We conclude that  $\Psi$  satisfies (i)-(iii), which finishes the proof.  $\square$

### 5.3. Some applications of the maximum principle

Recall that if  $(\mathcal{E}, \mathcal{F})$  is a regular Dirichlet form in  $L^2(M, \mu)$  then, for any open set  $\Omega$ ,  $(\mathcal{E}, \mathcal{F}(\Omega))$  is also a regular Dirichlet form in  $L^2(\Omega, \mu)$ . Denote by  $P_t^\Omega$  the heat semigroup of  $(\mathcal{E}, \mathcal{F}(\Omega))$ .

**Lemma 5.5.** *Assume that  $(\mathcal{E}, \mathcal{F})$  is a regular and local Dirichlet form. Let  $u(t, x)$  be a weak subsolution of the heat equation in  $(0, \infty) \times U$ , where  $U$  is an open subset of  $M$ . Assume further, for any  $t > 0$ ,  $u(t, \cdot)$  is bounded in  $M$  and is non-negative in  $U$ . If*

$$u(t, \cdot) \xrightarrow{L^2(U)} 0 \text{ as } t \rightarrow 0 \quad (5.6)$$

then the following inequality hold for all  $t > 0$  and almost all  $x \in U$ :

$$u(t, x) \leq (1 - P_t^U \mathbf{1}_U(x)) \sup_{0 < s \leq t} \|u(s, \cdot)\|_{L^\infty(U)}. \quad (5.7)$$

*Proof.* We first assume that  $U$  is precompact. Choose an open set  $W$  such that  $W \Subset U$ . Fix a real  $T > 0$  and set

$$m := \sup_{0 < s \leq T} \|u(s, \cdot)\|_{L^\infty(U)}. \quad (5.8)$$

We show that, for all  $0 < t \leq T$  and  $\mu$ -almost all  $x \in W$ ,

$$u(t, x) \leq m (1 - P_t^W \mathbf{1}_W(x)). \quad (5.9)$$

Let  $\zeta$  and  $\eta$  be cut-off functions<sup>3</sup> of the couples  $(W, U)$  and  $(U, M)$ , respectively. Consider the function

$$w := \zeta u - m [\eta - P_t^W \mathbf{1}_W]. \quad (5.10)$$

Then (5.9) will follow if we prove that  $w \leq 0$  in  $(0, T] \times W$ .

*Claim 1.* The  $w$  is a weak subsolution of the heat equation in  $(0, \infty) \times W$ .

<sup>3</sup>A cut-off function of the couple  $(W, U)$  is a function  $\zeta \in \mathcal{F} \cap C_0(M)$  such that  $0 \leq \zeta \leq 1$  in  $M$ ,  $\zeta = 1$  on an open neighborhood of  $\overline{W}$ , and  $\text{supp } \zeta \subset U$ . If  $(\mathcal{E}, \mathcal{F})$  is a regular Dirichler form then a cut-off function exists for any couple  $(W, U)$  provided  $U$  is open and  $\overline{W}$  is a compact subset of  $U$  (cf. [15, p.27])

Clearly,  $P_t^W \mathbf{1}_W$  is a weak solution of the heat equation in  $(0, \infty) \times W$ . Let us show that so is  $\zeta u$ . Indeed, the product  $\zeta u$  belongs to  $\mathcal{F}$  because both  $\zeta$  and  $u$  are in  $L^\infty \cap \mathcal{F}$ . For any test function  $\psi \in \mathcal{F}(W)$ , we have, using  $\zeta \psi \equiv \psi$ ,

$$\begin{aligned} \left( \frac{\partial(\zeta u)}{\partial t}, \psi \right) &= \left( \zeta \frac{\partial u}{\partial t}, \psi \right) = \left( \frac{\partial}{\partial t} u, \psi \right) \\ &= -\mathcal{E}(u, \psi) = -\mathcal{E}(\zeta u, \psi) + \mathcal{E}((\zeta - 1)u, \psi) \\ &= -\mathcal{E}(\zeta u, \psi), \end{aligned}$$

where we have used also that  $(\zeta - 1)u = 0$  in  $W$  and, hence,

$$\mathcal{E}((\zeta - 1)u, \psi) = 0,$$

by the locality of  $(\mathcal{E}, \mathcal{F})$ . Thus,  $\zeta u$  is a weak solution in  $(0, \infty) \times W$ .

Finally, the function  $\eta(x)$  considered as a function of  $(t, x)$ , is a weak supersolution of the heat equation in  $(0, \infty) \times W$ , since for any non-negative  $\psi \in \mathcal{F}(W)$

$$\mathcal{E}(\eta, \psi) = \lim_{t \rightarrow 0} t^{-1} (\eta - P_t \eta, \psi) = \lim_{t \rightarrow 0} t^{-1} (1 - P_t \eta, \psi) \geq 0,$$

whence it follows that  $w$  is a weak subsolution.

*Claim 2.* For every  $t \in (0, T]$ , we have  $(w(t, \cdot))_+ \in \mathcal{F}(W)$ .

By Remark 5.4, it suffices to prove that in  $(0, T] \times M$

$$w(t, \cdot) \leq m P_t^W \mathbf{1}_W, \quad (5.11)$$

because  $m P_t^W \mathbf{1}_W \in \mathcal{F}(W)$ . In  $M \setminus U$ , inequality (5.11) holds trivially because

$$\zeta = 0 = P_t^W \mathbf{1}_W \quad \text{in } M \setminus U$$

and, hence,  $w = -m\eta \leq 0$ . To prove (5.11) in  $U$ , observe that  $\eta = 1$  in  $U$  and  $0 \leq u \leq m$  in  $(0, T] \times U$ , whence

$$w = \zeta u - m + m P_t^W \mathbf{1}_W \leq u - m + m P_t^W \mathbf{1}_W \leq m P_t^W \mathbf{1}_W,$$

which was to be proved.

*Claim 3.* The function  $w$  satisfies the initial condition

$$w(t, \cdot) \xrightarrow{L^2(W)} 0 \quad \text{as } t \rightarrow 0. \quad (5.12)$$

Noticing that  $\eta = 1$  in  $W$ , we see that

$$\eta - P_t^W \mathbf{1}_W = \mathbf{1}_W - P_t^W \mathbf{1}_W \xrightarrow{L^2(W)} 0 \quad \text{as } t \rightarrow 0.$$

Combining with (5.6), we obtain (5.12).

By the parabolic maximum principle (cf. Prop. 5.2), we obtain from Claims 1-3 that  $w \leq 0$  in  $(0, T] \times W$ , thus proving (5.9).

Finally, let  $U$  be an arbitrary open subset of  $M$ . Let  $\{W_i\}_{i=1}^\infty$  and  $\{U_i\}_{i=1}^\infty$  be two increasing sequences of precompact open sets, both of which exhaust  $U$ , and such that  $W_i \Subset U_i$  for all  $i$ . For each  $i$ , we have by (5.9) with  $t = T$  that in  $W_i$

$$u \leq \left[ 1 - P_t^{W_i} \mathbf{1}_{W_i} \right] \sup_{0 < s \leq t} \|u(s, \cdot)\|_{L^\infty(U_i)}. \quad (5.13)$$

Replacing by the monotonicity in the right hand side  $U_i$  by  $U$ , and noticing that

$$P_t^{W_i} \mathbf{1}_{W_i} \xrightarrow{\text{a.e.}} P_t^U \mathbf{1}_U \quad \text{as } i \rightarrow \infty,$$

we obtain (5.7) by letting  $i \rightarrow \infty$  in (5.13).  $\square$

**Corollary 5.6.** *Assume that  $(\mathcal{E}, \mathcal{F})$  is regular and strongly local. Let  $U \subset \Omega$  be two open subsets of  $M$ . Then the following inequality holds for all  $t > 0$  and  $\mu$ -almost all  $x \in U$ :*

$$1 - P_t^\Omega \mathbf{1}_\Omega(x) \leq (1 - P_t^U \mathbf{1}_U(x)) \sup_{0 < s \leq t} \|1 - P_s^\Omega \mathbf{1}_\Omega\|_{L^\infty(U)}. \quad (5.14)$$

*Proof.* Approximating  $U$  by precompact open subsets, it suffices to prove the claim in the case when  $U \Subset \Omega$ . Let  $\varphi$  be a cut-off function of the couple  $(U, \Omega)$ . Then we can replace the term  $1 - P_t^\Omega \mathbf{1}_\Omega(x)$  in the both sides of (5.14) by the the function

$$u(t, x) = \varphi(x) - P_t^\Omega \mathbf{1}_\Omega(x).$$

Clearly, for any  $t > 0$ , the function  $u(t, \cdot)$  is bounded in  $M$ , non-negative in  $U$ , and satisfies the initial condition (5.6). Let us verify that  $u(t, x)$  is a weak solution of the heat equation in  $(0, \infty) \times U$ . It suffices to show that the function  $\varphi(x)$  as a function of  $(t, x)$  is a weak solution in  $(0, \infty) \times U$ . Indeed, since  $\varphi$  is constant in a neighborhood of  $\bar{U}$ , the strong locality of  $(\mathcal{E}, \mathcal{F})$  yields that  $\mathcal{E}(\varphi, \psi) = 0$  for any  $\psi \in \mathcal{F}(U)$ , which finishes the proof.  $\square$

## 6. Upper bounds in the local case

### 6.1. Exponential tail

Following [21], we give an analytical approach of how to obtain the exponential tail of the heat kernel upper bound on the doubling space. This is a modification of the argument of [27]. For an alternative approach see [20] (and [25] for the case of infinite graphs).

The following is a key technical lemma.

**Lemma 6.1.** *Let  $(\mathcal{E}, \mathcal{F})$  be a regular, strongly local Dirichlet form in  $L^2(M, \mu)$ . Let  $\rho : [0, \infty) \rightarrow [0, \infty)$  be an increasing function. Assume that there exist  $\varepsilon \in (0, \frac{1}{2})$  and  $\delta > 0$  such that, for any ball  $B \subset M$  of radius  $r$  and any positive  $t$  such that  $\rho(t) \leq \delta r$ ,*

$$P_t \mathbf{1}_{B^c} \leq \varepsilon \text{ in } \frac{1}{4}B. \quad (6.1)$$

Then, for any  $t > 0$  and any ball  $B$  of radius  $r > 0$ ,

$$P_t \mathbf{1}_{B^c} \leq C \exp\left(-c't \Psi\left(\frac{ct}{t}\right)\right) \text{ in } \frac{1}{4}B, \quad (6.2)$$

where  $C, c, c' > 0$  are constants depending on  $\varepsilon, \delta$ , and function  $\Psi$  is defined by

$$\Psi(s) := \sup_{\lambda > 0} \left\{ \frac{s}{\rho(1/\lambda)} - \lambda \right\} \quad (6.3)$$

for all  $s \geq 0$ .

**Remark 6.2.** Letting  $\lambda \rightarrow 0$  in (6.3), one sees that  $\Psi(s) \geq 0$  for all  $s \geq 0$ . It is also obvious from (6.3) that  $\Psi(s)$  is increasing in  $s$ .

**Remark 6.3.** If  $\rho(t) = t^{1/\beta}$  for  $\beta > 1$ , then

$$\Psi(s) = \sup_{\lambda > 0} \left\{ s\lambda^{1/\beta} - \lambda \right\} = c_\beta s^{\beta/(\beta-1)}$$

for all  $s \geq 0$ , where  $c_\beta > 0$  depends only on  $\beta$  (the supremum is attained for  $\lambda = (s/\beta)^{\frac{\beta}{\beta-1}}$ ). The estimate (6.2) becomes

$$P_t \mathbf{1}_{B^c} \leq C \exp\left(-c\left(\frac{r^\beta}{t}\right)^{\frac{1}{\beta-1}}\right) \text{ in } \frac{1}{4}B.$$

**Remark 6.4.** If the heat semigroup  $P_t$  possesses the heat kernel  $p_t(x, y)$  then the condition (6.1) can be equivalently stated as follows: if  $\rho(t) \leq \delta r$  then, for almost all  $x \in M$ ,

$$\int_{B(x, r)^c} p_t(x, y) d\mu(y) \leq \varepsilon. \quad (6.4)$$

Indeed, for any ball  $B(x_0, r)$  and for almost all  $x \in B(x_0, r/4)$ , we have

$$P_t \mathbf{1}_{B(x_0, r)^c}(x) = \int_{B(x_0, r)^c} p_t(x, y) d\mu(y) \leq \int_{B(x, r/2)^c} p_t(x, y) d\mu(y),$$

so that (6.4) implies (6.1) (although with a different value of  $\delta$ ). Similarly, for almost all  $x \in B(x_0, r/2)$ ,

$$\int_{B(x, r)^c} p_t(x, y) d\mu(y) \leq \int_{B(x_0, r/2)^c} p_t(x, y) d\mu(y) = P_t \mathbf{1}_{B(x_0, r/2)^c}(x),$$

so that (6.6) implies (6.4), for almost all  $x \in B(x_0, r/8)$ . Covering  $M$  by a countable family of balls of radius  $r/8$ , we obtain that (6.4) holds for almost all  $x \in M$ .

In the same way, the condition (6.2) is equivalent to the following: for all  $\lambda, t, r > 0$  and for almost all  $x \in M$ ,

$$\int_{B(x,r)^c} p_t(x,y) d\mu(y) \leq C \exp\left(-c't \Psi\left(\frac{cr}{t}\right)\right). \quad (6.5)$$

Hence, Lemma 6.1 in the presence of the heat kernel can be stated as follows: if (6.4) holds for some  $\varepsilon \in (0, 1/2)$ ,  $\delta > 0$  and all  $r, t > 0$  such that  $\rho(t) \leq \delta r$  then (6.5) holds for all  $r, t > 0$ .

*Proof of Lemma 6.1.* Let us first show that the hypothesis (6.1) implies that there exist  $\varepsilon \in (0, 1)$  and  $\delta > 0$  such that, for any ball  $B$  of radius  $r > 0$  and for any positive  $t$  such that  $\rho(t) \leq \delta r$ ,

$$P_t^B \mathbf{1}_B \geq 1 - \varepsilon \quad \text{in } \frac{1}{4}B. \quad (6.6)$$

Indeed, applying [21, Proposition 4.7] we obtain that, for all  $t$  and almost everywhere in  $M$ ,

$$P_t^B \mathbf{1}_{\frac{1}{2}B} \geq P_t \mathbf{1}_{\frac{1}{2}B} - \sup_{0 < s \leq t} \left\| P_s \mathbf{1}_{\frac{1}{2}B} \right\|_{L^\infty\left(\left(\frac{3}{4}B\right)^c\right)}. \quad (6.7)$$

For any  $x \in \frac{1}{4}B$ , we have that  $B(x, r/4) \subset \frac{1}{2}B$  (see Fig. 6).

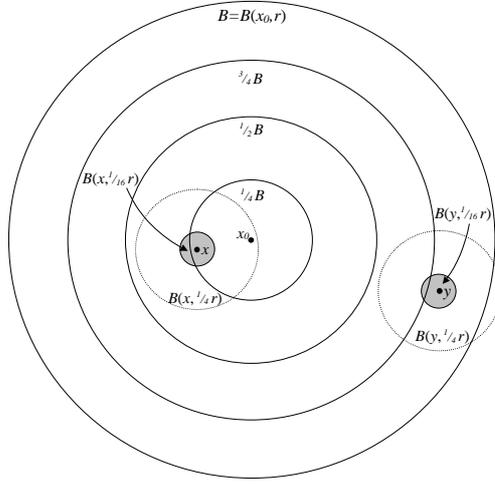


FIGURE 6. Illustration to the proof of (6.1) $\Rightarrow$ (6.6)

Using the identity  $P_t \mathbf{1} = 1$  we obtain, for any  $x \in \frac{1}{4}B$ ,

$$P_t \mathbf{1}_{\frac{1}{2}B} = 1 - P_t \mathbf{1}_{\left(\frac{1}{2}B\right)^c} \geq 1 - P_t \mathbf{1}_{B(x, r/4)^c}.$$

Applying (6.1) for the ball  $B(x, r/4)$ , we obtain that

$$P_t \mathbf{1}_{B(x, r/4)^c} \leq \varepsilon \quad \text{in } B(x, r/16),$$

provided  $t$  is so small that

$$\rho(t) \leq \delta \frac{r}{4}. \quad (6.8)$$

It follows that, for any  $x \in \frac{1}{4}B$ ,

$$P_t \mathbf{1}_{\frac{1}{2}B} \geq 1 - \varepsilon \quad \text{in } B(x, r/16),$$

whence

$$P_t \mathbf{1}_{\frac{1}{2}B} \geq 1 - \varepsilon \quad \text{in } \frac{1}{4}B. \quad (6.9)$$

On the other hand, for any  $y \in \left(\frac{3}{4}B\right)^c$ , we have  $\frac{1}{2}B \subset B(y, r/4)^c$  (see Fig. 6), whence

$$P_s \mathbf{1}_{\frac{1}{2}B} \leq P_s \mathbf{1}_{B(y, r/4)^c}.$$

Applying (6.1) for the ball  $B(y, r/4)$  at time  $s$ , we obtain if (6.8) holds then, for all  $0 < s \leq t$ ,

$$P_s \mathbf{1}_{B(y, r/4)^c} \leq \varepsilon \quad \text{in } B(y, r/16).$$

It follows that, for any  $y \in (\frac{3}{4}B)^c$ ,

$$P_s \mathbf{1}_{\frac{1}{2}B} \leq \varepsilon \text{ in } B(y, r/16),$$

whence

$$P_s \mathbf{1}_{\frac{1}{2}B} \leq \varepsilon \text{ in } \left(\frac{3}{4}B\right)^c. \quad (6.10)$$

Combining (6.7), (6.9) and (6.10), we obtain that, under condition (6.8),

$$P_t^B \mathbf{1}_B \geq P_t^B \mathbf{1}_{\frac{1}{2}B} \geq 1 - 2\varepsilon \text{ in } \frac{1}{4}B, \quad (6.11)$$

which is equivalent to (6.6).

Now we show that (6.6) implies (6.2). The proof will be split into 5 steps.

*Step 1.* Assuming that

$$\rho(t) \leq \delta r \quad (6.12)$$

and that  $B$  is a ball of radius  $r$ , rewrite (6.6) in the form

$$1 - P_t^B \mathbf{1}_B \leq \varepsilon \text{ in } \frac{1}{4}B. \quad (6.13)$$

For any positive integer  $k$ , set  $B_k = kB$  and we will prove that

$$1 - P_t^{B_k} \mathbf{1}_{B_k} \leq \varepsilon^k \text{ in } \frac{1}{4}B_k. \quad (6.14)$$

Since  $M$  is separable, there is a dense countable set of points in  $B_k$ . Let  $\{b_j\}$  be a sequence of balls of radii  $r$  centered at those points. Clearly,  $b_j \subset B_{k+1}$  and the family  $\{\frac{1}{4}b_j\}$  covers  $B_k$  (see Fig. 7).

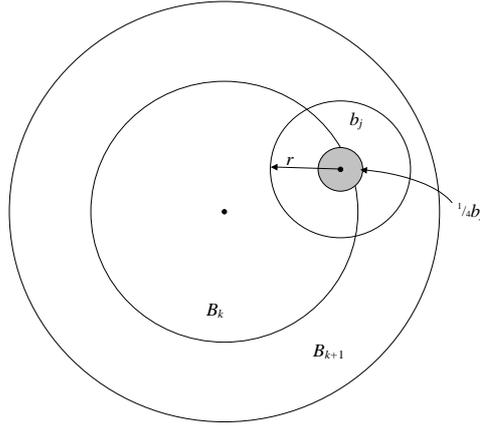


FIGURE 7. Balls  $\{B_k\}$  and  $\{b_j\}$

Due to (6.12), inequality (6.6) is valid for any ball  $b_j$ , that is, for all  $0 < s \leq t$ ,

$$P_s^{B_{k+1}} \mathbf{1}_{B_{k+1}} \geq P_s^{b_j} \mathbf{1}_{b_j} \geq 1 - \varepsilon \text{ in } \frac{1}{4}b_j.$$

It follows that

$$P_s^{B_{k+1}} \mathbf{1}_{B_{k+1}} \geq 1 - \varepsilon \text{ in } B_k.$$

Applying the inequality (5.14) of Corollary 5.6 with  $\Omega = B_{k+1}$  and  $U = B_k$ , we obtain that the following inequality holds in  $B_k$ :

$$\begin{aligned} 1 - P_t^{B_{k+1}} \mathbf{1}_{B_{k+1}} &\leq \left(1 - P_t^{B_k} \mathbf{1}_{B_k}\right) \sup_{0 < s \leq t} \|1 - P_s^{B_{k+1}} \mathbf{1}_{B_{k+1}}\|_{L^\infty(B_k)} \\ &\leq \varepsilon \left(1 - P_t^{B_k} \mathbf{1}_{B_k}\right). \end{aligned}$$

Iterating in  $k$  and using (6.13), we obtain (6.14).

It follows from (6.14) that

$$P_t \mathbf{1}_{B_k^c} \leq 1 - P_t \mathbf{1}_{B_k} \leq 1 - P_t^{B_k} \mathbf{1}_{B_k} \leq \varepsilon^k \text{ in } \frac{1}{4}B_k. \quad (6.15)$$

Although (6.15) has been proved for any integer  $k \geq 1$ , it is trivially true also for  $k = 0$ , if we define  $B_0 := \emptyset$ .

*Step 2.* Fix  $t > 0$ ,  $x \in M$  and consider the function

$$E_{t,x} = \exp\left(c \frac{d(x, \cdot)}{\rho(t)}\right), \quad (6.16)$$

where the constant  $c > 0$  is to be determined later on. Set

$$r = \delta^{-1} \rho(t),$$

and we will prove that

$$P_t(E_{t,x}) \leq C \text{ in } B(x, r/4), \quad (6.17)$$

where  $C$  is a constant depending on  $\varepsilon, \delta$ . Set as before  $B_k = B(x, kr)$ ,  $k \geq 1$ , and  $B_0 = \emptyset$ . Using (6.16) and (6.15), we obtain that in  $B(x, r/4)$ ,

$$\begin{aligned} P_t(E_{t,x}) &= \sum_{k=0}^{\infty} P_t(\mathbf{1}_{B_{k+1} \setminus B_k} E_{t,x}) \\ &\leq \sum_{k=0}^{\infty} \|E_{t,x}\|_{L^\infty(B_{k+1})} P_t(\mathbf{1}_{B_{k+1} \setminus B_k}) \\ &\leq \sum_{k=0}^{\infty} \exp\left(c \frac{(k+1)r}{\rho(t)}\right) P_t(\mathbf{1}_{B_k^c}) \\ &\leq \sum_{k=0}^{\infty} \exp(c(k+1)\delta^{-1}) \varepsilon^k. \end{aligned}$$

Choosing  $c < \delta \log \frac{1}{\varepsilon}$  we obtain that this series converges, which proves (6.17).

*Step 3.* Let us prove that, for all  $t > 0$  and  $x \in M$ ,

$$P_t E_{t,x} \leq C_1 E_{t,x}, \quad (6.18)$$

for some constant  $C_1 = C(\varepsilon, \delta)$ . Observe first that, for all  $y, z \in M$ , we have by the triangle inequality

$$\begin{aligned} E_{t,x}(y) &= \exp\left(c \frac{d(x, y)}{\rho(t)}\right) \\ &\leq \exp\left(c \frac{d(x, z)}{\rho(t)}\right) \exp\left(c \frac{d(z, y)}{\rho(t)}\right) = E_{t,x}(z) E_{t,z}(y), \end{aligned}$$

which can also be written in the form of a function inequality:

$$E_{t,x} \leq E_{t,x}(z) E_{t,z}.$$

It follows that

$$P_t(E_{t,x}) \leq E_{t,x}(z) P_t(E_{t,z}). \quad (6.19)$$

By the previous step, we have

$$P_t(E_{t,z}) \leq C \text{ in } B(z, r), \quad (6.20)$$

where  $r = \frac{1}{4} \delta^{-1} \rho(t)$ . For all  $y \in B(z, r)$ , we have

$$E_{t,z}(y) \leq \exp\left(\frac{cr}{\rho(t)}\right) = \exp(c\delta^{-1}/4) =: C',$$

whence

$$E_{t,x}(z) \leq E_{t,x}(y) E_{t,z}(y) \leq C' E_{t,x}(y).$$

Letting  $y$  vary, we can write

$$E_{t,x}(z) \leq C' E_{t,x} \text{ in } B(z, r).$$

Combining this with (6.19) and (6.20), we obtain

$$P_t(E_{t,x}) \leq CC' E_{t,x} \text{ in } B(z, r).$$

Since  $z$  is arbitrary, covering  $M$  by a countable sequence of balls like  $B(z, r)$ , we obtain that (6.18) holds on  $M$  with  $C_1 = CC'$ .

*Step 4.* Let us prove that, for all  $t > 0$ ,  $x \in M$ , and for any positive integer  $k$ ,

$$P_{kt}(E_{t,x}) \leq C_1^k \text{ in } \frac{1}{4}B, \quad (6.21)$$

where  $B = (x, \delta^{-1}\rho(t))$ . Indeed, by (6.18)

$$P_{kt}(E_{t,x}) = P_{(k-1)t}P_t(E_{t,x}) \leq C_1P_{(k-1)t}E_{t,x}$$

which implies by iteration that

$$P_{kt}(E_{t,x}) \leq C_1^{k-1}P_tE_{t,x}.$$

Combining with (6.17) and noticing that  $C \leq C_1$ , we obtain (6.21).

*Step 5.* Fix a ball  $B = B(x_0, r)$  and observe that (6.2) is equivalent to the following: for all  $t, \lambda > 0$ ,

$$P_t \mathbf{1}_{B^c} \leq C \exp\left(c'\lambda t - \frac{cr}{\rho(1/\lambda)}\right) \text{ in } \frac{1}{2}B, \quad (6.22)$$

where  $C, c, c' > 0$  are constants depending on  $\varepsilon, \delta$ . In what follows, we fix also  $t$  and  $\lambda$ .

Observe first that, for any  $x \in \frac{1}{2}B$ ,

$$P_t \mathbf{1}_{B^c} \leq P_t \mathbf{1}_{B(x, r/2)^c}.$$

Hence, it suffices to prove that, for any  $x \in \frac{1}{2}B$ ,

$$P_t \mathbf{1}_{B(x, r/2)^c} \leq C \exp\left(c'\lambda t - \frac{cr}{\rho(1/\lambda)}\right) \quad (6.23)$$

in a (small) ball around  $x$ . Covering then  $\frac{1}{2}B$  by a countable family of such balls, we then obtain (6.22).

Changing  $t$  to  $t/k$  in (6.21), we obtain that

$$P_t(E_{t/k, x}) \leq C_1^k \text{ in } B(x, \sigma_k)$$

where  $\sigma_k = \frac{1}{4}\delta^{-1}\rho(t/k)$ . Since

$$E_{t/k, x} \geq \exp\left(c\frac{r}{\rho(t/k)}\right) \text{ in } B(x, r)^c$$

and, hence,

$$\mathbf{1}_{B(x, r)^c} \leq \exp\left(-\frac{cr}{\rho(t/k)}\right) E_{t/k, x},$$

we obtain that the following inequality holds in  $B(x, \sigma_k)$

$$P_t \mathbf{1}_{B(x, r)^c} \leq \exp\left(-\frac{cr}{\rho(t/k)}\right) P_t(E_{t/k, x}) \leq \exp\left(c'k - \frac{cr}{\rho(t/k)}\right)$$

where  $c' = \log C_1$ . Given  $\lambda > 0$ , choose an integer  $k \geq 1$  such that

$$\frac{k-1}{t} < \lambda \leq \frac{k}{t}.$$

Then we obtain the following inequality in  $B(x, \sigma_k)$

$$P_t \mathbf{1}_{B(x, r)^c} \leq \exp\left(c'(\lambda t + 1) - \frac{cr}{\rho(1/\lambda)}\right), \quad (6.24)$$

which finishes the proof.  $\square$

## 6.2. Consequences of two-sided estimates (local case)

Now we are able to specify the local case in the statement of Theorem 4.9.

Given two points  $x, y \in M$ , a *chain* connecting  $x$  and  $y$  is any finite sequence  $\{x_k\}_{k=0}^n$  of points in  $M$  such that  $x_0 = x, x_n = y$ . We say that a metric space satisfies the *chain condition* if there is a constant  $C > 0$  such that for any positive integer  $n$  and for all  $x, y \in M$  there is a chain  $\{x_k\}_{k=0}^n$  connecting  $x$  and  $y$ , such that

$$d(x_k, x_{k+1}) \leq C \frac{d(x, y)}{n} \text{ for all } k = 0, 1, \dots, n-1. \quad (6.25)$$

For example, the geodesic distance on any length space satisfies the chain condition. On the other hand, the combinatorial distance on a graph does not satisfy it.

**Theorem 6.5.** *Assume that the metric measure space  $(M, d, \mu)$  satisfies the chain condition and that  $\mu$  satisfies (VD) and (RVD). Let  $(\mathcal{E}, \mathcal{F})$  be a regular, strongly local and conservative Dirichlet form, and let  $\{p_t\}$  be the associated heat kernel such that, for all  $t > 0$  and almost all  $x, y \in M$ ,*

$$\frac{C'_1}{V(x, t^{1/\beta})} \Phi\left(C_1 \frac{d(x, y)}{t^{1/\beta}}\right) \leq p_t(x, y) \leq \frac{C'_2}{V(x, t^{1/\beta})} \Phi\left(C_2 \frac{d(x, y)}{t^{1/\beta}}\right) \quad (6.26)$$

where  $C_1, C'_1, C_2, C'_2$  are positive constants,  $\alpha, \alpha'$  are the exponents from (VD) and (RVD), respectively, and  $\beta > \alpha - \alpha'$ . Then  $\beta \geq 2$  and the following inequality holds:

$$c'_1 \exp\left(-c_1 s^{\beta/(\beta-1)}\right) \leq \Phi(s) \leq c'_2 \exp\left(-c_2 s^{\beta/(\beta-1)}\right) \quad (6.27)$$

for some positive constants  $c_1, c'_1, c_2, c'_2$  and all  $s > 0$ .

*Proof.* We split the proof into five steps.

*Step 1.* By Lemma 4.3 and the lower bound of  $p_t$ , we obtain

$$\Phi(s) \leq c(1+s)^{-(\alpha'+\beta)} \quad \text{for all } s > 0.$$

Therefore, using the upper bound of  $p_t$ , we obtain (similar to (4.15)) that

$$\begin{aligned} \int_{B(x, r)^c} p_t(x, y) d\mu(y) &\leq c \int_{\frac{1}{2}r/t^{1/\beta}}^{\infty} s^{\alpha-1} \Phi(s) ds \\ &\leq c' \int_{\frac{1}{2}r/t^{1/\beta}}^{\infty} s^{\alpha-\alpha'-\beta-1} ds. \end{aligned}$$

Due to the condition  $\beta > \alpha - \alpha'$ , the integral in the right hand side converges and, hence, the right hand side can be made arbitrarily small provided  $rt^{-1/\beta}$  is large enough. We conclude by Lemma 6.1 (cf. Remark 6.4) with  $\rho(t) = t^{1/\beta}$  that, for all  $r, t > 0$  and for almost all  $x \in M$ ,

$$\int_{B(x, r)^c} p_t(x, y) d\mu(y) \leq C \exp\left(-c't \Psi\left(\frac{cr}{t}\right)\right), \quad (6.28)$$

where

$$\Psi(s) := \sup_{\lambda > 0} \left\{ s\lambda^{1/\beta} - \lambda \right\}. \quad (6.29)$$

*Step 2.* Let us prove that  $\beta > 1$ . If  $\beta < 1$  then it follows from (6.29) that  $\Psi \equiv \infty$ . Substituting into (6.28) and letting  $r \rightarrow 0$ , we obtain that, for almost all  $x \in M$ ,

$$\int_{M \setminus \{x\}} p_t(x, y) d\mu(y) = 0.$$

It follows from the stochastic completeness that there is a point  $x \in M$  of a positive measure, which contradicts Remark 4.4.

Assume now that  $\beta = 1$ . Then by (6.29)

$$\Phi(s) = \begin{cases} 0, & 0 \leq s \leq 1 \\ \infty, & s > 1, \end{cases}$$

which implies that, for all  $t < cr$  and for almost all  $x \in M$ ,

$$\int_{B(x, r)^c} p_t(x, y) d\mu(y) = 0,$$

that is,  $p_t(x, y) = 0$  for all  $t < cd(x, y)$  and almost all  $x, y \in M$ . Together with (6.26), this yields the following bounds of the heat kernel for all  $t > 0$  and almost all  $x, y \in M$ :

$$\frac{C^{-1}}{V(x, t^{1/\beta})} \Phi\left(\frac{d(x, y)}{t^{1/\beta}}\right) \leq p_t(x, y) \leq \frac{C}{V(x, t^{1/\beta})} \tilde{\Phi}\left(\frac{d(x, y)}{t^{1/\beta}}\right) \quad (6.30)$$

where

$$\tilde{\Phi}(s) = \begin{cases} \Phi(s), & s \leq c^{-1} \\ 0, & s > c^{-1}. \end{cases}$$

Clearly, the functions  $\Phi$  and  $\tilde{\Phi}$  satisfy the hypotheses of Theorem 4.7. We conclude by Theorem 4.7 that  $\beta = \beta^*$  whereas by Lemma 3.1  $\beta^* \geq 2$ , which contradicts to  $\beta = 1$ .

*Step 3.* Using that  $\beta > 1$ , let us show that the heat kernel satisfies the following upper bound

$$p_t(x, y) \leq \frac{C}{V(x, t^{1/\beta})} \exp\left(-c \left(\frac{d(x, y)}{t^{1/\beta}}\right)^{\beta/(\beta-1)}\right). \quad (6.31)$$

Setting in (6.29)  $\lambda = (s/\beta)^{\frac{\beta}{\beta-1}}$  we obtain as in Remark 6.3

$$\Psi(s) = c_\beta s^{\beta/(\beta-1)}$$

so that (6.28) becomes

$$\int_{B(x, r)^c} p_t(x, y) d\mu(y) \leq C \exp\left(-c \left(\frac{r}{t^{1/\beta}}\right)^{\beta/(\beta-1)}\right). \quad (6.32)$$

On the other hand, by the upper bound in (6.26), we have, for all  $t > 0$  and almost all  $x, y \in M$ ,

$$p_t(x, y) \leq \frac{C}{V(x, t^{1/\beta})}. \quad (6.33)$$

Setting  $r = \frac{1}{2}d(x, y)$ , we obtain that from (6.32) and (6.33)

$$\begin{aligned} p_{2t}(x, y) &= \int_M p_t(x, z)p_t(z, y) d\mu(z) \\ &\leq \int_{B(x, r)^c} p_t(x, z)p_t(z, y) d\mu(z) + \int_{B(y, r)^c} p_t(x, z)p_t(z, y) d\mu(z) \end{aligned} \quad (6.34)$$

$$\begin{aligned} &\leq \frac{C}{V(y, t^{1/\beta})} \int_{B(x, r)^c} p_t(x, z) d\mu(z) + \frac{C}{V(x, t^{1/\beta})} \int_{B(y, r)^c} p_t(y, z) d\mu(z) \\ &\leq \left(\frac{C}{V(y, t^{1/\beta})} + \frac{C}{V(x, t^{1/\beta})}\right) \exp\left(-c \left(\frac{r}{t^{1/\beta}}\right)^{\beta/(\beta-1)}\right). \end{aligned} \quad (6.35)$$

By (VD) we have

$$\frac{V(x, t^{1/\beta})}{V(y, t^{1/\beta})} \leq C \left(\frac{d(x, y) + t^{1/\beta}}{t^{1/\beta}}\right)^\alpha = C \left(1 + \frac{r}{t^{1/\beta}}\right)^\alpha.$$

Absorbing the polynomial function of  $r/t^{1/\beta}$  into the exponential term in (6.35), we obtain (6.31).

*Step 4.* Now we can prove that  $\beta \geq 2$ . Indeed, we have the estimate (6.30) where this time

$$\tilde{\Phi}(s) = \exp\left(-cs^{\beta/(\beta-1)}\right).$$

Since the estimate (6.30) satisfies the hypotheses Theorem 4.7, we obtain  $\beta \geq 2$  by the same argument as in Step 2.

*Step 5.* The lower bound in (6.30) implies that, for all  $t > 0$  and almost all  $x, y \in M$ , such that  $d(x, y) \leq s_0 t^{1/\beta}$ ,

$$p_t(x, y) \geq \frac{c}{V(x, t^{1/\beta})}. \quad (6.36)$$

Let us show that this implies the following lower bound

$$p_t(x, y) \geq \frac{c}{V(x, t^{1/\beta})} \exp\left(-C \left(\frac{d(x, y)}{t^{1/\beta}}\right)^{\beta/(\beta-1)}\right), \quad (6.37)$$

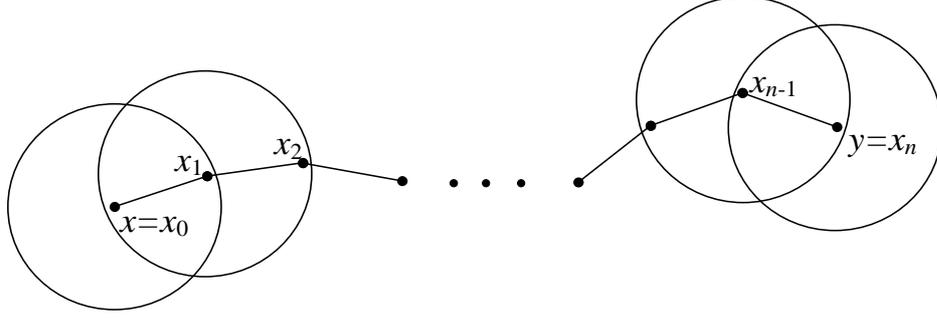
for all  $t > 0$  and almost all  $x, y \in M$ . Iterating the semigroup identity, we obtain for any positive integer  $n$  and real  $r > 0$

$$\begin{aligned} p_t(x, y) &= \int_M \dots \int_M p_{\frac{t}{n}}(x, z_1) p_{\frac{t}{n}}(z_1, z_2) \dots p_{\frac{t}{n}}(z_{n-1}, y) d\mu(z_{n-1}) \dots d\mu(z_1) \\ &\geq \int_{B(x_1, r)} \dots \int_{B(x_{n-1}, r)} p_{\frac{t}{n}}(x, z_1) p_{\frac{t}{n}}(z_1, z_2) \dots p_{\frac{t}{n}}(z_{n-1}, y) d\mu(z_{n-1}) \dots d\mu(z_1), \end{aligned} \quad (6.38)$$

where  $\{x_i\}_{i=0}^n$  is a chain connecting  $x$  and  $y$  that satisfies (6.25) (see Fig. 8).

Denote for simplicity  $z_0 = x$  and  $z_n = y$ . Setting

$$r = \frac{d(x, y)}{n} \quad (6.39)$$

FIGURE 8. Chain  $\{x_i\}$ 

and noticing that  $z_i \in B(x_i, r)$ ,  $0 \leq i \leq n-1$ , we obtain by the triangle inequality and (6.25)

$$d(z_i, z_{i+1}) \leq d(x_i, x_{i+1}) + 2r \leq C' \frac{d(x, y)}{n}$$

where  $C' = C + 2$ . Next, we would like to use (6.36) to estimate from below  $p_{t/n}(z_i, z_{i+1})$ . For that, the following condition must be satisfied:

$$d(z_i, z_{i+1}) \leq s_0 \left( \frac{t}{n} \right)^{1/\beta},$$

which will follow from

$$C' \frac{d(x, y)}{n} \leq s_0 \left( \frac{t}{n} \right)^{1/\beta}.$$

Absorbing the constants  $C'$  and  $s_0$  into one, we see that the latter condition is equivalent to

$$n \geq c \left( \frac{d(x, y)}{t^{1/\beta}} \right)^{\frac{\beta}{\beta-1}}. \quad (6.40)$$

If  $d(x, y) \leq s_0 t^{1/\beta}$  then (6.37) follows immediately from (6.36). Assume in the sequel that  $d(x, y) > s_0 t^{1/\beta}$  and choose  $n$  to be the least positive integer satisfying (6.40), that is

$$n \asymp \left( \frac{d(x, y)}{t^{1/\beta}} \right)^{\frac{\beta}{\beta-1}}. \quad (6.41)$$

This and (6.39) clearly imply that

$$r \asymp \left( \frac{t}{n} \right)^{1/\beta}. \quad (6.42)$$

Then we have by (6.36) and (VD)

$$p_{\frac{t}{n}}(z_i, z_{i+1}) \geq \frac{c}{V(z_i, (t/n)^{1/\beta})} \geq \frac{c}{V(z_i, r)}. \quad (6.43)$$

Since by (VD)

$$\frac{V(z_i, r)}{V(x_i, r)} \leq C \left( \frac{d(z_i, x_i) + r}{r} \right)^\alpha \leq C 2^\alpha,$$

it follows from (6.43) that

$$p_{\frac{t}{n}}(z_i, z_{i+1}) \geq \frac{c}{V(x_i, r)}.$$

Using (6.38), (6.42), (VD), (RVD), and (6.41), we obtain

$$\begin{aligned}
p_t(x, y) &\geq \int_{B(x_1, r)} \dots \int_{B(x_{n-1}, r)} \frac{c^n d\mu(z_{n-1}) \dots d\mu(z_1)}{V(x, r) V(x_1, r) \dots V(x_{n-1}, r)} \\
&= \frac{c^n}{V(x, r)} \geq c' \frac{c^n}{V(x, (t/n)^{1/\beta})} \\
&= \frac{c'}{V(x, t^{1/\beta})} \frac{c^n V(x, t^{1/\beta})}{V(x, (t/n)^{1/\beta})} \geq c' \frac{c^n n^{\alpha'/\beta}}{V(x, t^{1/\beta})} \\
&\geq \frac{c'}{V(x, t^{1/\beta})} \exp(-Cn) \\
&\geq \frac{c'}{V(x, t^{1/\beta})} \exp\left(-C \left(\frac{d(x, y)^\beta}{t}\right)^{\frac{1}{\beta-1}}\right).
\end{aligned}$$

Comparing (6.26) with (6.31) and (6.37), we obtain (6.27).  $\square$

**Corollary 6.6.** *Under the hypotheses of Theorem 6.5, we have  $\mathcal{E}(u) \asymp E_\beta(u)$  for all  $u \in L^2(M)$ . Consequently,  $\mathcal{F} = W^{\beta/2, 2}$ .*

*Proof.* Indeed, by Lemma 4.2 we have  $\mathcal{E}(u) \geq cE_\beta(u)$ . Using the upper bound (6.31) and Lemma (4.5), we obtain  $\mathcal{E}(u) \leq CE_\beta(u)$ , which finishes the proof.  $\square$

## References

- [1] D.G. Aronson, Non-negative solutions of linear parabolic equations Ann. Scuola Norm. Sup. Pisa. Cl. Sci. (4) **22** (1968), 607–694. Addendum **25** (1971), 221–228.
- [2] M.T. Barlow, *Diffusions on fractals*, Lect. Notes Math. **1690**, Springer, 1998, 1-121.
- [3] M.T. Barlow and R.F. Bass, Brownian motion and harmonic analysis on Sierpinski carpets, Canad. J. Math. (4) **51** (1999), 673-744.
- [4] M.T. Barlow, R.F. Bass, Z.-Q. Chen and M. Kassmann, Non-local Dirichlet forms and symmetric jump processes, Trans. Amer. Math. Soc. **361** (2009), 1963-1999.
- [5] M. Barlow, T. Coulhon and T. Kumagai, Characterization of sub-Gaussian heat kernel estimates on strongly recurrent graphs, Comm. Pure Appl. Math., **58** (2005), 1642-1677.
- [6] M.T. Barlow and E.A. Perkins, Brownian motion on the Sierpinski gasket, Probab. Theory. Related Fields **79** (1988), 543-623.
- [7] A. Bendikov and L. Saloff-Coste, On-and-off diagonal heat kernel behaviors on certain infinite dimensional local Dirichlet spaces, Amer. J. Math. **122** (2000), 1205–1263.
- [8] G. Carron, 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), 205–232.
- [9] E.A. Carlen and S. Kusuoka and D.W. Stroock, Upper bounds for symmetric Markov transition functions, Ann. Inst. Henri. Poincaré-Probab. Statist. **23**(1987), 245-287.
- [10] I. Chavel, *Eigenvalues in Riemannian geometry*, Academic Press, New York, 1984.
- [11] T. Coulhon, Ultracontractivity and Nash type inequalities, J. Funct. Anal. **141** (1996), 510-539.
- [12] T. Coulhon, Heat kernel and isoperimetry on non-compact Riemannian manifolds, in: Heat kernels and analysis on manifolds, graphs, and metric spaces, 65–99, Contemp. Math., **338** Amer. Math. Soc., Providence, RI, 2003.
- [13] M. Cowling and S. Meda, Harmonic analysis and ultracontractivity. Trans. Amer. Math. Soc. **340** (1993), 733–752.
- [14] E.B. Davies, *Heat kernels and spectral theory*, Cambridge University Press, 1989.
- [15] M. Fukushima and Y. Oshima Y. and M. Takeda, *Dirichlet forms and symmetric Markov processes*, De Gruyter, Studies in Mathematics, 1994.
- [16] M. Fukushima and T. Shima, On a spectral analysis for the Sierpiński gasket. Potential Anal. **1** (1992), 1–35.
- [17] A. Grigor'yan, Heat kernel upper bounds on a complete non-compact manifold, Rev. Mat. Iberoamericana **10**(1994), 395–452.

- [18] A. Grigor'yan, Estimates of heat kernels on Riemannian manifolds, London Math. Soc. Lecture Note Ser., **273** (1999), 140-225.
- [19] A. Grigor'yan, Heat kernels on metric measure spaces, to appear in "Handbook of Geometric Analysis No.2" (ed. L. Ji, P. Li, R. Schoen, L. Simon), Advanced Lectures in Math., IP.
- [20] A. Grigor'yan and J. Hu, Off-diagonal upper estimates for the heat kernel of the Dirichlet forms on metric spaces, *Invent. Math.* **174** 2008, 81–126.
- [21] A. Grigor'yan and J. Hu, Upper bounds of heat kernels on doubling spaces, preprint 2008.
- [22] A. Grigor'yan, J. Hu and K.-S. Lau, Heat kernels on metric-measure spaces and an application to semilinear elliptic equations, *Trans. Amer. Math. Soc.* **355** (2003), 2065–2095.
- [23] A. Grigor'yan, J. Hu and K.-S. Lau, Equivalence conditions for on-diagonal upper bounds of heat kernels on self-similar spaces, *J. Func. Anal.* **237** (2006), 427-445.
- [24] A. Grigor'yan and T. Kumagai, On the dichotomy in the heat kernel two sided estimates. In: *Analysis on Graphs and its Applications* (P. Exner et al. (eds.)), Proc. of Symposia in Pure Math. **77**, pp. 199–210, Amer. Math. Soc. 2008.
- [25] A. Grigor'yan and A. Telcs, Sub-Gaussian estimates of heat kernels on infinite graphs, *Duke Math. J.* **109**(2001), 451–510.
- [26] B.M. Hambly and T. Kumagai, Transition density estimates for diffusion processes on post critically finite self-similar fractals, *Proc. London Math. Soc.* (3)**79** (1999), 431-458.
- [27] W. Hebisch and L. Saloff-Coste, On the relation between elliptic and parabolic Harnack inequalities, *Ann. Inst. Fourier (Grenoble)* **51** (2001), 1437-1481.
- [28] A. Jonsson, Brownian motion on fractals and function spaces, *Math. Zeit.* **222** (1996), 495-504.
- [29] J. Jorgenson and W. Lynne, ed. *The ubiquitous heat kernel*. Contemporary Math. **398**, AMS, 2006.
- [30] J. Kigami, *Analysis on fractals*, Cambridge University Press, Cambridge, 2001.
- [31] J. Kigami, Local Nash inequality and inhomogenous of heat kernels, *Proc. London Math. Soc.* **89** (2004), 525–544.
- [32] P.Li and S.T. Yau, On the parabolic kernel of the Schrödinger operator, *Acta Math.* **156**(1986), 153-201.
- [33] J. Nash, Continuity of solutions of parabolic and elliptic equations, *Amer. J. Math.* **80** (1958), 931-954.
- [34] K. Pietruska-Pałuba, On function spaces related to fractional diffusion on  $d$ -sets, *Stochastics and Stochastics Reports*, **70** (2000), 153-164.
- [35] F.O. Porper and S.D. Eidel'man, Two-side estimates of fundamental solutions of second-order parabolic equations and some applications, *Russian Math. Surveys*, **39** (1984), 119-178.
- [36] L. Saloff-Coste, *Aspects of Sobolev-type inequalities*, Cambridge University Press, Cambridge, 2002.
- [37] L. Saloff-Coste, A note on Poincaré, Sobolev, and Harnack inequalities, *Internat. Math. Res. Notices* 1992, no. 2, 27–38.
- [38] R. Schoen and S.-T Yau, *Lectures on Differential Geometry*, International Press, 1994.
- [39] D.W. Stroock, Estimates on the heat kernel for the second order divergence form operators, in: *Probability theory. Proceedings of the 1989 Singapore Probability Conference held at the National University of Singapore, June 8-16 1989*, ed. L.H.Y. Chen, K.P. Choi, K. Hu and J.H. Lou, Walter De Gruyter, 1992, 29-44.
- [40] M. Tomisaki, Comparison theorems on Dirichlet norms and their applications, *Forum Math.* **2** (1990), 277–295.
- [41] N.Th. Varopoulos, Hardy-Littlewood theory for semigroups, *J. Funct. Anal.* **63** (1985), 240-260.
- [42] N. Th. Varopoulos and L. Saloff-Coste and T. Coulhon, *Analysis and geometry on groups*, Cambridge Tracts in Mathematics, **100**. Cambridge University Press, Cambridge, 1992.
- [43] F.-Y. Wang, Functional inequalities for empty essential spectrum, *J. Funct. Anal.* **170** (2000), 219-245.

Alexander Grigor'yan

Fakultät für Mathematik, Universität Bielefeld, Postfach 100131, 33501 Bielefeld, Germany.

e-mail: [grigor@math.uni-bielefeld.de](mailto:grigor@math.uni-bielefeld.de)

Jiixin Hu

Department of Mathematical Sciences, Tsinghua University, Beijing 100084, China.

e-mail: [hujiaxin@mail.tsinghua.edu.cn](mailto:hujiaxin@mail.tsinghua.edu.cn)

Ka-Sing Lau

Department of Mathematics, the Chinese University of Hong Kong, Shatin, N.T., Hong Kong.

e-mail: [kslau@math.cuhk.edu.hk](mailto:kslau@math.cuhk.edu.hk)