# Heat kernel upper bounds on fractal spaces

Alexander Grigor'yan
Imperial College
London SW7 2AZ, England
and
Institute of Control Sciences of RAS
Moscow, Russia
email: a.grigoryan@imperial.ac.uk

June 23, 2004

## Contents

# 1 Introduction and statements

Let $(M, d)$ be a locally compact separable metric space and $\mu$ be a Radon measure on $M$ with full support. Let $(\mathcal{E}, \mathcal{F})$ be a local regular Dirichlet form in $L^2(M, \mu)$ and $\{X_t\}_{t \geq 0}$ be an associated diffusion process on $M$. Denote by $\mathbb{P}_x$ and $\mathbb{E}_x$ respectively the probability measure and expectation associated with this process starting at the initial point $x \in M$.

Let us assume that the process $\{X_t\}$ has the *transition density* $p_t(x, y)$ with respect to the measure $\mu$; that is, for all $x \in M$, $t > 0$, and any Borel set $A \subset M$

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

For simplicity, assume further that $p_t(x, y)$ is a continuous function of $x, y \in M$ for all $t > 0$. The function $p_t(x, y)$ is called also the *heat kernel* of the form $(\mathcal{E}, \mathcal{F})$ or of the process $\{X_t\}$.

We have in mind two kind of examples of the above setting. Firstly, let $M$ be a Riemannian manifold. Then let $d$ be the geodesic distance, $\mu$ be the Riemannian volume, and $\mathcal{E}$ be the canonical energy form given by

$$\mathcal{E}[f] = \int_M |\nabla f|^2\, d\mu,$$

and $\mathcal{F} = \overset{o}{H}{}^1(M, \mu)$ (that is, $\mathcal{F}$ is the closure of $C_0^\infty$ in $W^{1,2}(M, \mu)$). In this case, $\{X_t\}$ is the standard Brownian motion on $M$, and the heat kernel $p_t(x, y)$ exists and is a smooth function in $(t, x, y)$. There is also a vast literature devoted to upper and lower bounds of the heat kernel in connection with the geometry of $M$ (see, for example, [6], [8], [13], [25], [27], [28], [29]).

Secondly, let $M$ be one of fractal spaces described, for example, in [1]. Normally, $d$ is an extrinsic distance, $\mu$ is a Hausdorff measure, and the energy form $(\mathcal{E}, \mathcal{F})$ is constructed by using graph approximations of $M$ and a scaling limit. On large classes of fractals, it was proved that the heat kernel exists and is a continuous function of $(t, x, y)$. Furthermore, on such fractals the heat kernel admit nice upper and lower bounds (see, for example, [1], [19], [21]).

Returning to the abstract setting, for any $x \in M$ and $r > 0$, set

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

and let $V(x, r) := \mu(B(x, r))$ be the volume of the ball $B(x, r)$. We will assume throughout that $0 < V(x, r) < \infty$. The aim of this paper is to provide equivalent conditions for the following *upper estimate* of the heat kernel, for a given parameter $\beta > 1$ called the *walk dimension*:

$(UE_\beta)$ : There is a constant $C > 0$ such that, for all $x, y \in M$, and for all $t > 0$,

$$p_t(x, y) \leq \frac{C}{V\left(x, t^{1/\beta}\right)} \exp\left(-\left(\frac{d^\beta(x, y)}{Ct}\right)^{\frac{1}{\beta - 1}}\right), \tag{1.1}$$

The form of the estimate $(UE_\beta)$ is motivated by the following two classes of examples.

1. If $M$ is a geodesically complete Riemannian manifold with non-negative Ricci curvature then the heat kernel satisfies $(UE_\beta)$ with $\beta = 2$ (see [23], [10], [26]). If $M = \mathbb{R}^n$ with the standard Euclidean structure then $(UE_2)$ holds because in $\mathbb{R}^n$ we have

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

and $V(x, r) = c_n r^n$.

2. On a large class of fractal spaces, one has $V(x,r) \simeq r^{\alpha}$, and the estimate $(UE_{\beta})$ (as well as a matching lower bound) holds with some $\beta > 2$ (see, for example, [2], [3], [4]).

In the case of a Riemannian manifold, the necessary and sufficient condition for $(UE_2)$ in terms of a *Faber-Krahn inequality* were proved in [11] (see below for more detail). In a general setting, Kigami [22] proved the necessary and sufficient conditions for $(UE_{\beta})$ in terms of a local form of a Nash inequality and a mean exit time estimate. The present paper is largely motivated by this result of Kigami. Our purpose here is threefold. Firstly, we use a Faber-Krahn inequality instead of a local Nash inequality to match the aforementioned result of [11]. Secondly, we improve the argument of Kigami to get rid of some additional technical assumptions. Thirdly, we prove a new equivalences for $(UE_{\beta})$.

In order to state the results, let us introduce notation and terminology. We say that the process $\{X_t\}_{t\geq 0}$ (or the heat kernel $p_t$) is *stochastically complete* if

$$\mathbb{P}_x(X_t \in M) = \int_M p_t(x,y)\, d\mu(y) \equiv 1 \quad \text{for all } x \in M \text{ and } t > 0. \tag{1.2}$$

For any open set $\Omega \subset M$ define the *exit time*

$$\tau_{\Omega} := \inf\{t > 0 : X_t \notin \Omega\} \tag{1.3}$$

(here $X_t \notin \Omega$ means that either $X_t \in \Omega^c := M \setminus \Omega$ or $X_t$ is in the *cemetery*, in the case if $\{X_t\}$ is stochastically incomplete). We will frequently consider the *mean exit time* from the center of a ball, which is $\mathbb{E}_x \tau_{B(x,r)}$.

For any open set $\Omega \subset M$ set

$$\mathcal{F}(\Omega) := \{f \in \mathcal{F} : f = 0 \text{ in } M \setminus \Omega\} \tag{1.4}$$

and define the *spectral gap* of $\Omega$ by

$$\lambda_{\min}(\Omega) := \inf_{f \in \mathcal{F}(\Omega)\setminus\{0\}} \frac{\mathcal{E}[f]}{\|f\|_2}, \tag{1.5}$$

where $\|f\|_2$ is the norm of $f$ in $L^2(M,\mu)$. In fact, $\lambda_{\min}(\Omega)$ is the bottom of the spectrum of the generator $H_{\Omega}$ of the Dirichlet form $(\mathcal{E}, \mathcal{F}(\Omega))$ in $L^2(\Omega,\mu)$.

Here and throughout we denote by $C$ and $c$ positive constants, whose values may change at each occurrence. Our results are quantitative in the sense that the constants in the conclusions depend only on the constants in the hypotheses. Consider the following hypotheses that in general may be true or not, with a fixed parameter $\beta > 1$.

$(VD)$ : The *volume doubling* property: for all $x \in M$ and $r > 0$, $V(x,r)$ is finite, positive, and

$$V(x,2r) \leq CV(x,r).$$

This condition is equivalent to the following: there exists $\alpha > 0$ such that, for all $x \in M$ and $0 < r \leq R$,

$$\frac{V(x,R)}{V(x,r)} \leq C\left(\frac{R}{r}\right)^{\alpha} \tag{1.6}$$

(see Lemma 11.1 below).

$(E_{\beta})$ : The *mean exit time estimate*: for all $x \in M$ and $r > 0$,

$$cr^{\beta} \leq \mathbb{E}_x \tau_{B(x,r)} \leq Cr^{\beta}.$$

For example, in $\mathbb{R}^n$ one has $\mathbb{E}_x \tau_{B(x,r)} = cr^2$, that is, $(E_\beta)$ holds with $\beta = 2$. The latter is true also for any complete non-compact manifold of non-negative Ricci curvature. On all fractal spaces mentioned above, one has $(E_\beta)$ with $\beta \geq 2$.

$(P_\beta)$ The *exit probability estimate*: for all $x \in M$ and $r > 0$,

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

for some $\varepsilon \in (0,1)$ and $\delta > 0$.

Note that $(E_\beta) \Longrightarrow (P_\beta)$ (see Theorem 9.3). Many equivalent conditions to $(P_\beta)$ are stated in Theorem 9.1. In particular, if the process $X_t$ is stochastically complete then $(P_\beta)$ is equivalent to the following one: there exists $0 < \varepsilon < \frac{1}{2}$ and $C > 0$ such that, for all $x \in M$ and $t > 0$,

$$\int_{B\left(x, Ct^{1/\beta}\right)} p_t(x,y)\, d\mu(y) \geq 1 - \varepsilon.$$

$(FK_\beta)$ : The *Faber-Krahn inequality*: there exists $\nu > 0$ such that, for any ball $B \subset M$ of radius $r$ and for any non-empty open set $\Omega \subset B$,

$$\lambda_{\min}(\Omega) \geq \frac{c}{r^\beta} \left( \frac{\mu(B)}{\mu(\Omega)} \right)^\nu. \tag{1.7}$$

Since $\mu(B) \geq \mu(\Omega)$, the value of $\nu$ can be chosen to be arbitrarily small, for example, $\nu < 1$, which will be frequently assumed.

It is easy to see that $(FK_2)$ holds in $\mathbb{R}^n$. Indeed, for any bounded open set $\Omega$ in $\mathbb{R}^n$, a theorem of Faber and Krahn says that

$$\lambda_{\min}(\Omega) \geq \lambda_{\min}(\Omega^*)$$

where $\Omega^*$ is a "symmetrization" of $\Omega$, that is, a ball of the same volume as $\Omega$. If the radius of $\Omega^*$ is $\rho$ then we have

$$\lambda_{\min}(\Omega^*) = \frac{c}{\rho^2} = \frac{c'}{\mu(\Omega^*)^{2/n}},$$

which yields

$$\lambda_{\min}(\Omega) \geq \frac{c'}{\mu(\Omega)^{2/n}}.$$

We see that (1.7) holds with $\nu = 2/n$ and $\beta = 2$ because the terms $\mu(B)^\nu$ and $r^2$ cancel out. It is possible to prove that in fact $(FK_2)$ holds on any complete non-compact Riemannian manifold with non-negative Ricci curvature (see [10]). In this generality, one cannot get rid of the term $\mu(B)^\nu$ in (1.7).

$(DUE_\beta)$ : A *diagonal upper estimate* of the heat kernel: for all $x \in M$ and all $t > 0$,

$$p_t(x,x) \leq \frac{C}{V\left(x, t^{1/\beta}\right)}. \tag{1.8}$$

Using the semigroup property, the symmetry of the heat kernel, and the Cauchy-Schwarz inequality, it is easy to show that (1.8) is equivalent to the estimate

$$p_t(x,y) \leq \frac{C}{\sqrt{V\left(x, t^{1/\beta}\right) V\left(y, t^{1/\beta}\right)}}, \tag{1.9}$$

for all $x, y \in M$ and $t > 0$. The estimate (1.9) will also be referred to as $(DUE_\beta)$.

$(\Phi UE_\beta)$ : An *upper estimate* with a function $\Phi$: for all $x, y \in M$ and all $t > 0$,

$$p_t(x,y) \le \frac{C}{V\left(x, t^{1/\beta}\right)} \Phi\left(\frac{d(x,y)}{t^{1/\beta}}\right), \tag{1.10}$$

where $\Phi(s)$ is a decreasing positive function on $[0, +\infty)$ such that

$$\int^{\infty} s^{\alpha-1} \Phi(s)\, ds < \infty, \tag{1.11}$$

and $\alpha$ is the exponent from (1.6).

For example, $(UE_\beta)$ can be stated in the form $(\Phi UE_\beta)$ with the function $\Phi(s) = \exp(-cs^{\frac{\beta}{\beta-1}})$, which obviously satisfies (1.11). Therefore, $(UE_\beta) \Rightarrow (\Phi UE_\beta)$ but a priori $(\Phi UE_\beta)$ is a weaker condition than $(UE_\beta)$. Since $\Phi$ is a bounded function, $(\Phi UE_\beta) \Rightarrow (DUE_\beta)$ by the symmetry of the heat kernel.

The following theorem is the main result of this paper.

**Theorem 1.1** *Let $(M, d)$ is a locally compact separable metric space, $\mu$ be a Radon measure on $M$ with full support, and $(\mathcal{E}, \mathcal{F})$ be a regular Dirichlet form in $L^2(M, \mu)$. Assume in addition that:*

*(a)* *$(M, d)$ is connected and* $\operatorname{diam}(M) = \infty$.

*(b)* *Measure $\mu$ satisfies the volume doubling property $(VD)$.*

*(c)* *The form $(\mathcal{E}, \mathcal{F})$ is local.*

*(d)* *The process $\{X_t\}_{t \ge 0}$ is stochastically complete.*

*(e)* *The diffusion process $\{X_t\}_{t \ge 0}$ associated with $(\mathcal{E}, \mathcal{F})$ admits a continuous heat kernel $p_t(x, y)$.*

*Then, for any $\beta > 1$, the following equivalences take place*

$$\begin{aligned}
(UE_\beta) \;&\Leftrightarrow\; (\Phi UE_\beta) & (1.12)\\
&\Leftrightarrow\; (DUE_\beta) + (P_\beta) \Leftrightarrow (FK_\beta) + (P_\beta) & (1.13)\\
&\Leftrightarrow\; (DUE_\beta) + (E_\beta) \Leftrightarrow (FK_\beta) + (E_\beta)\,. & (1.14)
\end{aligned}$$

In fact, Theorem 1.1 is a particular case of a more general Theorem 12.1 where neither existence nor continuity of the heat kernel is assumed. In fact, the existence of the heat kernel follows from $(FK_\beta) + (P_\beta)$ or $(FK_\beta) + (E_\beta)$. In this generality one cannot guarantee the continuity of the heat kernel, which makes all argument much more involved. Theorem 12.1 is proved in Section 12 after some preparation in the preceding sections.

In the setting of Riemannian manifolds, it was proved in [11, Proposition 5.2] that, for $\beta = 2$,

$$(UE_2) \Leftrightarrow (DUE_2) \Leftrightarrow (FK_2)\,, \tag{1.15}$$

so that in this case the hypotheses $(E_\beta)$ and $(P_\beta)$ can be dropped. However, in general $(DUE_\beta)$ is not equivalent to $(UE_\beta)$ so that the hypotheses $(E_\beta)$ or $(P_\beta)$ cannot be got rid of[1]. A weak replacement for (1.15) is the equivalence (1.12).

[1] As was pointed out by the referee, a counterexample is obtained by taking a direct product of two spaces with different values of the walk dimension $\beta$.

Kigami proved in [22] that

$$(UE_\beta) \Leftrightarrow (DUE_\beta) + (E_\beta) \Leftrightarrow (Nash) + (E_\beta),$$

where $(Nash)$ refers to a so called *local Nash inequality,* assuming in addition that

$$\inf_{x\in M} V(x,r) > 0 \quad \text{for some } r > 0. \tag{1.16}$$

The present paper is largely motivated by this result of Kigami. Our purpose here is threefold. Firstly, we use the Faber-Krahn inequality in (1.14) instead of the local Nash inequality to match (1.15). Secondly, we improve the argument of Kigami to get rid of the additional assumption (1.16) and of the continuity of the heat kernel. Thirdly, we prove new equivalences (1.12) and (1.13).

For the equivalence (1.12) it is very essential that the process $\{X_t\}$ is a diffusion (which is equivalent to the locality of $(\mathcal{E},\mathcal{F})$). Indeed, let for example $\{X_t\}$ be the symmetric stable process in $\mathbb{R}^n$ of index $\beta \in (0,2)$; that is $\{X_t\}$ is generated by $(-\Delta)^{\beta/2}$, where $\Delta$ is the Laplace operator. It is known that its heat kernel satisfies the following estimate:

$$p_t(x,y) \simeq \frac{1}{t^{n/\beta}} \frac{1}{\left(1+\frac{d(x,y)}{t^{1/\beta}}\right)^{n+\beta}},$$

that is $(\Phi UE_\beta)$ with $\Phi(s) = \frac{1}{(1+s)^{n+\beta}}$. Although this function satisfies (1.11) (note that here $\alpha = n$) and all other hypotheses of Theorem 1.1 are satisfied, too, except for the locality, the estimate $(UE_\beta)$ is obviously not true.

Acknowledgments. The author thanks Martin Barlow, Laurent Saloff-Coste, and Andras Telcs for useful discussions on the subject.

## 2 Preliminaries

Unless otherwise stated, here and in the rest of this paper let $(M,d)$ be a locally compact separable metric space, $\mu$ be a Radon measure on $M$ with full support, and $(\mathcal{E},\mathcal{F})$ be a regular[2] Dirichlet form in $L^2 = L^2(M,\mu)$.

It is well known that such a form has a *generator*, which will be denoted by $H$ and which is a positive definite self-adjoint operator in $L^2$. The domain $\mathrm{dom}(H)$ is a dense subspace of $\mathcal{F}$, and for all $f \in \mathcal{F}$ and $g \in \mathrm{dom}(H)$, we have

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

The operator $H$ determines a *heat semigroup* $\{P_t\}_{t\geq 0}$ by

$$P_t = \exp(-tH),$$

so that $P_t$ is a bounded self-adjoint operator in $L^2$ (and even $\|P_t\| \leq 1$). In addition, the semigroup $\{P_t\}_{t\geq 0}$ is strongly continuous in $L^2$ and is Markovian. The latter means that $f \geq 0$ implies $P_t f \geq 0$ and $f \leq 1$ implies $P_t f \leq 1$. and The Markovian properties of $P_t$ allow to extend $P_t$ from $L^1 \cap L^2$ to a bounded operator in $L^1$ and then, by duality, to a bounded operator in $L^\infty$.

[2]The form $(\mathcal{E},\mathcal{F})$ is called regular if $\mathcal{F} \cap C_0(M)$ is dense both in $\mathcal{F}$ and in $C_0(M)$.

**Definition.** A family $\{p_t\}_{t>0}$ of $\mu \times \mu$-measurable functions on $M \times M$ is called a *heat kernel* of the form $(\mathcal{E}, \mathcal{F})$ if $p_t$ is an integral kernel of the operator $P_t$, that is, for any $t > 0$ and for any $f \in L^2$,

$$P_t f(x) = \int_M p_t(x,y) f(y) \, d\mu(y) \quad \text{for } \mu\text{-a.a. } x \in M. \tag{2.1}$$

Clearly, of a heat kernel exists then, for any $t > 0$, $p_t(\cdot,\cdot)$ is uniquely defined on $M \times M$ up to a change on a set of measure 0. It is easy to see that a heat kernel satisfies the following properties, for all $t, s > 0$ and for $\mu$-a.a. $x, y \in M$:

- The positivity: $p_t(x,y) \geq 0$ (follows from $P_t f \geq 0$ for $f \geq 0$).
- The total mass inequality:
$$\int_M p_t(x,z) d\mu(z) \leq 1 \tag{2.2}$$
(follows from $P_t f \leq 1$ for $f \leq 1$).
- The symmetry: $p_t(x,y) = p_t(y,x)$ (follows from the self-adjointness of $P_t$, that is $(P_t f, g)_{L^2} = (f, P_t g)_{L^2}$).
- The semigroup property:
$$p_{s+t}(x,y) = \int_M p_s(x,z) p_t(z,y) d\mu(z) \tag{2.3}$$
(follows from $P_{t+s} = P_t P_s$).

By [9, Theorem 7.2.1], any regular Dirichlet form $(\mathcal{E}, \mathcal{F})$ admits an associated Hunt process $\left\{\{X_t\}_{t\geq 0}, \{\mathbb{P}_x\}_{x\in M}\right\}$ where $\mathbb{P}_x$ is a probability measure defined on the space of paths started at the point $x \in M$. By [9, Theorem 7.2.2], if the form $(\mathcal{E}, \mathcal{F})$ is local (which will be sometimes assumed) then the process $\{X_t\}$ is a *diffusion*, that is, the path $t \mapsto X_t$ is continuous almost surely.

The *transition function* $\mathcal{P}_t(x, B)$ of the Hunt process is defined by

$$\mathcal{P}_t(x,B) = \mathbb{P}_x(X_t \in B),$$

where $t > 0$, $x \in M$, and $B$ is a Borel subset of $M$. Hence, $\mathcal{P}_t(x,\cdot)$ is a probability measure on $X$ (possibly, with added cemetery), for any $x \in M$ and $t > 0$. Respectively, $\mathcal{P}_t$ acts as a semigroup on the space of bounded (or non-negative) Borel functions[3] by

$$\mathcal{P}_t f(x) = \int_X f(y) \mathcal{P}_t(x, dy) = \mathbb{E}_x f(X_t),$$

for all $x \in M$ and $t > 0$, where $\mathbb{E}_x$ is expectation associated with $\mathbb{P}_x$.

The relation of the Hunt process with the Dirichlet from is given by the identity

$$P_t f(x) = \mathcal{P}_t f(x) \quad \text{for } \mu\text{-a.a. } x \in M \text{ and all } t > 0, \tag{2.4}$$

for all bounded Borel functions $f$; in other words, we have

$$P_t f(x) = \mathbb{E}_x f(X_t) \quad \text{for } \mu\text{-a.a. } x \in M \text{ and all } t > 0. \tag{2.5}$$

[3] If the process $\{X_t\}$ is not stochastically complete then the value of $f$ at the cemetery is assumed to be 0.

Note that $P_t f(x)$ is defined for $\mu$-a.a. $x \in M$, whereas $\mathcal{P}_t f(x)$ is defined for *all* $x \in M$. Hence, the semigroup $\mathcal{P}_t$ contains some extra information compared to the semigroup $P_t$ although $\mathcal{P}_t$ and $P_t$ are identical as semigroups in $L^\infty$.

The identity (2.4) allows to extend $\mathcal{P}_t f$ to Borel functions $f$ from $L^2$. Indeed, assuming $f \geq 0$, set

$$\mathcal{P}_t f = \lim_{n\to\infty} \mathcal{P}_t (f \wedge n)$$

and observe that the limit is monotone and is finite $\mu$-a.a. $x \in M$ because so is $P_t f$.

For any open set $\Omega \subset M$, define $\mathcal{F}(\Omega)$ by (1.4) so that $(\mathcal{E}, \mathcal{F}(\Omega))$ is a regular Dirichlet form in $L^2(\Omega, \mu)$ (see [9, Theorem 4.4.3, p.154]). Hence, all notions defined for the form $(\mathcal{E}, \mathcal{F})$ make sense also for the form $(\mathcal{E}, \mathcal{F}(\Omega))$, in particular, the generator $H_\Omega$ and the heat semigroup $P_t^\Omega$. If $P_t^\Omega$ has a heat kernel then it is called the *Dirichlet heat kernel* of $\Omega$ and is denoted by $p_t^\Omega(x,y)$ (this terminology is motivated by the fact that in a classical setting $p_t^\Omega$ satisfies the Dirichlet boundary condition on $\partial\Omega$). It is frequently convenient to extend $p_t^\Omega(x,y)$ to entire $M$ so that $p_t^\Omega(x,y) = 0$ if $x$ or $y$ is outside $\Omega$.

In the general case when the heat kernel does not necessarily exists or is not a continuous function, we need to modify some of the conditions defined in Introduction as follows. Recall that $\beta > 1$ is a fixed real number.

$(UE_\beta)$ : A heat kernel exists and satisfies the estimate

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

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

$(DUE_\beta)$ A heat kernel exists and satisfies the estimate

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

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

$(\Phi UE_\beta)$ A heat kernel exists and satisfies the estimate

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

for $\mu$-a.a. $x, y \in M \setminus N$ and for all $t > 0$, where $\Phi(s)$ is a decreasing positive function on $[0, +\infty)$ satisfying (1.11).

**Definition.** A Borel set $N \subset M$ is called *negligible* for the process $X_t$ if $\mu(N) = 0$ and

$$\mathbb{P}_x (X_t \in N \text{ or } X_{t-} \in N \text{ for some } t \geq 0) = 0 \quad \text{for all } x \in M \setminus N.$$

$(E_\beta)$ : There exists a negligible set $N \subset M$ such that, for all $x \in M \setminus N$ and $r > 0$,

$$cr^\beta \leq \mathbb{E}_x \tau_{B(x,r)} \leq Cr^\beta.$$

$(P_\beta)$ There exists a negligible set $N \subset M$ such that, for all $x \in M \setminus N$ and $r > 0$,

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

for some $\varepsilon \in (0,1)$ and $\delta > 0$.

Finally, the conditions $(VD)$ and $(FK)$ remain unchanged. The statement of Theorem 1.1 is then slightly changed: the condition $(e)$ of the existence of a heat kernel is no longer needed as it is build into $(DUE_\beta)$, $(UE_\beta)$, and $(\Phi UE_\beta)$. This version of Theorem 1.1 is stated below in Section 12.

# 3 Heat semigroups and heat kernels

Denote by $\|\cdot\|_p$ the norm in $L^p = L^p(M,\mu)$. Also, denote by esup essential supremum; in particular, we have $\|f\|_\infty = \operatorname{esup}_M |f|$.

**Lemma 3.1** *If a heat kernel $p_t(x,y)$ exists then, for any measurable set $U \subset M$, the function*

$$t \mapsto \operatorname*{esup}_{x,y\in U} p_t(x,y)$$

*is non-increasing on $(0,+\infty)$.*

**Proof.** We have

$$\operatorname*{esup}_{x,y\in U} p_t(x,y) = \sup_{f,g} \int_U p_t(x,y) f(x) g(y)\, d\mu(x)\, d\mu(y) = \sup_{f,g} (P_t f, g), \tag{3.1}$$

where $(\cdot,\cdot)$ is the inner product in $L^2(M,\mu)$ and the supremum is taken over non-negative functions $f,g \in L^1 \cap L^2(U,\mu)$ such that $\|f\|_1 = \|g\|_1 = 1$. The symmetry of $P_t$ and the semigroup property imply

$$(P_t f, g) = \left(P_{t/2} f, P_{t/2} g\right) \le \|P_{t/2} f\|_2 \|P_{t/2} g\|_2 = (P_t f, f)^{1/2} (P_t g, g)^{1/2},$$

whence

$$\sup_{f,g} (P_t f, g) \le \sup_f (P_t f, f).$$

Since the opposite inequality is trivial, we have in fact

$$\sup_{f,g} (P_t f, g) = \sup_f (P_t f, f).$$

Finally, $(P_t f, f) = \|P_{t/2} f\|_2^2$ is non-increasing in $t$, whence the claim follows. ■

**Definition.** We say that a semigroup $P_t$ is $L^p \to L^q$ *ultracontractive* (where $1 \le p < q \le +\infty$) if there exists a positive decreasing function $\gamma(t)$ on $(0,+\infty)$ (called the *rate* function) such that, for all $t > 0$,

$$\|P_t f\|_q \le \gamma(t) \|f\|_p \quad \text{for all } f \in L^p \cap L^2. \tag{3.2}$$

It is easy to see that if $P_t$ is $L^p \to L^q$ ultracontractive then $P_t$ is also $L^{q^*} \to L^{p^*}$ ultracontractive with the same rate function, where $p^*$ and $q^*$ are the Hölder conjugates to $p$ and $q$, respectively. Indeed, any $f \in L^{q^*} \cap L^2$, we have by (3.2)

$$\|P_t f\|_{p^*} = \sup_{g \in L^p \cap L^2 \setminus \{0\}} \frac{(P_t f, g)}{\|g\|_p} = \sup_{g \in L^p \cap L^2 \setminus \{0\}} \frac{(f, P_t g)}{\|g\|_p} \le \sup_{g \in L^p \cap L^2 \setminus \{0\}} \|f\|_{q^*} \frac{\|P_t g\|_q}{\|g\|_p} \le \gamma(t) \|f\|_{q^*},$$

whence the claim follows.

In particular, $P_t$ is $L^1 \to L^2$ ultracontractive if and only if it is $L^2 \to L^\infty$ ultracontractive. In this case, we say that $P_t$ is ultracontractive.

The next lemma relates the ultracontractivity of $P_t$ with the existence of a heat kernel satisfying a uniform upper bound. This fact is well known but there hardly exists a reference with a detailed proof matching our setting (see [5], [1, Propositions 4.13, 4.14], [8, Lemma 2.1.2], [29] for proofs in various settings). So, we give a full proof here.

**Lemma 3.2** *The semigroup $P_t$ is ultracontractive with the rate function $\gamma(t)$ if and only if $P_t$ has a heat kernel $p_t(x,y)$ satisfying for all $t>0$ the estimate*

$$\operatorname*{esup}_{x,y\in M} p_t(x,y) \le \gamma(t/2)^2. \tag{3.3}$$

**Proof.** If a heat kernel exists and satisfies (3.3) then we have by (3.1) and (3.3)

$$(P_{2t}f,g) \le \operatorname*{esup}_{x,y\in M} p_{2t}(x,y)\,\|f\|_1\|g\|_1 \le \gamma(t)^2\,\|f\|_1\|g\|_1,$$

for all $f,g\in L^2\cap L^1$. Taking $f=g$ and noticing that $(P_{2t}f,f)=\|P_tf\|_2^2$, we obtain

$$\|P_tf\|_2 \le \gamma(t)\,\|f\|_1,$$

that is, $P_t$ is $L^1\to L^2$ ultracontractive.

Conversely, if $P_t$ is $L^1\to L^2$ ultracontractive then $P_t$ is also $L^2\to L^\infty$ ultracontractive, that is, for all $f\in L^2$ and $t>0$,

$$\|P_tf\|_\infty \le \gamma(t)\|f\|_2. \tag{3.4}$$

Fix $t>0$. For any $f\in L^2$, $P_tf$ is an element of $L^\infty\cap L^2$ and, hence, is defined for *$\mu$-almost* all $x$. We would like to choose a *pointwise* function realization of $P_tf(x)$ while keeping the linearity of $P_t$. Denote by $\mathcal{L}^\infty$ the set of all bounded measurable functions on $M$ defined pointwise. Then $\mathcal{L}^\infty$ is a Banach space with the sup-norm (in contrast to $L^\infty$ where the norm is the essential supremum esup).

We claim that there exists a linear operator[4] $\widetilde{P}_t : L^2\to\mathcal{L}^\infty$ such that, for any $f\in L^2$,

$$\widetilde{P}_tf = P_tf \quad \mu\text{-a.e.} \tag{3.5}$$

and

$$\sup\left|\widetilde{P}_tf\right| \le \gamma(t)\|f\|_2. \tag{3.6}$$

Observe that, for any $\varphi\in L^\infty$, there exists a norm preserving realization of $\varphi$ in $\mathcal{L}^\infty$, that is, a function $\varphi'\in\mathcal{L}^\infty$ such that

$$\varphi'=\varphi \ \ \mu\text{-a.e.} \quad \text{and} \quad \sup|\varphi'| = \operatorname{esup}|\varphi|.$$

Indeed, fix any pointwise realization of $\varphi$ and observe that the set

$$E(\varphi) := \{x\in M : |\varphi(x)| > \operatorname{esup}|\varphi|\}$$

has $\mu$-measure 0. Then define $\varphi'(x)$ to be equal to $\varphi(x)$ outside $E(\varphi)$, and to vanish on $E(\varphi)$.

[4] If $M$ is a Riemannian manifold then $\widetilde{P_tf}$ can be defined as a continuous realization of $P_tf$, because of the hypoellipticity of the Laplace operator. In general, one can only ensure that $P_tf$ has a *quasi-continuous* realization (see Lemma 7.1). If no point in $M$ is polar (as happens on many fractal spaces) then any quasi-continuous function is continuous, so that $\widetilde{P_tf}$ can again be defined as a continuous realization of $P_tf$. However, this does not work in general, although the estimates we obtain in this paper can help establishing the continuity of $P_tf$ *a posteriori*.

Let $\{v_k\}_{k=1}^{\infty}$ be an orthonormal basis in $L^2$, and let $V$ be the set of all finite linear combinations of functions $v_k$ with rational coefficients. Define $\widetilde{P}_t v_k$ to be a norm-preserving realization of $P_t v_k$ in $\mathcal{L}^{\infty}$; then extend $\widetilde{P}_t$ to the whole space $V$ by linearity. In particular, we have, for all $f \in V$,

$$\widetilde{P}_t f = P_t f \quad \mu\text{-a.e.} \tag{3.7}$$

The set $V$ is countable. Since each set $E(\widetilde{P}_t f)$ has measure 0, the union $U$ of all sets $E(\widetilde{P}_t f)$ over all $f \in V$ has also measure 0. Now we modify the definition of $\widetilde{P}_t f$ for every $f \in V$ by setting $\widetilde{P}_t f$ to be zero on $U$ (and not changing it outside $U$). Clearly, the linearity and (3.7) are preserved, but we acquire in addition that

$$\sup \left| \widetilde{P}_t f \right| = \operatorname{esup} |P_t f|,$$

which together with (3.4) implies

$$\sup \left| \widetilde{P}_t f \right| \le \gamma(t) \|f\|_2 \quad \text{for all } f \in V. \tag{3.8}$$

Hence, $\widetilde{P}_t$ is a bounded linear mapping from $(V, \|\cdot\|_2)$ to $\mathcal{L}^{\infty}$. Since $V$ is dense in $L^2$, $\widetilde{P}_t$ uniquely extends to a bounded linear mapping from $L^2$ to $\mathcal{L}^{\infty}$, which satisfies (3.5) and (3.6).

Fix $x \in M$, and consider a linear functional on $L^2$ defined by

$$f \mapsto \widetilde{P}_t f(x).$$

By (3.6), this is a bounded linear functional in $L^2$. By the Riesz representation theorem, there exists a function $p_{t,x} \in L^2$ such that, for any $f \in L^2$,

$$\widetilde{P}_t f(x) = (p_{t,x}, f)_{L^2} = \int_M p_{t,x}(y) f(y) d\mu(y). \tag{3.9}$$

The function $p_t(x, y) := p_{t,x}(y)$ will be the heat kernel if we prove that it is measurable jointly in $x, y \in M$. To that end, we use again the orthonormal basis $\{v_k\}_{k=1}^{\infty}$ in $L^2$. For any index $k$, the function $u_k := \widetilde{P}_t v_k$ is in $L^2$ and hence is measurable. On the other hand, we have by (3.9), for all $x \in M$,

$$u_k(x) = (p_{t,x}, v_k)_{L^2}.$$

The Parseval identity yields

$$\|p_{t,x}\|_2^2 = \sum_k |u_k(x)|^2,$$

whence it follows that the function $x \mapsto \|p_{t,x}\|_2$ is measurable.

It follows from (3.8) and (3.9) that

$$\|p_{t,x}\|_2 \le \gamma(t). \tag{3.10}$$

Therefore, for any compact set $K \subset M$,

$$\int_K \left( \int_M p_t(x, y)^2 d\mu(y) \right) d\mu(x) = \int_K \|p_{t,x}\|_2^2 d\mu(x) \le \gamma^2(t) \mu(K) < \infty,$$

and, by Fubini's theorem, $p_t(x, y) \in L^2(K \times M)$. By the local compactness of $M$, it follows that $p_t(x, y)$ is jointly measurable in $x, y$.

By the semigroup property and the symmetry of the heat kernel, we have, for $\mu$-a.a. $x, y \in M$,

$$p_t(x, y) = \int_M p_{t/2}(x, z) p_{t/2}(z, y) d\mu(z) \le \|p_{t/2,x}\|_2 \|p_{t/2,y}\|_2 \le \gamma(t/2)^2,$$

which was to be proved. ∎

## 4 The Dirichlet heat kernel

The results of this section are known in the setting of manifolds (see for example [18], [11], [13], [7]). Here we have modified the argument to adjust to the present singular setting. Recall that the spectral gap $\lambda_{\min}(\Omega)$ is defined by (1.5).

**Lemma 4.1** *Let $U \subset M$ be an open set such that $\mu(U) < \infty$. Assume that, for all non-empty open sets $\Omega \subset U$,*

$$\lambda_{\min}(\Omega) \geq a\mu(\Omega)^{-\nu}, \tag{4.1}$$

*where $a$ and $\nu$ are positive constants. Then, for any non-negative function $u \in \mathcal{F}(U) \setminus \{0\}$,*

$$\mathcal{E}[u] \geq c_\nu a \|u\|_2^{2+2\nu} \|u\|_1^{-2\nu}, \tag{4.2}$$

*where $c_\nu$ is a positive constant depending only on $\nu$.*

**Proof.** Assume first $u \in \mathcal{F}(U) \cap C_0(U)$. By the Markov property, for any $s \geq 0$ we have $(u-s)_+ \in \mathcal{F}(U)$ and

$$\mathcal{E}[u] \geq \mathcal{E}\left[(u-s)_+\right]. \tag{4.3}$$

The set $U_s := \{x \in U : u(x) > s\}$ is open. Since $(u-s)_+$ vanishes outside $U_s$, we obtain that $(u-s)_+ \in \mathcal{F}(U_s)$ whence by (1.5)

$$\mathcal{E}\left[(u-s)_+\right] \geq \lambda_{\min}(U_s) \int_{U_s} (u-s)_+^2 \, d\mu. \tag{4.4}$$

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

$$(u-s)_+^2 \geq u^2 - 2su,$$

which implies upon integration

$$\int_U (u-s)_+^2 d\mu \geq B - 2sA. \tag{4.5}$$

On the other hand, we have

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

and the assumption (4.1) yields

$$\lambda_{\min}(U_s) \geq a\mu(U_s)^{-\nu} \geq a\left(\frac{s}{A}\right)^\nu. \tag{4.6}$$

Combining (4.3), (4.4), (4.5), and (4.6), we obtain

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

Taking here $s = \frac{B}{4A}$, we finish the proof.

Consider now the general case $u \in \mathcal{F}(U)$. By the regularity of $(\mathcal{E}, \mathcal{F}(U))$, there exists a sequence $\{u_n\} \in \mathcal{F}(U) \cap C_0(U)$ such that

$$\|u_n - u\|_2 \longrightarrow 0 \quad \text{and} \quad \mathcal{E}[u_n - u] \longrightarrow 0. \tag{4.7}$$

Since $\mu(U) < \infty$, the Cauchy-Schwarz inequality yields

$$\|u_n - u\|_1 \leq \sqrt{\mu(U)} \|u_n - u\|_2 \to 0. \tag{4.8}$$

For each $u_n$, (4.2) holds by the previous argument. Passing to the limit as $n \to \infty$ we obtain (4.2) for $u$. ■

The next lemma is a modification of the Nash argument [24], which allows to obtain a heat kernel upper bound from the Nash type inequality (4.2) (see also [5], [11, Theorem 2.1]).

**Lemma 4.2** *Under the hypotheses of* Lemma 4.1, *a heat kernel* $p_t^U$ *exists and satisfies the inequality*

$$\operatorname*{esup}_{x,y\in U} p_t^U(x,y) \le C(at)^{-1/\nu} \tag{4.9}$$

*for all* $t > 0$*, where* $C = C(\nu)$*.*

**Proof.** Lemma 4.1 says that, for all non-negative $u \in \mathcal{F}(U) \setminus \{0\}$,

$$\mathcal{E}[u] \ge ca\|u\|_2^{2+2\nu}\|u\|_1^{-2\nu}. \tag{4.10}$$

Let $f \in L^2(U,\mu)$ be non-negative and $\|f\|_1 = 1$. Set $u_t = P_t^U f$ for all $t > 0$ and observe that $u_t \in \operatorname{dom}(H) \subset \mathcal{F}$ where $H = H_U$ is the generator of the form $(\mathcal{E}, \mathcal{F}(U))$. Moreover, we have $\frac{du_t}{dt} = -Hu_t$, whence

$$\left(\frac{du_t}{dt}, u_t\right) = -(Hu_t, u_t) = -\mathcal{E}[u_t].$$

On the other hand, differentiating the function $J(t) := \|u_t\|_2^2$ we obtain

$$\frac{dJ}{dt} = \frac{d}{dt}(u_t, u_t) = 2\left(\frac{du_t}{dt}, u_t\right) = -2\mathcal{E}[u_t]. \tag{4.11}$$

It follows from (2.2) that $\|u_t\|_1 \le 1$. Combining (4.10) and (4.11), we obtain the differential inequality

$$\frac{dJ}{dt} \le -caJ^{1+\nu},$$

whence $J(t) \le C(at)^{-1/\nu}$. Consequently, the semigroup $P_t^U$ is $L^1 \to L^2$ ultracontractive with the estimate

$$\|P_t^U\|_{1\to 2}^2 \le C(at)^{-1/\nu},$$

whence, by Lemma 3.2, $P_t^U$ has a heat kernel satisfying (4.9). ∎

# 5 The transition function and local ultracontractivity

For any open set $\Omega \subset M$, the Hunt process $\left\{\{X_t^\Omega\}_{t\ge 0}, \{\mathbb{P}_x^\Omega\}_{x\in M}\right\}$ associated with the Dirichlet form $(\mathcal{E}, \mathcal{F}(\Omega))$ is obtained from $X_t$ by killing the latter outside $\Omega$. The transition function $\mathcal{P}_t^\Omega$ of the process $X_t^\Omega$ is given by

$$\mathcal{P}_t(x,B) = \mathbb{P}_x^\Omega(X_t \in A) = \mathbb{P}_x(t < \tau_\Omega \text{ and } X_t \in B)$$

where $\tau_\Omega$ is the first exit time of the process $X_t$ from $\Omega$ defined by (1.3) (see [9, (4.1.2)]). Consequently, we have

$$\mathcal{P}_t^\Omega f(x) = \mathbb{E}_x^\Omega(f(X_t)) = \mathbb{E}_x\left(\mathbf{1}_{\{t<\tau_\Omega\}} f(X_t)\right), \tag{5.1}$$

for all $x \in M$, $t > 0$, and a bounded (or non-negative) Borel function $f$. For the heat semigroup $P_t^\Omega$ of the form $(\mathcal{E}, \mathcal{F}(\Omega))$, we have then

$$P_t^\Omega f(x) = \mathbb{E}_x\left(\mathbf{1}_{\{t<\tau_\Omega\}} f(X_t)\right) \quad \text{for } \mu\text{-a.a. } x \in M. \tag{5.2}$$

Clearly, the semigroup $P_t^\Omega$ is dominated by $P_t$, that is, $P_t^\Omega f \le P_t f$ for any non-negative function $f$. In particular, if $P_t$ is ultracontractive then $P_t^\Omega$ is also ultracontractive.

**Definition.** A sequence $\{\Omega_n\}$ of subsets of $M$ is called *exhausting* if $\Omega_n \subset \Omega_{n+1}$ and $\cup_n \Omega_n = M$.

**Lemma 5.1** *If $\{\Omega_n\}_{n=1}^{\infty}$ is an exhausting sequence of open sets in $M$ and a Dirichlet heat kernel $p_t^{\Omega_n}$ exists for any $n$ then the sequence $\{p_t^{\Omega_n}\}$ increases, and $\lim_{n\to\infty} p_t^{\Omega_n}$ determines a heat kernel $p_t$ of $P_t$.*

**Proof.** Since $\tau_\Omega$ is monotone in $\Omega$, it is obvious from (5.2) that $p_t^\Omega$ in monotone in $\Omega$. By (5.2) we have, for any non-negative Borel function $f$ on $M$,

$$\mathbb{E}_x\left(\mathbf{1}_{\{t<\tau_{\Omega_n}\}} f(X_t)\right) = \int_{\Omega_n} p_t^{\Omega_n}(x,y) f(y)\, d\mu(y)$$

for $\mu$-a.a. $x \in M$ and all $t > 0$. Letting $n \to \infty$ and setting $p_t := \lim_{n\to\infty} p_t^{\Omega_n}$, we obtain

$$\mathbb{E}_x(f(X_t)) = \int_M p_t(x,y) f(y)\, d\mu(y).$$

It follows from (2.5) that $p_t$ is a heat kernel of $P_t$. ■

**Definition.** We say that the semigroup $P_t$ is *locally ultracontractive* if there exists an exhausting sequence of open sets $\{\Omega_k\}_{k=1}^{\infty}$ such that the semigroup $P_t^{\Omega_k}$ is ultracontractive for any $k = 1, 2, ....$

It follows from Lemmas 3.2 and 5.1 that if $P_t$ is locally ultracontractive then a heat kernel exists.

**Lemma 5.2** *If a heat kernel $p_t(x,y)$ exists and is locally bounded (that is, belongs to $L^\infty_{loc}(M \times M)$) then the semigroup $P_t$ is locally ultracontractive.*

**Remark.** As we will see in Lemma 8.1, if the semigroup $P_t$ is locally ultracontractive then a heat kernel exists.

**Proof.** Since $M$ is locally compact and separable, there exists an exhausting sequence $\{\Omega_k\}_{k=1}^{\infty}$ of precompact open sets. By hypothesis, we have, for any $k$,

$$\gamma_k(t) := \operatorname*{esup}_{x,y\in\Omega_k} p_t(x,y) < \infty.$$

By Lemma 3.1, the function $\gamma_k(t)$ is decreasing. Using the fact that $P_t^{\Omega_k}$ is dominated by $P_t$ and the first part of the proof of Lemma 3.2, we obtain that $P_t^{\Omega_k}$ is ultracontractive with the rate function $\gamma_k$. ■

# 6 Mean exit time and the spectral gap

Let $f$ be a non-negative Borel function on $M$ and $\varphi(t)$ be a non-negative continuous function on $[0, +\infty)$. Multiplying (5.1) by $\varphi(t)$ and integrating in $t$, we obtain, for any open set $\Omega \subset M$ and all $x \in M$,

$$\int_0^\infty \varphi(t)\, \mathcal{P}_t^\Omega f(x)\ dt = \mathbb{E}_x \int_0^{\tau_\Omega} \varphi(t) f(X_t)\, dt. \tag{6.1}$$

In particular, for $\varphi \equiv 1$, we obtain

$$\int_0^\infty \mathcal{P}_t^\Omega f(x)\ dt = \mathbb{E}_x \int_0^{\tau_\Omega} f(X_t)\, dt, \tag{6.2}$$

whence it follows, for $f \equiv 1$, that

$$\mathbb{E}_x\tau_\Omega = \int_0^\infty \mathcal{P}_t^\Omega 1(x)\ dt. \tag{6.3}$$

For any open set $\Omega \subset M$, set

$$\overline{E}(\Omega) = \operatorname*{esup}_{x\in\Omega} \mathbb{E}_x\tau_\Omega. \tag{6.4}$$

**Lemma 6.1** *For any non-empty open set $\Omega \subset M$, we have*

$$\lambda_{\min}(\Omega) \geq \frac{1}{\overline{E}(\Omega)}. \tag{6.5}$$

**Remark.** This inequality is well known in the setting of random walks on graphs and diffusions on manifolds. Here we give a proof in a full generality.

**Proof.** Let $H = H_\Omega$ be the generator of the form $(\mathcal{E}, \mathcal{F}(\Omega))$ in $L^2(\Omega, \mu)$. For any $T > 0$ and consider the following operator

$$G_T = \int_0^T e^{-tH} dt = \varphi_T(H),$$

where

$$\varphi_T(\lambda) = \int_0^T e^{-t\lambda} dt = \frac{1 - e^{-T\lambda}}{\lambda}.$$

Since the function $\varphi_T$ is bounded and continuous on $[0, +\infty)$, the operator $G_T$ is a bounded self-adjoint operator in $L^2$. Since the function $\varphi_T$ is decreasing, we obtain by the spectral mapping theorem

$$\varphi_T(\lambda_{\min}(\Omega)) = \varphi_T(\inf \operatorname{spec}(H)) = \sup \operatorname{spec}(G_T).$$

Note that

$$\sup \operatorname{spec}(G_T) = \|G\|_{2\to 2},$$

where $\|\cdot\|_{p\to p}$ stands for the operator norm of an operator in $L^p(\Omega, \mu)$. We will prove below that, for all $T > 0$,

$$\|G\|_{2\to 2} \leq \overline{E}(\Omega), \tag{6.6}$$

Assuming that much, we obtain from the above three lines

$$\varphi_T(\lambda_{\min}(\Omega)) \leq \overline{E}(\Omega).$$

Letting $T \to \infty$ and observing that $\varphi_T(\lambda) \to 1/\lambda$, we obtain (6.5).

To verify (6.6), recall that the operator $e^{-tH} = P_t^\Omega$ can be considered as a bounded operator in $L^\infty$. Therefore, the operator $G_T$ also extends to a bounded operator in $L^\infty$. Since $P_t^\Omega$ and $\mathcal{P}_t^\Omega$ coincide as operators in $L^\infty$, we see that, for any bounded Borel function $f$,

$$G_T f = \int_0^T \left(\mathcal{P}_t^\Omega f\right) dt \quad \mu\text{-a.e.}.$$

Therefore, for $\mu$-a.a. $x \in \Omega$, we obtain

$$|G_T f(x)| \leq \int_0^\infty \mathcal{P}_t^\Omega |f|(x)\, dt = \mathbb{E}_x \int_0^{\tau_\Omega} |f|(X_t)\, dt \leq (\mathbb{E}_x \tau_\Omega) \sup |f|,$$

that is, using (6.4),

$$\operatorname*{esup}_\Omega |G_T f| \leq \overline{E}(\Omega) \sup |f|.$$

This implies, for any $g \in L^1 \cap L^2(\Omega, \mu)$,

$$\|G_T g\|_1 = \inf_{f \in C_0(\Omega)\setminus\{0\}} \frac{(G_T g, f)}{\|f\|_\infty} = \inf_{f \in C_0(\Omega)\setminus\{0\}} \frac{(g, G_T f)}{\|f\|_\infty} \leq \overline{E}(\Omega)\|g\|_1$$

that is,

$$\|G_T\|_{1\to 1} \leq \overline{E}(\Omega). \tag{6.7}$$

Since $P_t^{\Omega}$ is a positivity preserving operator, so is also $G_T$, that is $f \geq 0$ implies $G_T f \geq 0$, for any Borel function $f$. In particular, for any $s \in \mathbb{R}$ we have $G_T (f+s)^2 \geq 0$, that is

$$G_T f^2 + 2sG_T f + s^2 G_T 1 \geq 0,$$

whence

$$(G_T f)^2 \leq G_T 1 G_T f^2 \leq \overline{E}(\Omega) G_T f^2.$$

Therefore,

$$\|G_T f\|_2^2 \leq \overline{E}(\Omega) \|G_T f^2\|_1 \leq \overline{E}(\Omega)^2 \|f^2\|_1 = \overline{E}(\Omega)^2 \|f\|_2^2,$$

that is

$$\|G_T\|_{2\to 2} \leq \overline{E}(\Omega),$$

which was to be proved. ■

# 7 Negligible sets

This section is mostly based on the book by Fukushima, Oshima, Takeda [9] and complements some statements from this book. Our main result here is Lemma 7.5.

**Definition.** For any set $E \subset M$, the capacity cap (also called 1-capacity) is defined by

$$\operatorname{cap}(E) := \inf_{\varphi} \mathcal{E}_1[\varphi] \tag{7.1}$$

where $\mathcal{E}_1[\varphi] := \mathcal{E}[\varphi] + \|\varphi\|_2^2$ and $\varphi$ varies over all functions from $\mathcal{F}$ such that $\varphi \geq 1$ in an open neighborhood of $E$ (see [9, p.64]).

Clearly, we have $\operatorname{cap}(E) \geq \mu(E)$. Also, it is obvious from the definition that $\operatorname{cap}(E)$ is monotone function of $E$.

**Definition.** A function $f$ on $M$ is called *quasi-continuous* if, for any $\varepsilon > 0$, there exists an open set $E$ such that $\operatorname{cap}(E) < \varepsilon$ and $f$ is continuous in $M \setminus E$ (see [9, p.67]).

**Lemma 7.1** ([9, Theorem 4.2.3, p.144]) *For any Borel function $f \in L^2$ and for any $t > 0$, the function $x \mapsto \mathcal{P}_t f(x)$ is quasi-continuous on $M$.*

**Definition.** A Borel set $N \subset M$ is called *negligible* if $\mu(N) = 0$ and

$$\mathbb{P}_x (X_t \in N \text{ or } X_{t-} \in N \text{ for some } t \geq 0) = 0 \quad \text{for all } x \in M \setminus N.$$

**Lemma 7.2** *If $\operatorname{cap}(E) = 0$ then there is a negligible set $N \supset E$.*

**Proof.** Indeed, by [9, Proof of Theorem 4.2.1(ii), p.142], any set of capacity 0 is "exceptional", and by [9, Theorem 4.1.1, p.137], any exceptional set is contained in a Borel "properly exceptional" set, which, by the above definition, is negligible. ■

**Lemma 7.3** *Let $S$ be a countable family of quasi-continuous functions on $M$. Then there exists a negligible set $N \subset M$ such that, for any non-empty open set $U \subset M$ and any $f \in S$,*

$$\sup_{U\setminus N} f = \operatorname*{esup}_U f \quad \textit{and} \quad \inf_{U\setminus N} f = \operatorname*{einf}_U f. \tag{7.2}$$

**Proof.** By [9, Theorem 2.1.2, p.67], if $S$ is a countable family of quasi-continuous functions on $M$ then there exists a *regular nest* $\{F_k\}_{k=1}^{\infty}$ such that $f|_{F_k}$ is a continuous function for any $f \in S$. The term "a regular nest" means that $\{F_k\}$ is an increasing sequence of closed sets in $M$ such that $\operatorname{cap}(M \setminus F_k) \to 0$ as $k \to \infty$, and $\mu(F_k \cap \Omega) > 0$ for any open set $\Omega$ intersecting $F_k$. The set $E = M \setminus \cup_k F_k$ has capacity 0. By Lemma 7.2, the set $E$ of capacity 0 is contained in a negligible set $N$. Let us show that, for this set $N$,

$$\sup_{U \setminus N} f = \operatorname*{esup}_U f,$$

and the same will be for inf. Since $\mu(N) = 0$, the inequality

$$\sup_{U \setminus N} f \geq \operatorname*{esup}_U f$$

is trivial.

Fix $f \in S$, a non-empty open set $U \subset M$, and prove the opposite inequality. Since $M \setminus N$ is covered by $\cup_k F_k$, we see that

$$\sup_{U \setminus N} f \leq \lim_{k \to \infty} \sup_{U \cap F_k} f.$$

Since $\mu(U) > 0$ and $\mu(N) = 0$, $U \setminus N$ is non-empty and hence the intersection $U \cap F_k$ is non-empty for large enough $k$. Fix $\varepsilon > 0$ and find a point $x \in U \cap F_k$ such that

$$\sup_{U \cap F_k} f \leq f(x) + \varepsilon. \tag{7.3}$$

Since any $f \in S$ is continuous on $F_k$, there exists an open set $\Omega$ such that $x \in \Omega \subset U$ and

$$\inf_{\Omega \cap F_k} f \geq f(x) - \varepsilon. \tag{7.4}$$

Since $\mu(\Omega \cap F_k) > 0$, we have

$$\inf_{\Omega \cap F_k} f \leq \operatorname*{esup}_{\Omega \cap F_k} f. \tag{7.5}$$

From (7.3), (7.4), (7.5), we obtain

$$\sup_{U \cap F_k} f \leq \operatorname*{esup}_{\Omega \cap F_k} f + 2\varepsilon \leq \operatorname*{esup}_U f + 2\varepsilon,$$

whence the claim follows by letting $\varepsilon \to 0$ and $k \to \infty$. ■

The next lemma is a modification of the argument in [1, after Remark 4.13].

**Lemma 7.4** *Assume that the heat semigroup $P_t$ is $L^2 \to L^\infty$ ultracontractive, that is, for any $f \in L^2$ and $t > 0$,*

$$\|P_t f\|_\infty \leq \gamma(t) \|f\|_2 \tag{7.6}$$

*where $\gamma(t)$ is a positive decreasing function on $(0, +\infty)$. Then there exists a negligible set $N \subset M$ such that for all $x \in M \setminus N$, all Borel functions $f \in L^2$, and all $t > 0$,*

$$|\mathcal{P}_t f(x)| \leq \widetilde{\gamma}(t) \|f\|_2, \tag{7.7}$$

*where $\widetilde{\gamma}(t) = \lim_{s \to t-} \gamma(s)$.*

**Proof.** Let $S_0$ be the functional family on $M$ that consists of all finite linear combinations with rational coefficients of indicator functions $\mathbf{1}_B$, where $B$ runs over a countable family of precompact open sets in $M$, which form a base of topology of $M$. Note that $S_0$ is a countable set, $S_0$ is a vector space over $\mathbb{Q}$, and $S_0$ is dense in $L^2$.

Let us show the closure $\widetilde{S_0}$ of $S_0$ under pointwise limits coincides with the set of all Borel functions on $M$. Indeed, observe first that $\widetilde{S_0}$ is a linear space over $\mathbb{R}$. Next, for any open $\Omega \subset M$, the function $\mathbf{1}_\Omega$ is in $\widetilde{S_0}$ because $\Omega$ is a union of elements from the base. Hence, also functions $\mathbf{1}_{\Omega_1} - \mathbf{1}_{\Omega_2}$ are in $\widetilde{S_0}$ for any two open sets $\Omega_1, \Omega_2$. Therefore, for any continuous function $f$ on $M$, all the functions of the form $\mathbf{1}_{\{a \leq f < b\}}$ are in $\widetilde{S_0}$, for all real $a, b$, and so are the functions

$$\sum_{k=0}^{n-1} \alpha_k \mathbf{1}_{\{a_k \leq f < \alpha_{k+1}\}} \tag{7.8}$$

for any sequence $\alpha_0 < \alpha_1 < ... < \alpha_n$ of reals. Since functions of the form (7.8) tend to $f$ pointwise for appropriate choice of partition, we see that any continuous function is in $\widetilde{S_0}$. Hence, all Borel functions are also in $\widetilde{S_0}$.

Let $S$ be the family of all functions of the form $\mathcal{P}_t f$ where $f \in S_0$ and $t$ is a positive rational. Then $S$ is countable, and every function from $S$ is quasi-continuous by Lemma 7.1. Let $N$ be a negligible set that exists for the family $S$ by Lemma 7.3, so that, for all functions $f \in S_0$ and for all $t \in \mathbb{Q}_+$, we have

$$\sup_{M \setminus N} |\mathcal{P}_t f| = \operatorname{esup}_M |\mathcal{P}_t f| = \|\mathcal{P}_t f\|_\infty. \tag{7.9}$$

In particular, for any $x \in M \setminus N$, we obtain

$$|\mathcal{P}_t f(x)| \leq \|\mathcal{P}_t f\|_\infty \leq \gamma(t) \|f\|_2.$$

Since $S_0$ is dense in $L^2$, we conclude by the Riesz representation theorem that, for all $x \in M \setminus N$ and $t \in \mathbb{Q}_+$, there exists $p_{t,x} \in L^2$ such that, for all $f \in S_0$,

$$\mathcal{P}_t f(x) = \int_M p_{t,x} f \, d\mu. \tag{7.10}$$

It is clear that

$$\|p_{t,x}\|_2 \leq \gamma(t).$$

From the positivity preserving of $\mathcal{P}_t$, we conclude that $p_{t,x} \geq 0$ $\mu$-a.e. and from $\mathcal{P}_t 1 \leq 1$ we obtain that

$$\int_M p_{t,x} d\mu \leq 1. \tag{7.11}$$

Since $\mathcal{P}_t(x, \cdot)$ is a probability measure, the both sides of (7.10) will survive passage to a limit in $f$ provided the sequence of functions is bounded and converges pointwise. Hence, we conclude that (7.10) holds for all bounded Borel functions $f$. Taking monotone increasing limits, we obtain (7.10) for all non-negative Borel functions. Hence, for any Borel function $f \in L^2$, for any $x \in M \setminus N$, and $t \in \mathbb{Q}_+$, we have

$$|\mathcal{P}_t f(x)| \leq \mathcal{P}_t |f| \leq \|p_{t,x}\|_2 \|f\|_2 \leq \gamma(t) \|f\|_2.$$

To finish the proof, we need to extend this inequality to real $t$. Let $t > 0$ be real, and let $s < t$ be a positive rational. Then we have

$$\mathcal{P}_t f(x) = \mathcal{P}_s(\mathcal{P}_{t-s} f)(x) = \int_M p_{s,x}(\mathcal{P}_{t-s} f) \, d\mu,$$

and it follows from (7.11) and (7.6) that

$$|\mathcal{P}_t f(x)| \le \|\mathcal{P}_{t-s} f\|_\infty \le \gamma(t-s)\|f\|_2.$$

Passing to the limit as $s \to 0$ we finish the proof. ■

**Lemma 7.5** *Assume that the semigroup $P_t$ is locally ultracontractive (that is, there exists a exhausting sequence of open sets $\{\Omega_k\}_{k=1}^\infty$ in $M$ such that the semigroup $P_t^{\Omega_k}$ is ultracontractive for any $\Omega_k$). Then there exists a negligible set $N \subset M$ such that, for any non-negative Borel function $f$ on $M$, for any non-empty open set $U \subset M$, and for any $t > 0$,*

$$\sup_{U \setminus N} \mathcal{P}_t f = \operatorname*{esup}_U \mathcal{P}_t f. \tag{7.12}$$

**Proof.** It suffices to prove that

$$\sup_{U \setminus N} \mathcal{P}_t f \le \operatorname*{esup}_U \mathcal{P}_t f \tag{7.13}$$

because the opposite inequality trivially follows from $\mu(N) = 0$. Note that if (7.13) holds for some set $N$ then it holds also for any larger set $N$. Let us also observe that (7.13) survives increasing monotone limits in $f$; that is, if $\{f_k\}$ is an increasing sequence of functions for which (7.13) holds and if $f_k$ converges to $f$ pointwise then (7.13) holds also for the limit function $f$. Indeed, by the monotone convergence theorem, $\mathcal{P}_t f_k$ tends to $\mathcal{P}_t f$ pointwise. Hence, we obtain

$$\sup_{U \setminus N} \mathcal{P}_t f = \lim_{k\to\infty} \sup_{U \setminus N} \mathcal{P}_t f_k \le \lim_{k\to\infty} \operatorname*{esup}_U \mathcal{P}_t f_k \le \operatorname*{esup}_U \mathcal{P}_t f.$$

Note that it suffices to prove (7.13) assuming that $P_t$ is ultracontractive. Indeed, if we know that then applying it to $P_t^{\Omega_k}$ we obtain that there exists a negligible set $N_k$ such that, for any non-negative Borel function $f$ on $M$, for any non-empty open set $U \subset M$, and for any $t > 0$,

$$\sup_{\Omega_k \cap U \setminus N_k} \mathcal{P}_t^{\Omega_k} f \le \operatorname*{esup}_{\Omega_k \cap U} \mathcal{P}_t^{\Omega_k} f.$$

Taking $N = \cup N_k$, letting $k \to \infty$, and noticing that the sequence $\left\{\mathcal{P}_t^{\Omega_k} f\right\}$ is monotone increasing, we obtain (7.13).

Next, it suffices to assume that $f \in L^2 = L^2(M, \mu)$. Indeed, an arbitrary non-negative Borel function $f$ can be approximated by an increasing sequence $\{f_k\}$ of non-negative Borel functions $f_k \in L^2$, which converges to $f$ pointwise, whence the claim follows.

Hence, we assume in the sequel that $P_t$ is ultracontractive and $f$ is a non-negative Borel function from $L^2$. Let us first fix such a function $f$ and construct a set $N = N_f$ so that (7.13) holds for this particular $f$ and for all $U$ and $t$. Let $S_0$ be the functional family that consists of all functions of the form

$$\sum_{k=0}^{n-1} \alpha_k \mathbf{1}_{\{a_k \le f < \alpha_{k+1}\}},$$

where $n$ is any positive integer, and $0 < \alpha_1 < \alpha_2 < ... < \alpha_n$ are rational. Clearly, the set $S_0$ is countable, and there is a sequence of functions $\{g_k\} \subset S_0$ that increases and converges to $f$ pointwise.

Let $S$ be the family of all functions of the form $\mathcal{P}_t g$ where $g \in S_0$ and $t$ is a positive rational. Then $S$ is countable, and every function from $S$ is quasi-continuous by Lemma 7.1. Let $N = N_{\text{Lemma 7.3}}$ be a negligible set that exists for the family $S$ by Lemma 7.3, and let $N = N_{\text{Lemma 7.4}}$ be the negligible set that exists by Lemma 7.4; let us set

$$N = N_{\text{Lemma 7.3}} \cup N_{\text{Lemma 7.4}}.$$

In particular, we have

$$\sup_{U\setminus N} \mathcal{P}_t g \leq \operatorname*{esup}_U \mathcal{P}_t g \tag{7.14}$$

for all non-empty open sets $U$, all positive rationals $t$, and all $g \in S_0$. Let $\{g_k\}$ be an increasing sequence converging to $f$ pointwise. Applying (7.14) to $g = g_k$ and passing to the limit as $k \to \infty$, we obtain (7.13) for rational $t > 0$.

Let us now show that (7.13) holds also for all real $t > 0$. Before that, let us verify that

$$\sup_{M\setminus N} |\mathcal{P}_s f - \mathcal{P}_t f| \to 0 \quad \text{as } s \to t+. \tag{7.15}$$

Indeed, by Lemma 7.4 and $f \in L^2$, we obtain

$$\sup_{M\setminus N} |\mathcal{P}_s f - \mathcal{P}_t f| = \sup_{M\setminus N} |\mathcal{P}_t (\mathcal{P}_{s-t} f - f)| \leq \widetilde{\gamma}(t) \|\mathcal{P}_{s-t} f - f\|_2 \to 0 \quad \text{as } s \to t+.$$

Now fix a real $t > 0$ and let $\{s_k\}$ be a decreasing sequence of rationals converging to $t$. We already now that

$$\sup_{U\setminus N} \mathcal{P}_{s_k} f \leq \operatorname*{esup}_U \mathcal{P}_{s_k} f.$$

Passing to the limit as $k \to \infty$ and using (7.15), we obtain (7.13).

Now let us show that a set $N$ can be chosen so that (7.13) holds simultaneously for all non-negative Borel functions from $L^2$. There exists a countable dense subset of $L^2$. Replacing any function in this subset by its positive part and taking its Borel version, we obtain a countable set $S_1$ of non-negative Borel functions in $L^2$ which is dense in $L^2_+$. Choosing as above a negligible set $N = N_g$ for each function $g \in S_1$ and then setting $N = \bigcup_{g\in S_1} N_g$, we obtain a negligible set $N$ that serves all functions from $S_1$ simultaneously.

Let us show that, with this set $N$, (7.13) holds for *any* non-negative Borel function $f$ from $L^2$. Let $\{f_k\}$ be a sequence from $S_1$ that converges to $f$ in $L^2$. By the ultracontractivity of $P_t$, we obtain, for any $t > 0$,

$$\|\mathcal{P}_t f_k - \mathcal{P}_t f\|_\infty \leq \gamma(t) \|f_k - f\|_2 \to 0 \quad \text{as } k \to \infty. \tag{7.16}$$

By Lemma 7.4, we have, for any $x \in M \setminus N$,

$$|\mathcal{P}_t f_k(x) - \mathcal{P}_t f(x)| \leq \widetilde{\gamma}(t) \|f_k - f\|_2 \to 0 \quad \text{as } k \to \infty. \tag{7.17}$$

For any function $f_k$, for any open set $U$, for any $t > 0$, and for any $x \in U \setminus N$, we have

$$\mathcal{P}_t f_k(x) \leq \operatorname*{esup}_U \mathcal{P}_t f_k.$$

Letting $k \to \infty$ and using (7.16) and (7.17), we obtain

$$\mathcal{P}_t f(x) \leq \operatorname*{esup}_U \mathcal{P}_t f$$

whence (7.13) follows. ■

# 8 The transition density

**Definition.** A family $\{\widetilde{p}_t(x,y)\}_{t>0}$ of measurable functions on $M \times M$ is called a (symmetric) *transition density* of the process $\{X_t\}$ if the following properties are satisfied:

1. There exists a negligible set $N \subset M$ such that for all $x \in M \setminus N$, for all $t > 0$, and for all bounded Borel functions $f$ on $M$,

$$\mathcal{P}_t f(x) = \int_M \widetilde{p}_t(x,y) f(y) \, d\mu(y). \tag{8.1}$$

2. For all $x, y \in M$ and $t > 0$,

$$\widetilde{p}_t(x,y) = \widetilde{p}_t(y,x). \tag{8.2}$$

3. For all $x, y \in M$ and $t, s > 0$,

$$\widetilde{p}_{t+s}(x,y) = \int_M \widetilde{p}_t(x,y) \widetilde{p}_s(y,z) \, d\mu(z). \tag{8.3}$$

Clearly, if a transition density $\widetilde{p}_t(x,y)$ exists then its is also a heat kernel. However, unlike a heat kernel, which is defined *almost everywhere*, a transition density is defined *everywhere*.

**Lemma 8.1** *If the semigroup $P_t$ is locally ultracontractive then a transition density exists.*

**Proof.** Assume first that $P_t$ is ultracontractive with a rate function $\gamma(t)$. By Lemma 7.4, there exists a negligible set $N \subset M$ such that for all $x \in M \setminus N$, $t > 0$, and Borel functions $f \in L^2$,

$$|\mathcal{P}_t f(x)| \leq \widetilde{\gamma}(t) \|f\|_2,$$

where $\widetilde{\gamma}(t) = \gamma(t-0)$. By the Riesz representation theorem, for any $x \in M \setminus N$ and $t > 0$, there exists $p_{t,x} \in L^2$ such that

$$\mathcal{P}_t f(x) = (p_{t,x}, f),$$

for any Borel function $f \in L^2$. In fact, the function $p_t(x,y) = p_{t,x}(y)$ is a heat kernel. For all $x, y \in M \setminus N$, define a transition density by

$$\widetilde{p}_t(x,y) = (p_{s,x}, p_{t-s,y}), \tag{8.4}$$

where $0 < s < t$. It was shown in [30] that the right hand side in (8.4) does not depend on $s$ and that $\widetilde{p}_t(x,y)$ is indeed a transition density on $M \setminus N$. Extending $\widetilde{p}_t(x,y)$ to all $x, y \in M$ by setting it equal to 0 if $x \in N$ or $y \in N$, we conclude the proof.

Assume now that $P_t$ is locally ultracontractive, that is, $P_t^{\Omega_k}$ is ultracontractive for an exhaustive sequence $\{\Omega_k\}$ of open sets. By the above argument, a transition density $\widetilde{p}_t^{\Omega_k}$ exists for all $\Omega_k$, with a negligible set $N_k \subset \Omega_k$. Setting $N = \bigcup_k N_k$ and $\widetilde{p}_t = \lim_{k\to\infty} \widetilde{p}_t^{\Omega_k}$, we obtain a transition density on $M$. ■

**Corollary 8.2** *If a heat kernel exists and is in $L^\infty_{loc}$ then a transition density also exists and is in $L^\infty_{loc}$.*

**Proof.** Indeed, the existence of a transition density follows from Lemmas 5.2 and 8.1. The local boundedness follows from the fact the transition density is equal to the heat kernel $\mu$-a.e.. ■

**Lemma 8.3** *Assume that a transition density $\widetilde{p}_t(x,y)$ exists and is in $L^\infty_{loc}$. Then there exists a negligible set $N \subset M$ such that, for all non-empty open sets $U, V \subset M$ and $t > 0$,*

$$\sup_{\substack{x \in V \setminus N \\ y \in U \setminus N}} \widetilde{p}_t(x,y) = \operatorname*{esup}_{\substack{x \in V \\ y \in U}} \widetilde{p}_t(x,y). \tag{8.5}$$

**Proof.** Since $\widetilde{p}_t \geq 0$ $\mu$-a.e., the identity (8.1) extends to all non-negative Borel functions $f$. Set $\widetilde{p}_{t,x} := \widetilde{p}_t(x,\cdot)$ so that (8.1) takes the form

$$\mathcal{P}_t f(x) = \int_M \widetilde{p}_{t,x} f \, d\mu.$$

Changing here $t$ to $s$ and setting $f = \widetilde{p}_{t-s,y}$, where $y \in M$ and $0 < s < t$, we obtain, using also (8.2) and (8.3), that for all $x \in M \setminus N$,

$$\mathcal{P}_s \widetilde{p}_{t-s,y}(x) = \int_M \widetilde{p}_{s,x} \widetilde{p}_{t-s,y} d\mu = \widetilde{p}_t(x,y).$$

By Lemma 5.2, the semigroup $P_t$ is locally ultracontractive. Therefore, by 7.5, we have

$$\sup_{V \setminus N} \mathcal{P}_s f = \operatorname*{esup}_V \mathcal{P}_s f,$$

for any non-negative Borel function $f$ (we can assume that the negligible set $N$ from Lemma 7.5 is the same as the one from Definition 8; otherwise, just take the union of the two negligible sets). Hence, we conclude that, for any $y \in M$,

$$\sup_{x \in V \setminus N} \widetilde{p}_t(x,y) = \operatorname*{esup}_{x \in V} \widetilde{p}_t(x,y). \tag{8.6}$$

By symmetry, we have also, for any $x \in M$,

$$\sup_{y \in U \setminus N} \widetilde{p}_t(x,y) = \operatorname*{esup}_{y \in U} \widetilde{p}_t(x,y).$$

Taking in (8.6) supremum in $y$ we obtain

$$\begin{aligned}
\sup_{y \in U \setminus N} \sup_{x \in V \setminus N} \widetilde{p}_t(x,y) &= \sup_{y \in U \setminus N} \operatorname*{esup}_{x \in V} \widetilde{p}_t(x,y) \\
&= \sup_{y \in U \setminus N} \inf_{E, \mu(E)=0} \sup_{x \in V \setminus E} \widetilde{p}_t(x,y) \\
&\leq \inf_{E, \mu(E)=0} \sup_{y \in U \setminus N} \sup_{x \in V \setminus E} \widetilde{p}_t(x,y) \\
&= \inf_{E, \mu(E)=0} \sup_{x \in V \setminus E} \sup_{y \in U \setminus N} \widetilde{p}_t(x,y) \\
&= \operatorname*{esup}_{x \in V} \operatorname*{esup}_{y \in U} \widetilde{p}_t(x,y),
\end{aligned}$$

whence (8.5) follows (the opposite inequality is trivial because $\mu(N) = 0$). ■

**Corollary 8.4** *Assume that there exists a heat kernel $p_t(x,y)$ such that $p_t \in L^\infty_{loc}$ for all $t > 0$ and*

$$p_t(x,y) \leq F_t(x,y), \tag{8.7}$$

*for all $t > 0$ and for $\mu$-a.a. $x \in V$, $y \in U$ where $V, U$ are non-empty open subsets of $M$ and $F_t(x,y)$ is a continuous function on $V \times U$ for any $t > 0$. Then there exists a transition density $\widetilde{p}_t(x,y)$ and a negligible set $N \subset M$ such that*

$$\widetilde{p}_t(x,y) \leq F_t(x,y), \tag{8.8}$$

*for all $t > 0$ and $x \in V \setminus N$, $y \in U \setminus N$.*

**Proof.** By Corollary 8.2, a transition density $\widetilde{p}_t(x,y)$ also exists and, hence, satisfies (8.8) for all $t>0$ and for $\mu$-a.a. $x\in V$, $y\in U$. Let $N$ be a negligible set from Lemma 8.3, and let us prove that (8.8) holds for all $t>0$ and $x\in V\setminus N$, $y\in U\setminus N$. Fix $x_0\in V\setminus N$, $y_0\in U\setminus N$, and choose open sets $V_0\subset V$ and $U_0\subset U$ containing $x_0$ and $y_0$, respectively. Then, by Lemma 8.3,

$$\widetilde{p}_t(x_0,y_0)\leq \sup_{\substack{x\in V_0\setminus N\\ y\in U_0\setminus N}}\widetilde{p}_t(x,y)=\operatorname*{esup}_{\substack{x\in V_0\\ y\in U_0}}\widetilde{p}_t(x,y)\leq \operatorname*{esup}_{\substack{x\in V_0\\ y\in U_0}}F_t(x,y)=\sup_{\substack{x\in V_0\\ y\in U_0}}F_t(x,y).$$

Shrinking $V_0$ and $U_0$ to the points $x_0$ and $y_0$, respectively, and noticing that

$$\sup_{\substack{x\in V_0\\ y\in U_0}}F_t(x,y)\to F(x_0,y_0),$$

we conclude the proof. ■

# 9 The exit time

In this section, we assume that $(M,d)$ is a locally compact separable metric space, $\mu$ is a Radon measure on $M$ with full support, $(\mathcal{E},\mathcal{F})$ is a *local* regular Dirichlet form in $L^2(M,\mu)$, and $\left\{\{X_t\}_{t\geq0},\{\mathbb{P}_x\}_{x\in M}\right\}$ is the associated diffusion process. The main point of the next statement is to provide criteria for the estimate (9.3) of the exit time probability, which will be used later in Theorem 9.3.

**Theorem 9.1** *Assume that $\{X_t\}$ is a diffusion and $\{X_t\}$ is stochastically complete. Then, for any $\beta>1$ and for any negligible set $N\subset M$, the following are equivalent:*

*(i) There exists $0<\varepsilon<\frac{1}{2}$ and $\delta>0$ such that, for all $0<t\leq\delta r^\beta$ and $x\in M\setminus N$,*

$$\mathbb{P}_x\left(X_t\in B(x,r)^c\right)\leq\varepsilon.$$

*(ii) There exist $0<\varepsilon<1$ and $\delta>0$ such that, for all $0<t\leq\delta r^\beta$ and $x\in M\setminus N$,*

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

*(iii) There exist $\varepsilon>0$ such that, for all $r>0$ and $x\in M\setminus N$,*

$$\mathbb{E}_x\left(\tau_{B(x,r)}\wedge r^\beta\right)\geq\varepsilon r^\beta.$$

*(iv) There exist $0<\varepsilon<1$ and $\delta>0$ such that, for all $r>0$, $\lambda\geq\left(\delta r^\beta\right)^{-1}$ and $x\in M\setminus N$,*

$$\mathbb{E}_x\exp\left(-\lambda\tau_{B(x,r)}\right)\leq\varepsilon. \tag{9.1}$$

*(v) There exist $c,C>0$ such that, for all $r,\lambda>0$ and $x\in M\setminus N$,*

$$\mathbb{E}_x\exp\left(-\lambda\tau_{B(x,r)}\right)\leq C\exp\left(-c\lambda^{1/\beta}r\right). \tag{9.2}$$

*(vi) There exist $c,C>0$ such that, for all $r,t>0$ and $x\in M\setminus N$,*

$$\mathbb{P}_x\left(\tau_{B(x,r)}\leq t\right)\leq C\exp\left(-c\left(\frac{r^\beta}{t}\right)^{\frac{1}{\beta-1}}\right) \tag{9.3}$$

**Remark.** The hypothesis of stochastic completeness is only used in the proof of the implication $(i) \Longrightarrow (ii)$. As one can see from the proof, without stochastic completeness all the conditions $(ii) - (vi)$ are still equivalent and imply $(i)$.

**Proof.** $(i) \Longrightarrow (ii)$. Let us set

$$m_t(r) := \sup_{0<s\leq t} \sup_{z\in M\setminus N} \mathbb{P}_z\left(X_t \in B(z,r)^c\right).$$

Fix $x \in M \setminus N$ and set $A = B(x,r)$, $U = B(x,2r)$. By the stochastic completeness of $\{X_t\}$, we have

$$\mathbb{P}_x(X_t \in A) = 1 - \mathbb{P}_x(X_t \in A^c) \geq 1 - m_t(r), \tag{9.4}$$

whereas by Lemma 10.1

$$\mathbb{P}_x(X_t \in A) - \mathbb{P}_x^U(X_t \in A) \leq \sup_{0\leq s\leq t} \sup_{z\in\partial U\setminus N} \mathbb{P}_z(X_s \in A). \tag{9.5}$$

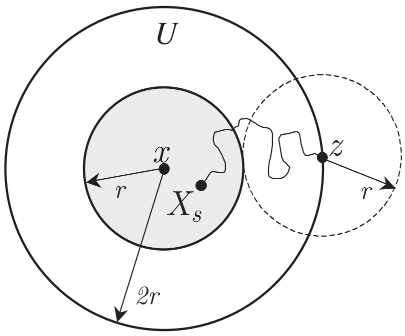


Figure 1: The ball $U = B(x,2r)$ and a point $z \in \partial U$.

For any $0 \leq s \leq t$ and $z \in \partial U \setminus N$ we have $d(z,A) \geq r$ and hence

$$\mathbb{P}_z(X_s \in A) \leq \mathbb{P}_z\left(X_s \in B(z,r)^c\right) \leq m_t(r)$$

(see Fig. 1). Subtracting (9.5) from (9.4), we obtain

$$\mathbb{P}_x^U(X_t \in A) \geq 1 - 2m_t(r),$$

whence it follows by (5.2) that

$$\mathbb{P}_x(\tau_U > t) = \mathbb{P}_x^U(X_t \in U) \geq 1 - 2m_t(r).$$

Therefore,

$$\mathbb{P}_x(\tau_U \leq t) \leq 2m_t(r), \tag{9.6}$$

which, by hypothesis, is bounded by $2\varepsilon$ provided $t \leq \delta r^\beta$. Noticing that $2\varepsilon < 1$ and renaming $2\varepsilon$ by $\varepsilon$, we finish the proof.

Note that the inequality (9.6) can also be extracted from [1, Lemma 3.9].

$(ii) \Longrightarrow (iii)$. Denoting $\tau = \tau_{B(x,r)}$, we have

$$\mathbb{E}_x\left(\tau \wedge r^\beta\right) \geq \mathbb{E}_x\left(\mathbf{1}_{\{\tau>\delta r^\beta\}}\tau \wedge r^\beta\right) \geq \mathbb{P}_x\left(\tau > \delta r^\beta\right)(\delta \wedge 1)\, r^\beta \geq (1-\varepsilon)(\delta \wedge 1)\, r^\beta,$$

whence the claim follows.

$(iii) \Longrightarrow (iv)$. For all non-negative reals $\lambda, \tau, t$, we have the elementary inequality

$$e^{-\lambda\tau} \leq 1 - \lambda e^{-\lambda t} (\tau \wedge t),$$

which is trivially verified by considering two cases $\tau \geq t$ and $\tau < t$. Therefore, for $\tau = \tau_{B(x,r)}$ and $t = r^\beta$, we obtain

$$\mathbb{E}_x e^{-\lambda\tau} \leq 1 - \lambda e^{-\lambda r^\beta} \mathbb{E}_x \left(\tau \wedge r^\beta\right) \leq 1 - \varepsilon r^\beta \lambda e^{-\lambda r^\beta}.$$

Setting $\lambda = r^{-\beta}$, we obtain

$$\mathbb{E}_x e^{-\lambda\tau} \leq 1 - \varepsilon e^{-1}.$$

The same inequality holds also for all $\lambda \geq r^{-\beta}$ and, hence, the condition $(iv)$ is satisfied with $\delta = 1$ (in fact, it is satisfied with *any* $\delta > 0$).

$(iv) \Longrightarrow (v)$. For a positive number $n$, set $\rho = r/n$, and consider the sequence of balls $B_k = B(x, k\rho)$, $k = 1, 2, ..., n$. Set $\tau_k = \tau_{B_k}$ and observe that, for any $x \in M \setminus N$,

$$\mathbb{E}_x e^{-\lambda\tau_{k+1}} = \mathbb{E}_x \left(e^{-\lambda\tau_k} e^{-\lambda(\tau_{k+1}-\tau_k)}\right) = \mathbb{E}_x \left(e^{-\lambda\tau_k} \mathbb{E}_{X_{\tau_k}} e^{-\lambda\tau_{k+1}}\right). \tag{9.7}$$

Since $X_{\tau_k} \in \partial B_k \setminus N$ with $\mathbb{P}_x$-probability 1 and, for any $y \in \partial B_k$, we have $\tau_{k+1} \geq \tau_{B(y,\rho)}$ and hence

$$\mathbb{E}_y e^{-\lambda\tau_{k+1}} \leq \mathbb{E}_y e^{-\lambda\tau_{B(y,\rho)}},$$

we obtain from (9.7), for any $x \in M \setminus N$,

$$\mathbb{E}_x e^{-\lambda\tau_{k+1}} \leq \mathbb{E}_x e^{-\lambda\tau_k} \sup_{y \in \partial B_k \setminus N} \mathbb{E}_y e^{-\lambda\tau_{B(y,\rho)}} \tag{9.8}$$

(see Fig. 2).

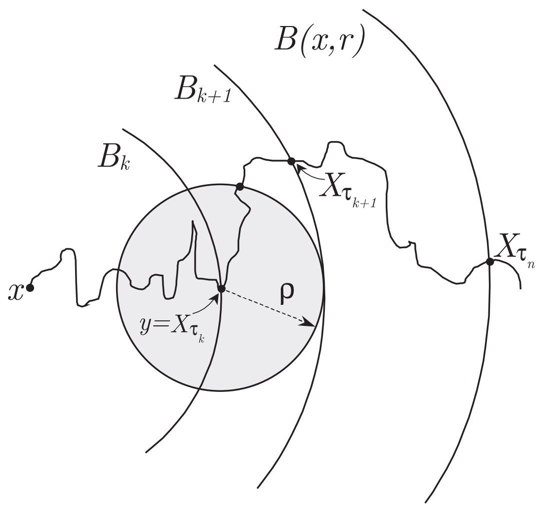


Figure 2: The stopping times $\tau_k$ and $\tau_{k+1}$

Now choose $n$ so that

$$\lambda \geq \frac{n^\beta}{\delta r^\beta} = \frac{1}{\delta \rho^\beta} \tag{9.9}$$

(if no positive integer $n$ satisfies (9.9) then (9.2) is trivially true for large enough $C$). Applying the hypothesis (9.1) in the ball $B(y, \rho)$, we obtain, for any $y \in M \setminus N$,

$$\mathbb{E}_y e^{-\lambda\tau_{B(y,\rho)}} \leq \varepsilon$$

whence by (9.8)

$$\mathbb{E}_x e^{-\lambda\tau_{k+1}} \leq \varepsilon \mathbb{E}_x e^{-\lambda\tau_k}.$$

Iterating this inequality and noticing that $\tau_{B(x,r)} = \tau_n$ we conclude that for $n$ satisfying (9.9),

$$\mathbb{E}_x e^{-\lambda\tau_{B(x,r)}} \leq \varepsilon^n = e^{-an},$$

where $a = \log \frac{1}{\varepsilon} > 0$. Taking the largest $n$ satisfying the inequality (9.9) and noticing that, for this $n$,

$$n \geq (\delta\lambda)^{1/\beta} r - 1,$$

we obtain

$$\mathbb{E}_x e^{-\lambda\tau_{B(x,r)}} \leq \exp\left(-a\delta^{1/\beta}\lambda^{1/\beta} r + a\right),$$

which was to be proved.

$(v) \Longrightarrow (vi)$. We have, for $\tau = \tau_{B(x,r)}$ and $\lambda > 0$,

$$\mathbb{P}_x (\tau \leq t) = \mathbb{P}_x \left(e^{-\lambda\tau} \geq e^{-\lambda t}\right) \leq e^{\lambda t}\mathbb{E}_x e^{-\lambda\tau} \leq C \exp\left(\lambda t - c\lambda^{1/\beta} r\right).$$

Choosing $\lambda = \left(\frac{cr}{2t}\right)^{\frac{\beta}{\beta-1}}$, we obtain the claim.

$(vi) \Longrightarrow (ii)$. We have

$$\mathbb{P}_x \left(\tau_{B(x,r)} \leq t\right) \leq C \exp\left(-c\left(\frac{r^\beta}{t}\right)^{\frac{1}{\beta-1}}\right).$$

If $t \leq \delta r^\beta$ and $\delta$ is small enough then the right hand side can be made arbitrarily small, whence the claim follows.

$(ii) \Longrightarrow (i)$ This is trivial because

$$\mathbb{P}_x (X_t \in B(x,r)^c) \leq \mathbb{P}_x \left(\tau_{B(x,r)} \leq t\right).$$

■

**Lemma 9.2** *If $f$ is a bounded non-negative Borel function on $M$ and*

$$u := \int_0^\infty \mathcal{P}_t f \, dt < \infty$$

*then, for any $\lambda \geq 0$,*

$$u = \int_0^\infty e^{-\lambda t}\mathcal{P}_t (f + \lambda u) \, dt. \tag{9.10}$$

**Remark.** If $M$ is a precompact open set on a Riemannian manifold and $\mathcal{P}_t$ is the transition function associated with the Dirichlet Laplacian on $M$ then the function $u$ solves the equation $\Delta u = -f$ with the vanishing Dirichlet boundary condition. Subtracting $\lambda u$ from the both sides, we obtain

$$\Delta u - \lambda u = -(f + \lambda u),$$

whence (9.10) follows. The proof below is an abstract version of this argument.

**Proof.** If $\lambda = 0$ then (9.10) is trivial, so we can assume $\lambda > 0$. For any $s > 0$, we have

$$u = \int_0^s \mathcal{P}_t f \, dt + \int_s^\infty \mathcal{P}_s \mathcal{P}_{t-s} f \, dt = \int_0^s \mathcal{P}_t f \, dt + \mathcal{P}_s \int_s^\infty \mathcal{P}_{t-s} f \, dt = \int_0^s \mathcal{P}_t f \, dt + \mathcal{P}_s u.$$

Multiplying this identity by $\lambda e^{-\lambda s}$ and integrating in $s$ from 0 to $\infty$, we obtain

$$\begin{aligned} u &= \int_0^\infty \lambda e^{-\lambda s}\left(\int_0^s \mathcal{P}_t f\,dt\right)ds + \int_0^\infty \lambda e^{-\lambda s}\mathcal{P}_s u\,ds \\ &= \int_0^\infty \left(\int_t^\infty \lambda e^{-\lambda s} ds\right)\mathcal{P}_t f\,dt + \int_0^\infty \lambda e^{-\lambda s}\mathcal{P}_s u\,ds \\ &= \int_0^\infty e^{-\lambda t}\mathcal{P}_t f\,dt + \int_0^\infty e^{-\lambda s}\mathcal{P}_s(\lambda u)\,ds, \end{aligned}$$

whence (9.10) follows. ∎

**Theorem 9.3** *If $\{X_t\}$ is a diffusion process and $(E_\beta)$ holds then there exist positive constant $c$ and $C$ (depending on the constants in the condition $(E_\beta)$) such that*

$$\mathbb{P}_x\left(\tau_{B(x,r)} \le t\right) \le C\exp\left(-c\left(\frac{r^\beta}{t}\right)^{\frac{1}{\beta-1}}\right), \tag{9.11}$$

*for all $r, t > 0$ and for all $x \in M\setminus N$, where $N$ is the negligible set from the condition $(E_\beta)$.*

**Remark.** The fact that $(E_\beta)$ implies (9.11) is basically due to M.Barlow and can be extracted from [1, Theorem 3.11]. Here we give a new proof of that. See also [17] for an alternative proof of Theorem 9.3, and [15, Proposition 7.1], [16, Lemma 6.3] for similar estimates for random walks on graphs.

**Proof.** Fix a point $x \in M\setminus N$ and $r > 0$, and set $U = B(x,r)$. Consider a function $u$ in $U$ defined by

$$u := \mathbb{E}_\cdot\tau_U = \int_0^\infty \mathcal{P}_t^U 1 dt.$$

The upper bound in the condition $(E_\beta)$ implies

$$\sup_{U\setminus N} u \le C' r^\beta. \tag{9.12}$$

Indeed, for any $y \in U\setminus N$ we have $U \subset B(y,2r)$ and, hence,

$$\mathbb{E}_y\tau_U \le \mathbb{E}_y\tau_{B(y,2r)} \le C(2r)^\beta.$$

Set

$$\lambda = \frac{1}{\sup_{U\setminus N} u}, \tag{9.13}$$

and notice that $\lambda u \le 1$ in $U\setminus N$; since the set $N$ is negligible, this implies that everywhere

$$\mathcal{P}_t^U(\lambda u) \le \mathcal{P}_t^U 1.$$

By Lemma 9.2 we obtain

$$u = \int_0^\infty e^{-\lambda t}\mathcal{P}_t^U(1+\lambda u)\,dt \le 2\int_0^\infty e^{-\lambda t}\mathcal{P}_t^U 1\,dt.$$

On the other hand, for any $\lambda > 0$, we have the identity

$$\mathbb{E}_y e^{-\lambda\tau_U} = 1 - \lambda\int_0^\infty e^{-\lambda t}\mathcal{P}_t^U 1(y)\,dt, \tag{9.14}$$

which follows from (6.1) with $\varphi(t) = \lambda e^{-\lambda t}$. Comparing the above two lines, we obtain, for $\lambda$ as in (9.13) and for all $y \in M$,

$$\mathbb{E}_y e^{-\lambda\tau_U} \leq 1 - \frac{\lambda}{2} u(y). \tag{9.15}$$

Using the lower bound in $(E_\beta)$ and (9.12), we obtain

$$u(x) \geq cr^\beta \geq c' \sup_{U\setminus N} u = \frac{c'}{\lambda},$$

which together with (9.15) yields

$$\mathbb{E}_x e^{-\lambda\tau_U} \leq 1 - \frac{c'}{2}.$$

Clearly, the last inequality remains true also for $\lambda \geq \frac{1}{\sup_{U\setminus N} u}$ and, in particular, for $\lambda \geq \frac{1}{cr^\beta}$. The proof is concluded by application of the part $(iv) \Rightarrow (vi)$ of Theorem 9.1. ■

# 10 Estimating heat kernel using exit probabilities

In this section, we assume that the form $(\mathcal{E}, \mathcal{F})$ is a local, and hence $\left\{\{X_t\}_{t\geq 0}, \{\mathbb{P}_x\}_{x\in M}\right\}$ is a diffusion process. As before, we denote by $\tau_U$ the first exit time from a set $U$ (see (1.3)). For any open set $U \subset M$, define the *exit probability function* $\psi^U(x,t)$ as follows:

$$\psi^U(x,t) = \mathbb{P}_x(\tau_U \leq t).$$

In words, $\psi^U(x,t)$ is the $\mathbb{P}_x$-probability that $X_t$ exits $U$ before time $t$. Our main result in this section is Theorem 10.4 that provides a certain upper bound of the heat kernel of $X_t$ using the exit probabilities.

The next lemma will not be directly used in the proof of Theorem 10.4, but it introduces the argument in a simpler setting. Besides, a weaker version of this lemma will be used in the proof of Theorem 1.1. Recall that $\mathcal{P}_t$ is the transition function defined in Section 5.

**Lemma 10.1** *Let $N \subset M$ be a negligible set for the diffusion $\{X_t\}$. Let $U$ be a non-empty open subset of $M$, and $f$ be a non-negative Borel function on $M$ such that $f \equiv 0$ in $M \setminus U$. Then, for all $x \in U \setminus N$ and all $t > 0$,*

$$\mathcal{P}_t f(x) \leq \mathcal{P}_t^U f(x) + \psi^U(x,t) \sup_{0<s\leq t} \sup_{z\in\partial U\setminus N} \mathcal{P}_s f(z). \tag{10.1}$$

**Remark.** One can always take here $N = \emptyset$, which is a perfectly good choice in the case when the function $z \mapsto \mathcal{P}_s f(z)$ is continuous. Otherwise, the supremum of $\mathcal{P}_s f(z)$ over the full boundary $\partial U$ may be not under control, so one has to reduce it by removing some singularities of this function.

We will use (10.1) for $f = \mathbf{1}_B$, where $B$ is a Borel subset of $U$, in which case we obtain

$$\mathbb{P}_x(X_t \in B) \leq \mathbb{P}_x^U(X_t \in B) + \psi^U(x,t) \sup_{0<s\leq t} \sup_{z\in\partial U\setminus N} \mathbb{P}_z(X_s \in B). \tag{10.2}$$

**Proof.** Without loss of generality, we can assume that the function $f$ is bounded and its support is compact (otherwise, approximate $f$ by an increasing sequence of bounded functions with compact supports). Writing $\tau = \tau_U$, we obtain by the strong Markov property

$$\begin{aligned}
\mathcal{P}_t f(x) &= \mathbb{E}_x f(X_t) \\
&= \mathbb{E}_x\left(\mathbf{1}_{\{\tau>t\}} f(X_t)\right) + \mathbb{E}_x\left(\mathbf{1}_{\{\tau\leq t\}} f(X_t)\right) \\
&= \mathbb{E}_x^U f(X_t) + \mathbb{E}_x\left(\mathbf{1}_{\{\tau\leq t\}} \mathbb{E}_{X_\tau} f(X_{t-\tau})\right) \\
&= \mathcal{P}_t^U f(x) + \mathbb{E}_x\left(\mathbf{1}_{\{\tau\leq t\}} \mathcal{P}_{t-\tau} f(X_\tau)\right). \qquad (10.3)
\end{aligned}$$

For any $x \in U$, we have $X_\tau \in \partial U$ with $\mathbb{P}_x$-probability 1. If $x \in U \setminus N$ then $\mathbb{P}_x (X_\tau \in N) = 0$ and hence $X_\tau \in \partial U \setminus N$ with $\mathbb{P}_x$-probability 1. Therefore, we obtain, for any $x \in U \setminus N$,

$$\mathcal{P}_{t-\tau} f (X_\tau) \leq \sup_{0 \leq s \leq t} \sup_{z \in \partial U \setminus N} \mathcal{P}_s f (z) \qquad \mathbb{P}_x\text{-a.s.},$$

whence

$$\mathbb{E}_x \left( \mathbf{1}_{\{\tau \leq t\}} \mathcal{P}_{t-\tau} f (X_\tau) \right) \leq \mathbb{P}_x (\tau \leq t) \sup_{0 \leq s \leq t} \sup_{z \in \partial U \setminus N} \mathcal{P}_s f (z) . \tag{10.4}$$

In fact, the range $0 \leq s \leq t$ for the variable $s$ can be replaced by $0 < s \leq t$ for the following reason: if $\tau = t$ (which corresponds to $s = 0$) then $\mathcal{P}_{t-\tau} f (X_\tau) = f (X_\tau) = 0$ because $f = 0$ on $\partial U$. Hence, (10.1) follows from (10.3) and (10.4) with the above improvement. ■

Let $\mathcal{A}_t$ be the $\sigma$-algebra of events in the space of all continuous paths $\omega : [0, +\infty) \to M$, which is generated by all events $\{\omega : \omega (s) \in B\}$ where $0 \leq s \leq t$ and $B$ is a Borel set in $M$. For any $\mathcal{A}_t$-measurable random variable $\xi$, define a random variable $\xi^t$ by

$$\xi^t (\omega) = \xi \left( \omega^t \right) ,$$

where $\omega^t := \omega (t - \cdot)$ . Although the path $\omega^t$ is defined only on $[0, t]$, the value $\xi \left( \omega^t \right)$ still makes sense because $\xi$ is $\mathcal{A}_t$-measurable.

**Lemma 10.2** *If $\xi$ is a non-negative $\mathcal{A}_t$-measurable random variable then, for all non-negative Borel functions $f, g$ on $M$,*

$$\int_M \mathbb{E}_x (\xi f (X_t)) \, g (x) \, d\mu (x) = \int_M \mathbb{E}_y \left( \xi^t g (X_t) \right) f (y) \, d\mu (y) . \tag{10.5}$$

**Proof.** It suffices to verify (10.5) for cylindrical random variables $\xi$, that is, for $\xi$ in the form

$$\xi = h_1 (X_{t_1}) \, h_2 (X_{t_2}) \ldots h_n (X_{t_n})$$

where $0 < t_1 < t_2 < \ldots < t_n < t$ and $h_k$ are non-negative Borel functions. For such $\xi$, (10.5) amounts to the identity

$$\int_M \mathbb{E}_x \left( f(X_t) h_n(X_{t_n}) \ldots h_1(X_{t_1}) g(X_0) \right) d\mu(x) = \int_M \mathbb{E}_y \left( g(X_t) h_1(X_{t-t_1}) \ldots h_n(X_{t-t_n}) f(X_0) \right) d\mu(y),$$

which was proved in [9, Lemma 4.1.2, p.135]. ■

The next statement is crucial for Theorem 10.4, and is a refinement of Lemma 10.1.

**Lemma 10.3** *Let $N \subset M$ be a negligible set for the diffusion $\{X_t\}$, and let $U$ and $V$ be two non-empty open subsets of $M$ such that either $U \subset V$ or $U \cap V = \emptyset$. Let $f, g$ be non-negative Borel functions on $M$, and let $f|_{M \setminus U} \equiv 0$. Then, for all $a, b, t > 0$ such that $a + b = t$, we have*

$$\int_V (\mathcal{P}_t f) \, g \, d\mu \leq \int_V (\mathcal{P}_t^V f) g \, d\mu \tag{10.6}$$

$$+ \sup_{b \leq s \leq t} \sup_{v \in \partial V \setminus N} \mathcal{P}_s f (v) \int_V \psi^V (x, a) \, g (x) \, d\mu (x) \tag{10.7}$$

$$+ \sup_{a \leq s \leq t} \sup_{u \in \partial U \setminus N} \mathcal{P}_s g (u) \int_U \psi^U (y, b) \, f (y) \, d\mu (y) . \tag{10.8}$$

**Remark.** If in addition $g|_{M\setminus V}=0$ then integration in all terms in (10.6)-(10.8) can be extended to entire space $M$. In this case, the left hand side of (10.6) is symmetric in $f,g$ (while the right hand side of (10.6) is always symmetric in $f,g$). Observe also, that if $U$ and $V$ are disjoint then the term $\int_V \left(\mathcal{P}_t^V f\right) g\, d\mu$ in (10.6) vanishes because $f=0$ in $V$ and hence $\mathcal{P}_t^V f=0$.

**Proof.** Without loss of generality, we can assume that both functions $f$ and $g$ are bounded and with compact support, so that all the terms in (10.6)-(10.8) are finite (in general, approximate each of the functions $f$ and $g$ by an increasing sequence of bounded functions with compact supports). Using $\mathcal{P}_t f(x)=\mathbb{E}_x f(X_t)$ and

$$\mathcal{P}_t^V f(x)=\mathbb{E}_x\left(\mathbf{1}_{\{\tau_V>t\}} f(X_t)\right),$$

we obtain

$$\mathcal{P}_t f(x)\le \mathcal{P}_t^V f(x)+\mathbb{E}_x\left(\mathbf{1}_{\{\tau_V\le a\}} f(X_t)\right)+\mathbb{E}_x\left(\mathbf{1}_{\{a\le\tau_V\le t\}} f(X_t)\right). \tag{10.9}$$

Assuming $x\in V\setminus N$, let us estimate the middle term in (10.9) similarly to (10.4), that is

$$\begin{aligned}\mathbb{E}_x\left(\mathbf{1}_{\{\tau_V\le a\}} f(X_t)\right) &= \mathbb{E}_x\left(\mathbf{1}_{\{\tau_V\le a\}}\mathbb{E}_{X_{\tau_V}} f\left(X_{t-\tau_V}\right)\right)\\ &\le \mathbb{P}_x(\tau_V\le a)\sup_{b\le s\le t}\sup_{v\in\partial V\setminus N}\mathcal{P}_s f(v),\end{aligned}$$

which implies

$$\int_V \mathbb{E}_x\left(\mathbf{1}_{\{\tau_V\le a\}} f(X_t)\right) g(x)\,d\mu(x)\le \sup_{b\le s\le t}\sup_{v\in\partial V\setminus N}\mathcal{P}_s f(v)\int_V \mathbb{P}_x(\tau_V\le a)\, g(x)d\mu(x). \tag{10.10}$$

To estimate the last term in (10.9), consider the random variable $\xi=\mathbf{1}_{\{a\le\tau_V\le t\}}$. For any continuous path $\omega$ that intersects $\partial V$ between times 0 and $t$, we have

$$\tau_U(\omega)\le t-\tau_V\left(\omega^t\right),$$

because by hypothesis $\partial V$ is outside $U$ (see Fig. 3 and 4).

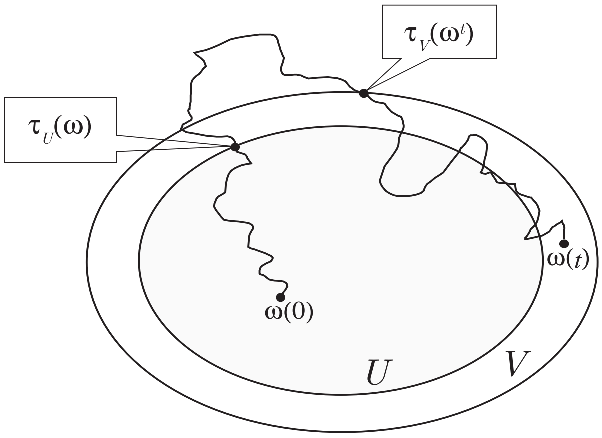


Figure 3: Case $U\subset V$.

Therefore, $a\le\tau_V(\omega^t)\le t$ implies $\tau_U(\omega)\le t-a=b$, which means that $\xi\left(\omega^t\right)\le\mathbf{1}_{\{\tau_U\le b\}}(\omega)$ and hence $\xi^t\le\mathbf{1}_{\{\tau_U\le b\}}$. Using (10.5) and $f|_{M\setminus U}=0$, we obtain

$$\begin{aligned}\int_M \mathbb{E}_x\left(\mathbf{1}_{\{a\le\tau_V\le t\}} f(X_t)\right) g(x)\,d\mu(x) &= \int_M \mathbb{E}_y\left(\xi^t g(X_t)\right) f(y)\,d\mu(y)\\ &\le \int_U \mathbb{E}_y\left(\mathbf{1}_{\{\tau_U\le b\}} g(X_t)\right) f(y)\,d\mu(y).\end{aligned}$$

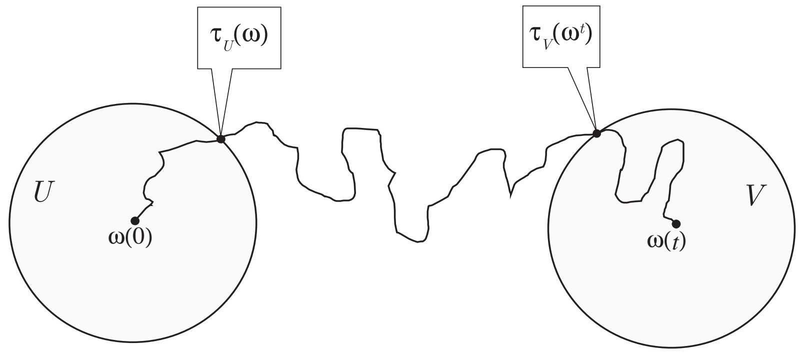


Figure 4: Case of disjoint $U$ and $V$.

Similarly to (10.10), we obtain, for any $y \in U \setminus N$,

$$\mathbb{E}_y \left( \mathbf{1}_{\{\tau_U \leq b\}} g\left(X_t\right) \right) \leq \mathbb{P}_y \left( \tau_U \leq b \right) \sup_{a \leq s \leq t} \sup_{u \in \partial U \setminus N} \mathcal{P}_s g\left(u\right),$$

whence it follows that

$$\int_V \mathbb{E}_x \left( \mathbf{1}_{\{a \leq \tau_V \leq t\}} f\left(X_t\right) \right) g(x) d\mu(x) \leq \sup_{a \leq s \leq t} \sup_{u \in \partial U \setminus N} \mathcal{P}_s g(u) \int_U \mathbb{P}_y \left( \tau_U \leq b \right) f(y) d\mu(y). \quad (10.11)$$

Finally, integrating (10.9) over $V$ against measure $g\left(x\right) d\mu(x)$ and using (10.10) and (10.11), we finish the proof. ■

**Theorem 10.4** *Assume that the diffusion $\{X_t\}$ has a heat kernel $p_t\left(x, y\right)$ which, for any $t > 0$, is a $L^{\infty}_{loc}$-function on $M \times M$. Let $U$ and $V$ be arbitrary non-empty open subsets of $M$ such that either $U \subset V$ or $U \cap V = \emptyset$, and $U'$ and $V'$ be open sets containing $\partial U$ and $\partial V$, respectively. Then, for* $\mu$-a.a. *$x \in V$, $y \in U$ and for all $t > 0$, we have*

$$p_t\left(x, y\right) \leq p_t^V\left(x, y\right) + \psi^V\left(x, \tfrac{t}{2}\right) \sup_{t/2 \leq s \leq t} \operatorname*{esup}_{v \in V'} p_s\left(v, y\right) + \psi^U\left(y, \tfrac{t}{2}\right) \sup_{t/2 \leq s \leq t} \operatorname*{esup}_{u \in U'} p_s\left(u, x\right). \quad (10.12)$$

**Remark.** If $V$ and $U$ are disjoint then $p_t^V\left(x, y\right) = 0$ so that we have

$$p_t\left(x, y\right) \leq \psi^V\left(x, \tfrac{t}{2}\right) \sup_{t/2 \leq s \leq t} \operatorname*{esup}_{v \in V'} p_s\left(v, y\right) + \psi^U\left(y, \tfrac{t}{2}\right) \sup_{t/2 \leq s \leq t} \operatorname*{esup}_{u \in U'} p_s\left(u, x\right). \quad (10.13)$$

Another important particular case, which will be used in the proof of Theorem 1.1, is when $U = V$.

**Remark.** If the heat kernel is continuous then (10.12) holds for *all* $x \in V$, $y \in U$, and the sets $V'$, $U'$ can be replaced by $\partial V$ and $\partial U$, respectively.

The estimates (10.12) and (10.13) were proved in [14] for diffusions on Riemannian manifolds, where the heat kernel is a smooth function. A particular case of (10.12) for $x = y$ was proved in [22, Lemma 4.5] in an abstract setting but still assuming that the heat kernel is continuous. Without continuity of the heat kernel, a major difficulty is to ensure the use of the essential supremum of the heat kernel rather than the supremum. The hypothesis that the heat kernel is locally bounded seems to be technical, but it is satisfied in all cases of interest when one may hope to use (10.12). Indeed, without local boundedness of the heat kernel, the terms in (10.12) containing the essential supremum of $p_s$, may be equal to $+\infty$.

**Proof.** By Lemma 5.2, the semigroup $P_t$ is locally ultracontractive. By Lemma 7.5, there exists a negligible set $N \subset M$ such that

$$\sup_{W \setminus N} \mathcal{P}_s h = \operatorname*{esup}_W \mathcal{P}_s h$$

for any non-negative Borel function $h$, any non-empty open set $W \subset M$, and for any $s > 0$.

Assuming that $f$ and $g$ are non-negative Borel functions supported in $U$ and $V$, respectively, we have

$$\sup_{\partial V \setminus N} \mathcal{P}_s f \leq \sup_{V' \setminus N} \mathcal{P}_s f = \operatorname*{esup}_{V'} \mathcal{P}_s f,$$

and, similarly,

$$\sup_{\partial U \setminus N} \mathcal{P}_s g \leq \operatorname*{esup}_{U'} \mathcal{P}_s g.$$

Substituting these inequalities into the estimate of Lemma 10.3, we obtain

$$\begin{aligned} \int_V (\mathcal{P}_t f)\, g \, d\mu \quad \leq \quad & \int_V \left(\mathcal{P}_t^V f\right) g \, d\mu \\ & + \sup_{b \leq s \leq t} \operatorname*{esup}_{v \in V'} \mathcal{P}_s f(v) \int_V \psi^V(x, a)\, g(x)\, d\mu(x) \\ & + \sup_{a \leq s \leq t} \operatorname*{esup}_{u \in U'} \mathcal{P}_s g(u) \int_U \psi^U(y, b)\, f(y)\, d\mu(y)\,. \end{aligned} \tag{10.14}$$

Since

$$\mathcal{P}_s f(v) = \int_U p_s(v, y)\, f(y)\, d\mu(y) \quad \mu\text{-a.a. } v \in M,$$

we obtain

$$\operatorname*{esup}_{v \in V'} \mathcal{P}_s f(v) \leq \int_U \operatorname*{esup}_{v \in V'} p_s(v, y)\ f(y)\, d\mu(y)\,.$$

Estimating similarly $\mathcal{P}_s g(u)$, we obtain from (10.14)

$$\begin{aligned} \iint\limits_{V \times U} p_t(x,y) g(x) f(y) d\mu(x) d\mu(y) \quad \leq \quad & \iint\limits_{V \times U} p_t^V(x, y)\, g(x) f(y) d\mu(x) d\mu(y) \\ & + \iint\limits_{V \times U} \sup_{b \leq s \leq t} \operatorname*{esup}_{v \in V'} p_s(v, y)\ \psi^V(x, a)\, g(x) f(y) d\mu(x) d\mu(y) \\ & + \iint\limits_{V \times U} \sup_{a \leq s \leq t} \operatorname*{esup}_{u \in U'} p_s(u, x)\ \psi^U(y, b)\, g(x) f(y) d\mu(x) d\mu(y). \end{aligned}$$

Setting $a = b = t/2$ and noticing that the functions of the form $g(x) f(y)$ span all $L^1(V \times U, \mu \times \mu)$, we finish the proof. ■

**Example 10.5** Here is a typical example of application of Theorem 10.4. Assume for simplicity that the heat kernel $p_t(x, y)$ is continuous in $x, y$ for any $t > 0$, and that we are given the following two conditions:

- For some $\nu > 0$ and all $t > 0$,

$$\sup_{x, y \in M} p_t(x, y) \leq C t^{-\nu}. \tag{10.15}$$

- For any $\varepsilon > 0$ there exists $\delta > 0$ such that for all $x \in M$ and $0 < t \leq \delta r^\beta$,

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

Then by Theorem 9.1$(ii) \Longrightarrow (vi)$ we have, for all positive $t, r$, that

$$\psi^{B(x,r)}(x,t) \leq C \exp\left(-c\left(\frac{r^\beta}{t}\right)^{\frac{1}{\beta-1}}\right). \tag{10.16}$$

Fix two distinct points $x, y \in M$, set $r = \frac{1}{2}d(x,y)$ and observe that, by Theorem 10.4

$$p_t(x,y) \leq \psi^{B(x,r)}\left(x, \tfrac{t}{2}\right) \sup_{\substack{t/2 \leq s \leq t \\ v \in \partial B(x,r)}} p_s(v,y) + \psi^{B(y,r)}\left(y, \tfrac{t}{2}\right) \sup_{\substack{t/2 \leq s \leq t \\ u \in \partial B(y,r)}} p_s(u,x). \tag{10.17}$$

Substituting (10.15) and (10.16) into (10.17), we obtain

$$p_t(x,y) \leq Ct^{-\nu} \exp\left(-c\left(\frac{r^\beta}{t}\right)^{\frac{1}{\beta-1}}\right).$$

# 11 Volume doubling

In this section, we assume that $(M,d)$ is a metric space and $\mu$ is a Borel measure on $M$. The following lemmas are well-known in the setting of complete manifolds (see for example [18], [10], [26]).

**Lemma 11.1** *If $(VD)$ holds on $M$ then there exists a positive constant $\alpha$ such that for all $x, y \in M$ and $0 < r \leq R$*

$$\frac{V(x,R)}{V(y,r)} \leq C\left(\frac{R + d(x,y)}{r}\right)^\alpha. \tag{11.1}$$

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

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

whence

$$\frac{V(x,R)}{V(x,r)} \leq C^n = 2^{n \log_2 C} \leq C\left(\frac{R}{r}\right)^{\log_2 C}. \tag{11.2}$$

If $x \neq y$ then $B(x,R) \subset B(y, R+d)$ where $d = d(x,y)$, and by

$$\frac{V(x,R)}{V(y,r)} \leq \frac{V(y,R+d)}{V(y,r)} \leq C\left(\frac{R+d}{r}\right)^{\log_2 C}.$$

■

**Lemma 11.2** *If $(M,d)$ is connected and satisfies $(VD)$ then there exist positive constants $\alpha', c$ such that for all $x \in M$ and $0 < r \leq R$*

$$\frac{V(x,R)}{V(x,r)} \geq c\left(\frac{R}{r}\right)^{\alpha'}, \tag{11.3}$$

*provided $B(x,R)^c$ is non-empty.*

**Remark.** If in addition $\operatorname{diam} M = \infty$ then $B(x,R)^c$ is always non-empty and, hence, (11.3) holds for all $x \in M$ and $0 < r \leq R$.

**Proof.** The condition $B(x,R)^c \neq \emptyset$ implies that $B(x,\rho') \setminus B(x,\rho) \neq \emptyset$ for all $0 < \rho < R$ and $\rho' > \rho$. Indeed, if $B(x,\rho') \setminus B(x,\rho) = \emptyset$ then $M$ 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 assumption that $B(x,R)^c$ is non-empty.

If $0 < \rho \leq R/2$ then by the above the annulus $B\left(x, \frac{5}{3}\rho\right) \setminus B\left(x, \frac{4}{3}\rho\right)$ is non-empty. Let $y$ be a point in this annulus. Then by (11.1) $V(x,\rho) \leq CV(y,\rho/3)$ whence

$$V(x,2\rho) \geq V(x,\rho) + V(y,\rho/3) \geq (1+\varepsilon)\, V(x,\rho), \tag{11.4}$$

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

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

$$n = \lfloor \log_2 \frac{R}{r} \rfloor \geq \log_2 \frac{R}{r} - 1.$$

For any $0 \leq k \leq n-1$ we have $2^k r \leq R/2$ whence by (11.4)

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

Iterating this inequality, we obtain

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

which was to be proved. ■

**Corollary 11.3** *If $(M,d)$ is connected and satisfies $(VD)$ then $\mu(M) = \infty$ if and only if* $\operatorname{diam}(M) = \infty$.

**Proof.** If $\operatorname{diam}(M) < \infty$ then $M$ is a ball of a finite radius, and $\mu(M) < \infty$ by $(VD)$. If $\operatorname{diam}(M) = \infty$ then $B^c(x,R)$ is non-empty for any ball $B(x,R)$. In this case, (11.3) implies that $V(x,R) \to \infty$ as $R \to \infty$, that is $\mu(M) = \infty$. ■

# 12 The main result

Here we state and prove our main result, which is more general than Theorem 1.1 from Introduction. In addition to conditions introduced in Section 2, consider one more condition as follows.

$(\overline{E}_\beta)$ : There are positive constants $C$ and $\nu$ such that, for any ball $B$ in $M$ of radius $r$ and for any non-empty open set $\Omega \subset B$,

$$\overline{E}(\Omega) := \operatorname*{esup}_{x \in \Omega} \mathbb{E}_x \tau_\Omega \leq C r^\beta \left(\frac{\mu(\Omega)}{\mu(B)}\right)^\nu. \tag{12.1}$$

## 12.1 Statement and the flowchart of the proof

**Theorem 12.1** *Let $(M,d)$ is a locally compact separable metric space, $\mu$ be a Radon measure on $M$ with full support, and $(\mathcal{E},\mathcal{F})$ be a regular Dirichlet form in $L^2(M,\mu)$. Assume in addition that:*

(*a*) $(M,d)$ *is connected and* $\operatorname{diam}(M)=\infty$.

(*b*) *Measure* $\mu$ *satisfies the volume doubling property* $(VD)$.

(*c*) *The form* $(\mathcal{E},\mathcal{F})$ *is local.*

(*d*) *The process* $\{X_t\}_{t\geq 0}$ *is stochastically complete.*

*Then, for any* $\beta>1$, *we have the following equivalences:*

$$\begin{aligned}(UE_\beta) &\Leftrightarrow (\Phi UE_\beta)\\ &\Leftrightarrow (DUE_\beta)+(P_\beta)\Leftrightarrow \left(\overline{E}_\beta\right)+(P_\beta)\Leftrightarrow (FK_\beta)+(P_\beta)\\ &\Leftrightarrow (DUE_\beta)+(E_\beta)\Leftrightarrow \left(\overline{E}_\beta\right)+(E_\beta)\Leftrightarrow (FK_\beta)+(E_\beta)\,.\end{aligned}$$

The proof is covered by the chains of implications as on the following diagram:

$$\boxed{\begin{array}{c}(DUE_\beta)\\(E_\beta)\end{array}} \overset{\text{T.9.3}}{\Longrightarrow} \boxed{\begin{array}{c}(DUE_\beta)\\(P_\beta)\end{array}} \overset{?}{\Longrightarrow} (UE_\beta) \overset{\text{trivial}}{\Longrightarrow} (\Phi UE_\beta) \overset{?}{\Longrightarrow}$$

$$\boxed{\begin{array}{c}(DUE_\beta)\\(P_\beta)\end{array}} \overset{?}{\Longrightarrow} \boxed{\begin{array}{c}\left(\overline{E}_\beta\right)\\(P_\beta)\end{array}} \overset{\text{T.9.1}}{\Longrightarrow} \boxed{\begin{array}{c}\left(\overline{E}_\beta\right)\\(E_\beta)\end{array}} \overset{\text{L.6.1}}{\Longrightarrow} \boxed{\begin{array}{c}(FK_\beta)\\(E_\beta)\end{array}} \overset{\text{T.9.3}}{\Longrightarrow}$$

$$\boxed{\begin{array}{c}(FK_\beta)\\(P_\beta)\end{array}} \overset{?}{\Longrightarrow} \boxed{\begin{array}{c}(DUE_\beta)\\(P_\beta)\end{array}} \overset{\text{from previous line}}{\Longrightarrow} \boxed{\begin{array}{c}(DUE_\beta)\\(E_\beta)\end{array}}\,.$$

The implications marked by '?' are yet to be proved. The other implications follow from already known results as it is indicated on the diagram. Indeed, Theorem 9.3 obviously yields $(E_\beta)\Longrightarrow(P_\beta)$, which is used twice on the diagram, and Lemma 6.1 yields $\left(\overline{E}_\beta\right)\Longrightarrow(FK)$. Let us explain the implication

$$\left(\overline{E}_\beta\right)+(P_\beta)\overset{\text{T.9.1}}{\Longrightarrow}\left(\overline{E}_\beta\right)+(E_\beta)\,.$$

Indeed, by Theorem 9.1(*ii*) $\Rightarrow$ (*iii*), the condition $(P_\beta)$ implies that there exist $\varepsilon>0$ such that, for all $x\in M\setminus N$ (where $N$ is a negligible set),

$$\mathbb{E}_x\left(\tau_{B(x,r)}\wedge r^\beta\right)\geq \varepsilon r^\beta,$$

whence it follows that $\mathbb{E}_x\tau_{B(x,r)}\geq\varepsilon r^\beta$, which is the lower bound in the condition $(E_\beta)$. The upper bound in $(E_\beta)$ follows trivially from $\left(\overline{E}_\beta\right)$ by taking $\Omega=B$ in (12.1).

Note also that the last implication

$$(DUE_\beta)+(P_\beta)\Longrightarrow(DUE_\beta)+(E_\beta)$$

holds because, by virtue of the previous implications on the diagram, we have

$$(DUE_\beta)+(P_\beta)\Longrightarrow(E_\beta)\,.$$

The rest of the diagram amounts to the following implications that will be proved below:

- $(FK_\beta)+(P_\beta)\Longrightarrow(DUE_\beta)$
- $(DUE_\beta)+(P_\beta)\Longrightarrow(UE_\beta)$
- $(DUE_\beta)\Longrightarrow\left(\overline{E}_\beta\right)$

- $(\Phi UE_\beta) \Longrightarrow (P_\beta)$ (the fact that $(\Phi UE_\beta) \Longrightarrow (DUE_\beta)$ is trivial).

Let us explain how Theorem 12.1 implies Theorem 1.1 from Introduction. Apart from having additional equivalence in Theorem 12.1, the main distinction is that in Theorem 1.1 one assumes a priori that a transition density exists and is a continuous function, whereas in Theorem 12.1 the existence of a heat kernel has to be proved and the conditions used are supposed to holds *almost* everywhere or outside a negligible set. In the case of Theorem 1.1, a heat kernel admits a continuous version, which greatly simplifies the arguments in the preceding sections. In particular, all essential supremums can be replaced by supremums and all negligible sets in all the hypotheses and statements can be set to be empty (hence, there is no need in the results of Section 7 whatsoever). Following the same line of argument yields Theorem 1.1.

The condition $(a)$ is used only in the proof of the implication $(DUE_\beta) \Longrightarrow \left(\overline{E}_\beta\right)$ via Lemma 11.2 and Corollary 11.3. The condition $(c)$ (the locality of the form $(\mathcal{E}, \mathcal{F})$) is explicitly used in Theorem 9.1, in all the statements of Section 10, and hence in all the results that use them. In particular, the condition $(c)$ is essential for obtaining $(UE_\beta)$ but the implication $(FK)+(P_\beta) \Longrightarrow (DUE_\beta)$ goes without it. The condition $(d)$ of stochastic completeness is used only in the proof of the implication $(\Phi UE_\beta) \Longrightarrow (P_\beta)$ via Theorem 9.1$(i) \Longrightarrow (ii)$.

## 12.2 Proof of $(FK) + (P) \Rightarrow (DUE)$

The following lemma is a modification of the iteration argument of Kigami [22, proof of Theorem 2.9]. This argument is enhanced, simplified, and generalized here to get rid of the hypothesis (1.16) and of the continuity of the heat kernel, which were used in [22].

**Lemma 12.2** *Let the form $(\mathcal{E}, \mathcal{F})$ be local and let the following two conditions are satisfied.*

- *For any ball $B = B(x_0, r)$ on $M$, a heat kernel $p_t^B$ exists and satisfies the estimate*

$$\operatorname*{esup}_{x,y\in B} p_t^B(x,y) \le \Psi_t(x_0, r), \tag{12.2}$$

  *for all $x_0 \in M$ and $r, t > 0$, where $\Psi_t(x_0, r)$ is a positive function that satisfies the following doubling condition:*

$$\Psi_{t'}(x_0, r') \le K \Psi_t(x_0, r) \tag{12.3}$$

  *for all $r \le r' \le 2r$ and $t/2 \le t' \le t$ and some constant $K$.*

- *There exists a positive, strictly monotone increasing function $\varphi(t)$ on $(0, +\infty)$ such that*

$$\int_0 \varphi(s) \frac{ds}{s} < \infty \tag{12.4}$$

  *and, for a negligible set $N \subset M$ and for all $x \in M \setminus N$, $t > 0$, and $r \ge \varphi(t)$,*

$$\mathbb{P}_x\left(\tau_{B(x,r)} \le t\right) \le \frac{1}{4K}. \tag{12.5}$$

*Then a heat kernel $p_t$ exists and satisfies the following estimate, for all $x_0 \in M$ and $t > 0$,*

$$\operatorname*{esup}_{x,y\in B(x_0,\varphi(t))} p_t(x,y) \le 2K\Psi_t(x_0, \varphi(t)).$$

**Proof.** Let $W \subset U \subset U' \subset \Omega$ be open subsets of $M$, and $\overline{U} \subset U'$. If $\Omega$ is a ball then, by (12.2), the heat kernel $p_t^\Omega$ is bounded. Applying the inequality (10.12) of Theorem 10.4 in $\Omega$ (instead of $M$) we obtain

$$p_t^\Omega(x,y) \le p_t^U(x,y) + \psi^U\left(x, \tfrac{t}{2}\right) \sup_{s/2 \le s \le t} \operatorname*{esup}_{u \in U'} p_s^\Omega(u,y) + \psi^U\left(y, \tfrac{t}{2}\right) \sup_{t/2 \le s \le t} \operatorname*{esup}_{u \in U'} p_s^\Omega(u,x),$$

for all $t > 0$ and $\mu$-a.a. $x, y \in U$, whence it follows that

$$\operatorname*{esup}_{x,y \in W} p_t^\Omega(x,y) \le \operatorname*{esup}_{x,y \in U} p_t^U(x,y) + 2 \operatorname*{esup}_{x \in W} \psi^U\left(x, \tfrac{t}{2}\right) \sup_{t/2 \le s \le t} \operatorname*{esup}_{x,y \in U'} p_s^\Omega(x,y). \tag{12.6}$$

Fix $x_0 \in M$, $0 < r < \rho' < \rho < R$, and set $W = B(x_0, r)$, $U = B(x_0, \rho')$, $U' = B(x_0, \rho)$, $\Omega = B(x_0, R)$. By (12.2) we have

$$\operatorname*{esup}_{x,y \in U} p_t^U(x,y) \le \operatorname*{esup}_{x,y \in U'} p_t^{U'}(x,y) \le \Psi_t(x_0, \rho). \tag{12.7}$$

By (12.5) we have, for all $x \in W \setminus N$,

$$\psi^U\left(x, \tfrac{t}{2}\right) \le \psi^{B(x, \rho' - r)}\left(x, \tfrac{t}{2}\right) \le \frac{1}{4K}, \tag{12.8}$$

provided $\rho' - r \ge \varphi\left(\frac{t}{2}\right)$. This will be the case if

$$\rho - r \ge \varphi(t) \tag{12.9}$$

and if $\rho'$ is sufficiently close to $\rho$. Assuming that and writing for simplicity

$$\operatorname*{esup}_V p_t^\Omega := \operatorname*{esup}_{x,y \in V} p_t^\Omega(x,y),$$

we obtain from above, for $\varepsilon := \frac{1}{2K}$,

$$\begin{aligned} \operatorname*{esup}_{B(x_0,r)} p_t^\Omega &\le \Psi_t(x_0, \rho) + \varepsilon \sup_{t/2 \le s \le t} \operatorname*{esup}_{B(x_0,\rho)} p_s^\Omega \\ &\le \Psi_t(x_0, \rho) + \varepsilon \operatorname*{esup}_{B(x_0,\rho)} p_{t/2}^\Omega, \end{aligned} \tag{12.10}$$

where we have used also the fact that, by Lemma 3.1, the function $s \mapsto \operatorname{esup}_V p_s^\Omega$ is non-increasing.

For a fix $t > 0$, set $t_n = t/2^n$, $n \ge 0$, and

$$r_n = \varphi(t_0) + \varphi(t_1) + ... + \varphi(t_{n-1}), \quad n \ge 1. \tag{12.11}$$

It follows from (12.11) that

$$r_n \le 2 \int_0^{2t} \varphi(s) \frac{ds}{s} =: I(t) < \infty.$$

Assume that $R \ge I(t)$ so that all the balls $B_n = B(x_0, r_n)$ are in $\Omega$ (see Fig. 5).

Using $r_{n+1} - r_n = \varphi(t_n)$ and observing that this condition matches (12.9), we obtain from (12.10)

$$\operatorname*{esup}_{B_n} p_{t_n}^\Omega \le \Psi_{t_n}(x_0, r_{n+1}) + \varepsilon \operatorname*{esup}_{B_{n+1}} p_{t_{n+1}}^\Omega.$$

Since $\varphi(t)$ is increasing in $t$, (12.11) implies

$$r_{n+1} = r_n + \varphi(t_n) \le r_n + \varphi(t_{n-1}) \le 2r_n,$$

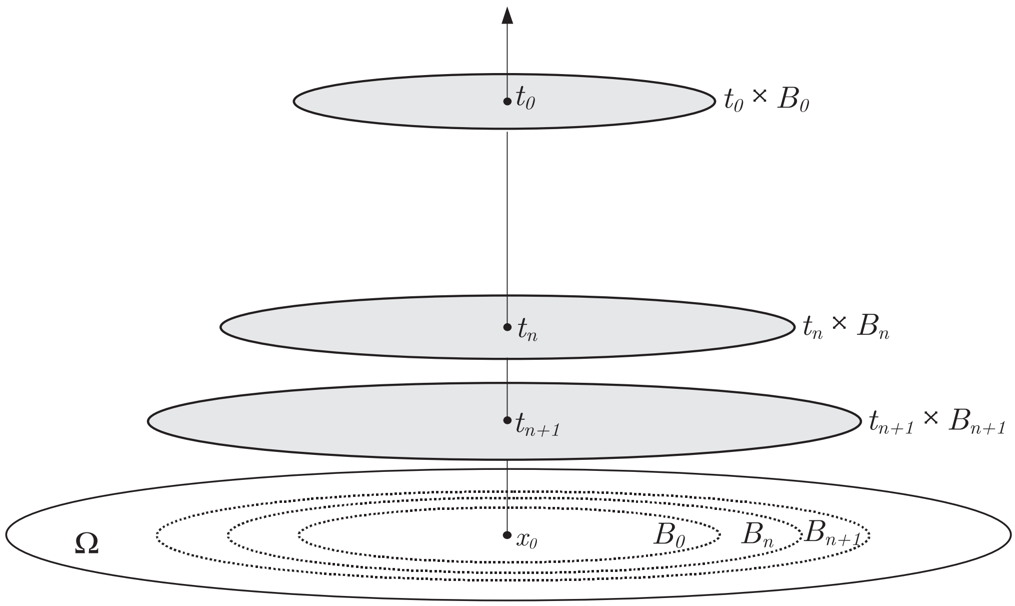


Figure 5: The sequences of times $\{t_n\}_{n=0}^{\infty}$ and balls $\{B_n\}_{n=0}^{\infty}$.

which, by the doubling property of $\Psi$, yields

$$\Psi_{t_n}(x_0, r_{n+1}) \leq K\Psi_{t_{n-1}}(x_0, r_n)$$

and hence

$$\Psi_{t_n}(x_0, r_{n+1}) \leq K^n \Psi_{t_0}(x_0, r_1) \leq K^{n+1}\Psi_{t_0}(x_0, r_0).$$

Therefore, we obtain

$$\operatorname*{esup}_{B_n} p_{t_n}^{\Omega} \leq K^{n+1}\Psi_t(x_0, r_0) + \varepsilon \operatorname*{esup}_{B_{n+1}} p_{t_{n+1}}^{\Omega},$$

whence it follows by iteration that

$$\operatorname*{esup}_{B_0} p_t^{\Omega} \leq K\Psi_t(x_0, r_0)\left(1 + K\varepsilon + (K\varepsilon)^2 + ...\right) + \varepsilon^n \operatorname*{esup}_{B_n} p_{t_n}^{\Omega}. \tag{12.12}$$

Since $\varepsilon = \frac{1}{2K}$, the above geometric series converges. Applying (12.2) for the ball $\Omega = B(x_0, R)$, we obtain

$$\operatorname*{esup}_{B_n} p_{t_n}^{\Omega} \leq \operatorname*{esup}_{\Omega} p_{t_n}^{\Omega} \leq \Psi_{t_n}(x_0, R) \leq K^n \Psi_t(x_0, R)$$

and hence

$$\lim_{n\to\infty} \varepsilon^n \operatorname*{esup}_{B_n} p_{t_n}^{\Omega} = 0.$$

Letting $n \to \infty$ in (12.12) we obtain

$$\operatorname*{esup}_{B_0} p_t^{\Omega} \leq 2K\Psi_t(x_0, r_0). \tag{12.13}$$

Finally, letting $R \to \infty$ and noticing that, by Lemma 5.1, $p_t^{\Omega} \to p_t$, we obtain that a heat kernel $p_t(x, y)$ exists and satisfies the same estimate, which was to be proved. ■

Now let prove that $(FK_\beta) + (P_\beta) \Longrightarrow (DUE_\beta)$. Let $B$ be a ball of radius $r > 0$ on $M$. Let us restate $(FK_\beta)$ as follows: for any non-empty open set $\Omega \subset B$,

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

where

$$a = \frac{c}{r^\beta}\mu(B)^\nu .$$

Therefore, by Lemma 4.2, a heat kernel $p_t^B$ exists and satisfies the upper bound

$$p_t^B(x,y) \le C(at)^{-1/\nu} ,$$

that is

$$p_t^B(x,y) \le \frac{C}{\mu(B)}\left(\frac{r^\beta}{t}\right)^{1/\nu}, \tag{12.14}$$

for $\mu$-a.a. $x, y \in B$ and all $t > 0$. Hence, the first condition of Lemma 12.2 is satisfied with the function

$$\Psi_t(x,r) = \frac{C}{V(x,r)}\left(\frac{r^\beta}{t}\right)^{1/\nu} .$$

By hypothesis $(P_\beta)$, the second condition of Lemma 12.2 is satisfied with the function $\varphi(t) = Ct^{1/\beta}$. Hence, Lemma 12.2 yields

$$\operatorname*{esup}_{x,y\in B(x_0,\varphi(t))} p_t(x,y) \le \frac{C}{V\left(x_0, t^{1/\beta}\right)},$$

whence $(DUE_\beta)$ follows.

## 12.3 Proof of $(DUE) + (P) \Rightarrow (UE)$

Here, we assume that $(DUE_\beta)$ and $(P_\beta)$ hold with some $\beta > 1$ and prove that, for $\mu$-a.a. $x, y \in M$ and $t > 0$,

$$p_t(x,y) \le \frac{C}{V\left(x, t^{1/\beta}\right)} \exp\left(-\left(\frac{d^\beta(x,y)}{Ct}\right)^{\frac{1}{\beta-1}}\right). \tag{$UE_\beta$}$$

The method of obtaining the off-diagonal upper bound of the heat kernel from the on-diagonal one using an estimate of exit probabilities, goes back to Barlow [1, Theorem 3.11], where the case $V(x,r) \simeq r^\alpha$ was covered[5]. For a general volume function $V(x,r)$ satisfying $(VD)$ the proof is more complicated – different versions can be found in [16, Theorem 6.2] and [22, Theorem 2.9]. The new proof that we present here seems to be simpler than the previous ones, although at expense of using Theorem 10.4.

Fix two distinct points $x_0, y_0 \in M$ and set $r = \frac{1}{2}d(x_0,y_0)$. By $(DUE_\beta)$, $p_t$ is locally bounded; applying inequality (10.13) of Theorem 10.4 with $V = B(x_0,r)$ and $U = B(y_0,r)$ we obtain, for $\mu$-a.a. $x \in B(x_0,r)$ and $y \in B(y_0,r)$,

$$\begin{aligned} p_t(x,y) \quad &\le \quad \psi^{B(x_0,r)}\left(x,\tfrac{t}{2}\right) \sup_{t/2\le s\le t}\ \operatorname*{esup}_{v\in B(x_0,2r)} p_s(v,y) &(12.15)\\ &\quad +\psi^{B(y_0,r)}\left(y,\tfrac{t}{2}\right) \sup_{t/2\le s\le t}\ \operatorname*{esup}_{u\in B(y_0,2r)} p_s(u,x) &(12.16)\end{aligned}$$

(see Fig. 6).

For any $x \in B(x_0,r/2)\setminus N$ we have by $(P_\beta)$ and Theorem 9.1$(ii) \Longrightarrow (vi)$,

$$\psi^{B(x_0,r)}\left(x,\tfrac{t}{2}\right) \le \psi^{B(x,r/2)}\left(x,\tfrac{t}{2}\right) \le C\exp\left(-\left(\frac{r^\beta}{Ct}\right)^{\frac{1}{\beta-1}}\right), \tag{12.17}$$

[5] It is interesting to mention that in the case when $M$ is a Riemannian manifold then $(DUE_2) \Leftrightarrow (UE_2)$ (see [11, Proposition 5.2] or [12, Theorem 3.1]).

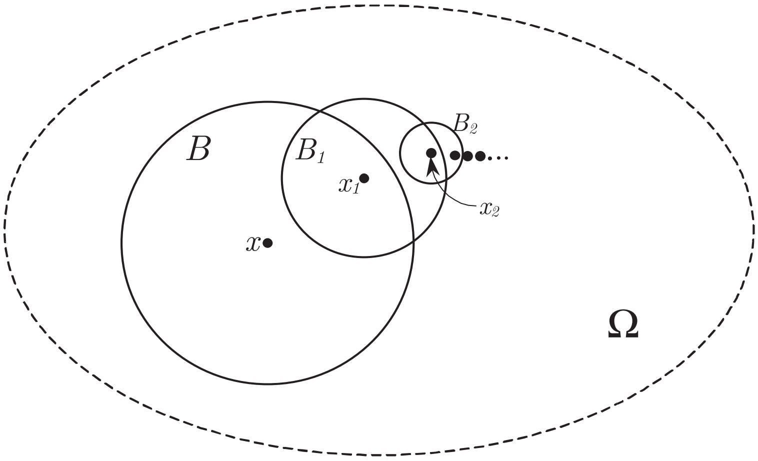


Figure 6: Illustration to the proof

and, similarly, for any $y \in B(y_0, r/2) \setminus N$,

$$\psi^{B(y_0,t)}\left(y, \tfrac{t}{2}\right) \leq C \exp\left(-\left(\frac{r^\beta}{Ct}\right)^{\frac{1}{\beta-1}}\right). \tag{12.18}$$

By $(DUE_\beta)$ we have, for $\mu$-a.a. $v, y \in M$ ,

$$p_s(v, y) \leq \frac{C}{\sqrt{V\left(v, s^{1/\beta}\right) V\left(y, s^{1/\beta}\right)}}. \tag{12.19}$$

Using (11.1) we obtain, for all $x, z \in B(x_0, Cr)$ and any $\varepsilon > 0$,

$$\frac{V\left(x, s^{1/\beta}\right)}{V\left(z, s^{1/\beta}\right)} \leq C\left(1 + \frac{r}{s^{1/\beta}}\right)^\alpha \leq C_\varepsilon \exp\left(\varepsilon\left(\frac{r^\beta}{s}\right)^{\frac{1}{\beta-1}}\right). \tag{12.20}$$

Applying (12.20) for $z = v$ and for $z = y$, and substituting into (12.19), we obtain, for $\mu$-a.a. $x, y, v \in B(x_0, Cr)$,

$$p_s(v, y) \leq \frac{C_\varepsilon}{V\left(x, s^{1/\beta}\right)} \exp\left(\varepsilon\left(\frac{r^\beta}{s}\right)^{\frac{1}{\beta-1}}\right). \tag{12.21}$$

Taking here supremum in $s \in \left(\frac{t}{2}, t\right)$ amounts to replacing $s$ by $t$ and to changing constants.

Finally, substituting (12.17), (12.18), (12.21), and a similar upper bound for $p_s(u, x)$ into (12.15)-(12.16), we obtain, for $\mu$-a.a. $x \in B(x_0, r/2)$ and $y \in B(y_0, r/2)$,

$$p_t(x, y) \leq \frac{C_\varepsilon}{V\left(x, t^{1/\beta}\right)} \exp\left(\varepsilon\left(\frac{r^\beta}{t}\right)^{\frac{1}{\beta-1}} - \left(\frac{r^\beta}{Ct}\right)^{\frac{1}{\beta-1}}\right).$$

Choosing $\varepsilon$ small enough and noticing that $d(x, y) \leq 3r$, we obtain $(UE_\beta)$. ■

## 12.4 Proof of $(DUE) \Rightarrow (\overline{E})$

Let us first prove that $(DUE_\beta)$ implies the following estimate: for any ball $B = B(x_0, r)$ on $M$, for $\mu$-a.a. $x, y \in B$, and for all $t > 0$,

$$p_t^B(x, y) \leq \frac{C}{\mu(B)}\left(\frac{r}{t^{1/\beta}}\right)^\alpha, \tag{12.22}$$

where $\alpha$ is the exponent from (11.1). Write for simplicity

$$\operatorname{esup} p_t^B := \operatorname*{esup}_{x,y\in B} p_t^B(x,y) ,$$

and observe the following property of the function $t \mapsto \operatorname{esup} p_t^B$: if the inequality

$$\operatorname{esup} p_t^B \le \frac{K}{\mu(B)}\left(\frac{r}{t^{1/\beta}}\right)^{\alpha} \tag{12.23}$$

holds for $t = s$ then (12.23) holds also for $t = 2s$ provided

$$s \ge T := 2K^{\beta/\alpha} r^{\beta} \tag{12.24}$$

(here $K > 1$ is a constant to be specified below). Indeed, by the semigroup property, we have for $\mu$-a.a. $x, y \in B$

$$p_{2s}^B(x,y) = \int_B p_s^B(x,z)\, p_s^B(z,y)\, d\mu(z) \le \left(\operatorname{esup} p_s^B\right)^2 \mu(B),$$

whence by (12.23) and (12.24)

$$\operatorname{esup} p_{2s}^B \le \frac{K^2}{\mu(B)}\left(\frac{r}{s^{1/\beta}}\right)^{2\alpha} \le \frac{K^2}{\mu(B)}\left(\frac{r}{(T/2)^{1/\beta}}\right)^{\alpha}\left(\frac{r}{(2s)^{1/\beta}}\right)^{\alpha} = \frac{K}{\mu(B)}\left(\frac{r}{(2s)^{1/\beta}}\right)^{\alpha},$$

which was claimed.

Assume for a moment that we have proved (12.23) for $t = T$. Then by induction the above property of the function $t \mapsto \operatorname{esup} p_t^B$ yields that (12.23) holds for all $t = 2^n T$, where $n$ is a non-negative integer. Since by Lemma 3.1 the function $t \mapsto \operatorname{esup} p_t^B$ is non-increasing, we obtain for $2^n T \le t < 2^{n+1} T$ that

$$\operatorname{esup} p_t^B \le \operatorname{esup} p_{2^n T}^B \le \frac{K}{\mu(B)}\left(\frac{r}{(2^n T)^{1/\beta}}\right)^{\alpha} \le \frac{K 2^{\alpha/\beta}}{\mu(B)}\left(\frac{r}{t^{1/\beta}}\right)^{\alpha}.$$

Therefore, if we prove that there exists $K$ such that (12.23) holds for $0 < t \le T$ then we can conclude that (12.22) holds for all $t > 0$.

Consider first the case $0 < t \le r^{\beta}$. By $(DUE_\beta)$ we have, for $\mu$-a.a. $x, y \in M$ and $t > 0$,

$$p_t(x,y) \le \frac{C_0}{\sqrt{V\left(x, t^{1/\beta}\right) V\left(y, t^{1/\beta}\right)}} \tag{12.25}$$

(the argument below is sensitive to constant factors, so we use individual notation for different constants as $C_0$). Observe that, by $(VD)$ and (11.1), for any $x \in B$ and $0 < t \le r^{\beta}$,

$$\frac{V(x_0, r)}{V\left(x, t^{1/\beta}\right)} \le C_1 \left(\frac{r}{t^{1/\beta}}\right)^{\alpha}. \tag{12.26}$$

Since $p_t^B \le p_t$, (12.25) and (12.26) imply that, for $\mu$-a.a. $x, y \in B$ and $0 < t \le r^{\beta}$,

$$p_t^B(x,y) \le \frac{C_0 C_1}{V(x_0, r)}\left(\frac{r}{t^{1/\beta}}\right)^{\alpha}, \tag{12.27}$$

that is, (12.23) holds for $0 < t \le r^{\beta}$ provided $K \ge C_0 C_1$.

Consider now the remaining case $r^\beta < t \leq T$. We have, for any $x \in B$,

$$\frac{1}{V\left(x, t^{1/\beta}\right)} = \frac{V\left(x_0, T^{1/\beta}\right)}{V\left(x, t^{1/\beta}\right)} \frac{V\left(x_0, r\right)}{V\left(x_0, T^{1/\beta}\right)} \frac{1}{V\left(x_0, r\right)}.$$

Since $t \leq T$, we obtain using $(VD)$, (11.1), and (12.24), that

$$\frac{V\left(x_0, T^{1/\beta}\right)}{V\left(x, t^{1/\beta}\right)} \leq C_1 \left(\frac{T}{t}\right)^{\alpha/\beta} = C_1 2^{a/\beta} K \left(\frac{r}{t^{1/\beta}}\right)^{\alpha}.$$

Since $r < T^{1/\beta}$, Lemma 11.2 yields

$$\frac{V\left(x_0, r\right)}{V\left(x_0, T^{1/\beta}\right)} \leq C_2 \left(\frac{r}{T^{1/\beta}}\right)^{\alpha'} \leq C_2 K^{-\alpha'/\alpha}$$

Hence, (12.25) implies, for $\mu$-a.a. $x, y \in B$,

$$p_t^B(x,y) \leq C_0 C_1 C_2 2^{\alpha/\beta} K^{-\alpha'/\alpha} \frac{K}{V\left(x_0, r\right)} \left(\frac{r}{t^{1/\beta}}\right)^{\alpha},$$

whence (12.23) follows, provided $K$ is chosen large enough to satisfy

$$C_0 C_1 C_2 2^{\alpha/\beta} K^{-\alpha'/\alpha} \leq 1.$$

Let us now show that (12.22) implies $\left(\overline{E}_\beta\right)$. Select a countable dense sequence $\{x_n\} \subset M$ and call a ball $B(x,r)$ *selected* if $x = x_n$ for some $n$ and if its radius $r$ is rational. Hence, the family of selected balls is countable. Recall that $\left(\overline{E}\right)$ means that for any ball $B = B(x,r)$ and for any open set $\Omega \subset B(x,r)$

$$\mathbb{E}_x \tau_\Omega \leq C \left(\frac{\mu(\Omega)}{\mu(B)}\right)^{\beta/\alpha} r^\beta, \tag{12.28}$$

for all $x \in \Omega \setminus N$, where $N$ is a negligible set. Clearly, it suffices to prove (12.28) for selected balls $B$.

By (12.22) the heat kernel $p_t^B$ is bounded. Hence, by Corollary 8.2, a transition density $\widetilde{p}_t(x,y)$ exists and satisfies the same upper bound (12.22). By Lemma 8.3, for any ball $B$ there exists a negligible set $N_B \subset M$ such that

$$\sup_{x \in B \setminus N_B} \sup_{y \in B \setminus N_B} \widetilde{p}_t(x,y) = \operatorname*{esup}_B \widetilde{p}_t^B(x,y).$$

Let $N$ be the union of all sets $N_B$ where $B$ is a selected balls, so that $N$ is also a negligible set. Hence, for any selected ball $B$ of radius $r$, we have

$$\sup_{x \in B \setminus N} \sup_{y \in B \setminus N} \widetilde{p}_t^B(x,y) \leq \frac{C}{\mu(B)} \left(\frac{r}{t^{1/\beta}}\right)^{\alpha}.$$

By (6.3), we obtain, for any non-empty open set $\Omega \subset B$ and for all $x \in \Omega \setminus N$, $T > 0$,

$$\begin{aligned}
\mathbb{E}_x \tau_\Omega &\leq \int_0^T \mathcal{P}_t^\Omega 1(x)\, d\mu + \int_T^\infty \mathcal{P}_t^B 1(x)\, d\mu \\
&\leq T + \int_T^\infty \int_\Omega \widetilde{p}_t^B(x,y)\, d\mu(y)\, dt \\
&\leq T + \int_T^\infty \frac{C\mu(\Omega)}{\mu(B)} \left(\frac{r}{t^{1/\beta}}\right)^{\alpha} dt \\
&\leq T + \frac{C\mu(\Omega)}{\mu(B)} r^\alpha T^{1-\alpha/\beta}
\end{aligned}$$

(note that $\alpha$ is the exponent from $(VD)$, which can be taken arbitrarily large; we have assumed here that $\alpha > \beta$). Finally, choosing

$$T = \left(\frac{\mu(\Omega)}{\mu(B)}\right)^{\beta/\alpha} r^{\beta},$$

we obtain (12.28).

## 12.5 Proof of $(\Phi UE) \Rightarrow (P)$

Since the heat kernel $p_t$ is locally bounded, by Corollary 8.2 a transition density $\widetilde{p}_t$ exists and also satisfies $(\Phi UE_\beta)$. By Corollary 8.4, there exists a negligible set $N$ such that for all $x, y \in M \setminus N$ and $t > 0$,

$$\widetilde{p}_t(x,y) \leq \frac{C}{V\left(x, t^{1/\beta}\right)} \Phi\left(\frac{d(x,y)}{t^{1/\beta}}\right).$$

Let us show that, for any $\varepsilon > 0$ there exists $\delta > 0$ such that if $0 < t \leq \delta r^{\beta}$ then, for all $x \in M \setminus N$,

$$\mathbb{P}_x\left(X_t \in B(x,r)^c\right) \leq \varepsilon. \tag{12.29}$$

Indeed, assuming $x \in M \setminus N$ and setting $r_k = 2^k r$ we obtain

$$\begin{aligned}
\mathbb{P}_x\left(X_t \in B(x,r)^c\right) &= \int_{B(x,r)^c} \widetilde{p}_t(x,y)\, d\mu(y) = \sum_{k=0}^{\infty} \int_{B(x,r_{k+1}) \setminus B(x,r_k)} \widetilde{p}_t(x,y)\, d\mu(y) \\
&\leq \sum_{k=0}^{\infty} V(x, r_{k+1}) \frac{C}{V\left(x, t^{1/\beta}\right)} \Phi\left(\frac{r_k}{t^{1/\beta}}\right) \\
&\leq \sum_{k=0}^{\infty} C\left(\frac{r_k}{t^{1/\beta}}\right)^{\alpha} \Phi\left(\frac{r_k}{t^{1/\beta}}\right) \\
&\leq C \int_{\frac{1}{2} r/t^{1/\beta}}^{\infty} s^{\alpha-1} \Phi(s)\, ds.
\end{aligned} \tag{12.30}$$

By the hypothesis (1.11), the right hand side of (12.30) can be made smaller than $\varepsilon$ provided $r/t^{1/\beta}$ is sufficiently large, which was claimed. By the part $(i) \Rightarrow (ii)$ of Theorem 9.1, (12.29) with $\varepsilon < 1/2$ implies $(P_\beta)$. ■

Hence, we have finished the proof of Theorems 1.1 and 12.1. Note that the hypothesis of stochastic completeness was used only once in the proof, namely in the proof of Theorem 9.1$(i) \Longrightarrow (ii)$ (see eq. (9.4)). The latter is exactly the part of Theorem 9.1, which was used in the above argument.

The proof of the implication $(UE_\beta) \Rightarrow (E_\beta)$ by Kigami [22] also uses the stochastic completeness although not explicitly stated. Let us present an example, showing that without the stochastic completeness the implication $(UE_\beta) \Rightarrow (E_\beta)$ is not true.

**Example 12.3** Consider in $\mathbb{R}$ the process $\{X_t\}$ generated by the operator $H = -\frac{d^2}{dx^2} + Q(x)$ where $Q \in C^\infty(\mathbb{R})$ is a positive function. To be precise, we consider $\mathbb{R}$ with the Euclidean distance $d$, the Lebesgue measure $\mu$, and the Dirichlet form

$$\mathcal{E}(f,g) = \int_{\mathbb{R}} \left(f'g' + Qfg\right) d\mu$$

in the domain $\mathcal{F} = \overset{o}{H}{}^1(\mathbb{R})$. All the hypotheses $(a)$-$(d)$ of Theorem 1.1 are satisfied. On the other hand, the associated diffusion process $\{X_t\}$ is not stochastically complete because of the killing term $Q$. Since $Q > 0$, the heat kernel of this process admits the upper bound

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

that is $(UE_2)$ holds.

Let us verify that the lower bound in $(E_2)$ fails in general. For example, take $Q(x) = x^2$. Then the heat kernel of $\{X_t\}$ is given by the explicit expression

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

In particular, noticing that $\frac{1}{\sinh 2t} + \tanh t \geq 1$, we obtain

$$p_t(0,x) \leq \frac{1}{(2\pi \sinh 2t)^{1/2}} \exp\left(-\frac{1}{2}x^2\right)$$

whence

$$\int_0^\infty \int_{\mathbb{R}} p_t(0,x)\, dx\, dt < \infty.$$

By (6.3), the function $r \mapsto \mathbb{E}_0 \tau_{B(0,r)}$ is bounded, which makes the lower bound $\mathbb{E}_0 \tau_{B(0,r)} \geq cr^2$ impossible.

# 13 Appendix: Resistance metric and $(FK)$

Here we show an example of derivation of a Faber-Krahn inequality $(FK)$ directly from the volume properties of the balls in a resistance metric. Define the *resistance* $\mathcal{R}(x,y)$ between points $x, y \in M$ by

$$\mathcal{R}(x,y) := \sup_{f \in \mathcal{F} \cap C_0(M)} \frac{|f(x) - f(y)|^2}{\mathcal{E}[f]}.$$

Of course, it may well happen that $\mathcal{R}(x,y) = \infty$ but for many examples of fractal spaces, it is known that $\mathcal{R}(x,y) < \infty$. It is easy to see that in this case $\sqrt{\mathcal{R}(x,y)}$ is a metric on $M$. Define the corresponding balls

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

and set $\mathcal{V}(x,r) := \mu(\mathcal{B}(x,r))$.

**Theorem 13.1** *Let $M$ be a locally compact connected topological space, $\mu$ be a Radon measure with full support on $M$, and $(\mathcal{E}, \mathcal{F})$ be a regular Dirichlet form in $L^2(M,\mu)$. Assume that $\sqrt{\mathcal{R}(x,y)}$ is a metric on $M$ compatible with the topology of $M$, and* $\operatorname{diam} M = \infty$. *Assume also that the volume function $\mathcal{V}(x,r)$ satisfies $(VD)$, that is*

$$\mathcal{V}(x,2r) \leq C\mathcal{V}(x,r),$$

*for all $x \in M$ and $r > 0$. Then, for any ball $\mathcal{B}(x,r)$ and any non-empty open set $\Omega \subset \mathcal{B}(x,r)$,*

$$\lambda_{\min}(\Omega) \geq \frac{c}{r\mathcal{V}(x,r)} \left(\frac{\mathcal{V}(x,r)}{\mu(\Omega)}\right)^\nu, \tag{13.1}$$

*for some positive constants $\nu$ and $c$.*

**Proof.** The following argument is close to [20, Lemma 4.2]. Recall that, by the variational property,

$$\lambda_{\min}(\Omega) = \inf_{f \in \mathcal{F} \cap C_0(\Omega)} \frac{\mathcal{E}[f]}{\|f\|_2^2}.$$

Clearly, one can restrict here to those $f$ for which $\sup|f| = 1$. Take any function $f \in \mathcal{F} \cap C_0(\Omega)$ with $\sup|f| = 1$ and let $x_0 \in \Omega$ be a point such that $|f(x_0)| = 1$. Let $\rho$ be the largest radius such that $\mathcal{B}(x_0, \rho) \subset \Omega$. The fact that the support $\operatorname{supp} f$ is a compact subset of $\Omega$ implies that the ball $\mathcal{B}(x_0, \rho)$ is not covered by $\operatorname{supp} f$; let $y_0 \in \mathcal{B}(x_0, \rho) \setminus \operatorname{supp} f$ (see Fig. 7).

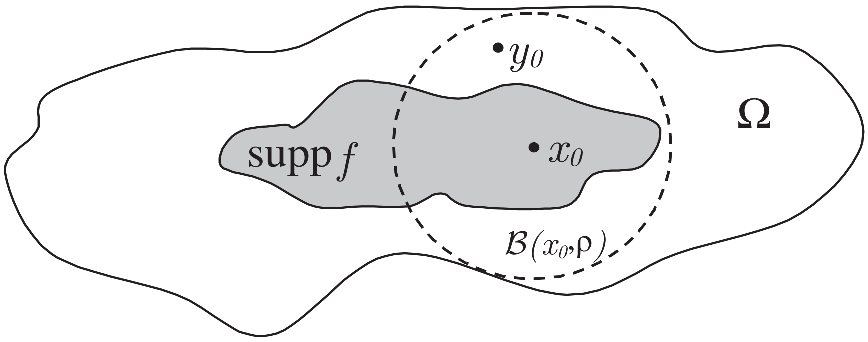


Figure 7: Points $x_0$ and $y_0$.

Then we have

$$\rho > \mathcal{R}(x_0, y_0) \geq \frac{|f(x_0) - f(y_0)|^2}{\mathcal{E}[f]} = \frac{1}{\mathcal{E}[f]},$$

whence $\mathcal{E}[f] \geq \rho^{-1}$. Since $\|f\|_2^2 \leq \mu(\Omega)$, we obtain

$$\frac{\mathcal{E}[f]}{\int_\Omega f^2 d\mu} \geq \frac{1}{\rho\mu(\Omega)}$$

and hence

$$\lambda_{\min}(\Omega) \geq \frac{1}{\rho\mu(\Omega)}.$$

On the other hand, we have

$$\mathcal{V}(x_0, \rho) \leq \mu(\Omega)$$

and, by the doubling property and Lemma 11.1,

$$\frac{\mathcal{V}(x, r)}{\mathcal{V}(x_0, \rho)} \leq C \left(\frac{r}{\rho}\right)^\alpha,$$

whence

$$\frac{1}{\rho} \geq \frac{c}{r} \left(\frac{\mathcal{V}(x, r)}{\mu(\Omega)}\right)^{1/\alpha}$$

and

$$\lambda_{\min}(\Omega) \geq \frac{c}{r\mu(\Omega)} \left(\frac{\mathcal{V}(x, r)}{\mu(\Omega)}\right)^{1/\alpha} = \frac{c}{r\mathcal{V}(x, r)} \left(\frac{\mathcal{V}(x, r)}{\mu(\Omega)}\right)^{1+1/\alpha}.$$

■

**Corollary 13.2** *Under the hypotheses of* Theorem 13.1, *if in addition* $\mathcal{V}(x, r) \leq Cr^N$ *for all* $x \in M$ *and* $r > 0$ *then* $M$ *satisfies* $(FK_\beta)$ *with* $\beta = 1 + N$.

**Proof.** Indeed, the inequality (13.1) of Theorem 13.1 implies

$$\lambda_{\min}(\Omega) \geq \frac{c}{r^{1+N}} \left( \frac{\mathcal{V}(x,r)}{\mu(\Omega)} \right)^{\nu},$$

which was to be proved. ■

# References


[1] **Barlow M.T.**, Diffusions on fractals, *in:* "Lectures on Probability Theory and Statistics, Ecole d'été de Probabilités de Saint-Flour XXV - 1995", Lecture Notes Math. 1690, Springer, 1998. 1-121.

[2] **Barlow M.T.**, Heat kernels and sets with fractal structure, *in:* "Heat kernels and analysis on manifolds, graphs, and metric spaces", *Contemporary Mathematics*, **338** (2003) 11-40.

[3] **Barlow M.T., Bass R.F.**, Brownian motion and harmonic analysis on Sierpinski carpets, *Canad. J. Math.*, **54** (1999) 673-744.

[4] **Barlow M.T., Perkins E.A.**, Brownian motion on the Sierpinski gasket, *Probab. Th. Rel. Fields*, **79** (1988) 543–623.

[5] **Carlen E.A., Kusuoka S., Stroock D.W.**, Upper bounds for symmetric Markov transition functions, *Ann. Inst. H. Poincaré, Prob. et Stat.*, (1987) suppl. au no.2, 245-287.

[6] **Chavel I.**, "Eigenvalues in Riemannian geometry", Academic Press, New York, 1984.

[7] **Chung F.R.K., Grigor'yan A., Yau S.-T.**, Higher eigenvalues and isoperimetric inequalities on Riemannian manifolds and graphs, *Comm. Anal. Geom*, **8** (2000) no.5, 969-1026.

[8] **Davies E.B.**, "Heat kernels and spectral theory", Cambridge University Press, 1989.

[9] **Fukushima M., Oshima Y., Takeda M.**, "Dirichlet forms and symmetric Markov processes", Studies in Mathematics 19, De Gruyter, 1994.

[10] **Grigor'yan A.**, The heat equation on non-compact Riemannian manifolds, (in Russian) *Matem. Sbornik*, **182** (1991) no.1, 55-87. Engl. transl. *Math. USSR Sb.*, **72** (1992) no.1, 47-77.

[11] **Grigor'yan A.**, Heat kernel upper bounds on a complete non-compact manifold, *Revista Matemática Iberoamericana*, **10** (1994) no.2, 395-452.

[12] **Grigor'yan A.**, Gaussian upper bounds for the heat kernel on arbitrary manifolds, *J. Diff. Geom.*, **45** (1997) 33-52.

[13] **Grigor'yan A.**, Estimates of heat kernels on Riemannian manifolds, *in:* "Spectral Theory and Geometry. ICMS Instructional Conference, Edinburgh 1998", Ed. E.B. Davies and Yu. Safarov, London Math. Soc. Lecture Note Series 273, Cambridge Univ. Press, 1999. 140-225.

[14] **Grigor'yan A., Saloff-Coste L.**, Heat kernel on manifolds with ends, in preparation.

[15] **Grigor'yan A., Telcs A.**, Sub-Gaussian estimates of heat kernels on infinite graphs, *Duke Math. J.*, **109** (2001) no.3, 452-510.

[16] **Grigor'yan A., Telcs A.**, Harnack inequalities and sub-Gaussian estimates for random walks, *Math. Ann.*, **324** no.3, (2002) 521-556.

[17] **Grigor'yan A., Telcs A.**, Two-sided estimates of heat kernels in metric measure spaces, in preparation

[18] **Gushchin A.K.**, On the uniform stabilization of solutions of the second mixed problem for a parabolic equation, (in Russian) *Matem. Sbornik*, **119(161)** (1982) no.4, 451-508. Engl. transl. *Math. USSR Sb.*, **47** (1984) 439-498.

[19] **Hambly B., Kumagai T.**, Transition density estimates for diffusion processes on post critically finite self-similar fractals, *Proc. London Math. Soc. (3)*, **78** (1999) 431–458.

[20] **Hambly B., Kumagai T.**, Heat kernel estimates for symmetric random walks on a class of fractal graphs and stability under rough isometries, preprint

[21] **Kigami J.**, "Analysis on fractals", Cambridge University Press, 2001.

[22] **Kigami J.**, Local Nash inequality and inhomogeneity of heat kernels, preprint 2003.

[23] **Li P., Yau S.-T.**, On the parabolic kernel of the Schrödinger operator, *Acta Math.*, **156** (1986) no.3-4, 153-201.

[24] **Nash J.**, Continuity of solutions of parabolic and elliptic equations, *Amer. J. Math.*, **80** (1958) 931-954.

[25] **Rosenberg S.**, "The Laplacian on a Riemannian manifold", London Mathematical Society Student Texts 31, Cambridge University Press, 1997.

[26] **Saloff-Coste L.**, A note on Poincaré, Sobolev, and Harnack inequalities, *Duke Math J.*, **65** no.3, *Internat. Math. Res. Notices*, **2** (1992) 27-38.

[27] **Saloff-Coste L.**, "Aspects of Sobolev inequalities", LMS Lecture Notes Series 289, Cambridge Univ. Press, 2002.

[28] **Schoen R., Yau S.-T.**, "Lectures on Differential Geometry", Conference Proceedings and Lecture Notes in Geometry and Topology 1, International Press, 1994.

[29] **Varopoulos N.Th., Saloff-Coste L., Coulhon T.**, "Analysis and geometry on groups", Cambridge Tracts in Mathematics, 100, Cambridge University Press, Cambridge, 1992.

[30] **Yan J.-A.**, A formula for densities of transition functions, *in:* "Sem. Prob. XXII", 92-100. Lecture Notes Maths. 1321,