# Sub-Gaussian estimates of heat kernels on infinite graphs


Alexander Grigor'yan*
Imperial College
180 Queen's Gate
London SW7 2BZ
United Kingdom
a.grigoryan@ic.ac.uk

Andras Telcs†
International Business School
Tárogató u.2-4
H-1021 Budapest
Hungary
atelcs@tech.ibs-b.hu


22 June 2000


**Abstract**

We prove that a two sided sub-Gaussian estimate of the heat kernel on an infinite weighted graph takes place if and only if the volume growth of the graph is uniformly polynomial and the Green kernel admits a uniform polynomial decay.


## Contents




*Supported by the EPSRC research fellowship B/94/AF/1782
†Partially supported by a visiting grant of the London Mathematical Society.

# 1 Introduction

Consider the heat equation

$$\frac{\partial f}{\partial t} = \Delta f, \tag{1.1}$$

where $f = f(t,x)$ is a function of $t > 0$ and $x \in \mathbb{R}^n$, and $\Delta$ is the Laplace operator in $\mathbb{R}^n$. The fundamental solution to (1.1) is given by the classical Gauss-Weierstrass formula

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

The function $p_t(x,y) = f(t, x-y)$ is called *the heat kernel* of the Laplace operator.

In the past 30-35 years, there have been numerous works devoted to estimates of heat kernels in various settings (see, for example, books and surveys [4], [15], [16], [24], [34], [50], [53], [63], [64]). These are parabolic equations with variable coefficients, the heat equation on Riemannian manifolds, the discrete heat equation on graphs, and the heat semigroups on general metric measure spaces including fractal-like sets. Despite of the high diversity of the underlying spaces and equations, in many important cases the heat kernel is naturally defined and, moreover, admits the so-called *Gaussian estimates.*

For any metric measure space $M$ with distance $d$ and measure $\mu$, denote by $B(x,r)$ the open metric ball of radius $r$ centered at $x$, and by $V(x,r)$ its measure $\mu$. Suppose first that $M$ is either a discrete group or a Lie group, with properly defined $d, \mu$ and the heat kernel $p_t(x,y)$. Assume that the volume growth of $M$ is polynomial, that is, for some $\alpha > 0$,

$$V(x,r) \simeq r^\alpha \tag{1.2}$$

(here the sign $\simeq$ means that the ratio of both sides of (1.2) stays between two positive constants). Then the heat kernel on $M$ admits the following Gaussian estimate (see [62], [36])

$$p_t(x,y) \simeq t^{-\alpha/2} \exp\left(-\frac{d^2(x,y)}{ct}\right) \tag{1.3}$$

(where the positive constant $c$ may be different for the upper and lower bounds). The heat kernel in $\mathbb{R}^n$ obviously satisfies (1.3) with $\alpha = n$.

Suppose now that $M$ is a complete manifold with non-negative Ricci curvature. Then the following estimate of Li and Yau [46] is well-known

$$p_t(x,y) \simeq \frac{1}{V(x,\sqrt{t})} \exp\left(-\frac{d^2(x,y)}{ct}\right). \tag{1.4}$$

In particular, if $V(x,r) \simeq r^\alpha$ then the heat kernel satisfies again the estimate (1.3).

As we see, for groups of the polynomial growth and for non-negatively curved manifolds, the heat kernel is fully determined (up to constant factors) by the volume growth function. In other words, the potential theory on such spaces is characterized by a single parameter $\alpha$ - the exponent of the volume growth.

The presence of the Gaussian estimates (1.3) or (1.4) reflects certain properties of the space $M$. In particular, (1.4) implies that the Markov process $X_t$ with the transition density $p_t(x,y)$ has the diffusion speed of the order $t^{\frac{1}{2}}$. The latter means that the process $X_t$ started at a point $x$ first exits the ball $B(x,R)$ at the time $t \simeq R^2$.

The development of Markov processes on fractals and the fractal like graphs ( [10], [7], [29], [35], [39], [40], [41], [43], [44], [57], [65] etc.) has led to construction of homogeneous metric spaces $M$ where the process $X_t$ has the diffusion speed of the order $t^{\frac{1}{\beta}}$, with some $\beta > 2$. Such a process $X_t$ is referred to as *subdiffusive*, and is characterized by *two* parameters $\alpha$ and $\beta$, which determine *sub-Gaussian estimates* of the heat kernel:

$$p_t(x,y) \simeq t^{-\alpha/\beta} \exp\left(-\left(\frac{d^\beta(x,y)}{ct}\right)^{\frac{1}{\beta-1}}\right). \tag{1.5}$$

Here $\alpha$ is the exponent of the volume growth as in (1.2). The Gaussian estimate (1.3) is a particular case of (1.5) for $\beta = 2$.

Barlow and Bass [7] showed that the sub-Gaussian estimates (1.5) with $\beta > 2$ can take place not only on singular spaces such as fractals but also on smooth Riemannian manifolds, for a certain range of time. Similar estimates hold for random walks on certain fractal-like graphs [8], [38]. It has become apparent that a large and interesting class of homogeneous spaces features sub-Gaussian estimates of the heat kernel. The potential theory on such spaces is determined by the *two* parameters and hence, cannot be recovered only from the volume growth[1].

A natural question arises:

> *How to characterize those spaces which admit sub-Gaussian estimates (1.5) of the heat kernel?*

If $M$ is a complete non-compact Riemannian manifold then the validity of the Gaussian estimate (1.3) is known to be equivalent to the following two conditions: the volume growth (1.2) and the Poincaré inequality

$$\lambda_1^{(N)}(B(x,r)) \geq \frac{c}{r^2}, \tag{1.6}$$

where $\lambda_1^{(N)}(B)$ is the first non-zero eigenvalue of the Neumann boundary value problem in the ball $B$ (see [51], [30]; similar results are known also for graphs [27] and for abstract local Dirichlet spaces [55]). It may be tempting to conjecture that by replacing in (1.6) $r^2$ by $r^\beta$, one obtains equivalent conditions for sub-Gaussian estimates. However, this conjecture is false. At the present time, no similar characterization of the spaces with sub-Gaussian estimates seems to be known. All examples of spaces where (1.5) is proved are fractal-like spaces featuring a self-similarity structure.

The purpose of this paper is to provide a new approach to obtaining sub-Gaussian estimates of the heat kernel. Our point of departure is the understanding that, apart from the uniform volume growth $V(x,r) \simeq r^\alpha$, we have to introduce additional hypotheses, which would contain the second parameter $\beta$ and provide the necessary homogeneity of the space (just the uniform volume growth is not enough for the latter).

[1] The parameters $\alpha$ and $\beta$ must satisfy the inequalities $2 \leq \beta \leq \alpha + 1$, which seem to be the *only* constraint on $\alpha$ and $\beta$. We are indebted to Martin Barlow for providing us with the evidence for the latter.

Let $g(x,y)$ be the Green kernel on $M$, that is

$$g(x,y)=\int_0^\infty p_t(x,y)dt.$$

Recall that, in $\mathbb{R}^n$, $g(x,y)=c_n\,|x-y|^{-(n-2)}$ if $n>2$ and $g\equiv\infty$ if $n\leq 2$.

Our general result says the following:

*Given the parameters $\alpha>\beta\geq 2$, the two-sided sub-Gaussian estimate*

$$p_t(x,y)\simeq t^{-\alpha/\beta}\exp\left(-\left(\frac{d^\beta(x,y)}{ct}\right)^{\frac{1}{\beta-1}}\right) \tag{1.7}$$

*holds if and only if*

$$V(x,r)\simeq r^\alpha \quad \text{and} \quad g(x,y)\simeq d(x,y)^{-(\alpha-\beta)}. \tag{1.8}$$

We do not specify here the ranges of the variables $x,y,t,r$ because they are different for different settings. In the present paper, we treat the case when the underlying space is a graph, and the time is also discrete. However, the graph case contains already all difficulties. We present the proof in the way that only minimal changes are required to pass to a general setting of abstract metric spaces, which will be dealt with elsewhere. The exact statements are given in the next section. Note that our result is new even for the Gaussian case $\beta=2$.

The hypothesis (1.8) consists of two conditions of different nature. The first one is a geometric condition of the volume growth whereas the second is an estimate of a fundamental solution to an elliptic equation. Neither of them separately implies the heat kernel bounds (1.7). Surprisingly enough, the exponent $\beta$ which provides the scaling of the space and time variables for a *parabolic* equation, can be recovered from an *elliptic* equation, although combined with the volume growth.

The paper is arranged as follows. In Section 2, we state the main result - Theorem 2.1. In Section 3, we introduce the necessary tools such as the discrete Laplace operator, its eigenvalues, the mean exit time, etc. In Section 4 we describe the scheme of the proof of Theorem 2.1 as well as some consequences. In particular, we mention some other conditions equivalent to (1.7). The actual proof of Theorem 2.1 consists of many steps which are considered in details in the rest Sections 5 - 15.

NOTATION

The letters $c,C$ are reserved for positive constants not depending on the variables in question. They may be different on different occurrences, even within the same formula. All results of the paper are quantitative in the sense that the constants in conclusions depends only on the constants in hypotheses.

The relation $f\simeq g$ means that the ratio of the functions $f$ and $g$ is bounded from above and below by positive constants, for the specified range of the variables. If one of those functions contains a sub-Gaussian factor $\exp\left(-\left(\frac{d^\beta}{ct}\right)^{\frac{1}{\beta-1}}\right)$ then the constant $c$ in exp may be different for the upper and lower bounds (cf. (1.7)).

We use a number of lettered formulas such as $(UE)$, $(LE)$ etc., to refer to the most important and frequently used conditions. In Appendix, we provide a complete list of all such formulas.

## 2 Statement of the main result

Throughout the paper, $\Gamma$ denotes an infinite, connected, locally finite graph. If $x, y \in \Gamma$ then we write $x \sim y$ provided $x$ and $y$ are connected by an edge. The graph is always assumed non-oriented, that is $x \sim y$ is equivalent to $y \sim x$. We do not exclude loops so that $x \sim x$ is possible. If $x \sim y$ then $\overline{xy}$ denotes the edge connecting $x$ and $y$. The distance $d(x, y)$ is the minimal number of edges in any edge path connecting $x$ and $y$.

Assume that graph $\Gamma$ is endowed by a *weight* $\mu_{xy}$ which is a symmetric non-negative function on $\Gamma \times \Gamma$ such that $\mu_{xy} > 0$ if and only if $x \sim y$. Given $\mu_{xy}$, we define also a measure $\mu$ on vertices by

$$\mu(x) := \sum_{y \sim x} \mu_{xy}$$

and

$$\mu(A) := \sum_{x \in A} \mu(x),$$

for any finite set $A \subset \Gamma$. The couple $(\Gamma, \mu)$ is called *a weighted graph*. Here $\mu$ refers both to the weight $\mu_{xy}$ and to the measure $\mu$.

Any graph $\Gamma$ admits a *standard weight* which is defined by $\mu_{xy} = 1$ for all edges $\overline{xy}$. For such a weight, $\mu(x)$ is equal to the degree of the vertex $x$, that is the number of its neighbors.

Any weighted graph has a natural *Markov operator* $P(x, y)$ defined by

$$P(x, y) := \frac{\mu_{xy}}{\mu(x)}. \tag{2.1}$$

Clearly, we have

$$\sum_{y \in \Gamma} P(x, y) = 1 \tag{2.2}$$

and

$$P(x, y)\mu(x) = P(y, x)\mu(y). \tag{2.3}$$

For the Markov operator $P$, there is an associated random walk $X_n$, jumping at each time $n \in \mathbb{N}$ from a current vertex $x$ to a neighboring vertex $y$ with probability $P(x, y)$. The process $X_n$ is Markov and reversible with respect to measure $\mu$. If $\mu$ is the standard weight on $\Gamma$ then $X_n$ is called *a simple random walk* on $\Gamma$.

Conversely, given a countable set $\Gamma$ with a measure $\mu$ and a Markov operator $P(x, y)$ on $\Gamma$ satisfying (2.3), the identity (2.1) uniquely determines a symmetric weight $\mu_{xy}$ on $\Gamma \times \Gamma$. Then one defines edges $\overline{xy}$ as those pairs of vertices for which $\mu_{xy} \neq 0$, and obtains a weighted graph $(\Gamma, \mu)$. One has to assume in addition that the resulting graph $\Gamma$ is connected and locally finite.

Let $P_n$ denote the $n$-th convolution power of the operator $P$. Alternatively, $P_n(x, y)$ is the transition function of the random walk $X_n$, i.e.

$$P_n(x, y) = \mathbb{P}_x(X_n = y).$$

Define also the transition density of $X_n$, or *the heat kernel*, by

$$p_n(x, y) := \frac{P_n(x, y)}{\mu(y)}.$$

As obviously follows from (2.3), $p_n(x, y) = p_n(y, x)$.

The only a priori assumption which we normally make about the transition probability is the following:

$$P(x,y) \geq p_0, \quad \forall x \sim y, \tag{$p_0$}$$

where $p_0$ is a positive constant. Due to (2.2), the hypothesis $(p_0)$ implies that the degree of each vertex $x \in \Gamma$ is uniformly bounded from above. The latter is in fact equivalent to $(p_0)$, provided $X_n$ is a simple random walk.

By *sub-Gaussian* heat kernel estimates on graphs we will mean the following inequalities:

$$p_n(x,y) \leq C n^{-\alpha/\beta} \exp\left(-\left(\frac{d(x,y)^\beta}{Cn}\right)^{\frac{1}{\beta-1}}\right) \tag{$UE$}$$

and

$$p_n(x,y) + p_{n+1}(x,y) \geq c n^{-\alpha/\beta} \exp\left(-\left(\frac{d(x,y)^\beta}{cn}\right)^{\frac{1}{\beta-1}}\right), \quad n \geq d(x,y), \tag{$LE$}$$

where $x, y$ are arbitrary points on $\Gamma$ and $n$ is a positive integer.

Let us comment on the differences between $(UE)$ and $(LE)$. First observe that $p_n(x,y) = 0$ whenever $n < d(x,y)$ (indeed, the random walk cannot get from $x$ to $y$ in a number of steps smaller than $d(x,y)$). Therefore, the restriction $n \geq d(x,y)$ in $(LE)$ is necessary. We could assume the same restriction in $(UE)$ but if $p_n(x,y) = 0$ then $(UE)$ is true anyway. Another difference - using $p_n + p_{n+1}$ in $(LE)$ in place of $p_n$ in $(UE)$ - is due to the parity problem. Indeed, if the graph $\Gamma$ is bipartite (for example, $\mathbb{Z}^D$) then $p_n(x,y) = 0$ whenever $n$ and $d(x,y)$ have different parities. Therefore, the lower bound for $p_n$ cannot hold in general, and we state it for $p_n + p_{n+1}$ instead. Alternatively, one could say that the lower bound holds either for $p_n$ or for $p_{n+1}$. The structure of the graph may cause one of $p_n$, $p_{n+1}$ to be small (or even vanish) but it is not possible to decide a priori which of these two terms admits the lower bound (see Section 14 for more details).

Denote by $B(x,R)$ a ball on $\Gamma$ of radius $R$ centered at $x$, and by $V(x,R)$ its measure, that is

$$B(x,R) := \{y \in \Gamma : d(x,y) < R\}, \quad V(x,R) := \mu(B(x,R)).$$

We say that the graph $(\Gamma, \mu)$ has *the regular volume growth* of degree $\alpha$ if

$$V(x,R) \simeq R^\alpha, \quad \forall x \in \Gamma,\ R \geq 1. \tag{$V$}$$

The *Green kernel* of $(\Gamma, \mu)$ is defined by

$$g(x,y) := \sum_{n=0}^{\infty} p_n(x,y).$$

Assuming that $\alpha > \beta$, the estimates $(UE)$ and $(LE)$ imply, upon summation in $n$,

$$g(x,y) \simeq d(x,y)^{-\gamma}, \quad \forall x \neq y \tag{$G$}$$

where $\gamma = \alpha - \beta$. It turns out that $(G)$ together with the volume growth condition $(V)$ is sufficient to recover the heat kernel estimates $(UE)$ and $(LE)$, as is stated in the following main theorem.

**Theorem 2.1** *Let $\alpha > \beta > 1$ and $\gamma = \alpha - \beta$. For any infinite connected weighted graph $(\Gamma, \mu)$ satisfying $(p_0)$, the following equivalence holds*

$$(V) + (G) \Longleftrightarrow (UE) + (LE).$$

**Remark 2.1** Under the hypotheses $(V)$ and $(G)$, some partial heat kernel estimates were obtained by the second author [60].

It is well known that a simple random walk in $\mathbb{Z}^D$ admits the Gaussian estimate

$$cn^{-D/2}\exp\left(-\frac{d^2(x,y)}{cn}\right)\le p_n(x,y)\le Cn^{-D/2}\exp\left(-\frac{d^2(x,y)}{Cn}\right), \tag{2.4}$$

the lower bound being subject to the restrictions $n\equiv d(x,y)(\operatorname{mod}2)$ and $d(x,y)\le n$. Similar Gaussian estimates were proved also for more general graphs, under various assumptions (see [36], [52], [22], [27]). It is easy to see that (2.4) is equivalent to $(UE)+(LE)$ for $\alpha=D$ and $\beta=2$ (see Section 14 for the parity matters).

Barlow and Bass [8] constructed a family of graphs - graphical Sierpinski carpets (resembling in the large scale the multi-dimensional Sierpinski carpet), which are characterized by the two parameter $\alpha$ and $\beta$, and heat kernels on those graphs satisfy the sub-Gaussian estimates $(UE)$ and $(LE)$. In general, the parameters $\alpha$ and $\beta$, for which $(UE)$ and $(LE)$ may take place, must satisfy the following inequalities

$$2\le\beta\le\alpha+1, \tag{2.5}$$

which can be seen as follows. By [9, Theorem 2.1], the lower bound in $(V)$ implies the on-diagonal upper bound $p_n(x,x)\le Cn^{-\alpha/(\alpha+1)}$. By the result of [47], the upper bound in $(V)$ implies the on-diagonal lower bound $p_n(x,x)\ge c\,(n\log n)^{-\alpha/2}$. Comparing these estimates with the on-diagonal lower and upper bounds implied by $(LE)$ and $(UE)$, we obtain (2.5) (cf. [4, Theorem 3.20 and Remark 3.22], [57] as well as Lemma 5.4 below).

The sub-Gaussian estimates for different $\alpha$ and $\beta$ are related as follows. Consider the right hand side of $(UE)$ and $(LE)$ as a function of $\alpha$ and $\beta$. It is easy to see that it decreases as $\beta$ and $\alpha/\beta$ simultaneously increase (assuming $d(x,y)\ge n$). In particular, $(UE)$ gets stronger (and $(LE)$ gets weaker) on increasing of $\alpha$ (while keeping the same $\beta$) whereas in general there is no monotonicity in $\beta$.

The estimates $(UE)$ and $(LE)$ were proved by Jones [38] for the graphical Sierpinski gasket. The latter is a graph which is obtained from an equilateral triangle by a fractal-like construction (see Fig. 1). The reason for a subdiffusive behaviour of the random walk on such graphs is that they contain plenty of "holes" of all sizes, which causes the random walk to spend more time on circumventing the obstacles rather than on moving away from the origin.

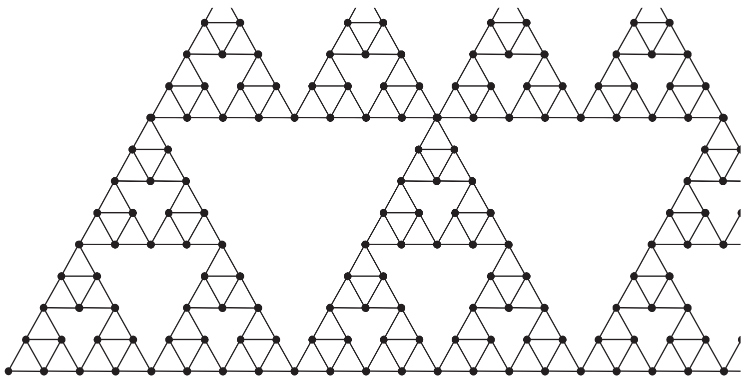

**Figure 1** A fragment of the graphical Sierpinski gasket

It is possible to show that $(V)$ and $(G)$ imply $\beta\ge2$ (see Lemma 5.4). The assumption $\alpha>\beta$ is necessary to ensure the finiteness of the Green function. It is known that either $g(x,y)$ is finite for all $x, y$ or $g \equiv \infty$. In the first case the graph $(\Gamma, \mu)$ is called *transient* and the second case - *recurrent* (for example, $\mathbb{Z}^D$ is transient if $D \geq 3$ and recurrent otherwise). Hence, Theorem 2.1 serves only transient graphs.

The question of finding equivalent conditions for the sub-Gaussian estimates $(UE)$ and $(LE)$ is equally interesting for recurrent graphs. By the way, the graph on Fig. 1 is recurrent[2]. Indeed, the volume function on this graph obviously admits the estimate

$$V(x, r) \leq Cr^2,$$

which implies the recurrence (see [18], [64]). Alternatively, one can see directly that $\alpha < \beta$ because the parameters $\alpha$ and $\beta$ for Sierpinski gasket are $\alpha = \frac{\log 3}{\log 2}$ and $\beta = \frac{\log 5}{\log 2}$ (see [4]).

Some hints on the recurrent case are given below in Section 4.

## 3 Preliminaries

If $P$ is the Markov operator of a weighted graph $(\Gamma, \mu)$ and $I$ is the identity operator then $\Delta := P - I$ is called the *Laplace operator* of $(\Gamma, \mu)$. For any set $A \subset \Gamma$, denote by $\overline{A}$ the set containing all vertices of $A$ and all their neighbors. If a function $f$ is defined on $\overline{A}$ then $\Delta f$ is defined on $A$ and

$$\Delta f(x) = \sum_{y \sim x} P(x, y) f(y) - f(x) = \frac{1}{\mu(x)} \sum_{y \in \Gamma} (\nabla_{xy} f) \, \mu_{xy} \, , \tag{3.1}$$

where

$$\nabla_{xy} f := f(y) - f(x).$$

Note that although the summation in the second sum in (3.1) runs over all vertices $y$, the summand is non-vanishing only if $y \sim x$.

The following is a discrete analogue of the Green formula: for any finite set $A$ and for all functions $f$ and $g$ defined on $\overline{A}$,

$$\sum_{x \in A} \Delta f(x) g(x) \mu(x) = \sum_{x \in A, y \notin A} (\nabla_{xy} f) \, g(x) \mu_{xy} - \frac{1}{2} \sum_{x, y \in A} (\nabla_{xy} f) \, (\nabla_{xy} g) \, \mu_{xy} \, . \tag{3.2}$$

We say that a function $v$ is *harmonic* in set $A$ if $v$ is defined in $\overline{A}$ and $\Delta v = 0$ in $A$. Similarly, we say that a function $v$ is *superharmonic* if $\Delta v \leq 0$. Observe that the inequality $\Delta v \leq 0$ is equivalent to

$$v(x) \geq \sum_{y \sim x} P(x, y) v(y).$$

The latter implies, in particular, that the infimum of a family of superharmonic functions is again superharmonic.

For any non-empty set $A \subset \Gamma$, let $c_0(A)$ be the set of functions on $\Gamma$ whose support is finite and is in $A$. Denote by $\Delta_A$ the Laplace operator with the vanishing Dirichlet boundary condition on $A$, i.e.

$$\Delta_A f(x) := \begin{cases} \Delta f, & x \in A, \\ 0, & x \notin A. \end{cases}$$

[2] Plenty of examples of transient graphs and fractals with sub-Gaussian heat kernel bounds can be found in [4], [7], [8].

The operator $\Delta_A$ is symmetric with respect to the measure $\mu$ and is non-positive definite. Moreover, it is essentially self-adjoint in $L^2(A,\mu)$.

For a finite set $A$, denote by $|A|$ its cardinality. If $A$ is finite and non-empty then the operator $-\Delta_A$ has $|A|$ non-negative eigenvalues which we enumerate in the increasing order and denote as follows:

$$\lambda_1(A) \le \lambda_2(A) \le \dots \le \lambda_{|A|}(A).$$

It is known that all eigenvalues $\lambda_i(A)$ lie in the interval $[0,2]$ and that $\lambda_1(A) \in [0,1]$ (see for example [19], [22, Section 3.3]). The smallest eigenvalue $\lambda_1(A)$ admits the variational definition

$$\lambda_1(A) = \inf_{f \in c_0(A)} \frac{-(\Delta f, f)}{(f,f)} = \inf_{f \in c_0(A)} \frac{\frac{1}{2}\sum_{x \sim y} (\nabla_{xy} f)^2 \mu_{xy}}{\sum_x f^2(x)\mu(x)}, \tag{3.3}$$

where

$$(f,g) := \sum_{x \in \Gamma} f(x) g(x) \mu(x).$$

If $A = B(x,R)$ then we write for simplicity

$$\lambda(x,R) := \lambda_1(B(x,R)).$$

Given a non-empty set $A \subset \Gamma$, let $X_n^A$ be the random walk on $(\Gamma,\mu)$ with the killing condition outside $A$. Its Markov operator $P^A(x,y)$ is defined by

$$P^A(x,y) := \begin{cases} P(x,y), & x,y \in A, \\ 0, & \text{otherwise.} \end{cases}$$

The transition function $P_n^A(x,y)$ of $X_n^A$ is defined inductively: $P_0^A(x,y) = \delta_{xy}$ and

$$P_{n+1}^A(x,y) = \sum_{z \in \Gamma} P_n^A(x,z) P^A(z,y) = \sum_{z \in \Gamma} P^A(x,z) P_n^A(z,y). \tag{3.4}$$

As easily follows from (3.4), the function $u_n(x) = P_n^A(x,y)$ satisfies in $A \times \mathbb{N}$ the discrete heat equation

$$u_{n+1} - u_n = \Delta_A u_n\,. \tag{3.5}$$

The heat kernel $p_n^A(x,y)$ of $X_n^A$ is defined by

$$p_n^A(x,y) := \frac{P_n^A(x,y)}{\mu(y)}.$$

As follows from (2.1), $p^A$ is symmetric in $x$ and $y$. In particular, the kernel $p_n^A(x,y)$ satisfies the heat equation (3.5) both in $(n,x)$ and $(n,y)$. If $f(x)$ is a function on $A$ then the function

$$u_n(x) := P_n^A f(x) = \sum_{y \in A} p_n^A(x,y) f(y) \mu(y)$$

solves in $A \times \mathbb{N}$ the heat equation (3.5) with the initial data $u_0 = f$ and the boundary data $u_n(x) = 0$ if $x \notin A$.

The Green function of $X_n^A$ is defined by

$$G_A(x,y) := \sum_{n=0}^{\infty} P_n^A(x,y).$$

The alternative definition is that the function $G_A(x, y)$ is the infimum of all positive fundamental solutions of the Laplace equation in $A$. If the Green function is finite then, for any $y \in A$, we have $\Delta_A G_A(\cdot, y) = -\delta_y$. The opposite case, when $G_A(x, y) \equiv +\infty$, is equivalent to the recurrence of the process $X_n^A$.

The Green kernel $g_A(x, y)$ is defined by

$$g_A(x, y) = \frac{G_A(x, y)}{\mu(y)} = \sum_{n=0}^{\infty} p_n^A(x, y).$$

Clearly, the Green kernel is symmetric in $x$, $y$. Therefore, if $g_A$ is finite then $g_A$ is superharmonic in $A$ with respect to both $x$ and $y$, and is harmonic away from the diagonal $x = y$. Observe that if $\mu(x) \simeq 1$ (which in particular follows from $(V)$) then $G_A(x, y) \simeq g_A(x, y)$ and $p_n^A(x, y) \simeq P_n^A(x, y)$.

It is easy to see that the kernels $p_n^A(x, y)$ and $g_A(x, y)$ increase on enlarging of $A$ and tend to the global kernels $p_n(x, y)$ and $g(x, y)$ (defined in Section 2) as an increasing sequence of sets $A$ exhausts $\Gamma$.

If $A$ is finite and non-empty then it makes sense to consider the Dirichlet problem in $A$

$$\begin{cases} \Delta u = f & \text{in } A, \\ u = h & \text{in } \overline{A} \setminus A, \end{cases} \tag{3.6}$$

where $f$ and $h$ are given function on $A$ and $\overline{A} \setminus A$ respectively. As follows easily from the maximum principle, the solution $u$ exists and is unique. For a finite set $A$, $c_0(A)$ is identified with all functions on $A$ extended by 0 outside $A$. Then the equation

$$\Delta_A u = f,$$

where $u$ and $f$ are in $c_0(A)$, is equivalent to the Dirichlet problem (3.6) with $h = 0$. Its solution is given by means of the Green operator $G_A$ as follows:

$$u(x) = -G_A f(x) = -\sum_y G_A(x, y) f(y). \tag{3.7}$$

In other words, we have $G_A = (-\Delta_A)^{-1}$.

For any set $A \subset \Gamma$ and a point $x \in \Gamma$, define *the mean exit time* $E_A(x)$ by

$$E_A(x) := \sum_{y \in A} G_A(x, y). \tag{3.8}$$

As follows from the above discussion, the function $E_A(x)$ solves the following boundary value problem in $A$:

$$\begin{cases} \Delta u = -1 & \text{in } A, \\ u = 0 & \text{outside } A. \end{cases} \tag{3.9}$$

Denote by $T_A$ the first exit time from set $A$ for the process $X_n$, that is,

$$T_A := \min\{k : X_k \notin A\}.$$

We claim that $E_A(x) = \mathbb{E}_x(T_A)$, which justifies the term "mean exit time" for $E_A$. Indeed, $T_A$ coincides with the cardinality of all $n = 0, 1, 2, ...$ for which $X_n^A$ is in $A$, that is,

$$T_A = \sum_{n=0}^{\infty} \mathbf{1}_{\{X_n^A \in A\}},$$

whence

$$\mathbb{E}_x(T_A) = \sum_{n=0}^{\infty} \mathbb{P}_x\left(X_n^A \in A\right) = \sum_{n=0}^{\infty}\sum_{y\in A} P_n^A(x,y) = \sum_{y\in A} G_A(x,y) = E_A(x).$$

If $A = B(x,R)$ then we use a shorter notation

$$E(x,R) := E_{B(x,R)}(x).$$

Another function associated with the exit time is *the exit probability* defined by

$$\Psi_n^A(x) := \mathbb{P}_x\{X_k \notin A \text{ for some } k \le n\} = \mathbb{P}_x\{T_A \le n\}. \tag{3.10}$$

In other words, $\Psi_n^A(x)$ is the probability that the random walk $X_k$ started at $x$ will at least once exit $A$ by time $n$. Alternatively, $\Psi_n^A(x)$ can be defined as the solution $u_n(x)$ to the following initial boundary value problem in $A \times \mathbb{N}$

$$\begin{cases} u_{n+1} - u_n = \Delta u_n, & \\ u_0(x) = 0, & x \in A, \\ u_n(x) = 1, & x \notin A \text{ and } n \ge 0. \end{cases} \tag{3.11}$$

If $A = B(x,R)$ then we will use the shorter notation

$$\Psi_n(x,R) := \Psi_n^{B(x,R)}(x).$$

In conclusion of this section, we prove two useful consequences of the condition $(p_0)$:

$$P(x,y) \ge p_0, \quad \forall x \sim y. \tag{$p_0$}$$

**Proposition 3.1** *If $(p_0)$ holds then, for all $x \in \Gamma$ and $R > 0$ and for some $C = C(p_0)$,*

$$V(x,R) \le C^R \mu(x). \tag{3.12}$$

**Remark 3.1** Inequality (3.12) implies that, for a bounded range of $R$, $V(x,R) \simeq \mu(x)$.

**Proof.** Let $x \sim y$. Since $P(x,y) = \frac{\mu_{xy}}{\mu(x)}$ and $\mu_{xy} \le \mu(y)$, the hypothesis $(p_0)$ implies $p_0\mu(x) \le \mu(y)$. Similarly, $p_0\mu(y) \le \mu(x)$. Iterating these inequalities, we obtain, for arbitrary $x$ and $y$,

$$p_0^{d(x,y)}\mu(y) \le \mu(x). \tag{3.13}$$

Another consequence of $(p_0)$ is that any point $x$ has at most $p_0^{-1}$ neighbors. Therefore, any ball $B(x,R)$ has at most $C^R$ vertices inside. By (3.13), any point $y \in B(x,R)$ has measure at most $p_0^{-R}\mu(x)$, whence (3.12) follows. ■

**Proposition 3.2** *Assume that the hypothesis $(p_0)$ holds on $(\Gamma,\mu)$. Let function $v$ be non-negative in $\overline{A}$ and superharmonic in $A$. Then, for all points $x, y \in A$, such that $x \sim y$, we have $v(x) \simeq v(y)$.*

**Proof.** Indeed, the superharmonicity of $v$ implies

$$v(x) \ge \sum_{z\sim x} P(x,z)v(z) \ge P(x,y)v(y),$$

whence $v(x) \ge p_0 v(y)$ by $(p_0)$. In the same way, $v(y) \ge p_0 v(x)$ whence the claim follows. ■

# 4 Outline of the proof and its consequences

The proof of Theorem 2.1 consists of many steps. Here we describe the logical order of these steps. The rest of the paper is arranged in the way that each section treats a certain topic corresponding to one or more steps in the proof of Theorem 2.1.

Apart from the conditions $(V)$, $(G)$, $(UE)$ and $(UE)$ described in Section 2, we introduce here some more lettered conditions which will be widely used in the proof.

We say that *the Faber-Krahn inequality* holds on $(\Gamma, \mu)$ if, for some positive exponent $\nu$,

$$\lambda_1(A) \geq c\mu(A)^{-1/\nu}, \tag{$FK$}$$

for all non-empty finite sets $A \subset \Gamma$. In particular, $(FK)$ holds in $\mathbb{Z}^D$ with $\nu = D/2$. If $\Gamma$ is infinite and connected and if $\mu$ is the standard weight on $\Gamma$ then $(FK)$ automatically holds with $\nu = 1/2$ (see [9, Prop. 2.5]). We will be interested in $(FK)$ with $\nu = \alpha/\beta$ where $\alpha$ and $\beta$ are the parameters from $(UE)$ and $(LE)$, in which case we have $\nu > 1$.

An easy consequence of $(UE)$ is the *diagonal upper estimate*

$$p_n(x, x) \leq Cn^{-\alpha/\beta}, \tag{$DUE$}$$

for all $x \in \Gamma$ and $n \geq 1$.

Consider the following estimates for the mean exit time and the exit probability:

$$E(x, R) \simeq R^{\beta} \tag{$E$}$$

for all $x \in \Gamma$, $R \geq 1$, and

$$\Psi_n(x, R) \leq C \exp\left(-\left(\frac{R^{\beta}}{Cn}\right)^{\frac{1}{\beta-1}}\right), \tag{$\Psi$}$$

for all $x \in \Gamma$, $R > 0$ and $n \geq 1$. For example, $(E)$ and $(\Psi)$ hold in $\mathbb{Z}^D$ with $\beta = 2$.

The part $(V)+(G) \Longrightarrow (UE)$ of Theorem 2.1 is proved by the following chain of implications.

$$\begin{array}{c}
\boxed{(V) \quad + \quad (G)} \\
\underbrace{\begin{array}{ll}
\Downarrow_{Prop.5.5} & \Downarrow_{Prop.6.3} \\
(FK) & (E) \\
\Downarrow_{Prop.5.1} & \Downarrow_{Prop.7.1} \\
(DUE) & (\Psi)
\end{array}} \\
\Downarrow_{Prop.8.1} \\
\boxed{(UE)}
\end{array}$$

The relations between the exponents $\alpha, \beta, \gamma$ and $\nu$ involved in all conditions are as follows:

$$\alpha - \beta = \gamma \quad \text{and} \quad \alpha/\beta = \nu.$$

Given $(DUE)$ and $(\Psi)$, one obtains easily the full upper bound $(UE)$ using the approach of Barlow and Bass [6] (see Section 8). The method of obtaining the Faber-Krahn inequality $(FK)$ from $(V)$ and $(G)$ is based on ideas of Carron [14]. The implication $(FK) \Longrightarrow (DUE)$ is a discrete modification of the approach of the first author [31]. The implication $(V)+(G) \Longrightarrow (E)$ was originally proved by the second author [57], and here we give a simpler proof for that.

The crucial part of the proof of the upper estimate $(UE)$ is the implication $(E) \Longrightarrow (\Psi)$. The following nearly Gaussian estimate is true *always*, without assuming $(E)$ or anything else:

$$\Psi_n(x, R) \le C \frac{V(x, R)}{\mu(x)} \exp\left(-\frac{R^2}{Cn}\right) \tag{4.1}$$

(see [56] and [32, p.355]). However, (4.1) is not good enough for us even if neglecting the factor $V(x, R)$ in front of the exponential. Indeed, the range of $n$ for which we will apply $(\Psi)$, is $n > R$ (see the proof of Proposition 8.1). Assuming $\beta > 2$, we have in this range

$$\left(\frac{R^\beta}{n}\right)^{\frac{1}{\beta-1}} > \frac{R^2}{n},$$

so that $(\Psi)$ is stronger than (4.1).

We provide here an entirely new argument for $(E) \Longrightarrow (\Psi)$, which is based on investigation of solutions of the equation $\Delta v = \lambda v$. The function $v$ can be estimated by comparing it to $\Delta u = -1$ (and the latter is related to the mean exit time). On the other hand, the function $(1 + \lambda)^n v(x)$ satisfies the discrete heat equation and, hence, can be compared to $\Psi_n^A(x)$ by using the parabolic comparison principle (see Section 7 for details). Another proof of $(E) \Longrightarrow (\Psi)$ can be obtained by using the probabilistic method of Barlow and Bass [5], [6], [7].

Before we consider the proof of the lower bound $(LE)$, let us introduce the following conditions.

The *near-diagonal lower estimate*

$$p_n(x, y) + p_{n+1}(x, y) \ge c n^{-\alpha/\beta}, \quad \text{if} \quad d(x, y) \le \delta n^{1/\beta}, \tag{$NLE$}$$

for some positive constant $\delta$. Obviously, $(NLE)$ is equivalent to $(LE)$ in the range $d(x, y) \le \delta n^{1/\beta}$.

As an intermediate step, we will use the following *diagonal lower estimate* for the killed random walk:

$$p_{2n}^{B(x,R)}(x, x) \ge c n^{-\alpha/\beta}, \quad \text{if } n \le \varepsilon R^\beta, \tag{$DLE$}$$

for some positive constant $\varepsilon$.

We say that *the Harnack inequality* holds on $(\Gamma, \mu)$ if, for any ball $B(x, 2R) \subset \Gamma$ and for any non-negative function $u$ in $\overline{B(x, 2R)}$ which is harmonic in $B(x, 2R)$,

$$\max_{B(x,R)} u \le H \min_{B(x,R)} u\,, \tag{$H$}$$

for some constant $H \ge 1$. The Harnack inequality reflects certain homogeneity of the graph. For example, it holds for $\mathbb{Z}^D$ with the standard weight but fails on the connected sum of two copies of $\mathbb{Z}^D$ as well as on a binary tree.

The scheme of the proof of $(V) + (G) \Longrightarrow (LE)$ is shown on the diagram below. From the previous diagram, we know already that the conditions $(FK)$ and $(E)$ follow from $(V) + (G)$, as well as the implications $(FK) \Longrightarrow (DUE)$ and $(E) \Longrightarrow (\Psi)$.

$$\boxed{(V) \quad + \quad (G)}$$

$$\Downarrow_{Prop. 5.5, 6.3}$$

$$\underbrace{\overbrace{\begin{array}{ccc} (FK) & (E) & (G) \\ \Downarrow_{Prop. 5.1} & \Downarrow_{Prop. 7.1} & \Downarrow_{Prop. 10.1} \\ (DUE) & \underbrace{(\Psi) + (V)} & (H) \\ \Downarrow_{Prop. 12.3} & \Downarrow_{Prop. 9.1} & \Downarrow_{Prop. 11.2} \\ [deriv] & (DLE) + (E) & [osc] \end{array}}}$$

$$\Downarrow_{Prop. 13.1}$$

$$\underbrace{(NLE) + (V)}$$

$$\Downarrow_{Prop. 13.2}$$

$$\boxed{(LE)}$$

The central point in the diagram is Proposition 13.1 where $(NLE)$ is obtained from $(DUE)$, $(DLE)$, $(E)$ and $(H)$. The proof goes through the intermediate steps which are denoted here by $[osc]$ and $[deriv]$. The former refers to the oscillation inequality (11.7) obtained from $(H)$ in Propositions 11.1 and 11.2, and the latter refers to the upper estimate (12.5) for $|p_{n+2} - p_n|$ obtained from $(DUE)$ in Proposition 12.3.

The idea of obtaining $(NLE)$ by means of an elliptic Harnack inequality seems to have appeared independently in the papers by Auscher [2], [3] and Barlow – Bass [6], [7], [8] (see also a recent preprint by Hebisch and Saloff-Coste [37]). Basically, one views the heat equation for the heat kernel as an elliptic equation with a non-trivial right-hand side:

$$\Delta(p_n + p_{n+1}) = f, \quad \text{where } f = p_{n+2} - p_n.$$

The elliptic Harnack inequality and the upper bound for $E(x, r)$ allow to estimate the oscillation of $p_n + p_{n+1}$ via $f$ (in the continuous setting, this argument is classical and is due to Moser [48]). On the other hand, the on-diagonal upper bound for $p_n$ implies a suitable estimate for the discrete time derivative $p_{n+2} - p_n$ (the fact that estimates of the time derivatives of the heat kernel can be obtained from $(DUE)$ is well-known in the context of manifolds and goes back to Cheng – Li – Yau [17] and Davies [25], [26]; see also [33]). Having an upper bound for the oscillation of $p_n + p_{n+1}$ and the on-diagonal lower bound for $p_n + p_{n+1}$, one gets $(NLE)$. The chaining argument for the implication $(NLE) + (V) \Longrightarrow (LE)$ is nowadays standard and goes back to Aronson [1].

The method of obtaining $(DLE)$ from $(\Psi)$ and $(V)$ used in Proposition 9.1, is well known. Its various modifications can be found in [6], [11], [21], [23], [47], [54] and possibly in other places.

The claim that the Green kernel estimate $(G)$ implies the elliptic Harnack inequality $(H)$ would not surprise experts. In the context of the uniformly elliptic operators in $\mathbb{R}^D$, this was first observed by Landis [45, p.145-146] and then was elaborated by Krylov – Safonov [42] and Fabes – Stroock [28]. However, this claim becomes rather non-trivial for arbitrary graphs (and manifolds) because of topological difficulties. We provide here a new, simple and general proof of the implication $(G) \Longrightarrow (H)$, which is based on a potential theoretic approach of Boukricha [12].

Finally, the converse implication $(UE) + (LE) \Longrightarrow (V) + (G)$ is quite straightforward and is proved in Proposition 15.1.

As a consequence of the above diagrams, we see that the following equivalence takes places:

$$(FK)+(V)+(E)+(H) \Longleftrightarrow (UE)+(LE).$$

It is possible to show that this equivalence is true also for recurrent manifolds. Furthermore, the Faber-Krahn inequality $(FK)$ turns out to follow from $(V)+(E)+(H)$ so that

$$(V)+(E)+(H) \Longleftrightarrow (UE)+(LE). \tag{4.2}$$

The condition $(H)$ ensures here a necessary homogeneity of the graph whereas $(V)$ and $(E)$ provide the exponents $\alpha$ and $\beta$, respectively.

Another consequence of the proof is that

$$(V)+(UE)+(H) \Longleftrightarrow (UE)+(LE) \tag{4.3}$$

(see Remark 15.1). There is a number of conditions given in terms of capacities, eigenvalues etc., which can replace $(E)$ or $(UE)$ in (4.2) and (4.3), respectively. In the presence of $(V)$ and $(H)$, the purpose of the other condition is to recover the exponent $\beta$ in $(UE)$ and $(LE)$. Note that if $\beta = 2$ then $(UE)$ in (4.3) can be replaced by $(DUE)$ (cf. [37]).

The complete proofs of (4.2), (4.3) and other related statements will be given elsewhere.

# 5 The Faber-Krahn inequality and on-diagonal upper bounds

Recall that *a Faber-Krahn inequality* holds on $(\Gamma, \mu)$ if there are constants $c > 0$ and $\nu > 0$ such that, for all non-empty finite sets $A \subset \Gamma$,

$$\lambda_1(A) \geq c\mu(A)^{-1/\nu} \tag{$FK$}$$

We discuss here relationships between eigenvalues estimates like $(FK)$ and estimates of the Green kernel, heat kernel and volume growth. The outcome will be the following implications

$$(V)+(G) \Longrightarrow (FK) \Longrightarrow (DUE),$$

which are contained in Propositions 5.5 and 5.1, respectively, and which constitute a part of the proof of Theorem 2.1.

**Proposition 5.1** *Let $(\Gamma, \mu)$ satisfy $(p_0)$, and let $\nu$ be a positive number. Then the following conditions are equivalent:*

*(a) The Faber-Krahn inequality $(FK)$.*

*(b) The on-diagonal heat kernel upper bound, for all $x \in \Gamma$ and $n \geq 1$,*

$$p_n(x,x) \leq Cn^{-\nu}. \tag{$DUE$}$$

*(c) The estimate of the level sets of the Green kernel, for all $x \in \Gamma$ and $t > 0$,*

$$\mu\{y : g(x,y) > t\} \leq Ct^{-\frac{\nu}{\nu-1}} \tag{5.1}$$

*provided $\nu > 1$.*

The analogue of Proposition 5.1 for manifolds was proved by Carron [14]. The equivalence $(a) \Longleftrightarrow (b)$ was proved also in [31] for heat kernels on manifolds, and in [20, Proposition V.1] for random walks satisfying in addition the condition $\inf_x P(x,x) > 0$.

We will provide detailed proof only for the implications $(a) \Longrightarrow (b)$ and $(c) \Longrightarrow (a)$ which we use in this paper. The implication $(b) \Longrightarrow (c)$ can be proved in the following way. By theorem of Varopoulos [61], $(DUE)$ implies a Sobolev inequality. Then one applies argument of [14, Proposition 1.14] (adapted to the discrete setting) to show that (5.1) follows from the Sobolev inequality.

Note that our proof of $(a) \Longrightarrow (b)$ goes through for any $\nu > 0$. If $\nu > 1$ then one could apply the approach of [14] using a Sobolev inequality as an intermediate step between $(a)$ and $(b)$. In general, we use instead a Nash type inequality which will be obtained in the following lemma.

**Lemma 5.2** *Let $(\Gamma,\mu)$ be a weighted graph (which is not necessarily connected). Assume that, for any non-empty finite set $A \subset \Gamma$,*

$$\lambda_1(A) \geq \Lambda(\mu(A)), \tag{5.2}$$

*where $\Lambda(\cdot)$ is a non-negative non-increasing function on $(0,\infty)$. Let $f(x)$ be a non-negative function on $\Gamma$ with finite support. Denote*

$$\sum_{x\in\Gamma} f(x)\mu(x) = a \quad \textit{and} \quad \sum_{x\in\Gamma} f^2(x)\mu(x) = b.$$

*Then, for any $s > 0$,*

$$\frac{1}{2}\sum_{x\sim y}(\nabla_{xy}f)^2\mu_{xy} \geq (b - 2sa)\,\Lambda(a/s). \tag{5.3}$$

**Proof.** If $b - 2sa < 0$ then (5.3) trivially holds. So, we can assume in the sequel that

$$s \leq \frac{b}{2a}. \tag{5.4}$$

Since $b \leq a \max f$, (5.4) implies $s < \max f$ and, therefore, the following set

$$A_s = \{x \in \Gamma : f(x) > s\}.$$

is non-empty (see Fig. 2).

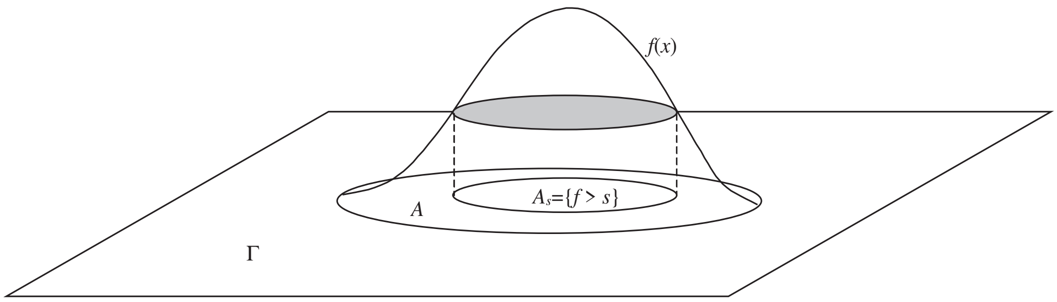


**Figure 2** Sets $A$ and $A_s$

Consider function $h = (f - s)_+$. This function belongs to $c_0(A_s)$ whence we obtain, by the variational property (3.3) of eigenvalues,

$$\frac{1}{2}\sum_{x\sim y}(\nabla_{xy}h)^2\mu_{xy} \geq \lambda_1(A_s)\sum_{x\in\Gamma} h^2(x)\mu(x). \tag{5.5}$$

Let us estimate all terms in (5.5) via $f$. We start with the obvious inequality

$$f^2 \le (f-s)_+^2 + 2sf \; = h^2 + 2sf,$$

which holds for any $s \ge 0$. It implies $h^2 \ge f^2 - 2sf$ whence

$$\sum_{x\in\Gamma} h^2(x)\mu(x) \ge b - 2sa. \tag{5.6}$$

The definition of $A_s$ implies $\mu(A_s) \le a/s$ whence, by (5.2)

$$\lambda_1(A_s) \ge \Lambda\left(\mu\left(A_s\right)\right) \ge \Lambda(a/s). \tag{5.7}$$

Clearly, we have also

$$\sum_{x\sim y}(\nabla_{xy}h)^2\mu_{xy} \le \sum_{x\sim y}(\nabla_{xy}f)^2\mu_{xy}.$$

Combining this with (5.7), (5.6) and (5.5), we obtain (5.3). ■

We will apply Lemma 5.2 for function $\Lambda(v) = cv^{-1/\nu}$. Choosing $s = \frac{b}{4a}$ in (5.3) we obtain

$$\frac{1}{2}\sum_{x\sim y}(\nabla_{xy}f)^2\mu_{xy} \ge c\, a^{-2/\nu}b^{1+1/\nu}. \tag{5.8}$$

This is a discrete version of the *Nash inequality* (cf. [49], [13]).

**Proof of $(a) \Longrightarrow (b)$ in Proposition 5.1.**

**STEP 1.** Let $f$ be a non-negative function on $\Gamma$ with finite support. Denote for simplicity

$$b = \sum_{x\in\Gamma} f^2(x)\mu(x) \quad \text{and} \quad b' = \sum_{x\in\Gamma}[Pf(x)]^2\mu(x),$$

where $P$ is the Markov operator of $(\Gamma,\mu)$. Then we have

$$b - b' = (f,f)_{L^2(\Gamma,\mu)} - (Pf,Pf)_{L^2(\Gamma,\mu)} = (f,(I-P_2)f)_{L^2(\Gamma,\mu)}.$$

Clearly $Q := P_2$ is also a Markov operator on $\Gamma$ reversible with respect to $\mu$, and it is associated with another structure of a weighted graph on the set $\Gamma$. Denote this weighted graph by $(\Gamma^*,\mu^*)$. As a set, $\Gamma^*$ coincides with $\Gamma$ and the measures $\mu$ and $\mu^*$ on vertices are the same. On the other hand, points $x, y$ are connected by an edge on $\Gamma^*$ if there is a path of length 2 from $x$ to $y$ in $\Gamma$, and the weight $\mu^*_{xy}$ on edges of $\Gamma^*$ is defined by

$$\mu^*_{xy} = Q(x,y)\mu(x).$$

Denote by $\Delta^*$ the Laplace operator of $(\Gamma^*,\mu^*)$. Then $\Delta^* = P_2 - I$ and, by the Green formula (3.2),

$$b - b' = -\sum_{x\in\Gamma} f(x)\,\Delta^* f(x)\mu(x) = \frac{1}{2}\sum_{x,y\in\Gamma}(\nabla_{xy}f)^2\mu^*_{xy}. \tag{5.9}$$

**STEP 2.** If $A$ is a non-empty finite subset of $\Gamma$ then [22, Lemma 4.3] says that[3]

$$\lambda_1^*(A) \ge \lambda_1(\overline{A}), \tag{5.10}$$

[3]The proof of (5.10) is based on the variational property (3.3) and on the fact that all eigenvalue of $-\Delta_A$ belong to the interval $[\lambda_1(A), 2-\lambda_1(A)]$.

where $\lambda_1^*(A)$ is the first eigenvalue of $-\Delta_A^*$. By the Faber-Krahn inequality $(FK)$ for the graph $(\Gamma, \mu)$, we obtain

$$\lambda_1^*(A) \geq c\mu(\overline{A})^{-1/\nu}. \tag{5.11}$$

Since $(p_0)$ and Proposition 3.1 imply

$$\mu(\overline{A}) \leq \sum_{x \in A} V(x, 2) \leq C \sum_{x \in A} \mu(x) = C\mu(A) = C\mu^*(A),$$

(5.11) yields $(FK)$ for the graph $(\Gamma^*, \mu^*)$.

**Remark 5.1** The only place where $(p_0)$ is used in the proof of $(a) \Longrightarrow (b)$ is to ensure that $\mu(\overline{A}) \leq C\mu(A)$. If this inequality holds for another reason then the rest of the proof goes in the same way.

**STEP 3.** For some fixed $y \in \Gamma$, denote $f_n(x) = p_n(x, y)$ and

$$b_n = \sum_{x \in \Gamma} f_n^2(x)\mu(x) = p_{2n}(y, y).$$

Then $f_{n+1} = Pf_n$ and we obtain by (5.9)

$$b_n - b_{n+1} = \frac{1}{2} \sum_{x,y \in \Gamma} (\nabla_{xy} f_n)^2 \mu_{xy}^*.$$

The graph $(\Gamma^*, \mu^*)$ satisfies $(FK)$ so that Lemma 5.2 can be applied. Since

$$\sum_{x \in \Gamma} f_n(x)\mu(x) = \sum_{x \in \Gamma} P_n(x, y) = 1\,,$$

(5.8) yields

$$\frac{1}{2} \sum_{x,y \in \Gamma} (\nabla_{xy} f_n)^2 \mu_{xy}^* \geq c\, b_n^{1+1/\nu}\,,$$

whence

$$b_n - b_{n+1} \geq c b_n^{1+1/\nu}. \tag{5.12}$$

In particular, we see that $b_n > b_{n+1}$.

Next we apply an elementary inequality

$$\nu(x - y) \geq \frac{x^\nu - y^\nu}{x^{\nu-1} + y^{\nu-1}}\,, \tag{5.13}$$

which is true for all $x > y > 0$ and $\nu > 0$. Taking $x = b_{n+1}^{-1/\nu}$ and $y = b_n^{-1/\nu}$, we obtain from (5.13) and (5.12)

$$\nu(b_{n+1}^{-1/\nu} - b_n^{-1/\nu}) \geq \frac{b_{n+1}^{-1} - b_n^{-1}}{b_{n+1}^{-(\nu-1)/\nu} + b_n^{-(\nu-1)/\nu}} = \frac{b_n - b_{n+1}}{b_{n+1}^{1/\nu} b_n + b_n^{1/\nu} b_{n+1}} \geq \frac{c b_n^{1+1/\nu}}{2 b_n^{1+1/\nu}} = \frac{c}{2},$$

whence

$$b_{n+1}^{-1/\nu} - b_n^{-1/\nu} \geq \frac{c}{2\nu} = \text{const}.$$

Summing up this inequality in $n$, we conclude $b_n^{-1/\nu} \geq cn$ and $b_n \leq Cn^{-\nu}$.

Since $b_n = p_{2n}(y,y)$, we have proved that, for all $y \in \Gamma$ and $n \geq 1$,

$$p_{2n}(y,y) \leq Cn^{-\nu}, \tag{5.14}$$

which is $(DUE)$ for all *even* times.

**STEP 4.** By the semigroup identity, we have, for any $0 < k < m$,

$$p_m(x,y) = \sum_{z \in \Gamma} p_{m-k}(x,z) p_k(z,y) \mu(z). \tag{5.15}$$

In particular, if $m = 2n$, $k = n$ and $y = x$ then

$$p_{2n}(x,x) = \sum_{z \in \Gamma} p_n^2(x,z) \mu(z). \tag{5.16}$$

On the other hand, (5.15), the Cauchy–Schwarz inequality and (5.16) imply

$$p_{2n}(x,y) = \sum_{z \in \Gamma} p_n(x,z) p_n(z,y) \mu(z) \leq \left[\sum_{z \in \Gamma} p_n^2(x,z) \mu(x)\right]^{\frac{1}{2}} \left[\sum_{z \in \Gamma} p_n^2(y,z) \mu(z)\right]^{\frac{1}{2}},$$

whence

$$p_{2n}(x,y) \leq p_{2n}(x,x)^{1/2} p_{2n}(y,y)^{1/2}. \tag{5.17}$$

Together with (5.14), this yields $p_{2n}(x,y) \leq Cn^{-\nu}$, for all $x, y \in \Gamma$. This implies $(DUE)$ also for *odd* times if we observe that, by (5.15) and (2.2),

$$p_{2n+1}(x,y) = \sum_{z \in \Gamma} p_{2n}(x,z) P(z,y) \leq \max_{z \in \Gamma} p_{2n}(x,z). \tag{5.18}$$

■

**Proof of $(c) \Rightarrow (a)$ in Proposition 5.1.** Let $A$ be a non-empty finite subset of $\Gamma$ and let $f \in c_0(A)$ be the first eigenfunction of $-\Delta_A$. We may assume that $f \geq 0$. Let us normalize $f$ so that $\max f = 1$ and let $x_0 \in A$ be the maximum point of $f$. The equation $-\Delta_A f = \lambda_1(A) f$ implies, by (3.7),

$$f(x) = \lambda_1(A) \sum_{y \in A} G_A(x,y) f(y)$$

whence, for $x = x_0$,

$$1 = \lambda_1(A) \sum_{y \in A} G_A(x_0,y) f(y) \leq \lambda_1(A) \sum_{y \in A} G_A(x_0,y)$$

and

$$\lambda_1(A) \geq \left(\max_{x \in A} \sum_{y \in A} G_A(x,y)\right)^{-1}. \tag{5.19}$$

On the other hand, for any $x \in A$,

$$\sum_{y \in A} G_A(x,y) = \sum_{y \in A} g_A(x,y) \mu(y) = \int_0^\infty \mu\left\{g_A(x,\cdot) > t\right\} dt.$$

Fix some $t_0 > 0$ and estimate the integral above using (5.1), $g_A \leq g$ and the fact that

$$\mu\left\{g_A(x,\cdot) > t\right\} \leq \mu(A).$$

Then we obtain

$$\sum_{y\in A} G_A(x,y) \le \int_0^{t_0} \mu(A)dt + \int_{t_0}^{\infty} Ct^{-\frac{\nu}{\nu-1}}\,dt = \mu(A)t_0 + Ct_0^{-\frac{1}{\nu-1}}.$$

Let us choose $t_0 \simeq \mu(A)^{-\frac{\nu-1}{\nu}}$ to equate the two terms on the right-hand side, whence

$$\sum_{y\in A} G_A(x,y) \le C\mu(A)^{1/\nu}. \tag{5.20}$$

Finally, (5.20) and (5.19) imply $(FK)$. ∎

The second result of this section will be preceded by two lemmas. We say that a weighted graph $(\Gamma,\mu)$ satisfies *the doubling volume condition* if

$$V(x,2R) \le CV(x,R), \quad \forall x\in\Gamma,\ R>0. \tag{$D$}$$

Clearly, $(D)$ is a weaker assumption than $(V)$.

**Lemma 5.3** *If $(\Gamma,\mu)$ satisfies $(D)$ then, for all $x\in\Gamma$ and $R>0$,*

$$\lambda(x,R) \le \frac{C}{R^2} \tag{5.21}$$

**Proof.** Let us apply the variational property (3.3) with the test function

$$f(y) = (R - d(x,y))_+ \in c_0(B(x,R)).$$

Since $|\nabla_{yz}f| \le 1$, (3.3) and $(D)$ imply

$$\lambda(x,R) \le \frac{\frac{1}{2}\sum_{y\sim z}(\nabla_{yz}f)^2\mu_{yz}}{\sum_y f^2(y)\mu(y)} \le \frac{CV(x,R)}{R^2V(x,R/2)} \le \frac{C'}{R^2},$$

which was to be proved. ∎

The next lemma was proved in [57] but we give here a shorter proof.

**Lemma 5.4** *Let $(\Gamma,\mu)$ satisfy $(p_0)$. If $(V)$ and $(G)$ hold, with some positive parameters $\alpha$ and $\gamma$, then*

$$\alpha - \gamma \ge 2\,.$$

**Proof.** By (5.19), we have

$$\lambda(x,R)^{-1} \le \max_{y\in B(x,R)} \sum_{z\in B(x,2R)} G(y,z). \tag{5.22}$$

By $(G)$ and Proposition 3.2, $G(y,y)$ is uniformly bounded from above. Using $(G)$ to estimate $G(y,z)$ for $y\ne z$ and $(V)$, we obtain

$$\begin{aligned}
\sum_{z\in B(y,2R)} G(y,z) &= G(y,y) + \sum_{i=-1}^{\lceil \log_2 R\rceil} \sum_{z\in B(y,2^{-i}R)\setminus B(y,2^{-i-1}R)} g(y,z)\mu(z) \\
&\le C + C\sum_{i=-1}^{\lceil \log_2 R\rceil} \left(2^{-i}R\right)^{-\gamma} V(y,2^{-i}R) \\
&\le C\left[1 + \sum_{i=-1}^{\lceil \log_2 R\rceil} \left(2^{-i}R\right)^{\alpha-\gamma}\right].
\end{aligned} \tag{5.23}$$

A straightforward computation of the sum (5.23) yields, for large $R$,

$$\sum_{z\in B(y,2R)} G(y,z) \leq C \begin{cases} R^{\alpha-\gamma}, & \alpha>\gamma, \\ \log_2 R, & \alpha=\gamma, \\ 1, & \alpha<\gamma. \end{cases} \tag{5.24}$$

Combining (5.22) and (5.24), we obtain

$$\lambda(x,R) \geq c \begin{cases} R^{-(\alpha-\gamma)}, & \alpha>\gamma, \\ (\log_2 R)^{-1}, & \alpha=\gamma, \\ 1, & \alpha<\gamma. \end{cases} \tag{5.25}$$

By Lemma 5.3, we have (5.21) which together (5.25) implies $\alpha-\gamma\geq 2$. ■

**Proposition 5.5** *Let $(\Gamma,\mu)$ satisfy $(p_0)$. If $(V)$ and $(G)$ hold, with some positive parameters $\alpha$ and $\gamma$, then the Faber-Krahn inequality $(FK)$ holds with the parameter $\nu=\frac{\alpha}{\alpha-\gamma}$.*

**Proof.** Note that, by Lemma 5.4, we have $\alpha>\gamma$ so that $\nu$ is positive and, moreover, $\nu>1$. Let us verify that

$$\mu\{y: g(x,y)>t\} \leq \operatorname{const} t^{-\alpha/\gamma}. \tag{5.26}$$

Then (5.1) would follow with $\nu=\frac{\alpha}{\alpha-\gamma}$, which implies $(FK)$, by Proposition 5.1.

The upper bound in $(G)$ and $(p_0)$ imply that, for all $x,y$ (including the case $x=y$ - see Proposition 3.2),

$$g(x,y) \leq C\min(1, d(x,y)^{-\gamma}). \tag{5.27}$$

If $t\geq C$ then the set $\{y: g(x,y)>t\}$ is empty, and (5.26) is trivially true.

Assume now $t\leq C$. Then (5.27) implies

$$\mu\{y: g(x,y)>t\} \leq \mu\{y: d(x,y)<(t/C)^{-1/\gamma}\} = V(x,(t/C)^{-1/\gamma}).$$

Since $R:=(t/C)^{-1/\gamma}\geq 1$, we can apply here the upper bound from $(V)$ and obtain (5.26). ■

# 6 The mean exit time and the Green kernel

The purpose of this section is to verify the part $(V)+(G)\Longrightarrow(E)$ of the proof of Theorem 2.1. Recall that $(E)$ stands for the condition

$$E(x,R)\simeq R^{\beta}, \quad \forall x\in\Gamma,\ R\geq 1. \tag{$E$}$$

Alongside with the mean exit time $E_A(x)$, consider and the *maximal* mean exit time $\overline{E}_A$ defined by

$$\overline{E}_A := \sup_y E_A(y). \tag{6.1}$$

If $A=B(x,R)$ then we write $\overline{E}(x,R):=\overline{E}_{B(x,R)}$. We will use also the following hypothesis:

$$\overline{E}(x,R)\leq CE(x,R), \quad \forall x\in\Gamma,\ R>0 \tag{$\overline{E}$}$$

**Proposition 6.1** *The upper bound in* $(E)$ *implies, for all* $x \in \Gamma$ *and* $R \geq 1$,

$$\overline{E}(x,R) \leq CR^{\beta}. \tag{6.2}$$

*The lower bound in* $(E)$ *implies*

$$\overline{E}(x,R) \geq cR^{\beta}. \tag{6.3}$$

*Consequently,* $(E)$ *implies* $(\overline{E})$ *and*

$$\overline{E}(x,R) \simeq R^{\beta}. \tag{6.4}$$

**Proof.** To show (6.2), let us observe that, for any point $y \in B(x,R)$, we have $B(x,R) \subset B(y,2R)$, whence

$$\overline{E}(x,R) = \sup_{y \in B(x,R)} E_{B(x,R)}(y) \leq \sup_{y \in B(x,R)} E_{B(y,2R)}(y) = \sup_{y \in B(x,R)} E(y,2R) \leq CR^{\beta}.$$

The lower bound (6.3) is obvious by $E \leq \overline{E}$. Finally, $(\overline{E})$ follows from $(E)$ and (6.4) if $R \geq 1$, and $(\overline{E})$ holds trivially if $R < 1$. ∎

**Proposition 6.2** *For any non-empty finite set* $A \subset \Gamma$, *we have*

$$\lambda_1(A) \geq (\overline{E}_A)^{-1}. \tag{6.5}$$

**Proof.** Indeed, this is a combination of (5.19) and definition of $\overline{E}$ (see (3.8) and (6.1)). ∎

The next statement was proved in [57].

**Proposition 6.3** *Let* $(\Gamma,\mu)$ *satisfy* $(p_0)$. *If* $(V)$ *and* $(G)$ *hold, with some positive parameters* $\alpha$ *and* $\gamma$, *then* $(E)$ *holds as well with* $\beta = \alpha - \gamma$.

**Proof.** Denote $A = B(x,R)$. Applying (3.8), the obvious inequality $g_A \leq g$, as well as $(V)$ and $(G)$, we obtain (cf. (5.23) and (5.24))

$$E(x,R) = \sum_{y \in A} g_A(x,y)\mu(y) \leq \sum_{y \in A} g(x,y)\mu(y) \leq CR^{\alpha-\gamma}.$$

Observe that, by Lemma 5.4, we know already that $\alpha > \gamma$.

For the lower bound of $E(x,R)$, let us prove that

$$g_A(x,y) \geq c\, d(x,y)^{-\gamma}, \quad \forall y \in B(x,\varepsilon R) \setminus \{x\} \tag{6.6}$$

provided $\varepsilon > 0$ is small enough. Consider the function

$$u(y) = g(x,y) - g_A(x,y)$$

which is harmonic in $A$. By the maximum principle, its maximum is attained at the boundary of $A$ whence, by $(G)$,

$$0 \leq u(y) \leq CR^{-\gamma}.$$

Therefore,

$$g_A(x,y) = g(x,y) - u(y) \geq c\, d(x,y)^{-\gamma} - CR^{-\gamma}. \tag{6.7}$$

If $R$ is large enough and if $d(x,y) \leq \varepsilon R$ with a small enough $\varepsilon$ then the second term in (6.7) is absorbed by the first one whence (6.6) follows.

Summing up (6.6) over $y$ we obtain (cf. (5.23) and (5.24))

$$E(x,R) = \sum_{y \in A} g_A(x,y)\mu(y) \geq \sum_{y \in B(x,\varepsilon R)\setminus\{x\}} g_A(x,y)\mu(y) \geq cR^{\alpha-\gamma}.$$

If $R$ is not big enough then the above argument does not work. However, in this case we argue as follows. If the random walk starts at $x$ then $T_{B(x,R)} \geq R$. Hence, we always have $E(x,R) = \mathbb{E}_x\left(T_{B(x,R)}\right) \geq R$ which yields the lower bound in $(E)$, provided $R \leq \text{const}$. ∎

Assuming that $(V)$ and $(E)$ hold, there are the following general relations between the exponents $\alpha$ and $\beta$: if the graph transient then $2 \leq \beta \leq \alpha$ and if it is recurrent then $2 \leq \beta \leq \alpha+1$ - see [57]. See also [58], [59] for various definitions of dimensions of graphs.

# 7 Sub-Gaussian term

The following statement is crucial for obtaining the off-diagonal upper bound of the heat kernel. It contains the part $(E) \Longrightarrow (\Psi)$ of the proof of Theorem 2.1.

**Proposition 7.1** *Assume that the graph $(\Gamma,\mu)$ possesses the property $(E)$. Then, for all $x \in \Gamma$, $R > 0$ and $n \geq 1$, we have*

$$\Psi_n(x,R) \leq C \exp\left(-\left(\frac{R^\beta}{Cn}\right)^{\frac{1}{\beta-1}}\right). \qquad (\Psi)$$

We start with the following lemma.

**Lemma 7.2** *Assume that the hypothesis $(\overline{E})$ holds on $(\Gamma,\mu)$. Let $A = B(x_0,r)$ be an arbitrary ball on $\Gamma$ and let $v$ be a function on $\overline{A}$ such that $0 \leq v \leq 1$. Suppose that $v$ satisfies in $A$ the equation*

$$\Delta v = \lambda v, \tag{7.1}$$

*where $\lambda$ is a constant such that*

$$\lambda \geq (\overline{E}_A)^{-1}. \tag{7.2}$$

*Then*

$$v(x_0) \leq 1 - \varepsilon\,, \tag{7.3}$$

*where $\varepsilon > 0$ depends on the constants in the condition $(\overline{E})$ (see Fig. 3).*

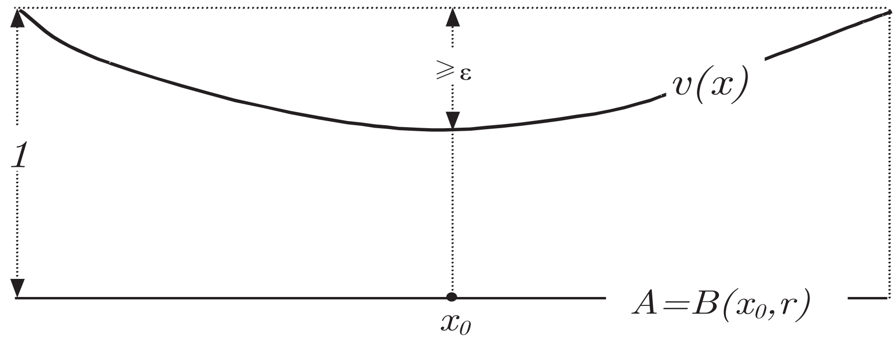


**Figure 3** The value of the function $v$ at the point $x_0$ does not exceed $1-\varepsilon$.

**Proof.** Denote for simplicity $u(x) = E_A(x)$ and recall that $u \in c_0(A)$ and $\Delta u = -1$ in $A$ (cf. (3.9)). Denote also

$$\lambda_0 := (\overline{E}_A)^{-1} = \frac{1}{\max u}.$$

Consider the function $w = 1 - \frac{\lambda_0}{2} u$. Then $\frac{1}{2} \leq w \leq 1$ and, in $A$,

$$\Delta w = \frac{\lambda_0}{2} \leq \lambda_0 w \leq \lambda w.$$

Since $v \leq 1$ and $w = 1$ outside $A$, the maximum principle for the operator $\Delta - \lambda$ implies that $v \leq w$ in $A$. In particular,

$$v(x_0) \leq w(x_0) = 1 - \frac{\lambda_0}{2} u(x_0) \leq 1 - \frac{u(x_0)}{2 \max u}.$$

The hypothesis $(\overline{E})$ yields

$$\frac{u(x_0)}{\max u} = \frac{E(x_0, r)}{\overline{E}(x_0, r)} \geq c,$$

whence (7.3) follows. ∎

**Lemma 7.3** *Assume that* $(\Gamma, \mu)$ *satisfies* $(E)$. *Let* $A = B(x_0, R)$ *be an arbitrary ball on* $\Gamma$, *and let* $v$ *be a function on* $\overline{A}$ *such that* $0 \leq v \leq 1$. *If* $v$ *satisfies in* $A$ *the equation (7.1) with a constant* $\lambda$ *such that*

$$CR^{-\beta} \leq \lambda < \overline{\lambda} \tag{7.4}$$

*then*

$$v(x_0) \leq \exp\left(-c\lambda^{1/\beta} R\right). \tag{7.5}$$

*Here* $\overline{\lambda}$ *is an arbitrary constant,* $C$ *is some constant depending on the condition* $(E)$, *and* $c > 0$ *is some constant depending on* $\overline{\lambda}$ *and on the condition* $(E)$.

**Proof.** The hypothesis $(E)$ implies $(\overline{E})$ and $\overline{E}(x, R) \simeq R^\beta$ (see Proposition 6.1). Choose the constant $C$ in (7.4) so big that the lower bound in (7.4) implies $\lambda \geq \overline{E}(x, R)^{-1}$. Then, by Lemma 7.2, we obtain $v(x_0) \leq 1 - \varepsilon$. If we have in addition

$$\lambda^{1/\beta} R \leq \text{const} \tag{7.6}$$

then (7.5) is trivially satisfied. In particular, if $R$ is in the bounded range then (7.6) is true because $\lambda$ is bounded from above by (7.4).

Hence, we may assume in the sequel that

$$R > C' \text{ and } \lambda > C'' R^{-\beta}, \tag{7.7}$$

with large enough constants $C'$ and $C''$ (in particular, $C'' \gg C$). The point of the present lemma is that it improves the previous one for this range of $R$ and $\lambda$. Choose a number $r$ from the equation $\lambda = Cr^{-\beta}$, where $C$ is the same constant as in (7.4). The above argument shows that Lemma 7.2 applies in any ball of radius $r$. Let $x_i$, $i \geq 1$, be a point in the ball $B(x_0, (r+1)i)$ in which $v$ takes the maximum value in this ball, and denote $m_i = v(x_i)$ (see Fig. 4). For $i = 0$, we set $m_0 = v(x_0)$.

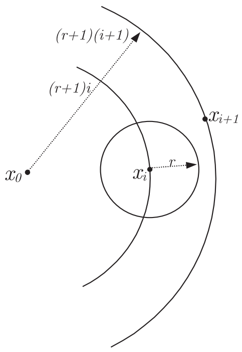


**Figure 4** The points $x_i$ where $v(x)$ takes the maximum values.

For each $i \geq 0$, consider the ball $A_i = B(x_i, r)$. Since

$$\overline{A}_i \subset B(x_i, r+1) \subset B(x_0, (r+1)(i+1)),$$

we have

$$\max_{\overline{A}_i} v \leq m_{i+1}.$$

Applying Lemma 7.2 to the function $v/m_{i+1}$ in the ball $A_i$, we obtain

$$m_i \leq (1-\varepsilon) m_{i+1}.$$

Iterating this inequality $k := \lfloor R/(r+1) \rfloor$ times and using $m_k \leq 1$, we conclude

$$v(x_0) = m_0 \leq (1-\varepsilon)^k. \tag{7.8}$$

By the conditions (7.7) and (7.4) and by the choice of $r$, we have

$$k \simeq \frac{R}{r} \simeq \lambda^{1/\beta} R,$$

so that (7.8) implies (7.5). ■

**Lemma 7.4** *Assume that $(\Gamma, \mu)$ satisfies $(E)$. Let $A = B(x_0, R)$ be an arbitrary ball in $\Gamma$ , and let $w_n(x)$ be a function in $\overline{A} \times \mathbb{N}$ such that $0 \leq w \leq 1$. Suppose that $w$ solves in $A \times \mathbb{N}$ the heat equation*

$$w_{n+1} - w_n = \Delta w_n \tag{7.9}$$

*with the initial data $w_0 \equiv 0$ in $A$ (see Fig. 5). Then, for all $n \geq 1$,*

$$w_n(x_0) \leq \exp\left(-c\left(\frac{R^\beta}{n}\right)^{\frac{1}{\beta-1}} + 1\right). \tag{7.10}$$

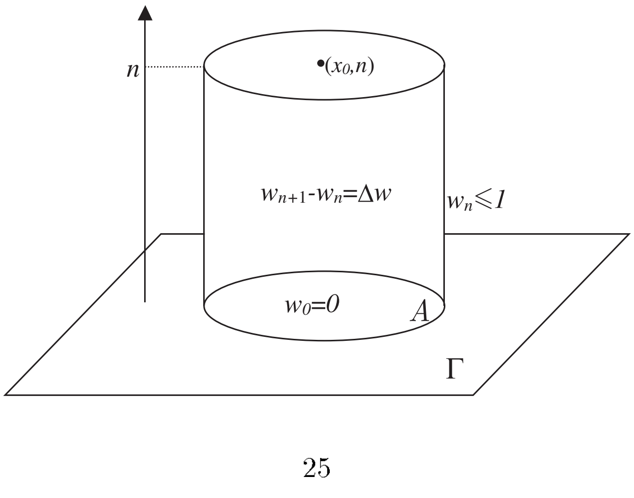

**Figure 5** The value of the function $w$ at the point $(x_0, n)$ is affected by the initial value $w = 0$ and by the boundary condition $w \leq 1$.

**Proof.** Consider first two trivial cases. If $R^\beta \leq Cn$ then (7.10) is true just by $w \leq 1$ provided $c$ is small enough. Since $\Delta w(x)$ depends only on the immediate neighbors of $x$, one gets by induction that $w_k(x) = 0$ for all $x \in B(x_0, R - k)$. Therefore, if $R > n$ then $w_n(x_0) = 0$, and (7.10) is true again.

Hence, we may assume in the sequel that, for a large enough $C$,

$$Cn^{1/\beta} < R \leq n. \tag{7.11}$$

Fix some $\lambda > 0$ and find a function $v(x)$ on $\overline{A}$ solving the boundary value problem

$$\begin{cases} \Delta v = \lambda v & \text{in } A, \\ v = 1 & \text{in } \overline{A} \setminus A. \end{cases}$$

The function $u_n(x) := (1 + \lambda)^n v(x)$ solves the heat equation (7.9) and satisfies the following boundary conditions: $u_n(x) \geq 1$ for $x \in \overline{A} \setminus A$ and $u_0(x) \geq 0$ for $x \in A$. By the parabolic comparison principle, we have $w \leq u$. Assume for a moment that $\lambda$ satisfies the hypothesis (7.4) of Lemma 7.3. Then we estimate $v(x_0)$ by (7.5) and obtain

$$w_n(x_0) \leq (1 + \lambda)^n v(x_0) \leq \exp\left(\lambda n - c\lambda^{1/\beta} R\right).$$

Now choose $\lambda$ from the condition $c\lambda^{1/\beta} R = 2\lambda n$, that is,

$$\lambda = \left(\frac{cR}{2n}\right)^{\frac{\beta}{\beta-1}}. \tag{7.12}$$

As follows from (7.11), this particular $\lambda$ satisfies (7.4). Therefore, the above application of Lemma 7.3 is justified, and we obtain

$$w_n(x_0) \leq \exp(-\lambda n) = \exp\left(-c'\left(\frac{R^\beta}{n}\right)^{\frac{1}{\beta-1}}\right),$$

finishing the proof. ∎

**Proof of Proposition 7.1.** Denote $A = B(x_0, R)$. By (3.11), the function $w_n(x) := \Psi_n^A(x)$ satisfies all the hypotheses of Lemma 7.4. Hence, $(\Psi)$ follows from (7.10). ∎

## 8 Off-diagonal upper bound of the heat kernel

Here we prove the following implication

$$(FK) + (E) \Longrightarrow (UE) \tag{8.1}$$

which will finish the proof of the heat kernel upper bound in Theorem 2.1. Indeed, together with the implications

$$(V) + (G) \overset{Prop.5.5}{\Longrightarrow} (FK)$$

and

$$(V) + (G) \overset{Prop.6.3}{\Longrightarrow} (E)$$

(8.1) yields the part $(V) + (G) \Longrightarrow (UE)$ of Theorem 2.1.

**Proposition 8.1** *On any graph* $(\Gamma, \mu)$*, we have*

$$(DUE) + (\Psi) \Longrightarrow (UE). \tag{8.2}$$

*In particular, if* $(p_0)$ *holds on* $(\Gamma, \mu)$ *then*

$$(FK) + (E) \Longrightarrow (UE) \tag{8.3}$$

**Proof.** By Proposition 5.1, $(p_0)$ and $(FK)$ imply the $(DUE)$. By Proposition 7.1, $(E)$ implies $(\Psi)$. Hence, the implication (8.3) is a consequence of (8.2).

To prove (8.2), let us fix some points $x, y \in \Gamma$ and denote $r = d(x, y)/2$. Since the balls $B(x, r)$ and $B(y, r)$ do not intersect, the semi-group identity (5.15) and the symmetry of the heat kernel imply, for any triple of non-negative integers $k$, $m$, $n$ such that $k + m = n$,

$$\begin{aligned} p_n(x, y) &\leq \sum_{z \notin B(x,r)} p_m(x, z) p_k(z, y) \mu(z) + \sum_{z \notin B(y,r)} p_m(x, z) p_k(z, y) \mu(z) \\ &\leq \sup_z p_k(z, y) \sum_{z \notin B(x,r)} P_m(x, z) + \sup_z p_m(x, z) \sum_{z \notin B(y,r)} P_k(y, z) \\ &= \sup_z p_k(y, z) \mathbb{P}_x\left(X_m \notin B(x, r)\right) + \sup_z p_m(x, z) \mathbb{P}_y\left(X_k \notin B(x, r)\right). \end{aligned}$$

As follows from the definition (3.10) of $\Psi$,

$$\mathbb{P}_x\left(X_m \notin B(x, r)\right) \leq \Psi_m(x, r).$$

Hence, we obtain the following general inequality, which is true for all reversible random walks:

$$p_n(x, y) \leq \sup_z p_k(y, z) \Psi_m(x, r) + \sup_z p_m(x, z) \Psi_k(y, r). \tag{8.4}$$

As follows from (5.17), the diagonal upper bound $(DUE)$ implies, for all $x, y \in \Gamma$,

$$p_n(x, y) \leq C n^{-\alpha/\beta}, \tag{8.5}$$

provided $n$ is even. Using inequality (5.18), we see that (8.5) holds also for odd $n$. Assuming $n \geq 2$, choosing $k \simeq m \simeq n/2$ and applying (8.5) and $(\Psi)$ to estimate the right-hand side of (8.4), we obtain $(UE)$. If $n = 1$ then $(UE)$ follows trivially from (8.5) and the fact that $p_n(x, y) = 0$ whenever $d(x, y) > n$. ■

# 9 On-diagonal lower bound

In this section, we prove the part $(\Psi) + (V) \Longrightarrow (DLE)$ of Theorem 2.1.

**Proposition 9.1** *Assume that the hypothesis* $(\Psi)$ *holds on* $(\Gamma, \mu)$*. For arbitrary* $x \in \Gamma$ *and* $R > 0$*, denote* $A = B(x, R)$*. Then the following on-diagonal lower bound is true*

$$p_{2n}^A(x, x) \geq \frac{c}{V(x, C n^{\frac{1}{\beta}})}, \tag{9.1}$$

*provided* $n \leq \varepsilon R^\beta$*, where* $\varepsilon$ *is a sufficiently small positive constant depending only on the constants from* $(\Psi)$*.*

*If in addition* $(V)$ *holds then*

$$p_{2n}^A(x, x) \geq c n^{-\alpha/\beta}, \quad \forall n \leq \varepsilon R^\beta. \tag{DLE}$$

**Remark 9.1** Since $p_{2n} \geq p_{2n}^A$ for any $A = B(x,R)$, inequality $(DLE)$ implies $p_{2n}(x,x) \geq cn^{-\alpha/\beta}$, for all positive integers $n$.

**Proof.** Let us fix some $r \in (0,R)$ and denote $B = B(x,r)$. Since $p^B \leq p^A$, it will suffice to prove (9.1) for $p^B$ instead of $p^A$, for some $r < R$. The semigroup identity (5.15) for $p^B$ and the Cauchy-Schwarz inequality imply

$$p_{2n}^B(x,x) = \sum_{z \in B} p_n^B(x,z)^2 \mu(z) \geq \frac{1}{\mu(B)} \left( \sum_{z \in B} p_n^B(x,z)\mu(z) \right)^2. \tag{9.2}$$

Let us observe that

$$\sum_{z \in B} p_n^B(\cdot,z)\mu(z) + \Psi_n^B(\cdot) = 1. \tag{9.3}$$

Indeed, the first term in (9.3) is the probability that the random walk $X_k$ stays in $B$ up to the time $k = n$ whereas $\Psi_n^B$ is the probability of the opposite event.

By the hypothesis $(\Psi)$, we have

$$\Psi_n^B(x) = \Psi_n(x,r) \leq C \exp\left( - \left( \frac{r^\beta}{Cn} \right)^{\frac{1}{\beta-1}} \right). \tag{9.4}$$

Choosing $r = Cn^{1/\beta}$ for large enough $C$ and assuming $n \leq \varepsilon R^\beta$ for sufficiently small $\varepsilon > 0$ (the latter ensures $r < R$) we obtain from (9.4) $\Psi_n(x,r) \leq \frac{1}{2}$ whence, by (9.3),

$$\sum_{z \in B} p_n^B(x,z)\mu(z) \geq \frac{1}{2}.$$

Therefore, (9.2) yields

$$p_{2n}^B(x,x) \geq \frac{1/4}{V(x,r)} = \frac{1/4}{V(x,Cn^{1/\beta})},$$

finishing the proof. ■

# 10 The Harnack inequality and the Green kernel

Recall that the weighted graph $(\Gamma,\mu)$ satisfies *the elliptic Harnack inequality* if, for all $x \in \Gamma, R > 0$ and for any non-negative function $u$ in $\overline{B(x,2R)}$ which is harmonic in $B(x,2R)$,

$$\max_{B(x,R)} u \leq H \min_{B(x,R)} u\,, \tag{$H$}$$

with some constant $H > 1$. In this section we establish that $(H)$ is implied by the condition $(G)$. Recall that the latter refers to

$$g(x,y) \simeq d(x,y)^{-\gamma}, \quad \forall x \neq y. \tag{$G$}$$

Consider the following *annulus* Harnack inequality for the Green kernel: for all $x \in \Gamma$ and $R > 1$,

$$\max_{y \in A(x,R)} g(x,y) \leq C \min_{y \in A(x,R)} g(x,y) \tag{$HG$}$$

where $A(x,R) := B(x,R) \setminus B(x,R/2)$.

**Proposition 10.1** *Assume that* $(p_0)$ *hold and the graph* $(\Gamma, \mu)$ *is transient. Then*

$$(G) \Longrightarrow (HG) \Longrightarrow (H).$$

Since the implication $(G) \Longrightarrow (HG)$ is obvious, we need to prove only the second implication. The main part of the proof is contained in the following lemma.

**Lemma 10.2** *Let* $U_0 \subset U_1 \subset U_2 \subset U_3$ *be a sequence of finite sets in* $\Gamma$ *such that* $\overline{U_i} \subset U_{i+1}$, $i = 0, 1, 2$. *Denote* $A = U_2 \setminus U_1$, $B = U_0$ *and* $U = U_3$. *Then, for any function* $u$ *which is non-negative in* $\overline{U_2}$ *and harmonic in* $U_2$, *we have*

$$\max_B u \le H \min_B u\,, \tag{10.1}$$

*where*

$$H := \max_{x \in B} \max_{y \in B} \max_{z \in A} \frac{G_U(y,z)}{G_U(x,z)} \tag{10.2}$$

*(see Fig. 6).*

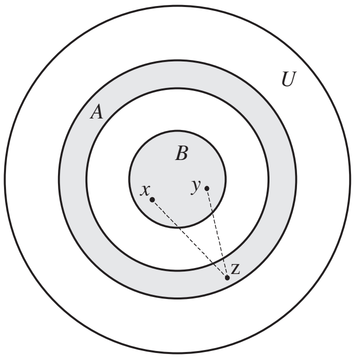


**Figure 6** The sets $B = U_0$, $A = U_2 \setminus U_1$ and $U = U_3$

**Remark 10.1** Note that no a priori assumption has been made about the graph $(\Gamma, \mu)$ (except for connectedness and unboundedness). If the graph is transient then, by exhausting $\Gamma$ by a sequence of finite sets $U$, we can replace $G_U$ in (10.2) by $G$. Note also that, without loss of generality, one can take $U_2 = \overline{U_1}$.

**Proof.** The following potential-theoretic argument is borrowed from [12]. We use the notation of Section 3. Given a non-negative harmonic function $u$ in $U_2$, denote by $S_u$ the following class of superharmonic functions:

$$S_u = \left\{ v : v \ge 0 \text{ in } \overline{U},\ \Delta v \le 0 \text{ in } U,\ \text{ and } v \ge u \text{ in } \overline{U_1} \right\}.$$

Define the function $w$ on $\overline{U}$ by

$$w(x) = \min \{ v(x) : v \in S_u \}\,. \tag{10.3}$$

Clearly, $w \in S_u$. Since the function $u$ itself is also in $S_u$, we have $w \le u$ in $\overline{U}$. On the other hand, by definition of $S_u$, $w \ge u$ in $\overline{U_1}$, whence we see that $u = w$ in $\overline{U_1}$ (see Fig. 7). In particular, it suffices to prove (10.1) for $w$ instead of $u$.

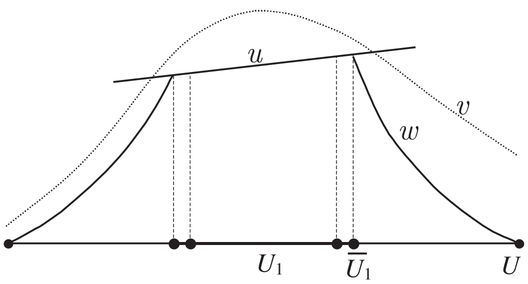


**Figure 7** The function $u$, a function $v \in S_u$ and the function $w = \min_{S_u} v$. The latter is harmonic in $U_1$ and in $U \setminus \overline{U_1}$.

Let us show that $w \in c_0(U)$. Indeed, let $v(x) = E_U(x)$. Then, by (3.9) and the strong minimum principle, $v$ is superharmonic and strictly positive in $U$. Hence, for a large enough constant $C$, we have $Cv \geq u$ in $\overline{U_1}$ whence $Cv \in S_u$ and $w \leq Cv$. Since $v = 0$ in $\overline{U} \setminus U$, this implies $w = 0$ in $\overline{U} \setminus U$ and $w \in c_0(U)$.

Denote $f := -\Delta w$ and observe that $f \geq 0$ in $U$. Since $w \in c_0(U)$, we have, for any $x \in U$,

$$w(x) = \sum_{z \in U} G_U(x,z) f(z). \tag{10.4}$$

Next we will prove that $f = 0$ outside $A$ so that the summation in (10.4) can be restricted to $z \in A$. Given that much, we obtain, for all $x, y \in B$,

$$\frac{w(y)}{w(x)} = \frac{\sum_{z \in A} G_U(y,z) f(z)}{\sum_{z \in A} G_U(x,z) f(z)} \leq H,$$

whence (10.1) follows.

We are left to verify that $w$ is harmonic in $U_1$ and outside $\overline{U_1}$. Indeed, if $x \in U_1$ then

$$\Delta w(x) = \Delta u(x) = 0,$$

because $w = u$ in $\overline{U_1}$. Let $\Delta w(x) \neq 0$ for some $x \in U \setminus \overline{U_1}$. Since $w$ is superharmonic, we have $\Delta w(x) < 0$ and

$$w(x) > Pw(x) = \sum_{y \sim x} P(x,y) w(y).$$

Consider the function $w'$ which is equal to $w$ everywhere in $\overline{U}$ except for the point $x$, and $w'$ at $x$ is defined to satisfy

$$w'(x) = \sum_{y \sim x} P(x,y) w'(y).$$

Clearly, $w'(x) < w(x)$, and $w'$ is superharmonic in $U$. Since $w' = w = u$ in $\overline{U_1}$, we have $w' \in S_u$. Hence, by the definition (10.3) of $w$, $w \leq w'$ in $\overline{U}$ which contradicts $w(x) > w'(x)$. $\blacksquare$

**Proof of Proposition 10.1.** Now we assume $(HG)$ and prove $(H)$. Given any ball $B(x_0, 2R)$ of radius $R > 4$ and a non-negative harmonic function $u$ in $B(x_0, 2R)$, define the sequence of radii $R_0 = R$, $R_1 = \frac{3}{2}R$ and $R_2 = 2R$ and denote $U_i = B(x_0, R_i)$ for $i = 0, 1, 2$ and $U_3 = \Gamma$. By Lemma 10.2, we have the inequality (10.1) which will imply $(H)$ provided we can show that the Harnack constant $H$ from (10.2) is bounded from above, uniformly in $x_0$ and $R$. Indeed, if $x, y \in B(x_0, R)$ and $z \in A = \overline{B(x_0, 2R)} \setminus B(x_0, \frac{3}{2}R)$ then both distances $d(z,x)$ and $d(z,y)$ are between $R/2$ and $7R/2$. By iterating $(HG)$ in the annuli centered at $z$, we obtain

$$\frac{G(y,z)}{G(x,z)} = \frac{g(z,y)}{g(z,x)} \leq \text{const},$$

whence we see that $H$ is indeed uniformly bounded from above.

The condition $R > 4$, which we have imposed above, ensures that $\overline{U_i} \subset U_{i+1}$, which is required for Lemma 10.2. If $R \leq 4$ then $(H)$ simply follows from $(p_0)$ and Proposition 3.2. ■

# 11 Oscillation inequalities

For any non-empty finite set $U$ and a function $u$ on $U$, denote

$$\operatorname*{osc}_U u := \max_U u - \min_U u$$

The purpose of this section is to prove the estimate (11.7) below which will provide the step $(H) \Longrightarrow [osc]$ of the prove of Theorem 2.1.

**Proposition 11.1** *Assume that the elliptic Harnack inequality $(H)$ holds on $(\Gamma, \mu)$. Then, for any $\varepsilon > 0$, there exists $\sigma = \sigma(\varepsilon, H) < 1$ such that, for any ball $B(x,R)$ and for any function $u$ defined in $\overline{B(x,R)}$ and harmonic in $B(x,R)$, we have*

$$\operatorname*{osc}_{B(x,\sigma R)} u \leq \varepsilon \operatorname*{osc}_{B(x,R)} u\,. \tag{11.5}$$

**Proof.** Fix a ball $B(x,R)$ and denote for simplicity $B_r = B(x,r)$. Let us prove that, for any $r \in (0, R/3]$,

$$\operatorname*{osc}_{B_r} u \leq (1-\delta) \operatorname*{osc}_{B_{3r}} u, \tag{11.6}$$

where $\delta = \delta(H) \in (0,1)$. Then (11.5) follows from (11.6) by iterating.

If $r \leq 1$ then the left hand side of (11.6) vanishes, and (11.6) is trivially satisfied. If $r > 1$ then $\overline{B_{2r}} \subset B_{3r}$, and the function $u - \min_{B_{3r}} u$ is non-negative in $\overline{B_{2r}}$ and harmonic in $B_{2r}$. Applying the Harnack inequality $(H)$ to this function, we obtain

$$\max_{B_r} u - \min_{B_{3r}} u \leq H \left( \min_{B_r} u - \min_{B_{3r}} u \right)$$

and

$$\operatorname*{osc}_{B_r} u \leq (H-1) \left( \min_{B_r} u - \min_{B_{3r}} u \right).$$

Similarly, we have

$$\operatorname*{osc}_{B_r} u \leq (H-1) \left( \max_{B_{3r}} u - \max_{B_r} u \right).$$

Summing up these two inequalities, we conclude

$$\operatorname*{osc}_{B_r} u \leq C \left( \operatorname*{osc}_{B_{3r}} u - \operatorname*{osc}_{B_r} u \right),$$

whence (11.6) follows. ■

**Proposition 11.2** *Assume that the elliptic Harnack inequality $(H)$ holds on $(\Gamma, \mu)$. Let $u \in c_0(B(x,R))$ satisfy in $B(x,R)$ the equation $\Delta u = f$. Then, for any positive $r < R$,*

$$\operatorname*{osc}_{B(x,\sigma r)} u \leq 2 \left( \overline{E}(x,r) + \varepsilon \overline{E}(x,R) \right) \max |f|\,, \tag{11.7}$$

*where $\sigma$ and $\varepsilon$ are the same as in Proposition 11.1.*

**Proof.** Denote for simplicity $B_r = B(x,r)$. By definition of the Green function, we have

$$u(y) = -\sum_{z\in B_R} G_{B_R}(y,z)f(z)$$

whence, using (3.8), we obtain

$$\max|u| \le \overline{E}(x,R)\max|f|\,.$$

Let $v\in c_0(B_r)$ solve the Dirichlet problem $\Delta v = f$ in $B_r$ (see Fig. 8). In the same way, we have

$$\max|v| \le \overline{E}(x,r)\max|f|\,.$$

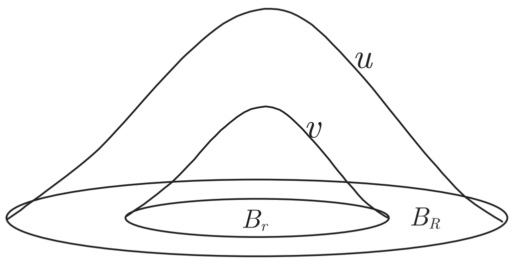


**Figure 8** The functions $u$ and $v$ in the case $f\le 0$.

The function $w = u - v$ is harmonic in $B_r$ whence, by Proposition 11.1,

$$\operatorname*{osc}_{B_{\sigma r}} w \le \varepsilon \operatorname*{osc}_{B_r} w.$$

Since $w = u$ on $\overline{B_r}\setminus B_r$, the maximum principle implies that

$$\operatorname*{osc}_{B_r} w = \operatorname*{osc}_{\overline{B_r}\setminus B_r} w = \operatorname*{osc}_{\overline{B_r}\setminus B_r} u \le 2\max|u|$$

Hence,

$$\operatorname*{osc}_{B_{\sigma r}} u \le \operatorname*{osc}_{B_{\sigma r}} v + \operatorname*{osc}_{B_{\sigma r}} w \le 2\max|v| + 2\varepsilon\max|u| \le 2\left(\overline{E}(x,r)+\varepsilon\overline{E}(x,R)\right)\max|f|\,,$$

which was to be proved. ■

# 12 Time derivative of the heat kernel

The purpose of this section is to show that an on-diagonal upper bound of the heat kernel implies a certain upper bound of its time derivative. Given a function $u_n(x)$ on $\Gamma\times\mathbb{N}$, by "the time derivative" of $u$ we mean the difference

$$\partial_n u := u_{n+2} - u_n\,.$$

Let us emphasize that this is *not* $u_{n+1} - u_n$.

The main result of this section is Proposition 12.3 which constitutes the part $(DUE) \Longrightarrow [deriv]$ of the proof of Theorem 2.1.

**Proposition 12.1** *Let $A$ be a non-empty finite subset of $\Gamma$ and $f$ be a function on $A$. Define*

$$u_n(x) = P_n^A f(x)$$

*Then, for all integers $1 \le k \le n$,*

$$\|\partial_n u\|_{L^2(A,\mu)} \le \frac{1}{k} \|u_{n-k}\|_{L^2(A,\mu)}.$$

**Proof.** The proof follows the argument from [17]. Let $\phi_1$,$\phi_2$, ..., $\phi_{|A|}$ be the eigenfunctions of the Laplace operator $-\Delta_A$ and $\lambda_1$,$\lambda_2$,...,$\lambda_{|A|}$ be the corresponding eigenvalues. Let us normalize $\phi_i$'s to form an orthonormal basis in $L^2(A,\mu)$. The function $f$ can be expanded in this basis

$$f = \sum_i c_i \phi_i.$$

Since $P^A = I - (-\Delta_A)$, we obtain

$$u_n = \sum_i \rho_i^n \phi_i \tag{12.1}$$

where $\rho_i := 1 - \lambda_i$ are eigenvalues of the Markov operator $P^A$.

In particular, we have

$$\|u_n\|^2_{L^2(A,\mu)} = \sum_i \rho_i^{2n}.$$

From (12.1), we obtain

$$u_n - u_{n+2} = \sum \left(1 - \rho_i^2\right) \rho_i^n \phi_i$$

and

$$\|u_n - u_{n+2}\|^2_{L^2(A,\mu)} = \sum_i \left(1 - \rho_i^2\right)^2 \rho_i^{2n}. \tag{12.2}$$

Note that $|\rho_i| \le 1$ and, hence, $\rho_i^2 \in [0,1]$. For any $a \in [0,1]$, we have

$$1 \ge (1 + a + a^2 + ... + a^k)(1-a) \ge k a^k (1-a)$$

whence

$$(1-a)\, a^k \le \frac{1}{k}.$$

Applying this inequality for $a = \rho_i^2$, we obtain from (12.2)

$$\|u_n - u_{n+2}\|^2_{L^2(A,\mu)} \le \frac{1}{k^2} \sum_i \rho_i^{2(n-k)} = \frac{1}{k^2} \|u_{n-k}\|^2_{L^2(A,\mu)}$$

which was to be proved. ∎

**Proposition 12.2** *Let $A$ be a non-empty finite subset of $\Gamma$. Then, for all $x, y \in A$,*

$$\left|\partial_n p^A(x,y)\right| \le \frac{1}{k} \sqrt{p^A_{2m}(x,x) p^A_{2(n-m-k)}(y,y)}, \tag{12.3}$$

*for all positive integers $n, m, k$ such $m + k \le n$.*

**Proof.** From the semigroup identity (5.15) for $p^A$, we obtain

$$\partial_n p^A(x,y) = \sum_{z\in A} p_m^A(x,z)\partial_{n-m}p^A(z,y)\mu(z),$$

whence

$$\left|\partial_n p^A(x,y)\right| \le \left\|p_m^A(x,\cdot)\right\|_{L^2(A,\mu)} \left\|\partial_{n-m}p^A(y,\cdot)\right\|_{L^2(A,\mu)}.$$

By Proposition 12.1,

$$\left\|\partial_{n-m}p^A(y,\cdot)\right\|_{L^2(A,\mu)} \le \frac{1}{k}\left\|p_{n-m-k}^A(y,\cdot)\right\|_{L^2(A,\mu)}$$

for any $1 \le k \le n-m$. Since

$$\left\|p_m^A(x,\cdot)\right\|_{L^2(A,\mu)}^2 = \sum_{z\in A} p_m^A(x,z)^2\mu(z) = p_{2m}^A(x,x),$$

we obtain (12.3). ∎

**Proposition 12.3** *Suppose that $(DUE)$ holds, that is, for all $x \in \Gamma$ and $n \ge 1$,*

$$p_n(x,x) \le Cn^{-\nu}. \tag{12.4}$$

*Then, for all $x, y \in \Gamma$ and $n \ge 1$,*

$$|\partial_n p(x,y)| \le Cn^{-\nu-1}. \tag{12.5}$$

**Proof.** Assume first $n > 3$. Then we can choose $k$ and $m$ in (12.3) so that $k \simeq m \simeq n/3$ and $n-m-k \simeq n/3$ . As follows from (12.4), for any non-empty finite set $A \subset \Gamma$,

$$p_{2m}^A(x,x) \le Cn^{-\nu} \quad \text{and} \quad p_{2(n-m-k)}^A(y,y) \le Cn^{-\nu},$$

whence, by Proposition 12.1,

$$\left|\partial_n p^A(x,y)\right| \le Cn^{-\nu-1}.$$

By letting $A \to \Gamma$, we obtain (12.5).

If $n \le 3$ then (12.5) follows from the trivial inequality $|\partial_n p| \le p_n + p_{n+2}$ and the fact that (12.4) implies a similar bound for $p_n(x,y)$ (cf. 5.17) and (5.18)). ∎

The phenomenon that the time derivative of the heat kernel decays as $n \to \infty$ faster than the heat kernel itself is well known in the context of manifolds (see, for example, [25], [26], [33]). The analogue of the time derivative in the setting of graphs is $\partial_n p$ rather than $p_{n+1} - p_n$. Indeed, in $\mathbb{Z}^D$ (as well as in any bipartite graph) $p_n(x,x) = 0$ if $n$ is odd. Therefore, the difference $p_{n+1}(x,x) - p_n(x,x)$ is equal either to $p_{n+1}(x,x)$ or to $-p_n(x,x)$, and, hence, decays as $n \to \infty$ at the same rate as $p_n(x,x)$.

## 13 Off-diagonal lower bound

An important intermediate step in proving the lower estimate $(LE)$ is a *near-diagonal lower estimate*

$$p_n(x,y) + p_{n+1}(x,y) \ge cn^{-\alpha/\beta}, \tag{$NLE$}$$

for all $x, y \in \Gamma$ and $n \geq 1$ such that

$$d(x, y) \leq \delta n^{1/\beta}. \tag{13.1}$$

In this section, we will finish the prove of the lower bound $(LE)$ in Theorem 2.1 as on the following diagram:

$$(V) + (G) \Longrightarrow (FK) + (V) + (E) + (H) \Longrightarrow (NLE) + (V) \Longrightarrow (LE).$$

The first implication here is given by Propositions 5.5, 6.3 and 10.1 whereas the other two will be proved below.

Let us recall that $(DLE)$ refers to the lower bound

$$p_{2n}^{B(x,R)}(x, x) \geq cn^{-\alpha/\beta}, \quad \forall n \leq \varepsilon R^{\beta}, \tag{$DLE$}$$

with some small enough $\varepsilon > 0$, and $(DUE)$ refers to the upper bound

$$p_n(x, x) \leq Cn^{-\alpha/\beta}. \tag{$DUE$}$$

Denote for simplicity by $(E \leq)$ the upper bound in $(E)$, that is

$$E(x, R) \leq CR^{\beta}, \quad \forall x \in \Gamma, R \geq 1. \tag{$E \leq$}$$

**Proposition 13.1** *For any graph* $(\Gamma, \mu)$ *we have*

$$(DUE) + (DLE) + (E \leq) + (H) \Longrightarrow (NLE). \tag{13.2}$$

*Consequently, if* $(p_0)$ *holds on* $(\Gamma, \mu)$ *then*

$$(FK) + (V) + (E) + (H) \Longrightarrow (NLE). \tag{13.3}$$

**Proof.** Let us first show how the second claim follows from the first one. Recall that, by Proposition 5.1, $(FK) \Longrightarrow (DUE)$, by Proposition 7.1, $(E) \Longrightarrow (\Psi)$ and, by Proposition 9.1, $(\Psi) + (V) \Longrightarrow (DLE)$. Hence, the hypotheses of (13.3) imply the hypotheses of (13.2).

To prove (13.2), fix $x \in \Gamma$, $n \geq 1$ and set

$$R = \left(\frac{n}{\varepsilon}\right)^{1/\beta}, \tag{13.4}$$

for a small enough positive $\varepsilon$. So far we assume only that $\varepsilon$ satisfies $(DLE)$ but later, one more upper bound on $\varepsilon$ will be imposed. Denote $A = B(x, R)$ and introduce the function

$$u(y) := p_n^A(x, y) + p_{n+1}^A(x, y).$$

By the hypothesis $(DLE)$, we have $u(x) \geq cn^{-\alpha/\beta}$. Let us show that

$$|u(x) - u(y)| \leq \frac{c}{2} n^{-\alpha/\beta}, \tag{13.5}$$

for all $y$ such that $d(x, y) \leq \delta n^{1/\beta}$, which would imply $u(y) \geq \frac{c}{2} n^{-\alpha/\beta}$, hence proving $(NLE)$.

The function $u(y)$ is in the class $c_0(A)$ and solves the equation $\Delta u(y) = f(y)$ where

$$f(y) := p_{n+2}^A(x, y) - p_n^A(x, y).$$

The on-diagonal upper bound $(DUE)$ implies, by Proposition 12.3,

$$\max_y |f(y)| \leq \frac{C}{n^{\alpha/\beta+1}}. \tag{13.6}$$

By $(H)$ and Proposition 11.2, we have, for any $0 < r < R$ and for some $\sigma \in (0,1)$,

$$\operatorname*{osc}_{B(x,\sigma r)} u \leq 2\left(\overline{E}(x,r) + \varepsilon^2 \overline{E}(x,R)\right) \max |f|. \tag{13.7}$$

By Proposition 6.1, $(E \leq)$ implies a similar upper bound for $\overline{E}$. Estimating $\max|f|$ by (13.6), we obtain from (13.7)

$$\operatorname*{osc}_{B(x,\sigma r)} u \leq C \frac{r^\beta + \varepsilon^2 R^\beta}{n^{\alpha/\beta+1}}.$$

Choosing $r$ to satisfy $r^\beta = \varepsilon^2 R^\beta$ and substituting from (13.4) $n = \varepsilon R^\beta$, we obtain

$$\operatorname*{osc}_{B(x,\sigma r)} u \leq C \frac{\varepsilon^2 R^\beta}{n^{\alpha/\beta+1}} = C\varepsilon n^{-\alpha/\beta},$$

which implies

$$\operatorname*{osc}_{B(x,\sigma r)} u \leq \frac{c}{2} n^{-\alpha/\beta}, \tag{13.8}$$

provided $\varepsilon$ is small enough.

Note that

$$\sigma r = \sigma \varepsilon^{2/\beta} R = \sigma \varepsilon^{2/\beta} \left(\frac{n}{\varepsilon}\right)^{1/\beta} = \sigma \varepsilon^{1/\beta} n^{1/\beta} = \delta n^{1/\beta}$$

where $\delta := \sigma \varepsilon^{1/\beta}$. Hence, (13.8) implies (13.5) provided $d(x,y) \leq \delta n^{1/\beta}$, which was to be proved. ∎

The final step in proving the part $(V) + (G) \Longrightarrow (LE)$ of Theorem 2.1 is covered by the following statement. Denote by $(V \geq)$ the lower bound in $(V)$, that is

$$V(x,R) \geq cR^\alpha, \quad \forall x \in \Gamma,\ R \geq 1. \tag{13.9}$$

**Proposition 13.2** *Assume that $(\Gamma, \mu)$ satisfies $(p_0)$. Then*

$$(NLE) + (V \geq) \Longrightarrow (LE).$$

We precede the proof with the following lemmas. Denote for simplicity

$$\tilde{P}_n = P_n + P_{n+1}, \tag{13.10}$$

where $P_n$ is the $n$-convolution power of the Markov operator $P$. In particular, we have

$$P_n P_m = P_{n+m}. \tag{13.11}$$

We need a replacement for this property for the operator $\tilde{P}_n$, which is stated below in Lemma 13.5.

**Lemma 13.3** *Assume that $(p_0)$ holds on $(\Gamma, \mu)$, Then, for all integers $n \geq l \geq 1$ such that*

$$n \equiv l \pmod 2, \tag{13.12}$$

*we have*

$$P_l(x,y) \leq C^{n-l} P_n(x,y), \tag{13.13}$$

*for all $x, y \in \Gamma$, with a constant $C = C(p_0)$.*

**Proof.** By the semigroup property (5.15), we have

$$P_{k+2}(x,y) = \sum_{z\in\Gamma} P_k(x,z)P_2(z,y) \geq P_k(x,y)P_2(y,y).$$

Using $(p_0)$, we obtain

$$P_2(y,y) = \sum_{z\sim y} P(y,z)P(z,y) \geq p_0 \sum_{z\sim y} P(y,z) = p_0$$

whence $P_{k+2}(x,y) \geq p_0 P_k(x,y)$. Iterating this inequality, we obtain (13.13) with $C = p_0^{-1/2}$. ■

**Lemma 13.4** *Assume that $(\Gamma,\mu)$ satisfies $(p_0)$. Then, for all integers $n \geq l \geq 1$ and all $x,y\in\Gamma$,*

$$\tilde{P}_l(x,y) \leq C^{n-l}\tilde{P}_n(x,y), \tag{13.14}$$

*where $C = C(p_0)$.*

**Remark 13.1** Note that no parity condition is required here in contrast to the condition (13.12) of Lemma 13.3.

**Proof.** This is an immediate consequence of Lemma 13.3 because both $P_l(x,y)$ and $P_{l+1}(x,y)$ can be estimated from above via either $P_n(x,y)$ or $P_{n+1}(x,y)$ depending on the parity of $n$ and $l$. ■

**Lemma 13.5** *Assume that $(\Gamma,\mu)$ satisfies $(p_0)$. Then, for all $n,m\in\mathbb{N}$ and $x,y\in\Gamma$, we have the following inequality*

$$\tilde{P}_n\tilde{P}_m(x,y) \leq C\tilde{P}_{n+m+1}(x,y), \tag{13.15}$$

*where $C = C(p_0)$.*

**Proof.** Observe that, by (13.10) and (13.11),

$$\tilde{P}_n\tilde{P}_m = (P_n + P_{n+1})(P_m + P_{m+1}) = P_{n+m} + 2P_{n+m+1} + P_{n+m+2}.$$

By Lemma 13.3, $P_{n+m}(x,y) \leq CP_{n+m+2}$ whence

$$\tilde{P}_n\tilde{P}_m\,(x,y) \leq C(P_{n+m+1} + P_{n+m+2}) = C\tilde{P}_{n+m+1}.$$

■

**Lemma 13.6** *Assume that $(\Gamma,\mu)$ satisfies $(p_0)$. Then, for all $x,y\in\Gamma$ and $k,m,n\in\mathbb{N}$ such that $n \geq km+k-1$, we have the following inequality*

$$\left(\tilde{P}_m\right)^k(x,y) \leq C^{n-km}\tilde{P}_n(x,y). \tag{13.16}$$

**Proof.** By induction, (13.15) implies

$$\left(\tilde{P}_m\right)^k(x,y) \leq C^{k-1}\tilde{P}_{km+k-1}(x,y).$$

From inequality (13.14) with $l = km+k-1$, we obtain

$$\tilde{P}_{km+k-1}(x,y) \leq C^{n-km-(k-1)}\tilde{P}_n(x,y)$$

whence (13.16) follows. ■

**Proof of Proposition 13.2.** Since

$$\tilde{P}_n(x,y) = (p_n(x,y) + p_{n+1}(x,y))\mu(y),$$

$(NLE)$ can be stated as follows:

$$\tilde{P}_n(x,y) \geq cn^{-\alpha/\beta}\mu(y), \quad \text{if } d(x,y) \leq \delta n^{1/\beta}. \tag{13.17}$$

The required $(LE)$ takes the form

$$\tilde{P}_n(x,y) \geq cn^{-\alpha/\beta}\mu(y)\exp\left[-\left(\frac{d^\beta(x,y)}{cn}\right)^{\frac{1}{\beta-1}}\right]. \tag{13.18}$$

To prove (13.18), fix $x, y \in \Gamma$, $n \geq d(x,y)$ and consider the following cases:

Case 1. $d(x,y) \leq \delta n^{1/\beta}$;

Case 2. $\delta n^{1/\beta} < d(x,y) \leq \varepsilon n$;

Case 3. $\varepsilon n < d(x,y) \leq n$.

Here $\delta$ is the constant from (13.17) and $\varepsilon > 0$ is a small constant to be chosen later. In the first case, (13.18) coincides with (13.17). In the third case, (13.18) becomes

$$\tilde{P}_n(x,y) \geq cn^{-\alpha/\beta}\mu(y)\exp(-Cn), \tag{13.19}$$

which can be deduced directly from $(p_0)$. Indeed, depending on the parity of $n$, there is a path from $x$ to $y$ of length either $n$ or $n+1$. The $\mathbb{P}_x$-probability that the random walk will follow this path is at least $p_0^{-(n+1)}$, whence

$$\tilde{P}_n(x,y) \geq \exp(-Cn).$$

This implies (13.19) using the fact that $\mu(y) \leq C$. The latter is proved as follows. Take in (13.17) $x \sim y$ and $n \simeq \delta^{-\beta}$. Then (13.17) implies

$$1 \geq \tilde{P}_n(x,y) \geq c\delta^\alpha \mu(y)$$

whence $\mu(y) \leq C$.

Consider the main second case. Denote $d = d(x,y)$, take a positive integer $k$ such that

$$k \leq d\,, \tag{13.20}$$

and define $m$ by

$$m = \lfloor \frac{n}{k} \rfloor - 1. \tag{13.21}$$

Since $k \leq d \leq \varepsilon n$, we see that $n/k \geq \varepsilon^{-1}$ and $m$ is positive. Since $n \geq k(m+1)$, Lemma 13.6 applies and yields

$$C^{n-mk}\tilde{P}_n(x,y) \geq \left(\tilde{P}_m\right)^k(x,y). \tag{13.22}$$

In order to estimate $\left(\tilde{P}_m\right)^k(x,y)$, observe that there exists a sequence $o_1$,$o_2$,...,$o_k$ of points on $\Gamma$ such that $x = o_1$, $y = o_k$ and, for all $i = 1, 2, ..., k-1$,

$$d(o_i, o_{i+1}) \leq \lceil \frac{d(x,y)}{k} \rceil =: r. \tag{13.23}$$

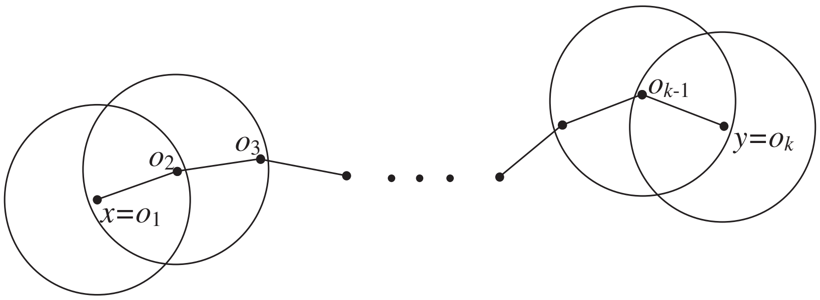


**Figure 9** The chain of balls $B(o_i, r)$

Clearly, we have

$$\left(\tilde{P}_m\right)^k (x,y) \geq \sum_{z_1 \in B(o_1,r)} \ldots \sum_{z_{k-1} \in B(o_{k-1},r)} \tilde{P}_m(x,z_1)\tilde{P}_m(z_1,z_2)\ldots\tilde{P}_m(z_{k-1},y). \tag{13.24}$$

Assume that we have in addition

$$3r \leq \delta m^{1/\beta}. \tag{13.25}$$

Since $d(z_{i-1}, z_i) \leq 3r$, each $\tilde{P}_m(z_{i-1}, z_i)$ can be estimated by (13.17) as follows:

$$\tilde{P}_m(z_{i-1}, z_i) \geq cm^{-\alpha/\beta}\mu(z_i).$$

The same applies to $\tilde{P}_m(x, z_1)$ and $\tilde{P}_m(z_{k-1}, y)$. Using the lower bound of the volume (13.9), we obtain from (13.22) and 13.24)

$$C^{n-mk}\tilde{P}_n(x,y) \geq (cm^{-\alpha/\beta})^{k-1}V(o_1,r)\ldots V(o_{k-1},r)\mu(y) \geq c^k m^{-(\alpha/\beta)k} r^{\alpha(k-1)}\mu(y).$$

Hence,

$$\tilde{P}_n(x,y) \geq c^{n-mk+k} m^{-(\alpha/\beta)k} r^{\alpha(k-1)} \geq c^k m^{-\alpha/\beta}\left(\frac{r}{m^{1/\beta}}\right)^{\alpha(k-1)}, \tag{13.26}$$

where we have used the fact that $n - mk + k \leq 3k$ which follows from (13.21).

Before we go further, let us specify the choice of $k$ to ensure that both (13.20) and (13.25) holds. Using definition (13.21) and (13.23) of $m$ and $r$, we see that (13.25) is equivalent to

$$C\frac{d}{k} \leq \delta \left(\frac{n}{k}\right)^{1/\beta}$$

or

$$k \geq C\delta^{-\frac{\beta}{\beta-1}}\left(\frac{d^\beta}{n}\right)^{\frac{1}{\beta-1}}. \tag{13.27}$$

Let $k$ be the minimal possible integer satisfying (13.27). By the hypothesis $d \geq \delta n^{1/\beta}$, we have

$$k \simeq \left(\frac{d^\beta}{n}\right)^{\frac{1}{\beta-1}}. \tag{13.28}$$

The condition (13.20) follows from the hypothesis $n \geq \varepsilon^{-1}d$ provided $\varepsilon$ is small enough.

From (13.28), (13.21) and (13.25), we obtain

$$m \simeq \left(\frac{n}{d}\right)^{\frac{\beta}{\beta-1}} \quad \text{and} \quad r \simeq \left(\frac{n}{d}\right)^{\frac{1}{\beta-1}}.$$

Hence, by (13.26) and $m \leq n/k$,

$$\tilde{P}_n(x,y) \geq c^k m^{-\alpha/\beta} \geq n^{-\alpha/\beta}k^{\alpha/\beta}\exp(-Ck) \geq n^{-\alpha/\beta}\exp(-C'k).$$

Substituting here $k$ from (13.28), we obtain (13.18). ■

## 14 Parity matters

Let us recall that $(LE)$ contains the estimate for $p_n + p_{n+1}$ rather than for $p_n$. In this section, we discuss to what extent it is possible to estimate $p_n$ from below. In general, there is no lower bound for $p_n(x,y)$ for the parity reason. Indeed, on any bipartite graph, the length of any path from $x$ to $y$ has the same parity as $d(x,y)$. Therefore, $p_n(x,y) = 0$ if $n \not\equiv d(x,y) (\mathrm{mod}\, 2)$.

We immediately obtain the following result for bipartite graphs.

**Proposition 14.1** *If $(\Gamma, \mu)$ is bipartite and satisfies $(LE)$ then*

$$p_n(x,y) \geq cn^{-\alpha/\beta} \exp\left(-\left(\frac{d(x,y)^\beta}{cn}\right)^{\frac{1}{\beta-1}}\right), \tag{14.1}$$

*for all $x, y \in \Gamma$ and $n \geq 1$ such that*

$$n \geq d(x,y) \text{ and } n \equiv d(x,y) (\mathrm{mod}\, 2). \tag{14.2}$$

**Proof.** Indeed, assuming (14.2), $n+1$ and $d(x,y)$ have different parities whence $p_{n+1}(x,y) = 0$, and (14.1) follows from $(LE)$. ■

If there is enough "mixing of parity" in the graph then one does get the lower bound regardless of the parity of $n$ and $d(x,y)$.

**Proposition 14.2** *Assume that graph $(\Gamma, \mu)$ satisfies $(p_0)$, $(LE)$ and the following "mixing" condition: there is an* ***odd*** *positive integer $n_0$ such that*

$$\inf_{x\in\Gamma} P_{n_0}(x,x) > 0. \tag{14.3}$$

*Then the lower bound (14.1) holds for* ***all*** *$n > n_0$ and $x, y \in \Gamma$ provided $n \geq d(x,y)$.*

For example if $n_0 = 1$ then the hypothesis (14.3) means than each point $x \in \Gamma$ has a loop edge $\overline{xx}$. If $n_0 = 3$ and there are no loops then (14.3) means that, for each point $x \in \Gamma$, there is an edge triangle $\overline{xy}$, $\overline{yz}$, $\overline{zx}$. This property holds, in particular, for the graphical Sierpinski gasket - see Fig. 1.

**Proof.** By (9.2), we obtain, for any positive integer $m$,

$$p_{2m}(x,x) \geq \frac{1}{V(x, m+1)}\left(\sum_{z\in B(x,m+1)} p_m(x,z)\mu(z)\right)^2 = \frac{1}{V(x,m+1)}.$$

The condition $(p_0)$ and Proposition 3.1 imply $V(x, m+1) \leq C^{m+1}\mu(x)$ whence

$$P_{2m}(x,x) = p_{2m}(x,x)\mu(x) \geq C^{-m-1}.$$

Since we will use this lower estimate only for bounded range of $m \leq m_0$, we can rewrite it as

$$P_{2m}(x,x) \geq c, \tag{14.4}$$

where $c = c(m_0) > 0$.

Assuming $n > n_0$, we have, by the semigroup property (5.15),

$$p_n(x,y) = \sum_{z\in\Gamma} p_{n-n_0}(x,z)P_{n_0}(z,y) \geq p_{n-n_0}(x,y)P_{n_0}(y,y) \tag{14.5}$$

and in the same way

$$p_n(x,y) \geq p_{n-n_0+1} P_{n_0-1}(y,y). \tag{14.6}$$

By the hypothesis (14.3), we can estimate $P_{n_0}(y,y)$ from below by a positive constant. Also $P_{n_0-1}(y,y)$ is bounded below by a constant as in (14.4). Hence, adding up (14.5) and (14.6), we obtain

$$p_n(x,y) \geq c(p_{n-n_0}(x,y) + p_{n-n_0+1}(x,y)). \tag{14.7}$$

The right-hand side of (14.7) can be estimated from below by $(LE)$ whence (14.1) follows. ■

Finally, let us show an example which explains why in general one cannot replace in $(LE)$ $p_n + p_{n+1}$ by $p_n$ even assuming the parity condition $n \equiv d(x,y) (\mathrm{mod}\, 2)$.

**Example 14.1** Let $(\Gamma,\mu)$ be $\mathbb{Z}^D$ with the standard weight $\mu_{xy} = 1$ for $x \sim y$, and let $D > 4$. We modify $\Gamma$ by adding one more edge $\xi$ of weight 1, which connects the origin $o = (0,0,...,0)$ to the point $(1,1,0,0,...,0)$, and denote the new graph by $(\Gamma',\mu')$.

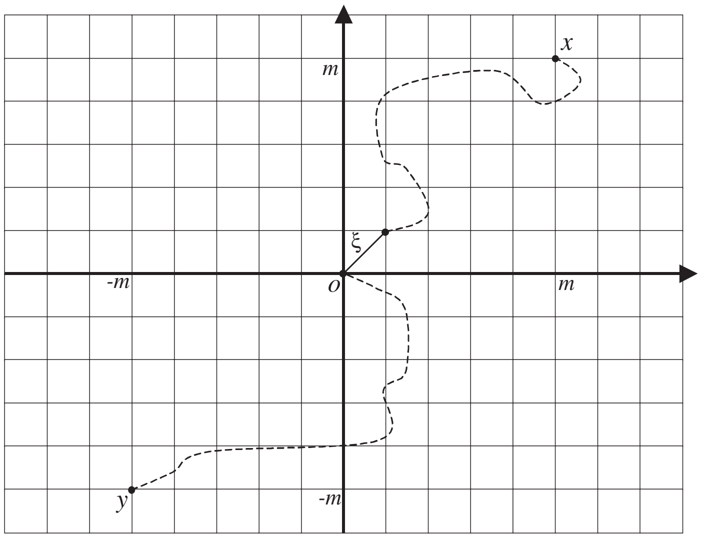


**Figure 10** Every path of odd length from $x$ to $y$ goes through $o$ and $\xi$.

Clearly, the volume growth and the Green kernel on $(\Gamma',\mu')$ are of the same order as on $(\Gamma,\mu)$, that is

$$V(x,r) \simeq r^D \quad \text{and} \quad g(x,y) \simeq d(x,y)^{2-D}.$$

Hence, for both graphs one has by Theorem 2.1

$$p_n(x,y) \leq Cn^{-D/2} \exp\left(-\frac{d^2(x,y)}{Cn}\right) \tag{14.8}$$

and a similar lower bound $(LE)$ for $p_n + p_{n+1}$. Since $\mathbb{Z}^D$ is bipartite, we have for $(\Gamma,\mu)$, by Proposition 14.1,

$$\begin{array}{c} p_n(x,y) \geq cn^{-D/2} \exp\left(-\frac{d^2(x,y)}{cn}\right) \\ \text{if } n \geq d(x,y) \text{ and } n \equiv d(x,y) (\mathrm{mod}\, 2) \end{array} \tag{14.9}$$

Let us show that $(\Gamma',\mu')$ does not satisfy (14.9). Fix some (large) odd integer $m$ and consider points $x = (m,m,0,0,...,0)$ and $y = -x$ (see Fig. 10).

The distance $d(x,y)$ on $\Gamma$ is equal to $4m$, whereas the distance $d'(x,y)$ on $\Gamma'$ is $4m-1$, due to the shortcut $\xi$. Denote $n=m^2$. Then $n \equiv d'(x,y) (\mathrm{mod}\, 2)$ and $n > d'(x,y)$. Let us estimate from above $p_n(x,y)$ on $(\Gamma', \mu')$ and show that it does not satisfy the lower bound (14.9). Since $n$ is odd and all odd paths from $x$ to $y$ have to go through the edge $\xi$, the strong Markov property yields

$$p_n(x,y) = \sum_{k=0}^{n} \mathbb{P}_x(\tau = k) p_{n-k}(o,y)\,, \tag{14.10}$$

where $\tau$ is the first time the random walk hits the point $o$. If $n-k < m$ then $p_{n-k}(o,y)=0$. If $n-k \geq m$ then we estimate $p_{n-k}(o,y)$ by (14.8) as follows

$$p_{n-k}(o,y) \leq \frac{C}{(n-k)^{D/2}} \leq \frac{C}{m^{D/2}}.$$

Therefore, (14.10) implies

$$p_n(x,y) \leq C m^{-D/2} \mathbb{P}_x \{\tau < \infty\}\,.$$

The $\mathbb{P}_x$-probability to hit $o$ is of the order $g(x,o) \simeq m^{2-D}$. Hence, we obtain

$$p_n(x,y) \leq C m^{-(3D/2-2)} = C n^{-(3D/4-1)} = o(n^{-D/2})$$

so that the lower bound (14.9) cannot hold.

A more careful argument shows that, in fact, $p_n(x,y) \simeq n^{-(D-1)}$.

## 15 Consequences of the heat kernel estimates

Here we prove the remaining part of Theorem 2.1 as stated in the next proposition.

**Proposition 15.1** *Assuming* $(p_0)$, *we have*

$$(LE) + (UE) \Longrightarrow (V) + (G).$$

**Proof.** The Green kernel is related to the heat kernel by

$$g(x,y) = \sum_{n=0}^{\infty} p_n(x,y). \tag{15.1}$$

Let $x \neq y$. Then $p_0(x,y)=0$, and the upper bound $(UE)$ for $p_n$ implies the upper bound for $g$ as follows:

$$g(x,y) \leq C \sum_{n=1}^{\infty} n^{-\alpha/\beta} \exp\left(-c\left(\frac{d^\beta}{n}\right)^{\frac{1}{\beta-1}}\right)\,,$$

where $d = d(x,y)$. By estimating the sum via an integral, we obtain $g(x,y) \leq C d^{-\gamma}$ with $\gamma = \alpha - \beta$. Similarly, one proves $g(x,y) \leq C d^{-\gamma}$ using $(LE)$ and the obvious consequence of (15.1)

$$g(x,y) \geq \frac{1}{2} \sum_{n=1}^{\infty} (p_n(x,y) + p_{n+1}(x,y)).$$

Let us prove the upper bound for the volume

$$V(x,R) \leq C R^{\alpha}, \tag{$V \leq$}$$

for any $x \in \Gamma$ and $R \geq 1$. Indeed, for any $n \in \mathbb{N}$, we have

$$\sum_{y \in \Gamma} p_n(x,y)\mu(y) \equiv 1 \tag{15.2}$$

whence

$$\sum_{y e B(x,R)} \left(p_n(x,y) + p_{n+1}(x,y)\right)\mu(y) \leq 2$$

and

$$V(x,R) \leq 2\left(\inf_{y \in B(x,R)} \left(p_n(x,y) + p_{n+1}(x,y)\right)\right)^{-1}.$$

Taking $n \simeq R^\beta$ and applying $(LE)$, we see that the inf is bounded below by $cn^{-\alpha/\beta} \simeq R^{-\alpha}$ whence $(V \leq)$ follows.

Let us prove the lower bound for the volume

$$V(x,R) \geq cR^\alpha. \tag{$V \geq$}$$

We first show that $(UE)$ and $(V \leq)$ imply the following inequality

$$\sum_{y \notin B(x,R)} p_n(x,y)\mu(y) \leq \frac{1}{2}, \quad \forall n \leq \varepsilon R^\beta, \tag{15.3}$$

provided $\varepsilon > 0$ is sufficiently small. Denoting $R_k = 2^k R$, we have

$$\begin{aligned}
\sum_{y \notin B(x,R)} p_n(x,y)\mu(y) &\leq C \sum_{y \notin B(x,R)} n^{-\alpha/\beta} \exp\left[-c\left(\frac{d(x,y)^\beta}{n}\right)^{\frac{1}{\beta-1}}\right] \\
&\leq C \sum_{k=0}^{\infty} \sum_{y \in B(x,R_{k+1}) \setminus B(x,R_k)} n^{-\alpha/\beta} \exp\left[-c\left(\frac{R_k^\beta}{n}\right)^{\frac{1}{\beta-1}}\right] \\
&\leq C \sum_{k=0}^{\infty} R_k^\alpha n^{-\alpha/\beta} \exp\left[-c\left(\frac{R_k^\beta}{n}\right)^{\frac{1}{\beta-1}}\right] \\
&= C \sum_{k=0}^{\infty} \left(\frac{2^k R}{n^{1/\beta}}\right)^\alpha \exp\left[-c\left(\frac{2^k R}{n^{1/\beta}}\right)^{\frac{\beta}{\beta-1}}\right].
\end{aligned} \tag{15.4}$$

If $R/n^{1/\beta}$ is large enough then the right hand side of (15.4) is majorized by a geometric series, and the sum can be made arbitrarily small, in particular, smaller than $1/2$.

From (15.2) and (15.3), we conclude

$$\sum_{y \in B(x,R)} p_n(x,y)\mu(y) \geq \frac{1}{2}, \tag{15.5}$$

whence

$$V(x,R) \geq \frac{1}{2}\left(\sup_{y \in B(x,R)} p_n(x,y)\right)^{-1}.$$

Finally, choosing $n = [\varepsilon R^\beta]$ and using the upper bound $p_n(x,y) \leq Cn^{-\alpha/\beta}$, we obtain $(V \geq)$.

This argument works only if $\varepsilon R^\beta \geq 1$. Let us now prove $(V \geq)$ for the opposite case when $\varepsilon R^\beta < 1$. To that end, define $R_0$ by $\varepsilon R_0^\beta = 1$. Then we have $R < R_o$. By the hypothesis $(p_0)$ and Proposition 3.1, we have $V(x, R_0) \leq C\mu(x)$. Combining with the lower bound $(V \geq)$ for $V(x, R_0)$, we obtain $\mu(x) \geq c > 0$. In particular, for any $R > 0$, we have $V(x, R) \geq c$, which implies $(V \geq)$ for the bounded range of $R$. ■

**Remark 15.1** Using similar argument, one can show also the following implication

$$(V) + (UE) + (H) \Longrightarrow (LE). \tag{15.6}$$

Indeed, as we have seen in the above proof, $(UE)$ implies $(G \leq)$ which together with $(V)$ is enough to obtain $(E \leq)$ (see Proposition 6.3). From $(UE)$ and $(V)$, one obtains the diagonal lower bound $p_{2n}(x, x) \geq cn^{-\alpha/\beta}$. Indeed, from (9.2) and (15.5) with $R = Cn^{1/\beta}$, we deduce

$$p_{2n}(x, x) \geq \frac{1}{V(x, R)} \left( \sum_{y \in B(x,R)} p_n(x, y) d\mu(y) \right)^2 \geq \frac{1}{4V(x, R)} \simeq n^{-\alpha/\beta}.$$

From this estimate, one gets $(DLE)$ (see [54]; the argument is similar to the proof of (6.6)). Also, $(DUE)$ follows trivially from $(UE)$. Hence, having $(DUE)$, $(DLE)$, $(E \leq)$ and $(H)$, we obtain $(NLE)$ by Proposition 13.1 and then deduce $(LE)$ from $(NLE) + (V)$ by Proposition 13.2.

The implication (15.6) yields that $(V) + (UE) + (H)$ is *equivalent* to either of our main conditions $(V) + (G)$ and $(UE) + (LE)$. Indeed, we have

$$(V) + (G) \Longrightarrow (V) + (UE) + (H) \Longrightarrow (UE) + (LE),$$

where the first implication follows by Theorem 2.1 and Proposition 10.1, and the second is the same as (15.6). We are left to close the circle by Theorem 2.1 or Proposition 15.1.

# 16 Appendix: the list of the lettered conditions

Here we provide a list the lettered conditions frequently used in the paper. The relations between the exponents $\alpha, \beta, \gamma, \nu$ are as follows:

$$\alpha > \beta \geq 2, \quad \gamma = \alpha - \beta \quad \text{and} \quad \nu = \alpha/\beta.$$

In all conditions, $n$ is an arbitrary positive integer, $R$ is an arbitrary positive real, $x, y$ are arbitrary points on $\Gamma$, subject to additional restrictions if any. The constants $C, c, \delta, \varepsilon, p_0$ are positive.

$$V(x, R) \simeq R^\alpha, \quad \forall R \geq 1 \tag{$V$}$$

$$E(x, R) \simeq R^\beta, \quad \forall R \geq 1 \tag{$E$}$$

$$g(x, y) \simeq d(x, y)^{-\gamma}, \quad x \neq y \tag{$G$}$$

$$V(x, 2R) \leq CV(x, R) \tag{$D$}$$

$$\overline{E}(x, R) \leq CE(x, R) \tag{$\overline{E}$}$$

$$\lambda_1(A) \geq c\mu(A)^{-1/\nu}, \quad \text{for all non-empty finite sets } A \subset \Gamma \tag{$FK$}$$

$$p_n(x,x) \le C n^{-1/\nu} \tag{$DUE$}$$

$$p_n(x,y) \le C n^{-\alpha/\beta} \exp\left[ -\left( \frac{d(x,y)^\beta}{Cn} \right)^{\frac{1}{\beta-1}} \right] \tag{$UE$}$$

$$(p_n + p_{n+1})(x,y) \ge c n^{-\alpha/\beta} \exp\left[ -\left( \frac{d(x,y)^\beta}{cn} \right)^{\frac{1}{\beta-1}} \right], \quad \text{if } n \ge d(x,y). \tag{$LE$}$$

$$p_{2n}^{B(x,R)}(x,x) \ge c n^{-\alpha/\beta}, \quad \text{if } n \le \varepsilon R^\beta \tag{$DLE$}$$

$$p_n(x,y) + p_{n+1}(x,y) \ge c n^{-\alpha/\beta}, \quad \text{if } d(x,y) \le \delta n^{1/\beta} \tag{$NLE$}$$

$$\Psi_n(x,R) := \mathbb{P}_x\left(T_{B(x,R)} \le n\right) \le C \exp\left[ -\left( \frac{R^\beta}{Cn} \right)^{\frac{1}{\beta-1}} \right] \tag{$\Psi$}$$

$$P(x,y) \ge p_0, \quad \text{if } x \sim y \tag{$p_0$}$$

$$\begin{gathered} \max\nolimits_{B(x,R)} u \le H \min\nolimits_{B(x,R)} u\,, \\ \text{for any function } u \text{ non-negative in } \overline{B}(x,2R) \text{ and harmonic in } B(x,2R). \end{gathered} \tag{$H$}$$

# References


[1] **Aronson D.G.**, Non-negative solutions of linear parabolic equations, *Ann. Scuola Norm. Sup. Pisa. Cl. Sci. (4)*, **22** (1968) 607-694. Addendum **25** (1971) 221-228.

[2] **Auscher P.**, Regularity theorems and heat kernel for elliptic operators, *J. London Math. Soc.*, **2(54)** (1996) 284-296.

[3] **Auscher P., Coulhon T.**, Gaussian lower bounds for random walks from elliptic regularity, *Ann. Inst. Henri Poincaré - Probabilités et Statistiques*, **35** no.5, (1999) 605-630.

[4] **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.

[5] **Barlow M.T., Bass R.F.**, The construction of the Brownian motion on the Sierpisnki carpet, *Ann. Inst. H. Poincare*, **25** (1989) 225-257.

[6] **Barlow M.T., Bass R.F.**, Transition densities for Brownian motion on the Sierpinski carpet, *Probab. Th. Rel. Fields*, **91** (1992) 307–330.

[7] **Barlow M.T., Bass R.F.**, Brownian motion and harmonic analysis on Sierpinski carpets, to appear in *Canad. J. Math.*

[8] **Barlow M.T., Bass R.F.**, Random walks on graphical Sierpinski carpets, preprint

[9] **Barlow M.T., Coulhon T., Grigor'yan A.**, Manifolds and graphs with slow heat kernel decay, preprint (1999)

[10] **Barlow M.T., Perkins A.**, Brownian motion on the Sierpinski gasket, *Probab. Th. Rel. Fields*, **79** (1988) 543–623.

[11] **Benjamini I., Chavel I., Feldman E.A.**, Heat kernel lower bounds on Riemannian manifolds using the old ideas of Nash, *Proceedings of London Math. Soc.*, **72** (1996) 215-240.

[12] **Boukricha A.**, Das Picard-Prinzip und verwandte Fragen bei Störung von harmonischen Räumen, *Math. Ann.*, **239** (1979) 247-270.

[13] **Carlen E.A., Kusuoka S., Stroock D.W.**, Upper bounds for symmetric Markov transition functions, *Ann. Inst. H. Poincaré, Prob. et Stat.*, (1987) suppl. auno.2, 245-287.

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

[15] **Chavel I.**, "Eigenvalues in Riemannian geometry" Academic Press, New York, 1984.

[16] **Chavel I.**, "Isoperimetric inequalities and heat diffusion on Riemannian manifolds" Lecture notes 1999.

[17] **Cheng S.Y., Li P., Yau S.-T.**, On the upper estimate of the heat kernel of a complete Riemannian manifold, *Amer. J. Math.*, **103** (1981) no.5, 1021-1063.

[18] **Cheng S.Y., Yau S.-T.**, Differential equations on Riemannian manifolds and their geometric applications, *Comm. Pure Appl. Math.*, **28** (1975) 333-354.

[19] **Chung F.R.K.**, "Spectral Graph Theory" CBMS Regional Conference Series in Mathematics 92, AMS publications, 1996.

[20] **Coulhon T.**, Ultracontractivity and Nash type inequalities, *J. Funct. Anal.*, **141** (1996) 510-539.

[21] **Coulhon T., Grigor'yan A.**, On-diagonal lower bounds for heat kernels on non-compact manifolds and Markov chains, *Duke Math. J.*, **89** (1997) no.1, 133-199.

[22] **Coulhon T., Grigor'yan A.**, Random walks on graphs with regular volume growth, *Geom. and Funct. Analysis*, **8** (1998) 656-701.

[23] **Coulhon T., Saloff-Coste L.**, Minorations pour les chaînes de Markov unidimensionnelles, *Prob. Theory Relat. Fields*, **97** (1993) 423-431.

[24] **Davies E.B.**, "Heat kernels and spectral theory" Cambridge University Press, 1989.

[25] **Davies E.B.**, Pointwise bounds on the space and time derivatives of the heat kernel, *J. Operator Theory*, **21** (1989) 367-378.

[26] **Davies E.B.**, Non-Gaussian aspects of heat kernel behaviour, *J. London Math. Soc.*, **55** (1997) no.1, 105-125.

[27] **Delmotte T.**, Parabolic Harnack inequality and estimates of Markov chains on graphs, *Revista Matemática Iberoamericana*, **15** no.1, (1999) 181-232.

[28] **Fabes E.B., Stroock D.W.**, A new proof of Moser's parabolic Harnack inequality via the old ideas of Nash, *Arch. Rat. Mech. Anal.*, **96** (1986) 327-338.

[29] **Goldstein S.**, Random walks and diffusion on fractals, *in:* "Percolation theory and ergodic theory on infinite particle systems", ed. H.Kesten, Springer, New York, 1987. 121-129.

[30] **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.

[31] **Grigor'yan A.**, Heat kernel upper bounds on a complete non-compact manifold, *Revista Matemática Iberoamericana*, **10** (1994) no.2, 395-452.

[32] **Grigor'yan A.**, Integral maximum principle and its applications, *Proc. Roy. Soc. Edinburgh*, **124A** (1994) 353-362.

[33] **Grigor'yan A.**, Upper bounds of derivatives of the heat kernel on an arbitrary complete manifold, *J. Funct. Anal.*, **127** (1995) no.2, 363-389.

[34] **Grigor'yan A.**, Estimates of heat kernels on Riemannian manifolds, *in:* "Spectral Theory and Geometry. ICMS Instructional Conference, Edinburgh 1998", ed. B.Davies and Yu.Safarov, London Math. Soc. Lecture Note Series 273, Cambridge Univ. Press, 1999. 140-225.

[35] **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.

[36] **Hebisch W., Saloff-Coste, L.**, Gaussian estimates for Markov chains and random walks on groups, *Ann. Prob.*, **21** (1993) 673–709.

[37] **Hebisch W., Saloff-Coste, L.**, On the relation between elliptic and parabolic Harnack inequalities, preprint

[38] **Jones O.D.**, Transition probabilities for the simple random walk on the Sierpinski graph, *Stoch. Proc. Appl.*, **61** (1996) 45-69.

[39] **Kigami J.**, Harmonic calculus for p.c.f. self-similar sets, *Trans. Amer. Math. Soc.*, **335** (1993) 721-755.

[40] **Kigami J.**, Harmonic calculus on limits of networks and its application to dendrites, *J. Funct. Anal.*, **128** (1995) 48-86.

[41] **Kigami J., Lapidus M.**, Weyl's spectral problem for the spectral distribution of Laplacians on p.c.f. self-similar fractals, *Comm. Math. Phys.*, **158** (1993) 93-125.

[42] **Krylov N.V., Safonov M.V.**, A certain property of solutions of parabolic equations with measurable coefficients, (in Russian) *Izv. Akad. Nauk SSSR*, **44** (1980) 81-98. Engl. transl. *Math. USSR Izvestija*, **16** (1981) 151-164.

[43] **Kumagai T.**, Estimates of transition densities for Brownian motion on nested fractals, *Prob. Th. Rel. Fields*, **96** (1993) 205-224.

[44] **Kusuoka S., Zhou X.Y.**, Dirichlet form on fractals: Poincaré constant and resistance, *Prob. Th. Rel. Fields*, **93** (1992) 169-196.

[45] **Landis E.M.**, "The second order equations of elliptic and parabolic type" (in Russian) Nauka, Moscow, 1971. Engl. transl. Transl. of Mathematical Monographs 171, AMS publications, 1998.

[46] **Li P., Yau S.-T.**, On the parabolic kernel of the Schrödinger operator, *Acta Math.*, **156** (1986) no.3-4, 153-201.

[47] **Lust-Piquard F.**, Lower bounds on $\|K^n\|_{1\to\infty}$ for some contractions $K$ of $L^2(\mu)$, with some applications to Markov operators, *Math.Ann.*, **303** (1995) 699-712.

[48] **Moser J.**, On Harnack's theorem for elliptic differential equations, *Comm. Pure Appl. Math.*, **14** (1961) 577-591.

[49] **Nash J.**, Continuity of solutions of parabolic and elliptic equations, *Amer. J. Math.*, **80** (1958) 931-954.

[50] **Porper F.O., Eidel'man S.D.**, Two-side estimates of fundamental solutions of second-order parabolic equations and some applications, (in Russian) *Uspekhi Matem. Nauk*, **39** (1984) no.3, 101-156. Engl. transl. *Russian Math. Surveys*, **39** (1984) no.3, 119-178.

[51] **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.

[52] **Saloff-Coste L.**, Isoperimetric inequalities and decay of iterated kernels for almost-transitive Markov chains, *Combinatorics, Prob. Comput.*, **4** (1995) 419-442.

[53] **Schoen R., Yau S.-T.**, "Lectures on Differential Geometry" Conference Proceedings and Lecture Notes in Geometry and Topology 1, International Press, 1994.

[54] **Stroock D.W.**, Estimates on the heat kernel for the second order divergence form operators, *in:* "Probability theory. Proceedings of the 1989 Singapore Probability Conference held at the National University of Singapore, June 8-16 1989", ed. L.H.Y. Chen, K.P. Choi, K. Hu and J.H. Lou, Walter De Gruyter, 1992. 29-44.

[55] **Sturm K-Th.**, Analysis on local Dirichlet spaces III. The parabolic Harnack inequality, *Journal de Mathématiques Pures et Appliquées*, **75** (1996) no.3, 273-297.

[56] **Takeda M.**, On a martingale method for symmetric diffusion process and its applications, *Osaka J. Math*, **26** (1989) 605-623.

[57] **Telcs A.**, Random walks on graphs, electric networks and fractals, *J. Prob. Theo. and Rel. Fields*, **82** (1989) 435-449.

[58] **Telcs A.**, Spectra of graphs and fractal dimensions I, *Prob. Theo. and Rel. Fields*, **85** (1990) 489-497.

[59] **Telcs A.**, Spectra of graphs and fractal dimensions II, *J. Theor. Prob.*, **8** (1995) 77-96.

[60] **Telcs A.**, Transition probability estimates for reversible Markov chains, to appear in *ECP*

[61] **Varopoulos N.Th.**, Hardy-Littlewood theory for semigroups, *J. Funct. Anal.*, **63** (1985) no.2, 240-260.

[62] **Varopoulos N.Th.**, Analysis on nilpotent groups, *J. Func. Anal.*, **66** (1986) 406-431.

[63] **Varopoulos N.Th., Saloff-Coste L., Coulhon T.**, "Analysis and geometry on groups" Cambridge University Press, Cambridge, 1992.

[64] **Woess W.**, Random walks on infinite graphs and groups - a survey on selected topics, *Bull. LMS*, **26** (1994) 1-60.

[65] **Zhou X.Y.**, Resistance Dimension, Random Walk Dimension and Fractal Dimension, *J. Theor. Prob.*, **4** (1993) 635-652.