Heat kernel estimates on weighted graphs

Bernd Metzger*
and
Peter Stollmann

Ludwig-Maximilians-Universität München and
Johann Wolfgang Goethe-Universität Frankfurt

Dedicated to J. Weidmann on the occasion of his sixtieth birthday

Abstract

We prove upper and lower heat kernel bounds for the Laplacian on weighted graphs which include the case that the weights have no strictly positive lower bound. Our estimates give rise to a very explicit probabilistic interpretation and can be formulated in terms of a weighted metric. Interestingly, this metric is not equivalent to the intrinsic metric.

MS classification: 39 A 12

The results

Heat kernel estimates are a tool of fundamental importance as they usually provide a link between different aspects of the Laplacian: spectral theory, geometry and probability theory. In this article we study heat kernels on graphs with weighted edges. Here the geometric and probabilistic features are to a large extent contained in the properties of the weight function. Different aspects of such a study have attracted considerable interest, see [1, 2, 3, 4, 7].
The present work is motivated by spectral theoretic applications, namely by the aim to prove absence of absolutely continuous spectrum for Schrödinger operators on graphs

*partly supported by the German Science Foundation, DFG

which contain enough barriers made up of edges with low conductance (low weight).
To this end we had to improve partly upon estimates by Davies [5, 6] and Pang [8]. The progress refers to a sensitivity of our estimates to the following geometry which is characteristic of the application we have in mind: given two points which are separated by edges of low weight the heat kernel should "feel" this barrier of low weight edges. Another feature is the new method of proof we introduce here. This method is rather elementary but powerful and gives rise to a most natural probabilistic interpretation.
Let us first fix notation. Let X be some countable set of vertices and $E \subset X \times X$ the set of edges of a directed graph; so every edge is of the form $e = (x, y)$ and we denote $x = ie, y = je$. Suppose we are given a weight $b : E \to (0, \infty)$ such that sup $b =: b_{max} < \infty$. We assume throughout that E is symmetric, i.e. for $e = (x, y) \in E$ we have $\bar{e} := (y, x) \in E$ and that $b(e) = b(\bar{e})$ for all $e \in E$. Moreover, edges of the form (x, x) are forbidden. Finally, let us assume that each vertex is the starting point of only a finite number of edges, and that we have a uniform upper bound

$$M = \sup_{x \in X} |\{e \in E; ie = x\}| < \infty.$$

The Laplacian on this weighted graph is the following operator on $l^2(X)$:

$$\Delta : l^2(X) \longrightarrow l^2(X),$$

$$\Delta f(x) = \sum_{e \in E, ie=x} b(e)\,(f(je) - f(x))$$

This operator is easily seen to be bounded, non-negative, and the generator of a strongly continuous Markovian semigroup of contractions. Therefore, it is the generator of a continuous time Markov chain on X with transition probabilities

$$p_t(x, y) = (e^{\Delta t} \delta_x)\,(y),$$

where $\delta_x \in l^2(X)$ denotes the unit mass at x. This continuous time Markov process $(X_t; t \geq 0)$ has the following characteristic properties: it jumps from x to one of its neighbours y with probability

$$\frac{b(x, y)}{\sum_{e \in E, ie=x} b(e)}. \tag{1}$$

Thus the weight b may be thought of as measuring the "conductance" of edges. For the holding time at a vertex x we have

$$\mathbb{P}\left[\inf\{s > t; X_s(\omega) \neq X_t(\omega)\} > t + h | X_t(\omega) = x\right] = \exp\left[\left(-\sum_{ie=x} b(e)\right) h\right]. \tag{2}$$

It is our purpose in this paper to prove heat kernel estimates which are sensitive to the following phenomenon: suppose two points x and y are connected only through edges with low weight. Then, $p_t(x,y)$ should be small. A precise version of this is the main result of our paper, for which we introduce a little more notation.

For $x, y \in X$, a finite sequence $\gamma = (e_1, .., e_m)$ is called a **path from** x **to** y, if $ie_1 = x$, $je_k = ie_{k+1}$, for $k = 1, ..., m-1$, and $je_m = y$; $m =: |\gamma|$ is called its length. By $\Gamma(x,y)$ we denote the set of all paths from x to y, and $e \in \gamma$ is shorthand for the fact that $e \in \{e_k; k = 1, ..., m\}$. Finally, denote by

$$d_0(x,y) = \min\{|\gamma|; \gamma \in \Gamma(x,y)\}$$

the combinatorial distance from x to y. Clearly, d_0 defines a metric, if we restrict ourselves to the case that the graph is connected, i.e. $\Gamma(x,y) \neq \emptyset$ for all $x, y \in X, x \neq y$. We will do so in the sequel without loosing any interesting phenomena: in the general case one can restrict attention to the connected components, as the Laplacian is a direct sum of the Laplacians of the connected components.

Theorem 1. *For $x, y \in X$, $t \geq 0$ we have*

$$\frac{e^{-b_{max}Mt}}{\sqrt{2\pi}} \cdot \sup_{\gamma \in \Gamma(x,y)} \left(\prod_{e \in \gamma} \frac{tb(e)}{|\gamma|} \right) \leq p_t(x,y) \leq e^{b_{max}Mt+1} \cdot \left(\sup_{\gamma \in \Gamma(x,y)} \prod_{e \in \gamma} \frac{b(e)}{b_{max}} \right) \cdot \left(\frac{eMb_{max}t}{d_0(x,y)} \right)^{d_0(x,y)}.$$

Although the estimate in Theorem 1 looks somewhat technical at first sight, it has a natural interpretation in terms of the underlying Markov process: Let us first concentrate on the upper bound. In view of (1) above, the second factor estimates, essentially, the probability that we have successive jumps along the edges of γ. This factor is small, if x and y are separated by a barrier of edges with low conductance. The third factor estimates the probability that we have at least $d_0(x,y)$ jumps during t units of time. In fact, using (2) above, we find that the probability of jumping from x to one of its neighbours in less than h units of time is given by

$$1 - \exp\left[\left(-\sum_{ie=x} b(e)\right) h\right] \sim \left(\sum_{ie=x} b(e)\right) h.$$

For each of the $d_0(x,y)$ jumps during the interval $[0,t]$ we have

$$h \sim \frac{t}{d_0(x,y)}.$$

Estimating

$$\sum_{ie=x} b(e) \leq M \cdot b_{max},$$

we arrive at a probability of at most

$$\left(\frac{M \cdot b_{max} \cdot t}{d_0(x,y)}\right)^{d_0(x,y)},$$

which is, more or less, the third factor. The lower bound has a similar structure in view of

$$\sup_{\gamma\in\Gamma(x,y)}\left(\prod_{e\in\gamma}\frac{b(e)\cdot t}{|\gamma|}\right)=\sup_{\gamma\in\Gamma(x,y)}\left[\left(\prod_{e\in\gamma}\frac{b(e)}{b_{max}}\right)\cdot\left(\frac{b_{max}\cdot t}{|\gamma|}\right)^{|\gamma|}\right].$$

We use a very direct scheme of proof, writing

$$e^{\Delta t}=\lim_{n\to\infty}(I+\frac{t}{n}\cdot\Delta)^n$$

and estimating the n-th power appearing on the r.h.s. above. To do so, we split Δ into its diagonal and off-diagonal part. Given a constant diagonal part, we can expand the power rather explicitly, as no commutator terms appear. Therefore, we compare Δ with operators with constant diagonal part in a first, rather elementary lemma

Lemma 2. *Denote by S the operator*

$$S:l^2(X)\longrightarrow l^2(X),$$

$$Sf(x)=\sum_{ie=x}b(e)f(je)$$

and by D_1 the operator of multiplication $-b_{max}M\cdot I$. Then we have, for $0<s<(b_{max}M)^{-1}$, $f\in l^2(X)$, $f\geq 0$ and $n\in\mathbb{N}$:

$$0\leq(I+s\cdot(D_1+S))^nf\leq(I+s\Delta)^nf\leq(I+s\cdot S)^nf$$

Proof. Denote by D the operator of multiplication

$$D(x)=-\sum_{e\in E,ie=x}b(e).$$

Note that $\Delta=D+S$ and that

$$D_1\leq D\leq 0$$

as functions. As the off-diagonal matrix elements of Δ, i.e. the matrix elements of S are positive, the latter inequality gives the assertion, as an easy induction argument shows. □

We now aim at estimates for the kernels of $(I+s(D_1+S))^n$ and $(I+sS)^n$. Let us first single out two preparatory inequalities which follow from Stirling's formula and some straightforward calculation:

Lemma 3. (a) *Let $k,n\in\mathbb{N}$, $0\leq k\leq n$. Then*

$$\binom{n}{k}\frac{1}{n^k}\geq\frac{1}{\sqrt{2\pi}}\exp\left(\frac{1}{12(n-1)}\right)\frac{1}{k^k}$$

(b) *Let $c > 0$, $m, n \in \mathbb{N}$, $n > m$. Then*

$$\sum_{k=m}^{n} \left(\begin{array}{c} n \\ k \end{array} \right) \left(\frac{c}{n}\right)^k \leq \left(\frac{c}{m}\right)^m \cdot e \left(1 + \frac{m}{n-m}\right)^{n-m} \cdot \sqrt{1 + \frac{m}{n-m}} \cdot$$

$$\exp\left(\frac{1}{12(n-1)}\right) \left(1 + \frac{c}{n}\right)^{n-m}$$

Proof of the Theorem. For a path $\gamma = (e_1, \ldots, e_m)$ we write $i\gamma = ie_1, j\gamma = je_m$ and

$$b(\gamma) = \prod_{k=1}^{m} b(e_k).$$

It is then easy to see that

$$S^k \delta_x = \sum_{\substack{i\gamma = x, \\ |\gamma| = k}} b(\gamma) \cdot \delta_{j\gamma}$$

and, therefore,

$$S^k \delta_x(y) = \sum_{\substack{\gamma \in \Gamma(x,y) \\ |\gamma| = k}} b(\gamma). \tag{3}$$

We want to estimate

$$p_t(x, y) = \lim_{n \to \infty} [(1 + \frac{t}{n}\Delta)^n \delta_x](y).$$

By Lemma 2, we have

$$\begin{aligned}
[(1 + \frac{t}{n}\Delta)^n \delta_x](y) &\leq [(1 + \frac{t}{n}S)^n \delta_x](y) \qquad (4)\\
&= \sum_{k=0}^{n} \binom{n}{k} \left[\left(\frac{t}{n}S\right)^k \delta_x\right](y) \\
&= \sum_{k=0}^{n} \binom{n}{k} \left(\frac{t}{n}\right)^k \left(\sum_{\substack{\gamma \in \Gamma(x,y) \\ |\gamma| = k}} b(\gamma)\right) \\
&\leq \sum_{k=0}^{n} \binom{n}{k} \left(\frac{t}{n}\right)^k M^k \sup_{\substack{\gamma \in \Gamma(x,y) \\ |\gamma| = k}} b(\gamma) \\
&\leq \sum_{k=d_0(x,y)}^{n} \binom{n}{k} \left(\frac{t}{n}\right)^k M^k \sup_{\substack{\gamma \in \Gamma(x,y) \\ |\gamma| = k}} b(\gamma) \\
&\leq \sum_{k=d_0(x,y)}^{n} \binom{n}{k} \left(\frac{t \cdot M \cdot b_{max}}{n}\right)^k \left(\sup_{\substack{\gamma \in \Gamma(x,y) \\ |\gamma| = k}} \prod_{e \in \gamma} \frac{b(e)}{b_{max}}\right) \\
&\leq \sup_{\gamma \in \Gamma(x,y)} \left(\prod_{e \in \gamma} \frac{b(e)}{b_{max}}\right) \sum_{k=d_0(x,y)}^{n} \binom{n}{k} \left(\frac{t \cdot M \cdot b_{max}}{n}\right)^k.
\end{aligned}$$

We can now estimate the latter sum by Lemma 3 (b) and perform the limit $n \longrightarrow \infty$ to obtain the upper estimate in Theorem 1. For the lower estimate, we start with n large enough, so that Lemma 2 applies:

$$\begin{aligned}
[(1+\frac{t}{n}\Delta)^n \delta_x](y) &\geq [(1-\frac{tMb_{max}}{n}+\frac{t}{n}S)^n\delta_x](y) \qquad (5)\\
&= \sum_{k=0}^{n}\binom{n}{k}\left(1-\frac{tMb_{max}}{n}\right)^{n-k}\left[\left(\frac{t}{n}S\right)^k\delta_x\right](y)\\
&= \sum_{k=0}^{n}\binom{n}{k}\left(1-\frac{tMb_{max}}{n}\right)^{n-k}\left(\frac{t}{n}\right)^k\left(\sum_{\substack{\gamma\in\Gamma(x,y)\\|\gamma|=k}} b(\gamma)\right),
\end{aligned}$$

where we used (3) above for the last equality. In virtue of Lemma 3 (a) we conclude that

$$\ldots \geq \left(1-\frac{tMb_{max}}{n}\right)^{n-k}\frac{1}{\sqrt{2\pi}}\exp\left[\frac{1}{12(n-1)}\right]\left(\sup_{\substack{\gamma\in\Gamma(x,y)\\|\gamma|=k}}\prod_{e\in\gamma}\frac{b(e)t}{|\gamma|}\right) \qquad (6)$$

for every $k \leq n$. Letting $n \longrightarrow \infty$ we arrive at the desired lower bound. $\square$

Remark 4. *Taking the k-th term in the sum (6) for $k = d_0(x,y)$ one gets*

$$p_t(x,y) \geq (2\pi)^{-1/2}e^{-b_{max}Mt}\left(\frac{tb_{max}}{d_0(x,y)}\right)^{d_0(x,y)}\cdot \sup_{\substack{\gamma\in\Gamma(x,y)\\|\gamma|=d_0(x,y)}}\left(\prod_{e\in\gamma}\frac{b(e)}{b_{max}}\right).$$

Let us now comment on a different way in which results of this type can be proven, the Davies' method. This method is based on the following strategy: for a weight function Φ, strictly positive and bounded on X, consider the operator $\Phi\Delta\Phi^{-1}$ on $l^2(X)$. A clever choice of the weight function is

$$\Phi(x) = \exp[-\rho(x)],$$

where ρ is a suitably chosen distance function. Taking ρ to be d_0, roughly speaking, Davies arrives at the following upper bound in [5], Corollary 12:

$$p_t(x,y) \leq \left(\frac{e\cdot k(x,y)\cdot t}{2d_0(x,y)}\right)^{d_0(x,y)},$$

where

$$k(x,y) = \max\{\sum_{ie=z} b(e); d_0(x,z) \leq d_0(x,y)\}.$$

In fact, Davies considers a more general case, including a weight on the vertices. One sees that this upper bound corresponds to the third factor in our estimate in Theorem 1 above. The second factor in our estimate provides extra decay in case that x and y are separated by a barrier of low weight edges. In [6] the same author proposes the use of metrics suggested by non-commutative geometry, which are reminiscent of the intrinsic metric of general Dirichlet forms. It turns out, however, that these metrics cannot be used in the general case considered here. In [8], Pang uses a weighted metric and logarithmic Sobolev inequalities. His results are restricted to a setting where the weight has a strictly lower bound. These results are sharp for the usual discrete Laplacian on the standard lattice $\mathbb{Z}$. In his diploma thesis [7], the first named author of the present paper used Davies' method and the following family of metrics:

$$d_\lambda(x,y) = \sup_{\gamma \in \Gamma(x,y)} \left(\lambda|\gamma| + \sum_{e\in\gamma} \log\left(\frac{Mb_{max}}{b(e)}\right)\right),$$

where $\lambda \in (0,\infty)$. He has found the following upper bound:

$$p_t(x,y) \le e^{b_{max}M^2 t} \cdot \left(\sup_{\gamma\in\Gamma(x,y)} \prod_{e\in\gamma} \frac{b(e)}{Mb_{max}} \right) \cdot \left(\frac{eM^2 b_{max} t}{2d_0(x,y)} \right)^{d_0(x,y)}.$$

With a minor modification of our proof above we can achieve almost the same result. In fact, in (4) above, we simply have to multiply and divide by M^k in the appropriate way. Starting from this expression (4) one can also derive an estimate which no longer involves b_{max}:

$$p_t(x,y) \le e^{Mt+1} \cdot \left(\sup_{\gamma\in\Gamma(x,y)} \prod_{e\in\gamma} b(e) \right) \cdot \left(\frac{eMt}{d_0(x,y)} \right)^{d_0(x,y)}.$$

By standard convergence arguments, this result extends to Laplacians with unbounded weights. However, situations are possible, in which the right hand side no longer is finite. Finally, let us remark that it is not hard to include a weight on the vertices of the graph, $a: X \to (0,\infty)$. In this case, one usually considers the Laplacian

$$\Delta f(x) = a(x)^{-1} \sum_{ie=x} b(e)(f(je) - f(ie)),$$

which is selfadjoint on $l^2(X,a)$. The condition one has to impose is

$$c_{max} = \sup_{e\in E} a(ie)^{-1} b(e) < \infty.$$

This leads to the estimates of Theorem 1 with b_{max} replaced by c_{max}.

References

[1] P. Auscher, T.Coulhon: Gaussian lower bounds for random walks from elliptic regularity, Preprint 1997

[2] T.Coulhon: Analysis on infinite graphs with regular volume growth, to appear in the Proceedings of the 1997 Conference on Random walks and Discrete Potential Theory, Symposia Math., Cambridge University Press

[3] T.Coulhon, A. Grigoryan: Random walks on graphs with regular volume growth, *G.A.F.A. Geom. funct. anal.* **Vol.8**, 656 - 701 (1998)

[4] T. Delmotte: Parabolic Harnack inequality and estimates of Markov chains on graphs, to appear in *Revista Mat. Iberoamericana*

[5] E.B. Davies: Large deviations for heat kernels on graphs, *J. London Math. Soc.* (2) **47**, 65 - 72 (1993)

[6] E.B. Davies: Analysis on Graphs and Noncommutative Geometry, *J. Funct. Anal.* **111**, 398 - 430 (1993)

[7] B. Metzger: Die innere Metrik des Laplaceoperators auf gewichteten Graphen und spektraltheoretische Anwendungen. Diplomarbeit, Frankfurt 1998

[8] M.H. Pang: Heat kernels of graphs, *J. London Math. Soc.* (2) **47**, 50 - 64 (1993)

Authors address:
Fachbereich Mathematik
Universität Frankfurt
D-60054 Frankfurt, Germany
e-mail:
bmetzger@math.uni-frankfurt.de stollman@math.uni-frankfurt.de