BULLETIN (New Series) OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 56, Number 4, October 2019, Pages 705–711
https://doi.org/10.1090/bull/1643
Article electronically published on August 3, 2018

Random walks and heat kernels on graphs, by Martin T. Barlow, London Mathematical Society Lecture Notes Series, Vol. 438, Cambridge University Press, Cambridge, 2017, xi+226 pp., ISBN 978-1-107-67442-4, US$80

Research on random walks has enjoyed long and continuous interest in both the theoretical and applied disciplines. In physics the study started with Karl Pearson's 1905 paper [11], where he asked about the probability that a person (or an ancient Scottish golfer, as some claim was the inspiration) hitting a ball to an exact distance l in a random angle will find himself after n hits at distance $r \pm dr$ from where he started. Lord Rayleigh [12,13] had already answered a more general question while studying sound waves in inhomogenous materials. He modeled the waves as sums of n isoperiodic vibrations of unit amplitude and random phases. One can justify the randomness in phases, since the waves change their direction when hitting scattering parts in the material but keep an approximately constant amplitude. Lord Rayleigh derived a density for "extremely large" n,

$$p_n(r)dr \sim \frac{2}{nl^2} e^{-\frac{r^2}{nl^2}} r dr,$$

answering Pearson's question.

A simple random walk is the stochastic process one obtains by summing i.i.d. (independent and identically distributed) random variables. In the case where the random walk is on $\mathbb{Z}^d$, there is a rich theory (see [8,14]) whose main message is that under diffusive scaling (multiplying by $\frac{1}{\sqrt{n}}$ at time n) the random walk converges to a Gaussian distribution. This is known as the central limit theorem (CLT). Moreover, the invariance principle (or functional CLT) tells us that as a stochastic process, random walk converges to Brownian motion.

The subject of the book under review is a generalization of the random walk notion called random walk on weighted graphs, allowing a much richer set of phenomena. The mathematical study of random walks on graphs interacts with many mathematical disciplines, e.g., harmonic analysis, geometry, functional analysis, PDE, and algebra. These interactions enrich the disciplines in both directions, providing powerful tools to study random walks and, in exchange, gaining understanding on some classical questions in those fields. Recently, applications of random walks on graphs were found in the study of deep neural nets (see [7] for some background). In this review we will explain the settings of the theory and delineate some elegant results arising from the interactions with other fields.

Notations and setting

A graph is a pair $\daleth = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V}$ is a finite or countably infinite set, called the vertex set, and $\mathcal{E}$ is a subset of $\{\{x, y\} : x, y \in \mathcal{V}\}$, called the edge set. We say that x and y are neighbors and denote it by $x \sim y$ if $\{x, y\} \in \mathcal{E}$. A path in $\daleth$ is a sequence $x_0, x_1, \ldots, x_n$ with $x_i \sim x_{i+1}$ for $0 \leq i \leq n-1$. Denote by $d(x, y)$ the length of the shortest path connecting x and y and by $B(x, r) = \{y \in \mathcal{V} : d(x, y) \leq r\}$.

2010 *Mathematics Subject Classification.* Primary 60-02, 60K35; Secondary 60D05.

Definition 1. $(\gimel, \mu)$ is a weighted graph with real weights (or conductances) $\{\mu_{xy}\}_{x,y\in\mathcal{V}}$ satisfying the following.

(i) $\mu_{xy} = \mu_{yx}$.
(ii) If $x \neq y$, $\mu_{xy} > 0$ for $x \sim y$ and $\mu_{xy} = 0$ for $x \nsim y$.

We abbreviate $\mu_x := \sum_{y\sim x} \mu_{xy}$.

For simplicity we make the following assumptions on the graphs.

(H1) $\gimel$ is locally finite, i.e., for every $x \in \mathcal{V}$, $|\{y \in \mathcal{V} : y \sim x\}| < \infty$.
(H2) $\gimel$ is connected, i.e., there is a path connecting every $x, y \in \mathcal{V}$.
(H3) $\gimel$ is of bounded degree, i.e., $\sup_{x\in\mathcal{V}} |\{y \in \mathcal{V} : y \sim x\}| < \infty$.
(H4) $(\gimel, \mu)$ has bounded weights, i.e., $\exists C_1 < \infty$ such that $C_1^{-1} \leq \mu_{xy} \leq C_1$ for all $x, y \in \mathcal{V}$, $x \sim y$.
(H5) $(\gimel, \mu)$ has controlled weights (also known as elliptic), i.e., $\exists C_2 < \infty$ s.t.

$$\frac{\mu_{xy}}{\mu_x} \leq \frac{1}{C_2},$$

for all $x, y \in \mathcal{V}$, $x \sim y$.

Note that (H3)+(H4) yield (H5).

Definition 2. A random walk on $(\gimel, \mu)$ is the discrete time Markov chain on the graph with transition probabilities

$$P(x, y) = \mathbf{P}[X_{n+1} = y | X_n = x] = \frac{\mu_{xy}}{\mu_x}.$$

We write $\mathbf{P}^x$ for the distribution of the random walk conditioned on $X_0 = x$, i.e.,

$$\mathbf{P}^x[X_n = \cdot] = \mathbf{P}[X_n = \cdot | X_0 = x].$$

Note that $P(x, y)$ is *reversible* with respect to μ, i.e.,

$$\mu_x P(x, y) = \mu_y P(y, x) \quad \forall x, y \in \mathcal{V}.$$

Definition 3. For $A \subset \mathcal{V}$ define the *first hitting time* of a set $A \subset \mathcal{V}$ by

$$\tau_A = \min\{n \geq 0 : X_n \in A\}. \tag{1}$$

Electric networks

A basic question about a random walk on a graph $\gimel$ is, Given that currently the walk is in position x, what is the probability it will ever hit position y, or what is the probability it will hit y before some other vertex z? One important way to answer this question is via the deep connection of random walks and electric networks. We will illustrate the connections and use of electric networks thorough the following gamblers ruin problem. A gambler plays a zero-sum game in which at every round he may win \$1 with probability p and loose \$1 with probability $q = 1 - p$. When the game starts the gambler has \$$k$ ($0 \leq k \leq n$), and he decides that he will fold once he either has \$$n$ or \$0. We wish to calculate the probability that the gambler does not leave the game broke. First let us see how we can model this question in the formalism of weighted graphs. A natural graph to choose is $\mathbb{Z} \cap [0, n]$ with edges between every two consecutive integers. The position of the random walk at time t models the amount of money the gambler has after t turns of the game. For normalization purposes, set $\mu_{01} = 1$. For any $1 \le x \le n-1$ assume that $\mu_{(x-1)x} = \xi$. Since the probability the gambler wins \$1 is p, we have

$$p = p(x, x+1) = \frac{\mu_{x(x+1)}}{\xi + \mu_{x(x+1)}} \Longrightarrow \mu_{x(x+1)} = \xi \frac{p}{q}.$$

Thus for all $0 \le x \le n-1$, we have $\mu_{x(x+1)} = \left(\frac{p}{q}\right)^x$. The electric network approach (which we will justify shortly) claims that the weighted graph $(\mathbb{Z} \cap [0,n], \mu)$ can be viewed as a network with potential difference 1 between 0 and n, and with resistance $1/\mu_{x(x+1)}$ (conductance $\mu_{x(x+1)}$) on the edge $(x, x+1)$ for each $0 \le x \le n-1$. The advantage of this view is that we can use standard electric network techniques to reduce the part of the network between 0 and k to a single edge with conductance

$$\left(\sum_{i=0}^{k-1} \left(\frac{q}{p}\right)^i\right)^{-1},$$

and the part of the network between k and n to a single edge with conductance

$$\left(\sum_{i=k}^{n-1} \left(\frac{q}{p}\right)^i\right)^{-1}.$$

Now our network has only two edges and the probability to turn right, corresponding to the random walk hitting n before 0, is

$$\mathbf{P}^k(\tau_{\{n\}} < \tau_{\{0\}}) = \frac{\left(\sum_{i=k}^{n-1} \left(\frac{q}{p}\right)^i\right)^{-1}}{\left(\sum_{i=k}^{n-1} \left(\frac{q}{p}\right)^i\right)^{-1} + \left(\sum_{i=0}^{k-1} \left(\frac{q}{p}\right)^i\right)^{-1}} = \begin{cases} \frac{1-\left(\frac{q}{p}\right)^k}{1-\left(\frac{q}{p}\right)^n} & p \ne q, \\ \frac{k}{n} & p = q. \end{cases}$$

Next, we justify the previous analysis and show the connection between weighted graphs and electric networks. Let $A, Z \subset \mathcal{V}$, and define $v(x) = \mathbf{P}^x(\tau_A < \tau_Z)$. It is immediate that $\forall x \in A, v(x) = 1$ and $\forall x \in Z, v(x) = 0$ and for $x \in \mathcal{V} \setminus (A \cup Z)$ by conditioning on the first step and using the Markov property

$$\begin{aligned} v(x) &= \sum_{y \sim x} \mathbf{P}^x(\text{first step to } y)\mathbf{P}^x(\tau_A < \tau_Z | \text{first step to } y) \\ &= \sum_{y \sim x} \mathbf{P}^x(\text{first step to } y)\mathbf{P}^y(\tau_A < \tau_Z) = \frac{1}{\mu_x} \sum_{y \sim x} \mu_{xy} v(y). \end{aligned}$$

This means that $v(x)$ is harmonic on the weighted graph. By the uniqueness principle of harmonic functions [10, Section 2.1], we obtain that $v(x)$ is the voltage at the point x when we view the graph $(\beth, \mu)$ as a network such that between neighboring vertices x and y there is a conducting wire with resistance $1/\mu_{xy}$. Now that we have the voltage at any vertex of our graph, we can define the current

$$i_{xy} := \mu_{xy}(v(x) - v(y)).$$

The calculation we did for the gambler's ruin problem is a simple example of effective resistance. This concept is well documented in the literature, and we will not elaborate it here; see [1,6,10] for good expositions on electric networks and effective resistance.

Discrete harmonic analysis and potential theory

In this section we present the harmonic analysis setting for random walks that allowed many crossovers of techniques and results between probability theory and PDE.

We call $p_n(x,y) := \frac{\mathbf{P}^x(X_n=y)}{\mu_y}$ the discrete time heat kernel of the graph $\gimel$. We abbreviate $p(x,y) = p_1(x,y)$. Note that the reversibility of the walk implies that the heat kernel is symmetric:

$$(2)\qquad p_n(x,y) = \frac{\mathbf{P}^x[X_n = y]}{\mu_y} = \frac{\mu_x \mathbf{P}^x[X_n = y]}{\mu_x \mu_y} = \frac{\mu_y \mathbf{P}^y[X_n = x]}{\mu_x \mu_y} = p_n(y,x).$$

Define the space $L^2(\mathcal{V},\mu)$ of real valued functions with finite L^2-norms, with the inner product

$$\langle f, g \rangle = \sum_{x \in \mathcal{V}} f(x) g(x) \mu_x.$$

Define the operators

$$P_n f(x) = \sum_{y \in \mathcal{V}} p_n(x,y) f(y) \mu_y = \mathbf{E}^x[f(X_n)].$$

The discrete Laplacian Δ is acting on $L^2(\mathcal{V},\mu)$ by

$$(3)\qquad \Delta f(x) = \frac{1}{\mu_x} \sum_{y \in \mathcal{V}} \mu_{xy}(f(y) - f(x)) = (P_1 - I) f(x).$$

The Dirichlet energy form is defined by

$$E(f,g) = \frac{1}{2} \sum_{x \in \mathcal{V}} \sum_{y \in \mathcal{V}} \mu_{xy}(f(x) - f(y))(g(x) - g(y)),$$

whenever the sum converges absolutely.

There are deep connections between the analysis of the discrete Laplacian and Dirichlet energy, and the geometric isoperimetric inequalities (see [5, 15]). This was first discovered by Cheeger [2, 4] in the framework of manifolds and was later introduced in the discrete setting by Varopoulos [16].

Definition 4. For a subset $A \subset \gimel$, define the boundary weight of A as

$$\mu(A; \mathcal{V} \setminus A) = \sum_{x \in A} \sum_{y \in \mathcal{V} \setminus A} \mu_{xy}.$$

Note that for equal unit weights this corresponds to the cardinality of the edge boundary of A. We say that a finite weighted graph $(\gimel, \mu)$ satisfies the *relative isoperimetric inequality* (related to the Cheeger constant) if $\exists C_R < \infty$ such that for every finite set $A \neq \emptyset$ satisfying $\mu(A) \leq \frac{1}{2}\mu(\mathcal{V})$,

$$\frac{\mu(A; \mathcal{V} \setminus A)}{\mu(A)} \geq C_R.$$

The largest C_R satisfying the last equation is called the *relative isoperimetric constant* (or *Cheeger's constant* in some literature), and we denote it by $R_\gimel$. Positivity of $R_\gimel$ is related to the concepts of nonamenability and expander graphs.

An infinite weighted graph $(\daleth, \mu)$ satisfies the isoperimetric inequality I_α if there exists a $C_0 < \infty$ such that for every finite nonempty set $A \subset \mathcal{V}$,

$$\frac{\mu(A; \mathcal{V} \setminus A)}{\mu(A)^{1-\frac{1}{\alpha}}} \geq C_0^{-1}.$$

The property I_α is stronger than the following *Nash inequality*. There exists a constant C_N such that for all $f \in L^1 \cap L^2$,

$$\mathcal{E}(f,f) \geq C_N \|f\|_2^{2+d/\alpha} \|f\|_1^{-4/\alpha}.$$

Here the classical Poincaré inequality takes the following form.

Definition 5. The weighted graph $(\daleth, \mu)$ satisfies the (*weak*) *Poincaré inequality* (PI) if there is a $C_P < \infty$ and $\lambda \geq 1$ such that for all $x \in \mathcal{V}$, $R \geq 1$, and $f : B(x, \lambda R) \mapsto \mathbb{R}$,

$$\begin{aligned}\sum_{y \in B(x,R)} (f(y) - \bar{f}_{B(x,R)})^2 \mu_y &\leq C_P R^2 \frac{1}{2} \sum_{x \in B(x,\lambda R)} \sum_{y \in B(x,\lambda R)} \mu_{xy} (f(x) - f(y))^2 \\ &= C_P R^2 \mathcal{E}_{B(x,\lambda R)}(f,f),\end{aligned}$$

where $\bar{f}_{B(x,R)} = \mu_{B(x,R)}^{-1} \sum_{y \in B(x,R)} f(y)\mu_y$. We say that $(\daleth, \mu)$ satisfies the *strong Poincaré inequality* if the definition holds with $\lambda = 1$.

Due to the following lemma, in order to prove the validity of the Poincaré inequality it is enough to control the isoperimetric constant of balls.

Lemma 6. *Let* $(\daleth, \mu)$ *be a finite graph. Then for any* $f : \mathcal{V} \mapsto \mathbb{R}$,

$$\min_\lambda \sum_{y \in \mathcal{V}} |f(y) - \lambda|^2 \mu_y \leq 2 R_\daleth^{-2} \mathcal{E}(f,f),$$

where $R_\daleth$ *is the isoperimetric constant of* $\daleth$.

As we will see in the next section, the Poincaré inequality will provide the exponential heat kernel bounds which are widely used in probability theory and other fields.

Another important concept connected to isoperimetry and the Poincaré inequality is the spectrum of a weighted graph. Consider the graph *connectivity* matrix

$$A_{ij} := \frac{\mu_{ij}}{(\mu_i \mu_j)^{1/2}} = \frac{\mu_i^{1/2} P(i,j)}{\mu_j^{1/2}}.$$

Since A is a symmetric matrix, all its eigenvalues are real. Denote the eigenvalues by $\rho_1 \geq \rho_2 \geq \cdots \geq \rho_N$. Since they are the same as the eigenvalues of $P(i,j)$, we get that $\forall i$, $|\rho_i| \leq 1$. It is clear that $\rho_1 = 1$, since $\vec{1}$ is an eigenvector with eigenvalue 1, $\sum_y P(x,y) \cdot 1 = 1$. A more convenient matrix to work with is $-\Delta$, whose eigenvalues are $\lambda_k = 1 - \rho_k$. Many theorems rely on some kind of convergence of the discrete Laplacian to the continuous Laplace operator. For a finite connected graph, the second smallest eigenvalue λ_2 is important for mixing time and exit time calculations (see [9]). Cheeger's inequality states:

Theorem 7 (Cheeger's inequality)**.** *Let* $(\daleth, \mu)$ *be a finite graph, and let* λ_2 *be second smallest eigenvalue of* $-\Delta$ (*also called the spectral gap*)*. Then*

$$\frac{1}{2} R_\daleth^2 \leq \lambda_2 \leq 2 R_\daleth.$$

Definition 8. A weighted graph $(\gimel, \mu)$ is *amenable* if there is a sequence of subgraphs A_n with $R_{A_n} \to 0$ as $n \to \infty$.

A sufficient condition for a graph to be nonamenable is for the random walk on the graph to be ballistic; i.e., $\liminf_{n\to\infty} \frac{1}{n} d(x, X_n) > 0$, $\mathbf{P}^x$-a.s. for each $x \in \mathcal{V}$.

Heat kernel bounds

Bounds on the heat kernel (2) have strong connections to the mixing time of Markov chains and local limit theorems. We start with the maximality of the diagonal heat kernel.

Lemma 9. *For any $x, y \in \mathcal{V}$, $p_{2n}(x,y) \le p_{2n}(x,x)^{1/2} p_{2n}(y,y)^{1/2}$.*

Proof. For every $n, m \in \mathbb{Z}$,

$$p_{n+m}(x,y) = \sum_{z\in\mathcal{V}} p_n(x,z) p_m(y,z) \mu_z = \langle p_n(x,\cdot), p_m(y,\cdot) \rangle.$$

The lemma follows by taking $n = m$ and using the Cauchy–Schwarz inequality. ☐

For simple random walk on a vertex transitive graph[1] this lemma states that the vertex you are most likely to visit is your starting point: for every $x, y \in \mathcal{V}$,

$$p_{2n}(x,x) \ge p_{2n}(x,y).$$

Note that we write $2n$ because of parity issues. This can be avoided by introducing laziness for the random walk (probability to stay in place) or working with continuous time. The next theorem shows a connection between the Nash inequality or isoperimetric inequality and bounds on the diagonal heat kernel.

Theorem 10. *If $(\gimel, \mu)$ satisfies the Nash inequality with parameter α (or the isoperimetric inequality I_α), then there is a constant C_H such that for all $n \ge 1$ and $x, y \in \mathcal{V}$,*

$$p_n(x,y) \le \frac{C_H}{n^{\alpha/2}}.$$

Example. The $\mathbb{Z}^d$ isoperimetric inequality states that for every set $A \subset \mathbb{Z}^d$, the edge boundary of A is greater than $|A|^{(d-1)/d}$. Thus the graph $\mathbb{Z}^d$ with constant weights between neighbors (corresponding to simple random walk on $\mathbb{Z}^d$) satisfies I_d. By Theorem 10 we get the simple random walk diagonal heat kernel bound

$$p_n(x,y) \le \frac{C}{n^{d/2}}.$$

Theorem 10 gives no information on the decay as $d(x,y) \to \infty$. The next theorem is called the Carne–Varopoulos bound [3, 17].

Theorem 11. *Let $(\gimel, \mu)$ be a weighted graph. Then*

$$p_n(x,y) \le \frac{2}{(\mu_x \mu_y)^{1/2)}} e^{-\frac{d(x,y)^2}{2n}} \quad \forall x, y \in \mathcal{V}, \ n \ge 1.$$

---

[1] $G = (\mathcal{V}, \mathcal{E})$ is a vertex transitive graph, if for every $x, y \in \mathcal{V}$ there is a graph automorphism $f : \mathcal{V} \mapsto \mathcal{V}$ such that $f(x) = y$.

About the book

The book under review delineates very thoroughly the general theory of random walks on weighted graphs. The author's expertise in both probability and analysis is apparent in the exposition and the elegant proofs depicted in the book. There are many subjects covered in the book that we did not discuss in this review. We believe this book can be used both as a graduate topic course textbook and as a reference source for expert researchers.

References

[1] M. T. Barlow, *Random walks and heat kernels on graphs*, London Mathematical Society Lecture Note Series, vol. 438, Cambridge University Press, Cambridge, 2017. MR3616731
[2] P. Buser, *A note on the isoperimetric constant*, Ann. Sci. École Norm. Sup. (4) **15** (1982), no. 2, 213–230. MR683635
[3] T. K. Carne, *A transmutation formula for Markov chains* (English, with French summary), Bull. Sci. Math. (2) **109** (1985), no. 4, 399–405. MR837740
[4] J. Cheeger, *A lower bound for the smallest eigenvalue of the Laplacian*, Problems in analysis (Papers dedicated to Salomon Bochner, 1969), Princeton Univ. Press, Princeton, N. J., 1970, pp. 195–199. MR0402831
[5] F. R. K. Chung, *Spectral graph theory*, CBMS Regional Conference Series in Mathematics, vol. 92, Published for the Conference Board of the Mathematical Sciences, Washington, DC; by the American Mathematical Society, Providence, RI, 1997. MR1421568
[6] P. G. Doyle and J. L. Snell, *Random walks and electric networks*, Carus Mathematical Monographs, vol. 22, Mathematical Association of America, Washington, DC, 1984. MR920811
[7] B. Hanin. *Which neural net architectures give rise to exploding and vanishing gradients?*, arXiv:1801.03744 (2018).
[8] G. F. Lawler and V. Limic, *Random walk: a modern introduction*, Cambridge Studies in Advanced Mathematics, vol. 123, Cambridge University Press, Cambridge, 2010. MR2677157
[9] D. A. Levin and Y. Peres, *Markov chains and mixing times*, American Mathematical Society, Providence, RI, 2017. MR3726904
[10] R. Lyons and Y. Peres, *Probability on trees and networks*, Cambridge Series in Statistical and Probabilistic Mathematics, vol. 42, Cambridge University Press, New York, 2016. MR3616205
[11] K. Pearson, *The problem of the random walk*, Nature (London) **72** (1905), 294.
[12] J. W. Strutt, 3rd Baron Rayleigh, *On the electromagnetic theory of light*, The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science, **12** (1881), no 73, 81–101.
[13] J. W. Strutt, 3rd Baron Rayleigh, *The problem of random walk*, Nature (London) **72** (1905), 318–325.
[14] F. Spitzer, *Principles of random walk*, 2nd ed., Springer-Verlag, New York-Heidelberg, 1976. Graduate Texts in Mathematics, Vol. 34. MR0388547
[15] T. Sunada, *Discrete geometric analysis*, Analysis on graphs and its applications, Proc. Sympos. Pure Math., vol. 77, Amer. Math. Soc., Providence, RI, 2008, pp. 51–83, DOI 10.1090/pspum/077/2459864. MR2459864
[16] N. Th. Varopoulos, *Isoperimetric inequalities and Markov chains*, J. Funct. Anal. **63** (1985), no. 2, 215–239, DOI 10.1016/0022-1236(85)90086-2. MR803093
[17] N. Th. Varopoulos, *Long range estimates for Markov chains* (English, with French summary), Bull. Sci. Math. (2) **109** (1985), no. 3, 225–252. MR822826

Eviatar B. Procaccia
Department of Mathematics, Texas A&M University
Email address: procaccia@tamu.edu
www.math.tamu.edu/∼procaccia