Bounds On The Cover Time

Andrei Z. Broder Anna R. Karlin

October 15, 1988

SRC Research Report 32

Anna Karlin co-wrote this report while she was a postdoctoral fellow in the Department of Computer Science, Princeton University. The work was supported in part by ONR grant N00014-87-K-0467.

Abstract.

Consider a particle that moves on a connected, undirected graph G with n vertices. At each step the particle goes from the current vertex to one of its neighbors, chosen uniformly at random. The *cover time* is the first time when the particle has visited all the vertices in the graph starting from a given vertex.

In this paper, we present upper and lower bounds that relate the expected cover time for a graph to the eigenvalues of the Markov chain that describes the random walk above. An interesting consequence is that regular expander graphs have expected cover time $\Theta(n \log n)$.

1. Introduction.

Consider a particle moving on an undirected graph $G = (V, E)$ from vertex to vertex according to the following rule: the probability of a transition from vertex i, of degree d_i, to vertex j is $1/d_i$ if $(i, j) \in E$, and 0 otherwise. This stochastic process is a Markov chain; it is called a random walk on the graph G. In this paper we derive upper and lower bounds on the expected *cover time*, the time taken by a random walk on G to visit all vertices in V. (See [11] for a general reference on Markov chains.)

Formally, let $\{X_t\}$ be any discrete time Markov chain with state space S. (For the random walk $S = V$ above.) The *hitting time* H_{ij} for $i, j \in S$ is defined by

$$H_{ij} = \inf\{t \ : \ X_t = j \mid X_0 = i\}$$

and the *cover time* (or covering time) C_i, for $i \in S$ is defined as

$$C_i = \max_{j \in S} H_{ij}.$$

In other words, the hitting time H_{ij} is the first time that state j is reached starting from state i, and the cover time C_i is the first time every state in S is visited at least once starting from state i.

Computer scientists originally became interested in analyzing the expected cover times for graphs in an attempt to obtain bounds on the space complexity of *undirected st-connectivity* [6] (Given an undirected graph G and two specified vertices s and t in G, determine if there is a path connecting s and t). There are several other topics in computer science and graph theory that motivate the investigation of cover times:

- The study of the relations between combinatorial properties of graphs (such as cover times, expansion etc.) and algebraic properties of graphs (the eigenstructure of the associated adjacency matrix). Some recent papers in this long line of research are [1], [7], and [9].
- The exploration of the special characteristics of expander graphs such as short covering times, rapidly-mixing properties, the existence of short universal sequences for them, etc.
- The simulation of token rings on arbitrary networks using random redirection of tokens: There is a plethora of protocols and algorithms designed for use on token ring networks. It is desirable to be able to simulate these protocols on arbitrary networks

and most notably on some of the extremely high-speed networks with limited buffer capacity that are currently being built. An obvious approach for doing such a simulation is simply to route the token along a spanning tree of the network. A problem with this idea is that it is not fault-tolerant; when some nodes or links become faulty or flooded, an extensive reconfiguration of the spanning tree might be required while the old ring is not operative. These operations entail rather complicated and slow mechanisms. An alternative approach, randomly routing tokens, is efficient (if the expected cover times are small), extremely fault-tolerant, and suffers little loss of efficiency when several "colored" tokens are routed simultaneously.

There is a large body of previous work on cover times and related problems. First, there is the classical solution to the coupon collector's problem, which shows that the expected time to collect n coupons is $\Theta(n \log n)$. This is essentially equivalent to showing that the expected cover time for a random walk on the complete graph is $\Theta(n \log n)$. There are many other specific graphs for which the expected cover time has been computed. These include paths, cycles, trees [10], bar-bell graphs[1] [14], and d-dimensional cubes [4].

For arbitrary connected graphs, Aleliunas et al. [6] showed a general upper bound $\mathbf{E}(C_i) = O(|E|\,|V|)$, starting from any vertex i, where $|E|$ is the number of edges and $|V|$ is the number of vertices.

A superficial examination of these problems might lead one to conjecture that adding more edges to the graph would reduce the expected cover time. This is false. For example, the complete graph, K_n, on n vertices, has expected cover time $O(n \log n)$; the bar-bell graph B_n, on n vertices has expected cover time $\Omega(n^3)$; and the line graph L_n, consisting of a path of length $n-1$, has expected cover time $O(n^2)$. On the other hand $K_n \supset B_n \supset L_n$.

In this paper, we present relations between the expected cover time for graphs and the rate of convergence of the corresponding Markov chain to its stationary distribution, which in turn is determined by the second largest eigenvalue of its transition probability matrix. A preliminary version of part of this work has appeared in [8]. Most theorems refer to reversible Markov chains, that is, Markov chains for which $\pi_i P_{ij} = \pi_j P_{ji}$, where π is a stationary distribution, and P is the transition probability matrix. A random walk on a graph is a particular case of a reversible Markov chain: a stationary distribution is $\pi_i = d_i/\big(2\,|E|\big)$; if $(i,j) \in E$ then $\pi_i P_{ij} = \pi_j P_{ji} = 1/\big(2\,|E|\big)$; if $(i,j) \notin E$ then $\pi_i P_{ij} =$

[1] The bar-bell graph on n vertices consists of two cliques each of size $n/3$, connected by a path of length $n/3$.

$\pi_j P_{ji} = 0$.

Our main results are the following:

Theorem 5. *Let M be an irreducible reversible Markov chain on n states with transition probability matrix P. Let $1 = \lambda_1 > \lambda_2 \geq \cdots \geq \lambda_n$ be the eigenvalues of P. Let $\pi = (\pi_1, \pi_2, \ldots, \pi_n)$ be the eigenvector corresponding to the eigenvalue 1, normalized so that $\sum_i \pi_i = 1$. (If M is aperiodic then π is the stationary distribution of M.) We rename the states such that $\pi_1 \leq \pi_2 \leq \cdots \leq \pi_n$. Let $\epsilon > 0$ be a constant. Then the cover time C starting from any state satisfies*

$$\mathbf{E}(C) \leq \frac{1}{1-\lambda_2}\left((2+\epsilon)n\ln n - n\ln\pi_1 + \sum_{1\leq i\leq n}\Big(\sum_{1\leq j\leq i}\pi_i\Big)^{-1}\right)\big(1+o(1)\big).$$

∎

This theorem has several interesting consequences:

1. Any symmetric, irreducible Markov chain has expected cover time $O\big(n\log n/(1-\lambda_2)\big)$. (A Markov chain is symmetric if its transition probability matrix is symmetric.)
2. Any d-regular expander graph has expected cover time $O(n\log n)$.
3. Any connected graph has expected cover time $O\big(n^2\log n/(1-\lambda_2)\big)$.

We also obtain two lower bound results:

Theorem 13. *Let M be an irreducible and aperiodic, reversible Markov chain on n states with transition probability matrix P. Assume that P is doubly-stochastic and has eigenvalues $1 = \lambda_1 > \lambda_2 \geq \cdots \geq \lambda_n$. Let $\lambda_{\max} = \max_{2\leq i\leq n}|\lambda_i|$. Suppose that $\lambda_{\max} \leq 1 - n^{\epsilon-1}$ for $\epsilon > 0$. Then the cover time C, starting from any state, satisfies $\mathbf{E}(C) = \Omega(n\log n)$.* ∎

Corollary 16. *Let M be an irreducible and aperiodic, reversible Markov chain on n states, with transition probability matrix P. Let $1 = \lambda_1 \geq \cdots \geq \lambda_n$ be the eigenvalues of P. If $\lambda_2 \geq 1 - 1/(n\ln n)$ then the expected cover time starting from the stationary distribution of M is $\Omega(n\log n)$.* ∎

The two results above can be generalized to periodic Markov chains, but we omit the proofs. For random walks on graphs, "the chain is aperiodic" is equivalent to "the graph is not bipartite," and "the transition probability matrix is doubly-stochastic" is equivalent to "the graph is regular."

Aldous [5] has recently showed a general lower bound for random walks on graphs, $\mathbf{E}(C) = \Omega(n \log n)$ starting from the stationary distribution, but not necessarily from every state.

2. The upper bound.

In this section, we present an upper bound on the expected number of steps needed to visit all states in a reversible Markov chain with state space V. We decompose this process into a sum of expected first passage times, each from a set of states S that the random walk has already visited to a state in $V - S$. These expected first passage times are bounded by a function of the eigenvalues of the corresponding transition submatrix. The critical new result needed to exploit this bound is Lemma 1(b) below, which relates the largest eigenvalue of this submatrix to the second largest eigenvalue of the complete matrix.

Lemma 1. *Let M be an irreducible and aperiodic, reversible Markov chain on n states with transition probability matrix P. Let S be a subset of the states of size s. Let R be the submatrix of P corresponding to the set of indices $S \times S$. Then*

(a) All the eigenvalues of P and of R are real.

(b) Let the eigenvalues of P be $1 = \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$, let $\lambda_{\max} = \max_{2 \leq i \leq n} |\lambda_i|$, and let the eigenvalues of R be $\mu_1 \geq \mu_2 \geq \cdots \geq \mu_s$. Then

$$1 - \mu_1 \geq (1 - \lambda_{\max})\Big(1 - \sum_{i \in S} \pi_i\Big)$$

where $(\pi_1, \ldots, \pi_n)$ is the stationary distribution of M.

(c)

$$\mu_2 \leq \lambda_{\max}.$$

Proof: Part (a) is standard but we present its proof for completeness. Let D be the diagonal matrix given by $D_{ii} = 1/\pi_i$. Then $P = DA$ where A is given by $A_{ij} = \pi_i P_{ij}$ and is symmetric because of reversibility. It is easy to show that P has exactly the same eigenvalues as the symmetric matrix $B = D^{1/2} A D^{1/2}$. Indeed, let x be a left eigenvector of P corresponding to an eigenvalue λ. Then $xDA = \lambda x$ implies $(xD^{1/2})D^{1/2}AD^{1/2} = \lambda(xD^{1/2})$. Since B is symmetric, it has real eigenvalues and hence P does. Similarly, R has the same eigenvalues as the symmetric submatrix $Q \subset B$ induced by the set of indices $S \times S$.

We now prove part (b): A left eigenvector of B corresponding to the stationary distribution of P is $\sigma = (\sqrt{\pi_1}, \sqrt{\pi_2}, \ldots, \sqrt{\pi_n})$. Let $\tilde{\rho}$ be a positive left eigenvector of Q for eigenvalue μ_1 (such a $\tilde{\rho}$ exists by a weak analogue of the Perron-Frobenius theorem), that is $\tilde{\rho}Q = \mu_1\tilde{\rho}$. Pad $\tilde{\rho}$ to an n element vector by adding zeros for all indices not in S. Call the resulting vector ρ. We normalize ρ such that

$$(\rho, \sigma) = (\sigma, \sigma) = 1 \tag{1}$$

(This is always possible because ρ and σ are positive.)

We can write

$$\rho B = \mu_1 \rho + w \tag{2}$$

where $w = (w_1, \ldots, w_n)$ and $w_i = 0$ for all $i \in S$. The vector w satisfies

$$(w, \sigma) = 1 - \mu_1 \tag{3}$$

because $1 = (\rho, \sigma) = (\rho, \sigma B) = (\rho B, \sigma) = \mu_1(\rho, \sigma) + (w, \sigma)$.

By the variational principle, $\lambda_2 = \max_{x \perp \sigma} (xB, x)/(x, x)$. Define $z = \rho - \sigma$. Since z is orthogonal to σ, we obtain that $\lambda_2 \geq (zB, z)/(z, z)$. Expanding this inequality via (2) and (3), we obtain that

$$\begin{aligned}\lambda_2 &\geq \frac{((\rho - \sigma)B, \rho - \sigma)}{(\rho - \sigma, \rho - \sigma)} = \frac{(\mu_1\rho + w - \sigma, \rho - \sigma)}{(\rho - \sigma, \rho - \sigma)} \\ &= \frac{\mu_1(\rho, \rho) - 1 - \mu_1 - (1 - \mu_1) + 1}{(\rho, \rho) - 1} = \frac{\mu_1(\rho, \rho) - 1}{(\rho, \rho) - 1}.\end{aligned}$$

Hence,

$$1 - \lambda_2 \leq 1 - \frac{\mu_1(\rho, \rho) - 1}{(\rho, \rho) - 1} = \frac{(\rho, \rho)(1 - \mu_1)}{(\rho, \rho) - 1}$$

or

$$(1 - \mu_1) \geq (1 - \lambda_2)\left(1 - \frac{1}{(\rho, \rho)}\right).$$

We now need a lower bound on (ρ, ρ). From (1), $\sum_{i \in S} \rho_i \sqrt{\pi_i} = 1$, so by Cauchy's inequality,

$$1 = \sum_{i \in S} \rho_i \sqrt{\pi_i} \leq \sqrt{\sum_{i \in S} \rho_i^2 \sum_{i \in S} \pi_i}.$$

Hence, $(\rho, \rho) \geq 1/\sum_{i \in S} \pi_i$, so finally,

$$1 - \mu_1 \geq (1 - \lambda_2)\Big(1 - \sum_{i \in S} \pi_i\Big) \geq (1 - \lambda_{\max})\Big(1 - \sum_{i \in S} \pi_i\Big).$$

Part (c) is a consequence of the interlacing of eigenvalues of submatrices. (Once again, this claim is standard.) To prove the claim, we use another form of the variational formula, which states that if $\lambda_2(M)$ is the second largest eigenvalue of the symmetric matrix Q, and $L^{(1)}$ is any linear space of dimension 1, then

$$\lambda_2(Q) = \min_{L^{(1)}} \max_{x \perp L^{(1)}} \frac{(Qx, x)}{(x, x)}.$$

Let σ be the principal eigenvector of the symmetric matrix B and let the vector $\bar{\sigma}$ have components $\bar{\sigma}_i = \sigma_i$ if $i \in S$, and $\bar{\sigma}_i = 0$ if $i \notin S$. Also, let $\tilde{\sigma}$ be the s element vector corresponding to the components of σ with indices in S. Finally, let x denote an s element vector and y denote an n element vector. We have

$$\begin{aligned}\mu_2 = \min_{L^{(1)}} \max_{x \perp L^{(1)}} \frac{(Qx, x)}{(x, x)} &\le \max_{x \perp \tilde{\sigma}} \frac{(Qx, x)}{(x, x)} \\ = \max_{\substack{y \perp \bar{\sigma} \\ y_i = 0, i \notin S}} \frac{(By, y)}{(y, y)} &\le \max_{y \perp \sigma} \frac{(By, y)}{(y, y)} = \lambda_2 \le \lambda_{\max}.\end{aligned}$$

∎

The next lemma is standard, but presented for completeness.

Lemma 2. *Let M be an irreducible and aperiodic, reversible Markov chain on n states with transition probability matrix P, and stationary distribution $(\pi_1, \pi_2, \ldots, \pi_n)$. Let $1 = \lambda_1 > \lambda_2 \ge \cdots \ge \lambda_n$ be the eigenvalues of P. Let S be a subset of the states of size s. Let R be the submatrix of P corresponding to the set of indices $S \times S$. Let the eigenvalues of R be $\mu_1 \ge \mu_2 \ge \cdots \ge \mu_s$ and let ρ_1 be a positive eigenvector of R corresponding to μ_1. Then*

$$R_{jk}^{(t)} \le \frac{\rho_{1j}\rho_{1k}}{\pi_j \left(\sum_l \rho_{1l}\right)^2} \mu_1^t + O\left(\sqrt{\frac{\pi_k}{\pi_j}} \mu_{\max}^t\right)$$

where $\mu_{\max} = \max_{2 \le i \le s} |\mu_i|$, and

$$P_{jk}^{(t)} = \pi_k + O\left(\sqrt{\frac{\pi_k}{\pi_j}} \lambda_{\max}^t\right),$$

where $\lambda_{\max} = \max_{2 \le i \le s} |\lambda_i|$.

Proof: Without loss of generality assume $S = \{1, \ldots, s\}$. Let $\tilde{D}$ be the s by s diagonal matrix given by $\tilde{d}_{ii} = 1/\pi_i$. As in Lemma 1, we associate with R the symmetric matrix $Q = \tilde{D}^{-1/2} R \tilde{D}^{1/2}$, which has the same eigenvalues as R. Let $\beta_1, \ldots, \beta_s$ be a set of orthonormal eigenvectors of Q with corresponding eigenvalues $\mu_1, \ldots, \mu_s$. We can always take $\beta_1 = \rho_1 \tilde{D}^{1/2} / \|\rho_1 \tilde{D}^{1/2}\|$.

Applying the spectral decomposition,

$$\begin{aligned} Q_{jk}^{(t)} &= \beta_{1j}\beta_{1k}\mu_1^t + \sum_{2 \le l \le s} \beta_{lj}\beta_{lk}\mu_l^t \le \beta_{1j}\beta_{1k}\mu_1^t + \sum_{2 \le l \le s} |\beta_{lj}|\,|\beta_{lk}|\,\left|\mu_l^t\right| \\ &\le \beta_{1j}\beta_{1k}\mu_1^t + \mu_{\max}^t \sqrt{\sum_{2 \le l \le s} |\beta_{lj}|^2 \sum_{2 \le l \le s} |\beta_{lk}|^2} \le \beta_{1j}\beta_{1k}\mu_1^t + O(\mu_{\max}^t), \end{aligned}$$

where the last two inequalities are obtained by applying Cauchy's inequality and the orthonormality of the β's. Hence

$$Q_{jk}^{(t)} = \frac{\rho_{1j}\rho_{1k}}{\sqrt{\pi_j \pi_k}} \frac{1}{\sum_l \rho_{1l}^2 / \pi_l} \mu_1^t + O(\mu_{\max}^t).$$

By construction

$$R_{jk}^{(t)} = \sqrt{\frac{\pi_k}{\pi_j}} Q_{jk}^{(t)} = \frac{\rho_{1j}\rho_{1k}}{\pi_j \sum_l \rho_{1l}^2/\pi_l} \mu_1^t + O\left(\sqrt{\frac{\pi_k}{\pi_j}} \mu_{\max}^t\right).$$

Finally, by Cauchy's inequality

$$\sum_{1 \le l \le s} \rho_{1l} = \sum_{1 \le l \le s} \frac{\rho_{1l}}{\sqrt{\pi_l}} \sqrt{\pi_l} \le \sqrt{\sum_{1 \le l \le s} \frac{\rho_{1l}^2}{\pi_l} \sum_{1 \le l \le s} \pi_l} \le \sqrt{\sum_{1 \le l \le s} \frac{\rho_{1l}^2}{\pi_l}}$$

and the lemma follows in the general case.

For the case $s = n$, we take $\rho_{1l} = \pi_l$ and the inequality above becomes an equality.

∎

Lemma 3. *Let $M = (X_0, X_1, \ldots)$ be an irreducible and aperiodic, reversible Markov chain with transition probability matrix P. We rename the states such that $\pi_1 \le \pi_2 \le \cdots \le \pi_n$ is the stationary distribution of M. Let $1 = \lambda_1 > \lambda_2 \ge \cdots \ge \lambda_n$ be the eigenvalues of P. Let S be a subset of states and let R be the submatrix of P corresponding to the set of indices $S \times S$. Let $T(i, S)$ be the time*

$$T(i,S) = \inf_t (X_t \notin S \mid X_0 = i \in S).$$

Then

$$\mathbf{E}(T(i,S)) \le \left(\frac{(2+\epsilon)\ln n - \ln \pi_1}{1 - \lambda_{\max}} + \frac{1}{1-\mu_1}\right)(1 + o(1)),$$

where $\lambda_{\max} = \max_{2\le i\le n} |\lambda_i|$, $\epsilon > 0$, *and* μ_1 *is the largest eigenvalue of the matrix* R.

Proof: Let the eigenvalues of R be $\mu_1 \ge \mu_2 \ge \cdots \ge \mu_s$ and let ρ_1 be a positive eigenvector of R corresponding to μ_1. The value $T(i,S)$ is the time when a random walk starting at state $i \in S$ first moves to some state in $V - S$. The probability that a random walk starting at $i \in S$ does not leave S within the first t steps is $\sum_{j\in S} R_{ij}^{(t)}$. Let τ_1 and τ_2 be positive values to be determined later. Then

$$\begin{aligned}\mathbf{E}(T(i,S)) = \sum_{t\ge 0}\sum_{k\in S} R_{ik}^{(t)} &\le \tau_1 + \sum_{j\in S} R_{ij}^{(\tau_1)} \sum_{k\in S}\sum_{t\ge 0} R_{jk}^{(t)} \\ &\le \tau_1 + \sum_{j\in S} P_{ij}^{(\tau_1)} \sum_{k\in S}\sum_{t\ge 0} R_{jk}^{(t)}.\end{aligned}$$

By Lemma 2,

$$\begin{aligned}\mathbf{E}(T(i,S)) &\le \tau_1 + \sum_{j\in S}\left(\pi_j + O\left(\sqrt{\frac{\pi_j}{\pi_i}}\lambda_{\max}^{\tau_1}\right)\right)\left(\tau_2 + \sum_{k\in S}\sum_{t\ge\tau_2} R_{jk}^{(t)}\right) \\ &\le \tau_1 + \sum_{j\in S}\left(\pi_j + O\left(\sqrt{\frac{\pi_j}{\pi_i}}\lambda_{\max}^{\tau_1}\right)\right)\times \\ &\quad \left(\tau_2 + \sum_{k\in S}\sum_{t\ge\tau_2}\left(\frac{\rho_{1j}\rho_{1k}}{\pi_j\left(\sum_l \rho_{1l}\right)^2}\mu_1^t + O\left(\sqrt{\frac{\pi_k}{\pi_j}}\lambda_{\max}^t\right)\right)\right) \\ &\le \tau_1 + \tau_2 + \frac{\mu_1^{\tau_2}}{1-\mu_1} + O\left(\frac{s\lambda_{\max}^{\tau_2}}{1-\lambda_{\max}}\right) + O\left(\frac{\tau_2\sqrt{s}\lambda_{\max}^{\tau_1}}{\sqrt{\pi_1}}\right) \\ &\quad + O\left(\frac{\lambda_{\max}^{\tau_1}}{\pi_1}\frac{\mu_1^{\tau_2}}{1-\mu_1}\right) + O\left(\frac{s^{3/2}}{\sqrt{\pi_1}}\frac{\lambda_{\max}^{\tau_1+\tau_2}}{1-\lambda_{\max}}\right).\end{aligned}$$

Now set

$$\tau_1 = \left\lceil \log_{\lambda_{\max}}\left(\frac{\pi_1}{n}\right)\right\rceil$$

and

$$\tau_2 = \left\lceil \log_{\lambda_{\max}}\left(\frac{1}{n^{1+\epsilon}}\right)\right\rceil,$$

where $\epsilon > 0$. With these values

$$\mathbf{E}(T(i,S)) \le \left(\frac{(2+\epsilon)\ln n - \ln \pi_1}{1-\lambda_{\max}} + \frac{1}{1-\mu_1}\right)(1+o(1)).$$

∎

We now put together these lemmas to obtain the main theorem for chains that are irreducible, aperiodic, and reversible .

Theorem 4. *Let M be an irreducible and aperiodic, reversible Markov chain on n states with transition probability matrix P. We rename the states such that $\pi_1 \leq \pi_2 \leq \cdots \leq \pi_n$ is the stationary distribution of M. Let $1 = \lambda_1 > \lambda_2 \geq \cdots \geq \lambda_n$ be the eigenvalues of P and let $\lambda_{\max} = \max_{2 \leq i \leq n} |\lambda_i|$. Let $\epsilon > 0$ be a constant. Then the cover time C starting from any state satisfies*

$$\mathbf{E}(C) \leq \frac{1}{1-\lambda_{\max}}\left((2+\epsilon)n\ln n - n\ln\pi_1 + \sum_{1\leq i<n}\Big(\sum_{1\leq j\leq i}\pi_i\Big)^{-1}\right)(1+o(1)).$$

Proof: Let T_i be the first time that i states have been seen. Then $C = T_n$ and $\mathbf{E}(C) = \sum_{1\leq i\leq n-1}\mathbf{E}(T_{i+1}-T_i)$. Let S_i be any set of states of cardinality i and let $\mu_1(S_i)$ be the largest eigenvalue of the submatrix of P corresponding to the set of indices $S_i \times S_i$. Let $\mu_1(i) = \max_{S_i}\mu_1(S_i)$. Then via Lemma 3

$$\begin{aligned}\mathbf{E}(C) &\leq \sum_{1\leq i\leq n-1}\left(\frac{(2+\epsilon)\ln n - \ln\pi_1}{1-\lambda_{\max}} + \frac{1}{1-\mu_1(i)}\right)(1+o(1))\\ &\leq \frac{1}{1-\lambda_{\max}}\left((2+\epsilon)n\ln n - n\ln\pi_1 + \sum_{1\leq i<n}\Big(\sum_{1\leq j\leq i}\pi_j\Big)^{-1}\right)(1+o(1)),\end{aligned}$$

where the final inequality follows from Lemma 1(b). ∎

This theorem easily generalizes to Markov chains that are irreducible but not necessarily aperiodic as follows.

Theorem 5. *Let M be an irreducible reversible Markov chain on n states with transition probability matrix P. Let $1 = \lambda_1 > \lambda_2 \geq \cdots \geq \lambda_n$ be the eigenvalues of P. Let $\pi = (\pi_1, \pi_2, \ldots, \pi_n)$ be the eigenvector corresponding to the eigenvalue 1, normalized so that $\sum_i \pi_i = 1$. (If M is aperiodic then π is the stationary distribution of M.) We rename the states such that $\pi_1 \leq \pi_2 \leq \cdots \leq \pi_n$. Let $\epsilon > 0$ be a constant. Then the cover time C starting from any state satisfies*

$$\mathbf{E}(C) \leq \frac{1}{1-\lambda_2}\left((2+\epsilon)n\ln n - n\ln\pi_1 + \sum_{1\leq i\leq n}\Big(\sum_{1\leq j\leq i}\pi_j\Big)^{-1}\right)(1+o(1)).$$

Proof: Consider the Markov chain M' with transition probability matrix $P' = \frac{1}{2}(I+P)$. Clearly M' is irreducible, aperiodic, and reversible and $\mathbf{E}(C_i(M')) = 2\,\mathbf{E}(C_i(M))$ from any starting state i. Furthermore, if $1 = \lambda_1 > \lambda_2 \geq \cdots \geq \lambda_n$ are the eigenvalues of P, then P' has eigenvalues $1 = \lambda_1' \geq \lambda_2' \geq \cdots \lambda_n'$, where $\lambda_i' = \frac{1}{2}(1+\lambda_i)$. Finally, the stationary distribution of M' is π. Hence the result follows immediately by applying the previous theorem to M'. ∎

Corollary 6. *Let M be any symmetric, irreducible Markov chain with transition probability matrix P. Then its expected cover time starting from any state satisfies*

$$\mathbf{E}(C) \le \frac{(2+\epsilon) n \ln n}{1-\lambda_2}\big(1+o(1)\big),$$

where λ_2 is the second largest eigenvalue of P and $\epsilon > 0$.

Proof: We apply the previous theorem. In this case $\pi_i = 1/n$ for all i. ∎

We now apply this result to expander graphs. An (n,d,k)-*expander* is a d-regular graph $G(V,E)$ on n vertices, such that for every set $X \subset V$ with $|X| \le n/2$, its neighborhood satisfies $|\Gamma(X) - X| \ge k\,|X|$. (For a subset X of V, the *neighborhood* of X is defined as $\Gamma(X) = \{v \in V \mid (v,x) \in E \text{ for some } x \in X\}$.)

Alon [1] has proven that if G is an (n,d,k)-expander then the transition probability matrix for a random walk on G has second largest eigenvalue $\lambda_2 < 1 - k^2/\big(d(4+2k^2)\big)$.

Corollary 7. *Consider a random walk on a d-regular expander graph. Its expected cover time starting from any state satisfies $\mathbf{E}(C) = O(n \log n)$.* ∎

Corollary 8. *Let P be the transition probability matrix corresponding to a random walk on a connected graph G. The expected time to visit all states in G satisfies*

$$\mathbf{E}(C) \le \frac{n^2 \ln n}{1-\lambda_2}\big(1+o(1)\big),$$

where λ_2 is the second largest eigenvalue of P.

Proof: Let d_i be the degree of node i in G. Then $\pi_i = d_i / \sum_i d_i$ and therefore $\pi_i \ge 1/n^2$ for all i. ∎

In general the bound of Corollary 8 cannot be improved as the following example shows: Let I be the n by n identity matrix and let J be the n by n matrix consisting of all 1's. Consider the graph G on $2n$ vertices with adjacency matrix A, where

$$A = \begin{pmatrix} J & I \\ I & 0 \end{pmatrix}.$$

This graph has n vertices connected to form a complete subgraph, each connected to one other distinct vertex. It is not difficult to see that the cover time of this graph is $\Theta(n^2 \log n)$. Since $\lambda_{\max} = 1/\sqrt{n}$, Corollary 8 gives a tight bound for this graph.

3. Application to universal sequences.

Let G be a graph on n vertices. At each vertex v, let the edges incident with v be given the distinct labels $1, \ldots, d(v)$, where $d(v)$ is the degree of v. The labels at the two ends of an edge are not necessarily equal, that is, each edge is labeled twice. A sequence σ in $\{1, \ldots, n\}^*$ is said to *traverse* G from v if, by starting from v and following the sequence of edge labels σ, one covers all the vertices of G. (We make the convention that a label not present at the current vertex results in a null move.) Let $\mathcal{G}$ be a collection of graphs. We say that σ is *universal for* $\mathcal{G}$ if it traverses every graph in $\mathcal{G}$, from every starting point v.

The original motivation for obtaining bounds on the lengths of universal sequences comes from the attempt to prove bounds on the deterministic space complexity of the reachability problem for undirected graphs. Aleliunas et al.[6] have shown that there exist sequences of length $O(d^2 n^3 \log n)$ that are universal for all labeled graphs on n vertices with degree bounded by d. For d-regular graphs the bound can be improved to $O(dn^3 \log n)$ via the results in [10]. Using the results of Lemmas 1 and 3, we show that there exist sequences of length $O(dn^2 \log n)$ that are universal for graphs with λ_2 bounded away from 1.

We start by considering connected graphs that are not bipartite (and hence the Markov chain is aperiodic) and then generalize the result to arbitrary connected graphs.

Lemma 9. *Consider a random walk on a d-regular graph G on n vertices. Assume that G is not bipartite. Let the corresponding Markov chain have transition probability matrix P with eigenvalues $\lambda_1 \geq \cdots \geq \lambda_n$. Let $\lambda_{\max} = \max_{2 \leq i \leq n} \lambda_i$. Let $\epsilon > 0$. Then the cover time C starting from any state satisfies*

$$\mathbf{Pr}\left(C > \frac{4n \ln n}{1 - \lambda_{\max}} + t\right) \leq \exp\left(n \ln(t+n) - \frac{t(1-\lambda_{\max})}{n}\right)(1 + o(1))$$

Proof: Let v be any starting position. Let the random variable T_k be the first time that k vertices are seen, starting from v. Let S be any connected set of states of cardinality k and let $T(i, S)$ be the time that a random walk starting at state $i \in S$ first leaves S. Fix $\tau > 0$. Then

$$\mathbf{Pr}\left(T_{k+1} - T_k > \tau + t\right) \leq \max_{\substack{i,S \\ |S|=k; i \in S}} \mathbf{Pr}\left(T(i,S) > \tau + t\right).$$

Let S be the subset of vertices and i the state in S for which the right hand side above is maximized, and let R be the submatrix of P corresponding to the indices $S \times S$. Let μ_1

be the largest eigenvalue of the matrix R and let $\mu_{\max} = \max_{2\leq j\leq s}|\mu_j|$. Then following the proof of Lemma 3 (applying Lemma 2 and the fact that regular graphs have stationary probabilities all equal to $1/n$) we derive

$$\begin{aligned}\mathbf{Pr}\left(T_{k+1}-T_k>\tau+t\right) &= \sum_{j\in S} R_{ij}^{\tau+t} = \sum_{h\in S} R_{ih}^{\tau/2}\sum_{j\in S} R_{hj}^{\tau/2+t} \leq \sum_{h\in S} P_{ih}^{\tau/2}\sum_{j\in S} R_{hj}^{\tau/2+t}\\ &\leq \sum_{h\in S}\left(\pi_h + O(\lambda_{\max}^{\tau/2})\right)\sum_{j\in S}\left(\frac{\rho_{1h}\rho_{1j}}{\pi_h(\sum_l \rho_1 l)^2}\mu_1^{\tau/2+t} + O(\mu_{\max}^{\tau/2+t})\right).\end{aligned}$$

Setting $\tau = 4\ln n/(1-\lambda_{\max})$, we get $O(\lambda_{\max}^{\tau/2}) = O(1/n^2)$ and since $\mu_{\max} \leq \lambda_{\max}$ and $\mu_{\max} \leq \mu_1$, we obtain that

$$\begin{aligned}\mathbf{Pr}\left(T_{k+1}-T_k>\tau+t\right) &\leq \sum_{h\in S}\left(\frac{1}{n}+O\left(\frac{1}{n^2}\right)\right)\sum_{j\in S}\left(n\frac{\rho_{1h}\rho_{1j}}{(\sum_l \rho_1 l)^2}\mu_1^t + O\left(\frac{\mu_1^t}{n^2}\right)\right)\\ &\leq \mu_1^t(1+o(1)).\end{aligned}$$

Finally by applying Lemma 1(b),

$$\begin{aligned}\mathbf{Pr}\left(T_{k+1}-T_k>\tau+t\right) &\leq \left(1-(1-\lambda_{\max})\frac{n-k}{n}\right)^t\big(1+o(1)\big)\\ &\leq \left(1-\frac{1-\lambda_{\max}}{n}\right)^t\big(1+o(1)\big).\end{aligned}$$

Let $\alpha = \big(1-(1-\lambda_{\max})/n\big)$. Then

$$\mathbf{Pr}\left(T_{k+1}-T_k>\tau+t\right) \leq \left\langle x^t\right\rangle \frac{1}{1-\alpha x}\big(1+o(1)\big),$$

where the notation $\langle x^t\rangle f(x)$ means the coefficient of x^t in the formal power series expansion of f. Therefore

$$\begin{aligned}\mathbf{Pr}\left(C>\tau+t\right) &\leq \left\langle x^t\right\rangle \frac{1}{(1-\alpha x)^n}\big(1+o(1)\big) = \binom{n+t-1}{t}\alpha^t\big(1+o(1)\big)\\ &\leq (t+n)^n\alpha^t\big(1+o(1)\big)\\ &\leq \exp\left(n\ln(t+n)-\frac{t(1-\lambda_{\max})}{n}\right)\big(1+o(1)\big).\end{aligned}$$

∎

Theorem 10. *Let $\mathcal{G}$ be the collection of all d-regular graphs on n vertices with $\lambda_2 \leq 1-c$ for some constant $c > 0$. There exist universal sequences for $\mathcal{G}$ of length $O(dn^2 \log n)$.*

Proof: First, we generalize the previous lemma as in the proof of Theorem 5, to get rid of the requirement that G is not bipartite. We obtain that

$$\mathbf{Pr}\left(C > \frac{8n\ln n}{1-\lambda_2} + t\right) \leq \exp\left(n\ln(t+n) - \frac{t(1-\lambda_2)}{2n}\right)\bigl(1+o(1)\bigr).$$

Therefore the probability that a randomly chosen sequence $\tilde{\sigma}$ from $\{0,1,\ldots,d-1\}^*$ of length $8n\ln n/(1-\lambda_2) + \alpha dn^2 \ln n$ does not traverse any fixed graph $G \in \mathcal{G}$ is bounded by $\exp\bigl((2-\alpha cd/2)n\ln n + O(n\log\log n)\bigr)$. Since the number of graphs in $\mathcal{G}$ is at most n^{nd}, the probability that there exists a $G \in \mathcal{G}$ that $\tilde{\sigma}$ doesn't traverse is at most $\exp\bigl((2+d-\alpha cd/2)n\ln n + O(n\log logn)\bigr)$, which is less than 1 for an appropriately chosen value of α. Hence, there exists a sequence σ of length $O(dn^2\log n)$ that is universal for $\mathcal{G}$. ∎

4. The lower bound for rapidly-mixing chains.

In this section, we will show that a rapidly-mixing doubly-stochastic Markov chain on n states, requires $\Omega(n\log n)$ expected steps to visit all states. (A rapidly-mixing chain is a chain for which $\lambda_{\max} \leq 1-\epsilon$, for a constant ϵ.) For example, the Markov chain corresponding to a random walk on a regular expander graph satisfies these conditions.

We use the fact that a rapidly-mixing Markov chain gets close to the stationary distribution in time proportional to $\log n$. Since the Markov chain is doubly-stochastic, the stationary distribution is uniform. Therefore, if s states have been seen, the expected number of new states seen in the next k steps is close to $k(n-s)/n$, for appropriately chosen values of k. This is essentially what we expect if we were to apply a coupon collector's argument. By judiciously choosing k as a function of s and repeating the argument, one can prove an $\Omega(n\log n)$ lower bound on the time to see all n states.

Lemma 11. *Let M be a Markov chain on n states with transition probability matrix P. Assume that P is doubly-stochastic. Let the random variable T_s be the first time that s states are seen. Suppose τ is such that $\left|P_{ij}^{(t)} - \frac{1}{n}\right| \leq \frac{1}{n^3}$ for $t \geq \tau$. Then*

$$\mathbf{E}(T_{s+r} - T_s) \geq \frac{1}{2}\left(\frac{nr}{n-s}\right) - \frac{\tau s}{n-s} + O(1),$$

for integers s and r satisfying $0 \leq s < n$, and $0 \leq r \leq n-s$.

Proof: Suppose that the random walk has visited a set of states S of cardinality $|S| = s$. Define N_k to be the number of new states seen in the next k steps and $R_k(i)$ to be the number of steps spent in S out of k steps, starting in some state $i \in S$. Then $R_k(i) = \sum_{j \in S} \sum_{1 \le m \le k} X_{ij}^m$, where X_{ij}^m is an indicator random variable, which is 1 if the random walk has moved from state i to state j after m steps, and 0 otherwise. Therefore,

$$\mathbf{E}(R_k(i)) = \sum_{j \in S} \sum_{1 \le m \le k} P_{ij}^{(m)} \ge \sum_{j \in S} \sum_{\tau \le m \le k} P_{ij}^{(m)} \ge (k - \tau)s \left(\frac{1}{n} - \frac{1}{n^3} \right).$$

On the other hand

$$\mathbf{E}(N_k) \le k - \mathbf{E}(R_k(i)) \le \frac{n-s}{n} k + \frac{\tau s}{n} + \frac{(k-\tau)s}{n^3}. \tag{4}$$

Recall that T_r is the time needed to see r different states. Plainly

$$\mathbf{Pr}(T_{s+r} - T_s \le k) = \mathbf{Pr}(N_k \ge r) \le \frac{\mathbf{E}(N_k)}{r},$$

and hence

$$\mathbf{E}(T_{s+r} - T_s) = \sum_{k \ge 0} \mathbf{Pr}(T_{s+r} - T_s \ge k) \ge r + \sum_{r < k \le U} \left(1 - \frac{\mathbf{E}(N_k)}{r} \right)$$

where U is an upper limit to be determined. Replacing $\mathbf{E}(N_k)$ from equation (4), we obtain

$$\begin{aligned} \mathbf{E}(T_{s+r} - T_s) &\ge r + \sum_{r < k \le U} \left(1 - \frac{n-s}{n} \frac{k}{r} - \frac{\tau s}{nr} - \frac{(k-\tau)s}{rn^3} \right) \\ &\ge U - \frac{U(U+1)}{2} \left(\frac{n-s}{nr} + \frac{s}{rn^3} \right) - U \frac{\tau s}{nr}. \end{aligned}$$

Taking $U = \left\lfloor \frac{nr}{n-s} \right\rfloor$, we get

$$\mathbf{E}(T_{s+r} - T_s) \ge \frac{nr}{n-s} - \frac{1}{2} \left(\frac{nr}{n-s} \right) - \frac{\tau s}{n-s} + O(1).$$

∎

Lemma 12. *Let M be a Markov chain on n states with transition probability matrix P. Assume that P is doubly-stochastic. Let the random variable T_s be the first time that s states are seen. Suppose τ is such that $\left| P_{ij}^{(t)} - \frac{1}{n} \right| \le \frac{1}{n^3}$ for $t \ge \tau$. Then*
(a)

$$\mathbf{E}(T_{\lfloor s+r \rfloor} - T_{\lfloor s \rfloor}) \ge \frac{1}{2} \left(\frac{nr}{n-s} - \frac{nr}{(n-s)^2} - \frac{n}{n-s} \right) - \frac{\tau s}{n-s} + O(1),$$

for reals s and r satisfying $0 \le s < n$, and $0 \le r \le n-s$.

(b)

$$\mathbf{E}(T_{\lfloor\frac{(m+1)n}{m+2}\rfloor} - T_{\lfloor\frac{mn}{m+1}\rfloor}) \geq \frac{1}{2}\frac{n}{m+2} + O(\tau m),$$

for any positive integer m.

Proof: For part (a) we apply the previous lemma and the fact that

$$\frac{n\lfloor r\rfloor}{n-\lfloor s\rfloor} \geq \frac{n(r-1)}{n-s+1} \geq \frac{nr}{n-s+1} - \frac{n}{n-s} \geq \frac{nr}{n-s} - \frac{nr}{(n-s)^2} - \frac{n}{n-s}.$$

For part (b), we substitute in part (a) with $s = \frac{mn}{m+1}$ and $r = \frac{n}{(m+1)(m+2)}$. We obtain

$$\begin{aligned}\mathbf{E}(T_{\lfloor\frac{(m+1)n}{m+2}\rfloor} - T_{\lfloor\frac{mn}{m+1}\rfloor}) &\geq \frac{1}{2}\left(\frac{n}{m+2} - \frac{m+1}{m+2} - (m+1)\right) - \tau m + O(1) \\ &\geq \frac{1}{2}\frac{n}{m+2} + O(\tau m).\end{aligned}$$

∎

We now apply Lemma 12 to sets of increasing size to obtain the main result.

Theorem 13. *Let M be an irreducible and aperiodic, reversible Markov chain on n states with transition probability matrix P. Assume that P is doubly-stochastic and has eigenvalues $1 = \lambda_1 > \lambda_2 \geq \cdots \geq \lambda_n$. Let $\lambda_{\max} = \max_{2\leq i\leq n}|\lambda_i|$. Suppose that $\lambda_{\max} \leq 1 - n^{\epsilon-1}$ for $\epsilon > 0$. Then the cover time C, starting from any state, satisfies $\mathbf{E}(C) = \Omega(n\log n)$.*

Proof: From Lemma 2,

$$P_{ij}^{(t)} = \frac{1}{n} + O(\lambda_{\max}^t).$$

Since $(1-\frac{1}{n^{1-\epsilon}})^t \leq \frac{1}{n^3}$ for $t > 3n^{1-\epsilon}\ln n$, we have $\left|P_{ij}^{(t)} - \frac{1}{n}\right| \leq \frac{1}{n^3}$ for $t \geq n^{1-\delta}$, $\delta > 0$.

Applying the previous lemma

$$\begin{aligned}\mathbf{E}(C) &\geq \sum_{0\leq m\leq U}\mathbf{E}(T_{\lfloor\frac{(m+1)n}{m+2}\rfloor} - T_{\lfloor\frac{mn}{m+1}\rfloor}) \geq \sum_{0\leq m\leq U}\left(\frac{1}{2}\frac{n}{m+2} + O(\tau m)\right) \\ &\geq \frac{n\ln U}{2} + O(n + \tau U^2).\end{aligned}$$

for any integer $U > 0$ and $\tau \geq n^{1-\delta}$.

Choosing $U = \sqrt{n/\tau}$, we obtain

$$\mathbf{E}(C) \geq \frac{n\ln(n/\tau)}{4} + O(n).$$

Since we can take $\tau = n^{1-\delta}$, we obtain $\mathbf{E}(C) = \Omega(n\log n)$. ∎

5. The lower bound for slowly-mixing chains.

In order to show that the cover time for slowly-mixing reversible Markov chains (chains for which $\lim_{n\to\infty} \lambda_2 = 1$) is large, it suffices to show that there exists a pair of vertices i and j such that the expected first passage time from i to j is large. We obtain such a result by proving an identity that relates the weighted average over all pairs of vertices i and j of $\mathbf{E}(H_{ij})$, the expected hitting time, to the eigenvalues of the transition probability matrix.

Lemma 14. *Let M be an irreducible and aperiodic, reversible Markov chain on n states, with transition probability matrix P and stationary distribution $(\pi_1, \ldots, \pi_n)$. Let $P_{ij}(z) = \sum_{m\geq 0} p_{ij}^{(m)} z^m$ be the generating function for the probability that a particle in state i moves to state j in m steps. Then*

$$P_{ij}(z) = \delta_{ij} + \pi_j \frac{z}{1-z} + A_{ij}(z),$$

where

$$A_{ij}(z) = \sum_{2\leq k\leq n} \varphi_{ki}\psi_{kj} \frac{\lambda_k z}{(1-\lambda_k z)},$$

and $1 = \lambda_1 > \lambda_2 \geq \ldots \geq \lambda_n$ are the eigenvalues of P, and $\psi_1, \ldots, \psi_n$ and $\varphi_1, \ldots, \varphi_n$ are a biorthonormal basis of left and right eigenvectors corresponding to $\lambda_1, \ldots, \lambda_n$.

Proof: By the spectral theorem

$$P_{ij}(z) = \sum_{m\geq 0} \sum_{1\leq k\leq n} \varphi_{ki}\psi_{kj}(\lambda_k z)^m = \delta_{ij} + \sum_{1\leq k\leq n} \varphi_{ki}\psi_{kj} \frac{\lambda_k z}{(1-\lambda_k z)}.$$

Since the right eigenvector φ_1 can be taken as $(1, 1, \ldots, 1)$ and the corresponding left eigenvector ψ_1 is $(\pi_1, \ldots, \pi_n)$, we obtain the claimed result. ∎

Theorem 15. *Let M be an irreducible and aperiodic, reversible Markov chain on n states, with transition probability matrix P and stationary distribution $(\pi_1, \ldots, \pi_n)$. Then*

$$\sum_{\substack{1\leq i\leq n \\ 1\leq j\leq n}} \pi_i \pi_j \, \mathbf{E}(H_{ij}) = \sum_{2\leq k\leq n} \frac{1}{1-\lambda_k},$$

where the random variable H_{ij} is the hitting time (or first passage time) from i to j, and $1 = \lambda_1 > \lambda_2 \geq \ldots \geq \lambda_n$ are the eigenvalues of P.

Proof: We start from the well-known identity

$$H_{ij}(z) = \sum_{m\geq 0} \mathbf{Pr}(H_{ij} = m) z^m = \frac{P_{ij}(z)}{P_{jj}(z)}.$$

Because M is reversible,

$$\pi_i P_{ij}(z) = \pi_j P_{ji}(z).$$

Therefore, via the previous lemma, we obtain that

$$\sum_{\substack{1\le i\le n\\ 1\le j\le n}} \pi_i \pi_j H_{ij}(z) = \sum_{1\le j\le n} \frac{\pi_j}{P_{jj}(z)} \sum_{1\le i\le n} \pi_i P_{ij}(z) = \sum_{1\le j\le n} \frac{\pi_j^2}{P_{jj}(z)} \sum_{1\le i\le n} P_{ji}(z)$$

$$= \sum_{1\le j\le n} \frac{\pi_j^2}{P_{jj}(z)} \frac{1}{(1-z)} = \sum_{1\le j\le n} \frac{\pi_j^2}{(1-z) + \pi_j z + A_{jj}(z)(1-z)}.$$

Differentiating with respect to z and setting $z \leftarrow 1$, we obtain

$$\sum_{\substack{1\le i\le n\\ 1\le j\le n}} \pi_i \pi_j \, \mathbf{E}(H_{ij}) = \sum_{1\le j\le n} \big(1 - \pi_j + A_{jj}(1)\big) = n - 1 + \sum_{2\le k\le n} \frac{\lambda_k}{1-\lambda_k}.$$

∎

Corollary 16. *Let M be an irreducible and aperiodic, reversible Markov chain on n states, with transition probability matrix P. Let $1 = \lambda_1 > \lambda_2 \ge \cdots \ge \lambda_n$ be the eigenvalues of P. If $\lambda_2 \ge 1 - \frac{1}{n \ln n}$ then the expected cover time starting from the stationary distribution of M is $\Omega(n \log n)$.*

Proof: If $\lambda_2 \ge 1 - \frac{1}{n \ln n}$, then Theorem 15 implies that $\sum_{1\le i\le n} \sum_{1\le j\le n} \pi_i \pi_j \, \mathbf{E}(H_{ij}) = \Omega(n \log n)$. Hence, there is a j such that $\sum_{1\le i\le n} \pi_i \, \mathbf{E}(H_{ij}) = \Omega(n \log n)$. But the expected cover time starting from the stationary distribution is defined as

$$\mathbf{E}(C_\pi) = \sum_{1\le i\le n} \pi_i \, \mathbf{E}(C_i) \ge \sum_{1\le i\le n} \pi_i \, \mathbf{E}(H_{ij}) = \Omega(n \log n).$$

∎

The following corollary was also proven by Landau and Odlyzko [12] via a completely different approach.

Corollary 17. *Let P be the transition probability matrix of an irreducible and aperiodic, reversible Markov chain on n states, corresponding to a random walk on an undirected, connected graph G, with maximum degree d. Let $1 = \lambda_1 > \lambda_2 \ge \lambda_2 \ge \cdots \ge \lambda_n$ be the eigenvalues of P. Then*

$$\lambda_2 \le 1 - \frac{1}{n^2 d}.$$

Proof: Aleliunas et al.[6] show that for any undirected, connected graph with maximum degree d, $\mathbf{E}(C_i) \leq dn^2$, for all i. Thus we have

$$\sum_{2 \leq k \leq n} \frac{1}{1-\lambda_k} = \sum_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}} \pi_i \pi_j \, \mathbf{E}(H_{ij}) \leq \max_{i,j} \mathbf{E}(H_{ij}) \leq \max_i \mathbf{E}(C_i) \leq dn^2.$$

Clearly $1/(1-\lambda_k) \geq 0$, for $k \geq 2$. Hence, $1/(1-\lambda_2) \leq dn^2$, and the corollary follows. ∎

Mazo [14] considered a closely related measure to our average first passage time, denoted N, and defined by

$$N = \frac{1}{n(n-1)} \sum_{1 \leq i \leq n} \sum_{\substack{1 \leq j \leq n \\ j \neq i}} \mathbf{E}(H_{ij}).$$

Mazo was interested in determining which chains minimize and maximize the quantity N and showed via a rather difficult proof that for random walks on connected, undirected graphs, N is minimized for the complete graph. He also conjectured that it is maximized for the bar-bell graph. Using our formula, we prove both results for the quantity $N' = \sum_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}} \pi_i \pi_j \, \mathbf{E}(H_{ij})$.

Corollary 18. *Consider a a random walk on an undirected graph with n vertices and no self loops. Let*

$$N' = \sum_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}} \pi_i \pi_j \, \mathbf{E}(H_{ij}).$$

Then N' is minimized for the complete graph.

Proof: We have shown that

$$N' = \sum_{2 \leq k \leq n} \frac{1}{1-\lambda_k}. \tag{5}$$

Because the trace of the transition probability matrix is 0, and the trace is equal to the sum of the eigenvalues, we have $\sum_{2 \leq k \leq n} \lambda_k = -1$. But the function on the right-hand side of (5) is Schur-convex (see [13] pp. 54–58) and hence is minimized when $\lambda_k = -\frac{1}{n-1}$ for all k, $2 \leq k \leq n$. Since these are precisely the eigenvalues of the Markov chain corresponding to the random walk on the complete graph, the claim is proven. ∎

Corollary 19. *Consider a random walk on an undirected graph with n vertices. Let*

$$N' = \sum_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}} \pi_i \pi_j \, \mathbf{E}(H_{ij}).$$

Then $N' = O(n^3)$ for all graphs and $N' = \Theta(n^3)$ for the bar-bell graph.

Proof: Again we use the bound $\mathbf{E}(C_i) \le n^3$ for all i. Thus we have

$$\sum_{2 \le k \le n} \frac{1}{1-\lambda_k} = \sum_{\substack{1 \le i \le n \\ 1 \le j \le n}} \pi_i \pi_j \, \mathbf{E}(H_{ij}) \le \max_{i,j} \mathbf{E}(H_{ij}) \le \max_i \mathbf{E}(C_i) \le n^3$$

and so $N' = O(n^3)$ for all graphs. But for the bar-bell graph Landau and Odlyzko [12] have shown that $\lambda_2 = 1 - \frac{c}{n^3}$ (for some constant c), and the corollary follows. ∎

Acknowledgements. We wish to thank Alistair Sinclair who helped us improve the error term in Lemma 2, based on an idea presented in [15]. We are also indebted to Lyle Ramshaw for a thorough reading of the manuscript and many useful comments and corrections.

References

[1] N. Alon, "Eigenvalues and expanders," *Combinatorica*, **6**(1986), 83–96.

[2] D. Aldous, "Random walks on finite groups and rapidly mixing Markov chains," Seminaire de Probabilités XVII, Springer-Verlag, 1983, 243–297.

[3] D. Aldous, "On the time taken by random walks on finite groups to visit every state," *ZW*, **62**(1983), 361–374.

[4] D. Aldous, "Minimization algorithms and random walk on the d-cube," *Annals of Probability*, **11**(1983), 403–413.

[5] D. Aldous, "Lower bounds for covering times for reversible Markov Chains and random walks on graphs," manuscript, 1988.

[6] R. Aleliunas, R.M. Karp, R.J. Lipton, L. Lovasz and C. Rackoff, "Random walks, universal traversal sequences, and the complexity of maze problems," *Proceedings of the 20th IEEE Symposium on Foundations of Computer Science*, 1979, 218–233.

[7] R.B. Boppana, "Eigenvalues and Graph Bisection: An Average-case Analysis," *Proceedings of the 28th IEEE Symposium on Foundations of Computer Science*, 1987, 280–285.

[8] A.Z. Broder and A.R. Karlin, "Bounds on Cover Times," Proceedings of the 29th Symposium on Foundations of Computer Science.1988

[9] A.Z. Broder and E. Shamir, "On the second eigenvalue of random regular graphs," *Proceedings of the 28th IEEE Symposium on Foundations of Computer Science*, 1987, 286–294.

[10] J.D. Kahn, N. Linial, N. Nisan, and M.E. Saks, "On the cover time of random walks in graphs," manuscript, 1988.

[11] J.G. Kemeny and J.L. Snell, *Finite Markov Chains*, Van Nostrand, 1969.

[12] H.J. Landau and A.M. Odlyzko, "Bounds for Eigenvalues of certain stochastic matrices," *Linear Algebra and Its Applications*, **38**(1981), 5–15.

[13] A.W. Marshall and I. Olkin, *Inequalities: Theory of Majorization and Its Applications*, Academic Press, 1979.

[14] J.E. Mazo, "Some Extremal Markov Chains," *Bell System Technical Journal*, **61**(1982), 2065–2080.

[15] A. Sinclair and M. Jerrum, "Approximate counting, uniform generation and rapidly mixing Markov chains," Report CSR-241-87, Department of Computer Science, University of Edinburgh, 1987.