Chapter 6

Random Walks

6.1 Spectral gap

The first thing we need to do is to develop results that bound rate of convergence of one random walk on a random graph to equilibrium. Consider a Markov chain transition kernel $K(i,j)$ on $\{1,2,\ldots,n\}$ with reversible stationary distribution $\pi_i$, i.e., $\pi_i K(i,j) = \pi_j K(j,i)$. To measure convergence to equilibrium we will use the relative pointwise distance

$$\Delta(t) = \max_{i,j} \left| \frac{K^t(i,j)}{\pi_j} - 1 \right|$$

which is larger than the total variation distance

$$\Delta(t) \geq \max_i \sum_j \left| \frac{K^t(i,j)}{\pi_j} - 1 \right| \pi_j = \max_i \sum_j |K^t(i,j) - \pi_j|$$

Let $D$ be a diagonal matrix with entries $\pi_1, \pi_2, \ldots, \pi_n$ and $a = D^{1/2}K D^{-1/2}$. Since

$$a(i,j) = \pi_i^{1/2} K(i,j) \pi_j^{-1/2} = \pi_j^{1/2} K(j,i) \pi_i^{-1/2} = a(j,i)$$

matrix theory tells us that $a(i,j)$ has real eigenvalues $1 = \lambda_0 \geq \lambda_1 \geq \ldots \lambda_{n-1} \geq -1$. Let $\lambda_{\text{max}} = \max\{\lambda_1, |\lambda_{n-1}|\}$ be the eigenvalue with largest magnitude. The next result is from Sinclair and Jerrum (1989), but similar results can be found in many other places. The weaker result is useful for Cooper’s proof.

**Theorem 6.1.1.** Let $K$ be the transition matrix of an irreducible reversible Markov chain on $\{1,2,\ldots,n\}$ with stationary distribution $\pi$ and let $\pi_{\text{min}} = \min_j \pi_j$. Then

$$\Delta(t) \leq \frac{\lambda_{\text{max}}^t}{\pi_{\text{min}}} \ |K^t(i,j) - \pi_j| \leq \lambda_{\text{max}}^t \sqrt{\frac{\pi_j}{\pi_i}}$$
CHAPTER 6. RANDOM WALKS

Proof. Since $a$ is symmetric, we can select an orthonormal basis $e_m$, $0 \leq m < n$ of eigenvectors of $a$, and $a$ has spectral decomposition:

$$a = \sum_{m=0}^{n-1} \lambda_m e_m e_m^T$$

The matrix $B_m = e_m e_m^T$ has $B_m^2 = B_m$, and $B_mB_m = 0$ if $\ell \neq m$ so

$$a^t(i, j) = \sum_{m=0}^{n-1} \lambda_m^t e_m(i) e_m(j)$$

$e_0(i) = \pi_i^{1/2}$ so

$$K^t(i, j) = (D^{-1/2}a^tD^{1/2})_{i,j} = \pi_j + \sqrt{\frac{\pi_j}{\pi_i}} \sum_{m=1}^{n-1} \lambda_m^t e_m(i) e_m(j)$$  (6.1.1)  \[DC1\]

From this it follows that

$$\Delta(t) = \max_{i,j} \left| \sum_{m=1}^{n-1} \lambda_m^t e_m(i) e_m(j) \right| \leq \lambda_{\max} \max_{i,j} \sum_{m=1}^{n-1} |e_m(i)||e_m(j)|$$

The Cauchy-Schwarz inequality implies

$$\sum_{m=1}^{n-1} |e_m(i)||e_m(j)| \leq \left( \sum_{m=1}^{n-1} |e_m(i)|^2 \sum_{m=1}^{n-1} |e_m(j)|^2 \right)^{1/2} \leq 1$$  (6.1.2)  \[DC2\]

To see that $\sum_{m=1}^{n-1} |e_m(i)|^2 \leq 1$ note that if $\delta_i$ is the vector with 1 in the $i$th place and 0 otherwise then expanding in the orthonormal basis $\delta_i = \sum_{m=0}^{n-1} e_m(i) e_m$, so the desired result follows by taking the $L^2$ norm of both sides of the equation. To get the second result combine (6.1.1) and (6.1.2).

For some of our results we will consider continuous time chains. If jumps occur at rate one then there are a Poisson mean $t$ jumps by time $t$ so the transition probability is

$$H_t(x, y) = e^{-t} \sum_{m=0}^{\infty} \frac{t^m}{m!} K^m(x, y)$$  (6.1.3)  \[KtoH\]

If $\lambda_i$ is an eigenvalue of $K$ then $e^{-t(1-\lambda_i)}$ is an eigenvalue of $H_t$. Thus there are no negative eigenvalues to worry about and we have

Theorem 6.1.2. If $\beta = 1 - \lambda_1$ is the spectral gap of $K$ then

$$\frac{|H_t(i, j) - \pi(j)|}{\pi(j)} \leq e^{-\beta t} \quad |H_t(i, j) - \pi_j| \leq e^{-\beta t} \sqrt{\frac{\pi_j}{\pi_i}}$$

Proof. This follows immediately from (6.1.3) and Theorem 6.1.1.  \[cMCconv\]
6.2 Conductance

Given a reversible Markov transition kernel $K(x, y)$ we define the Dirichlet form by

$$\mathcal{E}(f, g) = \frac{1}{2} \sum_{x, y} (f(x) - f(y))(g(x) - g(y))\pi(x)K(x, y)$$  \hspace{1cm} (6.2.1) \textit{Dirform}

Introducing the inner product $< f, g >_\pi = \sum_x f(x)g(x)\pi(x)$, a little algebra shows

$$\mathcal{E}(f, f) = < f, (I - K)f >_\pi$$

If we define the variance by $\text{var}_{\pi}(f) = E_\pi(f - E_\pi f)^2$ then the spectral gap can be computed from the variational formula

$$1 - \lambda_1 = \min \{ \mathcal{E}(f, f) : \text{var}_{\pi}(f) = 1 \}$$  \hspace{1cm} (6.2.2) \textit{varforev}

To see this note that $\mathcal{E}(f, f)$ is not affected by subtracting a constant from $f$ so

$$1 - \lambda_1 = \min \{ < f, f >_\pi - < f, Kf >_\pi : E_\pi f = 0, < f, f >_\pi = 1 \}$$

and the result follows from the usual variational formula for $\lambda_1$ for the nonnegative symmetric matrix $a_{i,j} = \pi(i)K(i, j)$, i.e.,

$$\lambda_1 = \max \left\{ \sum_{i,j} x_i a_{i,j} x_j : \sum_i x_i^2 = 1 \right\}$$

Let $Q(x, y) = \pi(x)K(x, y)$, and define

$$h = \min_{\pi(S) \leq 1/2} \frac{Q(S, S^c)}{\pi(S)}$$

where $Q(S, S^c) = \sum_{x \in S, y \in S^c} Q(x, y)$. Since this is the size of the boundary of $S$ when edge $(x, y)$ is assigned weight $Q(x, y)$, we will sometimes write this as $|\partial S|$. Our next result is Lemma 3.3.7 in Saloff-Coste (1996). His $I = 2h$ so the constants are different. Saloff-Coste attributes the result to Diaconis and Stroock (1991), who in turn named the result Cheeger’s inequality in honor of the eigenvalue bound in differential geometry.

Theorem 6.2.1. The spectral gap has

$$\frac{h^2}{2} \leq 1 - \lambda_1 \leq 2h$$

Proof. Taking $f = 1_S$ in the variational formula (6.2.2) we have

$$\mathcal{E}(1_S, 1_S) = Q(S, S^c)$$
and \( \text{var}_\pi(1_S) = \pi(S)(1 - \pi(S)) \), so \( 1 - \lambda_1 \leq Q(S, S^c)/\pi(S)(1 - \pi(S)) \). The right-hand side is the same for \( S \) and \( S^c \), so we can restrict our attention to \( \pi(S) \leq 1/2 \). Since \( 1 - \pi(S) \geq 1/2 \), we have \( 1 - \lambda_1 \leq 2h \).

For the other direction, let \( F_t = \{ x : f(x) \geq t \} \) and let \( f_t \) be the indicator function of the set \( F_t \). Since only differences \( f(x) - f(y) \) appear in \( \mathcal{E}(f, f) \), defined in (6.2.1), we can without loss of generality suppose that the median of \( f \) is 0, i.e., \( \pi(F_t) \leq 1/2 \) for \( t > 0 \), and \( \pi(F_t^c) \leq 1/2 \) for \( t < 0 \). Our next step is to compute something that would be the Dirichlet form if we had squared the increment.

\[
\frac{1}{2} \sum_{x, y} |f(x) - f(y)| Q(x, y) = \sum_{f(x) > f(y)} (f(x) - f(y)) Q(x, y)
= \sum_{x, y} \int_{-\infty}^{\infty} 1_{\{f(y) < t < f(x)\}} Q(x, y) \, dt
= \int_{-\infty}^{\infty} |\partial F_t| \, dt + \int_{-\infty}^{0} |\partial F_t^c| \, dt
\geq h \left( \int_{0}^{\infty} \pi(F_t) \, dt + \int_{-\infty}^{0} \pi(F_t^c) \, dt \right) = h \pi(\|f\|)
\]

Continuing to suppose that the median of \( f \) is 0, let \( g = f^2 \text{sgn} \,(f) \), where \( \text{sgn} \,(x) = 1 \) if \( x > 0 \), \( \text{sgn}(x) = -1 \) if \( x < 0 \), and \( \text{sgn}(0) = 0 \). \( |g| = f^2 \) so the last inequality implies

\[
2h \pi(f^2) \leq \sum_{x, y} |g(x) - g(y)| Q(x, y) \leq \sum_{x, y} |f(x) - f(y)|(|f(x)| + |f(y)|) Q(x, y)
\]

To check the last inequality, we can suppose without loss of generality that \( f(x) > 0 \) and \( f(x) > f(y) \). If \( f(y) \geq 0 \) we have an inequality, while if \( f(y) < 0 \) we have \( f^2(x) + f^2(y) < (|f(x)| + |f(y)|)^2 \). Using the Cauchy-Schwarz inequality now the above is

\[
\leq \left( \sum_{x, y} (f(x) - f(y))^2 Q(x, y) \right)^{1/2} \cdot \left( \sum_{x, y} (|f(x)| + |f(y)|)^2 Q(x, y) \right)^{1/2}
\leq (2\mathcal{E}(f, f))^{1/2} (4\pi(f^2))^{1/2}
\]

Rearranging gives \( (2\mathcal{E}(f, f))^{1/2} \geq h(\pi(f^2))^{1/2} \). Squaring we have

\[
\mathcal{E}(f, f) \geq \frac{h^2}{2} \pi(f^2) \geq \frac{h^2}{2} E_\pi(f - E_\pi f)^2
\]

which proves the desired result. \( \square \)

Let \( G \) be a finite connected graph, \( d(x) \) be the degree of \( x \), and write \( x \sim y \) if \( x \) and \( y \) are neighbors. We can define a transition kernel by \( K(x, x) = 1/2 \), \( K(x, y) = 1/2d(x) \) if \( x \sim y \) and \( K(x, y) = 0 \) otherwise. The 1/2 probability of staying put means that we don’t have to
worry about periodicity or negative eigenvalues. Our $K$ can be written $(I + p)/2$ where $p$ is another transition probability, so all of the eigenvalues of $K$ are in $[0, 1]$, and $\lambda_{\text{max}} = \lambda_1$.

$\pi(x) = d(x)/D$ where $D = \sum_{y \in G} d(y)$, defines a reversible stationary distribution since $\pi(x)K(x, y) = 1/2D = \pi(y)K(y, x)$. Letting $e(S, S^c)$ is the number of edges between $S$ and $S^c$, and $\text{vol}(S)$ be the sum of the degrees in $S$, we have

$$h = \frac{1}{2} \min_{\pi(S) \leq 1/2} \frac{e(S, S^c)}{\text{vol}(S)}$$

When $d(x) \equiv d$, $h = \iota/2d$ where

$$\iota = \min_{|S| \leq n/2} \frac{e(S, S^c)}{|S|}$$

is the edge isoperimetric constant.

To illustrate the use of Theorem 6.2.1 and to show that one cannot get rid of the power 2 from the lower bound, consider random walk on the circle $\mathbb{Z}$ mod $n$ in which we stay put with probability $1/2$ and jump from $x$ to $x \pm 1$ with probability $1/4$ each. Taking $S = \{1, 2, \ldots n/2\}$ we see that

$$\iota = \frac{2}{n/2} = 4/n$$

To bound the spectral gap, we let $f(x) = \sin(\pi x/n)$. Since $\sin(a + b) = \sin a \cos b + \sin b \cos a$ we have

$$(I - K)f(x) = f(x)(1 - \cos(\pi/n))/2$$

and $1 - \lambda_1 \leq (1 - \cos(\pi/n))/2 \sim \pi^2/4n^2$ as $n \to \infty$. Using Theorem 6.1.1 gives an upper bound on the convergence time of order $O(n^2 \log n)$. However using the local central limit theorem for random walk on $\mathbb{Z}$ it is easy to see that $\Delta(t) \leq \epsilon$ at a time $K_\epsilon n^2$.

**Mixing times and the conductance profile.** Since we are lazy, we will suppose in what follows that the chain is as well: i.e., $K(x, x) \leq 1/2$ for all $x$, and we will not give the proofs of these more sophisticated results. Given an initial state $i$, it is possible to define stopping times $T$ so that $X_T$ has the stationary distribution. Define $H(i, \pi)$ the minimum value of $ET$ for all such stopping times and let $H = \max_i H(i, \pi)$. Define the mixing time

$$T_{\text{mix}} = \max_i \min\{t : d_{TV}(K^t(i, \cdot), \pi) < 1/e\}$$

The cutoff $1/e$ is somewhat arbitrary. The important thing is that it is small enough to allow us to conclude that if $t \geq T_{\text{mix}}$ then $d_{TV}(K^t(i, \cdot), \pi) < (2/e)^t$. Aldous (1988) has shown, see also Aldous, Lovász, and Winkler (1997), that $C_1H \leq T_{\text{mix}} \leq C_2H$. Define the conductance profile by

$$\Phi(x) = \min_{S, 0 < \pi(S) \leq x} \frac{Q(S, S^c)}{\pi(S)\pi(S^c)}$$

Lovász and Kannan (1999) have shown that

$$H \leq 32 \int_{\pi_{\text{min}}}^{1/2} \frac{dx}{x\Phi(x)^2}$$
Morris and Peres (2003) used their notion of evolving sets to sharpen this result to

$$\text{If } n \geq \int_{\pi(i) \land \pi(j)}^{4/\epsilon} \frac{4 \, dx}{x \Phi(x)^2} \text{ then } \left| \frac{K^n(i, j)}{\pi(i)} - 1 \right| \leq \epsilon$$

These results are useful for improving rate of convergence results in some examples. However in some of our favorite examples the worst conductance occurs for small sets, so we will instead use a recent result of Fountoulakis and Reed (2008).

**Theorem 6.2.2.** If $\Phi_c(x)$ be the minimum $Q(S, S^c)/\pi(S)\pi(S^c)$ over all connected sets $S$ with $x/2 \leq \pi(S) \leq x$ then

$$T_{mix} \leq 32 \int_{\pi_{min}}^{1/2} \frac{dx}{x \Phi_c(x)^2}$$
6.3 Fixed degree distribution, minimum degree 3

Gkantsidis, Mihail, and Saberi (2003) have proved the following:

**Theorem 6.3.1.** Consider a random graph with a fixed degree distribution in which the minimum degree is \( r \geq 3 \). There is a constant \( \alpha_0 > 0 \) so that with probability tending to 1 as \( n \to \infty \)

\[
\min_{\pi(S) \leq 1/2} \frac{e(S, S^c)}{\text{vol}(S)} \geq \alpha_0.
\]

From this and Theorem 6.1.1 it follows that the mixing time is \( \leq C \log n \). The diameters of these graphs are of \( O(\log n) \) so it cannot occur at a faster rate. The condition \( r \geq 3 \) is necessary since if there is a positive density of vertices of degree 2 then there will be paths of length \( O(\log n) \) in which each vertex has degree 2 and if we start in the middle of the path then the mixing time will be \( \geq O(\log^2 n) \). We will consider that case in the next section.

**Proof.** We say that a set of vertices \( S \) is bad if

\[
\frac{e(S, S^c)}{\text{vol}(S)} \leq \frac{\alpha}{D/r}.
\]

Our goal is to show that \( \bar{P}(\exists \text{ bad } S) \to 0 \). There are at most \( D/r \) sets that have volume \( k \).

Let \( f(m) \) be the number of ways of dividing \( m \) objects into pairs.

\[
f(m) = \frac{m!}{(m/2)!^2}.
\]

Let \( P(k, \ell) \) that there is a set \( S \) with \( \text{vol}(S) = k \) has \( e(S, S^c) = \ell \).

\[
P(k, \ell) \leq \binom{k}{\ell} \binom{D-k}{\ell} \ell! f(k-\ell) f(D-k-\ell) \frac{1}{f(D)}.
\]

To see this recall that in the random configuration model we pair the \( D \) half-edges at random, which can be done in \( f(D) \) ways. We pick \( \ell \) of \( k \) half-edges in \( S \) and \( \ell \) of those \( D-k \) in \( S^c \). The \( \ell \) which will make up \( e(S, S^c) \) can be paired in \( \ell! \) ways. Then the remaining \( k-\ell \) half-edges in \( S \) can be paired in \( f(k-\ell) \) ways and the \( D-k-\ell \) in \( S^c \) in \( f(D-k-\ell) \) ways.

To make it easier to compare with the argument in GMS we change values \( \ell = ak \). Taking into account the number of choices of \( S \), the probability of a bad set with volume \( k \) and \( e(S, S^c) = ak \) is

\[
\binom{D/r}{k/r} \binom{k}{ak} \binom{D-k}{ak} \frac{(ak)!f(k-ak)f(D-k-ak)}{f(D)}
\]

Their formula (10) is this with \( s! = (ak)! \) replaced by the larger \( f(2ak) \). They also have a factor \( ak \) to account for \( 1 \leq s \leq ak \).

To bound the binomial coefficients, the following lemma is useful

**Lemma 6.3.2.**

\[
\binom{n}{m} \leq \frac{n^m}{m!} \leq \frac{n^m}{m^m e^{-m}}
\]
Proof. The first inequality follows from \( n(n-1) \cdots (n-m+1) \leq n^m \). For the second we note that the series expansion of \( e^m \) has only positive terms so \( e^m > m^m/m! \).

From Lemma 6.3.2, we see that the three binomial coefficients in (6.3.3) are

\[
\leq \left( \frac{D e}{k} \right)^{k/r} \left( \frac{e}{\alpha} \right)^{2ak} \left( \frac{D-k}{k} \right)^{ak} \tag{6.3.4} \text{ part1}
\]

Here, to prepare for a later step, we have transferred part of the bound for the third term into the second.

To bound the \( f \)'s in (6.3.3) we use Stirling’s formula to conclude

\[
f(m) = \frac{m!}{(m/2)!2^{m/2}} \sim C \frac{m^{m+1/2}e^{-m}}{(m/2)^{m/2+1/2}e^{-m/2}2^{m/2}} = C(m/e)^{m/2}
\]

From this we see that the fraction in (6.3.3) is

\[
\leq C k^{1/2} (\alpha k/e)^{ak} (k(1-\alpha)/e)^{k(1-\alpha)/2} ((D-(1+\alpha)k)/e)^{(D-(1+\alpha)k)/2} / (D/e)^{D/2}
\]

\[
= C k^{1/2} (\alpha k)^{ak} D^{-ak} \left( \frac{k(1-\alpha)}{D} \right)^{k(1-\alpha)/2} \left( 1 - \frac{(1+\alpha)k}{D} \right)^{(D-(1+\alpha)k)/2} \tag{6.3.5} \text{ part2}
\]

since the exponents in the numerator sum to \( D/2 \).

Combining (6.3.4) and (6.3.5) gives an upper bound

\[
\leq C k^{1/2} \left( \frac{D e}{k} \right)^{k/r} \left( \frac{e^2}{\alpha} \right)^{ak} \left( \frac{D-k}{k} \right)^{ak} \cdot \left( \frac{k(1-\alpha)}{D} \right)^{k(1-\alpha)/2} \left( 1 - \frac{(1+\alpha)k}{D} \right)^{(D-(1+\alpha)k)/2}
\]

Ignoring the \( C k^{1/2} \)'s, the first term is the first term from (6.3.4), the second and third terms come from combining the second and third terms of (6.3.4) and with the first and second terms of (6.3.5), while the remainder of the formula comes from (6.3.5). Using \( \alpha > 0 \) and \( D-k < D \) and rearranging we have

\[
\leq C k^{1/2} e^{k/r} \left( \frac{e^2}{\alpha} \right)^{ak} \cdot \left( \frac{k}{D} \right)^{k(1-\alpha)/2-k/r} \left( 1 - \frac{(1+\alpha)k}{D} \right)^{(D-(1+\alpha)k)/2}
\]

Setting \( \beta = e^2/\alpha \) and \( \gamma = (1-\alpha)/2 - 1/r \) we have

\[
\leq C k^{1/2} e^{k/r} \beta^k \cdot \left( \frac{k}{D} \right)^{\gamma k} \left( 1 - \frac{(1+\alpha)k}{D} \right)^{(D-(1+\alpha)k)/2} \tag{6.3.6} \text{ GMS17}
\]
Comparing with formula (17) in Gkantsidis, Mihail, and Saberi (2003), we see that apart from the differences that result from our use of \((\alpha k)!\) instead of \(f(2\alpha k)\), they are missing the \(e^{\frac{k}{r}}\) and we have retained an extra term to compensate for the error.

Let

\[
G(k) = e^{\frac{k}{r}} \beta^\alpha \cdot \left(\frac{k}{D}\right)^\gamma \left(1 - \frac{(1 + \alpha)k}{D}\right)^{(D-(1+\alpha)k)/2}
\]

\(Ck^{1/2} \leq Cn^{1/2}\) so we can show \(h \geq \alpha_0\) by showing that for \(0 \leq \alpha \leq \alpha_0\)

\[
\sup_{1 \leq k \leq D/2} G(k) = o(n^{-5/2})
\]

because then we can sum our estimate over \(k \leq D/2\) and \(s = \alpha k\) with \(\alpha \leq \alpha_0\) and end up with a result that is \(o(1)\).

\[
\beta^\alpha = \exp(\eta k) \text{ where } \eta = \alpha \log(e^2/\alpha) \to 0 \text{ as } \alpha \to 0.
\]

Ignoring this term, and setting \(k = D/2, \alpha = 0\)

\[
G(D/2) = e^{D/2r} (1/2)^{1/2-1/r} D/2+D/4 = \left(e^{1/3}(1/2)^{2/3}\right)^{D/2}
\]

when \(r = 3\), the worst case. Since \(4 > e\), the quantity in parentheses is \(< 1\) when \(\alpha = 0\) and hence also when \(0 \leq \alpha \leq \alpha_0\), if \(\alpha_0\) is small.

To extend this result to other values of \(k\), let

\[
H(k) = \log G(k) = \frac{k}{r} + k\alpha \log \beta + k\gamma \log(k/D) + \frac{D - (1 + \alpha)k}{2} \log \left(1 - \frac{(1 + \alpha)k}{D}\right)
\]

Since \(G(k) = \exp(H(k))\), differentiating gives \(G'(k) = G(k)H'(k)\) where

\[
H'(k) = \frac{1}{r} + \alpha \log \beta + \gamma \log(k/D) + \gamma - \frac{(1 + \alpha)}{2} \log \left(\frac{D - (1 + \alpha)k}{D}\right)
\]

\[
+ \frac{D - (1 + \alpha)k}{2} \cdot \frac{D}{D - (1 + \alpha)k} \cdot \left(\frac{-(1 + \alpha)k}{2}\right)
\]

Differentiating again \(G''(k) = G(k)(H'(k)^2 + H''(k))\) where

\[
H''(k) = \frac{\gamma}{k} - \frac{(1 + \alpha)}{2} \cdot \frac{D}{D - (1 + \alpha)k} \cdot \left(\frac{-(1 + \alpha)k}{D}\right) > 0
\]

From the last calculation we see that \(G(k)\) is convex. We have control of the value for \(k = D/2\). It remains then to inspect the values for small \(k\). Dropping the last factor which is \(< 1\)

\[
G(k) \leq e^{k/r} \beta^\alpha \left(\frac{1}{\alpha D}\right)^\gamma
\]

When \(0 \leq \alpha \leq \alpha_0 \leq 1/24, \gamma \geq 7/48\) and hence \(G(24) \leq Cn^{-7/2}\). Since \(e(S, S') \geq 1\) there is nothing to prove for \(k \leq 1/\alpha_0 = 24\) and the proof is complete.
The bound on the mixing time that emerges from this argument is not very precise. In contrast we have the following result due to Lubetsky and Sly (2010)

**Theorem 6.3.3.** The mixing time for the lazy random walk on the random 3-regular graph is asymptotically $6 \log_2 n$.

To explain the intuition, note that the random 3-regular graph looks locally like the 3-regular tree in which one edge points back toward the root and the other two point away. The speed of the lazy random walk on this tree is $1/6$. The tree grows at rate $2^m$, so the diameter of the graph is $\log_2 n$. At times $t \leq (1 - \epsilon)6 \log_2 n$ we cannot be in equilibrium because the support of $p^t(x, y)$ is concentrated on $O(n^{1-\epsilon})$ vertices.

Lubetsky and Sly (2010) prove the converse and more. Let

$$d_n(t) = \max_{x \in G_n} \| P_x(X_t \in \cdot) - \pi(\cdot) \|_{TV}$$

If $t_n = 6 \log_2 n$ and $w_n = \sqrt{\log n}$ then

$$\lim_{\lambda \to \infty} \liminf_{n \to \infty} d_n(t_n - \lambda w_n) = 1$$
$$\lim_{\lambda \to \infty} \limsup_{n \to \infty} d_n(t_n + \lambda w_n) = 0$$

The first result follows from the central limit theorem for the random walk on the tree.

Q. For more general graphs generated by the configuration model we know that the diameter is $\log \bar{q} n$, where $\bar{q} = \sum_k kq_k$ is the mean of the size biased distribution, that the speed of the random walk on the Galton-Watson tree is

$$\nu = \sum_k \frac{k - 1}{k + 1} q_k$$

and the walk has a central limit theorem. Is this lower bound for the convergence time sharp? The fact that asymptotically the walk concentrates on a small part of the tree cases doubt on this conjecture.
6.4 Only degrees 2 and 3

We start with a random 3-regular graph \( H \) with \((1-a)n\) vertices, and produce a new graph \( G \) by replacing each edge by a path with a geometric number of edges with success probability \( r \), i.e., with probability \((1-r)^{-1}r\) we have \( j \) edges. The number of vertices of degree 2 in one of these paths has mean \((1/r)-1\) so if we pick \( r \) so that \((1-a)((1/r)-1) = a\), then we asymptotically have \( n \) vertices and \( p_2 = a \).

Our main result is that

**Theorem 6.4.1.** The mixing time of the lazy random walk on \( G \) is \( O(\log^2 n) \).

The lower bound follows from the fact that the longest path is \( O(\log n) \). To prove an upper bound, we will use Theorem 6.2.2. To begin we need a simple combinatorial result. This generalizes easily to graphs with bounded degree but it is not clear what to do for a more general graph.

**Lemma 6.4.2.** The number of connected subsets of \( H \) of size \( k \) containing a fixed vertex \( v_0 \) is \( \leq 3^{3k} \).

**Proof.** Given a connected set \( V \) of vertices of \( H \), define the set \( W = \{(x, y) : x, y \in V, x \neq y\} \). Note that if \((x, y) \in W\) then \((y, x) \in W\) and think of these as two oriented edges between \( x \) and \( y \). We will show that there is a Hamiltonian paths starting from \( v_0 \) that traverses each oriented edge at most once. The number of edges in \( W \) is at most \( 3k \). At each stage we have at most 3 choices so the number of such paths is \( \leq 3^{3k} \) this proves the desired result.

To construct the path start at \( v_0 \) and pick an outgoing edge. When we are at a vertex \( v \neq v_0 \) we have used one more incoming edge than outgoing edge so we have at least one way out. This procedure may terminate by coming back to \( v_0 \) at a time when there are no more outgoing edges. If so, and we have not exhausted the graph, then there is some vertex \( v_1 \) on the current path with an outgoing edge. Repeat the construction starting from \( v_1 \) using edges not in the current path. We will eventually come back to \( v_1 \). We can combine the two paths by using the old path from \( v_0 \) to the first visit to \( v_1 \), using the new path to go from \( v_1 \) to \( v_1 \), and then the old path to return from \( v_1 \) to \( v_0 \). Repeating this construction we will eventually exhaust all of the edges.

Let \( B \) be a connected subset of \( G \), and let \( A = B \cap H \). It is easy to see that \( A \) is a connected subset of \( H \). By the isoperimetric inequality for random regular graphs, there is an \( \alpha > 0 \) so that \(|\partial A| \geq \alpha|A|\), where \( \partial A \) is the set of edges \((x, y)\) with \( x \in A \) and \( y \notin A \). From the construction of the graph it is easy to see that \(|\partial B| = |\partial A|\).

It remains to see how big \(|B|/|A|\) can be. When \(|A| = 1\) we can have \(|B| = O(\log n)\). The key to the proof is to show that the ratio cannot be big when \(|A|\) is. Let \( X_i \) be i.i.d with \( P(X_i = j) = (1-r)^{-1}r \) and let \( S_m = X_1 + \cdots X_m \).

**Lemma 6.4.3.** There are constants \( \beta \) and \( \gamma \) so that

\[ P(S_m \geq \beta \log n + \gamma m) \leq n^{-2} (2/81)^m \]
Proof. The moment generating function
\[ \psi(\theta) = E e^{\theta X_i} = \sum_{j=0}^{\infty} (e^\theta (1 - r))^j r = \frac{r}{1 - e^\theta (1 - r)} \]
when \( e^\theta (1 - r) < 1 \). If we pick \( \theta > 0 \) so that \( e^\theta (1 - r) = 1 - r/2 \) then \( \psi(\theta) = 2 \). Markov’s inequality implies
\[ P(S_m \geq \beta \log n + \gamma m) \leq \psi(\theta)^m \exp(-\theta[\beta \log n + \gamma m]) \]
Letting \( \beta = 2/\theta \) and \( \gamma = 81/\theta \) the desired result follows.

If \( |A| = k \) then the number of edges adjacent to some point in \( A \) is \( \geq k + 2 \), the value for a tree and \( \leq 3k \). Since the number of connected sets of size \( k \) is \( \leq n2^k \) it follows that with probability \( 1 - O(n^{-1}) \) we have \( |B| \leq \beta \log n + 3\gamma |A| \) for connected sets \( B \). From this it follows that
\[ |\partial B| = |\partial A| \geq \alpha |A| \geq \frac{\alpha}{\gamma} (|B| - \beta \log n) \]
if \( |B| \geq 2\beta \log n \) then \( |\partial B|/|B| \geq c \), while for \( |B| \leq 2\beta \log n \), \( |\partial B|/|B| \geq 2/|B| \).

To evaluate \( \int_{1/3n}^{1/2} dx/(x^2 \Phi(x)^2) \) up to a constant factor we note that
\[ \int_{2\beta \log n/n}^{1/2} \frac{dx}{x} = O(\log n) \]
while changing variables \( y = nx \), \( dy = n \, dx \) shows
\[ \int_{1/3n}^{2\beta \log n/n} \frac{dx}{x(2/xn)^2} = \int_{1/3}^{2\beta \log n} \frac{dy}{(y/2)} = O(\log^2 n) \]
and completes the proof.

Exercise. Extend this proof to minimum degree 1, by attaching to \( G \) side trees with the correct subcritical branching process.

Remark. Fountoulakis and Reed (2008) have proved a result for Erdős-Renyi random graphs. They show that when the average degree \( d = np \) satisfies \( O(\sqrt{\log n}) \) then the mixing time is \( \Theta((\log n/d)^2) \), due to the presence of long paths. As the average degree grows then the diameter takes over and the mixing time becomes the diameter \( \Theta(\log n/\log d) \).
References


