Chapter 7
Symmetric Graphs and Chains

David Aldous
Department of Statistics
University of California
Berkeley, CA 94720

James Allen Fill
Department of Mathematical Sciences
The Johns Hopkins University
Baltimore, MD 21218-2692

January 31, 1994

In this Chapter we show how general results in Chapters 3, 4 and 6 can sometimes be strengthened when symmetry is present. Many of the ideas are just simple observations. Since the topic has a "discrete math" flavor our default convention is to work in discrete time, though as always the continuous-time case is similar. Note that we use the word "symmetry" in the sense of *spatial* symmetry (which is the customary use in mathematics as a whole) and not as a synonym for time-reversibility. Note also our use of "random flight" for what is usually called "random walk" on a group.

Biggs [6] contains an introductory account of symmetry properties for graphs, but we use little more than the definitions. I have deliberately not been overly fussy about giving weakest possible hypotheses. For instance many results for symmetric reversible chains depend only of the symmetry of mean hitting times (7), but I haven't spelt this out. Otherwise one can end up with more definitions than serious results! Instead, we focus on three different strengths of symmetry condition. Starting with the weakest, section 1 deals with symmetric reversible chains, a minor generalization of what is usually called "symmetric random walk on a finite group". In the graph setting, this specializes to random walk on a Cayley or vertex-transitive graph. Section 2 deals with random walk on an arc-transitive graph, encompassing what is usually called "random walk on a finite group with steps uniform on a conjugacy class". Section 3 deals with random walk on a distance-regular graph, which roughly corresponds to nearest-neighbor isotropic random walk on a discrete Gelfand pair.

This book focuses on inequalities rather than exact calculations, and the limitation of this approach is most apparent in this chapter. Group representation theory, though of course developed for non-probabilistic reasons, turns out to be very well adapted to the study of many questions concerning random walks on groups. I lack the space (and, more importantly, the knowledge) to give a worthwhile treatment here, and in any case an account which is both introductory and gets to interesting results is available in Diaconis [12]. In many concrete examples, eigenvalues are known by group representation theory, and so in particular our parameters τ_2 and τ_0 are known. See e.g. section 2.1. In studying a particular example, after investigating eigenvalues one can seek to study further properties of the chain by either

(i) continuing with calculations specific to the example; or

(ii) using general inequalities relating other aspects of the chain to τ_2 and τ_0.

The purpose of this Chapter is to develop option (ii). Of course, the more highly-structured the example, the more likely one can get stronger explicit results via (i). For this reason we devote more space to the weaker setting of section 1 than to the stronger settings of sections 2 and 3.

xxx scattering of more sophisticated math in Chapter 10.

1 Symmetric reversible chains

1.1 Definitions

Consider an irreducible transition matrix $\mathbf{P} = (p_{ij})$ on a finite state space I. A *symmetry* of $\mathbf{P}$ is a $1-1$ map $\gamma : I \to I$ such that

$$p_{\gamma(i)\gamma(j)} = p_{ij} \text{ for all } i, j.$$

The set Γ of symmetries forms a group under convolution, and in our (non-standard) terminology a *symmetric* Markov transition matrix is one for which Γ acts transitively, i.e.

$$\text{for all } i, j \in I \text{ there exists } \gamma \in \Gamma \text{ such that } \gamma(i) = j.$$

Such a chain need not be reversible; a *symmetric reversible chain* is just a chain which is both symmetric and reversible. A natural setting is where I is itself a group under an operation $(i,j) \to ij$ which we write multiplicitively. If μ is a probability distribution on I and $(Z_t; t \geq 1)$ are i.i.d. I-valued with distribution μ then

$$X_t = x_0 Z_1 Z_2 \dots Z_t \tag{1}$$

is the symmetric Markov chain with transition probabilities

$$p_{ij} = \mu(i^{-1} * j)$$

started at x_0. This chain is reversible iff

$$\mu(i) = \mu(i^{-1}) \text{ for all } i. \tag{2}$$

We have rather painted ourselves into a corner over terminology. The usual terminology for the process (1) is "random walk on the group I" and if (2) holds then it is a

$$\text{"symmetric random walk on the group } I\text{" .} \tag{3}$$

Unfortunately in this phrase, both "symmetric" and "walk" conflict with our conventions, so we can't use the phrase. Instead we will use "random flight on the group I" for a process (1), and "reversible random flight on the group I" when (2) also holds. Note that we always assume chains are irreducible, which in the case of a random flight holds iff the support of μ generates the whole group I. Just keep in mind that the topic of this section, symmetric reversible chains, forms a minor generalization of the processes usually described by (3).

On an graph $(\mathcal{V}, \mathcal{E})$, a *graph automorphism* is a $1-1$ map $\gamma : \mathcal{V} \to \mathcal{V}$ such that

$$(\gamma(w), \gamma(v)) \in \mathcal{E} \text{ iff } (w, v) \in \mathcal{E}.$$

The graph is called *vertex-transitive* if the automorphism group acts transitively on vertices. Clearly, random walk on a (unweighted) graph is a symmetric reversible chain iff the graph is vertex-transitive. We specialize to this case in section 1.8. A further specialization is to random walk on a *Cayley graph.* If $\mathcal{G} = (g_i)$ is a set of generators of a group I, which we always assume to satisfy

$$g \in \mathcal{G} \text{ implies } g^{-1} \in \mathcal{G}$$

then the associated Cayley graph has vertex-set I and edge-set

$$\{(v, vg) : v \in I, g \in \mathcal{G}\}.$$

A Cayley graph is vertex-transitive.

Finally, recall from Chapter 3 yyy that we can identify a reversible chain with a random walk on a weighted graph. With this identification, a symmetric reversible chain is one where the weighted graph is vertex-transitive, in the natural sense.

1.2 This section goes into Chapter 3

Lemma 1 *For an irreducible reversible chain, the following are equivalent.*
(a) $P_i(X_t = i) = P_j(X_t = j), i, j \in I, t \geq 1$
(b) $P_i(T_j = t) = P_j(T_i = t), i, j \in I, t \geq 1.$

Proof. In either case the stationary distribution is uniform – under (a), by letting $t \to \infty$, and under (b) by taking $t = 1$, implying $p_{ij} \equiv p_{ji}$. So by reversibility $P_i(X_t = j) = P_j(X_t = i)$ for $i \neq j$ and $t \geq 1$. But recall from Chapter 2 Lemma yyy that the generating functions $G_{ij}(z) = \sum_t P_i(X_t = j)z^j$ and $F_{ij}(z) = \sum_t P_i(T_t = j)z^j$ satisfy

$$F_{ij} = G_{ij}/G_{jj}. \tag{4}$$

For $i \neq j$ we have seen that $G_{ij} = G_{ji}$, and hence by (4)

$$F_{ij} = F_{ji} \text{ iff } G_{jj} = G_{ii},$$

which is the assertion of Lemma 1.

1.3 Elementary properties

Our standing assumption is that we have an irreducible symmetric reversible n-state chain. The symmetry property implies that the stationary distribution π is uniform, and also implies

$$P_i(X_t = i) = P_j(X_t = j), i, j \in I, t \geq 1. \tag{5}$$

But by Chapter 3 Lemma yyy, under reversibility (5) is equivalent to

$$P_i(T_j = t) = P_j(T_i = t), i, j \in I, t \geq 1. \tag{6}$$

And clearly (6) implies

$$E_iT_j = E_jT_i \text{ for all } i, j. \tag{7}$$

We make frequent use of these properties. Incidently, (7) is in general strictly weaker than (6): van Slijpe [36] p. 288 gives an example with a 3-state reversible chain.

We also have, from the definition of symmetric, that $E_\pi T_i$ is constant in i, and hence

$$E_\pi T_i = \tau_0 \text{ for all } i. \tag{8}$$

So by Chapter 4 yyy

$$\tau^* \leq 4\tau_0. \tag{9}$$

The formula for $E_\pi T_i$ in terms of the fundamental matrix (Chapter 2 yyy) can be written as

$$\tau_0/n = 1 + \sum_{t=1}^{\infty}(P_i(X_t = i) - 1/n). \tag{10}$$

Approximating τ_0 by the first few terms is what we call the *local transience heuristic*. See Chapter xxx for rigorous discussion.

Lemma 2 *(i)* $E_i T_j \geq \frac{n}{1+p(i,j)}$, $j \neq i$.
(ii) $\max_{i,j} E_i T_j \leq 2\tau_0$

Proof. (i) This is a specialization of Chapter 6 xxx.
(ii) For any i, j, k,

$$E_i T_j \leq E_i T_k + E_k T_j = E_i T_k + E_j T_k.$$

Averaging over k, the right side becomes $2\tau_0$.

Recall that a simple Cauchy-Schwartz argument (Chapter 3 yyy) shows that, for any reversible chain whose stationary distribution is uniform,

$$P_i(X_{2t} = j) \leq \sqrt{P_i(X_{2t} = i) P_j(X_{2t} = j)}.$$

So by (5), for a symmetric reversible chain, the most likely place to be after $2t$ steps is where you started:

Corollary 3 $P_i(X_{2t} = j) \leq P_i(X_{2t} = i)$, *for all* $i, j, \in I, t \geq 1$.

This type of result is nicer in continuous time, where the inequality holds for all times.

1.4 Hitting times

Here is our first non-trivial result, from Aldous [3].

Theorem 4 *Suppose a sequence of symmetric reversible chains satisfies* $\tau_2/\tau_0 \to 0$. *Then*

(a) For the stationary chain, and for arbitrary j, *we have* $T_j/\tau_0 \stackrel{d}{\to} \xi$ *and* var $(T_j/\tau_0) \to 1$, *where* ξ *has exponential(1) distribution.*

(b) $\max_{i,j} E_i T_j/\tau_0 \to 1$.

(c) If (i_n, j_n) *are such that* $E_{i_n} T_{j_n}/\tau_0 \to 1$ *then* $P_{i_n}(T_{j_n}/\tau_0 \in \cdot) \stackrel{d}{\to} \xi$.

Note that, because $\tau_2 \leq \tau_1 + 1$ and $\tau_0 \geq (n-1)^2/n$, the hypothesis "$\tau_2/\tau_0 \to 0$" is weaker than either "$\tau_2/n \to 0$" or "$\tau_1/\tau_0 \to 0$".

Part (a) is a specialization of Chapter 3 Proposition yyy and its proof. Parts (b) and (c) use refinements of the same technique. Part (b) implies

$$\text{if } \tau_2/\tau_0 \to 0 \text{ then } \tau^* \sim 2\tau_0.$$

Because this applies in many settings in this Chapter, we shall rarely need to discuss τ^* further.

xxx give proof

In connection with (b), note that

$$E_v T_w \leq \tau_1^{(2)} + \tau_0 \tag{11}$$

by definition of $\tau_1^{(2)}$ and vertex-transitivity. So (b) is obvious under the slightly stronger hypothesis $\tau_1/\tau_0 \to 0$.

Chapter 3 Proposition yyy actually gives information on hitting times T_A to more general subsets A of vertices. Because (Chapter 3 yyy) $E_\pi T_A \geq \frac{(1-\pi(A))^2}{\pi(A)}$, we get (in continuous time) a quantification of the fact that T_A has approximately exponential distribution when $|A| \ll n/\tau_2$ and when the chain starts with the uniform distribution:

$$\sup_t |P_\pi(T_A > t) - \exp(-t/E_\pi T_A)| \leq \frac{\tau_2 n}{|A|} \left(1 - \frac{|A|}{n}\right)^{-2}.$$

1.5 Cover times

Recall the cover time C from Chapter 6. By symmetry, in our present setting $E_i C$ doesn't depend on the starting place i, so we can write EC. In this section we combine results on hitting times with various forms of Matthews method to obtain asymptotics for cover times in the setting of a sequence of symmetric reversible chains. Experience, and the informal argument above (15), suggest the principle

$$EC \sim \tau_0 \log n, \text{ except for chains resembling random walk on the } n\text{-cycle}. \tag{12}$$

The results in this chapter concerning cover times go some way towards formalizing this principle.

Corollary 5 *For a sequence of symmetric reversible chains*

(a) $\frac{1-o(1)}{1+p^*} n \log n \leq EC \leq (2 + o(1))\tau_0 \log n$, *where* $p^* \equiv \max_{j \neq i} p_{i,j}$.

(b) If $\tau_2/\tau_0 \to 0$ *then* $EC \leq (1+o(1))\tau_0 \log n$.
(c) If $\tau_2/\tau_0 = O(n^{-\beta})$ *for fixed* $0 < \beta < 1$ *then*

$$EC \geq (\beta - o(1))\tau_0 \log n.$$

Proof. Using the basic form of Matthews method (Chapter 2 yyy), (a) follows from Lemma 2 and (b) from Theorem 4. To prove (c), fix a state j and $\varepsilon > 0$. Using (11) and Markov's inequality,

$$\pi\{i : E_i T_j \leq (1-\varepsilon)\tau_0\} \leq \frac{\tau_1^{(2)}}{\varepsilon \tau_0} \equiv \alpha, \text{ say.}$$

So we can inductively choose $\lceil \alpha^{-1} \rceil$ vertices i_k such that

$$E_{i_k} T_{i_l} > (1-\varepsilon)\tau_0;\ 1 \leq k < l \leq \lceil \alpha^{-1} \rceil.$$

By the extended form of Matthews method (Chapter 6 Corollary yyy)

$$EC \geq (1-\varepsilon)\tau_0 h_{\lceil \alpha^{-1} \rceil - 1}.$$

From Chapter 4 yyy, $\tau_1 \leq \tau_2(1+\log n)$ and so the hypothesis implies $\tau_1/\tau_0 = O(n^{\varepsilon-\beta})$. So the asymptotic lower bound for EC becomes $(1-\varepsilon)\tau_0(\beta - \varepsilon)\log n$, and since ε is arbitrary the result follows.

Since the only natural examples with $\tau_1/\tau_0 \not\to 0$ are variations of random walk on the n-cycle, for which $EC = \Theta(\tau_0)$ without the "$\log n$" term, we expect a positive answer to

Open Problem 6 *In the setting of Corollary 5, is* $EC \leq (1+o(1))\tau_0 \log n$ *without further hypotheses?*

Here is an artificial example to illustrate the bound in (c).

Example 7 *Two time scales.*

Take $m_1 = m_1(n), m_2 = m_2(n)$ such that $m_1 \sim n^{1-\beta}, m_1 m_2 \sim n$. The underlying idea is to take two continuous-time random walks on the complete graphs K_{m_1} and K_{m_2}, but with the walks run on different time scales. To set this up directly in discrete time, take state space $\{(x,y) : 1 \leq x \leq m_1, 1 \leq y \leq m_2\}$ and transition probabilities

$$\begin{aligned}(x,y) &\to (x',y) \quad \text{chance } (m_1-1)^{-1}\left(1 - \frac{1}{a m_1 \log m_1}\right),\ x' \neq x \\ &\to (x,y') \quad \text{chance } (m_2-1)^{-1}\frac{1}{a m_1 \log m_1},\ y' \neq y\end{aligned}$$

where $a = a(n) \uparrow \infty$ slowly. It is not hard to formalize the following analysis. Writing the chain as (X_t, Y_t), the Y-component stays constant for time $\Theta(am_1 \log m_1)$, during which time every x-value is hit, because the cover time for K_{m_1} is $\sim m_1 \log m_1$. And $m_2 \log m_2$ jumps of the Y-component are required to hit every y-value, so

$$EC \sim (m_2 \log m_2) \times (am_1 \log m_1) \sim an(\log m_1)(\log m_2). \qquad (13)$$

Now $\tau_2 \sim am_1 \log m_1$, and because the mean number of returns to the starting point before the first Y-jump is $\sim a \log m_1$ we can use the local transience heuristic (10) to see $\tau_0 \sim (a \log m_1) \times n$. So $\tau_2/\tau_0 \sim m_1/n \sim n^{-\beta}$, and the lower bound from (c) is

$$(\beta - o(1))(a \log m_1) n \log n.$$

But this agrees with the exact limit (13), because $m_2 \sim n^\beta$.

We now turn to sharper distributional limits for C. An (easy) background fact is that, for independent random variables (Z_i) with exponential, mean τ, distribution,

$$\frac{\max(Z_1, \ldots, Z_n) - \tau \log n}{\tau} \stackrel{d}{\to} \eta$$

where η has the extreme value distribution

$$P(\eta \leq x) = \exp(-e^{-x}), \ -\infty < x < \infty. \qquad (14)$$

Now the cover time $C = \max_i T_i$ is the *max* of the hitting times, and with the uniform initial distribution the T_i's have mean τ_0. So if the T_i's have approximately exponential distribution and are roughly independent of each other then we anticipate the limit result

$$\frac{C - \tau_0 \log n}{\tau_0} \stackrel{d}{\to} \eta. \qquad (15)$$

Theorem 4 has already given us a condition for limit exponential distributions, and we shall build on this result to give (Theorem 9) conditions for (15) to hold.

The extreme value distribution (14) has transform

$$E \exp(\theta\eta) = \Gamma(1 - \theta), \ -\infty < \theta < 1. \qquad (16)$$

Classical probability theory (see Notes) says that to prove (15) it is enough to show that transforms converge, i.e. to show

$$E \exp(\theta C/\tau_0) \sim n^{-\theta}\Gamma(1-\theta),\ -\infty < \theta < 1. \tag{17}$$

But Matthews method, which previously we have used on expectations, can just as well be applied to transforms. By essentially the same argument as in Chapter 2 Theorem yyy, Matthews [29] obtained

Proposition 8 *The cover time C in a not-necessarily-reversible Markov chain with arbitrary initial distribution satisfies*

$$\frac{\Gamma(n+1)\Gamma(1/f_*(\beta))}{\Gamma(n+1/f_*(\beta))} \leq E\exp(\beta C) \leq \frac{\Gamma(n+1)\Gamma(1/f^*(\beta))}{\Gamma(n+1/f^*(\beta))}$$

where

$$f^*(\beta) \equiv \max_{j\neq i} E_i \exp(\beta T_j)$$
$$f_*(\beta) \equiv \min_{j\neq i} E_i \exp(\beta T_j).$$

Substituting into (17), and using the fact

$$\frac{\Gamma(n+1)}{\Gamma(n+1-s_n)} \sim n^s \text{ as } n\to\infty,\ s_n \to s$$

we see that to establish (15) it suffices to prove that for arbitrary $j_n \neq i_n$ and for each fixed $-\infty < \theta < 1$,

$$E_{i_n} \exp(\theta T_{j_n}/\tau_0) \to \frac{1}{1-\theta}. \tag{18}$$

Theorem 9 *For a sequence of symmetric reversible chains, if*
(a) $\min_{j\neq i} E_i T_j = \tau_0(1-o(1))$
(b) $\tau_2/\tau_0 = o\left(\frac{1}{\log n}\right)$
then

$$\frac{C - \tau_0 \log n}{\tau_0} \ \stackrel{d}{\to}\ \eta.$$

Proof. By hypothesis (a) and Theorem 4 (b,c), for arbitrary $j_n \neq i_n$ we have $P_{i_n}(T_{j_n}/\tau_0 \in \cdot) \ \stackrel{d}{\to}\ \xi$. This implies (18) for $\theta \leq 0$, and also by Fatou's lemma implies $\liminf_n E_{i_n} \exp(\theta T_{j_n}/\tau_0) \geq \frac{1}{1-\theta}$for $0 < \theta < 1$. Thus it is sufficient to prove

$$\max_{j \neq i} E_i \exp(\theta T_j/\tau_0) \leq \frac{1+o(1)}{1-\theta}, \ 0 < \theta < 1. \tag{19}$$

The proof exploits some of our earlier general inequalities. Switch to continuous time. Fix $\beta > 0$. By conditioning on the position at some fixed time s,

$$E_i \exp(\beta(T_j - s)^+) \leq \max_x (nP_i(X_s = x)) \times E_\pi \exp(\beta T_j).$$

By Corollary 3 the *max* is attained by $x = i$, and so

$$E_i \exp(\beta T_j) \leq (nP_i(X_s = i)e^{\beta s}) \times E_\pi \exp(\beta T_j).$$

We now apply some general inequalities. Chapter 4 yyy says $nP_i(X_s = i) \leq 1 + n\exp(-s/\tau_2)$. Writing α_j for the quasistationary distribution on $\{j\}^c$, Chapter 3 (yyy) implies $P_\pi(T_j > t) \leq \exp(-t/E_{\alpha_j}T_j)$ and hence

$$E_\pi \exp(\beta T_j) \leq \frac{1}{1 - \beta E_{\alpha_j} T_j}.$$

But Chapter 3 Theorem yyy implies $E_{\alpha_j} T_j \leq \tau_0 + \tau_2$. So setting $\beta = \theta/\tau_0$, these inequalities combine to give

$$E_i \exp(\theta T_j/\tau_0) \leq (1 + n\exp(-s/\tau_2)) \times \exp(\theta s/\tau_0) \times \frac{1}{1 - \theta(1 + \tau_2/\tau_0)}.$$

But by hypothesis (b) we can choose $s = o(\tau_0) = \Omega(\tau_2 \log n)$ so that each of the first two terms in the bound tends to 1, establishing (19). Finally, the effect of continuization is to change C by at most $O(\sqrt{EC})$, so the asymptotics remain true in discrete time.

Remark. Presumably (c.f. Open Problem 6) the Theorem remains true without hypothesis (b).

In view of Chapter 6 yyy it is surprising that there is no obvious example to disprove

Open Problem 10 *Let V denote the last state to be hit. In a sequence of vertex-transitive graphs with $n \to \infty$, is it always true that V converges (in variation distance, say) to the uniform distribution?*

1.6 Product chains

In our collection of examples in Chapter 5 of random walks on graphs, the examples with enough symmetry to fit into the present setting have in fact extra symmetry, enough to fit into the arc-transitive setting of section 2. So in a sense, working at the level of generality of symmetric reversible chains merely serves to illustrate what properties of chains depend only on this minimal level of symmetry. But let us point out a general construction. Suppose we have symmetric reversible chains $X^{(1)}, \ldots, X^{(d)}$ on state spaces $I^{(1)}, \ldots, I^{(d)}$. Fix constants $a_1, \ldots, a_d$ with each $a_i > 0$ and with $\sum_i a_i = 1$. Then (c.f. Chapter 4 section yyy) we can define a "product chain" with state-space $I^{(1)} \times \ldots \times I^{(d)}$ and transition probabilities

$$(x_1, \ldots, x_d) \to (x_1, \ldots, x'_i, \ldots, x_d)\text{: probability } a_i P(X^{(i)}_1 = x'_i | X^{(i)}_0 = x_i).$$

This product chain is also symmetric reversible. But if the underlying chains have extra symmetry properties, these extra properties are typically lost when one passes to the product chain. Thus we have a general method of constructing symmetric reversible chains which lack extra structure. Example 14 below gives a case with distinct underlying components, and Example 11 gives a case with a non-uniform product. In general, writing $(\lambda^{(i)}_u : 1 \leq u \leq |I^{(i)}|)$ for the continuous-time eigenvalues of $X^{(i)}$, we have (Chapter 4 yyy) that the continuous-time eigenvalues of the product chain are

$$\lambda_{\mathbf{u}} = a_1 \lambda^{(1)}_{u_1} + \ldots + a_d \lambda^{(d)}_{u_d}$$

indexed by $\mathbf{u} = (u_1, \ldots, u_d) \in \{1, \ldots, |I^{(1)}|\} \times \ldots \times \{1, \ldots, |I^{(d)}|\}$. So in particular

$$\tau_2 = \max_i \frac{\tau^{(i)}_2}{a_i}$$

$$\tau_0 = \sum_{\mathbf{u} \neq (1,\ldots,1)} \frac{1}{a_1 \lambda^{(1)}_{u_1} + \ldots + a_d \lambda^{(d)}_{u_d}}$$

and of course these parameters take the same values in discrete time.

Example 11 *Coordinate-biased random walk on the d-cube.*

Take $I = \{0,1\}^d$ and fix $0 < a_1 \leq a_2 \leq \ldots \leq a_d$ with $\sum_i a_i = 1$. Then the chain with transitions

$$(b_1, \ldots, b_d) \to (b_1, \ldots, 1 - b_i, \ldots, b_d) : \text{probability } a_i$$

is the weighted product of two-state chains. Most of the calculations for simple symmetric random walk on the d-cube done in Chapter 5 Example yyy extend to this example, with some increase of complexity. In particular,

$$\tau_2 = \frac{1}{2a_1}$$

$$\tau_0 = \frac{1}{2} \sum_{\mathbf{u} \in I, \mathbf{u} \neq \mathbf{0}} \frac{1}{\sum_{i=1}^d u_i a_i}.$$

In continuout time we still get the product form for the distribution at time t:

$$P_b(X_t = b') = 2^{-d} \prod_i (1 + \eta_i \exp(-2a_i t)) \ ; \ \eta_i = 1 \text{ if } b'_i = b_i, \ = 0 \text{ if not.}$$

So in a sequence of continuous time chains with $d \to \infty$, the "separation" parameter $\tau_1^{(1)}$ of Chapter 3 section yyy is asymptotic to the solution t of

$$\sum_i \exp(-2a_i t) = -\log(1 - e^{-1}).$$

More elaborate calculations can be done to study τ_1 and the discrete-time version.

1.7 The cutoff phenomenon and the upper bound lemma

Chapter 2 yyy and Chapter 4 yyy discussed quantifications of notions of "time to approach stationarity" using variation distance. The emphasis in Chapter 4 yyy was on inequalities which hold up to universal constants. In the present context of symmetric reversible chains, one can seek to do sharper calculations. Thus for random walk on the d-cube (Chapter 5 Example yyy), with chances $1/(d+1)$ of making each possible step or staying still, writing $n = 2^d$ and $c_n = \frac{1}{4} d \log d$, we have (as $n \to \infty$) not only the fact $\tau_1 \sim c_n$ but also the stronger result

$$d((1+\varepsilon)c_n) \to 0 \text{ and } d((1-\varepsilon)c_n) \to 1, \text{ for all } \varepsilon > 0. \qquad (20)$$

We call this the *cutoff phenomenon*, and when a sequence of chains satisfies (20) we say the sequence has "variation cutoff at c_n". As mentioned at xxx, the general theory of Chapter 4 works smoothly using $\bar{d}(t)$, but in examples it is more natural to use $d(t)$, which we shall do in this chapter. Clearly,

(20) implies the same result for $\bar{d}$ and implies $\tau_1 \sim c_n$. Also, our convention in this chapter is to work in discrete time, whereas the Chapter 4 general theory worked more smoothly in continuous time. (Clearly (20) in discrete time implies the same result for the continuized chains, provided $c_n \to \infty$). Note that, in the context of symmetric reversible chains,

$$d(t) = d_i(t) = ||P_i(X_t \in \cdot) - \pi(\cdot)|| \text{ for each } i.$$

We also can discuss *separation distance* (Chapter 4 yyy) which in this context is

$$s(t) = 1 - n \min_j P_i(X_t = j) \text{ for each } i,$$

and introduce the analogous notion of *separation threshold.*

It turns out that these cut-offs automatically appear in sequences of chains defined by repeated products. An argument similar to the analysis of the d-cube (see [4] for a slightly different version) shows

Lemma 12 *Fix an aperiodic symmetric reversible chain with m states and with relaxation time $\tau_2 = 1/(1-\lambda_2)$. Consider the d-fold product chain with $n = m^d$ states and transition probabilities*

$$(x_1, \ldots, x_d) \to (x_1, \ldots, y_i, \ldots, y_d) : \textit{probability } \frac{1}{d}\ p_{x_i, y_i}.$$

As $d \to \infty$, this sequence of chains has variation cutoff $\frac{1}{2}\tau_2 d \log d$ and separation cut-off $\tau_2 d \log d$.

xxx discuss upper bound lemma
xxx heuristics
xxx mention later examples

1.8 Vertex-transitive graphs and Cayley graphs

So far we have worked in the setting of symmetric reversible chains, and haven't used any graph theory. We now specialize to the case of random walk on a vertex-transitive or Cayley graph $(\mathcal{V}, \mathcal{E})$. As usual, we won't write out all specializations of the previous results, but instead emphasize what extra we get from graph-theoretic arguments. Let d be the degree of the graph.

Lemma 13 *For random walk on a vertex-transitive graph,*
(i) $E_v T_x \geq n$ *if* $(v, x) \notin \mathcal{E}$
(ii) $\frac{2dn}{d+1} - d \geq E_v T_x \geq \frac{dn}{d+1}$ *if* $(v, x) \in \mathcal{E}$

Proof. The lower bounds are specializations of Lemma 2(i), i.e. of Chapter 6 xxx. For the upper bound in (ii),

$$\begin{aligned} n-1 &= \frac{1}{d}\sum_{y\sim x} E_yT_x \qquad (21)\\ &\geq \frac{1}{d}\left(E_vT_x+(d-1)\frac{dn}{d+1}\right) \text{ by the lower bound in (ii).} \end{aligned}$$

Rearrange.

xxx mention general lower bound $\tau_0 \geq (1-o(1))nd/(d-2)$ via tree-cover.

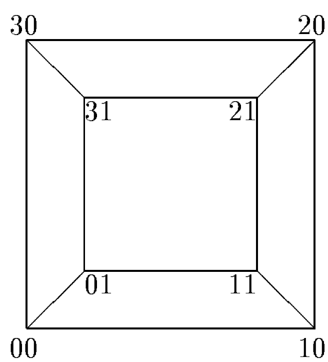
It is known (xxx ref) that a Cayley graph of degree d is d-edge-connected, and so Chapter 6 Proposition yyy gives

$$\tau^* \leq n^2\psi(d)/d$$

where $\psi(d)/d \approx \sqrt{2/d}$.

Example 14 *A Cayley graph where E_vT_w is not the same for all edges (v,w).*

Consider $Z_m \times Z_2$ with generators $(1,0),(-1,0),(0,1)$. The figure illustrates the case $m=4$.



Let's calculate $E_{00}T_{10}$ using the resistance interpretation. Put unit voltage at 10 and zero voltage at 00, and let a_i be the voltage at $i0$. By symmetry the voltage at $i1$ is $1-a_i$, so we get the equations

$$a_i = \frac{1}{3}(a_{i-1}+a_{i+1}+(1-a_i)),\ 1\leq i\leq m-1$$

with $a_0=a_m=0$. But this is just a linear difference equation, and a brief calculation gives the solution

$$a_i = \frac{1}{2} - \frac{1}{2}\frac{\theta^{m/2-i}+\theta^{i-m/2}}{\theta^{m/2}+\theta^{-m/2}}$$

where $\theta = 2 - \sqrt{3}$. The current flow is $1 + 2a_1$, so the effective resistance is $r = (1 + 2a_1)^{-1}$. The commute interpretation of resistance gives $2E_{00}T_{01} = 3nr$, and so

$$E_{00}T_{01} = \frac{3n}{2(1+2a_1)}$$

where $n = 2m$ is the number of vertices. In particular,

$$n^{-1}\ E_{00}T_{01} \to \gamma \equiv \frac{3}{1+\sqrt{3}} \text{ as } n \to \infty.$$

Using the averaging property (21)

$$n^{-1}\ E_{00}T_{10} \to \gamma' \equiv \frac{3\sqrt{3}}{2(1+\sqrt{3})} \text{ as } n \to \infty.$$

Turning from hitting times to mixing times, recall the Cheeger constant

$$\tau_c \equiv \sup_A c(A)$$

where A is a proper subset of vertices and

$$c(A) \equiv \frac{\pi(A^c)}{P_\pi(X_1 \in A^c | X_0 \in A)}.$$

For random walk on a Cayley graph one can use simple "averaging" ideas to bound $c(A)$. This is Proposition 15 below. The result in fact extends to vertex-transitive graphs by a covering graph argument - see xxx.

Consider a n-vertex Cayley graph with degree d and generators $\mathcal{G} = \{g_1, \ldots, g_d\}$, where $g \in \mathcal{G}$ implies $g^{-1} \in \mathcal{G}$. Then

$$P_\pi(X_1 \in A^c | X_0 \in A) = \frac{1}{d} \sum_{g \in \mathcal{G}} \frac{|Ag \setminus A|}{|A|}$$

where $Ag = \{ag : a \in A\}$. Lower bounding the sum by its maximal term, we get

$$c(A) \leq \frac{d}{n}\ \frac{|A|\ |A^c|}{\max_{g \in \mathcal{G}} |Ag \setminus A|}. \tag{22}$$

Proposition 15 *On a Cayley graph of degree d*

(i) $\tau_c \leq d\Delta$, *where* Δ *is the diameter of the graph.*

(ii) $c(A) \leq 2d\rho(A)$ *for all* A *with* $\rho(A) \geq 1$, *where*

$$\rho(A) \equiv \min_{v \in \mathcal{V}} \max_{w \in A} d(v, w)$$

is the radius of A.

Note that $\sup_A \rho(A)$ is bounded by Δ but not in general by $\Delta/2$ (consider the cycle), so that (ii) implies (i) with an extra factor of 2. Part (i) is from Aldous [2] and (ii) is from Babai [5].

Proof. (i) Fix A. Because

$$\frac{1}{n} \sum_{v \in \mathcal{V}} |A \cap Av| = |A|^2/n$$

there exists some $v \in \mathcal{V}$ such that $|A \cap Av| \leq |A|^2/n$, implying

$$|Av \setminus A| \geq |A||A^c|/n. \tag{23}$$

We can write $v = g_1 g_2 \ldots g_\delta$ for some sequence of generators (g_i) and some $\delta \leq \Delta$, and

$$|Av \setminus A| \leq \sum_{i=1}^{\delta} |Ag_1 \ldots g_i \setminus Ag_1 \ldots g_{i-1}| = \sum_{i=1}^{\delta} |Ag_i \setminus A|.$$

So there exists $g \in \mathcal{G}$ with $|Ag \setminus A| \geq \frac{1}{\Delta} \times |A||A^c|/n$, and so (i) follows from (22). For part (ii), fix A with $|A| \leq n/2$, write $\rho = \rho(A)$ and suppose

$$\max_{g \in \mathcal{G}} |Ag \setminus A| < \frac{1}{4\rho}|A|. \tag{24}$$

Fix v with $\max_{w \in A} d(w, v) = \rho$. Since $|Ag \setminus A| < \frac{1}{4\rho}|A|$ and

$$A \setminus Axg \subseteq (A \setminus Ag) \cup (A \setminus Ax)g$$

we have by induction

$$|A \setminus Ax| < \frac{1}{4\rho}|A| d(x, v). \tag{25}$$

Write $B^r \equiv \{vg_1 \ldots g_i; i \leq r, g_i \in \mathcal{G}\}$ for the ball of radius r about v. Since $(2\rho + 1)/(4\rho) < 1$, inequality (25) shows that $A \cap Ax$ is non-empty for each $x \in B^{2\rho+1}$, and so $B^{2\rho+1} \subseteq A^{-1}A$. But by definition of ρ we have $A \subseteq B^\rho$, implying $B^{2\rho+1} \subseteq B^{2\rho}$, which in turn implies $B^{2\rho}$ is the whole group. Now (25) implies that for every x

$$|Ax \setminus A| < \frac{1}{2}|A| \leq \frac{|A||A^c|}{n}.$$

But this contradicts (23). So (24) is false, i.e.

$$\max_{g \in \mathcal{G}} |Ag \setminus A| \geq \frac{1}{4\rho}|A| \geq \frac{1}{2\rho} \, \frac{|A||A^c|}{n}.$$

By complementation the final inequality remains true when $|A| > n/2$, and the result follows from (22).

1.9 Comparison arguments for eigenvalues

The "distinguished paths" method of bounding relaxation times (Chapter 4 yyy) can also be used to compare relaxation times of two random flights on the same group, and hence to bound one "unknown" relaxation time in terms of a second "known" relaxation time. This approach has been developed in great depth in

xxx ref Diaconis Saloff-Coste papers.

Here we give only the simplest of their results, from [13].

Consider generators $\mathcal{G}$ of a group I, and consider a reversible random flight with step-distribution μ supported on $\mathcal{G}$. Write $d(x, \mathrm{id})$ for the distance from x to the identity in the Cayley graph, i.e. the minimal length of a word

$$x = g_1 g_2 \ldots g_d;\ g_i \in \mathcal{G}.$$

For each x choose some minimal-length word as above and define $N(g, x)$ to be the number of occurences of g in the word. Now consider a different reversible random flight on I with some step-distribution $\tilde{\mu}$, not necessarily supported on $\mathcal{G}$. If we know $\tilde{\tau}_2$, the next result allows us to bound τ_2.

Theorem 16

$$\frac{\tau_2}{\tilde{\tau}_2} \leq K \equiv \max_{g \in \mathcal{G}} \frac{1}{\mu(g)} \sum_{x \in I} d(x, id) N(g, x) \tilde{\mu}(x).$$

xxx give proof – tie up with L^2 discussion

Perhaps surprisingly, Theorem 16 gives information even when the comparison walk is the "trivial" walk whose step-distribution $\tilde{\mu}$ is uniform on the group. In this case, both $d(x, \mathrm{id})$ and $N(g, x)$ are bounded by the diameter Δ, giving

Corollary 17 *For reversible flight with step-distribution μ on a group I,*

$$\tau_2 \leq \frac{\Delta^2}{\min_{g \in \mathcal{G}} \mu(g)},$$

where $\mathcal{G}$ is the support of μ and Δ is the diameter of the Cayley graph associated with $\mathcal{G}$.

When μ is uniform on $\mathcal{G}$ and $|\mathcal{G}| = d$, the Corollary gives the bound $d\Delta^2$, which improves on the bound $8d^2\Delta^2$ which follows from Proposition 15 and Cheeger's inequality (Chapter 4 yyy). The examples of the torus Z_N^d show that Δ^2 enters naturally, but one could hope for the following variation.

Open Problem 18 *Write $\tau_* = \tau_*(I, \mathcal{G})$ for the minimum of τ_2 over all symmetric random flights on I with step-distribution supported on $\mathcal{G}$. Is it true that $\tau_* = O(\Delta^2)$?*

2 Arc-transitivity

Example 14 shows that random walk on a Cayley graph does not necessarily have the property that $E_v T_w$ is the same for all edges (v, w). It is natural to consider some stronger symmetry condition which does imply this property. Call a graph *arc-transitive* if for each 4-tuple of vertices (v_1, w_1, v_2, w_2) such that (v_1, w_1) and (v_2, w_2) are edges, there exists an automorphism γ such that $\gamma(v_1) = w_1, \gamma(v_2) = w_2$. Arc-transitivity is stronger than vertex-transitivity, and immediately implies that $E_v T_w$ is constant over edges (v, w).

Lemma 19 *On a n-vertex arc-transitive graph,*
(i) $E_v T_w = n - 1$ for each edge (v, w).
(ii) $E_v T_w \geq n - 2 + d(v, w)$ for all $w \neq v$.

Proof. (i) follows from $E_v T_v^+ = n$. For (ii), write $N(w)$ for the set of neighbors of w. Then

$$E_v T_w = E_v T_{N(w)} + (n - 1)$$

and $T_{N(w)} \geq d(v, N(w)) = d(v, w) - 1$.

In particular, $\min_{w \neq v} E_v T_w = n - 1$, which gives the following bounds on mean cover time EC. The first assertion uses Matthews method for expectations (Chapter 2 yyy) and the second follows from Theorem 9.

Corollary 20 *On a n-vertex arc-transitive graph, $EC \geq (n - 1)h_{n-1}$. And if $\tau_0/n \to 1$ and $\tau_2 = o(n/\log n)$ then*

$$\frac{C - \tau_0 \log n}{\tau_0} \xrightarrow{d} \eta \tag{26}$$

Note that the lower bound $(n - 1)h_{n-1}$ is attained on the complete graph. It is not known whether this exact lower bound remains true for vertex-transitive graphs, but this would be a consequence of Chapter 6 Open Problem yyy. Note also that by xxx the hypothesis $\tau_0/n \to 1$ can only hold if the degrees tend to infinity.

Corollary 20 provides easily-checkable conditions for the distributional limit for cover times, in examples with ample symmetry, such as the card-shuffling examples in the next section. Note that

$$(26) \text{ and } \tau_0 = n\left(1 + \frac{b + o(1)}{\log n}\right) \text{ imply } \frac{C - n\log n - bn}{n} \stackrel{d}{\to} \eta.$$

Thus on the d-cube (Chapter 5 yyy) $\tau_0 = n\left(1 + \frac{1+o(1)}{d}\right) = n\left(1 + \frac{\log 2 + o(1)}{\log n}\right)$ and so

$$\frac{C - n\log n - n\log 2}{n} \stackrel{d}{\to} \eta.$$

2.1 Card-shuffling examples

These examples are formally random flights on the permutation group, though we shall describe them informally as models for random shuffles of a m-card deck. Write X_t for the configuration of the deck after t shuffles, and write $Y_t = f_1(X_t)$ for the position of card 1 after t shuffles. In most examples (and all those we discuss) Y_t is itself a Markov chain on $\{1, 2, \ldots, m\}$. Example 21, mentioned in Chapter 1 xxx, has become the prototype for use of group representation methods.

Example 21 *Card-shuffling via random transpositions.*

The model is

> Make two independent uniform choices of cards, and interchange the positions of the two cards.

With chance $1/m$ the same card is chosen twice, so the "interchange" has no effect. This model was studied by Diaconis and Shahshahani [14], and more concisely in the book Diaconis [12] Chapter 3D. The chain Y_t has transition probabilities

$$\begin{array}{lll} i \to j & \text{probability} & 2/m^2, \ j \neq i \\ i \to i & \text{probability} & 1 - \dfrac{2(m-1)}{m^2} \end{array}$$

This is essentially random walk on the complete m-graph (precisely: the continuized chains are deterministic time-changes of each other) and it is easy to deduce that (Y_t) has relaxation time $m/2$. So by the contraction principle xxx the card-shuffling process has $\tau_2 \geq m/2$, and group representation methods show

$$\tau_2 = m/2. \qquad (27)$$

Since the chance of being in the initial state after 1 step is $1/m$ and after 2 steps in $O(1/m^2)$, the local transience heuristic (10) suggests

$$\tau_0 = m!(1 + 1/m + O(1/m^2)) \qquad (28)$$

which can be verified by group representation methods (see Flatto et al [18]). The general bound on τ_1 in terms of τ_2 gives only $\tau_1 = O(\tau_2 \log m!) = O(m^2 \log m)$. In fact group group representation methods ([12]) show

$$\text{there is a variation cutoff at } \frac{1}{2} m \log m. \qquad (29)$$

Example 22 *Card-shuffling via random adjacent transpositions.*

The model is

> With probability $1/(m+1)$ do nothing. Otherwise, choose one pair of adjacent cards (counting the top and bottom cards as adjacent), with probability $1/(m+1)$ for each pair, and interchange them.

The chain Y_t has transition probabilities

$$\begin{array}{rcl} i \rightarrow & i+1 & \text{probability } 1/(m+1) \\ i \rightarrow & i-1 & \text{probability } 1/(m+1) \\ i \rightarrow & i & \text{probability } (m-1)/(m+1) \end{array}$$

with $i \pm 1$ counted modulo m. This chain is (in continuous time) just a time-change of random walk on the m-cycle, so has relaxation time

$$a(m) \equiv \frac{m+1}{2} \frac{1}{1 - \cos(2\pi/m)} \sim \frac{m^3}{4\pi^2}.$$

So by the contraction principle xxx the card-shuffling process has $\tau_2 \geq a(m)$, and (xxx unpublished Diaconis work) in fact

$$\tau_2 = a(m) \sim m^3/4\pi^2.$$

A coupling argument which we shall present in Chapter xxx gives an upper bound $\tau_1 = O(m^3 \log m)$ and (xxx unpublished Diaconis work) in fact

$$\tau_1 = \Theta(m^3 \log m).$$

The local transience heuristic (10) again suggests

$$\tau_0 = m!(1 + 1/m + O(1/m^2))$$

but this has not been studied rigorously.

Many variants of these examples have been studied, and we will mention a generalization of Examples 21 and 22 in Chapter xxx. Here is another example, from Diaconis and Saloff-Coste [13], which illustrates the use of comparison arguments.

Example 23 *A slow card-shuffling scheme.*

The model is: with probability 1/3 each, either
(i) interchange the top two cards
(ii) move the top card to the bottom
(iii) move the bottom card to the top.
This process is random walk on a certain Cayley graph, which (for $m \geq 3$) is not arc-transitive. Writing d for distances in the graph and writing

$$\beta = \max(d(\sigma, \mathrm{id}) : \sigma \text{ a transposition }),$$

it is easy to check that $\beta \leq 3m$. Comparing the present chain with the "random transpositions" chain (Example 21), denoted by $\tilde{}$, Theorem 16 implies

$$\frac{\tau_2}{\tilde{\tau}_2} \leq 3\beta^2.$$

Since $\tilde{\tau}_2 = m/2$ we get

$$\tau_2 \leq \frac{27m^3}{2}.$$

2.2 Cover times for the d-dimensional torus Z_N^d.

This is Example yyy from Chapter 5, with $n = N^d$ vertices, and is clearly arc-transitive. Consider asymptotics as $N \to \infty$ for d fixed. We studied mean hitting times in this example in Chapter 5. Here $\tau_0/n \not\to 1$, so we cannot apply Corollary 20. For $d = 1$ the graph is just the d-cycle, treated in Chapter 6 yyy. For $d \geq 3$, Chapter 5 yyy gave

$$E_0 T_i \sim n R_d \text{ as } N \to \infty,\ |i| \to \infty$$

where $|i|$ is Euclidean distance on the torus, i.e.

$$|(i_1, \ldots, i_d)|^2 = \sum_{u=1}^{d} (\min(i_u, N - i_u))^2.$$

So EC has the asymptotic upper bound $R_d n \log n$. Now if we apply the subset form of Matthews method (Chapter 6 yyy) to the subset

$$A = \{(j_1 m, \ldots, j_d m) : 1 \leq j_i \leq \frac{N}{m}\} \tag{30}$$

then we get a lower bound for EC asymptotic to

$$\log |A| \times n R_d.$$

By taking $m = m(n) \uparrow \infty$ slowly, this agrees with the upper bound, so we find

Corollary 24 *On the d-dimensional torus with $d \geq 3$,*

$$EC \sim R_d n \log n.$$

Perhaps surprisingly, the case $d = 2$ turns out to be the hardest of all explicit graphs for the purposes of estimating cover times. (Recall this case is the *white screen problem* Chapter 1 xxx.) Loosely, the difficulty is caused by the fact that $\tau_2 = \Theta(n \log n)$ – recall from Chapter 6 yyy that another example with this property, the balanced tree, is also hard. Anyway, for the case $d = 2$ the calculations in Chapter 5 yyy gave

$$E_0 T_i \sim n \left(\frac{2}{\pi} \log |i| \ + O(1) \right).$$

This leads to the upper bound in Corollary 25 below. For the lower bound, we repeat the $d \geq 3$ argument using a subset of the form (30) with $m \to \infty$, and obtain a lower bound asymptotic to

$$\frac{2}{\pi} \log m \ \times \ \log(n^2/m^2).$$

The optimal choice is $m \sim n^{1/2}$, leading to the lower bound below.

Corollary 25 *On the 2-dimensional torus Z_N^2,*

$$\left(\frac{1}{4\pi} - o(1)\right) n \log^2 n \leq EC \leq \left(\frac{1}{\pi} + o(1)\right) n \log^2 n.$$

Lawler [23] has improved the constant in the lower bound to $\frac{1}{2\pi}$ – see Notes. It is widely believed that the upper bound is in fact the limit.

Open Problem 26 *Prove that, on the 2-dimensional torus Z_N^2,*

$$EC \sim \frac{1}{\pi}\ n \log^2 n.$$

The usual distributional limit

$$\frac{C - \tau_0 \log n}{\tau_0} \ \stackrel{d}{\to}\ \eta$$

certainly fails in $d = 1$ (see Chapter 6 yyy). It has not been studied in $d \geq 2$, but the natural conjecture it that it is true for $d \geq 3$ but false in $d = 2$. Note that (by Chapter 6 yyy) the weaker concentration result

$$C/EC \ \stackrel{d}{\to}\ 1$$

holds for all $d \geq 2$.

2.3 Bounds for the parameters

In Chapter 6 we discussed upper bounds on parameters τ for regular graphs. One can't essentially improve these bounds by imposing symmetry conditions, because the bounds are attained (up to constants) by the n-cycles. But what if we exclude the n-cycles? Example 14 shows that one can invent vertex-transitive graphs which mimic the n-cycle, but it is not clear whether such arc-transitive graphs exist. So perhaps the next-worst arc-transitive graph is Z_m^2.

Open Problem 27 *Is it true that, over arc-transitive graphs excluding the n-cycles, $\tau^* = O(n \log n)$, $\tau_2 = O(n)$ and $\frac{\tau^*}{2\tau_0} = 1 + o(1)$?*

2.4 Group-theory set-up

Recall that the Cayley graph associated with a set $\mathcal{G}$ of generators of a group I has edges

$$\{(v, vg); v \in I, g \in \mathcal{G}\}$$

where we assume $\mathcal{G}$ satisfies
(i) $g \in \mathcal{G}$ implies $g^{-1} \in \mathcal{G}$.
To ensure that the graph is arc-transitive, it is sufficient to add the condition

(ii) for each pair g_1, g_2 in $\mathcal{G}$, there exists a group automorphism γ such that $\gamma(\mathrm{id}) = \mathrm{id}$ and $\gamma(g_1) = g_2$.
In words, "the stabilizer acts transitively on $\mathcal{G}$". This is essentially the general case: see [8] Prop. A.3.1.

As a related concept, recall that elements x, y of a group I are *conjugate* if $x = g^{-1}yg$ for some group element g. This is an equivalence relation which therefore defines *conjugacy classes*. It is easy to check that a conjugacy class must satisfy condition (ii). Given a conjugacy class C one can consider the uniform distribution μ_C on C and then consider the random flight with step distribution μ_C. Such random flights fit into the framework of section 2, and Example 21 and the torus Z_N^d are of this form. On the other hand, Example 22 satisfies (i) and (ii) but are not random flights with steps uniform on a conjugacy class.

3 Distance-regular graphs

A graph is called *distance-transitive* if for each 4-tuple v_1, w_1, v_2, w_2 with $d(v_1, w_1) = d(v_2, w_2)$ there exists an automorphism γ such that $\gamma(v_1) = w_1, \gamma(v_2) = w_2$. Associated with such a graph of diameter Δ are the *intersection numbers* $(a_i, b_i, c_i; 0 \leq i \leq \Delta)$ defined as follows. For each i choose (v, w) with $d(v, w) = i$, and define

$$\begin{aligned} c_i &= \text{ number of neighbors of } w \text{ at distance } \quad i-1 \quad \text{ from } v \\ a_i &= \text{ number of neighbors of } w \text{ at distance } \quad i \quad \text{ from } v \\ b_i &= \text{ number of neighbors of } w \text{ at distance } \quad i+1 \quad \text{ from } v. \end{aligned}$$

The distance-transitive property ensures that (a_i, b_i, c_i) does not depend on the choice of (v, w). A graph for which such intersection numbers exist is called *distance-regular*, and distance-regularity turns out to be strictly weaker than distance-transitivity. An encyclopedic treatment of such graphs has been given by Brouwer et al [8]. The bottom line is that there is almost a complete characterization (i.e. list of families and sporadic examples) of distance-regular graphs. Anticipating a future completion of the characterization, one could seek to prove inequalities for random walks on distance-regular graphs by simply doing explicit calculations with all the examples, but (to quote Biggs [7]) "this would certainly not find a place in *The Erdos Book* of ideal proofs". Instead, we shall just mention some properties of random walk which follow easily from the definitions.

Consider random walk (X_t) on a distance-regular graph started at v_0, and define $D_t = d(v_0, X_t)$. Then (D_t) is itself a Markov chain on states $\{0, 1, \ldots, \Delta\}$, and is in fact the birth-and-death chain with transition probabilities

$$p_{i,i-1} = c_i/r, \ p_{i,i} = a_i/r, \ p_{i,i+1} = b_i/r.$$

xxx b-and-d with holds

Finding exact t-step transition probabilities is tantamount to finding the orthogonal polynomials associated with the distance-regular graph – references to the latter topic can be found in [8], but we shall not pursue it.

3.1 Exact formulas

A large number of exact formulas can be derived by combining the standard results for birth-and-death chains in Chapter 5 section yyy with the standard renewal-theoretic identities of Chapter 2 section yyy. We present only the basic ones.

Fix a state $\mathbf{0}$ in a distance-regular graph. Let n_i be the number of states at distance i from $\mathbf{0}$. The number of edges with one end at distance i and the other at distance $i+1$ is $n_i b_i = n_{i+1} c_{i+1}$, leading to the formula

$$n_i = \prod_{j=1}^{i} \frac{b_{j-1}}{c_j}; \ 0 \leq i \leq \Delta.$$

The chain D_t has stationary distribution

$$\rho_i = n_i/n = n^{-1} \prod_{j=1}^{i} \frac{b_{j-1}}{c_j}; \ 0 \leq i \leq \Delta.$$

Switching to the notation of Chapter 5 yyy, the chain D_t is random walk on a weighted linear graph, where the weight w_i on edge $(i-1, i)$ is

$$w_i = \frac{n_{i-1} b_{i-1}}{2n} = \frac{n_i c_i}{2n}, \ 1 \leq i \leq \Delta$$

and total weight $w = 1$. This graph may have self-loops, but they don't affect the formulas. Clearly hitting times on the graph are related to hitting times of (D_t) by

$$E_v T_x = h(d(v,x)) \text{ , where } h(i) \equiv \tilde{E}_i T_0 \tag{31}$$

and where we write $\tilde{}$ to refer to expectations for D_t. Clearly $h(\cdot)$ is strictly increasing. Chapter 5 yyy gives the formula

$$h(i) = i + 2\sum_{j=1}^{i} \sum_{i=j+1}^{\Delta} w_i w_j^{-1}. \tag{32}$$

And Chapter 5 yyy gives the last equality in

$$\tau_0 = E_\pi T_{\mathbf{0}} = \tilde{E}_\rho T_0 = \frac{1}{2}\sum_{i=1}^{\Delta} w_i^{-1} (\sum_{j\geq i} \rho_j)^2. \tag{33}$$

Finally, Chapter 5 yyy gives

$$\tilde{E}_0 T_\Delta + \tilde{E}_\Delta T_0 = \sum_{i=1}^{\Delta} 1/w_i. \tag{34}$$

Thus if the graph has the property that there exists a unique vertex $\mathbf{0}^*$ at distance Δ from $\mathbf{0}$, then we can pull back to the graph to get

$$\frac{\tau^*}{2} = \max_{x \neq v} E_x T_v = E_{\mathbf{0}} T_{\mathbf{0}^*} = \frac{1}{2}\sum_{i=1}^{\Delta} 1/w_i. \tag{35}$$

If the graph lacks that property, we can use (31) to calculate $h(\Delta)$.

The general identities of Chapter 3 yyy can now be used to give formulas for quantities such as $P_x(T_y < T_z)$ or E_x (number of visits to y before T_z).

3.2 Examples

Many treatments of random walk on sporadic examples such as regular polyhedra have been given, e.g. [26, 27, 31, 34, 35, 36, 37], so I shall not repeat them here. Of infinite families, the complete graph was discussed in Chapter 5 yyy, and the complete bipartite graph is very similar. The d-cube also was treated in Chapter 5. Closely related to the d-cube is a model arising in several contexts under different names,

Example 28 c-*subsets of a* d-*set.*

The model has parameters (c,d), where $1 \leq c \leq d-1$. Formally, we have random walk on the distance-transitive graph whose vertices are the $\frac{d!}{c!(d-c)!}$ c-element subsets $A \subset \{1,2,\ldots,d\}$, and where (A,A') is an edge iff $|A \triangle A'| = 2$. More vividly, d balls $\{1,2,\ldots,d\}$ are distributed between a left urn and a right urn, with c balls in the left urn, and at each stage one ball is picked at random from each urn, and the two picked balls are interchanged. The induced birth-and-death chain is often called the *Bernouilli-Laplace diffusion model.* The analysis is very similar to that of the d-cube. See [15, 16] and [12] Chapter 3F for details on convergence to equilibrium and [11] for hitting and cover times.

3.3 Monotonicity properties

The one result about random walk on distance-regular graphs we wish to highlight is the monotonicity property given in Proposition 29 below. Part (ii) can be viewed as a strengthening of the monotonicity property for mean hitting times (by integrating over time and using the formula relating mean hitting times to the fundamental matrix).

Proposition 29 *For random walk* (X_t) *on a distance-regular graph in continuous time,* $P_v(X_t = w) = q(t, d(v,w))$*, where the function* $d \to q(t,d)$ *satisfies*
(i) $d \to q(t,d)$ *in non-increasing, for fixed* t.
(ii) $q(t,d)/q(t,0)$ *in non-decreasing in* t*, for fixed* d.

xxx proof – coupling – defer to coupling Chapter ??

Proposition 29 is a simple example of what I call a "geometric" result about a random walk. Corollary 3 gave a much weaker result in a more general setting. It's natural to ask for intermediate results, e.g.

Open Problem 30 *Does random walk on an arc-transitive graph have some monotonicity property stronger than that of Corollary 3?*

3.4 Extremal distance-regular graphs

Any brief look at examples suggests

Open Problem 31 *Prove that, over distance-regular graphs excluding the* n*-cycles,* $\tau_0 = O(n)$.

Of course this would imply $\tau^* = O(n)$ and $EC = O(n \log n)$. As mentioned earlier, one can try to tackle problems like this by using the list of known distance-regular graphs in [8]. Biggs [7] considered the essentially equivalent problem of the maximum value of $\max_{i,j} E_i T_j/(n-1)$, and found the value 195/101 taken on the cubic graph with 102 vertices, and outlined an argument that this may be the *max* over known distance-regular graphs.

xxx in same setting is $\tau_2 = O(\log n)$?

3.5 Gelfand pairs and isotropic flights

On a distance-regular graph, a natural generalization of our nearest-neighbor random walks is to *isotropic* random flight on the graph. Here one specifies a probability distribution $(s_0, s_1, \ldots, s_\Delta)$ for the step-length S, and each step moves to a random vertex at distance S from the previous vertex. Precisely, it is the chain with transition probabilities

$$p(v, w) = \frac{s_{d(v,w)}}{n_{d(v,w)}}. \tag{36}$$

The notion of isotropic random flight also makes sense in continuous space. For an isotropic random flight in R^d, the steps have some arbitrary specified random length S and a direction θ which is uniform and independent of S. A similar definition can be made on the d-dimensional sphere. The abstract notion which captures distance-regular graphs and their continuous analogs is a *Gelfand pair.* Isotropic random flights on Gelfand pairs can be studied in great detail by analytic methods. Brief accounts can be found in Letac [24, 25] and Diaconis [12] Chapter 3F, which contains an extensive annotated bibliography.

4 Notes on Chapter 7

Diaconis [12] Chapter 3 discusses random walks on groups, emphasizing use of the upper bound lemma to establish bounds on τ_1 and $d(t)$, and containing extensive references to previous work using group-theoretic methods. We have only mentioned reversible examples, but many natural non-reversible examples can also be handled by group representation methods. Also, in Example 21 and related examples, group representation methods give stronger information about $d(t)$ then we have quoted.

Elementary properties of hitting and cover times on graphs with symmetry structure have been noted by many authors, a particularly comprehensive treatment being given in the Ph.D. thesis Sbihi [33]. Less extensive treatments and specific elementary results can be found in many of the papers cited later, plus [30, 36, 37]

Section 1. The phrase "random flight" is classically used for R^d. I have used it (as did Takacs [34, 35]) in place of "random walk' to emphasize it is not necessarily a nearest-neighbor random walk.

Section 1.3. Other elementary facts about symmetric reversible chains are

$$E_\pi \min(T_i, T_j) = \frac{n}{2}(Z_{ii} + Z_{ij}).$$

$$P_i(X_{2t} = i) + P_i(X_{2t} = j) \geq 2/n.$$

Chapter 6 yyy showed that on any regular graph, $\max_{i,j} E_i T_j \leq 3n^2$. On a vertex-transitive graph the constant "3" can be improved to "2", by an unpublished argument of the author, but this is still far from the natural conjecture of 1/4.

Section 1.4. Another curious result from [3]is that for a symmetric reversible chain the first passage time cannot be concentrated around its mean:

$$\frac{\mathrm{var}\ _iT_j}{(E_iT_j)^2} \geq \frac{e-2}{e-1} - \frac{1}{E_iT_j}.$$

Section 1.5. Before Matthews method was available, a result like Corollary 5 (c) required a lot of work – see Aldous [1] for a result in the setting of non-reversible random flight on a group. The present version of Corollary 5 (c) is a slight polishing of ideas in Zuckerman [38] section 6.

The fact that (17) implies (15) is a slight variation of the usual textbook forms of the continuity theorem ([17] 2.3.4 and 2.3.11) for Fourier and Laplace transforms. By the same argument as therein, it is enough for the limit transform to be continuous at $\theta = 0$, which holds in our setting.

Matthews [28, 29] introduced Proposition 8 and used it to obtain the limiting cover time distribution for the d-cube and for card-shuffling examples. Devroye and Sbihi [11] applied it to generalized hypercubes and to Example 28. Our implementation in Theorem 9 and Corollary 20 reduces the need for ad hoc calculations in particular examples.

Section 1.6. Example 11 has been studied in the reliability literature (e.g. [22]) from the viewpoint of the exponential approximation for hitting times.

Section 1.7. The factor of 2 difference between the variation and separation cutoffs which appears in Lemma 12 is the largest possible – see Aldous and Diaconis [4].

Section 1.8. xxx walk-regular example – McKay paper.

Section 1.9. Diaconis and Saloff-Coste [13] give many other applications of Theorem 16. We mention some elsewhere; others include

xxx list.

Section 2. The name "arc-transitive" isn't standard: Biggs [6] writes "symmetric" and Brouwer et al [8] write "flag-transitive". Arc-transitivity is not necessary for the property "E_vT_w is constant over edges". For instance, a graph which is vertex-transitive and edge-transitive (in the sense of undirected edges) has the property, but is not necessarily arc-transitive [20]. Gobel and Jagers [19] observed that the property

$$E_vT_w + E_wT_v = 2(n-1) \text{ for all edges } (v,w)$$

(equivalently: the effective resistance across each edge is constant) holds for arc-transitive graphs and for trees.

Section 2.2. Sbihi [33] and Zuckerman [38] noted that the subset version of Matthews method could be applied to the d-torus to give Corollaries 24 and 25.

The related topic of the time taken by random walk on the infinite lattice Z^d to cover a ball centered at the origin has been studied independently – see Revesz [32] Chapter 22 and Lawler [23], who observed that similar arguments could be applied to the d-torus, improving the lower bound in Corollary 25. It is easy to see an informal argument suggesting that, for random walk on the 2-torus, when n^α vertices are unvisited the set of unvisited vertices has some kind of fractal structure. No rigorous results are known, but heuristics are given in Brummelhuis and Hilhorst [9].

Section 3.1. Deriving these exact formulas is scarcely more than undergraduate mathematics, so I am amazed to see that research papers have continued to be published in the 1980s and 1990s claiming various special or general cases as new or noteworthy.

Section 3.5. In the setting of isotropic random flight (36) with step-length distribution q, it is natural to ask what conditions on q and q' imply that $\tau(q) \geq \tau(q')$ for our parameters τ. For certain distributions on the d-cube, detailed explicit calculations by Karlin et al [21] establish an ordering of the entire eigenvalue sequences, which in particular implies this inequality for τ_2 and τ_0. Establishing results of this type for general Gelfand pairs seems an interesting project.

Miscellaneous. On a finite *field*, such as Z_p for prime p, one can consider "random walks" with steps of the form $x \to \alpha x + \beta$, with a specified joint distribution for (α, β). Chung et al [10] treat one example in detail.

References

[1] D.J. Aldous. On the time taken by random walks on finite groups to visit every state. *Z. Wahrsch. Verw. Gebiete*, 62:361–374, 1983.

[2] D.J. Aldous. On the Markov chain simulation method for uniform combinatorial distributions and simulated annealing. *Probab. Engineering Inform. Sci.*, 1:33–46, 1987.

[3] D.J. Aldous. Hitting times for random walks on vertex-transitive graphs. *Math. Proc. Cambridge Phil. Soc.*, 106:179–191, 1989.

[4] D.J. Aldous and P. Diaconis. Strong uniform times and finite random walks. *Adv. in Appl. Math.*, 8:69–97, 1987.

[5] L. Babai. Local expansion of vertex-transitive graphs and random generation in finite groups. In *Proc. 23rd ACM Symp. Theory of Computing*, pages 164–174, 1991.

[6] N. L. Biggs. *Algebraic Graph Theory.* Cambridge University Press, 1974.

[7] N. L. Biggs. Potential theory on distance-regular graphs. *Combin. Probab. Comput.*, 2:243–255, 1993.

[8] A.E. Brouwer, A.M. Cohen, and A. Neumaier. *Distance Regular Graphs.* Springer-Verlag, 1989.

[9] M.J.A.M. Brummelhuis and H.J. Hilhorst. Covering of a finite lattice by a random walk. *Physica A*, 176:387–408, 1991.

[10] F.R.K. Chung, P. Diaconis, and R. L. Graham. A random walk problem arising in random number generation. *Ann. Probab.*, 15:1148–1165, 1987.

[11] L. Devroye and A Sbihi. Random walks on highly-symmetric graphs. *J. Theoretical Probab.*, 3:497–514, 1990.

[12] P. Diaconis. *Group Representations in Probability and Statistics.* Institute of Mathematical Statistics, Hayward CA, 1988.

[13] P. Diaconis and L. Saloff-Coste. Comparison theorems for random walk on finite groups. *Ann. Probab.*, 21:2131–2156, 1993.

[14] P. Diaconis and M. Shahshahani. Generating a random permutation with random transpositions. *Z. Wahrsch. Verw. Gebiete*, 57:159–179, 1981.

[15] P. Diaconis and M. Shahshahani. Time to reach stationarity in the Bernouilli-Laplace diffusion model. *SIAM J. Math. Anal.*, 18:208–218, 1986.

[16] P. Donnelly, P. Lloyd, and A. Sudbury. Approach to stationarity of the Bernouilli-Laplace diffusion model. *Adv. in Appl. Probab.*, 26:715–727, 1994.

[17] R. Durrett. *Probability: Theory and Examples.* Wadsworth, Pacific Grove CA, 1991.

[18] L. Flatto, A.M. Odlyzko, and D.B. Wales. Random shuffles and group representations. *Ann. Probab.*, 13:154–178, 1985.

[19] F. Gobel and A.A. Jagers. Random walks on graphs. *Stochastic Process. Appl.*, 2:311–336, 1974.

[20] D.F. Holt. A graph which is edge-transitive but not arc-transitive. *J. Graph Theory*, 5:201–204, 1981.

[21] S. Karlin, B. Lindquist, and Y-C Yao. Markov chains on hypercubes: Spectral representations and several majorization relations. *Random Struct. Alg.*, 4:1–36, 1993.

[22] J. Keilson. *Markov Chain Models - Rarity and Exponentiality.* Springer-Verlag, 1979.

[23] G. Lawler. On the covering time of a disc by simple random walk in two dimensions. In *Seminar in Stochastic Processes 1992*, pages 189–208. Birkhauser, 1993.

[24] G. Letac. Problemes classiques de probabilite sur un couple de Gelfand. In *Analytic Methods in Probability Theory.* Springer-Verlag, 1981. Lecture Notes in Math. 861.

[25] G. Letac. Les fonctions spheriques d'un couple de Gelfand symmetrique et les chaines de Markov. *Adv. in Appl. Probab.*, 14:272–294, 1982.

[26] G. Letac and L. Takacs. Random walks on a 600-cell. *SIAM J. Alg. Disc. Math.*, 1:114–120, 1980.

[27] G. Letac and L. Takacs. Random walks on a dodecahedron. *J. Appl. Probab.*, 17:373–384, 1980.

[28] P.C. Matthews. *Covering Problems for Random Walks on Spheres and Finite Groups.* PhD thesis, Statistics, Stanford, 1985.

[29] P.C. Matthews. Covering problems for Markov chains. *Ann. Probab.*, 16:1215–1228, 1988.

[30] J. L. Palacios. Expected cover times of random walks on symmetric graphs. *J. Theoretical Probab.*, 5:597–600, 1992.

[31] J. L. Palacios. Fluctuation theory for the Ehrenfest urn model via electrical networks. *Adv. in Appl. Probab.*, 25:472–476, 1993.

[32] P. Revesz. *Random Walk in Random and Non-Random Scenery.* World Scientific, Singapore, 1990.

[33] A.M. Sbihi. *Covering Times for Random Walks on Graphs.* PhD thesis, McGill University, 1990.

[34] L. Takacs. Random flights on regular polytopes. *SIAM J. Alg. Disc. Meth.*, 2:153–171, 1981.

[35] L. Takacs. Random flights on regular graphs. *Adv. in Appl. Probab.*, 16:618–637, 1984.

[36] A.R.D. van Slijpe. Random walks on regular polyhedra and other distance-regular graphs. *Statist. Neerlandica*, 38:273–292, 1984.

[37] A.R.D. van Slijpe. Random walks on the triangular prism and other vertex-transitive graphs. *J. Comput. Appl. Math.*, 15:383–394, 1986.

[38] D. Zuckerman. A technique for lower bounding the cover time. *SIAM J. Disc.Math.*, 5:81–87, 1992.