The Annals of Applied Probability
1991, Vol. 1, No. 1, 62–87

EIGENVALUE BOUNDS ON CONVERGENCE TO STATIONARITY FOR NONREVERSIBLE MARKOV CHAINS, WITH AN APPLICATION TO THE EXCLUSION PROCESS[1]

BY JAMES ALLEN FILL

The Johns Hopkins University

We extend recently developed eigenvalue bounds on mixing rates for reversible Markov chains to nonreversible chains. We then apply our results to show that the d-particle simple exclusion process corresponding to clockwise walk on the discrete circle $\mathbf{Z}_p$ is rapidly mixing when d grows with p. The dense case $d = p/2$ arises in a Poisson blockers problem in statistical mechanics.

1. Introduction.

1.1. *Overview.* In this paper we establish bounds on the variation distance from stationarity for an ergodic but generally nonreversible Markov chain. We focus on the application of such bounds to a particular interacting (finite) particle system, namely, the simple exclusion process corresponding to clockwise walk on the discrete circle, to show that the process becomes nearly random in time polynomial in the number of points on the circle.

The simple exclusion process was introduced by Spitzer (1970) and is treated in Liggett (1985), which gives history and references (see Section VIII.6 there). We recall the definition of a (finite particle) exclusion process generated by a stochastic matrix $q = (q(x, y))_{x \in S, y \in S}$ on a finite or countably infinite set S. We superimpose an exclusion interaction on d otherwise independent continuous time Markov chains on S with jump matrix q as follows. The d particles are originally configured on S according to some specified initial distribution π_0 on the system's state space $\{0, 1\}^S$, with 1 at site $x \in S$ corresponding to the presence of a particle. The particles move on S according to the following rules:

1. a particle at $x \in S$ waits an $\mathrm{Exp}(1/d)$ time, after which
2. it chooses a $y \in S$ with probability $q(x, y)$;
3. if y is vacant, it moves to y, while if y is occupied, it remains at x.

In particular, there is always at most one particle per site. Note: We have scaled transition rates so that there is an average of one attempted particle

Received December 1989; revised August 1990.

[1]Research supported by the National Security Agency under Grant MDA904-89-H-2051.

AMS 1980 *subject classifications*. Primary 60J10; secondary 60J27, 60K35, 15A42.

Key words and phrases. Markov chains, reversibility, variation distance, rates of convergence, rapid mixing, Poincaré inequality, Cheeger's inequality, chi-square distance, interacting particle systems, exclusion process, Poisson blockers.

move per unit time. We could equally well treat a discrete time version of the process.

In Section 4 we analyze the exclusion process with S equal to the discrete circle $\mathbf{Z}_p$ (with addition modulo p) and $q(x, x+1) = 1$ for every x, corresponding to (deterministic) clockwise walk. Let the sites $0, \dots, p-1$ be placed clockwise about a circle. To give a fuller description of the process, we keep track of the individual particles. For this it is necessary only to arbitrarily designate one of the particles from the beginning as particle 1 and to record at each instant of time the position of particle 1 and the list of occupied sites. We can label the particles clockwise as $1, \dots, d$, beginning with particle 1. The state of the labelled exclusion process at any time is the ordered d-tuple of sites for particles $1, \dots, d$ at that time.

The labelled exclusion process for clockwise walk on $\mathbf{Z}_p$ is an irreducible Markov jump process on the state space $\mathscr{S}$ consisting of all clockwise rotations of strictly increasing d-tuples from $\{0, \dots, p-1\}$. The nonzero entries of the generator G are, for $i = 1, \dots, d$,

$$g((x_1, \dots, x_i, \dots, x_d), (x_1, \dots, x_i + 1, \dots, x_d)) = 1/d$$

provided $x_i + 1 \neq x_{i+1}$ (with subscript addition modulo d) and

$$g(\mathbf{x}, \mathbf{x}) = -\frac{1}{d}|\{i: x_i + 1 \neq x_{i+1}\}|.$$

Since $\sum_{\mathbf{x} \in \mathscr{S}} g(\mathbf{x}, \mathbf{y}) = 0$, the continuous time analogue of double stochasticity, the stationary distribution π is uniform over the $d\binom{p}{d}$ states of $\mathscr{S}$.

In Section 4 we show that the labelled particle system is rapidly mixing, that is, becomes nearly uniform in time polynomial in $p \ll |\mathscr{S}|$, when d grows with p. As this is the first bound, to the author's knowledge, on the rate of convergence to stationarity for an interacting particle system, we find the result interesting in its own right. Furthermore, in Section 5 we explain how the dense case $d = p/2$ arises in a statistical mechanics problem involving the passage across an interval of a particle blocked by the events of Poisson processes. For the dense case, our results are summarized in the following theorem, which shows that the process becomes nearly stationary in time at most of order p^7. The theorem is an immediate consequence of Proposition 4.4, the variation upper bound of Theorem 2.14 and the trivial bound $\binom{p}{p/2} < 2^p$.

THEOREM 1.1 (Rapid mixing for exclusion process). *Consider the labelled clockwise exclusion process with $p = 2d$. Let λ denote the second smallest eigenvalue of $-A(G)$, the negative of the additive reversiblization of the generator G. Then*

$$\frac{3/2}{p^6} \leq \lambda \leq \frac{12}{p(p^2-1)} \tag{1.1}$$

and consequently, for any initial configuration $\mathbf{x}$, *the variation distance*

$\|P_{\mathbf{x},\cdot}(t) - \pi\|$ *at time* $t = cp^7$ *satisfies*

$$\|P_{\mathbf{x},\cdot}(t) - \pi\| \le \left(\frac{p}{8}\right)^{1/2} \exp\left[-\frac{1}{2}(3c - \log 2)p\right]. \tag{1.2}$$

REMARK 1.2. We briefly consider two other possible approaches to this problem.

(a) *Coupling*. Couplings have been used by a number of authors to prove results about the exclusion process; see Liggett [(1985), Chapter VIII] for discussion and references. For simplicity return to the *unlabelled* clockwise process. We describe a rather natural coupling for studying convergence to stationarity; see Liggett [(1985), Section VIII.2] for related discussion of this graphical representation of the exclusion process.

Associate independent rate $1/d$ Poisson processes $(N_j(t))$ with the sites $j \in \mathbf{Z}_p$. At each event time of N_j, a particle is moved from j to $j+1$, provided that j is occupied and $j+1$ is vacant just prior to that time. Then versions of the exclusion process for all (distributions of) initial configurations are defined simultaneously on the same probability space. It is not hard to see that for any two versions of the process, the number of differences in occupied sites decreases monotonically to zero. Call the (random) time T at which the two configurations first agree the *coupling time* for the two versions. If one version of the process begins in a specified distribution π_0 and the other begins in stationarity [uniform over the $\binom{p}{d}$ configurations], then the tails of the coupling time distribution provide an upper bound on variation distance: see, for example, Lemma 5 in Chapter 4 of Diaconis (1988). But we do not know how to analyze this stopping time.

(b) *Strong stationary duality*. Diaconis and Fill (1990a, b) [see Fill (1990a) for continuous time] have introduced the notion of the *strong stationary dual* of a Markov chain. This is an absorbing Markov chain for which the time to absorption is distributed as the time to stationarity for the original chain. Such a dual exists when the state space $\mathscr{S}$ is partially ordered and the given chain possesses certain monotonicity properties: see Fill (1990b) for further explanation. For the linear exclusion process considered in Section 4.1, started in $(1, 2, \ldots, d)$, the monotonicity conditions are met and we can quite explicitly describe the dual. But the stopping time of interest again seems difficult to analyze.

1.2. *Organization*. The theory needed to extend the variation distance bounds of Diaconis and Stroock (1991) to nonreversible chains is developed in Section 2. In Section 3 we work a few simple examples. Section 4 develops results for the exclusion process and Section 5 describes the connection with the Poisson blockers problem.

2. Bounds for nonreversible chains. In this section we give a simple proof of an identity of Mihail (1989) and derive from it an upper bound on the variation distance from stationarity for an ergodic but not necessarily reversible Markov chain in terms of eigenvalues of two associated reversible chains. Using the Poincaré inequality we obtain bounds on the eigenvalues of these chains in terms of geometric quantities of associated edge weighted graphs.

2.1. *Multiplicative reversiblization and the variation upper bound.* The following setup will be used throughout. Let P be a given ergodic (i.e., irreducible and aperiodic) transition matrix on a finite state space $\mathscr{S}$. Let π be the (unique) stationary distribution and let $\tilde{P}$ denote the time reversal of P:

$$\tilde{P}(x,y) := \frac{\pi(y)P(y,x)}{\pi(x)}. \tag{2.1}$$

$\tilde{P}$ is also ergodic with stationary distribution π and its time reversal is P. Define the *multiplicative reversiblization* $M(P)$ of P by

$$M(P) := P\tilde{P}. \tag{2.2}$$

Since PR has stationary distribution π and the reversal of PR is $\tilde{R}\tilde{P}$ when both P and R are ergodic with the same stationary distribution π, it follows that $M(P)$ is indeed a reversible transition matrix. We claim that the eigenvalues of $M(P)$ are all real and nonnegative, so that $\beta_1(M) \in [0,1]$. Indeed, for any matrix M, define

$$S(M) = D^{1/2}MD^{-1/2} \quad \text{with } D := \operatorname{diag}(\pi).$$

Observe that $S(\cdot)$ is an algebraic isomorphism that preserves eigenvalues. We have

$$S(\tilde{P}) = D^{1/2}(D^{-1}P^tD)D^{-1/2} = (S(P))^t,$$

and so $S(M(P)) = S(P)(S(P))^t$ is nonnegative definite.

Denote the second largest eigenvalue of $M(P)$ by $\beta_1(M)$ (the largest eigenvalue being unity). Let P^n be the n-step transition matrix, so that $P^n(x,\cdot)$ is the law at time n of the chain P started in state x. Finally, let $\|\sigma - \pi\| := \max\{|\sigma(A) - \pi(A)|: A \subset \mathscr{S}\} = \frac{1}{2}\sum_{x \in \mathscr{S}}|\sigma(x) - \pi(x)|$ be the variation distance from a given probability mass function σ to π. With this notation we are prepared to state the main result of Section 2.

THEOREM 2.1 (Variation upper bound). *Let P be an ergodic transition matrix on a finite state space $\mathscr{S}$ and let π be the stationary distribution. Let $\beta_1(M)$ denote the second largest eigenvalue of the multiplicative reversiblization $M(P)$ of P. Then for any $x \in \mathscr{S}$,*

$$4\|P^n(x,\cdot) - \pi\|^2 \le (\beta_1(M))^n/\pi(x). \tag{2.3}$$

Theorem 2.1 is a corollary of Theorem 2.7, which handles arbitrary initial distributions π_0 in a manner that makes use of the closeness of π_0 to π.

REMARK 2.2. (a) If P is reversible ($\tilde{P} = P$), then $M(P) = P^2$ and $\beta_1(M) = |\beta|^2$, where β maximizes $|b|$ over all choices of eigenvalues $b \neq 1$ of P. In this case, Theorem 2.1 reproduces Proposition 3 in Diaconis and Stroock (1991).

(b) There is another natural way of reversiblizing P. Let

$$A(P) := (P + \tilde{P})/2 \tag{2.4}$$

denote the *additive reversiblization* of P. Like $M(P)$, $A(P)$ is indeed a reversible transition matrix. As will be seen from examples, $A(P)$ is often easier to analyze than $M(P)$. Corollary 2.9 gives a reformulation of Theorem 2.1 in terms of $A(P)$.

2.2. *An identity of Mihail.* We derive Theorem 2.1 from a simple identity first stated by Mihail (1989) in combinatorial language. Given a function ϕ on $\mathscr{S}$, let $\operatorname{Var}\phi := \operatorname{Var}\phi(X_\infty)$, where X_∞ is a random variable taking values in $\mathscr{S}$ with distribution π, and for any transition matrix R, let $R\phi$ be given by

$$(R\phi)(x) := \sum_y R(x,y)\phi(y). \tag{2.5}$$

Given functions ϕ and ψ on $\mathscr{S}$, define the Dirichlet form

$$\mathscr{E}_R(\phi,\psi) := \tfrac{1}{2}\sum_{x,y} (\phi(y) - \phi(x))(\psi(y) - \psi(x))\pi(x)R(x,y) \tag{2.6}$$

based on a given reversible R having stationary distribution π. We rephrase equation (3.22) in Mihail (1989) in the present probabilistic setting and give a simplified proof.

PROPOSITION 2.3 (Mihail's identity). *Let P be an ergodic transition matrix on a finite state space $\mathscr{S}$ and let π be the stationary distribution. Calculate variances with respect to π and let $\mathscr{E}_{M(P)}$ be the Dirichlet form* (2.6) *for the multiplicative reversiblization $M(P)$ of P. Then for any function ϕ on $\mathscr{S}$,*

$$\operatorname{Var}\phi = \operatorname{Var}(\tilde{P}\phi) + \mathscr{E}_{M(P)}(\phi,\phi), \tag{2.7}$$

with $\tilde{P}\phi$ defined by (2.5).

PROOF. Without loss of generality, ϕ has zero mean under π. Let $\langle \phi, \psi\rangle = \sum_x \pi(x)\phi(x)\psi(x)$ denote the $L^2(\pi)$ inner product and note that $\mathscr{E}_R(\phi,\psi) = \langle \phi, (I-R)\psi\rangle$ when R is reversible with respect to π. Thus

$$\begin{aligned}\mathscr{E}_{M(P)}(\phi,\phi) &= \langle \phi, (I - M(P))\phi\rangle = \operatorname{Var}\phi - \langle \phi, P\tilde{P}\phi\rangle \\ &= \operatorname{Var}\phi - \langle \tilde{P}\phi, \tilde{P}\phi\rangle = \operatorname{Var}\phi - \operatorname{Var}(\tilde{P}\phi),\end{aligned}$$

where we have used the fact that $\tilde{P}$ is the adjoint of P on $L^2(\pi)$. □

2.3. *Reversiblizations.* As the next lemma shows, when P is *strongly aperiodic*, in the sense that $P(x,x) \geq \frac{1}{2}$ for every x, the Dirichlet form $\mathscr{E}_{M(P)}$ appearing in Mihail's identity can be decomposed into Dirichlet forms based on two other reversible transition matrices associated with P.

LEMMA 2.4 (Dirichlet decomposition). *Let P be an ergodic transition matrix as in Proposition 2.3. Suppose additionally that P is strongly aperiodic. Let M(P) and A(P) denote the multiplicative and additive reversiblizations, respectively, of P. Then we have the following identity of Dirichlet forms*:

$$\mathscr{E}_{M(P)} = \mathscr{E}_{A(P)} + \tfrac{1}{4}\mathscr{E}_{M(2P-I)}.$$

PROOF. According to the hypotheses, $2P - I$ is a stochastic matrix with stationary distribution π. Moreover, by direct calculation, we have

$$M(P) = A(P) + \tfrac{1}{4}M(2P - I) - \tfrac{1}{4}I.$$

Since the diagonal entries of a transition matrix R do not affect the Dirichlet form $\mathscr{E}_R(\phi, \psi)$, the assertion follows. □

In many examples $A(P)$ is considerably simpler to analyze than $M(P)$ or $M(2P - I)$. The following immediate consequence of Lemma 2.4 is then helpful.

PROPOSITION 2.5. *Suppose that P in Proposition* 2.3 *is strongly aperiodic. Then for any* ϕ,

$$\operatorname{Var}\phi \ge \operatorname{Var}(\tilde{P}\phi) + \mathscr{E}_{A(P)}(\phi, \phi). \tag{2.8}$$

REMARK 2.6. Strong aperiodicity can always be arranged by introducing enough additional (constant) holding probability at each state. However, analysis of the chain so modified does *not* yield analysis of the original chain. In continuous time, on the other hand, multiplicative reversiblization of the transition function corresponds without reservation to additive reversiblization of the generator (see Section 2.7).

2.4. *Proof of the variation upper bound from Mihail's identity.* In this section we apply Mihail's identity to derive bounds on the "chi-square" distance from stationarity for a nonreversible Markov chain. Using the Cauchy–Schwarz inequality, we relate this distance to the more standard variation distance and, in particular, establish Theorem 2.1.

We begin by defining the chi-square distance. Given an ergodic P with stationary distribution π, consider the Markov chain with transition matrix P and an arbitrarily specified initial distribution π_0. Denote the distribution $\sum_x \pi_0(x)P^n(x, \cdot)$ at time n by π_n. We define the *chi-square distance* from stationarity at time n by

$$\chi_n^2 := \sum_x \frac{\left(\pi_n(x) - \pi(x)\right)^2}{\pi(x)}. \tag{2.9}$$

The choice of terminology is apparent from the similarity with the chi-square goodness-of-fit statistic.

Since χ_n^2 is simply the variance with respect to π (in the sense of Section 2.2) of the likelihood ratio $\rho_n := \pi_n/\pi$ and

$$(\tilde{P}\rho_n)(x) = \sum_y \tilde{P}(x,y)\frac{\pi_n(y)}{\pi(y)} = \sum_y \frac{\pi_n(y)P(y,x)}{\pi(x)} = \rho_{n+1}(x)$$

using (2.5) and (2.1), Mihail's identity implies

$$\chi_n^2 = \chi_{n+1}^2 + \mathscr{E}_{M(P)}(\rho_n, \rho_n). \tag{2.10}$$

But by the minimax characterization of $\beta_1(M)$ [see (2.19)], the second term on the right here is at least $(1-\beta_1(M))\chi_n^2$, so $\chi_{n+1}^2 \le \beta_1(M)\chi_n^2$. Using induction we obtain the following bound on chi-square distance:

$$\chi_n^2 \le (\beta_1(M))^n \chi_0^2. \tag{2.11}$$

Now by the Cauchy–Schwarz inequality,

$$\begin{aligned} 4\|\pi_n - \pi\|^2 &= \left(\sum_y |\pi_n(y) - \pi(y)|\right)^2 \\ &= \left(\sum_y \sqrt{\frac{\pi(y)}{\pi(y)}}\,|\pi_n(y) - \pi(y)|\right)^2 \le \chi_n^2. \end{aligned} \tag{2.12}$$

We have proved the following theorem, from which Theorem 2.1 follows simply.

THEOREM 2.7. *With the setup of Theorem* 2.1 *and* (2.9), *we may conclude for any initial distribution that*

$$4\|\pi_n - \pi\|^2 \le (\beta_1(M))^n \chi_0^2. \tag{2.13}$$

REMARK 2.8. Suppose for ease of discussion that π is the uniform distribution. From Theorem 2.1 and the triangle inequality, one easily derives the bound

$$4\|\pi_n - \pi\|^2 \le (\beta_1(M))^n |\mathscr{S}|.$$

Theorem 2.7 is always an improvement, and sometimes considerably so. An extreme but artificial example is the case $\pi_0 = \pi$. A more realistic renewal theory example is treated in Chang and Fill (1990).

When P is strongly aperiodic, Theorems 2.1 and 2.7 can be reformulated in terms of $A(P)$, whose second largest eigenvalue we denote by $\beta_1(A) \in [0,1)$.

COROLLARY 2.9. *If P is ergodic and strongly aperiodic, then*

$$4\|\pi_n - \pi\|^2 \le \chi_n^2 \le (\beta_1(A))^n \chi_0^2. \tag{2.14}$$

In particular, for any $x \in \mathscr{S}$,

$$4\|P^n(x,\cdot) - \pi\|^2 \le (\beta_1(A))^n / \pi(x). \tag{2.15}$$

PROOF. Equation (2.15) follows simply from (2.14), which in turn has at least two simple proofs. One proof proceeds as for Theorem 2.7, noting that the Dirichlet decomposition Lemma 2.4 and (2.10) together imply that $\chi_n^2 \geq \chi_{n+1}^2 + \mathscr{E}_{A(P)}(\rho_n, \rho_n)$. This last inequality may also be obtained by setting $\phi = \rho_n$ in Proposition 2.5. Another proof uses Dirichlet decomposition and the minimax characterization

$$(2.16) \qquad \beta_1 = 1 - \inf\{\mathscr{E}_R(\phi, \phi)/\mathrm{Var}(\phi)\colon \phi \text{ is not constant}\}$$

of the second largest eigenvalue of a π-reversible transition matrix R to derive

$$(2.17) \qquad \beta_1(M) \leq \beta_1(A) - \tfrac{1}{4}[1 - \beta_1(M(2P - I))] \leq \beta_1(A). \qquad \square$$

REMARK 2.10. Since $\beta_1(M) \leq \beta_1(A)$, the bound (2.14) is less sharp than (2.13), but not appreciably so. Indeed, since $S(A(P)) = (S(P) + S(P)^t)/2$ and $S(M(P)) = S(P)S(P)^t$, we have as a consequence of a theorem of Fan and Hoffman (1955) [see, for example, Theorem 9.F.4 in Marshall and Olkin (1979)] that $(\beta_1(A))^2 \leq \beta_1(M)$. Thus, application of Corollary 2.9 gives a bound no worse than $4\|\pi_n - \pi\|^2 \leq (\beta_1(M))^{n/2}\chi_0^2$. In applications, such a bound is of greatest interest when P (and hence all other quantities such as π, $\beta_1(M)$, etc.) depends on a parameter d (say) that measures the size of the state space, $\beta_1(M) = e^{-1/f_1(d)}$ and $\chi_0^2 = e^{f_2(d)}$, with $f_i(d) \to \infty$ as $d \to \infty$ for $i = 1, 2$. Then Theorem 2.7 says $4\|\pi_n - \pi\|^2 \leq e^{-c}$ for $n = f_1(d)f_2(d) + cf_1(d)$, while Corollary 2.9 gives the same bound for n twice as large. Either bound implies that order $f_1(d)f_2(d)$ steps are sufficient for near-stationarity when d is large. But, as discussed in Diaconis and Stroock [(1991), Section 1D], the truth typically is that order $f_1(d)$ steps are both necessary and sufficient, the discrepancy arising from bounding an entire spectrum using only the second largest eigenvalue [see also Diaconis (1988), Chapter 3].

2.5. *The Poincaré inequality.* In practice, one is seldom fortunate enough to know the exact value of $\beta_1(M)$ appearing in the upper bound Theorems 2.1 and 2.7. To simplify the exposition, we suppose for this subsection that R is a given transition matrix reversible with respect to a stationary distribution π satisfying $\pi(x) > 0$ for each $x \in \mathscr{S}$. The following comments then apply to $R = M(P)$ to give upper bounds on $\beta_1(M)$. [We do not need to consider whether $M(P)$ is ergodic.] Application to $R = A(P)$ likewise extends Corollary 2.9.

Our discussion here follows Section 1 of Diaconis and Stroock (1991). The eigenvalues of R are all real and lie in $[-1, 1]$. Call the eigenvalues $1 = \beta_0 \geq \beta_1 \geq \cdots \geq \beta_{|\mathscr{S}|-1} (\geq -1)$. The so-called minimax characterization of the eigenvalues of the real symmetric matrix

$$(2.18) \qquad S(R) := D^{1/2}RD^{-1/2} \quad \text{with } D := \mathrm{diag}(\pi)$$

implies that

$$(2.19) \qquad \beta_1 = 1 - \inf\{\mathscr{E}_R(\phi, \phi)/\mathrm{Var}(\phi)\colon \phi \text{ is not constant}\},$$

with $\mathscr{E}_R$ given by (2.6).

The *underlying graph* of the Markov chain R is an undirected graph (possibly containing loops) with vertex set $\mathscr{S}$ and $\{x, y\}$ an edge if and only if $Q(x, y) > 0$, where $Q(x, y) := \pi(x)R(x, y)$. Given vertices x and y in the same connected component, we randomly choose a (not necessarily geodesic) path (without repeated edges) and refer to it henceforth as the *canonical path* $\Gamma(x, y)$ from x to y. The canonical paths $\Gamma(x, y)$ can be chosen according to any joint distribution, although in all of our applications the various paths are chosen independently (and often deterministically). To simplify the exposition, we assume that our paths have no loops. Define the *length*

$$|\Gamma(x, y)| := \sum_{\vec{e} \in \Gamma(x, y)} 1, \tag{2.20}$$

of $\Gamma(x, y)$ to be the number of edges; here the sum is over oriented edges $\vec{e}$ in the path. Define $Q(\vec{e}) := Q(z, w)$ if $\vec{e} = (z, w)$. With

$$K := \max\left\{ (Q(\vec{e}))^{-1} E \sum_{(x, y): \Gamma(x, y) \ni \vec{e}} |\Gamma(x, y)| \pi(x) \pi(y) : \vec{e} \text{ is an oriented edge} \right\}, \tag{2.21}$$

if the graph is connected, that is, if R is irreducible and $K := \infty$ otherwise, we have the following bound on β_1 in terms of the geometric quantity K. The result is a slight generalization of Proposition 1′ in Diaconis and Stroock (1991) and is proved using (2.19) and the Cauchy–Schwarz inequality.

PROPOSITION 2.11 (Poincaré inequality). *The second largest eigenvalue β_1 of a Markov chain R reversible with respect to an everywhere positive stationary distribution π satisfies*

$$\beta_1 \le 1 - K^{-1} \tag{2.22}$$

with K the Poincaré constant defined at (2.21).

Most of the examples worked in this paper fit either the following specialization of Proposition 2.11 or its continuous time analogue. This special setup was also met in the present author's analysis of the problem of approximating a permanent described in Section 4 of Diaconis and Stroock (1991). Suppose that there exists a constant $r > 0$ such that $R(x, y) = r$ whenever $x \neq y$ and $R(x, y) > 0$ and that the uniform distribution π on $\mathscr{S}$ is a stationary distribution for R. Then $Q(\vec{e}) = r/|\mathscr{S}|$ if $\vec{e}$ is not a loop. Assuming R is irreducible, K equals $(r|\mathscr{S}|)^{-1}$ times the maximum expected sum of lengths of all canonical paths passing through any oriented edge $\vec{e}$ in the graph. This yields:

COROLLARY 2.12. *In the Poincaré inequality, if we have in addition that R is irreducible, π is uniform, $R(x, y) = r$ whenever $x \neq y$ and $R(x, y) > 0$ and no canonical path exceeds γ_* in length ($\equiv$ number of edges), then*

$$\beta_1 \le 1 - \frac{r|\mathscr{S}|}{\gamma_* b}, \tag{2.23}$$

with

$$b := \max_{\vec{e}} (\textit{expected number of canonical paths through } \vec{e}). \tag{2.24}$$

2.6. *Cheeger's inequality.* A standard alternative to the Poincaré inequality for bounding the eigenvalues of a Markov chain is Cheeger's inequality. Let P be a (not necessarily reversible) transition matrix with an everywhere positive stationary distribution. Define the *Cheeger constant* h by

$$\begin{aligned} h &\equiv h(P) \\ &:= \min\left\{ \frac{\sum_{x \in S} \sum_{y \in S^c} \pi(x) P(x,y)}{\pi(S)} : S \neq \varnothing \text{ and } \pi(S) \leq \frac{1}{2} \right\}. \end{aligned} \tag{2.25}$$

The following result for reversible chains, a discrete analogue of Cheeger's (1970) bound on the second smallest eigenvalue of the Laplacian on a compact Riemannian manifold, is stated and proved, for example, in Diaconis and Stroock (1991). See their paper, Sinclair and Jerrum (1989), Lawler and Sokal (1988) and Mihail (1989) for variants, further discussion and historical references.

PROPOSITION 2.13 (Cheeger's inequality). *The second largest eigenvalue β_1 of a transition matrix reversible with respect to an everywhere positive stationary distribution π satisfies*

$$1 - 2h \leq \beta_1 \leq 1 - \tfrac{1}{2} h^2. \tag{2.26}$$

Now let P be nonreversible. Cheeger's inequality may be applied to either $M(P)$ or $A(P)$. For $M(P)$, combining (2.12), (2.11) and the Cheeger upper bound yields

$$4\|\pi_n - \pi\|^2 \leq \chi_n^2 \leq \left(1 - \tfrac{1}{2} h^2(M)\right)^n \chi_0^2, \tag{2.27}$$

with $h(M) \equiv h(M(P))$, recapturing Theorem 3.3.1 in Mihail (1989). For $A(P)$, assume P is strongly aperiodic and use (2.14), the Cheeger upper bound on $\beta_1(A)$ and the observation [a minor variant of which is used in the proof of Theorem 2.3 in Lawler and Sokal (1988)] that $h(P) = h(\tilde{P}) = h(A(P))$ to conclude

$$4\|\pi_n - \pi\|^2 \leq \chi_n^2 \leq \left(1 - \tfrac{1}{2} h^2(P)\right)^n \chi_0^2. \tag{2.28}$$

The bound is improved in Section 3.3 of Mihail (1989) by removing the $\frac{1}{2}$.

Mihail's proof of (2.27) proceeds from (2.10) by showing *directly* that $\mathscr{E}_{M(P)}(\rho_n, \rho_n) \geq \frac{1}{2} h^2(M) \chi_n^2$, without reference to eigenvalues. The advantage of the present approach is that any eigenvalue bound for reversible chains, not just Cheeger's, is available for use in extending the variation bound (2.13). We have chosen to concentrate on the Poincaré inequality rather than Cheeger's because in all known applications where both can be computed, the Poincaré upper bound does better than the Cheeger bound. See Diaconis and Stroock (1991).

2.7. *Continuous time.* In this subsection we derive continuous-time analogues of the variation bounds in Theorems 2.1 and 2.7. Let $P_{xy}(t)$, $x \in \mathscr{S}$, $y \in \mathscr{S}$, $0 \le t < \infty$, be the transition function of an irreducible Markov jump process on a finite state space $\mathscr{S}$, with generator G and stationary distribution π. For each t, define the matrix $\tilde{P}(t) := D^{-1}P(t)^t D$ with $D := \text{diag}(\pi)$, just as in (2.1). Then $(\tilde{P}(t))$, the time reversal of $(P(t))$, is also an irreducible Markov jump process with stationary distribution π; its generator is $\tilde{G} := D^{-1}G^t D$. If $P = (P(t))$ and $R = (R(t))$ are both transition functions on $\mathscr{S}$, then so is $(P(t/2)R(t/2) = R(t/2)P(t/2))$, whose generator is the average of those for P and R. [To see this, recall $P(t) = e^{tG}$.] With $R = \tilde{P}$, we obtain the (multiplicative) reversiblization $(M(P(t/2)))$ of $(P(t))$, with generator

$$A(G) := (G + \tilde{G})/2. \tag{2.29}$$

[Note: $(A(P(t)))$ is *not* in general a transition function.] Denote the second smallest eigenvalue of $-A(G)$ by $\lambda \equiv \lambda_1(A) > 0$ (the smallest eigenvalue being 0). For a given initial distribution π_0, define π_t and the chi-square distance $\chi^2(t)$ just as in discrete time.

THEOREM 2.14. *In the present continuous-time setting we have*

$$4\|\pi_t - \pi\|^2 \le \chi^2(t) \le e^{-2t\lambda}\chi^2(0). \tag{2.30}$$

In particular, for any $x \in \mathscr{S}$,

$$4\|P_{x,\cdot}(t) - \pi\|^2 \le e^{-2t\lambda}/\pi(x). \tag{2.31}$$

PROOF (Sketch). Apply the discrete-time Theorem 2.7 to the Markov chain $P((k/n)t)$, $k = 0, 1, 2, \dots$ with t and n fixed to conclude

$$4\|\pi_t - \pi\|^2 \le \left[\beta_1(M(P(t/n)))\right]^n \chi^2(0).$$

But as $u \downarrow 0$,

$$M(P(u/2)) = e^{uA(G)} = I + uA(G) + O(u^2).$$

Thus, letting $n \to \infty$,

$$M(P(t/n))^n \to e^{2tA(G)}$$

and the second largest eigenvalue of this matrix is $e^{-2t\lambda}$. □

There is a Poincaré inequality for continuous time chains: *If λ is the second smallest eigenvalue of $-G$, where G is the generator of a reversible chain with stationary distribution π (> 0 everywhere) and K is defined as in the paragraph containing* (2.21) *with G replacing R throughout, then $\lambda \ge K^{-1}$.* Corollary 2.12 also extends to continuous time in the obvious fashion.

3. Examples. In this section we present four simple examples illustrating the theory developed in Section 2. A more substantial application to renewal theory is presented in Chang and Fill (1990). The first three examples treated here can be addressed by other techniques, but we find it instructive to see how well the eigenvalue bounds fare in problems where exact results are known. No other analysis is known for the fourth example.

EXAMPLE 3.1. *Clockwise random walk on the discrete circle* $\mathbf{Z}_p$. Consider the additive group $\mathbf{Z}_p$ of integers modulo p as p points placed about a circle. We treat the random walk which, at each step, holds at its present state with probability $r \in (0,1)$ and moves clockwise with probability $\bar{r} = 1 - r$. The transition matrix P is clearly ergodic and doubly stochastic, so the stationary distribution π is uniform. P is strongly aperiodic if and only if $r \geq \frac{1}{2}$. For simplicity, suppose that the walk starts deterministically, say at 0.

For any r we may (and for $r < \frac{1}{2}$ we must) use the multiplicative reversiblization $M(P)$ of P. Here $M(P) = P\tilde{P} = PP^t$ is simple symmetric random walk with holding probability $r^2 + \bar{r}^2 = 1 - 2r\bar{r}$. Thus

$$\begin{aligned}\beta_1(M) &= (1 - 2r\bar{r}) + 2r\bar{r}\cos(2\pi/p) = 1 - 2r\bar{r}(1 - \cos(2\pi/p)) \\ &= 1 - (1 + o(1))4\pi^2 r\bar{r}/p^2 = \exp\left[-(1 + o(1))4\pi^2 r\bar{r}/p^2\right].\end{aligned}$$

The asymptotics here are as $p \to \infty$, with r fixed. From the variation upper bound Theorem 2.1 it follows (cf. Remark 2.10) that order $p^2 \log p$ steps of the clockwise walk are sufficient for near-stationarity. In fact, Fourier analysis making careful use of the entire spectrum of P [as in Example 1 of Chapter 3C in Diaconis (1988)], or the simpler observation that π_n is the distribution of $Y_n \bmod p$ with $Y_n \sim$ binomial $(n, 2r\bar{r})$, shows that order p^2 steps are necessary and sufficient.

Let us examine how Poincaré fares for this example. Clearly $1 - \beta_1(M) = 2r\bar{r}(1 - \beta_1)$, where β_1 is the second largest eigenvalue for simple symmetric random walk without holding. For definiteness, suppose henceforth that p is odd. For x and y in $\mathbf{Z}_p$, choose $\Gamma(x,y)$ deterministically as the unique geodesic path from x to y. An easy calculation gives $K = \frac{1}{12}(p^2 - 1)$ as the Poincaré constant for the no-holding random walk. Thus

$$\beta_1(M) \leq 1 - 24r\bar{r}/(p^2 - 1),$$

which (asymptotically) gives the right order of magnitude and only misses in the p^{-2} term by a factor of $\pi^2/6 < 1.7$. As shown in Example 2.1 in Diaconis and Stroock (1991), Corollary 2.12 also gives the right order of magnitude and does about 50% worse than Proposition 2.11 in bounding the coefficient of p^{-2}.

If $r \geq \frac{1}{2}$, we may alternatively employ the additive reversiblization Corollary 2.9. Here $A(P)$ is simple symmetric random walk with holding probability r and so $\beta_1(A) = 1 - \bar{r}(1 - \cos(2\pi/p))$; the analysis proceeds as before, with a further degradation in the bounds on the coefficient of p^{-2}. For this example, the first inequality in (2.17) is by Dirichlet decomposition an equality. Indeed, $M(2P - I)$ is simple symmetric random walk with holding probability $1 - 4\bar{r}(2r - 1)$ and so the same ϕ achieves the infimum in (2.16) for each of the matrices $M(P)$, $A(P)$ and $M(2P - I)$.

EXAMPLE 3.2. *Simple asymmetric random walk on* $\mathbf{Z}_p$. If P is asymmetric simple random walk on $\mathbf{Z}_p$, the walk $M(P)$ is not simple and so Poincaré analysis becomes messy. However, $A(P)$ *is* a simple walk and can be analyzed as in the preceding example. We present a few of the details for a *continuous time* asymmetric walk mainly to illustrate the continuous time results of Section 2.7.

Consider a walk that starts at 0 (say) and moves counterclockwise and clockwise with respective rates μ and λ. The stationary distribution is uniform. The reversiblized chain moves in each direction at rate $(\mu + \lambda)/2$: recall (2.29). Thus $\lambda_1(A) = (\mu + \lambda)[1 - \cos(2\pi/p)]$ and so by Theorem 2.14,

$$4\|P_{0,\cdot}(t) - \pi\|^2 \le p \exp[-2t(\mu + \lambda)(1 - \cos(2\pi/p))].$$

Again order $p^2 \log p$ time units suffice, while actually order p^2 are necessary and sufficient. The Poincaré constant K for the reversiblized chain equals $(\mu + \lambda)^{-1}\frac{1}{12}(p^2 - 1)$.

EXAMPLE 3.3. *Top in at random card shuffle*. This example is discussed in Aldous and Diaconis (1986), Diaconis (1988) and Diaconis and Fill (1990a). Consider shuffling a deck of d cards by repeatedly removing the top card and inserting it at a random position. This may be formalized as a random walk on the symmetric group S_d with step distribution placing mass $1/d$ at each of the cycles id, (21), (321), (4321), ..., $(d\ d-1 \cdots 1)$. To simplify the analysis, we treat in place of this chain P_0 the strongly aperiodic chain $P = (I + P_0)/2$.

We use additive reversiblization and Poincaré. We can write $A(P) = (I + A(P_0))/2$; then $\beta_1(A) \doteq (1 + \beta_1)/2$, where for brevity $\beta_1 = \beta_1(A(P_0))$. Let $R = A(P_0)$. At first thought it would appear that we can apply Corollary 2.12, with $|\mathscr{S}| = d!$ and $r = (2d)^{-1}$. However, an interchange of the top two cards can result either from moving the top card below the second card (a step according to P) or from moving the second card above the top card (a step according to $\tilde{P}$). Thus while most of the nonzero off-diagonal entries of R equal $(2d)^{-1}$, some equal d^{-1}. Nevertheless, π is uniform and $Q(\vec{e}) \ge (2d)^{-1}/|\mathscr{S}|$ if $\vec{e}$ is not a loop, so from Proposition 2.11, we obtain

$$\beta_1 \le 1 - \frac{(2d)^{-1}|\mathscr{S}|}{\gamma_* b}$$

with γ_* and b as in Corollary 2.12.

Our canonical paths will be deterministic and very simple, but not always geodesic. We construct a path from a given initial order to a given final order using entirely top-in-to-somewhere transitions. Begin by moving the top card in the initial order to the bottom of the deck or to the next-to-bottom position so that the bottom two cards are then in their final relative order. Now move the top card to one of the bottom three positions so that the bottom three cards are then in their final relative order. Continue in this fashion until the deck is brought to its final order.

Clearly $\gamma_* = d - 1$. Now we bound b of (2.24) by first producing a bound [say, $b(j)$] on the number of canonical paths $\Gamma(x, y)$ for which any given edge $\vec{e}$ is the jth edge in the path, where $1 \le j \le d - 1$, and then summing over j.

Let an edge $\vec{e} = (z, w)$ and j be given. Then the bottom $j + 1$ cards of w are in their final (y-) relative order. So there are at most $d!/(j+1)!$ possibilities for y. Also, the top $d - j$ cards of z are precisely the cards in positions j through $d - 1$ (as measured from the top of the deck) in x, in the same order. So there are at most $j!$ possibilities for x. Thus $b(j) \le d!/(j+1) = |\mathscr{S}|/(j+1)$ and hence

$$b \le |\mathscr{S}| \sum_{j=1}^{d-1} (j+1)^{-1} < |\mathscr{S}| \log d.$$

We conclude from Proposition 2.11 that

$$\beta_1 < 1 - \frac{1}{2d^2 \log d}.$$

To get a lower bound on β_1, recall from (2.19) that $\beta_1 \ge 1 - \mathscr{E}_R(\phi,\phi)/\mathrm{Var}(\phi)$ [with $R = A(P_0)$] for any nonconstant ϕ. We use the natural notation

$$Q(\vec{e}) = Q(z,w), \qquad \phi(\vec{e}) = \phi(w) - \phi(z),$$

if $\vec{e} = (z, w)$. Let $\phi(x) = x(d)$, the number on the bottom card of the deck in order x. Then

$$\begin{aligned}\mathscr{E}_R(\phi,\phi) &= \frac{1}{2}\sum_{\vec{e}} \phi^2(\vec{e})Q(\vec{e}) \\ &= \frac{1}{2} \times \frac{1}{|\mathscr{S}|} \times \frac{1}{2d} \times \left(\sum_{\vec{e}\text{ top in}} \phi^2(\vec{e}) + \sum_{\vec{e}\text{ out to top}} \phi^2(\vec{e}) \right).\end{aligned}$$

But

$$\begin{aligned}\sum_{\vec{e}\text{ top in}} \phi^2(\vec{e}) &= \sum_{\vec{e}\text{ top in to bottom}} \phi^2(\vec{e}) \\ &= (d-2)! \sum_{u,v:\, u \ne v} (u-v)^2 = 2(d-2)!d^2\,\mathrm{Var}(\phi)\end{aligned}$$

and similarly for the other sum. Thus

$$\mathscr{E}_R(\phi,\phi) = \frac{1}{2} \times \frac{1}{|\mathscr{S}|} \times \frac{1}{2d} \times 4d^2(d-2)!\,\mathrm{Var}(\phi) = \frac{1}{d-1}\mathrm{Var}(\phi)$$

and therefore

$$\beta_1 \ge 1 - \frac{1}{d-1}.$$

We have shown

$$1 - \frac{1}{2(d-1)} \le \beta_1(A) < 1 - \frac{1}{4d^2 \log d}.$$

Suppose that the deck starts in natural order. Then we find that order $d^3(\log d)^2$ steps are sufficient for near-stationarity. The truth—via strong uniform times—is that order $d \log d$ steps are necessary and sufficient. The exact value of $\beta_1(A)$ is unknown, but it seems reasonable to conjecture that $1 - \beta_1(A)$ is of exact order d^{-1}.

EXAMPLE 3.4. *Another card shuffle.* Consider shuffling a deck of d cards by repeatedly removing a random (uniform on $\{1, \dots, d\}$) number of cards from the top and replacing them in reverse order. Here the Markov chain *is* reversible; let β_1 denote its second largest eigenvalue. An analysis similar to that for top-in-at-random shows that

$$1 - \frac{1}{d-1} \le \beta_1 \le 1 - \frac{1}{4d^3}$$

and hence that order $d^4 \log d$ steps are sufficient for mixing. This problem seems not to be amenable to Fourier analysis, strong stationary times or coupling. We conjecture that the truth, as for most simple shuffles, is that order $d \log d$ steps are both necessary and sufficient.

4. Application: The exclusion process.

4.1. *The exclusion process on a segment of the integers.* As we shall see in Section 4.2, analysis of the (d-particle) exclusion process for clockwise walk on the discrete circle $\mathbf{Z}_p$ reduces to that of the exclusion process for symmetric random walk on $\{1, \dots, p\}$. Thus, in the general notation of Section 1.1, we take $S = \{1, \dots, p\}$; $q(x, x-1) = \frac{1}{2}$, $x = 2, \dots, p$; $q(x, x+1) = \frac{1}{2}$, $x = 1, \dots, p-1$; $q(1,1) = \frac{1}{2}$; and $q(p,p) = \frac{1}{2}$ for the underlying random walk. The corresponding exclusion process is an irreducible jump process with a symmetric generator G, so the stationary distribution π is uniform over the $\binom{p}{d}$ states of $\mathscr{S}$. Note $g := G(\mathbf{x}, \mathbf{y}) = (2d)^{-1}$ is constant when $\mathbf{x} \ne \mathbf{y}$ and $G(\mathbf{x}, \mathbf{y}) > 0$. Let λ denote the second smallest eigenvalue of $-G$. We shall prove the following results.

PROPOSITION 4.1. *Let λ be the spectral gap for the exclusion process for symmetric random walk on $\{1, \dots, p\}$.*

(a) *Upper bound on λ*:

$$\lambda \le \frac{6}{dp(p+1)} \le \frac{6}{dp^2}.$$

(b) *Lower bound on λ for d fixed*:

$$\lambda \ge \frac{\binom{p}{d}}{2^d d^2 (p-d) p^{d+1}} = (1 + o(1)) \frac{1}{d! 2^d d^2 p^2}.$$

(c) *Lower bound on λ for d growing with p*:

$$\lambda \ge \frac{3}{d(p-1)^2 (p+2)(2p^2 + 5p + 6)} = (1 + o(1)) \frac{3/2}{dp^5}.$$

REMARK 4.2. (a) The exact bounds are valid for any values of p and d. The asymptotic estimate in (c) is valid uniformly in $0 \le d \le p$ as $p \to \infty$. The estimate in (b) is for constant d as $p \to \infty$.

(b) When d is fixed, (a) and (b) trap λ between two constant multiples of p^{-2}, but the ratio of these constants grows *very* rapidly with d; hence our interest in (c).

(c) When the exclusion interaction is removed, the resulting process is d independent random walks on $\{1, \dots, p\}$, for which $\lambda = d^{-1}(1 - \cos(\pi/p)) = (1 + o(1))(\pi^2/2)/(dp^2)$ uniformly in d. By comparison, this is of the same order as the upper bound in part (a) of the proposition.

(d) We do not know the right (i.e., exact) asymptotics for λ. But there is reason to believe that the Poincaré bound $\lambda \geq g|\mathscr{S}|/(\gamma_* b)$ is inadequate for large d, regardless of the choice of canonical paths. Indeed, for the process without exclusions as in the previous remark, this bound can give no better than $3/(d^2(p^2 - 1))$, off by a factor of about d. To see this, observe that γ_* is at least the diameter of the graph, namely, $d(p-1)$; moreover, a simple pigeonhole argument shows that $b \geq \frac{1}{6}p^d(p+1)$.

(e) The inadequacy of the bound in (b) of the proposition for large d stems not just from our crude analysis of b; b truly is large for our choice of paths. For example, when $d = p/2$, one can show $b/|\mathscr{S}| \geq (1 + o(1))(2\pi)^{-1/2}d^{-1/2}(3^{3/2}/4)^d$ by identifying an edge through which many canonical paths pass.

(f) Suppose d grows linearly with p. Combining (c) of the proposition with the variation bound Theorem 2.14 and standard asymptotic estimates, we conclude that order p^7 time is sufficient for near-stationarity, regardless of the distribution of the initial configuration; in particular, *the process is rapidly mixing*. We have not investigated, for any choice of initial distribution, the amount of time *necessary* for near-stationarity.

AN UPPER BOUND ON λ. We shall derive a lower bound on λ using the continuous time analogue of the Poincaré Corollary 2.12, but the paths we use will depend on whether d is regarded as fixed or as growing with p. On the other hand, we can get an upper bound on λ usable in both cases from the continuous time analogue

$$\lambda = \inf\{\mathscr{E}_G(\phi, \phi)/\mathrm{Var}(\phi) \colon \phi \text{ is not constant}\}$$

of (2.19) by choosing a "bad" ϕ. Indeed, let $\phi(x) = \sum_{i=1}^d x_i$. Then, with

$$Q(\vec{e}) = Q(z, w) = \pi(z)G(z, w) = \frac{1}{|\mathscr{S}|} \times \frac{1}{2d} \quad \text{when } \vec{e} = (z, w),$$

$$|\vec{E}| = \text{number of oriented edges} = \frac{2d(p-d)}{p}|\mathscr{S}|,$$

we have

$$\begin{aligned} \mathscr{E}_G(\phi, \phi) &= \frac{1}{2} \sum_{\vec{e}} \phi^2(\vec{e}) Q(\vec{e}) \\ &= \frac{1}{2} \times 1 \times \frac{1}{|\mathscr{S}|} \times \frac{1}{2d} \times |\vec{E}| = \frac{1}{2}\left(1 - \frac{d}{p}\right). \end{aligned}$$

For the variance, note that by a standard formula for sampling without replacement

$$\operatorname{Var}(\phi) = \frac{p-d}{p-1} \times d \times \operatorname{Var} U \quad (\text{where } U \sim \operatorname{unif}\{1,\dots,p\})$$
$$= \frac{p-d}{p-1} \times d \times \frac{1}{12}(p^2-1) = \frac{1}{12} d(p-d)(p+1).$$

Thus $\lambda \le 6/(dp(p+1))$.

A LOWER BOUND ON λ FOR d FIXED. In continuous time we have $\lambda \geq g|\mathscr{S}|/(\gamma_* b)$. Given vertices $\mathbf{x}$ and $\mathbf{y}$, we specify a (deterministic geodesic) two-stage canonical path from $\mathbf{x}$ to $\mathbf{y}$. Let $I \equiv I(\mathbf{x},\mathbf{y}) := \{i \in \{1,\dots,d\}: x_i \geq y_i\}$ and let $\bar{I}$ denote its complement in $\{1,\dots,d\}$. Writing $i_1 < i_2 < \cdots < i_l$ for the elements of I, our path from $\mathbf{x}$ to $\mathbf{y}$ first travels edges along which coordinate i_1 decreases from x_{i_1} to y_{i_1}, then likewise decrements coordinate i_2 from x_{i_2} to $y_{i_2},\dots,$ then decrements coordinate i_l from x_{i_l} to y_{i_l}. It is easy to see that each of the vertices along this first stage of the path lie in $\mathscr{S}$. For the second stage we similarly increment the coordinates in $\bar{I}$, but this time we proceed from right to left. Again all vertices along the path lie in $\mathscr{S}$. Since the paths are geodesic, γ_* equals the diameter of the graph, namely, the distance from $\mathbf{x} = (1,\dots,d)$ to $\mathbf{y} = (p-d+1,\dots,p)$, which is $d(p-d)$.

It seems hard to count the exact number of canonical paths through a given edge, but we can bound b crudely, as follows. Let us, for example, bound the number of paths through an edge $((z_1,\dots,z_j,\dots,z_d),(z_1,\dots,z_j-1,\dots,z_d))$. If a path $\Gamma(\mathbf{x},\mathbf{y})$ through this edge has $I(\mathbf{x},\mathbf{y}) = \{i_1,\dots,i_l\}$ with $i_1 < i_2 < \cdots < i_l$, then $i_k = j$ for some k and, for $i \neq j$, $z_i = y_i$ if $i = i_h$ with $h < k$, and $z_i = x_i$ if either $i = i_h$ with $h > k$ or $i \in \bar{I}$. This limits $(\mathbf{x},\mathbf{y})$ to (far) fewer than p^{d+1} possibilities. Since there are 2^{d-1} possibilities for the set I, the number of paths through the given edge is $\leq 2^{d-1}p^{d+1}$. The same bound applies to edges $((z_1,\dots,z_j,\dots,z_d),(z_1,\dots,z_j+1,\dots,z_d))$. Thus $b \leq 2^{d-1}p^{d+1}$.

Putting everything together,

$$\lambda \geq \frac{(2d)^{-1} \times \binom{p}{d}}{d(p-d) \times 2^{d-1}p^{d+1}}$$
$$= (1+o(1)) \frac{1}{d!2^d d^2 p^2} \quad \text{as } p \to \infty \text{ with } d \text{ fixed}.$$

A LOWER BOUND ON λ FOR d GROWING WITH p. *Canonical paths.* The lower bound we have just derived for λ is unsatisfactory when d grows with p. For such cases we rebuild our canonical paths $\Gamma(\mathbf{x},\mathbf{y}) \equiv \Gamma_{p,d}(\mathbf{x},\mathbf{y})$, $\mathbf{x} \in \mathscr{S}$, $\mathbf{y} \in \mathscr{S}$, inductively in a "semirandom" fashion. That is, certain parts of the paths will be specified deterministically, while other parts will employ randomization to more evenly distribute the numbers of paths through the various edges of the graph.

The induction, on p simultaneously for all values of d in $\{0, \dots, p\}$, is based trivially on the observation that $|\mathscr{S}| = 1$ if $d = 0$ or $d = p$. For the inductive step we may therefore assume that $0 < d < p$. Our canonical paths are of four types, depending on the values of x_d and y_d. The following example illustrates the sort of informal terminology we will use in describing the paths. The specification of a type 1 path means that, if $x_d = p$ and $y_d = p$ and the ordered vertices of $\Gamma_{p-1,d-1}((x_1, \dots, x_{d-1}), (y_1, \dots, y_{d-1}))$ are $(z_1^k, \dots, z_{d-1}^k)$, then $\Gamma_{p,d}(\mathbf{x}, \mathbf{y})$ is defined as the path whose vertices are $(z_1^k, \dots, z_{d-1}^k, p)$. Also, if $\mathbf{x} \in \mathscr{S}_{p,d}$, then $\mathbf{x} + j \in \mathscr{S}_{p,d} \cup \mathscr{S}_{p,d+1}$ is defined in the obvious manner, as is $\mathbf{x} - k \in \mathscr{S}_{p,d} \cup \mathscr{S}_{p,d-1}$.

1. If $x_d = p$ and $y_d = p$, use the $\binom{p-1}{d-1}$-path from $(x_1, \dots, x_{d-1}, p)$ to $(y_1, \dots, y_{d-1}, p)$.
2. If $x_d \le p - 1$ and $y_d \le p - 1$, use the $\binom{p-1}{d}$-path from $\mathbf{x}$ to $\mathbf{y}$.
3. If $x_d = p$ and $y_d \le p - 1$, pick a vacant site at random (uniformly) from $\mathbf{x}$ and call it J. (a) Move from $\mathbf{x}$ to $\mathbf{x}^* = (\mathbf{x} + J) - p \in \mathscr{S}_{p,d}$ deterministically by decrementing the ith coordinate in its turn, $i = 1, \dots, d$, from x_i to x_i^*. [For example, if $p = 8$, $d = 4$, $J = 2$ and $\mathbf{x} = (1, 4, 5, 8)$, then $\mathbf{x}^* = (1, 2, 4, 5)$ is reached by moving the second particle leftward from site 4 to site 2, then the third particle from site 5 to site 4 and finally the fourth particle from site 8 to site 5.] (b) Use the type 2 path from $\mathbf{x}^*$ to $\mathbf{y}$.
4. If $x_d \le p - 1$ and $y_d = p$, pick an occupied site at random (uniformly) from $\mathbf{x}$ and call it K. (a) Move from $\mathbf{x}$ to $\mathbf{x}^* = (\mathbf{x} + p) - K \in \mathscr{S}_{p,d}$ deterministically by incrementing the ith coordinate in its turn, $i = d, \dots, 1$, from x_i to x_i^*. (b) Use the type 1 path from $\mathbf{x}^*$ to $\mathbf{y}$.

It is not hard to show by induction that the paths described do not have any repeated edges. For example, none of the edges in part (a) of a type 3 path are repeated in part (b). Indeed, for edges across which the dth particle remains at site p, this is clear. The other type 3(a) edges move the dth particle leftward and the dth particle makes at most one change of direction in the course of a type 3 path.

Inductive analysis. We use induction to prove the following bounds on γ_* (no longer the diameter of the graph since our paths are not geodesic) and b. When these are combined with the Poincaré bound $\lambda \ge g|\mathscr{S}|/(\gamma_* b)$, we obtain part (c) of Proposition 4.1.

LEMMA 4.3. *Our semirandom canonical paths for the d-particle exclusion process for simple symmetric random walk on* $\{1, \dots, p\}$ *satisfy*

$$\gamma_* \le \tfrac{1}{2}(p-1)(p+2) \quad \textit{and} \quad b \le 2\binom{p}{d}\sum_{j=2}^{p} j^2.$$

PROOF. We may suppose $p > 1$ and $0 < d < p$. For $t = 1, 2, 3, 4$, let $\gamma_*^t(p, d)$ denote the greatest possible length of a $\binom{p}{d}$-path of type t; we may then take $\gamma_* \equiv \gamma_*(p, d) = \max_t \gamma_*^t(p, d)$. Clearly, $\gamma_*^1(p, d) \le \gamma_*(p-1, d-1)$,

$\gamma_*^2(p,d) \le \gamma_*(p-1,d)$, $\gamma_*^3(p,d) < p + \gamma_*^2(p,d) \le p + \gamma_*(p-1,d)$ and $\gamma_*^4(p,d) < p + \gamma_*^1(p,d) \le p + \gamma_*(p-1,d-1)$. Thus

$$\begin{aligned}\gamma_*(p,d) &< p + \max(\gamma_*(p-1,d-1), \gamma_*(p-1,d)) \\ &\le p + \tfrac{1}{2}(p-2)(p+1) = \tfrac{1}{2}(p-1)(p+2) \quad \text{by induction.}\end{aligned}$$

We will bound the maximum $b \equiv b(p,d)$ over all oriented edges $\vec{e}$ of the expected number $b(\vec{e})$ of canonical paths through an edge $\vec{e} = (\mathbf{z}, \mathbf{w})$ by considering four cases. In each case, we will show that

$$b(\vec{e}) \le B(p,d) := 2\binom{p}{d}\sum_{j=2}^{p} j^2. \tag{4.1}$$

For $t = 1,2,3,4$, let $b^t(\vec{e})$ denote the expected number of canonical paths of type t through $\vec{e}$; thus $b(\vec{e}) = \sum_{t=1}^{4} b^t(\vec{e})$. Likewise, we define $b^{3a}(\vec{e})$ to be the expected number of canonical paths of type 3 such that $\vec{e}$ belongs to part (a) of the path; $b^{3b}(\vec{e})$, $b^{4a}(\vec{e})$ and $b^{4b}(\vec{e})$ are defined similarly. As a preliminary, we prove that for *any* edge $\vec{e}$,

$$b^{3a}(\vec{e}) < p^2\binom{p}{d}, \tag{4.2}$$

$$b^{4a}(\vec{e}) < p^2\binom{p}{d}. \tag{4.3}$$

PROOF OF (4.2). For any $j = 1, \dots, p-1$, say that a vertex $\mathbf{x} \in \mathscr{S}$ is *j-feasible* if $x_d = p$ and, for all i, $x_i \ne j$. We will show

$$|\mathscr{S}_j(\vec{e})| \le p^2, \tag{4.4}$$

where $\mathscr{S}_j(\vec{e})$ is defined to be the set of j-feasible $\mathbf{x}$ such that the type 3(a) path from $\mathbf{x}$ to $(\mathbf{x} + j) - p$, call it $\gamma^{3a}(\mathbf{x}, (\mathbf{x}+j) - p)$, passes through $\vec{e}$. Granting (4.4) for the moment, it follows that

$$\begin{aligned}b^{3a}(\vec{e}) &= \sum_{(\mathbf{x},\mathbf{y}):\, x_d = p > y_d} (p-d)^{-1} \sum_{j:\, \forall i\, x_i \ne j} I_{\gamma^{3a}(\mathbf{x}, \mathbf{x}+j-p)}(\vec{e}) \\ &= \binom{p-1}{d} \times (p-d)^{-1} \sum_{j=1}^{p-1} |\mathscr{S}_j(\vec{e})| \\ &< \binom{p-1}{d} \times (p-d)^{-1} \times p^3 = p^2\binom{p}{d},\end{aligned}$$

which is (4.2). (4.3) is proved similarly.

To establish (4.4), fix j and suppose that i is the coordinate such that $w_i = z_i - 1$. Let $\mathbf{x} \in \mathscr{S}_j(\vec{e})$. Given the sites of the ith particle in both $\mathbf{x}$ and $\mathbf{x} + j - p$, one can determine $\mathbf{x}$ from $\vec{e}$. Indeed, it is not hard to verify that

$$\mathbf{x} = [(\mathbf{w} - w_i) + x_i + (\mathbf{x} + j - p)_i + p] - j.$$

Thus (4.4) is established. □

We are now prepared for the four-case proof of (4.1).

CASE A. $z_d = w_d = p$. We show that

$$b(\vec{e}) \le 2p^2\binom{p}{d} + \frac{p}{d}b(p-1, d-1)$$
$$\le 2p^2\binom{p}{d} + \frac{p}{d}B(p-1, d-1) = B(p, d).$$

First observe that any path through $\vec{e}$ is of type 1, 3 or 4.

Type 1: $b^1(\vec{e}) \le b(p-1, d-1)$.

Type 3: If a type 3 path passes through $\vec{e}$, then $\vec{e}$ belongs to part (a) of the path. Thus $b^3(\vec{e}) = b^{3a}(\vec{e}) < p^2\binom{p}{d}$.

Type 4: Part (a) of a type 4 path is handled by (4.3). As for part (b), observe first that for *any* edge $\vec{e}$, $db^{4b}(\vec{e})$ equals

$$\sum_{(\mathbf{x},\mathbf{y}):\, x_d < p = y_d} \; \sum_{k:\, x_i = k \text{ for some } i} P\{\text{type 1 path from } (\mathbf{x}+p) - k \text{ to } \mathbf{y} \text{ passes through } \vec{e}\}.$$

In this double sum, for each fixed $\mathbf{y}$ and any given $\mathbf{x}'$ with $x'_d = p$, the term

$$P\{\text{type 1 path from } \mathbf{x}' \text{ to } \mathbf{y}' \text{ passes through } \vec{e}\}$$

appears exactly $p - d$ times. Thus

$$b^{4b}(\vec{e}) = \frac{p-d}{d} \sum_{(\mathbf{x}',\mathbf{y}):\, x'_d = y_d = p} P\{\text{type 1 path from } \mathbf{x}' \text{ to } \mathbf{y} \text{ passes through } \vec{e}\}$$
$$= \frac{p-d}{d}b^1(\vec{e}) \le \frac{p-d}{d}b(p-1, d-1).$$

Combining this with (4.3),

$$b^4(\vec{e}) < p^2\binom{p}{d} + \frac{p-d}{d}b(p-1, d-1)$$

for any edge.

Adding our upper bounds for $b^1(\vec{e})$, $b^3(\vec{e})$ and $b^4(\vec{e})$, we obtain the asserted bound for $b(\vec{e})$ in Case A.

CASE B. $z_d \le p - 1$ and $w_d \le p - 1$. We show that

$$b(\vec{e}) \le 2p^2\binom{p}{d} + \frac{p}{p-d}b(p-1, d)$$
$$\le 2p^2\binom{p}{d} + \frac{p}{p-d}B(p-1, d) = B(p, d).$$

Any path through $\vec{e}$ is of type 2, 3 or 4.

Type 2: $b^2(\vec{e}) \le b(p-1, d)$.

Type 3: Just as we showed $b^{4b}(\vec{e}) \le ((p-d)/d)b(p-1, d-1)$, so one can show $b^{3b}(\vec{e}) \le (d/(p-d))b(p-1, d)$ for *any* edge $\vec{e}$. Combining this with (4.2),

$$b^3(\vec{e}) < p^2\binom{p}{d} + \frac{d}{p-d}b(p-1,d)$$

for any edge.

Type 4: If a type 4 path passes through $\vec{e}$, then $\vec{e}$ belongs to part (a) of the path. Thus $b^4(\vec{e}) = b^{4\mathrm{a}}(\vec{e}) < p^2\binom{p}{d}$.

Adding our upper bounds for $b^2(\vec{e})$, $b^3(\vec{e})$ and $b^4(\vec{e})$, we obtain the asserted bound for $b(\vec{e})$ in Case B.

CASE C. $z_d = p$ and $w_d \le p - 1$. Then $w_d = p - 1$, and any path passing through $\vec{e}$ is a type 3 path and does so in part (a). Therefore, by (4.2), $b(\vec{e}) = b^{3\mathrm{a}}(\vec{e}) < p^2\binom{p}{d} < B(p,d)$ in Case C.

CASE D. $z_d \le p - 1$ and $w_d = p$. As in Case C, $b(\vec{e}) < p^2\binom{p}{d} < B(p,d)$. □

4.2. *The exclusion process on a discrete circle.* In this subsection we use our knowledge (Proposition 4.1) about the exclusion process on $\{1,\dots,p\}$ to study the labelled exclusion process on $\mathbf{Z}_p$, as defined in Section 1.1. The symmetrized (i.e., reversiblized) chain is the labelled exclusion process corresponding to simple symmetric random walk on $\mathbf{Z}_p$; denote the generator of this symmetrized chain by G and the second smallest eigenvalue of $-G$ by λ. Here again $g := G(\mathbf{x},\mathbf{y}) = (2d)^{-1}$ is constant when $\mathbf{x} \ne \mathbf{y}$ and $G(\mathbf{x},\mathbf{y}) > 0$, so we will again be able to use the bound $\lambda \ge g|\mathscr{S}|/(\gamma_* b)$. Our paths for the symmetric circular process will be built in a straightforward fashion from those for the linear process. We shall establish the following results about λ. We omit a detailed proof of (a)—use $\phi(x) = \sum_{i=1}^d x_i$ as for the linear process.

PROPOSITION 4.4. *Let λ be the spectral gap for the d-particle labelled exclusion process for symmetric random walk on $\mathbf{Z}_p$ with $2 \le d \le p$. Set*

$$l_1(p,d) := \frac{\binom{p}{d}}{2\left[(2d-1)p - d^2\right]\left[d^2 p\binom{p}{d} + 2^{d-2}p^{d+1}\right]}$$

and

$$l_2(p,d) := \frac{3}{d\left[(p-2)(2p^2+p+3) + 3dp\right]\left[(p-2)(p+1) + 2d(p-1)\right]}.$$

(a) *Upper bound on λ*: $\lambda \le 6/(d(p^2-1)) = (1+o(1))(6/(dp^2))$.

(b) *Lower bound on λ for d fixed*: $\lambda \ge l_1(p,d) \ge (1+o(1))(d!2^d d + 4d^3)^{-1}p^{-2}$.

(c) *Lower bound on λ for d growing with p*:

$$\lambda \ge l_2(p,d) = (1+o(1))\frac{3/2}{dp^4(p+2d)}.$$

REMARK 4.5. (a) Suppose d grows linearly with p. Then from (c) it follows that order p^7 times units are sufficient for near-stationarity for any initial distribution; hence *the clockwise exclusion process is rapidly mixing*.

(b) How much time is necessary? We have not investigated this question carefully but shall make a few heuristic remarks. For simplicity, consider the case $d = p/2$ (p even).

First suppose that the initial configuration is deterministic. In the absence of interactions, it would take time of order p^3 just for particle 1 to become nearly random. We surmise that the exclusion interactions tend only to slow the convergence to stationarity; for a theorem of Arratia (1983) along these lines but in a somewhat different setting, see Theorem VIII.4.13 in Liggett (1985). It then follows that time at least of order p^3 is necessary for near-stationarity of the exclusion process.

Now suppose instead, for reasons discussed in Section 5, that the initial distribution is uniform over the d configurations (one for each possible position of particle 1) for which all the particles are at odd-numbered sites. In stationarity there is substantial probability of the event {at least $d/2 + cd^{1/2}$ particles among the sites $0, \dots, d-1$} when c is constant (and not too large). For the exclusion process to realize this event at time t, the particle initially at site $1 + p - 2cd^{1/2}$ must have already moved clockwise into (and perhaps out of) the set $\{0, \dots, d-1\}$ of sites. But this, we guess (again by comparison with the no-exclusion process), takes time at least of order $p^{3/2}$. Then in the present case time at least of order $p^{3/2}$ (and perhaps much greater) is necessary for the exclusion process to become nearly stationary.

CANONICAL PATHS. We now proceed to define the canonical paths for the labelled exclusion process corresponding to simple symmetric random walk on $\mathbf{Z}_p$. Here $\mathscr{S}$ consists of all clockwise rotations of strictly increasing d-tuples from $\{0, \dots, p-1\}$. We begin by specifying a canonical path between an ordered pair of configurations $\mathbf{x}, \mathbf{y}$ differing only by a rotation, say, $y_i = x_i + z$ for each i. Put $j = y_d$.

1. If $x_d = j$, then $\mathbf{y} = \mathbf{x}$ and the path is empty. Otherwise, let k be the first site clockwise from j such that site k is occupied by $\mathbf{x}$ but site $k+1$ is not. Call the particle at site k the *lead particle* and go on to step 2.
2. Perform a *unit* (*clockwise*) *rotation*, that is, move each particle clockwise one site as follows. First move the lead particle and then move the next particle reached in the counterclockwise direction. Continue in this fashion until the unit rotation is complete. [For example, if $p = 8$, $d = 4$, $j = 0$ and $\mathbf{x} = (6, 1, 2, 4)$, then rotate to $(7, 2, 3, 5)$ by first moving the lead particle 3 from site 2 to site 3, then particle 2 from site 1 to site 2, then particle 1 from site 6 to site 7 and finally particle 4 from site 4 to site 5.]
3. Repeat step 2 until the dth particle is at site j. [With j and $\mathbf{x}$ as in the previous example, $\mathbf{y} = (2, 5, 6, 0)$ is obtained from $\mathbf{x}$ by $z = 4$ unit rotations, each led by particle 3.]

For future reference, we observe that no element of $\mathscr{S}$ is the terminal vertex in too many paths of the above form. Let $j \in \mathbf{Z}_p$ and put $\mathscr{S}_j :=$

$\{\mathbf{y} \in \mathscr{S}: y_d = j\}$. For $\mathbf{y} \in \mathscr{S}_j$, define

$$A_j(\mathbf{y}) := \{\mathbf{x} \in \mathscr{S}: \mathbf{x} \text{ and } \mathbf{y} \text{ differ by a rotation}\}.$$

[Note that the sets $A_j(\mathbf{y})$, $\mathbf{y} \in \mathscr{S}_j$, partition $\mathscr{S}$.] Clearly,

$$|A_j(\mathbf{y})| = p.$$

Next we define a canonical path between an ordered pair $\mathbf{x}, \mathbf{y}$ of elements of $\mathscr{S}$ having $x_d = y_d = j$ (say). This is easy to describe. Consider the underlying graph for the labelled symmetric circular exclusion process. The subgraph induced by $\mathscr{S}_j$ is clearly isomorphic to the underlying graph of the linear $\binom{p-1}{d-1}$-exclusion process. Choose canonical paths isomorphic to those constructed in Section 4.1. (As before, we use entirely different schemes according as d grows with p or not.)

The canonical path between any pair $\mathbf{x}, \mathbf{y}$ of vertices in $\mathscr{S}$ is now defined as the concatenation of two segments as follows. Let $j = y_d$, $z = y_d - x_d$, and $\mathbf{x}^j = (x_1 + z, \dots, x_d + z = j)$:

initial segment: follow the canonical (rotational) path from $\mathbf{x}$ to $\mathbf{x}^j$;

main segment: follow the canonical path from $\mathbf{x}^j$ to $\mathbf{y}$.

REMARK 4.6. Technically, we have violated the conditions leading to the Poincaré inequality, since it is not hard to see that our "paths" can have repeated edges. However, the inequality $\lambda \geq g|\mathscr{S}|/(\gamma_* b)$ remains valid provided that in calculating b of (2.24) we count c canonical paths through $\vec{e}$ for each canonical path containing the edge $\vec{e}$ c times. Since the initial and main segments are (separately) free of repeated edges, the proof of the next lemma remains correct for this modified definition of b.

BOUNDS ON γ_* AND b. The next lemma gives bounds on γ_* and b for the labelled symmetric circular exclusion process in terms of bounds for the linear process. When the lemma is combined with the bounds from Section 4.1 (e.g., Lemma 4.3) and the Poincaré bound $\lambda \geq g|\mathscr{S}|/(\gamma_* b)$, we obtain parts (b) and (c) of Proposition 4.4.

LEMMA 4.7. *Let $\gamma_*(p, d)$ be the greatest possible length of a canonical path and $b(p, d)$ the largest expected number of canonical paths passing through any oriented edge for the d-particle labelled symmetric circular exclusion process with state space $\mathscr{S}$. Let $\gamma_*^0(p-1, d-1)$ and $b^0(p-1, d-1)$ denote the corresponding quantities for the linear $\binom{p-1}{d-1}$-exclusion process. Then*

$$\gamma_*(p, d) \leq d(p-1) + \gamma_*^0(p-1, d-1)$$

and

$$b(p, d) \leq dp|\mathscr{S}| + pb^0(p-1, d-1).$$

PROOF. Since the maximum length of a rotational canonical path is $d(p-1)$, the assertion about γ_* is clear. Next we bound the expected number $b(\vec{e})$ of canonical paths through any edge $\vec{e} = (\mathbf{z}, \mathbf{w})$. Suppose first that the canonical path from $\mathbf{x}$ to $\mathbf{y}$ passes through $\vec{e}$ in its initial segment. Once x_d and the lead particle in the rotation are identified, $\mathbf{x}$ can easily be reconstructed from $\vec{e}$. Hence the (deterministic) number of such paths does not exceed $dp|\mathscr{S}|$. Finally we consider main segments. For each j, the expected number of canonical paths through $\vec{e}$ between ordered pairs of vertices in $\mathscr{S}_j$ does not exceed $b^0(d-1, p-1)$. Since $|A_j(\mathbf{x}^j)| = p$ for any $\mathbf{x} \in \mathscr{S}$, we see that the expected number of canonical paths from $\mathscr{S}$ to $\mathscr{S}_j$ through $\vec{e}$ does not exceed $pb^0(d-1, p-1)$. But if the main segment of a path from $\mathbf{x} \in \mathscr{S}$ to $\mathbf{y} \in \mathscr{S}_j$ passes through $\vec{e}$, then we must have $j = z_d$ $(= w_d)$. The lemma is proved. □

5. Poisson blockers. In this final section we briefly describe how the dense $(p = 2d)$ clockwise exclusion process arises naturally in connection with a Poisson blockers problem from statistical mechanics, as related to the author by P. Doyle and the physicist D. Huse via L. A. Shepp. We show how our variation bound (1.2) yields estimates for a shortest path problem of natural interest.

To formally describe the blockers problem, associate independent unit rate Poisson processes $(N_j(t))$ with the sites $j \in \mathbf{Z}_p$ (p even). Let δ_{jk}, $k = 1, 2, \ldots,$ denote the event times of N_j. For each $t \geq 0$, set

$$\varepsilon_i := \begin{cases} 1, & \text{if } i \text{ is odd,} \\ 0, & \text{if } i \text{ is even} \end{cases}$$

and define

$$\nu_j(t) := \min\{l + \varepsilon_i : i \in \mathbf{Z}_p,\, l \in L_{ij}(t)\}. \tag{5.1}$$

Here $L_{ij}(t)$ is the random set of nonnegative integers l having the property that there exist (1) a sequence $j_0, j_1, \ldots, j_l$ of sites in $\mathbf{Z}_p$ with (a) $j_0 = i$, (b) $j_l = j$ and (c) each pair j_h, j_{h+1} consisting of neighboring sites; and (2) a corresponding increasing sequence $t_0, t_1, \ldots, t_l, t_{l+1}$ of times with (a) $t_0 = 0$ and (b) $t_{l+1} = t$; such that no time interval $[t_h, t_{h+1}]$ contains any $\delta_{j_h k}$.

In less formal terms, imagine the sites of $\mathbf{Z}_p$ arranged in a circle in the plane, with synchronized time axes protruding upward from each. In this space-time diagram, mark the jth time axis with a "blocker" δ each time an event from N_j occurs. A *path* from site i at time 0 to site j at time t moves upward along the various time axes, with a sideways switch from a time axis to either of its neighboring axes allowed at any time. A path, however, is forbidden to pass through a δ. Define the *length* l of such a path to be the number of switches and its *cost* to be l if i is even and $l + 1$ if i is odd. $L_{ij}(t)$ is then the set of lengths of all paths from $(i, 0)$ to (j, t) and $\nu_j(t)$ is the minimum cost of any path from $\mathbf{Z}_p$ at time 0 to site j at time t.

Observe that the nondecreasing function $\nu_j(t)$ of t always has the same parity as j and in fact that $\nu_{j+1}(t) = \nu_j(t) \pm 1$. If ν_j increases at time t, then of course $t = \delta_{jk}$ for some k. Conversely, if $t = \delta_{jk}$ for some k, then (with probability 1) no other $\nu_{j'}$ increases at time t, and ν_j increases (by 2) if and only if $\nu_j(t-) = \nu_{j-1}(t-) - 1 = \nu_{j+1}(t-) - 1$.

To relate this to the dense clockwise exclusion process, declare that a particle occupies site j at time t if and only if $\nu_j(t) = \nu_{j+1}(t) + 1$. We have seen that when ν_j increases at time t, the relations $\nu_{j-1} = \nu_j + 1$ and $\nu_j = \nu_{j+1} - 1$ are changed to the relations $\nu_{j-1} = \nu_j - 1$ and $\nu_j = \nu_{j+1} + 1$; thus a particle moves from site $j - 1$ to site j. The particle configurations therefore form a clockwise exclusion process (with exponential rate 1, rather than $1/d$, at each site). There are $d = p/2$ particles, which are initially situated at the odd vertices of $\mathbf{Z}_p$.

Turning our viewpoint around, the exclusion process we have generated has the graphical representation described in Remark 1.2 with the *same* unit rate Poisson processes N_j as here and $(\nu_j(t) - \varepsilon_j)/2$ simply counts the number of transition times through epoch t at which the vacant site j again becomes occupied.

Of natural interest is the marginal average behavior of one of the processes ν_j, say, $\nu := \nu_0$. Suppose that the exclusion process is in exact stationarity at the stopping time T; we claim that

$$E[\nu(T+t) - \nu(T)] = \frac{p}{p-1} \times \frac{1}{2}t, \qquad t \geq 0. \tag{5.2}$$

Indeed, the following argument can be made precise. At any epoch $u > T$, the probability that site $p - 1$ is occupied and site 0 is vacant is

$$\binom{p-2}{d-1} \Big/ \binom{p}{d} = \frac{p}{p-1} \times \frac{1}{4}.$$

Conditionally given this event, the probability of a jump from site $p - 1$ to site 0 in the time interval $(u, u + du]$ is du. Hence the expected number of reoccupations of site 0 in $(T, T + t]$, namely $\frac{1}{2} \times \text{LHS}(5.2)$, equals $p/(p-1) \times \frac{1}{4}t$.

Now we consider the exclusion process X started with an arbitrary distribution for the configuration of the $d = p/2$ particles. Let T be a maximal (i.e., fastest) time to coupling for the given chain X and the stationary chain Y. The existence of such a time is guaranteed by Griffeath (1975) [for further discussion, see Diaconis (1988), Chapter 4E]. Using the bound

$$P\{T > t\} \leq \left(\frac{p}{8}\right)^{1/2} \exp\left[-\frac{1}{2}(6tp^{-8} - \log 2)p\right]$$

of (1.2) (bear in mind that we have slowed down time by a factor of $d = p/2$ compared to Theorem 1.1) the following result can be derived. We omit the proof.

Proposition 5.1. *There exist universal constants $a > 0$ and $A > 0$ so that for all even p and all $t \geq 3$ the exclusion process X satisfies*

$$-ap^8 \leq E\nu(t) - \frac{p}{p-1} \times \frac{1}{2}t \leq Ap^8 \frac{\log t}{\log\log t}.$$

That is, $E\nu(t)/t$ converges as $t \to \infty$ to its stationary value $p/(p-1) \times \frac{1}{2}$ for each fixed p, with time of order $t = p^8 \log p/\log\log p$ sufficient for large p

to make the discrepancy

$$\left| \frac{E\nu(t)}{t} - \frac{p}{p-1} \times \frac{1}{2} \right|$$

small.

REMARK 5.2. For the rightward exclusion process on **Z** started with particles at the odd sites, Huse conjectures

$$E\nu(t) = \tfrac{1}{2}t + (1 + o(1))Ct^{1/3} \quad \text{as } t \to \infty$$

for some constant C.

Acknowledgments. The author thanks Beth Broman, Joe Chang, Persi Diaconis and Larry Shepp for helpful discussions, and Milena Mihail for providing a first draft of her Ph.D. dissertation.

REFERENCES

ALDOUS, D. and DIACONIS, P. (1986). Shuffling cards and stopping times. *Amer. Math. Monthly* **93** 333–348.

ARRATIA, R. (1983). The motion of a tagged particle in the simple symmetric exclusion system on Z. *Ann. Probab.* **11** 362–373.

CHANG, J. and FILL, J. A. (1990). Rates of convergence in renewal theory. Unpublished manuscript.

CHEEGER, J. (1970). A lower bound for the smallest eigenvalue of the Laplacian. In *Problems in Analysis* (R. C. Gunning, ed.) 195–199. Princeton Univ. Press.

DIACONIS, P. (1988). *Group Representations in Probability and Statistics.* IMS, Hayward, Calif.

DIACONIS, P. and FILL, J. A. (1990a). Strong stationary times via a new form of duality. *Ann. Probab.* **18** 1483–1522.

DIACONIS, P. and FILL, J. A. (1990b). Examples for the theory of strong stationary duality with countable state spaces. *Probab. Engrg. Inform. Sci.* **4** 157–180.

DIACONIS, P. and STROOCK. D. (1991). Geometric bounds for eigenvalues of Markov chains. *Ann. Appl. Probab.* **1** 36–61.

FAN, K. and HOFFMAN, A. (1955). Some metric inequalities in the space of matrices. *Proc. Amer. Math. Soc.* **6** 111–116.

FILL, J. A. (1990a). Strong stationary duality for continuous-time Markov chains. Unpublished manuscript.

FILL, J. A. (1990b). Strong stationary duality for partially ordered state spaces. Unpublished manuscript.

GRIFFEATH, D. (1975). A maximal coupling for Markov chains. *Z. Wahrsch. Verw. Gebiete* **31** 95–100.

LAWLER, G. F. and SOKAL, A. D. (1988). Bounds on the L^2 spectrum for Markov chains and Markov processes: A generalization of Cheeger's inequality. *Trans. Amer. Math. Soc.* **309** 557–580.

LIGGETT, T. M. (1985). *Interacting Particle Systems.* Springer, New York.

MARSHALL, A. W. and OLKIN, I. (1979). *Inequalities: Theory of Majorization and Its Applications.* Academic, New York.

MIHAIL, M. (1989). Combinatorial aspects of expanders. Ph.D. dissertation, Dept. Computer Science, Harvard Univ.

SINCLAIR, A. and JERRUM, M. (1989). Approximate counting, uniform generation and rapidly mixing Markov chains. *Inform. and Comput.* **82** 93–133.

SPITZER, F. (1970). Interaction of Markov processes. *Adv. in Math.* **5** 246–290.

DEPARTMENT OF MATHEMATICAL SCIENCES
JOHNS HOPKINS UNIVERSITY
BALTIMORE, MARYLAND 21218