Journal of Theoretical Probability, Vol. 16, No. 3, July 2003 (© 2003)

Shuffling Chromosomes

Rick Durrett[1]

Received March 18, 2002

The gene order of chromosomes can be rearranged by chromosomal inversions that reverse the order of segments. Motivated by a comparative study of two *Drosophila* species, we investigate the number of reversals that are needed to scramble the gene order when all reversals are equally likely and when the segments reversed are never more than L genes. In studying this question we prove some new results about the convergence to equilibrium of shuffling by transposition and the one dimensional simple exclusion process.

KEY WORDS: n-Reversal chain; exclusion process.

1. INTRODUCTION

To explain our motivation we begin with an example. Ranz *et al.*[7] did a comparative study of chromosome 2 of *Drosophila repleta* and chromosome arm 3R of *D. melanogaster*. If we number the 26 genes that they studied according to their order on the *D. repleta* chromosome then their order on *D. melanogaster* is given by

12 7 4 *2 3 21 20* 18 1 13 9 16 6 14 *26 25 24* 15 *10 11* 8 5 *23 22* 19 17

where we have used italics to indicate adjacencies that have been preserved in time. Since the divergence of these two species, this chromosome region has been subjected to many inversions that reverse a segment of the chromosome. Our two questions are: How many such reversals have occurred? Is the data consistent with the null model which supposes that all possible inversion have the same probability? To answer these questions we need to formulate and analyze some models.

[1] Department of Mathematics, Cornell University, Ithaca, New York 14853. E-mail: rtd1@cornell.edu

0894-9840/03/0700-0725/0 © 2003 Plenum Publishing Corporation

***n*-Reversal Chain.** Consider n markers on a chromosome, which we label with $1, 2, \ldots, n$, and that can be in any of the $n!$ possible orders. To these markers we add two others: one called 0 at the beginning and one called $n+1$ at the end. For convenience of description we connect adjacent markers by edges. For example, when $n=7$ the state of the chromosome might be

$$0-5-3-4-1-7-2-6-8$$

In our biological application the probability of an inversion in a given generation is small so, in contrast to the usual card shuffling problems, we will formulate the dynamics in continuous time. The labels 0 and $n+1$ never move. To shuffle the others, at times of a rate one Poisson process we pick two of the $n+1$ edges at random and invert the order of the markers in between. For example, if we pick the edges $5-3$ and $7-2$ the result is

$$0-5-7-1-4-3-2-6-8$$

If we pick $3-4$ and $4-1$ in the first arrangement there is no visible change. However, allowing this move will simplify the mathematical analysis and only amounts to a small time change of the dynamics in which one picks two markers $1 \leqslant i < j \leqslant n$ at random and reverses the segment with those endpoints.

It is clear that if the chromosome is shuffled repeatedly then all of the $n!$ orders for the interior markers will have equal probability. The basic question is how long does it take for the marker order to be randomized.

Theorem 1. Consider the state of the system at time $t = cn \ln n$ starting with all markers in order. If $c < 1/2$ then the total variation distance to the uniform distribution v goes to 1 as $n \to \infty$. If $c > 2$ then the distance goes to 0.

Lower Bound. To prove the first half of the result, we define an edge to be *conserved* if the markers at its two ends differ by exactly 1. It is easy to see that the expected number of conserved edges in equilibrium is 2. Suppose now that $t(\epsilon) = (1-\epsilon)\frac{n+1}{2}\ln(n+1)$. We say that an edge is *undisturbed* if it has not been involved in a reversal before time t. Let U be the total number of undisturbed edges at time $t(\epsilon)$. A simple computation shows that $EU = (n+1)^{\epsilon}$ and $\mathrm{Var}(U)/EU \to 1$ as $n \to \infty$. Letting A_ϵ be the event that there are at most $EU/2$ conserved edges and using Chebyshev's inequality it follows that for large n the total variation distance

$$\|p_{t(\epsilon)} - v\|_{TV} \geqslant |p_{t(\epsilon)}(A_\epsilon) - v(A_\epsilon)| \geqslant 1 - 9/EU \qquad \text{if } n \text{ is large.}$$

Upper Bound. To prove a result in the other direction, we will use the comparison techniques of Diaconis and Saloff-Coste.[2] We would like to thank Robin Pemantle for pointing out this argument. To set up for using their results, we let G be the group of permutations of $1,..., n$, which we think of as functions η from $\{1, 2,..., n\}$ onto $\{1, 2,..., n\}$. For $i<j$ let τ_{ij} be the transposition that exchanges i and j: $\tau_{i,j}(i)=j$, $\tau_{i,j}(j)=i$, and $\tau_{i,j}(k)=k$ otherwise. Again for $i<j$ let $\rho_{i,j}$ be the reversal that has $\rho_{i,j}(k)=j-(k-i)$ when $i\leqslant k\leqslant j$ and $\rho_{i,j}(k)=k$ otherwise.

The sets $\tilde{E}=\{\tau_{i,j}: i<j\}$ and $E=\{\rho_{i,j}: i<j\}$ are symmetric and generate G. Let $\tilde{q}$ be the uniform distributions on $\tilde{E}$. Let q be the measure that assigns mass $2/n(n+1)$ to each element of E and mass $2/(n+1)$ to the identity permutation, *id*. Introduce the Dirichlet forms defined by

$$\mathscr{E}(f, f)=\tfrac{1}{2}\sum_{x, y\in G}(f(x)-f(xy))^2\, q(y)$$

and $\tilde{\mathscr{E}}$ with q replaced by $\tilde{q}$. Using Theorem 1 on p. 2138 of Diaconis and Saloff-Coste,[2] we can show

$$\tilde{\mathscr{E}}(f, f)\leqslant A\mathscr{E}(f, f) \qquad \text{where} \quad A=4(n+1)/(n-1)$$

This result gives a comparison between L^2 norms: $\|p_t-\nu\|_2^2\leqslant\|\tilde{p}_t-\nu\|_2^2$. To transfer this result to the total variation norm we note that

$$2\,\|p_t-\nu\|_{TV}=\|p_t-\nu\|_1\leqslant(n!)^{1/2}\,\|p_t-\nu\|_2\equiv d_2(n)$$

and to quote Diaconis and Saloff-Coste:[2] all of their bounds as well as those in Diaconis[1] are bounds on $d_2(n)$.

Estimation. There are 6 conserved segments in our *Drosophila* data set. This means that at least $27-6=21$ edges have been disturbed, so at least 11 reversals have occurred. This lower bound is not sharp. In this example it can be shown that at least 14 reversals are needed to put the markers in order. Since undisturbed segments are necessarily conserved, we could use our computations for the lower bound to get a biased estimate of the number of reversals that have occurred. A better idea, which removes the bias, is to consider $\phi(\eta)=$ the number of conserved edges minus 2, and to check that ϕ is an eigenfunction of the chain with eigenvalue $(n-1)/(n+1)$. In our case $n=26$ and $\phi=4$ so solving

$$25\left(\frac{25}{27}\right)^m=4 \qquad \text{gives} \quad m=\frac{\ln(4/25)}{\ln(25/27)}=23.8$$

Ranz *et al.*[6] have recently enriched the comparative map so that 79 markers can be located in both species. Again numbering the markers on the *D. repleta* chromosome by their order on *D. melanogaster* we have:

36	*37*	17	40	*16*	*15*	*14*	63	*10*	*9*	55	28
13	51	22	79	39	70	66	*5*	*6*	*7*	35	64
33	*32*	*60*	*61*	18	65	62	12	1	11	23	20
4	52	68	29	48	3	21	53	8	43	72	*58*
57	*56*	19	49	34	59	30	77	31	67	44	2
27	38	50	*26*	*25*	76	69	41	24	75	71	78
73	47	54	45	74	42	46					

The number of conserved segment (again indicated with italics) is 11 so our moment estimate is

$$m = \frac{\ln(9/78)}{\ln(78/80)} = 85.3$$

This is comparable to one of the estimates of Ranz *et al.*[6] They used data on markers in four regions of chromosome arm 3R that had been mapped in detail to argue that there were approximately 6.32 ± 1.03 breakpoints per megabase and to calculate there were 177.07 ± 28.88 breakpoints in chromosome arm 3R, which would require 89 ± 14 reversals.

The number of reversals is surprising not only because it is large but also because there is still a strong correlation between the marker order in the two genomes. Spearman's rank correlation $\rho = 0.326$ which is significant at the $p = 0.001$ level. From the point of view of Theorem 1 this is not surprising: our lower bound on the mixing time predicts that $39.5 \ln 75 = 173$ reversals are needed to completely randomize the data. However, as Fig. 1 shows the rank correlation is randomized well before that time. In 10,000 runs the average rank correlation is only 0.0423 after 40 shuffles and only 4.3% of the runs had a rank correlation larger than 0.325. It is comforting to note that (conserved segments-2)/77 is almost a perfect exponential curve that ends with a value of $0.0777 \approx 0.0795 = (78/80)^{100}$.

One explanation for the results in the previous paragraph is that all chromosomal inversions may not be equally likely. To seek a biological explanation of the non-uniformity we note that the gene-to-gene pairing of homologous chromosomes implies that if one chromosome of the pair contains an inversion that the other does not, a loop will form in the region

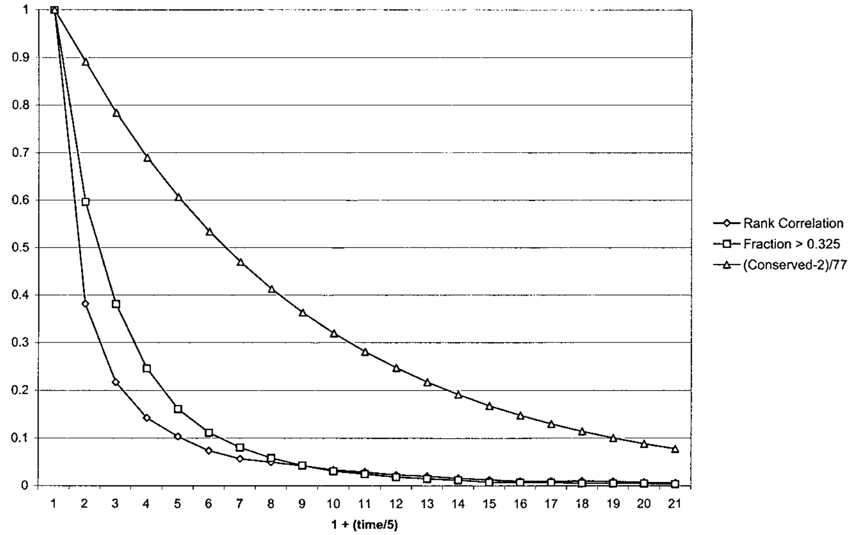

Fig. 1. *n*-Reversal chain.

in which the gene order is inverted. (See, e.g, p. 367 of Hartl and Jones.[5]) If a recombination occurs in the inverted region then the recombined chromosomes will contain two copies of some regions and zero of others, which can have unpleasant consequences. A simple way to take this into account is

***θ*-Reversal Model.** Inversions that reverse markers i to $i+j$ occur at rate

$$\theta^{j-1}(1-\theta)/n$$

The reasoning here is that the probability of no recombination decreases exponentially with the length of the segment reversed.

This model seems quite complicated to analyze, so we will make two simplifications. The first is that we will consider the markers 1, 2,..., n as lying on a circle and connect them by edges as before. For example, when $n=7$ and the markers start in order, we have

$$1-2-3-4-5-6-7-1$$

Departing slightly from our previous approach, we pick the location of the left marker of the segment to be inverted, i, uniformly over the set of possibilities, and then pick the location of the right marker to be $i+j$ with probability p_j, where the arithmetic is done modulo n. Some of our results will be obtained for this **p-reversal model**, but in most cases we will restrict our attention to the

L-Reversal Model. $p_j = 1/L$ for $1 \leqslant j \leqslant L$.

Our first result is an easy extension of the lower bound in Theorem 1.

Theorem 2. The amount of time for the p-reversal chain to reach equilibrium is at least $\frac{n}{2} \ln n$.

To prove a second lower bound we will use an idea of Wilson.[9] We would like to thank David Aldous and Laurent Saloff-Coste for telling us about his work. The first step in the analysis to note that a single marker performs a symmetric random walk on the circle. To compute the jump distribution we note that the marker at n will be moved to i if the left endpoint of the inversion is at $n-k$ and the right is at $i+k$ where $k \geqslant 0$ and $n-k > i+k$. Summing we have the rate for jumps by $+i$ is q_i/n where

$$q_i = \sum_{k \geqslant 0} p_{i+2k}$$

The condition $n-k > i+k$ does not appear in the sum since we suppose $p_m = 0$ for $m \geqslant n$. Symmetry implies that $q_{-i} = q_i$. When p_i is uniform on $1, 2,..., L$ this has almost a triangular distribution:

$$q_i = (1 + [(L-i)/2])/L \qquad \text{for} \quad 1 \leqslant i \leqslant L$$

where $[z]$ denotes the integer part of z. When p is geometric and $n(1-\theta)$ is large so we can ignore truncation of the infinite series

$$q_i = \theta^{i-1}(1-\theta)/(1-\theta^2)$$

Theorem 3. If the distribution p_i is fixed and $n \to \infty$ then convergence to equilibrium takes at least time

$$\frac{1}{2} \cdot \frac{n^3}{4\pi^2 \sum_i i^2 q_i} \ln n$$

In the L-shuffle if $L \to \infty$ and $(\ln L)/(\ln n) \to a \in [0, 1)$ then convergence to equilibrium takes at least time

$$\frac{(1-a)}{2} \cdot \frac{6}{\pi^2} \cdot \frac{n^3}{L^3} \ln n$$

Here and in Theorems 4 and 5 "takes time at least $cn^3 \ln n$" means that for any $\epsilon > 0$ the total variation distance at time $(c-\epsilon)\, n^3 \ln n$ tends to 1 as $n \to \infty$.

The key to the proof is the fact that $f(x) = \sin(2\pi x/n)$ is an eigenfunction for any symmetric random walk on the circle, which for the single particle walk has eigenvalue

$$-\lambda \equiv \sum_{i=1}^{n} \frac{2q_i}{n} [\cos(2\pi i/n) - 1] \sim -\frac{4\pi^2}{n^3} \sum_{i=1}^{n} i^2 q_i \qquad \text{as} \quad n \to \infty$$

Let $\eta_t(i)$ be the marker at position i at time t, and $X_t^j = \eta_t^{-1}(j)$ be the location of marker j at time t. Let

$$g(m) = \operatorname{sgn}\left(\frac{n+1}{2} - m\right)$$

and let

$$\Phi(\eta_t) = \sum_i g(\eta_t(i)) \sin(2\pi/n) = \sum_i g(j) \sin(\pi X_t^j/n)$$

Letting $\Phi_t = \Phi(\eta_t)$, it follows from the calculation above that

$$\frac{d}{dt} E\Phi_t = -\lambda E\Phi_t$$

When the markers start in order $\Phi_0 \approx 2n/\pi$. Since in equilibrium $\operatorname{Var}(\Phi_\infty) = O(n)$, this suggests that the chain cannot be in equilibrium until $E\Phi_t = O(\sqrt{n})$ which takes about $1/2\lambda$ units of time. To complete this outline we have to show that $\operatorname{Var}(\Phi_t) \leqslant Cn$. We get a result that is worse than for the L-reversal since we can only show that $\operatorname{Var}(\Phi_t) \leqslant CnL$.

To compare the first conclusion with Wilson's Theorem 4 we observe that when $1, 2, \ldots, n$ is an interval with reflecting boundary conditions then the first eigenfunction for the nearest neighbor random walk is $\cos(\pi(i-1/2)/n)$, so the walk on the circle equilibrates 4 times as fast. A second factor of 2 comes from the fact that Wilson's chain does nothing 1/2 of the time. As the next result suggests, the factor 1/2 in front of the first conclusion in Theorem 2 should not be there.

Theorem 4. Consider independent random walks on the circle that jump from x to $x+i$ at rate r_i and from x to $x-i$ at rate r_i. Then the time to reach equilibrium is asymptotically

$$\frac{n^3}{4\pi^2 \sum_i i^2 r_i} \ln n$$

Before turning to the task of getting upper bounds on the convergence time, we will take another look of the data in light of the developments above. Using Wilson's idea we can introduce a statistic for the shuffling a linear (not circular chromosome)

$$\sum_{i=1}^{79} g(\eta_t(i)) \cos(\pi(i-0.5)/n)$$

Here we have replaced the $\sin(2\pi i/n)$ by something that is an eigenfunction for the nearest neighbor random walk on $1, 2, \ldots, n$ with reflecting boundary conditions. Figure 2 shows the results of 10,000 simulations of the 23-reversal chain acting on 79 markers: plotting the logarithm of the rank correlation, Wilson's statistic and the number of conserved segments -2, all of which have been scaled to have maximum value 1. Note that even though we do not know if any of these quantities are eigenvectors in this case, the three curves are almost straight lines. The value at time 85 for the (conserved segments $-2)/77$ is 0.1224 which is a little larger than the value of 0.1168 for the data. The value of 23 was chosen so that the value of the rank correlation 0.312 matched the 0.326 of the data as closely as possible. A simulation of the 23-reversal chain gave a value of 0.355 at that time. Recalling that $(1+23)/2=12$ we see that in the simulated shuffle each event involves 12 markers on the average.

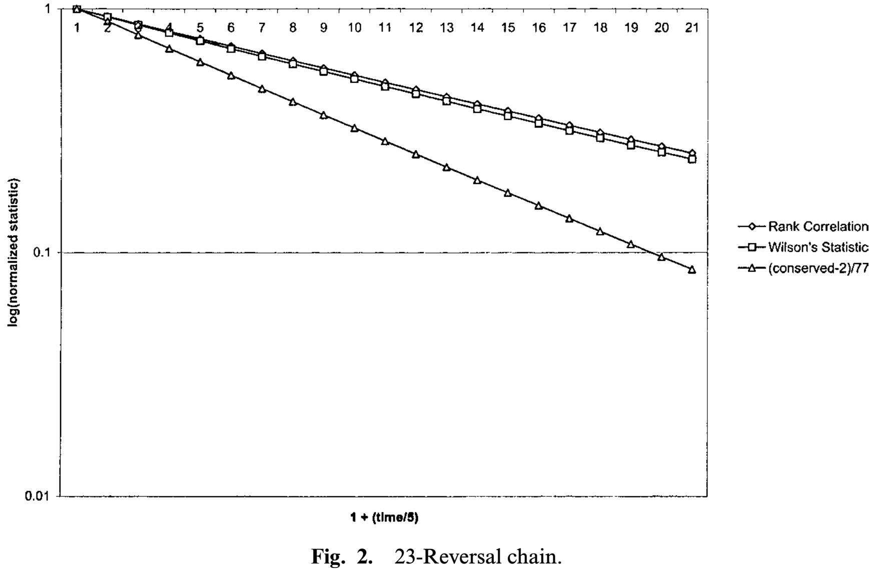

Fig. 2. 23-Reversal chain.

To get an upper bound on the time for the L-reversal to converge to equilibrium we will compare with

***p*-Transposition.** Pick an integer i uniformly on the circle then exchange the marker at i with the marker at $i+j$ with probability p_j.

***L*-Transposition.** $p_j = 1/L$ for $1 \leqslant i \leqslant L$.

Using the proof of Theorem 3 above one can show

Theorem 5. If the distribution in the p-transposition is fixed and $n \to \infty$ then convergence to equilibrium takes at least time

$$\frac{1}{2} \cdot \frac{n^3}{4\pi^2 \sum_i i^2 p_i} \ln n$$

In the L-transposition if $L \to \infty$ and $L/n \to 0$ then convergence to equilibrium takes at least time

$$\frac{1}{2} \cdot \frac{3}{4\pi^2} \cdot \frac{n^3}{L^2} \ln n$$

If $L/n \to a \in (0, 1]$ then convergence to equilibrium takes at least time

$$\frac{1}{2} \cdot \frac{na}{2 \int_0^a 1 - \cos 2\pi x \, dx} \ln n$$

Again the $1/2$'s should not be there. In support this we cite Theorem 4 and observe that when $a = 1$ the last lower bound is $(n/4) \ln n$. To gain some insight into the use of $g(m) = \operatorname{sgn}((m+1) - n/2)$ in Wilson's function and to give another application of Theorem 5, note that if we consider i with $\eta_t(i) \leqslant k$ to be occupied and others to be vacant, the p-transposition chain turns into the simple exclusion process starting with k particles. Theorem 5 gives rates of convergence to equilibrium that are uniform in k. In the opposite direction Wilson's argument uses the simple exclusion process with $k = [n/2]$ to bound the convergence of the p-transposition.

Using Theorem 1 on p. 2143 of Diaconis and Saloff-Coste[(2)] to compare the L-transposition chain to the n-transposition chain and the comparison of Dirichlet forms for the reversal and transposition chains introduced above, we can show

Theorem 6. The time required for the L-transposition to reach equilibrium is at most

$$\frac{2n^3}{(L+1)^2}\ln n$$

The time required for the L-reversal chain to reach equilibrium is at most

$$\frac{8n^3}{(L+1)^2}\ln n$$

Combined with Theorem 5 this shows that the convergence time for the L-transposition is of order $(n^3/L^2)\ln n$, a result that interpolates smoothly between the results of Diaconis–Shahshahani[3] and Wilson.[9] Comparing the second conclusion with Theorems 2 and 3 we see that the upper and lower bounds are not of the same order when $L = n^a$ with $0 < a < 1$. We believe that the lower bounds are the correct order of magnitude. In support of that guess we make another one.

Conjecture. The amount of time it takes for the two particle chain (X_t^1, X_t^2) to converge to equilibrium is $O(n \vee (n^3/L^2))$.

Why Is this True? The expected amount of time it takes for two adjacent particles to get separated is $n/2$. If we want to couple one of the markers to another independent one started in equilibrium this can be done in time $O(n^3/L^2)$. Run the two chains independently until they are within $L/2$. It is easy to see that this will occur before the difference chain exits from $(0, n)$ because if the particles are not within $L/2$ before the exit occurs, they will be after it does. When the two particles are within $L/2$ they can be forced to agree after their next jumps with positive probability. These two observations together suggest that one can couple the two particle chain to its equilibrium distribution in time $O(n \vee (n^3/L^2))$. However it is not easy to turn this idea into a proof.

The remainder of the paper is devoted to the proofs of Theorems 1 to 6, and is organized by the techniques being used: conserved edges, path comparisons, Wilson's method, and independent random walks. These four sections are independent and can be read in any order. We would like to thank Laurent Saloff-Coste for many helpful conversations while this paper was being written.

2. CONSERVED EDGES

Our first task is to show

Theorem 2.1. The amount of time for the n-reversal chain to reach equilibrium is at least $\frac{n+1}{2}\ln(n+1)$.

Lemma 2.1. The expected number of conserved edges in equilibrium is 2.

Proof. The first and last edges are conserved with probability $1/n$ each. The $n-1$ interior edges are conserved with probability $2/n$ each. □

Suppose now that $t(\epsilon)=(1-\epsilon)\frac{n+1}{2}\ln(n+1)$. We say that an edge is undisturbed if it has not been involved in a reversal before time t. Let $u_i=1$ if the ith edge is undisturbed, 0 otherwise, and let $U=u_1+\cdots+u_{n+1}$ be the total number of undisturbed edges.

Lemma 2.2. $EU=(n+1)^{\epsilon}$ and $\mathrm{Var}(U)/EU\to 1$ as $n\to\infty$.

Proof. Since edge i is disturbed at rate $2/(n+1)$

$$P(u_i=1)=\exp(-(2/n+1)\,t(\epsilon))=(n+1)^{-(1-\epsilon)}$$

and $EU=(n+1)\,P(u_i=1)=(n+1)^{\epsilon}$. To prove the second result we note that if $i\neq j$ the rate at which at least one of the edges is disturbed is $4/(n+1)-2/(n+1)\,n$ so

$$\begin{aligned}&P(u_i=1,u_j=1)-P(u_i=1)\,P(u_j=1)\\&\qquad=P(u_i=i)\,P(u_j=1)\left(\exp\left(\frac{2t(\epsilon)}{n(n+1)}\right)-1\right)\end{aligned}$$

Summing over i and j we have

$$\begin{aligned}\mathrm{Var}(U)&=nP(u_i=1)(1-P(u_i=1))\\&\qquad+P(u_i=1)^2\frac{n(n+1)}{2}\left(\exp\left(\frac{2t(\epsilon)}{n(n+1)}\right)-1\right)\end{aligned}$$

As $n\to\infty$, $P(u_i=1)\to 0$ so the first term $\sim nP(u_i=1)$. To see that the second term is smaller we note that $e^x-1\sim x$ so

$$\frac{(n+1)}{2}P(u_i=1)\left(\exp\left(\frac{2t(\epsilon)}{n(n+1)}\right)-1\right)\sim\frac{n^{\epsilon}}{2}\cdot\frac{(1-\epsilon)\ln(n+1)}{n}\to 0$$

and the desired result follows. □

Let p_t be the marker distribution at time t when the markers start in order, and let ν be the uniform distribution on the $n!$ orders. The next result is not very accurate since the proof uses Markov and Chebyshev's inequalities on quantities that should have approximate Poisson and normal distributions, but it does serve to establish a lower bound on the number of shuffles needed.

Lemma 2.3. If n is large then the total variation $\|p_{t(\epsilon)} - \nu\|_{TV} \geqslant 1 - 9/EU$.

Proof. Let A_ϵ be the set of configurations with at most $EU/2$ conserved edges. It follows from Lemma 2.2 and Markov's inequality that $(EU/2)\,\nu(A_\epsilon^c) \leqslant 2$ so $P(A_\epsilon^c) \leqslant 4/EU$. Using Lemma 2.2 with Chebyshev's inequality we have that for large n

$$(EU/2)^2\, p_{t(\epsilon)}(A_\epsilon) \leqslant \mathrm{Var}(U) \leqslant \tfrac{5}{4}\, EU$$

It follows that

$$\|p_t - \nu\|_{TV} = \sup_A\, |p_t(A) - \nu(A)| \geqslant |p_t(A_\epsilon) - \nu(A_\epsilon)| \geqslant 1 - \frac{9}{EU} \qquad \square$$

The proof of Theorem 2.1 generalizes easily to show

Theorem 2.2. The amount of time for the p-reversal chain to reach equilibrium is at least $\frac{n}{2} \ln n$.

Proof. As before we define an edge to be conserved if the markers at its two ends differ by exactly 1. Since each edge is conserved with probability $2/n$.

Lemma 2.4. The expected number of conserved edges in equilibrium is 2.

Let $t(\epsilon) = (1-\epsilon)(n/2) \ln n$. We say that an edge is undisturbed if it has not been involved in a reversal before time $t(\epsilon)$. Let $u_i = 1$ if the ith edge is undisturbed, 0 otherwise, and let $U = u_1 + \cdots + u_n$ be the total number of undisturbed edges.

Lemma 2.5. $EU = n^\epsilon$, $\mathrm{Var}(U)/EU \to 1$ as $n \to \infty$.

Proof. Since edge i is disturbed at rate $2/n$

$$P(u_i = 1) = \exp(-(2/n)\, t(\epsilon)) = n^{-(1-\epsilon)}$$

so $EU = nP(u_i = 1) = n^{\epsilon}$. To prove the second result we note that if $i \neq j$ the rate at which at least one of the edges is disturbed is $(2/n) - (p_{j-i} + p_{i-j})/n$ where the difference in the subscripts is done modulo n, so

$$\begin{aligned} &P(u_i = 1, u_j = 1) - P(u_i = 1)\ P(u_j = 1) \\ &\quad = P(u_i = i)\ P(u_j = 1) \left(\exp\left(\frac{2t(\epsilon)\ p_{j-i} + p_{j-i}}{n} \right) - 1 \right) \end{aligned}$$

Summing over i and j we have

$$\begin{aligned} \mathrm{Var}(U) &= nP(u_i = 1)(1 - P(u_i = 1)) \\ &\quad + P(u_1 = 1)^2\, n \sum_k \left(\exp\left(\frac{2t(\epsilon)\ p_k + p_{-k}}{n} \right) - 1 \right) \end{aligned}$$

As $n \to \infty$, $P(u_i = 1) \to 0$ so the first term $\sim nP(u_i = 1)$. To see that the second term is smaller we note

$$P(u_1 = 1) \sum_k \left(\exp\left(\frac{2t(\epsilon)\ p_k + p_{-k}}{n} \right) - 1 \right) \sim n^{-1+\epsilon} \cdot \frac{4t(\epsilon)}{n} \to 0$$

and the desired result follows. □

Repeating the proof of Lemma 2.3 shows that if n is large $\|p_{t(\epsilon)} - \nu\|_{TV} \geqslant 1 - 9/EU$ and the proof of Theorem 2.2 is complete. □

Returning to the n-reversal chain the final result in this section is.

Theorem 2.3. Let $\phi(\eta)$ be $-2+$ the number of conserved segments in the permutation η. ϕ is an eigenfunction with eigenvalue $(n-1)/(n+1)$.

Proof. For $1 \leqslant i \leqslant n-1$ let $\psi_i(\eta) = n-2$ if $i-(i+1)$ is conserved and $\psi_i(\eta) = -2$ if not. On one step the expected change in ψ_i

$$\text{if } i-(i+1) \text{ is conserved is } \frac{2(n-2)}{n(n+1)}(-n) = \left(\frac{-2}{n+1} \right)(n-2)$$

$$\text{if } i-(i+1) \text{ is not conserved is } \frac{4}{n(n+1)}(n) = \left(\frac{-2}{n+1} \right)(-2)$$

To check the first that in order to split up the markers one of the edges involved must be $i-(i+1)$ but the other one may not be $(i-1)-i$ or $(i+1)-(i+2)$. For the second we observe that exactly two reversals will bring the markers back together.

For $i=0$ and $i=n$ let $\psi_i(\eta)=n-1$ if $i-(i+1)$ is conserved and $\psi_i(\eta)=-1$ if not. On one step of the expected change in ψ_i

$$\text{if } i-(i+1) \text{ is conserved is } \frac{2(n-1)}{n(n+1)}(-n)=\left(\frac{-2}{n+1}\right)(n-1)$$

$$\text{if } i-(i+1) \text{ is not conserved is } \frac{2}{n(n+1)}(n)=\left(\frac{-2}{n+1}\right)(-1)$$

To check the first that in order to split up 0 and 1 one of the edges involved must be $0-1$ but the other one may not be $1-2$. For the second we observe that exactly one reversals will bring 1 back next to 0. The result now follows from the fact that $\phi(\eta)=\frac{1}{n}\sum_i \psi_i(\eta)$. □

3. PATH COMPARISONS

Here we prove the upper bounds in Theorems 1 and 6.

Proof of the Upper Bound in Theorem 1. For $i<j$ let τ_{ij} be the transposition that exchanges the markers at i and j let $\rho_{i,j}$ reverse the order of markers $i, i+1,..., j$. Let $\tilde{q}$ be the uniform distributions on $\tilde{E}=\{\tau_{i,j}: i<j\}$. Let q be the measure that assigns mass $2/n(n+1)$ to each element of $E=\{\rho_{i,j}: i<j\}$ and mass $2/(n+1)$ to the identity permutation, *id*. Define the Dirichlet forms by

$$\mathscr{E}(f,f)=\tfrac{1}{2}\sum_{x,\,y\in G}(f(x)-f(xy))^2\,q(y)$$

and $\tilde{\mathscr{E}}$ with q replaced by $\tilde{q}$.

Lemma 3.1.

$$\tilde{\mathscr{E}}(f,f)\leqslant 4\,\frac{n+1}{n-1}\,\mathscr{E}(f,f)$$

Proof. Given $y\in G$, let $|y|$ be the smallest k for which we can write $y=z_1z_2\cdots z_k$ with $z_i\in E$. Choose one such representation for each y and let $N(z,y)$ be the number of times z appears in that representation. Theorem 1 on p. 2138 of Diaconis and Saloff-Coste[2] shows $\tilde{\mathscr{E}}(f,f)\leqslant C\mathscr{E}(f,f)$ where

$$C=\max_{z\in E}\frac{1}{q(z)}\sum_{y\in G}|y|\,N(z,y)\,\tilde{q}(y) \tag{3.1}$$

To evaluate the constant C we begin by noting that if $y \in \tilde{E}$, $\tilde{q}(y) = 1/\binom{n}{2}$ and if $z \in E$, $q(z) = 1/\binom{n+1}{2}$ so $\tilde{q}(y)/q(z) = (n+1)/(n-1)$. To estimate $|y|$ observe that if $j = i+1$ or $j = i+2$ then $\tau_{i,j} = \rho_{i,j}$ while if $j \geqslant i+3$ then $\tau_{i,j} = \rho_{i,j}\rho_{i+1,j-1}$ so for all $y \in G$ with $\tilde{q}(y) > 0$ we have $|y| \leqslant 2$. On the other hand a given $z = \rho_{i,j}$ can only appear in the representation of $\tau_{i,j}$ and $\tau_{i-1,j+1}$ and it follows that $C \leqslant 4(n+1)/(n-1)$. □

To explain the interest in the comparison in Lemma 3.1, let p_t and $\tilde{p}_t$ be the distributions at time t of the rate 1 continuous time random walks with jump distributions q and $\tilde{q}$ starting from the identity permutation at time 0. Regarding p_t and the uniform distribution as vectors in $\mathbf{R}^g$ and using $\|z\|_2$ to denote the usual L^2 norm for such vectors, Lemma 5 on p. 2136 of Diaconis and Saloff-Coste(2) implies:

Lemma 3.2. If $\tilde{\mathscr{E}}(f, f) \leqslant C\mathscr{E}(f, f)$ then

$$\|p_t - \nu\|_2^2 \leqslant \|\tilde{p}_{t/C} - \nu\|_2^2 \tag{3.2}$$

To use this result we begin by observing that the total variation distance is 1/2 the L^1 norm so the Cauchy–Schwarz inequality implies

$$\|p_t - \nu\|_{TV} \leqslant g^{1/2}\,\|p_t - \nu\|_2 \tag{3.3}$$

where $g = n!$ is the number of elements in the group. To bound the right-hand side we will use results of Diaconis and Shahshahani described in Diaconis(1) which show that there is a constant $a > 0$ so that if $t = (n/2)\ln n + cn$ with $c \geqslant 0$ then

$$g^{1/2}\,\|\tilde{p}_t - \nu\|_2 \leqslant ae^{-2c} \tag{3.4}$$

Combining (3.2)–(3.4) gives the second half of the theorem. □

Proof of Theorem 6. To get an upper bound on the convergence time of the L-transposition chain we use Theorem 1 on p. 2143 of Diaconis and Saloff-Coste(2) which we state as follows:

Lemma 3.3. Let $\mathscr{G}$ be a connected graph on $\{1, 2, ..., n\}$ with edge set E. Define a probability on the symmetric group S_n by $p(id) = 1/n$, $p(i, j) = (n-1)/|E|\, n$ for $(i, j) \in E$ and $\rho(\eta) = 0$ otherwise. For each $x, y \in \mathscr{G}$ let $\gamma_{x,y}$ be a path from x to y in $\mathscr{G}$. Let γ be the length of the longest path, and let

$$b = \max_{e \in E} |\{(x, y): e \in \gamma_{x,y}\}|$$

be the maximum number of times an edge appears in this collection of paths. Let

$$k=\left(\frac{8\,|E|\,\gamma b}{(n-1)}+n\right)(\log n+c)$$

There is a universal constant $\alpha > 0$ so that

$$\|p^k - v\|_{TV} \leqslant \alpha e^{-c}$$

To apply this to the L-transposition, connect i to $i+j$ by an edge whenever $1 \leqslant j \leqslant L$ and the arithmetic is done modulo n. To construct the path from x to y, we suppose without loss of generality that $y = x+m$ with $m \leqslant n/2$. We will use cycles that consist of edges of length 1, 2, 3,..., L. We repeat this cycle until we are within $L(L+1)/2$ of the target site at which point we use edges of length 1 to complete the trip. It is easy to see that the length of the longest path has

$$\gamma \leqslant L\left[\frac{n/2}{L(L+1)/2}\right]+\frac{L(L+1)}{2}=\frac{n}{L+1}+o(n)$$

To bound b we note that edges of length 1 are used the most often and the number of times any edge of length 1 is used will achieve the upper bound. Consider for concreteness the edge from n to 1. There are two cases to consider: (i) this edge is in one of the cycles. (ii) This edge is one of the length one edges at the end. In case (i) x is at $n-m(L(L+1)/2)$ for some $0 \leqslant m \leqslant [n/L(L+1)]$ and y is some site between 1 and $x+n/2$. In case (ii) $n/2 < x \leqslant n$ and $1 \leqslant y \leqslant L(L+1)/2$. From this it follows that

$$b \leqslant \sum_{m=1}^{[n/L(L+1)]}\left(\frac{n}{2}-m\,\frac{L(L+1)}{2}\right)+n/2\cdot\frac{L(L+1)}{2}$$

The second term is of order n. The first is

$$\leqslant \sum_{k=1}^{[n/L(L+1)]+1} k\,\frac{L(L+1)}{2}=\frac{n^2}{4L(L+1)}+o(n^2)$$

Since $E = nL$, we have

$$\frac{8\,|E|\,\gamma b}{n-1}=8\cdot nL\cdot\frac{n}{L+1}\cdot\frac{n^2}{4L(L+1)}\cdot\frac{1}{n}+o(n^3)$$

This gives the result is for the discrete time L-transposition chain. If we run the continuous time chain to time $(1+2\epsilon)\,k$ then with probability $\geqslant 1-e^{-c(\epsilon)\,k}$ there have been at least $(1+\epsilon)\,k$ discrete time steps and the result for the L-transposition follows.

L-Reversal. To prove the result in this case, we use a comparison between the Dirichlet forms of the L-reversal and the n-transposition. This can be done by using the paths above and replacing edges of length $k>1$ by two inversions. Further details are left to the reader. □

4. WILSON'S METHOD

In this section we derive lower bounds on the convergence time of the p-reversal and the p-transposition chains beginning with the former. Recall that $q_i=\sum_{k\geqslant 0} p_{i+2k}$ gives is the rate for jumps of size i in the single marker chain and $q_{-i}=q_i$.

Theorem 3. If the distribution in the p-reversal is fixed and $n\to\infty$ then convergence to equilibrium takes at least time

$$\frac{1}{2}\cdot\frac{n^3}{4\pi^2\sum_i i^2 q_i}\ln n$$

In the L-reversal chain if $L\to\infty$ and $(\ln L)/(\ln n)\to a\in[0,1)$ then convergence to equilibrium takes at least time

$$\frac{(1-a)}{2}\cdot\frac{6}{\pi^2}\cdot\frac{n^3}{L^3}\ln n$$

Proof. The first step is the observation that $f(x)=\sin(2\pi x/n)$ is an eigenfunction for the single marker walk. To check this we note that if $X_0=x$ then

$$\frac{d}{dt}E\sin(2\pi X_t/n)\bigg|_{t=0}$$

$$=\sum_{i=1}^{n}\frac{q_i}{n}\left[\sin(2\pi(x+i)/n)+\sin(2\pi(x-i)/n)-2\sin(2\pi x/n)\right]$$

$$=\sum_{i=1}^{n}\frac{2q_i}{n}\left[\cos(2\pi i/n)-1\right]\sin(2\pi x/n)$$

where in the second step we have used the trigonometric identity

$$\sin(\alpha+\beta)=\sin(\alpha)\cos(\beta)+\sin(\beta)\cos(\alpha)$$

Thus f is an eigenfunction with eigenvalue

$$-\lambda=\sum_{i=1}^{n}\frac{2q_i}{n}[\cos(2\pi i/n)-1] \tag{4.1}$$

Let $\eta_t(i)$ be the marker at position i at time t, and $X_t^j=\eta_t^{-1}(j)$ be the location of marker j at time t. Let

$$g(m)=\operatorname{sgn}\left(\frac{n+1}{2}-m\right)$$

and let

$$\Phi(\eta_t)=\sum_i g(\eta_t(i))\sin(2\pi i/n)=\sum_j g(j)\sin(\pi X_t^j/n)$$

Letting $\Phi_t=\Phi(\eta_t)$, it follows from the calculation above that

$$\frac{d}{dt}E\Phi_t=-\lambda E\Phi_t \tag{4.2}$$

If the markers are initially in order, i.e., $\eta_0(i)=i$ for all i then

$$\Phi_0\approx 2n\int_0^{1/2}\sin(2\pi y)\,dy=\frac{2n}{\pi} \tag{4.3}$$

Our next step is to estimate the variance of Φ_t. The argument follows Lemma 5 of Wilson[9] but is simpler since time is continuous. To isolate this calculation from the rest of the proof we state it as

Lemma 4.1. Suppose Φ is an eigenfunction with eigenvalue $-\lambda$ for a chain that jumps from η to σ at rate $Q(n,\sigma)$. If we let

$$(\nabla\Phi)^2(\eta)=\sum_\sigma Q(\eta,\sigma)\{\Phi(\eta)-\Phi(\sigma)\}^2$$

and assume that $\operatorname{Var}(\Phi(\eta_0))=0$ then $\operatorname{Var}(\Phi(\eta_t))\leqslant\|(\nabla\Phi)^2\|_\infty/2\lambda$.

Proof. Our first step is to observe that

$$E(\Phi_{t+s}^2\mid\mathscr{F}_t)=\Phi_t^2+2\Phi_tE(\Phi_{t+s}-\Phi_t\mid\mathscr{F}_t)+E((\Phi_{t+s}-\Phi_t)^2\mid\mathscr{F}_t)$$

Letting $s \to 0$ and using our assumptions and notation we have

$$\frac{d}{ds} E(\Phi_{t+s}^2 \mid \mathscr{F}_t)\bigg|_{s=0} = -2\lambda\Phi_t^2 + (\nabla\Phi)^2\,(\eta_t)$$

Taking expected value gives

$$\frac{d}{dt} E(\Phi_t^2) = -2\lambda E(\Phi_t^2) + E(\nabla\Phi)^2\,(\eta_t)$$

Using the eigenfunction assumption again, it follows that

$$\frac{d}{dt}(E\Phi_t)^2 = 2E\Phi_t \frac{d}{dt} E\Phi_t = -2\lambda(E\Phi_t)^2$$

Subtracting this from the previous equation we have

$$\frac{d}{dt}\operatorname{Var}(\Phi_t) = -2\lambda \operatorname{Var}(\Phi_t) + E(\nabla\Phi)^2\,(\eta_t)$$

Solving the differential equation and noting that $\operatorname{Var}(\Phi_0) = 0$ we have

$$\operatorname{Var}(\Phi_t^2) = \int_0^t e^{-2\lambda(t-s)} E(\nabla\Phi)^2\,(\eta_s)\,ds$$

Integrating we have

$$\operatorname{Var}(\Phi_t^2) \leqslant \frac{\|(\nabla\Phi)^2\|_\infty}{2\lambda} \qquad \square$$

Lemma 4.2. Let $B = \|(\nabla\Phi)^2\|_\infty / 2\lambda$ and $t(\epsilon) = \frac{1}{\lambda}\ln(\Phi_0/2\sqrt{B/\epsilon})$. The total variation distance between the distribution at time $t(\epsilon)$ and equilibrium is at least $1 - 2\epsilon$.

Proof. The differential equation (4.2) implies that in equilibrium $E\Phi_\infty = 0$, so Chebyshev's inequality implies

$$P(\Phi_\infty \geqslant \sqrt{B/\epsilon}) \leqslant \epsilon$$

$E\Phi_{t(\epsilon)} = e^{-\lambda t(\epsilon)}\Phi_0 = 2\sqrt{B/\epsilon}$ so using Chebyshev's inequality again we have

$$P(\Phi_{t(\epsilon)} \leqslant \sqrt{B/\epsilon}) \leqslant \epsilon$$

and the desired result follows. □

Fixed Jump Distribution. Since $1-\cos x \sim x^2/2$ as $x \to 0$, our assumption implies

$$-\lambda = \sum_{i=1}^{n} \frac{2q_i}{n} [\cos(2\pi i/n) - 1] \sim -\frac{4\pi^2}{n^3} \sum_{i=1}^{n} i^2 q_i \tag{4.4}$$

To bound $\|(\nabla\Phi)^2\|_\infty$, we begin with the observation that

$$\left|\sin\left(2\pi\left(x+\frac{y}{n}\right)\right) - \sin\left(2\pi\left(x-\frac{y}{n}\right)\right)\right| = \left|\int_{2\pi(x-y/n)}^{2\pi(x+y/n)} \cos z\, dz\right| \leqslant 4\pi y/n \tag{4.5}$$

When the reversal $\rho_{i,i+j}$ has j even, e.g., $j=8$, where the picture might be

$$1 \quad 1 \quad 1 \quad 1 \quad ? \quad -1 \quad -1 \quad -1 \quad -1$$

Since the worst thing that can happen is that a 1 exchanges places with a -1, (4.5) implies that the change in Φ

$$|\Phi(\eta) - \Phi(\eta\rho_{i,i+j})| \leqslant \sum_{k=1}^{j/2} 2 \cdot \frac{4\pi k}{n} = \frac{\pi}{n}\, j(j+2)$$

When the reversal $\rho_{i,i+j}$ has j odd, e.g., $j=7$, where the picture might be

$$1 \quad 1 \quad 1 \quad 1 \quad -1 \quad -1 \quad -1 \quad -1$$

and (4.5) implies that the change in Φ

$$|\Phi(\eta) - \Phi(\eta\rho_{i,i+j})| \leqslant \sum_{k=0}^{(j-1)/2} 2 \cdot \frac{4\pi(2k+1)}{2n} = \frac{\pi}{n}\, (j+1)^2$$

since the sum of the first m odd integers, $1+3+\cdots+(2m-1) = m^2$. Since $(j+1)^2 > j(j+2)$ we have

$$|\Phi(\eta) - \Phi(\eta\rho_{i,i+j})| \leqslant \frac{\pi}{n}\, (j+1)^2$$

Using this in the definition we have

$$(\nabla\Phi)^2\,(\eta) \leqslant \sum_{i,j} \frac{p_j}{n} \frac{\pi^2}{n^2}\, (j+1)^4 = \frac{\pi^2}{n^2} \sum_j p_j (j+1)^4 \tag{4.6}$$

Comparing with (4.4), we have

$$B \leqslant Cn$$

Since $\Phi_0 \sim 2n/\pi$, using Lemma 4.2 we see that if $D = \ln(\epsilon^{1/2}/\pi C^{1/2})$ then at time $t = (1/2\lambda)(D + \ln n)$ the total variation distance is at least $1 - 2\epsilon$.

L-Reversal Chain. Under the assumption $L/n \to 0$ as $n \to \infty$ the computation in (4.4) is valid. In the uniform case, $q_i \approx (L-i)/2L$ for $i \leqslant L$, so if $L \to \infty$

$$\sum_{i=1}^{L} i^2 q_i \sim L^3 \int_0^1 x^2 \frac{(1-x)}{2}\, dx = L^3/24 \qquad \text{and} \qquad \lambda \sim \frac{\pi^2 L^3}{6n^3} \tag{4.7}$$

When p_i is uniform on 1, 2,..., L and $L \to \infty$, the sum on the right-hand side of (4.6) becomes

$$\frac{1}{L} \sum_{j=1}^{L} (j+1)^4 \sim L^4 \int_0^1 x^4\, dx = L^4/5 \tag{4.8}$$

Combining (4.6), (4.7), and (4.8) we have

$$B \leqslant CnL$$

Since $\Phi_0 \sim 2n/\pi$, using Lemma 4.2 we see that if $D = \ln(\epsilon^{1/2}/\pi C^{1/2})$ then at time $t = (1/2\lambda)(D + \ln(n/L))$ the total variation distance is at least $1 - 2\epsilon$. □

Theorem 5. If the distribution in the p-transposition is fixed and $n \to \infty$ then convergence to equilibrium takes at least time

$$\frac{1}{2} \cdot \frac{n^3}{4\pi^2 \sum_i i^2 p_i} \ln n$$

In the L-transposition if $L \to \infty$ and $L/n \to 0$ then convergence to equilibrium takes at least time

$$\frac{1}{2} \cdot \frac{3}{4\pi^2} \cdot \frac{n^3}{L^2} \ln n$$

If $L/n \to a \in (0, 1]$ then convergence to equilibrium takes at least time

$$\frac{1}{2} \cdot \frac{na}{\int_0^a 1 - \cos 2\pi x\, dx} \ln n$$

Proof. Fixed Distribution. Since $1-\cos x \sim x^2/2$ as $x \to 0$, our assumption implies

$$-\lambda = \sum_{i=1}^{n} \frac{2p_i}{n} [\cos(2\pi i/n) - 1] \sim -\frac{4\pi^2}{n^3} \sum_{i=1}^{n} i^2 p_i \tag{4.9}$$

To bound $\|(\nabla\Phi)^2\|_\infty$, we observe that (4.5) implies

$$|\Phi(\eta) - \Phi(\eta\lambda_{i,i+j})| \leqslant 2 \cdot \frac{2\pi j}{n}$$

Using this in the definition we have

$$(\nabla\Phi)^2(\eta) \leqslant \frac{16\pi^2}{n^2} \sum_j p_j j^2 \tag{4.10}$$

Comparing with (4.9), we have

$$B \leqslant Cn$$

Since $\Phi_0 \sim 2n/\pi$, using Lemma 4.2 we see that if $D = \ln(\epsilon^{1/2}/\pi C^{1/2})$ then at time $t = (1/2\lambda)(D + \ln n)$ the total variation distance is at least $1 - 2\epsilon$.

L-Transposition. Under the assumption $L/n \to 0$ as $n \to \infty$ the computation in (4.9) is valid. If $L \to \infty$ then

$$\sum_{i=1}^{L} i^2 p_i \approx L^2 \int_0^1 x^2\, dx = L^2/3 \qquad \text{and} \qquad \lambda \sim \frac{4\pi^2 L^2}{3n^3} \tag{4.11}$$

When p_i is uniform on 1, 2,..., L and $L \to \infty$, the right-hand side of (4.10) becomes

$$\frac{16\pi^2}{n^2} \cdot \frac{1}{L} \sum_{j=1}^{L} j^2 \sim \frac{16\pi^2}{n^2} \cdot L^2 \int_0^1 x^2\, dx = \frac{16\pi^2}{n^2} \cdot \frac{L^3}{3} \tag{4.12}$$

Comparing with (4.11), we have

$$B \leqslant 2n$$

and the proof is completed as in the previous case.

If $L/n \to a \in (0, 1]$ then using (4.9) we have

$$-\lambda = \frac{2}{nL} \sum_{i=1}^{L} [\cos(2\pi i/n) - 1] \sim \frac{2}{na} \int_0^a -1 + \cos(2\pi x)\, dx$$

Using (4.12) now we have $B \leqslant Cn$ and the rest is as before. □

5. INDEPENDENT PARTICLES

Consider first one random walk on the circle that jumps from x to $x+i$ at rate r_i and from x to $x-i$ at rate r_i and assume that $r_1 > 0$ so that the walk is irreducible. $d_n(t) = \| p_t(0, \cdot) - \nu \|_{TV} \downarrow 0$ continuously as $t \uparrow \infty$ so we can define $t(c, n)$ by $d_n(t(c, n)) = n^{-c}$. Let $\mu_i = p_{t(c,n)}(i, \cdot)$ and $\nu_i = \nu$ for $1 \leqslant i \leqslant n$.

Lemma 5.1. As $n \to \infty$ the total variation distance

$$e_n(c) = \| \mu_1 \times \mu_2 \times \cdots \times \mu_n - \nu_1 \times \nu_2 \times \cdots \times \nu_n \|_{TV}$$

tends to 0 for $c < 1$ and to 1 for $c > 1$.

Proof. A standard result, see, e.g., (6.2) on p. 139 of Durrett[4] implies

$$\| \mu_1 \times \mu_2 \times \cdots \times \mu_n - \nu_1 \times \nu_2 \times \cdots \times \nu_n \|_{TV} \leqslant \sum_i \| \mu_i - \nu_i \|_{TV} \tag{5.1}$$

From this it follows that if $c > 1$ then $e_n(c) \leqslant n^{1-c} \to 0$ as $n \to \infty$.

To prove a result in the other direction we need a converse to (5.1). We begin with the observation that we can define random variables X_i and Y_i with distributions μ_i and ν_i so that $P(X_i \neq X_j) = \| \mu_i - \nu_i \|_{TV}$. If (X_1, Y_1) and (X_2, Y_2) are independent then

$$\begin{aligned} P(X_1 = Y_1, X_2 = Y_2) &= P(X_1 = Y_1)\, P(X_2 = Y_2) \\ &= 1 - P(X_1 = Y_1) + P(X_1 = Y_1)\{1 - P(X_2 = Y_2)\} \end{aligned}$$

so we have $\| \mu_1 \times \mu_2 - \nu_1 \times \nu_2 \|_{TV} = \| \mu_1 - \nu_1 \|_{TV} + (1 - \| \mu_1 - \nu_1 \|_{TV}) \| \mu_2 - \nu_2 \|_{TV}$. From this and induction it follows that

$$\begin{aligned} &\| \mu_1 \times \mu_2 \times \cdots \times \mu_n - \nu_1 \times \nu_2 \times \cdots \times \nu_n \|_{TV} \\ &\qquad = \sum_{k=1}^{n} \| \mu_k - \nu_k \|_{TV} \prod_{j=1}^{k-1} (1 - \| \mu_j - \nu_j \|_{TV}) \end{aligned} \tag{5.2}$$

When $\| \mu_k - \nu_k \|_{TV} = n^{-c}$ with $c < 1$, we have

$$e_n(c) = \sum_{k=1}^{n} n^{-c}(1 - n^{-c})^{k-1} = 1 - (1 - n^{-c})^n$$

If $c < 1$ the right-hand side tends to 1 and the proof of Lemma 5.1 is complete. □

To complete the proof of Theorem 4 now it suffices to show

Lemma 5.2.

$$t(c, n) \sim \frac{cn^2}{4\pi^2 \sum_i i^2 r_i} \ln n$$

Proof. Repeating the first calculation of the proof of Theorem 3 in Section 4 it follows that for $0 \leqslant k < n$, $f_k(x) = \sin(2\pi kx/n)$ is an eigenfunction with eigenvalue

$$-\lambda_k = \sum_{i=1}^{n} 2r_i[\cos(2\pi ki/n) - 1] \approx -\frac{4\pi^2 k^2}{n^2} \sum_i i^2 r_i \tag{5.3}$$

the last approximation holding if k/n is small. Let $h_k(x)$ be $f_k(x)$ normalized to have $\sum_{x=1}^{n} h_k(x)^2 = 1$. When $k = 0$, $h_0(x) = 1\sqrt{n}$. If k/n is small

$$\sum_{i=1}^{n} \cos^2\left(\frac{2\pi ki}{n}\right) \approx n \int_0^1 \cos^2\left(\frac{2\pi kx}{n}\right) dx = \frac{n}{2}$$

since $\sin^2 + \cos^2 = 1$, so in this case $h_k(x) \approx \sqrt{2/n}\, f_k(x)$.

Since our chain is symmetric, a theorem on p. 243 of Riesz and Nagy's[8] *Functional Analysis* implies that we can write

$$p_t(x, y) = \sum_{k=0}^{n-1} e^{-\lambda_k t} h_k(x)\, h_k(y)$$

(Their result is for the powers of a matrix but generalizes easily to the form given here since the continuous time chain is the sum of a Poisson number of iterates of the one step transition matrix.) The $k = 0$ term corresponds to the stationary distribution so

$$p_t(0, y) - \nu(y) = \sum_{k=0}^{n-1} e^{-\lambda_k t} h_k(0)\, h_k(y) \tag{5.4}$$

To get a lower bound on the distance from equilibrium we note that

$$2\, \|p_t - \nu\|_{TV} = \|p_t - \nu\|_1 = \sup_{g:\, \|g\|_\infty = 1} \left| \sum_y (p_t(0, y) - \nu(y))\, g(y) \right|$$

with the sup achieved by $g(y)=\operatorname{sgn}(p_t(0, y)-v(y))$. Taking $g(y)=h_1(y)/h_1(0)$ and using the orthogonality of eigenfunctions, we have

$$\sum_y (p_t(0, y)-v(y))\, g(y)=\sum_y e^{-\lambda_1 t}h_1(y)^2=e^{-\lambda_1 t}$$

Combining this with the previous equation we have

$$\|p_t-v\|_{TV}\geqslant \tfrac{1}{2}\, e^{-\lambda_1 t}$$

To get an upper bound we use the Cauchy–Schwarz inequality to conclude

$$\|p_t-v\|_1\leqslant \sqrt{n}\, \|p_t-v\|_2$$

Orthogonality and (5.4) imply

$$\left\|\sqrt{n}\, p_t-\frac{1}{n}\right\|_2^2=\sum_{k=1}^{n-1} (\sqrt{n}\, h_k(0))^2\, e^{-2\lambda_k t}$$

We are interested in this result when $t=c(\ln n)/\lambda_1$. When k/n is small $\lambda_k\approx k^2\lambda_1$ so

$$e^{-2\lambda_k t}\approx n^{-2ck^2}$$

Pick an integer K so that $cK^2>1+2c$. Since $h_k(0)\leqslant 1$ for all k we have

$$\sum_{k=K}^{n-1} (\sqrt{n}\, h_k(0))^2\, e^{-2\lambda_k t}\leqslant n^2 e^{-2\lambda_K c}\leqslant n^{-4c}$$

The first part of the sum

$$\sum_{k=1}^{K-1} (\sqrt{n}\, h_k(0))^2\, e^{-2\lambda_k t}\sim 4n^{-2c}$$

and the proof of Lemma 5.2 is complete. □

ACKNOWLEDGMENT

This work was partially supported by NSF Grant DMS 98770066 and by a supplement to NIH Grant GM36431 (P.I. Chip Aquadro).

REFERENCES

1. Diaconis, P. (1988). *Group Representations in Probability and Statistics*, Institute of Mathematical Statistics Lecture Notes, Vol. 11.
2. Diaconis, P., and Saloff-Coste, L. (1993). Comparison techniques for random walks on finite groups. *Ann. Probab.* **21**, 2131–2156.
3. Diaconis, P., and Shahshahani, M. (1981). Generating a random permutation with random transpositions. *Z. Wahr.* **57**, 159–179.
4. Durrett, (1995). *Probability: Theory and Examples*, 2nd Edition, Duxbury Press, Belmont, California.
5. Hartl, D. L., and Jones, E. W. (2000). *Genetics: Analysis of Genes and Genomes*, Jones and Barlett, Sudbury, MA.
6. Ranz, J. M., Casals, F., and Ruiz, A. (2001). How malleable is the eukaryotic genome? Extreme rate of chromosomal rearrangement in the genus *Drosophila*. *Genome Research* **11**, 230–239.
7. Ranz, J. M., Segarra, S., and Ruiz, A. (1997). Chromosomal homology and molecular organization of Muller's element *D* and *E* in the *Drosophila repleta* species group. *Genetics* **145**, 281–295.
8. Riesz, F., and Sz.-Nagy, B. (1965). *Functional Analysis*. Translated from the second French edition, Frederick Ungar Publishing Co., New York.
9. Wilson (2001). *Mixing Times of Lozenge Tilings and Card Shuffling Markov Chains*, arXiv: math.PR/0102193.