The L-Reversal Shuffle: Reversibility, Azuma–Hoeffding Lower Bounds, and Dirichlet-Form Comparisons

Chen Jiayi

Supervisor: Sun Rongfeng
Examiner: N/A

Summary

This report highlights a surprising bridge between card shuffling mathematics and genome biology: the same probabilistic tools used to study how decks mix can model how genes reorder along a chromosome. We focus on the L-reversal process, which flips any contiguous block of up to L elements, and analyze its mixing time, the number of such inversions needed before the genome's order appears effectively random under the model.

Modeling genome rearrangements is medically important because inversions help drive cancers, congenital disorders, and pathogen evolution. Realistic models support variant interpretation, pipeline validation, and therapeutic risk assessment. The mixing time bound indicates how long to run randomized controls for significance testing, when observed structure reflects true biology rather than incomplete shuffling, and how local versus global rearrangements shape evolutionary trajectories. Together, L-reversal modeling and mixing time analysis provide a principled framework for translating combinatorial theory into clinically relevant insight.

In § 16.1.2 of *Markov Chains and Mixing times*, the text, by reference to Exercises 2.2 and 2.3, gives an upper bound on the expected time for the Markov chain Y to hit 0. In **Section 0.01** of this report, I prove an upper bound on the hitting time of 0 for the auxiliary chain Y_t. Subsequently, in **Section 0.02**, I apply this bound to complete the mixing-time upper bound for the lazy adjacent-transposition chain. Finally, **Section 0.04** uses this bound to obtain an upper bound on the mixing time of the L-reversal chain.

In **Section 0.03** of this report, I prove *Exercise 16.4* from *Markov Chains and Mixing times* by tracking a single tagged element with a Fourier test function, averaging over block locations and lengths, and invoking Proposition 4.5 to obtain a mixing time lower bound for the L-reversal chain that depends on the block length L.

Section 0.04 of this report compares the L-reversal chain with the adjacent transposition shuffle. Using Lemma 13.6 on Dirichlet forms together with Lemma 13.7 via the Rayleigh Ritz principle, I show that entrywise domination of the transition kernels yields a comparison of spectral gaps. Combining this with Theorem 12.4, which relates the spectral gap to the mixing time, gives an upper bound on the mixing time of the L-reversal chain that is within a constant factor of the adjacent transposition bound.

Motivated by § 16.2 of *Markov Chains and Mixing Times*, **Section 0.05** of this report studies the L-reversal chain under a weighted update rule that assigns a selection probability w_ℓ to each segment length $1 \leq \ell \leq L$, with $\sum_{\ell=1}^{L} w_\ell = 1$. The weight profile serves two purposes: it models scenarios in which some reversal lengths are chosen more frequently than others (for example, a bias toward short blocks), and it makes the dependence of the mixing time on the update rule explicit. Under these weighted dynamics, I prove a lower bound on the mixing

time whose order and dependence on n, L, and the weight profile are all specified.

By examining mixing times under modified dynamics and relating them to the random adjacent transposition shuffle, this report lays a foundation for further study of inversion–driven genome models and their potential medical applications.

Section 0.00. Basic probabilistic tools

In this report I repeatedly use a few basic tools from probability and Markov–chain theory. For completeness, I recall them here.

Markov's Inequality. Let X be a nonnegative random variable with $\mathbb{E}[X] < \infty$. Then for every $a > 0$,

$$\Pr\{X \geq a\} \leq \frac{\mathbb{E}[X]}{a}.$$

Hitting Time. Let $(X_t)_{t \geq 0}$ be a Markov chain on a finite state space $\mathcal{X}$ with transition matrix P. For a state $x \in \mathcal{X}$, the *hitting time* of x is

$$\tau_x := \min\{t \geq 0 : X_t = x\},$$

the first time at which the chain visits x.
More generally, for a subset $A \subseteq \mathcal{X}$, the hitting time of A is

$$\tau_A := \min\{t \geq 0 : X_t \in A\}.$$

Absorbing States. Let $(X_t)_{t \geq 0}$ be a Markov chain on a finite state space $\mathcal{X}$ with transition matrix P. A state $x \in \mathcal{X}$ is called *absorbing* if, once the chain enters x, it never leaves. Formally,

$$P(x, x) = 1 \quad \Longleftrightarrow \quad \Pr(X_{t+1} = x \mid X_t = x) = 1 \text{ for all } t \geq 0.$$

In terms of communicating classes, this means that the communicating class of x is the singleton $\{x\}$ and that x is essential: once the chain first visits x, it remains at x forever.

Reflecting Boundaries. A boundary state y is called *reflecting* if steps that would move the chain outside the state space are "bounced back" into the interior. For example, consider a simple random walk $(B_t)_{t \geq 0}$ on $\{0, 1, \ldots, n-1\}$:

$$\Pr(B_{t+1} = k + 1 \mid B_t = k) = \Pr(B_{t+1} = k - 1 \mid B_t = k) = \tfrac{1}{2}, \qquad 1 \leq k \leq n - 2.$$

A reflecting boundary at $n - 1$ can be implemented by

$$\Pr(B_{t+1} = n - 2 \mid B_t = n - 1) = 1,$$

so that the chain never leaves $\{0, \ldots, n-1\}$ but "reflects" when it tries to step from $n - 1$ to n.

Symmetric Group For a positive integer n, the *symmetric group* S_n is the set of all permutations of $\{1, \ldots, n\}$, with composition as the group operation. We view a permutation $\sigma \in S_n$ both as a bijection $\{1, \ldots, n\} \to \{1, \ldots, n\}$ and as the word $(\sigma(1), \ldots, \sigma(n))$. Throughout the report, the *uniform distribution* on S_n means the measure that assigns mass $1/n!$ to each $\sigma \in S_n$.

Mixing Time. Let $(X_t)_{t \geq 0}$ be an irreducible Markov chain on a finite state space $\mathcal{X}$ with transition matrix P and stationary distribution π. For $t \geq 0$ define the distance to stationarity

$$d(t) := \max_{x \in \mathcal{X}} \left\| P^t(x, \cdot) - \pi \right\|_{\mathrm{TV}},$$

where $\| \cdot \|_{\mathrm{TV}}$ denotes total variation distance. For $\varepsilon \in (0, 1)$ the *mixing time* at accuracy ε is

$$t_{\mathrm{mix}}(\varepsilon) := \min\{t \geq 0 : d(t) \leq \varepsilon\},$$

and the (standard) *mixing time* is

$$t_{\mathrm{mix}} := t_{\mathrm{mix}}(1/4).$$

Martingale. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $(\mathcal{F}_t)_{t \geq 0}$ a filtration (an increasing family of sub-σ-algebras of $\mathcal{F}$). A *martingale with respect to a filtration* $(\mathcal{F}_t)_{t \geq 0}$ is a stochastic process $(M_t)_{t \geq 0}$ such that

1. M_t is $\mathcal{F}_t$-measurable for each t;

2. $\mathbb{E}[|M_t|] < \infty$ for each t;

3. $\mathbb{E}[M_{t+1} \mid \mathcal{F}_t] = M_t$ for all $t \geq 0$.

Given an integrable random variable X_t and a filtration $(\mathcal{F}_s)_{s=0}^t$, the *Doob martingale* associated with X_t is the process

$$M_s := \mathbb{E}[X_t \mid \mathcal{F}_s], \qquad s = 0, 1, \ldots, t.$$

This is a martingale in the sense above.

Azuma–Hoeffding inequality. Let $(Y_s)_{s=1}^t$ be a martingale–difference sequence with respect to a filtration $(\mathcal{F}_s)_{s=0}^t$, i.e. $\mathbb{E}[Y_s \mid \mathcal{F}_{s-1}] = 0$ for all s, and assume the increments are uniformly bounded:

$$|Y_s| \leq B_s \qquad \text{almost surely, for } s = 1, \ldots, t.$$

Then for every $u > 0$,

$$\Pr\left\{ \sum_{s=1}^t Y_s \geq u \right\} \leq \exp\left(-\frac{u^2}{2 \sum_{s=1}^t B_s^2} \right), \qquad \Pr\left\{ \sum_{s=1}^t Y_s \leq -u \right\} \leq \exp\left(-\frac{u^2}{2 \sum_{s=1}^t B_s^2} \right).$$

Strong Markov property. Let $(X_t)_{t \geq 0}$ be a Markov chain with transition matrix P and let τ be a (finite) stopping time with respect to the natural filtration $(\mathcal{F}_t)_{t \geq 0}$ of X_t. Then, conditional on $\{\tau < \infty\}$ and on $X_\tau = x$, the process $(X_{\tau+t})_{t \geq 0}$ is a Markov chain with the same transition matrix P started from x. In particular, for every bounded function f and every $t \geq 0$,

$$\mathbb{E}\left[f(X_{\tau+t}) \,\middle|\, \mathcal{F}_\tau \right] = \mathbb{E}_x\left[f(X_t) \right]\big|_{x = X_\tau}.$$

Context: Random Adjacent Transposition and its lazy version. A classical Markov chain on the symmetric group S_n is the *random adjacent transposition* shuffle. At each step we choose a position $1 \leq i \leq n-1$ uniformly at random and transpose the cards in positions i and $i+1$. This defines an irreducible aperiodic walk on S_n whose stationary distribution is uniform. A common variant is the *lazy random adjacent transposition* chain, where at each step we stay put with probability $1/2$ and perform an adjacent transposition (with some $i, i+1$) with probability $1/2$. The laziness removes periodicity while preserving the same stationary distribution.

0.0.1 Upper bound on the Hitting time by coupling

Summary. In this section we prove an upper bound on the hitting timeof 0 for the comparison chain (Y_t). By the reduction from (X_t) to (Y_t) described in §16.1.2 of *Markov Chains and Mixing Times*, the hitting time of 0 for (X_t) is controlled by that of (Y_t), so our bound for (Y_t) immediately yields the same order of bound for the original chain (X_t).

Coupling on the doubled line. Let (Z_t) be the simple random walk on the doubled line

$$\{0, 1, \ldots, 2n-1\}, \qquad Z_0 = n,$$

with absorption at the boundary $\{0, 2n-1\}$.
Define the reflected process

$$B_t := \min\{Z_t, \ 2n-1-Z_t\},$$

so that B_t is absorbing at 0 and reflecting at $n-1$. Consequently the hitting time of 0 for B_t coincides in law with the boundary hitting time for Z_t:

$$\tau_B := \inf\{t \geq 0 : \ B_t = 0\} = \inf\{t \geq 0 : \ Z_t \in \{0, 2n-1\}\} = \tau_Z$$

Proposition 2.1.(Levin, Peres, and Wilmer 2009, Prop. 2.1). For the simple random walk on $\{0, \ldots, N\}$ absorbed at $\{0, N\}$, the expected absorption time from i is

$$\mathbb{E}_i[\tau_{0,N}] \ = \ i\,(N-i).$$

This is the standard gambler's ruin fact. Applying this with $N = 2n-1$ and $i = Z_0 = n$ gives,

$$\mathbb{E}_n[\tau_{0,2n-1}] \ = \ \mathbb{E}_n[\tau_Z] = n(n-1) = \mathbb{E}[\tau_B].$$

From moves to physical time (lazy chain). Let p be the probability that a given time step is a *move* (i.e., not a hold). For the lazy walk with hold probability $1/2$ and two symmetric neighbours (except at the reflecting top, where the move probability is smaller and thus can only decrease the hitting time, so it does not affect the *upper* bound we derive), we have

$$p = \frac{1}{n-1} + \frac{1}{n-1} = \frac{2}{n-1}.$$

Let $G_1, G_2, \ldots$ be the i.i.d. waiting times (in *time steps*) between consecutive moves; then $G_j \sim \mathrm{Geom}(p)$ and

$$\mathbb{E}[G_j] \;=\; \frac{1}{p} \;=\; \frac{n-1}{2}.$$

If K denotes the number of moves (i.e., rounds) required until absorption, then the total physical time is $\sum_{j=1}^{K} G_j$, and hence

$$\mathbb{E}\left[\sum_{j=1}^{K} G_j\right] = \mathbb{E}\left[\mathbb{E}\left[\sum_{j=1}^{K} G_j \,\middle|\, K\right]\right] = \mathbb{E}[K\,\mathbb{E}[G]] = \mathbb{E}[K]\,\mathbb{E}[G] \;=\; n(n-1)\cdot\frac{n-1}{2}.$$

When the delayed walk Y_t is started at $n-1$, it can be coupled with the reflected walk B_t started at $B_0 = n-1$, whose (non-lazy) move count matches that of the doubled–line walk Z_t started at $Z_0 = n$. Therefore,

$$\mathbb{E}_{n-1}[\tau_0^Y] \;=\; \mathbb{E}[K]\cdot\mathbb{E}[G] \;\le\; n(n-1)\cdot\frac{n-1}{2}$$

Moreover, $n-1$ is the farthest state from 0 (the unique "worst-case" starting point for hitting 0), so for any initial state i we have $\mathbb{E}_i[\tau_0^Y] \le \mathbb{E}_{n-1}[\tau_0^Y]$. Thus, the hitting time of 0 satisfies

$$\mathbb{E}_i\left[\tau_0^Y\right] \;\le\; \frac{n(n-1)^2}{2} \;\le\; \frac{n^2(n-1)}{2},$$

uniformly over all starting states i.

0.0.2   Upper bound on mixing time via coupling (§16.1.2)

Summary. This section recites the upper–bound on the mixing time proof for the (lazy) adjacent–transposition shuffle from §16.1.2. of *Markov Chains and Mixing Times.*

We couple two copies of the lazy adjacent–transposition chain, denoted σ_t and σ_t' ("left" and "right" decks). At each step choose an adjacent pair $(i, i+1)$ uniformly at random. Flip a fair coin to decide whether to perform the adjacent transposition at $(i, i+1)$ on the *left* deck (otherwise the left deck stays put). Now update the *right* deck as follows:

- If $\sigma_t(i) = \sigma_t'(i+1)$ or $\sigma_t(i+1) = \sigma_t'(i)$ (the two decks are "crossed" at $i, i+1$), perform the *opposite* update on the right deck (transpose if the left deck stayed still, and vice versa).

- Otherwise, perform the *same* update on the right deck as on the left deck.

Fix a card a, and let X_t be the (unsigned) distance between the positions of a in the two decks at time t. By construction,

$$|X_{t+1} - X_t| \le 1 \quad \text{and} \quad X_t = 0 \implies X_s = 0 \text{ for all } s \ge t.$$

Let $\tau_a := \min\{t \ge 0 : X_t = 0\}$ be the synchronization time of card a. The key point is that (X_t), although it is not itself a Markov chain, is stochastically dominated by a Markov chain Y_t

Domination by a delayed walk. Define Y_t to be the delayed random walk on $\{0, 1, \ldots, n-1\}$ with

$$
\begin{cases}
\text{hold with prob. } 1 - \dfrac{2}{n-1}, \\
\text{move } +1 \text{ with prob. } \dfrac{1}{n-1} \quad (\text{unless at } n-1), \\
\text{move } -1 \text{ with prob. } \dfrac{1}{n-1}, \\
\text{absorb at } 0.
\end{cases}
$$

One can couple so that $X_0 = Y_0$ and $X_t \le Y_t$ for all $t \ge 0$. Hence,

$$
\tau_a \le \tau_0^Y := \min\{t \ge 0 : Y_t = 0\}.
$$

In Section 0.0.1 we proved (by coupling to a reflected walk on the doubled line) that

$$
\mathbb{E}[\tau_0^Y] \le \frac{n^2(n-1)}{2} \le \frac{n^3}{2}. \tag{0.0.1}
$$

Moreover, this bound holds uniformly over the starting state $Y_0 \in \{0, 1, \ldots, n-1\}$.
By Markov's inequality and eq. (0.0.1),

$$
\Pr\{\tau_a > n^3\} \le \frac{\mathbb{E}[\tau_a]}{n^3} \le \frac{\mathbb{E}[\tau_0^Y]}{n^3} \le \frac{1}{2}.
$$

Thus, from any starting state, a block of n^3 shuffles synchronizes the fixed card a with probability at least $1/2$.

By the strong Markov property and the uniform bound above, we can iterate this estimate over successive blocks of length n^3: for every integer $k \ge 1$. Set

$$
T := 2 n^3 \left\lceil \log_2 n \right\rceil.
$$

From the previous bound we have, for each card a,

$$
\mathbb{P}\{\tau_a > T\} \le \left(\tfrac{1}{2}\right)^{2\lceil \log_2 n \rceil} \le \frac{1}{n^2}.
$$

Define the (coalescence) coupling time by

$$
\tau_{\mathrm{cpl}} := \inf\{t \ge 0 : \sigma_t = \sigma'_t\}.
$$

By construction, $\sigma_t = \sigma'_t$ holds iff all cards are synchronized, i.e., iff $X_t^{(a)} = 0$ for every a.
Hence

$$
\{\tau_{\mathrm{cpl}} \le t\} = \bigcap_a \{\tau_a \le t\} \qquad \text{and therefore} \qquad \tau_{\mathrm{cpl}} = \max_a \tau_a.
$$

Consequently,

$$
\mathbb{P}\{\tau_{\mathrm{cpl}} > T\} = \mathbb{P}\left\{\max_a \tau_a > T\right\} \le \sum_{a=1}^n \mathbb{P}\{\tau_a > T\} \le \frac{1}{n}.
$$

Invoking **Corollary 5.5** (Levin, Peres, and Wilmer 2009, Cor. 5.5),

$$d(t) \;=\; \max_x \left\| P^t(x, \cdot) - \pi \right\|_{\mathrm{TV}} \;\leq\; \max_{x,y} \Pr_{x,y}\{\tau_{\mathrm{cpl}} > t\}.$$

Since the tail bound $\Pr_{x,y}\{\tau_{\mathrm{cpl}} > T\} \leq 1/n$ holds uniformly over all starting states x, y, we have

$$d(T) \;\leq\; \max_{x,y} \Pr_{x,y}\{\tau_{\mathrm{cpl}} > T\} \;\leq\; \frac{1}{n}.$$

Therefore, for any fixed $\varepsilon > 0$ and all sufficiently large n (so that $1/n \leq \varepsilon$),

$$t_{\mathrm{mix}}(\varepsilon) \;\leq\; T \;=\; 2\,n^3 \left\lceil \log_2 n \right\rceil. \tag{0.0.2}$$

Context: *L*-reversal chain. Another classical shuffle on S_n is the *L*-reversal chain, which models reversals of contiguous blocks of genes. At each step:

- choose a starting position $i \in [n]$ uniformly;

- choose a block length $\ell \in \{1, \ldots, L\}$ uniformly;

- if $\ell = 1$, do nothing (a hold); otherwise reverse the circular block $(i, i+1, \ldots, i+\ell-1)$, with indices taken modulo n.

This includes the n segments of length 1, for which the move is the identity; hence the chain has hold probability $1/L$, which makes it aperiodic while preserving the uniform stationary distribution. This *L*-reversal chain is a standard toy model for inversion-driven genome rearrangements.

0.0.3 Fourier Mode Mixing-Time Lower Bound for the *L*-Reversal Chain

Summary. We track a single tagged element via a Fourier test function and average over random reversals to obtain the mixing time lower bound.

Consider the *L*-reversal chain on n positions described in the context paragraph above, and view the positions as arranged on the cycle $\mathbb{Z}_n$. Because the motion of a single labeled element is a Markov chain on this cycle, it is natural to analyze it in the discrete Fourier basis $\{e^{2\pi i k x/n}\}_{k=0}^{n-1}$. We focus on the first nontrivial mode $k = 1$, which leads to the following test function. Fix one labeled element and let $\mathrm{pos}_t \in \{0, \ldots, n-1\}$ be its position at time t (indices taken modulo n). Put $\omega = e^{2\pi i/n}$ and define

$$\phi_t := \omega^{\mathrm{pos}_t} \left(= e^{2\pi i \, \mathrm{pos}_t/n}\right).$$

Under the uniform stationary distribution π on permutations, we have $\mathbb{E}_\pi[\phi] = 0$ and $\mathrm{Var}_\pi(\phi) = 1$.

Condition on a chosen length ℓ and start a. Let $m = a + (\ell - 1)/2$ be the midpoint of the block. If the tagged element is not in the block then ϕ does not change. If it is in the block at position $j = \mathrm{pos}_t$, then it is reflected to $j' = 2m - j$.
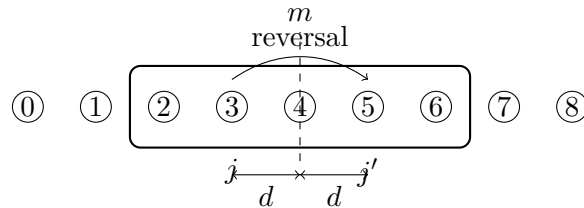

Figure 1: A reversal reflects the position j across the midpoint m of the block, sending it to j'. Both j and j' are at the same distance d from m, so $m - j = d = j' - m$, hence $j' = 2m - j$.

Therefore,

$$\phi_{t+1} = \omega^{j'} = \omega^{2m-j} = \omega^{2m} \overline{\phi_t}. \tag{0.0.3}$$

9

On the event that the chosen block hits the tagged element and has length ℓ, write the offset of the tagged element from the midpoint as

$$r := j - m,$$

so $m = j - r$ (note that $|r|$ is the distance d from Figure 1). Since the start a of the block is chosen uniformly, conditional on a hit the tagged element is equally likely to be in any of the ℓ positions of the block, and hence

$$r \in \left\{ -\frac{\ell - 1}{2}, \ldots, \frac{\ell - 1}{2} \right\}$$

is uniform on this set (for even ℓ, these are half integers; the formulas below are unchanged). Substituting $m = j - r$ into (0.0.3) gives

$$\phi_{t+1} = \omega^{2(j-r)-j} = \omega^j \, \omega^{-2r} = \phi_t \, \omega^{-2r}.$$

Given ϕ_t, the length ℓ, and the event that the block hits the tagged element, the only remaining randomness is the offset r, so to understand the evolution of the Fourier mode we compute the conditional expectation of ϕ_{t+1} by averaging over r:

$$\mathbb{E}[\phi_{t+1} \mid \phi_t, \text{ hit}, \ell] = \phi_t \cdot \frac{1}{\ell} \sum_{r=-(\ell-1)/2}^{(\ell-1)/2} \omega^{-2r}.$$

Let $z = \omega^{-2} = e^{-4\pi i/n}$. Using $\sum_{r=-k}^{k} z^r = \frac{\sin((2k+1)\theta/2)}{\sin(\theta/2)}$ for $z = e^{i\theta}$ with $k = (\ell - 1)/2$ and $\theta = -\frac{4\pi}{n}$ yields the real factor

$$\kappa_\ell := \frac{1}{\ell} \frac{\sin(2\pi\ell/n)}{\sin(2\pi/n)} \in \mathbb{R}.$$

Therefore,

$$\mathbb{E}[\phi_{t+1} \mid \phi_t, \text{ hit}, \ell] = \kappa_\ell \, \phi_t.$$

For a fixed ℓ, the probability that the random block hits the tagged element is $\Pr(\text{hit} \mid \ell) = \ell/n$. If there is no hit, ϕ does not change. Therefore,

$$\mathbb{E}[\phi_{t+1} \mid \phi_t, \ell] = \left(1 - \tfrac{\ell}{n}\right)\phi_t + \tfrac{\ell}{n} \kappa_\ell \, \phi_t = \left(1 - \tfrac{\ell}{n}(1 - \kappa_\ell)\right)\phi_t.$$

Averaging over $\ell \in \{1, \ldots, L\}$ gives

$$\mathbb{E}[\phi_{t+1} \mid \phi_t] = \left(1 - \frac{1}{nL} \sum_{\ell=1}^{L} \ell\,(1 - \kappa_\ell)\right)\phi_t.$$

A Taylor expansion of $\sin$ around 0 shows that there exist absolute constants $0 < c_1 \le c_2$ such that, for all $\ell \le n/4$,

$$c_1 \frac{\ell^2}{n^2} \le 1 - \kappa_\ell \le c_2 \frac{\ell^2}{n^2}. \tag{0.0.4}$$

10

Therefore,

$$\frac{1}{nL}\sum_{\ell=1}^{L}\ell(1-\kappa_\ell) \le \frac{c_2}{n^2} \cdot \frac{1}{nL}\sum_{\ell=1}^{L}\ell^3$$

$$= \frac{c_2}{n^3 L} \cdot \frac{L^2(L+1)^2}{4}$$

$$\le C\frac{L^3}{n^3},$$

for some absolute constant $C > 0$. Consequently,

$$\mathbb{E}[\phi_{t+1} \mid \phi_t] \ge \left(1 - C\frac{L^3}{n^3}\right)\phi_t.$$

Iterating gives

$$\mathbb{E}[\phi_t] \ge \left(1 - C\frac{L^3}{n^3}\right)^t \mathbb{E}[\phi_0],$$

that is, the first Fourier mode does not contract faster than a factor $1 - CL^3/n^3$ per step in expectation.

Recall **Proposition 4.5** (Levin, Peres, and Wilmer 2009, Prop. 4.5)

$$\|\mu - \nu\|_{\mathrm{TV}} = \frac{1}{2}\sup_{\max_x |f(x)| \le 1}\left(\sum_x f(x)\mu(x) - \sum_x f(x)\nu(x)\right). \tag{0.0.5}$$

Apply (0.0.5) with $\mu = \mu_t$ (the law at time t), $\nu = \pi$ (uniform), and the test function $g := \Re(\phi)$. Since $|\phi| = 1$, we have $|g| \le 1$; and by uniformity, $\mathbb{E}_\pi[g] = \Re(\mathbb{E}_\pi[\phi]) = 0$. Therefore,

$$\|\mu_t - \pi\|_{\mathrm{TV}} = \frac{1}{2}\sup_{\max|f|\le 1}\left|\mathbb{E}_{\mu_t}f - \mathbb{E}_\pi f\right| \ge \frac{1}{2}\left|\mathbb{E}_{\mu_t}g - \mathbb{E}_\pi g\right| = \frac{1}{2}\left|\mathbb{E}_{\mu_t}g\right| = \frac{1}{2}\left|\Re\mathbb{E}[\phi_t]\right|.$$

Choosing the start with $\mathbb{E}[\phi_0] = 1$ (e.g. the tag at a fixed position) yields

$$\|\mu_t - \pi\|_{\mathrm{TV}} \ge \frac{1}{2}\left(1 - C\frac{L^3}{n^3}\right)^t.$$

Thus, in order to have $\|\mu_t - \pi\|_{\mathrm{TV}} \le \varepsilon$, it is necessary that

$$\frac{1}{2}\left(1 - C\frac{L^3}{n^3}\right)^t \le \varepsilon \iff \left(1 - C\frac{L^3}{n^3}\right)^t \le 2\varepsilon.$$

Set $x := CL^3/n^3$. Taking logarithms gives

$$t\log(1-x) \le \log(2\varepsilon).$$

We have the general inequality

$$-\log(1-x) \le \frac{x}{1-x}$$

11

for all $x \in (0, 1)$, so for x sufficiently small (and hence for our $x = CL^3/n^3$ when n is large) we obtain
$$-\log(1 - x) \le 2x.$$
Using $\log(2\varepsilon) = -\log\big(1/(2\varepsilon)\big)$, this yields
$$t \ge \frac{\log\big(1/(2\varepsilon)\big)}{-\log(1 - x)} \ge \frac{1}{2x} \log \frac{1}{2\varepsilon} = \frac{n^3}{2CL^3} \log \frac{1}{2\varepsilon}.$$
Absorbing the factor $\log\big(1/(2\varepsilon)\big)$ into a constant $C_\varepsilon := \frac{1}{2C} \log \frac{1}{2\varepsilon}$ that depends only on ε, we obtain the mixing time lower bound
$$t_{\mathrm{mix}}(\varepsilon) \ge C_\varepsilon \frac{n^3}{L^3}.$$
Compared with the $\Omega(n \log n)$ bound of Proposition 16.2, this bound is stronger when $L \lesssim (n^2/\log n)^{1/3}$, and weaker when $L \gtrsim (n^2/\log n)^{1/3}$.

0.0.4 L-Reversal vs. Lazy Adjacent-Transposition (AT) Chain

Summary. We compare the L-reversal chain on S_n with the lazy adjacent–transposition (AT) chain **on the circle**, showing that mixing-time upper bounds for the AT chain transfer (up to constant factors) to the L-reversal chain.

Let $K^{(L)}$ be the transition matrix of the L–reversal chain on S_n, with uniform stationary distribution $\pi(\sigma) = 1/n!$. A length–2 reversal is exactly an adjacent transposition. In this model,
$$p_2 := \Pr\{\ell = 2\} = \frac{1}{L}.$$

Lemma 0.0.1 (Reversibility of the lazy AT chain)**.** *Let $\{s_i : 1 \le i \le n\}$ be the adjacent transpositions on the circle (indices taken modulo n), where s_i swaps the entries at positions i and $i + 1$. The lazy adjacent–transposition (AT) chain on S_n is the Markov chain with transition matrix $K_{\mathrm{AT,lazy}}$ given by*
$$K_{\mathrm{AT,lazy}}(\sigma, \tau) = \begin{cases} \dfrac{1}{2}, & \tau = \sigma, \\ \dfrac{1}{2n}, & \tau = s_i \sigma \text{ for some } i, \\ 0, & \text{otherwise.} \end{cases}$$
Then $K_{\mathrm{AT,lazy}}$ is reversible with respect to the uniform distribution $\pi(\sigma) = 1/n!$ on S_n.

Proof. Fix $\sigma, \tau \in S_n$. If $\tau = s_i \sigma$ for some i, then also $\sigma = s_i \tau$ (since $s_i^{-1} = s_i$), and
$$\pi(\sigma) K_{\mathrm{AT,lazy}}(\sigma, \tau) = \frac{1}{n!} \cdot \frac{1}{2n} = \frac{1}{n!} \cdot \frac{1}{2n} = \pi(\tau) K_{\mathrm{AT,lazy}}(\tau, \sigma).$$
If $\tau = \sigma$, the detailed balance identity reads $\pi(\sigma) \cdot \frac{1}{2} = \pi(\tau) \cdot \frac{1}{2}$, which is trivial. Otherwise $K_{\mathrm{AT,lazy}}(\sigma, \tau) = K_{\mathrm{AT,lazy}}(\tau, \sigma) = 0$. Hence $\pi(\sigma) K_{\mathrm{AT,lazy}}(\sigma, \tau) = \pi(\tau) K_{\mathrm{AT,lazy}}(\tau, \sigma)$ for all σ, τ, and the chain is reversible. $\square$

Lemma 0.0.2 (Reversibility of the L–reversal chain). *Let $K^{(L)}$ be the transition matrix of the L–reversal chain on S_n, with uniform stationary distribution $\pi(\sigma) = 1/n!$. Then $K^{(L)}$ is reversible with respect to π.*

Proof. In one step of the L–reversal chain we choose a length $\ell \in \{1, \ldots, L\}$ and a starting position $i \in \{1, \ldots, n\}$ uniformly, and apply the reversal $\rho_{i,i+\ell-1}$. Thus each admissible reversal $\rho_{i,j}$ is chosen with probability $1/(nL)$, independent of the current state. Fix $\sigma, \tau \in S_n$. If $\tau = \rho_{i,j}\sigma$ for some admissible $\rho_{i,j}$, then $\sigma = \rho_{i,j}\tau$ as well, since $\rho_{i,j}$ is an involution. Hence

$$K^{(L)}(\sigma, \tau) = \frac{1}{nL} = K^{(L)}(\tau, \sigma).$$

If τ is not obtained from σ by a single allowed reversal (including the trivial case $\tau = \sigma$), then the same is true with σ and τ interchanged, so again $K^{(L)}(\sigma, \tau) = K^{(L)}(\tau, \sigma)$. Thus $K^{(L)}$ is a symmetric matrix. Since $\pi(\sigma) = 1/n!$ for all σ, symmetry implies

$$\pi(\sigma)K^{(L)}(\sigma, \tau) = \pi(\tau)K^{(L)}(\tau, \sigma) \qquad \text{for all } \sigma, \tau \in S_n,$$

so the detailed balance condition holds and the L–reversal chain is reversible with respect to π. □

Off–diagonal domination $\Rightarrow$ spectral–gap comparison. Fix $\sigma \in S_n$ and an adjacent transposition s_i on the circle (indices taken modulo n). Conditional on choosing a length–2 move, the starting site is uniform on $\{1, \ldots, n\}$, so the probability of performing this specific swap s_i in one step of the L–reversal chain is

$$\Pr\{\text{do } s_i\} = \Pr\{\ell = 2\}\Pr\{\text{start at } i \mid \ell = 2\} = \Pr\{\ell = 2\}\frac{\Pr\{\text{start at } i, \ell = 2\}}{\Pr\{\ell = 2\}} = p_2 \cdot \frac{1}{n} = \frac{p_2}{n}.$$

Thus the corresponding off–diagonal entries are

$$K^{(L)}(\sigma, s_i\sigma) = \frac{p_2}{n}, \qquad K_{\text{AT,lazy}}(\sigma, s_i\sigma) = \frac{1}{2n},$$

and on every off–diagonal adjacent edge we have

$$\frac{K^{(L)}(\sigma, s_i\sigma)}{K_{\text{AT,lazy}}(\sigma, s_i\sigma)} = \frac{p_2/n}{1/(2n)} = 2p_2 \geq p_2\left(1 - \tfrac{1}{n}\right),$$

and on non–edges both sides are 0. Thus, entrywise for all $\sigma \neq \tau$,

$$K^{(L)}(\sigma, \tau) \geq p_2\left(1 - \tfrac{1}{n}\right)K_{\text{AT,lazy}}(\sigma, \tau). \tag{0.0.6}$$

By **Lemma 13.6**(Levin, Peres, and Wilmer 2009, Lem. 13.6). For a reversible chain with the uniform stationary law $\pi(\sigma) = 1/n!$ we have

$$\mathcal{E}_K(f, f) = \frac{1}{2n!}\sum_{\sigma \neq \tau}\left(f(\sigma) - f(\tau)\right)^2 K(\sigma, \tau),$$

13

so diagonal entries do not contribute.

Using the off–diagonal domination (0.0.6), we therefore obtain the energy comparison

$$\mathcal{E}_{K^{(L)}}(f,f) \geq p_2\left(1 - \tfrac{1}{n}\right)\mathcal{E}_{K_{\mathrm{AT,lazy}}}(f,f) \qquad \text{for all } f. \tag{0.0.7}$$

By **Lemma 13.7** (Levin, Peres, and Wilmer 2009, Lem. 13.7), the spectral gap of a reversible Markov chain with transition matrix P and stationary law π admits the variational representation

$$\gamma(P) = \min_{\substack{f \perp \pi 1 \\ \|f\|_2 = 1}} \mathcal{E}_P(f) = \min_{\substack{f \perp \pi 1 \\ f \neq 0}} \frac{\mathcal{E}_P(f)}{\|f\|_2^2},$$

where $\mathcal{E}_P$ is the corresponding Dirichlet form. Applying Lemma 13.7 to (0.0.7) yields the corresponding spectral-gap comparison

$$\gamma\left(K^{(L)}\right) \geq p_2\left(1 - \frac{1}{n}\right)\gamma(K_{\mathrm{AT,lazy}}).$$

For a reversible, irreducible chain with stationary law π, **Theorem 12.4**(Levin, Peres, and Wilmer 2009, Thm. 12.4) states that

$$t_{\mathrm{mix}}(\varepsilon) \leq \frac{1}{\gamma}\log\left(\frac{1}{\varepsilon\,\pi_{\mathrm{min}}}\right).$$

In our setting π is uniform on S_n, so $\pi_{\mathrm{min}} = 1/n!$. Hence

$$t_{\mathrm{mix}}^{(L)}(\varepsilon) \leq \frac{1}{\gamma\left(K^{(L)}\right)}\log\left(\frac{n!}{\varepsilon}\right). \tag{0.0.8}$$

From the previous step, we obtained the spectral-gap comparison. Plugging that into (0.0.8) gives the final bound

$$t_{\mathrm{mix}}^{(L)}(\varepsilon) \leq \frac{1}{p_2\left(1 - \frac{1}{n}\right)\gamma\left(K_{\mathrm{AT,lazy}}\right)}\log\left(\frac{n!}{\varepsilon}\right). \tag{0.0.9}$$

Circle vs. line: transferring the AT mixing bound. We now compare the lazy AT chain on the circle with the lazy AT chain on the line. On the line we choose $i \in \{1, \ldots, n-1\}$ uniformly and, with probability $1/2$, swap the entries in positions i and $i+1$ (otherwise we stay put); denote its transition matrix by $K_{\mathrm{AT,line}}$. On the circle we choose $i \in \{1, \ldots, n\}$ uniformly and, with probability $1/2$, swap the entries in positions i and $i+1$ modulo n; denote this transition matrix by $K_{\mathrm{AT,circ}}$. If $\tau = s_i\sigma$ is obtained from σ by a single adjacent transposition, then

$$K_{\mathrm{AT,line}}(\sigma, \tau) = \frac{1}{2(n-1)}, \qquad K_{\mathrm{AT,circ}}(\sigma, \tau) = \frac{1}{2n} = \left(1 - \frac{1}{n}\right)\frac{1}{2(n-1)} = \left(1 - \frac{1}{n}\right)K_{\mathrm{AT,line}}(\sigma, \tau).$$

If τ is not adjacent to σ, both transition probabilities are 0. Thus, for every off–diagonal pair $\sigma \neq \tau$ we have,

$$K_{\mathrm{AT,circ}}(\sigma, \tau) = \left(1 - \frac{1}{n}\right)K_{\mathrm{AT,line}}(\sigma, \tau), \tag{0.0.10}$$

By **Lemma 13.6**(Dirichlet form) and (0.0.10), for all f,

$$\mathcal{E}_{K_{\mathrm{AT,circ}}}(f,f) = \frac{1}{2n!} \sum_{\sigma \neq \tau} K_{\mathrm{AT,circ}}(\sigma,\tau) \left(f(\sigma) - f(\tau)\right)^2 = \left(1 - \frac{1}{n}\right) \mathcal{E}_{K_{\mathrm{AT,line}}}(f,f).$$

Hence, by **Lemma 13.7** (Rayleigh–Ritz characterization of the gap),

$$\gamma\left(K_{\mathrm{AT,circ}}\right) = \left(1 - \frac{1}{n}\right) \gamma\left(K_{\mathrm{AT,line}}\right). \tag{0.0.11}$$

Using (0.0.11) for AT (circle vs. line) and **Theorem 12.4** we get

$$t_{\mathrm{mix}}^{\mathrm{AT,circ}}(\varepsilon) \leq \frac{1}{\left(1 - \frac{1}{n}\right) \gamma\left(K_{\mathrm{AT,line}}\right)} \log\left(\frac{n!}{\varepsilon}\right) = \frac{1}{1 - \frac{1}{n}} t_{\mathrm{mix}}^{\mathrm{AT,line}}(\varepsilon). \tag{0.0.12}$$

For the L-reversal chain we already proved in (0.0.9) that

$$t_{\mathrm{mix}}^{(L)}(\varepsilon) \leq \frac{1}{p_2\left(1 - \frac{1}{n}\right)} \frac{1}{\gamma\left(K_{\mathrm{AT,circ}}\right)} \log\left(\frac{n!}{\varepsilon}\right) = \frac{1}{p_2\left(1 - \frac{1}{n}\right)} t_{\mathrm{mix}}^{\mathrm{AT,circ}}(\varepsilon).$$

Combining this with (0.0.12) yields the clean transfer

$$t_{\mathrm{mix}}^{(L)}(\varepsilon) \leq \frac{1}{p_2\left(1 - \frac{1}{n}\right)^2} t_{\mathrm{mix}}^{\mathrm{AT,line}}(\varepsilon).$$

Finally, invoking the AT-on-the-line bound proved earlier (see (0.0.2)), we obtain

$$t_{\mathrm{mix}}^{(L)}(\varepsilon) \leq \frac{2}{p_2\left(1 - \frac{1}{n}\right)^2} n^3 \left\lceil \log_2 n \right\rceil.$$

In particular, for $p_2 = 1/L$,

$$t_{\mathrm{mix}}^{(L)}(\varepsilon) = O\left(L n^3 \log n\right).$$

Context: Weighted _L_-reversal dynamics. In many applications (for example, models of genome rearrangements) it is natural to allow reversals of different lengths to occur with different frequencies. To capture such length–dependent bias, we consider a weighted version of the _L_-reversal chain of §16.2 of _Markov Chains and Mixing Times_: we still work on the circular index set $\{1, \ldots, n\}$ (indices modulo n), but we fix a weight profile $(w_\ell)_{\ell=1}^L$ with $w_\ell \geq 0$ and $\sum_{\ell=1}^L w_\ell = 1$, and choose the block length according to these weights.

0.0.5 A Mixing-Time bound for the Weighted _L_-Reversal Chain

Summary. We now analyse the weighted _L_-reversal chain from the preceding context paragraph. Its transition rule is as follows. Work on the circular index set $\{1, \ldots, n\}$ (indices modulo n). At each step:

1. pick a starting site $i \in \{1, \ldots, n\}$ uniformly;

2. pick a length $\ell \in \{1, \ldots, L\}$ with probability w_ℓ, where $\sum_{\ell=1}^L w_\ell = 1$;

3. if $\ell = 1$ (a length-1 block), _hold_; otherwise reverse the circular block $(i, i+1, \ldots, i+\ell-1)$.

Let $r_{i,\ell}$ denote the length-ℓ block reversal starting at i, and interpret $r_{i,1} = $ id as a hold. Then the transition probability is,

$$\Pr\{\sigma \to \sigma'\} = \sum_{\ell=1}^L \sum_{i=1}^n \Pr\{(i,\ell)\} \, \mathbf{1}\{r_{i,\ell}(\sigma) = \sigma'\} = \frac{1}{n} \sum_{\ell=1}^L w_\ell \sum_{i=1}^n \mathbf{1}\{r_{i,\ell}(\sigma) = \sigma'\},$$

Lemma 0.0.3 (Reversibility)**.** _The chain is reversible with respect to the uniform stationary distribution_ $\pi(\sigma) \equiv 1/n!$.

Proof. Each move $r_{i,\ell}$ is its own inverse and is chosen with probability $(1/n)w_\ell$ independently of the current state. Hence, for any permutations $\sigma, \tau \in S_n$,

$$\pi(\sigma)P(\sigma,\tau) = \frac{1}{n!} \cdot \frac{1}{n} \sum_{\ell=1}^L w_\ell \sum_{i=1}^n \mathbf{1}\{\, r_{i,\ell}(\sigma) = \tau \,\} = \frac{1}{n!} \cdot \frac{1}{n} \sum_{\ell=1}^L w_\ell \sum_{i=1}^n \mathbf{1}\{\, r_{i,\ell}(\tau) = \sigma \,\} = \pi(\tau)P(\tau,\sigma).$$

$\square$

For $\sigma \in S_n$ (viewed on the circle; indices taken modulo n), the edge between positions k and $k+1$ is _conserved_ in σ if the two elements that were at positions k and $k+1$ still end up adjacent after applying σ, i.e.

$$\sigma(k+1) - \sigma(k) \equiv \pm 1 \pmod{n}.$$

Edges and cuts. Place the n genes on a circular chromosome and let an _edge_ be the adjacency between two consecutive genes. Fix an edge e. We say that e is _cut in a given step_ if the chosen reversal block contains exactly one endpoint of e (so the adjacency represented by e is destroyed in that step). In particular, when $\ell = 1$ no edge is cut: an edge may be "touched" by choosing a site next to it, but it remains conserved because no cut occurs.
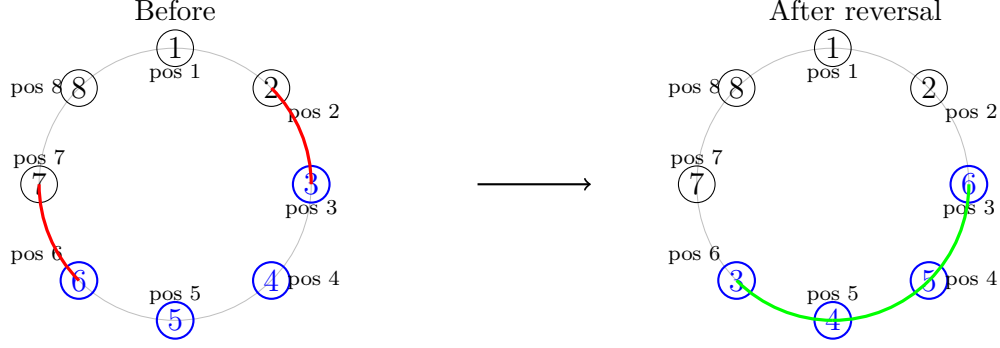
Figure 2: One step of the weighted L–reversal chain on a circle of $n = 8$ sites, starting from the identity permutation. On the left we choose a starting site $i = 3$ and a block of length $\ell = 4$ with probability w_4; the block positions $3, 4, 5, 6$ are highlighted in blue. The red edges show adjacencies that are cut by this reversal. On the right the positions are the same, but the labels on the highlighted arc are reversed; the green edges are the corresponding adjacencies in the new labeling, so these edges are *conserved* by the move.

Lemma 0.0.4. *Fix an edge e on the circle. Then*

$$\Pr\{e \text{ is cut in one step}\} = \frac{2}{n} \sum_{\ell=2}^{L} w_\ell = \frac{2}{n}(1 - w_1),$$

where $w_\ell = \Pr\{$we choose length $\ell\}$ and w_1 is the hold probability.

Proof. Condition on a fixed length $\ell \geq 2$. Let the chosen block be $B(i, \ell) = \{i, i+1, \ldots, i+\ell-1\}$ (indices mod n). The edge $e = (k, k+1)$ is cut *iff* it is one of the two boundary edges of the block. Thus, for any fixed $\ell \geq 2$, there are exactly two starting sites that cut e. Since i is uniform in $\{1, \ldots, n\}$, $\Pr\{e \text{ cut} \mid \ell\} = 2/n$.

Averaging over ℓ with weights w_ℓ (and noting that $\ell = 1$ is a hold and never cuts) gives

$$\Pr\{e \text{ is cut}\} = \sum_{\ell=2}^{L} w_\ell \cdot \frac{2}{n} = \frac{2}{n} \sum_{\ell=2}^{L} w_\ell = \frac{2}{n}(1 - w_1). \qquad \square$$

Let X_t be the number of *conserved* edges after t steps, i.e., the number of adjacencies that have not been cut up to and including time t. By definition $X_0 = n$.

Lemma 0.0.5. *With $q = \sum_{\ell=2}^{L} w_\ell = 1 - w_1$ (the non-hold probability), we have for all $t \geq 0$,*

$$\mathbf{E}[X_{t+1} \mid X_t] \geq X_t - 2q \qquad \text{and hence} \qquad \mathbf{E}[X_t] \geq n - 2qt.$$

Proof. During step $t+1$ one of two things happens.

- With probability w_1 (a hold), no edge is cut and $X_{t+1} = X_t$.

- With probability q (a non-hold), a single reversal is performed. This can cut at most the two boundary edges of the chosen block, so the number of *new* cuts is at most 2, and therefore $X_{t+1} \geq X_t - 2$.

Taking conditional expectation gives $\mathbf{E}[X_{t+1} \mid X_t] \geq w_1 X_t + q(X_t - 2) = X_t - 2q$. Iterating from $X_0 = n$ yields $\mathbf{E}[X_t] \geq n - 2qt$. $\qquad \square$

Doob martingale for $(X_s)_{s=0}^t$. Fix a time horizon t and let $(\mathcal{F}_s)_{s=0}^t$ be the natural filtration. Define

$$M_s := \mathbb{E}[X_t \mid \mathcal{F}_s], \qquad s = 0, 1, \ldots, t.$$

Lemma 0.0.6. *The process* $(M_s)_{s=0}^t$ *is a martingale. In particular,* $M_t = X_t$ *and* $M_0 = \mathbb{E}[X_t]$.

Proof. For each s, M_s is $\mathcal{F}_s$-measurable and $\mathbb{E}|M_s| \leq \mathbb{E}|X_t| < \infty$. By the tower property,

$$\mathbb{E}[M_s \mid \mathcal{F}_{s-1}] = \mathbb{E}\Big(\mathbb{E}[X_t \mid \mathcal{F}_s] \,\big|\, \mathcal{F}_{s-1}\Big) = \mathbb{E}[X_t \mid \mathcal{F}_{s-1}] = M_{s-1},$$

so (M_s) satisfies the martingale property. Moreover, $M_t = \mathbb{E}[X_t \mid \mathcal{F}_t] = X_t$ (measurability) and, since $\mathcal{F}_0$ is trivial, $M_0 = \mathbb{E}[X_t \mid \mathcal{F}_0] = \mathbb{E}[X_t]$. $\qquad\square$

Hence

$$X_t - \mathbb{E}[X_t] \;=\; M_t - M_0.$$

Define the martingale increments

$$\Delta M_s := M_s - M_{s-1}, \qquad s = 1, \ldots, t.$$

Then $(\Delta M_s)_{s=1}^t$ is a martingale–difference sequence and

$$M_t - M_0 = \sum_{s=1}^t (M_s - M_{s-1}) = \sum_{s=1}^t \Delta M_s.$$

In our setting, one step of the chain changes at most two adjacencies, so each increment is uniformly bounded:

$$|\Delta M_s| = |M_s - M_{s-1}| \leq 2, \qquad s = 1, \ldots, t.$$

Thus we may apply the Azuma–Hoeffding inequality to the sum $\sum_{s=1}^t \Delta M_s = M_t - M_0$ with bounds $B_s = 2$, giving

$$\Pr\big\{X_t - \mathbb{E}[X_t] \leq -u\big\} = \Pr\{M_t - M_0 \leq -u\} \;\leq\; \exp\left(-\frac{u^2}{2\sum_{s=1}^t B_s^2}\right) \;=\; \exp\left(-\frac{u^2}{8t}\right), \qquad u > 0. \tag{0.0.13}$$

By Lemma 0.0.5, we may choose $c \in (0, \frac{1}{4})$ and set $t = cn/q$ so that

$$u \;:=\; \mathbb{E}[X_t] - \frac{n}{2} \;\geq\; n - 2qt - \frac{n}{2} \;=\; \left(\frac{1}{2} - 2c\right)n \;>\; 0.$$

Now, whenever $X_t < n/2$ we have

$$X_t - \mathbb{E}[X_t] \;<\; \frac{n}{2} - \mathbb{E}[X_t] \;=\; -u,$$

so the event $\{X_t < n/2\}$ is contained in $\{X_t - \mathbb{E}[X_t] \leq -u\}$. Hence

$$\Pr_{\mathrm{id}}(X_t < n/2) \;\leq\; \Pr_{\mathrm{id}}\big\{X_t - \mathbb{E}[X_t] \leq -u\big\}. \tag{0.0.14}$$

18

Define

$$A := \left\{ \sigma \in S_n : X(\sigma) \ge n/2 \right\}, \qquad \text{so that} \qquad \{X_t \ge n/2\} = \{\sigma_t \in A\}.$$

Combining $(0.0.13)$ and $(0.0.14)$, we obtain

$$\Pr_{\mathrm{id}}(\sigma_t \notin A) = \Pr_{\mathrm{id}}(X_t < n/2) \le \exp\left(-\frac{(\frac{1}{2} - 2c)^2}{8c}\, qn\right).$$

Under stationarity π (uniform on S_n), each edge is conserved with probability $2/(n-1)$. Indeed, fix k and look at the pair $(\sigma(k), \sigma(k+1))$ under the uniform measure on S_n: once $\sigma(k)$ is chosen, $\sigma(k+1)$ is uniform over the remaining $n-1$ values, and exactly two of them are neighbours of $\sigma(k)$ on the n–cycle. Writing $X = \sum_{k=1}^{n} I_k$ with $I_k = \mathbf{1}\{\text{edge } k \text{ conserved}\}$ and using linearity of expectation, we obtain

$$\mathbb{E}_\pi[X] = \sum_{k=1}^{n} \Pr_\pi(\text{edge } k \text{ conserved}) = n \cdot \frac{2}{n-1} = 2 + o(1).$$

Hence, by Markov's inequality,

$$\pi(A) = \Pr_\pi\{X \ge n/2\} \le \frac{\mathbb{E}_\pi[X]}{n/2} \le \frac{4}{n}(1 + o(1)).$$

Therefore, the total variation distance at time t (from the identity) satisfies

$$\left\| P^t(\mathrm{id}, \cdot) - \pi \right\|_{\mathrm{TV}} \ge \left| \Pr_{\mathrm{id}}(\sigma_t \in A) - \pi(A) \right| \ge 1 - \exp\left(-\frac{(\frac{1}{2} - 2c)^2}{8c}\, qn\right) - \frac{4}{n}.$$

Take

$$c = \frac{1}{8} \qquad \Longrightarrow \qquad \frac{\left(\frac{1}{2} - 2c\right)^2}{8c} = \frac{(1/4)^2}{8 \cdot (1/8)} = \frac{1}{16}.$$

For $t = \dfrac{c\,n}{q} = \dfrac{n}{16\,q}$ our previous estimate gives

$$\left\| P^t(\mathrm{id}, \cdot) - \pi \right\|_{\mathrm{TV}} \ge 1 - \exp\left(-\tfrac{1}{16}\, qn\right) - \frac{4}{n}.$$

For all n large enough so that $qn \ge 16 \log 8$ and $n \ge 32$, the right–hand side is at least $1 - \frac{1}{8} - \frac{1}{8} = \frac{3}{4}$. Hence,

$$\left\| P^t(\mathrm{id}, \cdot) - \pi \right\|_{\mathrm{TV}} \ge \frac{3}{4}$$

this implies that at time $t = n/(16q)$ the chain is still at distance at least $3/4$ from stationarity. Therefore,

$$t_{\mathrm{mix}}\left(\tfrac{1}{4}\right) \ge \frac{n}{16q}. \tag{0.0.15}$$

Bibliography

Levin, David A., Yuval Peres, and Elizabeth L. Wilmer (2009). *Markov Chains and Mixing Times*. 2nd ed. American Mathematical Society.