DOI: 10.1214/ECP.v8-1071

Elect. Comm. in Probab. **8** (2003) 77–85

ELECTRONIC COMMUNICATIONS in PROBABILITY

MIXING TIME OF THE RUDVALIS SHUFFLE

DAVID BRUCE WILSON
Microsoft Research, One Microsoft Way, Redmond, WA 98052, U.S.A.
email: dbwilson@microsoft.com

Submitted November 1, 2002, *accepted in final form* January 16, 2003

AMS 2000 Subject classification: 60J10, 60C05
Keywords: Markov chain, card shuffling, mixing time

Abstract

We extend a technique for lower-bounding the mixing time of card-shuffling Markov chains, and use it to bound the mixing time of the Rudvalis Markov chain, as well as two variants considered by Diaconis and Saloff-Coste. We show that in each case $\Theta(n^3 \log n)$ shuffles are required for the permutation to randomize, which matches (up to constants) previously known upper bounds. In contrast, for the two variants, the mixing time of an individual card is only $\Theta(n^2)$ shuffles.

1. Introduction

In earlier work (Wilson, 2001) we derived upper and lower bounds on the mixing time of a variety of Markov chains, including Markov chains on lozenge tilings, card shuffling, and exclusion processes. The mixing time of a Markov chain is the time it takes to approach its stationary distribution, which is often measured in total variation distance (defined below). In this article we focus on the method for lower bounding the mixing time, and extend its applicability to the Rudvalis card shuffling Markov chain (defined below) and related shuffles.
Let P_x^{*t} denote the distribution of the Markov chain started in state x after it is run for t steps, and let μ denote the stationary distribution of the Markov chain. The total variation distance between distributions P^{*t} and μ is defined by

$$\left\|P_x^{*t} - \mu\right\|_{\text{TV}} = \max_A \left|P_x^{*t}(A) - \mu(A)\right| = \frac{1}{2}\sum_y \left|P_x^{*t}(y) - \mu(y)\right| = \frac{1}{2}\left\|P^{*t} - \mu\right\|_1,$$

and the mixing time is the time it takes for $\max_x \|P_x^{*t} - \mu\|_{\text{TV}}$ to become small, say smaller than ε. See (Aldous and Fill, 2005; Diaconis, 1988, 1996) for further background.
Arunas Rudvalis proposed the following shuffle: with probability 1/2 move the top card to the bottom of the deck, and with probability 1/2 move it to the second position from the bottom. Hildebrand (1990) showed that the Rudvalis shuffle mixes in $O(n^3 \log n)$ time. Diaconis and Saloff-Coste (1995) studied a variation, the shift-or-swap shuffle, which at each step either moves the top card to the bottom of the deck or exchanges the top two cards, each move with probability 1/2. Diaconis and Saloff-Coste (1993) also studied a symmetrized version of the Rudvalis shuffle, which at each step does one of four moves each with probability 1/4: move top card to bottom, move bottom card to top, exchange top two cards, or do nothing. In each

case a $O(n^3 \log n)$ upper bound on the mixing time is known, but order $n^3 \log n$ lower bounds were not known.

To lower bound the mixing time, one finds a set A of states such that $P^{*t}(A)$ is close to 1 and $\mu(A)$ is close to 0. The approach taken in (Wilson, 2001) uses an eigenvector Φ of the Markov chain. If X_t denotes the state of the Markov chain at time t, then $\mathbb{E}[\Phi(X_{t+1}) \mid X_t] = \lambda\Phi(X_t)$. To obtain a good lower bound, we need $\lambda < 1$ but $\lambda \approx 1$. Since $\lambda < 1$, in stationarity $\mathbb{E}[\Phi(X)] = 0$, but since $\lambda \approx 1$, it takes a long time before $\mathbb{E}[\Phi(X_t)] \approx 0$. If furthermore the eigenvector is "smooth" in the sense that $\mathbb{E}[|\Phi(X_{t+1}) - \Phi(X_t)|^2 \mid X_t]$ is never large, then we can bound the variance $\Phi(X_t)$, showing that it is with high probability confined to a small interval about its expected value. Then provided that $\mathbb{E}[\Phi(X_t)]$ is large enough, we can reliably distinguish $\Phi(X_t)$ from $\Phi(X)$ in stationarity, which implies that the Markov chain has not yet mixed by time t. Saloff-Coste (2002) gives an exposition of this and related ideas.

For the Rudvalis card shuffling Markov chain and its variants, there are a few difficulties when directly applying this approach to lower bound the mixing time. The eigenvectors that one wants to use are complex-valued rather than real-valued, and $\Phi(X_t)$ is no longer confined to a small interval around $\mathbb{E}[\Phi(X_t)]$. Instead what happens is that $\Phi(X_t)$ is with high probability confined to a narrow annulus centered at 0. $\mathbb{E}[\Phi(X_t)]$ becomes too small too quickly, and $\mathrm{Var}[\Phi(X_t)]$ remains too large to be useful.

To lower bound the mixing time we want to in effect work with $|\Phi(X_t)|$ and forget about $\arg\Phi(X_t)$. To do this we start by lifting the Markov chain to a larger state space, and let us denote the state at time t of the lifted chain by (X_t, Y_t). (The mixing time of the lifted Markov chain will upper bound the mixing time of the original chain, so at the outset it is not clear that we can lower bound the mixing time of the original chain by considering its lifted version.) We find an eigenvector Ψ on the lifted chain such that for all x, y_1 and y_2, $|\Psi(x, y_1)| = |\Psi(x, y_2)|$, so that $|\Psi(X_t)|$ is well-defined. If we show that $\Psi(X_t, Y_t)$ is with high probability close to $\mathbb{E}[\Psi(X_t, Y_t)]$, which in turn is far from 0, it will follow that $|\Psi(X_t)|$ is with high probability confined to a small interval far from 0, making it statistically distinguishable from $|\Psi(X)|$ in stationarity, implying that the Markov chain has not mixed by time t.

In the following sections we carry out these ideas to obtain lower bounds on the mixing time that match (to within constants) the previously obtained upper bounds. Specifically, we show

Theorem 1. *For any fixed $\varepsilon > 0$, after $\frac{1-o(1)}{8\pi^2} n^3 \log n$ shuffles of the Rudvalis shuffle, $\frac{1-o(1)}{2\pi^2} n^3 \log n$ shuffles of the shift-or-swap shuffle, or $\frac{1-o(1)}{\pi^2} n^3 \log n$ shuffles for the symmetrized Rudvalis shuffle, the distribution of the state of the deck has variation distance $\geq 1 - \varepsilon$ from uniformity.*

2. Lifting the Shuffles

When the top card is placed at the bottom of the deck, the position of any given card is cyclically shifted left, so we will call this move "shift-left", and similarly "shift-right" is the move which places the bottom card on top of the deck. The move which exchanges the top and bottom cards will be called "swap", and the move which does nothing will be called "hold". Thus the moves of the Rudvalis Markov chain are "shift-left" and "swap & shift-left", while the moves of the variation considered by Diaconis and Saloff-Coste are "shift-left" and "swap", and the moves of the symmetrized version are "shift-left", "shift-right", "swap", and "hold".

The state X_t of the Markov chain is the permutation giving the order of the cards at time t. Let $\diamondsuit$ (the diamond-suit symbol) denote a particular card of interest, and let $X_t(\diamondsuit)$ denote the location of card $\diamondsuit$ within the deck at time t, where the positions are numbered from 1 to n starting from the top of the deck. When the Markov chain does a shift-left or shift-right, while the position of a card $\diamondsuit$ will change, all the cards get moved together, so it does not have such a large randomizing effect on the permutation. We will track the position of card $\diamondsuit$, but we should also track the amount of shifting. So when we lift the Markov chain to (X_t, Y_t), the lifted Markov chain will also keep track of

$$Y_t = \#\text{ shift-left's} - \#\text{ shift-right's} \bmod n.$$

For the Rudvalis shuffle $Y_t = t \bmod n$ deterministically, whereas for the other two variations, Y_t will be a random number between 1 and n which approaches uniformity in $O(n^2)$ time. Recall that we need an eigenvector Ψ of the lifted chain (X_t, Y_t) such that $|\Psi(X_t, Y_t)|$ is a function of X_t alone. For a given card $\diamondsuit$, let

$$\Psi_\diamondsuit(X_t, Y_t) = v(X_t(\diamondsuit)) \exp(Z_t(\diamondsuit) 2\pi i/n),$$

where

$$Z_t(\diamondsuit) = X_t(\diamondsuit) - X_0(\diamondsuit) + Y_t \bmod n,$$

and $v()$ is a function, to be determined later, which makes $\Psi_\diamondsuit$ an eigenvector. Initially $Z_0(\diamondsuit) = 0$, and the only time that the $Z_t(\diamondsuit)$ changes is when the card $\diamondsuit$ gets transposed. The dynamics of $(X_t(\diamondsuit), Z_t(\diamondsuit))$ (mod n) are summarized by

$$(X_{t+1}(\diamondsuit), Z_{t+1}(\diamondsuit)) = \begin{cases} (X_t(\diamondsuit), Z_t(\diamondsuit)) & \text{if move was “hold”} \\ (X_t(\diamondsuit) - 1, Z_t(\diamondsuit)) & \text{if move was “shift-left”} \\ (X_t(\diamondsuit) + 1, Z_t(\diamondsuit)) & \text{if move was “shift-right”} \\ (X_t(\diamondsuit) - 1, Z_t(\diamondsuit) - 1) & \text{if move was “swap” and } X_t(\diamondsuit) = 1 \\ (X_t(\diamondsuit) + 1, Z_t(\diamondsuit) + 1) & \text{if move was “swap” and } X_t(\diamondsuit) = n \\ (X_t(\diamondsuit), Z_t(\diamondsuit)) & \text{if move was “swap” and } \diamondsuit \text{ elsewhere} \end{cases}$$

We define

$$\Psi(X_t, Y_t) = \sum_{\diamondsuit=1}^{n} \Psi_\diamondsuit(X_t, Y_t).$$

If we increment y while holding x fixed, then $\Psi(x, y)$ gets multiplied by the phase factor $\exp(2\pi i/n)$, so we have an eigenvector satisfying our requirement that $|\Psi(X_t, Y_t)|$ be a function of X_t alone.

3. The Lower Bound Lemma

The lower bounding lemma that we shall use is similar to Lemma 4 of (Wilson, 2001), but with the modifications described in the introduction. Saloff-Coste (2002) also gives a generalization of Lemma 4 from (Wilson, 2001) that may be used when the eigenvalues are complex, but the extension below seems to be better suited for the shuffles considered here.

Lemma 2. *Suppose that a Markov chain X_t has a lifting (X_t, Y_t), and that Ψ is an eigenfunction of the lifted Markov chain: $\mathbb{E}[\Psi(X_{t+1}, Y_{t+1}) \mid (X_t, Y_t)] = \lambda \Psi(X_t, Y_t)$. Suppose that $|\Psi(x, y)|$ is a function of x alone, $|\lambda| < 1$, $\Re(\lambda) \geq 1/2$, and that we have an upper bound R on $\mathbb{E}[|\Psi(X_{t+1}, Y_{t+1}) - \Psi(X_t, Y_t)|^2 \mid (X_t, Y_t)]$. Let $\gamma = 1 - \Re(\lambda)$. Then when the number of Markov chain steps t is bounded by*

$$t \leq \frac{\log \Psi_{\max} + \frac{1}{2} \log \frac{\gamma \varepsilon}{4R}}{-\log(1 - \gamma)},$$

the variation distance of X_t (the state of the original Markov chain) from stationarity is at least $1 - \varepsilon$.

The proof of this modified lemma is similar to the proof of Lemma 4 in (Wilson, 2001), but for the reader's convenience we give the modified proof. In the following sections we determine the functions $v()$ for the Markov chains which give the requisite eigenfunction Ψ, and then use Lemma 2 to obtain the mixing time bounds stated in Theorem 1.

Proof of Lemma 2. Let $\Psi_t = \Psi(X_t, Y_t)$, and $\Delta\Psi = \Psi_{t+1} - \Psi_t$. By induction

$$\mathbb{E}[\Psi_t \mid (X_0, Y_0)] = \Psi_0 \lambda^t.$$

By our assumptions on λ, in equilibrium $\mathbb{E}[\Psi] = 0$.
We have $\mathbb{E}[\Delta\Psi \mid (X_t, Y_t)] = (\lambda - 1)\Psi_t$ and

$$\begin{aligned} \Psi_{t+1}\Psi_{t+1}^* &= \Psi_t\Psi_t^* + \Psi_t\Delta\Psi^* + \Psi_t^*\Delta\Psi + |\Delta\Psi|^2 \\ \mathbb{E}[\Psi_{t+1}\Psi_{t+1}^* \mid (X_t, Y_t)] &= \Psi_t\Psi_t^*[1 + (\lambda-1)^* + (\lambda - 1)] + \mathbb{E}[|\Delta\Psi|^2 \mid X_t] \\ &\leq \Psi_t\Psi_t^*[2\Re(\lambda) - 1] + R \end{aligned}$$

and so by induction,

$$\mathbb{E}[\Psi_t\Psi_t^*] \leq \Psi_0\Psi_0^*[2\Re(\lambda) - 1]^t + \frac{R}{2 - 2\Re(\lambda)},$$

then subtracting $\mathbb{E}[\Psi_t]\mathbb{E}[\Psi_t]^*$,

$$\operatorname{Var}[\Psi_t] \leq \Psi_0\Psi_0^* \left[[2\Re(\lambda) - 1]^t - (\lambda\lambda^*)^t\right] + \frac{R}{2 - 2\Re(\lambda)}.$$

Since $(1-\lambda)(1-\lambda^*) \geq 0$, we have $\lambda\lambda^* \geq 2\Re(\lambda) - 1$, and by assumption $2\Re(\lambda) - 1 \geq 0$. Hence $(\lambda\lambda^*)^t \geq [2\Re(\lambda) - 1]^t$, and we have for each t

$$\operatorname{Var}[\Psi_t] \leq \frac{R}{2 - 2\Re(\lambda)} = \frac{R}{2\gamma}.$$

From Chebychev's inequality,

$$\Pr\left[|\Psi_t - \mathbb{E}[\Psi_t]| \geq \sqrt{R/(2\gamma\varepsilon)}\right] \leq \varepsilon.$$

As $\mathbb{E}[\Psi_\infty] = 0$, if $\mathbb{E}[\Psi_t] \geq \sqrt{4R/(\gamma\varepsilon)}$, then the probability that $|\Psi_t|$ deviates below $\sqrt{R/(\gamma\varepsilon)}$ is at most $\varepsilon/2$, and the probability that $|\Psi|$ in stationarity deviates above this threshold is at most $\varepsilon/2$, so the variation distance between the distribution at time t and stationarity must be at least $1 - \varepsilon$. If we take the initial state to be the one maximizing Ψ_0, then

$$\mathbb{E}[|\Psi_t|] = |\Psi_{\max}||\lambda|^t \geq |\Psi_{\max}|(\Re(\lambda))^t = |\Psi_{\max}|(1-\gamma)^t \geq \sqrt{4R/(\gamma\varepsilon)}$$

when

$$t \leq \frac{\log\left[\Psi_{\max} \div \sqrt{\frac{4R}{\gamma\varepsilon}}\right]}{-\log(1-\gamma)}. \qquad \square$$

4. The Rudvalis Shuffle

The first shuffle we consider is the original Rudvalis Markov chain. It will be instructive to consider a slight generalization, where the swap & shift-left move takes place with probability p, and the shift-left move takes place with probability $1-p$. We shall assume that $0 < p < 1$ and that p is independent of n. The particular values of p that we are interested in are $p = 1/2$ (for the original Rudvalis chain) and $p = 1/3$.

We need to find an eigenvector for the random walk that a single card takes under this shuffle. We remark that this random walk is similar in nature to (but distinct from) a class of random walks, known as daisy chains, for which Wilmer (1999) obtained eigenvalues and eigenvectors. From other work of Wilmer (2002), it readily follows that the position of a single card takes order n^3 steps to randomize, and that the precise asymptotic distance from stationarity of the card's position after cn^3 shuffles is given by an explicit expression involving theta functions.

Lemma 3. *The random walk followed by a card $\diamondsuit$ under the lifted Rudvalis shuffle has an eigenvector of the form*

$$\Psi_\diamondsuit(x,z) = v(x)e^{2\pi i z/n}$$

where $v(x)$ is the x^{th} number in the list

$$\lambda^{n-2},\dots,\lambda^2,\lambda,1,\chi\ ,$$

the eigenvalue is

$$\lambda = 1 - \frac{p}{1-p}\frac{4\pi^2}{n^3} + O(1/n^4),$$

and

$$\chi = 1 + \frac{p}{1-p}\frac{2\pi i}{n} + O(1/n^2).$$

Proof. Let $w = \exp(2\pi i/n)$. If at time t card $\diamondsuit$ is in any location between 2 and $n-1$, then

$$\Psi_\diamondsuit(X_{t+1},Y_{t+1}) = \lambda\Psi_\diamondsuit(X_t,Y_t)$$

deterministically. To ensure that

$$\mathbb{E}[\Psi_\diamondsuit(X_{t+1},Y_{t+1}) \mid (X_t,Y_t)] = \lambda\Psi_\diamondsuit(X_t,Y_t)$$

when $X_t(\diamondsuit) = 1$, we require

$$pw^{-1} + (1-p)\chi = \lambda^{n-1}$$
$$\chi = \frac{\lambda^{n-1} - pw^{-1}}{1-p},$$

and for when $X_t(\diamondsuit) = n$ we need

$$pw\chi + (1-p) = \lambda\chi$$
$$\chi = \frac{1-p}{\lambda - pw}.$$

Given these two equations, $\Psi_\diamondsuit$ will be an eigenvector with eigenvalue λ. Thus,

$$f(\lambda) = \lambda^n - pw\lambda^{n-1} - pw^{-1}\lambda - 1 + 2p = 0.$$

To identify a root of this polynomial, we use Newton's method: $z_{k+1} = z_k - f(z_k)/f'(z_k)$, starting with $z_0 = 1$. By Taylor's theorem,

$$|f(z_{k+1})| \le \frac{1}{2}\max_{0\le u\le 1}|f''(uz_k + (1-u)z_{k+1})| \times \left|\frac{f(z_k)}{f'(z_k)}\right|^2.$$

If $|z-1| \le 1/n^2$, then $f'(z) = (1-p)n + O(1)$ and $f''(z) = (1-p)n^2 + O(n)$. Consequently, if $|z_k - 1| \le 1/n^2$ and $|z_{k+1} - 1| \le 1/n^2$, then

$$|f(z_{k+1})| \le \frac{1+O(1/n)}{2} \frac{1}{1-p} |f(z_k)|^2.$$

Since $f(z_0) = p(2 - w - w^{-1}) = p4\pi^2/n^2 + O(1/n^4)$, for large enough n we have by induction that $|f(z_k)| \le (1-p)(p/(1-p)4\pi^2/n^2)^{2^k}$, $|z_{k+1} - z_k| \le (p/(1-p)4\pi^2/n^2)^{2^k}/(n+O(1))$, and $|z_{k+1} - z_0| \le O(1/n^3)$. Thus, for large enough n, the sequence $z_0, z_1, z_2, \ldots$ converges to a point λ, which by continuity, must be a zero of f. Since $z_1 = 1 - p/(1-p)4\pi^2/n^3 + O(1/n^4)$ and $|\lambda - z_1| = O(1/n^5)$, we conclude that the polynomial f has a root at

$$\lambda = 1 - \frac{p}{1-p}\frac{4\pi^2}{n^3} + O(1/n^4). \qquad \square$$

It is easy to check that $\Psi_{\max} = n + O(1/n)$. Next we evaluate R for this eigenvector.

$$\frac{\Psi_\diamondsuit(X_{t+1}, Y_{t+1}) - \Psi_\diamondsuit(X_t, Y_t)}{w^{Z_t(\diamondsuit)}} = \begin{cases} (\lambda - 1)\lambda^{X_t(\diamondsuit)} = O(1/n^3) & \text{if } 2 \le X_t(\diamondsuit) \le n-1 \\ \chi - \lambda^{n-2} = O(1/n) & \text{if } X_t(\diamondsuit) = 1, \text{ shift-left} \\ w^{-1} - \lambda^{n-2} = O(1/n) & \text{if } X_t(\diamondsuit) = 1, \text{ swap \& shift-left} \\ 1 - \chi = O(1/n) & \text{if } X_t(\diamondsuit) = n, \text{ shift-left} \\ w\chi - \chi = O(1/n) & \text{if } X_t(\diamondsuit) = n, \text{ swap \& shift-left} \end{cases}$$

Adding up these contributions over the various cards $\diamondsuit$, we find

$$|\Psi(X_{t+1}, Y_{t+1}) - \Psi(X_t, Y_t)| \le O(1/n)$$
$$R = \mathbb{E}[|\Psi(X_{t+1}, Y_{t+1}) - \Psi(X_t, Y_t)|^2 \mid (X_t, Y_t)] \le O(1/n^2).$$

Plugging λ, $\Psi_{\max}$, and R into the Lemma 2 gives, for fixed values of ε, a mixing time lower bound of

$$(1 - o(1)) \frac{1-p}{p} \frac{1}{8\pi^2} n^3 \log n.$$

5. The Shift-or-Swap Shuffle

At this point there are two ways we can approach the shift-or-swap shuffle. We can either take a direct approach in the same manner as in the previous section, or we can do a comparison with the Rudvalis shuffle with $p = 1/3$.

If we take the direct approach, then we let $v(x)$ denote the x^{th} element of the list

$$(2\lambda - 1)^{n-2}, \ldots, (2\lambda - 1)^2, 2\lambda - 1, 1, \chi.$$

The constraints on χ are

$$\chi = \frac{2\lambda}{1 + w^{-1}} (2\lambda - 1)^{n-2}$$

and

$$\chi = \frac{1 + w(2\lambda - 1)^{n-2}}{2\lambda}.$$

As in section 4, we solve for λ and find that $\lambda = 1 - (1 + o(1))\pi^2/n^3$, compute $\Psi_{\max} = \Theta(n)$ and $R = O(1/n^2)$, and obtain the mixing time lower bound of $\frac{1-o(1)}{2\pi^2} n^3 \log n$ shuffles.

Alternatively, we can couple the shift-or-swap shuffle with the Rudvalis shuffle. Whenever the shift-or-swap shuffle makes a shift, the number of swap's since the previous shift will be odd with probability 1/3. If it is odd, then this is equivalent to a swap-&-shift-left move, and if it is even, then it is equivalent to a shift-left move. This explains why we were interested in the case $p = 1/3$ in the previous section. After t steps, with high probability $(1 + o(1))t/2$ shift moves occured, which means that the state of the deck is what it would be after $(1 + o(1))t/2$ Rudvalis shuffles (with $p = 1/3$), possibly with an extra swap move. The lower bound for the shift-or-swap shuffle does not follow from the lower bound itself for the Rudvalis shuffle, but it does follow from what we showed about $|\Psi|$ for the Rudvalis shuffle.

6. Symmetrized Version of the Rudvalis Shuffle

When analyzing the symmetrized version of the Rudvalis shuffle, it will be convenient to have symmetric coordinates, so we re-index the card locations to run from $-(n-1)/2$ up to $(n-1)/2$, and the swaps occur at locations $-(n-1)/2$ and $(n-1)/2$.

Lemma 4. *The random walk followed by a card $\diamondsuit$ under the lifted symmetrized Rudvalis shuffle has an eigenvector of the form*

$$\Psi_\diamondsuit(x, z) = v(x) e^{2\pi i z/n}$$

where

$$\begin{aligned} v(x) = \frac{1+\delta}{2} e^{i\theta x} + \frac{1-\delta}{2} e^{-i\theta x} &= \cos(\theta x) + i\delta \sin(\theta x), \\ \theta &= (1 + o(1))\sqrt{2}\pi n^{-3/2}, \\ \delta &= (1 + o(1)) \frac{1}{\sqrt{2} n^{1/2}}, \end{aligned}$$

and both δ and θ are real. The eigenvalue is

$$\lambda = \frac{1 + \cos\theta}{2} = 1 - \frac{\pi^2 + o(1)}{2} n^{-3}.$$

Proof. When $x \neq \pm(n-1)/2$, we can readily compute the eigenvalue λ to be

$$\begin{aligned} \lambda &= \frac{\frac{1}{4}v(x+1) + \frac{1}{2}v(x) + \frac{1}{4}v(x-1)}{v(x)} \\ &= \frac{1}{2} \frac{\cos(\theta x)\cos\theta + i\delta\sin(\theta x)\cos\theta}{\cos(\theta x) + i\delta \sin(\theta x)} + \frac{1}{2} \\ &= \frac{1 + \cos\theta}{2}. \end{aligned}$$

In order for our guessed eigenvector to be correct, there is also a constraint at $x = (n-1)/2$:

$$\begin{aligned} \frac{v(\frac{n-1}{2})}{4} + \frac{v(\frac{n-3}{2})}{4} + (1+w)\frac{v(-\frac{n-1}{2})}{4} &= \lambda = \frac{v(\frac{n-1}{2})}{2} + \frac{v(\frac{n-3}{2})}{4} + \frac{v(\frac{n+1}{2})}{4} \\ (1+w)v(-(n-1)/2) &= v((n-1)/2) + v((n+1)/2) \\ (1+w)(1+\delta)e^{-i\theta(n-1)/2} + (1+w)(1-\delta)e^{i\theta(n-1)/2} &= (1+\delta)e^{i\theta(n-1)/2} + (1-\delta)e^{-i\theta(n-1)/2} + \\ &\quad\; (1+\delta)e^{i\theta(n+1)/2} + (1-\delta)e^{-i\theta(n+1)/2} \\ (w + 2\delta + w\delta)e^{-i\theta(n-1)/2} + (w - 2\delta - w\delta)e^{i\theta(n-1)/2} &= (1+\delta)e^{i\theta(n+1)/2} + (1-\delta)e^{-i\theta(n+1)/2} \\ w\cos\tfrac{\theta(n-1)}{2} - \cos\tfrac{\theta(n+1)}{2} &= \delta\left[(2+w)i\sin\tfrac{\theta(n-1)}{2} + i\sin\tfrac{\theta(n+1)}{2}\right]. \end{aligned}$$

The corresponding constraint at $x = -(n-1)/2$ is obtained by replacing w with $1/w$ and replacing θ with $-\theta$. Since these substitutions give the complex-conjugate of the above equation, the constraints at $x = \pm(n-1)/2$ are equivalent.

Equating the real parts of this equation gives

$$\delta = \frac{\cos\frac{2\pi}{n}\cos\frac{\theta(n-1)}{2} - \cos\frac{\theta(n+1)}{2}}{-\sin\frac{2\pi}{n}\sin\frac{\theta(n-1)}{2}},$$

and equating the imaginary parts gives

$$\delta = \frac{\sin\frac{2\pi}{n}\cos\frac{\theta(n-1)}{2}}{2\sin\frac{\theta(n-1)}{2} + \cos\frac{2\pi}{n}\sin\frac{\theta(n-1)}{2} + \sin\frac{\theta(n+1)}{2}}.$$

Cross-multiplying and performing trigonometric simplifications gives

$$\begin{aligned} -\sin^2(\tfrac{2\pi}{n})\frac{\sin(\theta(n-1))}{2} &= \cos^2(\tfrac{2\pi}{n})\frac{\sin(\theta(n-1))}{2} - \frac{\sin(\theta(n+1))}{2} + \cos(\tfrac{2\pi}{n})\sin\theta \\ &\quad + \cos(\tfrac{2\pi}{n})\sin(\theta(n-1)) - 2\sin\tfrac{\theta(n-1)}{2}\cos\tfrac{\theta(n+1)}{2} \end{aligned}$$

so that

$$\left(\frac{1}{2}+\cos\frac{2\pi}{n}\right)\sin(\theta(n-1)) - \frac{1}{2}\sin(\theta(n+1)) - \sin(\theta n) + \left(1+\cos\frac{2\pi}{n}\right)\sin(\theta) = 0. \tag{1}$$

Equation (1) is exact, but to estimate a solution, we perform a series expansion in θ

$$\begin{aligned} 0 = &\left[\left(\frac{1}{2}+\cos\frac{2\pi}{n}\right)(n-1) - (n+1)/2 - n + 1 + \cos\frac{2\pi}{n}\right]\theta \\ &- \left[\left(\frac{1}{2}+\cos\frac{2\pi}{n}\right)(n-1)^3 - (n+1)^3/2 - n^3 + 1 + \cos\frac{2\pi}{n}\right]\frac{\theta^3}{6} + O(n^4\theta^5) \\ 0 = &-\left[\frac{2\pi^2}{n} + O(1/n^2)\right]\theta + \left[6n^2 + O(n)\right]\frac{\theta^3}{6} + O(n^4\theta^5). \end{aligned}$$

While $\theta = 0$ is a solution, our expression for δ has a singularity at $\theta = 0$, so we seek a different solution. Ignoring the error terms suggests $\theta \doteq \sqrt{2}\pi n^{-3/2}$. Since the function in (1) is real-valued, we can appeal to the intermediate value theorem to show that there is in fact a root at

$$\theta \doteq \sqrt{2}\pi n^{-3/2}.$$

For this value of θ we have

$$\lambda = \frac{1+\cos\theta}{2} \doteq 1 - \frac{\pi^2}{2}n^{-3},$$

and (using the second equation for δ)

$$\delta \doteq \frac{2\pi/n}{2n\theta/2 + n\theta/2 + n\theta/2} \doteq \frac{1}{\sqrt{2}n^{1/2}}. \qquad \square$$

Again $\Psi_{\max} = (1+o(1))n$. Next we estimate R. If there is a shift-left, then provided card $\diamondsuit$ is not in position $-(n-1)/2$, we have

$$\begin{aligned} \Delta\Psi_\diamondsuit &= (\cos\theta - 1)[\cos(\theta x) + i\delta\sin(\theta x)]e^{2\pi i z/n} + \sin\theta[\sin(\theta x) - i\delta\cos(\theta x)]e^{2\pi i z/n} \\ &= O(\theta^2) + O(\theta^2 x) + O(\theta\delta) = O(n^{-3}) + O(n^{-2}) + O(n^{-2}) = O(n^{-2}). \end{aligned}$$

If card $\diamondsuit$ is in position $-(n-1)/2$, then

$$\Delta\Psi_\diamondsuit = 2i\delta\sin(\theta(n-1)/2)e^{2\pi i z/n} = O(n^{-1}).$$

Adding up these contributions over the different cards, we find $\Delta\Psi = O(n^{-1})$. Likewise $\Delta\Psi = O(n^{-1})$ if the move was a shift-right. For transposes, $\Delta\Psi_\diamondsuit$ is nonzero for only two cards, and for these it is $O(n^{-1})$. Thus in all cases we have $|\Delta\Psi|^2 \leq O(n^{-2})$, and so $R \leq O(n^{-2})$. Plugging our values of λ, Ψ_{max}, and R into Lemma 2, we obtain, for fixed values of ε, a lower bound on the mixing time of $\frac{1-o(1)}{\pi^2} n^3 \log n$ shuffles.

7. Remarks

We have seen how to extend the lower bound technique used in (Wilson, 2001) to shuffles that are much slower than what the position of a single card would indicate. Interestingly, for the shift-or-swap and symmetrized-Rudvalis shuffles, the spectral gap for the lifted shuffle is smaller than the spectral gap of the shuffle itself, so it is curious that we obtained a lower bound for these shuffles by considering their lifted versions.

In an early draft, we lower bounded the mixing time of the original Rudvalis shuffle without considering its lifted version, and this earlier approach might be considered simpler. But it is not clear how to lower bound the symmetrized Rudvalis shuffle without lifting it, and our current approach has the advantage that the analyses for all three shuffles treated here are similar. The original Rudvalis shuffle and its lifting are isomorphic, and the earlier analysis is effectively a special case of the present analysis where the lifting is not explicit.

We suspect that the constants in the lower bounds of Theorem 1 are tight. Rudvalis asked if his shuffle was the slowest shuffle evenly supported on two generators; the lower bounds given here suggest that the shift-or-swap shuffle (for odd n) is slower by a factor of 4.

Acknowledgements

The author thanks Laurent Saloff-Coste and Persi Diaconis for calling attention to these shuffles, and Laurent Saloff-Coste for comments on an earlier draft.

References

David J. Aldous and James A. Fill. *Reversible Markov Chains and Random Walks on Graphs.* Book in preparation, `http://www.stat.berkeley.edu/~aldous/RWG/book.html`, 2005.

Persi Diaconis. *Group Representations in Probability and Statistics.* Institute of Mathematical Statistics, 1988.

Persi Diaconis. The cutoff phenomenon in finite Markov chains. *Proceedings of the National Academy of Sciences, USA*, 93:1659–1664, 1996.

Persi Diaconis and Laurent Saloff-Coste. Comparison techniques for random walk on finite groups. *The Annals of Probability*, 21(4):2131–2156, 1993.

Persi Diaconis and Laurent Saloff-Coste. Random walks on finite groups: a survey of analytic techniques. In *Probability measures on groups and related structures, XI (Oberwolfach, 1994)*, pages 44–75. World Sci. Publishing, 1995.

Martin V. Hildebrand. *Rates of Convergence of Some Random Processes on Finite Groups.* PhD thesis, Harvard University, 1990.

Laurent Saloff-Coste. Lower bound in total variation for finite Markov chains: Wilson's lemma, 2002. Manuscript.

Elizabeth L. Wilmer. *Exact Rates of Convergence for Some Simple Non-Reversible Markov Chains.* PhD thesis, Harvard University, 1999.

Elizabeth L. Wilmer. A local limit theorem for a family of non-reversible Markov chains, 2002. arXiv:math.PR/0205189.

David B. Wilson. Mixing times of lozenge tiling and card shuffling Markov chains, 2001. To appear in *The Annals of Applied Probability.* arXiv:math.PR/0102193.