OPTIMAL COUPLINGS FOR CARD SHUFFLING

JENNIFER THOMPSON
JESUS COLLEGE, CAMBRIDGE

ADVISOR: YEVGENIY KOVCHEGOV
OREGON STATE UNIVERSITY

ABSTRACT. We look at mixing times of Markov Chains - speed of convergence towards an equilibrium distribution. We wish to find such mixing times using the coupling method, as this is the most general method available. So we try to find good couplings, where the coupling time matches other known measures of mixing time, such as cover time.

We work on this by primarily considering the shuffling times of a deck of cards, as these are easily solved by cover time which provides a way to measure our progress.

1. INTRODUCTION

Given any irreducible and aperiodic Markov Chain $(X_t)_{t=0}^{\infty}$, we expect it to tend to an equilibrium distribution π over time, irrespective of its initial distribution. An intuitive way to measure this is with Total Variation, a metric on probability distributions.

Here are some definitions that we will use:

Definition 1.1. *A Markov Chain is an infinite sequence of random variables $(X_t)_{t=0}^{\infty}$ in a state space Ω such that $\mathbb{P}(X_{t+1} = j | X_t = i) = P_{ij}$, independent of time t. The values P_{ij} form the transition matrix P of the Markov chain.*

Definition 1.2. *The period of a chain is the gcd$(\mathcal{T}(x))$ where $\mathcal{T}(x) := \min\{t \geq 1 : (P^t)_{x,x} > 0\}$. A chain is Aperiodic if all states have period 1.*

Definition 1.3. *A chain P is Irreducible if $\forall x, y \in \Omega$, there exists a $t \geq 0$ such that $(P^t)_{x,y} > 0$. That is, from any one state we can reach any other state.*

Definition 1.4. *The Total Variation Distance between two probability distributions μ and ν on Ω is*

$$||\mu - \nu||_{TV} := \frac{1}{2} \sum_{x \in \Omega} |\mu(x) - \nu(x)|$$

Theorem 1.5. *If a chain $(X_n)_{n=0}^{\infty}$ is irreducible and aperiodic over a finite state space then $\exists \pi :$ $||P^t \mu - \pi||_{TV} \to 0$.*

Date: August 13, 2010.
This work was done during the Summer 2010 REU program in Mathematics at Oregon State University.

How fast is this convergence? Using our metric, we would look at the time for total variation to become arbitrarily small. We call this the mixing time.

Definition 1.6. *Let $d(t) = max_{x\in\Omega}\|P^t(x,\cdot)-\pi\|$. The Mixing Time of a Markov chain is as*

$$\tau_{mix}(\varepsilon) := min\{t : d(t) \leq \varepsilon\}$$

This can be hard to evaluate, so we look instead at other times, such as coupling time. To use the coupling method, we introduce a second Markov Chain $(Y_t)_{t=0}^{\infty}$, which has the same transition probabilities as $(X_t)_{t=0}^{\infty}$, but Y_t is distributed by the equilibrium distribution π for all t.

We run these Markov Chains, not necessarily independently, until the meet at time τ_c. This is called the coupling time. After τ_c, we run the chains together. We see that X_t must have the equilibrium distribution for $t \geq \tau_c$, since $X_t = y_t$ after coupling.

Definition 1.7. *The Coupling Time of any two chains X_t, Y_t with shared transition probabilities is $\tau_c := min\{t : X_t = Y_t\}$.*

From this, it intuitively makes sense that $\tau_{mix}(\varepsilon)$ should be bounded above by τ_c. The question is whether we can find a way to control the relationship between X_t and Y_t for $t < \tau_c$ in such a way that mixing time is equal to coupling time. A coupling that achieves this is called an optimal coupling.

Theorem 1.8. *The Coupling Inequality, from* [1]. *For all x, $\|P^t(x,\cdot)-\mu\|_{TV} \leq \mathbb{P}\,[\tau_c > t]$.*

Corollary 1.9. $\tau_{mix}(\varepsilon) \leq \mathbb{E}\,\tau_c/\varepsilon$.

Proof. Take $t = \tau_c/\varepsilon$. Then $\varepsilon = \mathbb{E}\,\tau_c/t \geq \mathbb{P}\,[\tau_c > t]$, by Markov's inequality. Applying the Coupling Inequality we have $\varepsilon \geq \|P^t(x,\cdot)-\mu\|_{TV}\ \forall x$ which implies $d(t) < \varepsilon$. Thus $\tau_{mix}(\varepsilon) < t$. □

We seek optimal couplings for card shuffling techniques on a deck of n cards labelled $\{1,\dots n\}$. Such shufflings are Markov Chains on S_n, the permutations of $\{1,\dots n\}$. The equilibrium distribution is always uniform over S_n.

Shuffling techniques we might consider are Random Transpositions, Random-to-Random, Top-to-Random, and the much more complex Thorp Shuffle.

Yevgeniy Kovchegov and Robert Burton found an optimal coupling of time $n\log n$ in the case of the Random Transposition shuffle [2].

Here, we find an optimal coupling for the Random-to-Random shuffle, and a non-optimal coupling of time $n^2\log n$ for the Top-to-Random shuffle.

2. Random-to-Random Shuffling

We consider the shuffling algorithm whereby a random card is removed from the deck and reinserted at a random position.

This is irreducible and aperiodic, as $P_{x,x} = 1/n > 0$, so we have convergence to equilibrium.

Define τ_{cover}, to be the time when each card has been chosen at least once. For $t > \tau_{\text{cover}}$, each card has been randomly placed, and so the deck is mixed. Finding $\mathbb{E}\tau_{\text{cover}}$ is a simple problem to solve,known as the Coupon Collector. We see the time to choose the $(r+1)^{th}$ card after choosing the r^{th} card is distributed as Geom($n-r/n$). Summing the expectation for these marginal times yields $\mathbb{E}[\tau_{\text{cover}}] = n\log n + O(n)$.

2.1. **First Coupling Method.** Take deck A to be the original deck, and deck B to be a deck with the stationary distribution.

First consider the obvious coupling:

(1) Choose a card at random from deck A and remove it.
(2) Locate the same card from deck B an remove it.
(3) Reinsert both at the same random location in each deck.

What is the coupling time?

Let

$$a_t = \text{ cards from the top of the decks that are matched at time } t$$
$$b_t = \text{ cards from the bottom of the decks that are matched at time } t.$$

Then define $R_t = a_t + b_t$. R_t is a Markov Chain on $\{0,1,2,\dots,n\}$ with transition matrix P The decks are coupled when $R_t = n$.

Let C be the unmatched region and D be the matched region.

We find the transition probabilities:

$$P_{ii+1} = \mathbb{P}[\text{Card is chosen from } C \text{ and placed in } D] = \frac{n-i}{n}\frac{i+2}{n}$$
$$P_{ii-1} = \mathbb{P}[\text{Card is chosen from } D \text{ and placed in } C] = \frac{i}{n}\frac{n-i-1}{n}$$
$$P_{ii} = 1 - P_{ii+1} - P_{ii-1}.$$

Let $t_i = \mathbb{E}\left[\min\{t : R_t = n | R_0 = i\}\right]$, the expected time to hit n from i. We want to find $t_0 = \mathbb{E}\, T_{\text{coupling}}$.

The t_i form the minimal non negative solution to

$$t_i = 1 + (Pt)_i \ \ \forall i \neq n, \ \ t_n = 0.$$

In this case the recurrence relation is

$$t_i = P_{ii+1}t_{i+1} + P_{ii}t_i + P_{ii-1}t_{i-1} + 1$$

for $0 < i < n$. Substituting in the values above, gathering t_i on the left and multiplying by n^2 yields:

$$((n-i)(i+2)+i(n-i-1))t_i = (n-i)(i+2)t_{i+1}+i(n-i-1)t_{i-1}+n^2.$$

Let $j = n-i$, $S_j = t_{n-j}$. Then we have

$$(j(n-j+2)+(n-j)(j-1))S_j = j(n-j+2)S_{j-1}+(n-j)(j-1)S_{j+1}+n^2$$

with the boundary condition $S_0 = 0$, which does not have a solution in a nice form.

2.2. **Coupling by order of cards rather than position.** The previous method focused on the absolute position of cards within the deck. Now we will try instead to preserve the ordering of the cards.

We introduce a second deck, and couple them as follows.

- Randomly choose $i \in [1,n]$.
- Remove card with label i from each deck.
- Randomly reinsert card i in deck A.
- (1) If the new location if i in A is the top of A, then insert i on the top of B.
  (2) If the new location of i in A is below card j, insert i below j in B.

A matching is a pair of cards, one from deck A, the other from deck B, with the same label.

At time t, we have matchings between the two decks, as in Figure 1.

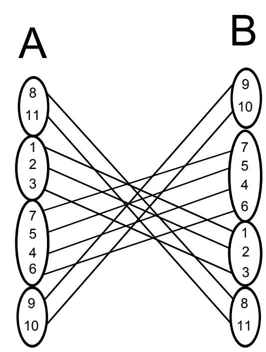

FIGURE 1. One configuration of matchings between two decks of eleven cards.

At time 0, these matchings are not in any particular order, but we try to group them into bunches, where a bunch is a set of aligned matchings - they are in the same order in both decks.

How does the coupling method act on the matchings? At time t, we randomly chose a matching and add it a particular bunch, with probability proportional to the size of that bunch - except for the bunch at the top, to which we assign additional probability $1/n$.

Coupling occurs when the bunches no longer cross over one another - that is, there is only one bunch left. Call this time T.

Then by the Law of Total Probability

$$\mathbb{E}\,T = \sum_i \mathbb{P}\,[\text{ Bunch } i \text{ is final bunch}]\mathbb{E}\,[T \mid \text{Bunch } i \text{ is final bunch}]$$

But

$$\mathbb{E}\,[T \mid \text{Bunch 0 is final bunch}] \leq \mathbb{E}\,[T \mid \text{Bunch } i \text{ is final bunch}]$$

so

$$\mathbb{E}\,T \leq \mathbb{E}\,[T \mid \text{Bunch 0 is not the final bunch}] = \mathbb{E}\,[T \mid \text{Bunch } n \text{ is final bunch}].$$

So, we write

$$s(x) = \mathbb{E}[T \mid \text{final bunch, } n\text{, is size } x \text{ at time } 0].$$

Then we have the boundary condition $s(n) = 0$ and the recurrence:

$$S(x) = 1 + \frac{x(n-x+1)}{n^2}s(x-1) + \frac{(n-x)x}{n^2}s(x+1) + \left(1 - \frac{2x(n-x)+x}{n^2}\right)s(x).$$

We rearrange this to the ODE-friendly form

$$-n^2 = x(n-x)[s(x+1) - 2s(x) + s(x-1)] + x[s(x-1) - s(x)]$$

which corresponds to the ODE

$$-n^2 = x(n-x)s''(x) - xs'(x).$$

So solve this, we try to rewrite as

$$\frac{n^2}{x} = ((x-n)s')'$$

so then

$$(x-n)s' = n^2 \log x - C$$

$$s(x) = n^2 \int_0^x \frac{\log t}{t-n}dt + C\log(n-x) + D$$

$$= -n^2\left(\mathrm{Li}_2\left(\frac{x}{n}\right) + \log x \log\left(1 - \frac{x}{n}\right)\right) + C\log(n-x) + D$$

where Li_2 is the polylogarithm function.

Applying $s(n) = 0$ yields

$$0 = -n^2(\mathrm{Li}_2(1) + \log n \log(0)) + C\log(0) + D.$$

However, $\log 0$ is singular, so we must have $C = n^2 \log n$.

Then the solution for s is

$$s(x) = n^2\left(-\mathrm{Li}_2\left(\frac{x}{n}\right) - \log x \log\left(1 - \frac{x}{n}\right) + \log n \log(n-x)\right) + D.$$

Substitute $y = {}^{x}\!/_{n}$,

$$\begin{aligned} s(y) &= n^2(-\mathrm{Li}_2(y) - \log(yn)\log(1-y) + \log n \log(n(1-y))) + D \\ &= n^2(-\mathrm{Li}_2(y) - (\log y + \log n)\log(1-y) + \log n(\log n + \log(1-y))) + D \\ &= n^2(-\mathrm{Li}_2(y) - \log y \log(1-y) + (\log n)^2) + D \end{aligned}$$

Returning to the boundary condition

$$s(y=1) = 0 = n^2(-\mathrm{Li}_2(1) - 0 + (\log n)^2) + D$$

which implies $n^2(\log n)^2 + D = n^2\pi^2/6$, and so the solution for s, with $0 < y < 1$ is

$$s(y) = n^2(\pi^2/6 - \mathrm{Li}_2(y) - \log y \log(1-y)).$$

$-\mathrm{Li}_2(y) - \log y \log(1-y)$ is bounded on $[0,1]$ and so the coupling method is $O(n^2)$.

Once more, this is not the desired coupling time. We also note that this analysis is missing the possibility for different bunches of matchings to merge, because all those matchings inbetween have disappeared.

2.3. **Random-to-Random Coupling with Three Bunches of Matchings.**

So far, the lower bound of $n \log n$ has proven elusive. Perhaps we will have better luck answering the question where the number of bunches small; where we are close to achieving coupling.

So in the method above, consider the stage where we have just 3 bunches of laces, as illustrated below.

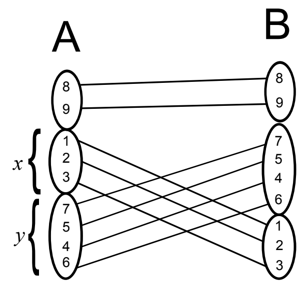

FIGURE 2. The case with 3 sets of aligned matchings.

Let X be the size of the second bunch, and Y be the size of the third bunch.

Then the decks are coupled when either the second or third bunch disappears.

We obtain the transition probabilities:

(1) $(x,y) \to (x+1,y)$ w.p. ${}^{(n-x-y)x}\!/_{n^2}$
(2) $(x,y) \to (x+1,y-1)$ w.p. ${}^{yx}\!/_{n^2}$
(3) $(x,y) \to (x,y-1)$ w.p. ${}^{y(n-x-y+1)}\!/_{n^2}$
(4) $(x,y) \to (x,y+1)$ w.p. ${}^{(n-x-y)y}\!/_{n^2}$

(5) $(x,y) \to (x-1,y)$ w.p. $x(n-x-y-1)/n^2$
(6) $(x,y) \to (x-1,y+1)$ w.p. xy/n^2
(7) $(x,y) \to (x,y)$ w.p. $x(x-1)+y(y-1)+(n-x-y)^2/n^2$

Then define the hitting time

$$h_{x,y} = \mathbb{E}_{x,y}\left[\text{Time till } X=0 \text{ or } Y=0\right].$$

So $h_{x,y}$ satisfies the reccurence

$$\begin{aligned} h_{x,y} = n^2 &+ (n-x-y)xh_{x+1,y} + yxh_{x+1,y-1} + y(n-x-y+1)h_{x,y-1} + (n-x-y)yh_{x,y+1} \\ &+ x(n-x-y+1)h_{x-1,y} + xyh_{x-1,y+1} + (x(x-1)+y(y-1)+(n-x-y)^2)h_{x,y}. \end{aligned}$$

This is rather intimidating to solve directly, so we try to transform it into a PDE. First, consider the following rearrangement:

$$\begin{aligned} -n^2 = {} & xy[h_{x+1,y-1} - 2h_{x,y} + h_{x-1,y+1}] \\ & +(n-x-y)x[h_{x+1,y} - 2h_{x,y} + h_{x-1,y}] \\ & +(n-x-y)y[h_{x,y+1} - 2h_{x,y} + h_{x,y-1}] \\ & +x[h_{x-1,y} - h_{x,y}] + y[h_{x,y-1} - h_{x,y}]. \end{aligned}$$

Then we define the difference operators:

$$\Delta_x h_{x,y} = h_{x+1,y} - h_{x,y}, \nabla_x h_{x,y} = h_{x,y} - h_{x-1,y}$$
$$\Delta_y h_{x,y} = h_{x,y+1} - h_{x,y}, \nabla_y h_{x,y} = h_{x,y} - h_{x,y-1}$$

So the recurrence can be written as

$$-n^2 = [xy(\Delta_x - \Delta_y)(\nabla_x - \nabla_y) + (n-x-y)x\Delta_x\nabla_x + (n-x-y)y\Delta_y\nabla_y - x\nabla_x - y\nabla_y]h_{x,y}.$$

We then transform this into a PDE using the correspondence

Difference operator	Differential operator
$\Delta_x\nabla_y + \Delta_y\nabla_x$	$2\frac{\partial^2}{\partial x \partial y}$
$\Delta_x\nabla_x$	$\frac{\partial^2}{\partial x^2}$
$\Delta_y\nabla_y$	$\frac{\partial^2}{\partial y^2}$
∇_x	$\frac{\partial}{\partial x}$
∇_y	$\frac{\partial}{\partial y}$

So the PDE approximation for h is

$$-n^2 = \left[xy\left(\frac{\partial^2}{\partial x^2} - 2\frac{\partial^2}{\partial x \partial y} + \frac{\partial^2}{\partial y^2}\right) + (n-x-y)x\frac{\partial^2}{\partial x^2} + (n-x-y)y\frac{\partial^2}{\partial y^2} - x\frac{\partial}{\partial x} - y\frac{\partial}{\partial y}\right]h$$

which simplifies to

$$-n^2 = \left[n\left(x\frac{\partial^2}{\partial x^2} + y\frac{\partial^2}{\partial y^2} \right) - \left(x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y} \right)^2 \right] h.$$

This appears to be another dead end; there is no obvious way to solve this PDE.

2.4. **A heuristic approach for two sets of matchings.** At this point, we begin to question whether this coupling really have expected time $O(n\log n)$. Can we adopt a heuristic approach to get some better idea of the mathematical proof?

Consider the case where we have just two sets of matchings, as illustrated in Figure 3.

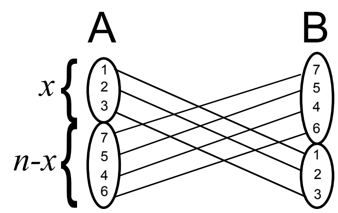

FIGURE 3. A case with just two sets of aligned matchings.

The first set has size x. In reality the transition probabilities are close to but not quite symmetrical, due to the $1/n$ probability of placing the cards on the top of the decks. However, for now we will ignore this case and use the transition probabilities:

$$P_{x,x-1} = \frac{x(n-x)}{n^2} = P_{x,x+1},\ P_{x,x} = 1 - 2\frac{x(n-x)}{n^2}.$$

Then writing $t_x = \mathbb{E}\,[T|X_0 = x]$, the recurrence relation $t_x = 1 + (Pt)_x \;\; \forall i \neq n, 0$ becomes

$$-\frac{n^2}{x(n-x)} = t_{x+1} - 2t_x + t_{x-1}.$$

Coupling occurs at time T when either $X_T = 0$, or $X_T = n$. Note that $\mathbb{P}\,[X_T = 0] = \frac{n-x}{n}$ and $\mathbb{P}\,[X_T = n] = \frac{x}{n}$. So, by the Law of Total Probability

$$\mathbb{E}\,[T] = \frac{n-x}{n}\mathbb{E}\,[T \mid X_T = 0] + \frac{x}{n}\mathbb{E}\,[T \mid X_T = n].$$

First consider the case where $x = O(1)$.

So we look at

$\mathbb{E}\,[T \mid X_T = 0] = (\#\text{ steps in a non-lazy symmetric walk})(\text{delay between steps}) = O(1)O(n) = O(n)$

and

$\mathbb{E}\,[T \mid X_T = n] = (\#\text{ steps in a non-lazy symmetric walk})(\text{delay between steps}) = O(n^2)O(1) = O(n^2)$, since most of the steps in the centre region occur where $x = O(n)$, so $P_{x,x} = O(1)$ and delay $= O(1)$.

Then we have

$$\mathbb{E}\ [T] = \frac{n - O(1)}{n} O(n) + \frac{O(1)}{n} O(n^2) = O(n).$$

2.5. **Expanding the heuristic approach to the general case.** We now look at the situation where we have a lot of ordered sets.

We note that if we have large ordered sets, then we can subdivide so that we have $O(n)$ sets of size $O(1)$.

So now we are looking at expiration times of each small bunch of matchings, $T_1, T_2, \dots, T_m$ where $m = O(n)$ and $\mathbb{E}\ T_i = O(n)$.

We model these with Exp(λ) random variables where $\frac{1}{\lambda} = O(n)$. Then we want $T_c = \max[T_i]$. Use the property of memorylessness ($\mathbb{P}\ [T_i > t + s | T_i > s] = \mathbb{P}\ [T_i > t]$), together with $\min_{i \in [1,m]}[T_i] \sim$ Exp($m\lambda$) to find $\mathbb{E} \max[T_i]$.

Split $\mathbb{E}\ T_c$ into a sequence as illustrated in Figure 4.

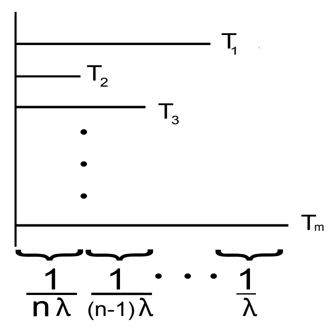

FIGURE 4. Illustration of $\mathbb{E} \max_{i \in [1,m]}[T_i] \sim m \log m$

Then $\mathbb{E} \max_{i \in [1,m]}[T_i] = \mathbb{E}\ \min_{i \in [1,m]}[T_i] + \mathbb{E}\ \min_{i \in [1,m-1]}[T_i] + \cdots + \mathbb{E}\ \min_{i \in [1,1]}[T_i] = \frac{1}{\lambda}(m^{-1} + (m-1)^{-1} + \cdots + 1^{-1}) = \frac{1}{\lambda}(\log m + O(1)) = O(n \log n)$, as required.

2.6. **Formalizing the heuristic approach - A Laces Approach.** We seek to formalize the above argument via path-coupling.

We introduce the metric $d(\sigma, \sigma')$, $d : S_n \times S_n \to \mathbb{Z}_{\geq 0}$, where $d(\sigma, \sigma')$ is the minimal number of nearest neighbour transpositions to traverse between the two permutations.

For example

Note that $\max(d) = \binom{n}{2}$, the permutation where every number crosses every other number.

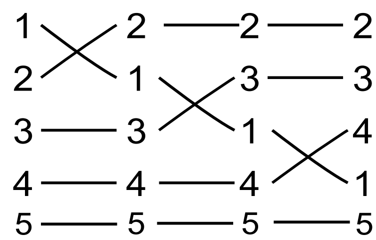

FIGURE 5. Minimal number of crossings between (1234) and the identity is $d((id),(1234)) = 3$.

We have the quantity $d_t = d(B_t, A_t)$, the distance between our two permutations at time t. We want to find the relationship $\mathbb{E}\, d_{t+1} = f(\mathbb{E}\, d_t)$.

Call the path taken by a card label a lace. Thus each lace in involved in a certain number of crossings. Say $r = \#$ crossings per lace. Then we have $d_t = {}^{nr}/_{2}$.

At each timestep we pick lace at random and remove it - subtracting r from d_t.

We have two cases for reinsertion:

(1) We add the new lace to the top of the deck - there are no new crossings. This has probability ${}^{1}/_{n}$.
(2) We add the new lace below lace j. This has probability ${}^{n-1}/_{n}$. Then the number of additional crossings is the number of crossings of lace j, as in Figure 6. Then

$$\mathbb{E}\,[\text{new crossings}] = \mathbb{E}\,[\text{average number of crossings for the remaining laces}] = \left(\frac{nr}{2} - r\right)\frac{1}{n-1}.$$

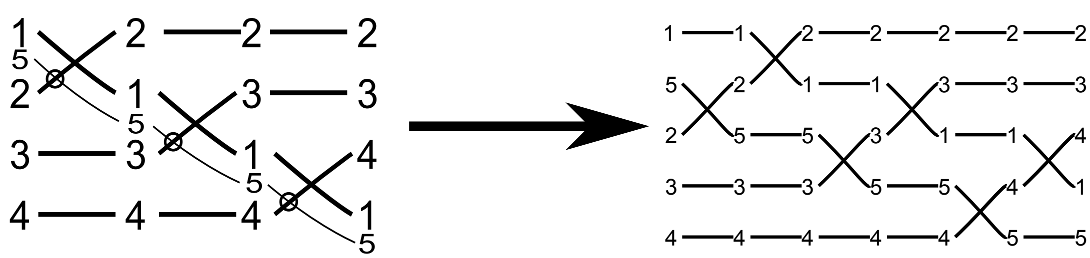

FIGURE 6. Inserting lace 5 below lace 1 incurs three additional crossings.

Then

$$\mathbb{E}\, d_{t+1} = \frac{nr}{2} - r + \left(\frac{n-1}{n}\right)\left(\frac{nr}{2} - r\right)\frac{1}{n-1}.$$

Now we can find f.

$$\frac{\mathbb{E}\, d_{t+1}}{\mathbb{E}\, d_t} = \frac{\left(\frac{n}{2} - 1\right)\left(1 + \frac{1}{n}\right)}{\frac{n}{2}} = \left(1 - \frac{2}{n}\right)\left(1 + \frac{1}{n}\right) = 1 - \frac{1}{n} - \frac{2}{n^2}.$$

So

$$\mathbb{E}\, d_{t+1} = \left(1 - \frac{1}{n} - \frac{2}{n^2}\right)^t \mathbb{E}\, d_0 < \left(1 - \frac{1}{n} - \frac{2}{n^2}\right)^t \binom{n}{2}.$$

Want to find t such that

$$\left(1 - \frac{1}{n} - \frac{2}{n^2}\right)^t \binom{n}{2} = \varepsilon.$$

Take logarithms:

$$t \log\left(1 - \frac{1}{n} - \frac{2}{n^2}\right) + 2\log n - \log 2 = \log \varepsilon.$$

Use the Taylor approximation for logarithms:

$$t\left[-\left(\frac{1}{n} + O\left(\frac{1}{n^2}\right)\right)\right] = -2\log n + \log 2 + \log \varepsilon.$$

Then

$$t = \frac{2\log n - \log 2\varepsilon}{\frac{1}{n} + O\left(\frac{1}{n^2}\right)} = \frac{2n\log n - n\log 2\varepsilon}{1 + O\left(\frac{1}{n}\right)} = 2n\log n + O(n).$$

Finally, we have to show how this relates to coupling time.

Use Markov's Inequality: for a random variable $X > 0$,

$$\mathbb{P}\left[X > a\right] \leq \frac{\mathbb{E}\, X}{a}.$$

We have

$$\|\mu_t - \pi\|_{TV} \leq \mathbb{P}\left[T_c > t\right] = \mathbb{P}\left[d_t \geq 1\right] \leq \frac{\mathbb{E}\, d_t}{1} \leq \varepsilon,$$

thus the deck is mixed at time $t = 2n\log n + O(n)$.

3. TOP-TO-RANDOM SHUFFLING

The Top-to-Random Shuffling is simple. At each stage, take the top card from the deck, and randomly insert it in one of the n positions in the deck.

Here cover time, time for the bottom card to reach the top, is $n\log n$. A sketch proof of this is very similar to the Coupon Collector [3], simply by noting that the time for the original bottom card to reach the $(r+1)^{th}$ position after reaching the r^{th} position is distributed as Geom(r/n), where the bottom position is 1, and the top position is n.

3.1. **First Coupling Method.** Here we find an upper bound of ${n^3}/{2}$ on the coupling time of a top to random coupling method.

Method.

- If the top cards are the same in both decks, randomly place them in the same position anywhere in the decks.
- Take the card from the top of deck A, with label i, and randomly place it at position j in deck A.
- Remove the top card from deck B,with label k. Then if $B_j = i$, place card k in position $j+1$ of deck B. If $B_{j+1} = i$ then place card k in position j of deck B. Otherwise, place card k in position j.

We see that a matching is achieved in the first two cases, with ${2}/{n}$ probability. However, this is only an improvement if the cards were not matched in the first place, that is $i \neq k$.

We want to look at the time it takes to achieve an additional matching, given that there are r matchings already. Call this time ρ_{r+1}.

Consider the worst case, which is that all r matched cards are positioned at the top of the deck. How long do we expect to wait until the $r+1^{th}$ card reaches the top?

Let $\alpha_k =$ Time for k^{th} card to reach $k-1^{th}$ position. Then $\alpha_k \sim Geom(\frac{n-k}{n})$, as the top card must be placed below the k^{th} position.

Then $\mathbb{E}\ \alpha_k = {n}/{n-k}$.

So the expected time till the $r+1^{th}$ card reaches the top is $\sum_{k=1}^{r} {n}/{n-k}$.

The expected number of attempts to make a matching with different cards on the top of each deck is ${n}/{2}$.

So $\mathbb{E}\rho_{r+1} \leq \frac{n}{2} \sum_{k=1}^{r} {n}/{n-k}$.

The expected coupling time is

$$\mathbb{E} \sum_{r=0}^{n-1} \rho_{r+1} \leq \sum_{r=0}^{n-1} \frac{n}{2} \sum_{k=1}^{r} \frac{n}{n-k}.$$

Then by changing the order of summation, this becomes

$$\frac{n^2}{2} \sum_{k=1}^{n-1} \sum_{r=k}^{n-1} \frac{1}{n-k} = \frac{n^2}{2} \sum_{k=1}^{n-1} \frac{n-k}{n-k} = \frac{n^2(n-1)}{2} \sim \frac{n^3}{2}$$

3.2. **Improved Top-to-Random Coupling.** We can better analyse the method above, to get a coupling time of $O(n^2 \log n)$.

To do this, let m_t be the number of matchings that have not yet been randomly placed so far, i.e. the number of matchings that have not reached the top after being matched.

If these m_t cards are at the top of the deck, then it takes expected time

$$\sum_{k=1}^{m_t} \frac{n}{n-k}$$

to randomly place these cards. We also have $\sum_t m_t = n$, as each of these matched pairs can only bother us once in the whole process.

Once this process is completed, or if the cards at the top are not any of these m_t cards, then the cards at the top have probability $\frac{n-r}{n}$ of being unmatched, where r is the number of matchings so far.

Then using the same process as the previous method for unmatched cards, we have

$$\mathbb{P}\left[\text{new matching}\right] = \frac{n-r}{n}\frac{2}{n} \text{ and } \mathbb{E}\left[\text{new matching}\right] = \frac{n}{n-r}\frac{n}{2}.$$

We want to find the maximum time the randomly place the m_t cards overall. This is

$$\max \sum_i \sum_{k=1}^{m_i} \frac{n}{n-k}.$$

Each summand $\leq n$, and there are n summands max $\Rightarrow \leq n^2$.

So

$$T_{\text{coupling}} \leq n^2 + \sum_{r=0}^{n-1} \frac{n^2}{2(n-r)} \leq n^2 + \frac{n^2 \log n}{2} = O(n^2 \log n).$$

4. Conclusion

We have been successful in finding path coupling method for random to random shuffling using a metric over the permutations of the deck. This coupling bounds the mixing time by $2n\log n + O(n)$, which is the right order, but still could be improved to $n \log n$.

We have found a coupling for the top-to-random shuffling, which previously had no known coupling. This coupling bounds the mixing time by $O(n^2 \log n)$. This could also be improved to $n \log n$.

REFERENCES

[1] Torgny Lindvall. *Lectures on the Coupling Method*. John Wiley and Sons, Inc, 1992.
[2] Robert Burton and Yevgeniy Kovchegov. Mixing times via super-fast coupling.
[3] David A. Levin, Yuval Peres, and Elizabeth L. Wilmer. *Markov Chains and Mixing Times*. American Mathematical Society, 2008.