Leonard J. Schulman | **Scribe: Farshad Moradi**
Notes for lecture 10, February 12th, 2003. Mixing Time: Exponential Decay

---

Theorem: If T is the mixing time of Markov chain A with stationary distribution π, then $|pA^{kT} - \pi| \leq 2^{-k}$

Proof: Using induction
For $k = 1$, the theorem holds. Assume for $k-1$, $|pA^{(k-1)T} - \pi| \leq 2^{-k+1}$, and $pA^{(k-1)T} = \pi + f^+ - f^-$, where

$$f^+ = \max_v \{(pA^{(k-1)T})_v - \pi_v, 0\},$$

similarly,

$$f^- = \max_v \{\pi_v - (pA^{(k-1)T})_v, 0\}$$

and $f = f^+ - f^-$, consequently,

$$|f| = |f^+| + |f^-|, \text{ and } |f^+|, |f^-| \leq l \text{ for some } l,$$

by induction,

$$|f| \leq 2l \leq 2^{-k+1}$$

it follows

$$(pA^{(k-1)T})A^T = (\pi + f^+ - f^-)A^T = (\pi + l(f^+/l) - l(f^-/l))A^T$$

But f^+/l and f^-/l are probability distributions, and $\pi A^T = \pi$, therefore

$$(pA^{(k-1)T})A^T = \pi + l(\pi + f^{++} - f^{+-}) - l(\pi + f^{-+} - f^{--})$$

or

$$pA^{kT} - \pi = l(f^{++} - f^{+-}) - l(f^{-+} - f^{--})$$

where

$$f^{++} = \max\{(\frac{f^+}{l}A^T)_v - \pi_v, 0\},$$

and similarly for f^{+-}, f^{-+}, f^{--}. We know that

$$|f^{++}| + |f^{+-}| \leq \frac{1}{2}$$

and

$$|f^{-+}| + |f^{--}| \leq \frac{1}{2}$$

and

$$l \leq 2^{-k}$$

Consequently,

$$|pA^{kT} - \pi| \leq 2^{-k}.$$

Examples of mixing time:

- Deck of n cards, take the top card and insert it in a random position (from n positions). Repeat. ("Top to random" shuffle.)

- Deck of n cards, pick a card randomly and put it on the top. Repeat. ("Random to top" shuffle.)

Question: what is the mixing time in each case?

Let $X_1, X_2, \ldots$ be successions of time sequence states from a Markov chain M. A *stopping time* T for M is a random variable ϵN such that $P(T = t)$ is independent of $X_{t+1}, X_{t+2}, \ldots$, conditioned on $X_1, X_2, \ldots, X_t$. Example: Imagine a random walk, the time that it first hits zero is a stopping time.

Definition: A strong stopping time for Markov chain M with stationary distribution π is a random variable T such that

$$P(X_t = x | T = t) = \pi_x$$

In other terms, the conditional state if the chain is stopped is the stationary distribution.
Corollary: $P(X_t = x | T \leq t) = \pi_x$ *iff* T is a strong stationary time.

Let T_x be a strong stationary time for chain M, started in state x. Let $\Delta_x(t)$ be the L_1 distance from π after time t.

Lemma: $\Delta_x(t) \leq 2P(T_x > t)$
Proof: Let

$$p_x(t) = e_x M^t$$

where e_x is the singular distribution on state x. and

$$p_x^{(t)}(A) = \sum_{y \epsilon A} p_x^{(t)}(y)$$

$$\begin{aligned} p_x^{(t)}(A) = P(X_t \epsilon A) = \\ & P(X_t \epsilon A \& T_x > t) + P(X_t \epsilon A \& T_x \leq t) = \\ & P(X_t \epsilon A | T_x > t) P(T_x > t) + \pi(A) P(T_x \leq t) = \\ & P(X_t \epsilon A | T_x > t) P(T_x > t) + \pi(A)(1 - P(T_x > t)) = \\ & \pi(A) + P(T_x > t)[P(X_t \epsilon A | T_x > t) - \pi(A)] \end{aligned}$$

but $P(X_t \epsilon A | T_x > t) - \pi(A)$ is less than or equal to one. Therefore,

$$p_x^{(t)}(A) - \pi(A) \leq p(T_x > t)$$

Reminder:

$$|p - q|_1 = 2 \max_A \sum_{x \epsilon A} (p_x - q_x)$$

$$\Delta_x(t) = |p_x(t) - \pi|$$

Consequently,

$$\Delta_x(t) \leq 2P(T_x > t)$$

Top-to-random shuffle: Let B be the initial bottom card. Let T equal one time-step after B reaches the top of the deck.
Claim: T is a strong stationary time.
To verify the claim, define times $K_1, K_2, \ldots, K_{n-1}$, where K_i denotes the first time at which there are i cards beneath B.
Note that $K_i - K_{i-1}$ has the same distribution as the gathering of the $(n - i + 1)$st coupon in the coupon collector problem.

Conclusion: the mixing time for top-to-random shuffle is $\Theta(n \log n)$.
(Actually it can be shown that time T that the last coupon is collected has the property that $P(T > n \ln n + cn) < e^{-c}$.)

Coupling: (with application to random-to-top shuffle)
Consider Markov chain M, and two M-chains $X_1, X_2, \ldots$, and $Y_1, Y_2, \ldots$. We can think of these chains as the positions of two independent "particles".
A *coupling process* is a pair of walks whose marginals are as in M, and there is a time t such that $X_t = Y_t \Rightarrow X_{t+1} = Y_{t+1}$. In other terms, when the two particles collide, they attach and never separate again, although the walk for each particle still obeys M.
We will be interested in pairs such that X starts in an arbitrary distribution, and Y starts either in an arbitrary distribution or in π, where π is the stationary distribution for M. Random variable C denotes the first time such that $X_t = Y_t$.

Lemma: Suppose Y starts in π. Then, $Mix(M) \leq 4E[C]$.
Markov inequality: for any $\lambda > 0, P(C > \lambda E[C]) \leq \frac{1}{\lambda}$.
Consequently,

$$P(C \leq \lambda E[C]) \leq 1 - \frac{1}{\lambda}$$

Let $f_t(v) = P(X_t = Y_t = v)$.

$$\sum_v f_{\lambda E[C]}(v) \geq 1 - \frac{1}{\lambda}$$

X starts in arbitrary p. Y starts (and therefore stays) in π . It follows that

$$f_t(v) = P(X_t = Y_t = v) \leq P(Y_t = v) \Rightarrow f_t(v) \leq \pi(v)$$

Similarly,

$$f_t(v) \leq p^{(t)}(v)$$

So we have the two functions $p^{(\lambda E[C])}(v)$ and $f_{\lambda E[C]}(v)$, such that the first one dominates the second one.

$$|p^{(\lambda E[C])} - f_{\lambda E[C]}| \leq \frac{1}{\lambda}$$

$$|\pi - f_{\lambda E[C]}| \leq \frac{1}{\lambda}$$

It results that

$$|p^{(\lambda E[C])} - \pi| \leq \frac{2}{\lambda}$$

for any λ. We can get the mixing time by simply choosing $\lambda = 4$.

random-to-top shuffle
Process X: Take the deck of cards, pick one card at random (uniformly), and then move it to the top.
Process Y: Start from the stationary distribution. Pick the same card (by label, not by position) as the one that was picked in X. Again, the picking is uniformly distributed, so Y will remain in the stationary distribution.
The two walks couple when all cards (identified by label) are picked at least once. (Those cards which have been picked at least once occupy a segment at the top of the deck, and the segment includes the same cards, in the same uniformly random order, in both decks.) The coupling time is therefore bounded by a coupon collector analysis.