Shuffling

Top-to-Random policy: take the top card and insert it at one of the \( n \) positions in the deck chosen uniformly at random.

3 cards, \( 3! \)

\[
\frac{1}{3} \rightarrow \frac{1}{3} \rightarrow \frac{1}{3}
\]

irreducible, aperiodic.

but it is not reversible.

If we insert the top card into (say) the middle of the deck, we cannot bring the card back to the top in one step.

However, notice that every permutation \( Y \) can be obtained in one step, from exactly \( n \) different permutations. Since every non-zero transition probability is \( \frac{1}{n} \), this implies that \( \sum \frac{P(X,Y)}{n} = 1 \).

\[
\pi(Y) = \sum_{X} \pi(X)P(X,Y)
\]

\[
\sum_{Y} P(X,Y) = 1
\]

In fact, \( \pi \) is uniform iff \( P \) is doubly stochastic.

Consider the following coupling:

Given two copies \( X_t \) and \( Y_t \) of the chain in different states, choose a position \( j \) uniformly at random from 1 to \( n \) and simultaneously move the top card into position \( j \) in both chains.

This is a valid coupling, because each chain individually acts as the original shuffling Markov chain.
Consider the reverse shuffling:
- pick a card $c$ from the deck uniformly at random.
- Move card $c$ to the top of the deck.

Define the coupling by making both $X_t$ and $Y_t$ choose the same card $c$ (which of course is not necessarily in the same position in both decks) and move it to the top.

Now, the key observation is the following: once a card has been chosen in the coupling, this card will be in the same position in both decks for the rest of time.

Try is therefore once again dominated by the coupon collector random variable for $n$ coupons. This leads to

$$T_{mix} \leq n \ln n + o(n)$$

and

$$T_2 \leq n \ln n + 2 \ln \ln n$$

### Random Transpositions

pick two cards $i$ and $j$ uniformly at random with replacement and switch cards $i$ and $j$.

Irreducible: every permutation can be expressed as a product of transpositions.
Aperiodic, since we may choose $i = j$, so the chain has self-loops.
Invertible: the random transpositions are invertible.

$$P(x,y) = P(y,x) \Rightarrow \text{uniform stationary distribution}.$$
It is easy to define a coupling using this second definition, make $X_t$ and $Y_t$ choose the same $c$ and $p$ at each step. This coupling ensures that the distance between $X$ and $Y$ is non-increasing. More explicitly, writing $d_t = d(X_t, Y_t)$ for the number of positions at which the two decks differ, we have the following case analysis:

1. If card $c$ is in the same position in both decks, then $d_{t+1} = d_t$.
2. If card $c$ is in different positions in the two decks, there are two possible subcases.
   a. If the card at position $p$ in both decks is the same, then $d_{t+1} = d_t$.
   b. Otherwise, $d_{t+1} = d_t - 1$.

Thus, we get a decrease in distance only in case 2(b), and this occurs with probability

$$\Pr\{d_{t+1} = d_t\} = \left(\frac{d_t}{n}\right)^2$$

There, the time for $d_t$ to decrease from value $d$ is stochastically dominated by a geometric random variable with mean $\left(\frac{n}{d}\right)^2$. This implies that $E[T_{xy}] \leq \sum_{i=1}^{n}\left(\frac{n}{i}\right)^2$, which is $O(n^3)$

$$\Pr\{T_{xy} > cn^2\} < \frac{E[T_{xy}]}{cn^2} = \frac{\sum_{i=1}^{n}\left(\frac{n}{i}\right)^2}{cn^2} = \frac{1}{c} \sum_{i=1}^{n} \frac{1}{i^2} = \frac{1}{2e}$$

$$\Rightarrow \tau_{max} \leq cn^2 \quad (\text{where } \frac{1}{c} \sum_{i=1}^{n} \frac{1}{i^2} = \frac{1}{2e})$$

Remarks. Actually, for the shuffle it is known that

$$\tau_{max} \sim \frac{1}{2} n \log n.$$

So our analysis in this case is off by quite a bit.

Exercise. Design a better coupling that gives $\tau_{max} = O(n \log n)$.

Background. Random walks on groups

1. A group is a set $G$ endowed with an associative operation $G \times G \to G$ and an identity id $G$ such that
for all $g \in G$,
(i) $\text{id} \cdot g = g$ and $g \cdot \text{id} = g$ 
(ii) there exists an inverse $g^{-1} \in G$ for which $g \cdot g^{-1} = g^{-1} \cdot g = \text{id}$.

2. Given a probability distribution $\mu$ on a group $(G, \cdot)$, we define the random walk on $G$ with increment distribution $\mu$ as follows: it is a Markov chain with state space $G$ and which moves by multiplying the current state on the left by a random element of $G$ selected according to $\mu$. Equivalently, the transition matrix $P$ of this chain has entries

$$P(g, h) = \mu(h)$$

for all $g, h \in G$.

3. (Proposition). Let $P$ be the transition matrix of a random walk on a finite group $G$ and let $\mu$ be the uniform probability distribution on $G$. Then $\mu$ is a stationary distribution for $P$.

Proof. Let $\mu$ be the increment distribution of the random walk. For any $g \in G$,

$$\sum_{h \in G} \mu(h) P(h, g) = \frac{1}{|G|} \sum_{k \in G} P(k, g) = \frac{1}{|G|} \sum_{k \in G} \mu(k) = \frac{1}{|G|} \mu(g)$$

For the first equality, we re-indexed by setting $x = gh^{-1}$.

4. Let $P$ be the transition matrix of a random walk on a group $G$ with increment distribution $\mu$ and let $\hat{P}$ be that of the walk on $G$ with increment distribution $\hat{\mu}$. Let $\hat{\mu}$ be the uniform distribution.
distribution on $G$. Then for any $t \geq 0$, 
\[ \| p^t(\text{id}, \cdot) - \pi \|_{TV} = \| \hat{p}^t(\text{id}, \cdot) - \pi \|_{TV} \]

**Proof:** Let $(X_t) = (\text{id}, X_t, \ldots)$ be a Markov chain with transition matrix $P$ and initial state $\text{id}$. We can write $X_t = G_1 G_2 \ldots G_t$, where the random elements $G_1, G_2, \ldots \in G$ are independent choices from the distribution $\mu$.

Similarly, let $(Y_t)$ be a chain with transition matrix $\hat{P}$, with increments $h_1, h_2, \ldots \in G$, chosen independently from $\mu$.

For any fixed elements $a_1, \ldots, a_t \in G$,
\[ P\{ G_1 = a_1, \ldots, G_t = a_t \} = \hat{P}\{ h_1 = a_1^{-1}, \ldots, h_t = a_t^{-1} \} \]
by the definition of $\hat{P}$.

Summing over all strings such that $a_1 a_2 \ldots a_t = a$ yields
\[ p^t(\text{id}, a) = \hat{p}^t(\text{id}, a^{-1}) \]

Hence,
\[ \sum_{a \in G} | p^t(\text{id}, a) - |G|^t | = \sum_{a \in G} | \hat{p}^t(\text{id}, a^{-1}) - |G|^t | = \sum_{a \in G} | \hat{p}^t(\text{id}, a) - |G|^t | \]