Coupling:

Finite space Ω

Pair of Distributions μ, ν on Ω.

Distribution ω on Product space $\Omega \times \Omega$ is a coupling of μ, ν if:

for all $i \in \Omega$, $\sum_{j \in \Omega} \omega(i,j) = \mu(i)$ $\left(\text{rows sum to } \mu\right)$

for all $j \in \Omega$, $\sum_{i \in \Omega} \omega(i,j) = \nu(i)$. $\left(\text{columns sum to } \nu\right)$

Example: $\Omega = \{1,2,3,4\}$

$\mu = \left(\frac{1}{2}, \frac{1}{4}, 0, \frac{1}{4}\right)$, $\nu = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}, 0\right)$

$$\omega = \begin{bmatrix} 1/3 & 1/12 & \frac{1}{12} & 0 \\ 0 & 1/8 & 1/8 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & \frac{1}{8} & \frac{1}{8} & 0 \end{bmatrix} \qquad \omega' = \begin{bmatrix} \frac{1}{3} & \frac{1}{12} & \frac{1}{12} & 0 \\ 0 & \frac{1}{4} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1/4 & 0 \end{bmatrix}$$

Both ω and ω' are couplings of μ, ν.

A sample from ω is a pair $(\sigma, \tau) \in \Omega^2$

If we just observe σ then $\sigma \sim \mu$

& just τ then $\tau \sim \nu$.

Coupling Lemma: For coupling ω of μ, ν, let $(\sigma, \tau) \sim \omega$.

Then: a) $d_{TV}(\mu, \nu) \leq \Pr(\sigma \neq \tau)$

b) $\exists$ coupling ω where $d_{TV}(\mu, \nu) = \Pr(\sigma \neq \tau)$.

So any coupling ω upper bounds the variation distance between μ & ν. And there's always an optimal coupling.

In the previous example, $\Pr_\omega(\sigma \neq \tau) = \frac{13}{24} \geq d_{TV}(\mu, \nu) = \frac{5}{12}$

& $\Pr_{\omega'}(\sigma \neq \tau) = \frac{5}{12} = d_{TV}(\mu, \nu)$

So ω' is optimal.

Proof of (a):

Note: for $\eta \in \Omega$, $\omega(\eta, \eta) \leq \mu(\eta)$ & $\omega(\eta, \eta) \leq \nu(\eta)$

Thus: $\omega(\eta, \eta) \leq \min\{\mu(\eta), \nu(\eta)\}$

Thus, $\Pr(\sigma = \tau) = \sum_\eta \omega(\eta, \eta) \leq \sum_\eta \min\{\mu(\eta), \nu(\eta)\}$

Therefore:
$$\Pr(\sigma \neq \tau) \geq 1 - \sum_\eta \min\{\mu(\eta), \nu(\eta)\}$$
$$= \sum_\eta \mu(\eta) - \min\{\mu(\eta), \nu(\eta)\}$$
$$= \sum_{\eta : \mu(\eta) \geq \nu(\eta)} \mu(\eta) - \nu(\eta)$$
$$= \max_{S \subset \Omega} \mu(S) - \nu(S) = d_{TV}(\mu, \nu). \qquad \blacksquare$$

Proof of (b):

Set $\omega(\eta,\eta) = \min\{\mu(\eta), \nu(\eta)\}$
& set off-diagonal entries of ω to be a product distribution of remaining.
HW exercise to complete. $\blacksquare$

Coupling of ~~Marov~~ Markov chain Defined by P on Ω.

Two chains $(X_t), (Y_t)$.

Define joint evolution (X_t, Y_t) where
X_t viewed in isolation behaves as P
& same for Y_t but they can correlate transitions.

So need that:

for all $i, j, k, l \in \Omega$,

$$\Pr(X_{t+1} = k \mid X_t = i, Y_t = j) = P(i,k)$$

$$\Pr(Y_{t+1} = l \mid X_t = i, Y_t = j) = P(j,l).$$

and if $X_t = Y_t$ then $X_{t+1} = Y_{t+1}$.

By the coupling lemma:

$$d_{TV}\left(P^t(X_0,\cdot), P^t(Y_0,\cdot)\right) \leq \Pr\left(X_t \neq Y_t\right)$$

Therefore, for $i, j \in \Omega$,

let $T^{i,j}_{couple} = \min\left\{t : \Pr\left(X_t \neq Y_t \mid X_0 = i, Y_0 = j\right) \leq \frac{1}{4}\right\}$

& $T_{couple} = \max\limits_{X_0, Y_0} T^{i,j}_{couple}$

Thus, $T_{mix} \leq T_{couple}$.

Example: Random walk on the hypercube

n-dimensional cube:

$$V = \{0,1\}^n = \text{all } n\text{-bit strings}, \ |V| = 2^n$$

$$E = \{(x,y) : x \ \& \ y \text{ differ in exactly 1 coordinate}\}$$

Markov chain:

From $X_t \in V$,

1. With prob. $\frac{1}{2}$, set $X_{t+1} = X_t$.
2. Else, pick $i \in \{1, \dots, n\}$ u.a.r. & flip $X_t(i)$

i.e. set $X_{t+1}(i) = 1 - X_t(i) = \begin{cases} 1 & \text{if } X_t(i) = 0 \\ 0 & \text{if } X_t(i) = 1 \end{cases}$

& $X_{t+1}(j) = X_t(j)$ for all $j \neq i$.

Equivalent Definition:

1. Pick $i \in \{1, \dots, n\}$ u.a.r. & $b \in \{0,1\}$ u.a.r.
2. Set $X_{t+1}(i) = b$ & $X_{t+1}(j) = X_t(j) \ \forall j \neq i$.

Consider 2 copies of random walk on hypercube (X_t) & (Y_t)

Use second form of chain: choose i, b u.a.r.

Product coupling:

if $X_t = Y_t$ choose same i, b.

else choose i, b for X_t & i', b' for Y_t independently.

Let $A_t = \{ j : X_t(j) = Y_t(j) \}$

$D_t = \{ j : X_t(j) \neq Y_t(j) \}$

If $i \in D_t$ or $i' \in D_t$ then with $\frac{1}{2}$ prob. agree on this bit afterwards

If $i \in A_t$ or $i' \in A_t$ then with $\frac{1}{2}$ prob. disagree after.

So if $|D_t| = 1$ then small chance to agree $|D_{t+1}| = 0$

but high change to disagree more $|D_{t+1}| > 1$.

Better coupling: (identity coupling)

Choose same i, b for X_t, Y_t.

if $i \in D_t$ then $|D_{t+1}| = |D_t| - 1$.

if $i \in A_t$ then $|D_{t+1}| = |D_t|$.

$$E\left[|D_{t+1}| \mid X_t \neq Y_t\right]$$

$$= D_t - \frac{D_t}{n} = D_t\left(1 - \frac{1}{n}\right)$$

$$\Pr(X_t \neq Y_t) \leq E\left[|D_t|\right] \leq |D_0|\left(1 - \frac{1}{n}\right)^t$$

$$\leq n e^{-t/n} \leq \frac{1}{4} \text{ for } t \geq n \ln(4n).$$

So $T_{mix} = O(n \log n)$.

Top-in-at-Random shuffle

n cards $\{1, 2, \ldots, n\}$
$\Omega = S_n =$ permutations on $\{1, \ldots, n\}$
$|\Omega| = n!$

Random walk:

Take top card & insert into a ~~#~~ random position.
(n choices)

Ergodic:

Aperiodic since for $\sigma \in \Omega$, $P(\sigma, \sigma) = 1/n$.
Irreducible since for $\sigma, \tau \in \Omega$,
to go $\sigma \rightsquigarrow \tau$: $O(n)$ moves to get bottom card to agree with τ
then induct.

Not symmetric, nor reversible.
What's stationary distribution π?
It's doubly stochastic:
every transition has prob. $1/n$ &
there are n transitions to every state.

<u>HW:</u> π is uniform iff P is doubly stochastic.

Hard to analyze this chain directly.
Let's look at the inverse chain:

Random-to-top:

1. Pick a card c u.a.r.
2. Move c to the top.

Coupling: X_t & Y_t choose the same card c.

After choose c_1 for $(X_0, Y_0) \to (X_1, Y_1)$ then same top card.

c_1 is always in same position in both.

After choose a card $c_2 \neq c_1$ then c_1, c_2 in same position in X_t, Y_t.

$\vdots$

So after each card is chosen ≥ 1 then we've coupled.

Coupon collectors: n coupons, each step get a random one.

T = time to get all at least once.

t_i = time to get i^{th} after collecting $(i-1)$.

$t_i = \text{Geometric}(p_i)$ where $p_i = 1 - \frac{i-1}{n}$

$$= \frac{n-i+1}{n}$$

$$E[T] = \sum_{i=1}^{n} E[t_i] = \frac{n}{n} + \frac{n}{n-1} + \cdots + \frac{n}{n}$$

$$= n\left(1 + \frac{1}{2} + \cdots + \frac{1}{n}\right)$$

$$\leq n(1 + \ln n).$$

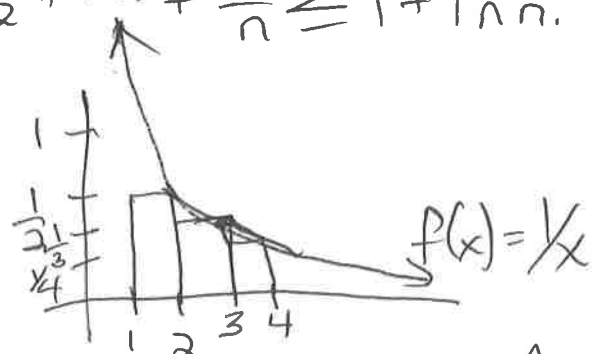
Markov's inequality: $\Pr(T > 4E[T]) \leq \frac{1}{4}$.

So mixing time is $O(n \log n)$.

Claim:

$$1 + \frac{1}{2} + \cdots + \frac{1}{n} \leq 1 + \ln n.$$

Proof:



Note: $\sum_{j=2}^{n} \frac{1}{j} \leq \int_{x=1}^{n} \frac{1}{x}\,dx = \ln x \Big|_{x=1}^{n} = \ln n$

Thus, $\frac{1}{2} + \cdots + \frac{1}{n} \leq \ln n$

$\blacksquare$