* Improving the coloring chain analysis

- One-step couplings
  - Recall our analysis of the graph-coloring MC
    - $(X_t, Y_t) \equiv$ Pick (V_t, C_t) u.a.r and apply to both
      node, color, if proper coloring
    - $d_t = d_t(X_t, Y_t) \equiv$ # of disagreeing vertices in X_t, Y_t
    - Now suppose we analyze d_t backwards from the point of coupling. Let $x, y \in \Omega$ $(\subseteq Q^{|V|})$ s.t $d(x,y) = 1$ (i.e., differ in 1 vertex).
    - Let X_1^x, Y_1^y be one step of the coupling
    - $\mathbb{P}[d(X_1^x, Y_1^y) = 0] = \mathbb{P}_{x,y}[d_1 = 0] \geq \frac{q - \Delta}{nq}$ ← Since the colors of all but one vertex is same

    (note: For general x, y, we had $\mathbb{P}_{x,y}[d_t \to d_t - 1] \geq d_t\left(\frac{q - 2\Delta}{nq}\right)$)

    $\mathbb{P}_{xy}[d_1 = 2] \leq \frac{2\Delta}{nq}$ (similar to $\frac{2 \cdot \Delta d_t}{nq}$)

    $\Rightarrow \mathbb{E}_{xy}[d_1 - 1] \leq \frac{3\Delta - q}{nq} \Rightarrow \underset{\substack{x,y \\ d(x,y)=1}}{\mathbb{E}}[d_1] \leq 1 - \left(\frac{q - 3\Delta}{nq}\right)$

  - Now for any x, y with $d(x,y) = r$, there is some sequence of colorings $x_0 = x, x_1, x_2, \ldots, x_r = y$ s.t $d(x_k, x_{k+1}) = 1$

- Since $d(\cdot,\cdot)$ is a metric, then by Δ inequality, for any x, y with $d(x,y) = r$

~~$d(x,y) \leq \sum_{i=} d(x$~~

$$\mathbb{E}\left[d(X_1^x, X_1^y)\right] \leq \sum_{k=1}^{r} \mathbb{E}\left[d(X_1^{x_{k-1}}, X_1^{x_k})\right]$$

$$\leq d(x,y)\left(1 - \frac{q-3\Delta}{nq}\right)$$

- Now for our coupling (X_t, Y_t), given any starting states x, y, let $X_{t-1} = x_{t-1}$, $Y_{t-1} = y_{t-1}$. Then (X_t^x, Y_t^y) (starting from x, y) has the same dist^n as $(X_1^{x_{t-1}}, Y_1^{y_{t-1}})$

$$\Rightarrow \mathbb{E}\left[d(X_t^x, Y_t^y) \mid X_{t-1}^x = x_{t-1}, Y_{t-1}^y = y_{t-1}\right] = \mathbb{E}\left[d(X_1^{x_{t-1}}, Y_1^{y_{t-1}})\right]$$

$$\leq d(x_{t-1}, y_{t-1})\left(1 - \frac{q-3\Delta}{nq}\right)$$

$$\Rightarrow \mathbb{E}\left[d(X_t^x, Y_t^y)\right] \leq \mathbb{E}\left[d(X_{t-1}^x, Y_{t-1}^y)\right]\left(1 - \frac{q-3\Delta}{nq}\right)$$

$$\leq d(x,y)\left(1 - \frac{q-3\Delta}{nq}\right)^t \leq n \exp\left(\frac{q-3\Delta}{nq}t\right)$$

$$\Rightarrow \mathbb{P}\left[d(X_t^x, Y_t^y) \geq 1\right] \leq n e^{-t\left(\frac{q-3\Delta}{nq}\right)}$$

$$\Rightarrow t_{mix}(1/2e) \leq \left(\frac{q}{q-3\Delta}\right) n(\ln(n) + \ln(2e))$$

* Idea for improving to $q \geq 2\Delta + 1$

- In the previous analysis, we lost a 2Δ factor in analyzing the 'bad move', i.e., $\mathbb{P}[d(X_1^x, Y_1^y) = 2] \leq \frac{2\Delta}{nq}$
- Consider an alternate coupling for all X_t, Y_t s.t. $d(X_t, Y_t) = 1$ (Let $v_0 \equiv$ disagreeing vertex)
  - Pick (V_t, C_t) u.a.r. If $V_t = v_0$ or $V_t \notin N(v_0)$, color with C_t if possible

    neighborhood
  - If $V_t \in N(v_0)$, then color X_t with C_t, and Y_t with C_t' s.t. $C_t' = \begin{cases} C_X & \text{if } C_t = C_Y \\ C_Y & \text{if } C_t = C_X \\ C_t & \text{if } C_t \notin \{C_X, C_Y\} \end{cases}$

    ($C_X \equiv$ color of v_0 in X_t, $C_Y \equiv$ color of v_0 in Y_t)
- Check this is still a valid coupling for (X_t, Y_t) when $d(X_t, Y_t) = 1$.
- Now let's try to redo our previous proof with this coupling

- As before, for good moves, we have (for $d(x,y)=1$)

$$\mathbb{P}[d(X_1^x, Y_1^y)=0] \geq \frac{q-\Delta}{nq}$$

- For bad moves, now we have an error only if $C_t = C_y$

as if $C_t = C_x$, we can not color in either chain!

$$\Rightarrow \mathbb{P}[d(X_1^x, Y_1^y)=2] \leq \frac{\Delta}{nq}$$

$\mathbb{P}[v_t \in N(v_0)]$, only one option for C_t

- Now we get $\mathbb{E}_{xy}[d_1] \leq 1 - \left(\frac{q-2\Delta}{nq}\right)$!

$d(x,y)=1$

However, we only defined the coupling for 'adjacent' X_t, Y_t (i.e, $d(X_t, Y_t)=1$)

- If we could guarantee that $\exists$ coupling (X_t, Y_t) that matches the one above for all x,y s.t $d(x,y)=1$, then we are done (extension thm)
  - this is what we get from $\boxed{\text{path coupling}}$

- Suppose we are given a (connected, undirected) graph $H(\Omega, E_0)$ on the state space Ω of a MC, with edge lengths $l(x,y)$ (s.t $l(x,y) \geq 1$ $\forall x,y \in E_0$). The path metric on Ω is

$$\rho(x,y) = \min\{\text{length of } \xi \mid \xi \text{ path from } x \to y \text{ in } H\}$$

(Assume H is a pre-metric, i.e. $\rho(x,y) = l(x,y)$ $\forall (x,y) \in E_0$)

- **Note** - H may be different from the graph G associated with MC (Ω, P)

- $\boxed{\text{Thm}}$. Suppose for every edge $(x,y) \in E_0$, $\exists$ a coupling (X_1, Y_1) for $P(x,\cdot), P(y,\cdot)$ s.t

$$\mathbb{E}_{x,y}\left[\rho(X_1,Y_1)\right] \leq \rho(x,y)\, e^{-\alpha} = \ell(x,y) e^{-\alpha}$$

Then for any $(x,y) \in \Omega$, we have

$$\mathbb{E}_{x,y}\left[\rho(X_1,Y_1)\right] \leq e^{-\alpha} \rho(x,y)$$

Pf 1 - Direct construction of coupling

- Given x, y arbitrary states in Ω, and a pre-metric, let $x = z_0 \to z_1 \to \dots \to z_r = y$ be a shortest path.
- We are given couplings $\mathbb{P}[X_1^{z_k}, Y_1^{z_{k+1}}]$ for all z_k, z_{k+1} (since $\rho(z_k, z_{k+1}) =$ edge in pre-metric)
- - Sample $(\hat{X}_1^{z_0}, \hat{X}_1^{z_1})$ from coupling $\mathbb{P}$ for $P(z_0,\cdot), P(z_1,\cdot)$
  - Conditioned on $\hat{X}_1^{z_1}$, sample $(\hat{X}_1^{z_1}, \hat{X}_1^{z_2})$ from next coupling
  - Repeat to get $\hat{X}_1^{z_2}, \hat{X}_1^{z_3}, \dots \hat{X}_1^{z_r} = \hat{X}_1^{y}$
  - $(\hat{X}_1^{z_0}, \hat{X}_1^{z_r}) = (X_1^x, X_1^y)$ is a coupling for $P(x,\cdot), P(y,\cdot)$

- To show this is a valid coupling, we show by induction that $\mathbb{P}[\hat{X}_1^{z_k} = w] = P(z_k, w)$
  - For $X_1^{z_0}, X_1^{z_1}$, this is true (since its a coupling)
  - Assume true for $\hat{X}_1^{z_k}$. Then

$$\mathbb{P}[\hat{X}_1^{z_{k+1}} = w] = \sum_{w' \in \Omega} \mathbb{P}[\hat{X}_1^{z_k} = w'] \cdot \frac{\mathbb{P}[\hat{X}_1^{z_k} = w', \hat{X}_1^{z_{k+1}} = w]}{\sum_{s \in \Omega} \mathbb{P}[\hat{X}_1^{z_k} = w', \hat{X}_1^{z_{k+1}} = s]}$$

$$= \sum_{w' \in \Omega} \underbrace{P(z_k, w')}_{\text{induction}} \cdot \frac{P(\hat{X}_1^{z_k} = w', \hat{X}_1^{z_{k+1}} = w)}{\underbrace{P(z_k, w')}_{(\hat{X}_1^{z_k}, \hat{X}_1^{z_{k+1}}) \text{ valid coupling}}}$$

$$= \sum_{w' \in \Omega} \mathbb{P}[\hat{X}_1^{z_k} = w', \hat{X}_1^{z_{k+1}} = w] = P(z_{k+1}, w)$$

- $$\mathbb{E}[\rho(X_1^x, X_1^y)] \leq \mathbb{E}\left[\sum_{k=0}^{n-1} \rho(X_1^{z_k}, X_1^{z_{k+1}})\right]$$

$$\leq \sum_{k=0}^{n-1} \mathbb{E}\left[\rho(X_1^{z_k}, X_1^{z_{k+1}})\right] \leq \sum_{k=0}^{n-1} e^{-\alpha} \rho(z_k, z_{k+1})$$

$$= e^{-\alpha} \sum_{k=0}^{n-1} \rho(z_k, z_{k-1}) = e^{-\alpha} \rho(x, y)$$

Corollary - If we have a path coupling as above, then

$$t_{mix}(\varepsilon) \leq \left\lceil \frac{\ln(D) - \ln(\varepsilon)}{\alpha} \right\rceil$$, where $D \equiv$ diameter of Ω under metric

(Note - $l(x,y) \geq 1$ for nbs (x,y))

Pf 2 (Via the Wasserstein distance)

- For any given metric $\rho(\cdot,\cdot)$ over Ω, the following quantity is called the Wasserstein metric or transportation metric (or Kantorovich metric) over distributions on Ω

$$d_\rho(\mu,\nu) = \inf\{\mathbb{E}[\rho(X,Y)] \mid (X,Y) \text{ is a coupling of } \mu,\nu\}$$

- Remarks on $d_\rho(\cdot,\cdot)$
  - If $\rho(x,y) = \mathbb{1}\{x \neq y\}$, then $d_\rho(\mu,\nu) = d_{TV}(\mu,\nu)$
  - For any X,Y,Z, we have $d_\rho(X,Z) \leq d_\rho(X,Y) + d_\rho(Y,Z)$ (Δ-inequality for the Wasserstein metric)
  - The set of distn (μ,ν) on $\Omega\times\Omega$ is a compact subset of $\mathbb{R}^{|\Omega|^2}$ (i.e, the $|\Omega|^2$ simplex).
  - The set of distn in $\Omega_1\times\Omega_2$ s.t the projection on $\Omega_1 = \mu$, on $\Omega_2 = \nu$ is a compact subset of the $|\Omega|^2$ simplex
  - $d_\rho(\mu,\nu) = \inf\Big\{\sum_{(x,y)\in\Omega\times\Omega} \rho(x,y)\underbrace{q(x,y)}_{\text{joint}} \;\Big|\; \underbrace{q(\cdot,\Omega)=\mu,\ q(\Omega,\cdot)=\nu}_{\text{marginals}}\Big\}$

Note - the fn $q \longmapsto \sum_{(x,y)\in\Omega\times\Omega} \rho(x,y)\,q(x,y)$ is continuous on set of couplings

$\Rightarrow$ There exists q^* s.t $\sum_{(x,y)\in\Omega\times\Omega} \rho(x,y)\, q^*(x,y) = d_\rho(\mu,\nu)$

This is called the <u>optimal coupling (X_ρ^*, Y_ρ^*)</u> for $d_\rho(\mu,\nu)$. Note $\mathbb{E}\left[\rho(X_\rho^*, Y_\rho^*)\right] = d_\rho(\mu,\nu)$

- Returning to path coupling, we are given a metric ρ over Ω. Now fix $x, y \in \Omega$ and let $(x = z_0, z_1, \dots, z_k = y)$ be a shortest path.

  - by Δ inequality for d_ρ, we have

$$d_\rho(P(x,\cdot), P(y,\cdot)) \le \sum_{k=0}^{n} d_\rho(P(x_k,\cdot), P(x_{k+1},\cdot))$$

  - Now if for any edge (a,b), we have $\ell(a,b) = \rho(a,b) \le 1$ and by assumption, $d_\rho(P(a,\cdot), P(b,\cdot)) \le e^{-\alpha}\rho(a,b)$

$$\Rightarrow d_\rho(P(x,\cdot), P(y,\cdot)) \le e^{-\alpha} \sum_{k=0}^{n} \rho(z_k, z_{k+1}) = e^{-\alpha}\rho(x,y)$$

  - Moreover, we know $\exists$ coupling (X_1^x, X_1^y) s.t

$$d_\rho(X_1^x, X_1^y) = \inf \mathbb{E}\left[\rho(X_1^x, Y_1^y)\right]$$

$$\Rightarrow \mathbb{E}\left[\rho(X_1^x, Y_1^y)\right] \le e^{-\alpha}\rho(x,y)$$

- **Balanced coupling** - If we have a pre-metric $\rho(\cdot,\cdot)$ on Ω, and a coupling (X,Y) on adjacent states x,y s.t $\mathbb{E}\left[\rho(X_1^x, X_1^y)\right] \leq (1-\alpha)\rho(x,y)$

  - If $\alpha < 1$, $t_{mix}(\varepsilon) = O(\alpha^{-1}\ln(D))$ where $D = \max_{x,y} \rho(x,y)$
  - If $\alpha = 1$, then $t_{mix}(\varepsilon) = O(D^2/\beta)$, where

$$\beta = \min_{\underbrace{x,y\in\Omega^2}_{\text{not only neighbors}}} \mathbb{E}\left[\left(\rho(X_1^x, X_1^y) - \rho(x,y)\right)^2\right]$$

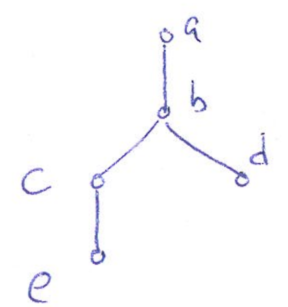
(<u>Note</u> - $\beta \geq \min_{x,y\in\Omega^2} \mathbb{P}\left[\,|\rho(X_1^x, X_1^y) - \rho(x,y)| \geq 1\right]$)

Pf via Martingale Optional Stopping Thm (later in course)

<u>Application</u> - Sampling linear extensions of a partial order

- Input - a partial order $\preceq$ on $V = \{1, 2, \ldots, n\}$
- Output - a linear extension (i.e. a total order $\leq$ on V that respects $\preceq$ in the sense $x \preceq y \Rightarrow x \leq y$) u.a.r

Ex



$\equiv$

$a \leq b \leq c \leq d \leq e$

$a \leq b \leq c \leq e \leq d$

$a \leq b \leq d \leq c \leq e$

- linear extension $\Omega \subseteq$ Set of permutations $\sigma(\cdot)$ on V

- Applications in combinatorics, ranking, approx sorting, decision theory, etc.

- Markov chain - wp 1/2, do nothing
  - wp 1/2, pick random $p \in \{1, 2, \dots, n-1\}$ u.a.r. and exchange $\sigma(p), \sigma(p+1)$ if possible (i.e., if new order is a valid extension)

- MC is aperiodic, symmetric (so π is uniform). Check that it is irreducible.

- Analysis using path coupling.
  - Pre-metric: $x, y \in \Omega$ are adjacent iff ~~$d(x,y) = 2$~~, they differ in 2 posns $1 \le i < j \le n$
    - distance between adjacent pair $= j-i$
  - natural extension to metric (may be complicated)

- Coupling for adjacent pairs - Given adjacent x, y differing in $i < j$
  - Case 1 - If $j \neq i+1$: Same as MC
  - Case 2 - If $j = i+1$: i) wp $\frac{1}{2(n-1)}$ - do nothing in X, swap $i, i+1$ in Y
    ii) wp $\frac{1}{2(n-1)}$ - do nothing in Y, swap $i, i+1$ in X

iii) wp $\frac{n-2}{2(n-1)}$ - do nothing in both

iv) wp $\frac{1}{2}$, choose p u.a.r. in $\{1,\dots,n-1\} \setminus \{i\}$ and swap $\sigma(p), \sigma(p+1)$ if possible.

- Analysis of coupling

- If $p \notin \{i-1, i, j-1, j\} \Rightarrow d(X_1^x, X_1^y) = d(x,y) = j-i$
- If $p = i-1$ or $j \Rightarrow d(X_1^x, X_1^y) \leq d(x,y)+1 = j-i+1$
  - Happens with prob $\left(\frac{1}{(n-1)\cdot 2}\right)\cdot 2 = \frac{1}{n-1}$
- If $p = i$ or $p = j-1$
  - If $j-i=1$, then $d(X_1^x, X_1^y) = 0 = d(x,y)-1$
  - If $j-i>1$: Note that $\sigma(i), \sigma(i+1)$ and $\sigma(j-1), \sigma(j)$ must be incomparable in $\preccurlyeq$ $\Rightarrow$ exchange is legal!
  - $d(X_1^x, X_1^y) = d(x,y)-1$ wp $\frac{1}{n-1}$
- Thus $\mathbb{E}[d(\underbrace{X_1^x, X_1^y}_{\text{adjacent } x,y})] \leq d(x,y)$
- By path coupling, $t_{mix}(\varepsilon) = O(\beta^{-1} D^2) = O(n^5)$, as $D \leq \binom{n}{2} = O(n^2)$, and $\beta = \min_{x,y\in\Omega} \mathbb{P}[|d(X_1^x, X_1^y) - d(x,y)| \geq 1] \leq \frac{1}{n}$