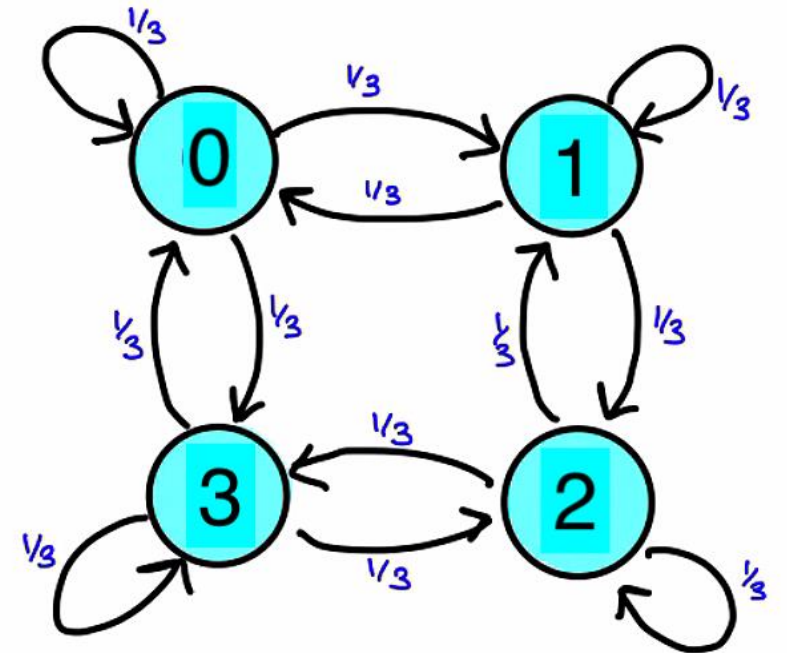


# Class 13

Markov Chains I

# Warm-Up

- What is the transition matrix for this Markov Chain?
- How would you compute  $\Pr[ X_t = 2 \mid X_0 = 0 ]$  for  $t=1, 2, \dots, 100$ ?
- As  $t \rightarrow \infty$ , what do you think this probability should tend to?



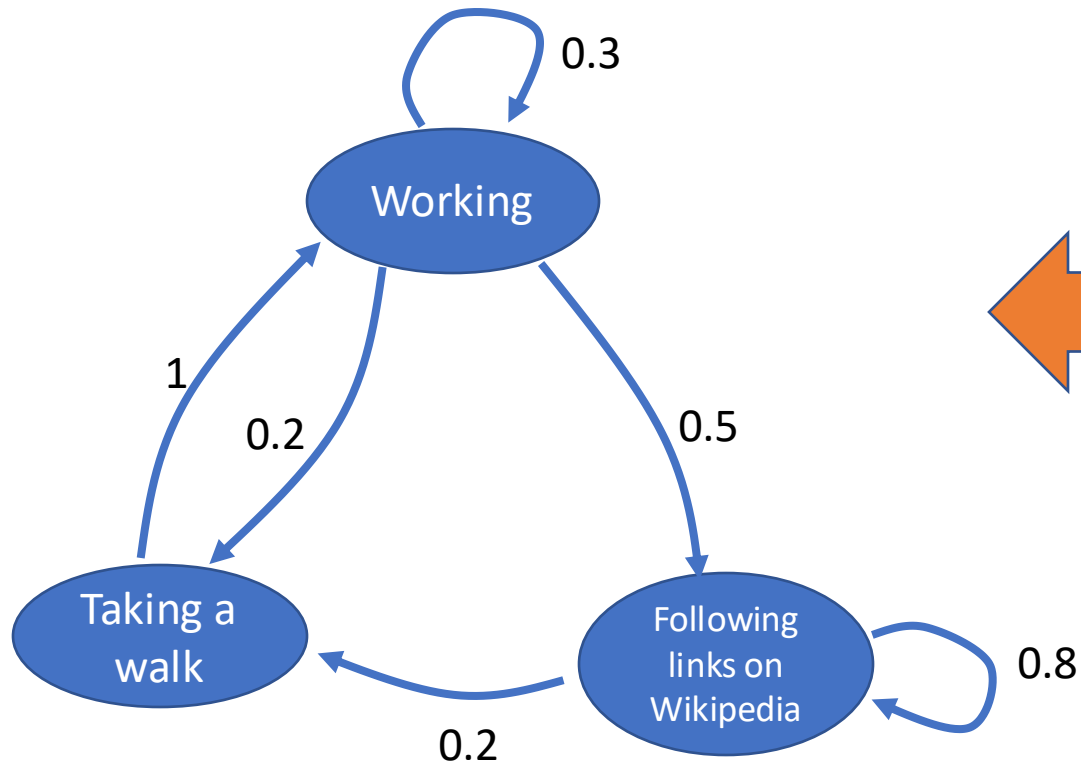
# Announcements

- HW5 due Friday!

# Recap

(Time homogeneous, finite) Markov Chains!

Memorylessness:  $\Pr[X_t = a \mid X_0, \dots, X_{t-1}] = \Pr[X_t = a \mid X_{t-1}]$



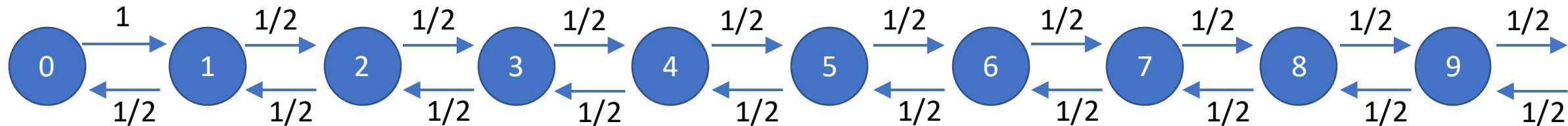
	Working	Walking	Wikipedia
Working	0.3	0.2	0.5
Walking	1	0	0
Wikipedia	0	0.2	0.8

$$\text{e.g., } \varphi = (x_1 \vee x_2) \wedge (\bar{x}_1 \vee x_3) \wedge \cdots \wedge (x_5 \vee \bar{x}_7)$$

# Recap

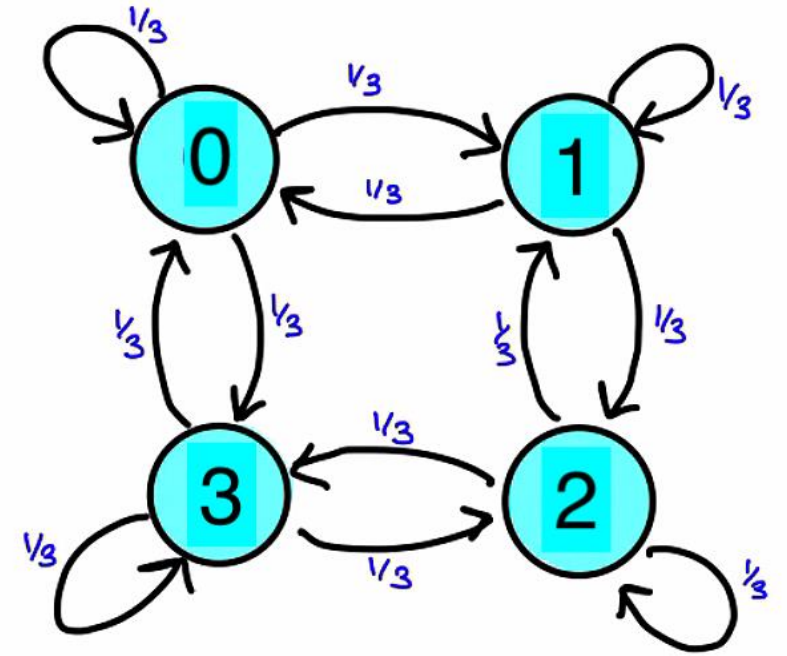
## Randomized algorithm for 2SAT!

- Algorithm:
  - While not done:
    - Find an unsatisfied clause, flip one of the variables at random.
- Analysis: This Markov chain hits  $n$  in no more than  $100n^2$  steps whp:



# Questions?

Mini-lectures, quiz, warm-up?



# Warm-Up

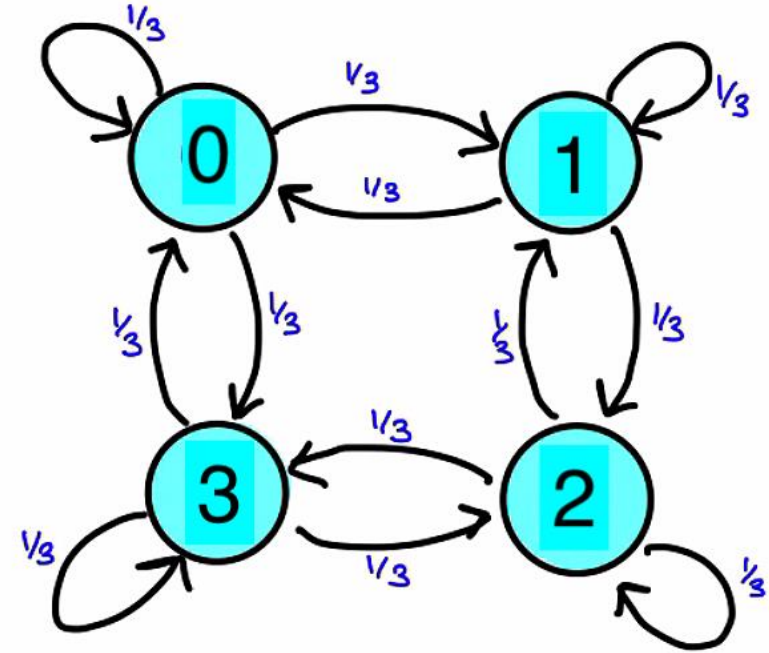
- Transition matrix?

$$P = \begin{pmatrix} 1/3 & 1/3 & 0 & 1/3 \\ 1/3 & 1/3 & 1/3 & 0 \\ 0 & 1/3 & 1/3 & 1/3 \\ 1/3 & 0 & 1/3 & 1/3 \end{pmatrix}$$

- $\Pr[X_t = 2 \mid X_0 = 0]$  for  $t=1, 2, \dots, 100$ ?

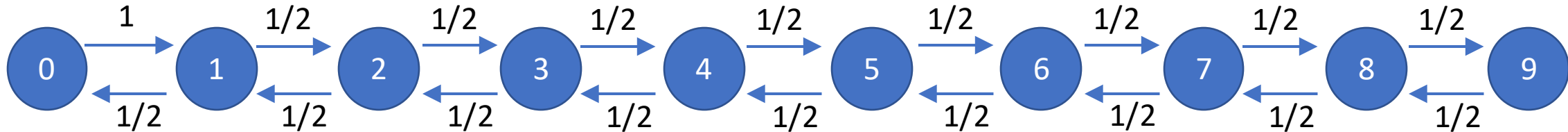
$$(1,0,0,0) \cdot P^t \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}$$

- As  $t \rightarrow \infty$ , what do you think this prob. should go to?  $\frac{1}{4}$ ?



# In the mini-lectures

- We saw one way to analyze a chain that looked like this:



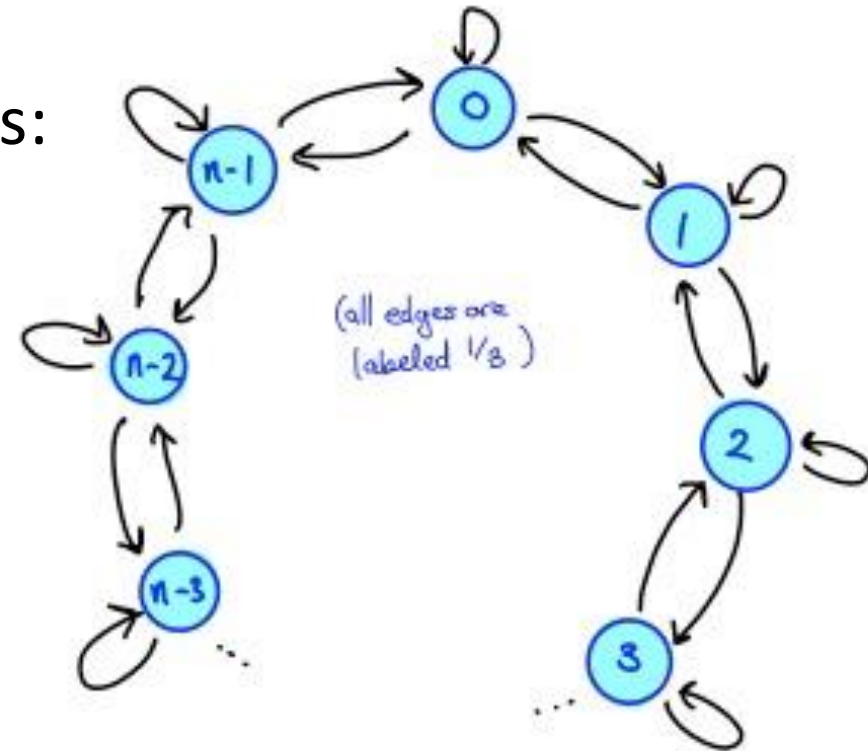
- Whp, it took  $\Theta(n^2)$  steps to reach  $n$ .

# Today

- We will see another way to use the transition matrix to analyze (certain nice) Markov chains.

- We'll analyze a Markov chain that looks like this:

- What should happen if the Markov chain runs for long enough?
- What's your intuition about how long "long enough" is?



## How long do you think it will take for the Markov chain to get "close" to uniform?

$O(n)$  steps

0%

$O(n^2)$  steps

0%

$O(n^c)$  steps, for some  $c \geq 3$

0%

$\exp(n)$  steps

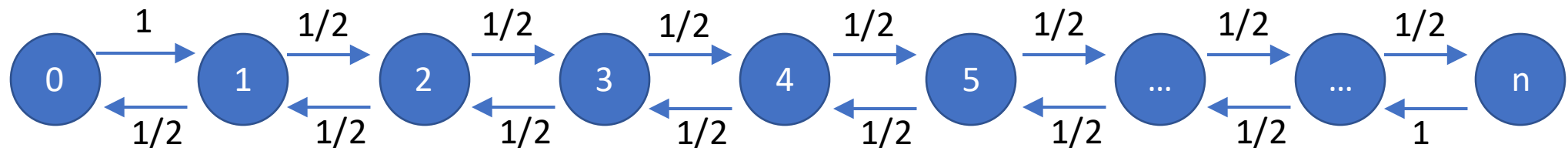
0%

No idea

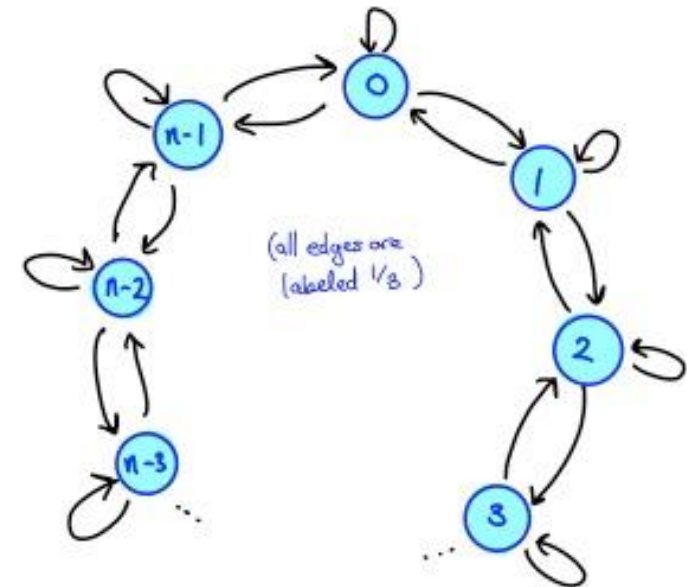
0%

# Intuition?

- Maybe  $O(n^2)$ ?



- This  $\uparrow$  took  $\Theta(n^2)$  to get to  $n$ ...
- Same logic should apply to the number of steps you need to “wrap around” the cycle once...
- You can't be close to uniform until you can at least get around the cycle...



# Today...

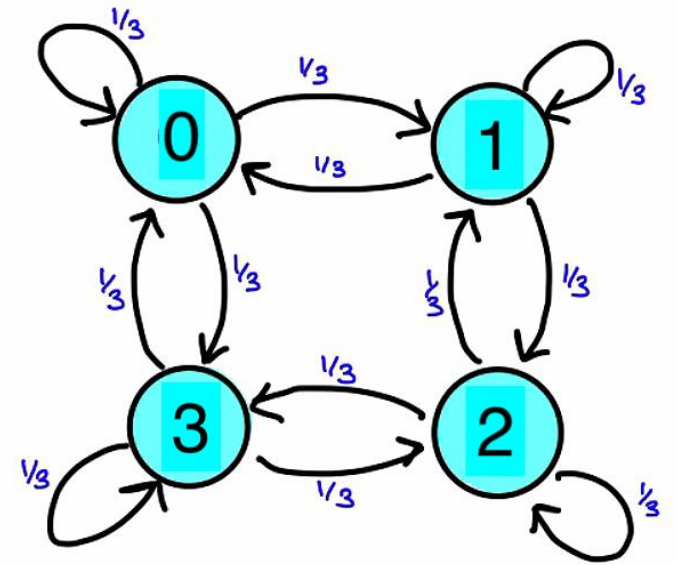
- We'll see a different way to analyze this: **spectral methods!**

# First group work!

- Show that:

$$F = \frac{1}{2} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -i & -1 & i \\ 1 & -1 & 1 & -1 \\ 1 & i & -1 & -i \end{pmatrix}$$

$$P = F \cdot \begin{pmatrix} 1 & & & \\ & 1/3 & & \\ & & -1/3 & \\ & & & 1/3 \end{pmatrix} \cdot F^*.$$



- Use this to find a nicer way of computing  $\Pr[ X_t = 2 \mid X_0 = 0 ]$

Show

$$P = F \cdot \begin{pmatrix} 1 & & & \\ & 1/3 & & \\ & & -1/3 & \\ & & & 1/3 \end{pmatrix} \cdot F^*.$$

$$F = \frac{1}{2} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -i & -1 & i \\ 1 & -1 & 1 & -1 \\ 1 & i & -1 & -i \end{pmatrix}$$

$$\begin{pmatrix} 1/3 & 1/3 & 0 & 1/3 \\ 1/3 & 1/3 & 1/3 & 0 \\ 0 & 1/3 & 1/3 & 1/3 \\ 1/3 & 0 & 1/3 & 1/3 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

$$\begin{pmatrix} 1/3 & 1/3 & 0 & 1/3 \\ 1/3 & 1/3 & 1/3 & 0 \\ 0 & 1/3 & 1/3 & 1/3 \\ 1/3 & 0 & 1/3 & 1/3 \end{pmatrix} \begin{pmatrix} 1 \\ -i \\ -1 \\ i \end{pmatrix} = \frac{1}{3} \cdot \begin{pmatrix} 1 \\ -i \\ -1 \\ i \end{pmatrix}$$

$$\begin{pmatrix} 1/3 & 1/3 & 0 & 1/3 \\ 1/3 & 1/3 & 1/3 & 0 \\ 0 & 1/3 & 1/3 & 1/3 \\ 1/3 & 0 & 1/3 & 1/3 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix} = \frac{-1}{3} \cdot \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix}$$

$$\begin{pmatrix} 1/3 & 1/3 & 0 & 1/3 \\ 1/3 & 1/3 & 1/3 & 0 \\ 0 & 1/3 & 1/3 & 1/3 \\ 1/3 & 0 & 1/3 & 1/3 \end{pmatrix} \begin{pmatrix} 1 \\ i \\ -1 \\ -i \end{pmatrix} = \frac{1}{3} \cdot \begin{pmatrix} 1 \\ i \\ -1 \\ -i \end{pmatrix}$$

Show

$$P = F \cdot \begin{pmatrix} 1 & & & \\ & 1/3 & & \\ & & -1/3 & \\ & & & 1/3 \end{pmatrix} \cdot F^*.$$

$$F = \frac{1}{2} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -i & -1 & i \\ 1 & -1 & 1 & -1 \\ 1 & i & -1 & -i \end{pmatrix}$$

$$\begin{pmatrix} 1/3 & 1/3 & 0 & 1/3 \\ 1/3 & 1/3 & 1/3 & 0 \\ 0 & 1/3 & 1/3 & 1/3 \\ 1/3 & 0 & 1/3 & 1/3 \end{pmatrix} \cdot \frac{1}{2} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -i & -1 & i \\ 1 & -1 & 1 & -1 \\ 1 & i & -1 & -i \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -i & -1 & i \\ 1 & -1 & 1 & -1 \\ 1 & i & -1 & -i \end{pmatrix} \cdot \begin{pmatrix} 1 & & & \\ & 1/3 & & \\ & & -1/3 & \\ & & & 1/3 \end{pmatrix}$$

$$P \cdot F = F \cdot \Lambda$$

$$P = F\Lambda F^*$$

A nicer way of computing  
 $\Pr[X_t = 2 \mid X_0 = 0]$

$$(1,0,0,0) \cdot P^t \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}$$

$$P^t = (F \Lambda F^*)^t$$

$$= F \Lambda F^* F \Lambda F^* F \Lambda F^* F \Lambda F^* \dots F \Lambda F^* F \Lambda F^*$$

$$= F \Lambda^t F^* = F \begin{pmatrix} 1 & & & \\ & (1/3)^t & & \\ & & (-1/3)^t & \\ & & & (1/3)^t \end{pmatrix} F^*$$

$$F = \frac{1}{2} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -i & -1 & i \\ 1 & -1 & 1 & -1 \\ 1 & i & -1 & -i \end{pmatrix}$$

$$P = F \cdot \begin{pmatrix} 1 & & & \\ & 1/3 & & \\ & & -1/3 & \\ & & & 1/3 \end{pmatrix} \cdot F^*$$

A nicer way of computing  
 $\Pr[X_t = 2 \mid X_0 = 0]$

$$F = \frac{1}{2} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -i & -1 & i \\ 1 & -1 & 1 & -1 \\ 1 & i & -1 & -i \end{pmatrix}$$

$$P = F \cdot \begin{pmatrix} 1 & & & \\ & 1/3 & & \\ & & -1/3 & \\ & & & 1/3 \end{pmatrix} \cdot F^*.$$

$$(1,0,0,0) \cdot F \begin{pmatrix} 1 & & & \\ & (1/3)^t & & \\ & & (-1/3)^t & \\ & & & (1/3)^t \end{pmatrix} F^* \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}$$

$$= \frac{1}{4} \cdot (1,1,1,1) \begin{pmatrix} 1 & & & \\ & (1/3)^t & & \\ & & (-1/3)^t & \\ & & & (1/3)^t \end{pmatrix} \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix}$$

$$= \frac{1}{4} - \frac{1}{2} \left(\frac{1}{3}\right)^t + \frac{1}{4} \left(-\frac{1}{3}\right)^t = \frac{1}{4} + O(3^{-t})$$

So indeed this converges to  $\frac{1}{4}$ !  
 And very quickly!

# Aside: Spectral analysis

- Say that a transition matrix  $P$  is symmetric.

# Aside: Spectral analysis

- Say that a transition matrix  $P$  is symmetric.
- Write  $P = V\Lambda V^*$  where  $V$  is unitary and  $\Lambda$  is diagonal with real values on the diagonal.

# Aside: Spectral analysis

- Say that a transition matrix  $P$  is symmetric.
- Write  $P = V\Lambda V^*$  where  $V$  is Hermitian and  $\Lambda$  is diagonal with real values on the diagonal.
- $P^t = V\Lambda^t V^*$

# Aside: Spectral analysis

- Say that a transition matrix  $P$  is symmetric.
- Write  $P = V\Lambda V^*$  where  $V$  is Hermitian and  $\Lambda$  is diagonal with real values on the diagonal.

- $P^t = V\Lambda^t V^*$

- What can we say about  $\Pr[X_t = j \mid X_0 = i] = e_i^T P^t e_j$ ?

$$= (e_i^T V) \begin{pmatrix} \lambda_1^t & & & \\ & \lambda_2^t & & \\ & & \ddots & \\ & & & \lambda_n^t \end{pmatrix} (V^* e_j)$$

- What is  $\lambda_1$ ? (The largest eigenvalue of  $P$ ?)

$$\begin{pmatrix} 1/3 & 1/3 & 0 & 1/3 \\ 1/3 & 1/3 & 1/3 & 0 \\ 0 & 1/3 & 1/3 & 1/3 \\ 1/3 & 0 & 1/3 & 1/3 \end{pmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

$\lambda_1 = 1$ , because every row/column is a probability distribution.

# Aside: Spectral analysis

- Say that a transition matrix  $P$  is symmetric.
- Write  $P = V\Lambda V^*$  where  $V$  is Hermitian and  $\Lambda$  is diagonal with real values on the diagonal.

- $P^t = V\Lambda^t V^*$

- What can we say about  $\Pr[X_t = j \mid X_0 = i] = e_i^T P^t e_j$ ?

$$= (e_i^T V) \begin{pmatrix} 1 & & & \\ & \lambda_2^t & & \\ & & \ddots & \\ & & & \lambda_n^t \end{pmatrix} (V^* e_j)$$

All of these have  
 $|\lambda_i| \leq 1$

$\Rightarrow$  As  $t \rightarrow \infty$ , all  
these get small\*

\*Assuming that  $|\lambda_i| < 1$  ...

# Aside: Spectral analysis

- Say that a transition matrix  $P$  is symmetric.
- Write  $P = V\Lambda V^*$  where  $V$  is Hermitian and  $\Lambda$  is diagonal with real values on the diagonal.

- $P^t = V\Lambda^t V^*$

- What can we say about  $\Pr[X_t = j \mid X_0 = i] = e_i^T P^t e_j$ ?

$$= (e_i^T V) \begin{pmatrix} 1 & & & \\ & \text{tiny} & & \\ & & \ddots & \\ & & & \text{tiny} \end{pmatrix} (V^* e_j)$$

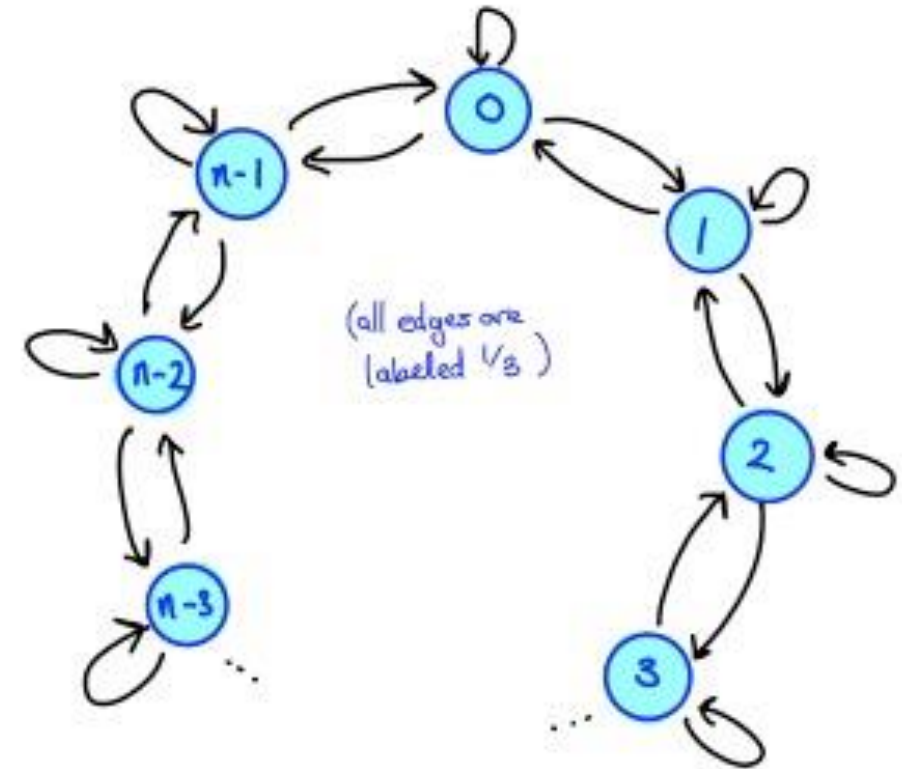
$$= v_{1,i} \cdot v_{1,j} + \text{tiny} \approx v_{1,i} \cdot v_{1,j}$$

So we can compute this! (Up to some small error term that we can hopefully quantify!)

# Back to group work!

Do the same thing for a large cycle.

- We'll do Part 1 together since the math is a bit gross.







# Rest of the group work! [See handout for more details and hints](#)

2. Come up with an expression for  $\Pr[X_t = 0 | X_0 = 0]$
3. Convince yourself that this  $\rightarrow \frac{1}{n}$  as  $t \rightarrow \infty$
4. How large does  $t$  have to be before

$$\Pr[X_t = 0 | X_0 = 0] = \frac{1+o(1)}{n} ?$$

$$P = F_n \Lambda F_n^*$$
$$F_n = \frac{1}{\sqrt{n}} \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & \omega & \omega^2 & \omega^3 & \dots & \omega^{n-1} \\ 1 & \omega^2 & \omega^4 & \omega^6 & & \\ 1 & \omega^3 & \omega^6 & \vdots & & \\ 1 & \vdots & & & \ddots & \\ 1 & \omega^{n-1} & & & & \omega^{(n-1)^2} \end{pmatrix}$$

$\Lambda$  is diagonal and  $\Lambda_{j,j} = \frac{1+2 \cos\left(\frac{2\pi j}{n}\right)}{3}$

$\Lambda$  is diagonal and  $\Lambda_{j,j} = \frac{1+2 \cos\left(\frac{2\pi j}{n}\right)}{3}$

Part 2:  $\Pr[X_t = 0 \mid X_0 = 0]$

$$\Pr[X_t = 0 \mid X_0 = 0] = e_0^T F_n \Lambda^t F_n^* e_0$$

$$\frac{1}{\sqrt{n}} (1, 1, 1, \dots, 1) \begin{pmatrix} \Lambda_{0,0}^t & & & & \\ & \Lambda_{1,1}^t & & & \\ & & \ddots & & \\ & & & \Lambda_{n-2,n-2}^t & \\ & & & & \Lambda_{n-1,n-1}^t \end{pmatrix} \frac{1}{\sqrt{n}} \begin{pmatrix} 1 \\ 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$$

$$= \frac{1}{n} \sum_{j=0}^{n-1} \left( \frac{1 + 2 \cos\left(\frac{2\pi j}{n}\right)}{3} \right)^t$$

## Part 3: What happens as $t$ gets big?

$$\Pr[X_t = 0 | X_0 = 0] = \frac{1}{n} \sum_{j=0}^{n-1} \left( \frac{1 + 2 \cos\left(\frac{2\pi j}{n}\right)}{3} \right)^t$$
$$\rightarrow \frac{1}{n} \cdot 1$$

When  $j = 0$ :  $\left( \frac{1 + 2 \cos\left(\frac{2\pi j}{n}\right)}{3} \right)^t = \left( \frac{1 + 2}{3} \right)^t = 1$

When  $j = 1, \dots, n-1$ :  $\left( \frac{1 + 2 \cdot (<1)}{3} \right)^t = \left( \frac{1 + (<2)}{3} \right)^t \rightarrow 0$  as  $t \rightarrow \infty$

## Part 4: How big does $t$ need to be?

Aka, how big does  $t$  need to be before  $\left(\frac{1+2 \cos\left(\frac{2\pi j}{n}\right)}{3}\right)^t$  is really small for all  $1 \leq j \leq n - 1$ ?

$$\left(\frac{1+2 \cos\left(\frac{2\pi j}{n}\right)}{3}\right)^t \approx \left(\frac{1}{3} \cdot \left(1 + 2 \left(1 - \frac{1}{2} \left(\frac{2\pi j}{n}\right)^2\right)\right)\right)^t$$

$$= \left(1 - \frac{1}{3} \left(\frac{2\pi j}{n}\right)^2\right)^t$$

$$\approx \exp\left(-\frac{4\pi^2 j^2}{3} \cdot \frac{t}{n^2}\right)$$

If this is going to be small  
for all  $j$ , we need  $t \gg n^2$

This matches out intuition  
from the beginning! 😊

# Take-aways

- If we can diagonalize the transition matrix, it can make analyzing a Markov chain easier.
  - This is called “spectral analysis.”
- If the second eigenvalue of a (symmetric) transition matrix is bounded away from 1, then the Markov chain “mixes” quickly.
  - We’ll see what it means to “mix quickly” more formally next time!