Lesson 3
Dynamical Systems
3-4: Markov Chains

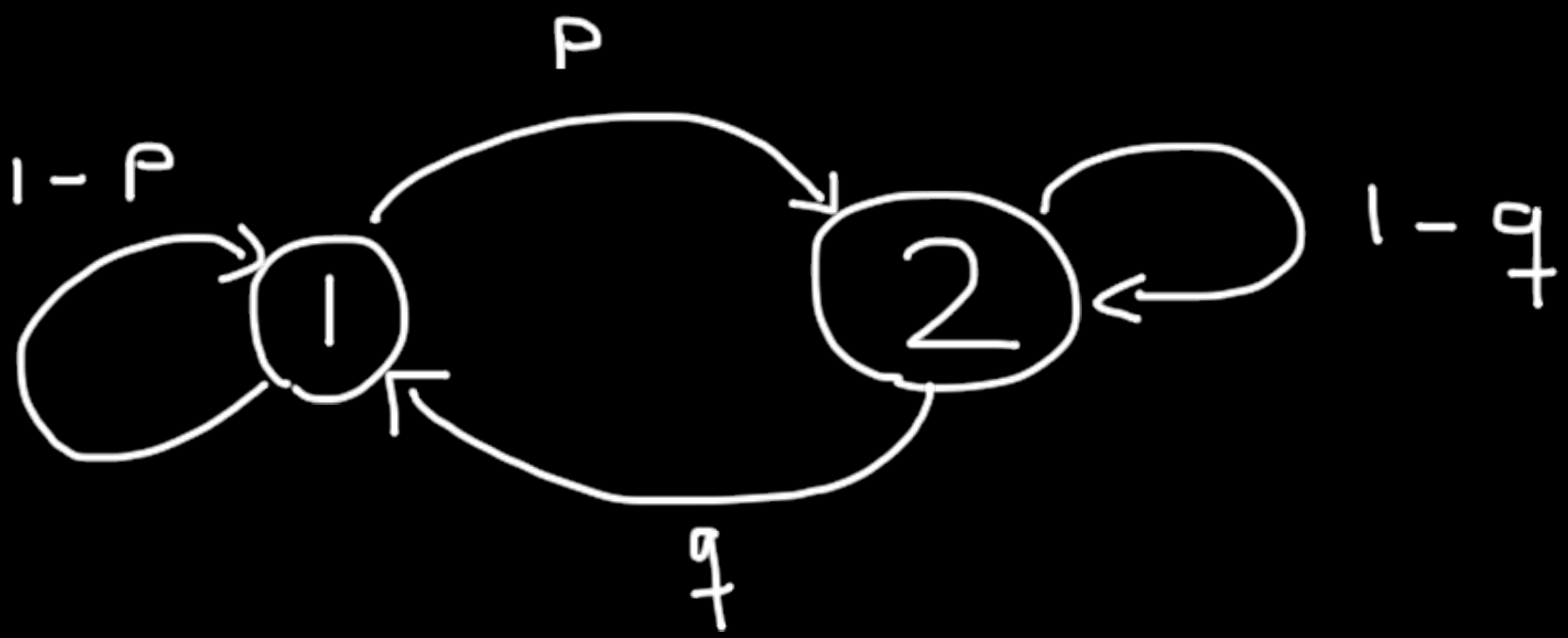
Jan Reimann

Math 574, Topics in Logic
Penn State, Spring 2014

Markov Chains

Markov Shift: Infinite paths through a finite state system (finite directed graph).

Markov Chains: Random walks through a finite directed graph with transition probabilities.

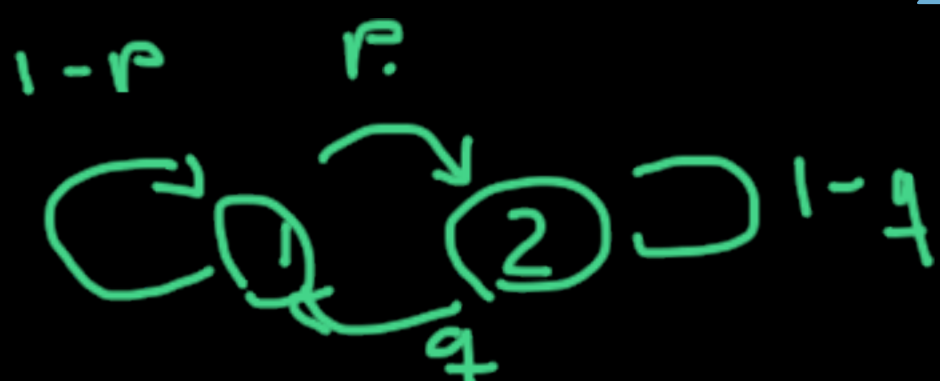


Probabilistic Transition Matrices

Mathematical description using probabilistic matrices:

Let $M = M^{(G)}$ be a stochastic $k \times k$-matrix, i.e. $M_{ij} \in [0,1]$ and for each $1 \leqslant i \leqslant k$,

$$\sum_j M_{ij} = 1.$$



$$\begin{bmatrix} 1-p & p \\ q & 1-q \end{bmatrix}$$

- M_{ij} represents the probability that we pass from state i to state j.
- $(M^n)_{ij}$ then yields the probability that we go from state i to state j in n steps.

Markov Chains

Given a stochastic matrix M, we can define a process (X_n) as follows:

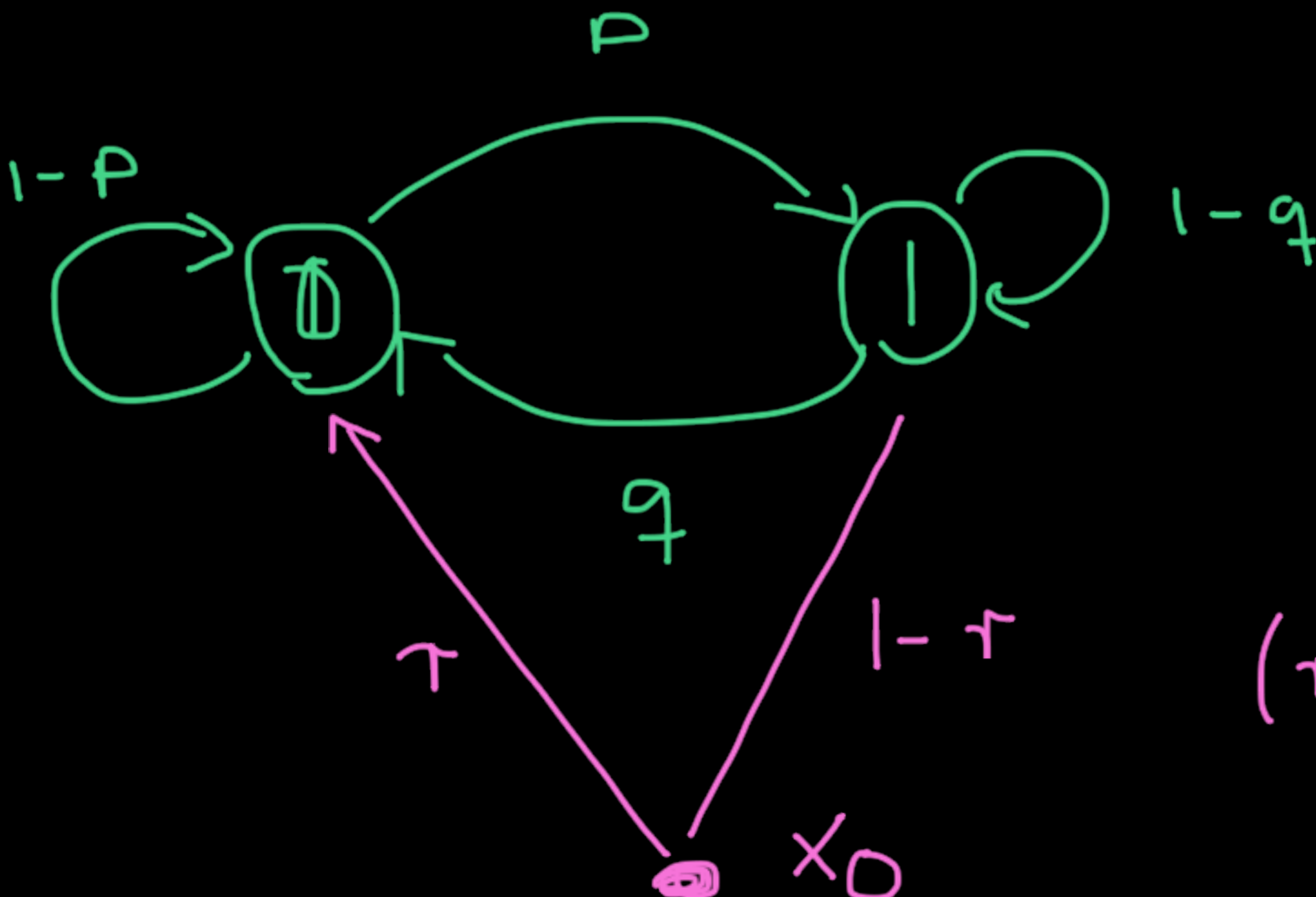
- X_0 is distributed according to some initial distribution

$$\text{Prob}(X_0 = i) \quad (i \in A), \qquad A = \{0, \ldots, k-1\}$$

- the joint distribution for $n \geqslant 1$ is given by

$$\begin{aligned}\text{Prob}(X_{n+1} = j \mid X_0 = i_0, \ldots, X_n = i_n) &= \text{Prob}(X_{n+1} = j \mid X_n = i_n) \\ &= M_{i_n, j}.\end{aligned}$$

(X_n) is called the Markov chain given by M.



(r, 1-r)

Markov Chains

Any (A-valued) process satisfying

$$\mathrm{Prob}(X_{n+1} = j \mid X_0 = i_0, \ldots, X_n = i_n) = \mathrm{Prob}(X_{n+1} = j \mid X_n = i_n)$$

is usually called a Markov chain.

- One can also consider Markov chains for processes taking values in $\mathbb{R}$.
- As with Markov shifts, one can also define k-step Markov chains for $k \geqslant 1$. Here, the memory of the process extends over k positions instead of just one:

$$\mathrm{Prob}(X_{n+1} = j \mid X_0 = i_0, \ldots, X_n = i_n) = \\ \mathrm{Prob}(X_{n+1} = j \mid X_{n-k+1} = i_{n-k+1}, \ldots, X_n = i_n)$$

PENNSTATE Eberly College of Science

Stationary Markov Chains

When is a Markov chain a stationary process?

- The transition probabilities have to be stationary (i.e. all be determined by the same transition matrix).
- The initial distribution has to be stationary, too. This means

$$\mathrm{Prob}(X_0 = j) = \mathrm{Prob}(X_1 = j), \quad \text{for all } j.$$

We have

$$\begin{aligned}\mathrm{Prob}(X_1 = j) &= \sum_{i=0}^{k-1} \mathrm{Prob}(X_1 = j \mid X_0 = i)\ \mathrm{Prob}(X_0 = i) \\ &= \sum_{i=0}^{k-1} M_{ij}\ \mathrm{Prob}(X_0 = i)\end{aligned}$$

PENNSTATE
Eberly College of Science

Stationary Markov Chains

Hence we want

j-th column

$$\text{Prob}(X_0 = j) = \sum_{i=0}^{k-1} M_{ij} \, \text{Prob}(X_0 = i)$$

If we write the initial distribution $\text{Prob}(X_0 = .)$ as a row vector $\vec{p} = [P(0) \ldots P(k-1)]$, then this can be written as

$$\boxed{\vec{p}M = \vec{p}.}$$

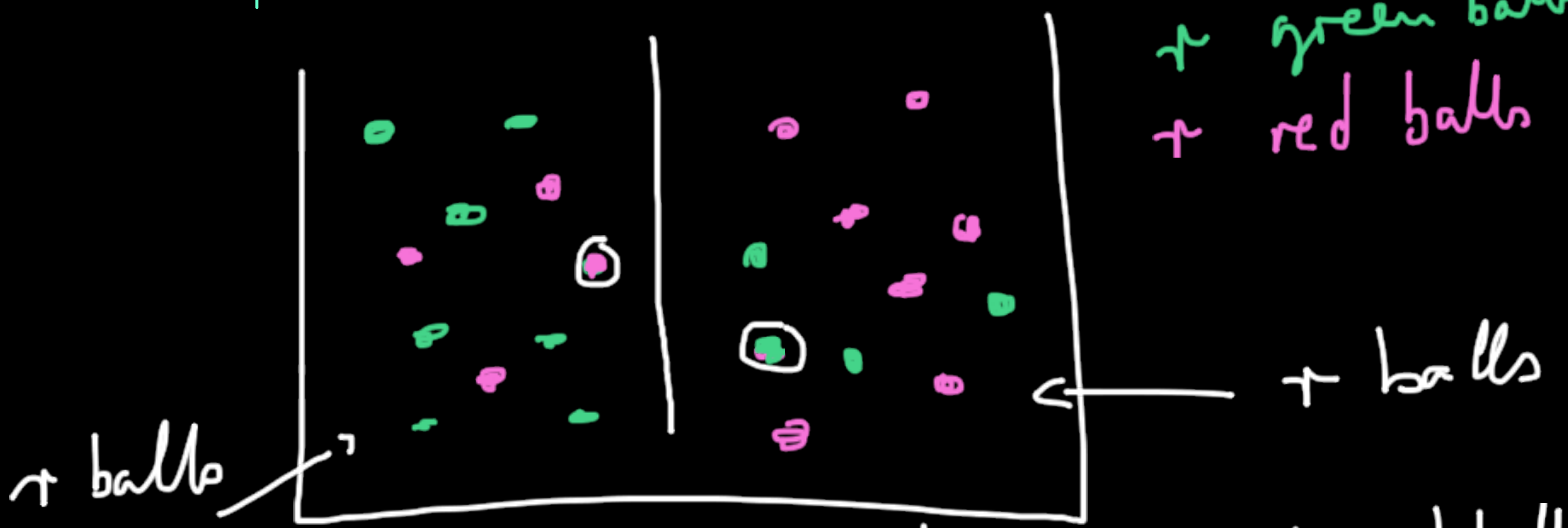
We call such an initial distribution stationary.

In the following, we will use the term stationary Markov chain to denote a Markov chain with stationary transition probabilities (given by a stochastic matrix M) and a stationary initial distribution.

PENNSTATE Eberly College of Science

Example: Diffusion

Bernoulli-Laplace model



state space = # red balls in left side $= \{0, \dots, r\}$

Transition probabilities:

$$M_{i,i-1} = \left(\frac{i}{r}\right)^2 \quad M_{i,i+1} = \left(\frac{r-i}{r}\right)^2 \quad M_{i,i} = \frac{i(r-i)}{r^2}$$

$M_{i,j} = 0$ for all other pairs i, j.

PENNSTATE
Eberly College of Science