INTRODUCTION

I MATHEMATICAL DEFINITION OF STOCHASTIC PROCESSES

We want to describe random processes evolving in time.

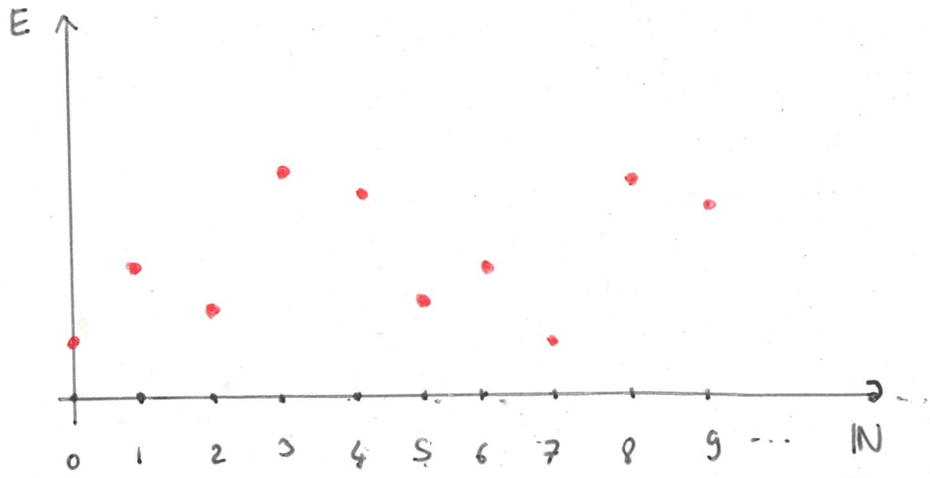
→ discrete time $I = \mathbb{N}$

→ continuous time $\mathbb{R}_+ = [0, +\infty)$

Framework: In this chapter, we fix a probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

> Def: Let $(E, \mathcal{E})$ be a measurable space.
>
> A discrete(-time) stochastic process with state space E is a collection of r.v. $(X_n)_{n \in \mathbb{N}}$ with values in E.

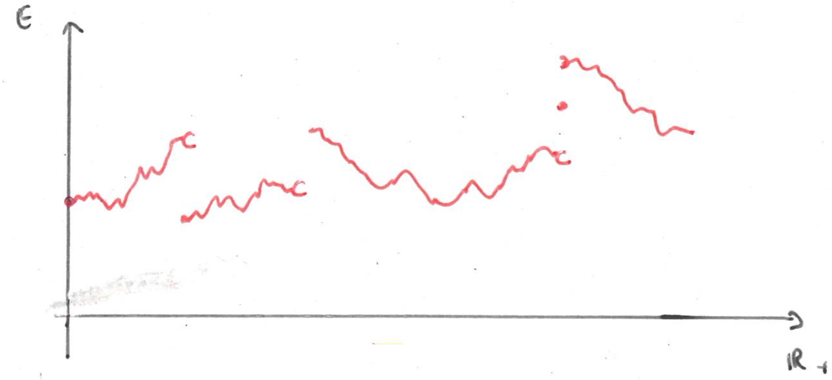
discrete stochastic process = "random sequence"



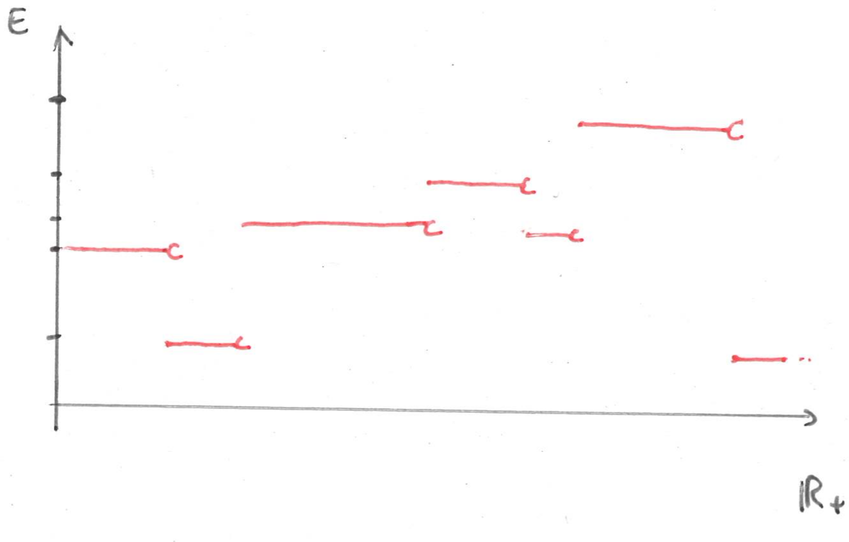
For fixed $\omega \in \Omega$, $(X_n(\omega))_{n \in \mathbb{N}}$ is a sequence of elements of E.

Def: A continuous (-time) stochastic process with state space E is a collection $(X_t)_{t \in \mathbb{R}_+}$ of r.v. with values in E.

continuous stochastic process = "random function"



In this class, we will consider jump processes (E finite or countable):



Rk: general stochastic process with state space E

→ collection $(X_t)_{t \in I}$ of r.v. with values in E

↳ A stochastic process is a collection of r.v. on the same probability space, nothing more!

"less complicated" $\longrightarrow$ more complicated

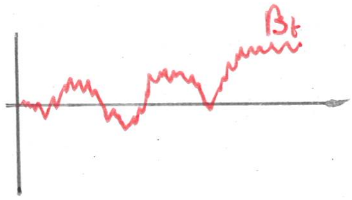
time space I	finite	infinite countable (e.g. $I = \mathbb{N}$)	uncountable ($I = \mathbb{R}_+$)
state space E	finite	infinite countable ($E = \mathbb{Z}$)	uncountable ($E = \mathbb{R}$)

Processes studied in this class:

- discrete time Markov Chains $I = \mathbb{N}$ E finite or countable
- Poisson processes / Renewal processes $I = \mathbb{R}_+$ $E = \mathbb{N}$
- Continuous-time Markov Chains $I = \mathbb{R}_+$ E finite or countable

Not in this class

- Brownian motion : $I = \mathbb{R}_+$ $E = \mathbb{R}$



Some questions about stochastic processes:

- Definition : a stochastic process is not always easy to define.
- Dependences : for $s, t \in I$ how do X_s and X_t depend on each other ?
- long time behaviour ? ($I = \mathbb{N}$ or $\mathbb{R}_+$)

  how does (X_t) look like for t large ?

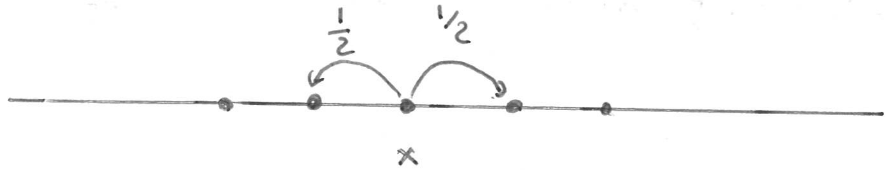
2 EXAMPLE I : THE SIMPLE RANDOM WALK ON $\mathbb{Z}^d$

State space $E = \mathbb{Z}^d$.

$x, y \in \mathbb{Z}^d$ are neighbors if they are at Euclidean distance 1.

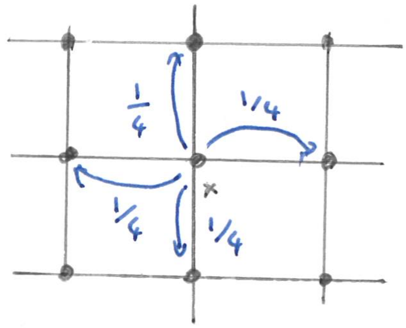
A particle starts at the origin and at each step, it jumps uniformly on one of its neighbors.

on $\mathbb{Z}$:

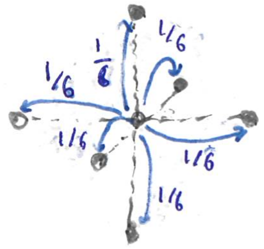


$$X_0 = 0 \quad \text{and} \quad P[X_{n+1} = y \mid X_n = x] = \begin{cases} \frac{1}{2} & \text{if } y \in \{x-1, x+1\} \\ 0 & \text{otherwise} \end{cases}$$

on $\mathbb{Z}^2$



on $\mathbb{Z}^3$



<u>Definition</u>? Let $(Z_n)_{n\geq 0}$ iid with $\mathbb{P}[Z_n = \pm e_i] = \frac{1}{2d}$

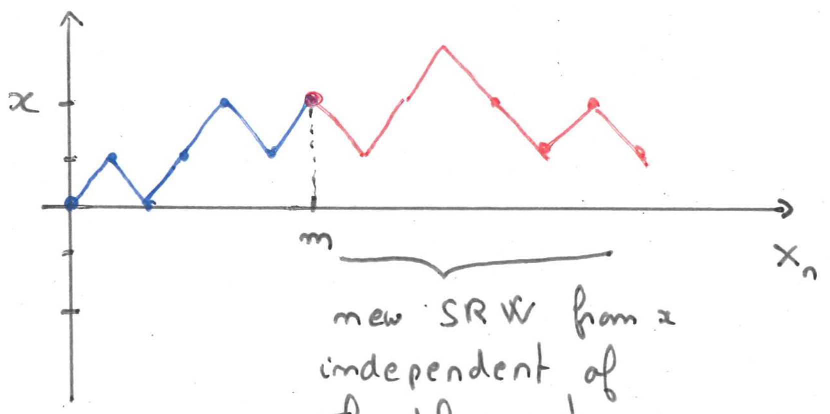
$$X_n := \sum_{k=1}^{n} Z_k$$

<u>Dependence</u>? $\forall m, n$ X_m and X_n are not independent (ex.)

Furthermore it satisfies the <u>Markov property</u>:

Condition on $X_m = x$, $(X_{m+n})_{n\geq 0}$ is a SRW

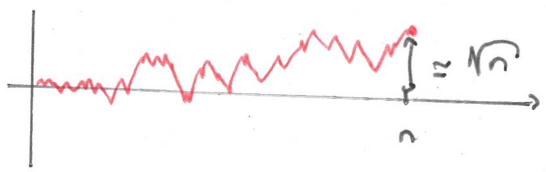
starting at x, independent of the first steps $X_1, \ldots, X_m$.



<u>Long-time behaviour</u>?

<u>d=1</u>: By the central limit theorem we have

$$\frac{X_n}{\sqrt{n}} = \frac{Z_1 + \cdots + Z_n}{\sqrt{n}} \xrightarrow[n]{(\text{law})} \mathcal{N}\left(0, \tfrac{1}{4}\right)$$

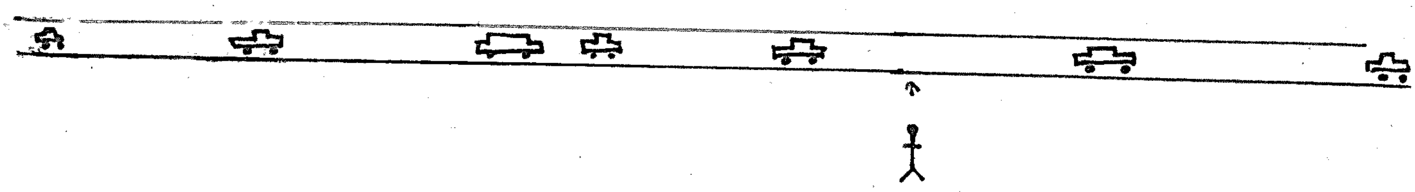


Is the state 0 visited infinitely many times?

Answer: (Polya's theorem)

$$\sum_{n=0}^{\infty} \mathbb{1}_{X_n = 0} \begin{cases} = +\infty \text{ a.s } & d = 1, 2 \\ < +\infty \text{ a.s } & d \geq 3 . \end{cases}$$

3 EXAMPLE 2 : POISSON PROCESS.



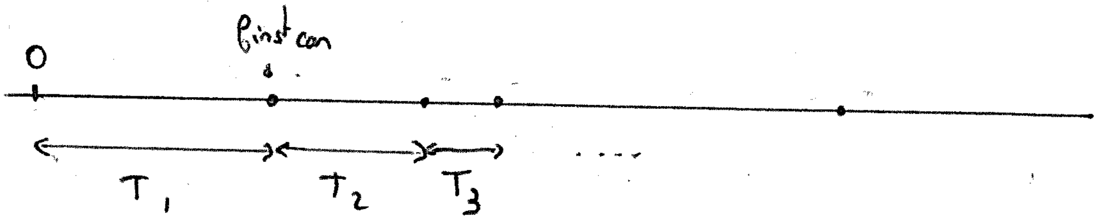
Goal: define and study

N_t = Number of cars passing at a point during a time interval $[0, t]$

Definition? consider T_1 = passage time of the first car

T_2 = time between the first & second car

$\vdots$



Define for every $t \geq 0$

$$N_t = \# \{ \text{arrivals before time } t \}$$
$$= \sum_{i=1}^{\infty} \mathbb{1}_{T_1 + \dots + T_i \leq t} \, .$$

$\rightarrow (N_t)_{t \geq 0}$ is a stochastic process, called Poisson Process (with intensity λ).

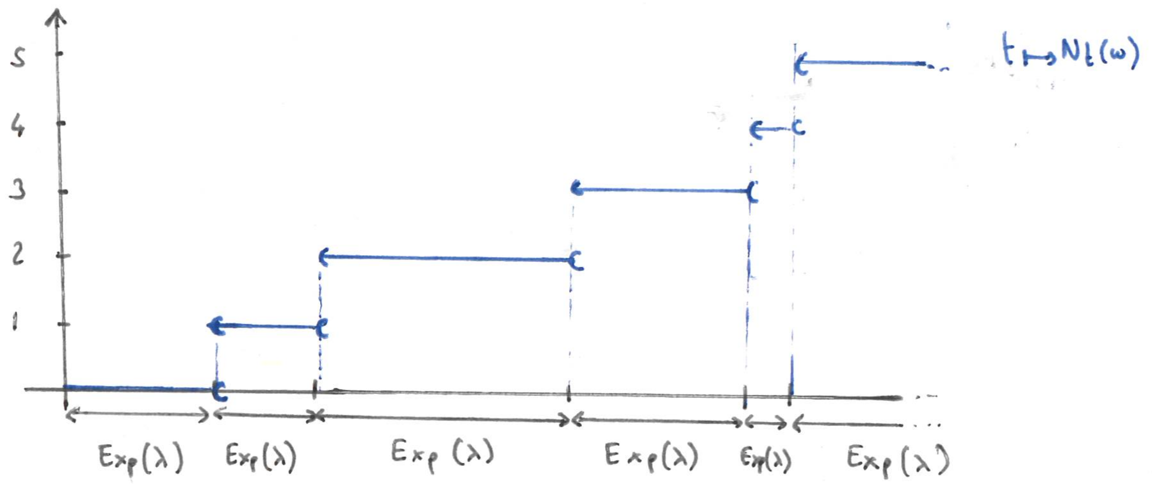


Fig: A possible trajectory of the Poisson process $(N_t)_t$.

Applications: arrival of customers in a queue, times at which telephone calls arrive at a call center, times at which claims arrive in an insurance company, times of emission for α particles by a radioactive source.

Hypotheses : • $(T_i)_{i\geq 1}$ are i.i.d. → "Right after the first car arrived, the time before the second arrival has the same law as T_1 and is independent, and so on..."

• $P[T_1 \geq t+s \mid T_1 \geq t] = P[T_1 \geq s]$.

"memoryless property": knowing that at time t, no car has arrived, the law of the remaining waiting time is the same as the original waiting time.

• The waiting times are "nice". $P[T_1 < \infty] = 1$, $t \mapsto P[T_1 \geq t]$ is continuous (no atom).
and $\forall t \; P[T_1 \geq t] > 0$.

Writing $g(t) = P[T_1 \geq t]$, we must have

$$\forall s,t \geq 0 \quad g(s) = P[T_1 \geq s] = \frac{P[T_1 \geq t+s]}{P[T_1 \geq t]} = \frac{g(t+s)}{g(t)}$$

Hence $\begin{cases} g(t+s) = g(t)\,g(s). & s,t \geq 0 \\ g(0) = 1. \end{cases}$

$\longrightarrow \exists \lambda > 0$ s.t. $\forall t. \quad g(t) = e^{-\lambda t}$

$T_1, \dots, T_i, \dots$ are iid exponential random variables. with some parameter $\lambda > 0$.

CHAPTER 1:

MARKOV CHAINS.

GENERALITIES.

Ref: [NORRIS] Markov Chains
[DURRETT] Probability : Theory and examples.
[SZNITMAN) Lecture notes 2017.

Framework: . $(\Omega, \mathcal{F}, \mathbb{P})$ probability space
. E finite or countable, non empty set, equiped with the σ-algebra $\mathcal{P}(E)$.

Idea: A Markov Chain is a discrete time stochastic process $(X_n)_{n\in\mathbb{N}}$ without memory: Given a time k the future $(X_{k+n})_{n\geq 0}$ depends only on the current position X_k and is independent of the past.

Motivation: . application in → Physics (e.g. evolution of a system in time)
→ Genetics (e.g. DNA sequences)
→ Computer science (e.g. Page Rank from Google) simulations ...
→ Mathematics (e.g. construction of measures ...) resolution of PDE's ...
→ Linguistics (e.g. original motivation of Markov 1856 - 1922)
→ Music (software for music generation)

Theoretical motivations:

→ easy to define (one of the simplest process besides i.i.d)

→ can be studied via algebraic tools (matrix theory ...) and analysis (operator theory)

Goals of the chapter

- Define Markov Chains. (MC) / weak Markov property
- Representation with transition probabilities

$$MC \longleftrightarrow \text{"matrix"} \ (P_{xy})_{x,y \in E}$$

- Existence theorem.
- Invariant / reversible distributions.
- Strong Markov properties, application to hitting times
- Application to Dirichlet problem.

I DEFINITIONS

Def: A sequence $(X_n)_{n \in \mathbb{N}}$ of r.v. with values in E is a homogeneous discrete-time Markov chain (MC) if

(i) $\forall n \geq 0 \quad \forall x_0, \dots, x_{n+1} \in E$

$$\mathbb{P}[X_{n+1} = x_{n+1} \mid X_0 = x_0, \dots, X_n = x_n] = \mathbb{P}[X_{n+1} = x_{n+1} \mid X_n = x_n]$$

"one-step Markov Property"

(ii) $\forall m, n \geq 0 \quad \forall x, y \in E$

$$\mathbb{P}[X_{n+1} = y \mid X_n = x] = \mathbb{P}[X_{m+1} = y \mid X_m = x]$$

"homogeneity"

⚠ there is a small abuse of notation, because the events in the conditioning may have zero probability. By convention, when we write $\mathbb{P}[A|B]$, we make the implicit assumption that $\mathbb{P}[B]>0$

e.g. (ii) corresponds to:

$$\forall m,n \geq 0 \quad \forall x,y \text{ satisfying } \mathbb{P}[X_m = x]>0 \ \mathbb{P}[X_n = x]>0$$

$$\mathbb{P}[X_{n+1} = y \mid X_n = x] = \mathbb{P}[X_{m+1} = y \mid X_m = x]$$

Rk: (i) $\iff$ $\begin{cases} \forall f: E \to \mathbb{R} \text{ bounded} \\ E[f(X_{n+1}) \mid X_0, \dots, X_n] = E[f(X_{n+1}) \mid X_n] \text{ a.s.} \end{cases}$

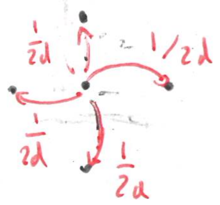
Example 1: If $(X_n)_{n\geq 0}$ are iid r.v. in E, then $(X_n)_{n\geq 0}$ is a M.C.

Example 2: Simple random walk on $\mathbb{Z}^d$

Let $(Z_n)_{n\geq 1}$ iid uniform in $\{\pm e_1, \dots, \pm e_d\}$. Then

$$X_n = \sum_{k=1}^{n} Z_k \qquad (X_0 = 0)$$

defines a MC on $\mathbb{Z}^d$.



Indeed: $\forall\, x_0, \dots, x_{n+1} \in \mathbb{Z}^d \qquad (x_0 = 0)$

$$\mathbb{P}[X_{n+1} = x_{n+1} \mid X_0 = x_0, \dots, X_n = x_n]$$

$$= \mathbb{P}[Z_{n+1} = x_{n+1} - x_n \mid Z_1 = x_1 - x_0, \dots, Z_n = x_n - x_{n-1}]$$

$$= \mathbb{P}[Z_{n+1} = x_{n+1} - x_n] \qquad \text{(by independence)}$$

and

$$\mathbb{P}[X_{n+1} = x_{n+1} \mid X_n = x_n] = \mathbb{P}[Z_{n+1} = x_{n+1} - x_n \mid Z_1 + \dots + Z_n = x_n]$$
$$= \mathbb{P}[Z_{n+1} = x_{n+1} - x_n]$$

This prove (i)

For (ii), simply use that $\mathbb{P}[X_{n+1} = y \mid X_n = x] = \mathbb{P}[Z_{n+1} = y - x]$
$$= \underbrace{\begin{cases} \frac{1}{2d} & \text{if } y = x \pm e_i \\ 0 & \text{otherwise} \end{cases}}_{\text{independent of } n}$$

2 TRANSITION PROBABILITIES

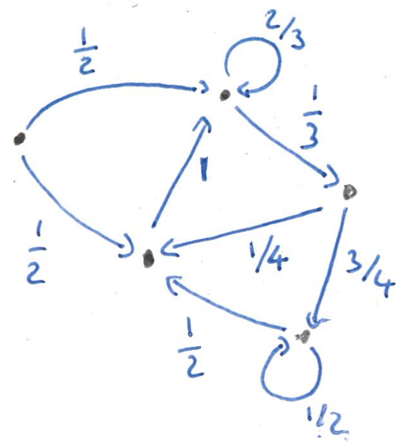
Def: A <u>transition probability</u> is a sequence $p = (p_{xy})_{x,y \in E}$ s.t.

- $\forall x, y \in E \quad p_{xy} \geq 0,$
- $\forall x \in E \quad \sum_{y \in E} p_{xy} = 1.$

Different representations of a transition probability

- weighted oriented graph (E finite or infinite)

vertices $= E$ weighted edges : $(x,y) \in E^2$ s.t $p_{xy} > 0$



→ the sum of the weights of the edges exiting a vertex is equal to 1.

- matrix (E finite)

for simplicity $E = \{1, \dots, N\}$

$$P = \begin{pmatrix} p_{11} & \cdots & p_{1N} \\ \vdots & & \vdots \\ \vdots & & \vdots \\ p_{N1} & \cdots & p_{NN} \end{pmatrix}$$

"Stochastic matrix"

- $p_{ij} \geq 0$ "non negative entries"
- $\sum_{j=1}^{N} p_{ij} = 1$

↳ each line sums to 1

- operator (E finite or infinite)

$\forall f \in L^\infty(E)$ define the function $Pf \in L^\infty(E)$ by

$$(Pf)(x) = \sum_{y \in E} p_{xy} f(y)$$

P positive ($\forall f \geq 0$ $Pf \geq 0$) and satisfies $P\mathbb{1} = \mathbb{1}$

constant function

Def: Let p be a transition probability, μ proba measure on E. A sequence of r.v. $(X_n)_{n\geq 0}$ with values in E is a Markov Chain with initial distribution μ and transition probability p ($MC(\mu, p)$) if

$$\forall x_0, \dots, x_n \in E \quad \mathbb{P}[X_0 = x_0, \dots, X_n = x_n] = \mu(x_0)\, p_{x_0 x_1} \cdots p_{x_{n-1} x_n}$$

p_{xy} = probability to jump from x to y.

Notation: $\mathcal{M} = \{$ proba. measure on $E \}$.

Prop: Let $(X_n)_{n\geq 0}$ be a sequence of r.v. with value in E. Then

$$\big((X_n)_{n\geq 0} \text{ is a MC}\big) \iff \big(\exists \mu \in \mathcal{M}\ \exists p\ (X_n)_{n\geq 0} \text{ is } MC(\mu, p)\big)$$

Proof:

$\boxed{\Rightarrow}$ Define μ = law of X_0 and set

$$p_{xy} = \begin{cases} \mathbb{P}[X_{n+1} = y \mid X_n = x] & \text{if } \exists n \ \ \mathbb{P}[X_n = x] > 0 \\ \mathbb{1}_{x=y} & \text{otherwise.} \end{cases}$$

By homogeneity, p_{xy} is well-defined.

Furthermore, for every $x_0, \dots, x_n \in E$ we have

$$\mathbb{P}[X_0 = x_0, \dots, X_n = x_n]$$

$$= \underbrace{\mathbb{P}[X_0 = x_0]}_{= \gamma(x_0)} \cdot \prod_{i=1}^{n} \underbrace{\mathbb{P}[X_i = x_i \mid X_0 = x_0, \dots, X_{i-1} = x_{i-1}]}_{= \mathbb{P}[X_i = x_i \mid X_{i-1} = x_{i-1}]}$$

$= \mathbb{P}[X_i = x_i \mid X_{i-1} = x_{i-1}]$ "by the 1-step Markov property"

$= p_{x_{i-1} x_i}$ "by def. of p"

$$= \gamma(x_0)\, p_{x_0 x_1} \cdots p_{x_{n-1} x_n}.$$

It remains to check that p is a transition probability.

Let $x \in E$. If there exists $n \geq 0$ s.t. $\mathbb{P}[X_n = x] > 0$, then

$$\sum_{y \in E} p_{xy} = \sum_{y \in E} \mathbb{P}[X_{n+1} = y \mid X_n = x] = 1.$$

Otherwise,

$$\sum_{y \in E} p_{xy} = \sum_{y \in E} \mathbb{1}_{x=y} = 1.$$

$\boxed{\Leftarrow}$ Assume $(X_n)_{n \geq 0}$ is $MC(\gamma, p)$.

Let $n \geq 0$ and $x_0, \dots, x_{n+1}$ s.t. $\gamma(x_0)\, p_{x_0 x_1} \cdots p_{x_{n-1} x_n} > 0$

We have $\mathbb{P}[X_{n+1} = x_{n+1} \mid X_0 = x_0 \dots X_n = x_n)$

$$= \frac{\mathbb{P}[X_0 = x_0 \dots X_{n+1} = x_{n+1}]}{\mathbb{P}[X_0 = x_0 \dots X_n = x_n]} = p_{x_n x_{n+1}}$$

Now let $n \geq 0$, $x, y \in E$ s.t. $\mathbb{P}[X_n = x] > 0$.

$$\mathbb{P}[X_{n+1} = y \mid X_n = x] = \sum_{u_0 \ldots u_{n-1} \in E} \mathbb{P}[X_{n+1} = y \mid X_0 = u_0, \ldots, X_{n-1} = u_{n-1}, X_n = x] \times \mathbb{P}[X_0 = u_0, \ldots, X_{n-1} = u_{n-1} \mid X_n = x]$$

$$= p_{xy} \sum_{u_0, \ldots u_{n-1}} \mathbb{P}[X_0 = u_0, \ldots, X_{n-1} = u_{n-1} \mid X_n = x]$$

$$= p_{xy}$$

This concludes (i) and (ii) in the def of MC. ■

<u>Why is the representation of MC with μ and P nice?</u>

Assume $E = \{1, \ldots, N\}$

Let $\mu \in \mathcal{M} \longleftrightarrow \mu = (\mu_1, \ldots, \mu_N)$

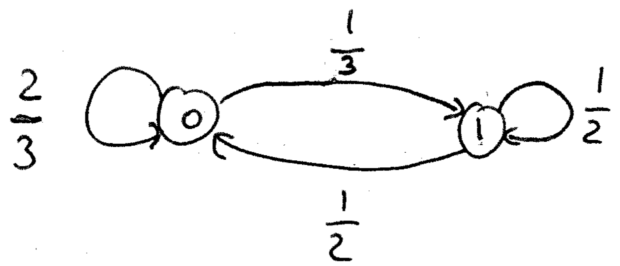
$$\mathbb{P}[X_n = j] = \sum_{i_0, \ldots i_{n-1}} \mathbb{P}[X_0 = i_0, \ldots, X_{n-1} = i_{n-1}, X_n = j]$$

$$= \sum_{i_0, \ldots, i_{n-1}} \mu_{i_0} P_{i_0 i_1} \cdots P_{i_{n-2} i_{n-1}} P_{i_{n-1} j}$$

$$= (\mu P^n)_j$$

Example : the weather Markov Chain

$0 =$ "cloudy day" $\qquad 1 =$ "sunny day"



$$P = \begin{pmatrix} \frac{2}{3} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

We start on a cloudy day. $X_0 = 0$ (ie. $\mu = \delta_0$)

What is the probability that the n-th day is sunny.

We have $\mathbb{P}[X_n = 1] = (\mu P^n)_1$

$$P = \underbrace{\begin{pmatrix} 1 & -\frac{2}{5} \\ 1 & \frac{3}{5} \end{pmatrix}}_{Q} \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{6} \end{pmatrix} \underbrace{\begin{pmatrix} \frac{3}{5} & \frac{2}{5} \\ -1 & 1 \end{pmatrix}}_{Q^{-1}}$$

$$P^n = Q \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{6^n} \end{pmatrix} Q^{-1}$$

$$(1, 0)\, P^n \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \frac{2}{5} - \frac{2}{5} \cdot \left(\frac{1}{6}\right)^n$$

3 EXISTENCE THEOREM

Question: For fixed p and μ, does there exist a stochastic process $(X_n)_{n\geq 0}$ which is a $MC(\mu, p)$?

Theorem.

Let p be a transition probability on E. There exist

- a measurable space $(\Omega, \mathcal{F})$
- a collection of probability measures $(P_x)_{x\in E}$
- a sequence of r.v. $(X_n)_{n\geq 0}$ on $(\Omega, \mathcal{F})$

such that, for every x we have

under P_x, $(X_n)_{n\geq 0}$ is $MC(\delta_x, p)$

Proof. We first fix a measure μ on E with $\mu(x)>0$ for every x. and construct a $MC(\mu, p)$ on some abstract probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

Consider
- X_0 a r.v with law μ
- $U_1, U_2, \dots$ iid uniform r.v. on $[0,1]$

One can construct a measurable function

$$\Phi : E \times [0,1) \longrightarrow E$$

such that $\forall x \in E \quad \mathbb{P}[\Phi(x, U_1) = y] = p_{xy}$

(order $E = \{x_1, x_2, \dots\}$ and then define for every i, j

$s_{i,j} = \sum_{k<j} p_{x_i x_k}$ and $\Phi(x_i, u) = x_j$ if $s_{ij} \leq u < s_{ij} + p_{x_i x_j}$)

Define by induction, for every $n \geq 0$

$$X_{n+1} = \Phi(X_n, U_{n+1})$$

Then we have for every $x_0, \dots, x_n \in E$

$$\mathbb{P}[X_0 = x_0, X_1 = x_1, \dots, X_n = x_n]$$

$$= \mathbb{P}[X_0 = x_0, \Phi(x_0, U_1) = x_1, \dots, \Phi(x_{n-1}, U_n) = x_n]$$

$$\overset{\text{indep.}}{=} \mu(x_0) \; p_{x_0 x_1} \cdots p_{x_{n-1} x_n}$$

Now define for every x $\quad P_x := \mathbb{P}[\,\cdot \mid X_0 = x]$

(well defined because $\mu(x) > 0$)

Then we have $\forall x \quad P_x[X_0 = x_0, \dots, X_n = x_n] = \delta_x(x_0)\, p_{x_0 x_1} \cdots p_{x_{n-1} x_n}$. ■

Framework for the rest of the chapter.

- E finite or countable
- p transition probability
- $(\Omega, \mathcal{F}, (P_x)_{x \in E})$ proba spaces.
- $(X_n)_{n \geq 0}$ r.v. s.t. X_n $MC(\delta_x, p)$ under P_x.

Notation: For every μ proba. measure on E

$$P_\mu := \sum_{x \in E} \mu(x) \cdot P_x$$

This implies $\forall x_0, \dots, x_n \in E$

$$P_\mu(X_0 = x_0, \dots, X_n = x_n) = \mu(x_0)\, p_{x_0 x_1} \cdots p_{x_{n-1} x_n}$$

$\hookrightarrow$ Under P_μ $(X_n)_{n \geq 0}$ is $MC(\mu, p)$.

Rk: Under P_μ, $(X_n)_{n \geq 0}$ is a MC. Hence it satisfies

$\forall x_0, \dots, x_n, x_{n+1} \in E$

$$P_\mu[X_{n+1} = x_{n+1} \mid X_0 = x_0 \dots X_n = x_n] = P_\mu[X_{n+1} = x_{n+1} \mid X_n = x_n]$$

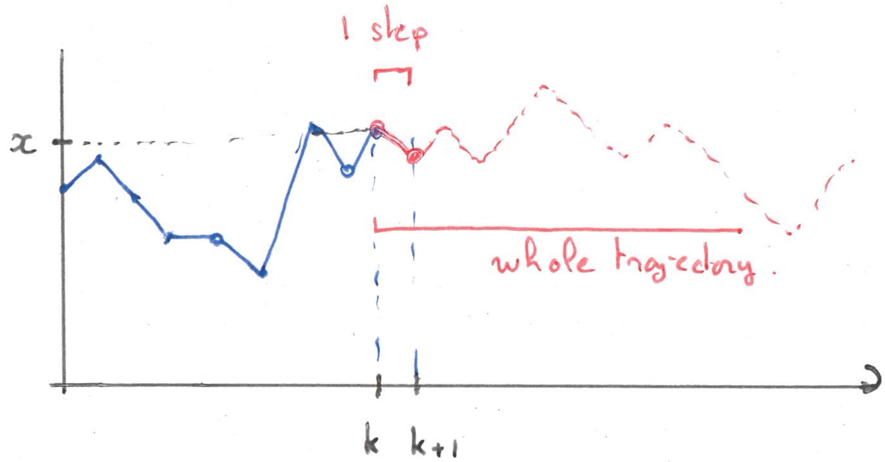
$$= P_{x_n}[X_1 = x_{n+1}]$$

"rephrasing of the 1-step Markov Property"

4 SIMPLE MARKOV PROPERTY.

The one-step Markov property says:

- Condition on $X_k = x$, X_{k+1} is sampled like the first step of a $MC(\delta_x, p)$, independent of $X_0, \dots, X_k$.



The Markov property will say that the whole trajectory is sampled according to a $MC(\delta_x, p)$ independent of the past.

<u>Not</u>: $\mathcal{F}_n = \sigma(X_0, \dots, X_n)$

<u>Thm [Markov property]</u> Let $\mu \in \mathcal{M}_1$.

Let $x \in E$, $k \in \mathbb{N}$. For every $f : E^{\mathbb{N}} \to \mathbb{R}$ meas. bounded, for every Z $\mathcal{F}_k$-measurable bounded, we have

$$(*) \quad \boxed{E_\mu\left[f((X_{k+n})_{n\geq 0}) \cdot Z \mid X_k = x\right] = E_x\left[f((X_n)_{n\geq 0})\right] \times E_\mu\left[Z \mid X_k = x\right]}$$

"Condition on $X_k = x$, $(X_{k+n})_{n\geq 0}$ is $MC(\delta_x, p)$, independent of $\mathcal{F}_k$."

Corollary:

Let $\mu \in \mathcal{M}$, $x \in E$, $k \in \mathbb{N}$. For every $f : E^{\mathbb{N}} \to \mathbb{R}$ meas. bounded

$$\boxed{E_\mu\left[f((X_{k+n})_{n\geq 0}) \mid X_k = x\right] = E_x\left[f((X_n)_{n\geq 0})\right]}$$

Proof: take $Z = 1$ in the Markov property:

Rk: The statement $(*)$ is equivalent to

$\forall\, x_0, \dots, x_{k-1}, x_k \in E$

$$(**) \quad E_\mu\left[f((X_{k+n})_{n\geq 0}) \mid X_0 = x_0, \dots, X_{k-1} = x_{k-1}, X_k = x_k\right] = E_{x_k}\left[f((X_n)_{n\geq 0})\right]$$

(exercise. Hint: take $Z = \mathbb{1}_{X_0 = x_0, \dots, X_{k-1} = x_{k-1}}$)

Proof: We prove $(**)$.

By approximating f by step functions and by linearity it suffices to prove $(**)$ for $f = \mathbb{1}_A$ where $A \subset E^{\mathbb{N}}$ measurable.

$$P_\mu\left[(X_{k+n})_{n\geq 0} \in A \mid X_0 = x_0, \dots, X_k = x_k\right] = P_{x_k}\left[(X_n)_{n\geq 0} \in A\right]$$

By Dynkin's lemma, it suffices to prove the statement above for A of the form

$$A = \{\omega \in E^{\mathbb{N}} : \omega_0 = y_0, \dots, \omega_N = y_N\} \qquad \begin{matrix} N \geq 0 \\ y_0, \dots, y_N \in E \end{matrix}$$

For such event A, we have

$$P_\mu[(X_{k+n})_{n\geq 0}\in A \mid X_0=x_0,\dots,X_k=x_k]$$

$$= P_\mu[X_k=y_0,\dots,X_{k+N}=y_N \mid X_0=x_0,\dots,X_k=x_k]$$

$$= \frac{\mu(x_0)\, p_{x_0x_1}\cdots p_{x_{k-1}x_k}\cdot \mathbb{1}_{x_k=y_0}\cdot p_{y_0y_1}\cdots p_{y_{N-1}y_N}}{\mu(x_0)\, p_{x_0x_1}\cdots p_{x_{k-1}x_k}}$$

$$= \delta_{x_k}(y_0)\, p_{y_0y_1}\cdots p_{y_{N-1}y_N} = P_{x_k}[(X_n)_{n\geq 0}\in A] \quad \blacksquare$$

5. n-STEP TRANSITION PROBABILITIES

Def: For every $n\geq 0$, for every $x,y\in E$, define

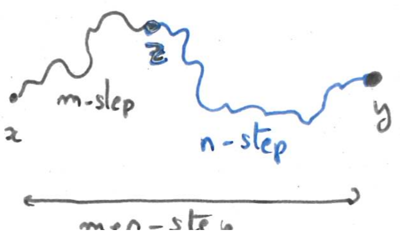
$$\boxed{p^{(n)}_{x,y} := P_x[X_n=y]}$$

"probability to reach y from x in n steps"

Prop. [Chapman-Kolmogorov]

For every $m,n\geq 0$, $x,y\in E$, we have

$$\boxed{p^{(m+n)}_{xy} = \sum_{z\in E} p^{(m)}_{xz}\, p^{(n)}_{zy}}$$



Proof: $$P_x[X_{m+n}=y] = \sum_{z\in E} \underbrace{P_x[X_{m+n}=y \mid X_m = z]}_{\overset{MP}{=} P_z[X_n=y]} P_x[X_m=z]$$

$$= \sum_{z\in E} P_x[X_m=z]\, P_z[X_n=y] \qquad \blacksquare$$

Prop: [matrix interpretation]

Assume $E=\{1,\dots,N\}$ for some $N\geq 1$. Write $P=(p_{ij})_{1\leq i,j\leq N}$.

Then the matrix $(p^{(n)}_{i,j})_{1\leq i,j\leq N}$ is the n-th power of P:

$$\forall n\geq 0 \quad \boxed{(p^{(n)}_{i,j})_{1\leq i,j\leq N} = P^n.}$$

Furthermore, for every distribution μ on E, and every function $f: E\longrightarrow \mathbb{R}$ we have

$$\forall n\geq 0 \quad \boxed{E_\mu[f(X_n)] = \mu P^n f}$$

where we identify μ with the row vector $\mu=(\mu(1),\dots,\mu(N))$ and f with the column vector $f=\begin{pmatrix} f(1)\\ \vdots \\ f(N)\end{pmatrix}$.

Proof: The first equation follows from $p^{(n+1)}_{ik} = \sum_j p^{(n)}_{ij} p_{jk}$ by induction. For the second equation, we use the def. of the expectation

$$E_\mu[f(X_n)] = \sum_{y\in E} f(y)\, P_\mu[X_n=y] = \sum_{\substack{x\in E\\ y\in E}} \mu(x) \underbrace{P_x[X_n=y]}_{=p^{(n)}_{xy}} f(y) \qquad \blacksquare$$

Dictionary probability $\leftrightarrow$ algebra for E finite

$E = \{1, \dots, N\}$

Probability	Linear Algebra
distribution μ on E $\leftrightarrow$	$\mu = (\mu_1, \dots, \mu_N) \in \mathbb{R}_+^N$ with $\sum_i \mu_i = 1$
measurable map $f : E \to \mathbb{R}$ $\leftrightarrow$	$f = \begin{pmatrix} f_1 \\ \vdots \\ f_N \end{pmatrix} \in \mathbb{R}^N$
Markov Chain $\leftrightarrow$	$\begin{cases} \text{row vector } \mu \quad \sum_{i=1}^N \mu_i = 1 \\ \text{stochastic matrix } P \end{cases}$
$E_\mu[f(X_n)]$ $\leftrightarrow$	$\mu P^n f$
law of X_n $\leftrightarrow$	μP^n
$\left(E_x[f(X_n)]\right)_{x \in E}$ $\leftrightarrow$	$P^n f$

See exercises for a general method to compute $P_x[X_n = y]$.

6 STATIONARY DISTRIBUTIONS

Motivation: write μ_n for the distribution of X_n under P_μ

$\hookrightarrow$ we have $\begin{cases} \mu_0 = \mu \\ \mu_{n+1} = \mu_n \cdot P \end{cases}$ "law of X_{n+1} = f (law of X_n)"

$\rightsquigarrow$ we expect that for n large "$\mu_n \simeq$ fixed point of $\lambda \mapsto \lambda P$"

Def: Let π be a distribution on E (= proba measure on E)
We say that π is stationary (for p) if

$$\boxed{\forall y \in E \quad \sum_x \pi(x)\, p_{xy} = \pi(y)}$$

Probabilistic interpretation:

If π is a stationary distribution, then for every $n \geq 0$

$$\boxed{P_\pi[X_n = x] = \pi(x)}$$

Equivalently, if $(X_n)_{n\geq 0}$ is a MC (π, p) then the law of X_n is π at every time $n \geq 0$.

This follows from the definition by induction:

- For $n=0$ $\quad P_\pi[X_0 = x] = \pi(x)$ by definition.

- For $n \geq 1$ $\quad P_\pi[X_n = y] = \sum_{x \in E} \underbrace{P_\pi[X_n = y \mid X_{n-1} = x]}_{\underset{MP}{=} p_{xy}} \underbrace{P_\pi[X_{n-1} = x]}_{\underset{\text{induction}}{=} \pi(x)}$

$$\underset{\pi \text{ stationary}}{=} \pi(y)$$

Algebraic interpretation: $\quad E = \{1, \dots, N\} \quad \pi = (\pi(1), \dots, \pi(N))$

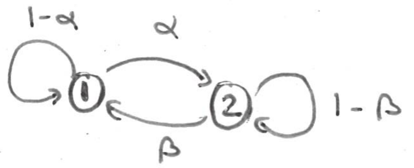
π stationary $\iff$ $\boxed{\pi P = \pi}$ $\quad$ "π is a left eigen vector associated to the eigen value 1"

Rk: Since $P\mathbb{1} = \mathbb{1}$ we know that 1 is an eigen value for P^t as well (because the spectrum of P is the same as the spectrum of P).
Therefore, there exists $v \in \mathbb{R}^N \setminus \{0\}$ s.t $P^t v = v$

ie $v^t P = v^t$ $\rightsquigarrow$ it is a priori not clear that there exists such a vector with ≥ 0 entries ...

Examples:

- "2-state" - MC $E = \{1,2\}$



$0 < \alpha, \beta \leq 1$

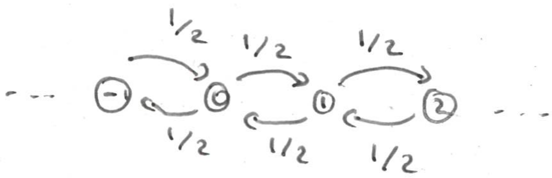
$$\pi = \left(\frac{\beta}{\alpha+\beta}, \frac{\alpha}{\alpha+\beta} \right)$$ unique stationary distrib.

- degegenerate 2-state MC $(\alpha = \beta = 0)$

Any $\pi = \alpha\,\delta_1 + (1-\alpha)\,\delta_2$, $\alpha \in [0,1]$ is stationary.

$\hookrightarrow$ infinitely many stationary distributions.

- SRW on $\mathbb{Z}$



π is stationary $\Rightarrow \forall x \;\; \pi(x) = \frac{1}{2}\pi(x-1) + \frac{1}{2}\pi(x+1)$

$\Rightarrow \forall x \;\; \pi(x) - \pi(x-1) = \pi(x+1) - \pi(x) \Rightarrow \pi$ is linear

$\hookrightarrow$ There is no stationary distribution.

Rk: for the simple random walk the constant measures

$\mu(x) = 1$ satisfy $\mu P = \mu$

$\hookrightarrow$ "invariant" measure (but not a probability measure)

7. REVERSIBILITY

Def: A distribution π on E is said to be reversible (for p) if

$$\boxed{\forall x, y \in E \quad \pi(x)\, p_{xy} = \pi(y)\, p_{yx}}$$ "detailed balance"

Rk: Why is it called reversibility ?

π reversible $\iff$ $\forall x, y \in E \quad P_\pi[X_0 = x, X_1 = y] = P_\pi[X_0 = y, X_1 = x]$

More generally, one can prove by induction that π is reversible iff:

$$\forall n \quad \forall x_0, \ldots, x_n \in E \quad P_\pi[X_0 = x_0, \ldots, X_n = x_n] = P_\pi[X_0 = x_n, \ldots, X_n = x_0]$$

$$\leadsto P_\pi[\; x_0 \qquad x_n \;] = P_\pi[\; x_0 \qquad x_n \;]$$

"If X_0 is sampled according to π then the probability of a given trajectory is equal to the probability of its reversed version"

Motivation: • criterion for stationarity (see next proposition)

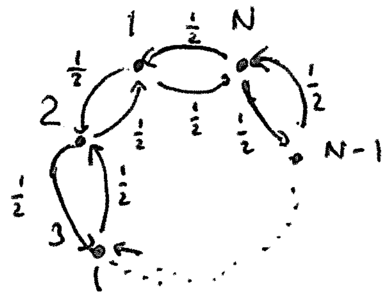
• appear often for dynamics in physics (see Ex. 3 below)

Prop. Let π be a reversible distribution. Then π is invariant.

Proof. For every $y \in E$, we have

$$\sum_{x\in E} \pi(x)\, p_{xy} \overset{(\text{Rev.})}{=} \sum_{x\in E} \pi(y)\, p_{yx} = \pi(y) \underbrace{\sum_{x\in E} p_{yx}}_{=1} \qquad \blacksquare$$

Example 1: Sym. RW on a torus $E = \{1, \dots, N\}$

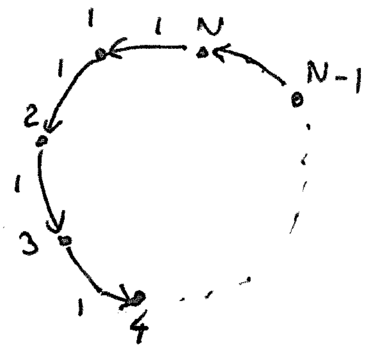


$p_{x\,x+1} = p_{x\,x-1} = \frac{1}{2}$

"mod N"

$\boxed{\text{The uniform measure } \pi(x) = \frac{1}{N} \text{ is reversible.}}$

Example 2: totally asymetric RW on the torus



$p_{x\,x+1} = 1$

$p_{x.\,x-1} = 0$

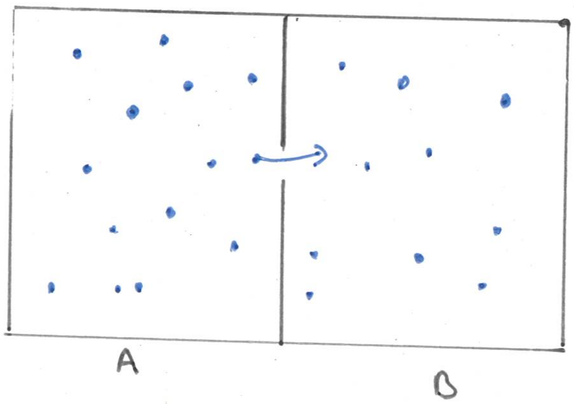
Assume $\exists$ a reversible proba π

then $\forall x \quad \pi(x) = \pi(x) \underbrace{p_{x\,x+1}}_{=1} \overset{(\text{Rev})}{=} \pi(x+1)\, p_{x+1\,x} = 0$ contradiction

$\longrightarrow$ $\boxed{\text{No Reversible distrib. (But uniform distrib. is invariant.)}}$

Example 3 : Ehrenfest model of Diffusion.

Two containers A and B are placed adjacent to each others and gas is allowed to pass through a small aperture joining them. A total of N gas molecules is distributed between the containers.



(A and B are assumed to be of the same size)

Model: We consider a discrete time $\mathbb{N}$, and we write

X_n = Number of particules in A at time n.

At each time, a uniformly chosen molecule passes through the aperture.

$\hookrightarrow (X_n)$ is a Markov Chain on the state space $E = \{1, \dots, N\}$ and the transition probability is given by

$$\boxed{\begin{array}{lll} P_{x\,x+1} = 1 - \dfrac{x}{N} & & 0 \leq x < N \\[2ex] P_{x\,x-1} = \dfrac{x}{N} & & 0 < x \leq N \\[2ex] P_{xy} = 0 & & y \notin \{x-1, x+1\} \end{array}}$$

Let us look for a reversible distribution. $\pi = (\pi_x)_x$

We must have $\quad \pi_x \, p_{x(x+1)} = \pi_{x+1} \, p_{(x+1)x}$

$$\text{ie} \quad \pi_{x+1} = \frac{N-x}{x+1} \pi_x$$

By induction we find $\forall x \in \{0, \dots, N\}$ $\pi_x = \dfrac{N \times \dots \times (N-x+1)}{1 \times \dots \times x} \pi_0 = \dbinom{N}{x} \pi_0$

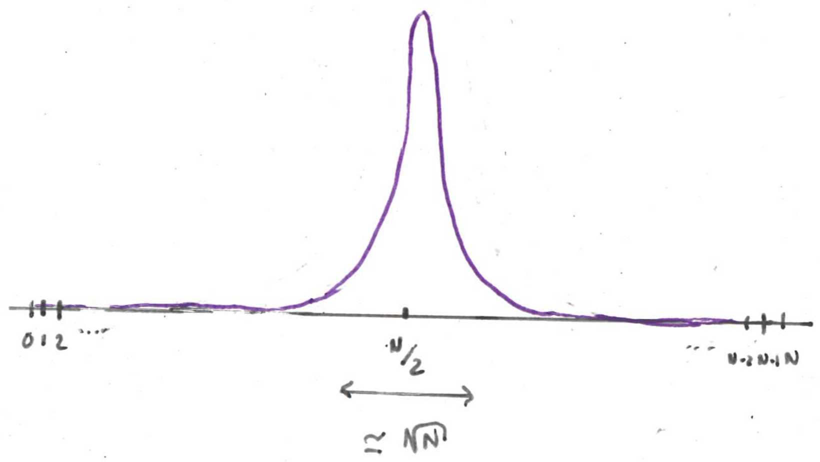
The condition $\sum_x \pi_x = 1$ imposes $\pi_0 = \left(\sum_x \binom{N}{x} \right)^{-1} = \dfrac{1}{2^N}$

We find $\forall x$ $\boxed{\pi_x = \dfrac{1}{2^N} \dbinom{N}{x}}$ $\quad$ "π is Binomial $(N, \frac{1}{2})$"

Conversely, one can check that the Binomial distribution is reversible.

At equilibrium(*), the number of particles in the container A is distributed like a Binomial $(N, \frac{1}{2})$.

(*) when X_{n+1} has the same law as X_n



distribution of the number of molecules in A.

8 COMMUNICATION CLASSES

<u>Rk.</u> graph theoretical notion.

($p \leftrightarrow$ weighted oriented graph)

Def: Let $x, y \in E$. Write

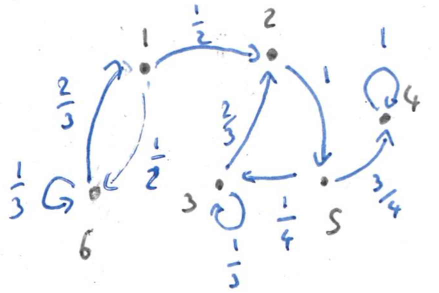
- $x \longrightarrow y$ if $\exists n \geq 0$ s.t. $p^{(n)}_{xy} > 0$ "y can be reached from x"
- $x \longleftrightarrow y$ if ($x \to y$ and $y \to x$) "x and y communicate"

<u>Rk:</u> Since $$p^{(n)}_{xy} = \sum_{z_1, \dots, z_{n-1} \in E} p_{xz_1} p_{z_1 z_2} \cdots p_{z_{n-2} z_{n-1}} p_{z_{n-1} y}$$

we have $(x \to y) \iff (\exists z_1, \dots, z_{n-1} \quad p_{xz_1}, \dots, p_{z_{n-1} y} > 0)$

seeing p as a oriented graph $x \to y$ corresponds to the existence of an oriented path from x to y.

<u>Example</u>



- $1 \to 3$
- $3 \not\to 1$
- $2 \leftrightarrow 3$

<u>Probabilistic interpretation:</u>

$$x \to y \underset{\text{by def}}{\iff} \exists n \geq 0 \quad P_x[X_n = y] > 0 \underset{\text{ex.}}{\iff} P_x[\exists n \geq 0 \; X_n = y] > 0$$

Prop. $\boxed{\leftrightarrow \text{ is an equivalence relation on } E.}$

Proof. • Since $p^{(0)}_{xx} = 1$, we have $x \leftrightarrow x$ for every $x \in E$.

• If $x \leftrightarrow y$ and $y \leftrightarrow z$, then there exist $m, n \geq 0$ such that $p^{(m)}_{xy}, p^{(n)}_{yz} > 0$. Therefore,

$$p^{(m+n)}_{xz} \underset{(CK)}{=} \sum_{u \in E} p^{(m)}_{xu} \, p^{(n)}_{uy} \geq p^{(m)}_{xy} \, p^{(n)}_{yz} > 0.$$

Hence $x \to z$, and $z \to x$ equivalently. ∎

Def: • The equivalence classes of $\leftrightarrow$ are called __communication classes__.

• The chain p is said to be __irreducible__ if there is a unique communication class

Rk: p irreducible $\iff \forall x, y \quad x \to y$

communication class $\rightsquigarrow$ strongly connected component of G

p irreducible $\rightsquigarrow$ G is strongly connected

I the example before, there are 3 communication classes

$\hookrightarrow C_1 = \{1, 6\} \quad C_2 = \{2, 3, 5\} \quad C_3 = \{4\}$

Motivation : we will see that

$$(p \text{ irreducible}) \Rightarrow (p \text{ has at most } 1 \text{ stationary distrib.})$$

↳ algebraic proof for E finite (see exercises)

↳ probabilistic proof (see next chapter).

Def: A communication class C is closed if for every $x, y \in E$

$$(x \in C \,,\ x \to y) \Rightarrow y \in C$$

Probabilistic interpretation: Let C be a communication class

$$\boxed{C \text{ is closed} \iff \forall x \in C \quad P_x[\forall n \geq 0 \ \ X_n \in C] = 1}$$

↳ once the Markov chain enters in a closed class, it stays in it forever.

Proof: C not closed $\iff \exists x \in C \ \exists y \in E \setminus C \quad x \to y$

$$\iff \exists x \in C \ \exists y \in E \setminus C \ \ P_x[\exists n \ X_n = y] > 0$$

$$\iff \exists x \in C \quad P_x[\exists n \ \ X_n \in E \setminus C] > 0$$

$E \setminus C$ at most countable

$$\iff \exists x \in C \quad P_x[\forall n \ \ X_n \in C] < 1 .$$

9 STRONG MARKOV PROPERTY.

$\mathcal{F}'_n = \sigma(X_0, \dots, X_n)$

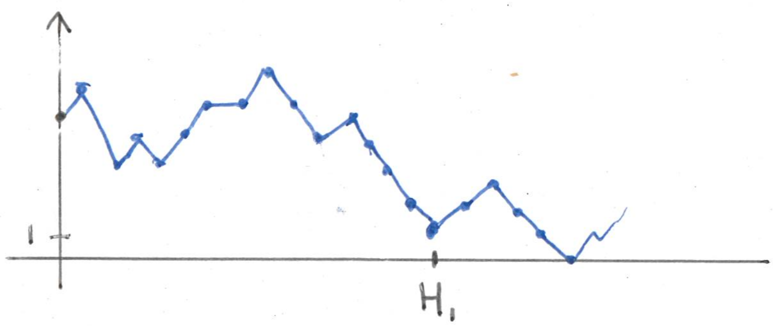
Def Let T be a r.v. with values in $\mathbb{N}$. We say that T is a $(\mathcal{F}'_n)$ - stopping time if

$$\forall n \in \mathbb{N} \quad \{T = n\} \in \mathcal{F}_n$$

Ex: Hitting times. Let $A \subset E$, $x \in E$. Define

$$H_A = \min\{n \geq 1 \ : X_n \in A\}$$

$$H_x = \min\{n \geq 1 \ : X_n = x\}$$



Def: [σ-algebra at a stopping time]

Let T be a stopping time. Define

$$\mathcal{F}_T = \{A \in \mathcal{F} : \forall n \in \mathbb{N} \ \{T = n\} \cap A \in \mathcal{F}_n\}$$

" intuition: $\mathcal{F}_T = \sigma(X_0, \dots, X_T)$ information up to time T "

Rk: • T is $\mathcal{F}_T$-measurable.

• (Z is $\mathcal{F}_T$-measurable) $\iff$ $Z \mathbb{1}_{T=n}$ $\mathcal{F}_n$ - meas. for every n.

Thm [Strong Markov property]

Let μ be a distribution on E. Let T be a $(\mathcal{F}_n)$-stopping time. Let $x \in E$. For every $f : E^{\mathbb{N}} \to \mathbb{R}$ meas. bounded, for every Z $\mathcal{F}_T$-measurable bounded, we have.

$$E_\mu\left[f((X_{T+n})_{n\geq 0}) \cdot Z \,\middle|\, T<\infty, X_T = x\right] = E_x\left[f((X_n)_{n\geq 0})\right] E_\mu\left[Z \,\middle|\, T<\infty, X_T = x\right]$$

"Condition on $\{T<\infty, X_T = x\}$, $(X_{T+n})_{n\geq 0}$ is MC (δ_x, P), independent of $\mathcal{F}_T$"

Proof: We prove the equation multiplied by $P[T<\infty, X_T = x]$

$$E_\mu\left[f((X_{T+n})_{n\geq 0}) Z \mathbb{1}_{T<\infty, X_T = x}\right]$$

$$= \sum_{k=0}^{\infty} E_\mu\left[f((X_{k+n})_{n\geq 0}) Z \mathbb{1}_{T=k} \mathbb{1}_{X_k = x}\right]$$

$$= \sum_{k=0}^{\infty} E_\mu\left[f((X_{k+n})_{n\geq 0}) \cdot \underbrace{Z \mathbb{1}_{T=k}}_{\in \mathcal{F}_k} \,\middle|\, X_k = x\right] P_\mu\left[X_k = x\right]$$

$$\overset{M.P}{=} E_x\left[f((X_n)_{n\geq 0})\right] \times \underbrace{\sum_{k=0}^{\infty} E_\mu\left[Z \mathbb{1}_{T=k}, X_k = x\right]}_{= E_\mu\left[Z \mathbb{1}_{T<\infty}, X_T = x\right]} \qquad \blacksquare$$

APPLICATION : REFLECTION PRINCIPLE FOR THE SRW.

Consider the SRW on $\mathbb{Z}$ $\left(E = \mathbb{Z} \text{ and } p_{xy} = \frac{1}{2} 1\!\!1_{|x-y|=1} \right)$

Let $k \geq 0$ even and $a \geq 1$ odd

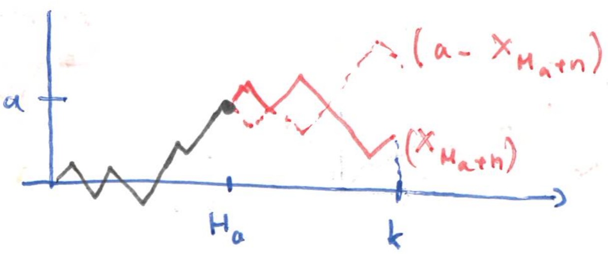
$$\boxed{P_0\left[\max_{0 \leq m \leq k} X_m \geq a\right] = P_0\left[|X_k| \geq a\right].}$$

Proof: Recall $H_a = \min\{n \geq 1 : X_n = a\}$

We have $P_0\left[\max_{0 \leq m \leq k} X_m \geq a\right] = P_0\left[H_a \leq k\right]$

$$= P_0\left[X_n > a\right] + P_0\left[H_a \leq k, X_k < a\right].$$

The law of $(X_{H_a+n})_{n \geq 0}$ is the same as $(a - X_{H_a+n})_{n \geq 0}$



By reflecting the last part of the trajectory, we can prove $P_0[X_n > a] = P_0[H_a \leq k, X_k < a]$

We have

$$P_0\left[H_a \leq k, X_k < a\right] = \sum_{m=0}^{k} P_0\left[X_k < a, H_a = m\right]$$

By the strong Markov property, we have

$$P_0\left[X_k < a, H_a = m\right] \overset{(*)}{=} P_a\left[X_{k-m} < a\right] P_0\left[H_a = m\right]$$

$$\underset{\text{Symmetry}}{=} P_a\left[X_{k-m} > a\right] P_0\left[H_a = m\right]$$

$$= P_0\left[X_k > a, H_a = m\right]$$

(to justify (*), one can use the strong Markov property as follows: $P_0[X_k < a, H_a = m] = P_0[X_k < a, H_a = m, H_a < \infty]$

$$= \underbrace{P_0[X_{H_a + k - m} < a, H_a = m \mid H_a < \infty, X_{H_a} = a]}P_0[H_a < \infty]$$

$$\overset{\text{StMP}}{=} P_a[X_{k-m} < a]\, P_0[H_a = m \mid H_a < \infty, X_{H_a} = a]$$

$$= P_a[X_{k-m} < a]\, P_0[H_a = m]\,)$$

Plugging the identity above in the sum, we get

$$P_0\left[H_a \leq k, X_k < a\right] = \sum_{m=0}^{k} P_0\left[X_k > a, H_a = m\right]$$

$$= P_0\left[X_k > a, H_a \leq k\right]$$

$$= P_0\left[X_k > a\right] \qquad \blacksquare$$

CHAPTER 2 : MARKOV CHAINS : LONG-TIME BEHAVIOUR

Framework:

- E finite or countable set,
- $p = (p_{xy})_{x,y \in E}$ transition probability,
- $(\Omega, \mathcal{F})$ measurable space equiped with $(P_x)_{x \in E}$ proba measures.
- $(X_n)_{n \geq 0}$ MC (δ_x, p) under P_x.

Not.: for $\mu = (\mu_x)_{x \in E}$ distribution on E, write $P_\mu = \sum_x \mu_x P_x$

Questions:
- Fix $x \in E$. Will $(X_n)_{n \geq 0}$ visit x infinitely many times?
- What is the distribution of X_n for n large?

I RECURRENCE / TRANSIENCE

Notation: For $x \in E$ $H_x = \min\{n \geq 1 : X_n = x\}$

Def. Let $x \in E$. We say that
- x is recurrent if $P_x[H_x < \infty] = 1$. "the chain always come back at x"
- x is transient if $P_x[H_x < \infty] < 1$. "the chain may never come back"

Notation: For $x \in E$, write $V_x = \sum_{n \geq 1} \mathbb{1}_{X_n = x}$.

"total number of visits of x"

Thm [DICHOTOMY THM]

Let $x \in E$.

- If x is recurrent, then $V_x = +\infty$ P_x-a.s.
- If x is transient, then $E_x[V_x] < \infty$ P_x-a.s

↳ it is impossible to have $P_x[V_x < \infty] > 0$. and $E_x[V_x] = \infty$.

Lemma: For every $i \geq 0$, $x \in E$, we have

$$P_x[V_x \geq i] = \rho_x^i \quad \text{where} \quad \rho_x = P_x[H_x < \infty].$$

Proof: Define for every $i \geq 0$, $T_i = \min\{n \geq 1 : \sum_{k=1}^{n} \mathbb{1}_{X_k = x} = i\}$

T_i = "time of the i-th visit of x"

(conventions: $\min \emptyset = +\infty$, $T_0 = 0$)

T_i is a stopping time because

$$\{T_i = n\} = \left\{ \sum_{k=1}^{n-1} \mathbb{1}_{X_k = x} = i-1,\ X_n = x \right\} \in \mathcal{F}_n.$$

For every $i \geq 1$, we have

$$P_x[V_x \geq i] = P_x[T_i < \infty, T_{i-1} < \infty]$$

$$= P_x[\underbrace{T_i < \infty}_{\{\exists n \geq 1 : X_{T_{i-1}+n} = x\}} \mid T_{i-1} < \infty, X_{T_{i-1}} = x] \underbrace{P_x[T_{i-1} < \infty]}_{P_x[V_x \geq i-1]}$$

$$\underset{\text{St. MP}}{\overset{\uparrow}{=}} P_x[T_1 < \infty] \, P_x[V_x \geq i-1]$$

$$= \rho_x P_x[V_x \geq i-1] \underset{\text{induction}}{=} \rho_x^i \qquad \blacksquare$$

Proof of the theorem:

If x is recurrent, we have

$$P_x[V_x = +\infty] = P_x\Big[\bigcap_{i \geq 1} \{V_x \geq i\}\Big] \overset{\uparrow}{=} \lim_{i \to \infty} P_x[V_x \geq i] \underset{\uparrow}{=} 1$$

$(\{V_x \geq i\} \supset \{V_x \geq i+1\})$ Lemma.

If x is transient, we have

$$E_x[V_x] = \sum_{i \geq 1} P_x[V_x \geq i]$$

$$= \sum_{i \geq 1} \rho_x^i = \frac{\rho_x}{1-\rho_x} < \infty \qquad \blacksquare$$

Rk: $E_x[V_x] = E_x\Big[\sum_{n \geq 0} 1\!\!1_{X_n = x}\Big] = \sum_n p_{xx}^{(n)}.$

We conclude this section with a useful consequence of the dichotomy theorem, when E is finite.

Prop. Assume E is finite. Then there exists a recurrent state $x \in E$.

Proof: Observe that $\sum_{x \in E} V_x = \sum_{x \in E} \sum_{n \geq 0} \mathbb{1}_{X_n = x}$

$$= \sum_{n \geq 0} \underbrace{\Big(\sum_{x \in E} \mathbb{1}_{X_n = x} \Big)}_{=1} = +\infty$$

Fix $y \in E$. We have $\sum_{x} E_y[V_x] = +\infty$.

Hence, there must exist $x \in E$ s.t. $E_y[V_x] = +\infty$.

Using that $V_x = V_x \mathbb{1}_{H_x < \infty}$, we find

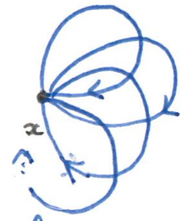
$$+\infty = E_y\left[V_x \mathbb{1}_{H_x < \infty}\right] \overset{\text{St.MP}}{=} \big(1 + E_x[V_x]\big) P_y[H_x < \infty]$$

$$\leq E_x[V_x]$$

Therefore, $E_x[V_x] = +\infty$, which concludes that x is recurrent.

Illustration:

x recurrent ($\rho_x = 1$)

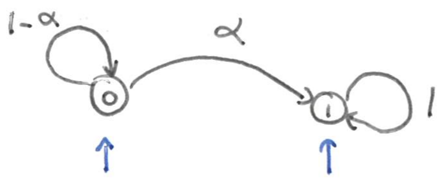


The chain always come back

x transient ($\rho_x < 1$)

The chain exits after a geometric number of visits.

Example: two-state MC. $\alpha \in (0, 1]$



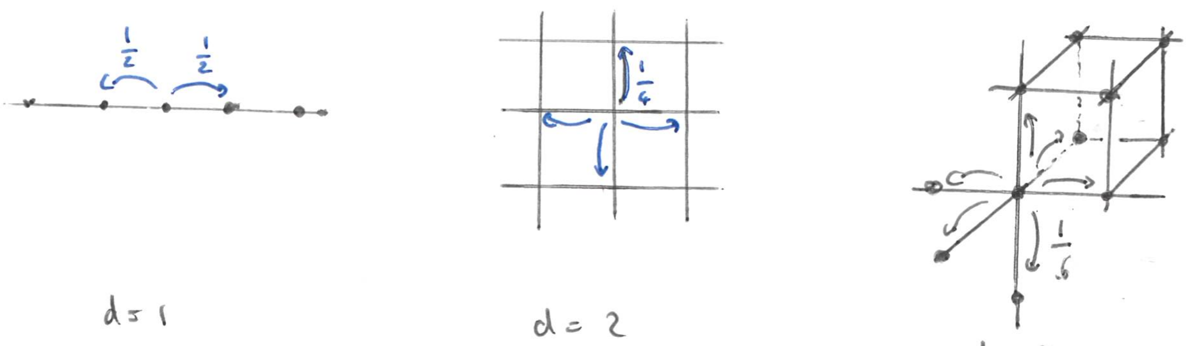
transient recurrent

2 RECURRENCE / TRANSIENCE FOR THE SRW ON $\mathbb{Z}^d$.

In this section we consider the SRW (simple random walk on $\mathbb{Z}^d$):

$E = \mathbb{Z}^d$, $d \geq 1$
$$p_{x,y} = \begin{cases} \frac{1}{2d} & \text{if } |x_1 - y_1| + \dots + |x_d - y_d| = 1 \\ 0 & \text{otherwise.} \end{cases}$$

Thm: For the SRW (ie the chain with transition p as above), every state is
- recurrent if $d = 1, 2$,
- transient if $d \geq 3$.



$d = 1$ $d = 2$ $d = 3$

Proof: Let $(Z_k)_{k\geq 1}$ be iid r.v. (on some $(\Omega, \mathcal{F}, \mathbb{P})$.)

with $\mathbb{P}[Z_1 = \pm e_i] = \frac{1}{2d}$ $\quad i = 1, \dots, d$.

Define $X_n = \sum_{k=1}^{n} Z_k$ $\quad \rightarrow (X_n)_{n\geq 0}$ is a $MC(\delta_0, p)$

$$E[V_0] = E\Big[\sum_{n\geq 1} \mathbb{1}_{X_n = 0}\Big] = \sum_{n\geq 1} \mathbb{P}[X_n = 0]$$

$$\mathbb{P}[X_n = x] = \mathbb{P}[Z_1 + \dots + Z_n = x] = \sum_{\delta_1 + \dots + \delta_n = x} \mathbb{P}[Z_1 = \delta_1] \dots \mathbb{P}[Z_n = \delta_n]$$

↳ not easy to calculate...

$$(\mathbb{P}[X_n = x])_{x \in \mathbb{Z}^d} \xleftrightarrow[\text{transform}]{\text{Fourier}} E[e^{iX_n}] = E[e^{iZ_1}]^n$$

easy to calculate

Define $\forall \xi \in \mathbb{T}^d := [-\pi, \pi)^d$ $\quad \varphi(\xi) = E[e^{i \xi \cdot Z_1}]$.

We have
$$\varphi(\xi) = \frac{1}{2d} \sum_{i=1}^{d} \left(e^{i \xi \cdot e_i} + e^{-i \xi \cdot e_i}\right)$$
$$= \frac{1}{d} \sum_{i=1}^{d} \cos(\xi_i).$$

Fix $n \geq 0$. By independence, the characteristic function of X_n is

$$\varphi_{X_n}(\xi) := E[e^{i \xi \cdot X_n}] = E[e^{i \xi \cdot Z_1 + \dots + i \xi \cdot Z_n}]$$
$$= \varphi(\xi)^n.$$

By Fourier inversion, we have

$$\mathbb{P}[X_n = 0] = \frac{1}{(2\pi)^d} \int_{\mathbb{T}^d} \varphi(\vartheta)^n d\vartheta$$

(this can be checked directly:

$$\int_{[0,2\pi)^d} \varphi(\vartheta)^n d\vartheta = \int_{[-\pi,\pi)^d} \sum_{x \in \mathbb{Z}^d} \mathbb{P}[X_n = x] e^{i\vartheta \cdot x} d\vartheta$$

$$\overset{\text{Fub.}}{=} \sum_{x \in \mathbb{Z}^d} \mathbb{P}[X_n = x] \underbrace{\int_{[0,2\pi)^d} e^{i\vartheta\cdot x} d\vartheta}_{= \begin{cases} (2\pi)^d & x = 0 \\ 0 & x \neq 0 \end{cases}} \Bigg)$$

Therefore,

$$(2\pi)^d \sum_{n \geq 0} \mathbb{P}[X_n = 0] = \sum_{n \geq 1} \int_{\mathbb{T}^d} \varphi(\vartheta)^n d\vartheta$$

monotone convergence

$$= \lim_{\substack{\alpha \to 1 \\ <}} \sum_{n \geq 1} \int_{\mathbb{T}^d} (\alpha \varphi(\vartheta))^n d\vartheta$$

Fubini.

$$= \lim_{\substack{\alpha \to 1 \\ <}} \int_{\mathbb{T}^d} \frac{1}{1 - \alpha \varphi(\vartheta)} d\vartheta$$

monotone convergence

$$= \int_{\mathbb{T}^d} \frac{1}{1 - \varphi(\vartheta)} d\vartheta .$$

Using $\forall\ \xi_i \in [-\pi,\pi)$ $\quad \frac{\xi_i^2}{4} \leq 1-\cos(\xi_i) \leq \frac{\xi_i^2}{2}$

we get $\quad \frac{1}{4d}\|\xi\|_2^2 \leq 1-\varphi(\xi) \leq \frac{1}{2d}\|\xi\|_2^2,$

And therefore

$$\sum \mathbb{P}[X_n=0] < \infty \iff \int_{B(0,1)} \frac{d\xi}{\|\xi\|_2^2} < \infty$$

$$\iff d > 2$$

(For the last equivalence, one can use a change of variable into polar coordinates $\int_{B(0,1)} \frac{d\xi}{\|\xi\|^2} = \int_{r=0}^{1} \frac{\text{"}Area(\partial B(0,r))\text{"}}{r^2}\,dr$

$$= Cte \int_{r=0}^{1} r^{d-3}\,dr.$$

or, one can use homogeneity: Define $A_i = B(0,\frac{1}{2^i}) \setminus B(0,\frac{1}{2^{i+1}})$

Using the change of variable $\psi = 2^i \xi$, we find

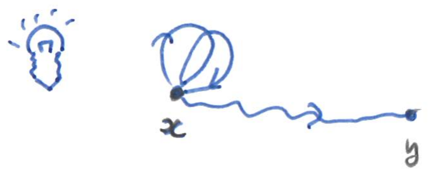
$$\int_{A_i} \frac{d\xi}{\|\xi\|^2} = \int_{A_0} \frac{2^{2i}\,d\psi}{\|\psi\|^2} \times (2^i)^{-d}\,d\psi = (2^i)^{2-d} \underbrace{\int_{A_0} \frac{d\psi}{\|\psi\|^2}}_{=:\,I_0}.$$

Therefore $\int_{B(0,1)} \frac{d\xi}{\|\xi\|^2} = \sum_{i=0}^{\infty} \int_{A_i} \frac{d\xi}{\|\xi\|^2} = I_0 \cdot \underbrace{\sum_{i=0}^{\infty} (2^i)^{2-d}}_{<\infty \text{ iff } d>2}$) $\blacksquare$

3 CLASSIFICATION OF STATES.

Thm: Let $x, y \in E$ such that $x \to y$.

If x is recurrent, then y is recurrent and $P_x[H_y < \infty] = P_y[H_x < \infty] = 1$. (in particular we have $x \leftrightarrow y$)



at each visit of x, the chain has >0 probability to hit y. since x is visited infinitely many times, y must also be visited infinitely many times.

Proof: Assume $y \neq x$. and x recurrent

Let $x, z_1, \dots, z_{k-1}, y$ distinct such that $p_{x z_1} \cdots p_{z_{k-1} y} > 0$

We have

$$0 = P_x[H_x = +\infty]$$

$$\geq P_x[X_1 = z_1, \dots, X_k = y, \forall n \geq 1 \quad X_{k+n} \neq x]$$

$$\overset{SiMP}{=} \underbrace{P_x[X_1 = z_1, \dots, X_k = y]}_{>0} \underbrace{P_y[\forall n \geq 1 \quad X_n \neq x]}_{= P_y[H_x = +\infty]}$$

Hence $P_y[H_x < \infty] = 1$.

Now, let us prove that y is recurrent. Define m, n s.t. $p^{(n)}_{xy}, p^{(m)}_{yx} > 0$.

We have $E_y[V_y] = \sum_{k\geq 1} p_{yy}^{(k)} \geq \sum_{k\geq 1} p_{yy}^{(m+k+n)}$

$$\overset{CK}{\geq} \underbrace{p_{yx}^{(m)}}_{>0} \underbrace{\left(\sum_{k\geq 1} p_{xx}^{(k)}\right)}_{=+\infty} \underbrace{p_{xy}^{(n)}}_{>0}$$

Therefore y is recurrent. It remains to prove $P_x[H_y < \infty] = 1$, which follows from $y \to x$ and y is recurrent, as before. ∎

Corollary 1: Let C be a communication class for p.
Either $\forall x \in C$, x is recurrent, ("C is recurrent")
or $\forall x \in C$, x is transient. ("C is transient")

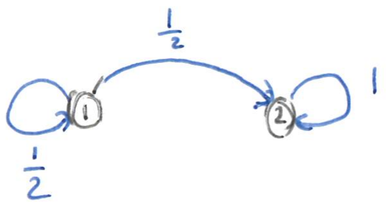
Proof: If $x \leftrightarrow y$ we have (x recurrent) $\Leftrightarrow$ (y recurrent). ∎

Corollary 2: A recurrent class is always closed.

Pf: Let C be a recurrent class. If $x \in C$ and $x \to y$ then we must have $y \to x$ (because x recurrent), therefore $y \in C$.

Corollary 2 give a criterion for transience: If $x \to y$ but $y \not\to x$, then x is transient.

Example:



$1 \to 2$ but $2 \not\to 1$. Hence 1 is transient.

In general one can always partition

$$\boxed{E = T \cup R_1 \cup R_2 \cup \dots}$$

where $T = \{x : x \text{ is transient}\}$ "T is the union of all the transient classes"

$R_1, R_2, \dots$ are the recurrent classes.

If the chain starts at $x \in R_i$, then $X_n \in R_i$ $\forall n \geq 1$ a.s.

If it starts at $x \in T$ $\xrightarrow{\text{case 1}}$ $\forall n$ X_n stays on T

$\xrightarrow{\text{case 2}}$ X_n moves at some time to some R_i and stays there.

4 POSITIVE / NULL RECURRENCE.

Notation: For $x \in E$ $\quad m_x := E_x[H_x]$

Def: Let x be a recurrent state [i.e. $P_x[H_x < \infty] = 1$]

We say that x is
- positive recurrent if $m_x < \infty$
- null recurrent if $m_x = +\infty$.

Thm: [density of visit times]

Let $x, y \in E$ s.t. $x \leftrightarrow y$. Then

$$\boxed{\lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} p_{xy}^{(k)} = \frac{1}{m_y}}$$

Rk: Write $V_y^{(n)} = \sum_{k=1}^{n} \mathbb{1}_{X_k = y}$ "visits of y before time n"

$$\frac{1}{n} \sum_{k=1}^{n} p_{xy}^{(k)} = E_x\left[\frac{V_y^{(n)}}{n}\right]$$ "expected proportion of time spent at y"

If y transient or null-recurrent : $(m_y = +\infty)$

$$\lim_{n \to \infty} E_x\left[\frac{V_y^{(n)}}{n}\right] = 0$$ "null density of visits"

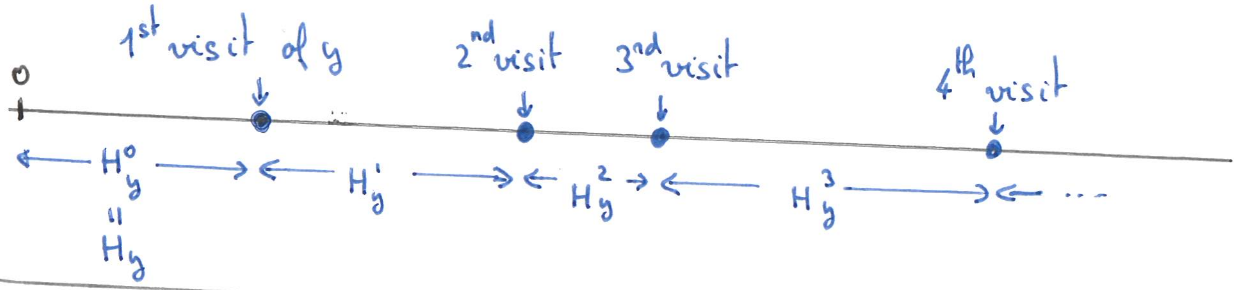
If y positive recurrent :

$$\lim_{n \to \infty} E_x\left[\frac{V_y^{(n)}}{n}\right] > 0$$ ">0 density of visits"

Def [inter-visit time]

Let $y \in E$. Define $H_y^0 = H_y$ and by induction

$$\forall i \geq 1 \quad H_y^i := \begin{cases} \min\{n \geq 1 : X_{H_y^0 + \dots + H_y^{i-1} + n} = y\} & \text{if } H_y^{i-1} < \infty \\ +\infty & \text{if } H_y^{i-1} = +\infty. \end{cases}$$



Lemma: Let $x, y \in E$ s.t $x \leftrightarrow y$. Assume that y is recurrent.

Then for every $j \geq 1$, $t_0, \dots, t_j \in \mathbb{N}$:

$$P_x[H_y^0 = t_0, \dots, H_y^j = t_j] = P_x[H_y = t_0] P_y[H_y = t_1] \dots P_y[H_y = t_j]$$

In particular, under P_x, $H_y^1, \dots, H_y^j, \dots$ are iid with law $P_x[H_y^j = t] = P_y[H_y = t]$

Proof: By induction on j. (for simplicity, we write $H^i = H_y^i$)

. By def. we have $\forall t \in \mathbb{N}$ $P_x[H^0 = t] = P_x[H_y = t]$.

Let $j \geq 0$ and assume that the equation holds.

First, observe that

$$P_x[H^0 < \infty, \dots, H^j < \infty] = \sum_{t_0, \dots, t_j} P_x[H^0 = t_0, \dots H^i = t_i]$$

$$= \underbrace{P_x[H_y < \infty)}_{=1} \cdot \underbrace{P_y[H_y < \infty]^i}_{=1} = 1.$$

Hence the stopping time $T = H_y^0 + \dots + H_y^j$ is finite P_x - a.s. and $X_T = y$ by definition. Hence, for every $t_0, \dots, t_{j+1} \in \mathbb{N}$, we have

$$P_x\left[H^0 = t_0, \dots, H^{j+1} = t_{j+1}\right] = P_x\left[\underbrace{H^0 = t_0, \dots,}_{\in \mathcal{F}_T} \underbrace{H^{j+1} = t_{j+1}}_{\min\{n \geq 1 : X_{T+n} = y\} = t_{j+1}} \mid T < \infty, X_T = y\right]$$

$$\overset{\text{St. MP}}{=} P_x\left[H^0 = t_0, \dots, H^j = t_j\right] P_y\left[\min\{n \geq 1 : X_n = y\} = t_{j+1}\right]$$

$$\overset{\text{induction}}{=} P_x\left[H_y = t_0\right] P_y\left[H_y = t_1\right] \dots P_y\left[H_y = t_{j+1}\right] \blacksquare$$

<u>Proof of the theorem.</u>

If y is transient, then, we have $E_y[V_y] < \infty$ and therefore, by the strong Markov property we also have $E_x[V_y] < \infty$. Hence, for every $n \geq 1$

$$\frac{E_x[V_y^{(n)}]}{n} \leq \frac{E_x[V_y]}{n} \xrightarrow[n \to \infty]{} 0 .$$

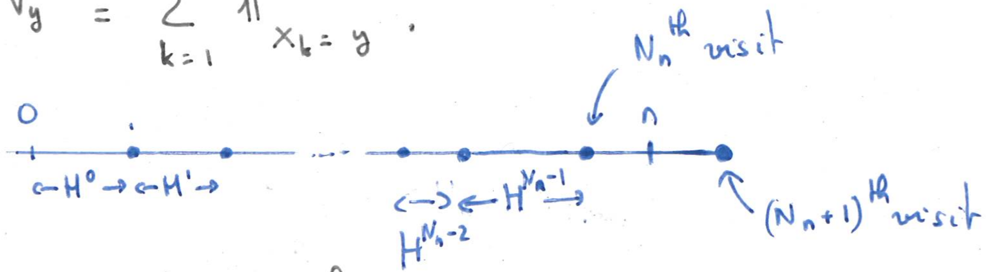
Now let us assume that y is recurrent. By the lemma, the random variables $H^1, H^2, \dots$ are iid under P_x and satisfy $E_x[H^1] = E_y[H_y] = m_y$.

Hence by the law of large number, and using $P_x[H^0 < \infty] = 1$, we have

$$\lim_{i\to\infty} \frac{H^0 + \dots + H^i}{i} = m_y \qquad P_x\text{-a.s.}$$

(this includes the case $m_y = +\infty$)

Write $N_n = V_y^{(n)} = \sum_{k=1}^{n} 1\!\!1_{X_k = y}$.



By definition, we have for every $n \geq 1$

$$H^0 + \dots + H^{N_n - 1} \leq n < H^0 + \dots + H^{N_n}$$

Hence for every $n \geq 1$

$$\frac{N_n}{\underbrace{H^0 + \dots + H^{N_n}}_{\xrightarrow[P_x\text{-a.s.}]{} \frac{1}{m_y}}} < \frac{V_y^{(n)}}{n} \leq \frac{N_n}{\underbrace{H^0 + \dots + H^{N_n - 1}}_{\xrightarrow[P_x\text{-a.s.}]{} \frac{1}{m_y}}}$$

And we can conclude that $E_x\left[\frac{V_y^{(n)}}{n}\right] \xrightarrow[n\to\infty]{} \frac{1}{m_y}$ by dominated cv. ■

Prop. [Classification of recurrent classes]

> Let R be a recurrent class. Then
>
> - either $\forall x \in R$ x is >0 recurrent, "R is a >0 rec. class"
> - or $\forall x \in R$ x is null recurrent. "R is a null rec. class"

Proof: Let $x, y \in E$ s.t. $x \leftrightarrow y$. Assume x >0 recurrent.

Fix $k \geq 0$ s.t. $p_{xy}^{(k)} > 0$

By Chapman-Kolmogorov, we have

$$\forall j \geq 1 \quad p_{xy}^{(k+j)} \geq p_{xx}^{(j)} \, p_{xy}^{(k)}$$

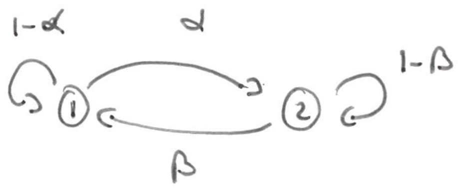
Hence

$$\underbrace{\frac{1}{n}\sum_{i=1}^{n} p_{xy}^{(i)}}_{\substack{\downarrow n\to\infty \\ \frac{1}{E_y[H_y]}}} \geq \underbrace{\left(\frac{1}{n}\sum_{j=1}^{n-k} p_{xx}^{(j)}\right)}_{\substack{\downarrow n\to\infty \\ \frac{1}{E_x[H_x]}}} \underbrace{p_{xy}^{(k)}}_{>0}$$

Therefore $\dfrac{1}{E_y[H_y]} > 0$ and y is >0 recurrent ∎

Examples:

- 2-state MC



$\beta > 0$

For $k \geq 1$ $P_1[H_1 \geq k] = P_1[X_1 = \dots = X_{k-1} = 2]$

$$= \alpha(1-\beta)^{k-1}$$

Hence $E_1[H_1] = \sum_k \alpha(1-\beta)^{k-1} < \infty$ $\quad$ 1 is >0-rec.

• <u>SRW on $\mathbb{Z}$</u>.

We have

$$P_0[H_0 \geq k] \geq P_0\Big[X_1 = -1,\ \max_{1 \leq m \leq k}(X_m) \leq -1\Big]$$

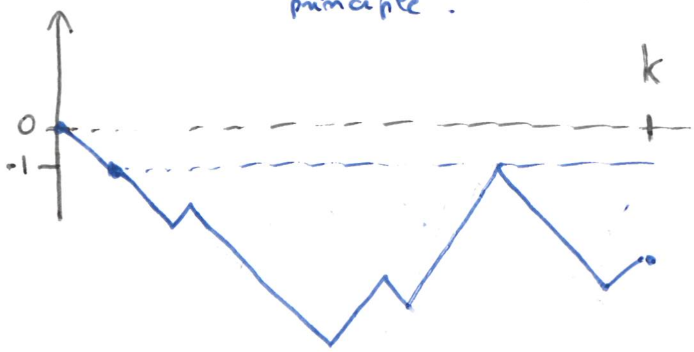
$$\overset{\text{SiMP}}{=} P_0[X_1 = -1]\ P_{-1}\Big[\max_{0 \leq m \leq k-1}(X_m) \leq -1\Big]$$

$$= \frac{1}{2} P_0\Big[\underbrace{\max_{0 \leq m \leq k-1}(X_m) \leq 0}\Big]$$

translation invariance

$$= P_0\big[|X_{k-1}| = 0\big] \quad \text{if } k \text{ is odd}$$

reflection principle.



Hence $E_0[H_0] \geq \sum_{\substack{k \text{ odd} \\ k \geq 1}} P_0\big[|X_{k-1}| = 0\big] = +\infty$

because the SRW on $\mathbb{Z}$ is recurrent.

Hence 0 is null-recurrent (and therefore every $x \in \mathbb{Z}$ is null-recurrent by irreducibility)

We finish this section with a simple condition ensuring positive recurrence

Prop. Any finite recurrent class is positive recurrent. In particular, if E is finite, then every recurrent state is >0 recurrent.

Proof: Let R be a finite recurrent class, $x \in R$. Since R is closed we have for every $n \geq 1$

$$1 = P_x[X_n \in R] = \sum_{y \in R} p_{xy}^{(n)} .$$

Hence
$$1 = \sum_{y \in R} \underbrace{\left(\frac{1}{n} \sum_{k=1}^{n} p_{xy}^{(k)} \right)}_{\xrightarrow[n\to\infty]{} \frac{1}{E_y[H_y]}}$$

Therefore, there must exist $y \in R$ s.t. $E_y[H_y] < \infty$, which implies that the class is >0 recurrent. $\blacksquare$

5 STATIONARY DISTRIBUTIONS FOR IRREDUCIBLE CHAINS.

Theorem: Assume that p is irreducible.

- If the chain is transient or null recurrent, then there is no stationary distribution.
- If the chain is >0 recurrent, then there exists a unique stationary distribution, given by

$$\boxed{\forall x \in E \quad \pi(x) = \frac{1}{E_x[H_x]}.}$$

Proof. Assume that the chain is transient or null recurrent

Assume for contradiction that there exists a stationary distribution π. For every $x \in E$, we have

$$\forall n \geq 1 \quad \pi(x) = \frac{1}{n} \sum_{k=1}^{n} P_\pi [X_k = x]$$

$$\overset{\text{Fubini}}{=} \sum_{y \in E} \pi(y) \underbrace{\frac{1}{n} \sum_{k=1}^{n} P_y [X_k = x]}_{\xrightarrow[n\to\infty]{} \frac{1}{E_x[H_x]} = 0}$$

Hence, by dominated convergence, we have

$$\forall x \in E \quad \pi(x) = \frac{1}{E_x[H_x]} = 0,$$

which contradicts $\sum_{x \in E} \pi(x) = 1$.

- Now assume that the chain is >0 recurrent.

The same calculation as before shows that the only possible candidate is given by

$$\forall x \quad \pi(x) := \frac{1}{E_x[H_x]} .$$

To conclude one needs to prove that the measure defined above is indeed a stationary distribution.

First, fix $k \geq 1$. We have

$$\forall y \in E \qquad \frac{1}{E_y[H_y]} = \lim_{n\to\infty} \frac{1}{n} \sum_{j=k}^{n} p_{yy}^{(j)}$$

$$\overset{CK}{=} \lim_{n\to\infty} \sum_{x\in E} \left(\frac{1}{n} \sum_{j=k}^{n} p_{yx}^{(j-k)} \right) p_{xy}^{(k)}$$

$$\underset{(1)}{\overset{\text{Fatou}}{\geq}} \sum_{x\in E} \liminf_{n\to\infty} \left(\frac{1}{n} \sum_{j=k}^{n} p_{yx}^{(j-k)} \right) p_{xy}^{(k)}$$

$$= \sum_{x\in E} \frac{1}{E_x[H_x]} p_{xy}^{(k)}$$

Similarly, we have, for a fixed x

$$1 = \lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} P_x[X_j \in E] = \lim_{n\to\infty} \sum_{y\in E} \frac{1}{n} \sum_{j=1}^{n} P_x[X_j = y]$$

$$\underset{(2)}{\overset{\text{Fatou}}{\geq}} \sum_{y\in E} \frac{1}{E_y[H_y]}$$

In order to conclude one needs to prove that the two inequalities above are equalities.

First by summing the first inequality over y we have

$$\sum_{y \in E} \frac{1}{E_y[H_y]} \geq \sum_{y \in E} \left(\sum_{x \in E} \frac{1}{E_x[H_x]} P_{xy}^{(k)} \right) = \sum_{x \in E} \frac{1}{E_x[H_x]}$$

And therefore the inequality (1) should be an equality and we have for every $k \geq 1$

$$\forall y \quad \frac{1}{E_y[H_y]} = \sum_{x \in E} \frac{1}{E_x[H_x]} P_{xy}^{(k)} \qquad (*)$$

We now use this equality to prove that (2) is also an equality. Fix $y \in E$. (By >0 recurrence, $\frac{1}{E_y[H_y]} > 0$)

We have
$$\frac{1}{E_y[H_y]} = \lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} \left(\sum_{x \in E} \frac{1}{E_x[H_x]} P_{xy}^{(k)} \right)$$

$$= \lim_{n \to \infty} \sum_{x \in E} \frac{1}{E_x[H_x]} \times \left(\frac{1}{n} \sum_{k=1}^{n} P_{xy}^{(k)} \right)$$

$$\overset{\text{(dominated convergence)}}{=} \sum_{x \in E} \frac{1}{E_x[H_x]} \times \frac{1}{E_y[H_y]}$$

Hence $\pi(x) = \frac{1}{E_x[H_x]}$ defines a distribution, which is stationary, by $(*)$.

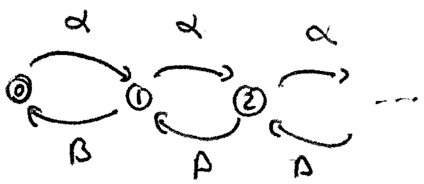
Applications

- The SRW on $\mathbb{Z}$ is null-recurrent.

(We already prove this result by showing $E_0[H_0] = +\infty$, here we give a second proof).

The SRW on $\mathbb{Z}$ is recurrent, but has no stationary distribution, hence it is null recurrent.

- Consider the reflected random walk on $\mathbb{N}$ $\quad \alpha + \beta = 1 \quad \alpha < \beta$



$E = \mathbb{N}$

$p_{i(i+1)} = \alpha \qquad i \geq 1$

$p_{i(i-1)} = \beta \qquad i \geq 1$

$p_{01} = 1$

One can check that $\lambda_i = \left(\frac{\alpha}{\beta}\right)^i$ $\; i \geq 1$ and $\lambda_0 = \alpha$ defines an invariant measure $\left(\forall i \geq 1 \quad \lambda_i = \lambda_{i-1}\alpha + \lambda_{i+1} \beta \right.$ and $\left. \lambda_0 = \lambda_1 \beta \right)$

Therefore $\pi_i = \dfrac{1}{\sum_{j \geq 0} \lambda_j} \lambda_i$ is a stationary distribution

The reflected random is > 0 recurrent if $\alpha < \beta$.

6 PERIODICITY.

Def. Let $x \in E$. The <u>period</u> of x is defined by

$$d_x = \gcd\{n \geq 1 : P_{xx}^{(n)} > 0\}.$$

(Convention: $\gcd(\emptyset) = +\infty$)

Prop. Let $x, y \in E$ s.t. $x \longleftrightarrow y$. Then $d_x = d_y$.

Proof: Let $x \neq y$. We prove that $d_y \mid d_x$.

First, let us fix $k, \ell \geq 0$ s.t. $P_{yx}^{(k)}, P_{xy}^{(\ell)} > 0$

Since $P_{yy}^{(k+\ell)} \geq P_{yx}^{(k)} P_{xy}^{(\ell)} > 0$ we have $\underline{d_y \mid k+\ell}$.

Now, we prove that d_y is a common divisor of $\{n \geq 1 \;\; P_{xx}^{(n)} > 0\}$. (this will imply $d_y \mid d_x$)

For every $n \geq 1$ satisfying $P_{xx}^{(n)} > 0$, we have

$$P_{yy}^{(k+n+\ell)} \geq P_{yx}^{(k)} P_{xx}^{(n)} P_{xy}^{(\ell)} > 0,$$

hence $d_y \mid k+\ell+n$. Since $d_y \mid k+\ell$, we also have $d_y \mid n$. ■

<u>Consequence</u>: if the chain p is irreducible, we have

$$\forall x, y \in E \quad d_x = d_y.$$

Def : We say that the chain p is aperiodic if

$$\forall x \in E \quad d_x = 1$$

Prop: Let $x \in E$. We have

$$(d_x = 1) \iff (\exists n_0 \geq 1 \text{ s.t. } \forall n \geq n_0 \ p_{xx}^{(n)} > 0)$$

We use the following lemma from number theory.

Lemma: Let $A \subset \mathbb{N} \setminus \{0\}$ stable under addition ($x, y \in A \Rightarrow x+y \in A$). Then

$$(\gcd(A) = 1) \iff (\exists n_0 \in \mathbb{N} \text{ s.t. } \{n \in \mathbb{N} : n \geq n_0\} \subset A)$$

Proof: $\boxed{\Leftarrow}$: follows from the fact that $\gcd(n_0, n_0+1) = 1$

$\boxed{\Rightarrow}$ Assume $\gcd(A) = 1$. Let $a \in A$ arbitrary and $a = \prod_{i=1}^{k} p_i^{\alpha_i}$ be its prime factorization ($k \geq 0$, $p_1, \dots, p_k$ primes $\alpha_1, \dots, \alpha_k \geq 1$)

Since $\gcd(A) = 1$, one can find $b_1, \dots, b_k \in A$ s.t. $\forall i \ p_i \nmid b_i$. This implies

$$\gcd(a, b_1, \dots, b_k) = 1.$$

Write $d = \gcd(b_1, \dots, b_k)$. By Bezout theorem, we can pick $u_1, \dots, u_k \in \mathbb{Z}$ s.t.

$$u_1 b_1 + \dots + u_k b_k = d$$

Now, choose and integer λ large enough s.t. $u_i + \lambda a \geq 0$ for every i and define

$$b = (u_1 + \lambda a) b_1 + \cdots + (u_k + \lambda a) b_k$$

$$= d + \lambda (b_1 + \cdots + b_k) a$$

The first expression shows that $b \in A$, and the second expression implies that $\gcd(a, b) = \gcd(a, d) = 1$

To summarize, we found $a, b \in A$ s.t.

$$\gcd(a, b) = 1.$$

Without loss of generality, we may assume $a < b$. Since $\gcd(a,b) = 1$, the set $B = \{b, 2b, \ldots, ab\}$ covers all the residue classes modulo a. Since $a < b$, this implies that $B + \{ka,\ k \in \mathbb{N}\}$ includes every number $\geq ab$ -

This concludes the proof of the lemma by choosing $n_0 = ab$ ∎

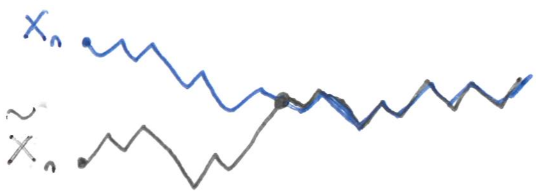
Proof of the proposition:

The set $A_x := \{n \geq 1 \text{ s.t. } p_{xx}^{(n)} > 0\}$ is stable under addition because $p_{xx}^{(m+n)} \geq p_{xx}^{(m)} p_{xx}^{(n)}$ for every $m, n \geq 1$.

The proof follows by applying the lemma to $A = A_x$ ∎

7 THE COUPLING METHOD.

Goal: define two Markov chains $(X_n) \sim MC(\mu, P)$ and $(\tilde{X}_n) \sim MC(\nu, P)$ on the same probability space such that $X_n = \tilde{X}_n$ for n large.



To achieve that we first consider two independent chains $(X_n), (Y_n)$. We show that the two chains meet a.s. (under some assumptions on P) at some random time T. And then, we ask that the chains follow the same trajectory for $t \geq T$.

In order to introduce a suitable probability space, we consider the product chain, defined below.

Def. [Product chain]

Define for every $w = (x, y)$ and $w' = (x', y')$ in E^2

$$\bar{P}_{w,w'} = P_{xx'} \, P_{yy'}.$$

Rk: $\displaystyle\sum_{w' \in E} \bar{P}_{w,w'} = \sum_{x', y' \in E} P_{xx'} \, P_{yy'} = 1$.

Therefore $\bar{P}$ is a transition probability on E^2.

Notation: Consider

- $(\Omega, \mathcal{F}, (P_w)_{w \in E^2})$ proba spaces.
- $(W_n)_{n \geq 0} = (X_n, Y_n)_{n \geq 0}$ r.v. on $(\Omega, \mathcal{F})$ s.t.

$\forall w \in E^2$ W_n is $MC(\delta_w, \bar{P})$ under P_w

[If μ, ν are distribution on E. $\mu \otimes \nu$ is a distribution on E^2 and we write.

$$P_{\mu \otimes \nu} = \sum_{(x,y) \in E^2} \mu(x) \nu(y) P_{(x,y)}$$]

Prop: Let μ, ν be two distributions on E.
Under $P_{\mu \otimes \nu}$, $(X_n)_{n \geq 0}$ is $MC(\mu, P)$ and $(Y_n)_{n \geq 0}$ is $MC(\nu, P)$ and they are independent.

Proof: For every $k \geq 0$ $x_0, \dots, x_k, y_0, \dots, y_k \in E$ we have

$$P_{\mu \otimes \nu}[X_0 = x_0, \dots, X_k = y_k, Y_0 = y_0, \dots, Y_k = y_k]$$

$$= P_{\mu \otimes \nu}[W_0 = (x_0, y_0) \dots W_k = (x_k, y_k)]$$

$$= \mu(x_0) P_{x_0 x_1} \cdots P_{x_{k-1} x_k} \nu(y_0) P_{y_0 y_1} \cdots P_{y_{k-1} y_k}$$

By summing over all possible $y_0, \dots, y_k \in E$, this implies that $(X_n)_n$ is $MC(\mu, P)$, and equivalently $(Y_n)_n$ is $MC(\nu, P)$.

To prove independence we need to show that for every measurable sets $A, B \subset E^{\mathbb{N}}$

$$P_{\mu \otimes \nu}[X \in A, Y \in B] = P_{\mu \otimes \nu}[X \in A]\, P_{\mu \otimes \nu}[Y \in B]$$

The computation above shows that it holds for sets of the form $A = \{(x_0, \dots, x_k)\} \times E^{\mathbb{N}}$ $B = \{(y_0, \dots, y_k)\} \times E^{\mathbb{N}}$.

Therefore, it holds for every cylindrical sets, and then for any measurable sets, by Dynkin's lemma. ∎

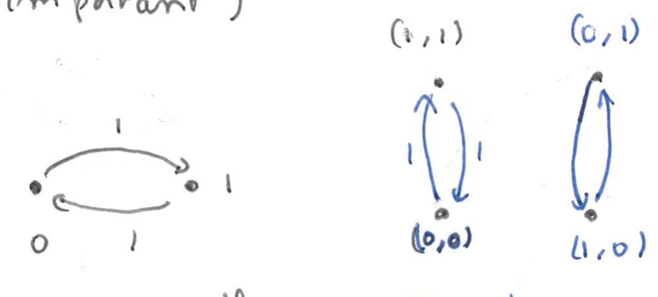
<u>Proposition:</u>

> If the chain p is irreducible, aperiodic and >0 recurrent, then $\bar{p}$ is also irreducible, aperiodic and >0 recurrent.

<u>Rk:</u> p irreducible $\not\Rightarrow$ $\bar{p}$ irreducible in general.

(aperiodicity is important)

for example



p irreducible $\qquad$ $\bar{P}$ not irreducible

<u>Proof.</u> Let $w = (x, y)$, $w' = (x', y') \in E^2$. By irreducibility, one can pick $k, \ell \geq 0$ s.t. $P^{(k)}_{xx'}, P^{(\ell)}_{yy'} > 0$.

For every $n \geq \max(k, \ell)$ we have

$$\bar{P}^{(n)}_{ww'} = P^{(n)}_{xx'} P^{(n)}_{yy'} \underset{CK}{\geq} P^{(k)}_{xx'} \underbrace{P^{(n-k)}_{x'x'}}_{\substack{>0 \\ \text{for } n \text{ large}}} P^{(\ell)}_{yy'} \underbrace{P^{(n-\ell)}_{y'y'}}_{\substack{>0 \\ \text{for } n \text{ large}}} > 0 .$$

for n large

This proves that $\bar{p}$ is irreducible aperiodic.

Since p is irr., >0 recurrent, it admits a stationary distribution π. For every $(y, y') \in E^2$, we have

$$\pi(y) \times \pi(y') = \sum_{x \in E} \pi(x)\, p_{xy} \sum_{x' \in E} \pi(x')\, p_{x'y'}$$

$$= \sum_{(x,x') \in E} \pi(x)\, \pi(x')\, p_{xy}\, p_{x'y'}$$

Hence $\pi \otimes \pi$ is stationary for $\bar{p}$, which implies that $\bar{p}$ is >0 recurrent. ∎

<u>Def.</u> $T := \min\{ n \geq 0 : X_n = Y_n \}$

<u>Rk:</u> $T = H_A$ where $A = \{ (x,y) \in E^2 : x = y \}$

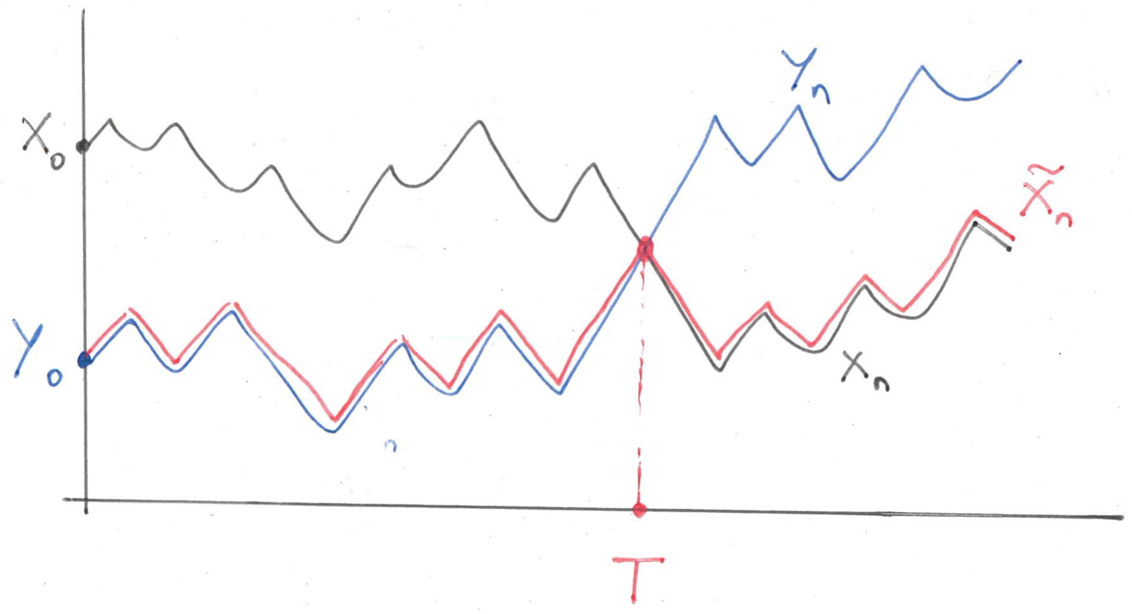
and therefore T is a stopping time.

<u>Prop:</u> For every μ, ν distributions on E.

$$\forall n \geq 0 \quad \sum_{x \in E} \left| P_\mu[X_n = x] - P_\nu[X_n = x] \right| \leq 2 P_{\mu \otimes \nu}[T > n]$$

<u>Proof:</u> We consider the product Markov Chain $W_n = (X_n, Y_n)$ under $P_{\mu \otimes \nu}$. Define for every n

$$\tilde{X}_n = \begin{cases} Y_n & \text{if } n < T \\ X_n & \text{if } n \geq T \end{cases}$$



We prove that $(\tilde{X}_n)$ is MC (ν, P) under $\mathbb{P} := P_{\gamma \otimes \nu}$

Let $n \geq 0$ and $x_0, \dots, x_n \in E$. By distinguishing between the possible values for T, we have

$$\mathbb{P}[\tilde{X}_0 = x_0, \dots, \tilde{X}_n = x_n] = \sum_{k \in \mathbb{N} \cup \{+\infty\}} \mathbb{P}[\tilde{X}_0 = x_0, \dots, \tilde{X}_n = x_n, T = k]$$

If $k > n$, the summand is equal to

$$\nu(x_0) P_{x_0 x_1} \cdots P_{x_{n-1} x_n} \times \mathbb{P}[T = k \mid Y_0 = x_0, \dots, Y_n = x_n]$$

If $k \leq n$, the summand is equal to

$$\mathbb{P}[\underbrace{Y_0 = x_0, \dots, Y_k = x_k, T = k}_{\in \mathcal{F}_T}, X_{T+1} = x_{k+1}, \dots, X_{T+n-k} = x_n]$$

$$\overset{\text{St M.P}}{=} \underbrace{\mathbb{P}[Y_0 = x_0, \dots, Y_k = x_k, T = k]}_{= \nu(x_0) P_{x_0 x_1} \cdots P_{x_{k-1} x_k} \mathbb{P}[T = k \mid Y_0 = x_0, \dots, Y_k = x_k]} \times \underbrace{P_{(x_k, x_k)}[X_1 = x_{k+1}, \dots, X_{n-k} = x_n]}_{= P_{x_k x_{k+1}} \cdots P_{x_{n-1} x_n}}$$

$$= \nu(x_0) P_{x_0 x_1} \cdots P_{x_{n-1} x_n} \times \mathbb{P}[T = k \mid Y_0 = x_0, \dots, Y_n = x_n]$$

For the last equality we use independence between $(X_n)_n$ and $(Y_n)_n$ to write $\mathbb{P}[T = k \mid Y_0 = x_0, \dots, Y_k = x_k]$ as $\mathbb{P}[\forall i < k \ X_i \neq x_i, X_k = x_k] = \mathbb{P}[T = k \mid Y_0 = x_0, \dots, Y_n = x_n]$

Finally using $\sum_{k \in \mathbb{N} \cup \{+\infty\}} P[T=k \mid Y_0 = x_0, \dots, Y_n = x_n] = 1$,

we obtain

$$\mathbb{P}[\tilde{X}_0 = x_0, \dots, \tilde{X}_n = x_n] = \nu(x_0)\, p_{x_0 x_1} \cdots p_{x_{n-1} x_n}.$$

We use the coupling between $(X_n)_n$ and $(\tilde{X}_n)_n$ to conclude the proof. For every $n \geq 0$

$$\sum_{x \in E} \left| P_\mu[X_n = x] - P_\nu[X_n = x] \right| = \sum_{x \in E} \left| \mathbb{P}[X_n = x] - \mathbb{P}[\tilde{X}_n = x] \right|$$

$$= \sum_{x \in E} \left| \cancel{\mathbb{P}[X_n = x, T \leq n]} + \mathbb{P}[X_n = x, T > n] - \cancel{P[\tilde{X}_n = x, T \leq n]} - \mathbb{P}[\tilde{X}_n = x, T > n] \right|$$

$$\leq \sum_{x \in E} \mathbb{P}[X_n = x, T > n] + \mathbb{P}[\tilde{X}_n = x, T > n] = 2\, \mathbb{P}[T > n] \quad \blacksquare$$

8 CONVERGENCE FOR IRREDUCIBLE APERIODIC CHAINS.

Thm Assume that the chain p is irreducible aperiodic and admits a stationary distribution π.
Then for every distribution μ on E and every $x \in E$

$$\boxed{\lim_{n \to \infty} P_\mu[X_n = x] = \pi(x)}$$

Rk: Equivalently

- Under P_μ $X_n \xrightarrow[n\to\infty]{(\text{law})} X_\infty$ where $X_\infty \sim \pi$.
- $\forall f : E \to \mathbb{R}$ bounded

$$\lim_{n\to\infty} E_\mu[f(X_n)] = \int_E f \, d\pi.$$

Proof: Consider the product chain $(X_n, Y_n)_{n\geq 0}$ introduced in the previous section. Since $\bar{p}$ is irreducible, >0 recurrent, the stopping time $T = \min\{n\geq 0 : X_n = Y_n\}$ is finite $P_{\mu\otimes\pi}$ - a.s. (indeed for a fixed $a\in E$, we have $T \leq H_{(a,a)} < \infty$ a.s.). For every $x\in E$

$$|P_\mu[X_n = x] - \pi(x)| = |P_\mu[X_n = x] - P_\pi[X_n = x]|$$

$$\leq 2 P_{\mu\otimes\pi}[T > n] \xrightarrow[n\to\infty]{} 0. \quad \blacksquare$$

Thm: Assume that the chain p is irreducible aperiodic, null recurrent or transient. Then for every distribution μ and every $x\in E$ we have

$$\boxed{\lim_{n\to\infty} P_\mu[X_n = x] = 0}$$

Lemma: Assume that the product chain $\bar{p}$ is irreducible recurrent. Then for every distribution μ and every $i \geq 0$, we have

$$\boxed{\lim_{n \to \infty} \left| P_\mu [X_n = x] - P_\mu [X_{n+i} = x] \right| = 0}.$$

Proof: Define $\mu_i(y) = P_\mu [X_i = y]$ ("$\mu_i = \mu P^i$")

Observe that

$$P_{\mu_i} [X_n = x] = \sum_y \mu_i(y) \, P_y [X_n = x]$$

$$\overset{\text{SiMP}}{=} \sum_y P_\mu [X_i = y] \, P_\mu [X_{n+i} = x \mid X_i = y]$$

$$= P_\mu [X_{n+i} = x]$$

Consider the product chain $(X_n, Y_n)_{n \geq 0}$ under $P_{\mu \otimes \mu_i}$ and define $T = \min\{n : X_n = Y_n\}$. ($T < \infty$ a.s. since $\bar{p}$ is rec.)

Hence $\left| P_\mu [X_n = x] - P_\mu [X_{n+i} = x] \right| = \left| P_\mu [X_n = x] - P_{\mu_i} [X_n = x] \right|$

$$\leq 2 P_{\mu \times \mu_i} [T > n] \xrightarrow[n \to \infty]{} 0 .$$

∎

Proof of the theorem.

Case 1: Assume that the chain $\bar{p}$ is transient.

Consider the product chain (X_n, Y_n) under $P_{\gamma\otimes\gamma}$

Since (x,x) is a transient state the last visit

time $L := \max\{n ; (X_n, Y_n) = (x,x)\}$ is $< \infty$

$P_{\gamma\otimes\gamma}$ a.s. Hence

$$P_\gamma[X_n = x]^2 = P_{\gamma\otimes\gamma}[X_n = x, Y_n = x]$$

$$\leq P_{\gamma\otimes\gamma}[L \geq n] \xrightarrow[n\to\infty]{} 0.$$

Case 2: Assume that $\bar{p}$ is recurrent.

Let $y \in E$. we wish to prove

$$p_{yx}^{(n)} \xrightarrow[n\to\infty]{} 0.$$

Fix $\varepsilon > 0$ and k s.t.

$$\frac{1}{k+1} \sum_{i\leq k} p_{xx}^{(i)} < \varepsilon$$

Define $H = \min\{j \geq n ; X_j = x\}$ (H is a stopping time)

For every $n \geq 0$ we have

$$\frac{1}{k+1} \sum_{i=0}^{k} P_y[X_{n+i} = x] \leq \frac{1}{k+1} \sum_{i=0}^{k} P_y[X_{H+i} = x]$$

$$\overset{SMP}{=} \frac{1}{k+1} \sum_{i=0}^{k} P_x[X_i = x] \leq \varepsilon$$

Now,

$$P_y[X_n = x] = \frac{1}{k+1} \sum_{i=0}^{k} P_y[X_n = x]$$

$$\leq \frac{1}{k+1} \sum_{i=0}^{k} \left| P_y[X_n = x] - P_y[X_{n+i} = x] \right|$$

$$+ \underbrace{\frac{1}{k+1} \sum_{i=0}^{k} P_y[X_{n+i} = x]}_{\leq \varepsilon}$$

Since $\bar{P}$ is irreducible and recurrent, the lemma concludes that

$$\limsup_{n \to \infty} P_y[X_n = x] \leq \varepsilon \qquad \blacksquare$$