V. Markov Chains.

Given two random variables X, Y defined on $(\Omega, \mathcal{F}, \mathbb{P})$, $\mathcal{A}$ a σ-algebra $\subset \mathcal{F}$, we say that

X and Y are independent conditionally to $\mathcal{A}$ if, $\forall F, G \to \mathbb{R}$ bounded measurable,

$$\mathbb{E}[F(X)G(Y) | \mathcal{A}] = \mathbb{E}[F(X) | \mathcal{A}]\, \mathbb{E}[G(Y) | \mathcal{A}].$$

This is different from Independence, but equivalent if $(X, Y) \perp \mathcal{A}$.

Previously we considered sequences $(X_n)_{n \geq 0}$ of independent random variables: $\forall n \geq 1$, X_n is independent from $(X_0, \dots, X_{n-1})$.

Weaker hypothesis: $\forall n \geq 1$, $\underbrace{X_{n+1}}_{\text{future}}$ and $\underbrace{(X_0, \dots, X_{n-1})}_{\text{past}}$ are independent conditionally to $\underbrace{X_n}_{\text{present}}$.

Chapter V. Markov chains.

1. Definitions and constructions.

We fix a countable set $\mathcal{X}$ (the state space), endowed with the σ-algebra $\mathcal{P}(\mathcal{X})$.

<u>Definition</u> A stochastic matrix on $\mathcal{X}$ is a family $(P(x,y))_{x,y \in \mathcal{X}}$ of nonnegative real numbers with:

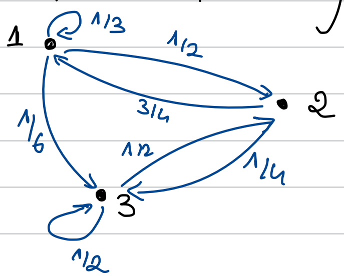
$$\sum_{y \in \mathcal{X}} P(x,y) = 1 \qquad \forall x \in \mathcal{X}.$$

Hence P is a (possibly infinite) matrix where each row is a probability measure on $\mathcal{X}$.

<u>example</u> $\mathcal{X} = [\![1, 3]\!]$ $\qquad P = \begin{pmatrix} \frac{1}{3} & \frac{1}{2} & \frac{1}{6} \\ 3/4 & 0 & 1/4 \\ 0 & 1/2 & 1/2 \end{pmatrix}$

We can represent a stochastic matrix by a labelled oriented graph:

$\mathcal{G}_P$



For each vertex $x \in \mathcal{X}$, the sum of labels of outgoing arrows is 1.

remarks: ① Given a probability measure $\pi = (\pi(x))_{x\in\mathcal{X}}$ represented as a row vector, we can consider

$$\pi P = \text{matrix product}$$

$$\pi P(y) = \sum_{x\in\mathcal{X}} \pi(x) P(x,y).$$

Then, πP is again a probability measure:

$$\sum_{y\in\mathcal{X}} \pi P(y) = \sum_{x,y\in\mathcal{X}} \pi(x)\, P(x,y) = \sum_{x\in\mathcal{X}} \pi(x) = 1.$$

② Therefore, even if $|\mathcal{X}| = +\infty$, we can take a product PQ of two stochastic matrices on $\mathcal{X}$, and this is again a stochastic matrix.

(in general, one has to be careful with products of infinite matrices, but here there are no problems of associativity because all terms are nonnegative)

Definition Given $\mathcal{X}$ and | a stochastic matrix P on $\mathcal{X}$
| a probability measure π_0 on $\mathcal{X}$ (row vector),

a Markov chain on $\mathcal{X}$ is a sequence of random variables $(X_n)_{n\in\mathbb{N}}$ defined on a space $(\Omega, \mathcal{F}, \mathbb{P})$, with values in $\mathcal{X}$, such that, for any states $x_0, x_1, \dots, x_n$:

$$\mathbb{P}[X_0 = x_0, X_1 = x_1, \dots, X_n = x_n] = \pi_0(x_0) P(x_0,x_1) P(x_1,x_2) \dots P(x_{n-1},x_n). \quad (*)$$

We then say that π_0 is the initial measure and P the transition matrix.

<u>remarks</u> ① The formula determines the law of $(X_n)_{n\in\mathbb{N}}$, viewed as a random element of $\left(\mathcal{X}^{\mathbb{N}}, \mathcal{P}(\mathcal{X})^{\otimes \mathbb{N}}\right)$.
(probability of cylinders).

We denote $\mathbb{P}_{(\pi_0, P)}$ this law, and $\mathbb{E}_{(\pi_0, P)}$ the expectation of observables with respect to this law.

② This is not the standard definition!

(∗) is equivalent to:

1) X_0 has law π_0

and 2) $\forall x_0, \dots, x_{n+1}$, if $\{X_0 = x_0, \dots, X_n = x_n\}$ has positive probability, then
$$\mathbb{P}[X_{n+1} = x_{n+1} \mid X_0 = x_0, \dots, X_n = x_n] = P(x_n, x_{n+1}).$$

Indeed: if (∗), then the case $n=0$ yields 1), and 2) is a ratio:
$$\mathbb{P}[X_{n+1} = x_{n+1} \mid X_0 = x_0, \dots, X_n = x_n] = \frac{\pi_0(x_0) P(x_0, x_1) \dots P(x_n, x_{n+1})}{\pi_0(x_0) P(x_0, x_1) \dots P(x_{n-1}, x_n)}$$
$$= P(x_n, x_{n+1}).$$

Conversely, if 1) and 2), one recovers (∗) since
$$\mathbb{P}[X_0 = x_0, \dots, X_n = x_n] = \mathbb{P}[X_0 = x_0]\, \mathbb{P}[X_1 = x_1 \mid X_0 = x_0] \dots \mathbb{P}[X_n = x_n \mid X_0 = x_0, \dots, X_{n-1} = x_{n-1}]$$

If $\mathcal{F}_n = \sigma(X_0, \dots, X_n)$, then the definition with conditional probabilities say that for a Markov chain with law $\mathbb{P}_{(\pi_0, P)}$,
$$\text{law of } X \text{ conditionally to } \mathcal{F}_n = P(X_n, \cdot).$$

(∗) is a much simpler definition: why bother with ratios ???

③ Set π_n = law of X_n. One has
$$\pi_n(x) = \sum_{x_0, x_1, \dots, x_{n-1}} \mathbb{P}[X_0 = x_0, \dots, X_{n-1} = x_{n-1}, X_n = x]$$

$$= \sum_{x_0, \dots, x_{n-1}} \pi_0(x_0) P(x_0, x_1) \dots \quad P(x_{n-1}, x)$$

$$= \pi_0 P^n(x). \qquad \text{So, } \pi_n = \underbrace{\pi_0 \cdot P^n}_{\text{matrix product}}.$$

Consider a bounded map $f : \mathcal{X} \to \mathbb{R}$. If we represent it by a column vector, then:

$$\mathbb{E}[f(X_{n+1}) \mid \mathcal{F}_n] = \sum_{y \in \mathcal{X}} P(X_n, y) f(y) = (Pf)(X_n).$$

$$\mathbb{E}[f(X_n)] = \sum_{x \in \mathcal{X}} \pi_n(x) f(x) = \pi_n \cdot f = \pi_0 \cdot P^n \cdot f.$$

Hence, many computations on Markov chains rely on matrices and basic linear algebra.

existence of Markov chains? construction?

intuitively:
- one picks X_0 at random according to π_0.
- if X_n is constructed, one chooses X_{n+1} by looking at the X_n-th row of P.

so this is a random walk on G_P.

Theorem (representation of Markov chains).

Consider
- a sequence of iid random variables $(\xi_n)_{n \geq 1}$ with values in $(E, \mathcal{E})$
- X_0 independent from $(\xi_n)_{n \geq 1}$, with law π_0 on $\mathcal{X}$.
- $F : \mathcal{X} \times E \to \mathcal{X}$ measurable.

We define $(X_n)_{n \in \mathbb{N}}$ by induction: $X_{n+1} = F(X_n, \xi_{n+1})$.

Then, $(X_n)_{n \in \mathbb{N}}$ is a Markov chain on $\mathcal{X}$ with transition matrix

$$P(x, y) = \mathbb{P}[f(x, \xi_\cdot) = y]$$

Conversely given π_0 and P, one can find F $(\Omega, \mathcal{F}, \mathbb{P})$, X_0 and $(\xi_n)_{n \geq 1}$ such that the construction above yields a MC with law $\mathbb{P}_{(\pi_0, P)}$.

Proof : For the first part, we compute

$$\mathbb{P}[X_0 = x_0, X_1 = x_1, \dots, X_n = x_n] = \mathbb{P}[X_0 = x_0, f(X_0, \xi_1) = x_1, \dots f(X_{n-1}, \xi_n) = x_n]$$

$$= \mathbb{P}[X_0 = x_0, f(x_0, \xi_1) = x_1, \dots, f(x_{n-1}, \xi_n) = x_n]$$

$$=_{(\text{independance})} \mathbb{P}[X_0 = x_0]\, \mathbb{P}[f(x_0, \xi_1) = x_1] \cdots \mathbb{P}[f(x_{n-1}, \xi_n) = x_n]$$

$$=_{(\text{same law for the } \xi_n\text{'s})} \pi_0(x_0) P(x_0, x_1) \dots P(x_{n-1}, x_n)$$

Conversely, we consider $(\xi_n)_{n \geq 0}$ a sequence of i.i.D uniform varibles in $[0,1)$

$(\Omega, \mathcal{F}, \mathbb{P}) = ([0,1), \mathcal{B}([0,1)), \text{Leb})^{\otimes \mathbb{N}}$.
ξ_n = the coordinate maps.

- construction of X_0 : we enumerate the elements of $\mathcal{X} = \{x_1, x_2, \dots\}$

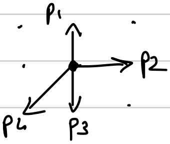
$X_0 = x_k$ where k is the unique integer such that $\sum_{j=1}^{k-1} \pi_0(x_j) \leq \xi_0 < \sum_{j=1}^{k} \pi_0(x_j)$

- map F: $F(x, \xi) = x_k$ where k is such that $\sum_{j=1}^{k-1} P(x, x_j) \leq \xi < \sum_{j=1}^{k} P(x, x_j)$.

By the direct sense, $(X_n)_{n \in \mathbb{N}}$ defined by induction by
$X_{n+1} = F(X_n, \xi_{n+1})$ is a MC.
Moreover, X_0 has law π_0, and $P(x,y) = \mathbb{P}[F(x, \xi) = y]$ $\forall x, y$. $\square$

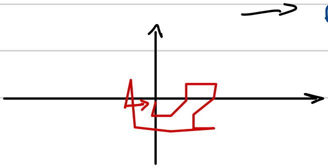
example : $\mathcal{X} = \mathbb{Z}^d$.

μ = some law on a finite subset of $\mathbb{Z}^d$



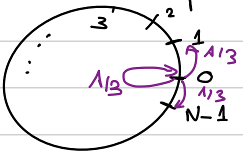
$X_n = \xi_1 + \xi_2 + \dots + \xi_n$ is a Markov chain of $\mathbb{Z}^d$.

$\longrightarrow$ random walk on $\mathbb{Z}^d$ with steps distributed as μ.



what do the trajectories look like?

An instructive example : the random walk on $\mathbb{Z}/N\mathbb{Z}$.



$$P(k,k) = P(k,k+1) = P(k,k-1) = \frac{1}{3}$$

By invariance by rotation, the starting point is not important, we take $X_0 = 0 \mod N$.

Construction : $(\xi_n)_{n\geq 1}$ iid with $\mathbb{P}[\xi_\cdot = 1] = \mathbb{P}[\xi_\cdot = -1] = \mathbb{P}[\xi_\cdot = 0] = \frac{1}{3}$

$X_n = \xi_1 + \xi_2 + \cdots + \xi_n \mod N$.

law of X_n? $\quad P = P_N = 1/3 \begin{pmatrix} 1 & 1 & & & 1 \\ 1 & 1 & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & 1 \\ 1 & & & 1 & 1 \end{pmatrix}_{N\times N}$

and $\pi_n = \underbrace{\delta_N}_{(0,\dots,0,1)} (P_N)^n$. We need to diagonalise the transition matrix.

$$P_N = \frac{1}{3}\left(I_N + C_N + C_N^{-1}\right), \quad C_N = \begin{pmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 1 & & & 0 \end{pmatrix}.$$

Given $v = (v(1), \dots v(N)) \in \mathbb{C}^N$, its discrete Fourier transform is $\hat{v}$ with

$$\hat{v}(k) = \frac{1}{\sqrt{N}} \sum_{j=1}^{N} v(j)\, e^{i\frac{jk2\pi}{N}}.$$

Lemma 1 $\quad v \to \hat{v}$ is invertible: $v(\ell) = \frac{1}{\sqrt{N}} \sum_{k=1}^{N} \hat{v}(k)\, e^{-\frac{ik\ell 2\pi}{N}}$.

Indeed, $\text{RHS} = \frac{1}{N} \sum_{j,k \in \mathbb{Z}/N\mathbb{Z}} v(j)\, e^{ik\frac{(j-\ell)2\pi}{N}}$

and with j fixed, $\sum_k \left(e^{\frac{i(j-\ell)2\pi}{N}}\right)^k$ is a geometric sum, $= \begin{cases} N & \text{if } j = \ell \\ \frac{1-z^N}{1-z} = 0 & \text{if } j \neq \ell \end{cases}$

So $RHS = \sum_j v(j)\, \mathbb{1}_{(j=\ell)} = v(\ell)$.

<u>Lemma 2</u> $\widehat{vC_N}(k) = e^{i\frac{k2\pi}{N}} \hat{v}(k)$.

Indeed, $vC_N(j) = v(j-1)$, so $\widehat{vC_N}(k) = \frac{1}{\sqrt{N}} \sum_j v(j-1) e^{i\frac{jk2\pi}{N}}$

$$= e^{i\frac{k2\pi}{N}} \frac{1}{\sqrt{N}} \sum_{j'\in\mathbb{Z}/N\mathbb{Z}} v(j') e^{i\frac{j'k2\pi}{N}}.$$

$$= e^{\frac{ik2\pi}{N}} \hat{v}(k).$$

Similarly, $\widehat{vC_N^{-1}}(k) = e^{-\frac{ik2\pi}{N}} \hat{v}(k)$, so

$$\widehat{vP_N}(k) = \frac{1 + 2\cos\frac{2k\pi}{N}}{3} \hat{v}(k) = \lambda_k \hat{v}(k)$$

We then compute: $\pi_n(\ell) = \frac{1}{\sqrt{N}} \sum_{k\in\mathbb{Z}/N\mathbb{Z}} \hat{\pi}_n(k)\, e^{-\frac{ik\ell 2\pi}{N}}$.

$$= \frac{1}{\sqrt{N}} \sum_k \hat{\pi}_0(k) (\lambda_k)^n e^{-\frac{ik\ell 2\pi}{N}}$$

And $\hat{\pi}_0(k) = \frac{1}{\sqrt{N}}$ $\forall k$. So: $\pi_n(\ell) = \frac{1}{N} \sum_{k\in\mathbb{Z}/N\mathbb{Z}} (\lambda_k)^n e^{-\frac{ik\ell 2\pi}{N}}$.

Notice that unless $k = 0 \bmod N$, $-\frac{1}{3} \leq \lambda_k < 1$.

So: $\forall \ell \in \mathbb{Z}/N\mathbb{Z}$, $\pi_n(\ell) \xrightarrow[n\to+\infty]{} \frac{1}{N}$. convergence to the uniform distribution.

This is a particular case of a general result...

2. Markov property.

In the sequel, P is fixed, and we abbreviate $\mathbb{P}_{(\pi_0, P)} = \mathbb{P}_{\pi_0}$. If $\pi_0 = \delta_{x_0}$, we also abbreviate $\mathbb{P}_{\delta_{x_0}} = \mathbb{P}_{x_0}$.

<u>Theorem</u> Suppose $(X_n)_{n\in\mathbb{N}}$ is a MC with law $\mathbb{P}_{\pi_0}$. Conditionally to $\mathcal{F}_{m\geq 1}$, $(X_{n+m})_{n\in\mathbb{N}}$ is a MC with law $\mathbb{P}_{X_m}$.

Set $Y_n = X_{n+m}$.

<u>Proof</u>: We have to prove that for any states $x_0, x_1, \dots, x_m, y_0, y_1, \dots, y_n$

$$\mathbb{P}_{\pi_0}[Y_0 = y_0, \dots, Y_n = y_n \mid X_0 = x_0, \dots, X_m = x_m] = \mathbb{1}_{y_0 = x_m} P(y_0, y_1) \cdots P(y_{n-1}, y_n)$$

assuming that the conditioning event has positive probability.

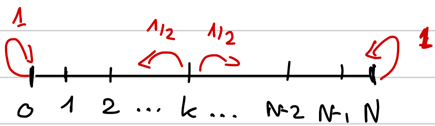
This is trivial: $p = \dfrac{\mathbb{P}[A \cap B]}{\mathbb{P}[B]}$ with:

$$\mathbb{P}[A \cap B] = \mathbb{1}_{y_0 = x_m} \pi_0(x_0) P(x_0, x_1) \cdots P(x_{m-1}, x_m) P(y_0, y_1) \cdots P(y_{n-1}, y_n)$$

$$\mathbb{P}[B] = \pi_0(x_0) P(x_0, x_1) \cdots P(x_{m-1}, x_m). \qquad \square$$

This seems quite abstract... it allows one to write recurrence relation for certain observables of Markov chains.

<u>example</u>: Probability of ruin



$T = \inf(\{n \in \mathbb{N} \mid X_n \in \{0, N\}\})$.

What is $f(k) = \mathbb{P}_k[T < \infty \text{ and } X_T = 0]$?

Notice that $f(0) = 1$ and $f(N) = 0$.

Suppose $0 < k < N$. T is a measurable map $\mathcal{X}^{\mathbb{N}} \to [0, +\infty]$.

$$f(k) = \mathbb{P}_k[T((X_n)_{n\in\mathbb{N}}) < \infty \text{ and } X_{T((X_n)_n)} = 0]$$

$$= \mathbb{P}_k[T((Y_n)_{n\in\mathbb{N}}) < \infty \text{ and } Y_{T((Y_n)_n)} = 0]$$

rewrite in terms of $Y_n = X_{n+1}$.

$$= \mathbb{P}_k[X_1 = k-1]\, \mathbb{P}_k[\dots (Y_n)_{n\in\mathbb{N}} \mid X_1 = k-1]$$

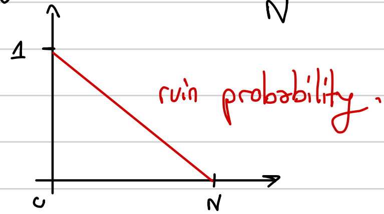
conditioning according to X_1 $\quad + \mathbb{P}_k[X_1 = k+1]\, \mathbb{P}_k[\dots (Y_n)_{n\in\mathbb{N}} \mid X_1 = k-1]$

$$= \frac{1}{2} \mathbb{P}_{k-1}[\dots (X_n)_{n\in\mathbb{N}}] + \frac{1}{2} \mathbb{P}_{k+1}[\dots (X_n)_{n\in\mathbb{N}}]$$

$$= \frac{1}{2} f(k-1) + \frac{1}{2} f(k+1).$$

The recurrence relation rewrites as $f(k+1)-f(k) = f(k)-f(k-1) \to$ constant slope.

So, $f(k) =$ affine function $= 1 - \frac{k}{N}$.



One can also compute $g(k) = \mathbb{E}_k[T]$

$g(0) = g(N) = 0$.

If $X_0 \notin \{0, N\}$, then $T((X_n)_{n\in\mathbb{N}}) = T((Y_n)_{n\in\mathbb{N}}) + 1$.

$$g(k) = \mathbb{E}_k[T((X_n)_{n\in\mathbb{N}})] = \mathbb{E}_k[\mathbb{E}_k[T((Y_n)_{n\in\mathbb{N}}) + 1 \mid \mathcal{F}_1]]$$

$$= \mathbb{E}_k[1 + \mathbb{E}_{X_1}[T((\tilde{X}_n)_{n\in\mathbb{N}})]] = \mathbb{E}_k[1 + g(X_1)]$$

$$= 1 + \frac{1}{2}(g(k-1) + g(k+1)).$$

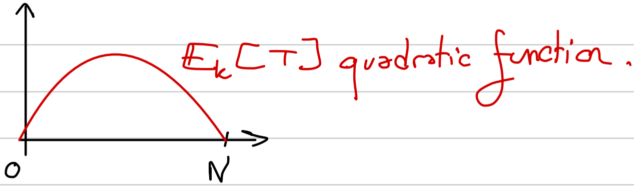
So, $\underbrace{g(k+1)-g(k)}_{\Delta g(k)} = g(k)-g(k-1) - 2$;

$$\Delta g(k) = \Delta g(0) - 2k$$

$$\sum_{k=0}^{N-1} \Delta g(k) = g(N) - g(0) = 0 = N\Delta g(0) - N(N-1)$$

$\Delta g(0) = N-1$, $\Delta g(k) = N-1-2k$.

We conclude that $g(k) = \sum_{j=0}^{k-1} \Delta g(j) = k(N-1) - k(k-1) = k(N-k)$.



Translating by a time $m \geq 1$ a Markov chain keeps the structure of Markov chains (weak or simple Markov property).

Is it also true if one translates by a random time T ?.

Theorem (Strong Markov property).

Consider a Markov chain $(X_n)_{n\in\mathbb{N}}$ with law $\mathbb{P}_{(\pi_0, P)}$.

a stopping time T w.r.t. $(\mathcal{F}_n = \sigma(X_0, \dots, X_n))_{n\in\mathbb{N}}$.

We define the associated σ-algebra :

$$\mathcal{F}_T = \left\{ A \in \mathcal{F} \mid \forall n \in \mathbb{N},\ A \cap \{T = n\} \in \mathcal{F}_n \right\}.$$

0) T is a $\mathcal{F}_T$ measurable random variable.

1) If $T < +\infty$ a.s., then the law of $(X_{n+T})_{n\in\mathbb{N}}$ conditionally to $\mathcal{F}_T$ is $\mathbb{P}_{(X_T, P)}$.

This means that for any bounded measurable function $f : \mathcal{X}^{\mathbb{N}} \to \mathbb{R}$ and any $\mathcal{F}_T$-measurable variable Z,

$$\mathbb{E}_{\pi_0}\left[f((X_{n+T})_{n\in\mathbb{N}}) Z \right] = \mathbb{E}_{\pi_0}\left[\mathbb{E}_{X_T}\left[f((\tilde{X}_n)_n) \right] Z \right].$$

2) For a stopping time T which might be equal to $+\infty$, this stays true by adding adequate indicator functions :

$$\mathbb{E}_{\pi_0}\left[\mathbb{1}_{T<+\infty} f((X_{n+T})_{n\in\mathbb{N}}) Z \right] = \mathbb{E}_{\pi_0}\left[\mathbb{1}_{T<+\infty} Z \, \mathbb{E}_{X_T}\left[f((\tilde{X}_n)_n) \right] \right].$$

Proof 0) Each event $\{T = m\}$ satisfies :

$$\{T = m\} \cap \{T = n\} = \begin{cases} \{T = n\} & \text{if } m = n \\ \emptyset & \text{if } m \neq n \end{cases} \in \mathcal{F}_n.$$

This is also obviously true for $\{T = +\infty\}$. So, T^{-1}(any subset of $\mathbb{N} \cup \{+\infty\}$) is in $\mathcal{F}_T$, and T is $\mathcal{F}_T$ measurable

1) and 2). Easy if one expands $\mathbb{1}_{(T<+\infty)} = \sum_{k=0}^{+\infty} \mathbb{1}_{(T=k)}$.

Note that if Z is $\mathcal{F}_T$-measurable, then for any k, $\mathbb{1}_{(T=k)} Z$ is $\mathcal{F}_k$ measurable.
Indeed, if $0 \notin B$, then

$$\{\mathbb{1}_{(T=k)} Z \in B\} = \{T=k\} \cap \underbrace{\{Z \in B\}}_{\in \mathcal{F}_T} \in \mathcal{F}_k.$$

and if $0 \in B$, then

$$\{\mathbb{1}_{(T=k)} Z \in B\} = \{T \neq k\} \cup \{T=k\} \cap \{Z \in B\}.$$

Then, by the simple Markov property,

$$\mathbb{E}_{\pi_0}[\mathbb{1}_{(T<+\infty)} Z f((X_{n+T})_n)]$$
$$= \sum_{k=0}^{\infty} \mathbb{E}_{\pi_0}[\mathbb{1}_{(T=k)} Z f((X_{n+T})_n)]$$
$$= \sum_{h=0}^{\infty} \mathbb{E}[\mathbb{1}_{(T=k)} Z \, \mathbb{E}_{X_k}[f((\tilde{X}_n)_n)]]$$
$$= \mathbb{E}[\mathbb{1}_{T<+\infty} Z \, \mathbb{E}_{X_T}[f((\tilde{X}_n)_n)]]. \qquad \square$$

<u>example</u>: Consider a sequence $(\xi_n)_{n\geq 1}$ of Bernoulli ± 1 iid variables with $\mathbb{P}[\xi_n = 1] = 1 - \mathbb{P}[\xi_n = -1] = p$, $p \in (0,1)$.

By the representation theorem, $X_n = \xi_1 + \dots + \xi_n$ is a Markov chain on $\mathbb{Z}$.
It's transition matrix is:

$$P(k, k+1) = p; \quad P(k, k-1) = 1-p.$$

For $k \geq 1$, set $T_k = \inf(\{n \in \mathbb{N} \mid X_n = k\})$.

<u>Lemma</u>: T_k has the law of a sum of k independent random variables distributed as T_1.

$$\forall \ell \in \mathbb{N}, \quad \mathbb{P}[T_k = \ell] = \sum_{\ell_1 + \dots + \ell_k = \ell} \mathbb{P}[T_1 = \ell_1] \dots \mathbb{P}[T_k = \ell_k].$$

Indeed, if $k=2$, then:

$$\mathbb{P}_0[T_k = \ell] = \mathbb{E}_0[\mathbb{1}_{T_k((X_n)_n) = \ell}]$$

(if T_k is finite, then T_1 is finite, smaller than T_k)

$$= \mathbb{E}_0[\mathbb{1}_{T_k((X_n)_n) = \ell} \mathbb{1}_{T_1((X_n)_n) \leq \ell}]$$

$$= \mathbb{E}_0\left[\mathbb{E}_0[\mathbb{1}_{T_k((X_{n+T_1})_n) + T_1((X_n)_n) = \ell} | \mathcal{F}_{T_1}] \mathbb{1}_{T_1((X_n)_n) \leq \ell}\right]$$

$$= \sum_{\ell_1 = 1}^{\ell} \mathbb{E}_0\left[\mathbb{E}_0[\mathbb{1}_{T_k((X_{n+T_1})_n) + \ell_1 = \ell} | \mathcal{F}_1] \mathbb{1}_{T_1((X_n)_n) = \ell_1}\right]$$

$$\underset{\text{(strong Markov)}}{=} \sum_{\ell_1 = 1}^{\ell} \mathbb{E}_0\left[\mathbb{E}_1[\mathbb{1}_{T_k((\tilde{X}_n)) = \ell - \ell_1}] \mathbb{1}_{T_1((X_n)_n) = \ell_1}\right]$$

$$= \sum_{\ell_1 = 1}^{\ell} \mathbb{P}_1[T_k = \ell - \ell_1] \mathbb{P}_0[T_1 = \ell_1]$$

$$\underset{\text{(translation invariance)}}{=} \sum_{\ell_1 = 1}^{\ell} \mathbb{P}_0[T_{k-1} = \ell - \ell_1] \mathbb{P}_0[T_1 = \ell_1]$$

and then use induction on ℓ.

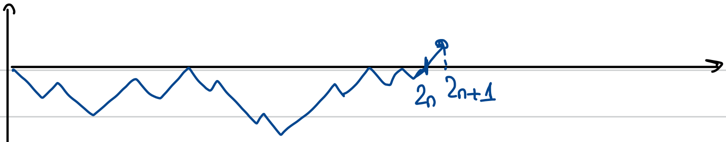
As a consequence, $\mathbb{P}_0[T_k < +\infty] = (\mathbb{P}_0[T_1 < +\infty])^k \quad \forall k \geq 1$.

What is the law of T_1? the value of $\mathbb{P}_0[T_1 < +\infty]$?

This can be solved by combinatorics:

- T_1 can only equal to an odd integer (or $+\infty$):
  $\forall n, \ X_n \equiv n \ [2]$ a.s.
  So $\mathbb{P}_0[T_1 = 2n] \leq \mathbb{P}[X_{2n} = 1] = 0 \quad \forall n \in \mathbb{N}$.

- $$\mathbb{P}_0[T_1 = 2n+1] = \mathbb{P}_0[X_1, \dots, X_{2n} \leq 0 \text{ and } X_{2n+1} = 1]$$
  $$= \mathbb{P}_0[\underbrace{X_1, \dots, X_{2n-1} \leq 0, X_{2n} = 0}_{\text{event in } \mathcal{F}_n}, X_{2n+1} = 1]$$
  $$= \mathbb{P}_0[X_1, \dots, X_{2n-1} \leq 0, X_{2n} = 0] \underbrace{P(0,1)}_{=p}$$



Suppose given some sequence $x_1, \dots, x_{2n}$ of negative integers, with $x_{2n} = 0$, $|x_i - x_{i-1}| = 1$ $(x_0 = 0)$

$\longrightarrow$ negative excursion.

Since $x_{2n} - x_0 = \sum_{i=1}^{2n} x_i - x_{i-1} = 0$, there are n positive steps $x_i - x_{i-1} = 1$

n negative steps $x_i - x_{i-1} = -1$

So, $\mathbb{P}_0[X_1 = x_1, \dots, X_{2n} = x_{2n}] = \mathbb{P}_0[\xi_1 = x_1 - x_0, \xi_2 = x_2 - x_1, \dots, \xi_{2n} = x_{2n} - x_{2n-1}]$

$$= \prod_{i=1}^{2n} \mathbb{P}[\xi_i = x_i - x_{i-1}]$$

$$= p^n (1-p)^n.$$

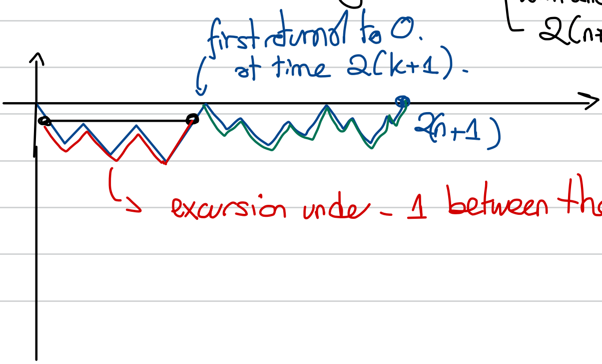
$$\Rightarrow \mathbb{P}_0[T_1 = 2n+1] = p^{n+1}(1-p)^n \; \# \underbrace{\left\{ \begin{array}{l} \text{number of negative excursions} \\ \text{with length } 2n \end{array} \right\}}_{C_n.}$$

C_n is the n-th Catalan number. $C_0 = 1$, $C_1 = 1$, $C_2 = 2$, $C_3 = 5$, ...

The C_n satisfy:

$$C_{n+1} = \sum_{k=0}^{n} C_k C_{n-k}.$$

Indeed, there is a bijection $\left\{ \begin{array}{l} \text{excursions} \\ \text{with length} \\ 2(n+1) \end{array} \right\} \longrightarrow \bigsqcup_{k=0}^{n} \left\{ \begin{array}{l} \text{exc.} \\ \text{with length} \\ 2k \end{array} \right\} \times \left\{ \begin{array}{l} \text{excursions} \\ \text{with length} \\ 2(n-k) \end{array} \right\}$



$\hookrightarrow$ excursion under -1 between the times 1 and $2k+1$.

Set $C(z) = \sum_{n=0}^{\infty} C_n z^n$. Since $C_n \leq 2^{2n}$, $C(z)$ is a power series with radius of converge $R \geq \frac{1}{4}$.

$$z\,(C(z))^2 = \sum_{n=0}^{\infty} \left(\sum_{k=0}^{n} C_k C_{n-k}\right) z^{n+1}$$
$$= \sum_{n=0}^{\infty} C_{n+1} z^{n+1} = C(z) - 1.$$

So, $z\,(C(z))^2 - C(z) + 1 = 0$.

$$C(z) = \frac{1 \pm \sqrt{1-4z}}{2z}.$$ By analyticity, $\forall |z| < 1$, $C(z)$ is one of the two functions.

But the $+$ solution has a singularity at $z = 0$

$$\Rightarrow C(z) = \frac{1-\sqrt{1-4z}}{2z}.$$

From there one can recover $C_n = \frac{1}{n+1}\binom{2n}{n}$.

$$\mathbb{P}[T_1 < +\infty] = \sum_{n=0}^{+\infty} C_n\, p^{n+1}(1-p)^n = p \cdot C(p(1-p))$$
$$= p\,\frac{1-|1-2p|}{2p(1-p)}$$
$$= \frac{1-|1-2p|}{2(1-p)} = \begin{cases} 1 & \text{if } p \geq \frac{1}{2} \\ \frac{p}{1-p} & \text{if } p < \frac{1}{2} \end{cases}$$

If $p \geq \frac{1}{2}$, then $T_k < +\infty$ a.s. $\forall k \geq 1$, so $\sup_{n\in\mathbb{N}} X_n = +\infty$ a.s.

If $p < \frac{1}{2}$, then $\mathbb{P}_0[\sup_{n\in\mathbb{N}} X_n \geq k] = \mathbb{P}_0[T_k < +\infty]$
$$= (\mathbb{P}_0[T_1 < +\infty])^k = \left(\frac{p}{1-p}\right)^k.$$

So, in this case, $\sup_{n\in\mathbb{N}} X_n$ follows a geometric law with parameter $1 - \frac{p}{1-p} = \frac{1-2p}{1-p}$.

3. Recurrent and transient states.

P fixed stochastic matrix on a state space $\mathcal{X}$.

V_x = number of visits of x = $|\{n \geq 1 \mid X_n = x\}|$

for $(X_n)_{n\in\mathbb{N}}$ Markov chain with law $\mathbb{P}_{(\pi_0, P)}$

<u>Theorem</u> Under $\mathbb{P}_x$, either $V_x = +\infty$ a.s., or V_x is finite a.s. and follows a geometric distribution.

One says that x is recurrent in the first case and transient in the second case.

<u>Proof</u>: Set $T_x^+ = \inf(\{n \geq 1 \mid X_n = x\})$ = first return time to x.

If $k \geq 1$, then:

$$\mathbb{P}_x[V_x \geq k] = \mathbb{P}_x[T_x^+((X_n)_n) < +\infty \text{ and } V_x((X_n)_n) \geq k]$$

$$= \mathbb{P}_x[T_x^+((X_n)_n) < +\infty \text{ and } V_x((X_{n+T_x^+})_n) \geq k-1]$$

$$= (\text{Strong Markov w.r.t. } \mathcal{F}_{T_x^+})$$

$$\mathbb{P}_x[T_x^+((X_n)_n) < +\infty]\ \mathbb{P}_x[V_x \geq k-1]$$

If $\mathbb{P}_x[T_x^+ < +\infty] = 1$:

then $V_x \geq k$ a.s. $\forall k$ by induction, so $V_x = +\infty$ a.s.

Otherwise: set $\mathbb{P}_x[T_x^+ < +\infty] = 1 - p_x$, $p_x > 0$.

$\mathbb{P}_x[V_x \geq k] = (1-p_x)^k$.

$V_x \sim \mathrm{Geom}(p_x)$ on $\mathbb{N}$. $\square$.

So, we can partition $\mathcal{X} = R \sqcup T$

<u>numerical criterion</u>

x recurrent $\iff \sum_{n=1}^{\infty} P^n(x,x) = +\infty$.

x transient $\iff \sum_{n=1}^{\infty} P^n(x,x) < +\infty$.

Proof: Recall that if $V \sim \mathrm{Geom}(p)$ on $\mathbb{N}$, then $\mathbb{E}[V] = \frac{1}{p} - 1 < +\infty$

However, $\sum_{n=1}^{\infty} P^n(x,x) = \sum_{n=1}^{\infty} \mathbb{P}_x[X_n = x]$

$$= \sum_{n=1}^{\infty} \mathbb{E}_x[\mathbb{1}_{(X_n=x)}]$$

$$= \mathbb{E}_x\Big[\sum_{n=1}^{\infty} \mathbb{1}_{X_n=x}\Big] = \mathbb{E}_x[V_x] \quad \square.$$

example random walk on $\mathbb{Z}$ with parameter $p \in (0,1)$.

$P(k, k+1) = p$; $P(k, k-1) = 1-p$.

$P^n(x,x)$ does not depend on x, so the states are either all recurrent or all transient.

$P^n(0,0) = ?$ This is 0 if n is odd.

Otherwise, $P^{2n}(0,0) = \sum_{\substack{\text{admissible path} \\ 0 \to x_1 \to x_2 \to \cdots \to x_{2n-1} \to 0}} p^{\#\text{positive steps}} (1-p)^{\#\text{negative steps}}$

An admissible path has n positive and n negative steps.
then choosing an admissible path amounts to choose where to put the n positive steps among the $2n$ steps $\to \binom{2n}{n}$ possibilities.

So $P^{2n}(0,0) = \binom{2n}{n}(p(1-p))^n$.

- If $p \neq \frac{1}{2}$, then $p(1-p) < \frac{1}{4}$, whereas $\binom{2n}{n} \leq 2^{2n} = 4^n$.
So the series is convergent $\Rightarrow$ all states are transient.

- If $p = \frac{1}{2}$, then $\binom{2n}{n} \sim_{\text{Stirling}} \frac{(2n)^{2n}\sqrt{2\pi 2n}}{n^{2n} \; 2\pi n} = \frac{4^n}{\sqrt{\pi n}}$.

So the general term of the series $\sum P^{2n}(0,0)$ is $O\left(\frac{1}{\sqrt{n}}\right)$
$\Rightarrow$ divergent series $\Rightarrow$ all states are recurrent.

4. Classification of states

$\mathcal{X}$, P fixed.

Definition One says that x communicates with y if $\exists n \geq 1 \mid P^n(x,y) > 0$
$\iff$ there is a path $x \to x_1 \to \dots \to x_{n-1} \to y$ in G_P

notation: $x \rightsquigarrow y$. This is transitive: if $x \rightsquigarrow y$ and $y \rightsquigarrow z$, then $x \rightsquigarrow z$.

Theorem If x is recurrent and $x \rightsquigarrow y$, then y is recurrent and $y \rightsquigarrow x$.

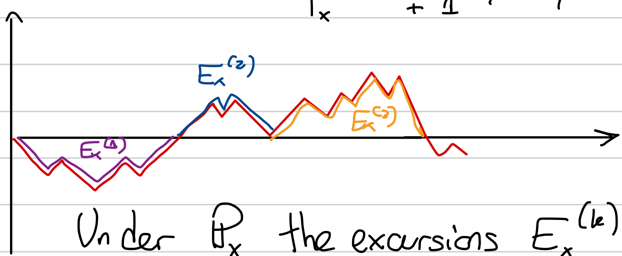
Lemma Given x recurrent, consider the sequence of return times:

$$\left.\begin{aligned} T_x^{(1)} &= \inf\{n \geq 1 \mid X_n = x\} \\ T_x^{(k+1)} &= \inf\{n \geq 1 + T_x^{(k)} \mid X_n = x\} \end{aligned}\right) \text{ finite a.s. under } \mathbb{P}_x$$

The k-th excursion from x to x is:

$$E_x^{(k)} = \left(X_{T_x^{(k-1)}+1}, \dots, X_{T_x^{(k)}}\right) \in \bigsqcup_{n \geq 1} \mathcal{X}^n.$$

$(T_x^{(0)} = 0)$



Under $\mathbb{P}_x$, the excursions $E_x^{(k)}$ are iiD variables in $\bigsqcup_{n \geq 1} \mathcal{X}^n$

Indeed, note that $T_x^{(k+1)}((X_n)_n) = T_x^{(k)}((X_{n+T_x^{(1)}})_n) + T_x^{(1)}$

$$E_x^{(k+1)}((X_n)_n) = E_x^{(k)}((X_{n+T_x^{(1)}})_n).$$

Fix sets $B_1, \dots, B_k \subset \bigsqcup_{n \geq 1} \mathcal{X}^n$

$$\mathbb{P}_x[E_x^{(1)}((X_n)_n) \in B_1 \dots E_x^{(k)}((X_n)_n) \in B_k]$$

$$= \mathbb{E}_x\Big[\mathbb{1}_{E_x^{(1)}((X_n)_n) \in B_1} \mathbb{E}_x\Big[\mathbb{1}_{E_x^{(1)}((X_{n+T_1^{(x)}})_n) \in B_2} \dots \mathbb{1}_{E_x^{(k-1)}((X_{n+T_1^{(x)}})_n) \in B_k} \,\Big|\, \mathcal{F}_{T_1^{(x)}}\Big]\Big]$$

$= $ (Strong Markov property + rewriting)

$$\mathbb{P}_x\left[E_x^{(1)} \in B_1\right] \mathbb{P}_x\left[E_x^{(1)} \in B_2 \dots E_x^{(k-1)} \in B_k\right]$$

$= $ (induction) $\displaystyle\prod_{j=1}^{k} \mathbb{P}_x\left[E_x^{(1)} \in B_j\right]$ $\square$.

<u>Proof of the theorem</u> Consider the events:

$$A_k = \{y \in E_x^{(k)}\}, \quad k \geq 1.$$

If $x \to x_1 \to \dots \to x_{n-1} \to y$ is a possible path with minimal length, then

$$\mathbb{P}_x[A_1] \geq \mathbb{P}_x[X_1 = x_1, \dots, X_{n-1} = x_{n-1}, X_n = y]$$
$$\geq P(x, x_1) P(x_1, x_2) \dots P(x_{n-1}, y) = p > 0.$$

By the lemma, the A_k are independent, all with the same positive probability.

$\implies$ an infinity of A_k happen $\mathbb{P}_x$-a.s.

$\implies$ $V_y = +\infty$ $\mathbb{P}_x$-a.s. (in particular, $T_y < +\infty$ $\mathbb{P}_x$-a.s.).

Then, $y \rightsquigarrow x$: otherwise, we would have:

$$\mathbb{P}_x[V_x = +\infty] = \mathbb{E}_x\left[\mathbb{1}_{V_x((X_n)_n) = +\infty}\right]$$
$$= \mathbb{E}_x\left[\mathbb{E}_x\left[\mathbb{1}_{V_x((X_{n+T_y})_n) = +\infty} \mid \mathcal{F}_{T_y}\right]\right]$$
$$= \mathbb{E}_y\left[\mathbb{1}_{V_x((\tilde{X}_n)_n) = +\infty}\right] = \mathbb{P}_y[V_x = +\infty] = 0.$$

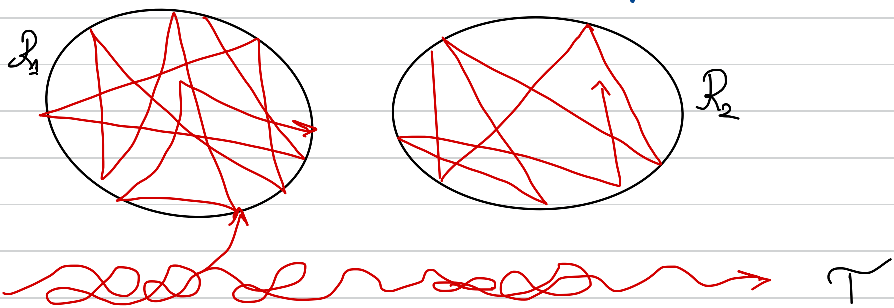
Also, $\underset{+\infty}{\underset{\parallel}{\mathbb{E}_x[V_y]}} = \mathbb{E}_x\left[\mathbb{E}_x\left[1 + V_y((X_{n+T_y})_n) \mid \mathcal{F}_{T_y}\right]\right]$

$$= 1 + \mathbb{E}_y[V_y]$$

so $\mathbb{E}_y[V_y] = +\infty$ and y is recurrent $\square$.

Consequently, the restriction of $\rightsquigarrow$ to the set $\mathcal{R}$ of recurrent states is an equivalence relation.

We split $\mathcal{X} = \mathcal{T} \sqcup \bigsqcup_{i \in I} \mathcal{R}_i$ ↪ recurrence classes = equivalence classes of recurrent states



qualitative aspect of the trajectories under $\mathbb{P}_x$:

- if $x \in \mathcal{R}_i$ then the set of visited states is exactly $\mathcal{R}_i$, and all the y's in $\mathcal{R}_i$ are visited an infinite number of times.
- if $x \in \mathcal{T}$: either $(X_n)_{n\in\mathbb{N}}$ stays in $\mathcal{T}$ and visits these transient states each a finite number of times
  or, at some time T, X_n hits a recurrence class $\mathcal{R}_i$, which is then explored infinitely.

Particular cases

1) One says that $(\mathcal{X}, P)$ is irreducible if $x \rightsquigarrow y$ $\forall x,y \in \mathcal{X}$.
   Then, either all states are transient, or all states are recurrent (in the same class).

2) Suppose that $\mathcal{X}$ is finite. Then, there is at least one recurrent state.

example Consider a finite graph $G=(V,E)$ which is connected. A natural stochastic matrix on V is:

$$P_G(x,y) = \frac{1}{\deg x} \mathbb{1}_{(x\sim y)}.$$

Then, the corresponding Markov chain is necessarily recurrent, and all vertices are a.s. visited infinitely often.

Set $V_{x,n} = \text{card}\{k \mid 1\le k\le n,\ X_k = x\}$.

Can one estimate the frequence of visits $\frac{V_{x,n}}{n}$?

5. Invariant measures.

$\mathcal{X}$, P fixed.

Definition An invariant measure for P is a positive measure on $\mathcal{X}$, not necessarily with finite mass, such that $\pi P = \pi$.

Thus, $\forall x\in\mathcal{X}$, $\pi(x) = \sum_{v\in\mathcal{X}} \pi(v) P(v,x)$.

example random walk with parameter p on $\mathbb{Z}$.

To find the invariant measures, we have to solve the degree 2 linear recurrence:

$$\pi(k) = \pi(k-1)\,p + \pi(k+1)(1-p).$$

$$X^2(1-p) - X + p = 0 \ ; \quad \Delta = 1-4p(1-p) = (1-2p)^2$$

$$x = \frac{1\pm(1-2p)}{2(1-p)} = 1 \text{ or } \frac{p}{1-p}.$$

1st case: $p=\frac{1}{2}$. The formal solutions are $\pi(k) = a+bk$. To have a solution which is non negative, we need $b=0$ and $a\ge 0$.

Conclusion: the only invariant measures are the multiple of the counting measure $\nu = \sum_{k\in\mathbb{Z}} \delta_k$.

2nd case : $p \neq \frac{1}{2}$. Then, the solutions are of the form $\pi(k) = a + b\left(\frac{p}{1-p}\right)^k$

and $a, b \geq 0$ gives non-negative solutions $\Rightarrow$ a two-dimensional cone of invariant measures.

example random walk on a connected finite graph $G = (V, E)$

$$P_G(x,y) = \frac{1}{\deg x} \mathbb{1}_{(x,y)\in E}.$$

Lemma Suppose that π is reversible with respect to P:

$$\forall x, y, \quad \pi(x) P(x,y) = \pi(y) P(y,x).$$

Then, π is invariant for P.

Proof: Taking the sum over y of the identity gives

$$\pi(x) = (\pi P)(x) \qquad \forall x \in \mathcal{X} \qquad \square.$$

⚠ The inverse implication is false. Consider e.g. $\mathcal{X} = [\![1,3]\!]$,

$P = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$ and $\pi = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$.

In our example, the equation of reversibility yields:

$\forall x \leftrightarrow y$ in G, $\dfrac{\pi(x)}{\deg x} = \dfrac{\pi(y)}{\deg y}$. So by connexity, $\pi(x) = c \cdot \deg x$ $\forall x \in V$.

One can choose the constant c to have an invariant probability measure, also called stationary law.

Take $c = \sum_{y\in V} \deg y = 2 \operatorname{card} E$. $\pi(x) = \dfrac{1}{2|E|} \deg x$.

Proposition If $(\mathcal{X}, P)$ is irreducible and π is invariant and nonzero, then $\pi(x) > 0$ $\forall x \in \mathcal{X}$.

Proof: Fix x such that $\pi(x) > 0$, and y another state. If $\pi P = \pi$, then $\pi P^n = \pi$ $\forall n \geq 0$.

There exists n such that $P^n(x,y) > 0$, and then

$$\pi(y) = (\pi P^n)(y) = \sum_{v \in \mathcal{X}} \pi(v) P^n(v,y) \geq \pi(x) P^n(x,y) > 0 \qquad \square$$

Theorem Consider an irreducible recurrent Markov chain $(\mathcal{X}, P)$. There exists a nonzero invariant measure π, and it is unique up to multiplication by a positive real number.

Proof: • existence.

We fix $x_0 \in \mathcal{X}$; by assumption, under $\mathbb{P}_{x_0}$, $T_{x_0}^+ < +\infty$ a.s.

We set $\pi_{x_0}(y) = \pi(y) = \mathbb{E}_{x_0}\left[\sum_{k=1}^{T_{x_0}^+} \mathbb{1}(X_k = y)\right]$

$=$ expected number of visits of y during the excursion $\mathcal{E}_{x_0}^{(1)}$.

⋆ cyclic lemma: a.s. under $\mathbb{P}_{x_0}$, $\sum_{k=1}^{T_{x_0}^+} \mathbb{1}(X_k = y) = \sum_{k=0}^{T_{x_0}^+ - 1} \mathbb{1}(X_k = y)$

⋆ finiteness: $\pi(y) = \mathbb{E}_{x_0}[V_{y, T_{x_0}^+}]$

$$= \mathbb{E}_{x_0}[V_{y, T_{x_0}^+}((X_n)_n)\, \mathbb{1}_{T_y < T_{x_0}^+}]$$

Strong Markov property $\rightarrow$ $= \mathbb{E}_{x_0}\left[\mathbb{E}_y\left[1 + \text{number of visits of } y \text{ before hitting } x_0\right] \mathbb{1}_{T_y < T_{x_0}^+}\right]$

$$= \mathbb{P}_{x_0}[T_y < T_{x_0}^+]\ \mathbb{E}_y[1 + \text{number of visits of } y \text{ before hitting } x_0].$$

Since excursions from y are iid, the expectation is the one of a geometric variable with parameter $\mathbb{P}_y[T_{x_0} < T_y^+]$.

So, $\pi(y) = \dfrac{\mathbb{P}_{x_0}[T_y < T_{x_0}^+]}{\mathbb{P}_y[T_{x_0} < T_y^+]} < +\infty$.

★ invariance : let us compute $(\pi P)(y)$:

$$\pi P(y) = \sum_{v \in \mathcal{X}} \pi(v) P(v,y) = \sum_{v \in \mathcal{X}} \mathbb{E}_{x_0}\Big[\sum_{k=0}^{T_{x_0}^+ - 1} \mathbb{1}_{X_k = v}\Big] P(v,y)$$

$$= \sum_{v \in \mathcal{X}} \sum_{k=0}^{+\infty} \mathbb{P}_{x_0}[\underbrace{k < T_{x_0}^+, X_k = v}_{\in \mathcal{F}_k}]\, P(v,y).$$

$$= \sum_{v \in \mathcal{X}} \sum_{k=0}^{+\infty} \mathbb{P}_{x_0}[k+1 \leq T_{x_0}^+, X_k = v, X_{k+1} = y]$$

$$= \sum_{k=0}^{+\infty} \mathbb{P}_{x_0}[k+1 \leq T_{x_0}^+, X_{k+1} = y] = \sum_{\ell=1}^{\infty} \mathbb{P}_{x_0}[\ell \leq T_{x_0}^+, X_\ell = y]$$

$$= \mathbb{E}_{x_0}\Big[\sum_{\ell=1}^{T_{x_0}^+} \mathbb{1}_{X_\ell = y}\Big] = \pi(y).$$

- unicity. Consider another invariant measure υ; up to multiplication by a scalar, one can assume $\upsilon(x_0) = 1 = \pi_{x_0}(x_0)$.

Take then $y \neq x_0$. We have :

$$\upsilon(y) = (\upsilon P)(y) = \sum_{x_1 \neq x_0} \upsilon(x_1) P(x_1, y) + \underbrace{P(x_0, y)}_{\mathbb{E}_{x_0}[\mathbb{1}_{1 \leq T_{x_0}^+, X_1 = y}]}.$$

$$= \sum_{x_2, x_1 \neq x_0} \upsilon(x_2) P(x_2, x_1) P(x_1, y) + \sum_{x_1 \neq x_0} P(x_0, x_1) P(x_1, y) + \mathbb{E}_{x_0}[\mathbb{1}_{1 \leq T_{x_0}^+, X_1 = y}]$$

$$= \sum_{x_2, x_1 \neq x_0} \upsilon(x_2) P(x_2, x_1) P(x_1, y) + \sum_{k=1}^{2} \mathbb{E}_{x_0}[\mathbb{1}_{k \leq T_{x_0}^+, X_k = y}]$$

By induction on n :

$$\upsilon(y) = \sum_{k=1}^{n} \mathbb{E}_{x_0}[\mathbb{1}_{k \leq T_{x_0}^+, X_k = y}] + \text{remainder}.$$

So $\upsilon(y) \geq \sum_{k=1}^{\infty} \mathbb{E}_{x_0}[\mathbb{1}_{k \leq T_{x_0}^+, X_k = y}] = \pi_{x_0}(y).$

Then, $\nu - \pi \geq 0$ is invariant and vanishes at x_0, so $\nu - \pi = 0$ $\square$.

<u>remark</u>: Therefore, if $p \neq \frac{1}{2}$, the random walk with parameter p on $\mathbb{Z}$ is transient, because it is irreducible with a 2 dimensional cone of invariant measures.

The theorem says nothing about the total mass of an invariant measure for an irreducible recurrent Markov chain. Note however that:

$$\pi_{x_0}(\mathcal{X}) = \sum_{x \in \mathcal{X}} \mathbb{E}_{x_0}\left[\sum_{k=1}^{T_{x_0}^+} \mathbb{1}_{X_k = x}\right] = \mathbb{E}_{x_0}[T_{x_0}^+].$$

<u>Theorem</u>: In this setting, there are two possible situations:

1) all the invariant measures have infinite total mass $\pi(\mathcal{X}) = +\infty$

$\iff \forall x \in \mathcal{X}, \ \mathbb{E}_x[T_x^+] = +\infty$ (although $T_x^+ < +\infty$ a.s. under $\mathbb{P}_x$).

null recurrence.

2) all the invariant measures have finite total mass $\pi(\mathcal{X}) < +\infty$.

$\iff \forall x \in \mathcal{X}, \ \mathbb{E}_x[T_x^+] < +\infty$.

positive recurrence.

In the second case, there is a unique invariant probability measure π. Moreover,

$$\pi(x) = \frac{\pi_{x_0}(x)}{\mathbb{E}_{x_0}[T_{x_0}^+]}, \qquad \forall x \in \mathcal{X}.$$

Taking $x = x_0$ yields: $\pi(x_0) = \dfrac{1}{\mathbb{E}_{x_0}[T_{x_0}^+]}$.

Actually, for an irreducible transition matrix:

P is positive recurrent $\iff$ P admits a finite invariant measure (criterion of recurrence!).

Indeed, $\Rightarrow$ is the definition of positive recurrence.

Conversely, suppose that $\pi P = \pi$ with $\pi(\mathcal{X}) < +\infty$. With x_0, we can assume without loss of generality that $\pi(x_0) = 1$.

Then, the previous proof shows without knowing beforehand the recurrence that:

$$\pi(x) \geq \mathbb{E}_{x_0}[\text{number of visits of } x \text{ before } T_{x_0}^+].$$

$\Rightarrow +\infty > \pi(\mathcal{X}) \geq \mathbb{E}_{x_0}[T_{x_0}^+]$ so $T_{x_0}^+ < +\infty$ a.s. under $\mathbb{P}_{x_0}$ $\square$

<u>examples</u>: 1) the symmetric random walk with $p = 1/2$ on $\mathbb{Z}$ is null recurrent ($\pi = \sum_{k \in \mathbb{Z}} \delta_k$ infinite invariant measure)

2) any irreducible Markov chain on a finite state space is positive recurrent.

6. Ergodic theorem

Consider an irreducible Markov chain $(\mathcal{X}, P)$, and the empirical measures $\nu_n = \frac{1}{n} \sum_{k=1}^{n} \delta_{X_k}$ of $(X_n)_{n \in \mathbb{N}} \sim \mathbb{P}_{(\pi_0, P)}$

π_0 arbitrary initial measure.

<u>Theorem</u>. • If the Markov chain is transient or null recurrent, then
$$\nu_n(x) \xrightarrow[\text{a.s.}]{} 0 \quad \forall x \in \mathcal{X}.$$

• If the Markov chain is positive recurrent with stationary law π, then
$$\nu_n(x) \xrightarrow[\text{a.s.}]{} \pi(x) \quad \forall x \in \mathcal{X}.$$

(asymptotics of visit frequencies).

<u>Proof</u>: The transient case is trivial since $\{k \in \mathbb{N} \mid X_k = x\}$ is finite a.s.

Suppose in the sequel that the chain is recurrent.

By decomposition of the initial measure $\pi_0 = \sum_{y \in \mathcal{X}} \pi_0(y)\, \delta_y$, it suffices to prove the result when $\pi_0 = \delta_y$.

Recall that the excursions $E_y^{(k)}$, $k \geq 1$ are iid, with expected length $\mathbb{E}[\ell(E_y^{(1)})] = \mathbb{E}[T_y^{(1)}]$

Denote $V_{y,x}^{(k)} = \sum_{n=T_y^{(k-1)}+1}^{T_y^{(k)}} \mathbb{1}_{(X_n = x)}$ the number of visits of x along the kth excursion $E_y^{(k)}$.

- The random variables $V_{y,x}^{(k)}$ are i.i.d. under $\mathbb{P}_y$.
- They are integrable with mean $\Pi_y(x)$.
- If the chain is positive recurrent, then $\Pi_y(.) = \frac{\pi(.)}{\pi(y)}$ with π unique invariant probability measure.

Suppose the chain positive recurrent. Then, for $T_y^{(k)} \leq n < T_y^{(k+1)}$,

$$\frac{k}{T_y^{(k+1)}} \frac{V_{y,x}^{(1)} + \cdots + V_{y,x}^{(k)}}{k} \leq \upsilon_n(x) \leq \frac{V_{y,x}^{(1)} + \cdots + V_{y,x}^{(k+1)}}{k+1} \frac{k+1}{T_y^{(k)}}$$

When n goes to infinity, k also goes to infinity (because the times $T_y^{(k)} - T_y^{(k-1)}$ are finite).
By the law of large numbers, both sides converge almost surely towards:

$$\frac{1}{\mathbb{E}[T_y^{(1)}]} \mathbb{E}[V_{y,x}^{(1)}] = \pi(y) \frac{\pi(x)}{\pi(y)} = \pi(x).$$

Suppose now the chain null recurrent. The same inequality holds, this time with $\frac{V_{y,x}^{(1)} + \cdots + V_{y,x}^{(k)}}{k} \xrightarrow[a.s.]{} \Pi_y(x) < +\infty$,

and $\frac{T_y^{(k)}}{k} \xrightarrow[a.s.]{} +\infty$ (generalisation of the law of large numbers). So, $\upsilon_n(x) \xrightarrow[a.s.]{} 0$ in this case.

<u>Lemma</u>: Consider i.i.d. nonnegative random variables $X_1, \ldots, X_n, \ldots$ such that $\mathbb{E}[X_1] = \int_\Omega X_1(\omega)\, \mathbb{P}[d\omega] = +\infty$
Then, $\frac{X_1 + \cdots + X_n}{n} \xrightarrow[a.s.]{} +\infty$.

Indeed, by the monotone convergence theorem, $\lim_{M \to +\infty} \mathbb{E}[X_1 \mathbb{1}_{X_1 \leq M}] = +\infty$. Fix $L > 0$ and M such that $\mathbb{E}[X_1 \mathbb{1}_{X_1 \leq M}] \geq L$. The variables $Y_i = X_i \mathbb{1}_{X_i \leq M}$ are i.i.d. integrable, so by the law of large numbers, $\liminf_{n \to +\infty} \frac{X_1 + \cdots + X_n}{n} \geq \lim_{n \to +\infty} \frac{Y_1 + \cdots + Y_n}{n} = \mathbb{E}[Y_1] \geq L$. Now make L go to $+\infty$. $\square$.

Fix $f : \mathcal{X} \to \mathbb{R}$ bounded function. If the chain is irreducible positive recurrent, then

$$\underbrace{\frac{1}{n}\sum_{k=1}^{n} f(X_k)}_{\text{temporal mean}} = \upsilon_n(f) = \sum_{x\in\mathcal{X}} f(x)\,\upsilon_n(x) \underset{a.s.}{\longrightarrow} \underbrace{\sum_{x\in\mathcal{X}} \pi(x) f(x)}_{\text{spatial mean.}}$$

link with Birkhoff ergodic theorem?

Consider an irreducible, positive recurrent Markov chain with stationary measure π.

<u>lemma</u> If $(X_n)_{n\in\mathbb{N}} \sim \mathbb{P}_{(P,\pi)}$, then $(X_{n+1})_{n\in\mathbb{N}} \sim \mathbb{P}_{(P,\pi)}$.

Indeed, $\mathbb{P}_\pi[X_1 = x_1, \dots X_{n+1} = x_{n+1}] = \sum_{x_0\in\mathcal{X}} \mathbb{P}_\pi[X_0 = x_0, \dots, X_{n+1} = x_{n+1}]$

$$= \sum_{x_0\in\mathcal{X}} \pi(x_0) P(x_0,x_1)\dots P(x_n,x_{n+1})$$

$$\underset{(\text{invariance of }\pi)}{=} \pi(x_1) P(x_1,x_2)\dots P(x_n,x_{n+1}).$$

Therefore, $T : \begin{array}{rcl} \mathcal{X}^{\mathbb{N}} & \to & \mathcal{X}^{\mathbb{N}} \\ (x_n)_{n\in\mathbb{N}} & \mapsto & (x_{n+1})_{n\in\mathbb{N}} \end{array}$ is a transformation of the probability space $(\mathcal{X}^{\mathbb{N}}, \mathcal{P}(\mathcal{X})^{\otimes\mathbb{N}}, \mathbb{P}_\pi)$.

<u>lemma</u> (to be proved later) T is ergodic.

By the Birkhoff ergodic theorem, for any $F \in L^1(\mathcal{X}^{\mathbb{N}}, \mathcal{P}(\mathcal{X})^{\otimes\mathbb{N}}, \mathbb{P}_\pi)$,

$$\frac{F((X_n)_n) + F((X_{n+1})_n) + \dots + F((X_{n+m-1})_n)}{m} \underset{\substack{m\to+\infty \\ a.s. \\ L^1}}{\longrightarrow} \mathbb{E}_\pi[F((X_n)_n)]$$

In particular, if $f \in L^1(\mathcal{X}, \mathcal{P}(\mathcal{X}), \pi)$, then:

$$\frac{f(X_1)+\dots+f(X_m)}{m} \underset{m\to+\infty}{\longrightarrow} \sum_{x\in\mathcal{X}} \pi(x) f(x) \quad \text{a.s. and in } L^1.$$

7. Period of a Markov chain and convergence to the stationary law.

Recall that given $(X_n)_{n\in\mathbb{N}} \sim \mathbb{P}_{(P,\pi_0)}$, the marginal law of X_n is $\pi_n = \pi_0 P^n$.

We saw with the random walk on $\mathbb{Z}/N\mathbb{Z}$ that in some situations $\pi_n \to$ some limiting law π.

<u>lemma</u> if $\left|\begin{array}{l} \pi_n \to \pi, \\ \text{the chain is irreducible} \end{array}\right.$ then π is an invariant law, and therefore the Markov chain is positive recurrent.

Indeed, $\pi = \lim_{n\to+\infty} \pi_{n+1} = \lim_{n\to+\infty} \pi_n P = \pi P$.

However, positive recurrence is not a sufficient condition for convergence of the marginal laws.

example : $\mathcal{X} = [\![1,2]\!]$, $P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, $\pi_0 = \delta_1$. Then $\pi_{2n} = \delta_1$ and $\pi_{2n+1} = \delta_2$, so there is no convergence although the chain is irreducible positive recurrent.

$\leadsto$ we need to take care of possible periodicity of the chain.

Definition The period of a state x of a (recurrent) Markov chain is $h(x) = \gcd(R(x))$, where

$$R(x) = \{\text{possible return times to } x\} = \{n \in \mathbb{N} \mid P^n(x,x) > 0\}.$$

If $n_1, n_2 \in R(x)$, then $P^{n_1+n_2}(x,x) \geq P^{n_1}(x)\, P^{n_2}(x) > 0$, so $n_1 + n_2 \in R(x)$.

$\Longrightarrow$ $R(x)$ is a subset of $\mathbb{N}$ stable by addition.

lemma If $R \neq \{0\}$ is a subset of $\mathbb{N}$ stable by addition and $h = \gcd(R)$, then $R \subset h\mathbb{N}$, and $h\mathbb{N} \setminus R$ is finite.

example : $R = 3\mathbb{N} + 5\mathbb{N} = \{0, 3, 5, 6, 8, 9, 10, \dots\}$ differs from $\mathbb{N} = \gcd(3,5).\mathbb{N}$ by the finite set $\{1, 2, 4, 7\}$.

proof. By definition, h divides every element of R, so $R \subset h\mathbb{N}$. Conversely, there exists a Bezout relation:

$$h = k_1 a_1 + \dots + k_r a_r - \ell_1 b_1 - \dots - \ell_s b_s, \quad \text{the } a_i \text{ and } b_j \in R. \quad \text{the } k_i \text{ and } \ell_j \in \mathbb{N}.$$

By stability by addition : $a = k_1 a_1 + \dots + k_r a_r$ and $b = \ell_1 b_1 + \dots + \ell_s b_s$ belong to R.
$h = a - b$. We claim that all multiples of h starting from ab are in R.

Indeed, if $n = ab + kh$, let us divide k by a: $k = qa + r$, $r < a$.
We then have
$$\begin{aligned} n &= ab + qah + r(a-b) \\ &= (a-r)b + (qh+r)a \in \mathbb{N}a + \mathbb{N}b \subset R. \end{aligned}$$
$\square$.

Suppose now the Markov chain irreducible.

lemma Then, $h(x)$ does not depend on $x \in \mathcal{X}$.

proof: Given $x \neq y$, there exists $\ell, m \geq 1$ such that $P^\ell(x,y) > 0$ and $P^m(y,x) > 0$.

For k large enough, $kh(y)$ and $(k+1)h(y) \in R(y)$.
Then, $kh(y) + \ell + m$ and $(k+1)h(y) + \ell + m \in R(x)$.

So, the difference $h(y)$ is divided by $h(x)$: $h(x) \mid h(y)$. By symmetry, $h(y) \mid h(x)$, so $h(x) = h(y)$ $\square$.
In the irreducible case, we can thus define the period $h = h(P) = h(x)$ of the Markov chain.

Definition The Markov chain is aperiodic if $h(P) = 1$.

$$\iff \forall x,\ P^n(x,x) > 0 \text{ for } n \geq n_x \text{ large enough.}$$
$$\iff \forall x,y,\ P^n(x,y) > 0 \text{ for } n \geq n_{x,y} \text{ large enough.}$$

Theorem Consider an irreducible Markov chain $(X_n)_{n\in\mathbb{N}} \sim \mathbb{P}_{(P,\pi_0)}$.

- If the chain is transient or null recurrent, then $\forall x \in \mathcal{X}$, $\pi_n(x) \underset{n\to+\infty}{\longrightarrow} 0$.
- If the chain is positive recurrent and aperiodic, then $\forall x \in \mathcal{X}$, $\pi_n(x) \underset{n\to+\infty}{\longrightarrow} \pi(x)$. ($\pi$ = stationary law)

(we shall deal with the non-aperiodic case later).

Proof: The transient case is trivial: $\pi_n(x) = \mathbb{P}_{\pi_0}[X_n = x] \leq \mathbb{P}_{\pi_0}[\text{the Markov chain visits } x \text{ after time } n]$
$\longrightarrow 0$ because x is visited a finite number of times a.s.

Consider now the positive recurrent case.
Let π_0 and ν_0 be two arbitrary initial laws. We consider $(X_n)_{n\in\mathbb{N}} \sim \mathbb{P}_{\pi_0}$ and $(Y_n)_{n\in\mathbb{N}} \sim \mathbb{P}_{\nu_0}$ independent.
Then,

$$\begin{aligned}\mathbb{P}[(X_0,Y_0)=(x_0,y_0),\dots,(X_n,Y_n)=(x_n,y_n)] &= \mathbb{P}[X_0 = x_0,\dots,X_n=x_n]\,\mathbb{P}[Y_0=y_0,\dots,Y_n=y_n] \\ &= \pi_0(x_0)P(x_0,x_1)\dots P(x_{n-1},x_n)\ \nu_0(y_0)P(y_0,y_1)\dots P(y_{n-1},y_n) \\ &= \pi_0\otimes\nu_0(x_0,y_0)\ Q((x_0,y_0),(x_1,y_1))\dots Q((x_{n-1},y_{n-1}),(x_n,y_n)).\end{aligned}$$

with $Q((a,b),(x,y)) = P(a,x)\,P(b,y)$.

So, the sequence of pairs is a Markov chain on $\mathcal{X}^2$ with law $\mathbb{P}_{(Q,\pi_0\otimes\nu_0)}$. We claim that it is still irreducible positive recurrent aperiodic:

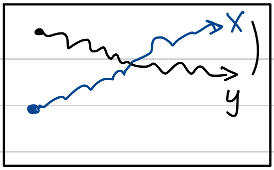
- irreducible aperiodic: $\forall (a,b),(x,y)$, $P^n(a,x)$ and $P^n(b,y)$ are positive for n large enough, so $Q^n((a,b),(x,y)) = P^n(a,x)\,P^n(b,y) > 0$.
- positive recurrent: $\pi\otimes\pi$ is a probability measure invariant by Q.

Set $T = \inf\{n\in\mathbb{N} \mid X_n = Y_n\}$. Because the pair MC visits every state of $\mathcal{X}^2$ a.s., it visits the diagonal, so $T < +\infty$ a.s.

We introduce the total variation distance between two probability measures μ and ν on $\mathcal{X}$:

$$d_{TV}(\mu,\nu) = \sup_{A\subset\mathcal{X}} |\mu(A) - \nu(A)|.$$

Lemma If $m \leq n$, then conditionally to $\{T = m\}$, X_n and Y_n have the same law.



exchangeable after time T.

$\mathcal{X}$

Indeed, $\mathbb{P}_{\pi_0\otimes\nu_0}[T=m \text{ and } X_n=x, Y_n=y] = \sum_{z,\, x_0\neq y_0,\dots, x_{m-1}\neq y_{m-1}} \pi_0(x_0)\nu_0(y_0)P(x_0,x_1)\dots P(x_{m-1},z)P(y_0,y_1)\dots P(y_{m-1},z)\, P^{n-m}(z,x)\, P^{n-m}(z,y)$

and this expression is symmetric with respect to (x,y).

Therefore, given $A\subset\mathcal{X}$,

$$\pi_n(A)-\nu_n(A) = \mathbb{P}[X_n\in A]-\mathbb{P}[Y_n\in A] = \underbrace{\mathbb{P}[X_n\in A \text{ et } T\leq n] - \mathbb{P}[Y_n\in A \text{ et } T\leq n]}_{=0.} + \mathbb{P}[X_n\in A \text{ et } T>n] - \mathbb{P}[X_n\in A \text{ et } T>n]$$

is the difference of two terms $\in[0,\mathbb{P}[T>n]]$

$\Longrightarrow d_{TV}(\pi_n,\nu_n)\leq \mathbb{P}[T>n]$.

With $\nu_0=\pi$, $\nu_n=\pi$ $\forall n\in\mathbb{N}$, so $d_{TV}(\pi_n,\pi)\leq \mathbb{P}_{\pi_0\otimes\pi}[T>n] \xrightarrow[n\to+\infty]{} 0$ and the convergence to the stationary law is proven.

Consider finally the null recurrent case.

__lemma__: Suppose P recurrent irreducible on $\mathcal{X}$ with period $h\geq 2$. Then, there is a partition $\mathcal{X}=\mathcal{X}_1\sqcup\mathcal{X}_2\sqcup\dots\sqcup\mathcal{X}_h$ of $\mathcal{X}$ such that P^h induces on each $\mathcal{X}_i$ a recurrent irreducible Markov chain.

__proof__: Fix a base point $x_0\in\mathcal{X}$ and denote

$$\mathcal{X}_i = \{x\in\mathcal{X} \mid \exists n\equiv i[h] \text{ with } P^n(x_0,x)>0\}.$$

Since the chain is irreducible, $\mathcal{X}=\bigcup_{i=1}^h \mathcal{X}_i$. Moreover, for any $x\in\mathcal{X}$, if $n\equiv i[h]$ and n is large enough, then $P^n(x_0,x)>0$.

Suppose that $x\in\mathcal{X}_i\cap\mathcal{X}_j$. Fix $m>0$ such that $P^m(x,x_0)>0$. For k large enough, $P^{i+kh}(x_0,x)>0$ and $P^{j+kh}(x_0,x)>0$. Then, $i+kh+m$ and $j+kh+m\in R(x_0)$, so h divides the difference: this forces $i=j$.

So, the union is disjoint.

Fix x in some class $\mathcal{X}_i$. By assumption of recurrence, $\sum_{k\in R(x)} P^k(x,x)=+\infty$, and $R(x)\subset h\mathbb{N}$,

So $\sum_{n\in\mathbb{N}}(P^h)^n(x,x)=+\infty$ and x is recurrent for P^h.

Finally, if $x \overset{P^h}{\rightsquigarrow} y$, then $y\in\mathcal{X}_i$, so the restriction of P^h to $\mathcal{X}_i$ induces an irreducible recurrent chain $\square$.

Notice now that $P^h|_{\mathfrak{X}_i}$ has period 1. Suppose that one can prove that for a null recurrent aperiodic chain, the marginal laws go to 0. Then, in the general case $h \geq 1$, $\Pi_{nh} = \Pi_0 (P^h)^n \xrightarrow[n\to+\infty]{} 0$. But this is sufficient to prove that $\Pi_m = \Pi_{nh+r} = \underbrace{\Pi_{nh}}_{\to 0} P^r \to 0$.
$\hookrightarrow 0 \leq r < h$, a finite number of cases.

So, let us consider a null recurrent aperiodic chain. By the same argument as before, $(X_n, Y_n)_{n\in\mathbb{N}} \sim \mathbb{P}_{\Pi_0} \otimes \mathbb{P}_{\upsilon_0}$ is also irreducible aperiodic.

1st case: the pair Markov chain is transient. Then, $\underbrace{\mathbb{P}_{\Pi_0\otimes\Pi_0}[(X_n, Y_n) = (x,x)]}_{= (\Pi_n(x))^2} \xrightarrow[n\to+\infty]{} 0$.

2nd case: the pair Markov chain is recurrent. By compacity, if $\Pi_0 = \delta_y$ and $\upsilon_0 = \delta_z$, then:

1) $\exists (n_k)_{k\in\mathbb{N}}$ subsequence such that $P^{n_k}(y,x) \to \rho(x)$ $\forall x \in \mathfrak{X}$.
2) By the same argument as before, $d_{TV}(P^n(y,\cdot), P^n(z,\cdot)) \xrightarrow[n\to+\infty]{} 0$, so we have also $P^{n_k}(z,x) \to \rho(x)$.

For any $k \in \mathbb{N}$, $\sum_{x\in\mathfrak{X}} P^{n_k}(y,x) = 1$.
$$\Longrightarrow \sum_{x\in\mathfrak{X}} \rho(x) \leq 1. \quad \text{(Fatou's lemma).}$$

Then,
$$(\rho P)(x) = \sum_{y\in\mathfrak{X}} \rho(y) P(y,x) = \sum_{y\in\mathfrak{X}} \lim_{k\to+\infty} P^{n_k}(z,y) P(y,x) \leq \liminf_{k\to+\infty} \sum_{y\in\mathfrak{X}} P^{n_k}(z,y) P(y,x)$$
$$\leq \liminf_{k\to+\infty} P^{n_k+1}(z,x) = \liminf_{k\to+\infty} \sum_{y\in\mathfrak{X}} P(z,y) P^{n_k}(y,x) \underset{\substack{\uparrow \\ \text{dominated convergence}}}{=} \sum_{y\in\mathfrak{X}} P(z,y)\rho(x) = \rho(x).$$

Since the action of stochastic matrices preserves the total mass of measure, we also have $\sum_{x\in\mathfrak{X}} \rho P(x) = \sum_{x\in\mathfrak{X}} \rho(x)$, so above we cannot have a strict inequality.

Conclusion: $\rho P = \rho$. Then, we cannot have $\rho \neq 0$, because this would imply the positive recurrence of the chain.
So any convergent subsequence of $P^n(y,x)$ is 0, and this implies: $\forall y, x$, $\lim_{n\to+\infty} P^n(y,x) = 0$.
By taking linear combinations, $\lim_{n\to+\infty} \Pi_n(x) = 0$.

$\square$.

Two remaining details:

① what happens in the positive recurrent periodic ($h \geq 2$) case?
The proof above shows that P^h is irreducible recurrent on each class $\mathfrak{X}_i$.

Denote π the stationary law, $\pi = \sum_{i=1}^{h} \pi_i$ its decomposition with respect to $\mathcal{X} = \bigsqcup_{i=1}^{h} \mathcal{X}_i$.

<u>lemma</u> if $x \in \mathcal{X}_i$ and $P(x,y) > 0$, then $y \in \mathcal{X}_{i+1 \bmod h}$, and conversely, if $y \in \mathcal{X}_{i+1}$, then $P(x,y) > 0$.
(proof similar to the previous arguments for the set partition $\mathcal{X} = \bigsqcup_{i=1}^{h} \mathcal{X}_i$).

This implies that $\pi_i P$ is supported by $\mathcal{X}_{i+1}$. Moreover, $\sum_{i=1}^{h} \pi_i = \pi = \pi P = \sum_{i=1}^{h} \pi_i P$,
therefore $\pi_i P = \pi_{i+1 \bmod h}$.
$\Rightarrow \pi_i(\mathcal{X}_i) = \frac{1}{h}$, $h\pi_i$ is the stationary measure of $P^h|_{\mathcal{X}_i}$.

Consider then an arbitrary initial measure π_0 on P, decomposed as $\pi_0 = \sum_{i=1}^{h} \pi_0|_{\mathcal{X}_i}$. By the aperiodic case, $(\pi_0|_{\mathcal{X}_i}) P^{hm} \xrightarrow[m \to +\infty]{} \pi_0(\mathcal{X}_i)\pi_i$.

Then, if $n = hm + r$ with $0 \le r < h$, we have $\pi_0 P^n = \sum_{i=1}^{h} (\pi_0|_{\mathcal{X}_i}) P^n \xrightarrow[n \to +\infty]{} \sum_{i=1}^{h} \pi_0(\mathcal{X}_i)\pi_i P^r$.

$$\pi_0 \frac{(P^n + P^{n+1} + \dots + P^{n+h-1})}{h} \xrightarrow[n \to +\infty]{} \sum_{i=1}^{h} \pi_0(\mathcal{X}_i)\pi = \pi.$$

Conclusion: if $h \ge 2$, $\frac{\pi_n + \pi_{n+1} + \dots + \pi_{n+h-1}}{h} \xrightarrow[n \to +\infty]{} \pi$ invariant law.

② ergodicity of a positive recurrent Markov chain.
Consider an invariant event A for $T : (x_n)_{n \in \mathbb{N}} \to (x_{n+1})_{n \in \mathbb{N}}$.
For any $B \in \mathcal{P}(\mathcal{X})^{\otimes \mathbb{N}}$, any $\varepsilon > 0$, there exists $n_{\varepsilon,B} = n$ and $B_n \in \mathcal{F}_n$ such that $\mathbb{P}_\pi[B \,\Delta\, B_n] \le \varepsilon$.
Then, $A_n \cap T^{-k}(A_n)$ approximates A:

$$\mathbb{P}_\pi[A] = \mathbb{P}_\pi[A \cap T^{-k}(A)] \le \mathbb{P}_\pi[(A \Delta A_n) \cap T^{-k}(A)] + \mathbb{P}_\pi[A_n \cap T^{-k}(A)]$$
$$\le \varepsilon + \mathbb{P}_\pi[A_n \cap T^{-k}(A \Delta A_n)] + \mathbb{P}_\pi[A_n \cap T^{-k}(A_n)] \le 2\varepsilon + \mathbb{P}_\pi[A_n \cap T^{-k}(A_n)],$$

and similarly $\mathbb{P}_\pi[A_n \cap T^{-k}(A_n)] \le 2\varepsilon + \mathbb{P}_\pi[A]$.
Thus, $|\mathbb{P}_\pi[A] - \mathbb{P}_\pi[A_n \cap T^{-k}(A_n)]| \le 2\varepsilon$.
We now claim that if A_n is fixed, then $\lim_{k \to +\infty} \frac{1}{h} \sum_{j=0}^{h-1} \mathbb{P}_\pi[A_n \cap T^{-k-j}(A_n)] = (\mathbb{P}_\pi[A_n])^2$.
This will imply in the end $\mathbb{P}_\pi[A] = (\mathbb{P}_\pi[A])^2$, so $\mathbb{P}_\pi[A] \in \{0,1\}$ and the ergodicity.
The event A_n fixes $X_0 = x_0, \dots, X_n = x_n$ for $(x_0, \dots, x_n)$ in some fixed set $S \subset \mathcal{X}^{n+1}$.

Then, for $k \geq n$,

$$\mathbb{P}_\pi[A_n \cap T^{-k}(A_n)] = \sum_{\substack{x_0,\dots,x_n\in S\\ y_0,\dots,y_n \in S}} \pi(x_0)P(x_0,x_1)\dots P(x_{n-1},x_n)P^{k-n}(x_n,y_0)P(y_0,y_1)\dots P(y_{n-1},y_n)$$

$$\frac{1}{h}\sum_{j=0}^{h-1}\mathbb{P}_\pi[A_n \cap T^{-k-j}(A_n)] = \sum_{\substack{x_0,\dots,x_n\in S\\ y_0,\dots,y_n \in S}} \pi(x_0)\dots P(x_{n-1},x_n)\left(\frac{1}{h}\sum_{j=0}^{h-1}P^{k-n+j}(x_n,y_0)\right)P(y_0,y_1)\dots P(y_{n-1},y_n)$$

$$\xrightarrow[n\to+\infty]{} \sum \pi(x_0)P(x_0,x_1)\dots P(x_{n-1},x_n)\pi(y_0)P(y_0,y_1)\dots P(y_{n-1},y_n).$$

convergence to the stationary law

$$= (\mathbb{P}_\pi[A_n])^2 \ !$$

$\square$.