A classical example of time-reversible M.C.:

the Ehrenfest chain (1907 Paul & Tatiana Ehrenfest)
to describe the movements of molecules

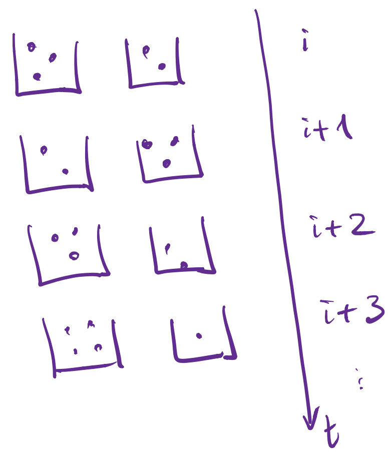
→ Toy model of gas behaviour in two containers

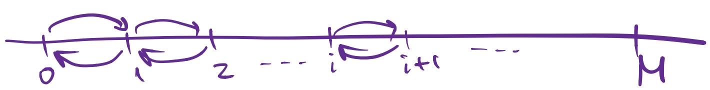
→ Consider M balls distributed in 2 urns.

At each step select a ball at random and move it to the other urn.

$X_n = \#$ balls in urn #1.

$S = \{0, 1, \dots, M\}$.





The chain is ergodic (why?) finite state space, aperiodic irreducible, positive recurrent ⇒ ergodic

What is the stationary distribution?

$$P(X_n = i+1 \mid X_{n-1} = i) = \frac{M-i}{M}.$$

$$P(X_n = i-1 \mid X_{n-1} = i) = \frac{i}{M}.$$

for $i = 0, 1, \dots, M$.

$$\pi_i p_{i,i+1} = \pi_{i+1} p_{i+1,i}$$

$$\pi_i \frac{M-i}{\cancel{M}} = \pi_{i+1} \cdot \frac{i+1}{\cancel{M}}$$

$$\frac{\pi_i}{\pi_{i+1}} = \frac{i+1}{M-i}$$

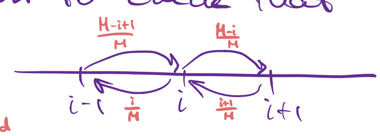
Intuition for time-reversibility: State 0 only receives "jumps" from state 1, and only gives "jumps" to 1, so the # jumps $0 \to 1$ and #jumps $1 \to 0$ are equal.

Similarly, between every 2 consecutive jumps $i \to i+1$, there must be exactly one jump $i+1 \to i$

$\Rightarrow$ #jumps $i \to i+1$ will equal #jumps $i+1 \to i$ (± 1 but that doesn't matter in the limit).

$\Rightarrow$ we get detailed balance.

Let's use our time-reversibility criterion to check that the Ehrenfest chain is time-reversible.



Check for a loop $i \to i+1 \to \cdots \to i+k$
$\nwarrow \cdots \leftarrow i+k-1 \swarrow$

it's the same loop clockwise and counter-clockwise.

clockwise: $p_{i,i+1} \cdots p_{i+k-1,i+k} \cdot p_{i+k,i+k-1} \cdots p_{i+1,i}$

counter-clockwise: $p_{i,i+1} \cdots p_{i+k-1,i+k} \cdot p_{i+k,i+k-1} \cdots p_{i+1,i}$

same

To calculate π, we can either find a solution to the detailed balance equations ($x_i p_{ij} = x_j p_{ji}$), or guess and check a solution that works.

each ball should spend half the time in each urn. $\Rightarrow$ #balls is $\text{Bin}(M, \frac{1}{2})$...

Solving the detailed balance equations:

$$\pi_i p_{ij} = \pi_j p_{ji}$$

do only for $j = i \pm 1$

otherwise $p_{ij} = 0$.

$$\sum \pi_i = 1.$$

$$\pi_i p_{i,i+1} = \pi_{i+1} p_{i+1,i}.$$

$$\pi_i \frac{M-i}{\cancel{M}} = \pi_{i+1} \cdot \frac{i+1}{\cancel{M}} \qquad \forall i = 0, 1, \dots, M-1.$$

$$\pi_{i+1} = \frac{M-i}{i+1} \pi_i \qquad \forall i = 0, 1, \dots, M-1$$

$$\pi_1 = \frac{M}{1}\pi_0 \quad , \quad \pi_2 = \frac{M-1}{2}\pi_1 = \frac{(M-1)}{2}\cdot\frac{M}{1}\cdot\pi_0$$

$$\pi_3 = \frac{M-2}{3}\pi_2 = \frac{M-2}{3}\cdot\frac{M-1}{2}\cdot\frac{M}{1}\pi_0$$

$$\dots \quad \pi_i = \frac{M-i+1}{i}\pi_{i-1} = \dots = \frac{M-i+1}{i}\cdot\frac{M-i+2}{i-1}\cdots\frac{M-2}{3}\cdot\frac{M-1}{2}\cdot\frac{M}{1}\pi_1$$

$$= \frac{M!}{i!\,(M-i)!}\pi_0 = \binom{M}{i}\pi_0$$

Also $\sum_{i=0}^{M}\pi_i = 1$

$$\Rightarrow \quad \sum_{i=0}^{M}\binom{M}{i}\pi_0 = 1 \quad \Rightarrow \quad \pi_0\Big(\underbrace{\sum_{i=0}^{M}\binom{M}{i}}_{2^M}\Big) = 1$$

$$\Rightarrow \pi_0 = \frac{1}{2^M} \quad \text{and} \quad \boxed{\pi_i = \binom{M}{i}\frac{1}{2^M}}.$$

$$\Rightarrow \pi_i \sim \mathrm{Bin}\left(M, \tfrac{1}{2}\right).$$

Branching process (Ross 4.7, but we go into more depth)

- We conclude our introduction to discrete-time M.C. with the so-called branching process.
- Such processes originated from Galton & Watson (1874) who studied the dynamics and extinction of family names of British nobility
  ↳ lead to important applications in population genetics, ecology etc.

$Z_0 = 1$

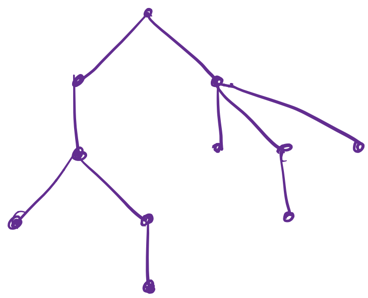
$Z_1 = 2$

$Z_2 = 4$

$Z_3 = 3$

$Z_4 = 1$

$Z_n = 0, n \geq 5$



- A population evolves in genealogies
- Each individual has a random number of offspring (with same distribution) and independently the individuals. $\triangledown$
- this process can be modeled as follows.

<u>Definition</u>: A branching process (or Galton-Watson process) is a sequence of random variables $(Z_n)_{n \geq 0}$ defined by

$$\begin{cases} Z_0 = 1 \\ Z_{n+1} = \sum_{i=1}^{Z_n} Y_{n,i}, \end{cases}$$

where $Y_{n,i}$ are i.i.d. random variables on $\mathbb{N}$ whose law is called the reproduction law of the process. Z_n is the # of individuals in the n-th generation, and $Y_{n,i}$ is the # of offspring of the i-th individual of the n-th generation.

<u>Q</u>: The original question asked by Galton and Watson was: what is the probability of extinction of the population?
More generally, how can we find $P(Z_n = k)$?

<u>Remarks</u>: $\cdot \{0\}$ is a recurrent class, also absorbing. Assuming

$P(Y_{n,i} > 0) = p_0 > 0$, $P_{i0} = p_0^i > 0$, so 0 is accessible from all $i > 0$.

$\Rightarrow i$ is transient $\forall i > 0$.

- Since any finite set of transient states can only be visited finitely many times, either Z_n is eventually 0, or $Z_n \to \infty$.

To study the extinction probability (or the survival probability), we will use Generating functions.

<u>Definition</u>: Let X be a random variable on $\mathbb{N}$. We call the generating function of X the power series:

$$G_X(s) = \mathbb{E}[s^X] = \sum_{k=0}^{\infty} P(X=k)\, s^k.$$

<u>Remarks</u>: • The radius of convergence of G_X is ≥ 1 since

$$G_X(1) = \sum_{k=0}^{\infty} P(X=k) = 1. \Rightarrow G_X \text{ is well-defined } \forall s \in [-1,1].$$

• $G_X(0) = P(X=0)$, $G_X'(s) = \sum_{k=1}^{\infty} P(X=k)\, k\, s^{k-1} \Rightarrow G_X'(0) = P(X=1)$.

$$G_X''(s) = \sum_{k=2}^{\infty} P(X=k)\, k(k-1)\, s^{k-2} \Rightarrow G_X''(0) = \frac{P(X=2)}{2!}$$

$$\ldots \qquad G_X^{(k)}(0) = \frac{P(X=k)}{k!}.$$

$\Rightarrow$ Two random variables with the same generating functions have the same laws.

<u>Theorem</u>: If X and Y are independent random variables on $\mathbb{N}$, then $\forall t \in [-1,1]$, $G_{X+Y}(t) = G_X(t) \cdot G_Y(t)$.

<u>Proof</u>: $G_{X+Y}(t) = \mathbb{E}\left[t^{X+Y}\right] = \mathbb{E}\left[t^X \cdot t^Y\right] \underset{\text{independence}}{=} \mathbb{E}[t^X]\,\mathbb{E}[t^Y]$

$$= G_X(t)\, G_Y(t).$$

<u>Proposition</u>: Let $X_1, X_2, \ldots$ be i.i.d. r.v.'s with generating function G_X. Let N be a r.v. with generating function G_N, independent of $X_1, X_2, \ldots$. Let $T = X_1 + \cdots + X_N$. Then, $G_T(s) = G_N(G_X(s))$.

Proof: $G_T(s) = \mathbb{E}[s^T] = \mathbb{E}[\mathbb{E}[s^T | N]]$

$$= \sum_{n=0}^{\infty} \mathbb{E}[s^T | N=n] \, P(N=n) = \sum_{n=0}^{\infty} \mathbb{E}[s^{X_1+\dots+X_n}] \, P(N=n)$$

$$= \sum_{n=0}^{\infty} (G_X(s))^n P(N=n) = G_T(G_X(s)).$$

Application to the Branching process:

$$Z_{n+1} = Y_{n,1} + Y_{n,2} + \dots + Y_{n,Z_n}.$$

$$\Rightarrow G_{Z_{n+1}}(s) = G_{Z_n}(G_Y(s)) = G_{Z_{n-1}}(G_Y(G_Y(s)))$$

$$= G_{Z_{n-1}}(G_Y \circ G_Y(s)) = \dots = G_{Z_1}(\underbrace{G_Y \circ \dots \circ G_Y}_{n}(s))$$

$Z_1 = Y_{0,1} \Rightarrow$ $$= \underbrace{G_Y \circ G_Y \circ \dots \circ G_Y}_{n+1.}(s) =: G_{n+1}(s)$$

Will show $P(\text{extinction}) = \lim_{n \to \infty} G_n(0).$