Lectures on Markov processes

Alexandre V. Morozov

Let us consider a system that can be in any of the N_S states from some state space S. Let X be a random variable that takes values in S. Then the state of the system as it evolves in time can be described by

$$\{ X(t_1), X(t_2), \dots \}$$

In other words, the state of the system at time t is given by a random function $X(t)$:

$t \to N_S$ values in S

The function $X(t)$ is called a <u>stochastic process</u>.

In general, $X(t)$ has several components (i.e., is a vector) because states are described by vectors.

For example, in 3D space:

$$\vec{X}(t) = (X(t), Y(t), Z(t))$$

states are described by $\vec{x}(t) = (x(t), y(t), z(t))$

We can define a joint prob. $p(\vec{x}_1, t_1; \vec{x}_2, t_2; \dots)$ that the system is in state $\vec{x}_1$ @ t_1, $\vec{x}_2$ @ t_2, etc.

In the simplest case,

$$p(\vec{x}_1, t_1; \vec{x}_2, t_2; \dots) = \prod_i \underbrace{p(\vec{x}_i, t_i)}_{\text{independence}}$$

but in general this is <u>not</u> the case.

Consider conditional probabilities:

$$p(\vec{x}_1, t_1; \vec{x}_2, t_2; \dots | \vec{y}_1, \tau_1; \vec{y}_2, \tau_2; \dots) =$$

$$= \frac{p(\vec{x}_1, t_1; \vec{x}_2, t_2; \dots; \vec{y}_1, \tau_1; \vec{y}_2, \tau_2; \dots)}{p(\vec{y}_1, \tau_1; \vec{y}_2, \tau_2; \dots)}$$

$P(A,B) = P(A|B)P(B)$

Use $t_1 > t_2 > \dots > \tau_1 > \tau_2 > \dots$ without any loss of generality

<u>Markov processes</u> are stochastic processes such that

$$p(\vec{x}_1, t_1; \vec{x}_2, t_2; \dots | \vec{y}_1, \tau_1; \vec{y}_2, \tau_2; \dots) =$$

$$= p(\vec{x}_1, t_1; \vec{x}_2, t_2; \dots | \vec{y}_1, \tau_1).$$

But then

$$p(\vec{x}_1, t_1; \vec{x}_2, t_2; \vec{y}_1, \tau_1) = \underbrace{p(\vec{x}_1, t_1; \vec{x}_2, t_2 | \vec{y}_1, \tau_1)}_{\substack{p(\vec{x}_1, t_1 | \vec{x}_2, t_2; \vec{y}_1, \tau_1) \times \\ \times p(\vec{x}_2, t_2 | \vec{y}_1, \tau_1)}} \times$$

$$\times\, p(\vec{y}_1, \tau_1) = $$

$$= \underbrace{p(\vec{x}_1, t_1 | \vec{x}_2, t_2)}_{\text{Markov property (assumption)}} p(\vec{x}_2, t_2 | \vec{y}_1, \tau_1)\, p(\vec{y}_1, \tau_1).$$

This argument can be extended to joint probabilities of any order

---

Now, consider the time-evolution equation:

$$p(\vec{x}_1, t_1) = \int d\vec{x}_2\, p(\vec{x}_1, t_1 | \vec{x}_2, t_2)\, p(\vec{x}_2, t_2)$$

$$\sum_B P(A|B)\, P(B) = \sum_B P(A, B) = P(A)$$

Likewise,

$$p(\vec{x}_1, t_1 | \vec{x}_3, t_3) = \int d\vec{x}_2\, p(\vec{x}_1, t_1 | \vec{x}_2, t_2; \vec{x}_3, t_3) \times$$

$$\times\, p(\vec{x}_2, t_2 | \vec{x}_3, t_3) \overset{\text{Markov}}{=}$$

$t_1 > t_2 > t_3$

$$= \int d\vec{x}_2\, p(\vec{x}_1, t_1 | \vec{x}_2, t_2)\, p(\vec{x}_2, t_2 | \vec{x}_3, t_3).$$

Chapman - Kolmogorov (CK) equation

Thus, a Markov process is defined by:

(1) The CK equation:

$$p(\vec{x}_1,t_1|\vec{x}_3,t_3) = \int d\vec{x}_2\, p(\vec{x}_1,t_1|\vec{x}_2,t_2) \times$$
$$\times\, p(\vec{x}_2,t_2|\vec{x}_3,t_3)$$

(2) The time-evolution equation:

$$p(\vec{x}_1,t_1) = \int d\vec{x}_2\, p(\vec{x}_1,t_1|\vec{x}_2,t_2)\, p(\vec{x}_2,t_2)$$

(3) Normalization:

$$\int d\vec{x}\, p(\vec{x},t) = 1, \qquad \int d\vec{x}_1\, p(\vec{x}_1,t_1|\vec{x}_2,t_2) = 1$$

$$\sum_A P(A) = 1 \qquad\qquad \sum_A P(A|B) = 1$$

<u>Homogeneous process</u>: $p(\vec{x}_1,t_1|\vec{x}_2,t_2)$ depends only on t_1-t_2

<u>Stationary process</u>: homogeneous + $p(\vec{x},t)$ is indep. of t: $p(\vec{x},t) \to p(\vec{x})$

It is convenient to use the bra-ket notation for homogeneous processes:
define a vector $|x\rangle$ for each state $x \in S$. This defines a vector space $\mathcal{V}$ of dim. N_S spanned by all $|x\rangle$.

Then $\sum_x p(x,t)\,|x\rangle = \underbrace{|p(t)\rangle}_{\substack{\text{a time-dependent} \\ \text{vector in } \mathcal{J}}} \Rightarrow p(x',t) = \langle x'|p(t)\rangle$

Moreover, define $p(x_1,t_1|x_2,t_2) =$

$$= \langle x_1|\underbrace{\Pi(t_1-t_2)}_{\substack{\text{matrix } \Pi; \\ \text{homogeneity assumed}}}|x_2\rangle$$

In this notation, we have:

(1) The CK equation:

$$\langle x_1|\Pi(t_1-t_3)|x_3\rangle = \sum_{x_2} \langle x_1|\Pi(t_1-t_2)|x_2\rangle \times$$
$$\times \langle x_2|\Pi(t_2-t_3)|x_3\rangle$$

$$\hookrightarrow \Pi(t_1)\,\Pi(t_2) = \Pi(t_1+t_2)$$

(2) The time-evolution equation:

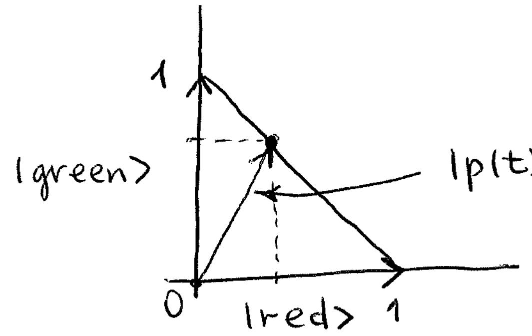
$$\langle x_1|p(t_1)\rangle = \sum_{x_2} \langle x_1|\Pi(t_1-t_2)|x_2\rangle\langle x_2|p(t_2)\rangle,$$
or

$$|p(t_1)\rangle = \Pi(t_1-t_2)\,|p(t_2)\rangle$$

(3) Normalization:

$$\sum_x \langle x|p(t)\rangle = 1, \quad \sum_{x_1} \langle x_1|\Pi(t)|x_2\rangle = 1.$$

Ex. $N_s = 2$ $\qquad \begin{cases} |\text{red}\rangle = (1, 0) \\ |\text{green}\rangle = (0, 1) \end{cases}$ red ball, green ball



$$|p(t)\rangle = p(\text{red}, t)\,|\text{red}\rangle + p(\text{green}, t)\,|\text{green}\rangle$$

Note that any valid transformation of $|p(t)\rangle$ has to preserve normalization. This is exactly what's accomplished by the time-evolution equation (note that

$$\sum_{x_1} \langle x_1 | \Pi(t) | x_2 \rangle = 1 \text{).}$$

Since $\Pi(t_1)\Pi(t_2) = \Pi(t_1 + t_2)$ due to the CK equation, we have $\Pi(t) = \underbrace{P}_{\text{constant operator}}{}^{t/\underbrace{\Delta t}_{\text{time scale}}}$

For the discrete time process, $t = m\Delta t$ and

$$|p(t + m\Delta t)\rangle = P^m |p(t)\rangle$$

Thus the matrix elements $\langle x_1 | P | x_2 \rangle$ are transition probabilities from x_2 to x_1 in time Δt.

Consider a system with $\mathcal{N}_s$ states, then the discrete-time dynamics is given by the <u>master equation</u>:

$$\underbrace{p_i(t+1)}_{\Delta t=1 \text{ here}} = p_i(t) + \overbrace{\sum_{j\in \underbrace{nn_i}_{\substack{\text{nearest neighbors}\\ \text{of state } i}}} P(j\to i)\, p_j(t)}^{\text{gain term}} -$$

$i = 1, \ldots, \mathcal{N}_s$

$$\underbrace{- p_i(t) \sum_{j\in nn_i} P(i\to j)}_{\text{loss term}} \;=$$

introduce

$$\begin{cases} P(j\to i) = 0, & j \notin nn_i \\ P(i\to j) = 0, & j\notin nn_i \end{cases}$$

$$= p_i(t) + \sum_{j\neq i} P(j\to i)\, p_j(t) - p_i(t) \sum_{j\neq i} P(i\to j) =$$

$$= p_i(t) \left[1 - \sum_{j\neq i} P(i\to j)\right] + \sum_{j\neq i} P(j\to i)\, p_j(t) .$$

$$\boxed{\text{Note that } \sum_{i=1}^{\mathcal{N}_s} p_i(t) = 1}$$

Introducing $|p(t)\rangle = \begin{pmatrix} p_1(t) \\ p_2(t) \\ \vdots \\ p_{\mathcal{N}_s}(t) \end{pmatrix}$,

we immediately see that

$|p(t+1)\rangle = P\, |p(t)\rangle$, where

$$\underbrace{P}_{\text{transition matrix}} = \begin{pmatrix} 1 - \sum_{j=1, j\neq 1}^{\mathcal{N}_s} P(1\to j) & P(2\to 1) & P(3\to 1) & \ldots & P(\mathcal{N}_s\to 1) \\ P(1\to 2) & 1 - \sum_{\substack{j=1\\ j\neq 2}}^{\mathcal{N}_s} P(2\to j) & \ldots & & \\ P(1\to 3) & \ldots & & & \\ \vdots & \ldots & & & \\ P(1\to \mathcal{N}_s) & & & & \end{pmatrix}$$

Note that each column in $\mathbb{P}$ sums to 1.0 and that we can define

$$\underbrace{\mathbb{P}(i\to i)}_{\text{diagonal elements of } \mathbb{P}} = 1 - \sum_{j=1,\, j\neq i}^{N_s} \mathbb{P}(i\to j) \;,$$

the probability to remain in state i after a time step. Note also that

$$\sum_{j\in nn_i} \mathbb{P}(i\to j) \leq 1$$

follows.

Continuous time limit

Consider $|p(t+\Delta t)\rangle = \mathbb{P}\,|p(t)\rangle$, then

$$\frac{|p(t+\Delta t)\rangle - |p(t)\rangle}{\Delta t} = \frac{1}{\Delta t}(\mathbb{P} - \underset{\substack{\uparrow \\ \text{unit matrix}}}{\mathbb{1}})\,|p(t)\rangle$$

In the $\Delta t \to 0$ limit, we have

$$(+) \qquad \partial_t |p(t)\rangle = \mathbb{W}\,|p(t)\rangle \;, \quad \text{where}$$

$$\mathbb{W} = \lim_{\Delta t\to 0^+} \frac{1}{\Delta t}(\mathbb{P} - \mathbb{1})$$

Thus,

$$\langle x|\mathbb{W}|y\rangle = \begin{cases} \lim\limits_{\Delta t\to 0^+} \dfrac{1}{\Delta t}\langle x|\mathbb{P}|y\rangle \;, & x\neq y \\[2ex] \lim\limits_{\Delta t\to 0^+} \left[-\dfrac{1}{\Delta t}\sum\limits_{z\neq x}\langle z|\mathbb{P}|x\rangle\right] = -\sum\limits_{z\neq x}\langle z|\mathbb{W}|x\rangle \;, & x = y \end{cases}$$

$$\langle x|\mathbb{P}|x\rangle = 1 - \sum_{y\neq x}\langle y|\mathbb{P}|x\rangle$$

Now, apply $\langle x|$ to Eq. (+):

$$\partial_t p(x,t) = \langle x|W|\underbrace{p(t)\rangle}_{\sum_y p(y,t)|y\rangle} = \sum_y \langle x|W|y\rangle p(y,t) \ominus$$

$$\ominus \sum_{y\neq x} \langle x|W|y\rangle p(y,t) + \underbrace{\langle x|W|x\rangle}_{-\sum_{y\neq x}\langle y|W|x\rangle} p(x,t) =$$

$$= \underbrace{\sum_{y\neq x} \langle x|W|y\rangle p(y,t)}_{\text{gain}} \underbrace{-\Big(\sum_{y\neq x}\langle y|W|x\rangle\Big) p(x,t)}_{\text{loss}} .$$

We can interpret $\underset{x\neq y}{\langle x|W|y\rangle}$ elements as <u>rates</u> from state y to state x.

The formal solution of Eq. (+) is given by

$$|p(t)\rangle = e^{tW}|p(0)\rangle = \Pi(t)\,|p(0)\rangle .$$

Indeed, recall that $\Pi(t) = P^{t/\Delta t}$ in the discrete case:

$$\underbrace{\Pi(t)}_{\text{continuous case}} = \lim_{\Delta t\to 0+} P^{t/\Delta t} = \lim_{\Delta t\to 0+} (1+\Delta t\, W)^{t/\Delta t} = e^{tW} .$$

Explicitly,

$$W = \begin{pmatrix} -\sum_{j=2}^{N_s} W(1\to j) & W(2\to 1) & \cdots \\ W(1\to 2) & -\sum_{j=1, j\neq 2}^{N_s} W(2\to j) & \cdots \\ W(1\to 3) & \cdots & \cdots \\ \vdots & \cdots & \\ W(1\to N_s) & \cdots & \end{pmatrix}$$

Each column in W sums to 0.0.

Eq'n (+) can be generalized to

$$\partial_t |p(t)\rangle = W(t)|p(t)\rangle$$ if the rates are time-dependent.

The formal solution is:

$$|p(t)\rangle = \mathcal{T}\{e^{\int_0^t d\tau\, W(\tau)}\}|p(0)\rangle,$$ where

$$\langle x|\mathcal{T}\{e^{\int_0^t d\tau\, W(\tau)}\}|y\rangle =$$

$$= \langle x|y\rangle + \int_0^t d\tau \langle x|W(\tau)|y\rangle +$$

$$+ \frac{1}{2}\int_0^t d\tau_1 \int_0^t d\tau_2 \langle x|\mathcal{T}\{W(\tau_2)W(\tau_1)\}|y\rangle + \ldots$$

Here, $\mathcal{T}\{\ldots\}$ is the time-ordering operator:

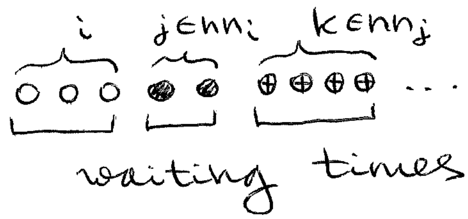
$$\mathcal{T}\{A(t_1)B(t_2)\} = \begin{cases} A(t_1)B(t_2) & \text{if } t_1 \geq t_2 \\ B(t_2)A(t_1) & \text{if } t_1 < t_2 \end{cases}$$

Time-ordering is needed since

$$W(t_1)W(t_2) \neq W(t_2)W(t_1) \text{ in general}$$

Discrete-time embedding

In general, stochastic processes we have discussed so far have a finite probability of staying in the current state and a finite probability of leaving. This translates into waiting in the same state followed by jumps:



The prob. to remain in state x after one very small step Δt is given by

$$\langle x|\Pi(\Delta t)|x\rangle = \langle x|e^{\Delta t W}|x\rangle \simeq \langle x|\mathbb{1}+\Delta t W|x\rangle =$$

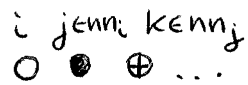
$$= 1 + \Delta t \underbrace{\langle x|W|x\rangle}_{<0}$$

The prob. for the system to remain in state x over a finite time interval t is given by $[\Delta t = \frac{t}{m}]$

$$\lim_{m\to\infty}\left(1+\frac{t}{m}\langle x|W|x\rangle\right)^m = e^{t\langle x|W|x\rangle}.$$

Thus, waiting times are distributed exponentially with $\tau = -\frac{1}{\langle x|W|x\rangle}$.

Now, let us consider a "projected" stochastic process:



Here, the jump occurs at every time step by construction and therefore $\langle x|P|x\rangle = 0$ <u>unless</u> state x is absorbing: $\langle x|W|x\rangle = 0$.

altogether,
$$\langle x|P|x\rangle = \begin{cases} 0, & \langle x|W|x\rangle \neq 0 \\ 1, & \langle x|W|x\rangle = 0 \end{cases}$$

Furthermore,
$$\underbrace{\langle x|P|y\rangle}_{\text{projected}} = \lim_{\Delta t \to 0^+} \frac{\overbrace{\langle x|\Pi(\Delta t)|y\rangle}^{\text{prob. to transition } y \to x}}{\underbrace{1 - \langle x|\Pi(\Delta t)|x\rangle}_{\text{prob. to leave to any } y \in nn_x}} \stackrel{\bigcirc}{=}$$

$$\stackrel{\bigcirc}{=} \lim_{\Delta t \to 0^+} \frac{\langle x|P^{orig}|y\rangle/\Delta t}{\langle x|\mathbb{I} - \underbrace{P^{orig}}_{\text{unprojected}}|x\rangle/\Delta t} = \underbrace{-\frac{\langle x|W|y\rangle}{\langle x|W|x\rangle}}_{\geq 0}.$$

Classification of discrete-time processes

$$|p(t+m\Delta t)\rangle = \mathbb{P}^m |p(t)\rangle .$$

If it exists, the steady-state (ss) distribution is defined by

$$|p^{ss}\rangle = \lim_{m\to\infty} \mathbb{P}^m |p(0)\rangle .$$

State y is accessible from state x:
$x \to y$ if $\langle y|\Pi(t)|x\rangle > 0$ for finite t.

If both $x \to y$ and $y \to x$, x & y are in the same communication class.

It is possible to break all N_s states into non-overlapping communication classes.
If there is a single communication class, the stochastic process is irreducible, otherwise it is reducible. A class is called absorbing if it is impossible to leave it.

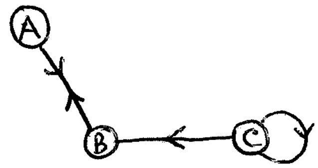
Markov processes on finite spaces can be represented by graphs: states $\equiv$ nodes,
directed edges $x \to y \equiv \langle y|P|x\rangle > 0$

$$x \to x \equiv \langle x|P|x\rangle > 0$$

$\uparrow$ 'sticky' state

Ex.

$$P = \begin{array}{c} \begin{array}{ccc} A & B & C \end{array} \\ \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1/2 \\ 0 & 0 & 1/2 \end{pmatrix} \end{array} \begin{array}{c} A \\ B \\ C \end{array}$$



Communication classes: $\underbrace{\{A, B\}}_{\text{absorbing class}}$ $\{C\}$

State C is sticky

<u>Periodicity</u>: the period of state x, $d(x)$, is an integer s.t. all closed paths $x \to \ldots \to x$ are multiples of $d(x)$.

If $d(x) > 1$, the state is *periodic*.

If $d(x) = 1$, the state is *aperiodic*.

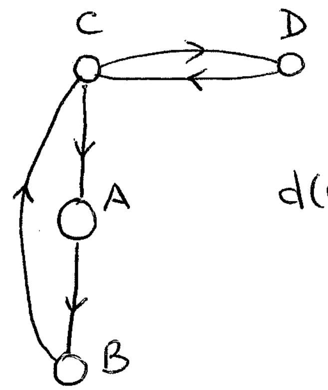
If $\langle x | P^m | x \rangle = 0$, $\forall m \Rightarrow d(x) = 0$ by definition

[Periodicity is the same for all nodes in the same communication class.]

If $\underbrace{\langle x | P | x \rangle > 0}_{\text{Sticky State}} \Rightarrow d(x) = 1$

In the example above, $d(A) = d(B) = 2$ and $d(C) = 1$.

Consider



$d(C) = 1 \Rightarrow d(\forall \text{ state}) = 1$
since there is a single communication class

→ State x is <u>transient</u> if there is a non-zero prob. to never return to x after starting from it.

→ State x is <u>recurrent</u> otherwise:

<u>positive recurrent</u> : mean return time is finite

<u>null recurrent</u> —"— is infinite

In the 3-node example above, A & B are positive recurrent while C is transient.

If a class is aperiodic and pos. recurrent, it is <u>strongly ergodic</u>.

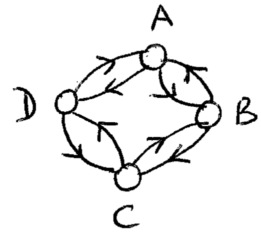
If a class is aperiodic and null recurrent, it is <u>weakly ergodic</u>.

The above classifications are useful in classifying SS behavior.

For finite systems (with a single comm. class), there are only 2 possibilities:

(a) pos. recurrent + periodic

Ex.



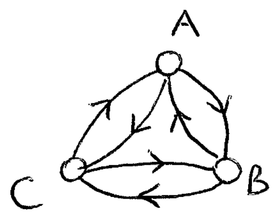
$\mathcal{N}_S = 4$ (even)

$d(A) = 2$, etc.

Oscillatory SS

(b) pos. recurrent + aperiodic

Ex.



$\mathcal{N}_S = 3$ (odd)

$d(A) = 1$, etc.

Unique SS

For infinite systems, SS does not exist unless the system is effectively confined.

For example, 1D symmetric random walk (RW) is null recurrent with (aperiodic) or without (periodic) sticking in the $\mathcal{N}_S \to \infty$ limit.

In both cases, there is no SS.

However, if the symmetric RW is in a square or parabolic potential well, it has a unique SS with sticking.

Spectral expansions

Consider eigenvectors & eigenvalues of the matrix $\mathbf{P}$ describing a discrete-time process. Usually, P is not symmetric: left & right eigenvectors $\langle \lambda_j |$ & $| \lambda_i \rangle$ are not the same. Eigenvectors: $\lambda_1, \ldots, \lambda_{N_s}$

may be complex

$$\begin{cases} P | \lambda_i \rangle = \lambda_i | \lambda_i \rangle , \\ \langle \lambda_j | P = \langle \lambda_j | \lambda_j . \end{cases}$$

Moreover, $\underbrace{\langle \lambda_j | \lambda_i \rangle = \delta_{ij}}_{\text{orthonormality}}$

Spectral expansion: $P = \sum_{j=1}^{N_s} \lambda_j | \lambda_j \rangle \langle \lambda_j |$, yielding

$$P^m = \sum_{j=1}^{N_s} \lambda_j^m | \lambda_j \rangle \langle \lambda_j |$$

Eigenvalues: $|\lambda_j| \leq 1$, $\forall j$

at least one λ (λ_1 if ranked by $|\lambda_j|$) is always $= 1$: $\underline{\underline{\lambda_1 = 1}}$

Suppose $\lambda_1 = 1$, $|\lambda_j| < 1$ $j = 2, \ldots, N_s$

Then, if P is not symmetric, the left eigenvector corresponding to λ_1 is

$$\underbrace{\langle \lambda_1 |}_{\text{each element } = 1} = \sum_x \langle x |$$

The right eigenvector is the steady state, as can be seen below:

$$\overset{\uparrow}{|p^{ss}\rangle}$$

$$P^m = \sum_x |p^{ss}\rangle\langle x| + \sum_{j=2}^{N_s} \lambda_j^m |\lambda_j\rangle\langle\lambda_j| .$$

Clearly, $\lim_{m\to\infty} P^m = \sum_x |p^{ss}\rangle\langle x|$, s.t.

$$\langle y | \Pi(\infty) | x \rangle = \underbrace{p^{ss}(y)}_{\text{steady-state prob. of state } y}$$

If P is symmetric, $|\lambda_i\rangle$ and $\langle\lambda_i|$ are simply transposes of one another and

$$|\langle x | \lambda_1 \rangle|^2 = p^{ss}(x) .$$

In the continuous case,

$$e^{tW} = \sum_{j=1}^{N} e^{\mu_j t} |\mu_j\rangle\langle\mu_j| \quad \text{and}$$

in 'well-behaved' cases, $\mu_1 = 0$ and

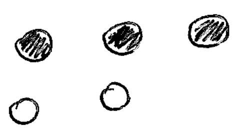
$|\mu_i| < 0 \qquad i = 2, \dots, N_s$

The left eigenvector is again $\underbrace{\sum_x \langle x|}_{\text{all 1's}}$, s.t.

$$\lim_{t\to\infty} e^{tW} = \sum_x |p^{ss}\rangle\langle x| \,,$$ similar to the discrete case

---

Ex.



3 red balls
2 white balls

State space: $\{ A[RR],\ B[WWR];$ $\quad s_3$
$\quad A[RW],\ B[RRW];$ $\quad s_2$
$\quad A[WW],\ B[RRR] \}$ $\quad s_1$

$N_s = 3$

Transitions between states:
pick a ball out of A and a ball out of B at random & interchange them.

$$\mathbb{P} = \begin{array}{c} \\ \\ \\ \end{array}\begin{array}{c} \begin{array}{ccc} s_1 & s_2 & s_3 \end{array} \\ \begin{pmatrix} 0 & 1/3 & 1/3 \\ 1 & \frac{1}{2} & 2/3 \\ 0 & 1/6 & 0 \end{pmatrix} \end{array} \begin{array}{c} \\ s_1 \\ s_2 \\ s_3 \end{array}$$

eigenvalues: $\lambda_1 = 1, \quad \lambda_2 = 1/6, \quad \lambda_3 = -1/3$

eigenvectors:
$$\begin{cases} \langle \lambda_1| = \overline{\underline{1 \ 1 \ 1}} \\ \langle \lambda_2| = \overline{\underline{-\tfrac{3}{2} \ -\tfrac{1}{4} \ 1}} \\ \langle \lambda_3| = \overline{\underline{3 \ -1 \ 1}} \end{cases}$$

$$|\lambda_1\rangle = \begin{pmatrix} 1/10 \\ 6/10 \\ 3/10 \end{pmatrix} \qquad |\lambda_2\rangle = \begin{pmatrix} -4/15 \\ -4/15 \\ 8/15 \end{pmatrix}$$

$$|\lambda_3\rangle = \begin{pmatrix} 1/6 \\ -1/3 \\ 1/6 \end{pmatrix}$$

As $m \to \infty$, the steady-state prob. is given by $|p^{ss}\rangle = |\lambda_1\rangle = \begin{pmatrix} 1/10 \\ 6/10 \\ 3/10 \end{pmatrix}$

For ex., the prob. to find two white balls in container A is $\underline{\underline{1/10}}$.

Detailed balance

Consider SS in the continuous system:

$$(+) \qquad \sum_{y \neq x} \left[\langle x|W|y\rangle p^{ss}(y) - \langle y|W|x\rangle p^{ss}(x) \right] = 0$$

Eq. (+) is trivially satisfied if

$$(++) \qquad \underbrace{\langle x|W|y\rangle p^{ss}(y) = \langle y|W|x\rangle p^{ss}(x)}_{\text{detailed balance condition}}.$$

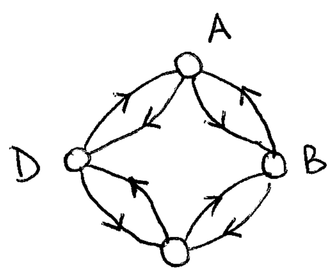
Since $\langle x|W|y\rangle p(y,t)$ is the $y \to x$ probability flux at time t,

Eq. (++) implies that all fluxes balance pairwise $\Rightarrow$ no net flux (or current) in the system.

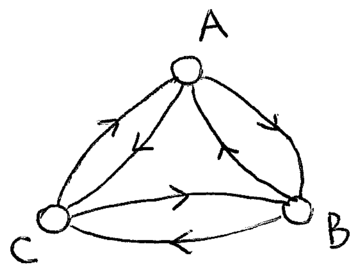
However, it is possible to have a SS with a net current.

Systems with detailed balance exhibit time reversibility: the prob. of the forward transition is equal to the prob. of the reverse transition.

HW analyze $N_s = 3$ and $N_s = 4$ circular RWs using spectral methods:



$N_s = 4$



$N_s = 3$

(a) what are the eigenvalues and eigenvectors in both cases?

(b) Compute SS for both systems. which system exhibits a unique SS with detailed balance? which system supports a net current at SS?

Self-adjoint symmetry

Detailed balance can be used to find a spectral decomposition for W or e^{tW}:

convenient

define a matrix $V = \Lambda^{-1} W \Lambda$, where

$$\langle x|\Lambda|y\rangle = \delta_{xy}\sqrt{p^{ss}(x)} .$$

Λ, Λ^{-1} are diagonal

Then $\langle x|V|y\rangle = \sum_{z,w} \langle x|\Lambda^{-1}|z\rangle\langle z|W|w\rangle\langle w|\Lambda|y\rangle =$

$$= \sum_{z,w} \frac{\delta_{xz}}{\sqrt{p^{ss}(x)}} \langle z|W|w\rangle \delta_{wy}\sqrt{p^{ss}(y)} =$$

$$= \sqrt{\frac{p^{ss}(y)}{p^{ss}(x)}} \langle x|W|y\rangle .$$

Now, $\langle y|V|x\rangle = \sqrt{\frac{p^{ss}(x)}{p^{ss}(y)}} \underbrace{\langle y|W|x\rangle}_{\langle x|W|y\rangle \frac{p^{ss}(y)}{p^{ss}(x)}}$ $=$

by detailed balance

$$= \sqrt{\frac{p^{ss}(y)}{p^{ss}(x)}} \langle x|W|y\rangle =$$

$$= \langle x|V|y\rangle .$$

Thus, V is symm. by construction.

Now, $\begin{cases} |\lambda\rangle_v = \sum\limits_x a_\lambda(x)\,|x\rangle, \\ \langle\lambda|_v = \sum\limits_x a_\lambda(x)\,\langle x| \end{cases}$

$$\langle\lambda|\lambda'\rangle_v = \sum_x a_\lambda(x)\,a_{\lambda'}(x) = \underbrace{\delta_{\lambda\lambda'}}_{\text{normal'n}}$$

Then $\underline{\underline{e^{tV} = \sum\limits_\lambda e^{\lambda t}\,|\lambda\rangle_v\langle\lambda|_v}}$

Since $e^{tW} = \Lambda\, e^{tV}\Lambda^{-1}$,

$$\langle x|e^{tW}|y\rangle = \langle x|\Lambda e^{tV}\Lambda^{-1}|y\rangle =$$

$$= \sum_\lambda e^{\lambda t}\,\langle x|\Lambda|\lambda\rangle_v\,\langle\lambda|_v\Lambda^{-1}|y\rangle =$$

$$= \underline{\underline{\sum_\lambda e^{\lambda t}\sqrt{\frac{p^{ss}(x)}{p^{ss}(y)}}\,a_\lambda(x)\,a_\lambda(y)}}.$$

Finally, $p(x,t) = \sum\limits_y \langle x|e^{tW}|y\rangle\, p(y,0) =$

$$= \sum_y p(y,0)\sum_\lambda \sqrt{\frac{p^{ss}(x)}{p^{ss}(y)}}\,a_\lambda(x)\,a_\lambda(y)\,e^{\lambda t} =$$

$$= \underline{\underline{\sum_\lambda \sqrt{p^{ss}(x)}\,a_\lambda(x)\,e^{\lambda t}\sum_y \frac{a_\lambda(y)}{\sqrt{p^{ss}(y)}}\,p(y,0)}}.$$

$\underline{\underline{\text{Ex.}}}$

$$\begin{tikzcd} \overset{A}{\bigcirc} \arrow[r, bend left, "a"] & \overset{B}{\bigcirc} \arrow[l, bend left, "b"] \end{tikzcd}$$

detailed balance trivially obeyed:

$a, b > 0$

$$W = \begin{array}{cc} \begin{array}{cc} A & B \end{array} & \\ \begin{pmatrix} -a & b \\ a & -b \end{pmatrix} & \begin{array}{c} A \\ B \end{array} \end{array}$$

$$p_a^{ss}\, a = p_b^{ss}\, b, \text{ or}$$

$$\frac{p^{ss}_a}{p^{ss}_b} = \frac{b}{a}.$$

$$p_a^{ss} + p_b^{ss} = 1 \Rightarrow |p^{ss}\rangle = \frac{1}{a+b}\begin{pmatrix} b \\ a \end{pmatrix}$$

Then
$$\Lambda = \frac{1}{\sqrt{a+b}}\begin{pmatrix} \sqrt{b} & 0 \\ 0 & \sqrt{a} \end{pmatrix} \quad \text{and}$$

$$V = \Lambda^{-1} W \Lambda = \begin{pmatrix} \frac{1}{\sqrt{b}} & 0 \\ 0 & \frac{1}{\sqrt{a}} \end{pmatrix}\begin{pmatrix} -a & b \\ a & -b \end{pmatrix}\begin{pmatrix} \sqrt{b} & 0 \\ 0 & \sqrt{a} \end{pmatrix} =$$

$$= \begin{pmatrix} \frac{1}{\sqrt{b}} & 0 \\ 0 & \frac{1}{\sqrt{a}} \end{pmatrix}\begin{pmatrix} -a\sqrt{b} & b\sqrt{a} \\ a\sqrt{b} & -b\sqrt{a} \end{pmatrix} = \underbrace{\begin{pmatrix} -a & \sqrt{ab} \\ \sqrt{ab} & -b \end{pmatrix}}_{\text{symm. matrix}}$$

For V, $\lambda_1 = 0$ & $\lambda_2 = -a-b$ and

$$\begin{cases} |\lambda_1\rangle_V = \frac{1}{\sqrt{a+b}}\begin{pmatrix} \sqrt{b} \\ \sqrt{a} \end{pmatrix}, \\ |\lambda_2\rangle_V = \frac{1}{\sqrt{a+b}}\begin{pmatrix} -\sqrt{a} \\ \sqrt{b} \end{pmatrix} \end{cases} \qquad \begin{aligned} \langle\lambda_1|\lambda_1\rangle_V &= 1, \\ \langle\lambda_2|\lambda_2\rangle_V &= 1, \\ \langle\lambda_1|\lambda_2\rangle_V &= 0 \end{aligned}$$

Finally,
$$e^{tW} = \begin{pmatrix} \frac{b}{a+b} + \frac{a}{a+b} e^{-(a+b)t} & \sqrt{\frac{b}{a}}\left(\frac{\sqrt{ab}}{a+b} - \frac{\sqrt{ab}}{a+b} e^{-(a+b)t}\right) \\ \sqrt{\frac{a}{b}}\left(\frac{\sqrt{a}}{\sqrt{a+b}}\frac{\sqrt{b}}{\sqrt{a+b}} - \frac{\sqrt{a}\sqrt{b}}{a+b} e^{-(a+b)t}\right) & \frac{a}{a+b} + \frac{b}{a+b} e^{-(a+b)t} \end{pmatrix} \Leftrightarrow$$

$$\Leftrightarrow \frac{1}{a+b}\begin{pmatrix} b+a e^{-(a+b)t} & b(1-e^{-(a+b)t}) \\ a(1-e^{-(a+b)t}) & a+b e^{-(a+b)t} \end{pmatrix}.$$

$$|p(t)\rangle = e^{tW}|p(0)\rangle$$

In the $t\to\infty$ limit,

$\alpha+\beta=1$

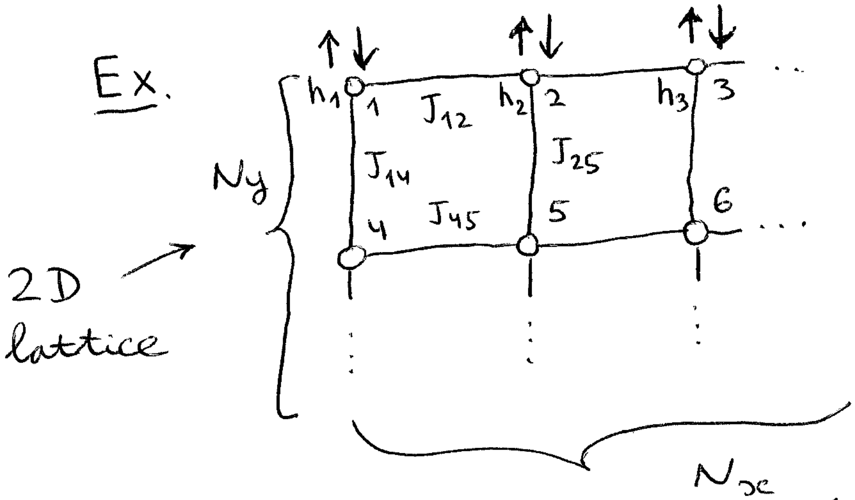
$$|p^{ss}\rangle = \frac{1}{a+b}\begin{pmatrix} b & b \\ a & a \end{pmatrix}\begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \frac{1}{a+b}\begin{pmatrix} b \\ a \end{pmatrix}$$

as expected; regardless of $\underline{\underline{\alpha \,\&\, \beta}}$.

[Case study: stochastic sampling of energy / fitness landscapes]

Goal: estimate Boltzmann-weighted averages of various quantities of interest by random sampling.

Ex.



$\overset{= N_x N_y}{N}$ spins represented by discrete variables: $\sigma_k \in \{\overset{\downarrow}{-1}, \overset{\uparrow}{1}\}$

$k = 1, \dots, N$

Spin configuration: $\sigma = \{\sigma_1, \sigma_2, \dots, \sigma_N\}$

The energy of a spin configuration is given by

$$E(\sigma) = -\sum_{\langle ij \rangle} \underbrace{J_{ij}}_{\text{spin couplings}} \sigma_i \sigma_j - \sum_j \underbrace{h_j}_{\text{external magnetic field}} \sigma_j$$

$\langle ij \rangle$: sum over all pairs of nearest neighbors; each pair is counted once

Common models of couplings: ➀ J_{ij} are randomly sampled from a prob. distribution

➁ $J_{ij} = \underbrace{J}_{= \text{const}} > 0$

In general, $J_{ij} > 0$ induces ferromagnetic order: spins which are all up or all down will have lower $E(\sigma)$.

→ $J_{ij} = 0$ (no couplings, indep. spins)

Magnetic field: → $h_j > 0$ spins up favored
→ $h_j < 0$ spins down favored

special case:
$h_j = h$ ← const magnetic field
→ $h_j = 0$ no magnetic field

In the canonical ensemble of statistical mechanics, it can be shown that at thermal equilibrium with temperature T, the prob. of spin configurations is given by

$$P(\sigma) = \frac{e^{-E(\sigma)/k_B T}}{\sum_{\sigma} e^{-E(\sigma)/k_B T}}$$

, where k_B is the Boltzmann constant and $Z = \sum_{\sigma} e^{-E(\sigma)/k_B T}$ called is the normalization constant called the partition function. $e^{-E(\sigma)/k_B T}$ is called the Boltzmann factor.

We are interested in thermal averages

$$\langle f \rangle = \sum_{\sigma} f(\sigma) P(\sigma) ,$$ where

$f(\sigma)$ is some function of spins.

For ex., if $f(\sigma) = \sigma_i \Rightarrow \langle f \rangle = \underbrace{\langle \sigma_i \rangle}_{\text{thermal average of spin } i}$

If $f(\sigma) = \sigma_i \sigma_j \Rightarrow \langle f \rangle = \underbrace{\langle \sigma_i \sigma_j \rangle}_{\text{correlation function for spins } i,j}$

~~How~~ to compute thermal averages by random sampling? If the stochastic process is in steady state,

$$p_i^{ss} = \frac{e^{-\beta E_i}}{Z} \qquad \beta = \frac{1}{k_B T}$$

$i = 1, \ldots, 2^N$ labels spin configurations

will guarantee that

$$\sum_{i=1}^{M} f_i = \langle f \rangle$$

M ← # steps in the stochastic process at SS

We need to choose a move set - a natural one is to flip one randomly chosen spin.

move set ← a way to generate a new spin configuration

Then the master equation is

$$p_i(t+1) = p_i(t) + \sum_{j \in nn_i} P(j \to i)\, p_j(t) - p_i(t) \sum_{j \in nn_i} P(i \to j)$$

$t+1$ ← discrete time steps

$$\boxed{\sum_{j \in nn_i} P(i \to j) \le 1}$$

at SS,

$$\sum_{j \in nn_i} P(j \to i)\, p_j^{ss} = p_i^{ss} \sum_{j \in nn_i} P(i \to j)$$

This eq'n is trivially satisfied if detailed balance holds:

$$P(j \to i)\, p_j^{ss} = P(i \to j)\, p_i^{ss} \quad \forall j \in nn_i \text{ , or}$$

$$\frac{p_j^{ss}}{p_i^{ss}} = \frac{P(i \to j)}{P(j \to i)} = e^{-\beta(E_j - E_i)} \quad \text{for Boltzmann probs.}$$

Consider $P(i \to j) = \underbrace{T(i \to j)}_{\text{prob. of transition}} \underbrace{A(i \to j)}_{\text{prob. of acceptance}}$

Focus on the case when the set of nearest neighbors nn_i has the same size N_B for $\forall i$:

$$T(i \to j) = T(j \to i) = \frac{1}{\underbrace{N_B}_{\text{prob. to choose a neighbor}}}$$

Then $\dfrac{P(i \to j)}{P(j \to i)} = e^{-\beta(E_j - E_i)}$ yields

$$\frac{A(i \to j)}{A(j \to i)} = e^{-\beta(E_j - E_i)} \qquad (+)$$

In general,

$$A(i \to j) = \begin{cases} e^{-\beta(1-\overset{\text{const}}{\overset{\shortparallel}{c}})(E_j - E_i)}, & E_j > E_i \\ e^{-\beta c(E_j - E_i)}, & E_j \leq E_i \end{cases}$$

satisfies Eq. (+):

$$\frac{A(i \to j)}{A(j \to i)} \underset{\substack{\uparrow \\ E_j > E_i}}{=} \frac{e^{-\beta(1-c)(E_j - E_i)}}{e^{-\beta c(E_i - E_j)}} = e^{-\beta(E_j - E_i)}$$

$$\frac{A(i \to j)}{A(j \to i)} \underset{\substack{\uparrow \\ E_j < E_i}}{=} \frac{e^{-\beta c(E_j - E_i)}}{e^{-\beta(1-c)(E_i - E_j)}} = e^{-\beta(E_j - E_i)}$$

If $c > 0$, $e^{-\beta c(E_j - E_i)} > 1$ if $E_j < E_i$

and if $E_j < E_i$ for $\forall j \in nn_i$:

$$\sum_{j \in nn_i} P(i \to j) = \frac{1}{\mathcal{N}_B} \sum_{j \in nn_i} \underbrace{e^{-\beta c(E_j - E_i)}}_{>1} > 1,$$ not allowed

If $c < 0$, $$A(i \to j) = \begin{cases} e^{-\beta(1+|c|)(E_j - E_i)}, & E_j > E_i \\ e^{\beta|c|(E_j - E_i)}, & E_j \leq E_i \end{cases}$$

$e^{\beta|c|(E_j - E_i)} < 1$ for $E_j < E_i$ $\Rightarrow$

$\Rightarrow \sum_{j \in nn_i} P(i \to j) \leq 1$ satisfied

However, the acceptance prob. ↓ as $|c| \uparrow$, leading to slower approach to the SS. Thus, it makes sense to choose $|c| = 0 \Rightarrow$ fastest kinetics that still respects the prob. normalization.

This leads to:

$$A(i \to j) = \begin{cases} e^{-\beta(E_j - E_i)}, & E_j > E_i \\ 1, & E_j \leq E_i \end{cases}$$

↗ Metropolis MC

Another prescription would be to normalize the acceptance probs. explicitly:

$$A(i \to j) = \frac{e^{-\beta E_j}}{e^{-\beta E_i} + e^{-\beta E_j}} \leq 1$$

gibbs sampling ↗

Then $\dfrac{A(i \to j)}{A(j \to i)} = \dfrac{e^{-\beta E_j}}{e^{-\beta E_i}} = e^{-\beta(E_j - E_i)}$.

In general, $A(i \to j) = \dfrac{e^{-\beta(c_1 E_i + c_2 E_j)}}{Z}$, where

$$Z = e^{-\beta(c_1 E_i + c_2 E_j)} + e^{-\beta(c_1 E_j + c_2 E_i)}.$$

Detailed balance: $\dfrac{A(i \to j)}{A(j \to i)} = \dfrac{e^{-\beta(c_1 E_i + c_2 E_j)}}{e^{-\beta(c_1 E_j + c_2 E_i)}} =$

$$= e^{-\beta[\underbrace{(c_2 - c_1)}_{\overset{!}{=}1} E_j + \underbrace{(c_1 - c_2)}_{\overset{!}{=}-1} E_i]}.$$

Thus, $\begin{cases} C_2 = \alpha + 1, \\ C_1 = \alpha . \end{cases}$

But then $$A(i \to j) = \frac{\cancel{e^{-\beta\alpha(E_i+E_j)}}\, e^{-\beta E_j}}{\cancel{e^{-\beta\alpha(E_i+E_j)}}\, e^{-\beta E_j} + \cancel{e^{-\beta\alpha(E_i+E_j)}}\, e^{-\beta E_i}} =$$

$$= \frac{e^{-\beta E_j}}{e^{-\beta E_i} + e^{-\beta E_j}}$$ as before.

Hence we can set $\alpha = 0$ from the beginning, recovering standard Gibbs sampling.

(HW) Consider a RW on an arbitrary energy landscape according to Metropolis MC rules. assume that $T(i \to j) = T(j \to i) = \frac{1}{N_B}$. Consider a projected RW which consists only of <u>accepted</u> steps:

● ● ● ○ ○ ⊕ ⊕ ⊕ ... original RW

$\Downarrow$

● ○ ⊕ ... projected RW

What are the steady-state probabilities p_i^{ss} for the <u>projected</u> RW?

Entropy and the approach to SS

How does the system approach the steady state? Consider an irreducible process on a finite state space: no transient states, $\Rightarrow \exists$ unique steady state $p^{ss}_{(x)}$.

Define a function $f(k)$ s.t. $f(k) \geq 0$ and $f''(k) > 0$ for $\forall k \geq 0$: $f(k)$ is a non-negative convex function.

Now, define a function

$$H(t) = -\sum_x p^{ss}(x)\, f(x,t) \quad , \quad \text{where}$$

$$f(x,t) = f\left(+\frac{p(x,t)}{p^{ss}(x)}\right) \equiv f(k_{x,t})$$

Note that $H(t) \leq 0$, $\forall t$ by construction.

Consider $\partial_t H(t) = -\sum_x p^{ss}(x)\, f'(k_{x,t})\, \partial_t k_{x,t} =$

$= -\sum_x f'(k_{x,t})\, \partial_t p(x,t) = -\sum_x f'(k_{x,t}) \times$

$\times \left[\sum_y \left(\langle x|W|y\rangle p(y,t) - \langle y|W|x\rangle p(x,t)\right)\right] =$

$$= -\sum_{x,y} f'(k_{x,t}) \langle x|W|y\rangle p(y,t) + \underset{x\leftrightarrow y}{+} \sum_{x,y} f'(k_{y,t}) \langle x|W|y\rangle p(y,t) \;=$$

$$(\equiv)\; -\sum_{x,y} \left[f'(k_{x,t}) - f'(k_{y,t}) \right] \langle x|W|y\rangle \underbrace{p(y,t)}_{= \; +k_{y,t}\, p^{ss}(y)} =$$

$$= -\sum_{x,y} \left[k_{y,t} f'(k_{x,t}) - k_{y,t} f'(k_{y,t}) \right] \times p^{ss}(y) \langle x|W|y\rangle .$$

Now, for any function $a_{x,t}$:

$$(1) \quad \sum_{x,y} a_{x,t} \langle x|W|y\rangle p^{ss}(y) = \sum_x a_{x,t} \sum_y \langle x|W|y\rangle p^{ss}(y)$$

$$(2) \quad \sum_{x,y} a_{y,t} \langle x|W|y\rangle p^{ss}(y) = \sum_y a_{y,t} \sum_x \underbrace{p^{ss}(y) \langle x|W|y\rangle}_{= \; p^{ss}(x)\langle y|W|x\rangle}$$

by detailed balance

Clearly, (2) $\to$ (1) under $x \leftrightarrow y$, yielding

$$\sum_{x,y} (a_{y,t} - a_{x,t}) \langle x|W|y\rangle p^{ss}(y) = 0 .$$

↑ add this to the expression for $\partial_t H$ and use

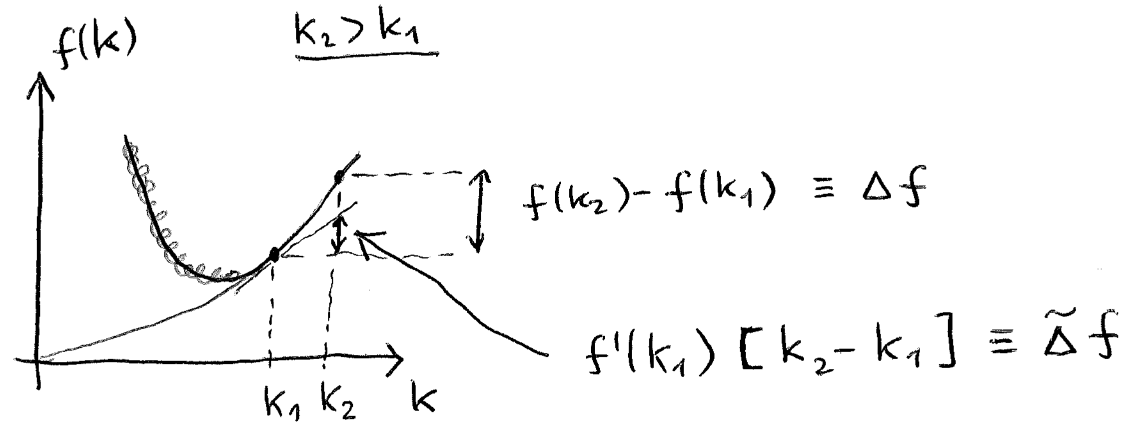
$$a_{x,t} = f(k_{x,t}) - k_{x,t} f'(k_{x,t}) :$$

$$\partial_t H = -\sum_{x,y} \left[k_{y,t} f'(k_{x,t}) - k_{y,t} f'(k_{y,t}) \right] p^{ss}(y) +$$

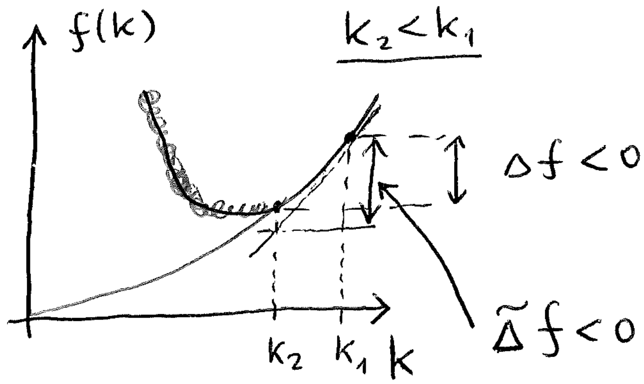
$$+ \sum_{x,y} \left[\left(f(k_{y,t}) - k_{y,t} f'(k_{y,t}) \right) - \left(f(k_{x,t}) - k_{x,t} f'(k_{x,t}) \right) \right] \langle x|W|y\rangle p^{ss}(y) =$$

$$= \sum_{x,y} \left\{ \left[f(k_{y,t}) - f(k_{x,t}) \right] + (k_{x,t} - k_{y,t}) f'(k_{x,t}) \right\} \times \langle x|W|y\rangle p^{ss}(y) .$$

Now, consider $f(k_2) - f(k_1) - [k_2 - k_1] f'(k_1)$ and recall that $f(k)$ is non-negative & convex:



Clearly, $\Delta f > \tilde{\Delta} f$



$\Delta f - \tilde{\Delta} f > 0$ again

~~etc~~

Thus, the quantity in $\{ \ldots \}$ is always $\geq 0 \Rightarrow \partial_t H \geq 0$ as well.

But $H(t) \leq 0$, $\forall t \Rightarrow$ must have $H(t) \to 0$ as $t \to \infty$

i.e., nnb) $[\partial_t H \to 0$ as $t \to \infty]$

For x, y s.t. $\langle x | W | y \rangle \neq 0$, $\partial_t H = 0$ implies that $k_{x,t} = k_{y,t} \Rightarrow k_{x,t} = \text{const}$, $\forall x$

for all states in the single communication class (irreducibility assumed).

But $k_{x,t} = + \frac{p(x,t)}{p^{ss}(x)}$ $\Rightarrow$ we must have

$$p(x,\infty) = \alpha \, p^{ss}(x)$$

$\uparrow$ const, must be $=1$ by normalization

Thus, $\lim_{t\to\infty} p(x,t) = p^{ss}(x)$ corresponds to the max of $H(t)$: $\lim_{t\to\infty} H(t) = 0$.

In physics, we use $f(k) = k \log k$, or

$$f(k_{x,t}) = + \frac{p(x,t)}{p^{ss}(x)} \log \frac{p(x,t)}{p^{ss}(x)},$$ yielding

$$\left[H(t) = - \sum_x p(x,t) \log\left(\frac{p(x,t)}{p^{ss}(x)} \right). \right]$$

Note that $k \log k \geq 0$ if $k \geq 0$;

$f'(k) = \log k + 1$, $f''(k) = \frac{1}{k} > 0$ if $k \geq 0$.

$H(t)$ defined above is extensive: for indep. subsystems

$$H_{1+2}(t) = - \sum_{x \in S_1} \sum_{y \in S_2} p_1(x,t)\, p_2(y,t) \times$$

$$\times \log\left(\frac{p_1(x,t)\, p_2(y,t)}{p_1^{ss}(x)\, p_2^{ss}(y)} \right) =$$

$$= - \sum_{x \in S_1} p_1(x,t) \log\left(\frac{p_1(x,t)}{p_1^{ss}(x)} \right) \ominus$$

$$\ominus \sum_{y \in S_2} p_2(y,t) \log\left(\frac{p_2(y,t)}{p_2^{ss}(y)}\right) = H_1(t) + H_2(t).$$

Finally, note that

$$H(t) = \underbrace{-D_{KL}(p(x,t) \,\|\, p^{ss}(x))}_{\text{Kullback-Leibler distance } (\geq 0) \text{ between } p(x,t) \ \& \ p^{ss}(x)}$$

The deterministic limit

Consider $\langle F \rangle_t = \sum_{\substack{x \\ x \in S}} F(x)\, p(x,t)$

$\uparrow$ a time-dependent average of some state function $F(x)$

Now, $\partial_t \langle F \rangle_t = \sum_x F(x)\, \partial_t p(x,t) =$

$$= \sum_x F(x) \left[\sum_y \{ \langle x|W|y\rangle p(y,t) - \langle y|W|x\rangle p(x,t) \} \right] =$$

$x \leftrightarrow y$ in the 2nd term

$$\stackrel{\downarrow}{=} \sum_{x,y} [F(x) - F(y)] \langle x|W|y\rangle p(y,t) =$$

$$= \sum_y m(y)\, p(y,t), \quad \text{where}$$

$$m(y) = \sum_x [F(x) - F(y)] \langle x|W|y\rangle$$

are the "jump moments":

the change in the value of $F(y)$ due to jumps from $y \to x$, weighted by the jump rates $\langle x | W | y \rangle$.

Thus, $\partial_t \langle F \rangle_t = \sum_y m(y) p(y,t) = \langle m \rangle_t$

Next, consider $\langle m \rangle_t = \langle m(F(y), y) \rangle_t$

"$\langle m(y) \rangle_t$" dependence through $\langle x | W | y \rangle$

Expand $m(F(y), y)$ around $\langle y \rangle_t$:

$$m(F(y), y) \simeq m(F(\langle y \rangle_t), \langle y \rangle_t) + (y - \langle y \rangle_t) \left.\frac{dm}{dy}\right|_{y=\langle y \rangle_t} + \frac{1}{2}(y - \langle y \rangle_t)^2 \left.\frac{d^2 m}{dy^2}\right|_{y=\langle y \rangle_t} + \dots$$

But then

$$\langle m \rangle_t = m(\langle y \rangle_t) + \frac{1}{2} \langle (y - \langle y \rangle_t)^2 \rangle \left.\frac{d^2 m}{dy^2}\right|_{\langle y \rangle_t} + \dots$$

Likewise, $\langle F(y) \rangle_t \simeq F(\langle y \rangle_t) + \frac{1}{2} \langle (y - \langle y \rangle_t)^2 \rangle \left.\frac{d^2 F}{dy^2}\right|_{\langle y \rangle_t} + \dots$

To the lowest order,

$$\partial_t F(\langle y \rangle_t) = m(\langle y \rangle_t)$$

a generally non-linear equation for $\langle y \rangle_t$

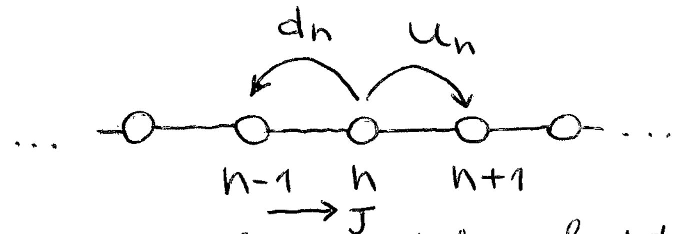
It's clear from the above that to the lowest order we just need an equation for the evolution of $\langle y \rangle_t$ itself rather than $F(\langle y \rangle_t)$ - the dynamics should be consistent for any $F(\cdot)$.

That is given by

$$\partial_t \langle y \rangle_t = m(\langle y \rangle_t) = \sum_x (x - \langle y \rangle_t) \underbrace{\langle x | W | \langle y \rangle_t \rangle}_{\text{need some approx. for the 'average rate'}}$$

Markov processes on 1D lattices

Consider a continuous-time stochastic process where jumps are only allowed between nearest neighbors on the lattice:



Assume that the lattice is finite: $n = 0, 1, \dots, N$

Then
$$\begin{cases} \langle n+1 | W | n \rangle = u_n, \\ \langle n-1 | W | n \rangle = d_n, \\ \langle n | W | n \rangle = -u_n - d_n \end{cases}$$
all other elements of W are $\emptyset$

Thus, W is tri-diagonal.

The master equation is given by:

$$\partial_t p_n(t) = \underbrace{u_{n-1}\, p_{n-1}(t) + d_{n+1}\, p_{n+1}(t)}_{\text{gain}} - \underbrace{(u_n + d_n)\, p_n(t)}_{\text{loss}},$$

where $p_n(t)$ = prob. to be at site n at time t

Define operators:

$$\begin{cases} a^+ f(n) = f(n+1) & \text{creation} \\ a\, f(n) = f(n-1) & \text{annihilation} \end{cases}$$

Then $\partial_t p_n(t) = [(a^+-1)d_n + (a-1)u_n]\,p_n(t)$,

where $a^+ d_n p_n(t) = d_{n+1}\,p_{n+1}(t)$, etc.

At SS we have:

$$[(a^+-1)d_n + (a-1)u_n]\,p_n^{ss} = 0 \;, \text{ or}$$

$$(a^+-1)(d_n - a u_n)\,p_n^{ss} = 0 .$$

$\uparrow$ $a^+ a\, f(n) = f(n)$

The quantity $(d_n - a u_n)\,p_n^{ss} \equiv -J$ is a quantity that does not depend on n.

$J = u_{n-1}\,p_{n-1}^{ss} - d_n\,p_n^{ss}$ is in fact a current from $n-1$ to n (or $-$current from n to $n-1$).

Now, $n=0$ is at the boundary. If the boundary does not allow any current to/from it (no sinks or sources):

$$u_0\,p_0^{ss} - d_1\,p_1^{ss} = 0 \;(=J).$$

But then $u_{n-1}\,p_{n-1}^{ss} - d_n\,p_n^{ss} = 0 \;,\; \forall n$
since J is indep. of $\underline{\underline{n}}$.

Hence $\underbrace{u_{n-1}\,p_{n-1}^{ss} = d_n\,p_n^{ss}}_{\text{detailed balance, time reversibility @ SS}}$ $\quad (+)$

Eq. (+) is a recursion relation which yields: $$p_n^{ss} = \underbrace{\frac{u_{n-1} u_{n-2} \dots u_0}{d_n d_{n-1} \dots d_1} p_0^{ss}}.$$

p_0^{ss} can be obtained by normalization:
$$\sum_{n=0}^{N} p_n^{ss} = 1.$$

The Poisson process

It is a special case of the one-step process described above:

$$\begin{cases} \langle n+1 | W | n \rangle = \alpha, \\ \langle n-1 | W | n \rangle = 0, \\ \langle n | W | n \rangle = -\alpha \end{cases}$$

all other elements of W are $\emptyset$

$$\partial_t p_n(t) = \alpha [p_{n-1}(t) - p_n(t)]$$

This eq'n is satisfied by the Poisson distribution:

$$\underbrace{p_n(t) = \frac{(\alpha t)^n}{n!} e^{-\alpha t}} \quad \leftarrow \sum_{n=0}^{\infty} p_n(t) = 1$$

Indeed, $\partial_t p_n(t) = \frac{n\alpha (\alpha t)^{n-1}}{n!} e^{-\alpha t} - \alpha p_n(t) =$

$= \alpha p_{n-1}(t) - \alpha p_n(t)$, as desired.

Initial conditions: $p_n(0) = \delta_{n,0}$
[no particles @ $t=0$]

Note that $\langle n \rangle_t = \underset{\substack{\uparrow \\ \text{rate}}}{\alpha} t \nwarrow$ time interval

This result can be obtained directly from the moment equation:

$$\partial_t \left(\sum_{n=0}^{\infty} n\, p_n(t) \right) = \alpha \Big[\underbrace{\sum_{n=1}^{\infty} n\, p_{n-1}(t)} - \overbrace{\sum_{n=0}^{\infty} n\, p_n(t)}^{\langle n \rangle_t} \Big] \ominus$$

$$\sum_{n=1}^{\infty} (n-1)\, p_{n-1}(t) + \sum_{n=1}^{\infty} p_{n-1}(t) =$$

$$= \langle n \rangle_t + 1$$

$$\ominus\ \alpha \left[\langle n \rangle_t + 1 - \langle n \rangle_t \right] = \alpha .$$

Thus, $\partial_t \langle n \rangle_t = \alpha \Rightarrow \langle n \rangle_t = \alpha t$ as before

---

Symmetric random walk

$$\begin{cases} \langle n+1 | W | n \rangle = 1, \\ \langle n-1 | W | n \rangle = 1, \\ \langle n | W | n \rangle = -2 \end{cases}$$

$n = -\infty, \dots, -1, 0, 1, \dots, +\infty$ infinite lattice

$$\partial_t p_n(t) = p_{n+1}(t) + p_{n-1}(t) - 2\, p_n(t),$$

$$p_n(0) = \delta_{n,0} .$$

Use the generating function:

$$F(z,t) = \sum_{n=0}^{\infty} z^n p_n(t)$$

to obtain

$$\partial_t F(z,t) = (z+z^{-1}-2)\,F(z,t).$$

This is solved by

$$\underbrace{F(z,t) = C(z)\, e^{t(z+z^{-1}-2)}} \qquad (\ddagger)$$

$$\underbrace{F(z,0) = 1}_{``z^0=1} \implies F(z,t) = e^{t(z+z^{-1}-2)}.$$

contour is a unit circle

Since $p_n(t) = \frac{1}{2\pi i}\oint dz\, z^{-n-1} F(z,t) =$

$$= e^{-2t}\frac{1}{2\pi i}\oint dz\, z^{-n-1} e^{\underbrace{t(z+z^{-1})}_{2t\frac{z+z^{-1}}{2}}} =$$

$$= e^{-2t}\underbrace{I_n(2t)}.$$

modified Bessel function of the 1st kind

Indeed, $\frac{1}{2\pi i}\oint dz\, z^{-n-1}\sum_{n'=0}^{\infty} z^{n'} p_{n'}(t) \overset{z=e^{i\alpha}}{=}$

$$= \frac{1}{2\pi i}\int_0^{2\pi} d\alpha\,(iz)\, z^{-n-1}\sum_{n'=0}^{\infty} e^{i\alpha n'} p_{n'}(t) =$$

$$= \frac{1}{2\pi}\sum_{n'=0}^{\infty} p_{n'}(t)\underbrace{\int_0^{2\pi} d\alpha\, e^{i\alpha(n'-n)}}_{2\pi\delta_{n,n'}} =$$

$= p_n(t)$, as claimed above.

Thus,

$$p_n(t) = \frac{e^{-2t}}{2\pi} \int_0^{2\pi} d\alpha \, e^{-in\alpha} e^{2t\cos\alpha} .$$

In the $t \to \infty$ limit (strictly speaking, the diffusion limit : $t \to \infty$, $n \to \infty$ s.t. $\frac{n^2}{t}$ = finite), $\alpha \ll 1$ and

$$\cos\alpha \simeq 1 - \frac{\alpha^2}{2} .$$

indeed, $f(\alpha) = in\alpha - 2t\cos\alpha$,

$$\left.\frac{df}{d\alpha}\right|_{\alpha^*} = 0 = in + 2t\sin\alpha^*, \quad \text{or}$$

$$\underbrace{\sin\alpha^* = -\frac{in}{2t}},$$

Since $n/t \to 0$ as $t \to \infty$, $\alpha^* \to 0$ as well

Then
$$p_n(t) \simeq \frac{1}{2\pi} \int_{-\infty}^{\infty} d\alpha \, e^{-in\alpha - t\alpha^2} = \frac{1}{\sqrt{4\pi t}} e^{-\frac{n^2}{4t}} .$$

extend the limits

This is a Gaussian distribution with $\mu = 0$ & $\sigma^2 = 2t$. Note that

$$\int_{-\infty}^{\infty} dn \, p_n(t) = 1 , \quad \forall t .$$