IX Introduction to martingales

The non scientific meaning of "martingale" is twofold:

- for horses, a martingale refers to a device controlling the horse's head height: a little rope below the neck prevents the horse to throw head so high that it would hurt the rider
- in the context of betting games, it refers to a strategy to "control randomness" (in the same way as the horse is controlled ?), and to lead to a sure gain. As such, it only is a dream ! (see below, a "no dream" theorem exists).

For the mathematicians, martingales are an important and common tool; it is fair to say that "a probabilist is someone searching for martingales" (Marc Yor). Given its relevance in applied math (eg actuarial), it is quite surprising that this very notion is often ignored by physicists. Martingale theory allows to show that martingales, in the sense of "winning bet strategy", do not exist. This tool is of particular interest for first passage properties, extreme value statistics, and has implications in stochastic thermodynamics.

1°) Definition

Let $X(t)$ be a stochastic process. $M(t)$ is a martingale with respect to $X(t)$ if

$$\langle |M(t)| \rangle < \infty$$

$$\langle M(t) \,|\, X(t_1) \cdots X(t_n) \rangle = M(t_n) \qquad \text{for all } t_1 \leqslant t_2 \leqslant \cdots \leqslant t_n \leqslant t$$

Martingales are not restricted to Markov processes. They can be real or complex. Below, we do not attempt at rigor, but seek clarity: the goal is to give an idea of what this tool allows. Note that $\langle M(t) \rangle = \langle M(t_n) \rangle = \langle M(t_1) \rangle \ldots$

Examples

(*) The Wiener process is a martingale (and so is the discrete version, the symmetric random walk). For $t > t_n$: $X(t) = X(t_n) + \underbrace{\int_{t_n}^{t} \sqrt{2D}\, \xi(t')\, dt'}_{B_{t-t_n}}$

$$\langle X(t) \,|\, X(t_n) \rangle = X(t_n) \quad \Rightarrow \text{martingale}$$

and also, $\langle X(t) \rangle = $ cst, does not depend on t.

$\langle B \rangle = 0$

"Coup de boule" : headbutt

Ask "martingale?" in the street

a) Dunno!

b) Horse

c) Winning strategy (games / politics)

d) mathematician

If you ask a physicist: may know c), perhaps b), but presumably not d)

Martingales introduced by P. Lévy (1934), named by Ville (1939) and developped by J. Doob (See his book "Stochastic processes", 1952).

In Doob's bibliography, apart from Doob himself, W. Doblin is the third most cited mathematician (after Kolmogorov and Levy). Yet in 1952, Doblin's work on stochastic calculus in his "Pli cacheté", was unknown to the community.

(*) With Wiener $(X(t))$ again, $X^2(t) - 2Dt$ is a martingale.

$$X^2(t) = X^2(t_n) + B^2_{t-t_n} + 2X(t)\,B_{t-t_n}$$

$$\langle X^2(t) \mid X(t_n) \rangle = X^2(t_n) + 2D(t-t_n) + 0$$

$$\langle X^2(t) - 2Dt \mid X(t_n) \rangle = X^2(t_n - 2Dt_n) \quad \square$$

(*) The same trick applies to a Poisson distributed variable $N_t \rightsquigarrow \mathcal{P}(\lambda t)$

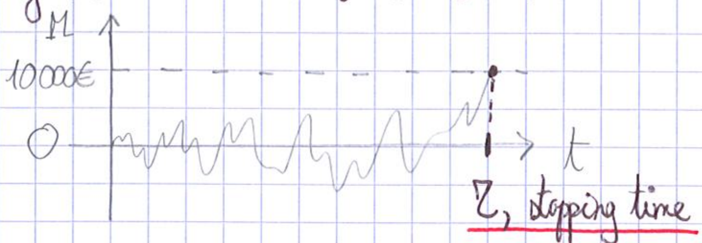
Then $N_t - \lambda t$ is a martingale (of 0 mean).

(*) Games of chance. Suppose a gambler with fortune f_1 plays some game of chance once and that his fortune after the game is f_2 → this is a r.variable and the game is considered *fair* if $\langle f_2 \mid f_1 \rangle = f_1$ ... and so on: $\langle f_n \mid f_{n-1} \rangle = f_n$.

More generally, if y_n represents the past history info avaible to the gambler (y_n is itself a collection of random variables), the fairness condition is

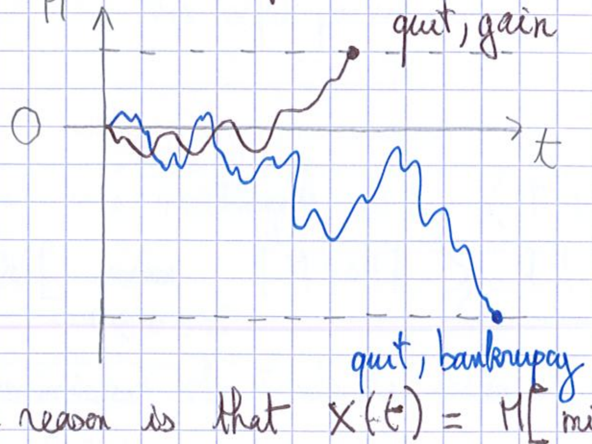
$\langle f_n \mid y_{n-1} \rangle = f_{n-1}$: a fair game is a martingale

Consider a fair game where on average, after each play, the gain M is 0. M is a martingale. Can a gambler make fortune playing such a game, with a clever strategy? One may think of defining a threshold, say 10 000 €, and quit the game when threshold is reached.



With a random walk having no bias, the threshold will be met with probability 1: $\langle M(\tau) \rangle = 10\,000$ € → the gambler makes profit!

But as such, the game is unrealistic: we need to add a constraint, that either the duration is finite, or that the gambler cannot accept arbitrary level of debt. Assume for instance that the gambler is not infinitely rich, and has a bankrupcy level (⇒ forced to quit). The game stops when one of the two thresholds is reached. We define again τ as the corresponding stopping time. One can show that $\langle M(\tau) \rangle = 0$



and the reason is that $X(t) = M[\min(t, \tau)]$ is a martingale.

Thus, a gambler that cannot foresee the future (and cannot cheat), and has finite budget, cannot make money by quitting a fair game at intelligently chosen moments ... meaning that the mathematical theory of martingales precludes "real-life" martingales.

2°) Main properties of martingales

Martingales enjoy a number of properties in terms of $\begin{cases}\text{convergence}\\ \text{inequalities}\\ \text{stopping times}\end{cases}$.

We only address briefly the first aspect, with a theorem stating that a positive martingale M_n converges with probability 1 to a limit M_∞ ; this limit can itself be random (ie, there may be no self averaging)

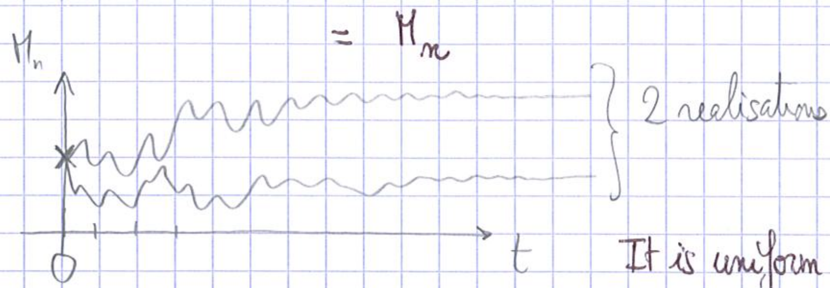
This is the case with Polya's urn : an urn contains W_0 white balls and B_0 black balls. A ball is drawn at random from the urn, and put back with another ball of the same color. After n draws, there are $W_n + B_n = W_0 + B_0 + n$ balls

$$M_n = \frac{W_n}{W_n + B_n} \quad \text{is a martingale}$$

$$\langle M_{n+1} \mid W_n, B_n \rangle = \frac{W_n + 1}{W_n + B_n + 1} \underbrace{\frac{W_n}{W_n + B_n}}_{\text{proba that a W is drawn}} + \frac{W_n}{W_n + B_n + 1} \underbrace{\left(1 - \frac{W_n}{W_n + B_n}\right)}_{\text{proba draw a Black}}$$

$$= \frac{W_n}{W_n + B_n + 1} \; \frac{W_n + 1 + B_n}{W_n + B_n}$$

$$= M_n$$



The law of M_∞ strongly depends on the initial composition of the urn

It is uniform on $[0,1]$ for $W_0 = B_0 = 1$

We will meet below other useful positive martingales.

a) Martingale inequality

Let $M(s)$ be a martingale s.t $\langle M^2(s) \rangle$ exists at all $0 \leqslant s \leqslant t$

$$\Pr\left[\max_{0 \leqslant s \leqslant t} |M(s)| \geqslant \lambda \right] \leqslant \frac{1}{\lambda^2} \langle M^2(t) \rangle$$

This is reminiscent of Tchebychev inequality for a r.v. X, with $\langle X \rangle = 0$:

$$\Pr\left[|X| \geqslant \lambda \right] \leqslant \frac{1}{\lambda^2} \langle X^2 \rangle .$$

This follows from a result applying to so-called sub-martingales ($S(t)$ is a submartingale if $\langle S(t) \mid X(t_1), \dots, X(t_n) \rangle \geq S(t_n)$)

thus each martingale $M(t)$ is a submartingale as well, and for any convex-up function φ, Jensen's inequality means $\langle \varphi(M(t)) \rangle \geq \varphi(\langle M(t) \rangle)$

$\Rightarrow \varphi(M)$ is a submartingale.

The inequality result for submartingales is as follows:

Let S be a positive submartingale : $\Pr\left[\max_{0 \leq s \leq t} S(s) \geq \lambda\right] \leq \frac{1}{\lambda} \langle S(t) \rangle$

If M is a martingale, we have just seen that $M^2 = S$ is a submartingale ; $|M(s)| \geq \lambda \Leftrightarrow S(s) \geq \lambda^2$

$$\Rightarrow \Pr\left[\max_{0 \leq s \leq t} M^2(s) \geq \lambda^2\right] \leq \frac{1}{\lambda^2} \langle M^2(t) \rangle .$$

Application to Brownian Motion (Wiener process again $X(s)$) ; C arbitrary constant

$$X(s) \geq C \Leftrightarrow e^{\lambda X(s)} \geq e^{\lambda C} \quad \text{for any } \lambda \geq 0$$

Since $X(s)$ is a martingale, $e^{\lambda X}$ is a submartingale :

$$\Pr\left[\max_{0 \leq s \leq t} X(s) \geq C\right] = \Pr\left[\max_{0 \leq s \leq t} e^{\lambda X(s)} \geq e^{\lambda C}\right]$$

$$\leq e^{-\lambda C} \left\langle e^{\lambda X(t)} \right\rangle = e^{-\lambda C} e^{\lambda^2 \langle X^2(t) \rangle / 2}$$

since $X(t)$ is gaussian

$$\leq e^{-\lambda C} e^{2Dt\lambda^2/2}$$

and we minimize bound with respect to λ

$$\frac{d}{d\lambda}\left(-\lambda C + Dt\lambda^2\right) = 0 \Rightarrow \lambda = C/2Dt$$

$$-C\frac{C}{2Dt} + Dt\frac{C^2}{4D^2t^2} = \frac{-C^2}{4Dt}$$

$$\leq \exp\left(-\frac{C^2}{4Dt}\right)$$

This bound is not ridiculous : the exact calculation yields :

$$\Pr\left[\max_{0 \leq s \leq t} X(s) \geq C\right] = 1 - \operatorname{erf}\left(\frac{C}{\sqrt{4Dt}}\right) \underset{C \to \infty}{\sim} \frac{\sqrt{4Dt}}{\sqrt{\pi}\, C} e^{-\frac{C^2}{4Dt}}$$

$$\text{since } \operatorname{erf}(z) = \frac{2}{\sqrt{\pi}} \int_0^z e^{-u^2} du \underset{\infty}{\sim} 1 - \frac{1}{z\sqrt{\pi}} e^{-z^2}$$

2°| b) Martingale stopping theorem

Let τ be a stopping time for $X(t)$, ie a random variable such that for each t, the occurrence or non-occurrence of $\tau = t$ depends on the values of $X(s)$ for $s \leq t$ (the player looks at the sequence up to time t, and decides to quit or not according to a predefined criterion, like gain / bankrupcy threshold reached, see above)

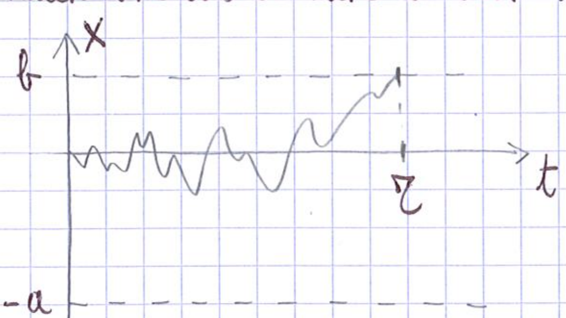
{Doob theorem} If $M(t)$ is a martingale, τ a stopping time s.t. $\langle \tau \rangle < \infty$

and Xenophon (4 centuries BCE) then $\langle M(\tau) \rangle = \langle M(0) \rangle$

3°| Application of the stopping theorem to first passage properties

We consider the symmetric random walk on $\mathbb{Z}$, proba $1/2$, $1/2$ to go left/right, starting from $X_0 = 0$. The position after n steps, X_n, is a martingale.

How much time does it take to exit the segment $[-a, b]$? $a > 0$, $b > 0$



The stopping time τ can be infinite, but $\langle \tau \rangle$ is finite due to the 2 bounds. Note that if $a \to \infty$ or $b \to \infty$, $\langle \tau \rangle = +\infty$ and Doob then does not apply.

Here, Doob's theorem applies and this yields the splitting probabilities, ie the proba $\Pr(X_\tau = -a)$ to leave the interval through $-a$, or $\Pr(X_\tau = +b)$, the proba to reach first the other boundary at $X = b$. From Doob:

$$\langle X(\tau) \rangle = 0 = -a \Pr[X_\tau = -a] + b \Pr[X_\tau = b]$$

Since $\Pr[X_\tau = -a] + \Pr[X_\tau = b] = 1$, we find

$$\Pr[X_\tau = -a] = \frac{b}{a+b} \quad ; \quad \Pr[X_\tau = b] = \frac{a}{a+b}$$

Other interesting results follow invoking other martingales constructed from $X(t)$.

$\longrightarrow$ $X_n^2 - 2Dn$ is a martingale. Here we have $D = 1/2$, such that $\langle X_n^2 \rangle = n$.

Hence we consider $X_n^2 - n$. Doob tells us that

$$\langle X_\tau^2 \rangle - \langle \tau \rangle = 0$$

while

$$\langle X_\tau^2 \rangle = a^2 \Pr[X_\tau = -a] + b^2 \Pr[X_\tau = b] = \frac{a^2 b + a b^2}{a+b}$$
$$= ab$$

$$\langle \tau \rangle = a\,b$$

This is a result we already met (0, x_0, L) Starting from x_0, the mean first passage time at 0 or L is $\underbrace{x_0}_{a}\underbrace{(L-x_0)}_{b}$ for $D = 1/2$

or, more generally: $\dfrac{1}{2D}\, x_0\,(L - x_0)$

→ To go further, a new martingale is in order:

$$\boxed{M_n = \frac{1}{(\cosh\lambda)^n}\, e^{-\lambda X_n}}$$

$$\langle M_n \mid X_{n-1} \rangle = (\cosh\lambda)^{-n}\left[e^{-\lambda(X_{n-1}+1)}\,\frac{1}{2} + e^{-\lambda(X_{n-1}-1)}\,\frac{1}{2}\right]$$

$$= (\cosh\lambda)^{-n}\,(\cosh\lambda)\; e^{-\lambda X_{n-1}}$$

$$= M_{n-1} \qquad : \quad M_n \text{ is indeed a martingale}$$

We apply once more Doob's theorem for stopping time:

$$\langle M_\tau \rangle = M_0 = 1 = \left\langle \frac{1}{(\cosh\lambda)^\tau}\, e^{-\lambda X_\tau} \right\rangle$$

To simplify, we take $a = b$ (−a, 0, a; start)

Due to symmetry, the law of exit time τ does not depend on the fact that the walker exits through a, or $-a$; the joint distribution of τ and $X_\tau = \pm a$, factorizes:

$$\left\langle \frac{1}{(\cosh\lambda)^\tau} \right\rangle \underbrace{\left\langle e^{-\lambda X_\tau} \right\rangle}_{\cosh(\lambda a)} = 1$$

$$\left\langle \frac{1}{(\cosh\lambda)^\tau} \right\rangle = \frac{1}{\cosh(\lambda a)}$$

and to put this in a Laplace transform form, take $\cosh\lambda = e^{s}$:

$$\left\langle e^{-s\tau} \right\rangle = \frac{1}{\cosh\left[a \,\operatorname{argcosh}(e^{s})\right]}$$

This is interesting: we obtained the Laplace transform $\int_0^\infty e^{-s\tau}\, p_a(\tau)\, d\tau$ at a rather modest cost.

<u>Sanity check</u>: take $a = 1$: (−1, ↓, 1) in one step, a boundary is reached and thus: $\tau = 1$, and the pdf is $p_{a=1}(\tau) = \delta(\tau - 1)$, for which

$$\left\langle e^{-s\tau} \right\rangle = e^{-s} \quad \text{and indeed:} \quad \cosh\left[\operatorname{argcosh} e^{s}\right] = e^{s}$$

Also, by taking the successive derivatives $\dfrac{d^n}{ds^n}\left\langle e^{-s\tau} \right\rangle \Big|_{s=0}$, we get $(-1)^n \langle \tau^n \rangle$

Rk 1 : although it was forbidden to take $a \to \infty$ before applying Doob thm, it now is possible, and we should thereby reach the continuum limit (ie Wiener process), provided we rescale τ by a^2 (diffusive scaling), to yield a well behaved limit

$$\left\langle e^{-\mu \tau / a^2} \right\rangle = \frac{1}{\cosh\left[a \operatorname{argcosh}\left(e^{\mu/a^2}\right)\right]} \xrightarrow[a \to \infty]{} \frac{1}{\cosh\sqrt{2\mu}}$$

$$\operatorname{argcosh} x \underset{1}{\sim} \sqrt{2(x-1)}$$

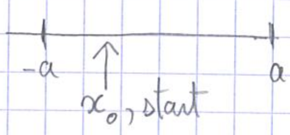
We can make an inverse Laplace transform $\mathcal{L}^{-1}$ by using a Mittag-Leffler representation:

$$\frac{1}{\cosh\sqrt{2p}} = \sum_{k=0}^{\infty} \frac{4\pi(2k+1)(-1)^k}{8p + \pi^2(2k+1)^2} \quad \text{of form } \sum_k \frac{\alpha_k}{p + r_k} \xrightarrow{\mathcal{L}^{-1}} \alpha_k e^{-r_k \tau}$$

$$\xrightarrow{\mathcal{L}^{-1}} \sum_{k=0}^{\infty} \frac{\pi}{2}(2k+1)(-1)^k \exp\left[-\frac{\tau}{a^2}\frac{(2k+1)^2\pi^2}{8}\right] = p_a(\tau)$$

pdf of first passage time at $X = \pm a$

Rk 2 : we can recover these results with Feynman-Kac (in the continuum limit)



$$Q_f(x_0) = \left\langle e^{-\mu\tau/a^2} \right\rangle_{x(0) = x_0} \quad \text{ie Wiener}$$

$$\dot{x} = \xi(t) \; ; \; \left\{ \begin{array}{l} \langle \xi(t) \rangle = 0 , \; \langle \xi(t)\xi(t') \rangle = \delta(t-t') \\ \text{ie } D = 1/2 \end{array} \right.$$

We choose the functional $\Omega = \frac{\mu}{a^2}\int_0^\tau dt' = \frac{\mu\tau}{a^2}$ so that here, $V(x) = \mu/a^2$ in the notation of chapter VII.

We have shown that the first passage functional $Q_f(x_0)$ obeys the back. eq

$$\frac{1}{2}\frac{d^2 Q_f}{dx_0} - \frac{\mu}{a^2} Q_f(x_0) = 0$$

with boundary condition $Q_f(\pm a) = 1$; and $Q(x_0) = Q(-x_0)$

$$\Rightarrow Q_f(x_0) = \frac{\cosh(x_0\sqrt{2\mu}/a)}{\cosh\sqrt{2\mu}}$$

Finally, we are interested in $Q_f(0) = \frac{1}{\cosh\sqrt{2\mu}}$; same as with martingales

$\longrightarrow$ What about the biased random walk with left/right proba q and p ?

Show that, X_n denoting the position after n steps, ie

$$X_n = \sum_{i=1}^{n} \eta_i \qquad \eta_i \begin{cases} +1 & , \text{proba } p \\ -1 & , \quad " \quad q \end{cases}$$

Then $M_n \equiv \left(\frac{q}{p}\right)^{X_n}$ is a martingale

$X_n - n(p-q)$ is also a martingale

From this, the splitting probabilities follow, and also the mean escape time from the interval $[-a, b]$ → exercise

4°) Application to stochastic thermodynamics

In our previous study, ch $\underline{\text{VIII}}$, a martingale is hiding.

For proving the Crooks' identity, we showed that for a trajectory γ that starts at equilibrium from x_0 at $t=t_0$, and its reversed $\check{\gamma}$, starting also from equilibrium at $x = x_f$ at $t = t_f$:

$$\frac{P_r(\gamma)}{P_r(\check{\gamma})} = e^{\beta W(\gamma) - \beta \Delta F} = e^{S_{tot}/k_B} \quad \text{where } TS_{tot} = W(\gamma) - \Delta F$$

To see that this relation is still correct, in terms of the entropy production S_{tot}, for an arbitrary trajectory γ, not starting from equilibrium, we will make use of the above "Crooks" relation. In general:

$$\frac{P_r(\gamma)}{P_r(\check{\gamma})} = \frac{P(x_0,t_0)\, P_r(\gamma|x_0)}{P(x_f,t_f)\, P_r(\check{\gamma}|x_f)} \quad \text{where } \gamma \text{ starts at } x_0 \text{ at } t_0\text{, ends at } x_f \text{ at } t_f$$

$$= \frac{P(x_0,t_0)\, Z_0\, e^{\beta U(x_0,\lambda(0))}\, Z_0^{-1} e^{-\beta U(x_0,\lambda(0))}\, P_r(\gamma|x_0)}{P(x_f,t_f)\, Z_f\, e^{\beta U(x_f,\lambda(t_f))}\, Z_f^{-1} e^{-\beta U(x_f,\lambda(t_f))}\, P_r(\check{\gamma}|x_f)}$$

$e^{\beta \Delta F}$ Crooks calculation → $\exp(+\beta W(\gamma) - \beta \Delta F)$

$$= \frac{P(x_0,t_0)}{P(x_f,t_f)}\, e^{-\beta\underbrace{[U(x_f,\lambda(t_f)) - U(x_0,\lambda(t_0))]}_{\Delta U = W(\gamma)+Q}}\, e^{\beta W(\gamma)}$$

and we remember our definition of the system's stochastic entropy

$$S_{syst}(t) = -k_B \log p(x(t), t)$$

$$= e^{\Delta S_{syst}/k_B}\, e^{-\beta W(\gamma) - \beta Q}\, e^{\beta W(\gamma)}$$

$$S_{tot} = \Delta S_{syst} + \Delta S_{thermostat}, \text{ entropy creation}$$
$$= \Delta S_{syst} - \frac{Q}{T}$$

$$\boxed{\frac{P_r(\gamma)}{P_r(\check{\gamma})} = e^{S_{tot}/k_B}}$$

We are now in a position to show that **in a non-equilibrium steady state** (NESS), e^{-S_{tot}/k_B} **is a martingale**.

For simplicity, we consider a discrete time dynamics

$\gamma_{[0,t]}$ denotes the whole trajectory where $x_0 \to x_1 \to \cdots \to x_s \to x_{s+1} \to \cdots \to x_t$ and $\gamma_{[0,s]}$ is the sub-trajectory up to time s included, where x_0 is drawn according to the time-independent distribution in the NESS: $p(x_0)$.

$$\left\langle e^{-S_{tot}(t)/k_B} \,\middle|\, \gamma_{[0,t]} \right\rangle = \sum_{\gamma_{[s+1,t]}} e^{-S_{tot}(t)/k_B} \Pr\left[\gamma_{[s+1,t]}\right]$$

$$= \sum_{\gamma_{[s+1,t]}} \frac{\Pr[\check{\gamma}_{[0,t]}]}{\Pr[\gamma_{[0,t]}]} \Pr\left[\gamma_{[s+1,t]}\right]$$

$$\Pr\left[\gamma_{[s+1,t]}\right] = \frac{\Pr[\gamma_{[0,t]}]}{\Pr[\gamma_{[0,s]}]}$$

$$= \frac{1}{\Pr[\gamma_{[0,s]}]} \sum_{\gamma_{[s+1,t]}} \Pr\left[\check{\gamma}_{[0,t]}\right] \qquad (*)$$

so that what remains is to "marginalize" $\Pr[\check{\gamma}_{[0,t]}]$. It is at this point that we need to invoke the NESS condition, that played no role up to now.

Assume indeed that we are not in a NESS, so that the transition probabilities $M_{x_i \to x_j}$ do depend on time (say through some $\lambda(t)$, as in the previous chapter):

$$\Pr[\gamma_{[0,t]}] = p(x_0, t=t_0)\, M_{x_0\to x_1}(\lambda_0)\, M_{x_1 \to x_2}(\lambda_2) \cdots M_{x_{t-1}\to x_t}(\lambda_{t-1})$$

Summing over all trajectories means $\sum$ or $\int$ over all $x_0, x_1, \cdots x_t$ and we have telescopic simplification: $\sum_{x_t} M_{x_{t-1}\to x_t}(\lambda_{n-1}) = 1$ … down to $\sum_{x_0}$ → we indeed find $\sum_{\gamma_{[0,t]}} \Pr[\gamma_{[0,t]}] = 1$. Yet, the summation in $(*)$ is of a different nature:

$$\sum_{\gamma_{[s+1,t]}} \Pr[\check{\gamma}_{[0,t]}] = \sum_{x_s, x_{s+1} \cdots x_t} \boxed{p(x_t, t)\, M_{x_t \to x_{t-1}}(\lambda_{t-1}) \cdots M_{x_1\to x_0}(\lambda_0)}$$

and no telescopic simplification occurs. We need here to invoke the NESS condition, meaning that $p(x_t, \cancel{t})$ and $M_{x_k \to x_{k-1}}(\cancel{\lambda_{k-1}})$. Then

$$\boxed{} = p(x_t)\, p(x_{t-1}|x_t) \cdots p(x_0|x_1) = p(x_0, x_1, \cdots x_t)$$

$$\Rightarrow \sum_{x_s, x_{s+1}, \cdots x_t} p(x_t, x_{t-1} \cdots x_0) = p(x_s, \cdots x_0) = p(x_s)\, M_{x_s \to x_{s-1}} \cdots M_{x_1 \to x_0}$$

$$= \Pr\left[\check{\gamma}_{[0,s]}\right]$$

Thus $\left\langle e^{-S_{tot}(t)/k_B} \,\middle|\, \gamma_{[0,s]} \right\rangle = \dfrac{\Pr[\check{\gamma}_{[0,s]}]}{\Pr[\gamma_{[0,s]}]}$

$$= e^{-S_{tot}(s)/k_B}$$

which proves the martingale property (Chetrite / Gupta, 2011, Neri / Roldan / Julicher, 2017)

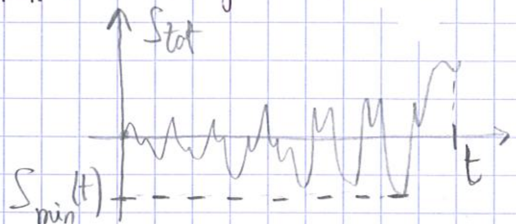
Since we deal with a positive martingale, we can apply the inequality

$$\Pr\left[\max_{0 \leq s \leq t} M(s) \geq \lambda\right] \leq \frac{1}{\lambda}\langle M(t)\rangle \quad ; \quad M(t) = \exp\left(-S_{tot}(t)/k_B\right)$$

$$\Pr\Big[\underbrace{\max_{0 \leq s \leq t} e^{-S_{tot}(s)/k_B} \geq \lambda}_{\min\limits_{0 \leq s \leq t} S_{tot}(s) \leq -k_B \log\lambda}\Big] \leq \frac{1}{\lambda} \underbrace{\left\langle e^{-S_{tot}(t)/k_B}\right\rangle}_{= e^{-S_{tot}(0)/k_B} = 1} \quad \text{by martingale pty}$$

$$\underbrace{\min_{0 \leq s \leq t} S_{tot}(s)}_{\equiv S_{min}(t)} \leq -k_B \log\lambda \qquad \leq \frac{1}{\lambda}$$

$S_{min}(t)$ is defined as the minimum entropy production in a time trajectory:



$S_{tot}(t)$ fluctuates, we known from Jensen inequality and $\left\langle e^{-S_{tot}(t)/k_B}\right\rangle = 1$ that $\langle S_{tot}(t)\rangle \geq 0$ but it is possible that $S_{min}(t) \leq 0$.

What can we say about the mean value of $S_{min}(t)$? Take $\begin{cases} \lambda = \exp(s/k_B) \\ s = +k_B \log\lambda \end{cases}$

$$\Pr\left[S_{min}(t) \leq -s\right] \leq \frac{1}{\lambda} = e^{-s/k_B}$$

$$\Rightarrow \Pr\left[S_{min}(t) \geq -s\right] \geq 1 - e^{-s/k_B}$$

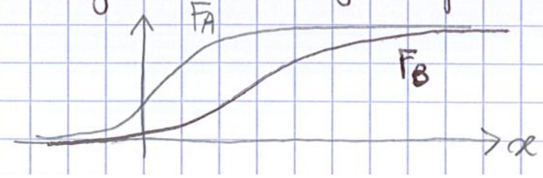
$$\Leftrightarrow \Pr\left[-S_{min}(t) \leq s\right] \geq 1 - e^{-s/k_B} \equiv F_S(s), \text{ cumulative distribution}$$

of an exponentially distributed r.v. S with $\langle S\rangle = k_B$

$\equiv F_{-S_{min}}(s)$, cumulative dist of $-S_{min}$

We are thus in a situation of stochastic dominance: $F_A(x)$, $F_B(x)$ being two cumulatives, $F_A(x) \geq F_B(x), \forall x \Rightarrow \langle x\rangle_A \leq \langle x\rangle_B$

as may be shown by integration by parts: $\langle x\rangle_A = \int_{-\infty}^{+\infty} p_A(x)\, x\, dx$



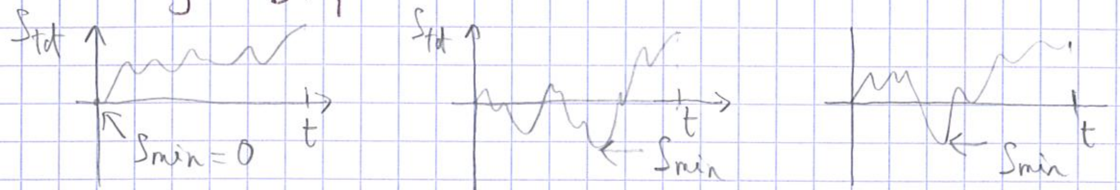
$$F_A(x) = \int_{-\infty}^{x} p_A(x')\, dx'$$

Applying this here:

$$\langle - S_{min}(t) \rangle \leqslant \langle S \rangle = k_B$$

$$\Rightarrow \langle S_{min}(t) \rangle \geqslant - k_B$$

This is a remarkably universal result: averaging over all trajectories, in a non-equilibrium steady state, the minimum entropy production over an arbitrary time window is bounded by $-k_B$, from below.



3 trajectories, each with its $S_{min}(t)$, t being fixed

See the discussion in Peliti / Pigolotti, "Stochastic Thermo", 2021.

NB The martingale constructed here bears similarities with the likelihood ratio used in statistics. Consider random variables $x_1, x_2, \dots$ with pdf $p_n(x_1, \dots, x_n)$. We consider another pdf $q_n(x_1, \dots, x_n)$, and we assume for simplicity that p_n and $q_n \neq 0$ always. Then, with

$$M_n = \frac{q_n(x_1, \dots, x_n)}{p_n(x_1, \dots, x_n)}$$

$M_1, M_2, \dots M_n$ define a random sequence, that is a martingale:

$$\langle M_n \mid x_1, \dots, x_{n-1} \rangle = \int dx_n \, M_n \, p_n(x_1, \dots, x_n) \frac{1}{p_{n-1}(x_1, \dots, x_{n-1})}$$

$$= \int dx_n \frac{q_n(x_1, \dots, x_n)}{p_{n-1}(x_1, \dots, x_{n-1})}$$

$$= \frac{q_{n-1}(x_1, \dots, x_{n-1})}{p_{n-1}(x_1, \dots, x_{n-1})}$$

$$= M_{n-1}$$