V Markov processes : master and Fokker-Planck equations

Goal : introduce tools and vocabulary for a general analysis
switch from trajectory view like with Langevin, to a partial diff. equad view for pdf.

1° Definitions

We study a stochastic process $X(t)$. It is customary to order time $t_1 \leq t_2 \leq \dots \leq t_n$
and we denote $P_n(x_1,t_1 ; x_2,t_2 ; \dots x_n,t_n) \equiv \langle \delta(X(t_1)-x_1)\,\delta(X(t_2)-x_2) \dots \delta(X(t_n)-x_n) \rangle$ the joint pdf.

We have $\langle x(t_1)x(t_2)\rangle = \int dx_1\,dx_2\, x_1 x_2\, P_2(x_1,t_1;x_2,t_2)$

$P_2(x_1,t_1;x_3,t_3) \underset{\forall t_2}{=} \int dx_2\, P_3(x_1,t_1;x_2,t_2;x_3,t_3)$

We also introduce conditional proba-density, such as

$$P_{1|1}(y_2,t_2 \,|\, y_1,t_1) := \frac{P_2(y_1,t_1;y_2,t_2)}{P_1(y_1,t_1)}$$, that is a transition probab.

$$P_{\ell|m}(m+1,\dots \ell \,|\, 1,\dots m) = \frac{P_{\ell+m}(1,\dots \ell+m)}{P_m(1,\dots m)}$$ where $x_i,t_i \to i$

A process is Gaussian if all P_n (and thus all $P_{\ell|m}$) are (multivariate) gaussians.
It is entirely specified by $\langle x(t)\rangle$ and $\langle x(t_1)x(t_2)\rangle$

2° Markov processes

a) Definition and consequences

A particularly important class : loosely speaking, their future evolution does only depend on the present, not on the past. For all sequences $\tau_1 \leq \tau_2 \dots \leq \tau_m \leq t_1 \leq \dots \leq t_n$

$$P_{n|m}(x_1,t_1;\dots x_n,t_n \,|\, y_1,\tau_1;y_2,\tau_2 \dots y_m,\tau_m) = P_{n|1}(x_1,t_1 \dots x_n,t_n \,|\, y_m,\tau_m)$$

The evolution after τ_m does not depend on $y_1, y_2 \dots y_{m-1}$ — No memory

Thus $P_3(1;2;3) = \underbrace{P_{1|2}(3|1;2)}_{\text{Markov} \to P_{1|1}(3|2)} P_2(1;2) = P_{1|2}(3|1,2)\,P_{1|1}(2|1)\,P_1(1)$ where "i" $\equiv (x_i,t_i)$

$\Rightarrow P_3(1;2;3) = P_{1|1}(3|2)\,P_{1|1}(2|1)\,P_1(1)$ (*) and so forth for all P_n.

A Markov process is fully characterized by $P_{1|1}(x_2,t_2|x_1,t_1)$ and $P_1(x,t)$

NB if the process is stationary, $P_1(x)$, does not depend on time, and we often have $P_{1|1}(x_2;t_2|x_1,t_1) \xrightarrow[t_2 \to \infty]{} P_1(x_2)$. The process is then specified by the transition proba $P_{1|1}$ alone

Let us integrate (*) over x_2 : $\underbrace{\int P_3(1,2,3)\,dx_2}_{P_2(1,3)} = \int dx_2\, P_{1|1}(3|2)\,P_{1|1}(2|1)\,P_1(1)$ and $P_2(1,3)/P_1(1) = P_{1|1}(3|1)$

$$\Rightarrow \boxed{P_{1|1}(x_3,t_3|x_1,t_1) = \int dx_2\; P_{1|1}(x_3,t_3|x_2,t_2)\,P_{1|1}(x_2,t_1|x_1,t_1)}$$

↳ CHAPMAN-KOLMOGOROV equation, derived non rigorously by Louis Bachelier in 1900, before Markov processes were "invented" (1906)

The CK equat imposes a constraint on $P_{1|1}$, the transition proba; it admits a large number of solutions.

2°) f-) Examples of Markovian processes

(*) Back to the Wiener process (with $D = 1/2$) : $\dot{x} = \eta(t)$, $\langle \eta(t)\rangle = 0$, $\langle \eta(t)\eta(t+\tau)\rangle = \delta(\tau)$

$x(t)$ obeys a first order diff. equation, thus, to know $x(t)$ for $t>0$, only $x(0)$ and $\eta(t)$ for $t \geq 0$ are required. We do not need $x(t)$ for $t<0 \Rightarrow$ markovian. It is also gaussian

$$P_{1|1}(x_2,t_2|x_1,t_1) = \frac{1}{\sqrt{2\pi(t_2-t_1)}}\exp\left[-\frac{(x_2-x_1)^2}{2(t_2-t_1)}\right]$$

$$P_1(x,t=0) = \delta(x)$$

Note that here, $P_1(x,t) = P_{1|1}(x,t|0,0) = \frac{1}{\sqrt{2\pi t}}\,e^{-x^2/2t}$; not stationary

(*) the most widely known example of a stationary marko proces is the Ornstein-Uhlenbeck process. It corresponds to the Langevin eq $\dot{v} = -v + \eta(t)$

We have seen $\langle v(0)v(\tau)\rangle = e^{-|\tau|}$, and

$$P_{1|1}(v_2,t_2|v_1,t_1) = \frac{1}{\sqrt{2\pi(1-e^{-2\tau})}}\exp\left[-\frac{(v_2 - v_1 e^{-\tau})^2}{2(1-e^{-2\tau})}\right] \; ; \; \tau = t_2 - t_1 \geq 0$$

$$P_1(v) = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{v^2}{2}\right)$$

One can check that Chapman-Kolmogorov is obeyed.

Ornstein-Uhlenbeck process is Gaussian, stationary, markovian. Doob's theorem states that it is the only proces with these 3 properties (leaving aside the completely random process where $P_n(1,\dots,n) = \prod_{i=1}^{n} P_1(i)$).

(*) the discrete random walks considered so far are markovian ; so is the PageRank algorithm etc.

Ex of a non-markovian process : when $v(t)$ obeys Ornstein-Uhlenbeck, and is thus markovian, $\dot{x} = v(t)$ so that $x(t)$ is generated by a "noise" $v(t)$ that has autocorrelation $\langle v(t)v(t')\rangle = e^{-|t-t'|}$, and is not a $\delta(t-t')$.

↳ coloured noise, with memory. The evolution of x depend on x_0 and v_0. Thus $x(t)$ is not markovian. This can also be seen from the fact that x obeys a second order differential equation, not a first order.

Rk 1 : yet, if time is sampled over time interval $\Delta t \gg \gamma^{-1} = 1$ here, then the memory of v_0 is lost and $x(t)$ becomes markovian (the transition proba only depends on initial and final positions)

Rk 2 add an external force $F(x)$: $m\dot{v} = -\gamma m v + F(x) + R(t)$ and $v(t)$ is no longer markovian; $x(t)$ is non markovian as well. Yet, the joint process $\{x(t), v(t)\}$ obeys a first order diff equation. Thus, if $R(t)$ is δ-correlated, $\{x(t), v(t)\}$ is a Markov process

Conclusion : it is pointless to ask if a physical phenomenon is of Markov type or not. One has to specify the variables used for description. What is interesting is to find the set of variables - if it exists - that makes the description markovian.

2°) c) The master equation

We consider a Markov process with homogeneous transition proba (time translation invariant) It obeys

$$P(x, t+\Delta t) = \int dy\ P(x, \Delta t | y)\, P(y, t)$$

Not very convenient : often better to have a differential form.

$$P(x,t+\Delta t) - P(x,t) = \int dy\ P(x,\Delta t|y) P(y,t) - P(x,t) \underbrace{\int dy\ P(y,\Delta t|x)}_{1}$$

We denote $w(x|y) = \lim_{\Delta t \to 0} \frac{1}{\Delta t} P(x, \Delta t | y)$, the transition rate (and assume the limit exists) :

$$\frac{\partial P(x,t)}{\partial t} = \int dy \left[w(x|y) P(y,t) - w(y|x) P(x,t) \right] \quad : \text{Master equation}$$

↳ "équation maîtresse" / "équation pilote". It may seem general, accounting for a gain / loss balance, but its validity is limited to Markovian processes

Doob allows to retrieve a few results. Take $\dot{v} = -\gamma v + R(t)/m$, and Fourier transform

$$-i\omega \hat{v}(\omega) + \gamma \hat{v}(\omega) = \frac{1}{m} \hat{R}(\omega)$$

↳ correlation time is τ_c, not necessarily 0

$$\hat{v}(\omega) = \frac{1}{m(\gamma - i\omega)} \hat{R}(\omega)$$

Thus, $\hat{v}(\omega)$ has the same statistical properties as $\hat{R}(\omega)$. Since $R(t)$ is gaussian, then its Fourier Transform also is: $\hat{R}(\omega)$ gaussian $\Rightarrow$ $\hat{v}(\omega)$ is gaussian

In terms of spectral densities

$$S_v(\omega) = \left| \frac{1}{m(\gamma - i\omega)} \right|^2 S_R(\omega) = \frac{1}{m^2(\gamma^2 + \omega^2)} S_R(\omega)$$

If one wants that $v(t)$ is markovian, since it is also stationary and gaussian, it is required, since it has exp autocorrelation (from Doob), that its spectral density is Lorentzian: $S_v(\omega) \propto \frac{1}{\gamma^2 + \omega^2} \Rightarrow S_R(\omega) = \text{cte}$

From Wiener-Khinchine, $\langle R(0) R(t) \rangle \propto \delta(t) \Rightarrow \tau_c = 0$; R has to be white noise.

Otherwise $v(t)$ is not markovian

$w(x|y)$: jump $y \to x$

the Markovian hypothesis hides in the initial condition, eg $P(x) = \delta(x - x_0)$ at $t = t_0 \rightarrow$ the transition proba is in general a $P_{1|2}(x, t+\Delta t | x_0, t_0; y, t)$

Ex the Poisson process $N(t)$ takes integer values (it is a counting process), with $N(0) = 0$. the events that are counted are independent. In a short time Δt, there is a proba $\lambda \Delta t$ that an event occurs, and that $N(t)$ is increased by 1.

We define $P_n(t) = Pr[N(t) = n]$. We have

$$\frac{d}{dt} P_n(t) = \lambda P_{n-1}(t) - \lambda P_n(t) \quad , n \geq 1$$

this comes from $P_n(t+\Delta t) = (1 - \lambda \Delta t) P_n(t) + \lambda \Delta t \, P_{n-1}(t)$

And $\frac{d}{dt} P_0(t) = -\lambda P_0$. This can be solved by recurrence :

$$P_n(t) = e^{-\lambda t} \frac{(\lambda t)^n}{n!} \quad ; \quad \lambda \text{ is the rate of the process}$$

Applications are many :

- death from horse kicks in Prussian army (~ 1900)
- γ emission and radioactive decay
- financial markets shocks
- for a phone company : number of phone calls, or service requests to web platform...
- appears generally when counting rare events. Rare means that

$Pr(1 \text{ event in time } \Delta t) = \lambda \Delta t + o(\Delta t)$

$Pr(2 \text{ events or more in time } \Delta t) = o(\Delta t)$

(For instance, a binomial $\mathcal{B}(n, p)$ for $n \to \infty$, $p \to 0$ ($\rightarrow$ rare event), $np \to \Lambda$ tends to a Poisson distrib with param $\Lambda \rightarrow e^{-\Lambda} \frac{\Lambda^n}{n!}$)

3) Markov chains

These are Markov processes where x and t are discrete, and the transition proba is time trans. invariant (homogeneous). If x takes only finite $\#$ values, N the chain is finite (what we assume here). Notation $P(x,t) \rightarrow P_i(t)$ and we take $t = 0, 1, 2, \ldots$ The transition probabilities can be viewed as $N \times N$ matrix

$$P(i, t+1 | j, t) \equiv M_{ij} \quad ; \quad M \text{ is a stochastic matrix}$$

$= Pr[j \to i]$

$$(\text{i.e. } M_{ij} \geq 0, \quad \sum_{i=1}^{N} M_{ij} = 1, \ \forall j)$$

$$P_i(t+1) = \sum_j M_{ij} P_j(t)$$

Thus $\vec{P}(t+1) = M\vec{P}(t)$ where $\vec{P} = \begin{vmatrix} P_1 \\ \vdots \\ P_N \end{vmatrix}$ and $\vec{P}(n) = M^n \vec{P}(0)$ ($t=n$, $t=0$)

and we have to study the properties of M^n → spectral properties of M important

a) Properties of stochastic matrices : ⚠ in general non symmetric (not self adjoint)

it may thus have complex eigenvalues ($\sum_i M_{ij} = 1$; $\sum_j M_{ij} \neq 1$ in general)

Yet, $\sum_i \pi_{ij} = 1$, $\forall j$ means $(1, \dots 1) \times M = (1, \dots 1)$

so that $(1, \dots 1)$ is a left-eigenvector, with eigenvalue 1.

$\Rightarrow \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$ is a right eigenvector of ${}^tM$, " " "

and we remember that a matrix and its transpose have the same spectrum

$\Rightarrow$ M admits an eigenvector $\vec{P}^*$, with eigenvalue 1 : $\vec{P}^* = M\vec{P}^*$

this yields a stationary solution. Besides (Unknown at this point.)

The largest eigenvalue is ≤ 1 (then, it is one). This is clear on intuitive grounds, otherwise $M^n\vec{P}(0)$ would "explode", and not be normalizable.

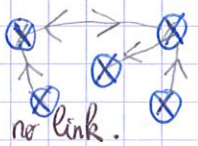
Proof : Let λ be an eigenvalue, with eigenvector $\vec{v}$

$$\sum_j M_{ij} v_j = \lambda v_j \Rightarrow |\lambda v_i| = |\lambda||v_i| \leq \sum_j M_{ij} |v_j| \quad \text{since } M_{ij} \geq 0$$

$$\Rightarrow |\lambda| \sum_i |v_i| \leq \sum_j \underbrace{\left(\sum_i M_{ij}\right)}_{1} |v_j| = \sum_j |v_j| \Rightarrow |\lambda| \leq 1$$

Is it possible that there be several stationary distributions ? In other words, is the eigenvalue $\lambda = 1$ (the largest) degenerate ? In a large number of cases, it is not : consequence of the Perron - Frobenius theorem.

The matrix M, $N \times N$, defines an oriented graph



N points with oriented links

$M_{ij} > 0$ means $\exists$ a link from i to j ; $M_{ij} = 0$ means no link.

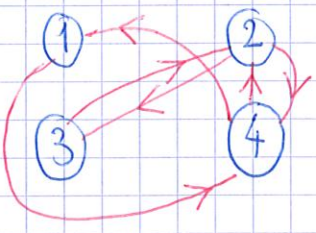
If between any two points, there exists a path made of oriented links, (from $i \to j$ and $j \to i$), the matrix is called irreducible, the theorem applies (beyond stochastic matrices, but for matrices with ≥ 0 entries) and it states that the largest eigenvalue is non degenerate, and that the corresponding eigenvector has all components all of the same sign (that we can take ≥ 0). This is fortunate for the probabilistic interpretation! Indeed, the eigenvector is a set, here, of probabilities.

Note that if all $M_{ij} > 0$, $\forall i,j$, M is irreducible and Perron - Frobenius applies

Note (ex) that Chapman-Kolmogorov becomes: $M^{m+n} = M^m M^n$

In other words, M is not always diagonalizable (on $\mathbb{R}$), but can always be decomposed in Jordan blocks.

Ex: irreducible matrix with some 0 elements:



$$M = \begin{pmatrix} 0 & 0 & 0 & >0 \\ 0 & 0 & >0 & >0 \\ 0 & >0 & 0 & 0 \\ >0 & >0 & 0 & 0 \end{pmatrix}$$

Consequence for the time evolution : we order the eigenvalues s.t.

$$\lambda_1 = 1 > |\lambda_2| \geq |\lambda_3| \cdots \geq |\lambda_N|$$

and denote $\vec{v}_\lambda$ the corresponding eigenvectors, for $\lambda = \lambda_1, \lambda_2, \cdots \lambda_N$

$$\vec{P}(t=0) = \sum_\lambda a_\lambda \vec{v}_\lambda \quad , \text{ by projection}$$

$$\vec{P}(t=n) = M^n \vec{P}(0) = \sum_\lambda a_\lambda \lambda^n \vec{v}_\lambda \underset{n\to\infty}{\sim} a_1 \vec{v}_1 + a_{\lambda_2} \lambda_2^n \vec{v}_{\lambda_2} + \cdots$$

($\vec{P}^*$)

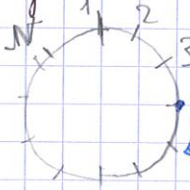
If $\vec{v}_1$, associated to $\lambda_1 = 1$, is normalized,

then since $\vec{P}$ also is normalized, and $\vec{P} \to a_1 \vec{v}_1 = \vec{P}^*$ i.e. $a_1 = 1$ necessarily.

More interesting : we see that $\vec{P}$ converges to the steady state $\vec{P}^*$, exponentially fast at long times: $\lambda_2^n = [\mathrm{sgn}(\lambda_2)]^n e^{n \log|\lambda_2|} = (\pm 1)^n e^{-n/\tau}$

$$\vec{P}(t=n) \underset{n\to\infty}{\simeq} \vec{P}^* + (\pm 1)^n \exp(-n/\tau) \quad ; \quad \tau = -\frac{1}{\log|\lambda_2|} > 0 \quad (|\lambda_2| < 1)$$

Example of a finite chain : random walk on a ring



$$M = \begin{pmatrix} 0 & 1-p & 0 & \cdots & p \\ p & 0 & 1-p & & 0 \\ 0 & p & 0 & & \vdots \\ \vdots & \ddots & \ddots & & 1-p \\ 1-p & 0 & \cdots & p & 0 \end{pmatrix}$$

circulant matrix. use F.T to diagonalize

Here, the long-time state is simple: $P_i^* = \frac{1}{N}$, for all sites i.

3°) f) Detailed balance and the different sorts of Markov chains

Take N finite first. The stationary state $\vec{P}^*$ obeys

$$P_i^* = \sum_j M_{ij} P_j^* \quad \Rightarrow \quad \sum_j M_{ij} P_j^* = \sum_j M_{ji} P_i^*$$

In equilibrium statistical physics, it is possible to show that as a consequence of the reversibility of Hamilton's eq. of motion, a stronger condition holds (Wigner, 1954)

$$M_{ij} P_j^* = M_{ji} P_i^*$$

, called detailed balance ("équilibre en détail" / "bilan détaillé")

When this condition holds, P^* is said to be an equilibrium state

If not, the stationary state P^* " " " " a non-equilibrium steady state (NESS)

↳ such states are characterized by existence of non zero fluxes

Recall M is N×N matrix. When $N \to \infty$, it is possible that $\vec{P}^*$ be non normalizable (like for r.w on a ring, $P_i^* = \frac{1}{N} \overset{N\to\infty}{\to} 0$). The process is then said to be transient.

Rk : when M is symmetric, its study is simpler. But it is not in gal. Yet, if detailed balance holds :

$$\underbrace{M_{ij}\sqrt{\frac{P_j^*}{P_i^*}}}_{\tilde{M}_{ij}} = \underbrace{M_{ji}\sqrt{\frac{P_i^*}{P_j^*}}}_{\tilde{M}_{ji}}$$

$\tilde{M}$ is symmetric → convenient!

Note that M and $\tilde{M}$ have same spectrum } Easier to manipulate, real spectrum.

$$\tilde{M}_{ij}\tilde{v}_j = \lambda \tilde{v}_j \Rightarrow M_{ij}\underbrace{\sqrt{P_j^*}\,\tilde{v}_j}_{v_i} = \lambda \sqrt{P_i^*}\,\tilde{v}_i$$

ie $M\vec{v} = \lambda\vec{v} \Rightarrow M$ has real spectr.

3e) c) An important application of Markov chains : Monte Carlo simulations

In stat. phys, and more broadly in statistics, one often needs to sample a variable from a known distribution, for instance to compute the mean of an observable $\mathcal{O}$: $\langle \mathcal{O} \rangle = \sum_{\mathcal{C}} \mathcal{O}(\mathcal{C}) \frac{1}{Z} e^{-\beta \mathcal{H}(\mathcal{C})}$, for equilibrium of Gibbs type $p(\mathcal{C}) = Z^{-1} \exp(-\beta \mathcal{H}(\mathcal{C}))$ → difficult even with a computer.

Take eg a spin $\frac{1}{2}$ system in 3d, size $1000 \times 1000 \times 1000 = 10^9$.

There are $2^{(10^9)}$ states, $\gg \simeq 10^{10^8}$, astronomical. The curse of dimensionality looms : the vast majority of states do not contribute (or very weakly), to the result.

Trick : define a suitable Markov chain, obeying detailed balance, such that the steady state (equilibrium reached at long times) is $P^*(\mathcal{C}) \propto e^{-\beta \mathcal{H}(\mathcal{C})}$

What is interesting is that we seek to get this P^*, while being unable to compute $Z = \sum_{\mathcal{C}} e^{-\beta \mathcal{H}(\mathcal{C})}$, but it turns out that we do not need to know Z to define the probability transitions for the chain :

$$M(\mathcal{C} \to \mathcal{C}')\, e^{-\beta \mathcal{H}(\mathcal{C})} \frac{1}{Z} = M(\mathcal{C}' \to \mathcal{C})\, e^{-\beta \mathcal{H}(\mathcal{C}')} \frac{1}{Z} \quad (*)$$

There are many possible choices for M. A popular one is METROPOLIS ALGORITHM

$$M(\mathcal{C} \to \mathcal{C}') = \min\left[1, e^{-\beta[\mathcal{H}(\mathcal{C}') - \mathcal{H}(\mathcal{C})]}\right] \equiv M_{\mathcal{C}'\mathcal{C}}$$

↳ if $\mathcal{H}(\mathcal{C}) < \mathcal{H}(\mathcal{C}')$, energy ↗, $M(\mathcal{C} \to \mathcal{C}') = e^{-\beta \mathcal{H}(\mathcal{C}') + \beta \mathcal{H}(\mathcal{C})} < 1$

and $M(\mathcal{C}' \to \mathcal{C}) = 1 \Rightarrow (*)$ obeyed

↳ if $\mathcal{H}(\mathcal{C}) > \mathcal{H}(\mathcal{C}')$, energy ↘, $M(\mathcal{C} \to \mathcal{C}') = 1$ (move always accepted)

and $M(\mathcal{C}' \to \mathcal{C}) = \exp\left[-\beta(\mathcal{H}(\mathcal{C}) - \mathcal{H}(\mathcal{C}'))\right] < 1$, $(*)$ obeyed also.

This dynamics converges towards Gibbs equilibrium ($M_{\mathcal{C}\mathcal{C}'} > 0$, irreducible)

and with a sufficiently long walk (n steps)

$$\frac{1}{n}\sum_{\alpha=1}^{n} \mathcal{O}(\alpha) \longrightarrow \langle \mathcal{O} \rangle = \frac{1}{Z}\sum_{\mathcal{C}} \mathcal{O}(\mathcal{C})\, e^{-\beta \mathcal{H}(\mathcal{C})}$$

↳ mean over all states visited, generated by the Markov chain

There is a problem if $\mathcal{H}$ is very large for some states → large energy barriers to overcome, ergodicity problem (difficult equilibration).

Monte Carlo is a method of enormous importance in various fields of basic science & engineering

4°) The Fokker-Planck equation

Back to the continuum ($x \in \mathbb{R}$). The master eq is an integro-differential equation, not convenient to handle → we will now transform it into a PDE, easier to deal with.

a) Derivation

$$\partial_t P(x,t) = \int dy \left[w(x|y) P(y,t) - w(y|x) P(x,t) \right]$$

We rewrite $w(x \mid x-\eta)$ as $\tilde{w}(\eta ; x-\eta)$ to have the jump length explicit

(end, start; jump, start)

$$\Rightarrow \partial_t P(x,t) = \int d\eta \left[\tilde{w}(\eta, x-\eta) P(x-\eta, t) - \tilde{w}(\eta; x) P(x,t) \right]$$

make Taylor exp of this funct of $x-\eta$ in powers of η

$$= \int d\eta \sum_{n=1}^{\infty} (-1)^n \eta^n \frac{\partial^n}{\partial x^n} \left[\tilde{w}(\eta, x) P(x,t) \right] \frac{1}{n!}$$

$$\Rightarrow \partial_t P(x,t) = \sum_{n=1}^{\infty} (-1)^n \frac{\partial^n}{\partial x^n} \left[M_n(x) P(x,t) \right]$$

with $M_n(x) = \frac{1}{n!} \int d\eta\, \eta^n \tilde{w}(\eta, x) = \frac{1}{n!} \lim_{\Delta t \to 0} \frac{\langle (\Delta x)^n \rangle}{\Delta t}$

(jump, start)

$$= \lim_{\Delta t \to 0} \frac{1}{n!\,\Delta t} \int d\eta\, \eta^n P(x+\eta, \Delta t \mid x) = \lim_{\Delta t \to 0} \frac{1}{n!\,\Delta t} \int d(\Delta x) (\Delta x)^n P(x+\Delta x, \Delta t \mid x)$$

(end, start)

This is the Kramers-Moyal expansion.

It relies on fact that $\tilde{w}(\eta, x)$ is a sharp function of η, the jump amplitude, that quickly goes to 0 when $\eta \nearrow$. Thus, the moments $\langle (\Delta x)^n \rangle = \langle \eta^n \rangle$ exist.

In many cases we have/will meet, we furthermore have that $M_n = 0$ for $n > 2$.

Thus $\partial_t P(x,t) = -\partial_x \left[M_1(x) P \right] + \partial_x^2 \left[M_2(x) P \right]$; FOKKER-PLANCK eq

We will see this explicitly when computing the moments $\langle (\Delta x)^n \rangle$. This feature is deeply connected to the central limit theorem. Take indeed the dynamics of a

Langevin equation (such as $\dot{x} = \mu F + \sqrt{2D}\,\xi(t)$) with a correlation time τ_c for $\xi(t)$. ⚠ $\lim_{\Delta t \to 0}$ above has to be understood such that $\Delta t \gg \tau_c$ always! We can formally take $\tau_c = 0$ here, before considering $\tau_c \to 0$. Then, the distribution of jumps, on time scale Δt, is Gaussian. Moments $\langle (\Delta x)^3 \rangle$, $\langle (\Delta x)^5 \rangle$ vanish, and $\langle (\Delta x)^2 \rangle = O(\Delta t) \propto \Delta t$. But all cumulants > 2 vanish:

$c_4 = 0 \Rightarrow \langle (\Delta x)^4 \rangle = 3 \langle (\Delta x)^2 \rangle^2 \propto (\Delta t)^2$ and thus $M_4 = 0$.

Same argument with $c_6 = 0 \Rightarrow \langle (\Delta x)^6 \rangle \propto (\Delta t)^3 \Rightarrow M_6 = 0$.

Remembering here the Marcinkiewicz theorem (1939), stating that a cumulant generating function cannot be a polynomial of degree > 2 (ie that all cumulants beyond 2 vanish and we have a Gaussian, or there is an ∞ number of $\neq 0$ cumulants), we can understand the Pawula theorem stating that the Kramers-Moyal expansion either stops at order 2, or features an ∞ number of terms. Indeed, if we are in the Marcinkiewicz situation where Δx is non Gaussian and there are ∞ nonzero cumulants, then $\log \langle e^{k \Delta x} \rangle$ is of $O(\Delta t)$ and all $\neq 0$ cumulants are $O(\Delta t)$. Thus, the corresponding $M_n \neq 0$.

Proof of Pawula thm: follows from Cauchy-Schwarz inequality

$$\int \left[f(x) g(y) - f(y) g(x) \right]^2 P(x) P(y) \geq 0$$

$$\Rightarrow \left(\int f(x) g(x) P(x) dx \right)^2 \leq \int f^2(x) P(x) dx \int g^2(x) P(x) dx$$

We take $f(x) = (x - x')^n \equiv (\Delta x)^n$ for a given x'

$$g(x) \equiv (\Delta x)^{n+m}$$

$$\Rightarrow \langle (\Delta x)^{2n+m} \rangle^2 \leq \langle (\Delta x)^{2n} \rangle \langle (\Delta x)^{2n+2m} \rangle$$

$$\Rightarrow (M_{2n+m})^2 \leq M_{2n} M_{2n+2m}$$

Thus, if $M_{2p} = 0$, all higher order $= 0$: $M_{2p+1}, M_{2p+2} \dots = 0$

But we also "go down". Take for instance $M_6 = 0$.

$$M_4^2 \leq M_6 M_2 \Rightarrow M_4 = 0 \quad (n=1, m=2)$$

$$M_3^2 \leq M_4 M_2 \Rightarrow M_3 = 0 \quad (n=1, m=1)$$

We cannot go below: $n \geq 1$ and $m \geq 1$, since $M_0 = \lim_{\Delta t \to 0} \frac{\langle 1 \rangle}{\Delta t} \to \infty$.

Thus, if some $M_{2n} = 0$ for $n \geq 2$, all $M_n = 0$ down to $n = 3 \Rightarrow$ yields Fokker-Planck

Alternatively, if no $M_{2n} = 0$, we have os terms remaining in the Kramers-Moyal expansion.

For a vectorial variable $\vec{x}$, the Fokker Planck eq reads:

$$\partial_t P(\vec{x},t) = -\vec{\nabla}_{\vec{x}} \cdot \left[\vec{M}_1(\vec{x}) P \right] + \vec{\nabla}_{\vec{x}} \vec{\nabla}_{\vec{x}} : \left[\overleftrightarrow{M}_2(\vec{x}) P(\vec{x},t) \right]$$

(matrix)

$$\hookrightarrow \lim_{\Delta t \to 0} \frac{\langle \Delta \vec{x} \rangle}{\Delta t} \; ; \quad \hookrightarrow \sum_{i,j} \partial_{x_i} \partial_{x_j} \left(M_{2,ij} P \right)$$

For FP, the first two moments of the jump distribution only suffice to characterize $P(x,t)$. This is an important simplification compared to the master equation.

Yet, FP remains difficult to solve analytically, even if a # of interesting results can be obtained (see below).

4° f) A few remarks on the F.P. equation

F.P. has the form of a drift-diffusion equation: take for simplicity M_1 and M_2 cst:

$$\partial_t P = -a_1 \partial_x P + a_2 \partial_x^2 P$$

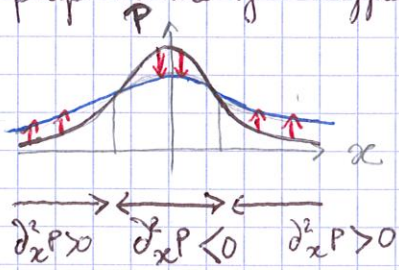
Let us discuss the two terms separately.

$a_2 = 0$, only drift: $(\partial_t + a_1 \partial_x) P = 0 \Rightarrow P(x,t) = \varphi(x - a_1 t)$

the density (of p, $\equiv$ pdf P) is advected at constant speed a_1. It drifts away

$a_1 = 0$, only diffusion $\partial_t P = a_2 \partial_x^2 P$, $\rightarrow$ normal diffusion

a_2 plays the role of a diffusion coefficient.



The density "spreads", if $a_2 > 0$

$a_2 < 0$ leads to a singularity.

Note that FP can be rewritten $\partial_t P + \partial_x J = 0$ with a current $J = a_1 P - a_2 \partial_x P$

(drift current; diffusive current)

so that $\partial_t \left(\int dx P \right) = 0$ if the current $J \to 0$ at infinity.

This makes probability / density conservation apparent.

This form of the current translates the energy (drift) / entropy (diffusion) competition: Newton against Boltzmann in the case of sedimentation.

For the solution to the diffusion equation, see exercises.

Note also that the diffusion equation was hiding in plain view in the basic r-w on $\mathbb{Z}$, with left/right proba of jump $p = 1/2$:

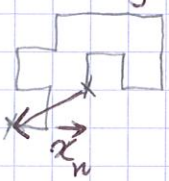
$$P_i(t+1) = \frac{1}{2} P_{i-1}(t) + \frac{1}{2} P_{i+1}(t)$$

(P_i : site)

$$\Rightarrow \underbrace{P_i(t+1) - P_i(t)}_{\parallel \partial_t P} = \underbrace{\frac{1}{2}}_{D} \Big[\underbrace{P_{i+1} - 2P_i + P_{i-1}}_{\text{discrete Laplacian } \partial_x^2 P} \Big]$$

We are indeed in a case where jumps $\eta_i = \pm 1 \Rightarrow \langle \eta_i^2 \rangle = 1$ so that with $x_n = \sum_{i=1}^{n} \eta_i$ the position of the walker : $\langle x_n^2 \rangle = V(x_n) = n V(\eta_1) = n$, to be compared to the normal diffusion law $2Dt$. Here $t = n \Rightarrow D = 1/2$.

Generalization to higher dimension is straightforward. r-w on cubic lattice (say), $\mathbb{Z}^d$: $\langle \vec{x}_n^{\,2} \rangle \propto n$ also (normal diffusion). The problem is much more difficult with **excluded volume** : an already visited site can no longer be visited $\Rightarrow$ **SELF AVOIDING WALK** : non markovian



Drawing all displacements $\rightarrow$ we have a simplified but relevant model of polymer, where monomers cannot overlap.

At variance with the ideal case (no excluded volume), $\langle \vec{x}_n^{\,2} \rangle$ grows faster than n, the number of monomers (ie the size/weight of polymer)

$$\langle \vec{x}_n^{\,2} \rangle \propto n^{3/4}, \quad d = 2$$
$$\propto n^{\zeta}, \quad d = 3 \qquad \zeta \simeq 0{,}59$$
$$\propto n^{1/2}, \quad d \geq 4$$

A classic argument due to Paul Flory (chemistry Nobel '74), yields $\zeta = \frac{3}{d+2}$, $d \leq 4$ (quite accurate but not exact when $d = 3$).

4°) c) Back to Brownian motion

Consider Langevin $\begin{cases} m\dot{v} = -\gamma m v + F_{ext}(x) + R(t) \\ \dot{x} = v \end{cases}$ and work with vector (x, v)

This process is Markovian (first order diff eq) and it makes sense to look for the Kramers-Rayal / Fokker-Planck eq. We need to introduce joint pdf $P(x, v, t)$

Note that although $x(t)$ and $v(t)$ are coupled at the trajectory level ($\dot{x} = v$), x and v are the 2 arguments of P. They are coupled in general : for a given x many v are possible. We will also see that at equilibrium, they decouple.

$$\partial_t P(x,v,t) = -\partial_x\left(M_1^{(x)} P\right) - \partial_v\left(M_1^{(v)} P\right) + \partial_x^2\left(M_2^{(xx)} P\right) + 2\,\partial_x\partial_v\left(M_2^{(xv)} P\right) + \partial_v^2\left(M_2^{(vv)} P\right)$$

The moments M_1, M_2 are obtained, for given x and v, by averaging the jumps (Δx, Δv, $\Delta x \Delta v$, $(\Delta x)^2 \dots$) over all trajectories possible (noise realizations)

$$\Delta x = v\,\Delta t + o(\Delta t)$$

$$\Delta v = -\gamma v\,\Delta t + F_{ext}/m\;\Delta t + \frac{1}{m}\underbrace{\int_t^{t+\Delta t} R(t')\,dt'}_{B_{\Delta t}}$$

$$\langle B_{\Delta t}^2\rangle = 2\Gamma m^2 \Delta t$$

$$\Rightarrow M_1^{(x)} = \lim_{\Delta t\to 0}\frac{1}{\Delta t}\langle \Delta x\rangle = v$$

$$M_1^{(v)} = \lim \frac{1}{\Delta t}\langle \Delta v\rangle = -\gamma v + \frac{F_{ext}}{m}$$

$$M_2^{(xx)} = \frac{1}{2}\lim \frac{1}{\Delta t}\langle (\Delta x)^2\rangle = \frac{1}{2}\lim v^2\Delta t = 0$$

$$M_2^{(xv)} = \frac{1}{2}\lim \frac{1}{\Delta t}\langle (\Delta x)(\Delta v)\rangle = \lim_{\Delta t} O((\Delta t)^2) = 0$$

$$M_2^{(vv)} = \frac{1}{2}\lim \frac{1}{\Delta t}\langle (\Delta v)^2\rangle = \frac{1}{2m^2}\lim \frac{1}{\Delta t}\langle B_{\Delta t}^2\rangle = \Gamma \quad \text{(diff coeff in } v\text{-space)}$$

and all moments of higher order vanish (check). $\Gamma = \gamma\frac{kT}{m}$

We thereby see that Kramers-Mayal reduces to Fokker-Planck:

KRAMERS

$$\underbrace{\left(\partial_t + v\partial_x + \frac{F_{ext}(x)}{m}\partial_v\right)}_{\text{free streaming}} P(x,v,t) = \underbrace{\partial_v\left[\gamma v P + \frac{\gamma kT}{m}\partial_v P\right]}_{\text{interaction with suspending fluid}}$$

$\parallel \frac{D}{Dt}$ in fluid mechanics; material derivative

"collisional contribution" $-\partial_v j$; $j = -\gamma v P - \Gamma\partial_v P$ current in velocity space.

Generalization to higher dimension:

$$\left(\partial_t + \vec{v}\cdot\vec{\nabla}_{\vec{r}} + \frac{\vec{F}_{ext}}{m}\cdot\vec{\nabla}_{\vec{v}}\right)P(\vec{r},\vec{v},t) = \gamma\,\vec{\nabla}_{\vec{v}}\cdot\left[\vec{v}P + \frac{\gamma kT}{m}\vec{\nabla}_{\vec{v}}P\right]$$

As it should be, the equilibrium distribution

$$P_{eq}(\vec{r},\vec{v}) = \frac{1}{Z}\,e^{-\beta U_{ext}(\vec{r})}\,e^{-\frac{1}{2}\beta m v^2}, \quad \text{when } \vec{F}_{ext} = -\vec{\nabla}U_{ext}$$

is a solution to the Kramers eq. (it nullifies seperately the free streaming op and $v + \frac{\gamma kT}{m}\partial_v$, seen as an operator).

In the long-time limit ($\gamma t\to\infty$ ⇔ overdamped limit where $\gamma^{-1}\to 0$), we have seen that $x(t)$ (non Markov in general), becomes markovian, with Langevin eq

$$\dot{x} = \frac{1}{m\gamma}F_{ext} + \frac{1}{m\gamma}R(t)$$

$$\langle \Delta x\rangle = \frac{1}{m\gamma}F_{ext}(x)\,\Delta t + o(\Delta t) \Rightarrow M_1^{(x)} = \mu F_{ext} \quad (\mu = \frac{1}{m\gamma})$$

$$M_2^{(xx)} = \frac{1}{2}\lim_{\Delta t\to 0}\frac{\langle(\Delta x)^2\rangle}{\Delta t} = \frac{1}{2}\,2D = D$$

taken after $\gamma^{-1}\to 0$; in practice, $\Delta t\to 0$ but $\gg \gamma^{-1}$.

We thus obtain the **SMOLUCHOWSKI** equation for the density $P(x,t)$

$$\partial_t P(x,t) = -\partial_x \left[\mu F_{ext}(x) P(x,t) \right] + D \partial_x^2 P(x,t) = -\partial_x j$$

where the current has a familiar drift-diffusion form

$$j = \mu F_{ext} P \underbrace{- D \partial_x P}_{\text{Fick's law}} \quad ; \quad \text{// Einstein's argument ch III, 3°) d)}$$

$P(x,t)$ can be viewed as the particle density

Take again a conservative force $F_{ext} = -\partial_x U_{ext}$. The equilibrium distribution

$$P_{eq}(x) = \frac{1}{Z} e^{-\beta U_{ext}} \quad \text{is a solution to Smoluchowski,}$$

provided $Z = \int dx\, e^{-\beta U_{ext}(x)} < +\infty$, ie that $U_{ext}(x)$ grows sufficiently fast when $|x| \to \infty$, for the particle to be confined. If $Z = +\infty$, equilibrium is never reached, the pdf keeps spreading, as for free diffusion (ie when $F_{ext} = 0$).

Note that equilibrium $\Rightarrow j = 0$ (no current). This also holds for $d > 1$, for $\vec{F}_{ext} = -\vec{\nabla} U_{ext}$. If on the other hand, $\vec{F}_{ext} = \text{curl}\, \vec{A}$, then a stationnary solution is $P(\vec{x}) = \text{cste} = \frac{1}{\text{Volume}}$. Indeed $\vec{j} = \frac{1}{V} \mu\, \text{curl}\, A - \vec{0} = \frac{\mu}{V} \text{curl}\, \vec{A}$ fulfills $\text{div}\, \vec{j} = 0$. Here, $\vec{j} \neq \vec{0} \Rightarrow$ this is not an equilibrium solution.

4) d) Unicity of the long-time solution

We will show that 2 solutions (with different initial conditions) of the FP eq

$$\partial_t P = L P \quad \text{where} \quad L P = -\partial_x (M_1(x) P) + \partial_x^2 (M_2(x) P)$$

become the same at late times. This relies on $M_2(x) > 0$. Take two such solutions $P_1(x,t)$ and $P_2(x,t)$ and study the evolution of their Kullback-Leibler "distance" $D(t) = \int P_2 \log(P_2/P_1)\, dx$. Notation $R(x,t) = \frac{P_2(x,t)}{P_1(x,t)}$

$$\frac{dD}{dt} = \int dx\, (\partial_t P_2) \log \frac{P_2}{P_1} + \int dx\, P_2 \left[\frac{\partial_t P_2}{P_2} - \frac{\partial_t P_1}{P_1} \right]$$

$$\int dx\, \partial_t P_2 = \partial_t \int P_2 = 0$$

$$= \int dx\, (\log R) \partial_t P_2 - \int dx\, R \partial_t P_1$$

$$= \underbrace{\int dx\, (\log R) L P_2}_{\int dx\, (L^+ \log R) P_2} - \int dx\, R L P_1 \qquad \text{as found with integration by parts}$$

$$\text{with} \quad L^+ = M_1(x) \partial_x + M_2(x) \partial_x^2$$

$$L^+ \log R = M_1 \partial_x \log R + M_2 \partial_x^2 \log R$$

$$= M_1 \frac{1}{R} \partial_x R + M_2 \partial_x \left(\frac{1}{R} \partial_x R \right)$$

$$= \frac{1}{R} L^+ R - M_2 \partial_x R\, \partial_x \frac{1}{R} = \frac{L^+ R}{R} - M_2 \left(\partial_x \log R \right)^2$$

$$\frac{dD}{dt} = \int dx \left(\frac{1}{R} P_2\right)_{P_1} L^+ R - \int dx\, P_2 M_2 \left(\partial_x \log R\right)^2 - \int dx\, RLP_1$$

$$= \int dx\, RLP_1 - \int dx\, P_2 M_2(x) \partial_x \left(\log R\right)^2 - \int dx\, RLP_1$$

$$\leqslant 0$$

Since $D \geqslant 0$ and $\frac{dD}{dt} \leqslant 0$, $D \to cst \Rightarrow \frac{dD}{dt} \to 0$ for $t \to \infty$.

Thus $\log R$ becomes independent from $x \Rightarrow R \to cst$ ie $P_2 \to C P_1$.

But $\int P_2\, dx = \int P_1\, dx \Rightarrow C = 1 \Rightarrow P_2/P_1 \to 1$ for $t \to \infty$

Thus : the long time solution is unique.

The result carries over to higher dimension, as long as the matrix $\overleftrightarrow{M}_2$ is definite positive (in general, it is only positive : $\sum_{ij} \langle \Delta x_i \Delta x_j \rangle v_i v_j \geqslant 0$, $\forall \vec{v}$)

Indeed, we find in the vectorial case

$$\frac{dD}{dt} = -\int d\vec{x}\, P_2(\vec{x}) \left(\overleftrightarrow{M}_2\right)_{ij} \frac{\partial \ln R}{\partial x_i} \frac{\partial \ln R}{\partial x_j} \leqslant 0.$$

The proof proposed here bears similarities with **Boltzmann's H theorem**

Application to Kramers' equation : $\overleftrightarrow{M}_2 = \begin{pmatrix} 0 & 0 \\ 0 & \Gamma \end{pmatrix}$ positive but not definite

Hence the argument does not apply as such.

Yet : there is a way, not direct. We learn from the fact that $D \to cst$ that $\partial_v \ln R \to 0$: all solutions have same v dependence. We know that the eq solution is in $\exp(-\beta m v^2/2)$, Thus $P(x,v,t) \overset{t\to\infty}{\sim} f(x,t)\, e^{-\beta m v^2/2}$

Plug this into Kramers : $\partial_t f + v \partial_x f + \frac{F}{m}\left(-\beta f v m\right) = 0$

But f does not depend on $v \Rightarrow \partial_t f = 0$ and $\partial_x f = \beta F f = -\beta f \partial_x U_{ext}$

$\Rightarrow f(x) \propto e^{-\beta U_{ext}(x)}$

Hence unicity ... here also ... at late time (equilibrium in the present case).

Equilibrium or not ?

We saw at the level of 1D Langevin equation that equilibrium in one ext potential $U(x)$ requires $D = \mu k T$, relating diffusion coefficient and mobility. In other words, for any D and μ, there will be a temp that ensures equality. This is wrong in higher dimensions. Consider an overdamped dynamics :

$$P = \frac{1}{Z} e^{-\beta H} \Rightarrow \vec{\nabla} P = \beta \vec{F} P$$

$$\vec{J} = \left(\overleftrightarrow{\mu} \vec{F} - \beta \overleftrightarrow{D} \vec{F} \right) P$$

Even if $\overleftrightarrow{D} = \overleftrightarrow{\mu} kT$, with non conservative forces, no equilibrium is possible

Take $\overleftrightarrow{D} = D \overleftrightarrow{I}$: $\vec{j} = D \left[\beta \vec{F}(\vec{x}) P - \vec{\nabla} P \right]$ and $\vec{curl}\, \vec{F} \neq \vec{0}$

$$\Rightarrow \frac{\vec{j}}{P} = D \beta \vec{F}(\vec{x}) - D \vec{\nabla} \ln P$$

$$\Rightarrow \vec{curl} \left(\frac{\vec{j}}{P} \right) = D \beta \vec{curl}\, \vec{F} \neq \vec{0}$$

and thus, it is not possible to get $\vec{j} = \vec{0}$: non conservative force $\Rightarrow$ no equilibrium

(which we recover from the contrapont : equil $\Rightarrow$ conservative force : indeed $\vec{j} = \vec{0}$ means $\vec{F} = kT \vec{\nabla} \ln P_{eq}$, ie $\vec{F}$ is conservative)

$$\mathcal{L}_x^{KM} = \sum_{n=1}^{\infty} (-1)^n \partial_x^n \left[M_n P \right] \quad , \quad \left(\mathcal{L}_x^{KM} \right)^+ = \sum_{n=1}^{\infty} (-1)^n (-1)^n M_n \partial_x^n P$$

$$\dot{\vec{x}} = \overleftrightarrow{\mu}\,\vec{F}(\vec{x}) + \vec{\xi}(t) \quad ; \quad \langle \xi_i(t)\rangle = 0$$

$\hookrightarrow$ mobility matrix $\qquad \langle \xi_i(t)\xi_j(t')\rangle = 2D_{ij}\,\delta(t-t')$

where i,j denote cartesian coordinates. Fokker-Planck reads $D_{ij}(\vec{x})$, ot, additive noise

$$\partial_t P(\vec{x},t) = -\partial_i\left(\mu_{ij} F_j P\right) + \partial_i\partial_j\left(D_{ij} P\right)$$ with Einstein's convention summation

Take $\vec{F} = -\vec{\nabla} U$; condition for Boltzmann's equilibrium to be a solution ($P \propto e^{-\beta U}$) is

$$\boxed{D_{ij} = kT\,\mu_{ij}}$$ the current $\vec{J} = +\overleftrightarrow{\mu}\,\vec{F}P - \overleftrightarrow{D}\,\vec{\nabla}P = \vec{0}$

If this relation does not hold, Boltzmann's eq is not a solution (and thus not the long-time solution) → not all dynamics are compatible with equilibrium

e) The backward Fokker-Planck equation

For a given initial condition ($x = x_0$ at $t = t_0$), we have written the FP equation:

$$\partial_t P(x,t|x_0,t_0) = -\partial_x\left[M_1(x)P\right] + \partial_x^2\left[M_2(x)P\right] \equiv \mathcal{L}_x P$$

For certain types of problems (first passage, absorbing boundary conditions), we need to integrate over x, and know the x_0, t_0 dependance, through some c_{t_0}.

To compute $\partial_{t_0} P(x,t|x_0,t_0)$, we could go back to Chapman-Kolmogorov and repeat calculations. It is not necessary. We can indeed show

$$\boxed{\partial_{t_0} P(x,t|x_0,t_0) = -M_1(x_0)\partial_{x_0}P - M_2(x_0)\partial_{x_0}^2 P = -\mathcal{L}_{x_0}^{\dagger} P}$$

Indeed, from Chapman-Kolmogorov:

$$P(x,t|x_0,t_0) = \int dy\; P(x,t|y,t_1)\,P(y,t_1|x_0,t_0) \quad , \text{ true } \forall\, t_1,\ t_0 \le t_1 \le t$$

$$\Rightarrow \frac{\partial P}{\partial t_1}(x,t|x_0,t_0) = 0 = \int dy\Big[\partial_{t_1}P(x,t|y,t_1)\,P(y,t_1|x_0,t_0) + P(x,t|y,t_1)\,\partial_{t_1}P(y,t_1|x_0,t_0)\Big]$$

$$\Rightarrow 0 = \int dy\Big[\partial_{t_1}P(x,t|y,t_1)\,P(y,t_1|x_0,t_0) + \{\mathcal{L}_y^{\dagger}P(x,t|y,t_1)\}\,P(y,t_1|x_0,t_0)\Big]\ \mathcal{L}_y P(y,t_1|x_0,t_0)$$

$$\Rightarrow \partial_{t_1} P(x,t|y,t_1) = -\mathcal{L}_y^{\dagger} P(x,t|y,t_1)$$

↳ Backward FP (a.k.a backward Kolmogorov eq). The "forward" and "backward" descriptions are equivalent. Note that one can retrieve the sign $-M_2\partial_{x_0}^2 P$ with basic diffusion ($M_1 = 0$, $M_2 = D = cst$), since the solution is in $\varphi(t-t_0) \propto \exp\left(-\frac{(x-x_0)^2}{4D(t-t_0)}\right)$

Hence $\partial_t \varphi = -\partial_{t_0}\varphi = D\partial_x^2\varphi = D\partial_{x_0}^2\varphi$.

Note also that the above applies to the Kramers-Moyal expansion

$$-\partial_{t_0} P(x,t|x_0,t_0) = \left(\mathcal{L}_{x_0}^{KM}\right)^{\dagger} = \sum_{n=1}^{\infty} M_n(x_0)\,\partial_{x_0}^n P(x,t|x_0,t_0)$$

5) Solving the Fokker-Planck equation : connexion to quantum mechanics

For specific processes, like Ornstein-Uhlenbeck, specific solutions can be found.
We will rather present a general method, based on **spectral analysis**, in the framework of the Smoluchowski eq, with a force $F = -\partial_x U$

$$\partial_t P = -\partial_x (\mu F P) + D\, \partial_x^2 P \equiv \mathcal{L}_{FP} P \qquad D = \mu k T$$

Noise is here additive (ie D does not depend on x), but method generalises to $D(x)$, ie to multiplicative noise. Since $D = \mu k T$, we have the equilibrium solution

$$P_{eq}(x) = \frac{1}{Z} e^{-\beta U(x)}$$, assuming $Z < \infty$ (U is "confining").

It is useful here to remember what we found for Markov chains, where $\vec{P}(t+1) = \overleftrightarrow{M}\, \vec{P}(t)$, where we mentionned the connexion {eq exists} $\Leftrightarrow$ {detailed balance holds} : $M_{ij} P_j^{eq} = M_{ji} P_i^{eq}$ (no summation over repeated ind.)

$$\Leftrightarrow M_{ij} \sqrt{P_j^{eq}/P_i^{eq}} = M_{ji} \sqrt{P_i^{eq}/P_j^{eq}}$$

which brings a new operator / matrix $\widetilde{M}_{ij} = \frac{1}{\sqrt{P_i^{eq}}} M_{ij} \sqrt{P_j^{eq}} = \widetilde{M}_{ji}$ (*),
symmetric, thus more convenient to handle than the original M.

$$\vec{P}(t+1) - \vec{P}(t) = \left(\overleftrightarrow{M} - \overleftrightarrow{I}\right) \vec{P}(t) \qquad // \quad \partial_t P = \mathcal{L}_{FP} P \Rightarrow \boxed{\mathcal{L}_{FP} \leftrightarrow \overleftrightarrow{M} - \overleftrightarrow{I}}$$

In particular, since all eigenvalues of $\overleftrightarrow{M}$ are ≤ 1, we expect those of $\mathcal{L}_{FP}$ to be ≤ 0
The main difficulty with $\mathcal{L}_{FP}$ is that it is not self-adjoint

$$\mathcal{L}_{FP} \neq \mathcal{L}_{FP}^{+} = \mu F(x) \partial_x + D \partial_x^2$$

Defining a self-adjoint operator from $\mathcal{L}_{FP}$ (if it exists) is // to defining a symmetric matrix $\widetilde{M}$ from M ; (*) suggests to define a new function $g(x,t)$ s.t.

$$P(x,t) = \sqrt{P^{eq}(x)}\, g(x,t)$$

This works ! The evolution equation for g is simpler than that for P :

$$\partial_t g = -H g \quad , \quad \text{with } H^{+} = H \text{, self-adjoint.}$$

$$\frac{1}{D} H = -\partial_x^2 + \underbrace{\left(\frac{1}{2}\partial_x \beta U\right)^2 - \frac{1}{2}\partial_x^2 \beta U}_{\equiv V_{eff}(x)}$$

Schrödinger eq in imaginary time in a potential $V_{eff}(x)$

Proof: we inject $P(x,t) = Z^{-1/2} e^{-\beta U/2}\, g(x,t)$ into F-P.

$$e^{-\beta U/2}\partial_t g = \mu(\partial_x^2 U)e^{-\beta U/2} g + \mu(\partial_x U)\partial_x\left(e^{-\beta U/2} g\right) + D\left[g\,\partial_x^2 e^{-\beta U/2} + 2(\partial_x g)\partial_x\left(e^{-\beta U/2}\right) + (\partial_x^2 g)e^{-\beta U/2}\right]$$

$\left.\begin{array}{l}\tilde{U} = \beta U \\ \mu = \beta D\end{array}\right\}$

$$\frac{1}{D} e^{-\tilde{U}/2}\partial_t g = (\partial_x^2 \tilde{U})e^{-\tilde{U}/2} g + (\partial_x \tilde{U})\left[-\frac{1}{2}(\partial_x \tilde{U}) g\, e^{-\tilde{U}/2} + e^{-\tilde{U}/2}\partial_x g\right] + \frac{1}{2} g\,\partial_x\left(-(\partial_x \tilde{U})e^{-\tilde{U}/2}\right) - 2(\partial_x g)\frac{1}{2}(\partial_x U)e^{-\tilde{U}/2} + e^{-\tilde{U}/2}\partial_x^2 g$$

$$\partial_t g \;\frac{1}{D} = \partial_x^2 g + g\,\partial_x^2 \tilde{U} - \frac{g}{2}(\partial_x \tilde{U})^2 + \frac{1}{2} g\left(-\partial_x^2 \tilde{U} + \frac{1}{2}(\partial_x \tilde{U})^2\right).$$

$$= \partial_x^2 g + g\left[-\frac{1}{4}(\partial_x \tilde{U})^2 + \frac{1}{2}\partial_x^2 \tilde{U}\right] \quad ✓$$

This mapping to a quantum mechanics problem is convenient. We denote $\psi_\lambda(x)$ the eigenfunctions of H: $H\psi_\lambda = \lambda\psi_\lambda$.

$$g(x,t) = \sum_\lambda c_\lambda(t)\,\psi_\lambda(x) \quad \text{and} \quad \partial_t g = -Hg$$

$$\Rightarrow\; g(x,t) = \sum_\lambda c_\lambda(0)\, e^{-\lambda t}\,\psi_\lambda(x)$$

Besides: $\frac{H}{D} = Q^+ Q$ with $Q = \partial_x + \frac{1}{2}\partial_x \tilde{U}$, simple structure

$$Q^+ = -\partial_x + \frac{1}{2}\partial_x \tilde{U}$$

Proof: $Q^+ Q\, g = \left(\partial_x + \frac{1}{2}\partial_x \tilde{U}\right)\left(+\partial_x g + \frac{g}{2}\partial_x \tilde{U}\right)$

$$= -\partial_x^2 g - \frac{1}{2}(\partial_x g)\partial_x \tilde{U} - \frac{1}{2} g\,\partial_x^2 \tilde{U} + \frac{1}{2}\partial_x \tilde{U}\,\partial_x g + \frac{1}{4} g(\partial_x \tilde{U})^2 \quad ✓$$

Consequence: all eigenvalues of H are positive; indeed, using bra-ket notation

$$\langle \psi_\lambda | Q^+ Q | \psi_\lambda \rangle = \| Q|\psi_\lambda\rangle \|^2 = \lambda \langle \psi_\lambda | \psi_\lambda \rangle = \lambda \|\psi_\lambda\|^2$$

ground state of H: and there exists a 0 eigenvalue, associated to $\psi_0(x) \propto e^{-\tilde{U}(x)/2}$ since $Q\psi_0 = 0$. With proper normalization: $\psi_0(x) = \frac{1}{\sqrt{Z}} e^{-\tilde{U}/2} = \sqrt{P_{eq}(x)}$

Under "Perron-Frobenius-like" conditions, this 0 eigenvector is non-degenerate, which guarantees existence, unicity of the steady state (equilibrium), + convergence

If we single out in the summation the 0 eigenvalue from rest of spectrum:

$$P(x,t) = \sqrt{P_{eq}(x)}\; g(x,t) = \sqrt{P_{eq}}\left[\sqrt{P_{eq}} + \sum_{\lambda>0} c_\lambda(0)\, e^{-\lambda t}\,\psi_\lambda(x)\right]$$

$$= P_{eq}(x) + \sum_{\lambda>0} c_\lambda(0)\,\psi_0(x)\,\psi_\lambda(x)\exp(-\lambda t)$$

which gives the asymptotic relaxation rate toward equilibrium as $\tau = 1/\lambda_1$ where λ_1 is the energy of the first excited state of Hamiltonian H.

To loop the loop, it is instructive to analyze the form of $P(x,t)$ when the initial condition is $P(x,t) = \delta(x-x_0)$, ie we work with $P(x,t|x_0)$.

$$P(x,t|x_0) = \sum_\lambda \tilde{c}_\lambda(x_0)\,\psi_0(x)\,\psi_\lambda(x)\,e^{-\lambda t}$$

ie $c_\lambda(0)$ depends on x_0 and is denoted $\tilde{c}_\lambda(x_0)$

We take advantage of completeness

$$\sum_\lambda \psi_\lambda(x)\,\psi_\lambda(x') = \delta(x-x')$$

$$P(x,0|x_0) = \delta(x-x_0) = \sum_\lambda \tilde{c}_\lambda(x_0)\,\psi_0(x)\,\psi_\lambda(x) = \sum_\lambda \psi_\lambda(x_0)\,\psi_\lambda(x)$$

and we see that $\tilde{c}_\lambda(x_0) = \frac{1}{\psi_0(x_0)}\psi_\lambda(x_0)$ is a solution: $\delta(x-x_0)$

$$P(x,0|x_0) = \sum_\lambda \frac{1}{\psi_0(x_0)}\psi_\lambda(x_0)\,\psi_0(x)\,\psi_\lambda(x) = \frac{\psi_0(x)}{\psi_0(x_0)}\sum_\lambda \psi_\lambda(x_0)\psi_\lambda(x) = \delta(x-x_0)$$

$$\Rightarrow \boxed{P(x,t|x_0) = \frac{\psi_0(x)}{\psi_0(x_0)}\sum_\lambda \psi_\lambda(x)\,\psi_\lambda(x_0)\,e^{-\lambda t}}$$

What we see is that $P(x,t|x_0)\,\psi_0(x_0)/\psi_0(x)$ is invariant if $x \leftrightarrow x_0$:

$$P(x,t|x_0)\frac{\psi_0(x_0)}{\psi_0(x)} = P(x_0,t|x)\frac{\psi_0(x)}{\psi_0(x_0)} \qquad \text{and since } \psi_0^2(x) = P_{eq}(x)$$

$$\boxed{P(x,t|x_0)\,P_{eq}(x_0) = P(x_0,t|x)\,P_{eq}(x)} \quad \text{ie detailed-balance holds}$$

Indeed, $P(x,t|x_0)$ can be viewed as a purely equilibrium property, although is also is relevant to describe processes outside equilibrium. We remember here Wigner's result showing that equilibrium coarse-grained dynamics obeys detailed bal.

Rk the connexions between operators are interesting. One can show that

$$\boxed{H = -\,e^{\beta U/2}\,\mathcal{L}_{FP}\,e^{-\beta U/2}} \qquad // \quad \tilde{M}_{ij} = \sqrt{(P_i^{eq})^{-1}}\,M_{ij}\,\sqrt{P_j^{eq}}$$

given also that

$$\boxed{\mathcal{L}_{FP} = +D\,\partial_x\left[e^{-\beta U}\,\partial_x\,e^{\beta U}\right]} \qquad e^{\beta U}\,e^{-\beta U/2}$$

we get the expressions of Q / Q^+

$$\frac{H}{D} = \underbrace{-\,e^{\beta U/2}\,\partial_x\left[e^{-\beta U/2}}_{Q^+}\underbrace{e^{-\beta U/2}\,\partial_x\,e^{\beta U/2}}_{Q}\right]; \quad Q = \partial_x + \frac{1}{2}\partial_x(\beta U)$$

Besides, H and $\mathcal{L}_{FP}$ have the same spectrum. While it is not obvious that $\mathcal{L}_{FP}$ is diagonalizable since $\mathcal{L}_{FP} \neq \mathcal{L}_{FP}^+$, we have $H = H^+$

$$H|\psi_\lambda\rangle = \lambda|\psi_\lambda\rangle \Rightarrow -\,e^{\beta U/2}\,\mathcal{L}_{FP}\,e^{-\beta U/2}|\psi_\lambda\rangle = \lambda|\psi_\lambda\rangle$$

ie $\mathcal{L}_{FP}\underbrace{e^{-\beta U/2}|\psi_\lambda\rangle}_{\text{right eigenvector of } \mathcal{L}_{FP}} = -\lambda\,e^{-\beta U/2}|\psi_\lambda\rangle$, with eigenvalue $-\lambda$.

We can remember the mapping between a stochastic dynamics governed by the Fokker-Planck equation and a quantum dynamics governed by Schrödinger equation. In this analogy, equilibrium ↔ ground state, dynamics involve excited states.

For an observable $\mathcal{O}$

$$\langle \mathcal{O} \rangle_{eq} = \frac{1}{Z} \int \mathcal{O}(x)\, e^{-\beta U(x)}\, dx = \int dx\, \langle x|\psi_0\rangle^2\, \mathcal{O}(x)$$
$$= \int dx\, \langle \psi_0|x\rangle\, \mathcal{O}(x)\, \langle x|\psi_0\rangle$$
$$= \langle \psi_0 | \hat{\mathcal{O}} | \psi_0 \rangle \quad \text{with} \quad \hat{\mathcal{O}} = \int dx\, |x\rangle\langle x|\, \mathcal{O}(x)$$

Averages at equilibrium ↔ averages in the ground state for the quantum problem.

6° Relation between Langevin and Fokker-Planck equations : back to the Itô-Stratonovich dilemma

Consider a Langevin eq with additive noise

$$\dot{\vec{x}} = \mu \vec{F}(\vec{x}) + \vec{R}(t)$$
$$\langle \vec{R}(t)\, \vec{R}(t+t') \rangle = 2 \overleftrightarrow{D}\, \delta(t') \quad \text{where } \overleftrightarrow{D} \text{ is a constant matrix.}$$

Then $\partial_t P(\vec{x},t) = -\mu \vec{\nabla}_{\vec{x}} \cdot \left[\vec{F} P \right] + \overleftrightarrow{D} : \vec{\nabla}_x \vec{\nabla}_x P$

and there is an equivalence between the 2 descriptions, Langevin & F-P.

Yet, in our derivation of F-P from the master eq, we met formally the case $\overleftrightarrow{D}(\vec{x})$. In 1d → $\partial_t P = -\partial_x (\mu F P) + \partial_x^2 (D(x) P)$.

What is the corresponding Langevin eq ?

The answer is not unique. To see it, consider a multiplicative noise situation

$$\dot{x} = A(x) + C(x) R(t) \quad \text{where } \langle R(t) R(t') \rangle = 2D \delta(t-t')$$

We know that a Langevin eq acquires a meaning only when it is discretized. The problem is that there is not a unique way to discretize it : we are back to the Itô / Stratonovich choices, which here, are no longer equivalent.

Indeed, we have to give a value to

$$\int_t^{t+\Delta t} C(x(t'))\, R(t')\, dt' \quad , \text{ which is closely related to the question}$$

of computing $\langle C(x(t))\, R(t) \rangle$ → Itô and Strato give ≠t results.

Should we choose $C(x(t)) \int_t^{t+\Delta t} R(t')\, dt'$, or $\frac{1}{2}\left[C(x(t)) + C(x(t+\Delta t)) \right] \int_t^{t+\Delta t} R(t')\, dt'$, or something else ? This is not the same → lead to ≠t Fokker-Planck.

Note that a F-P equation is not poly-interpretable (this a simple PDE).

The ambiguity is only present at Langevin level.

We generalize Itō and Stratonovich choices by introducing a parameter q such that

$$x(t+\Delta t) = x(t) + A(x(t))\Delta t + C\left[q\,x(t) + (1-q)\,x(t+\Delta t)\right] \int_t^{t+\Delta t} R(t')\,dt'$$

$q=1$ corresponds to Itō, and $q=\frac{1}{2}$ to Stratonovich. Here, we cannot use the $B_{\Delta t}$ rules of stochastic calculus learned earlier, because they hold for a Langevin eq with additive noise. We have to compute explicitly the moments $\langle \Delta x \rangle$ and $\langle (\Delta x)^2 \rangle$ to construct the Fokker-Planck equation:

$$\Delta x = x(t+\Delta t) - x(t) = A(x(t))\Delta t + C\left[q\,x(t) + (1-q)\,x(t+\Delta t)\right] B_{\Delta t}$$

$$C\left[q\,x(t) + (1-q)\,x(t+\Delta t)\right] \simeq C\left\{q\,x(t) + (1-q)\left[\underbrace{A(x(t))\Delta t + C(x(t))B_{\Delta t}}_{x(t+\Delta t)}\right]\right\}$$

$$\simeq C\left[x(t) + (1-q)A(x(t))\Delta t + (1-q)C(x(t))B_{\Delta t}\right]$$

$$\simeq C\left[x(t)\right] + \left[(1-q)A(x(t))\Delta t + (1-q)C(x(t))B_{\Delta t}\right] C'(x(t))$$

$$\Rightarrow \langle \Delta x \rangle = A(x)\Delta t + \Big\langle \big(\text{“} \text{“} \text{“} \text{“} \big) B_{\Delta t} \Big\rangle$$

$$= A(x)\Delta t + 0 + 0 + (1-q)C(x)C'(x)\underbrace{\langle B_{\Delta t}^2 \rangle}_{2D\Delta t} + o(\Delta t)$$

$$= A(x)\Delta t + 2D\,C(x)C'(x)\,\Delta t\,(1-q)$$

$$\langle (\Delta x)^2 \rangle \simeq C^2\left[x(t)\right] \langle B_{\Delta t}^2 \rangle = 2D\,C^2(x)\,\Delta t$$

$$\Rightarrow \partial_t P(x,t) = -\partial_x \left\{ \left[A(x) + 2(1-q)D\,C(x)C'(x)\right] P \right\} + D\,\partial_x^2 \left[C^2(x) P\right]$$

This F.P eq can be rewritten

$$\boxed{\partial_t P(x,t) = -\partial_x \left[A(x) P\right] + D\,\partial_x \left[C^{2(1-q)}\, \partial_x \left(C^{2q} P \right) \right]}$$

$D\,\partial_x^2 \left[C^2(x) P\right]$ for Itō ($q=1$) ; $D\,\partial_x \left[C(x)\,\partial_x (C(x) P) \right]$ for Stratonovich ($q=1/2$)

Itō and Stratonovich no longer lead to the same physics: it is because they lead to different discrete equations, while in ch IV, the discrete equations were the same.

Yet, we recover a feature already met: Stratonovich is compatible with basic calculus.

Indeed: $\dot{x} = A(x) + C(x)\,R(t)$

$\hookrightarrow \dfrac{\dot{x}}{C(x)} = \dfrac{A(x)}{C(x)} + R(t)$ which is an additive noise Langevin eq for $\bar{x}(t)$ where $d\bar{x} = \dfrac{1}{C(x)}\,dx$, with a drift $\bar{A}(\bar{x}) = \dfrac{A(x)}{C(x)}$

From $\dot{\bar{x}} = \bar{A}(\bar{x}) + R(t)$, we deduce the F.P for the pdf of $\bar{x}$: $\bar{P}(\bar{x},t)$

$$\partial_t \bar{P}(\bar{x},t) = -\partial_{\bar{x}} \left[\bar{A}\,\bar{P}\right] + D\,\partial_{\bar{x}}^2 \left[\bar{P}\right]$$

which can be recast into a PDE for $P(x,t)$ where $P(x,t)dx = \overline{P}(\overline{x},t)\,d\overline{x}$

$$\overline{P}(\overline{x},t) = C(x)\,P(x,t) \qquad \text{and} \qquad \partial_{\overline{x}} = C(x)\partial_x$$

thus $\partial_t\left[C(x)P(x,t)\right] = -C(x)\partial_x\left[\frac{A}{C}CP\right] + DC(x)\partial_x\left[C(x)\partial_x(CP)\right]$

$$\partial_t P(x,t) = -\partial_x\left(C(x)P\right) + D\,\partial_x\left[C(x)\partial_x\left(C(x)P\right)\right]$$

which is Stratonovich result. We are no longer surprised to see that Itô type of calculations are not compatible with standard calculus. In particular, the change of variables above ($x \to \overline{x}$) belongs to std calculus. We also recover that in a physical process where noise $R(t)$ has a finite although small correlation time τ_c, Stratonovich route is more natural, and even, that taking $\tau_c \to 0^+$ yields Stratonovich solution. Situations where Itô applies naturally are found outside physics but are less frequent in physics (see nevertheless the discussion in van Kampen, IX · 5)

Thus, when a "prescription" has been chosen (ie, a q-value → "une grille de lecture de l'équation de Langevin"), there is an equivalence between Langevin ⇔ Fokker-Planck. This equivalence is lost in dimensions > 1, see Risken **3.4.1**