Stochastic Equations and Processes in Physics and Biology

Andrey Pototsky

Swinburne University of Technology, Melbourne, Australia

I. PROBABILITY: BASIC CONCEPTS AND DEFINITIONS

For a random variable X, we define the average (mean) $\langle X\rangle$ according to

$$\langle X\rangle = \lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n} X_i, \tag{1}$$

where X_i are particular realizations of X, i.e. numbers, recorded in an experiment, and n is the total number of random experiments.

Probability: discrete case

In case when X is a discrete variable, the probability $P(X)$ is defined as the limit of the occurrence frequency

$$P(X_i) = \lim_{n\to\infty}\frac{1}{n}\#\{X = X_i\}, \tag{2}$$

where $\#\{X = X_i\}$ denotes the total number of experiments with the outcome $X = X_i$. The probability is normalized

$$\sum_{X_i\in\Omega} P(X_i) = 1, \tag{3}$$

where Ω denotes the set of all possible values of X.

Clearly, the average $\langle X\rangle$, is linked to the probability via

$$\langle X\rangle = \sum_{X_i\in\Omega} X_i P(X_i), \tag{4}$$

(prove it).

Similarly, the average of any given function $Y = f(X)$ is found as

$$\langle Y\rangle = \sum_{X_i\in\Omega} f(X_i)P(X_i). \tag{5}$$

Probability: continuous case

If the variable X changes contiguously in some interval $X \in [a, b]$, one defines the *probability density* $\rho(x)$

$$\rho(x) = \lim_{n\to\infty,\ \delta x\to 0}\frac{\#\{X\in[x, x+\delta x]\}}{n\,\delta x}. \tag{6}$$

In other words, the probability to find X in a very narrow interval $[x, x+\delta x]$ is given by

$$P(X\in[x, x+\delta x]) = \rho(x)\delta x. \tag{7}$$

The normalization in the continuous case is written as

$$\int_a^b \rho(x)\delta x = 1. \tag{8}$$

With this definition of $\rho(x)$, the average $\langle X \rangle$ is found as

$$\langle X \rangle = \int_a^b x\rho(x)\,dx. \tag{9}$$

(prove it). As before, for any $Y = f(x)$, we have

$$\langle Y \rangle = \int_a^b f(x)\rho(x)\,dx. \tag{10}$$

Distribution function

Alongside with the distribution density $\rho(x)$, one also considers the distribution function $F(x)$, defined as the probability that X is smaller then x, i.e.

$$F(x) = P(X \leq x). \tag{11}$$

For a continuous variable, the relation between $\rho(x)$ and $F(x)$ is as follows

$$F(x) = \int_{-\infty}^x \rho(x)\,dx, \qquad \rho(x) = F'(x). \tag{12}$$

Note that $F(x)$ and $\rho(x)$ are defined on the whole line $x \in (-\infty, +\infty)$.

Change of variables

Let us answer the following question: if $X \in [a, b]$ is continuous random variable with known distribution density, what would be the distribution density of a new variable $Y = f(X)$?

One way to find the unknown distribution density $p(y)$ is to use the expression for the average

$$\langle Y \rangle = \int_a^b f(x)\rho(x)\,dx. \tag{13}$$

On the other hand, we should have

$$\langle Y \rangle = \int_A^B y\,p(y)\,dy, \tag{14}$$

with $A = f(a)$ and $B = f(b)$. Changing the variables according to $dy = f'(x)dx$, we have

$$\langle Y \rangle = \int_a^b f(x)\rho(x)\,dx = \int_A^B y\rho(f^{-1}(y))\,\frac{dy}{f'(f^{-1}(y))}, \tag{15}$$

where we have used the notation $x = f^{-1}(y)$. Consequently,

$$p(y) = \rho(f^{-1}(y)) \frac{1}{f'(f^{-1}(y))}. \tag{16}$$

Examples

Q1. You are writing a computer code for generating random numbers $q \geq 0$, with a given distribution density $\rho(q)$. The only built-in generator of random numbers that is available is the generator of uniformly distributed random numbers x from the interval $x \in [0, 1]$.

1. Find the conversion formula $q = q(x)$

2. Give the answer for the exponential distribution $\rho(q) = \alpha\, e^{-\alpha q}, \quad \alpha > 0$

answer:

1. $x = \int_{-\infty}^{q} \rho(q)\, dq = F(q), \quad q = F^{-1}(x)$

2. $q = -\alpha^{-1} \ln(1 - x)$

Q2: Relativistic Brownian motion: electrons in Graphene

Charge carriers in graphene behave like two-dimensional relativistic mass less particles. It is known that the equilibrium distribution of the momentum of electrons is given by the relativistic Jüttner distribution (*Dunkel and Hänggi, Phys. Reports (2009)*)

$$\rho(p_x, p_y) = C \exp\left(-\frac{\sqrt{p_x^2 + p_y^2}}{kT}\right) \tag{17}$$

Generate (p_x, p_y) using two uniformly distributed random variables $u_1, u_2 \in [0, 1]$.

solution

The normalization condition yields

$$\begin{aligned} 2\pi C(kT)^2 \int_0^\infty e^{-x} x\, dx = 1, \quad C = (2\pi k^2 T^2)^{-1} \\ u_1 = F(x = \sqrt{p_x^2 + p_y^2}/kT) = 1 - (1 + x)e^{-x}, \\ p_0 = \sqrt{p_x^2 + p_y^2} = kT\left(-1 - LambertW\left(\frac{u_1 - 1}{e}\right)\right), \\ (p_x, p_y) = p_0(\cos 2\pi u_2, \sin 2\pi u_2). \end{aligned} \tag{18}$$

Q3: Ideal gas of active particles. Consider a particle, which moves with a constant absolute velocity v_0 in a direction that changes randomly in time. For a gas of such active particles with a given concentration ρ_0, the distribution of the direction of motion is uniform.

1. Find the average number of hits per unit time with a surface area S, assuming that the interaction with the surface does not change the velocity distribution.

2. Find the distribution of the absolute relative velocity $U =\mid \vec{u}_1 - \vec{u}_2 \mid$ under the same assumption.

answer:

1. $\langle N \rangle = S\rho_0 \int_0^{v_0} v_x p(v_x)\, dv_x = S\rho_0 \frac{v_0}{\pi}$ $(2D)$, $\langle N \rangle = S\rho_0 \frac{v_0}{4}$ $(3D)$, $\langle N \rangle = S\rho_0 \frac{v_0}{2}$ $(1D)$

2. $p(U = 0) = 0.5,\ p(U = 2v_0) = 0.5,\ (1D),\ U = \sqrt{2v_0^2 + 2v_0 u_x},\ p(U) = (\pi v_0)^{-1}(1 - (U/2v_0)^2)^{-1/2}$ $(2D)$, $p(U) = (U/2v_0^2)$ $(3D)$,

Conditional and joint probability

The whole probability space Ω can be divided into various sub-spaces $A, B, C, D, \ldots \in \Omega$. Every such subspace has its own probability, say $P(A)$. For any two $A \in \Omega$ and $B \in \Omega$ introduce the so-called *joint* probability $P(A \cap B)$, defined as the probability of the simultaneous occurrence of A and B.

Later, we will also be using the notation $P(A \cap B) = P(A; B)$.

If $P(B) \neq 0$, we define the conditional probability $P(A \mid B)$

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}. \tag{19}$$

Visualize with a diagram.

Example

You know that your neighbor has two children, one of which is a boy, what is the probability that the other one is also a boy?

answer: $p = 1/3$.

II. STOCHASTIC PROCESSES AND RANDOM WALK

Loosely speaking, any stochastic process is a probabilistic time series. Consider a system, whose dynamics is described by a certain time-dependent random variable $x(t)$.

Introduce the joint probability to observe the values of $x_1, x_2, \ldots$ at the respective times $t_1 > t_2 > \ldots$

$$p(x_1, t_1; x_2, t_2; \ldots) \tag{20}$$

For any two moments of time $t_1 > t_2$ we can define the conditional probability

$$p(x_1, t_1|x_2, t_2) = \frac{p(x_1, t_1; x_2, t_2)}{p(x_2, t_2)} \quad (21)$$

Note that

$$p(x_1, t_1) = \int_\Omega dx_2\, p(x_1, t_1; x_2, t_2), \quad (22)$$

where we integrate over the entire probability space Ω.

Ensemble average

The integration over x, as in the above equation, is associated with the averaging over the *ensemble* of different realizations of the stochastic process $x(t)$. We will denote the ensemble average by $\langle \ldots \rangle$. Thus, the conditional time-dependent average of x is given by

$$\langle x(t)|x_0, t_0\rangle = \int dx\, x\, p(x, t|x_0, t_0). \quad (23)$$

Visualize with a diagram

Markov processes:

For any $t_1 > t_2 > t_3 > t_4$, the probability at times t_1 and t_2 only conditionally depends on the state at time t_3, i.e.

$$p(x_1, t_1; x_2, t_2|x_3, t_3; x_4, t_4) = p(x_1, t_1; x_2, t_2|x_3, t_3). \quad (24)$$

As a consequence of this property, we have

$$p(x_1, t_1; x_2, t_2|x_3, t_3) = p(x_1, t_1|x_2, t_2)p(x_2, t_2|x_3, t_3). \quad (25)$$

Indeed

$$\begin{aligned} p(x_1, t_1|x_2, t_2)p(x_2, t_2|x_3, t_3) &= p(x_1, t_1|x_2, t_2; x_3, t_3)p(x_2, t_2|x_3, t_3) \\ &= \frac{p(x_1, t_1; x_2, t_2; x_3, t_3)}{p(x_2, t_2; x_3, t_3)} \frac{p(x_2, t_2; x_3, t_3)}{p(x_3, t_3)} \\ &= p(x_1, t_1; x_2, t_2|x_3, t_3). \end{aligned} \quad (26)$$

The Chapman-Kolmogorov equation

From Eq. (25), one easily obtains

$$p(x_1, t_1|x_3, t_3) = \int_\Omega dx_2 \frac{p(x_1, t_1; x_2, t_2; x_3, t_3)}{p(x_3, t_3)} = \int_\Omega dx_2\, p(x_1, t_1|x_2, t_2)p(x_2, t_2|x_3, t_3). \quad (27)$$

the relation between the transitional probabilities (1|3) and (1|2), (2|3) is known as the Chapman-Kolmogorov equation.

Stationary processes

Process $x(t)$ is called stationary if for any ϵ, $x(t+\epsilon)$ has the same statistics as $x(t)$. Important properties of a stationary process:

$$\begin{aligned} \langle x(t) \rangle &= const \\ \langle x(t)x(t') \rangle &= f(t-t'). \end{aligned} \tag{28}$$

Ergodic processes

For an ergodic process, the averaging over time is equivalent to the averaging over the ensemble. *Note that ergodicity is stronger than stationarity.* Example:

$$x(t) = A, \;\; A \text{ is uniformly distributed in } [0;1] \tag{29}$$

Then, any realization is a straight line $x(t) = A_i$, but A_i are different for different realizations, implying that $\langle x(t) \rangle = 0.5$, whereas the average over time for every single realization is given by $\int dt\, x(t) = A_i$.

Simple random walk in one dimension

Consider a walker that moves along a line and makes just one step at a time of a fixed length Δ either to the left or to the right with equal probability. Additionally assume that the time between two subsequent steps is δt.

This rule sets up the conditional probability $p(x, t+1|y, t)$ as follows

$$p(x, t+\delta t|y, t) = \begin{cases} \frac{1}{2}, & \text{if} \quad | x - y |= \Delta \\ 0, & \text{otherwise} \end{cases} \tag{30}$$

For the sake of argument, say $x = 0$ at $t = 0$, implying that $p(x, 0) = \delta_{x,0}$. Using the Chapmap-Kolmogorov equation, determine the conditional probabilities $p(x, t|x = 0, 0)$ for all subsequent times $t > 0$. Visualize the results graphically for $t = \delta t, 2\delta t, 3\delta t, 4\delta t, 5\delta t$.

Generally, one can write the so-called *master equation*

$$p(x, t+\delta t|x_0, t_0) = \frac{1}{2}(p(x-\Delta, t|x_0, t_0) + p(x+\Delta, t|x_0, t_0)). \tag{31}$$

Consider now the so-called continuous limits, when δt, $\Delta \to 0$. In this case, we can approximately write

$$
\begin{aligned}
p(x,t+\delta t|x_0,t_0) &\approx p(x,t|x_0,t_0)+\partial_t p(x,t|x_0,t_0)\delta t = \frac{1}{2}(p(x-\Delta,t|x_0,t_0)+p(x+\Delta,t|x_0,t_0)),\\
\partial_t p(x,t|x_0,t_0) &= \frac{\Delta^2}{2\delta t}\frac{[(p(x-\Delta,t|x_0,t_0)+p(x+\Delta,t|x_0,t_0)-2p(x,t|x_0,t_0)]}{\Delta^2}\\
&\approx \frac{\Delta^2}{2\delta t}\frac{\partial^2 p(x,t|x_0,t_0)}{\partial x^2}. \qquad (32)
\end{aligned}
$$

Thus, the random walk boils down to a diffusion of a point along the line. The continuous limit holds if $D=\frac{\Delta^2}{2\delta t}$ remains constant as δt, $\Delta \to 0$. The quantity D is called the diffusion coefficient.

The solution of the diffusion equation is

$$
p(x,t|x_0,t_0)=\frac{1}{\sqrt{2\pi D(t-t_0)}}\exp\left(-\frac{(x-x_0)^2}{2D(t-t_0)}\right). \qquad (33)
$$

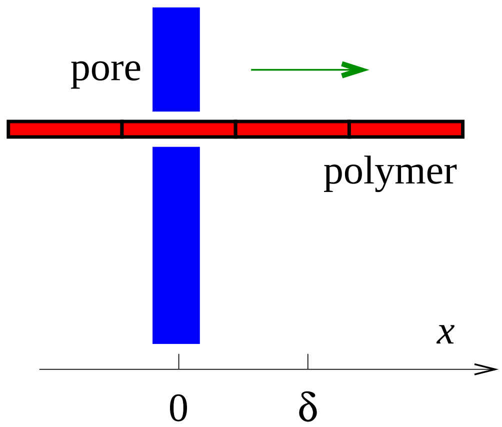

FIG. 1: Schematic diagram of the translocation ratchet. Shown is the pore with a segmented polymer inside, moving to the right. The length of each segment is δ.

Example: translocation ratchet, translocation of polymers

following C. S. Peskin et al, "Cellular Motion and Thermal Fluctuations: The Brownian Ratchet", Biophysical Journal **65** *316-324 (1993)*

Consider a protein, passing through a translocation pore. It is modeled by a long rod, which diffuses along the x axis. The rod is made of identical sections (segments) of the length δ (see Fig. 1). The pore acts as a perfect ratchet, only allowing the motion to the right. The flux of the rod, driven by the force f is given by

$$\phi = -\mu f \rho - D\partial_x \rho, \tag{34}$$

where ρ is the density, μ is the mobility of the rod and D is the diffusion coefficient. Note that the force f acts against the translocation direction.

The Fokker-Planck equation

$$\partial_t \rho + \partial_x \phi = 0 \tag{35}$$

is supplemented by the following boundary conditions

$$\phi(0,t) = \phi(\delta,t), \quad \rho(\delta,t) = 0. \tag{36}$$

These boundary conditions imply that we consider the motion of one segment between $x = 0$ and $x = \delta$. Each time a segment of the rod reaches the point $x = \delta$, it is removed from the system and is replaced by the new section at the point $x = 0$. In the stationary state, the flux ϕ is constant and the density ρ_s can be found exactly

$$\rho_s(x) = Ae^{(-\mu f x/D)} - \frac{\phi}{\mu f}, \tag{37}$$

with some unknown A and the unknown flux ϕ. From the second BC, we get

$$A = \frac{\phi}{\mu f} e^{(\mu f \delta/D)}. \tag{38}$$

The first BC holds automatically, because $\phi = const$.

From the normalization condition

$$\int_0^\delta \rho_s(x)\, dx \;=\; N, \quad \text{or} \tag{39}$$

$$\frac{\phi}{\mu f}\left[\frac{D}{\mu f}\left(e^{\mu f \delta/D} - 1\right) - \delta\right] \;=\; N \tag{40}$$

Finally, the average translocation velocity $v = (\text{flux}) \times \delta/N$

$$v = \frac{D}{\delta}\frac{\omega^2}{(e^\omega - 1) - \omega}, \tag{41}$$

with $\omega = \delta \mu f/D$. Interestingly, for vanishing forcing $f = 0$, the translocation velocity is $v = 2D/\delta$. The velocity v decreases monotonically with ω and reaches zero $v = 0$ at a certain stall load f_c.

For imperfect translocation ratchet as well as the polymerization ratchets, see the original paper.

Example: discrete walk on the line

Find the conditional distribution $p(i, t|0, 0)$ for a symmetric walker on an infinite line.

answer:

$$p(i, t|0, 0) = \left(\frac{1}{2}\right)^t C_t^{(i+t)/2} = \left(\frac{1}{2}\right)^t \frac{t!}{\left(\left(\frac{t}{2} - \frac{i}{2}\right)!\left(\left(\frac{t}{2} + \frac{i}{2}\right)!\right.\right.}. \tag{42}$$

The classical ruin problem, mean exit time.

Consider an asymmetric random walker that starts the motion in $x > 0$. The probability for the walker to move to the right and to the left are given by p and $q = 1 - p$, respectively. Assume that the walker is removed from the system as soon as he reaches either $x = 0$ or $x = a$. Question: what is the probability q_x of reaching $x = 0$ first?

Solution

The master equation is given by

$$q_x = pq_{x+1} + qq_{x-1} \tag{43}$$

with the boundary conditions $q_0 = 1, \;\; q_a = 0.$

The general solution of the difference equation is

$$\begin{aligned} q_x &= C_1 + C_2(q/p)^x, \quad \text{if} \quad q \neq p, \\ q_x &= C_1 + C_2 x \quad \text{if} \quad q = p. \end{aligned} \tag{44}$$

From the boundary conditions, we obtain

$$\begin{aligned} q_x &= \frac{(q/p)^a - (q/p)^x}{(q/p)^a - 1}, \quad \text{if} \quad q \neq p, \\ q_x &= 1 - x/a \quad \text{if} \quad q = p. \end{aligned} \tag{45}$$

What happens if $a \to \infty$?

Average duration of the game (Mean exit time)

Let D_x be the mean exit time if the walker starts in x. Then we can write

$$D_x = pD_{x+1} + qD_{x-1} + 1, \quad \text{with} \quad D_0 = D_a = 0. \tag{46}$$

The general solution of the non-homogeneous equation is

$$\begin{aligned} D_x &= \frac{x}{q-p} + C_1 + C_2(q/p)^x, \quad \text{if} \quad q \neq p, \\ D_x &= C_1 + C_2 x - x^2 \quad \text{if} \quad q = p. \end{aligned} \tag{47}$$

The solution that satisfies the boundary conditions is

$$\begin{aligned} D_x &= \frac{x}{q-p} - \frac{q}{q-p}\frac{1-(q/p)^x}{1-(q/p)^a}, \quad \text{if} \quad q \neq p, \\ D_x &= x(a-x) \quad \text{if} \quad q = p. \end{aligned} \tag{48}$$

III. BASIC CONCEPTS AND DEFINITIONS

Autocorrelation function and the Wiener-Khinchin theorem

Consider a time series $x(t)$ (signal). Assuming that this signal is known over an infinitely long interval $[-T, T]$, with $T \to \infty$, we can build the following function

$$G(\tau) = \lim_{T\to\infty} \frac{1}{T} \int_0^T dt x(t) x(t+\tau), \tag{49}$$

known as the *autocorrelation function* of the signal $x(t)$ (ACF).

ACF is an even function

Proof:

$$\begin{aligned} G(-\tau) &= \lim_{T\to\infty} \frac{1}{T} \int_0^T dt x(t) x(t-\tau) = \left\{ t-\tau = y |_{-\tau}^{T-\tau},\ \ dt = dy \right\} \\ &= \lim_{T\to\infty} \frac{1}{T} \int_{-\tau}^{T-\tau} dy x(\tau+y) x(y) = \lim_{T\to\infty} \frac{1}{T} \left[\int_{-\tau}^{0} (\ldots) + \int_0^T (\ldots) - \int_{T-\tau}^{T} (\ldots) \right] = G(\tau) \end{aligned} \tag{50}$$

The last equality holds in the limit $T \to \infty$.

The wiener-Khinchin theorem

This theorem plays a central role in the stochastic series analysis, since it relates the Fourier transform of $x(t)$ to the ACF.

Introduce a new function $S(\omega)$ according to

$$S(\omega) \ = \ \lim_{T\to\infty} \frac{1}{2\pi T} \mid \hat{x}(\omega) \mid^2, \tag{51}$$

where the forward Fourier transform of $x(t)$ is given by

$$\hat{x}(\omega) \ = \ \int_0^T dt e^{-i\omega t} x(t). \tag{52}$$

Note that if for real signal $x(t)$, the Fourier transform obeys the following symmetry

$$\hat{x}(-\omega) \ = \ \hat{\omega}^* \tag{53}$$

The function $S(\omega)$ is called *the power spectral density of* $x(t)$ (psd).

The following lines relate $S(\omega)$ to $G(\tau)$ (see Fig.2 for details)

$$\begin{aligned} S(\omega) &= \lim_{T\to\infty} \frac{1}{2\pi T} \int_0^T dt e^{-i\omega t} x(t) \int_0^T dt' e^{i\omega t'} x^*(t') \\ &= \{(t,t') \to (t', \tau = t - t')\} \end{aligned} \tag{54}$$

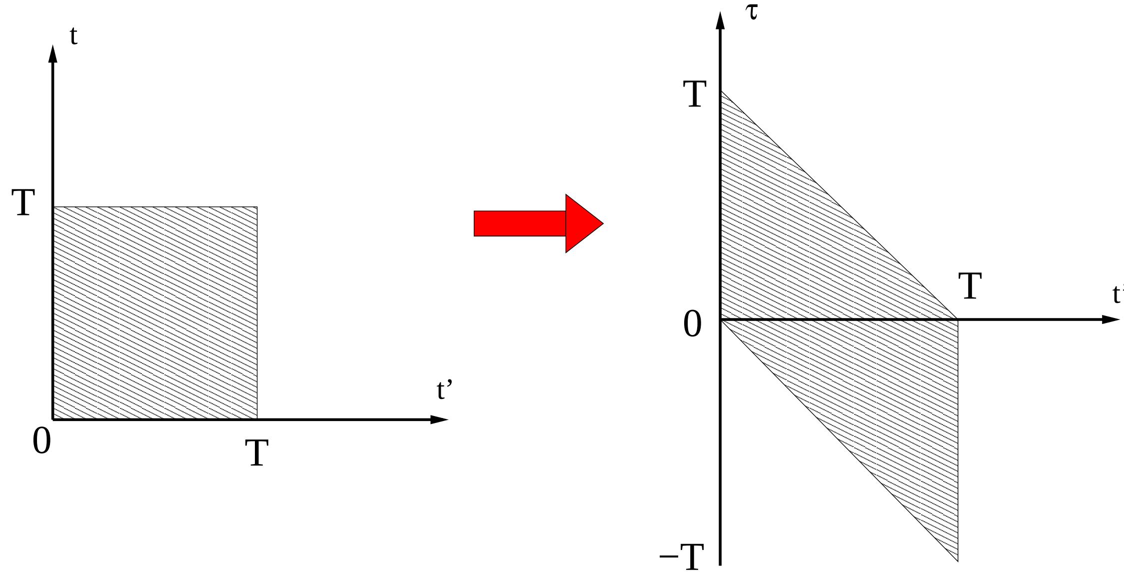

FIG. 2: Transformation of the integration domain under the change of variables $\{(t,t') \to (t', \tau = t - t')\}$.

$$
\begin{aligned}
&= \lim_{T\to\infty} \frac{1}{2\pi T}\left[\int_0^T d\tau e^{-i\omega\tau}\int_0^{T-\tau} dt' x(t')x^*(t'+\tau) + \int_{-T}^0 d\tau e^{-i\omega\tau}\int_{-\tau}^{T} dt' x(t')x^*(t'+\tau)\right] \\
&= \frac{1}{2\pi T}\left[\int_{-T}^0 d\tau e^{-i\omega\tau}(TG(\tau) + O(1/T)) + \int_0^T d\tau e^{-i\omega\tau}(TG(\tau)+O(1/T))\right] \\
&= \frac{1}{2\pi}\lim_{T\to\infty}\int_{-T}^{T} d\tau e^{-i\omega\tau} G(\tau).
\end{aligned}
$$

Because the ACF is an even function, we can write

$$S(\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} G(\tau)e^{\pm i\omega\tau}d\tau = \frac{1}{2\pi}\int_{-\infty}^{\infty} G(\tau)\cos\omega\tau d\tau = \frac{1}{\pi}\int_0^{\infty} G(\tau)\cos\omega\tau d\tau \tag{55}$$

Consequently, $S(-\omega) = S(\omega)$ and using the following representation of the delta-function

$$\delta(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{\pm i\omega x} d\omega, \tag{56}$$

we conclude that

$$G(\tau) = \int_{-\infty}^{\infty} e^{i\omega\tau} S(\omega) d\omega. \tag{57}$$

What happens in case of a deterministic signal?

Let $x(t) = A\cos\Omega t$, then the Fourier transform

$$\hat{x}(\omega) = \int_0^T dt e^{-i\omega t} A\cos\Omega t = A\frac{T}{2}\delta_{\Omega,\omega} + O(1). \tag{58}$$

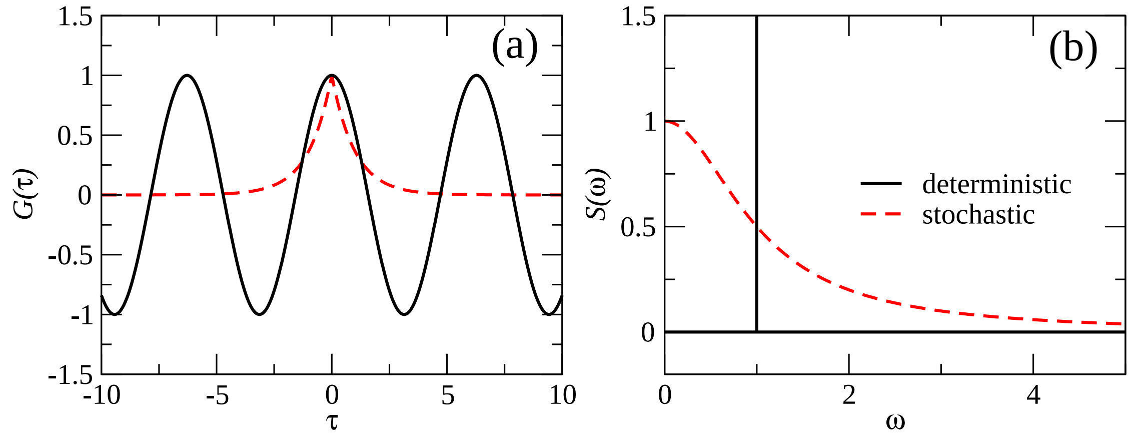

FIG. 3: (a) ACF $G(\tau)$ of a deterministic (dashed line) and a stochastic (solid line) signals. (b) The corresponding psd $S(\omega)$.

The psd

$$S(\omega) = \lim_{T\to\infty} \frac{1}{2\pi T}\hat{x}(\omega)\hat{x}^*(\omega) = \lim_{T\to\infty} \frac{A^2 T}{8\pi}\delta^2_{\Omega,\omega} + O(1) = O(T)\delta_{\Omega,\omega}. \tag{59}$$

Consequently, $S(\omega = \Omega)$ diverges as $T \to \infty$.

The ACF is found as

$$\begin{aligned} G(\tau) &= \lim_{T\to\infty} \frac{A^2}{T}\int_0^T dt\ \cos\Omega t\cos\Omega(t+\tau) \qquad (60)\\ &= \lim_{T\to\infty} \frac{A^2}{T}\int_0^T dt\ \left(\cos^2\Omega t\cos\Omega\tau - \sin\Omega t\cos\Omega t\sin\Omega\tau\right) = \frac{A^2\cos\Omega\tau}{2} + O(1/T) \end{aligned}$$

Stochastic *vs* deterministic.

The ACF and the psd functions can be used to distinguish between stochastic and deterministic signals. Thus, the ACF of a deterministic signal is periodic, whereas the ACF of a stochastic signal typically decays exponentially with τ (see Fig. 3(a)). The power spectral function $S(\omega)$ of a deterministic signal consists of delta peaks, whereas the psd of a stochastic signal typically shows a power law decay (Fig. 3(b)).

Computation of Power spectral density

For stationary processes, the function $S(\omega)$ can be computed by taking the ensemble average of the square of the Fourier transform of $x(t)$. Thus we have

$$\langle\hat{x}(\omega)\hat{x}^*(\omega')\rangle = \int dt\,dt' e^{-i\omega t}e^{i\omega' t'}\langle x(t)x(t')\rangle, \tag{61}$$

where the integrals are taken in the limit of infinitely large intervals $T \to \infty$. Because for a stationary process

$$\langle x(t)x(t')\rangle = G(t-t'), \tag{62}$$

consequently

$$\begin{aligned}\langle \hat{x}(\omega)\hat{x}^*(\omega')\rangle &= \int dt\, dt' e^{-i\omega t} e^{i\omega' t'} G(t-t') = \{(t,t') \to (t', \tau = t-t')\} \\ &= \int dt' e^{i(\omega'-\omega)t'} \int d\tau e^{-i\omega'\tau} G(\tau) \\ &= 2\pi \int dt' e^{i(\omega'-\omega)t'} S(\omega') = (2\pi)^2 S(\omega')\delta(\omega-\omega').\end{aligned} \tag{63}$$

White noise and stochastic differential equations

The above result allows us to compute $S(\omega)$ for systems, described by linear stochastic differential equations. Any such equation can be written in the form

$$\dot{\boldsymbol{x}} = f(\boldsymbol{x}) + D\boldsymbol{\xi}(t), \tag{64}$$

where $\boldsymbol{x}$ is a vector variable, which characterizes the state of a system, $f(\boldsymbol{x})$ is a linear function, D is the noise strength (intensity), and, $\boldsymbol{\xi}(t)$ is the so called *white noise*. It can be introduced as a completely uncorrelated time series with the ACF given

$$\langle \xi(t)\xi(t')\rangle = \delta(t-t'). \tag{65}$$

Now it is easy to compute $S(\omega)$ for $\boldsymbol{x}(t)$ in Eq. (64). To this end, one needs to take the Fourier transform of both sides in Eq. (64) and use Eq. (63).

Examples

We will compute $S(\omega)$ for the following standard processes:

(1): The one-dimensional Wiener process is described by the equation

$$\dot{x} = \xi(t). \tag{66}$$

Thus, we have in the Fourier space

$$i\omega\hat{x}(\omega) = \hat{\xi}(\omega), \tag{67}$$

Consequently

$$\langle \hat{x}(\omega)\hat{x}^*(\omega')\rangle = \frac{\langle \hat{\xi}(\omega)\hat{\xi}^*(\omega')\rangle}{(i\omega)(-i\omega')} \tag{68}$$

But since by the definition of the white noise

$$\langle\hat{\xi}(\omega)\hat{\xi}^*(\omega')\rangle = (2\pi)^2\left(\frac{1}{2\pi}\int G_\xi(\tau)d\tau e^{-i\omega\tau}\right)\delta(\omega-\omega') = (2\pi)\delta(\omega-\omega'), \tag{69}$$

we obtain

$$\begin{aligned}\langle\hat{x}(\omega)\hat{x}^*(\omega')\rangle &= \frac{2\pi\delta(\omega-\omega')}{\omega^2} = (2\pi)^2 S(\omega)\delta(\omega-\omega')\\ S(\omega) &= \left(\frac{1}{2\pi}\right)\frac{1}{\omega^2}.\end{aligned} \tag{70}$$

If we formally try to compute the stationary ACF $G(\tau)$ using the Wiener-Khinchin theorem, we obtain a contradictory result

$$G(\tau) = \int_{-\infty}^{\infty}\frac{1}{2\pi\omega^2}e^{i\omega\tau}d\omega \sim \tau. \tag{71}$$

The contradiction comes to surface if we now try to compute $S(\omega)$ using the $G(\tau)$. Clearly, the function $f(x) = x$ is not a square-integrable on the interval $(-\infty,\infty)$, implying that its Fourier transform does not exist. Physically, it means that the Wiener process is not a stationary process, as we will see in more details later.

(2): 2D plasma in magnetic field

Consider a charged particle, confined to move on a plane in a constant magnetic field B, which is perpendicular to the plane of motion. The equation of motion of a single particle is

$$\begin{aligned}\dot{\boldsymbol{r}} &= \boldsymbol{v}\\ \dot{\boldsymbol{v}} &= \frac{q}{m}\boldsymbol{v}\times\boldsymbol{B} - \gamma\boldsymbol{v} + \sqrt{2\gamma\frac{kT}{m}}\boldsymbol{\xi}(t),\end{aligned} \tag{72}$$

where γ is the damping coefficient and T is the absolute temperature. We are interested in the power spectrum of the x-component of the velocity v_x. On the ground of the symmetry reasons, it is clear that $S_{v_x}(\omega) = S_{v_y}(\omega)$. The scalar form of the equations of motion is

$$\begin{aligned}\dot{v_x} &= \frac{qB}{m}v_y - \gamma v_x + \sqrt{2\gamma\frac{kT}{m}}\xi_x(t),\\ \dot{v_y} &= -\frac{qB}{m}v_x - \gamma v_y + \sqrt{2\gamma\frac{kT}{m}}\xi_y(t).\end{aligned} \tag{73}$$

In what follows we assume that the sources of the white noise ξ_x and ξ_y are uncorrelated, implying that $\langle\xi_x(t)\xi_y(t')\rangle = 0$. Taking the Fourier transform, we obtain

$$\begin{aligned}i\omega\hat{v}_x &= \frac{qB}{m}\hat{v}_y - \gamma\hat{v}_x + \sqrt{2\gamma\frac{kT}{m}}\hat{\xi}_x,\\ i\omega\hat{y} &= -\frac{qB}{m}\hat{v}_x - \gamma\hat{v}_y + \sqrt{2\gamma\frac{kT}{m}}\hat{\xi}_y.\end{aligned} \tag{74}$$

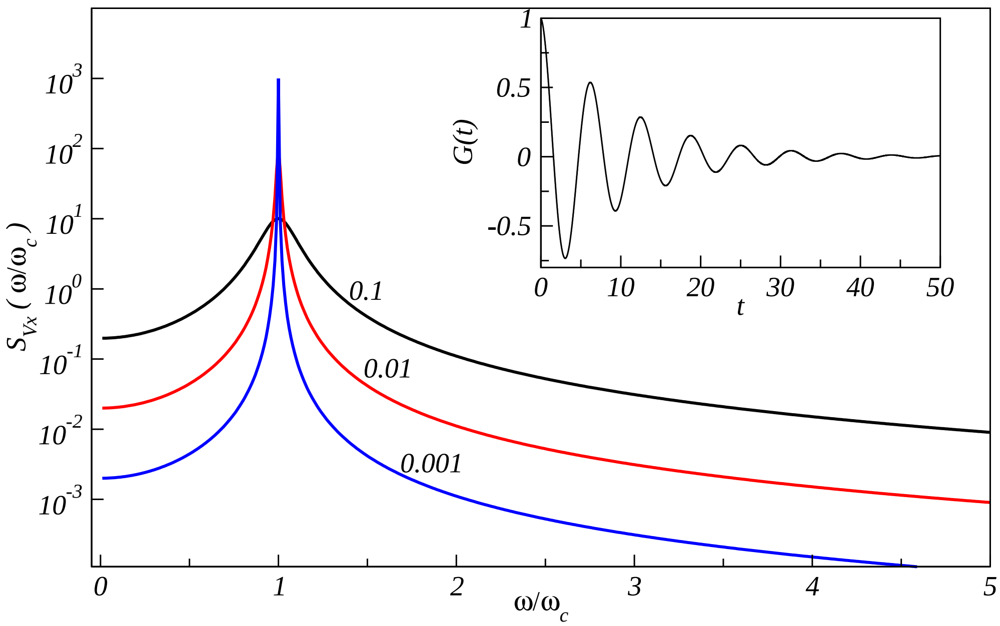

FIG. 4: Psd of the 2D plasma for different γ as in the legend and $kT/m = \omega_c = 1$. Inset: the ACF function for $\gamma = 0.1$.

Finally, making use of $\langle \xi_x(\omega)\xi_y(\omega')\rangle = 0$, we get

$$\begin{aligned} S_{v_x}(\omega) &= S_{v_y}(\omega) = \frac{1}{2\pi}\frac{2kT\gamma}{m}\frac{\omega^2+\gamma^2+\omega_c^2}{(\omega_c^2+\gamma^2-\omega^2)^2+4\omega^2\gamma^2} \\ &= \frac{1}{2\pi}\frac{2kT}{m}\frac{\gamma(\omega^2+\gamma^2)(\omega^2+\gamma^2+\omega_c^2)}{\gamma^2(\omega^2+\gamma^2+\omega_c^2)^2+\omega^2(\omega^2+\gamma^2-\omega_c^2)^2}, \end{aligned} \tag{75}$$

where $\omega_c = qB/m$ is the cyclotron frequency.

The stationary ACF is found as

$$G(\tau) = \frac{kT}{m}e^{-\gamma|\tau|}\cos\omega_c\tau. \tag{76}$$

The psd and the ACF are shown in Fig. 4 for different values of γ as in the legend.

IV. SOLVING STOCHASTIC ODES

The Wiener Process.

Central role in stochastic calculus plays the Wiener process, which has been shortly discussed above. Consider a continuous stochastic process $W(t)$, with the conditional probability $p(w,t|w_0,t_0)$, satisfying the following (diffusion) equation

$$\partial_t p(w,t|w_0,t_0) = \frac{1}{2}\partial_w^2 p(w,t|w_0,t_0). \tag{77}$$

It is well known that the solution of this equation with the initial condition given by the delta function $p(w,t|w_0,t_0) = \delta(w-w_0)$, at $t = t_0$, is the Gaussian

$$p(w,t|w_0,t_0) = \frac{1}{\sqrt{2\pi(t-t_0)}} \exp\left(-\frac{(w-w_0)^2}{2(t-t_0)}\right). \tag{78}$$

Using this result, we conclude that

$$\begin{aligned} \langle W(t)\rangle &= \int_{-\infty}^{\infty} dw\, w\, p(w,t|w_0,t_0) = w_0, \\ \langle (W(t)-w_0)^2\rangle &= t - t_0. \end{aligned} \tag{79}$$

Important is to observe that random variables $\Delta W_i = W(t_i) - W(t_{i-1})$, with any $t_i > t_{i-1}$ are independent. Indeed, using the Markov property, we can write the joint distribution as

$$p(w_n,t_n;w_{n-1},t_{n-1};\ldots;w_0,t_0) \tag{80}$$

$$= [p(w_n,t_n|w_{n-1},t_{n-1})p(w_{n-1},t_{n-1}|w_{n-2},t_{n-2})\ldots p(w_1,t_1|w_0,t_0)]\, p(w_0,t_0). \tag{81}$$

Now

$$\begin{aligned} p(w_n,t_n;w_{n-1},t_{n-1};\ldots;w_0,t_0) &= \prod_{i=0}^{n}\left\{\frac{1}{\sqrt{2\pi(t_i-t_{i-1})}}\exp\left(-\frac{(w_i-w_{i-1})^2}{2(t_i-t_{i-1})}\right)\right\} p(w_0,t_0) \\ &= p(\Delta w_n;\Delta w_{n-1}\ldots;w_0). \end{aligned} \tag{82}$$

Using the last equation, we obtain

$$\langle \Delta W_i^2\rangle = t_i - t_{i-1} \tag{83}$$

$$\langle W(t)W(s)|[w_0,t_0]\rangle = \langle (W(t)-W(s))W(s)\rangle + \langle W(s)^2\rangle = \min(t-t_0, s-s_0) + w_0^2.$$

Basics of the Ito and Stratonovich stochastic calculus

The starting point for the discussion is the relation between the Wiener process and the white noise

$$\Delta W_i = W(t_i) - W(t_{i-1}) = \xi(t_{i-1})\Delta t. \tag{84}$$

Then the solution of a general stochastic equation

$$\dot{x} = f(x) + g(x)\xi(t), \tag{85}$$

is represented via the integral

$$x(T) = x(t_0) + \int_{t_0}^{T} \dot{x}\, dt = \int_{t_0}^{T} [f(x(t))dt + g(x(t))\xi(t)dt] = \int_{t_0}^{T} [f(x(t))dt + g(x(t))dW(t)] \tag{86}$$

This result raises the question of the interpretation of a stochastic integral of the form

$$\int_{t_0}^{T} g(x(t))dW(t). \tag{87}$$

This integral is understood in the *mean-square* limit sense. More precisely, a sequence of random variables $X_n(\omega)$ is said to converge to $X(\omega)$ in the sense of the mean-square limit, if

$$\lim_{n\to\infty} \int d\omega\, p(\omega)[X_n(\omega) - X(\omega)]^2 = 0 \tag{88}$$

Ito stochastic integral

Consider

$$\int_{t_0}^{T} W(t)dW(t). \tag{89}$$

Partitioning the interval $[t_0, T]$ into the subintervals by t_i, $i = 1, \ldots, n$, the Ito integral is defined as the mean-square limit of the sum

$$\begin{aligned} \sum_{i=1}^{n} W_{i-1}(W_i - W_{i-1}) &= \sum_{i=1}^{n} W_{i-1}\Delta W_i = \frac{1}{2}\sum_{i=1}^{n}\Big[(W_{i-1} + \Delta W_i)^2 - (W_{i-1})^2 - (\Delta W_i)^2\Big] \\ &= \frac{1}{2}[W(T)^2 - W(t_0)^2] - \frac{1}{2}\sum_{i=1}^{n}(\Delta W_i)^2. \end{aligned} \tag{90}$$

Now compute the mean-square limit of the sum in the last equation. Using Eq. (83), we get

$$\langle\sum_{i=1}^{n}(\Delta W_i)^2\rangle = \sum_{i=1}^{n}(t_i - t_{i-1}) = T - t_0. \tag{91}$$

Additionally, we have

$$\begin{aligned} &\langle\Big[\sum_{i=1}^{n}(\Delta W_i)^2 - (T - t_0)\Big]^2\rangle \\ &= \langle\sum_i \Delta W_i^4 + 2\sum_{i<j}\Delta W_i^2\Delta W_j^2 - 2(t - t_0)\sum_i \Delta W_i^2 + (T - t_0)^2\rangle. \end{aligned} \tag{92}$$

But because ΔW_i are independent Gaussian variables, it holds

$$\begin{aligned} \langle\Delta W_i^2 \Delta W_j^2\rangle &= (t_i - t_{i-1})(t_j - t_{j-1}), \\ \langle\Delta W_i^4\rangle &= 3(t_i - t_{i-1})^2. \end{aligned} \tag{93}$$

Finally

$$\begin{aligned} &\langle\Big[\sum_{i=1}^{n}(\Delta W_i)^2 - (T - t_0)\Big]^2\rangle \\ &= 2\sum_i (t_i - t_{i-1})^2 + \sum_{\text{all } i,j}[(t_i - t_{i-1} - \frac{(T - t_0)}{n})(t_j - t_{j-1} - \frac{(T - t_0)}{n})] = 2\frac{(T - t_0)^2}{n} \to 0. \end{aligned} \tag{94}$$

This completes the proof and we obtain

$$\int_{t_0}^{T} W(t)dW(t) = \frac{1}{2}\left[W(T)^2 - W(t_0)^2 - (T - t_0)\right]. \tag{95}$$

Stratonovich interpretation

The integral Eq. (89) can also be approximated by taking the mid point in every subinterval $[t_{i-1}, t_i]$

$$\sum_{i=1}^{n} \frac{1}{2}\left(W_{i-1} + W_i\right)\left(W_i - W_{i-1}\right) = \frac{1}{2}\sum_{i=1}^{n}(W_{i-1}\Delta W_i + W_i \Delta W_i) \tag{96}$$
$$= \frac{1}{2}\left\{\frac{1}{2}(W(T)^2 - W(t_0)^2 - \sum_i \Delta W_i^2) - \frac{1}{2}\sum_i((W_i - \Delta W_i^2)^2 - W_i^2 - \Delta W_i^2)\right\}$$
$$= \frac{1}{2}\left\{\frac{1}{2}(W(T)^2 - W(t_0)^2 - \sum_i \Delta W_i^2) + \frac{1}{2}(W(T)^2 - W(t_0)^2 + \sum_i \Delta W_i^2)\right\} = \frac{1}{2}(W(T)^2 - W(t_0)^2).$$

Change of variables: the Ito formula

Recalling that according to Eq. (83),

$$\langle \Delta W_i^2 \rangle = t_i - t_{i-1} = \Delta t, \tag{97}$$

we will shortly write in what follows $dW^2 = dt$.

Then it is easy to see how the differentiation chain rule for an arbitrary function $F(x)$ changes in case when $x(t)$ satisfies the stochastic differential equation Eq. (85)

$$\begin{aligned} dF[x(t)] &= F[x(t) + dx(t)] - F[x(t)] \\ &= F'[x(t)]dx(t) + \frac{1}{2}F''[x(t)]dx(t)^2 + \ldots \\ &= F'[x(t)]\left\{f(x)dt + g(x)dW(t)\right\} + \frac{1}{2}F''[x(t)]\left\{f(x)dt + g(x)dW(t)\right\}^2. \end{aligned} \tag{98}$$

Retaining only terms of the order of $dW(t) \sim \sqrt{dt}$ and dt, we obtain

$$dF[x(t)] = \left\{f(x)F'[x] + \frac{1}{2}g(x)^2 F''[x]\right\}dt + g(x)F'[x]dW(t). \tag{99}$$

Equivalence of an Ito sde to a Stratonovich sde

Consider a Stratonovich sde

$$dx(t) = \dot{x}dt = \alpha(x)dt + \beta(x)dW(t). \tag{100}$$

The solution of this equation is represented as a sum of a regular and a stochastic Stratonovich integrals

$$x(T) = \int_{t_0}^{T}\alpha(x(t))dt + (S)\int_{t_0}^{T}\beta(x(t))dW(t) \tag{101}$$
$$= \int_{t_0}^{T}\alpha(x(t))dt + \sum_i \beta\left(\frac{1}{2}(x_i + x_{i-1})\right)\Delta W_i = \int_{t_0}^{T}\alpha(x(t))dt + \sum_i \beta\left(x_{i-1} + \frac{1}{2}\Delta x_i\right)\Delta W_i,$$

where $\Delta x_i = x_i - x_{i-1}$. But

$$\beta\left(x_{i-1}+\frac{1}{2}\Delta x_i\right) = \beta(x_{i-1}) + \frac{\partial\beta}{\partial x}\frac{1}{2}\Delta x_i + \frac{1}{2}\frac{\partial^2\beta}{\partial x^2}\left(\frac{1}{2}\Delta x_i\right)^2 + \dots \tag{102}$$

Our aim is to express the r.h.s. of Eq. (101) using the Ito interpretation. To this end, we set

$$\Delta x_i = a(x_{i-1})\Delta t + b(x_{i-1})\Delta W_i, \tag{103}$$

with $a(x)$ and $b(x)$ different from $\alpha(x)$ and $\beta(x)$. Combining Eq. (103) and Eq. (102), we obtain

$$\begin{aligned} b(x_{i-1})\partial_x\beta(x_{i-1})\Delta W_i.\beta\left(x_{i-1}+\frac{1}{2}\Delta x_i\right) &= \beta(x_{i-1}) + \left[a(x_{i-1})\partial_x\beta(x_{i-1}) + \frac{1}{4}b^2(x_{i-1})\partial_x^2\beta(x_{i-1})\right]\frac{1}{2}\Delta t \\ &+ \frac{1}{2}b(x_{i-1})\partial_x\beta(x_{i-1})\Delta W_i. \end{aligned} \tag{104}$$

Therefore, neglecting the terms of the order of $dt\,dW$ and dt^2, and keeping in mind that $\Delta W_i^2 = \Delta t$, we conclude

$$(S)\int_{t_0}^{T}\beta(x(t))dW(x(t)) = (I)\int_{t_0}^{T}\beta(x(t))dW(t) + \frac{1}{2}(I)\int_{t_0}^{T}b(x(t))\partial_x\beta(x(t))dt \tag{105}$$

This shows that a Stratonovich sde

$$dx = \alpha(x)dt + \beta(x)dW(t) \tag{106}$$

is equivalent to an Ito sde

$$dx = a(x)dt + b(x)dW(t), \tag{107}$$

where

$$a(x) = \alpha(x) + \frac{1}{2}\beta(x)\partial_x\beta(x), \quad b(x) = \beta(x). \tag{108}$$

Connection between Fokker-Planck equation and a stochastic sde

For an arbitrary function $f(x)$, where $x(t)$ satisfies the Ito sde Eq. (107), we obtain using the Ito formula

$$\langle df(x(t))\rangle = \left(a(x)f'(x) + \frac{1}{2}b^2(x)f''(x)\right)dt, \tag{109}$$

where we have used the fact that $\langle dW(t)\rangle = 0$ and further assumed that $\langle b(x(t))f'(x(t))dW(t)\rangle = 0$. Then

$$\begin{aligned} \frac{\langle df(x(t))\rangle}{dt} &= \langle\frac{df(x(t))}{dt}\rangle = \frac{d}{dt}\langle f(x(t))\rangle \\ &= \int dx f(x)\partial_t p(x,t|x_0,t_0) = \int dx\left[a(x)\partial_x f + \frac{1}{2}b(x)^2\partial_x^2 f\right]p(x,t|x_0,t_0), \end{aligned} \tag{110}$$

where $p(x,t|x_0,t_0)$ is the conditional probability. Using integration by parts and natural boundary conditions at $\pm\infty$, we get

$$\int dx f(x)\partial_t p = \int dx f(x)\left[-\partial_x(a(x)p) + \frac{1}{2}\partial_x^2(b(x)^2 p)\right]. \tag{111}$$

Because $f(x)$ is arbitrary, we are left with the equation for $p(x,t|x_0,t_0)$, which is called the Fokker-Planck equation, corresponding to the Ito sde Eq. (107)

$$\partial_t p = -\partial_x\left[a(x)p - \frac{1}{2}\partial_x(b(x)^2 p)\right] = -\partial_x J(x), \tag{112}$$

where $J(x)$ is the probability current.

Using our previous results, it is easy to show that if Eq. (107) i treated in the Stratonovich interpretation, then the corresponding Fokker-Planck equation becomes

$$\partial_t p = -\partial_x\left[a(x)p - \frac{1}{2}b(x)\partial_x(b(x)p)\right] = -\partial_x J(x). \tag{113}$$

More generally, if a system is described by a set of stochastic equations

$$\dot{x}_i = f_i(x_1,x_2,\ldots,x_n) + \sum_{j=1}^{n} g_{ij}(x_1,x_2,\ldots,x_n)\xi_j(t), \quad (i=1,2,\ldots,n), \tag{114}$$

where $\xi_k(t)$ represent sources of uncorrelated white Gaussian noise. Than in the Stratonovich interpretation, we have

$$\partial_t p = -\partial_i(f_i p) + \frac{1}{2}\partial_i\left(g_{im}\partial_k g_{km} p\right). \tag{115}$$

In the last equation we have used Einstein's summation convention over a pair of repeated indexes.

Example: active Brownian particle

How to use the Fokker-Planck equation to compute the evolution equation for the moments $\langle x^m(t)\rangle$.

Numerical solution of stochastic equations

In practice, stochastic equations can be solved using methods with much lower accuracy then the solution of the corresponding deterministic equations would require.

Thus, the $\sqrt{dt}$-accurate Euler scheme for equation Eq. (107) in the Ito interpretation is given by

$$\begin{aligned} x(t+dt) &= x(t) + a(x(t))dt + b(x(t))dW(t), \\ dW(t) &= N(0,1)\sqrt{dt}, \end{aligned} \tag{116}$$

where $N(0,1)$ is a normally distributed random number with variance one and zero mean.

The corresponding numerical scheme for the Stratonovich interpretation is

$$\begin{aligned} y &= x(t) + a(x(t))dt + b(x(t))dW(t), \\ x(t+dt) &= x(t) + \frac{1}{2}\left(x(t) + x(y)\right)dt + \frac{1}{2}\left(b(x(t)) + b(y)\right)dW(t), \\ dW(t) &= N(0,1)\sqrt{dt}, \end{aligned} \tag{117}$$

where $dW(t)$ is *the same* in the first and in the second equations.

The advantage of this method is the fact that it is explicit. We now show that this explicit algorithm is indeed equivalent to the implicit algorithm, which corresponds to the definition of the Stratonovich interpretation.

Thus, according to the definition of the Stratonovich integration, we should have had the following implicit scheme

$$\begin{aligned} \bar{x} &= \frac{x(t+dt) + x(t)}{2}, \\ x(t+dt) &= x(t) + a(\bar{x})dt + b(\bar{x})dW(t), \\ dW(t) &= N(0,1)\sqrt{dt}. \end{aligned} \tag{118}$$

For simplicity set $a(x) = 0$ and use the Newton's method to solve the algebraic equation

$$0 = x(t+dt) - x(t) - b\left(\frac{x(t) + x(t+dt)}{2}\right)dW(t). \tag{119}$$

The iteration scheme of the Newton's method for an algebraic equation $g(y) = 0$ is

$$y^{(n+1)} = y^{(n)} - \frac{g(y^{(n)})}{g'(y^{(n)})}. \tag{120}$$

By setting $y^{(n)} = x(t)$ and $y^{(n+1)} = x(t+dt)$, we obtain in our case

$$\begin{aligned} x(t+dt) &= x(t) - \frac{-b(x(t))dW(t)}{1 - \frac{1}{2}b'(x(t))dW(t)} \\ &\approx x(t) + b(x(t))dW(t) + \frac{1}{2}b(x(t))b'(x(t))dt + O(dt\,dW(t)). \end{aligned} \tag{121}$$

This result is equivalent to Eq. (117). Indeed

$$\begin{aligned} x(t+dt) &= x(t) + \frac{1}{2}\left(b(x(t)) + b(x(t) + b(x(t))dW(t))\right)dW(t) \\ &= x(t) + b(x(t))dW(t) + \frac{1}{2}b(x(t))b'(x(t))dt + O(dt\,dW(t)). \end{aligned} \tag{122}$$