I. 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

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

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

ACF is an even function

Proof:

$$\begin{aligned} G(-\tau) &= \lim_{T\to\infty} \frac{1}{T} \int_0^T dt x(t) x(t-\tau) = \Big\{ t-\tau = y|_{-\tau}^{T-\tau},\ \ dt = dy \Big\} \qquad (2) \\ &= \lim_{T\to\infty} \frac{1}{T} \int_{-\tau}^{T-\tau} dy x(\tau+y) x(y) = \lim_{T\to\infty} \frac{1}{T} \left[\int_{-\tau}^{0} (\dots) + \int_0^T (\dots) - \int_{T-\tau}^{T} (\dots) \right] = G(\tau) \end{aligned}$$

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

$$S(\omega) \;=\; \lim_{T\to\infty} \frac{1}{2\pi T} \mid \hat{x}(\omega) \mid^2, \tag{3}$$

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

$$\hat{x}(\omega) \;=\; \int_0^T dt e^{-i\omega t} x(t). \tag{4}$$

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

$$\hat{x}(-\omega) \;=\; \hat{\omega}^* \tag{5}$$

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.1 for details)

$$\begin{aligned} S(\omega) &= \lim_{T\to\infty} \frac{1}{2\pi T} \int_0^T dt e^{-i\omega t} x(t) \int_0^T dt' e^{i\omega t'} x^*(t') \\ &= \{(t, t') \to (t', \tau = t - t')\} \qquad (6) \end{aligned}$$

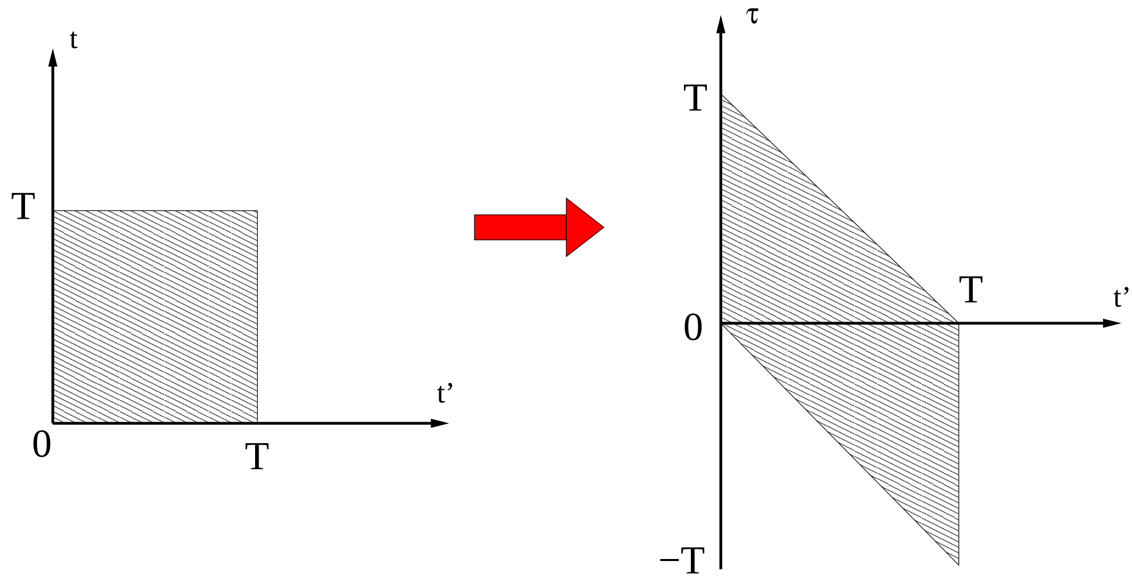

FIG. 1: Transformation of the integration domain under the change of variables $\{(t, t') \to (t', \tau = t - t')\}$.

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

Because the ACF is an even function, we can write

$$S(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} G(\tau) e^{\pm i\omega\tau} d\tau = \frac{1}{2\pi} \int_{-\infty}^{\infty} G(\tau) \cos\omega\tau d\tau = \frac{1}{\pi} \int_0^{\infty} G(\tau) \cos\omega\tau d\tau \tag{7}$$

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

$$\delta(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{\pm i\omega x} d\omega, \tag{8}$$

we conclude that

$$G(\tau) = \int_{-\infty}^{\infty} e^{i\omega\tau} S(\omega) d\omega. \tag{9}$$

What happens in case of a deterministic signal?

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

$$\hat{x}(\omega) = \int_0^T dt e^{-i\omega t} A \cos\Omega t = A\frac{T}{2}\delta_{\Omega,\omega} + O(1). \tag{10}$$

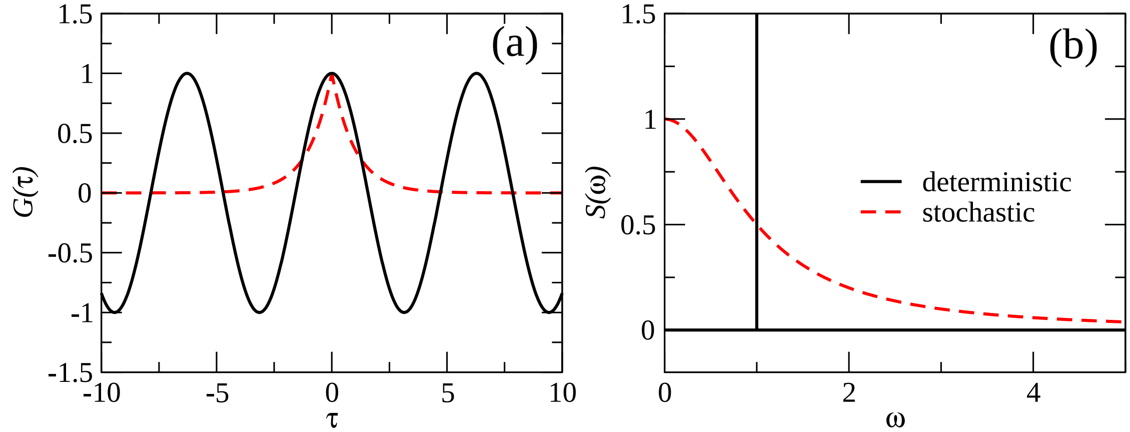

FIG. 2: (a) ACF $G(\tau)$ of a deterministic (dashed line) and a stochastic (solid line) signals. (b) The corresponding psd $S(\omega)$.

The psd

$$S(\omega) = \lim_{T\to\infty} \frac{1}{2\pi T}\hat{x}(\omega)\hat{x}^*(\omega) = \lim_{T\to\infty} \frac{A^2 T}{8\pi}\delta^2_{\Omega,\omega} + O(1) = O(T)\delta_{\Omega,\omega}. \quad (11)$$

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

The ACF is found as

$$\begin{aligned} G(\tau) &= \lim_{T\to\infty} \frac{A^2}{T}\int_0^T dt\ \cos\Omega t\cos\Omega(t+\tau) \quad (12)\\ &= \lim_{T\to\infty} \frac{A^2}{T}\int_0^T dt\ \left(\cos^2\Omega t\cos\Omega\tau - \sin\Omega t\cos\Omega t\sin\Omega\tau\right) = \frac{A^2\cos\Omega\tau}{2} + O(1/T) \end{aligned}$$

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. 2(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. 2(b)).

Stochastic processes

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 $t_1 > t_2 > \ldots$

$$p(x_1, t_1; x_2, t_2; \ldots) \quad (13)$$

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

$$p(x_1, t_1|x_2, t_2) = \frac{p(x_1, t_1; x_2, t_2)}{p(x_2, t_2)} \tag{14}$$

Note that

$$p(x_1, t_1) = \int_\Omega dx_2 \, p(x_1, t_1; x_2, t_2), \tag{15}$$

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

$$\langle x(t)|x_0, t_0\rangle = \int dx \, x \, p(x, t|x_0, t_0). \tag{16}$$

Similarly, the conditional (non-stationary) ACF can be computed as

$$\langle x(t)x(t')|x_0, t_0\rangle = \int dx \, dx' \, x \, x' p(x, t; x', t'|x_0, t_0) = \int dx \, dx' \, x \, x' p(x, t|x', t')p(x', t'|x_0, t_0). \tag{17}$$

The last equality only holds for Markovian processes (see below).

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.

$$p(x_1, t_1; x_2, t_2|x_3, t_3; x_4, t_4) = p(x_1, t_1; x_2, t_2|x_3, t_3). \tag{18}$$

As a consequence of this property, we have

$$p(x_1, t_1; x_2, t_2|x_3, t_3) = p(x_1, t_1|x_2, t_2)p(x_2, t_2|x_3, t_3). \tag{19}$$

Indeed

$$\begin{aligned} p(x_1, t_1|x_2, t_2)p(x_2, t_2|x_3, t_3) &= p(x_1, t_1|x_2, t_2; x_3, t_3)p(x_2, t_2|x_3, t_3) \\ &= \frac{p(x_1, t_1; x_2, t_2; x_3, t_3)}{p(x_2, t_2; x_3, t_3)} \frac{p(x_2, t_2; x_3, t_3)}{p(x_3, t_3)} \\ &= p(x_1, t_1; x_2, t_2|x_3, t_3). \end{aligned} \tag{20}$$

This proves the earlier result for non-stationary ACF Eq. (17).

The Chapman-Kolmogorov equation

From Eq. (19), one easily obtains

$$p(x_1,t_1|x_3,t_3) = \int_\Omega dx_2\, \frac{p(x_1,t_1;x_2,t_2;x_3,t_3)}{p(x_3,t_3)} = \int_\Omega dx_2\, p(x_1,t_1|x_2,t_2)p(x_2,t_2|x_3,t_3). \quad (21)$$

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:

$$\begin{aligned} \langle x(t)\rangle &= const \\ \langle x(t)x(t')\rangle &= f(t-t'). \end{aligned} \quad (22)$$

Stationary ACF can be obtained from Eq. (17) by assuming that the initial moment of time t_0 is in the remote infinity

$$\begin{aligned} G(\tau) &= \lim_{t_0\to-\infty} \langle x(t)x(t')|x_0,t_0\rangle \quad (23) \\ &= \lim_{t_0\to-\infty} \int xx'dx\,dx'\,p(x,t;x',t'|x_0,t_0) = \lim_{t_0\to-\infty} \int xx'dx\,dx'\,p(x,t|x',t')p(x',t'|x_0,t_0) \\ &= \int xx'dx\,dx'\,p(x,t|x',t')p_s(x') \quad (24) \end{aligned}$$

Example: The random telegraph process.

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:

$$x(t) = A, \;\; A \text{ is uniformly distributed in } [0;1] \quad (25)$$

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 $\int dt\, x(t) = A_i$.

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

$$\langle \hat{x}(\omega)\hat{x}^*(\omega')\rangle = \int dt\,dt' e^{-i\omega t} e^{i\omega' t'} \langle x(t)x(t')\rangle, \quad (26)$$

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

$$\langle x(t)x(t')\rangle = G(t-t'), \tag{27}$$

consequently

$$\begin{aligned}\langle \hat{x}(\omega)\hat{x}^*(\omega')\rangle &= \int dt\,dt' e^{-i\omega t}e^{i\omega' t'}G(t-t') = \{(t,t') \to (t', \tau = t-t')\} \\ &= \int dt' e^{i(\omega'-\omega)t'} \int d\tau e^{-i\omega'\tau}G(\tau) \\ &= 2\pi \int dt' e^{i(\omega'-\omega)t'} S(\omega') = (2\pi)^2 S(\omega')\delta(\omega-\omega').\end{aligned} \tag{28}$$

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

$$\dot{\boldsymbol{x}} = f(\boldsymbol{x}) + D\boldsymbol{\xi}(t), \tag{29}$$

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

$$\langle \xi(t)\xi(t')\rangle = \delta(t-t'). \tag{30}$$

Now it is easy to compute $S(\omega)$ for $\boldsymbol{x}(t)$ in Eq. (29). To this end, one needs to take the Fourier transform of both sides in Eq. (29) and use Eq. (28).

Examples

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

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

$$\dot{x} = \xi(t). \tag{31}$$

Thus, we have in the Fourier space

$$i\omega\hat{x}(\omega) = \hat{\xi}(\omega), \tag{32}$$

Consequently

$$\langle \hat{x}(\omega)\hat{x}^*(\omega')\rangle = \frac{\langle \hat{\xi}(\omega)\hat{\xi}^*(\omega')\rangle}{(i\omega)(-i\omega')} \tag{33}$$

But since by the definition of the white noise

$$\langle \hat{\xi}(\omega)\hat{\xi}^*(\omega')\rangle = (2\pi)^2 \left(\frac{1}{2\pi}\int G_\xi(\tau) d\tau e^{-i\omega\tau}\right)\delta(\omega-\omega') = (2\pi)\delta(\omega-\omega'), \tag{34}$$

we obtain

$$\begin{aligned}\langle \hat{x}(\omega)\hat{x}^*(\omega')\rangle &= \frac{2\pi\delta(\omega-\omega')}{\omega^2} = (2\pi)^2 S(\omega)\delta(\omega-\omega') \\ S(\omega) &= \left(\frac{1}{2\pi}\right)\frac{1}{\omega^2}.\end{aligned} \tag{35}$$

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

$$G(\tau) = \int_{-\infty}^{\infty} \frac{1}{2\pi\omega^2} e^{i\omega\tau} d\omega \sim \tau. \tag{36}$$

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): Overdamped particle, the Ornstein-Uhlenbeck process The equation of motion is

$$\dot{x} = -\alpha x + D\xi(t). \tag{37}$$

Proceeding as before, we obtain

$$\langle \hat{x}(\omega)\hat{x}^*(\omega')\rangle = \frac{D^2(2\pi)\delta(\omega-\omega')}{\alpha^2+\omega^2}, \tag{38}$$

Consequently

$$S(\omega) = \frac{D^2}{2\pi(\alpha^2+\omega^2)}. \tag{39}$$

This type of the power spectrum is known as the Lorentzian. It appears in many applications, such as: chemical reactions, bistable systems, random telegraph process (details later). The stationary ACF does exist and is given by

$$G(\tau) = \int_{-\infty}^{\infty} \frac{D^2}{2\pi(\alpha^2+\omega^2)} e^{i\omega\tau} d\omega = \frac{D^2}{2\alpha} e^{-\alpha\tau}. \tag{40}$$

The last result shows that $\langle x^2\rangle = G(0) = D^2/2\alpha$.

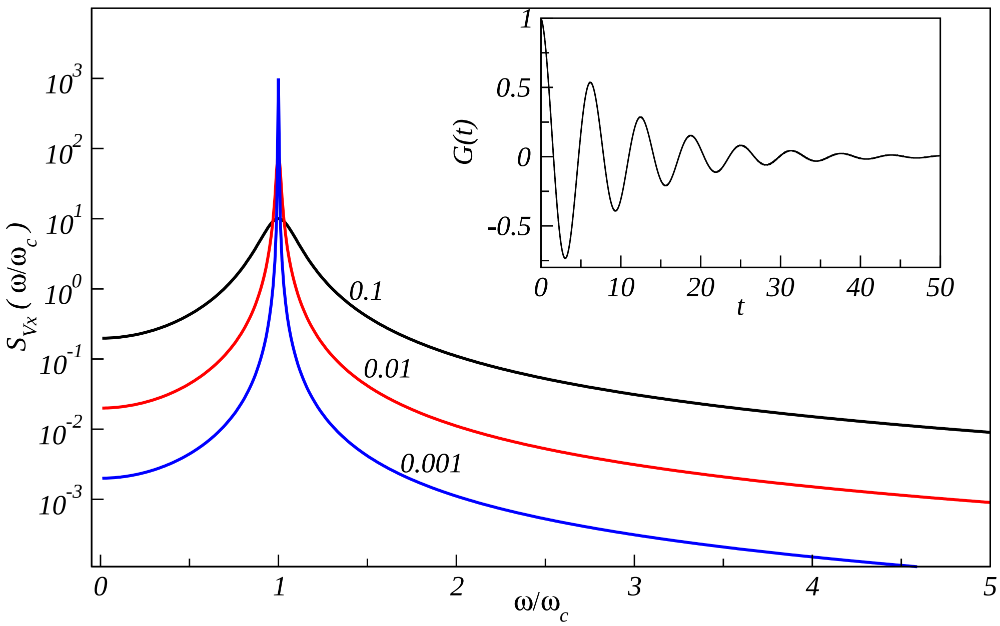

FIG. 3: Psd of the 2D plasma for different γ as in the legend and $kT/m = \omega_c = 1$. Inset: the ACF function for $\gamma = 0.1$.

(3): 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

$$\begin{aligned} \dot{\boldsymbol{r}} &= \boldsymbol{v} \\ \dot{\boldsymbol{v}} &= \frac{q}{m}\boldsymbol{v} \times \boldsymbol{B} - \gamma \boldsymbol{v} + \sqrt{2\gamma \frac{kT}{m}}\boldsymbol{\xi}(t), \end{aligned} \tag{41}$$

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

$$\begin{aligned} \dot{v_x} &= \frac{qB}{m}v_y - \gamma v_x + \sqrt{2\gamma \frac{kT}{m}}\xi_x(t), \\ \dot{v_y} &= -\frac{qB}{m}v_x - \gamma v_y + \sqrt{2\gamma \frac{kT}{m}}\xi_y(t). \end{aligned} \tag{42}$$

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

$$\begin{aligned} i\omega\hat{v}_x &= \frac{qB}{m}\hat{v}_y - \gamma\hat{v}_x + \sqrt{2\gamma \frac{kT}{m}}\hat{\xi}_x, \\ i\omega\hat{y} &= -\frac{qB}{m}\hat{v}_x - \gamma\hat{v}_y + \sqrt{2\gamma \frac{kT}{m}}\hat{\xi}_y. \end{aligned} \tag{43}$$

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

$$\begin{aligned} S_{v_x}(\omega) &= S_{v_y}(\omega) = \frac{1}{2\pi}\frac{2kT\gamma}{m}\frac{\omega^2+\gamma^2+\omega_c^2}{(\omega_c^2+\gamma^2-\omega^2)^2+4\omega^2\gamma^2} \\ &= \frac{1}{2\pi}\frac{2kT}{m}\frac{\gamma(\omega^2+\gamma^2)(\omega^2+\gamma^2+\omega_c^2)}{\gamma^2(\omega^2+\gamma^2+\omega_c^2)^2+\omega^2(\omega^2+\gamma^2-\omega_c^2)^2}, \end{aligned} \tag{44}$$

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

The stationary ACF is found as

$$G(\tau) = \frac{kT}{m}e^{-\gamma|\tau|}\cos\omega_c\tau. \tag{45}$$

The psd and the ACF are shown in Fig. 3 for different values of γ as in the legend.

II. 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

$$\partial_t p(w,t|w_0,t_0) = \frac{1}{2}\partial_w^2 p(w,t|w_0,t_0). \tag{46}$$

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

$$p(w,t|w_0,t_0) = \frac{1}{\sqrt{2\pi(t-t_0)}}\exp\left(-\frac{(w-w_0)^2}{2(t-t_0)}\right). \tag{47}$$

Using this result, we conclude that

$$\begin{aligned} \langle W(t)\rangle &= \int_{-\infty}^{\infty} dw\, w\, p(w,t|w_0,t_0) = w_0, \\ \langle (W(t)-w_0)^2\rangle &= t-t_0. \end{aligned} \tag{48}$$

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

$$p(w_n,t_n;w_{n-1},t_{n-1};\ldots;w_0,t_0) \tag{49}$$

$$= [p(w_n,t_n|w_{n-1},t_{n-1})p(w_{n-1},t_{n-1}|w_{n-2},t_{n-2})\ldots p(w_1,t_1|w_0,t_0)]\,p(w_0,t_0). \tag{50}$$

Now

$$p(w_n, t_n; w_{n-1}, t_{n-1}; \dots; w_0, t_0) = \prod_{i=0}^{n} \left\{ \frac{1}{\sqrt{2\pi(t_i - t_{i-1})}} \exp\left(-\frac{(w_i - w_{i-1})^2}{2(t_i - t_{i-1})} \right) \right\} p(w_0, t_0)$$
$$= p(\Delta w_n; \Delta w_{n-1} \dots; w_0). \quad (51)$$

Using the last equation, we obtain

$$\langle \Delta W_i^2 \rangle = t_i - t_{i-1} \quad (52)$$
$$\langle W(t)W(s)|[w_0, t_0] \rangle = \langle (W(t) - W(s))W(s) \rangle + \langle W(s)^2 \rangle = \min(t - t_0, s - s_0) + w_0^2.$$

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

$$\Delta W_i = W(t_i) - W(t_{i-1}) = \xi(t_{i-1})\Delta t. \quad (53)$$

Then the solution of a general stochastic equation

$$\dot{x} = f(x) + g(x)\xi(t), \quad (54)$$

is represented via the integral

$$x(T) = x(t_0) + \int_{t_0}^{T} \dot{x}\, dt = \int_{t_0}^{T} [f(x(t))dt + g(x(t))\xi(t)dt] = \int_{t_0}^{T} [f(x(t))dt + g(x(t))dW(t)] \quad (55)$$

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

$$\int_{t_0}^{T} g(x(t))dW(t). \quad (56)$$

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

$$\lim_{n\to\infty} \int d\omega\, p(\omega)[X_n(\omega) - X(\omega)]^2 = 0 \quad (57)$$

Ito stochastic integral

Consider

$$\int_{t_0}^{T} W(t)dW(t). \quad (58)$$

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

$$\begin{aligned}\sum_{i=1}^{n} W_{i-1}(W_i - W_{i-1}) &= \sum_{i=1}^{n} W_{i-1}\Delta W_i = \frac{1}{2}\sum_{i=1}^{n}\left[(W_{i-1} + \Delta W_i)^2 - (W_{i-1})^2 - (\Delta W_i)^2\right] \\ &= \frac{1}{2}[W(T)^2 - W(t_0)^2] - \frac{1}{2}\sum_{i=1}^{n}(\Delta W_i)^2. \end{aligned} \tag{59}$$

Now compute the mean-square limit of the sum in the last equation. Using Eq. (52), we get

$$\langle\sum_{i=1}^{n}(\Delta W_i)^2\rangle = \sum_{i=1}^{n}(t_i - t_{i-1}) = T - t_0. \tag{60}$$

Additionally, we have

$$\begin{aligned}&\langle\left[\sum_{i=1}^{n}(\Delta W_i)^2 - (T - t_0)\right]^2\rangle \\ &= \langle\sum_{i}\Delta W_i^4 + 2\sum_{i<j}\Delta W_i^2\Delta W_j^2 - 2(t - t_0)\sum_{i}\Delta W_i^2 + (T - t_0)^2\rangle.\end{aligned} \tag{61}$$

But because ΔW_i are independent Gaussian variables, it holds

$$\begin{aligned}\langle\Delta W_i^2\Delta W_j^2\rangle &= (t_i - t_{i-1})(t_j - t_{j-1}), \\ \langle\Delta W_i^4\rangle &= 3(t_i - t_{i-1})^2.\end{aligned} \tag{62}$$

Finally

$$\begin{aligned}&\langle\left[\sum_{i=1}^{n}(\Delta W_i)^2 - (T - t_0)\right]^2\rangle \\ &= 2\sum_{i}(t_i - t_{i-1})^2 + \sum_{\text{all } i,j}[(t_i - t_{i-1} - \frac{(T - t_0)}{n})(t_j - t_{j-1} - \frac{(T - t_0)}{n})] = 2\frac{(T - t_0)^2}{n} \to 0.\end{aligned} \tag{63}$$

This completes the proof and we obtain

$$\int_{t_0}^{T} W(t)dW(t) = \frac{1}{2}\left[W(T)^2 - W(t_0)^2 - (T - t_0)\right]. \tag{64}$$

Stratonovich interpretation

The integral Eq. (58) can also be approximated by taking the mid point in every subinterval $[t_{i-1}, t_i]$

$$\begin{aligned}&\sum_{i=1}^{n}\frac{1}{2}\left(W_{i-1} + W_i\right)\left(W_i - W_{i-1}\right) = \frac{1}{2}\sum_{i=1}^{n}(W_{i-1}\Delta W_i + W_i\Delta W_i) \\ &= \frac{1}{2}\left\{\frac{1}{2}(W(T)^2 - W(t_0)^2 - \sum_{i}\Delta W_i^2) - \frac{1}{2}\sum_{i}((W_i - \Delta W_i^2)^2 - W_i^2 - \Delta W_i^2)\right\} \\ &= \frac{1}{2}\left\{\frac{1}{2}(W(T)^2 - W(t_0)^2 - \sum_{i}\Delta W_i^2) + \frac{1}{2}(W(T)^2 - W(t_0)^2 + \sum_{i}\Delta W_i^2)\right\} = \frac{1}{2}(W(T)^2 - W(t_0)^2).\end{aligned} \tag{65}$$

Change of variables: the Ito formula

Recalling that according to Eq. (52),

$$\langle \Delta W_i^2 \rangle = t_i - t_{i-1} = \Delta t, \tag{66}$$

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. (54)

$$\begin{aligned} dF[x(t)] &= F[x(t) + dx(t)] - F[x(t)] \\ &= F'[x(t)]dx(t) + \frac{1}{2}F''[x(t)]dx(t)^2 + \ldots \\ &= F'[x(t)]\left\{f(x)dt + g(x)dW(t)\right\} + \frac{1}{2}F''[x(t)]\left\{f(x)dt + g(x)dW(t)\right\}^2. \end{aligned} \tag{67}$$

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

$$dF[x(t)] = \left\{f(x)F'[x] + \frac{1}{2}g(x)^2F''[x]\right\}dt + g(x)F'[x]dW(t). \tag{68}$$

Equivalence of an Ito sde to a Stratonovich sde

Consider a Stratonovich sde

$$dx(t) = \dot{x}dt = \alpha(x)dt + \beta(x)dW(t). \tag{69}$$

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

$$\begin{aligned} x(T) &= \int_{t_0}^{T} \alpha(x(t))dt + (S)\int_{t_0}^{T} \beta(x(t))dW(t) \qquad (70) \\ &= \int_{t_0}^{T} \alpha(x(t))dt + \sum_i \beta\left(\frac{1}{2}(x_i + x_{i-1})\right)\Delta W_i = \int_{t_0}^{T} \alpha(x(t))dt + \sum_i \beta\left(x_{i-1} + \frac{1}{2}\Delta x_i\right)\Delta W_i, \end{aligned}$$

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

$$\beta\left(x_{i-1} + \frac{1}{2}\Delta x_i\right) = \beta(x_{i-1}) + \frac{\partial \beta}{\partial x}\frac{1}{2}\Delta x_i + \frac{1}{2}\frac{\partial^2 \beta}{\partial x^2}\left(\frac{1}{2}\Delta x_i\right)^2 + \ldots \tag{71}$$

Our aim is to express the r.h.s. of Eq. (70) using the Ito interpretation. To this end, we set

$$\Delta x_i = a(x_{i-1})\Delta t + b(x_{i-1})\Delta W_i, \tag{72}$$

with $a(x)$ and $b(x)$ different from $\alpha(x)$ and $\beta(x)$. Combining Eq. (72) and Eq. (71), we obtain

$$\begin{aligned} b(x_{i-1})\partial_x\beta(x_{i-1})\Delta W_i.\beta\left(x_{i-1} + \frac{1}{2}\Delta x_i\right) &= \beta(x_{i-1}) + \left[a(x_{i-1})\partial_x\beta(x_{i-1}) + \frac{1}{4}b^2(x_{i-1})\partial_x^2\beta(x_{i-1})\right]\frac{1}{2}\Delta t \\ &+ \frac{1}{2}b(x_{i-1})\partial_x\beta(x_{i-1})\Delta W_i. \end{aligned} \tag{73}$$

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

$$(S)\int_{t_0}^{T}\beta(x(t))dW(x(t)) = (I)\int_{t_0}^{T}\beta(x(t))dW(t) + \frac{1}{2}(I)\int_{t_0}^{T}b(x(t))\partial_x\beta(x(t))dt \tag{74}$$

This shows that a Stratonovich sde

$$dx = \alpha(x)dt + \beta(x)dW(t) \tag{75}$$

is equivalent to an Ito sde

$$dx = a(x)dt + b(x)dW(t), \tag{76}$$

where

$$a(x) = \alpha(x) + \frac{1}{2}\beta(x)\partial_x\beta(x), \quad b(x) = \beta(x). \tag{77}$$

Connection between Fokker-Planck equation and a stochastic sde

For an arbitrary function $f(x)$, where $x(t)$ satisfies the Ito sde Eq. (76), we obtain using the Ito formula

$$\langle df(x(t))\rangle = (a(x)f'(x) + \frac{1}{2}b^2(x)f''(x))\, dt, \tag{78}$$

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

$$\begin{aligned}\frac{\langle df(x(t))\rangle}{dt} &= \langle\frac{df(x(t))}{dt}\rangle = \frac{d}{dt}\langle f(x(t))\rangle \\ &= \int dx f(x)\partial_t p(x,t|x_0,t_0) = \int dx\left[a(x)\partial_x f + \frac{1}{2}b(x)^2\partial_x^2 f\right]p(x,t|x_0,t_0),\end{aligned} \tag{79}$$

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

$$\int dx f(x)\partial_t p = \int dx f(x)\left[-\partial_x(a(x)p) + \frac{1}{2}\partial_x^2(b(x)^2 p)\right]. \tag{80}$$

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. (76)

$$\partial_t p = -\partial_x\left[a(x)p - \frac{1}{2}\partial_x(b(x)^2 p)\right] = -\partial_x J(x), \tag{81}$$

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

Using our previous results, it is easy to show that if Eq. (76) i treated in the Stratonovich interpretation, then the corresponding Fokker-Planck equation becomes

$$\partial_t p = -\partial_x \left[a(x)p - \frac{1}{2} b(x) \partial_x (b(x)p) \right] = -\partial_x J(x). \tag{82}$$

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

$$\dot{x}_i = f_i(x_1, x_2, \dots, x_n) + \sum_{j=1}^{n} g_{ij}(x_1, x_2, \dots, x_n) \xi_j(t), \quad (i = 1, 2, \dots, n), \tag{83}$$

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

$$\partial_t p = -\partial_i (f_i p) + \frac{1}{2} \partial_i \left(g_{im} \partial_k g_{km} p \right). \tag{84}$$

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

Example: active Brownian particle

The rotation of the direction vector $\boldsymbol{p}$ is described by the equation

$$\dot{\boldsymbol{p}} = \boldsymbol{\eta}(t) \times \boldsymbol{p}, \tag{85}$$

where $\boldsymbol{\eta}$ is a 3-dimensional vector with components given by uncorrelated white Gaussian noise, i.e $\langle \eta_i(t) \eta_k(t') \rangle = 2D\delta(t - t')$, and the indexes $(i, k) = (1, 2, 3)$, corresponding to (x, y, z).

We want to show that if Eq. (85) is treated in the Stratonovich interpretation, than it gives rise to the following Smoluchowski equation for the probability density $\rho(\boldsymbol{p})$

$$\partial_t \rho = D \boldsymbol{R}^2 \rho, \tag{86}$$

where $\boldsymbol{R} = \boldsymbol{p} \times \boldsymbol{\nabla}_p$.

Recall that the cross product between any two vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ can be written with the help of the antisymmetric unity tensor E_{ijk}, which is equal either to ± 1 for all different (ijk) $(= +1$ for $(ijk) = (123))$, or it is equal to 0 if any two indexes (ijk) are identical.

$$\boldsymbol{a} \times \boldsymbol{b} = E_{ijk} a_j b_k \boldsymbol{e}_i, \tag{87}$$

where $\boldsymbol{e}_i$ denotes the unity vector, pointing along the i-axes.

Eq. (86) can be rewritten in terms of E_{ijk}

$$\partial_t \rho = D\left(E_{ijk} p_j \partial_{p_k} E_{ims} p_m \partial_{p_s}\right)\rho, \tag{88}$$

Than in the Stratonovich interpretation, we have

$$\partial_t \rho = \frac{1}{2}\partial_i\left(\sigma_{im}\partial_k \sigma_{km}\rho\right) \tag{89}$$

By comparing Eq. (85) with Eq. (83), we see that

$$\sigma_{ik} = E_{ikm} p_m. \tag{90}$$

Now we can rewrite Eq. (89) for the case of Eq. (85)

$$\partial_t \rho = 2D\frac{1}{2}\partial_{p_i}\left(E_{ims} p_s \partial_{p_k} E_{kmj} p_j \rho\right) \tag{91}$$

Finally, after noticing that $\partial_{p_i} p_k = \delta_{ik}$ and after the appropriate permutation of indexes in Eq. (91), we conclude that Eq. (91) and Eq. (88) are identical.

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

Example: diffusion coefficient of a classical Brownian particle.

Consider one-dimensional motion of a classical particle m in contact with a fluctuating environment

$$\begin{aligned} \dot{x} &= \frac{p}{m} \\ \dot{p} &= -\alpha p + \sqrt{2D}\xi(t), \end{aligned} \tag{92}$$

where α is the damping coefficient and D characterizes the strength of fluctuations. The corresponding Fokker-Planck equation is given by

$$\partial_t \rho(x,p,t) = -\partial_x\left(\frac{p}{m}\rho\right) + \partial_p[\alpha p \rho + D\partial_p \rho]. \tag{93}$$

The stationary solution is x-independent

$$\rho_s = \sqrt{\frac{\alpha}{2\pi D}}\exp\left(-\frac{\alpha p^2}{2D}\right). \tag{94}$$

In order to recover the Maxwell's velocity distribution at temperature T

$$\rho_s(v) = \sqrt{\frac{m}{2\pi kT}}\exp\left(-\frac{mv^2}{2kT}\right), \tag{95}$$

one needs to impose the Einstein's condition

$$D = \alpha m k T. \tag{96}$$

We are interested in the diffusion coefficient of the particle

$$D_\infty = \lim_{t\to\infty} \frac{\langle (x(t)-x(0))^2\rangle}{2t}. \tag{97}$$

In order to find $\langle x(t)\rangle$, we use the Fokker-Planck equation, integration by parts and natural BCs at $\pm\infty$

$$\partial_t\langle x(t)\rangle \;=\; \int dxdp\, x\partial_t\rho(x,t) = \int dxdp\, x\left[\partial_x\left(-\frac{p}{m}\rho\right) + \partial_p[\alpha p\rho + D\partial_p\rho]\right] = \frac{\langle p(t)\rangle}{m}. \tag{98}$$

Similarly, we find the equations for $\langle p(t)\rangle$, $\langle x(t)p(t)\rangle$, $\langle p(t)^2\rangle$ and $\langle x(t)^2\rangle$

$$\begin{aligned}
\partial_t\langle p(t)\rangle &= \int dxdp\, p\partial_t\rho(x,t) = \int dxdp\, p\left[\partial_x\left(-\frac{p}{m}\rho\right) + \partial_p[\alpha p\rho + D\partial_p\rho]\right] = -\alpha\langle p(t)\rangle, \\
\partial_t\langle x(t)p(t)\rangle &= \int dxdp\, px\left[\partial_x\left(-\frac{p}{m}\rho\right) + \partial_p[\alpha p\rho + D\partial_p\rho]\right] = \frac{\langle p(t)^2\rangle}{m} - \alpha\langle x(t)p(t)\rangle \\
\partial_t\langle x(t)^2\rangle &= \int dxdp\, x^2\left[\partial_x\left(-\frac{p}{m}\rho\right) + \partial_p[\alpha p\rho + D\partial_p\rho]\right] = \frac{2\langle x(t)p(t)\rangle}{m} \\
\partial_t\langle p(t)^2\rangle &= \int dxdp\, p^2\left[\partial_x\left(-\frac{p}{m}\rho\right) + \partial_p[\alpha p\rho + D\partial_p\rho]\right] = -2\alpha\langle p(t)^2\rangle + 2D.
\end{aligned} \tag{99}$$

Eqs. (98,99) represent a closed system of linear equations for the first five moments, which can be solved analytically

$$\langle p(t)\rangle \;=\; p_0 e^{-\alpha(t-t_0)} \tag{100}$$

$$\begin{aligned}
\langle p(t)^2\rangle &= \frac{D}{\alpha}\left(1-e^{-2\alpha(t-t_0)}\right) + p_0^2 e^{-2\alpha(t-t_0)} \\
\langle x(t)\rangle &= x_0 + \frac{p_0}{\alpha m}\left(1-e^{-\alpha(t-t_0)}\right) \\
\langle x(t)p(t)\rangle &= \frac{D}{\alpha^2 m} - \left(\frac{p_0^2}{\alpha m} - \frac{D}{\alpha^2 m}\right)e^{-2\alpha(t-t_0)} + \left((xp)_0 - \frac{2D}{\alpha^2 m} + \frac{p_0^2}{\alpha m}\right)e^{-\alpha(t-t_0)} \\
\langle x(t)^2\rangle &= \frac{2Dt}{(\alpha m)^2} + C_1 e^{-\alpha(t-t_0)} + C_2 e^{-2\alpha(t-t_0)}.
\end{aligned} \tag{101}$$

The terms proportional to C_1 and C_2 vanish in the limit $t\to\infty$ in Eq. (97) and we finally obtain

$$D_\infty = \frac{D}{(\alpha m)^2} = \frac{kT}{\alpha m}. \tag{102}$$

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. (76) in the Ito interpretation is given by

$$\begin{aligned} x(t+dt) &= x(t) + a(x(t))dt + b(x(t))dW(t), \\ dW(t) &= N(0,1)\sqrt{dt}, \end{aligned} \tag{103}$$

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

$$\begin{aligned} y &= x(t) + a(x(t))dt + b(x(t))dW(t), \\ x(t+dt) &= x(t) + \frac{1}{2}\left(x(t) + x(y)\right)dt + \frac{1}{2}\left(b(x(t)) + b(y)\right)dW(t), \\ dW(t) &= N(0,1)\sqrt{dt}, \end{aligned} \tag{104}$$

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

$$\begin{aligned} \bar{x} &= \frac{x(t+dt) + x(t)}{2}, \\ x(t+dt) &= x(t) + a(\bar{x})dt + b(\bar{x})dW(t), \\ dW(t) &= N(0,1)\sqrt{dt}. \end{aligned} \tag{105}$$

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

$$0 = x(t+dt) - x(t) - b\left(\frac{x(t) + x(t+dt)}{2}\right)dW(t). \tag{106}$$

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

$$y^{(n+1)} = y^{(n)} - \frac{g(y^{(n)})}{g'(y^{(n)})}. \tag{107}$$

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

$$\begin{aligned} x(t+dt) &= x(t) - \frac{-b(x(t))dW(t)}{1 - \frac{1}{2}b'(x(t))dW(t)} \\ &\approx x(t) + b(x(t))dW(t) + \frac{1}{2}b(x(t))b'(x(t))dt + O(dt\, dW(t)). \end{aligned} \tag{108}$$

This result is equivalent to Eq. (104). Indeed

$$\begin{aligned} x(t+dt) &= x(t) + \frac{1}{2}\left(b(x(t)) + b(x(t) + b(x(t))dW(t))\right) dW(t) \\ &= x(t) + b(x(t))dW(t) + \frac{1}{2}b(x(t))b'(x(t))dt + O(dt\, dW(t)). \end{aligned} \tag{109}$$