Computational Methods in Nonlinear Physics
4. Stochastic Processes and Stochastic Differential Equations

P.I. Hurtado

Departamento de Electromagnetismo y Física de la Materia, and Instituto Carlos I de Física Teórica y Computacional. Universidad de Granada. E-18071 Granada. Spain

E-mail: phurtado@onsager.ugr.es

Abstract. These notes correspond to the first part (20 hours) of a course on Computational Methods in Nonlinear Physics within the Master on Physics and Mathematics (*FisyMat*) of University of Granada. In this fourth chapter we review the basic concepts of what a stochastic process is, and its description in terms of stochastic differential equations and/or Fokker-Planck equations. Moreover, we also discuss numerical methods to generate sample trajectories for a given Langevin stochastic differential equation.

Keywords: Computational physics, probability and statistics, Monte Carlo methods, stochastic differential equations, Langevin equation, Fokker-Planck equation, molecular dynamics.

References and sources

[1] R. Toral and P. Collet, *Stochastic Numerical Methods*, Wiley (2014).
[2] P.L. Krapivsky, S. Redner and E. Ben-Naim, *A Kinetic View of Statistical Physics*, Cambridge University Press (2010).

Contents

In this chapter, we review the basic concepts of what a stochastic process is. Our aim is not to be rigorous on the mathematical side but rather to focus on the physical insights behind the concepts. The name *"stochastic process"* is usually associated with a trajectory in phase space which is random enough to demand a probabilistic description. A paradigmatic example is that of the Brownian motion.

1. A pedagogical example: Brownian motion

1.1. Einstein description

- The botanist Robert Brown discovered in 1827 that pollen particles in suspension execute random movements (see Fig. 1) which he even interpreted initially as some sort of life.
- It is not so well known that L. Boltzmann knew as early as 1896 the reason for this erratic movement when he wrote *"... very small particles in a gas execute motions which result from the fact that the pressure on the surface of the particles may fluctuate."*
- However, it was A. Einstein in 1905 who successfully introduced the first mathematical treatment of the erratic movement of the Brownian particles. Rather than focusing on the (complicated) trajectory of a single particle, Einstein introduced a probabilistic description valid for an ensemble of Brownian particles.
- First, Einstein introduced the concept of a *coarse-grained* description defined by a time τ such that different parts of the trajectory separated by a time τ or larger can be considered independent. No attempt is made to characterize the dynamics at a timescale smaller than this coarse-grain time τ. Instead, we consider snapshots of the system taken at time intervals τ (see right panel of Fig. 1).
- The second concept, probabilistic in nature, introduced by Einstein is that of the probability density function (or pdf) $f(\boldsymbol{\Delta})$, for the three-dimensional distance $\boldsymbol{\Delta} = (\Delta_x, \Delta_y, \Delta_z)$ traveled by the Brownian particle in a fixed time interval τ.

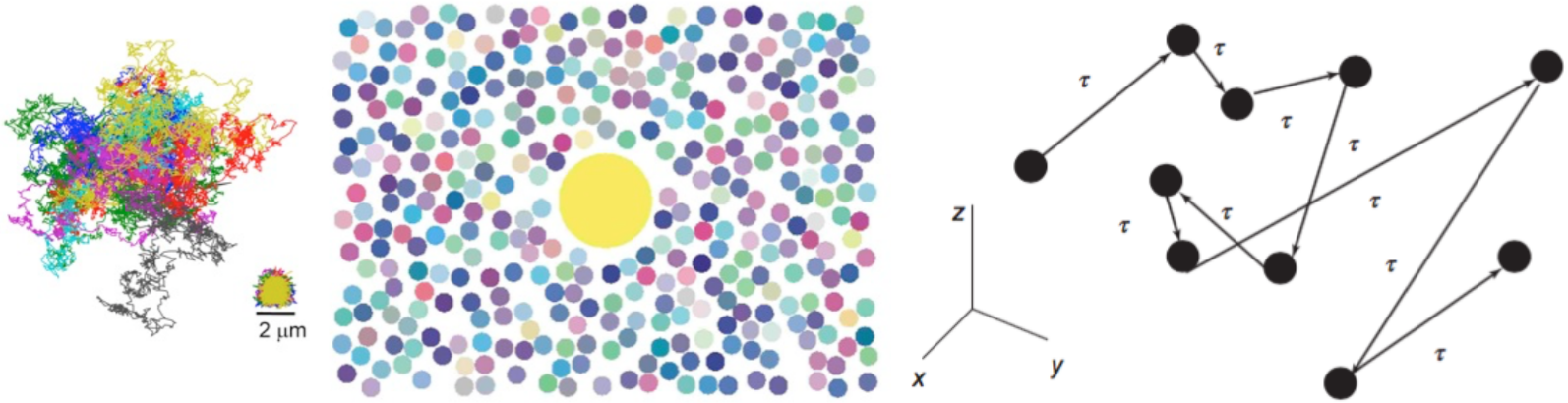

Figure 1. Sketch of Brownian motion. Left panel: random trajectory of a mesoscopic particle immersed in a molecular fluid. The many random collisions of the fluid's microscopic particles with the mesoscopic object (middel panel) induce in the latter an erratic, intermitent motion in the latter. Right panel: Coarse-grained description of the mesoscopic particle position at time intervals τ.

- The only assumption one needs to make about the function $f(\boldsymbol{\Delta})$ (besides the general condition of nonnegativity and normalization) comes from the fact that the collisions of the fluid molecules and the Brownian particle occur with the same probability in any direction[1]. The absence of preferred directions translates to a symmetry condition for $f(\boldsymbol{\Delta})$, as

$$f(-\boldsymbol{\Delta}) = f(\boldsymbol{\Delta})\,. \tag{1}$$

- The third step in this description is to consider an ensemble of N Brownian particles in a large enough system. Also, we focus on large spatial scales, much larger than the size of a Brownian particle, so that we can define a density of the particles $n(\boldsymbol{x}, t)$ such that $n(\boldsymbol{x}, t)\, d\boldsymbol{x}$ is the number of particles in the interval $(\boldsymbol{x}, \boldsymbol{x} + d\boldsymbol{x})$ at time t.

† We also disregard the effect of gravity in the Brownian particle, which would lead to a preferred direction in the movement.

- From the assumption that the parts of the trajectories separated a time interval τ are statistically independent, it follows that the number of particles at location $\boldsymbol{x}$ at time $t+\tau$ will be given by the number of particles at location $\boldsymbol{x}-\boldsymbol{\Delta}$ at time t multiplied by the probability that the particle jumps from $\boldsymbol{x}-\boldsymbol{\Delta}$ to $\boldsymbol{x}$ in an elementary time step τ, which is $f(\boldsymbol{\Delta})$, and integrated for all the possible $\boldsymbol{\Delta}$ values, i.e.

$$n(\boldsymbol{x}, t+\tau) = \int_{\mathbb{R}^3} n(\boldsymbol{x}-\boldsymbol{\Delta}, t)\, f(\boldsymbol{\Delta}) d\boldsymbol{\Delta}\,. \tag{2}$$

  This is the basic evolution equation for the number density $n(\boldsymbol{x}, t)$. From the technical point of view, it is a continuity equation expressing particle conservation, i.e. that Brownian particles can neither be created, nor can they disappear as a result of the collisions with the fluid molecules.

- By Taylor expansion of the above expression

$$n(\boldsymbol{x}, t+\tau) \approx n(\boldsymbol{x}, t) + \tau \frac{\partial n(\boldsymbol{x}, t)}{\partial t}\,, \tag{3}$$

$$n(\boldsymbol{x}-\boldsymbol{\Delta}, t) \approx n(\boldsymbol{x}, t) - \boldsymbol{\Delta}\cdot\boldsymbol{\nabla} n(\boldsymbol{x}, t) + \frac{\boldsymbol{\Delta}^2}{2}\boldsymbol{\nabla}^2 n(\boldsymbol{x}, t)\,. \tag{4}$$

  Therefore

$$\int_{\mathbb{R}^3} n(\boldsymbol{x}-\boldsymbol{\Delta}, t)\, f(\boldsymbol{\Delta}) d\boldsymbol{\Delta} \approx n(\boldsymbol{x}, t) \int_{\mathbb{R}^3} f(\boldsymbol{\Delta}) d\boldsymbol{\Delta} - \boldsymbol{\nabla} n(\boldsymbol{x}, t)\cdot \int_{\mathbb{R}^3} \boldsymbol{\Delta}\, f(\boldsymbol{\Delta}) d\boldsymbol{\Delta} + \frac{1}{2}\boldsymbol{\nabla}^2 n(\boldsymbol{x}, t) \int_{\mathbb{R}^3} \boldsymbol{\Delta}^2\, f(\boldsymbol{\Delta}) d\boldsymbol{\Delta}\,, \tag{5}$$

  and making use of the normalization of the pdf $f(\boldsymbol{\Delta})$ and the symmetry relation $f(-\boldsymbol{\Delta}) = f(\boldsymbol{\Delta})$, we simplify the first term in the right-hand side of this equation and get rid of the second one. Hence we finally obtain the diffusion equation for the particle density field as

$$\frac{\partial n(\boldsymbol{x}, t)}{\partial t} = D\boldsymbol{\nabla}^2 n(\boldsymbol{x}, t)\,, \tag{6}$$

where the *diffusion coefficient* D is given in terms of the pdf $f(\boldsymbol{\Delta})$ as

$$D = \frac{1}{2\tau}\int_{\mathbb{R}^3} \boldsymbol{\Delta}^2\, f(\boldsymbol{\Delta}) d\boldsymbol{\Delta} = \frac{\langle \boldsymbol{\Delta}^2 \rangle}{2\tau}\,. \tag{7}$$

- To solve this partial differential equation, we first need an initial condition. If initially all Brownian particles are located at the origin, we thus have $n(\boldsymbol{x}, 0) = N\delta(\boldsymbol{x})$. The diffusion equation can be solved via different methods. One of the most common is via Fourier transform of the particle density field. In particular, we define

$$n(\boldsymbol{k}, t) = \int_{\mathbb{R}^3} \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}} n(\boldsymbol{x}, t)\, d\boldsymbol{x}\,, \tag{8}$$

and derive it with respect to time to obtain (after integrating by parts twice and neglecting boundary terms at ∞)

$$\frac{\partial n(\boldsymbol{k}, t)}{\partial t} = -D\boldsymbol{k}^2 n(\boldsymbol{k}, t) \qquad \Rightarrow \qquad n(\boldsymbol{k}, t) = N\mathrm{e}^{-D\boldsymbol{k}^2 t}\,, \tag{9}$$

where we have already used the initial condition mentioned above. Notice that $n(\boldsymbol{k}, t)$ is a Gaussian curve, so its inverse Fourier transform is also a Gaussian, namely

$$n(\boldsymbol{x}, t) = \frac{N}{(4\pi D t)^{3/2}} \mathrm{e}^{-\boldsymbol{x}^2/4Dt}\,. \tag{10}$$

- Another, very instructive way of solving the diffusion equation is via dimensional analysis. For that, first note that the $n(\boldsymbol{x}, t)$ is a particle density field so it has units of inverse volume, $[n(\boldsymbol{x}, t)] = L^{-d}$ in dimension d, with L the unit of length. Moreover, the diffusion constant D defined above has units $[D] = L^2/T$, with T the unit of time, and obviously $[t] = T$. The only nontrivial quantity that can be defined with units of length in the problem is $\sqrt{Dt}$, so $(Dt)^{d/2} n(\boldsymbol{x}, t)$ is a dimensionless quantity which should depend only on other dimensionless

quantities. From variables $\boldsymbol{x}$, t, D we can form a single dimensionless quantity $\boldsymbol{x}/\sqrt{Dt}$. Therefore the most general dependence of the density field on the basic variables that is allowed by dimensional analysis is

$$n(\boldsymbol{x},t) = \frac{1}{(Dt)^{d/2}}\rho(\boldsymbol{\zeta})\,, \qquad \boldsymbol{\zeta} = \frac{\boldsymbol{x}}{\sqrt{Dt}}\,. \tag{11}$$

- The density depends on a single scaling variable rather than on two basic variables $\boldsymbol{x}$ and t. This remarkable feature greatly simplifies analysis of the typical partial differential equations that describe nonequilibrium systems. Equation (11) is often referred to as the scaling ansatz. Finding the right scaling ansatz for a physical problem often represents a large step toward a solution. For the diffusion equation, substituting in the ansatz (11) reduces this partial differential equation to

$$2\boldsymbol{\nabla}^2\rho(\boldsymbol{\zeta}) + \boldsymbol{\zeta}\cdot\boldsymbol{\nabla}\rho(\boldsymbol{\zeta}) + d\rho(\boldsymbol{\zeta}) = 0\,. \tag{12}$$

  Invoking both the symmetry of $\rho(\boldsymbol{\zeta})$ around $\boldsymbol{\zeta} = 0$ and normalization, and integrating twice the above equation, we obtain $\rho(\boldsymbol{\zeta}) = (4\pi)^{-d/2}\mathrm{e}^{-\boldsymbol{\zeta}^2/4}$ for arbitrary dimension d, and using that $\boldsymbol{\zeta} = \boldsymbol{x}/\sqrt{Dt}$ we recover the Gaussian solution for $n(\boldsymbol{x},t)$ described above.

- Whatever the solution method, from the Gaussian solution $n(\boldsymbol{x},t)$ it follows that the average position of the Brownian particle is $\langle\boldsymbol{x}(t)\rangle = 0$ and that the average square position increases linearly with time, namely $\langle\boldsymbol{x}(t)^2\rangle = 6Dt$ in $d = 3$ dimensions. Interestingly, this prediction has been successfully confirmed in experiments and contributed to the acceptance of the atomic theory.

1.2. Langevin description

- However successful Einstein's approach was, it is very phenomenological and cannot yield, for instance, an explicit expression for the diffusion coefficient in terms of microscopic quantities.

- Paul Langevin (1908) initiated a different approach which, in some ways, can be considered complementary to the previous one. In his approach, Langevin focused on the trajectory of a single Brownian particle and wrote down Newton's equation Force = mass × acceleration.

- The trajectory of the Brownian particle is highly erratic and therefore its description would demand a peculiar kind of force. Langevin considered two types of forces acting on the Brownian particle: the usual friction forces which, according to Stokes law, would be proportional to the velocity, and a sort of *"fluctuating"* force $\boldsymbol{\xi}(t)$, which represents the *"erratic"* force that comes from the action of the fluid molecules on the Brownian particle.

- The equation of motion for the Brownian particle becomes then

$$m\frac{d\boldsymbol{v}}{dt} = -6\pi\eta R\boldsymbol{v} + \boldsymbol{\xi}\,, \tag{13}$$

  where η is the viscosity coefficient and R is the radius of the Brownian particle (which is assumed to be spherical). Multiplying both sides of by $\boldsymbol{x}$, one gets

$$\frac{m}{2}\frac{d^2\boldsymbol{x}^2}{dt^2} - m\left(\frac{d\boldsymbol{x}}{dt}\right)^2 = -3\pi\eta R\frac{d\boldsymbol{x}^2}{dt} + \boldsymbol{x}\cdot\boldsymbol{\xi}\,. \tag{14}$$

- Langevin made two assumptions about the fluctuating force $\boldsymbol{\xi}(t)$: that it has mean 0 (collisions do not push the Brownian particle in any preferred direction) and that it is uncorrelated to the actual position of the Brownian particle (the action of the molecules of fluid on the Brownian particle is the same no matter the location of the Brownian particle), that is

$$\langle\boldsymbol{\xi}(t)\rangle = 0 \qquad ; \qquad \langle\boldsymbol{x}\cdot\boldsymbol{\xi}(t)\rangle = \langle\boldsymbol{x}\rangle\cdot\langle\boldsymbol{\xi}(t)\rangle = 0\,. \tag{15}$$

- Taking the averages with respect to all realizations of the random force $\boldsymbol{\xi}(t)$ in the above differential equation and using the previous conditions on the statistics of $\boldsymbol{\xi}(t)$, one gets

$$\frac{m}{2}\frac{d^2\langle \boldsymbol{x}^2\rangle}{dt^2} = m\langle \boldsymbol{v}^2\rangle - 3\pi\eta R\frac{d\langle \boldsymbol{x}^2\rangle}{dt}\,, \tag{16}$$

  which is an equation for the average square position of the Brownian particle.

- Langevin assumed that we are now in a regime in which thermal equilibrium between the Brownian particle and the surrounding fluid has been reached. In particular, this implies that, according to the equipartition theorem, the average kinetic energy of the Brownian particle is $\langle m\boldsymbol{v}^2/2\rangle = 3kT/2$ (k is Boltzmann's constant and T is the fluid temperature).

- One can now solve the previous equation and find that, after some transient time, the asymptotic mean square displacement is given

$$\langle \boldsymbol{x}^2(t)\rangle = \frac{kT}{\pi\eta R}t\,. \tag{17}$$

  This is nothing but Einstein's diffusion law, but now we have an explicit expression for the diffusion coefficient in terms of other macroscopic variables

$$D = \frac{kT}{6\pi\eta R}\,. \tag{18}$$

  In this way, by measuring the Brownian particle diffusion coefficient and the fluid's viscosity (two macroscopic observables) we can obtain an indirect measurement of the size R of the fluid's molecules!

- Langevin's random force $\boldsymbol{\xi}(t)$ is an example of a stochastic process. It is time to proceed to a more precise definition of a stochastic process. The natural machinery is that of probability theory.

2. Stochastic processes

- In Chapter 1, we introduced the concept of a random variable $\hat{\mathbf{x}}$ resulting from a probabilistic experiment. We now define a *stochastic process* as a family $\hat{\mathbf{x}}(t)$ of random variables depending on some continuous, real parameter t. In most applications, t is a physical time and the stochastic process can be thought as performing multiple probabilistic experiments one at each time instant.

- The trajectory followed by the system depends on the outcome of each probabilistic experiment. As a consequence, the knowledge of the initial condition x_0 at time t_0 is not enough to determine the position in phase space of the system at a later time t_1. Instead, the trajectory, and therefore the final state of the system, acquires a probabilistic nature. To fully determine the final state, it is also necessary to know the outcome of all the successive probabilistic experiments between the initial time t_0 and the final time t_f . In fact, each possible set of successive outcomes determines a possible trajectory for the system, all starting at x_0 at time t_0 but ending at different locations at time t_f. The stochastic process can be seen as the collection of all these possible trajectories.

- Arguably, the most well-known example of a stochastic process is that of the random walk. The probabilistic experiment is now a series of binary results representing, for instance, the outcome of repeatedly tossing a coin:

$$(0, 0, 1, 1, 0, 1, 0, 0, 0, 1, 1, 1, 1, 0, 0, 1, \ldots) \tag{19}$$

  where 1 means *"heads"* and 0 means *"tails"*. Consider that the tossing takes place at given times $0, \tau, 2\tau, \ldots$ To the outcome of this set of probabilistic experiments, we associate a one-dimensional function $x(t)$ which starts at $x(0) = 0$ and that moves to the left (right) at time $k\tau$ an amount a $(-a)$ if the k-th result of the tossed coin was 0 (1). In the intermediate times between two consecutive tossings, namely in the times between k and $(k + 1)$, the system just remains in the same location. Fig. 2 shows a typical trajectory.

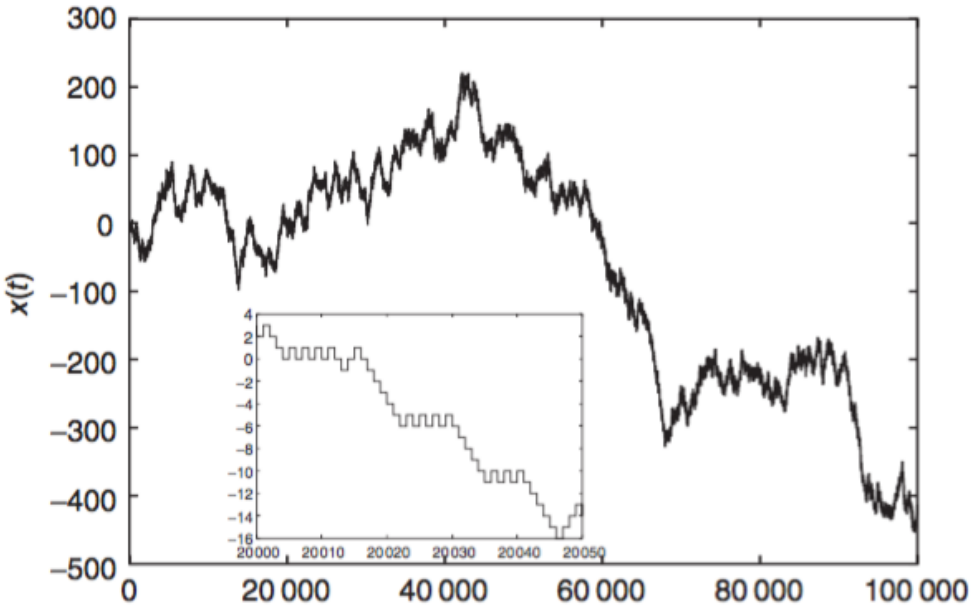

Figure 2. Example of a random walk trajectory. We have taken $\tau = a = 1$ and plotted the resulting trajectory after a large number of steps (10^5). In the inset, we see the fine detail with the discrete jumps occurring at times which are multiples of τ.

- What does one mean by characterizing a stochastic process? Since it is nothing but a continuous family of random variables, a stochastic process will be completely characterized when we give the joint probability density function for the arbitrary set $\{\hat{\mathbf{x}}(t_1), \hat{\mathbf{x}}(t_2), \ldots, \hat{\mathbf{x}}(t_m)\}$, that is, when we give the function $f(x_1, \ldots, x_m; t_1, \ldots, t_m)$ for an arbitrary m. This function is such that

$$f(x_1, \ldots, x_m; t_1, \ldots, t_m) dx_1 \ldots dx_m \tag{20}$$

represents the probability that the random variable $\hat{\mathbf{x}}(t_1)$ takes values in the interval $(x_1, x_1 + dx_1)$, the random variable $\hat{\mathbf{x}}(t_2)$ takes values in the interval $(x_2, x_2 + dx_2)$, and so on [1]. We can see here a generalization of the successions of random variables presented in Chapter 1.

† In a different language, we can say that a complete characterization of the trajectory is obtained by giving the functional probability density function $f([x(t)])$.

- As t is a continuous variable, we have here formally a nonnumerable, infinite number of random variables $\hat{\mathbf{x}}(t)$. However, when we extract from this a finite set $\hat{\mathbf{x}}(t_1), \ldots, \hat{\mathbf{x}}(t_m)$, we can use the same definitions and results as in the case of a succession of random variables.

- A stochastic process is said to be a Markov process if the rather general conditional probability

$$f(x_m; t_m | x_1, \ldots, x_{m-1}; t_1, \ldots, t_{m-1}) \equiv \frac{f(x_1, \ldots, x_m; t_1, \ldots, t_m)}{f(x_1, \ldots, x_{m-1}; t_1, \ldots, t_{m-1})} \tag{21}$$

  is equal to the two-times conditional probability

$$f(x_m; t_m | x_{m-1}; t_{m-1}) \equiv \frac{f(x_{m-1}, x_m; t_{m-1}, t_m)}{f(x_{m-1}; t_{m-1})} \tag{22}$$

  for all times $t_m > t_{m-1} > \ldots t_1$. This Markov property means (loosely speaking) that the probability of the next event depends only on the present state of the system and not on the way it reached its present situation (*history*).

- The Markov property allows us to compute the m-times pdf as

$$f(x_1, \ldots, x_m; t_1, \ldots, t_m) = f(x_m; t_m | x_{m-1}; t_{m-1}) f(x_{m-1}; t_{m-1} | x_{m-2}; t_{m-2}) \ldots f(x_2; t_2 | x_1; t_1) f(x_1; t_1)\,. \tag{23}$$

- The random walk constitutes an example of a Markov process, as the probability of having a particular value of the position at time $(k+1)\tau$ depends only on the particle's location at time $k\tau$ and not on the way it got to this location.

- A particular case of a Markov process is a completely independent process in which an arbitrary set of random variables at different times are independent, and then we are able to write

$$f(x_1, \ldots, x_m; t_1, \ldots, t_m) = f(x_m; t_m) f(x_{m-1}; t_{m-1}) \ldots f(x_2; t_2) f(x_1; t_1)\,. \tag{24}$$

- Another important example of a random process is that of a Gaussian process in which the m-times pdf admits an explicit Gaussian-like form, namely

$$f(x_1, \ldots, x_m; t_1, \ldots, t_m) = \sqrt{\frac{|A|}{(2\pi)^m}} \exp\left[-\frac{1}{2}\sum_{i,j=1}^{m}(x_i - b_i)A_{ij}(x_j - b_j)\right] \tag{25}$$

where $x_i \equiv x(t_i)$, $b_i = \langle x(t_i)\rangle$, and A is a matrix with determinant $|A|$ and components such that $(A^{-1})_{ij} = \langle x(t_i)x(t_j)\rangle - \langle x(t_i)\rangle\langle x(t_j)\rangle$. For Gaussian processes, the explicit form of the pdf above is rarely written down, rather the process is fully characterized by giving the mean value $\langle \hat{\mathbf{x}}(t)\rangle$ and the correlation function $\langle \hat{\mathbf{x}}(t)\hat{\mathbf{x}}(t')\rangle$.

3. Stochastic differential equations

- A stochastic differential equation is a differential equation that contains a stochastic process $\hat{\eta}(t)$: that is, an equation of the form

$$\frac{d\hat{\mathbf{x}}(t)}{dt} = G\left(\hat{\mathbf{x}}(t), t, \hat{\eta}(t)\right) \tag{26}$$

  where G is a given function that depends, in general, on the variable $x(t)$, on the time t, and on the stochastic process $\hat{\eta}(t)$.

- A stochastic differential equation can be seen as a family of ordinary differential equations, one for each outcome of all the successive probabilistic experiments associated with the stochastic process $\hat{\eta}(t)$. As a consequence, for any given initial condition x_0 at time t_0, one has a family of possible trajectories.

- Therefore, $\hat{\mathbf{x}}(t)$, which is the collection of all these possible trajectories, has to be viewed also as a stochastic process and this is why we label it with the "hat" symbol. However, $\hat{\mathbf{x}}(t)$ is not an arbitrary stochastic process, rather it depends on $\hat{\eta}(t)$ in a specific manner determined by the stochastic differential equation, and, as a consequence, the statistical properties of $\hat{\mathbf{x}}(t)$ depend on the statistical properties of $\hat{\eta}(t)$.

- Strictly speaking, *"solving the stochastic differential equation"* means to provide the complete characterization of the stochastic process $\hat{\mathbf{x}}(t)$, namely to give all the m-times pdfs $f(x_1, \ldots, x_m; t_1, \ldots, t_m)$, in terms of the statistical properties of $\hat{\eta}(t)$. However, one has to understand that a complete characterization of a general stochastic process implies the knowledge of a function of an arbitrary number of parameters and is very difficult to carry out in practice.

- On many occasions, one is happy if one can give just the one-time pdf $f(x;t)$ and the two-times pdf $f(x_1, x_2; t_1, t_2)$. In terms of those, it is possible to compute the trajectory averages

$$\langle \hat{\mathbf{x}}(t)^n \rangle = \int_{-\infty}^{\infty} dx\, x^n f(x;t) \tag{27}$$

and the time correlations

$$\langle \hat{\mathbf{x}}(t_1)^n \hat{\mathbf{x}}(t_2)^m \rangle = \int_{-\infty}^{\infty} dx\, x_1^n x_2^m f(x_1, x_2; t_1, t_2)\,. \tag{28}$$

- In general, the function G can depend on the stochastic process $\hat{\eta}(t)$ in an arbitrary way. However, many systems of interest can be described by stochastic differential equations in which $\hat{\eta}(t)$ appears linearly, namely

$$\frac{d\hat{\mathbf{x}}(t)}{dt} = q\left(\hat{\mathbf{x}}(t)\right) + g\left(\hat{\mathbf{x}}(t)\right)\hat{\eta}(t)\,. \tag{29}$$

  This kind of stochastic differential equations, which are the only ones to be considered in this Chapter, are called Langevin equations.

- In this case (Langevin equation), the independent stochastic process $\hat{\eta}(t)$ is usually referred to as noise, a notation that comes from the early days of radio broadcasting when the random fluctuations in the electrical signals taking place in the emitter during the propagation in the atmosphere or at the receiver device led to noises that were actually heard on top of the radio emission.

- Terminology: Following this notation, the term $g\left(\hat{\mathbf{x}}(t)\right)\hat{\eta}(t)$ in Eq. (95) is referred to as the noise term or *diffusion term* whereas $q\left(\hat{\mathbf{x}}(t)\right)$ is the deterministic term or *drift term*. One distinguishes the case in which the function $g\left(\hat{\mathbf{x}}(t)\right)$ is a constant, in which case the noise is said to be additive. Otherwise, the noise is said to be multiplicative.

- For the sake of simplicity in the notation, from now on we will drop the "hats" from the stochastic process and therefore we write the Langevin differential equation as

$$\frac{dx(t)}{dt} = q\left(x(t)\right) + g\left(x(t)\right)\eta(t)\,. \tag{30}$$

- We have already encountered an example of Langevin differential equation in the first Section, the equation introduced by Langevin himself to describe the movement of a Brownian particle. In this case, the independent stochastic process is the random force that acts on the Brownian particle and models the collisions of the water molecules and it appears in the equation as an additive noise, and the deterministic term is the drag induced by the water viscosity. And, somehow, we have already "solved" this Langevin equation when we determined some statistical properties of the movement of the Brownian particle, such as the mean square displacement.

4. White noise

- We proceed now to characterize in a more detailed way the stochastic process $\xi(t)$ that appears in the Langevin equation for the Brownian motion. To do this, we first start with the characterization of another process we have already encountered, the one-dimensional random walk.

- Starting at $x = 0$ at time $t = 0$, the location of the random walker after tossing the coin n times is given by the number of steps taken in the sense that x increases n_1 (number of "heads") minus the number of steps taken in the opposite sense $n_0 = n - n_1$ (number of "tails"),

$$x(n\tau) = (n_1 - n_0)a = (2n_1 - n)a\,. \tag{31}$$

- The probability of having n_1 "heads" after n throws is given by the binomial expression

$$P(n_1) = \binom{n}{n_1} 2^{-n}\,, \tag{32}$$

i.e., a binomial $\hat{\mathbf{N}}_B(\frac{1}{2}, n)$. Hence, the probability that the walker is at a location $x = ra$ after a time $t = n\tau$ is

$$P\left(x(n\tau) = ra\right) = \binom{n}{\frac{n+r}{2}} 2^{-n}\,. \tag{33}$$

- Using now the binomial distribution properties we derived in Chapter 1, in particular the first two central moments ($\langle \hat{\mathbf{N}}_B \rangle = Mp$ and $\langle \hat{\mathbf{N}}_B^2 \rangle = M^2p^2 + Mp(1-p)$ for a $\hat{\mathbf{N}}_B(p, M)$ binomial random variable), we find that

$$\langle x(n\tau) \rangle = (2\langle n_1 \rangle - n)a = 0\,, \tag{34}$$
$$\langle x(n\tau)^2 \rangle = (4\langle n_1^2 \rangle - 4\langle n_1 \rangle n + n^2)a^2 = na^2\,, \tag{35}$$

where we have used that $\langle n_1 \rangle = \frac{n}{2}$ and $\langle n_1^2 \rangle = \frac{n(n+1)}{4}$.

- We also studied in Chapter 1 that in the limit $n \gg 1$ the binomial distribution can be well approximated by a Gaussian distribution (of mean Mp and variance $Mp(1-p)$ for a $\hat{\mathbf{N}}_B(p, M)$ variable), so the associated cumulative distribution function in this limit reads

$$P\left(x(n\tau) \leq ra\right) = \frac{1}{2} + \mathrm{erf}\left(\frac{r}{\sqrt{n}}\right) . \tag{36}$$

- We now take the continuum limit $n \to \infty$, $\tau \to 0$, $r \to \infty$, $a \to 0$ while preserving a finite value for $t = n\tau$, $x = ra$, and $D = a^2/\tau$. In this limit, the random walk process is called the Wiener process $W(t)$ and can be written as

$$P\left(W(t) \leq x\right) = \frac{1}{2} + \mathrm{erf}\left(\frac{x}{\sqrt{Dt}}\right) . \tag{37}$$

  which is the probability distribution function of a Gaussian variable with zero mean and variance Dt. The corresponding probability density function is

$$f(x;t) = \frac{1}{\sqrt{2\pi Dt}} \exp\left(-\frac{x^2}{2Dt}\right) . \tag{38}$$

- The Gaussian Wiener process inherits the Markovian character of the random walk process and can be characterized by giving its mean value and the two-times correlation function:

$$\langle W(t) \rangle = 0 , \tag{39}$$

$$\langle W(t_1) W(t_2) \rangle = D \min(t_1, t_2) . \tag{40}$$

  A plot of a typical realization of the Wiener process is shown in Figure 3.

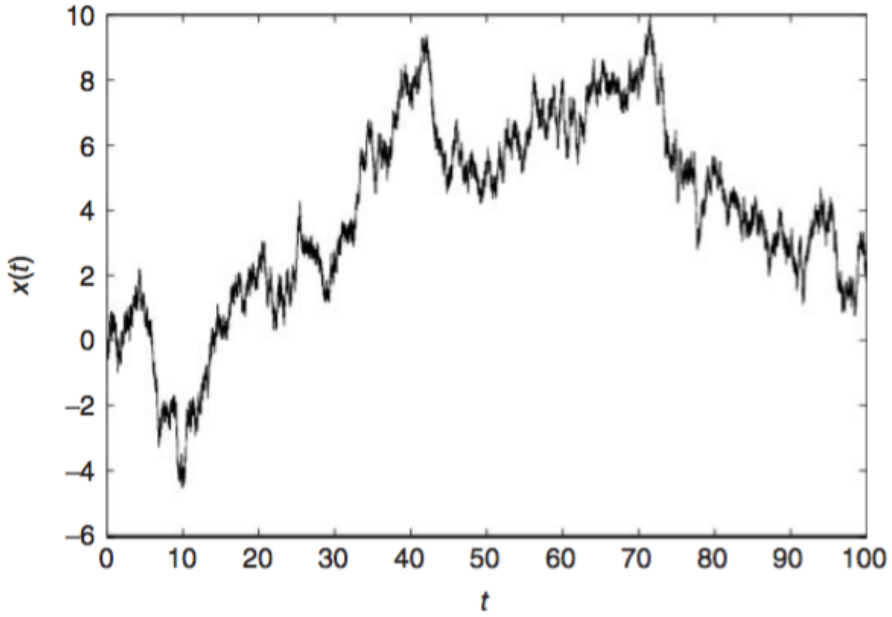

Figure 3. Typical realization of the Wiener process generated using the random walk with $\tau = 10^{-4}$ and $a = 10^{-2}$.

- To understand the correlation at different times, it is sufficient to realize that the increments in a Wiener process, $W(t) - W(0)$, have a Gaussian distribution with zero mean and variance Dt, and that non-overlapping increments are independent. Suppose now that $t_1 < t_2$, and let's use that $W(t_2) = [W(t_2) - W(t_1)] + W(t_1)$ and $W(t_1) = [W(t_1) - W(t_0)]$ since $W(t_0) = 0$, so that

$$\langle W(t_1)W(t_2)\rangle = \langle W(t_1)[W(t_2) - W(t_1)]\rangle + \langle W(t_1)^2\rangle = \langle W(t_1)\rangle\langle [W(t_2) - W(t_1)]\rangle + Dt_1 = Dt_1\,. \qquad (41)$$

  Similarly, if $t_2 < t_1$ we then find, $\langle W(t_1)W(t_2)\rangle = Dt_2$, and hence in general $\langle W(t_1)W(t_2)\rangle = D\min(t_1, t_2)$.

- The random walk process was a sequence of step functions. As a consequence, the Wiener process is continuous but it does not have a well-defined first derivative[1].

† In fact, it is a fractal of dimension 1/2.

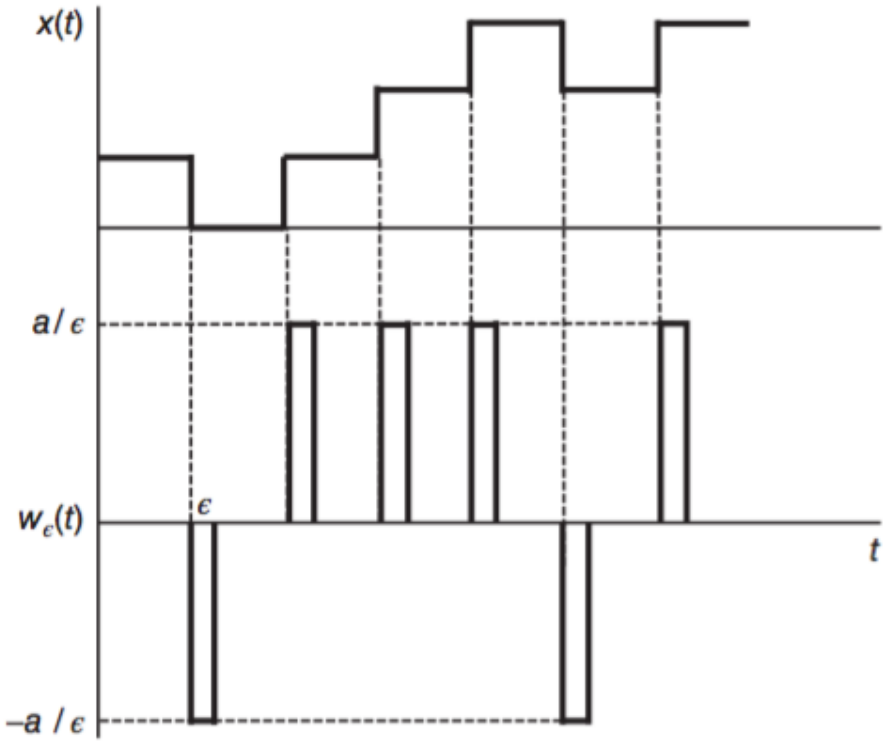

Figure 4. Random walk and its derivative.

- Still, we will define now the white noise random process $\xi(t)$ as the derivative of the Wiener process.
- As we just mentioned that the Wiener process does not have a well-defined derivative, it is not surprising that the result depends on the way the derivative is performed. We first go back to the discrete random walk process $x(t)$ and define a new stochastic process w_ϵ as

$$w_\epsilon = \frac{x(t+\epsilon) - x(t)}{\epsilon} . \tag{42}$$

- A sketch of the process w_ϵ is given in Figure 4. When taking the continuum limit defined above ($n \to \infty$, $\tau \to 0$, $r \to \infty$, $a \to 0$ with finite $t = n\tau$, $x = ra$, and $D = a^2/\tau$), the process $w_\epsilon(t)$ tends to a Gaussian process since

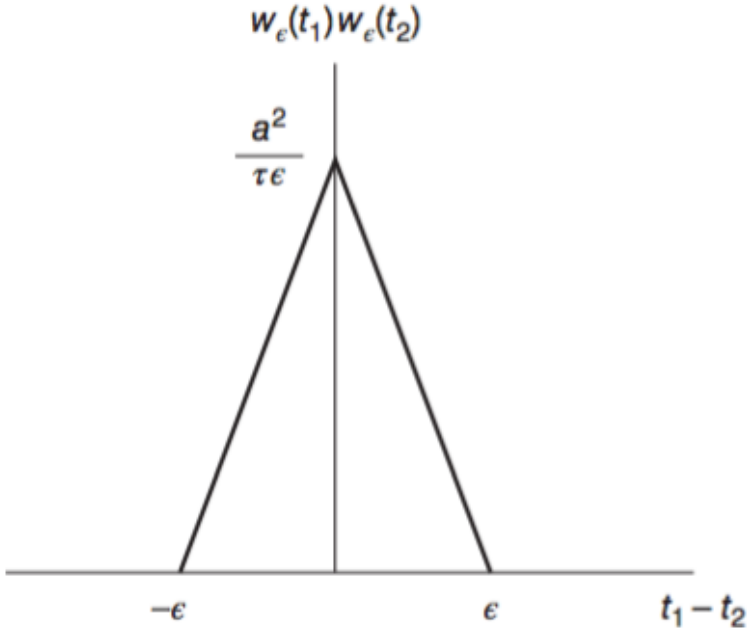

Figure 5. Correlation function for the derivative of the random walk process.

it is a linear combination of Gaussian processes. Therefore, it is sufficiently defined by its mean and correlations:

$$\langle w_\epsilon(t)\rangle = 0\,, \tag{43}$$

$$\langle w_\epsilon(t_1)w_\epsilon(t_2)\rangle = \begin{cases} 0\,, & t_1 - t_2 < -\epsilon\,, \\ a^2/(\tau\epsilon)(1 + (t_1 - t_2)/\epsilon)\,, & -\epsilon \le t_1 - t_2 < 0\,, \\ a^2/(\tau\epsilon)(1 - (t_1 - t_2)/\epsilon)\,, & 0 \le t_1 - t_2 \le \epsilon\,, \\ 0\,, & t_1 - t_2 > \epsilon\,. \end{cases} \tag{44}$$

The shape of the correlation function is shown in Figure 5.

- In the limit $\epsilon \to 0$, the process $w(t) = \lim_{\epsilon\to 0} w_\epsilon(t)$ becomes the derivative of the random walk process and the correlation function above becomes a delta function

$$\langle w(t_1)w(t_2)\rangle = (a^2/\tau)\delta(t_1 - t_2)\,. \tag{45}$$

If we go now to the continuum limit, the random walk tends to the Wiener process and its derivative can be written as $w(t) = D^{1/2}\xi(t)$, where $\xi(t)$ is a Markovian, Gaussian process of mean and correlations

$$\langle \xi(t) \rangle = 0\,, \tag{46}$$

$$\langle \xi(t_1)\xi(t_2) \rangle = \delta(t_1 - t_2)\,, \tag{47}$$

known as white noise. It is the derivative of the Wiener process (for $D = 1$):

$$\xi(t) = \frac{dW(t)}{dt}\,. \tag{48}$$

- The white noise can be understood as a series of pulses each of which is very short but very intense, in a way that their effect is finite. The pulses are independent among them, which represents the perturbations acting on the system in random directions so that the average of all the perturbations is zero.

- The name comes from the fact that its power spectral density (the Fourier transform of the correlation function) is flat: that is, it is the same at all frequencies.

5. Stochastic integrals: Itô and Stratonovich interpretations

- In the previous section, we introduced the white noise stochastic process as the derivative of the Wiener process.
- As the Wiener process does not have a well-defined derivative, the precise interpretation of the white noise process depends on the way the derivative is performed. The most widely used interpretations are those of Itô and Stratonovich.
- To illustrate this, let us consider the integral

$$\int_t^{t'} f(x(s))\xi(s)\,ds \tag{49}$$

where $f(x(t))$ is an arbitrary function of the stochastic process $x(t)$ whose dynamics is given by a Langevin equation of the form

$$\frac{dx(t)}{dt} = q\left(x(t)\right) + g\left(x(t)\right)\xi(t)\,. \tag{50}$$

- The integral above is sometimes also written as

$$\int_t^{t'} f(x(s))\,dW \tag{51}$$

although here we will mainly use the notation $\xi(s)ds$ rather than dW.

- As the integral depends on a stochastic process, each realization of the stochastic process will lead to a different value for the integral, thus integrals of this form are called stochastic integrals.
- Now, we first proceed to compute the integral above for $t' = t + h$ in the limit $h \to 0$. In the so-called Itô interpretation, the result of the integral is

$$\int_t^{t+h} f(x(s))\xi(s)\,ds = f(x(t))\left[W(t+h) - W(t)\right]\,, \tag{52}$$

namely the function is evaluated at the initial time. In the so-called Stratonovich interpretation, the result of the integral is

$$\int_t^{t+h} f(x(s))\xi(s)\,ds = f\left(\frac{x(t)+x(t+h)}{2}\right)[W(t+h)-W(t)]\ . \tag{53}$$

That is, the function is evaluated at a point which is the average of the value of $x(t)$ at the initial and final times.

- In some sense, the existence of the different definitions is related to the definition of the following expression

$$\int_0^\infty dt\,\delta(t) = \begin{cases} 1\,, & \text{Itô} \\ 1/2\,, & \text{Stratonovich} \end{cases} \tag{54}$$

- For a finite integration time, we divide the integration interval $[t,t']$ into N subintervals and take the limit $N\to\infty$. In Itô calculus, this leads to

$$\int_t^{t'} f(x(s))\xi(s)\,ds = \lim_{N\to\infty}\sum_{i=1}^{N} f(x(t_{i-1}))\,[W(t_i)-W(t_{i-1})] \tag{55}$$

where $t_i = t+ih$ with $h=(t'-t)/N$. In Stratonovich calculus, the integral is given

$$\int_t^{t'} f(x(s))\xi(s)\,ds = \lim_{N\to\infty}\sum_{i=1}^{N} f\left(\frac{x(t_{i-1})+x(t_i)}{2}\right)[W(t_i)-W(t_{i-1})]\ . \tag{56}$$

- In general, one could have evaluated the stochastic integral as

$$\int_t^{t'} f(x(s))\xi(s)\,ds = \lim_{N\to\infty}\sum_{i=1}^{N} f\left(\alpha x(t_{i-1})+(1-\alpha)x(t_i)\right)[W(t_i)-W(t_{i-1})]\ . \tag{57}$$

with any arbitrary α such that $0\le\alpha\le 1$. The Itô interpretation corresponds to $\alpha=1$, whereas the Stratonovich interpretation corresponds to $\alpha=1/2$.[1]

† The choice $\alpha=0$ corresponds to Klimontovich interpretation of the stochastic integral.

- In the past, there has been much argument on which was the *"correct"* interpretation. There is no such thing. It is just a matter of convention. This means that a Langevin stochastic differential equation with white noise is not completely defined unless we specify the interpretation considered. Once the interpretation has been fixed, then any expression such as the integral above has a unique meaning and its result can be uniquely determined.

- In many cases of interest, the Stratonovich interpretation turns out to be more *"natural"* since it commutes with the continuum limit described in previous sections ($n \to \infty$, $\tau \to 0$, $r \to \infty$, $a \to 0$ with finite $t = n\tau$, $x = ra$, and $D = a^2/\tau$). In this Chapter, unless otherwise stated, we will always use the Stratonovich interpretation.

- In addition, the Stratonovich interpretation allows the use of the familiar rules of calculus, such as the change of variables in an integration step. Instead, in the Itô interpretation, in some instances, one cannot use the ordinary rules of calculus and instead use specific ones, which tend to be more cumbersome.

- One advantage of the Itô interpretation is that the stochastic process at time t, $x(t)$, is not correlated with the white noise acting on the system at the same time $\xi(t)$: that is, $\langle x(t)\xi(t)\rangle = 0$. This is not the case with the Stratonovich interpretation.

- To illustrate the differences between the Itô and Stratonovich interpretations, let us consider the simple integral

$$\int_{t_a}^{t_b} W(s)\xi(s)\,ds\,. \tag{58}$$

  In Itô calculus, this leads to

$$\begin{aligned}\int_{t_a}^{t_b} W(s)\xi(s)\,ds &= \lim_{N\to\infty}\sum_{i=1}^{N} W(t_{i-1})\left[W(t_i) - W(t_{i-1})\right] \\ &= \frac{1}{2}\left[W(t_b)^2 - W(t_a)^2\right] - \frac{1}{2}\lim_{N\to\infty}\sum_{i=1}^{N}\left[W(t_i) - W(t_{i-1})\right]^2\,. \end{aligned} \tag{59}$$

In Stratonovich calculus, this leads to

$$\int_{t_a}^{t_b} W(s)\xi(s)\,ds = \lim_{N\to\infty}\sum_{i=1}^{N}\frac{1}{2}\left[W(t_i)+W(t_{i-1})\right]\left[W(t_i)-W(t_{i-1})\right]$$
$$= \frac{1}{2}\lim_{N\to\infty}\sum_{i=1}^{N}\left[W(t_i)^2 - W(t_{i-1})^2\right] = \frac{1}{2}\left[W(t_b)^2 - W(t_a)^2\right]\,. \tag{60}$$

- Thus, while Stratonovich calculus leads to a result for the integral which is what one would expect from ordinary calculus, in the Itô interpretation there is a nonintuitive additional term.

- There is a simple relation between the Langevin stochastic differential equations written in both interpretations. The rule is that the Langevin equation

$$\frac{dx(t)}{dt} = q_I(x(t)) + g_I(x(t))\xi(t)\,. \tag{61}$$

in the Itô sense is equivalent to the Langevin equation

$$\frac{dx(t)}{dt} = q_I(x(t)) - \frac{1}{2}g_I(x)g_I'(x) + g_I(x(t))\xi(t)\,. \tag{62}$$

in the Stratonovich sense. Therefore, the equivalent Stratonovich equation to a given Itô equation has an additional drift term.

- Notice that, when the noise is additive, $g(x)$ is a constant, and then the Langevin equation is the same in both interpretations.

6. Numerical integration of stochastic differential equations with Gaussian white noise

- As already described in the pedagogical example of Brownian motion of Section 1, stochastic processes admit two complementary descriptions: the first one, in terms of stochastic differential equations, comes from Langevin's treatment of the Brownian motion that focuses on trajectories; the second one, in terms of the Fokker-Planck equation, focuses on probabilities as in Einstein's treatment of the Brownian motion.

- In principle, both ways allow the characterization of the stochastic process, such as calculating averages or correlations. We will study in detail the Fokker-Planck equation in next sections.

- However, in many instances, looking at the individual trajectories allows one to extract valuable information that is not so easy to obtain from the time evolution of the probability density function. In fact, the generation and visualization of some representative trajectories is usually quite helpful in providing physical understanding of what is going on.

- Therefore, here we focus on numerical methods to generate trajectories. To be more precise, consider that we have a given Langevin stochastic differential equation and we would like to obtain with a computer several representative trajectories.

- As the noise will be different in each trajectory, even if we start always from the same initial condition, all the trajectories will be different and, in principle, there will be an infinite number of them.

- Therefore, we cannot aim at generating all possible trajectories, but just a finite number of them. Still, if properly done, this finite number of trajectories can be sufficient to obtain averages or correlations with a certain degree of accuracy.

- Because for a given realization of the noise a stochastic equation becomes an ordinary differential equation, one may naively consider that it can be numerically integrated using any standard method, such as the Euler method,

the popular fourth-order Runge-Kutta method, or a predictor-corrector method. Therefore, one may question the need for this section altogether.

- The point is that these standard methods cannot be used to integrate typical stochastic differential equations. To illustrate this, let us consider the equation

$$\dot{x}(t) = q(x) + g(x)\xi(t) \tag{63}$$

where, as usual, $\xi(t)$ is a Gaussian white noise of zero mean and δ-type correlations.

- The standard methods to integrate the ordinary differential equations $\dot{x}(t) = q(x)$ assume that the functions $q(x)$ are well behaved to a certain extent. For example, the Euler method assumes that these functions are differentiable. Runge-Kutta methods or predictor-corrector methods assume that they are differentiable to a higher order.

- However, as described in previous sections, the white noise is not a differentiable function. Even for a single realization of the white noise term, $\xi(t)$ is highly irregular and not differentiable even at first order. As discussed above, it can be thought of as a series of Dirac delta functions spread all over the real axis.

- Therefore the stochastic contribution is a non-analytical function and it is necessary to use a different kind of algorithms. The basic idea behind the construction of these new algorithms is to integrate: whereas the values of the white noise or its derivatives are not well defined, the integrals are (for instance, the first integral is the Wiener process, which is a continuous function).

6.1. A simple example

- Let us start with a simple stochastic differential equation

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

- In fact, as the right-hand side does not depend on x, this equation can be solved exactly, the solution being

$$x(t) = x(0) + \int_0^t f(s)\; ds + \int_0^t \xi(s)\; ds \tag{65}$$

which can be written as

$$x(t) = x(0) + F(t) + W(t)\,, \tag{66}$$

where $F(t) = \int_0^t f(s)\; ds$ and $W(t)$ is the Wiener process.

- This expression indicates that $x(t)$ is a Gaussian process whose mean and correlation are given by

$$\langle x(t)\rangle = x(0) + F(t)\,, \tag{67}$$

$$\langle x(t)x(t')\rangle = [x(0) + F(t)][x(0) + F(t')] + \min(t, t')\,. \tag{68}$$

- In what follows, we disregard that the equation can be solved and instead focus on a numerical solution in which we generate trajectories: that is, we want to obtain $x(t)$ at discrete time intervals

$$x(t+h) = x(t) + \int_t^{t+h} \dot{x}(s)\; ds = x(t) + \int_t^{t+h} f(s)\; ds + \int_t^{t+h} \xi(s)\; ds \tag{69}$$

- Introducing now

$$f_h(t) = \int_t^{t+h} f(s)\; ds \qquad \text{and} \qquad w_h(t) = \int_t^{t+h} \xi(s)\; ds = W(t+h) - W(t) \tag{70}$$

we can write

$$x(t+h) = x(t) + f_h(t) + w_h(t)\,. \tag{71}$$

- In this equation, as $w_h(t)$ is the difference of the Wiener process at two different times, it is a Gaussian process and can be fully characterized by giving the mean and correlations

$$\langle w_h(t)\rangle = \int_t^{t+h} \langle \xi(s)\rangle \, ds = 0\,, \tag{72}$$

$$\langle w_h(t) w_h(t')\rangle = \int_t^{t+h}\int_{t'}^{t'+h} \langle \xi(s)\xi(u)\rangle \, ds \, du = \int_t^{t+h}\int_{t'}^{t'+h} \delta(s-u) \, ds \, du\,. \tag{73}$$

- To evaluate the integral, we make use of the properties of the Dirac δ-function. We can assume, without loss of generality, that $t' > t$. If $t' > t + h$, the integral is 0, as there is no overlap in the integration intervals, and the delta function vanishes. If $t \leq t' < t + h$, the double integral equals the length of the overlap interval, that is

$$\langle w_h(t) w_h(t')\rangle = \int_{t'}^{t+h} ds = t - t' + h\,. \tag{74}$$

  In particular, for $t = t'$ one has $\langle w_h(t)^2\rangle = h$.

- In numerical calculations, time takes always discrete values as multiples of the integration time step. If we consider discrete times $t = t_i = ih$, $t' = t_j = jh$, and note that the condition $t \leq t' < t + h$ above implies that $i = j$ in the discrete case, the correlation becomes

$$\langle w_h(t_i) w_h(t_j)\rangle = h\delta_{ij}\,. \tag{75}$$

- We introduce now a set of independent Gaussian random variables $\{u_i\}$ of zero mean and variance 1:

$$\langle u_i\rangle = 0 \qquad ; \qquad \langle u_i u_j\rangle = \delta_{ij} \tag{76}$$

  in terms of which we can write

$$w_h(t_i) = \sqrt{h} u_i\,. \tag{77}$$

- The set of independent Gaussian variables u_i can be generated using the methods discussed in Chapter 1. Finally, the recurrence relation to generate numerically trajectories of the stochastic process defined above is

$$x(t_0) = x_0\,, \tag{78}$$
$$x(t_{i+1}) = x(t_i) + f_h(t_i) + \sqrt{h}u_i\,. \tag{79}$$

- For small h, the deterministic contribution $f_h(t)$ can be approximated as $f_h(t) = hf(t)$, implying that for small h the deterministic contribution is of order h^1 and successive contributions go as h^2, h^3, and so on. In contrast the stochastic contribution is of order $h^{1/2}$, an order that never appears in methods for numerical integration of ordinary differential equations.

- A consequence of this scaling with the time step is that at very small times the stochastic contribution, being of order $h^{1/2}$, dominates over the deterministic one, of order h. Over the course of an integration for a large period, the stochastic contributions on average will be generally of less importance than the deterministic ones.

- The previous recurrence equation can be readily implemented in a program that allows the generation of a numerical trajectory. Averages and correlations can be obtained by integrating many trajectories with independent values for the noise. Besides, if the initial condition is given by a probability distribution, it is also necessary to sample this distribution of initial conditions.

6.2. General stochastic differential equations

- We now apply the same ideas to a stochastic differential equation of the general form

$$\dot{x}(t) = q(x) + g(x)\xi(t) \tag{80}$$

- We discretize the time $t = t_i = t_0 + ih$, where t_0 is the time of the initial condition, $i = 0, 1, 2, \ldots$, and h is the integration time step.

- Our goal is to obtain a recurrence relation that provides the value of $x(t_{i+1})$ as a function of $x(t_i)$. Proceeding as before, we write

$$x(t_{i+1}) = x(t_i) + \int_{t_i}^{t_{i+1}} q(x(s))\, ds + \int_{t_i}^{t_{i+1}} g(x(s))\xi(s)\, ds\,. \tag{81}$$

- Now we assume that the functions $q(x(s))$ and $g(x(s))$ are differentiable functions and expand them in a Taylor series around $x = x(t_i)$

$$q(x(s)) = q(x(t_i)) + \left.\frac{dq(x)}{dx}\right|_{x(t_i)} (x(s) - x(t_i)) + O\Big[(x(s) - x(t_i))^2\Big]\,, \tag{82}$$

$$g(x(s)) = g(x(t_i)) + \left.\frac{dg(x)}{dx}\right|_{x(t_i)} (x(s) - x(t_i)) + O\Big[(x(s) - x(t_i))^2\Big]\,. \tag{83}$$

- Substitution of these expansions into the above recurrence leads to

$$\begin{aligned} x(t_{i+1}) = x(t_i) + hq(x(t_i)) + q'(x(t_i)) \int_{t_i}^{t_{i+1}} (x(s) - x(t_i))\, ds + hO\Big[(x(s) - x(t_i))^2\Big] \\ + w_h(t_i)g(x(t_i)) + g'(x(t_i)) \int_{t_i}^{t_{i+1}} (x(s) - x(t_i))\xi(s)\, ds + w_h(t_i)O\Big[(x(s) - x(t_i))^2\Big] \end{aligned} \tag{84}$$

where we used the notation $f'(x(t_i)) \equiv \left.\frac{df(x)}{dx}\right|_{x(t_i)}$ for any function f.

- As in the simple previous example, we can write $w_h(t_i) = \sqrt{h}u_i$ (with u_i Gaussian variables of zero mean and unit variance) and is of order $h^{1/2}$.

- In the previous recurrence, we can identify the deterministic contribution (terms involving $q(x,t)$) and the noise contribution (terms involving $g(x,t)$). Now, we want to keep all the terms up to order h.

- On the deterministic part, the third (the integral) and the fourth terms of the first line are of higher order and can be disregarded.

- On the stochastic part, the term $w_h(t_i)g(x(t_i))$ is only of order $h^{1/2}$, so we may need to keep also some other terms. To do so, we need to evaluate the integral present in the second term of the stochastic part, namely

$$\int_{t_i}^{t_{i+1}} (x(s) - x(t_i))\xi(s)\, ds\,. \tag{85}$$

- The evaluation of this integral requires the knowledge of $x(s)$ at a time s between t_i and t_{i+1} which is not known *a priori*. As we do not know $x(s)$, a way to proceed is to estimate $x(s)$ using an approximation. To zeroth order, $x(s) = x(t_i)$, which does not give any contribution in the integral above. The next order is $h^{1/2}$, and to this order one has

$$x(s) - x(t_i) = g(x(t_i)) \int_{t_i}^{s} \xi(u)\, du + O[h]\,. \tag{86}$$

- Using this approximation in the integral above, we have

$$\int_{t_i}^{t_{i+1}} (x(s) - x(t_i))\xi(s)\, ds = g(x(t_i)) \int_{t_i}^{t_{i+1}} \int_{t_i}^{s} \xi(u)\xi(s)\, ds\, du + w_h(t_i)O[h]\,. \tag{87}$$

- Since $w_h(t)$ is of order $h^{1/2}$, the last term is of order $h^{3/2}$ and hence negligible in our approximation. The double integral can be readily evaluated as

$$\int_{t_i}^{t_{i+1}} \int_{t_i}^{s} \xi(u)\xi(s)\, ds\, du = \int_{t_i}^{t_{i+1}} [W(s) - W(t_i)]\, \xi(s)\, ds = \int_{t_i}^{t_{i+1}} W(s)\xi(s)\, ds - W(t_i) \int_{t_i}^{t_{i+1}} \xi(s)\, ds \tag{88}$$

As discussed in previous sections about Stratonovich calculus, the first integralis given by $[W(t_{i+1})^2 - W(t_i)^2]/2$. Therefore

$$\int_{t_i}^{t_{i+1}} \int_{t_i}^{s} \xi(u)\xi(s)\, ds\, du = \frac{1}{2}[W(t_{i+1})^2 - W(t_i)^2] - W(t_i)[W(t_{i+1}) - W(t_i)] = \frac{1}{2}[W(t_{i+1}) - W(t_i)]^2 = \frac{w_h(t_i)^2}{2} \tag{89}$$

- Therefore, the second term of the stochastic part of the right-hand side of Eq. (84) is given by

$$g'(x(t_i)) \int_{t_i}^{t_{i+1}} (x(s) - x(t_i))\xi(s)\, ds = \frac{g'(x(t_i))}{2} w_h(t_i)^2 + O[h^{3/2}]\,. \tag{90}$$

- We note that the corrections given by the last term on the right-hand side of Eq. (84) are of higher order

$$w_h(t_i) O\Big[\big(x(s) - x(t_i)\big)^2\Big] = O[h^{3/2}] \tag{91}$$

and therefore can be disregarded.

- Collecting all the results, up to order h, Eq. (84) can be written as

$$x(t_{i+1}) = x(t_i) + hq(x(t_i)) + g(x(t_i))h^{1/2}u_i + \frac{1}{2}g(x(t_i))g'(x(t_i))hu_i^2 + O[h^{3/2}]\,. \tag{92}$$

This recurrence relation is known in the literature as the *Milshtein algorithm*.

- If the noise is additive, $g(x(t))$ does not depends on x, and $g(x(t)) = D$ with D being a constant, then the resulting algorithm is called the Euler or Euler-Maruyama algorithm

$$x(t+h) = x(t) + hq(x(t)) + \sqrt{Dh}u(t) + O[h^{3/2}]\,. \tag{93}$$

- While the Euler-Maruyama algorithm is suitable for additive noise only, it is sometimes used for stochastic differential equations with multiplicative noise

$$x(t_{i+1}) = x(t_i) + hq(x(t_i)) + g(x(t_i))h^{1/2}u_i + O[h]\,. \tag{94}$$

  Notice that the error at each time step is of order $O[h]$, which is pretty bad. Therefore, we strongly advise against using the Euler-Maruyama algorithm for stochastic differential equations with multiplicative noise.

- We note, however, that, while the Euler-Maruyama algorithm is not suitable for stochastic differential equations understood in the Stratonovich sense, it is appropriate for Itô stochastic differential equations.

- Coming back to Stratonovich calculus, we would like to note that, in general, the Milshtein algorithm can be readily used in stochastic differential equations in which the drift or the diffusion terms depend explicitly on the time.

- When integrating ordinary differential equations, one typically uses algorithms in which the error is of much higher order than $O[h^{3/2}]$, such as, for example, a fourth-order Runge-Kutta algorithm, where the error at each time step is of order $O[h^5]$. Therefore, it is natural to ask for a higher order method than the Milshtein algorithm for stochastic processes.

- In principle, one can try to extend the previous calculations and obtain such development. However, for a stochastic equation of the general form, as soon as one goes beyond the Milshtein method, the development involves more random processes, which are not Gaussian and which have nontrivial correlations among them. As a consequence, in general, it is very difficult to generate these terms appropriately and the resulting algorithms are more cumbersome.

- We now provide a program to generate a trajectory of a stochastic differential eq. using the Milshtein algorithm:

```
! Program implementing the Milshtein method for stochastic differential eqs
program Milshtein
implicit none
double precision :: x,t,x_0,t_0,h,h_sqrt,uh
double precision :: q,g,gprime,ran_g
integer :: i_step,n_step,i_write,n_write
external q,g,gprime
x_0=1.d0
t_0=0.d0
h=0.01d0
n_write =200
n_step=10
h_sqrt=sqrt(h)
x=x_0
t=t_0
do i_write=1,n_write
   do i_step=1,n_step
      uh=h_sqrt*ran_g()
      x=x+uh*g(t,x)+h*q(t,x)+0.5d0*g(t,x)*gprime(t,x)*uh**2
      t=t+h
   enddo
   write (20,*) t,x
enddo
end program Milshtein
```

7. The Fokker-Planck equation

- So far, we have focused on trajectories to describe stochastic processes following the initial Langevin approach for Brownian motion.
- The alternative approach introduced by Einstein focuses on probabilities rather than in trajectories and, as it happens with the Langevin approach, it can be extended way beyond the Brownian motion.
- In what follows, we are going to determine an equation for one-time probability distribution for a stochastic process described by a Langevin equation with white noise. This is called the Fokker-Planck equation and it is a generalization of the diffusion equation obtained by Einstein to describe the Brownian process.
- To be more precise, we want to find an equation for the one-time probability distribution function $f(x,t)$ for a stochastic process $x(t)$ which arises as a solution of a stochastic differential equation with a Gaussian white noise $\xi(t)$, defined such that $\langle \xi(t) \rangle = 0$ and $\langle \xi(t)\xi(t') \rangle = \delta(t-t')$, i.e.

$$\frac{dx(t)}{dt} = q(x(t)) + g(x(t))\xi(t)\,. \tag{95}$$

7.1. Heuristic derivation of the Fokker-Planck equation

- We first derive a general form of the Fokker-Planck equation by using heuristic arguments. In subsequent sections we turn to an alternative, more formal derivation.
- In order to simplify equations, we now introduce the following notation associated to the time discretization, already used before: $t_i = t_0 + ih$, $x_i = x(t_i)$.
- Integrating the above Langevin equation, we can write the following recurrence relation for x_{i+1} in terms of x_i, see Eq. (81),

$$x_{i+1} = x_i + \int_{t_i}^{t_{i+1}} q(x(s))\; ds + \int_{t_i}^{t_{i+1}} g(x(s))\xi(s)\; ds\,. \tag{96}$$

- Using now a general discretization scheme with $0 \leq \alpha \leq 1$, see Eq. (57), and noting that $h \ll 1$, we can write

$$\int_{t_i}^{t_{i+1}} g(x(s))\xi(s)\, ds \approx g(\alpha x_i + (1-\alpha)x_{i+1})[W(t_{i+1}) - W(t_i)]$$
$$= g(\alpha x_i + (1-\alpha)x_{i+1})\omega_h(t_i) = \sqrt{h}u_i\; g(\alpha x_i + (1-\alpha)x_{i+1})\,, \tag{97}$$

where we have used that $\omega_h(t_i) \equiv W(t_{i+1}) - W(t_i) = \sqrt{h}u_i$, with u_i independent Gaussian random variables with mean $\langle u_i \rangle = 0$ and variance $\langle u_i u_j \rangle = \delta_{ij}$, see Eq. (76) in previous Section. For consistency reasons, we can also write

$$\int_{t_i}^{t_{i+1}} q(x(s))\, ds \approx hq(\alpha x_i + (1-\alpha)x_{i+1})\,. \tag{98}$$

- We hence have the recurrence

$$x_{i+1} = x_i + hq_\alpha^i + \sqrt{h}u_i g_\alpha^i\,, \tag{99}$$

where we have defined $q_\alpha^i \equiv q(\alpha x_i + (1-\alpha)x_{i+1})$ and $g_\alpha^i \equiv g(\alpha x_i + (1-\alpha)x_{i+1})$.

- Since $h \ll 1$, we can now introduce the Taylor expansion

$$q_\alpha^i \equiv q(\alpha x_i + (1-\alpha)x_{i+1}) = q(x_i + (1-\alpha)(hq_\alpha^i + \sqrt{h}u_i g_\alpha^i)) = q(x_i) + O(\sqrt{h})\,, \tag{100}$$
$$g_\alpha^i \equiv g(\alpha x_i + (1-\alpha)x_{i+1}) = g(x_i + (1-\alpha)(hq_\alpha^i + \sqrt{h}u_i g_\alpha^i)) = g(x_i) + (1-\alpha)\sqrt{h}u_i g(x_i)g'(x_i) + O(h)\,. \tag{101}$$

where we hace used the above recurrence relation. We hence arrive at

$$x_{i+1} = x_i + \sqrt{h}u_i g(x_i) + h\left[q(x_i) + (1-\alpha)u_i^2 g(x_i)g'(x_i)\right] + O(h^{3/2})\,. \tag{102}$$

Note that this equation reduces to Eq. (92) for $\alpha = 1/2$, i.e. for the Stratonovich interpretation.

- In order to find an equation for the probability $f(x_{i+1}, t_{i+1})$ that the stochastic process takes a value x_{i+1} at time t_{i+1}, we have to sum over all possible previous-step values x_i and all possible noise realizations u_i, weighted by their respective probabilities, and restricted to those pairs leading to x_{i+1}. Mathematically,

$$f(x_{i+1}, t_{i+1}) = \left\langle \int dx_i \; f(x_i, t_i) \; \delta\left[\left(x_i + \sqrt{h} u_i g(x_i) + h\left[q(x_i) + (1-\alpha) u_i^2 g(x_i) g'(x_i)\right]\right) - x_{i+1}\right]\right\rangle_{u_i} \tag{103}$$

  where $\langle \cdot \rangle_{u_i}$ denotes averaging over the Gaussian white noise, and the Dirac δ-function constrains the (continuous) sum over (x_i, u_i) to those pairs leading to x_{i+1}.

- We may use now that $h \ll 1$ to Taylor-expand the Dirac δ-function

$$\begin{aligned}
\delta\left[\left(x_i + \sqrt{h} u_i g(x_i) \; + \; h\left[q(x_i) + (1-\alpha) u_i^2 g(x_i) g'(x_i)\right]\right) - x_{i+1}\right] &= \delta(x_i - x_{i+1}) \\
&+ \left(\sqrt{h} u_i g(x_i) + h\left[q(x_i) + (1-\alpha) u_i^2 g(x_i) g'(x_i)\right]\right) \frac{\partial}{\partial x_i} \delta(x_i - x_{i+1}) \\
&+ \frac{h}{2} u_i^2 g(x_i)^2 \frac{\partial^2}{\partial x_i^2} \delta(x_i - x_{i+1}) + O(h^{3/2}) \, .
\end{aligned} \tag{104}$$

- Plugging this expansion into the equation for $f(x_{i+1}, t_{i+1})$ we obtain

$$\begin{aligned}
f(x_{i+1}, t_{i+1}) = & \overbrace{\left\langle \int dx_i \; f(x_i, t_i) \; \delta(x_i - x_{i+1}) \right\rangle_{u_i}}^{f(x_i, t_{i+1})} + \sqrt{h} \overbrace{\left\langle \int dx_i \; f(x_i, t_i) \; u_i \; g(x_i) \right\rangle_{u_i}}^{0 \text{ since } \langle u_i \rangle = 0} \\
& + h \left\langle \int dx_i \; f(x_i, t_i) \; q(x_i) \frac{\partial}{\partial x_i} \delta(x_i - x_{i+1}) \right\rangle_{u_i} \\
& + h(1-\alpha) \left\langle \int dx_i \; f(x_i, t_i) \; u_i^2 \; g(x_i) g'(x_i) \frac{\partial}{\partial x_i} \delta(x_i - x_{i+1}) \right\rangle_{u_i} \\
& + \frac{h}{2} \left\langle \int dx_i \; f(x_i, t_i) \; u_i^2 \; g(x_i)^2 \frac{\partial^2}{\partial x_i^2} \delta(x_i - x_{i+1}) \right\rangle_{u_i}
\end{aligned} \tag{105}$$

- Integrating by parts, neglecting boundary terms, and taking into account that $\langle u_i \rangle = 0$ and $\langle u_i^2 \rangle = 1$, we hence arrive at

$$f(x_{i+1}, t_{i+1}) = f(x_{i+1}, t_{i+1}) + h \left[-\frac{\partial}{\partial x_{i+1}} \left(q(x_{i+1})\ f(x_{i+1}, t_{i+1}) \right) \right.$$
$$\left. - \ (1-\alpha) \frac{\partial}{\partial x_{i+1}} \left(g(x_{i+1})\ g'(x_{i+1})\ f(x_{i+1}, t_{i+1}) \right) + \frac{1}{2} \frac{\partial^2}{\partial x_{i+1}^2} \left(g(x_{i+1})^2\ f(x_{i+1}, t_{i+1}) \right) \right] \tag{106}$$

- Taking the limit $h \to 0$, and simplifying the notation, we finally obtain

$$\frac{\partial f(x,t)}{\partial t} = -\frac{\partial}{\partial x} \left[(q(x) + (1-\alpha) g(x) g'(x))\ f(x,t) \right] + \frac{1}{2} \frac{\partial^2}{\partial x^2} \left[g(x)^2 f(x,t) \right] . \tag{107}$$

which is nothing but the general form ($0 \leq \alpha \leq 1$) of the Fokker.Planck equation associated to the Langevin equation (95).

7.2. Alternative derivation I: Deterministic dynamics and sampling over initial conditions

- We now turn to an alternative, more formal derivation of the Fokker-Planck equation. In what follows we adopt the Stratonovich interpretation ($\alpha = 1/2$) of Langevin equation, unless otherwise specified.

- Let us consider first the corresponding deterministic equation

$$\frac{dx(t)}{dt} = q(x(t)) , \qquad x(t=0) = x_0 . \tag{108}$$

The solution of this equation is a (deterministic) function $x(t) = F(t, x_0)$. Nevertheless, for our purposes, we can see $x(t)$ as a random variable whose probability density function $\rho(x;t)$ gives the probability to find the system

at a given location x in phase space at a time t. As the trajectory followed by the system is uniquely determined once the initial condition is given, this probability is zero everywhere except at the location $x = F(t, x_0)$. Therefore, the probability density function is a delta function, given by

$$\rho(x;t) = \delta(x - F(t, x_0))\,. \tag{109}$$

- We can now transform $x(t)$ into a stochastic process by simply letting the initial condition x_0 to become a random variable. We have now an ensemble of trajectories, each one starting from a different initial condition x_0. In this case, the probability density function $\rho(x,t)$ is obtained by averaging the above pdf over the distribution of the initial conditions

$$\rho(x;t) = \langle \delta(x - F(t, x_0)) \rangle_{x_0}\,. \tag{110}$$

  In fact, $\rho(x,t)$ can be seen as a density function.

- To visualize it, we assume that at time $t = 0$ we have unleashed a large number of particles N with one particle at each of the initial conditions we are considering in our ensemble. We let the particles evolve according to the dynamics and, after a time t, $\rho(x,t)dx$ measures the fraction of particles located in the interval $(x, x + dx)$

$$\rho(x,t)dx = \frac{\#\text{ particles in } (x, x+dx) \text{ at time } t}{\text{total } \#\text{ of particles}} = \frac{n(x,t)}{N}dx\,. \tag{111}$$

- Assuming there are no sinks nor sources of particles, then the number of particles at a given location changes as a result of the number of particles crossing the borders. Let $J(x,t)$ be the flux of particles:

$$J(x,t)dx = \frac{\#\text{ particles that cross the point } x \text{ in the interval } (t, t+dt)}{N}\,. \tag{112}$$

  J has a direction. For convenience, we consider $J > 0$ if the particle moves to the right. Then

$$\rho(x;t+dt)dx - \rho(x;t)dx = J(x,t)dt - J(x+dx,t)dt\,. \tag{113}$$

- The first term on the left-hand side (LHS) is the final number of particles at x, whereas the second is the initial one. Thus, the LHS corresponds to the variation of the number of particles at x from time t to time $t + dt$. On the right-hand side (RHS), the first term is the number of particles that have entered or left during the interval $(t, t + dt)$ through the left boundary, whereas the second term measures the number of particles that crossed the right boundary.

- Therefore, one has

$$\frac{\rho(x; t + dt) - \rho(x; t)}{dt} = \frac{J(x, t) - J(x + dx, t)}{dx} \tag{114}$$

  which in the continuous limit corresponds to

$$\frac{\partial \rho(x; t)}{\partial t} = -\frac{\partial J(x, t)}{\partial x} \tag{115}$$

  which is nothing but the continuity equation.

- As particles move following a deterministic dynamics given by $\frac{dx(t)}{dt} = q(x(t))$, trajectories cannot cross and therefore in a one-dimensional system they cannot advance from one to the other. The particles that will cross the point x in the interval $(t, t + dt)$ are all those located in the interval $dx = q(x)dt$. Therefore,

$$J(x, t) = \frac{n(x, t)}{N}\frac{dx}{dt} = \rho(x; t)q(x)\,. \tag{116}$$

  Replacing this in the above continuity equation, one gets the Liouville equation

$$\frac{\partial \rho(x; t)}{\partial t} = -\frac{\partial}{\partial x}\left[q(x)\rho(x; t)\right]\,. \tag{117}$$

7.3. Alternative derivation II: Full stochastic dynamics

- We consider now the full stochastic differential equation

$$\frac{dx(t)}{dt} = q(x(t)) + g(x(t))\xi(t)\,. \tag{118}$$

- We can repeat the above argument for a given realization of the noise term. The probability density function $f(x;t)$ will be the average of $\rho(x;t)$ with respect to the noise distribution:

$$f(x;t) = \langle \rho(x;t) \rangle_\xi = \langle \delta(x - x(t,x_0)) \rangle_{x_0,\xi}\,. \tag{119}$$

- $\rho(x;t)$ satisfies the continuity equation. Now the current is given by

$$J(x,t) = \rho(x;t)\frac{dx(t)}{dt} = \rho(x;t)\left[q(x(t)) + g(x(t))\xi(t)\right]\,. \tag{120}$$

Therefore

$$\frac{\partial \rho(x;t)}{\partial t} = -\frac{\partial}{\partial x}\left[(q(x) + g(x(t))\xi(t))\rho(x;t)\right]\,. \tag{121}$$

- Taking now the average over the noise term, we get

$$\begin{aligned}\frac{\partial f(x;t)}{\partial t} &= \left\langle \frac{\partial \rho(x;t)}{\partial t} \right\rangle_\xi = -\frac{\partial}{\partial x}\left\langle \left[q(x) + g(x(t))\xi(t)\right]\rho(x;t)\right\rangle_\xi \\ &= -\frac{\partial}{\partial x}\left[q(x)\left\langle \rho(x;t)\right\rangle_\xi\right] - \frac{\partial}{\partial x}\left[g(x(t))\left\langle \xi(t)\rho(x;t)\right\rangle_\xi\right]\,.\end{aligned} \tag{122}$$

The averages on the first term of the right-hand side (RHS) can be easily performed, $\langle \rho(x;t) \rangle = f(x;t)$.

- However, averages on the second term of the RHS are more cumbersome, since one has to keep in mind that the density distribution ρ depends on $x(t)$, which itself depends functionally on the noise through the Langevin dynamics.

- In fact, to be precise, the average on the second term of the RHS should be written as $\langle \xi(t)\rho(x[\xi(t)];t)\rangle$. This average can be obtained by using Novikov's theorem, which establishes that, for any Gaussian stochastic process $\xi_G(t)$ with zero mean, $\langle \xi_G(t)\rangle = 0$, and for any functional of the noise $\mathcal{F}[\xi_G(t)]$ one has

$$\langle \xi_G(t)\mathcal{F}[\xi_G(t)]\rangle_{\xi_G} = \int_0^t ds\, \langle \xi_G(t)\xi_G(s)\rangle_{\xi_G} \left\langle \frac{\delta \mathcal{F}}{\delta \xi_G(s)} \right\rangle_{\xi_G} \tag{123}$$

  where the last term is the functional derivative of $\mathcal{F}[\xi_G(t)]$ with respect to $\xi_G(s)$. We note that Novikov's theorem is quite general; it only requires the noise to be Gaussian but does not require it to be white, namely, the noise correlations need not be a delta function.

- In our case, we have a Gaussian white noise, for which the correlation is a delta function. This allows us to evaluate the integral easily:

$$\langle \xi(t)\rho(x[\xi(t)];t)\rangle = \int_0^t ds\, \delta(t-s) \left\langle \frac{\delta \rho(x[\xi(t)];t)}{\delta \xi(s)} \right\rangle_\xi = \frac{1}{2} \left\langle \frac{\delta \rho(x[\xi(t)];t)}{\delta \xi(s)}\Big|_{s=t} \right\rangle_\xi , \tag{124}$$

  where we have used the Stratonovich convention to perform the integral over the δ-function. By using the chain rule in functional calculus, this can be computed as follows

$$\left\langle \frac{\delta \rho(x[\xi(t)];t)}{\delta \xi(s)}\Big|_{s=t} \right\rangle_\xi = \left\langle \frac{\delta x(t)}{\delta \xi(s)}\Big|_{s=t} \frac{\partial \rho(x;t)}{\partial x(t)} \right\rangle_\xi = -\frac{\partial}{\partial x} \left\langle \frac{\delta x(t)}{\delta \xi(s)}\Big|_{s=t} \rho(x;t) \right\rangle_\xi . \tag{125}$$

- To evaluate the functional derivative of the stochastic process $x(t)$ with respect to the noise, we use a formal solution of the Langevin stochastic differential equation, namely

$$x(t) = x_0 + \int_0^t ds\, q(x(s)) + \int_0^t ds\, g(x(s))\xi(s)\,. \tag{126}$$

Then

$$\frac{\delta x(t)}{\delta \xi(s)}\Big|_{s=t} = g(x(t)) \tag{127}$$

and

$$\left\langle \frac{\delta \rho(x[\xi(t)];t)}{\delta \xi(s)}\Big|_{s=t} \right\rangle_\xi = -\frac{\partial}{\partial x}\left[g(x(t))\left\langle \rho(x;t)\right\rangle_\xi\right] = -\frac{\partial}{\partial x}\left[g(x(t))f(x;t)\right]\,. \tag{128}$$

- Using this result in Eq. (122) above, we finally get the Fokker-Planck equation for the probability density function:

$$\frac{\partial f(x;t)}{\partial t} = -\frac{\partial}{\partial x}\left[q(x)f(x;t)\right] + \frac{1}{2}\frac{\partial}{\partial x}\left[g(x)\frac{\partial}{\partial x}\left[g(x)f(x;t)\right]\right]\,. \tag{129}$$

- The procedure we have used to derive the Fokker-Planck equation for a single variable can be extended to the case of several variables. Consider a set of stochastic variables $\boldsymbol{x} = (x_1, x_2, \ldots, x_N)$ whose dynamics is given by the set of Langevin equations to be considered in the Stratonovich interpretation

$$\frac{dx_i}{dt} = q_i(\boldsymbol{x},t) + \sum_{j=1}^{N} g_{ij}(\boldsymbol{x},t)\xi_j(t) \qquad\qquad i = 1,\ldots,N \tag{130}$$

where we allow for the drift terms $q_i(\boldsymbol{x},t)$ and diffusion terms $g_{ij}(\boldsymbol{x},t)$ to explicitly depend on the time. $\xi_j(t)$ are uncorrelated Gaussian white noises with zero mean, that is

$$\langle \xi_i(t)\rangle = 0 \qquad ; \qquad \langle \xi_i(t)\xi_j(s)\rangle = \delta_{ij}\delta(t-s) \qquad \forall i,j \in [1,N] \tag{131}$$

- By using a straightforward extension of the method used for one variable, one can prove that the one-time probability density function $f(\boldsymbol{x};t)$ satisfies the following multivariate Fokker-Planck equation

$$\frac{\partial f(\boldsymbol{x};t)}{\partial t} = -\sum_{i=1}^{N} \frac{\partial}{\partial x_i} \left[q_i(\boldsymbol{x},t) f(\boldsymbol{x};t)\right] + \frac{1}{2} \sum_{i,j,k=1}^{N} \frac{\partial}{\partial x_i} \left[g_{ik}(\boldsymbol{x},t) \frac{\partial}{\partial x_j} \left[g_{jk}(\boldsymbol{x},t) f(\boldsymbol{x};t)\right]\right] . \tag{132}$$

- It can also be written in the form of a continuity equation

$$\frac{\partial f(\boldsymbol{x};t)}{\partial t} + \sum_{i=1}^{N} \frac{\partial}{\partial x_i} J_i(\boldsymbol{x},t) = 0 , \tag{133}$$

  where the probability currents $J_i(\boldsymbol{x},t)$ are given by

$$J_i(\boldsymbol{x},t) = q_i(\boldsymbol{x},t) f(\boldsymbol{x};t) - \frac{1}{2} \sum_{j,k=1}^{N} \left[g_{ik}(\boldsymbol{x},t) \frac{\partial}{\partial x_j} \left[g_{jk}(\boldsymbol{x},t) f(\boldsymbol{x};t)\right]\right] . \tag{134}$$

- If $q_i(\boldsymbol{x},t)$ and $g_{ij}(\boldsymbol{x},t)$ do not depend explicitly on time, i.e. $q_i(\boldsymbol{x},t) = q_i(\boldsymbol{x})$ and $g_{ij}(\boldsymbol{x},t) = g_{ij}(\boldsymbol{x})$, then the Fokker-Planck equation is called *homogeneous*.

7.4. Stationary solution to the Fokker-Planck equation

- In general, solutions of the Fokker-Planck equation are difficult to be found analytically.
- An exception is the stationary solution in the case of a single variable. In this case, imposing

$$\left.\frac{\partial f(\boldsymbol{x};t)}{\partial t}\right|_{f^{\rm st}(x)} = 0 \tag{135}$$

one has that in the steady state the current must fulfill

$$\frac{\partial}{\partial x} J^{\rm st}(x) = 0\,, \tag{136}$$

That is, $J^{\rm st}(x)$ must be a constant.

- The simplest situation is when this constant is zero, which means that in the stationary state there is no flux of probability, and then one has

$$q(x) f^{\rm st}(x) - \frac{g(x)}{2}\frac{d}{dx}\left[g(x) f^{\rm st}(x)\right] = 0\,. \tag{137}$$

This can be rewritten as

$$\frac{2q(x) - g(x)g'(x)}{g(x)^2} dx = \frac{df^{\rm st}(x)}{f^{\rm st}(x)}\,. \tag{138}$$

- Integrating on both sides, and considering that a and b are the lower and upper boundaries for the variable x, one has

$$\int_a^x dy\, \frac{2q(y) - g(y)g'(y)}{g(y)^2} = \log\left[f^{\rm st}(x)\right]\Big|_a^x \tag{139}$$

Then $f^{\rm st}(x) = C\Psi(x)$ where

$$\Psi(x) = \exp \int_a^x dy\, \frac{2q(y) - g(y)g'(y)}{g(y)^2} = \frac{|g(a)|}{|g(x)|} \exp \int_a^x dy\, \frac{2q(y)}{g(y)^2} \tag{140}$$

and C is a normalization constant such that $\int_a^b f^{\rm st}(x)\, dx = 1$.

- Let's now consider the case in which the stationary current is a nonzero constant, $J^{\rm st}(x) = J$. This situation, in which there is a stationary flux of probability, is usually encountered in systems with periodic boundary conditions, for which $f^{\rm st}(a) = f^{\rm st}(b)$, with a and b being the left and right boundaries of the system, respectively.

- J is not arbitrary; its value is determined by the normalization of $f^{\rm st}(x)$ and by the boundary value of the probability density function. In this case, one has

$$q(x)f^{\rm st}(x) - \frac{g(x)}{2}\frac{d}{dx}\left[g(x)f^{\rm st}(x)\right] = J\,. \tag{141}$$

which is an equation similar to the one above but with an additional inhomogeneous term.

- It can be shown that the current is given by

$$J = \left[\frac{g(b)^2}{\Psi(b)} - \frac{g(a)^2}{\Psi(a)}\right] \frac{f^{\rm st}(a)}{\displaystyle\int_a^b \frac{dy}{\Psi(y)}} \tag{142}$$

and the stationary solution is

$$f^{\rm st}(x) = C \frac{\dfrac{g(b)^2}{\Psi(b)} \displaystyle\int_a^x \frac{dy}{\Psi(y)} + \frac{g(a)^2}{\Psi(a)} \int_x^b \frac{dy}{\Psi(y)}}{\dfrac{g(x)^2}{\Psi(x)} \displaystyle\int_a^b \frac{dy}{\Psi(y)}} \tag{143}$$

where C is a normalization constant.