Part I

Basic concepts

IOP Publishing

Nonequilibrium Statistical Mechanics
Basic concepts, models and applications
Alessandro Sarracino, Andrea Puglisi and Angelo Vulpiani

Chapter 1

From Brownian motion to the Langevin equation

> Multa videbis enim plagis ibi percita caecis // commutare viam retroque repulsa reverti // nunc huc nunc illuc in cunctas undique partis; // scilicet hic a principiis est omnibus rerum[a].
>
> De rerum natura, II 129–132
> Lucretius

This chapter focuses on Brownian motion, a phenomenon of immense significance in both physics and mathematics. The theoretical study and the experimental investigation of Brownian motion allowed us to understand the physical reality of atoms. Moreover, the Langevin equation, introduced to describe Brownian motion, laid the foundation for the theory of stochastic processes and it is still one of the pillars for the study of nonequilibrium statistical mechanics.

1.1 From the surprising behaviour of pollen grains to the existence of atoms

In 1827 the botanist Robert Brown, during his study on an Australian plant (*Clarkia Pulchella*), discovered a phenomenon, now called Brownian motion, which had a rather important role in physics [1]. Using a microscope, Brown observed that pollen grains, which are small at macroscopic level, but large (say a few microns) compared

[a] 'For there you will see how many things set in motion by unseen blows change their course and beaten back return back again, now this way, now that way, in all directions. You may be sure that all take their restlessness from the first-beginnings.' Trans. W H D Rouse, revised by Martin F Smith (Harvard University Press: Cambridge, MA and London).

doi:10.1088/978-0-7503-6229-0ch1

with the molecules, suspended in water show a rapid and irregular movement. At the beginning, such a phenomenon was considered only a sort of curiosity; however, after some decades the relevance of Brownian motion and its connection with thermodynamics was realized. For instance, the Italian physicist Cantoni in 1868 claimed that Brownian motion is a *'beautiful and direct experimental demonstration of the fundamental principles of the mechanical theory of heat'* [2]. It is remarkable that the main actors of the building of statistical mechanics (Clausius, Maxwell, Boltzmann and Gibbs) did not show interest in Brownian motion.

The scientist who fully understood the relevance of Brownian motion and was able to develop a theory which constituted the basic point for the acceptance of the atomic structure of matter was Einstein [3]. In his *annus mirabilis* of 1905, Einstein introduced a groundbreaking theory. Essentially, he proposed that the motion of pollen grains in a liquid is governed by two key principles: (a) Stokes' law, which describes the frictional force acting on a body moving through a fluid, and (b) the equipartition of kinetic energy, which ensures that energy is evenly distributed among the system's degrees of freedom, including both the fluid particles and the grain undergoing Brownian motion. The ingenious and bold hypothesis is the validity of both the above assumptions; this in spite of the fact that the mass of the pollen grains, or more generally of a colloidal particle, is much larger that the mass of the molecules.

Once the phenomenon is formalized, it is possible to show that the particle performs an irregular motion whose (Gaussian) statistical features can be computed; namely, one has—for the average over many realizations

$$\langle (x(t) - x(0))^2 \rangle \simeq 2Dt,$$

where $x(t)$ is one component of the position of the particle at time t, and D is the diffusion coefficient whose value is given by the Einstein–Smoluchowski law:

$$D = \frac{k_B T}{6\pi\eta r}, \tag{1.1}$$

where η is the viscosity of the fluid, r the radius of the particle, and k_B the Boltzmann constant, i.e. the ratio of the constant of gases R and the Avogadro number N_A, $k_B = R/N_A$.

In the following we will discuss the great relevance of the Einstein–Smoluchowski relation, which allows us to link the experimentally measurable values of the coefficients D, T, η and r (which are macroscopic quantities), with the Avogadro number.

1.1.1 Langevin's approach

In 1908, a few years after Einstein's paper, Langevin was able to reproduce the results in a simple and elegant way by introducing the first example of the so-called stochastic differential equation [4]. Starting from the basic equation of the mechanics $F = ma$, Langevin, following the ideas of Einstein, split the force in two parts, a systematic one (due to the friction between the colloidal particle and the

fluid) and a random one related to the collisions of the (fast) molecules of the liquid with the colloidal particle.

In presenting Langevin's approach, we follow his original treatment which is rather heuristic; later we shall discuss the problems for the building of a consistent mathematical theory of stochastic differential equations. For the sake of simplicity, we consider the one-dimensional case:

$$m\frac{d}{dt}v = f_T\,, \tag{1.2}$$

where f_T is the force acting on the particle.

In the case of a spherical particle of radius r, with velocity v in a fluid with viscosity η, we have the friction force, given by the Stokes law, $f_{\text{Stokes}} = -6\pi\eta r v$. In addition, we consider the random force given by the collisions of the fluid molecules with the particle, $F_R(t)$ (in modern terms this is white noise). Therefore, we have the evolution equation:

$$m\frac{d}{dt}v = -6\pi\eta r v + F_R(t). \tag{1.3}$$

Let us note that without the random force we have $v(t) = v(0)e^{-\frac{t}{\tau}}$, with

$$\tau = \frac{m}{6\pi\eta r}. \tag{1.4}$$

For colloidal particles of size of the order of a few microns, in a standard fluid we have $\tau \sim 10^{-6}$–10^{-7} s.

Next, let us write equation (1.3) in the form

$$\frac{d^2}{dt^2}x = -\frac{1}{\tau}\frac{d}{dt}x + \frac{1}{m}F_R(t); \tag{1.5}$$

multiplying by x, and using the identities

$$\begin{aligned} x\frac{d}{dt}x &= \frac{1}{2}\frac{d}{dt}x^2, \\ x\frac{d^2}{dt^2}x + \left(\frac{d}{dt}x\right)^2 &= \frac{1}{2}\frac{d^2}{dt^2}x^2, \end{aligned} \tag{1.6}$$

one obtains

$$\frac{d^2}{dt^2}x^2 = -\frac{1}{\tau}\frac{d}{dt}x^2 + \frac{2}{m}xF_R(t) + 2\left(\frac{d}{dt}x\right)^2. \tag{1.7}$$

Let us now perform an average: for instance one can imagine considering a large number of independent colloidal particles, obtaining

$$\frac{d^2}{dt^2}\langle x^2\rangle = -\frac{1}{\tau}\frac{d}{dt}\langle x^2\rangle + \frac{2}{m}\langle xF_R(t)\rangle + 2\left\langle\left(\frac{d}{dt}x\right)^2\right\rangle. \tag{1.8}$$

The term $\langle xF_R(t)\rangle$ is originated by the collisions of the molecules with the particle and it is natural to assume that it is zero. Since we are interested in the long-time behaviour of the particle, we can assume that, as a consequence of the many collisions, the colloidal particle is in thermal equilibrium with the fluid:

$$\frac{1}{2}m\langle \dot{x}^2\rangle = \frac{1}{2}k_{\mathrm{B}}T. \tag{1.9}$$

Therefore, we have

$$\frac{d^2}{dt^2}\langle x^2\rangle = -\frac{1}{\tau}\frac{d}{dt}\langle x^2\rangle + 2\frac{k_{\mathrm{B}}T}{m}, \tag{1.10}$$

whose solution is rather easy: assuming $\langle x^2\rangle|_{t=0} = 0$ and $\frac{d}{dt}\langle x^2\rangle|_{t=0} = 0$[1], we have

$$\langle x(t)^2\rangle = 2\frac{k_{\mathrm{B}}T\tau}{m}\left[t + \tau\left(e^{-\frac{t}{\tau}} - 1\right)\right]. \tag{1.11}$$

For $t \gg \tau$, we have

$$\langle x(t)^2\rangle \simeq 2\frac{k_{\mathrm{B}}T\tau}{m}t, \tag{1.12}$$

and therefore, using (1.4), we obtain the expression of D given in equation (1.1).

1.1.2 The relevance of Brownian motion

It is worth noting that Einstein's interest in Brownian motion was not for its own sake; in his autobiographical notes he wrote: *'My major aim in this was to find facts which would guarantee as much as possible the existence of atoms …'*. To prove (or disprove) the existence of atoms was indeed Einstein's purpose, as clearly stated in his paper on Brownian motion: *'If the movement discussed here can actually be observed (together with the laws relating to it that one would expect to find), then classical thermodynamics can no longer be looked upon as applicable with precision to bodies even of dimensions distinguishable in a microscope: an exact determination of actual atomic dimensions is then possible. On the other hand, had the prediction of this movement proven to be incorrect, a weighty argument would be provided against the molecular-kinetic conception of heat'* [3].

After the theoretical work of Einstein (and Smoluchowski) and some experiments on the diffusion of colloidal particles performed by Svedberg, the conclusive experimental contribution was given by Perrin with his study on the sedimentation and the diffusion [5], see figure 1.1. Perrin was able to determine N_A from the measurement of D: the agreement of N_A with values obtained independently definitively closed the heated controversy about the physical existence of atoms

[1] This is not really important in the asymptotic limit, however, it comes from the physical fact that the motion of the Brownian particle at very short time ($t \ll \tau$) is ballistic, i.e. it is straight with some finite velocity $x(t \to 0) \approx v_0 t$; therefore, one expects $\langle x^2\rangle|_{t\to 0} \approx \langle v_0^2\rangle t^2$. The solution in equation (1.11) shows that $\langle v_0^2\rangle = \frac{k_{\mathrm{B}}T}{m}$.

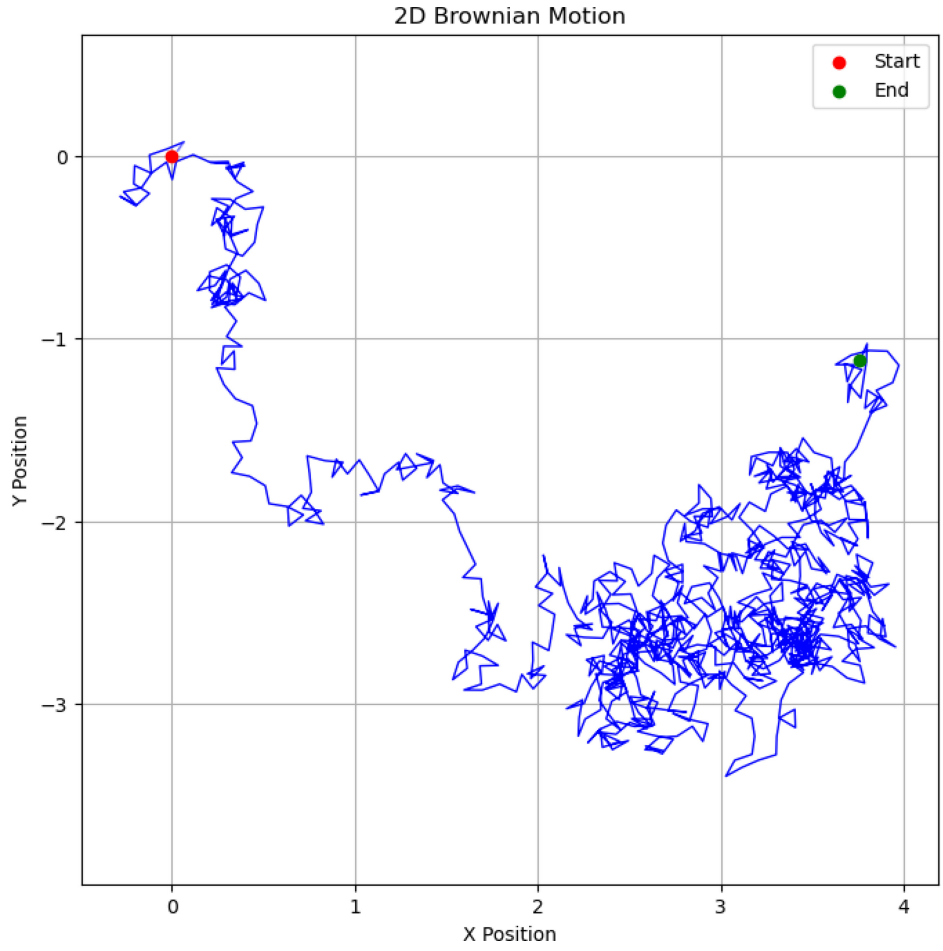

Figure 1.1. A simple example of Brownian motion, not very different from what was observed by Perrin in his renowned experiment. The figure has been generated numerically: it is actually a kind of random walk in two dimensions, where each step has a fixed length and a uniformly distributed (in $(0, 2\pi)$) random angle.

between Boltzmann, on one side, and Mach and Ostwald, on the other, who regarded atoms as useful, but merely as a mathematical tool for constructing a coherent description of Nature.

In a lecture in Paris, in 1911, Arrhenius, summarizing the works of Einstein and Perrin, declared: *'after this, it does not seem possible to doubt that the molecular theory entertained by the philosophers of antiquity, Leucippus and Democritus, has attained the truth at least in essentials'*.

1.2 Toward a mathematical theory

The reader can have the feeling that in the previous discussion some non-trivial aspects have been hidden: actually, we just repeated the straightforward argument in

the original Langevin's paper. Now we shall discuss the delicate aspects in the mathematical treatment, introducing the Fokker–Planck equation and its link with the stochastic differential equations.

In order to convince the reader that such a careful mathematical treatment is necessary, we show what happens if one tries to compute the mean squared velocity of the Brownian particle using the same argument as Langevin used for the mean squared displacement. In fact, an equation for the mean squared velocity is immediately obtained by multiplying by v the Langevin equation (1.3), and taking an average, which results in:

$$\frac{1}{2}m\frac{d}{dt}\langle v^2\rangle = -6\pi\eta r\langle v^2\rangle + \langle F_R(t)v(t)\rangle. \tag{1.13}$$

Now, if one assumes that $\langle F_R(t)v(t)\rangle = 0$ as Langevin assumed for $\langle x(t)F_R(t)\rangle$, the result $\langle v^2\rangle \to 0$ for $t \gg \tau$ is obtained, which is in paradoxical contradiction with Einstein and Langevin assumptions of energy equipartition. Actually, the force $F_R(t)$ is a subtle mathematical function whose properties have been clarified in the first decades of the 20th century, giving rise to the theory of stochastic processes.

1.2.1 A discrete time version of the Langevin equation

In order to understand the mathematical difficulties in the treatment of stochastic differential equations, it can be useful, at least at a pedagogical level, to study a discrete time version, due to Lorentz, see [6], of the Langevin equation:

$$v_{n+1} = av_n + bw_n, \tag{1.14}$$

where n is a non-negative integer, $v_n = v(t_n)$, with $t_n = n\Delta t$, $a = 1 - \frac{\Delta t}{\tau}$ and b is a constant whose value will be discussed in the following; the $\{w_n\}$ are independent identically distributed Gaussian variables with zero mean and unitary variance which are indicated with $\mathcal{N}(0, 1)$. Clearly the terms bw_n and av_n correspond to the random force and the Stokes friction, respectively.

Let us assume that at the initial time the probability density of v_0 is Gaussian with mean value $\langle v_0\rangle$ and variance σ_0^2, i.e. $\mathcal{N}(\langle v_0\rangle, \sigma_0^2)$. Since the w_n are independent of v_n, using the well known result that a linear combination of Gaussian variables is Gaussian, we have that at time n the probability density is $\mathcal{N}(\langle v_n\rangle, \sigma_n^2)$. It is quite easy, from (1.14), to show that

$$\begin{aligned} \langle v_n\rangle &= a^n\langle v_0\rangle, \\ \langle v_{n+1}^2\rangle &= a^2\langle v_n^2\rangle + b^2. \end{aligned} \tag{1.15}$$

Noting that $a < 1$, we have that $\langle v_n\rangle = a^n\langle v_0\rangle$ tends to zero for $n \to \infty$. For the variance $\sigma_n^2 = \langle v_n^2\rangle - \langle v_n\rangle^2$, we have

$$\sigma_{n+1}^2 = a^2\sigma_n^2 + b^2, \tag{1.16}$$

whose solution is:

$$\sigma_n^2 = \langle v^2 \rangle + a^{2n}(\sigma_0^2 - \langle v^2 \rangle), \tag{1.17}$$

where $\langle v^2 \rangle = \frac{b^2}{1-a^2}$; therefore, for $n \to \infty$, $\sigma_n^2 \to \sigma_\infty^2 = \langle v^2 \rangle$. So we have that starting with an arbitrary Gaussian distribution after a short transient we have always a Gaussian with zero mean and variance $\langle v^2 \rangle$.

Let us now consider the position $x_n = x(n\Delta t)$ which evolves with the rule

$$x_{n+1} = x_n + v_n \Delta t.$$

It is easy to show that for large n one has

$$\langle (x_n - x_0)^2 \rangle = n\Delta t^2 \langle v^2 \rangle + 2n\Delta t^2 \sum_{j=1}^{\infty} \langle v_j v_0 \rangle.$$

Since $\langle v_j v_0 \rangle = \langle v^2 \rangle a^j$, we can determine the diffusion coefficient as the limit for $n \to \infty$ and $\Delta t \to 0$, of the ratio

$$\frac{\langle (x_n - x_0)^2 \rangle}{2n\Delta t},$$

obtaining the result $D = \langle v^2 \rangle \tau$, i.e. equation (1.1).

Let us consider again σ_∞^2; for small Δt, we have

$$\sigma_\infty^2 = \langle v^2 \rangle = \frac{b^2}{1-a^2} \simeq \frac{b^2}{2\Delta t}\tau,$$

and therefore

$$b \simeq \sqrt{\frac{2\langle v^2 \rangle}{\tau}}\sqrt{\Delta t}\,. \tag{1.18}$$

This allows us to understand the mathematical issue of the continuous limit; writing

$$\frac{v_{n+1} - v_n}{\Delta t} = -\frac{1}{\tau} v_n + \sqrt{\frac{2\langle v^2 \rangle}{\tau}} \frac{1}{\sqrt{\Delta t}} w_n, \tag{1.19}$$

and taking the limit $\Delta t \to 0$, we have the formal expression

$$\frac{d}{dt} v = -\frac{v}{\tau} + \sqrt{\frac{2\langle v^2 \rangle}{\tau}}\,\xi, \tag{1.20}$$

where $\xi = \lim_{\Delta t \to 0} (w_n / \sqrt{\Delta t})$ is a divergent quantity!

From a mathematical point of view it is better to avoid the derivative and it is more appropriate to write

$$dv = -\frac{v}{\tau} dt + \sqrt{\frac{2\langle v^2 \rangle}{\tau}}\, dW, \tag{1.21}$$

where W is the so-called Wiener process, i.e. a Gaussian process with $\langle W(t)\rangle = 0$, $\langle dW(t)\rangle = 0$ and $\langle dW(t)^2\rangle = dt$ (see the following for more details)[2].

1.2.2 The Fokker–Planck equation

The Langevin equation involves a noisy term whose treatment can appear ambiguous; for such a reason it is desirable to introduce an equation for the time evolution of the probability distribution of the stochastic process associated with the Langevin equation [7–9].

Let us briefly describe how to obtain the Fokker–Planck equation ruling the evolution of the probability density. We start recalling the evolution rule in a Markov chain for the probability to be in the state i at time $t + 1$:

$$P_i(t+1) = \sum_j P_j(t) P_{j\to i},$$

where $P_j(t)$ is the probability to be in the state j at time t and $P_{j\to i}$ is the transition probability form j to i. In the case of continuous states we have

$$P(x, t+\Delta t) = \int P(y, t) W(y, t \to x, t+\Delta t)\, dy, \tag{1.22}$$

where now $P(x, t)$ is the probability density at time t, and $W(y, t \to x, t + \Delta t)$ is the transition rate starting from y at time t to be in x at $t + \Delta t$.

Let us introduce the increment $\Delta x(\Delta t) = z = x - y$ and consider the following conditions: (a) small Δt, and (b) small Δx; the precise meaning will be clear in the following. We consider the case where W is function of x, z and Δt, i.e. $W(y, t \to x, t + \Delta t) = \psi(x - z, z, \Delta t)$. Equation (1.22) can be written in the form

$$P(x, t+\Delta t) = \int P(x - z, t)\psi(x - z, z, \Delta t) dz.$$

For small z we can expand $P(x - z, z, \Delta t)\psi(x - z, z, \Delta t)$ around x:

$$\begin{aligned} P(x, t+\Delta t) &= \int P(x - z, t)\psi(x - z, z, \Delta t) dz \\ &= P(x, t) \int \psi(x, z, \Delta t) dz - \frac{\partial}{\partial x}\left[P(x, t) \int \psi(x, z, \Delta t) z dz \right] \\ &+ \frac{1}{2}\frac{\partial^2}{\partial x^2}\left[P(x, t) \int \psi(x, z, \Delta t) z^2 dz \right] - \cdots \end{aligned} \tag{1.23}$$

Let us note that $\int \psi(x, z, \Delta t) dz = 1$, while $\int \psi(x, z, \Delta t) z dz = E(\Delta x | x)$ is the mean value of Δx starting from x, in a time interval Δt; in a similar way

[2] Some readers may wonder what is the meaning of an equation with differentials instead of derivatives. Naively, writing $dx = a(t)df(t)$ seems identical to writing $x(t) - x(t_0) = \int_{t_0}^{t} a(t')df(t')$. The meaning of the latter is a so-called Riemann–Stieltjes integral, which is—roughly speaking—the continuous limit ($N \to \infty$) of an integral sum $\sum_{i=0}^{N-1} a_i[f(t_i + dt) - f(t_i)]$ where $t_i = t_0 + idt$, $dt = (t - t_0)/N$ and a_i is $a(t)$ evaluated at some time $t \in [t_i, t_{i+1}]$. For the differential of the Wiener process, dW, however, this notion must be carefully adapted for two reasons: (1) it is a stochastic function, (2) even if continuous, it is non-differentiable anywhere. This point will be briefly discussed in the following.

$\int \psi(x, z, \Delta t) z^2 dz = E(\Delta x^2|x)$ and so on. Therefore, since $P(x, t + \Delta t) = P(x, t) + \partial_t P(x, t)\Delta t$, we have:

$$\begin{aligned} \frac{\partial}{\partial t} P(x, t)\Delta t = &- \frac{\partial}{\partial x}[P(x, t)E(\Delta x(\Delta t)|x)] \\ &+ \frac{1}{2}\frac{\partial^2}{\partial x^2}[P(x, t)E(\Delta x^2(\Delta t)|x)] - \cdots \end{aligned} \tag{1.24}$$

Let us now assume that

$$\lim_{\Delta t \to 0} \frac{E(\Delta x|x)}{\Delta t} = a(x), \ \lim_{\Delta t \to 0} \frac{E(\Delta x^2|x)}{\Delta t} = b(x), \tag{1.25}$$

$$\lim_{\Delta t \to 0} \frac{E(\Delta x^n|x)}{\Delta t} = 0 \text{ for } n \geqslant 3. \tag{1.26}$$

The above assumptions are the mathematical relations to express the condition that Δx is small when Δt is small. With the above assumptions we have the Fokker–Planck equation:

$$\frac{\partial}{\partial t} P(x, t) = -\frac{\partial}{\partial x}[a(x)P(x, t)] + \frac{1}{2}\frac{\partial^2}{\partial x^2}[b(x)P(x, t)]. \tag{1.27}$$

Let us now consider the case in N dimensions:

$$\mathbf{x} = (x_1, x_2, \ldots, x_N) \in \mathbb{R}^N. \tag{1.28}$$

Repeating the same argument presented for the one-dimensional case we have the multi-dimensional Fokker–Planck equation:

$$\frac{\partial P(\mathbf{x}, t)}{\partial t} = -\sum_{n=1}^{N} \frac{\partial}{\partial x_n}[a_n(\mathbf{x})P(\mathbf{x}, t)] + \frac{1}{2}\sum_{n,n'} \frac{\partial^2}{\partial x_n \partial x_{n'}}[b_{nn'}(\mathbf{x})P(\mathbf{x}, t)], \tag{1.29}$$

where

$$\begin{aligned} a_n(\mathbf{x}) &= \lim_{\Delta t \to 0} \frac{\langle \Delta x_n|\mathbf{x}\rangle}{\Delta t}, \\ b_{nn'}(\mathbf{x}) &= \lim_{\Delta t \to 0} \frac{\langle \Delta x_n \Delta x_{n'}|\mathbf{x}\rangle}{\Delta t}, \\ \lim_{\Delta t \to 0} \frac{\langle \Delta x_n^k \Delta x_{n'}^{k'}|\mathbf{x}\rangle}{\Delta t} &= 0 \quad \text{if } k + k' > 2. \end{aligned} \tag{1.30}$$

1.2.3 Relation between Fokker–Planck and Langevin equations

The procedure to obtain the Fokker–Planck equation can appear a sort of dirty game, a truncation of the Taylor series performed with an ad hoc assumption. So it is rather natural to wonder whether the conditions (equations 1.25, 1.26) hold for some interesting stochastic processes. Consider the Langevin equation

$$dx = a(x)dt + \sqrt{b(x)}\,dW, \tag{1.31}$$

where W is the so-called Wiener process, introduced above, i.e. a Gaussian process with the following properties (assuming $W(t_0) = w_0$)

$$\langle W(t)\rangle = w_0,\ \langle W(t)W(t')\rangle = \min\{t - t_0, t' - t_0\} + w_0^2,\ \langle [W(t) - w_0]^2\rangle = t - t_0,$$

which imply for the infinitesimal increment $dW = W(t + dt) - W(t)$:

$$\langle dW\rangle = 0,\ \langle dW^2\rangle = dt.$$

It is easy to show the validity of the following relation

$$W(t) = \int_0^t \xi(t')dt', \tag{1.32}$$

where $\xi(t)$ is the so-called white noise, i.e. a Gaussian process with

$$\langle \xi(t)\rangle = 0,\ \langle \xi(t)\xi(t')\rangle = \delta(t - t').$$

In the physics literature often one finds the equation

$$\frac{d}{dt}x = a(x) + \sqrt{b(x)}\,\xi, \tag{1.33}$$

which is formally equivalent to (1.31) once one identifies ξ with dW/dt. On the other hand W is not differentiable, and therefore, strictly speaking, the expression $\xi = dW/dt$ is meaningless; however, following an established tradition, sometimes we shall write the stochastic differential equations in the form (1.33).

Now, for simplicity, we consider only the case of constant b, postponing the discussion of the general case with $b(x)$ to the following. We have

$$\Delta x(\Delta t) = x(t + \Delta t) - x(t) = \int_t^{t+\Delta t} a[x(t')]dt' + \sqrt{b}\int_t^{t+\Delta t} \xi(t')dt'; \tag{1.34}$$

noting that the mean value of the second integral is zero, for small Δt one has

$$E(\Delta x|x) = \int_t^{t+\Delta t} a[x(t')]dt' = a[x(t)]\Delta t \implies \lim_{\Delta t\to 0}\frac{E(\Delta x|x)}{\Delta t} = a(x). \tag{1.35}$$

Repeating the computation for $\Delta x(\Delta t)^2$ we obtain

$$\left(\int_t^{t+\Delta t} a[x(t')]dt'\right)^2 + b\iint_t^{t+\Delta t} \xi(t')\xi(t'')dt'dt'' + 2\sqrt{b}\iint_t^{t+\Delta t} a[x(t')]\xi(t'')dt'dt''.$$

The average of the third integral is zero[3], the average of the second one can be easily computed using the property $\langle \xi(t')\xi(t'')\rangle = \delta(t - t')$, and therefore we have

[3] This is actually one of the most delicate points of the theory. The result of this integral depends upon the particular choice of stochastic integration, as discussed in the following [8]. In the Ito integration scheme the average of a stochastic integral, working only with the so-called 'non-anticipating functions', $\int a(t')dW(t')$ is always zero.

$$E(\Delta x^2|x) = a^2(x)\Delta t^2 + b\Delta t \implies \lim_{\Delta t\to 0}\frac{E(\Delta x^2|x)}{\Delta t} = \lim_{\Delta t\to 0} a^2(x)\Delta t + b = b. \quad (1.36)$$

In a similar way it is easy to verify that

$$\lim_{\Delta t\to 0}\frac{E(\Delta x^n|x)}{\Delta t} = 0 \text{ for } n > 2.$$

Therefore, we have that there exists a stochastic process, the Langevin equation, whose time evolution is given by the Fokker–Planck equation.

1.2.3.1 A mathematical digression

Let us open a brief mathematical parenthesis, for simplicity of notation in the one-dimensional case. We can write the solution of equation (1.31) as

$$x(t) = x(0) + \int_0^t a[x(t')]\,dt' + \int_0^t \sqrt{b(x(t'))}\,dW(t').$$

So we have to treat integrals of the form

$$I(t) = \int_0^t f[x(t')]\,dW(t'),$$

where f depends on W through $x(t)$ which solves the stochastic differential equation where $W(t)$ appears. At variance with the usual integrals, such a problem is rather delicate. Following the procedure of the theory of the Riemann–Stieltjes integral, it seems natural to introduce $\Delta t = t/N$, $x_n = x(n\Delta t)$, $W_n = W(n\Delta t)$ and define $I(t)$ as the limit for $N \to \infty$ of

$$I_I^N = \sum_{n=0}^{N-1} f(x_n)(W_{n+1} - W_n). \quad (1.37)$$

Another possibility is

$$I_S^N = \sum_{n=0}^{N-1} f\left(\frac{x_{n+1} + x_n}{2}\right)(W_{n+1} - W_n). \quad (1.38)$$

The first procedure is called Ito stochastic integral, and the second Stratonovich stochastic integral; actually the two protocols are not equivalent. This can be seen with an explicit example: in the case $f = W$, a simple computation shows that

$$I_S^N = \frac{W_N - W_0}{2} \to \langle I_S^N\rangle = \frac{t}{2},$$

while

$$\langle I_I^N\rangle = 0.$$

This means that when we consider the formula (1.31) in the case where $b(x)$ is not constant we have to specify how we are interpreting the stochastic integral; once this has been specified we have a well-defined mathematical theory, in particular we can

associate with a Langevin equation its correspondent Fokker–Planck equation. Of course there are infinitely many other interpretations of the stochastic integral, since the function $f(x)$ can be computed with other prescriptions. Stratonovich and Ito prescriptions are the most used, since each of them has an important advantage which we recall here, briefly, as a rule of thumb to decide between one and the other. The Stratonovich prescription has the advantage of having the usual differentiation rule $dy(x) = y'(x)dx$ but the disadvantage of having (in general) a complicated result for $\langle f[x(t')]dW(t')\rangle$, since $f(x)$ is computed in x evaluated in the middle of the time interval and therefore could be entangled with the Wiener increment. The Ito prescription, in contrast, has the advantage of always giving $\langle f[x(t')]dW(t')\rangle = 0$, since—in the set of non-anticipating functions, which is sufficient for physical situations— is computed at the beginning of the time interval and is independent of the successive Wiener increment; however, it has the disadvantage of having an unusual differentiation rule, i.e. in general $dy(x) \neq y'(x)dx$. More precisely, in one dimension, if $dx = adt + bdW$, one has

$$\begin{aligned} dy(x,\, t) &= \frac{\partial y}{\partial t}dt + \frac{\partial y}{\partial x}dx + \frac{1}{2}\frac{\partial^2 y}{\partial x^2}dx^2 \\ &= \frac{\partial y}{\partial t}dt + \frac{\partial y}{\partial x}(adt + bdW) + \frac{1}{2}\frac{\partial^2 y}{\partial x^2}b^2dt, \end{aligned} \tag{1.39}$$

where in the last passage we have kept for dx^2 only the term b^2dt, while the others are of order $O(dt^{3/2})$ or higher.

Note that with the Ito prescription one can solve the conundrum of the mean squared velocity mentioned at the beginning of this section. In fact, starting from the correct writing of the Langevin equation, equation (1.3) in differential form

$$dv = -\frac{1}{\tau}vdt + bdW, \tag{1.40}$$

with $\tau = m/(6\pi\eta r)$ and some noise amplitude b which has to be determined, if one calls $y(v) = v^2$, then the Ito prescription says that

$$\langle dy\rangle = \langle 2vdv + dv^2\rangle = \left\langle 2v\left[-\frac{v}{\tau}dt + bdW\right] + b^2dt\right\rangle = \left(b^2 - 2\frac{y}{\tau}\right)dt, \tag{1.41}$$

after using the Ito property $\langle vdW\rangle = 0$. The final equation leads at large times to $\langle v^2\rangle \to \tau b^2/2$ and this immediately sets $b^2 = 2\frac{k_B T}{m\tau}$. In contrast, if the Stratonovich prescription is used, the term dv^2 should not appear, but the term $\langle vdW\rangle$ could not be put to zero: anyway, after more complicate calculations, the result of the two integration schemes would be the same.

A final remark on the Ito–Stratonovich prescription is how to go from one scheme to the other. Let us forget for a moment the deterministic part of the Langevin equation and consider the equation $dx = b(x)dW$. If interpreted according to Stratonovich, it means

$$dx = b\left(\frac{x(t) + x(t+dt)}{2}\right)dW(t) = b\left(x(t) + \frac{dx}{2}\right)dW(t)$$
$$= \left\{b[x(t)] + \frac{1}{2}\frac{\partial b}{\partial x}dx\right\}dW(t) \tag{1.42}$$
$$= b[x(t)]dW(t) + \frac{1}{2}\frac{\partial b}{\partial x}b[x(t)]dt,$$

which is now to be integrated with the Ito scheme, since b is computed in $x(t)$. In the last passage we have again used the fact that $dxdW = b(x)dWdW \approx b(x)dt$. In general, therefore, these two stochastic equations—the first integrated with the Stratonovich prescription, the second with the Ito one,

$$dx_S = adt + bdW, \tag{1.43a}$$

$$dx_I = \left(a + \frac{1}{2}\frac{\partial b}{\partial x}b\right)dt + bdW, \tag{1.43b}$$

have the same solution, i.e. $dx_S = dx_I$. The knowledge of this transformation rule makes the Fokker–Planck equation for the Stratonovich equation easy to derive. We discuss this point in the general N-dimensional case below.

In N dimensions one has

$$dx_n = a_n(\mathbf{x})dt + \sum_j \sigma_{nj}(\mathbf{x})\, dW_j(t), \tag{1.44}$$

where the $\{W_j\}$ are independent Wiener processes, i.e. $\langle W_j\, W_i(t')\rangle = 0$ for $i \neq j$. Using the Ito rule we have the following Fokker–Planck equation

$$\frac{\partial P(\mathbf{x},\, t)}{\partial t} = -\sum_{n=1}^{N}\frac{\partial}{\partial x_n}[a_n(\mathbf{x})P(\mathbf{x},\, t)] + \frac{1}{2}\sum_{n,n'}\frac{\partial^2}{\partial x_n \partial x_{n'}}[D_{nn'}(\mathbf{x})P(\mathbf{x},\, t)] \tag{1.45}$$

where $D_{nn'}(\mathbf{x}) = \sum_k \sigma_{nk}(\mathbf{x})\sigma_{n'k}(\mathbf{x})$.

In order to get the corresponding Stratonovich equation we can exploit the generalization to N dimensions for the Ito–Stratonovich connection shown in equation (1.43):

$$dx_n^S = a_n(\mathbf{x})dt + \sum_j \sigma_{nj}(\mathbf{x})\, dW_j(t), \tag{1.46a}$$

$$dx_n^I = \left[a_n(\mathbf{x}) + \frac{1}{2}\sum_{jk}\frac{\partial b_{nj}}{\partial x_k}b_{kj}\right]dt + \sum_j \sigma_{nj}(\mathbf{x})\, dW_j(t), \tag{1.46b}$$

which leads to the following Fokker–Planck equation for the Stratonovich equation (1.46)

$$\frac{\partial P(\mathbf{x},\, t)}{\partial t} = -\sum_{n=1}^{N}\frac{\partial}{\partial x_n}[a_n(\mathbf{x})P(\mathbf{x},\, t)] + \frac{1}{2}\sum_{k,j,i}\frac{\partial}{\partial x_i}\left[\sigma_{ik}(\mathbf{x})\frac{\partial}{\partial x_j}(\sigma_{jk}(\mathbf{x})P(\mathbf{x},\, t))\right]. \tag{1.47}$$

Let us note that in the case of constant σ_{ik} we obtain the same Fokker–Planck equation.

We remark that when equation (1.44) is a model of a physical system, the integration prescription (e.g. Ito or Stratonovich) is part of the model, i.e. it cannot be chosen independently, otherwise one would have the paradox of a model that describes a single physical system but predicts infinite different behaviours. In contrast, a Fokker–Planck equation is unambiguous and therefore can be the description of a physical system. Once a Fokker–Planck is understood as a correct model for the system, then for the sake of calculations, it can be useful to put it in the corresponding Langevin form, and the exact form depends upon the choice of the integration prescription.

1.2.4 Beyond the Fokker–Planck equation

We can now wonder why to take into account only two terms in the Taylor expansion, i.e. why equation (1.26) holds. Actually one can consider the Kramers–Moyal expansion [8], containing all the terms:

$$\frac{\partial}{\partial t}P(x,\,t) = \sum_{n=1}^{\infty}\frac{\partial^n}{\partial x^n}[D_n(x)P(x,\,t)],$$

where

$$D_n(x) = \frac{(-1)^n}{n!}\lim_{\Delta t\to 0}\frac{E(\Delta x^n|x)}{\Delta t}.$$

The Fokker–Planck equation is recovered in the case where $D_n(x) = 0$ for $n \geqslant 3$. In particular, one can show a theorem due to Pawula [9] which states that either the sequence becomes zero at the third term, or all its even terms are positive. In physical situations, the possibility to truncate the expansion to $n = 2$ is related to the existence of a small parameter that guarantees the rapid convergence of the series and the fact that the sum of the terms for $n \geqslant 3$ is really negligible. A fundamental example is the so-called van Kampen size expansion which takes as small parameter the ratio between the size (or mass) of the molecules and the size (mass) of the Brownian particle [7, 10]. It becomes reasonable, then, that in the limit of a very large (massive) particle, its dynamics is described by observables which are continuous in time and therefore obey a stochastic differential equation. This programme is described in section 6.3.2.

More in general, if one considers non-Gaussian noises, instead of the Fokker–Planck equation we have that the evolution of the probability distribution contains also an integral part. Let us discuss for notation simplicity a one-dimensional example

$$\frac{dx}{dt} = a(x)dt + f(t),$$

where now the noise f is

$$f(t) = \sqrt{2D_2}\,\xi(t) + \zeta(t),$$

where ξ is a Gaussian white noise, and $\zeta(t)$ a Poissonian noise

$$\zeta(t) = \sum_j z_j \delta(t - t_j).$$

The z_j are independent random variables distributed according to a symmetric $\rho(z)$, while the intervals $(t_j - t_{j-1})$ are distributed according to a Poissonian distribution with parameter λ. In such a stochastic process the evolution equation for the probability density is

$$\frac{\partial}{\partial t}P(x, t) = \sum_{n=1}^{2} \frac{\partial^n}{\partial x^n}[D_n(x)P(x, t)] + \lambda \int (P(x - z, t) - P(x, t))\rho(z)\, dz.$$

In the above equation the Fokker–Planck part corresponds to the contribution of the Gaussian part of the noise, while the integral part includes the contribution of the Kramers–Moyal expansion for $n > 2$ and is given by the non-Gaussian noise terms. The integral part shown above is a particular case of the integral that appears in the so-called master equations, describing discontinuous Markov processes (e.g. in a discrete space or in continuous space but with *jumps* [8]), and which are discussed in detail in section 2.1.1.

1.3 Again about the Langevin equation

1.3.1 The Langevin equation from the microscopic dynamics

The Langevin equation for the dynamics of the colloidal particle originally had been introduced in a phenomenological way, but a derivation of Brownian motion from the first principles is an interesting general problem. In chapter 6 we shall discuss this formidable task, which can be accomplished in a rigorous way only in few cases. Here, we briefly discuss the general idea of the procedure to derive the Langevin equation from molecular dynamics.

Consider a system of colloidal particles suspended in a liquid (or in a gas). At the microscopic level we have the canonical coordinates $(\mathbf{Q}_i, \mathbf{P}_i)$ and $(\mathbf{q}_n, \mathbf{p}_n)$ of colloidal particles and solvent molecules, respectively; m is the mass of a solvent molecule, M is the mass of a colloidal particle, and we assume $M \gg m$.

The evolution of such a system is ruled by the Hamiltonian equations; if we are interested in the colloidal subsystem alone, we must remove somehow the degrees of freedom of the solvent particles. Since, in comparison with the solvent molecules, the colloidal particles have a much larger mass, they have a much slower evolution. As a consequence of this timescale separation between the two subsystems, and because of the huge number of the solvent particles, we can conjecture that the fast solvent dynamics can be consistently decoupled from the slow colloid dynamics, by approximating its effects on the big suspended particles by means of an effective force. This latter may be decomposed into a systematic part, of viscous type, and a truly stochastic fluctuating part, described by white noise. In such a limit we recover a stochastic differential equation for the colloidal subsystem.

1.3.2 Langevin equation as a tool for numerical computations

If we are interested in a numerical study of $P(\mathbf{x}, t)$ we can follow two paths. The first one is rather obvious: one can study the Fokker–Planck equation using some numerical methods for partial differential equations. Another possibility is to study the Langevin equation looking at the time evolution of a large number of trajectories and then determine $P(\mathbf{x}, t)$. The procedure is the following:

1. at $t = 0$ we consider $\mathcal{N} \gg 1$ trajectories with initial conditions
$$\mathbf{x}^{(1)}(0), \cdots, \mathbf{x}^{(\mathcal{N})}(0),$$
distributed according to some initial condition $P(\mathbf{x}, 0)$;
2. we follow each trajectory which evolves according to the Langevin equation, where for each particle one has to use independent noisy terms;
3. from the $\mathbf{x}^{(1)}(t),\ldots, \mathbf{x}^{(\mathcal{N})}(t)$ we obtain a proxy of $P(\mathbf{x}, t)$, which approaches the correct value as $\mathcal{N} \to \infty$.

The above procedure has several advantages: at variance with the numerical treatment of the Fokker–Planck equation which can be rather difficult, with several troubles such as numerical instabilities, the approach based on the Langevin equation is quite easy, because it is enough to have a method to solve numerically the Langevin equation. There exist efficient algorithms, based on the generalization to the stochastic case of the usual approach: time is discretized and then one obtains $\mathbf{x}(t + \Delta t)$ starting from $\mathbf{x}(t)$, using a stochastic version of the Euler or Runge–Kutta method. It is easy to see that the Lorentz model, equation (1.14) for small Δt, can be interpreted as a numerical algorithm for the Langevin equation for the velocity.

In order to clarify the difference between the Fokker–Planck or the Langevin approach to the study of a stochastic process, we can mention the problem of the spreading of a pollutant in a steady velocity field $\mathbf{u}(\mathbf{x})$. The pollutant concentration $C(\mathbf{x}, t)$ evolves according to

$$\partial_t C(\mathbf{x}, t) = -\nabla \cdot (\mathbf{u}(\mathbf{x})C(\mathbf{x}, t)) + D\Delta C(\mathbf{x}, t),$$

where Δ denotes the Laplacian, which is nothing but the Fokker–Planck equation associated with the equation for the evolution of a test particle advected by the velocity field and molecular noise:

$$\frac{d}{dt}\mathbf{x} = \mathbf{u}(\mathbf{x}) + \sqrt{2D}\,\xi.$$

For this problem the approach in terms of Fokker–Planck equation corresponds to the Eulerian point of view, while the Langevin equation is the analogous of the Lagrangian description.

1.4 Further remarks

We conclude this chapter with some digressions which we consider useful for students as well as for curious scientists.

1.4.1 A pedagogical parenthesis on the Lorentz model

Let us consider again the Lorentz model (1.14), where now w_n are independent variables distributed according to a generic $P_w(w)$ which is regular enough, i.e. its cumulants exist. We can show that starting from an arbitrary probability distribution for v_0, $P_v(v;0)$, then $P_v(v;n) \to P^{(\infty)}(v)$, where $P^{(\infty)}(v)$ can be determined in terms of $P_w(w)$; in particular $P^{(\infty)}(v)$ is Gaussian if and only if $P_w(w)$ is Gaussian.

Let us recall the definition of the characteristic function of a density probability

$$\Phi_x(t) = \langle e^{itx} \rangle = \sum_{k=0}^{\infty} \frac{(it)^k}{k!} \langle x^k \rangle, \tag{1.48}$$

with

$$\ln \Phi_x(t) = \sum_{k=1}^{\infty} \frac{(it)^k}{k!} c_k \tag{1.49}$$

where $c_1, c_2, \ldots$ are the cumulants of the distribution; we recall that the first cumulants can be written in terms of the moments as:

$$c_1 = \langle x \rangle, \; c_2 = \langle x^2 \rangle - \langle x \rangle^2, \; c_3 = \langle (x - \langle x \rangle)^3 \rangle, \; c_4 = \langle (x - \langle x \rangle)^4 \rangle - 3c_2^2.$$

For a Gaussian distribution the cumulants c_k with $k > 2$ are zero.

Now let us introduce

$$\Phi_{v_n}(t) = \langle e^{iv_n t} \rangle, \; \Phi_{w_n}(t) = \langle e^{iw_n t} \rangle.$$

Reminding that v_n and w_n are independent, for $v_{n+1} = av_n + bw_n$ we have

$$\Phi_{v_{n+1}}(t) = \langle e^{iv_{n+1}t} \rangle = \langle e^{iav_n t} e^{ibw_n t} \rangle = \langle e^{iav_n t} \rangle \langle e^{ibw_n t} \rangle = \Phi_{v_n}(at) \Phi_{w_n}(bt), \tag{1.50}$$

and taking the logarithm

$$\ln \Phi_{v_{n+1}}(t) = \ln \Phi_{v_n}(at) + \ln \Phi_{bw_n}(t). \tag{1.51}$$

Denoting by $C_k^{(n)}$ and c_k the k-th cumulant of v_n and w_n, respectively, we have

$$\sum_{k=1}^{\infty} \frac{(it)^k}{k!} C_k^{(n+1)} = \sum_{k=1}^{\infty} \frac{a^k (it)^k}{k!} C_k^{(n)} + \sum_{k=1}^{\infty} \frac{b^k (it)^k}{k!} c_k. \tag{1.52}$$

This allows us to find a recursive rule for each k:

$$C_k^{(n+1)} = a^k C_k^{(n)} + b^k c_k. \tag{1.53}$$

Repeating the analysis previously done for σ_n in section 1.2.1, we have

$$C_k^{\infty} = \frac{b^k c_k}{1 - a^k}. \tag{1.54}$$

Therefore, if $P_w(w)$ is Gaussian we have

$$C_1^{(n)} \to \frac{b\langle w \rangle}{1 - a} \qquad C_2^{(n)} \to \frac{b^2 \sigma_w^2}{1 - a^2} \qquad C_{k \geqslant 3}^{(n)} \to 0,$$

and therefore $P^{(\infty)}(v)$ must be Gaussian. In the case where $P_w(w)$ is not Gaussian, we have $P(v;\, n) \to P^{(\infty)}(v)$ which is not Gaussian in general, but, using (1.54), it can be determined by $P_w(w)$. In physical models a and b should scale with Δt and therefore one may have cases where a Gaussian limit distribution is recovered. For instance, the case considered in section 1.2.1, where $a = 1 - \Delta t/\tau$ and $b \sim \Delta t^{1/2}$ (assuming $\langle w \rangle = 0$), corresponds to $C_k^\infty \to 0$ for $\Delta t \to 0$ when $k > 2$, that yields a Gaussian distribution for $P^{(\infty)}(v)$. A general discussion of the continuous-time limit of discrete stochastic processes can be found in [8]. A particularly interesting case of this problem is analysed in section 6.3.2.

1.4.2 Einstein's approach to Brownian motion

Let us briefly recall the original idea of Einstein who based his theory on:

- the Stokes law for the friction force acting on a sphere of radius r with velocity v in a fluid with viscosity η; only for notation simplicity we consider the one-dimensional case: $F_S = -6\pi r\eta v$;
- the Van't Hoff law for the osmotic pressure.

As already mentioned, the simple and ingenious idea was to assume that the statistical mechanics holds even for macroscopic objects whose size is on the order of a few microns. Let us consider particles of mass M in a liquid in the presence of a constant gravity field along the z-axis. From the barotropic formula (as well as the Van't Hoff law), we have:

$$\rho(z) = \rho(0)e^{-Mgz/k_{\mathrm{B}}T}, \tag{1.55}$$

where ρ is the density. Each particle feels the force

$$F_S - Mg,$$

which is zero for

$$v = v^* = -\frac{Mg}{6\pi\eta r}.$$

At the height z we have a down forward current density

$$j_-(z) = v^*\rho(z) = -\frac{Mg}{6\pi\eta r}\rho(z). \tag{1.56}$$

Since the density is not constant we have an up forward current density $j_+(z)$ due to the motion of particles from a region of larger density to a region of lower density:

$$j_+(z) = -D\frac{\partial}{\partial z}\rho(z,\, t).$$

From equation (1.55) one has

$$j_+(z) = D\frac{Mg}{k_{\mathrm{B}}T}\rho(z,\, t), \tag{1.57}$$

and in the case of stationarity one has a balance of the two fluxes, i.e. $j_+ + j_- = 0$. Therefore, using equations (1.56) and (1.57) one has

$$D = \frac{k_B T}{6\pi\eta r}.$$

It is straightforward to check that the result does not change if instead of the gravity force $-Mg$ one considers colloidal particles in a generic potential $U(z)$ and therefore a force $f(z) = -dU(z)/dz$.

1.4.3 Remarks on Perrin's experiments

Jean Baptiste Perrin was a sophisticated experimental physicist, and had the great merit to verify in a very clever way the theoretical prediction of Einstein, and determine the value of the Avogadro number from the diffusion coefficient of the Brownian motion. In his laboratory he was able to overcome enormous practical difficulties, in particular, at that time it is was not so easy to have little spheres of a given mass and/or radius. Studying, with the help of a microscope, the statistical features of grains with diameters in the range [0.2–5.5] microns, after a long series of experiments he was able to verify the relation (1.1), and was awarded the Nobel Prize in Physics in 1926, the official motivation being *'for his work on the discontinuous structure of matter'*.

Once one assumes, according to Einstein, that the same statistical laws are valid for molecules as well as for colloidal particles, one has a powerful method to determine the Avogadro number. For instance Perrin verified the validity of the barotropic formula for grains of size of the order of micron and mass M (order 10^{-12}–10^{-11} g) in water at temperature T, studying the density of the grains at varying height with a microscope. In the case of high dilution, the interaction among the colloidal particles can be neglected and we can assume the barotropic formula for perfect gases in an external field, equation (1.55). Let us note that the value of the quantity $\zeta = k_B T/(Mg)$ decreases with the mass M; for colloidal particles ζ is small enough so that the behaviour of $\rho(z)$ can be observed in the laboratory. The experimental results of Perrin gave further evidence of Einstein's idea about the validity of the statistical mechanics of macroscopic objects, as well as another way to determine the Avogadro number from the value of ζ.

1.4.4 The legacy of Brownian motion

Brownian motion is not only an important historical and technical aspect of modern physics, but it has been also the starting point for the development of the mathematical theory of stochastic processes; in addition, Brownian motion has played a role in the recent progress in biophysics [11] and finance [12].

The Langevin equation had been the first non-trivial example of stochastic process. Such a mathematical field, now largely used in physics, chemistry, biology and applied sciences, was developed in a systematic way in the 30s of the last century, mainly by Kolmogorov with the formalization of the Fokker–Planck and master equations for the continuous-time Markov processes [13]. A seminal work,

which started from the mathematical description of Brownian motion, is due to Wiener who introduced the idea of path integral which plays an essential role for the Feynman formulation of quantum mechanics [14].

A few years before the paper of Einstein, the French mathematician Louis Jean-Baptiste Alphonse Bachelier in his doctoral dissertation *Théorie de la Spéculation* (1900) proposed that the price of a stock behaves as a particle performing Brownian motion. Recently, the stochastic processes, in particular the differential stochastic equations, have received vast attention from the financial community. The celebrated Black and Scholes theory [15] for the option pricing is, from a mathematical point of view, nothing but an application of the Langevin equation. It is interesting to note that, despite their completely different origins, the phenomenon of Brownian motion and the option pricing in finance are described by the same mathematical formalism.

At first glance the explanation of Brownian motion from kinetic theory can appear a perfect example of the success of reductionism: an explanation of complicated visible phenomena in terms of simple invisible elements. Actually, this point of view cannot be considered correct: Brownian motion theory was derived by Einstein without him being really familiar with the details of the experiments that had been performed for several decades, and the real interest of Einstein was not to describe Brownian motion from the detailed knowledge of the motion of atoms. He started from the bold assumption that particles, even if of macroscopic size, had to obey the laws of statistical mechanics, and then he found an expression for the diffusion coefficient which can be compared with the experiments. In the Einstein approach to Brownian motion, we do not find a path from the microscopic to the macroscopic world; on the contrary, different levels of description were mixed up several times. The force acting on pollen grains was assumed to be made up of two contributions: one of macroscopic character and one of microscopic origin; in particular, the macroscopic term was given by the Stokes law, assumed as a phenomenological result and not derived from the microscopic realm.

It is interesting to note that, sometimes, Brownian motion, one of the major chapters of modern physics, has been so poorly understood, and even misconstrued as incompatible with the second law of thermodynamics, by leading philosophers of science, for instance Popper [16] and Feyerabend [17], who stated that Brownian motion is a serious problem for the second law.

We conclude with a few words about present times: today modern technologies allow scientists easy access to the study of thermal fluctuations even in small systems such as colloidal suspensions. This kind of study is rather important in nanoscience, and fluctuations can have a crucial role in noise-assisted transport mechanisms, also called Brownian motors [18] (see chapter 8). As an example, molecular motors in biophysics use chemical energy to lock in the fluctuations due to thermal noise, and move the molecular machine forward.

Therefore, we can say that the paper of Einstein on Brownian motion, which sometimes had been considered the less relevant among those of the *annus mirabilis*, with its huge impact on the physics of colloidal particles, other forms of soft matter, and of biophysical systems, had its revenge on the subatomic world.

References

[1] Robert B 1827 A brief account of microscopical observations *Phil. Mag.* **4** 161

[2] Cantoni G 1867 Su alcune condizioni fisiche dell'affinità e sul moto Browniano *Il Nuovo Cimento (1855–1868)* **27** 156–67

[3] Einstein A 1905 Über die von der molekularkinetischen theorie der wärme geforderte bewegung von in ruhenden flüssigkeiten suspendierten teilchen *Ann. Phys., Lpz.* **4** 549–60

[4] Langevin P 1908 Sur la théorie du mouvement Brownien *Compt. Rendus* **146** 530–3

[5] Perrin J 1909 Mouvement Brownien et réalité moléculaire *Ann. Chim. Phys.* **18** 1–114

[6] Pauli W 1973 *Pauli Lectures on Physics, Volume 4: Statistical Mechanics* (Cambridge: Cambridge University Press)

[7] van Kampen N G 1992 *Stochastic Processes in Physics and Chemistry* (Amsterdam: Elsevier)

[8] Gardiner C 2009 *Stochastic Methods* (Berlin: Springer)

[9] Risken H 1996 *Fokker-Planck equation* (Berlin: Springer)

[10] van Kampen N G 1961 A power series expansion of the master equation *Can. J. Phys.* **39** 551–67

[11] Frey E and Kroy K 2005 Brownian motion: a paradigm of soft matter and biological physics *Ann. Phys., Lpz.* **517** 20–50

[12] Karatzas I, Shreve S E, Karatzas I and Shreve S E 1998 *Methods of Mathematical Finance* (Berlin: Springer)

[13] Kolmogoroff A 1931 Über die analytischen methoden in der wahrscheinlichkeitsrechnung *Math. Ann.* **104** 415–58

[14] Wiener N 1976 *Collected Works* **vol 1** (Cambridge, MA: MIT Press)

[15] Black F and Scholes M 1974 From theory to a new financial product *J. Finance* **29** 399–412

[16] Popper K R 1957 Irreversibility; or, entropy since 1905 *Br. J. Phil. Sci.* **8** 151–5

[17] Feyerabend P K 2010 *Against Method* (London: Verso Books)

[18] Reimann P 2002 Brownian motors: noisy transport far from equilibrium *Phys. Rep.* **361** 57–265