1 Brownian Motion, Langevin and Fokker–Planck Equations

In this chapter the reader can find the basic ingredients of elementary kinetic theory and of the mathematical approach to discrete and continuous stochastic processes, all that is necessary to establish a solid ground for nonequilibrium processes concerning the time evolution of physical systems subject to a statistical description. In fact, from the first sections we discuss problems where we deal with the time evolution of average quantities, such as in the elementary random walk model of diffusion. We also illustrate the bases of transport phenomena that allow us to introduce the concept of transport coefficients, which will be reconsidered later in the framework of a more general theory (see Chapter 2). Then we focus on the theory of Brownian motion, as it was originally formulated by Albert Einstein, and how this was later described in terms of the Langevin and of the Fokker–Planck equations, specialized to a Brownian particle. The peculiar new ingredient that was first introduced in the Langevin formulation is noise, which epitomizes the effect of incoherent fluctuations of the Brownian particle due to the interaction with the solvent particles, which are subject to a thermal motion. Averaging over thermal noise allows one to obtain a statistical inference on the diffusive behavior of a Brownian particle. The Fokker–Planck formulation tells us that we can obtain an equivalent description by considering the evolution of the space-time probability function of a Brownian particle, rather than averaging over its noisy trajectories.

The mathematical formulation of stochastic processes in discrete space and time (Markov chains) is illustrated, together with many examples and applications, including random walk processes and the Monte Carlo procedure. This important mathematical theory provides us with the tools for a general formulation of stochastic processes in continuous space and time. This is not at all a straightforward step, since the presence of noise needs a suitable mathematical procedure, when passing to a continous time description. In particular, we have to establish a consistent relation between infinitesimal time and noise increments, which allows for two possible different formulations of continuous time Langevin–like equations. In this general framework we can derive also the Fokker–Planck equation for general stochastic processes, rather than for the mere description of the diffusive motion of the Brownian particle. We discuss some interesting applications of this equation to point out the physical importance of this general formulation of stochastic processes and, specifically, its relevance for nonequilibrium processes. In the last part of this chapter we introduce a description of those stochastic processes that do not exhibit a standard diffusive behavior. More precisely, we discuss the so-called continuous time random walk model and we focus our considerations on processes named Lévy flights and Lévy walks, which play an increasing importance in the modern applications of nonequilibrium processes.

https://doi.org/10.1017/9781107278974.002 Published online by Cambridge University Press

1.1 Introduction

The idea that thermodynamics could be related to a mechanical theory of matter dealing with a large number of particles, i.e., atoms and molecules, was speculated on from the very beginning of kinetic theory on the middle of the nineteenth century. In a historical perspective, we could say that such an idea was a natural consequence of the formulation of the first principle of thermodynamics by the German natural philosopher Julius Robert von Mayer, establishing the equivalence between mechanical work and heat. This was checked experimentally in the famous experiment by James Prescott Joule and many contemporary physicists, among which Rudolf Clausius, August Karl Krönig, William Thomson (Lord Kelvin), James Clerk Maxwell, and Ludwig Eduard Boltzmann devoted a good deal of their efforts to develop the foundations of kinetic theory.

The reader should consider that these scientists were assuming the validity of the atomic hypothesis, despite no direct experimental evidence of the existence of atoms and molecules available at that time. Accordingly, the reader should not be surprised that such a nowadays "obvious" concept was strongly opposed by a large part of the scientific community in the last decades of the nineteenth century, as a reaction to a mechanistic foundation of science that, on the one hand, supported a materialistic and, apparently, deterministic basis of natural phenomena and, on the other hand, raised serious conceptual paradoxes, most of which related to the time reversibility of mechanical laws. In fact, the other cornerstone of thermodynamics is the second principle, which amounts to establishing the irreversibility of thermodynamic processes, due to the natural tendency of thermodynamic systems to evolve toward a well-defined equilibrium state in the absence of energy supplied by some external source.

The mechanistic approach to thermodynamics was pushed to its extreme consequences in the work by Ludwig Eduard Boltzmann. His celebrated transport equation represents a breakthrough in modern science and still today we cannot avoid expressing our astonishment about the originality and deep physical intuition of the Austrian physicist. Despite being inspired by a specific model, namely the ideal gas, the main novelty of Boltzmann's equation was that it represents the evolution of a distribution function, rather than the trajectories of individual particles in the gas. Boltzmann realized quite soon that the only way to describe the behaviour of a large number of particles (a mole of a gas contains an Avogadro number of particles, approximately equal to $N_A \simeq 6.022 \times 10^{23}$) was to rely on a statistical approach, where the laws of probability had to be merged into the description of physical laws. We want to point out that the success of Boltzmann's equation is not limited to establishing the foundations of equilibrium statistical mechanics. In fact, it also provides a description of the evolution toward equilibrium by the derivation of hydrodynamic equations associated with the conservation of mechanical quantities, i.e., number, momentum, and energy of particles. They are found to correspond to the continuity equation and to two more phenomenological equations, i.e., the Navier–Stokes and the heat ones. These equations provide a mathematical basis for the theory of transport phenomena and a physical definition of transport coefficients in terms of basic quantities of kinetic theory, such as the mean free path, the average speed of particles, the heat capacity, etc.

https://doi.org/10.1017/9781107278974.002 Published online by Cambridge University Press

On top of that, since 1827 the experiment performed by English botanist Robert Brown describing the phenomenon known as Brownian motion challenged the scientific community. In fact, a pollen particle suspended in water (or any similar solvent) was found to exhibit an erratic motion that, apparently, could not be reconciled with any standard mechanical description. Even assuming the atomistic hypothesis and modeling the motion of the pollen particle as a result of collisions with the atoms of the solvent seemed to fail to provide a convincing explanation. In fact, at the microscopic level one might argue that elastic collisions with the atoms of the solvent could transmit a ballistic motion to the pollen particle. However, the conclusion would be that the combined effect of all of these collisions, occurring for symmetry reasons in any direction, vanishes to zero. On the contrary, the experimental observation of the erratic motion of the pollen particle indicated that the distance of the particle from its original position grew over sufficiently long time intervals as the square root of time, thus showing the diffusive nature of its motion. Repeating many times the same experiment, where the pollen particle, the solvent, and the temperature of the solvent are the same, the particle in each realization follows different paths, but one can perform a statistical average over these realizations that enforces the conclusion that the particle exhibits a diffusive motion.

The universal character of this phenomenon was confirmed by the experimental observations that a diffusive behavior was found also when the type of Brownian particle, the solvent, and the temperature were changed, yielding different values of the proportionality constant between time and the average squared distance of the particle from its initial position. A convincing explanation of Brownian motion had to wait for the fundamental contribution of Albert Einstein, which appeared in 1905, the same year as his contributions on the theories of special relativity and the photoelectric effect. Einstein's phenomenological theory of Brownian motion, relying on simple physical principles, inspired the French scientist Paul Langevin, who proposed a mechanistic approach. The basic idea was to write a Newton-like ordinary differential equation where, for the first time, a force was attributed a stochastic nature. In fact, the microscopic forces exerted by the solvent particles through elastic collisions with the Brownian particle are represented as uncorrelated fluctuations in space and time, whose square amplitude is assumed to be proportional to the thermal energy; according to kinetic theory, this amounts to the solvent temperature T, provided the Brownian particle is at thermodynamic equilibrium with the solvent. Some years later Adriaan Daniël Fokker and Max Planck proposed an alternative formulation of the Brownian particle problem, based on a partial differential equation, describing the evolution of the probability distribution of finding a Brownian particle at position $\mathbf{x}$ at time t, in the same spirit of Boltzmann's equation for an ideal gas. In fact, the Fokker–Planck equation was derived as a master equation, where the rate of change in time of the distribution function depends on favorable and unfavorable processes, described in terms of transition rates between different space-time configurations of the Brownian particle. Making use of some simplifying assumptions, this equation was cast into a form where the diffusive nature of the problem emerges naturally, while it allows one to obtain an explicit solution of the problem.

On the side of mathematics, at the end of the nineteenth century the Russian Andrej Andreevič Markov developed a new mathematical theory concerning stochastic processes,

https://doi.org/10.1017/9781107278974.002 Published online by Cambridge University Press

nowadays known as Markov chains. The original theory takes advantage of some simplifying assumptions, like the discreteness of space and time variables as well as of the numerability of the possible states visited by the stochastic process. It was the first time a dynamical theory was assumed to depend on random uncorrelated events, typically obeying the laws of probability. Despite the scientific motivations of Markov, which were quite different from those that moved the above-mentioned physicists to tackle the problem of Brownian motion, some decades later a more general theory of stochastic processes in continuous space and time emerged from the fruitful combination of these different scientific pathways. This allowed the scientific community to unveil the great potential contained in this theory, which could be applied to a wide spectrum of mathematical and physical problems concerning the evolution in time of statistical systems and thus providing the conceptual foundations of nonequilibrium statistical mechanics. Typical modern aspects of this field of physics are contained in the theory of continuous time random walk, discussed at the end of this chapter. It provides the mathematical tools for describing a wide range of stochastic processes, which overtake the limits of standard diffusive behavior, allowing for subdiffusive and superdiffusive regimes. These have been recently recognized as almost ubiquitous in nature, since they have been found to characterize a wide range of phenomena of interest not only for physics, but also for biology, chemistry, geology, finance, sociology, etc. All the following chapters will be devoted to an illustration of the many aspects concerning nonequilibrium statistical mechanics and their relevance for physical science. The introductory and pedagogical character of this book cannot allow us to account for the interdisciplinary potential of this approach, which overtakes, by far, any other domain of modern physics.

1.2 Kinetic Theory

1.2.1 The Ideal Gas

The basic model for understanding the mechanical foundations of thermodynamics is the ideal gas of Boltzmann. It is a collection of N identical particles of mass m that can be represented geometrically as tiny homogeneous spheres of radius r. One basic assumption of the ideal gas model is that we are dealing with a diluted system; i.e., the average distance δ between particles is much larger than their radius,

$$\delta = \left(\frac{1}{n}\right)^{\frac{1}{3}} \gg r, \tag{1.1}$$

where $n = N/V$ is the density of particles in the volume V occupied by the gas.[1] In the absence of external forces particles move with constant velocity[2] until they collide

[1] For a real gas of hydrogen molecules at room temperature (300 K) and atmospheric pressure (1 atm), $\delta \sim 10^{-6}$ m and $r \sim 10^{-10}$ m.

[2] One could argue that at least gravity should be taken into account, but its effects are generally negligible in standard conditions. An example where gravity has relevant, measurable effects will be studied in Section 1.4: it is the Brownian motion of colloidal particles; see Fig. 1.6.

https://doi.org/10.1017/9781107278974.002 Published online by Cambridge University Press

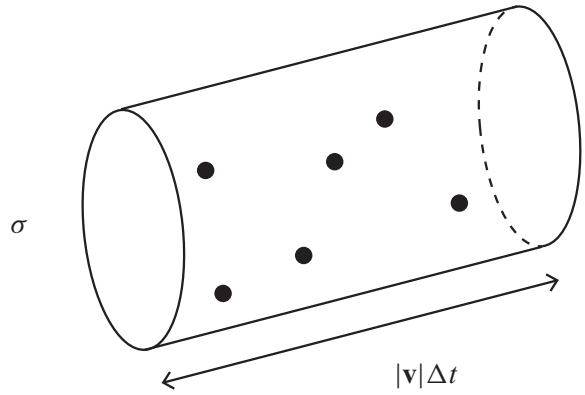

Fig. 1.1 Illustration of the concept of cross section. The black dots in the cylinder spanned by the cross section σ represent the centers of molecules hit in the time interval Δt by a molecule moving at speed $\mathbf{v}$.

pairwise, keeping their total momentum and energy constant (elastic collisions[3]). It can be easily realized that in such a diluted system multiple collisions are such rare events that they can be neglected for practical purposes.

Now we want to answer the following question: what is the rate of these collisions and the average distance run by a particle between subsequent collisions? We can estimate these quantities by considering that a particle moving with velocity $\mathbf{v}$ in a time interval Δt can collide with the particles that are contained in a cylinder of basis $\sigma = 4\pi r^2$ (called cross section) and height $|\mathbf{v}|\Delta t$; see Fig. 1.1. For the sake of simplicity we can assume that all the particles inside the cylinder are at rest with respect to the moving particle, so that we can estimate the number of collisions as

$$\mathcal{N}_{coll} = n\sigma|\mathbf{v}|\Delta t. \tag{1.2}$$

Accordingly, the number of collisions per unit time is given by the expression

$$\frac{\mathcal{N}_{coll}}{\Delta t} = n\sigma|\mathbf{v}| \tag{1.3}$$

and the average time between collisions reads

$$\tau \equiv \frac{\Delta t}{\mathcal{N}_{coll}} = \frac{1}{n\sigma|\mathbf{v}|}. \tag{1.4}$$

A quantitative estimate of τ can be obtained by attributing to $|\mathbf{v}|$ the value $\langle v\rangle$ of the equilibrium average of the modulus of the velocity of particles, v, in the ideal gas, according to Maxwell's distribution (see Fig. 1.2),

$$P(v) = \frac{4}{\sqrt{\pi}}\left(\frac{m}{2T}\right)^{3/2} v^2 \exp\left(-\frac{mv^2}{2T}\right), \tag{1.5}$$

[3] This hypothesis amounts to assuming that the particles of the gas are rigid spheres, that they do not suffer any deformation in the collision process. In fact, in a real gas the energy transferred to the internal degrees of freedom of the molecules can be practically neglected in standard conditions.

https://doi.org/10.1017/9781107278974.002 Published online by Cambridge University Press

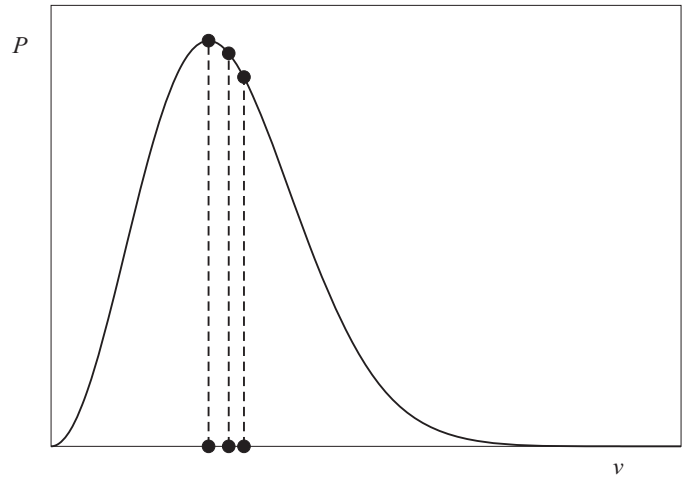

Fig. 1.2 The Maxwell distribution, Eq. (1.5). We indicate, from left to right, the most likely velocity v_{max}, the average velocity $\langle v \rangle$, and the square root of the average square velocity, $\langle v^2 \rangle^{1/2}$, whose expressions are given in Eq. (1.6).

where T is the temperature of the ideal gas at equilibrium. Using such distribution, we obtain the expressions

$$v_{\max} = \sqrt{\frac{2T}{m}}, \qquad \langle v \rangle = \sqrt{\frac{8T}{\pi m}} = \frac{2}{\sqrt{\pi}} v_{\max}, \qquad \langle v^2 \rangle^{1/2} = \sqrt{\frac{3T}{m}} = \sqrt{\frac{3}{2}} v_{\max}, \tag{1.6}$$

for the most likely velocity, the average velocity and the square root of the average square velocity, respectively.

We can now rewrite (1.4) as

$$\tau = \frac{1}{n\sigma \langle v \rangle} \tag{1.7}$$

and determine the average distance run by a particle between two collisions, i.e., its mean free path, by the expression

$$\lambda = \langle v \rangle \tau = \frac{1}{n\sigma}. \tag{1.8}$$

This formula corresponds to the case of a single moving particle colliding with target particles that are supposed to be immobile. But this is not the case, because in reality the target particles also move and a better estimate of τ and λ can be obtained using the formula

$$\tau = \frac{1}{n\sigma \langle v_r \rangle}, \tag{1.9}$$

where v_r is the modulus of the relative velocity v_r, which follows the distribution

$$P_r(v_r) = \sqrt{\frac{2}{\pi}} \left(\frac{m}{2T} \right)^{3/2} v_r^2 \exp\left(-\frac{m v_r^2}{4T} \right). \tag{1.10}$$

This formula is a consequence of the general observation that the sum (or the difference) of two Gaussian variables is a Gaussian variable whose variance is the sum of their variances.

https://doi.org/10.1017/9781107278974.002 Published online by Cambridge University Press

In this case, $\mathbf{v}_r = \mathbf{v}_1 - \mathbf{v}_2$, with $\mathbf{v}_{1,2}$ satisfying the Maxwell distribution (1.5) and the doubling of the variance explains why the exponent $(mv^2/2T)$ in Eq. (1.5) now becomes $(mv_r^2/4T)$. Then, the prefactor changes accordingly, in order to keep $P_r(v_r)$ normalized.

With Eq. (1.10) at hand, we can evaluate

$$\langle v_r \rangle = \sqrt{\frac{16T}{\pi m}} = \sqrt{2}\langle v \rangle \tag{1.11}$$

and obtain

$$\tau = \frac{1}{\sqrt{2} n \sigma \langle v \rangle}, \tag{1.12}$$

from which we can evaluate the mean free path,

$$\lambda = \langle v \rangle \tau = \frac{1}{\sqrt{2} n \sigma}. \tag{1.13}$$

It is worth noting that the ratio between λ and τ gives $\langle v \rangle$, not $\langle v_r \rangle$, because one particle travels an average distance λ in time τ.

We can finally use the formula (1.13) to evaluate the mean free path for a gas at room temperature and pressure. In this case λ is typically $O(10^{-7}\text{m})$, which is three orders of magnitude larger than the typical size r of a particle, $O(10^{-10}\text{m})$.

1.2.2 Random Walk: A Basic Model of Diffusion

We consider an ideal gas at thermal equilibrium with a heat bath at temperature T. If we fix our attention on one particle, we observe that collisions with the other particles produce a stepwise irregular trajectory, i.e., a sort of random walk. Beyond this qualitative observation we would like to obtain a quantitative description of this random walk. In principle the problem could be tackled by applying the laws of classical mechanics. In practice such a program is unrealistic, because one should know not only the initial velocity of the particle under examination, but also the velocities of all the particles that it will collide with. Such a computation is practically unfeasible, if we have to deal with a very large number of particles, like those contained in a mole of a gas.

In order to overcome such a difficulty we can introduce a suitable model, based on simplifying hypotheses. We assume that in between two collisions the observed particle keeps constant the modulus of its velocity, v. Moreover, the distance run by the particle between two collisions is also assumed to be constant and equal to ℓ. Finally, the direction along which the particle moves after a collision is completely uncorrelated with the one it was moving along before the collision. The latter hypothesis amounts to assuming that collisions can actually be considered as random events, thus contradicting the fully mechanical, i.e., deterministic, origin of the problem.[4] Without prejudice of generality, we

[4] We want to point out that in this way we introduce a statistical concept into the description of a purely mechanical process. This conceptual step has been at the origin of a long-standing debate in the scientific community over more than a century. Nowadays, it has been commonly accepted and it is a cornerstone of modern science. Anyway, this basic assumption still today relies more on its effectiveness in predicting observed phenomena, rather than on its logical foundations. On the other hand, the need of a stochastic approach

https://doi.org/10.1017/9781107278974.002 Published online by Cambridge University Press

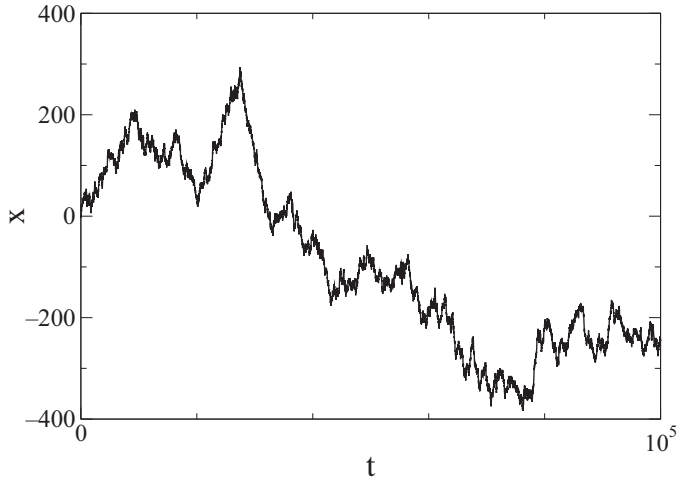

Fig. 1.3 Plot of a one-dimensional random walk, corresponding to $\ell = 1$.

assume that the selected particle at time $t = 0$ is at the origin of a Cartesian reference frame and we call $\mathbf{X}(t)$ the vector that identifies its position at time t. After having gone through N collisions, we can write

$$\mathbf{X} = \sum_{i=1}^{N} \mathbf{x}_i, \tag{1.14}$$

where $\mathbf{x}_i$ is the ith segment run by the particle after the ith collision ($|\mathbf{x}_i| = \ell, \forall i$), whose direction is random, i.e. uniformly distributed in the solid angle 4π. It is intuitive to conclude that as $N \to \infty$ the average value $\mathbf{X}/N \to 0$. For $d = 1$, in Fig. 1.3 we plot the resulting space-time trajectory, corresponding to $\ell = 1$.

As for the square displacement, we can write

$$X^2 \equiv \mathbf{X} \cdot \mathbf{X} = \sum_{i=1}^{N} \sum_{j=1}^{N} \mathbf{x}_i \cdot \mathbf{x}_j = \sum_{i=1}^{N} \sum_{j=1}^{N} \ell^2 \cos(\theta_{ij}) \tag{1.15}$$

where θ_{ij} is the angle in between the directions of segments $\mathbf{x}_i$ and $\mathbf{x}_j$. Since ℓ is a constant we can write the previous expression in a more convenient form,

$$X^2 = \ell^2 \sum_{i=1}^{N} \left(\sum_{j=1}^{N} \cos(\theta_{ij}) \right). \tag{1.16}$$

If $j = i$, then $\theta_{ij} = 0$, i.e. $\cos(\theta_{ij}) = 1$, and the previous equation can be written

$$X^2 = \ell^2 \sum_{i=1}^{N} \left(1 + \sum_{j \neq i} \cos(\theta_{ij}) \right). \tag{1.17}$$

could be also justified by invoking the contribution of dynamical details, such as the finite size of the particles or their internal rotational or vibrational degrees of freedom, that are usually neglected.

https://doi.org/10.1017/9781107278974.002 Published online by Cambridge University Press

The values taken by $\cos(\theta_{ij})$ can be thought as random numbers, distributed in the interval $[-1, +1]$. If we compute the average of X^2 over a very large number of different realizations (replicas) of this random walk the sum $\sum_{j\neq i} \cos(\theta_{ij})$ is negligible and one can finally write

$$\langle X^2 \rangle = \ell^2 N, \tag{1.18}$$

where the symbol $\langle \ \ \rangle$ denotes the average over replicas. Notice that the larger the number of replicas, the better the statistical estimate $\langle X^2 \rangle$ for any N.

This result can be generalized by assuming the less strict hypothesis that the length of runs in between subsequent collisions is distributed according to some normalized distribution $g(\ell)$. A "real" example in two dimensions is discussed later on, in the context of Brownian motion (see Section 1.4, Fig. 1.5). If $\mathbf{x}_i = \ell_i \hat{\mathbf{x}}_i$, where $\hat{\mathbf{x}}_i$ is the unit vector in the direction of $\mathbf{x}_i$, we can write

$$\langle \mathbf{x}_i \cdot \mathbf{x}_j \rangle = \langle \ell_i \ell_j \rangle \langle \hat{\mathbf{x}}_i \cdot \hat{\mathbf{x}}_j \rangle = \langle \ell_i^2 \rangle \delta_{ij} \tag{1.19}$$

and

$$\langle X^2 \rangle = \langle \ell^2 \rangle N. \tag{1.20}$$

If $g(\ell)$ is a Poisson distribution, which corresponds to independent random events,

$$g(\ell) = \frac{1}{\lambda} \exp\left(-\frac{\ell}{\lambda}\right), \tag{1.21}$$

where λ is the mean free path defined in Eq. (1.13), we have to substitute ℓ^2 with $\langle \ell^2 \rangle = 2\lambda^2$ in Eq. (1.18), thus obtaining

$$\langle X^2 \rangle = 2\lambda^2 N. \tag{1.22}$$

Notice that $\lambda N = L = \langle v \rangle t$ is the total length run by the particle after N collisions, so we can write

$$\langle X^2 \rangle = 2\lambda \langle v \rangle t. \tag{1.23}$$

This relation indicates that in the random walk, the particle that was at the origin at time $t = 0$ is found at time t at an average distance from the origin, $\sqrt{\langle X^2 \rangle}$, that grows proportionally to $\sqrt{t}$. The proportionality constant between $\langle X^2 \rangle$ and t is usually written as $2\lambda \langle v \rangle = 2dD$, where D is the diffusion coefficient of the random walk in d space dimensions. Diffusion in real situations is quite a slow process. For instance, if we consider air molecules at $T = 20^\circ$C we have $\langle v \rangle \sim 450$ m/s, while $\lambda \sim 0.06\,\mu$m. Accordingly, a diffusing air molecule in these conditions runs a distance of 1 m in approximately 5 h and a distance of 10 m in approximately 20 days (a quasi-static situation, if convective or turbulent motions do not occur).

https://doi.org/10.1017/9781107278974.002 Published online by Cambridge University Press

1.3 Transport Phenomena

The random walk model of a particle in an ideal gas is the appetizer of the general problem of transport processes. They concern a wide range of phenomena in hydrodynamics, thermodynamics, physical chemistry, electric conduction, magnetohydrodynamics, etc. They typically occur in physical systems (gases, liquids, or solids) made of many particles (atoms or molecules) in the presence of inhomogeneities. Such a situation can result from nonequilibrium conditions (e.g., the presence of a macroscopic gradient of density, velocity, or temperature), or simply from fluctuations around an equilibrium state.

The kinetic theory of transport phenomena provides a unified description of these apparently unlike situations. It is based on the assumption that even in nonequilibrium conditions gradients are small enough to guarantee that local equilibrium conditions still hold. In particular, the kinetic approach describes the natural tendency of the particles to transmit their properties from one region to another of the fluid by colliding with the other particles and eventually establishing global or local equilibrium conditions.

The main success of the kinetic theory is the identification of the basic mechanism underlying all the above-mentioned processes: the transport of a microscopic quantity (e.g., the mass, momentum or energy of a particle) over a distance equal to the mean free path λ of the particles, i.e. the average free displacement of a particle after a collision with another particle (see Eq. (1.13)). By this definition we are implicitly assuming that the system is a fluid, where each particle is supposed to interact with each other by collisions and propagate freely between successive collisions, the same conditions that we have discussed for the ideal gas model in Section 1.2.1.

Here we assume that we are dealing with a homogeneous isotropic system, where λ, the mean free path, is the same at any point and in any direction in space. Without prejudice of generality we consider a system where a uniform gradient of the quantity $A(\mathbf{x})$ is established along the z-axis, and $A(x,y,z) = A(x',y',z) = A(z)$ for any x, x', y, and y'. In particular, we assume that $A(z)$ is a microscopic quantity, which slowly varies at constant rate along the coordinate z of an arbitrary Cartesian reference frame. We consider also a unit surface S_1 located at height z and perpendicular to the z-axis; see Fig. 1.4(a). Any particle crossing the surface S_1 last collided at an average distance $\pm\lambda$ along the z-axis, depending on the direction it is moving. The net transport of the quantity $A(z)$ through S_1 amounts to the number of crossings of S_1 from each side in the unit time. Consistently with the assumption of local equilibrium we attribute the same average velocity $\langle v\rangle$ to all particles crossing S_1. Isotropy and homogeneity of the system imply also that one-third of the particles move on average along the z-axis, half of them upward and half downward. Accordingly, S_1 is crossed along z in the unit time interval by $\frac{1}{6}n\langle v\rangle$ particles in each direction.

The net flux of $A(z)$ through S_1 is given by

$$\Phi(A) = \frac{1}{6}\langle v\rangle\,[n(z-\lambda)A(z-\lambda\,) - n(z+\lambda)A(z+\lambda\,)\,]\,. \tag{1.24}$$

Since n and A vary weakly on the scale λ, one can use a first-order Taylor expansion and rewrite Eq. (1.24) as

https://doi.org/10.1017/9781107278974.002 Published online by Cambridge University Press

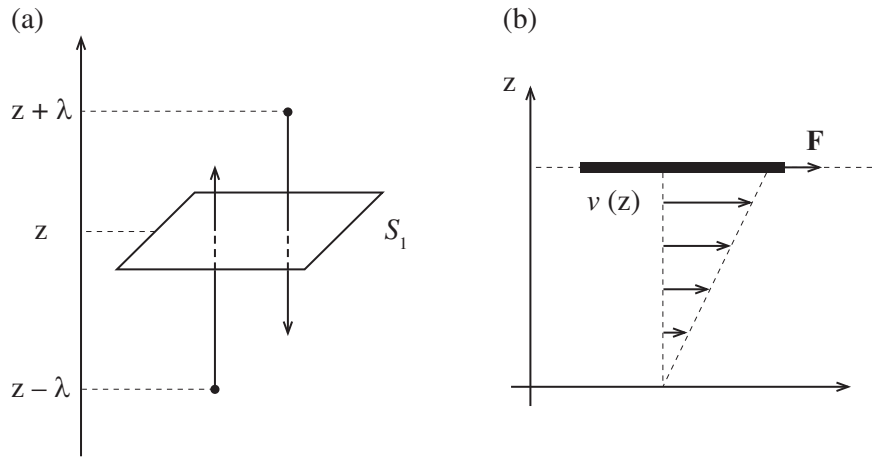

Fig. 1.4 (a) The surface S_1, normal to the $\hat{z}$ axis, is crossed by particles from both sides. Assuming particles move along the $\hat{z}$ axis, their most recent collision occurred at height $z - \lambda$ ($z + \lambda$) if their speed is positive (negative). (b) A gradient of velocity along the z-axis in a liquid, produced by a plate at the liquid surface that is constrained to move at a finite speed by the application of a force **F**.

$$\Phi(A) = -\frac{1}{3}\,\langle v\rangle\,\lambda\,\frac{\partial(nA)}{\partial z}. \tag{1.25}$$

This calculation can be performed more carefully by introducing explicitly the Maxwell distribution function of particle velocities at equilibrium. Nonetheless, one recovers the same result.

The simplest case is the transport of mass, because of a density gradient along the z-axis, in which case $A(z)$ is a constant and

$$\Phi(n) = -\frac{1}{3}\,\langle v\rangle\,\lambda\,\frac{\partial n}{\partial z} = -D\,\frac{\partial n}{\partial z}, \tag{1.26}$$

where the quantity $D = \frac{1}{3}\langle v\rangle\lambda$ defines the diffusion coefficient of particles inside the fluid. This expression is equal to the definition of D through Eq. (1.23), $D = \lambda\langle v\rangle/d$, because in the calculation here above $d = 3$. In a real physical situation D depends both on the diffusing substance and the medium of diffusion. At room temperature a gas in air typically has $D \simeq 0.3\,\mathrm{cm^2 s^{-1}}$; the diffusion coefficient of a liquid in water is typically of the order $D \simeq 10^{-5}\,\mathrm{cm^2 s^{-1}}$; a gas in a solid has a much smaller diffusivity, of the order $D \simeq 10^{-9}\,\mathrm{cm^2 s^{-1}}$.

Other cases of physical interest correspond to the situations where a gradient of velocity or temperature is present and the density n is assumed to be constant, so $\Phi(A) = -\frac{1}{3}\,n\langle v\rangle\,\lambda\,\frac{\partial A}{\partial z}$. If there is a gradient of velocity, we assume that the fluid flows with constant macroscopic velocity $v(z)$ parallel to the (x,y)-plane. In such a situation there is a net transport of kinetic momentum $mv(z)$ (m is the mass of a particle), yielding a shear stress $\Phi(mv(z))$ between the fluid layers laying on the (x,y) plane (see Fig. 1.4(b)):

https://doi.org/10.1017/9781107278974.002 Published online by Cambridge University Press

$$\Phi(mv(z)) = -\frac{1}{3}\, n\, m \langle v \rangle\, \lambda\, \frac{\partial v(z)}{\partial z} = -\eta\, \frac{\partial v(z)}{\partial z}, \tag{1.27}$$

where the quantity $\eta = \frac{1}{3}\, n\, m\, \langle v \rangle\, \lambda$ defines the viscosity of the fluid. By substituting Eq. (1.13) into the previous expression, one finds $\eta = \frac{m\langle v \rangle}{\sqrt{2}\, 3\sigma}$. This implies that the viscosity of an ideal fluid is independent of its density, i.e. of the pressure. This counterintuitive conclusion was first derived by Maxwell and its experimental verification sensibly contributed to establish in the scientific community a strong consensus on the atomistic approach of kinetic theory. It is worth stressing that such a conclusion does not hold when dealing with very dense fluids. At room temperature, diluted gases typically have η of order 10 μPa·s, while in water and blood η is of the order of few millipascal-seconds and honey at room temperature has $\eta \approx 1$ Pa·s.

It remains to consider the case when $A(z)$ is the average kinetic energy of particles $\bar{\epsilon}(z)$. At equilibrium, the energy equipartition condition yields the relation $n\bar{\epsilon}(z) = \rho\, C_V\, T(z)$, where $\rho = mn$ is the mass density of particles, C_V is the specific heat at constant volume and $T(z)$ is the temperature at height z. The net flux of kinetic energy $\Phi(\bar{\epsilon})$ can be read as the heat transported through the fluid along the z-axis,

$$\Phi(\bar{\epsilon}) = -\frac{1}{3} n \langle v \rangle \lambda \frac{\partial \bar{\epsilon}}{\partial z} = -\frac{1}{3}\, \rho\, C_V\, \langle v \rangle\, \lambda\, \frac{\partial T(z)}{\partial z} = -\kappa\, \frac{\partial T(z)}{\partial z}, \tag{1.28}$$

where the quantity $\kappa = \frac{1}{3}\, \rho\, C_V\, \langle v \rangle\, \lambda = \frac{m C_V \langle v \rangle}{\sqrt{2}\, 3\sigma}$ defines the heat conductivity. Also κ is found to be independent of n. The variability of κ in real systems is less pronounced than for other kinetic coefficients: in fact, a very good conductor like silver has $\kappa \simeq 400\, \mathrm{Wm^{-1}K^{-1}}$, while for cork, an effective heating insulator, it drops down to 4×10^{-2} in the same units.

One can conclude that the transport coefficients, i.e. the diffusion constant D, the viscosity η and the heat conductivity κ, are closely related to each other and depend on a few basic properties of the particles, like their mass m, their average velocity $\langle v \rangle$, and their mean free path λ. For example, by comparing the definitions of κ and η one finds the remarkable relation

$$\frac{\kappa}{\eta} = \alpha\, C_V, \tag{1.29}$$

with $\alpha = 1$. In real systems the constant α takes different values, which depend on the presence of internal degrees of freedom (e.g., $\alpha = \frac{5}{2}$ for realistic models of monoatomic gases).

The conceptual relevance of the relation (1.29) is that it concerns quantities that originate from quite different conditions of matter. In fact, on the left-hand side we have the ratio of two transport coefficients associated with macroscopic nonequilibrium conditions, while on the right-hand side we have a typically equilibrium quantity, the specific heat at constant volume. After what has been discussed in this section, this observation is far from mysterious: by assuming that even in the presence of a macroscopic gradient of physical quantities equilibrium conditions set in locally, the kinetic theory provides a unified theoretical approach for transport and equilibrium observables.

https://doi.org/10.1017/9781107278974.002 Published online by Cambridge University Press

1.4 Brownian Motion

The basic example of transport properties emerging from fluctuations close to equilibrium is Brownian motion. This phenomenon had been observed by Robert Brown in 1827: small pollen particles (typical size, 10^{-3} cm) in solution with a liquid exhibited an erratic motion, whose irregularity increased with decreasing particle size. The long scientific debate about the origin of Brownian motion lasted over almost a century and raised strong objections to the atomistic approach of the kinetic theory (see Section 1.2.1). In particular, the motion of the pollen particles apparently could not be consistent with an explanation based on collisions with the molecules of the liquid, subject to thermal fluctuations. In fact, pollen particles have exceedingly larger mass and size with respect to molecules and the amount of velocity acquired by a pollen particle in a collision with a molecule is typically 10^{-6} m/s. In Fig. 1.5 we reproduce some original data by Jean Baptiste Perrin.

The kinetic approach also implies that the number of collisions per second of the Brownian particle with liquid molecules is gigantic and random. Thermal equilibrium conditions and the isotropy of the liquid make equally probable the sign of the small velocity variations of the pollen particle. On observable time scales the effect of the very many collisions should average to zero and the pollen particle should not acquire any net displacement from its initial position in the fluid. This way of reasoning is actually wrong, as Albert Einstein argued in his theory of Brownian motion.

Einstein's methodological approach, based on the combination of simple models with basic physical principles, was very effective: he first assumed that the Brownian macroscopic particles should be considered as "big mass molecules," so that the system

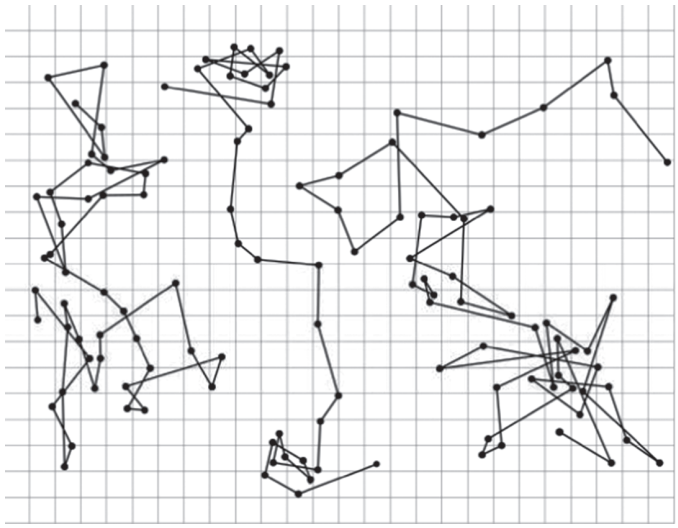

Fig. 1.5 Three tracings of the motion of colloidal particles of radius 0.53 μm, as seen under the microscope, are displayed. Successive positions every 30 s are joined by straight line segments (the mesh size is 3.1 μm). Reproduced from J. B. Perrin, *Les Atomes* (Paris: Librairie Félix Alcan, 1913). Wikimedia Commons.

https://doi.org/10.1017/9781107278974.002 Published online by Cambridge University Press

composed by the Brownian particles and the solvent, in which they are suspended, can be considered a mixture of two fluids at thermal equilibrium. This implies the validity of Van t' Hoff's law,

$$P(\mathbf{x}) = Tn(\mathbf{x}), \tag{1.30}$$

where P is the osmotic pressure between the fluids, T is the equilibrium temperature of the solution and n is the volumetric density of Brownian particles (the "solute" of the binary mixture). However, according to classical kinetic theory one should not be allowed to write an equation of state like (1.30) for a collection of the mesoscopic Brownian particles; they should be better described by the laws of dynamics, rather than by those of thermodynamics. Einstein's opposite point of view can be justified by considering that, despite their dimension, the Brownian particles are not subject to standard macroscopic forces: they experience microscopic forces originated by thermal fluctuations of the fluid molecules and in this sense they are equivalent to particles in solution.

Einstein guessed that Brownian motion looks random on a much larger time scale than the one needed for dissipating the energy acquired through collisions with molecules. In practice, this amounts to assuming that the dissipation mechanism should be described by the macroscopic Stokes's law

$$\mathbf{F} = m\frac{d\mathbf{v}}{dt} = -6\pi\eta R\mathbf{v}, \tag{1.31}$$

where $\mathbf{F}$ is the friction force proportional to the velocity $\mathbf{v}$ of the Brownian particle of radius R and η is the viscosity of the solvent. Therefore, the energy dissipation process of Brownian particles has to occur on the time scale

$$t_d = \frac{m}{6\pi\eta R}, \tag{1.32}$$

which must be much larger than the time τ between two collisions; see Eq. (1.7). If we consider a Brownian particle whose mass density is close to that of the solvent (e.g., water) and with radius $R \simeq 10^{-4}$ cm, at room temperature and pressure one has $t_d = O(10^{-7}\text{ s})$, while $\tau = O(10^{-11}\text{ s})$.

On a time scale much larger than t_d the Brownian particles are expected to exhibit a diffusive motion, as a consequence of the many random collisions with the solvent molecules. On the other hand, the kinetic theory associates diffusion with transport of matter in the presence of a density gradient (see Section 1.3). We should therefore induce such a gradient by a force. The Einstein approach will be exemplified by considering an ensemble of Brownian particles within a solvent, in a closed container and subject to the gravity. If particles were macroscopic (and heavier than the solvent), they would sit immobile at the bottom of the container, the collisions with the atoms of the solvent having no effect. In the case of Brownian particles we expect some diffusion, hindered by gravity. More precisely, we expect some profile of the density $n(z)$ of Brownian particles, with $\partial_z n < 0$ and $n(z) \to 0$ for $z \to +\infty$. This profile is the outcome of a stationary state that we are now going to describe in terms of equilibrium between currents and equilibrium between forces.

https://doi.org/10.1017/9781107278974.002 Published online by Cambridge University Press

If $F_0 = -mg$ is the force of gravity acting on a Brownian particle of mass m, its velocity satisfies the equation[5]

$$m\frac{dv}{dt} = F_0 - 6\pi\eta R v, \tag{1.33}$$

which leads to the asymptotic, sedimentation speed $v_0 = F_0/6\pi\eta R$. In the steady state, the resulting current $J_0 = n(z)v_0$ must be counterbalanced by the diffusion current, giving a vanishing total current,

$$-D\partial_z n(z) + n(z)v_0 = 0. \tag{1.34}$$

In terms of equilibrium between forces, the force related to the diffusive motion of Brownian particles is due to the osmotic pressure (1.30). Therefore, equilibrium implies

$$-\partial_z P + n(z)F_0 = 0. \tag{1.35}$$

Using the Van t' Hoff law and the explicit expression of the sedimentation speed, from Eqs. (1.34) and (1.35) we find the same equation for the density profile,

$$-\partial_z n(z) + \frac{F_0}{6\pi\eta RD} n(z) = 0 \tag{1.36}$$

$$-\partial_z n(z) + \frac{F_0}{T} n(z) = 0, \tag{1.37}$$

which implies the remarkable relation, called Einstein's formula,

$$D = \frac{T}{6\pi\eta R} \equiv \frac{T}{\tilde{\gamma}}, \tag{1.38}$$

where we have defined the friction coefficient (see Eq. (1.33)), $\tilde{\gamma} = 6\pi\eta R$. Perrin showed that this formula provides very good quantitative agreement with experimental data.

The solution of the equation for $n(z)$ gives an exponential profile for the density of Brownian particles,

$$n(z) = n(0)\exp\{(-mg/T)z\}, \tag{1.39}$$

as attested by the experimental results reported in Fig. 1.6.

In summary, Einstein's theory of Brownian motion is based on the description of the Brownian particle in both microscopic and macroscopic terms: the microscopic description is employed when we use the van t' Hoff law to describe Brownian particles as a diluted gas in a solvent, while the macroscopic description is introduced via the Stokes law.

1.4.1 The Langevin Equation for the Brownian Particle

A mechanical approach that is valid at both $t < t_d$ (ballistic regime) and $t > t_d$ (diffusive regime) was later proposed by Paul Langevin. For the first time deterministic and stochastic forces were introduced into the same equation. The deterministic component is the friction term $-\tilde{\gamma}\mathbf{v}$, acting on the Brownian particle: $\tilde{\gamma}$ is the friction coefficient, which amounts to $\tilde{\gamma} = 6\pi\eta R$ in the Stokes regime and $\mathbf{v}$ is the three-dimensional velocity vector of

[5] We can deal with relations among scalar quantities, because all of them are directed along the z-axis.

https://doi.org/10.1017/9781107278974.002 Published online by Cambridge University Press

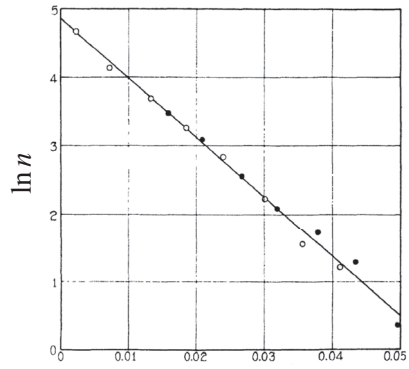

Fig. 1.6 Height distribution of the density of gold sols, showing an exponentially decreasing function. From N. Johnston and L. G. Howell, Sedimentation equilibria of colloidal particles, *Physical Review*, **35** (1930) 274-282.

the Brownian particles, whose components are denoted by v_i, $i = 1, 2, 3$. The stochastic component associated with the effect of collisions with the solvent molecules is a random, time-dependent force $\tilde{\boldsymbol{\eta}}(t)$, whose components are analogously denoted by $\tilde{\eta}_i$, $i = 1, 2, 3$. Symmetry arguments indicate that each one of these components can be assumed to be independent, isotropic, uncorrelated in time (at least for $t \gg \tau$) and Gaussian, since it results from the combination of a very large number of collisions, which can be approximately considered as independent events (see Appendix A). Accordingly, the average of the stochastic force is null and its time-correlation function can be approximated by a Dirac delta for $t \gg \tau$:

$$\langle \tilde{\eta}_i(t) \rangle = 0 \tag{1.40a}$$

$$\langle \tilde{\eta}_i(t) \tilde{\eta}_j(t') \rangle = \tilde{\Gamma} \delta_{ij} \delta(t - t'), \tag{1.40b}$$

where the brackets $\langle \ \rangle$ indicate the statistical average, $\tilde{\Gamma}$ is a suitable dimensional constant, and δ_{ij} is a Kronecker delta. The Langevin equation is a Newtonian equation containing the stochastic force $\tilde{\boldsymbol{\eta}}$. For each component of the three-dimensional velocity vector it reads

$$m \frac{dv_i(t)}{dt} = -\tilde{\gamma}\, v_i(t) + \tilde{\eta}_i(t), \tag{1.41}$$

which is rewritten, for the sake of simplicity, as

$$\frac{dv_i(t)}{dt} = -\gamma\, v_i(t) + \eta_i(t), \tag{1.42}$$

where we have introduced the symbol $\gamma = \tilde{\gamma}/m = 1/t_d$ and the stochastic force per unit mass, $\boldsymbol{\eta} = \tilde{\boldsymbol{\eta}}/m$. The latter satisfies the same equations as (1.40), with $\tilde{\Gamma}$ replaced by $\Gamma = \tilde{\Gamma}/m^2$.

Eq. (1.42) can be formally integrated, yielding the solution

$$v_i(t) = \exp(-\gamma t) \left[v_i(0) + \int_0^t d\tau \exp(\gamma \tau)\, \eta_i(\tau) \right]. \tag{1.43}$$

https://doi.org/10.1017/9781107278974.002 Published online by Cambridge University Press

Accordingly, $v_i(t)$ is a stochastic function with average value

$$\langle v_i(t)\rangle \;=\; v_i(0)\exp\left(-\gamma t\right) \tag{1.44}$$

and average squared value

$$\begin{aligned}\langle v_i^2(t)\rangle &= \exp\left(-2\gamma t\right)\left[v_i^2(0) \;+\; \int_0^t d\tau \int_0^t d\tau' \exp\left(\gamma\left(\tau+\tau'\right)\right)\langle \eta_i(\tau)\eta_i(\tau')\rangle\right] \\ &= \exp\left(-2\gamma t\right)\left[v_i^2(0) \;+\; \int_0^t d\tau\,\Gamma\exp\left(2\gamma\tau\right)\right] \\ &= v_i^2(0)\,\exp\left(-2\gamma t\right) + \frac{\Gamma}{2\gamma}\left[1 \;-\; \exp\left(-2\gamma t\right)\right].\end{aligned} \tag{1.45}$$

Above we first used Eq. (1.40a) to cancel out the terms that are linear in the noise, then we explicitly used the noise correlation function (1.40b).

As time t grows the average velocity vanishes exponentially with rate γ, while the average squared velocity approaches the value

$$\lim_{t\to\infty}\langle v_i^2(t)\rangle \;=\; \frac{\Gamma}{2\gamma}. \tag{1.46}$$

In order to bridge this purely mechanical approach with Einstein theory it remains to establish a relation with thermodynamics. This is naturally obtained by attributing a thermal origin to the fluctuations of the stochastic force. In particular, if we assume that the Brownian particles and the solvent are in thermal equilibrium at temperature T, the energy equipartition principle establishes that, in the limit of very large times, the average kinetic energy per degree of freedom is proportional to the temperature,

$$\lim_{t\to\infty}\left\langle \frac{1}{2}\,m\,v_i^2(t)\right\rangle \;=\; \frac{1}{2}\,T. \tag{1.47}$$

By comparison with (1.46) one obtains the formula,

$$\Gamma \;=\; \frac{2\,\gamma\,T}{m}, \tag{1.48}$$

or equivalently

$$\tilde{\Gamma} = 2\tilde{\gamma}T, \tag{1.49}$$

which is a basic example of a fluctuation–dissipation relation.

From an experimental point of view, the fluctuations of the velocity of the particle cannot be measured because of the erratic trajectories, while it is easy to measure the mean square displacement of the Brownian particle from its initial position:

$$\begin{aligned}\langle\left(x_i(t)-x_i(0)\right)^2\rangle &= \left\langle\left[\int_0^t v_i(\tau)d\,\tau\right]^2\right\rangle \\ &= \int_0^t d\,\tau\int_0^t d\,\tau'\langle v_i(\tau)v_i(\tau')\rangle.\end{aligned} \tag{1.50}$$

https://doi.org/10.1017/9781107278974.002 Published online by Cambridge University Press

The expression between brackets in the last integral is the velocity correlation function of the Brownian particle, which can be computed from (1.43), similarly to the derivation of Eq. (1.45). We obtain

$$\langle v_i(\tau)v_i(\tau')\rangle = v_i^2(0)e^{-\gamma(\tau+\tau')} + \Gamma e^{-\gamma(\tau+\tau')}\int_0^{\tau} dt_1 \int_0^{\tau'} dt_2 e^{\gamma(t_1+t_2)}\delta(t_1-t_2). \tag{1.51}$$

The double integral depends on which variable between τ and τ' is the smallest; if $\tau < \tau'$, we must first integrate over t_2, obtaining

$$\begin{aligned}\Gamma e^{-\gamma(\tau+\tau')}\int_0^{\tau} dt_1 \int_0^{\tau'} dt_2 e^{\gamma(t_1+t_2)}\delta(t_1-t_2) &= \Gamma e^{-\gamma(\tau+\tau')}\int_0^{\min(\tau,\tau')} dt_1 e^{2\gamma t_1} \\ &= \frac{\Gamma}{2\gamma}\Big(e^{-\gamma|\tau-\tau'|} - e^{-\gamma(\tau+\tau')}\Big),\end{aligned}$$

a result that is valid in any case. So, Eq. (1.51) becomes

$$\langle v_i(\tau)v_i(\tau')\rangle = v_i^2(0)e^{-\gamma(\tau+\tau')} + \frac{\Gamma}{2\gamma}\Big(e^{-\gamma|\tau-\tau'|} - e^{-\gamma(\tau+\tau')}\Big). \tag{1.52}$$

Notice that for $\tau' = \tau$, the previous correlator simplifies to Eq. (1.45). Finally, substituting into (1.50) we obtain

$$\langle (x_i(t) - x_i(0))^2\rangle = \left(v_i^2(0) - \frac{\Gamma}{2\gamma}\right)\frac{(1-e^{-\gamma t})^2}{\gamma^2} + \frac{\Gamma}{\gamma^2}t - \frac{\Gamma}{\gamma^3}\Big(1-e^{-\gamma t}\Big), \tag{1.53}$$

where we have used the following relations

$$\begin{aligned}\int_0^t d\tau \int_0^t d\tau' e^{-\gamma(\tau+\tau')} &= \left(\frac{1-e^{-\gamma t}}{\gamma}\right)^2 \\ \int_0^t d\tau \int_0^t d\tau' e^{-\gamma|\tau-\tau'|} &= 2\int_0^t d\tau \int_0^{\tau} d\tau' e^{-\gamma(\tau-\tau')} \\ &= \frac{2}{\gamma}t - \frac{2}{\gamma^2}\Big(1-e^{-\gamma t}\Big).\end{aligned} \tag{1.54}$$

We can now evaluate (1.53) in the limits $t \ll t_d$ and $t \gg t_d$. In the former case, $\gamma t \ll 1$ and $1 - e^{-\gamma t} \simeq \gamma t - \gamma^2 t^2/2$; in the latter case, $\gamma t \gg 1$ and $1 - e^{-\gamma t} \simeq 1$. Therefore, in the two limits the mean square displacement of the Brownian particle has the following expressions

$$\langle (x_i(t) - x_i(0))^2\rangle = \begin{cases} v_i^2(0)t^2 & t \ll t_d \quad \text{ballistic regime} \\ \dfrac{\Gamma}{\gamma^2}t & t \gg t_d \quad \text{diffusive regime.}\end{cases} \tag{1.55}$$

Using (1.48) and the Einstein relation (1.38), we obtain $\Gamma/\gamma^2 = 2T/(m\gamma) = 2D$, so that $\langle (x_i(t) - x_i(0))^2\rangle = 2Dt$ in the diffusive regime. More generally, in d spatial dimensions we have

$$\lim_{t\to\infty} \frac{\langle (\mathbf{x}(t) - \mathbf{x}(0))^2\rangle}{t} = 2dD. \tag{1.56}$$

https://doi.org/10.1017/9781107278974.002 Published online by Cambridge University Press

It is straightforward to generalize the Langevin equation (1.41) to the case where the Brownian particle is subject to a conservative force $F(x) = -U'(x)$,

$$m\ddot{x}(t) = -U'(x) - \tilde{\gamma}\dot{x}(t) + \tilde{\eta}(t). \tag{1.57}$$

This problem will be reconsidered in Section 1.6.3 in the context of the Ornstein–Uhlenbeck processes and in Section 1.6.5 where detailed balance conditions are considered.

1.4.2 The Fokker–Planck Equation for the Brownian Particle

The statistical content of the Langevin approach emerges explicitly from the properties attributed to the random force in Eq. (1.40b). The statistical average $\langle\ \rangle$ can be interpreted as the result of the average over many different trajectories of the Brownian particle obtained by different realizations of the stochastic force components η_i in Eq. (1.42). By taking inspiration from Boltzmann kinetic theory one can get rid of mechanical trajectories and describe from the very beginning the evolution of Brownian particles by a distribution function $p(\mathbf{x}, t)$, such that $p(\mathbf{x}, t)d^3x$ represents the probability of finding the Brownian particle at time t in a position between $\mathbf{x}$ and $\mathbf{x} + d^3x$. The evolution is ruled by the transition rate $W(\mathbf{x}', \mathbf{x})$, which represents the probability per unit time and unit volume that the Brownian particle "jumps" from $\mathbf{x}$ to $\mathbf{x}'$.

In this approach any reference to instantaneous collisions with the solvent molecules is lost and one has to assume that relevant time scales are much larger than t_d. As proposed by Fokker and Planck, one can write the evolution equation for $p(\mathbf{x}, t)$ as a master equation, i.e. a balance equation where the variation in time of $p(\mathbf{x}, t)$ emerges as the result of two competing terms: a gain factor, due to jumps of particles from any position $\mathbf{x}'$ to $\mathbf{x}$, and a loss factor, due to jumps from $\mathbf{x}$ to any $\mathbf{x}'$. Thus

$$\frac{\partial p(\mathbf{x}, t)}{\partial t} = \int_{-\infty}^{+\infty} d^3x' \left[p(\mathbf{x}', t)W(\mathbf{x}, \mathbf{x}') - p(\mathbf{x}, t)W(\mathbf{x}', \mathbf{x}) \right], \tag{1.58}$$

where $d^3x'W(\mathbf{x}, \mathbf{x}')$ ($d^3x'W(\mathbf{x}', \mathbf{x})$) represents the probability per unit time that the particle jumps from a neighborhood of $\mathbf{x}'$ to $\mathbf{x}$ (or vice versa). If we define $\boldsymbol{\chi} = \mathbf{x}' - \mathbf{x}$ and we use the notation $W(\mathbf{x}; \boldsymbol{\chi}) = W(\mathbf{x}', \mathbf{x})$, we obtain

$$\frac{\partial p(\mathbf{x}, t)}{\partial t} = \int d^3x' \left[p(\mathbf{x}', t)W(\mathbf{x}'; -\boldsymbol{\chi}) - p(\mathbf{x}, t)W(\mathbf{x}; \boldsymbol{\chi}) \right] \tag{1.59}$$

$$= \int d^3\chi \left[p(\mathbf{x} - \boldsymbol{\chi}, t)W(\mathbf{x} - \boldsymbol{\chi}; \boldsymbol{\chi}) - p(\mathbf{x}, t)W(\mathbf{x}; \boldsymbol{\chi}) \right], \tag{1.60}$$

where in the second equality we passed from the variable $\mathbf{x}'$ to $\boldsymbol{\chi}$ and in the first integral of (1.60) we have substituted $\boldsymbol{\chi}$ with $-\boldsymbol{\chi}$.

Since large displacements of the Brownian particles are very infrequent, one can reasonably assume that the rate functions $W(\mathbf{x} - \boldsymbol{\chi}; \boldsymbol{\chi})$ and $W(\mathbf{x}; \boldsymbol{\chi})$ are significantly different from zero only for very small $\boldsymbol{\chi}$. This allows one to introduce a formal Taylor series expansion in Eq. (1.60) around $\boldsymbol{\chi} = 0$. By considering only terms up to the second order one obtains

https://doi.org/10.1017/9781107278974.002 Published online by Cambridge University Press

$$\frac{\partial p(\mathbf{x},t)}{\partial t} = \int d^3\chi \left[-\nabla \left(p(\mathbf{x},t) W(\mathbf{x};\chi) \right) \cdot \boldsymbol{\chi} + \frac{1}{2} \sum_{i,j} \frac{\partial^2}{\partial x_i \, \partial x_j} \left(p(\mathbf{x},t) W(\mathbf{x};\boldsymbol{\chi}) \right) \chi_i \chi_j \right]. \tag{1.61}$$

Eq. (1.61) has been found with the Brownian particle in mind, but the variable $\mathbf{x}$ can be any vectorial quantity defining the state of a physical system, not only the spatial position of a particle. For this reason it has a range of application going well beyond the motion of pollen particles in a solvent and it is worth making a further step before specializing to the Brownian motion. We can formally define the quantities

$$\alpha_i(\mathbf{x}) = \int d^3\chi \; W(\mathbf{x};\boldsymbol{\chi}) \; \chi_i \tag{1.62}$$

$$\beta_{ij}(\mathbf{x}) = \int d^3\chi \, W(\mathbf{x};\boldsymbol{\chi}) \chi_i \chi_j \tag{1.63}$$

so that Eq. (1.61) can be written as

$$\frac{\partial p(\mathbf{x},t)}{\partial t} = -\sum_i \frac{\partial}{\partial x_i} \left(\alpha_i(\mathbf{x}) p(\mathbf{x},t) \right) + \frac{1}{2} \sum_{i,j} \frac{\partial^2}{\partial x_i \partial x_j} \left(\beta_{ij}(\mathbf{x}) p(\mathbf{x},t) \right). \tag{1.64}$$

This is the celebrated Fokker–Planck equation and it will be used in different contexts of this book.

Let us now come back to the Brownian motion, in which case $\mathbf{x}$ is the spatial position of the particle and the quantities $\boldsymbol{\alpha}$ and β_{ij} have a simple interpretation as average quantities. The quantity $\boldsymbol{\alpha}$ is the average displacement of a Brownian particle per unit time;[6] i.e. its average velocity is

$$\boldsymbol{\alpha} = \frac{\langle \Delta \mathbf{x} \rangle}{\Delta t}. \tag{1.65}$$

In the absence of external forces it vanishes, but in the presence of a constant force $\mathbf{F}_0$ (e.g., gravity) the average velocity is equal to the sedimentation speed $\mathbf{v}_0 = \mathbf{F}_0/\tilde{\gamma}$ (see below Eq. (1.33)), resulting from the balance of the external force acting on the Brownian particle with the viscous friction force produced by the solvent.

The quantity β_{ij} amounts to the average squared displacements per unit time,

$$\beta_{ij} = \frac{\langle \Delta x_i \Delta x_j \rangle}{\Delta t}. \tag{1.66}$$

In a homogeneous, isotropic medium it is diagonal and proportional to the diffusion constant D,

$$\beta_{ij} = 2\delta_{ij} D. \tag{1.67}$$

For the Brownian particle, we finally obtain

$$\frac{\partial p(\mathbf{x},t)}{\partial t} = -\mathbf{v}_0 \cdot \nabla p(\mathbf{x},t) + D\nabla^2 p(\mathbf{x},t). \tag{1.68}$$

[6] Note that W is a transition rate, i.e. a probability per unit time and also per unit volume.

https://doi.org/10.1017/9781107278974.002 Published online by Cambridge University Press

In the absence of external forces the viscous term can be neglected and Eq. (1.68) reduces to the diffusion equation,

$$\frac{\partial p(\mathbf{x},t)}{\partial t} = D\nabla^2 p(\mathbf{x},t), \tag{1.69}$$

whose solution is given, for $d = 1$, by Eq. (D.11),

$$p(x,t) = \frac{1}{\sqrt{4\pi Dt}} \exp\left(-\left[\frac{(x-x_0)^2}{4Dt}\right]\right). \tag{1.70}$$

The average squared displacement is given by the variance of the Gaussian distribution, i.e.

$$\langle x^2\rangle - \langle x\rangle^2 = 2\,D\,t. \tag{1.71}$$

Comparing with Eq. (1.55), we can conclude that the Fokker–Planck approach predicts the same asymptotic diffusive behavior as the Langevin approach.

1.5 Discrete Time Stochastic Processes

As it happens for a Brownian particle, the evolution in time of a great many systems in nature is affected by random fluctuations, usually referred to as "noise." These fluctuations can be originated by many different physical processes, such as the interaction with thermalized particles of the solvent (as in the case of the Brownian particle), absorption and emission of radiation, chemical reactions at molecular level, different behaviors of agents in a stock market or of individuals in a population, etc. When we introduce the concept of "noise" in these phenomena we are assuming that part of their evolution is ruled by "unknown" variables, whose behavior is out of our control, i.e. they are unpredictable, in a deterministic sense. On the other hand, we can assume that such fluctuations obey general statistical rules, which still allow us to obtain the possibility of predicting the evolution of these systems in a probabilistic sense. An everyday example of such a situation is a closed bottle thrown into the sea. During its navigation the bottle is continuously subject to small-scale random fluctuations of different duration and amplitude, due to the waves interfering on the sea surface. This notwithstanding, the bottle is expected to follow over a large-scale distance some macroscopic current flow, which may eventually lead it to a populated seaside. This is at least what the castaway, who threw the bottle in the sea with a message inside, expects will eventually occur with some not entirely negligible probability.

The mathematics that has been set up for describing these kind of phenomena is the theory of stochastic processes. Many great scientists, such as Marian Smoluchowski, Louis Bachelier, Max Planck, Norbert Wiener, George Uhlenbeck, Albert Enstein, and Paul Langevin (as mentioned in the previous section), contributed to this branch of mathematics. Its foundations were established by the systematic work of the Russian mathematician Andrej Markov, who introduced the concept of the so-called Markov chains.

At variance with the previous section, here we want to introduce the reader to a more rigorous mathematical formulation of the concepts that have been described in the previous

https://doi.org/10.1017/9781107278974.002 Published online by Cambridge University Press

section by intuitive or heuristic arguments. First of all we shall deal with stochastic processes in discrete space and time. They are particularly useful to establish the basic mathematical tools that will be later extended to continuous space and time, into the form of stochastic differential equations. Eventually, we shall obtain a consistent formulation of the general Fokker–Planck formalism, which deals with the evolution of probability distributions.

1.5.1 Markov Chains

Markov chains are stochastic processes discrete in time and in the state space, where the value assumed by each stochastic variable depends on the value taken by the same variable at the previous instant of time. Let us translate these concepts into a mathematical language. We assume that the stochastic variable $x(t)$ takes values at each instant of time t over a set of N states, $S = \{s_1, s_2, \ldots, s_{N-1}, s_N\}$. For the sake of simplicity, we also assume that $x(t)$ is measured at equal finite time intervals, so that also time becomes a discrete variable, equivalent, in some arbitrary time unit, to the ordered sequence of natural numbers, $t = 1, 2, \ldots, n, \ldots$. The basic quantity we want to deal with is the probability, $p(x(t) = s_i)$, that $x(t)$ is in state s_i at time t. If $p(x(t) = s_i)$ does not depend on the previous history of the stochastic process, we are dealing with the simple case of a sequence of independent events, like tossing a coin. In general, one could expect that, if some time correlation (i.e. memory) is present in the evolution of $x(t)$, then $p(x(t) = s_i)$ could depend also on the previous history of the stochastic process. In this case it is useful to introduce the conditional probability

$$\Omega\left(x(t) = s_i \,|\, x(t-1) = s_{i_1}, x(t-2) = s_{i_2}, \ldots, x(t-n) = s_{i_n}\right) \tag{1.72}$$

that $x(t)$ is in the state s_i at time t, given the evolution of the stochastic process backward in time up to $t - n$. In this case we say that the stochastic process has memory n: the case $n = 1$ defines a Markov process, where

$$\Omega\left(x(t) = s_j \,|\, x(t-1) = s_i\right) \tag{1.73}$$

is the transition probability in a unit time step from s_i to s_j; see Fig. 1.7.[7]

In general, $\Omega\left(x(t) = s_j \,|\, x(t-1) = s_i\right)$ is a function of time, but here we limit ourselves to consider stationary Markov processes, where this transition rate is independent of time. In this case we can deal with many interesting problems that can be discussed and solved with basic mathematical ingredients. The time-dependent case is certainly more general

[7] An example of a non-Markovian process is a self-avoiding random walk. We can think of a random walker moving at each time step between nearby sites of a regular square lattice, choosing with equal probability any available direction. On the other hand, the walker cannot move to any site already visited in the past. The time evolution of the random process modifies the probability of future events, or said differently, the walker experiences a persistent memory effect, that typically yields long-time correlations. Such a situation is peculiar to many interesting phenomena concerning other domains of science, such as sociology or economics. For instance, the price X of a stock at time t is usually more properly represented as an outcome of a non-Markovian process, where this quantity depends on the overall economic activity $a(t)$ up to time t, i.e. $X(a(t))$ rather than simply $X(t)$.

https://doi.org/10.1017/9781107278974.002 Published online by Cambridge University Press

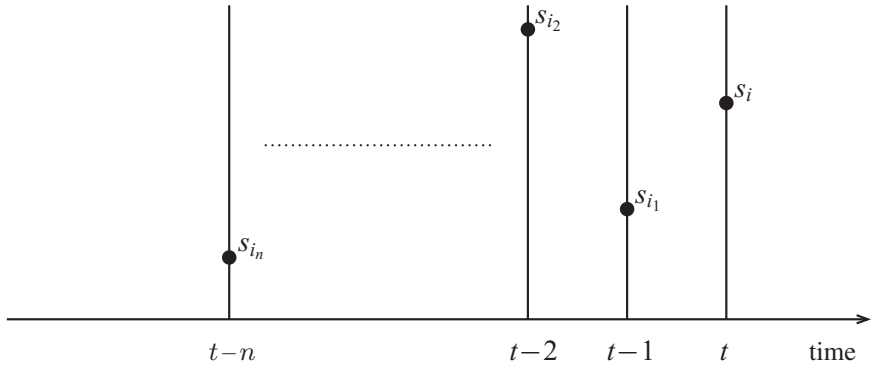

Fig. 1.7 The configuration space at time t is represented by a vertical axis crossing the time axis in t. In the most general case, the probability to be in s_i at time t depends on the previous history at times $t-1, t-2, \ldots, t-n$.

and of major interest, but even a short introduction would be hardly accessible for the audience to which this textbook is addressed.

Making use of the shorthand notations $p_i(t) \equiv p(\,x(t) = s_i\,)$ and $W_{ji} \equiv \Omega(\,x(t+1) = s_j \,|\, x(t) = s_i\,)$, these quantities must obey the following conditions:

$$p_i(t) \geq 0, \qquad \forall i, t \tag{1.74}$$

$$\sum_i p_i(t) = 1, \qquad \forall t \tag{1.75}$$

$$W_{ij} \geq 0, \qquad \forall i,j \tag{1.76}$$

$$\sum_i W_{ij} = 1, \qquad \forall j\,. \tag{1.77}$$

We can define the stochastic dynamical rule of the Markov chain as

$$p_j(t+1) = \sum_i W_{ji} p_i(t), \tag{1.78}$$

so one can easily realize that the W_{ij} can be viewed as the entries of an $N \times N$ matrix W, called stochastic matrix. Accordingly, Eq. (1.78) can be rewritten in vector form as

$$\mathbf{p}(t+1) = W\mathbf{p}(t), \tag{1.79}$$

where $\mathbf{p}(t) = (\,p_1(t), p_2(t), \ldots, p_j(t), \ldots, p_N(t)\,)$ is the column vector of the probability. This matrix relation can be generalized to obtain

$$\mathbf{p}(t+n) = W^n\mathbf{p}(t), \tag{1.80}$$

where W^n, the nth power of W, is also a stochastic matrix, since it satisfies the same properties (1.76) and (1.77) of W. This is easily proved by induction, assuming W^n is a stochastic matrix and showing that W^{n+1} is also stochastic. In fact,

$$(W^{n+1})_{ij} = \sum_k (W^n)_{ik} W_{kj} \geq 0, \qquad \forall i,j \tag{1.81}$$

https://doi.org/10.1017/9781107278974.002 Published online by Cambridge University Press

because each term of the sum is the product of nonnegative quantities, and

$$\sum_i (W^{n+1})_{ij} = \sum_{i,k} (W^n)_{ik} W_{kj} \tag{1.82}$$

$$= \sum_k \left(\sum_i (W^n)_{ik} \right) W_{kj} \tag{1.83}$$

$$= \sum_k W_{kj} \tag{1.84}$$

$$= 1. \tag{1.85}$$

The matrix relation (1.80) also leads to another important relation concerning the stochastic matrix, the Chapman–Kolmogorov equation,

$$p(t+n) = W^n p(t) = W^n W^t p(0) = W^{t+n} p(0). \tag{1.86}$$

It extends to stochastic processes the law valid for deterministic dynamical systems, where the evolution operator from time 0 to time $(t+n)$, $\mathcal{L}^{t+n}$, can be written as the composition of the evolution operator from time 0 to time t with the evolution operator from time t to time $t+n$, namely $\mathcal{L}^{t+n} = \mathcal{L}^n \circ \mathcal{L}^t$.

As usual for any $N \times N$ matrix it is useful to solve the eigenvalue problem

$$\det (W - \lambda I) = 0,$$

where I is the identity matrix and λ is a scalar quantity, whose values solving the eigenvalue equation are called the spectrum of W.

Since W is not a symmetric matrix, its eigenvalues are not necessarily real numbers.[8] Moreover, one should distinguish between right and left eigenvectors of W. We denote the right ones as $\mathbf{w}^{(\lambda)} = (w_1^{(\lambda)}, w_2^{(\lambda)}, \ldots, w_j^{(\lambda)}, \ldots, w_N^{(\lambda)})$, so that we can write

$$W\, \mathbf{w}^{(\lambda)} = \lambda\, \mathbf{w}^{(\lambda)}. \tag{1.87}$$

The spectrum of W has the following properties, whose proof is given in Appendix B:

(a) $|\lambda| \leq 1$;
(b) there is at least one eigenvalue $\lambda = 1$;
(c) $\mathbf{w}^{(\lambda)}$ is either an eigenvector with eigenvalue 1, or it fulfills the condition $\sum_j w_j^{(\lambda)} = 0$.

We conclude this section by introducing some definitions:

Definition 1.1 A state s_j is accessible from a state s_i if there is a finite value of time t such that $(W^t)_{ji} > 0$.

Definition 1.2 A state s_j is persistent if the probability of returning to s_j after some finite time t is 1, while it is transient if there is a finite probability of never returning to s_j for any finite time t. As a consequence of these definitions a persistent state will be visited infinitely many times, while a transient state will be discarded by the evolution after a sufficiently long time.

[8] Notice that the asymmetry of W implies that, in general, $\sum_j W_{ij} \neq 1$, at variance with (1.77).

https://doi.org/10.1017/9781107278974.002 Published online by Cambridge University Press

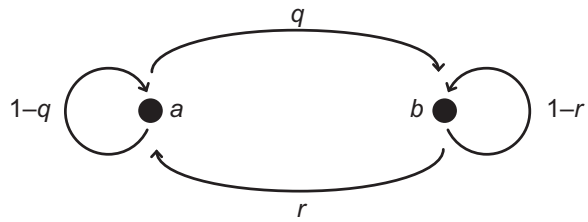

Fig. 1.8 Diagram of allowed transitions for a two-state Markov chain.

Definition 1.3 A Markov chain is irreducible when all the states are accessible from any other state.

Definition 1.4 A Markov chain is periodic when the return times T_j on a state s_j are all (integer) multiples of a period T, i.e. $(W^T)_{jj} > 0$.

1.5.2 Useful Examples of Markov Chains

The Two-State Case

We consider a Markov chain made of two states, $S = \{a, b\}$. The stochastic matrix has the form

$$W = \begin{pmatrix} 1-q & r \\ q & 1-r \end{pmatrix}, \tag{1.88}$$

because condition (1.77) implies that the sum of elements of each column is equal to one. Accordingly, q is the probability rate per unit time of passing from a to b and $(1-q)$ of remaining in a, while r is the probability rate per unit time of passing from b to a and $(1-r)$ of remaining in b (see Fig. 1.8).

Let us denote with $p_a(t)$ the probability of observing the system in state a at time t: relation (1.75) yields $p_b(t) = 1 - p_a(t)$ at any t. By applying the stochastic evolution rule (1.78) one obtains the equation

$$p_a(t+1) = (1-q)p_a(t) + r(1-p_a(t)) = r + (1-r-q)p_a(t). \tag{1.89}$$

This equation implies also the similar equation for $p_b(t)$, which can be obtained from (1.89) by exchanging $p_a(t)$ with $p_b(t)$ and r with q. By simple algebra[9] one can check that the explicit solution of (1.89) is

$$p_a(t) = \alpha + (1-r-q)^t \left(p_a(0) - \alpha\right), \qquad \alpha = \frac{r}{r+q}, \tag{1.90}$$

where $p_a(0)$ is the initial condition, i.e. the probability of observing the state a at time 0.

There are two limiting cases: (i) $r = q = 0$, no dynamics occurs; (ii) $r = q = 1$, dynamics oscillates forever between state a and state b. In all other cases, $|1-r-q| < 1$

[9] If we start iterating Eq. (1.89), we find $p_a(1) = r + \beta p_a(0)$, $p_a(2) = r(1+\beta) + \beta^2 p_a(0)$, $p_a(3) = r(1+\beta+\beta^2) + \beta^3 p_a(0)$, …, with $\beta = 1-r-q$. We can make the ansatz $p_a(t) = r\sum_{\tau=0}^{t-1}\beta^\tau + \beta^t p_a(0)$. Evaluating the summation explicitly, which is equal to $(1-\beta^t)/(1-\beta)$, we find the expression given in Eq. (1.90).

https://doi.org/10.1017/9781107278974.002 Published online by Cambridge University Press

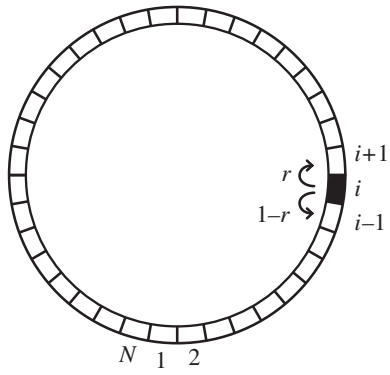

Fig. 1.9 Anisotropic diffusion on a ring, i.e., a linear segment of N sites with periodic boundary conditions.

and in the limit $t \to \infty$, $p_a \to \alpha$ and $p_b \to (1-\alpha)$. More precisely, $p_a(t)$ converges exponentially fast to α with rate $\tau = -1/\ln|r+q-1|$, i.e. the dynamics approaches exponentially fast a stationary state.[10] This simple Markov chain is irreducible and its states are accessible and persistent.

Notice that also the following relation holds

$$W_{ba}p_a(\infty) = W_{ab}p_b(\infty). \tag{1.91}$$

As we are going to discuss in more generality in Section 1.5.4, Eq. (1.91) is the detailed balance condition that establishes a sort of time reversibility of the stochastic process. Actually, Eq. (1.91) tells us that, in the stationary state, the probability of being in state a and passing in a unit time step to state b is equal to the probability of being in b and passing in a unit time step to state a.

Random Walk on a Ring

The state space is the collection of the nodes of a one-dimensional lattice, i.e. $S = \{1, 2, \ldots, i, \ldots, N\}$, with periodic boundary conditions: a random walker moves along the ring by jumping in a unit time step from site i to site $i+1$ with rate r or to site $i-1$ with rate $1-r$, for any i (see Fig. 1.9). This Markov chain is described by the $N \times N$ tridiagonal stochastic matrix

$$W = \begin{pmatrix} 0 & 1-r & 0 & 0 & 0 & \cdots & 0 & 0 & r \\ r & 0 & 1-r & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & r & 0 & 1-r & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & r & 0 & 1-r & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 1-r & 0 & 0 & 0 & 0 & \cdots & 0 & r & 0 \end{pmatrix} \tag{1.92}$$

[10] We are allowed to speak about a stationary state, rather than an asymptotic evolution, because, after a time $t \gg \tau$, we make an exponentially small error in approximating $p_a(t)$ with α.

https://doi.org/10.1017/9781107278974.002 Published online by Cambridge University Press

and Eq. (1.78) reads

$$p_i(t+1) = (1-r)p_{i+1}(t) + rp_{i-1}(t). \tag{1.93}$$

If a stationary state can be eventually attained, all probabilities should be independent of time and (1.93) simplifies to

$$p_i = (1-r)p_{i+1} + rp_{i-1}, \tag{1.94}$$

where we write p_i rather than $p_i(\infty)$. A straightforward solution of this equation is $p_i =$const, independently of i. Due to (1.75) we finally obtain $p_i = 1/N$. This result could be conjectured also on the basis of intuitive arguments: in the long run the random walker will have lost any memory of its initial state and, since due to the lattice symmetry all sites are equivalent, the stationary state will correspond to an equal probability of visiting any site. As in the previous example, the Markov chain is irreducible and all its states are accessible and persistent.

The spectrum of W provides us with more detailed mathematical information about this problem. Eq. (1.87) reads

$$(1-r)w_{k+1} + rw_{k-1} = \lambda w_k, \tag{1.95}$$

which has the solution $w_k = \omega^k$, where ω should be determined imposing a periodic boundary condition, which requires $w_{N+1} = w_1$, i.e. $\omega^N = 1$. We therefore have N independent solutions,

$$\omega_j = \exp\left(\frac{2\pi\, i}{N} j\right), \qquad j = 0, 1, \ldots, N-1. \tag{1.96}$$

The corresponding eigenvalues λ_j are determined by Eq. (1.95), which rewrites

$$\lambda_j = (1-r)\omega_j + r\omega_j^{-1} = \cos\left(\frac{2\pi}{N} j\right) + i(1-2r)\sin\left(\frac{2\pi}{N} j\right), \quad j = 0, 1, \ldots, N-1. \tag{1.97}$$

Notice that for $r = 1/2$ the eigenvalues λ_j are all real, at variance with the case $r \neq 1/2$, where, apart $\lambda_0 = 1$, they are complex.

The corresponding eigenvectors are

$$\mathbf{w}^{(\lambda_j)} = \frac{1}{\sqrt{N}}(1, \omega_j, \omega_j^2, \ldots, \omega_j^{N-1})^T, \tag{1.98}$$

where we have introduced the normalization factor, $1/\sqrt{N}$, such that $\mathbf{w^T w} = 1$. For $j = 0$ we have $\lambda_0 = 1$, independently of r, thus showing that the stationary solution is the same for the asymmetric and symmetric random walk on a ring. On the other hand, for $r \neq 1/2$ there is a bias for the walker to move forward ($r > 1/2$) or backward ($r < 1/2$). In fact, the walker moves with average velocity $v = r - (1-r) = 2r - 1$ and the time it takes to visit all lattice sites is $O(N/v)$.[11]

Making reference to the basic model discussed in Section 1.2.2, we can guess that for $r = 1/2$ the walker performs a diffusive motion, and the time it takes to visit all lattice

[11] This is true for v fixed and diverging N, because for N fixed the ballistic time N/v should be compared with the diffusive time, of order N^2.

https://doi.org/10.1017/9781107278974.002 Published online by Cambridge University Press

sites is $O(N^2)$. In this symmetric case some additional considerations are in order. In fact, if N is an even number and at $t = 0$ the random walker is located at some even site $2j$, i.e. $p_{2j}(0) = 1$, at any odd time step, t_o, $p_{2k}(t_o) = 0$, while at any even time step, t_e, $p_{2k+1}(t_e) = 0$. A similar statement holds if the random walker at $t = 0$ is on an odd site, by exchanging odd with even times and vice versa. This does not occur if N is odd and in this case $\lim_{t\to+\infty} p_i(t) = 1/N$. Conversely, if N is even this limit does not exist, but the following limit exists

$$\lim_{t\to+\infty} \frac{1}{t} \sum_{\tau=1}^{t} p_i(\tau) = \frac{1}{N}. \tag{1.99}$$

This is the ergodic average of $p_i(t)$, that, by this definition, recovers the expected value of the stationary probability. As a final remark, we want to observe that for even N and $r = 1/2$ we have $\lambda_{N/2} = -1$, whose eigenvector is made by an alternation of 1 and -1 (see Eq. (1.98)). The existence of this peculiar eigenvalue and eigenvector can be related to the nonexistence of the limit $\lim_{t\to+\infty} p_i(t) = 1/N$.

Random Walk with Absorbing Barriers

The problem of the random walk with absorbing barriers has been widely investigated, because of its relevance for game theory and in several diffusion problems occurring in material physics. The random walker has an absorbing barrier if there is a state i such that in the corresponding Markov chain $W_{ii} = 1$ and, accordingly, $W_{ji} = 0$ for any $j \neq i$. We can sketch the problem as follows. A random walker moves on a lattice of N sites with fixed ends at states 1 and N, which are absorbing barriers (see Fig. 1.10). The stochastic dynamics on the lattice, apart at sites 1 and N, is the same as that of the random walker on a ring discussed in the previous example. The corresponding stochastic matrix has the same form of (1.92), apart from the first and the last columns, which change into $(1, 0, \ldots, 0, 0)$ and $(0, 0, \ldots, 0, 1)$, respectively, thus yielding

$$W = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & 1-r & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & r & 0 & 1-r & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & r & 0 & 1-r & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 1 \end{pmatrix}. \tag{1.100}$$

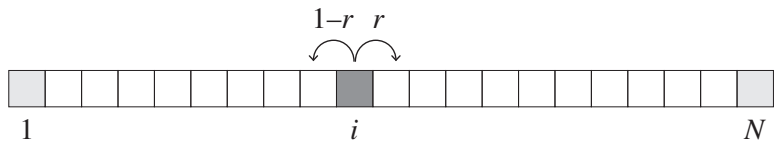

Fig. 1.10 Anisotropic diffusion on a segment with absorbing boundary conditions in $i = 1$ and $i = N$. In the main text we also use the quantity $s = (1 - r)/r$ as a parameter to characterize the asymmetry.

https://doi.org/10.1017/9781107278974.002 Published online by Cambridge University Press

Let us assume that at time 0 the walker is at site j ($1 < j < N$): due to the structure of the stochastic matrix we can state that the walker will be eventually absorbed (with probability 1) either at 1 or N. We can conclude that the Markov chain is irreducible, but, apart from the persistent states 1 and N, all the other states are transient.

We can also ask the more interesting question: which is the probability that a random walker starting at j will reach 1 without being first absorbed in N? In order to answer this question we can introduce the time-independent conditional probability of being trapped at 1 starting at time 0 from site j, $\mathfrak{p}_j$. It obeys the equation ($j = 2, \ldots, N-1$)

$$\begin{aligned} \mathfrak{p}_j &= W_{j+1,j}\ \mathfrak{p}_{j+1} + W_{j-1,j}\ \mathfrak{p}_{j-1} \\ &= r\ \mathfrak{p}_{j+1} + (1-r)\ \mathfrak{p}_{j-1}, \end{aligned} \tag{1.101}$$

which exhibits the same formal structure of (1.94), but with boundary conditions $\mathfrak{p}_1 = 1$ and $\mathfrak{p}_N = 0$. The first equality in (1.101) stems from the basic rule that the joint probability of independent events is the product of the probabilities of each single event: actually, the first (second) term on the right-hand side is the product of the probability rate of passing from j to $j+1$ ($j-1$) in a unit time step with the probability of being first absorbed in 1 starting from $j+1$ ($j-1$).

In order to solve Eq. (1.101) we cannot use the same solution of (1.94), because of the presence of the absorbing barriers that impose the boundary conditions $\mathfrak{p}_1 = 1$ and $\mathfrak{p}_N = 0$. However, we use the same ansatz $\mathfrak{p}_j = \psi^j$. By substituting into (1.101) one obtains

$$\psi^j = r\psi^{j+1} + (1-r)\psi^{j-1}, \tag{1.102}$$

or $1 = r\psi + (1-r)\psi^{-1}$, whose solutions are $\psi_1 = 1$, i.e. $\mathfrak{p}_j^{(1)} = 1$ and $\psi_2 = (1-r)/r \equiv s$, i.e. $\mathfrak{p}_j^{(2)} = s^j$. These equations coincide for $r = 1/2$, i.e. $s = 1$, in which case (1.101) reads

$$\mathfrak{p}_j = \frac{1}{2}(\mathfrak{p}_{j+1} + \mathfrak{p}_{j-1}). \tag{1.103}$$

The solution $\mathfrak{p}_j^{(1)} = 1$ still holds, but the second solution now reads $\mathfrak{p}_j^{(2)} = j$, as one can immediately check.

Since Eq. (1.101) is a finite-difference linear equation in the variable $\mathfrak{p}_j$, its general solution has to be a linear combination of $\mathfrak{p}_j^{(1)}$ and $\mathfrak{p}_j^{(2)}$. For $r \neq 1/2$ we have

$$\mathfrak{p}_j = As^j + B \tag{1.104}$$

with A and B constants to be determined by the boundary conditions,

$$\mathfrak{p}_1 = As + B = 1 \tag{1.105}$$

$$\mathfrak{p}_N = As^N + B = 0, \tag{1.106}$$

whose solution is

$$A = \frac{1}{s(1-s^{N-1})}, \qquad B = -\frac{s^{N-1}}{(1-s^{N-1})}. \tag{1.107}$$

Replacing in Eq. (1.104) we obtain

$$\mathfrak{p}_j = \frac{s^{j-1} - s^{N-1}}{1-s^{N-1}}. \tag{1.108}$$

https://doi.org/10.1017/9781107278974.002 Published online by Cambridge University Press

For $r = 1/2$ ($s = 1$) we have instead

$$\mathfrak{p}_j = Aj + B \tag{1.109}$$

and boundary conditions are $\mathfrak{p}_1 = A + B = 1$ and $\mathfrak{p}_N = AN + B = 0$, i.e.

$$A = -\frac{1}{N-1}, \qquad B = \frac{N}{N-1}. \tag{1.110}$$

Replacing in Eq. (1.109) we obtain

$$\mathfrak{p}_j = \frac{N-j}{N-1}. \tag{1.111}$$

Let us now consider the interesting limit $N \to \infty$. For $r > 1/2$ ($s < 1$) one has $A = 1/s$, $B = 0$ and

$$\mathfrak{p}_j = s^{j-1}, \tag{1.112}$$

i.e. the probability of reaching the absorbing barrier at 1 is exponentially small with j, because the random walker has a bias toward the other absorbing barrier at infinity. For $r < 1/2$, $A = 0$, $B = 1$, and $\mathfrak{p}_j = 1$, i.e. the absorbing barrier at 1 will be eventually reached with probability 1, because the random walker has a bias to move toward 1. The same conclusion can be drawn for $r = 1/2$, because, even if this case is unbiased, any site j is at a finite distance from the absorbing barrier at 1. Therefore, we can conclude

$$\mathfrak{p}_j = 1, \qquad r \leq \frac{1}{2} \quad \text{and} \quad N \to \infty. \tag{1.113}$$

This stochastic process describes a basic problem of game theory. Imagine that you are a player with an initial capital j playing a unit of your capital on "red" or "black" for each run at a roulette table. You would like to know what the probability is that you will eventually reach the situation of being left with one single unit of capital (the absorbing barrier at 1) before reaching a capital $N > j$, that may convince you to stop playing (perhaps because your profit is adequate to your expectation). What we have discussed before allows you to conclude that if the game is biased against the player, i.e. you are in the case $r < 1/2$, you have a probability quite close to 1 of being eventually ruined. For instance, we can compute the probability $\mathfrak{p}_j$, with $j = 500$ units of capital, of eventually reaching a capital of a single unit, before having doubled your capital. Despite the unfavorable bias in playing "red" or "black" corresponding to just $r = 18/37$ and $s = 19/18$ (the zero of the roulette determines the bias favorable to the casino), $\mathfrak{p}_j \approx 1 - O(10^{-12})$; i.e. you have a probability $O(10^{-12})$ of reaching the expected capital of 1000 units, before being ruined. You can realize that this strategy of playing roulette is certainly time consuming, but practically will lead you to ruin.

1.5.3 Ergodic Markov Chains

An important class of Markov chains are the ergodic ones. Let us consider a Markov chain with a finite state space, i.e. $S = \{s_1, s_2, \ldots, s_N\}$: it is ergodic if it is irreducible, nonperiodic and all states are persistent (see Definitions 1.1–1.4 at the end of Section 1.5.1). The main property of ergodic Markov chains is that they determine a unique invariant, i.e. stationary,

https://doi.org/10.1017/9781107278974.002 Published online by Cambridge University Press

probability distribution. This is given by the eigenvector $\mathbf{w}^{(1)}$, which is the solution of the eigenvalue problem with $\lambda = 1$:

$$W\mathbf{w}^{(1)} = \mathbf{w}^{(1)}. \tag{1.114}$$

The spectral property (b), proved in Appendix B, guarantees that such an eigenvector exists. For an ergodic Markov chain it is also unique, because each state is accessible (i.e., the matrix is irreducible) and persistent, i.e. it will be revisited with probability 1 after a finite lapse of time. If these properties do not hold, in general one is faced with several peculiar scenarios, e.g., the reduction of the stochastic matrix into blocks, whose number equals the degeneracy of the eigenvalue $\lambda = 1$.

The stationary probability, $\mathbf{w}^{(1)}$, of an ergodic matrix will be eventually attained exponentially fast on a time scale, τ, independent of the initial conditions, namely

$$\mathbf{p}(t) \approx \mathbf{w}^{(1)} + \mathbf{A}e^{-t/\tau}, \tag{1.115}$$

where $\mathbf{A}$ is a suitable vector with constant components, that sum up to zero to fulfill condition (1.75). In the limit $t \to \infty$, $\mathbf{w}^{(1)}$ is a true dynamical state of the stochastic process, and, accordingly, it has to obey conditions (1.74) and (1.75), i.e. all its components are nonnegative and it is normalized.

Since all states of an ergodic Markov chain are persistent we would like to know the typical return time $\langle T_j \rangle$ to j (i.e. the average value of the return time T_j in the limit $t \to \infty$). The answer to this question is given by the Kac lemma, according to which

$$\langle T_j \rangle = \frac{1}{w_j^{(1)}}. \tag{1.116}$$

A simple argument to explain this result is the following. Let us denote with $T_j^{(n)}$, with $n = 1, 2, \ldots$, the nth return time to j. The total time $\mathcal{T}_M$ needed for the ergodic Markov chain to come back M times to j is given by

$$\mathcal{T}_M = \sum_{n=1}^{M} T_j^{(n)}. \tag{1.117}$$

On the other hand, this means that the fraction of time $\phi_j(\mathcal{T}_M)$ spent by the stochastic process at j in the time $\mathcal{T}_M$ is given by

$$\phi_j(\mathcal{T}_M) = \frac{M}{\mathcal{T}_M}. \tag{1.118}$$

According to the interpretation of $\mathbf{w}^{(1)}$, $\phi_j(\mathcal{T}_M)$ has to converge to $w_j^{(1)}$ in the limit $M \to \infty$ (which is equivalent to the limit $t \to \infty$), and one can write

$$\langle T_j \rangle = \lim_{M\to\infty} \frac{1}{M}\, \mathcal{T}_M = \lim_{M\to\infty} \frac{1}{\phi_j(\mathcal{T}_M)} = \frac{1}{w_j^{(1)}}. \tag{1.119}$$

This relation points out that ergodicity amounts to the equivalence between ensemble and time averages.

Another important result is that the spectral properties of an ergodic Markov chain determine the time scale of convergence to the stationary probability. In fact, according

https://doi.org/10.1017/9781107278974.002 Published online by Cambridge University Press

to the results of Appendix B, the eigenvalues $\lambda^{(j)}$ of the stochastic matrix representing an ergodic Markov chain can be ordered as

$$\lambda^{(1)} = 1 > |\lambda^{(2)}| \geq |\lambda^{(3)}| \geq \cdots \geq |\lambda^{(N)}|. \tag{1.120}$$

Let us come back to Eq. (1.115) to explain its origin. According to the projection theorem of linear algebra, any probability on the state space at time t, $\mathbf{p}(t)$, can be written as a suitable linear combination of the eigenvectors $\mathbf{w}^{(k)}$ of the stochastic matrix, which form an orthonormal basis:[12]

$$\mathbf{p}(t) = \sum_{k=1}^{N} a_k(t)\mathbf{w}^{(k)}, \qquad a_j(t) \in \mathbb{R} \tag{1.121}$$

where $a_k(t) = \mathbf{p}(t) \cdot \mathbf{w}^{(k)}$.

Let us consider the evolution of the ergodic Markov chain from time 0 to time t,

$$\mathbf{p}(t) = W^t\mathbf{p}(0) = W^t \sum_{k=1}^{N} a_k(0)\mathbf{w}^{(k)} = \sum_{k=1}^{N} a_k(0)(\lambda^{(k)})^t\mathbf{w}^{(k)} \equiv \sum_{k=1}^{N} a_k(t)\mathbf{w}^{(k)}. \tag{1.122}$$

Apart from $a_1(0)$, which does not change in time, because $\lambda^{(1)} = 1$, all the other coefficients evolve in time as

$$a_k(t) = a_k(0)(\lambda^{(k)})^t = (\pm)^t a_k(0)e^{-t/\tau_k}, \tag{1.123}$$

where $(\pm)$ is the sign of the eigenvalue $\lambda^{(k)}$ and

$$\tau_k = -\frac{1}{\ln|\lambda^{(k)}|}. \tag{1.124}$$

Therefore, for $k > 1$, $a_k(t)$ eventually vanish exponentially with rate τ_k.

Making use of property (1.120), we can conclude that the overall process of relaxation to equilibrium from a generic initial condition is dominated by the longest time scale, i.e. the one corresponding to the eigenvalue $\lambda^{(2)}$, so that in (1.115) we have

$$\tau = \tau_2 = -\frac{1}{\ln|\lambda^{(2)}|}. \tag{1.125}$$

1.5.4 Master Equation and Detailed Balance

The dynamical rule of the Markov chain, Eq. (1.78), can be rewritten as

$$p_i(t+1) = p_i(t) - p_i(t) + \sum_j p_j(t)W_{ij} = p_i(t) - p_i(t)\sum_j W_{ji} + \sum_j p_j(t)W_{ij}, \tag{1.126}$$

where we have used the condition (1.77). The previous formula can be recast in the form of a master equation,

$$p_i(t+1) - p_i(t) = \sum_{i \neq j} \left(W_{ij}p_j(t) - W_{ji}p_i(t)\right). \tag{1.127}$$

[12] For the sake of simplicity we exclude any degeneracy of the eigenvalues.

https://doi.org/10.1017/9781107278974.002 Published online by Cambridge University Press

This equation tells us that the variation of the probability of being in state s_i in a unit time step can be obtained from the positive contribution of all transition processes from any state s_j to state s_i and from the negative contribution of all transition processes from state s_i to any other state s_j.

This form is particularly useful to define the conditions under which one can obtain a stationary probability, i.e. all p_i are independent of time t: in this case the left-hand side of (1.127) vanishes and the stationarity condition reads

$$\sum_{i \neq j} \left(W_{ij} p_j - W_{ji} p_i \right) = 0, \quad \forall i, \tag{1.128}$$

where the p_js are the components of the stationary probability.

Notice that (1.128) is verified if the following stronger condition holds

$$W_{ij} p_j - W_{ji} p_i = 0, \quad \forall i, j, \tag{1.129}$$

which is called the detailed balance condition. A Markov chain whose stochastic matrix elements obey (1.129) is said to be reversible and it can be shown that it is also ergodic (see Appendix C), with $\mathbf{p} \doteq \mathbf{w}^{(1)}$ representing the so-called equilibrium probability.[13]

We have already discussed in Section 1.5.2 the simple example of a two-state Markov chain obeying detailed balance. As one can immediately realize, detailed balance does not hold in the asymmetric ($r \neq 1/2$) random walk on a ring, because the stationary probability is $w_j^{(1)} = 1/N$ for all j, while $r = W_{j+1,j} \neq W_{j,j+1} = 1 - r$. In this example we are faced with an ergodic Markov chain that is not reversible. Only if the symmetry of the process is restored (i.e. $r = 1/2$) does the detailed balance condition hold and the invariant probability is an equilibrium one, although it is the same for the asymmetric case.

It is important to point out that the detailed balance condition can be used as a benchmark for the absence of an equilibrium probability in an ergodic irreversible Markov chain. Actually, in this case it is enough to find at least a pair of states for which (1.129) does not hold to conclude that the stationary probability $\mathbf{w}^{(1)}$ is not the equilibrium one. Many of the examples discussed in Section 1.5.2 belong to the class of stochastic processes that evolve to a stationary nonequilibrium probability. The difference between equilibrium and nonequilibrium steady states will be reconsidered in Section 3.3.

1.5.5 Monte Carlo Method

The Monte Carlo method is one of the most useful and widely employed applications of stochastic processes. The method aims at solving the problem of the effective statistical sampling of suitable observables by a reversible Markov chain.

More precisely, the problem we are discussing here is the following: how to estimate the equilibrium average of an observable O making use of a suitable, reversible Markov process. As in the rest of Section 1.5, we assume that the configuration space is made

[13] In the case of reversible Markov chains, the stationary probability is more properly called equilibrium probability, because it is equivalent to the concept of thermodynamic equilibrium in statistical mechanics, where the equilibrium probability is determined by the Hamiltonian functional, which engenders a time-reversible dynamics through the Hamilton equations (see Eqs. (2.27a) and (2.27b)).

https://doi.org/10.1017/9781107278974.002 Published online by Cambridge University Press

by N microscopic states, $S = \{s_1, s_2, \ldots, s_{N-1}, s_N\}$, whose equilibrium probabilities are $(p_1, p_2, \ldots, p_N)$. We can write

$$\langle O \rangle_{eq} = \sum_{j=1}^{N} O_j \, p_j \tag{1.130}$$

where $O_j \doteq O(s_j)$ is the value taken by O in state s_j. The problem of computing such an equilibrium average emerges when the number of states N is exceedingly large and we want to avoid sampling O over equilibrium states, whose probabilities may be very small. This is the typical situation of many models of equilibrium statistical mechanics. For instance, the Ising model (see Section 3.2.2), originally introduced to describe the ferromagnetic phase transition, is made by locally interacting binary spin variables, $\sigma = \pm 1$, located at the nodes of a lattice. If the lattice contains L nodes, the total number of states is $N = 2^L$: already for $L = 100$ one has $N \approx 10^{30}$, quite close to an astronomical number, which makes the computation of (1.130) practically unfeasible.[14]

In order to work out the Monte Carlo strategy, we have to assume from the very beginning that we know the equilibrium probabilities $(p_1, p_2, \ldots, p_N)$. In equilibrium statistical mechanics they are given by the Gibbs weight

$$p_i = \frac{e^{-\beta E_i}}{Z}, \tag{1.131}$$

where $\beta = T^{-1}$ is the so-called inverse reduced temperature, E_i is the energy of state s_i, and $Z = \sum_{i=1}^{N} e^{-\beta E_i}$ is the partition function. One could argue that there is something contradictory in the previous statement: in order to know p_i one has to also compute Z, which again is a sum over an astronomical number of states.[15] As we are going to show, the Monte Carlo procedure needs only local information, i.e. the ratios of the equilibrium probabilities $p_i/p_j = e^{-\beta(E_i - E_j)}$, which do not depend on Z.

The following step is how to define a reversible Markov chain that has the p_is as equilibrium probabilities. As we pointed out, a reversible Markov chain is also ergodic, i.e. $\langle O \rangle_{eq}$ can be approximated by a sufficiently long trajectory $(s_1, s_2, \ldots, s_n)$ in the state space of the Markov chain:

$$\frac{1}{n} \sum_{t=1}^{n} O(t) \approx \sum_{j=1}^{N} O_j \, p_j. \tag{1.132}$$

The Monte Carlo procedure is effective, because we can obtain a good estimate of $\langle O \rangle_{eq}$ also if $n \ll N$. In fact, because of ergodicity, we know (see Section 1.5.3) that the fraction of time spent by the stochastic process at state s_i is p_i for time going to infinity: the states s_k that are preferably visited by the trajectory of a stochastic process of finite length n are typically those with $p_k > 1/N$. Accordingly, the Monte Carlo method selects automatically,

[14] The value $L = 100$ is a very small number: in $d = 3$, a cube of only five nodes per side has more than 100 nodes!

[15] In statistical mechanics there are few exactly solvable models, e.g., the Ising model in $d = 1, 2$, where Z can be computed analytically in the thermodynamic limit $N \to \infty$.

https://doi.org/10.1017/9781107278974.002 Published online by Cambridge University Press

after some transient time depending on the initial state, the states corresponding to higher values of the equilibrium probability.

The main point is to estimate how the quality of the approximation depends on n. In fact, any practical implementation of the Monte Carlo procedure into a numerical algorithm is subject to various problems. For instance, the rule adopted to construct the "trajectory" of the stochastic process typically maintains some correlations among the sampled values $O(1), O(2), \ldots$. In particular, we can estimate the correlation time of the Monte Carlo process by measuring how the time autocorrelation function of the observable $O(t)$ decays in time, namely $\langle O(t)O(0)\rangle - \langle O(t)\rangle^2 \sim \exp(-t/\tau)$. If the process lasts over a time n we can assume that the number of statistically independent samples is of the order n/τ. According to the central limit theorem (see Appendix A), the error we make in the approximation (1.132) is $O(\sqrt{\tau/n})$: in practice, the Monte Carlo procedure is effective only if $n \gg \tau$.

Now, let us describe how the reversible Markov chain we are looking for can be explicitly defined for the Ising model, according to the most popular algorithm, called Metropolis. The steps of the algorithm are as follows:

1. We select the initial state by sampling a random spin configuration on the lattice.
2. We select at random with uniform probability $1/L$ a spin variable, say the one at node n_k, and compute its local interaction energy with the nearby spins, E_{n_k}.
3. We flip the spin variable, i.e. $\sigma_{n_k} \to -\sigma_{n_k}$, and we compute its local interaction energy with the nearby spins in this new configuration, E'_{n_k}.
4. If $E'_{n_k} < E_{n_k}$, the next state in the Markov process is the flipped configuration.
5. If $E'_{n_k} \geq E_{n_k}$, we "accept" the new configuration with probability

$$p^* = \exp(-\beta(E'_{n_k} - E_{n_k})).$$

   This means we extract a random number r, uniformly distributed in the interval $[0\,,\,1]$: if $r < p^*$, the next state is the flipped configuration; otherwise, it is equal to the old state.
6. We iterate the process starting again from step 2.

In general terms, in equilibrium statistical mechanics the detailed balance condition is satisfied if

$$W_{ji}e^{-\beta E_i} = W_{ij}e^{-\beta E_j}, \tag{1.133}$$

where we have made explicit the equilibrium probability of occupancy of a microscopic state, given by the Gibbs weight. Therefore, the transition rates must satisfy the relation

$$\frac{W_{ji}}{W_{ij}} = e^{-\beta(E_j - E_i)}. \tag{1.134}$$

The Metropolis algorithm is defined by

$$W_{ji} = \min\left\{1, e^{-\beta(E_j - E_i)}\right\}, \tag{1.135}$$

therefore it satisfies Eq. (1.134).

https://doi.org/10.1017/9781107278974.002 Published online by Cambridge University Press

1.6 Continuous Time Stochastic Processes

The stochastic processes described in Section 1.5 are defined on a discrete state space and in discrete time. On the other hand, many physical processes are better represented in continuous space and time. For instance, if we want to describe a fluid it is simpler to attribute continuous position $\mathbf{x}(t)$ and momentum $\mathbf{p}(t)$ coordinates to each fluid element (i.e. an infinitesimal portion of the fluid), rather than considering the same quantities for each particle in the fluid. A continuous time variable, on its side, allows us to take advantage of the powerful machinery of differential equations. By the way, we should keep in mind that in the previous sections we have first introduced the basic, discrete model of a diffusive random walk (Section 1.2.2), then we have passed to the continuous formulation of Brownian motion in terms of the Langevin equation (Section 1.4.1) and of the Fokker–Planck equation (Section 1.4.2). Here below we provide a continuous time version of stochastic processes, while keeping the discrete nature of the state space.

As a first step we derive the time continuous version of the master equation (1.127) making use of the Chapman–Kolmogorov relation (1.86) in continuous time t,[16]

$$W_{ij}^{t+\Delta t} = \sum_k W_{ik}^{\Delta t} W_{kj}^{t}. \tag{1.136}$$

If Δt is an infinitesimal increment of time, we can imagine that $W_{ik}^{\Delta t}$ vanishes with Δt if $i \neq k$ and that $W_{ii}^{\Delta t} \to 1$ in the same limit. We can therefore write, up to terms $O(\Delta t^2)$,

$$W_{ik}^{\Delta t} = \begin{cases} \mathcal{R}_{ik}\Delta t, & k \neq i \\ 1 - \mathcal{R}_{ii}\Delta t, & k = i. \end{cases} \tag{1.137}$$

The rates $\mathcal{R}_{ik}$ are not independent, because the normalization condition (1.77),

$$1 = \sum_i W_{ik}^{\Delta t} = 1 + \Delta t \left(-\mathcal{R}_{kk} + \sum_{i \neq k} \mathcal{R}_{ik} \right), \tag{1.138}$$

implies that

$$\mathcal{R}_{kk} = \sum_{i \neq k} \mathcal{R}_{ik}. \tag{1.139}$$

By substituting (1.137) into (1.136) and making use of (1.139) we can write

$$\begin{aligned} W_{ij}^{t+\Delta t} - W_{ij}^{t} &= \Delta t \left(\sum_{k \neq i} W_{kj}^{t} \mathcal{R}_{ik} - W_{ij}^{t} \mathcal{R}_{ii} \right) \\ &= \Delta t \left(\sum_{k \neq i} W_{kj}^{t} \mathcal{R}_{ik} - \sum_{k \neq i} W_{ij}^{t} \mathcal{R}_{ki} \right). \end{aligned} \tag{1.140}$$

[16] With respect to the case of discrete time, here W^t is not the elementary stochastic matrix raised to the power t. Rather, W^t should be understood as a transition matrix depending on the continuous parameter t.

https://doi.org/10.1017/9781107278974.002 Published online by Cambridge University Press

By dividing both members by Δt and performing the limit $\Delta t \to 0$ we obtain the continuous time master equation

$$\frac{d\,W_{ij}^t}{d\,t} = \sum_{k\neq i}(W_{kj}^t \mathcal{R}_{ik} - W_{ij}^t \mathcal{R}_{ki}). \tag{1.141}$$

We can now write it in terms of the probability $p_i(t)$ the system is in state i at time t, using the relation between $p_i(t)$ and W_{ij}^t,

$$p_i(t) = \sum_j W_{ij}^t p_j(0). \tag{1.142}$$

In fact, it is sufficient to derive Eq. (1.142) with respect to time,

$$\frac{d\,p_i(t)}{d\,t} = \sum_j \frac{d\,W_{ij}^t}{d\,t} p_j(0) \tag{1.143}$$

and replacing (1.141) here above, we obtain

$$\frac{d\,p_i(t)}{d\,t} = \sum_j \sum_{k\neq i}(\mathcal{R}_{ik} W_{kj}^t - \mathcal{R}_{ki} W_{ij}^t) p_j(0). \tag{1.144}$$

Finally, we can write the continuous time master equation

$$\frac{d\,p_i(t)}{d\,t} = \sum_{k\neq i}(\mathcal{R}_{ik} p_k(t) - \mathcal{R}_{ki} p_i(t)). \tag{1.145}$$

Similar to its time discrete version (1.127), this equation tells us that the variation in time of the probability of being in state s_i is obtained from the positive contribution of all transition processes from any state s_k to state s_i and from the negative contribution of all transition processes from state s_i to any other state s_k.

1.6.1 Stochastic Differential Equations

In Section 1.4.1 we introduced the Langevin equation for a Brownian particle (see Eq. (1.42)), on the basis of some heuristic arguments. The main conceptual weakness of this approach is the assumption that the components of the stochastic force, $\eta_i(t)$, acting on the Brownian particle are completely uncorrelated:

$$\langle \eta_i(t)\eta_j(t')\rangle = \Gamma \delta_{i,j}\delta(t-t'), \tag{1.146}$$

where Γ is given by the fluctuation–dissipation relation (1.48). In a physical perspective this is justified by assuming that the components of the stochastic force are independent of each other and their correlation time is practically negligible with respect to the typical time scale, $t_d = m/\tilde{\gamma} = 1/\gamma$, associated with the dynamics of the deterministic part, where m is the mass of the Brownian particle and $\tilde{\gamma}$ is the viscous friction.

Despite this assumption seeming quite plausible in physical phenomena (e.g., Brownian motion) usually modeled by stochastic processes, the mathematical structure of (1.42) is quite weird: the deterministic part is made of differentiable functions with respect to time,

https://doi.org/10.1017/9781107278974.002 Published online by Cambridge University Press

while $\eta_i(t)$ is intrinsically discontinuous at any time, due to (1.146). A way out of this problem is the introduction of Wiener processes, which can be considered the continuous time version of a random walk. The basic idea is quite simple: we use the integrated stochastic force to define a new stochastic process that can be made continuous in time,

$$W_i(t) = \int_0^t dt' \eta_i(t'). \tag{1.147}$$

The statistical average introduced in Section 1.4.1 has to be thought of as the average over the realizations of the stochastic process. If $\langle \eta_i(t) \rangle = 0$, then also $\langle W_i(t) \rangle = 0$, while (1.146) yields

$$\langle W_i(t) W_j(t') \rangle = \delta_{ij} \Gamma \min\{t, t'\}. \tag{1.148}$$

In particular,

$$\langle W_i^2(t) \rangle = \Gamma t, \tag{1.149}$$

meaning that the average squared amplitude of the Wiener process diffuses in time.

The Wiener process defined in (1.147) obeys the following properties:

- $W_i(t)$ is a stochastic process continuous in time and with zero average, $\langle W_i(t) \rangle = 0$.
- For any $t_1 < t_2 < t_3$ the increments $(W_i(t_2) - W_i(t_1))$ and $(W_i(t_3) - W_i(t_2))$ are independent quantities, following the same distribution.
- For any $t_1 < t_2$ the probability distribution of the increments $(W_i(t_2) - W_i(t_1))$ is a Gaussian with zero average and variance $\Gamma(t_2 - t_1)$, which is a consequence of the central limit theorem; see Appendix A.

Notice that $W_i(t)$ is not differentiable, meaning that we cannot define its time derivative, but we can define its infinitesimal increment $dW_i(t)$ for an arbitrary small time interval dt, namely,

$$dW_i(t) = W_i(t + dt) - W_i(t) = \int_t^{t+dt} dt' \eta_i(t'), \tag{1.150}$$

which implies $\langle dW_i(t) \rangle = 0$ and $\langle (dW_i(t))^2 \rangle = \Gamma dt$.

In order to simplify the notation, the infinitesimal increment of the Wiener process is usually redefined as

$$dW(t) \to \sqrt{\Gamma} \Omega(t) \sqrt{dt}, \tag{1.151}$$

where $\Omega(t)$ is a stochastic process with zero average and unit variance and $\sqrt{\Gamma}$ is extracted by the stochastic force. This relation is quite useful, because it attributes to the amplitude of the infinitesimal increment of the Wiener process a physical scale, given by the square root of the infinitesimal time increment dt.

With these ingredients we can write the general form of a well-defined stochastic differential equation,

$$dX_i(t) = a_i(X(t), t)dt + b_i(X(t), t)dW_i(t), \quad i = 1, 2 \ldots, n, \tag{1.152}$$

where $X_i(t)$, $a_i(X(t), t)$, and $b_i(X(t), t)$ are the components of generic vectors of functions in $\mathbb{R}^n$: each of these functions is differentiable with respect to its argument. In a physical

https://doi.org/10.1017/9781107278974.002 Published online by Cambridge University Press

perspective $a_i(X(t), t)$ is a generalized drift coefficient, while $b_i(X(t), t)$ is related to a generalized diffusion coefficient. For instance, the Langevin equation of the Brownian particle (see Section 1.4.1) is defined in $\mathbb{R}^3$ with $X_i(t) = v_i(t)$ (i.e., the ith component of the velocity of the Brownian particle), while $a_i(X(t), t) = -\gamma v_i(t)$ and $b_i(X(t), t)$ is proportional to $\sqrt{2\gamma T/m}$.

Let us further simplify the notation by overlooking the subscript i

$$dX(t) = a(X(t), t)\ dt + b(X(t), t)\ dW(t), \tag{1.153}$$

and integrate this general stochastic differential equation, continuous in both space and time. On a formal ground this task can be immediately accomplished, writing the solution of (1.153) as

$$X(t) = X(0) + \int_0^t a(X(t'), t')\ dt' + \int_0^t b(X(t'), t')\ dW(t'). \tag{1.154}$$

If we want to extract any useful information from this formal solution we have to perform its statistical average with respect to the Wiener process, as we have done for Brownian motion when passing from Eq. (1.43) to Eqs. (1.44) and (1.45) by averaging with respect to the components of the stochastic force $\eta_i(t)$.

Here, we have to face the further problem that the last integral is not uniquely defined when averaged over the stochastic process. In fact, according to basic mathematical analysis, the integral of a standard (Riemann) function $f(t)$ can be estimated by the Euler approximations

$$\int_0^t f(t')dt' \approx \sum_{i=0}^{N-1} f(t_i)\Delta t_i \approx \sum_{i=0}^{N-1} f(t_{i+1})\Delta t_i \tag{1.155}$$

where the support of t is subdivided into an ordered sequence of $N+1$ sampling times $(t_0 = 0, t_1, t_2 \ldots, t_{N-1}, t_N = t)$, with $\Delta t_i = t_{i+1} - t_i$. For the sake of simplicity let us assume a uniform sampling, i.e. $\Delta t_i = t/N,\ \forall i$. A theorem guarantees that in the limit $\Delta t \to 0$ both Euler approximations in (1.155) converge to the integral. Such a theorem does not hold for a Wiener process. The analogs of the Euler approximations for the stochastic integral are:

$$\mathcal{I}_I = \sum_{i=0}^{N-1} f(t_i)\Delta W(t_i), \tag{1.156}$$

$$\mathcal{I}_S = \sum_{i=0}^{N-1} f(t_{i+1})\Delta W(t_i). \tag{1.157}$$

Equations (1.156) and (1.157) are called the Itô and the Stratonovich[17] discretizations, respectively. For very small $\Delta W(t_i)$ (1.157) can be approximated as

$$\mathcal{I}_S = \sum_{i=0}^{N-1} [f(t_i) + \alpha \Delta W'(t_i) + O(\Delta W'^2(t_i))]\Delta W(t_i), \tag{1.158}$$

[17] After Kiyoshi Itô and Ruslan L. Stratonovich.

https://doi.org/10.1017/9781107278974.002 Published online by Cambridge University Press

where $\Delta W'$ is, in general, a different realization of the Wiener process with respect to ΔW, and $\alpha = \frac{\delta f}{\delta W'}(t_i)$ is the functional derivative of $f(t)$ with respect to W' at t_i (see Appendix L.1).

By performing the limit $\Delta W \to 0$ and averaging over the Wiener process, we obtain

$$\left\langle \lim_{\Delta W \to 0} \mathcal{I}_I \right\rangle = \left\langle \int_0^t f(t')\ dW(t') \right\rangle = 0 \tag{1.159}$$

while, using (1.158),

$$\left\langle \lim_{\Delta W \to 0} \mathcal{I}_S \right\rangle = \left\langle \int_0^t f(t')\ dW(t') \right\rangle + \int_0^t \int_0^t \alpha(t'') \left\langle dW(t') dW(t'') \right\rangle = \int_0^t \alpha(t') dt' \tag{1.160}$$

where we have used the relation $\langle dW(t) dW(t') \rangle = \delta(t - t')\ dt$, which is a consequence of Eq. (1.150).

Notice that, at variance with (1.159), the last expression in (1.160) does not vanish: the Itô and the Stratonovich formulations are not equivalent. Anyway, when dealing with models based on stochastic equations, we have to choose and explicitly declare which formulation we are adopting to integrate the equations. If there are no specific reasons, in many applications it is advisable using the Itô formulation, because of property (1.159), which makes calculations simpler. However, when noise is multiplicative rather than additive, the two formulations are known to differ and the Stratonovich one is preferable. A simple (but a bit artificial) way to obtain a model with multiplicative noise is to consider the Langevin equation for the Brownian particle, with the velocity variable v replaced by its kinetic energy $E = mv^2/2$, so that Eq. (1.41) becomes

$$\frac{dE}{dt} = -2\gamma E + \sqrt{\frac{2E}{m}} \tilde{\eta}(t). \tag{1.161}$$

In order to verify Eq. (1.49), it is necessary to use the Stratonovich prescription.

1.6.2 General Fokker–Planck Equation

We can use the general stochastic differential equation (1.153) to derive the general Fokker–Planck equation. Let us consider a function $f(X(t), t)$ that is at least twice differentiable with respect to X. We can write its Taylor series expansion as

$$df = \frac{\partial f}{\partial t} dt + \frac{\partial f}{\partial X} dX + \frac{1}{2} \frac{\partial^2 f}{\partial X^2} dX^2 + O(dt^2, dX^3). \tag{1.162}$$

Now we also assume that $X(t)$ obeys the general stochastic differential equation (1.153), which means that $dX = a dt + b dW$ (the arguments of the functions $a(X(t), t)$ and $b(X(t), t)$ have been overlooked). By substituting into (1.162) and discarding higher-order terms, we obtain the so-called Itô formula,

$$df = \left(\frac{\partial f}{\partial t} + a \frac{\partial f}{\partial X} + \frac{1}{2} b^2 \frac{\partial^2 f}{\partial X^2} \right) dt + b \frac{\partial f}{\partial X} dW(t), \tag{1.163}$$

where we have assumed $(dW)^2 = dt$. Formally, this is not correct, because $\langle (dW)^2 \rangle = dt$; however, we should consider that, in order to find the solution of (1.163), we have to

https://doi.org/10.1017/9781107278974.002 Published online by Cambridge University Press

integrate and average over the Wiener process. Accordingly, the previous assumption is practically correct, because of the averaging operation. Obviously, we cannot replace $dW(t)$ with its average value.

The important result is that this new stochastic equation is formally equivalent to (1.153). We can conclude that any function $f(X(t))$, at least twice differentiable, obeys the same kind of stochastic differential equation obeyed by $X(t)$. The Itô formula is quite useful, because it allows us to transform a stochastic differential equation into another one, whose solution is known by a suitable guess of the function f.

Now, let us assume that $f(t)$ does not depend explicitly on t, i.e. $\frac{\partial f}{\partial t} = 0$; we average (1.163) over the Wiener process and obtain

$$\frac{d}{d\,t}\,\langle f(X(t))\rangle = \left\langle a\,\frac{\partial f}{\partial\,X} + \frac{1}{2}\,b^2\,\frac{\partial^2 f}{\partial\,X^2}\right\rangle, \tag{1.164}$$

where we have used the condition $\langle dW(t)\rangle = 0$.

In order to specify the averaging operation we introduce explicitly the probability density $P(X,t)$, defined on the state space $\mathcal{S}$ of $X(t)$, where this stochastic variable takes the value X at time t. Given a function $f(X)$, its expectation value is

$$\langle f(X)\rangle = \int_{\mathcal{S}} d\,X\,P(X,t)\,f(X), \tag{1.165}$$

and we can rewrite (1.164) as

$$\begin{aligned}\frac{d}{dt}\langle f(X)\rangle &= \int_{\mathcal{S}} d\,X f(X)\frac{\partial\,P(X,t)}{\partial\,t}\\ &= \int_{\mathcal{S}} d\,X P(X,t) a\frac{\partial f}{\partial\,X} + \frac{1}{2}\int_{\mathcal{S}} d\,X P(X,t) b^2\frac{\partial^2 f}{\partial\,X^2}.\end{aligned} \tag{1.166}$$

Integrating by parts both integrals in the last line we obtain

$$\begin{aligned}\int_{\mathcal{S}} d\,X f(X)\frac{\partial\,P(X,t)}{\partial\,t} = &-\int_{\mathcal{S}} d\,X f(X)\,\frac{\partial\,(a(X,t)P(X,t))}{\partial\,X}\\ &+\frac{1}{2}\int_{\mathcal{S}} d\,X f(X)\,\frac{\partial^2(b^2(X,t)\,P(X,t))}{\partial\,X^2},\end{aligned} \tag{1.167}$$

where it has been assumed that the probability density $P(X,t)$ vanishes at the boundary of $\mathcal{S}$. The analogous assumption for the Brownian particle is that the distribution function of its position in space vanishes at infinity.

Since (1.167) holds for an arbitrary function $f(X)$, we can write the general Fokker–Planck equation for the probability density $P(X,t)$ as

$$\frac{\partial\,P(X,t)}{\partial\,t} = -\frac{\partial}{\partial\,X}\,(a(X,t)P(X,t)) + \frac{1}{2}\,\frac{\partial^2}{\partial\,X^2}\,(b^2(X,t)\,P(X,t)). \tag{1.168}$$

We point out that the previous equation has been obtained from the stochastic differential equation (1.153), which describes the evolution of each component of a vector of variables (e.g., coordinates) $X_i(t)$; see Eq. (1.152). In practice, in Section 1.6.1 we have assumed that all $X_i(t)$ are independent stochastic processes, so that the general Fokker–Plank equation has the same form for any $P(X_i,t)$. On the other hand, in principle one should deal with the

https://doi.org/10.1017/9781107278974.002 Published online by Cambridge University Press

general case, where $a_i(X,t)$ and $b_i(X,t)$ are different for different values of i. By assuming spatial isotropy for these physical quantities, one can overlook any dependence on i in the general Fokker–Plank equation. In practice, with these assumptions, all the components of the vector variable $X_i(t)$ are described by a single general Fokker–Plank equation, where $X(t)$ is a scalar quantity.

In a mathematical perspective, finding an explicit solution of (1.168) depends on the functional forms of the generalized drift coefficient $a(X,t)$ and of the generalized diffusion coefficient[18] $b^2/2$. On the other hand, as discussed in the following section, interesting physical examples correspond to simple forms of these quantities.

1.6.3 Physical Applications of the Fokker–Planck Equation

Stationary Diffusion with Absorbing Barriers

Eq. (1.168) can be expressed in the form of a conservation law, or continuity equation,

$$\frac{\partial P(X,t)}{\partial t} + \frac{\partial J(X,t)}{\partial X} = 0, \tag{1.169}$$

where

$$J(X,t) = a(X,t)P(X,t) - \frac{1}{2}\frac{\partial}{\partial X}(b^2(X,t)\,P(X,t)) \tag{1.170}$$

can be read as a probability current.

Since $X \in \mathbb{R}$, we can consider the interval $I = [X_1, X_2]$ and define the probability, $\mathcal{P}(t)$, that at time t the stochastic process described by (1.169) is in I,

$$\mathcal{P}(t) = \int_I P(X,t)dX. \tag{1.171}$$

By integrating both sides of (1.169) over the interval I one obtains

$$\frac{\partial \mathcal{P}(t)}{\partial t} = J(X_1,t) - J(X_2,t). \tag{1.172}$$

The current is positive if the probability flows from smaller to larger values of X, and various boundary conditions can be imposed on I. For instance, the condition of reflecting barriers, i.e. no flux of probability through the boundaries of I, amounts to $J(X_1,t) = J(X_2,t) = 0$ at any time t. Accordingly, the probability of finding the walker inside I is conserved, as a straightforward consequence of (1.172). Conversely, the condition of absorbing barriers implies that the walker reaching X_1 or X_2 will never come back to I, i.e. $P(X_1,t) = P(X_2,t) = 0$ at any time t. When dealing with stochastic systems defined on a finite interval (a typical situation of numerical simulations), it may be useful to impose periodic boundary conditions that correspond to $P(X_1,t) = P(X_2,t)$ and $J(X_1,t) = J(X_2,t)$.

[18] The mathematical problem of establishing conditions for the existence and the uniqueness of the solution of Eq. (1.168) demands specific assumptions of the space and time dependence of $a(X,t)$ and $b(X,t)$. They must be differentiable functions with respect to their arguments and they must be compatible with the physical assumption that $P(X,t)$ must rapidly vanish in the limit $X \to \pm\infty$ in such a way that the normalization condition $\int_{\mathbb{R}} P(X,t)dX = 1$ holds at any time t.

https://doi.org/10.1017/9781107278974.002 Published online by Cambridge University Press

In this case the probability $\mathcal{P}$ is conserved not because the flux vanishes at the borders, but because the inner and outer fluxes compensate.

Stationary solutions $P^*(X)$ of Eq. (1.169) must fulfill the condition[19]

$$\frac{d}{dX}\left(a(X)P^*(X)\right) - \frac{1}{2}\frac{d^2}{dX^2}\left(b^2(X)\,P^*(X)\right) = 0, \tag{1.173}$$

where we have to assume that also the functions $a(X)$ and $b(X)$ do not depend on t. For instance, in the purely diffusive case, i.e. $a(X) = 0$, $b^2(X) = 2D$, the general solution of (1.173) is

$$P^*(X) = C_1 X + C_2, \tag{1.174}$$

where C_1 and C_2 are constants to be determined by boundary and normalization conditions.

In the case of reflecting barriers at X_1 and X_2 in the interval I, there is no flux through the boundaries, so that the stationary solution must correspond to no flux in I, i.e.,

$$J^*(X) = -D\,\frac{dP^*(X)}{dX} = 0. \tag{1.175}$$

This condition implies $C_1 = 0$ and $P^*(X) = C_2 = |X_2 - X_1|^{-1}$, where the last expression is a straightforward consequence of the normalization condition $\int_{\mathbb{R}} P^*(X)dX = 1$. We can conclude that in this case the stationary probability of finding a walker in I is a constant, given by the inverse of the length of I.

In the case of absorbing barriers we obtain the trivial solution $P^*(X) = 0$ for $X \in I$, because for $t \to \infty$ the walker will eventually reach one of the absorbing barriers and disappear.

A more interesting case with absorbing barriers can be analyzed by assuming that at each time unit a new walker starts at site $X_0 \in I$; i.e. we consider a stationary situation where a constant flux of walkers is injected in I at X_0 (see Fig. 1.11). Due to the presence of absorbing barriers, the stationary solution of (1.173) with a source generating walkers at unit rate in X_0 is

$$P^*(X) = \begin{cases} C_1(X - X_1), & \text{for } X < X_0 \\ C_2(X_2 - X), & \text{for } X > X_0, \end{cases} \tag{1.176}$$

which fulfills the condition of absorbing barriers $P^*(X_1) = P^*(X_2) = 0$. Moreover, we have to impose the continuity of $P^*(X)$ at the source point X_0. This condition yields the relation

$$C_1(X_0 - X_1) = C_2(X_2 - X_0). \tag{1.177}$$

Because of the presence of the flux source, Eq. (1.169) should be written, more correctly, as

$$\partial_t P + \partial_x J = F\delta(X - X_0). \tag{1.178}$$

[19] Partial derivatives turn into standard derivatives, because the stationary solution is, by definition, independent of t.

https://doi.org/10.1017/9781107278974.002 Published online by Cambridge University Press

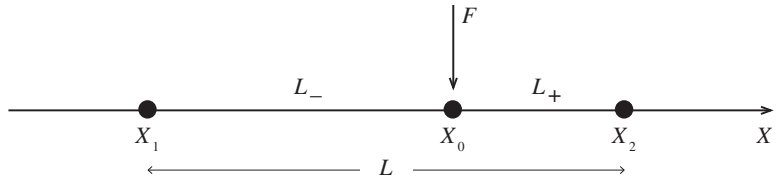

Fig. 1.11 Isotropic diffusion in a segment (X_1, X_2) of length L with absorbing boundaries at the interval extrema and a continuous injection of particles in $X = X_0$. The system relaxes to a steady state, where the stationary distribution of particle density varies linearly in each segment, because $P^*(X)$ is the solution of the equation $d^2P^*/dX^2 = 0$. Absorbing boundaries imply $P^*(X_1) = P^*(X_2) = 0$. In $X = X_0$, $P^*(X)$ is continuous, but its derivative has a jump depending on the flux F (see Eq. 1.178).

Therefore, the stationary solution must satisfy the relation

$$F = \int_{X_0^-}^{X_0^+} dX \partial_x J \tag{1.179}$$

$$= -D\left[\left.\frac{dP^*}{dX}\right|_{X_0^+} - \left.\frac{dP^*}{dX}\right|_{X_0^-}\right] \tag{1.180}$$

$$= D(C_2 + C_1). \tag{1.181}$$

The last equation, along with (1.177) gives

$$C_1 = \frac{F}{D}\frac{L_+}{L}, \qquad C_2 = \frac{F}{D}\frac{L_-}{L}, \tag{1.182}$$

where we have used the notations $L = X_2 - X_1$ for the total length of the interval and $L_- = X_0 - X_1$, $L_+ = X_2 - X_0$, for the left and right parts of the interval; see Fig. 1.11. The probability distributions and the currents are

$$P^*(X) = \begin{cases} \dfrac{F}{D}\dfrac{L_+}{L}(X - X_1) \\ \dfrac{F}{D}\dfrac{L_-}{L}(X_2 - X), \end{cases} \qquad J^*(X) = \begin{cases} -F\dfrac{L_+}{L} \equiv -J_-, & \text{for } X < X_0 \\ F\dfrac{L_-}{L} \equiv J_+, & \text{for } X > X_0, \end{cases} \tag{1.183}$$

where J_- and J_+ are the net fluxes of particles flowing in X_1 and X_2, respectively (and $J_- + J_+ = F$, because of matter conservation). Therefore, a particle deposited in X_0 has a probability

$$\Pi(X_1|X_0) = \frac{J_-}{F} = \frac{L_+}{L} = \frac{X_2 - X_0}{X_2 - X_1} \tag{1.184}$$

to be absorbed in X_1.

The same walker has an average exit time from I, $\mathcal{T}(X_0)_I$, which is equal to the number of walkers, $\mathcal{N}_w$, in the interval I divided by the total current flowing out (equal to F; see above). Since

$$\mathcal{N}_w = \int_{X_1}^{X_2} P^*(X)dX = \frac{F}{2D}L_-L_+, \tag{1.185}$$

https://doi.org/10.1017/9781107278974.002 Published online by Cambridge University Press

we have

$$\mathcal{T}(X_0)_I = \frac{\mathcal{N}_w}{F} = \frac{1}{2D}(X_0 - X_1)(X_2 - X_0). \tag{1.186}$$

The reader can easily realize that we have just reconsidered in the continuous approach of the Fokker–Planck equation the problem of the random walk with absorbing barriers discussed in Section 1.5.2.

A Stochastic Particle Subject to an Elastic Force

We want to describe in the Itô formulation the motion of a stochastic particle in the presence of a conservative force and thermal fluctuations. In the case of an elastic force and of a constant diffusion coefficient, $b(X(t), t) = \sqrt{2D}$, this is known as the Ornstein–Uhlenbeck process,

$$dX(t) = -kX(t)\ dt + \sqrt{2D}\ dW(t), \tag{1.187}$$

where k is the Hook elastic constant divided by the friction coefficient $\tilde{\gamma}$.[20] This equation can be solved by introducing the function $Y(t) = X(t)\, e^{kt}$ that obeys the stochastic equation

$$dY(t) = e^{kt}(dX(t) + kXdt) = \sqrt{2D}\, e^{kt}\, dW(t), \tag{1.188}$$

which can be integrated between time 0 and time t, thus yielding

$$Y(t) = Y(0) + \sqrt{2D} \int_0^t e^{ks}\, dW(s). \tag{1.189}$$

Coming back to the variable $X(t)$ this equation becomes

$$X(t) = X(0)\, e^{-kt} + \sqrt{2D} \int_0^t e^{-k(t-s)}\ d\,W(s), \tag{1.190}$$

where $X(0)$ is the initial position of the particle. By averaging over the Wiener process we finally obtain

$$\langle X(t) \rangle = X(0)\, e^{-kt} \tag{1.191}$$

and

$$\langle X^2(t) \rangle = \langle X(t) \rangle^2 + 2X(0)e^{-kt}\sqrt{2D} \int_0^t e^{-k(t-s)} \langle dW(s) \rangle \tag{1.192}$$

$$+ 2D \int_0^t e^{-k(t-s)} \int_0^t e^{-k(t-s')} \langle dW(s) dW(s') \rangle. \tag{1.193}$$

Using the relations $\langle dW(s) \rangle = 0$ and $\langle dW(s) dW(s') \rangle = \delta(s - s') ds$, we find

$$\langle X^2(t) \rangle - \langle X(t) \rangle^2 = \frac{D}{k} \left(1 - e^{-2kt} \right). \tag{1.194}$$

[20] The physical interpretation of this equation is quite different from Newtonian mechanics: the presence of the Wiener process on the right-hand side points out that the overall process is the result of a balance among a friction term on the left-hand side (a force proportional to a velocity) and a conservative elastic force in the presence of thermal fluctuations. The inertia term, i.e. the one proportional to the second derivative with respect to time of $X(t)$, is negligible on time scales larger than t_d.

https://doi.org/10.1017/9781107278974.002 Published online by Cambridge University Press

In the limit $t \to +\infty$ the average displacement vanishes, irrespectively of the initial condition, while the variance of the stochastic process converges to D/k. The stochastic particle diffuses around the origin and its average squared displacement is distributed according to a Gaussian with zero average, while its variance is inversely proportional to the Hook constant k and proportional to the amplitude of fluctuations, i.e. to the diffusion constant D. If we want to recover the result for the Brownian particle, we should take the limit $k \to 0$ in Eq. (1.194) before the limit $t \to \infty$, obtaining $\langle X^2(t)\rangle - \langle X(t)\rangle^2 = 2Dt$, as expected.

The Fokker–Planck equation of the Ornstein–Uhlenbeck process is

$$\frac{\partial P(X,t)}{\partial t} = k\frac{\partial}{\partial X}(XP(X,t)) + D\frac{\partial^2}{\partial X^2}P(X,t). \tag{1.195}$$

This equation can be solved via Fourier transformation and then using the method of characteristics. Here we limit to give the solution,

$$P(X,t) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{(X-\langle X\rangle)^2}{2\sigma^2}\right) \tag{1.196}$$

with

$$\langle X\rangle = X_0 \exp(-kt) \tag{1.197}$$

$$\sigma^2 = \frac{D}{k}[1-\exp(-2kt)]. \tag{1.198}$$

As discussed in Appendix D.2, $\langle X\rangle$ is the average value and σ^2 is the variance. In the limit $t \to \infty$, if $k \neq 0$ the two momenta become

$$\langle X\rangle = 0 \tag{1.199}$$

$$\langle X^2\rangle = \sigma^2 = \frac{D}{k}. \tag{1.200}$$

We can conclude that the solution of the Ornstein–Uhlenbeck process is a time-dependent Gaussian distribution, whose center and variance, in the limit $t \to +\infty$, tend to the origin and to a constant, respectively. The asymptotic constancy of σ^2 is the result of the balance between diffusion (which would tend to increase fluctuations) and the elastic restoring force (which suppresses fluctuations).

The Fokker–Planck Equation in the Presence of a Mechanical Force

Let us consider the Fokker–Planck equation for a particle subject to an external mechanical force $\mathbf{F}(\mathbf{x})$ generated by a conservative potential $U(\mathbf{x})$, i.e.,

$$\mathbf{F}(\mathbf{x}) = -\nabla U(\mathbf{x}). \tag{1.201}$$

Making use of the general formalism introduced in this section (see Eqs. (1.169) and (1.170)), we can write

$$\frac{\partial P(\mathbf{x},t)}{\partial t} = -\nabla\cdot\mathbf{J}, \tag{1.202}$$

https://doi.org/10.1017/9781107278974.002 Published online by Cambridge University Press

where

$$\mathbf{J} = \frac{\mathbf{F}}{\tilde{\gamma}} P(\mathbf{x}, t) - D \nabla P(\mathbf{x}, t). \tag{1.203}$$

Here D is the usual diffusion coefficient and the parameter $\tilde{\gamma}$ is the same phenomenological quantity introduced in Section 1.4.1.

Here we avoid reporting the derivation of the solution $P(\mathbf{x}, t)$ of Eq. (1.202) and we limit our analysis to obtaining the explicit expression for the equilibrium probability distribution $P^*(\mathbf{x})$. This problem can be easily solved by considering that it is, by definition, independent of time, and Eq. (1.202) yields the relation

$$\nabla \cdot \mathbf{J} = 0. \tag{1.204}$$

On the other hand, for physical reasons, any (macroscopic) current must vanish at equilibrium, and the only acceptable solution of this equation is $\mathbf{J} = 0$. Under this condition, Eq. (1.203) reduces to

$$\frac{\nabla P^*(\mathbf{x})}{P^*(\mathbf{x})} = -\frac{\nabla U(\mathbf{x})}{\tilde{\gamma}\, D}, \tag{1.205}$$

or, equivalently, to

$$\nabla \ln P^* = -\frac{\nabla U}{\tilde{\gamma}\, D}. \tag{1.206}$$

This partial differential equation can be integrated if $U(\mathbf{x})$ diverges (and therefore $P(\mathbf{x})$ vanishes) for large $|\mathbf{x}|$,

$$P^*(\mathbf{x}) = A \exp\left(-\frac{U(\mathbf{x})}{T}\right), \tag{1.207}$$

where A is a suitable normalization constant and we have used the Einstein relation $D = T/\tilde{\gamma}$ (see Eq. (1.38)). As one should have expected on the basis of general arguments, we have obtained the equilibrium Boltzmann distribution, where T is the equilibrium temperature.

1.6.4 A Different Pathway to the Fokker–Planck Equation

In Section 1.6.2 we have derived the Fokker–Planck equation by the Itô formula (1.163) for the stochastic differential equation (1.153). A useful alternative formulation of the Fokker–Planck equation can be obtained from the fully continuous version of the Chapman–Kolmogorov equation (1.136), which is defined by the integral equation[21]

$$W(X_0, t_0 | X, t + \Delta t) = \int_{\mathbb{R}} dY\, W(X_0, t_0 | Y, t)\, W(Y, t | X, t + \Delta t), \tag{1.208}$$

where the transition probability $W(X, t|X', t')$ from position X at time t to position X' at time t' exhibits a continous dependence on both space and time variables.[22] As in Section 1.6.2

[21] Until now, with a discrete set of states, it has been useful to define W_{ij}^t as the transition rate between state j and state i, so as to use a matrix notation. Now, with a continuous set of states, it is simpler to use the notation given in Eq. (1.208), so that time flows from left to right and it is easier to follow the calculations.

[22] As discussed at the end of Section 1.6.2, we can assume the condition of spatial isotropy, so that Eq. (1.208) holds for any component X_i ($i = 1, \ldots, N$) of the space vector variable $\mathbf{X} \in \mathbb{R}^N$.

https://doi.org/10.1017/9781107278974.002 Published online by Cambridge University Press

we can consider an arbitrary function $f(X(t))$, which has to vanish sufficiently rapidly for $X \to \pm\infty$, in such a way that

$$\int_{\mathbb{R}} dX f(X)\, \mathcal{O} W(X_0, t_0|X, t) < \infty, \tag{1.209}$$

where the operator $\mathcal{O} = \mathbb{I}, \partial_t, \partial_X, \partial_{XX}, \ldots, (\partial_X)^n, \ldots$, the last symbol being a shorthand notation for the nth-order derivative with respect to the variable X. In particular, we can define the partial derivative of $W(X_0, t_0|X, t)$ with respect to time through the relation

$$\int_{\mathbb{R}} dX f(X) \frac{\partial W(X_0, t_0|X, t)}{\partial t} = \lim_{\Delta t \to 0} \frac{1}{\Delta t} \left\{ \int_{\mathbb{R}} dX f(X) [W(X_0, t_0|X, t + \Delta t) - W(X_0, t_0|X, t)] \right\}. \tag{1.210}$$

Making use of (1.208), we can write the first integral on the right-hand side in the form

$$\int_{\mathbb{R}} dX f(X) W(X_0, t_0|X, t + \Delta t) = \int_{\mathbb{R}} dX f(X) \int_{\mathbb{R}} dY\, W(X_0, t_0|Y, t) W(Y, t|X, t + \Delta t). \tag{1.211}$$

In the limit $\Delta t \to 0$, the transition probability $W(Y, t|X, t + \Delta t)$ vanishes unless X is sufficiently close to Y, so that we can expand $f(X)$ in a Taylor series,

$$f(X) = f(Y) + f'(Y)\,(X - Y) + \frac{1}{2} f''(Y)\,(X - Y)^2 + O(X - Y)^3.$$

By substituting into (1.210) and neglecting higher-order terms, we obtain

$$\begin{aligned} \int_{\mathbb{R}} dX f(X) \frac{\partial W(X_0, t_0|X, t)}{\partial t} = \lim_{\Delta t \to 0} \frac{1}{\Delta t} \Bigg\{ \Bigg[\int_{\mathbb{R}} dX \int_{\mathbb{R}} dY \Bigg(f(Y) + f'(Y)\,(X - Y) \\ + \frac{1}{2} f''(Y)\,(X - Y)^2 \Bigg) \times W(X_0, t_0|Y, t) W(Y, t|X, t + \Delta t) \Bigg] \\ - \int_{\mathbb{R}} dX f(X) W(X_0, t_0|X, t) \Bigg\}. \end{aligned} \tag{1.212}$$

Because of the normalization condition

$$\int_{\mathbb{R}} dX\, W(Y, t|X, t') = 1, \tag{1.213}$$

the first term of the Taylor series expansion cancels with the last integral in (1.212), and we can finally rewrite (1.210) as

$$\begin{aligned} \int_{\mathbb{R}} dX f(X) \frac{\partial W(X_0, t_0|X, t)}{\partial t} = \lim_{\Delta t \to 0} \frac{1}{\Delta t} \Bigg\{ \int_{\mathbb{R}} dY \int_{\mathbb{R}} dX \left[f'(Y)\,(X - Y) + \frac{1}{2} f''(Y)\,(X - Y)^2 \right] \\ \times W(X_0, t_0|Y, t) W(Y, t|X, t + \Delta t) \Bigg\}. \end{aligned} \tag{1.214}$$

We can now define the quantities

$$a(Y, t) = \lim_{\Delta t \to 0} \frac{1}{\Delta t} \int_{\mathbb{R}} dX (X - Y)\, W(Y, t|X, t + \Delta t) \tag{1.215}$$

$$b^2(Y, t) = \lim_{\Delta t \to 0} \frac{1}{\Delta t} \int_{\mathbb{R}} dX (X - Y)^2\, W(Y, t|X, t + \Delta t) \tag{1.216}$$

https://doi.org/10.1017/9781107278974.002 Published online by Cambridge University Press

and rewrite (1.214) as

$$\int_{\mathbb{R}} dX f(X) \frac{\partial W(X_0, t_0|X, t)}{\partial t} = \int_{\mathbb{R}} dY \left[a(Y,t) f'(Y) + \frac{1}{2} b^2(Y,t) f''(Y) \right] W(X_0, t_0|Y, t). \tag{1.217}$$

The right-hand side of this integral equation can be integrated by parts, and by renaming the integration variable $Y \to X$ we obtain

$$\int_{\mathbb{R}} dX f(X) \left[\frac{\partial W}{\partial t} + \frac{\partial}{\partial X}(a(X,t)W) - \frac{1}{2}\frac{\partial^2}{\partial X^2}(b^2(X,t)W) \right] = 0, \tag{1.218}$$

where $W \equiv W(X_0, t_0|X, t)$. Due to the arbitrariness of $f(X)$, we can finally write the Fokker–Planck equation for the transition probability $W(X_0, t_0|X, t)$,

$$\frac{\partial W(X_0, t_0|X, t)}{\partial t} = -\frac{\partial}{\partial X}(a(X,t) W(X_0, t_0|X, t)) + \frac{1}{2}\frac{\partial^2}{\partial X^2}(b^2(X,t) W(X_0, t_0|X, t)). \tag{1.219}$$

This equation is also known as the forward Kolmogorov equation. The pedagogical way this equation has been derived can be made more precise using the Kramers–Moyal expansion: in Appendix E we illustrate this technique for deriving the backward Kolmogorov equation, which describes the evolution with respect to the initial point, rather than to the final one. The backward equation reads

$$\frac{\partial W(X_0, t_0|X, t)}{\partial t_0} = -a(X_0, t_0)\frac{\partial}{\partial X_0} W(X_0, t_0|X, t) - \frac{1}{2} b^2(X_0, t_0)\frac{\partial^2}{\partial X_0^2} W(X_0, t_0|X, t). \tag{1.220}$$

Notice that this equation is not symmetric to (1.219), thus showing that the forward and backward formulations of the Kolmogorov equation are not straightforwardly related to each other. Eq. (1.220) is useful for some applications, such as the one discussed in the following section.

We want also to point out that (1.219) tells us that the transition probability $W(X_0, t_0|X, t)$ obeys a Fokker–Planck equation that is equivalent to the general Fokker–Planck equation (1.168) for the probability density $P(X, t)$. This is a consequence of the relation

$$P(X,t) = \int_{\mathbb{R}} dY P(Y, t') \, W(Y, t', |X, t), \tag{1.221}$$

which is the space–time continuous version of Eq. (1.78). In fact, by deriving the previous equation with respect to t and using Eq. (1.219), one recovers Eq. (1.168). It is worth stressing that these equivalent formulations of the Fokker–Planck equation hold because we have assumed that $W(X_0, t_0|X, t)$ vanishes in the limit $t \to t_0$ if $|X - X_0|$ remains finite. If this is not true, Eqs. (1.219) and (1.220) are not applicable, while the Chapman–Kolmogorov equation (1.208) still holds.

First Exit Time and the Arrhenius Formula

As a first interesting application of what we have discussed in the previous section we want to reconsider the problem of the escape time from an interval $I = [X_1, X_2]$. At page 42 we considered a symmetric, homogeneous diffusion process with absorbing barriers in $X_{1,2}$.

https://doi.org/10.1017/9781107278974.002 Published online by Cambridge University Press

Here we will consider a more complicated process and different boundary conditions. Let us start by setting the problem in general terms.

We can use the transition probability $W(X_0, t_0|X, t)$ to define the probability $\mathbb{P}_{X_0}(t)$ of being still in I at time t after having started from $X_0 \in I$ at time t_0:

$$\mathbb{P}_{X_0}(t) = \int_{X_1}^{X_2} dY\, W(X_0, t_0|Y, t). \tag{1.222}$$

In other words, $\mathbb{P}_{X_0}(t)$ is the probability that the exit time $\mathcal{T}_I(X_0)$ is larger than t. It should be stressed that Eq. (1.222) is not valid for any boundary conditions, because if the "particle" is allowed to exit the interval and reenter, such relation between $\mathbb{P}_{X_0}(t)$ and $W(X_0, t_0|Y, t)$ is no longer true. In the following we are considering either reflecting boundary conditions or absorbing boundary conditions: in both cases, reentry is not allowed and Eq. (1.222) is valid.

Let us indicate with $\Pi(\mathcal{T})$ the probability density of the first exit time from I, starting at X_0 (to simplify the notation in the argument of $\Pi(\mathcal{T})$ we have overlooked the explicit dependence on I and X_0). By definition we have

$$\mathbb{P}_{X_0}(t) = \int_t^{+\infty} d\tau\, \Pi(\tau) \quad \rightarrow \quad \Pi(t) = -\frac{\partial \mathbb{P}_{X_0}(t)}{\partial t}. \tag{1.223}$$

The average exit time is given by the expression

$$\langle \mathcal{T}_I \rangle = \int_0^{+\infty} d\tau\, \Pi(\tau)\, \tau = -\int_0^{+\infty} d\tau \frac{\partial \mathbb{P}_{X_0}(\tau)}{\partial \tau}\, \tau = \int_0^{+\infty} d\tau \mathbb{P}_{X_0}(\tau), \tag{1.224}$$

where the last expression has been obtained integrating by parts and assuming that $\mathbb{P}_{X_0}(t)$ vanishes sufficiently rapidly for $t \to +\infty$.

Now, we consider the backward Kolmogorov equation for the transition probability $W(X_0, t_0|Y, t)$ (1.220), specialized to the diffusive case, $b^2(X, t) = 2D$, with a time-independent drift term $a(X)$, namely,

$$\frac{\partial W(X_0, t_0|Y, t)}{\partial t} = a(X_0)\frac{\partial}{\partial X_0} W(X_0, t_0|Y, t) + D\frac{\partial^2}{\partial X_0^2} W(X_0, t_0|Y, t), \tag{1.225}$$

where we have used $\partial_{t_0} W(X_0, t_0|Y, t) = -\partial_t W(X_0, t_0|Y, t)$, because time translational invariance implies that $W(X_0, t_0|Y, t)$ depends on $(t - t_0)$ only.

We can obtain an equation for $\mathbb{P}_{X_0}(t)$ by integrating both sides of (1.225) over Y, varying in the interval I:

$$\frac{\partial \mathbb{P}_{X_0}(t)}{\partial t} = a(X_0)\frac{\partial}{\partial X_0} \mathbb{P}_{X_0}(t) + D\frac{\partial^2}{\partial X_0^2} \mathbb{P}_{X_0}(t). \tag{1.226}$$

In the following we are going to replace the starting point X_0 with X, to make notations less cumbersome. Integrating further both sides of this equation over t and assuming $t_0 = 0$, we finally obtain an equation for $\langle \mathcal{T}_I(X) \rangle$:

$$a(X)\frac{\partial}{\partial X}\langle \mathcal{T}_I(X) \rangle + D\frac{\partial^2}{\partial X^2}\langle \mathcal{T}_I(X) \rangle = -1, \tag{1.227}$$

where the right-hand side comes from the conditions $\mathbb{P}_X(0) = 1$ and $\mathbb{P}_X(+\infty) = 0$.

https://doi.org/10.1017/9781107278974.002 Published online by Cambridge University Press

Eq. (1.227) can be applied to the example discussed at page 43 by setting $a(X) = 0$ (no drift) and $\langle \mathcal{T}_I(X_1) \rangle = \langle \mathcal{T}_I(X_2) \rangle = 0$ (absorbing boundaries), yielding the solution

$$\langle \mathcal{T}_I(X) \rangle = \frac{1}{2D}(X - X_1)(X_2 - X), \tag{1.228}$$

which is the same result reported in Eq. (1.186).

It is possible to obtain the general solution of Eq. (1.227), once we observe that defining

$$\Phi(X) = \exp\left(\frac{1}{D}\int a(X)dX\right), \tag{1.229}$$

such equation can be rewritten as

$$\frac{d}{dX}\left(\Phi(X)\,\frac{d\langle \mathcal{T}(X)\rangle}{dX}\right) = -\frac{1}{D}\,\Phi(X), \tag{1.230}$$

which can be integrated twice, first giving

$$\frac{d\langle \mathcal{T}(X)\rangle}{dX} = -\frac{1}{D\,\Phi(X)}\int^X dY\,\Phi(Y) + \frac{C_1}{\Phi(X)}, \tag{1.231}$$

then

$$\langle \mathcal{T}(X)\rangle = C_1\int^X dY\frac{1}{\Phi(Y)} - \frac{1}{D}\int^X dY\frac{1}{\Phi(Y)}\int^Y dZ\Phi(Z) + C_2, \tag{1.232}$$

with the integration constants to be determined using the appropriate boundary conditions in $X = X_{1,2}$.

Let us now suppose absorbing conditions at both extrema, $\langle \mathcal{T}(X_1)\rangle = \langle \mathcal{T}(X_2)\rangle = 0$. The former condition can be implemented by taking the lower limit of the two integrals $\int^X dY \cdots$ equal to X_1, which automatically implies $C_2 = 0$. As for the lower limit of the integral $\int^Y dZ$, it is irrelevant, because its variation leads to redefine the constant C_1, so we can take it equal to X_1 as well. Now, Eq. (1.232) is written as

$$\langle \mathcal{T}(X)\rangle = C_1\int_{X_1}^X dY\frac{1}{\Phi(Y)} - \frac{1}{D}\int_{X_1}^X dY\frac{1}{\Phi(Y)}\int_{X_1}^Y dZ\Phi(Z), \tag{1.233}$$

with the right boundary condition, $\langle \mathcal{T}(X_2)\rangle = 0$, implying

$$C_1\int_{X_1}^{X_2} dY\frac{1}{\Phi(Y)} - \frac{1}{D}\int_{X_1}^{X_2} dY\frac{1}{\Phi(Y)}\int_{X_1}^Y dZ\Phi(Z) = 0. \tag{1.234}$$

We can finally write

$$\langle \mathcal{T}(X)\rangle = \frac{A(X;X_1,X_2) - B(X;X_1,X_2)}{D\int_{X_1}^{X_2}\frac{dY}{\Phi(Y)}}, \tag{1.235}$$

where

$$A(X;X_1,X_2) = \int_{X_1}^{X_2}\frac{dY}{\Phi(Y)}\int_{X_1}^Y dZ\Phi(Z)\int_{X_1}^X\frac{dY'}{\Phi(Y')} \tag{1.236}$$

$$B(X;X_1,X_2) = \int_{X_1}^{X}\frac{dY}{\Phi(Y)}\int_{X_1}^Y dZ\Phi(Z)\int_{X_1}^{X_2}\frac{dY'}{\Phi(Y')}. \tag{1.237}$$

https://doi.org/10.1017/9781107278974.002 Published online by Cambridge University Press

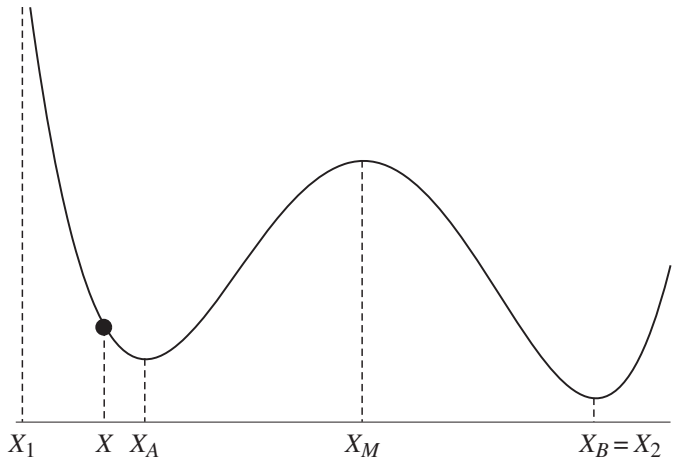

Fig. 1.12 Sketch of the double well potential $U(X)$. The value X_1 can be finite or can diverge to $-\infty$.

We conclude this section with an application of great physical relevance, the Arrhenius formula,[23] which describes the problem of the escape time of a stochastic diffusive process from an asymmetric double-well potential $U(X)$; see Fig. 1.12. In this case the generalized drift term is given by the expression $a(X) = -\frac{1}{\tilde{\gamma}} \frac{dU(X)}{dX}$, where $\tilde{\gamma}$ is the friction coefficient (see Eqs. (1.201)–(1.203)).

We want to ask the following question (see Fig. 1.12 for notations): if a diffusive walker starts at $t = 0$ close to the relative minimum X_A, what is the average time needed to overtake the barrier in X_M and to reach the absolute minimum in X_B? The question can be answered by computing the average time the particle takes to leave the interval $I = [X_1, X_2]$, with a reflecting barrier at X_1 and an absorbing barrier at $X_2 = X_B$. The left boundary condition yields

$$\lim_{X \to X_1^+} \frac{d\langle \mathcal{T}(X) \rangle}{dX} = 0 \tag{1.238}$$

and can be implemented (see Eq. (1.231)) taking $C_1 = 0$ and the lower limit of the integral equal to X_1, i.e.,

$$\frac{d\langle \mathcal{T}(X) \rangle}{dX} = -\frac{1}{D\,\Phi(X)} \int_{X_1}^{X} dY\, \Phi(Y). \tag{1.239}$$

By applying the right boundary condition, $\langle \mathcal{T}(X_B) \rangle = 0$, we obtain the final, explicit result,

$$\langle \mathcal{T}(X) \rangle = \frac{1}{D} \int_{X}^{X_B} dY e^{\frac{U(Y)}{T}} \int_{X_1}^{Y} dZ\, e^{-\frac{U(Z)}{T}}, \tag{1.240}$$

where we have used the Einstein relation to replace $\tilde{\gamma} D$ with T in the exponential functions.

[23] The Arrhenius formula has a great historical importance because it was originally formulated on the basis of heuristic arguments and successfully applied to the study of chemical reactions.

https://doi.org/10.1017/9781107278974.002 Published online by Cambridge University Press

The first integral is dominated by values of the variable Y close to the maximum of $U(X)$ in X_M, while the second integral is weakly dependent on Y for $Y \sim X_M$. Therefore, we can disregard such dependence and approximate the average exit time as follows:

$$\langle \mathcal{T}(X) \rangle \approx \frac{1}{D} \int_{X_1}^{X_M} dZ\, e^{-\frac{U(Z)}{T}} \int_{X}^{X_B} dY e^{\frac{U(Y)}{T}}. \tag{1.241}$$

The integral on Z is now a constant and is dominated by values of $U(Z)$ close to its minimum X_A, so we can use the Taylor series expansion ($\alpha_1 = U''(X_A)$)

$$U(Z) \simeq U(X_A) + \frac{1}{2}\alpha_1 (Z - X_A)^2 \tag{1.242}$$

and obtain the approximate estimate

$$\begin{aligned} \int_{X_1}^{X_M} dZ\, e^{-\frac{U(Z)}{T}} &\approx e^{-\frac{U(X_A)}{T}} \int_{X_1}^{X_M} dZ\, e^{-\frac{\alpha_1}{2T}(Z-X_A)^2} \\ &\approx e^{-\frac{U(X_A)}{T}} \int_{-\infty}^{+\infty} dZ\, e^{-\frac{\alpha_1}{2T}(Z-X_A)^2} \\ &= \sqrt{\frac{2\pi T}{\alpha_1}}\, e^{-\frac{U(X_A)}{T}}. \end{aligned} \tag{1.243}$$

In the second passage we are allowed to extend the limits of the integral to $\pm\infty$ because we make an exponentially small error.

By applying a similar argument, we can observe that the integral over Y in Eq. (1.241) is dominated by values of Y close to X_M, and we can use the Taylor series expansion

$$U(Y) \simeq U(X_M) - \frac{1}{2}\alpha_2 (Y - X_M)^2, \tag{1.244}$$

with $\alpha_2 = -U''(X_M) > 0$, to obtain the approximate estimate

$$\begin{aligned} \int_{X}^{X_B} dY\, e^{\frac{U(Y)}{T}} &\approx e^{\frac{U(X_M)}{T}} \int_{X}^{X_B} dY e^{-\frac{\alpha_2}{2T}(Y-X_M)^2} \\ &\approx e^{\frac{U(X_M)}{T}} \int_{-\infty}^{+\infty} dY\, e^{-\frac{\alpha_2}{2T}(Y-X_M)^2} \\ &= \sqrt{\frac{2\pi T}{\alpha_2}}\, e^{\frac{U(X_M)}{T}}. \end{aligned} \tag{1.245}$$

In conclusion, the average exit time from a metastable potential well[24] is essentially independent of the starting point X, assuming it is close to X_A. The Kramers approximation for the exit time can be finally written as

$$\langle \mathcal{T} \rangle = \frac{2\pi}{\sqrt{\alpha_1 \alpha_2}} \frac{T}{D} e^{\frac{\Delta U}{T}}, \tag{1.246}$$

[24] The absorbing boundary condition on the right means that on the time scale of the exit process the particle has a vanishing probability to come back. Physically, this means that the new minimum in X_B must be lower than the minimum in X_A.

https://doi.org/10.1017/9781107278974.002 Published online by Cambridge University Press

where $\Delta U = U(X_M) - U(X_A)$, meaning that this quantity essentially depends on the height of the barrier separating X_A from X_B: the higher the barrier, the longer the average exit time from X_A.

The Arrhenius formula is essentially equivalent to (1.246) and provides an estimate of the reaction rate $\mathcal{R}$ of two diffusing chemical species, separated by a potential barrier ΔE, corresponding to the energy involved in the reaction process:

$$\langle \mathcal{T}(X) \rangle^{-1} \sim \mathcal{R} \propto \exp\left(-\frac{\Delta E}{T}\right). \tag{1.247}$$

1.6.5 The Langevin Equation and Detailed Balance

We want to complete this section on continuous time stochastic processes by returning to the Einstein relation (1.49), $\tilde{\Gamma} = 2\tilde{\gamma}T$, and discussing how its validity is equivalent to imposing detailed balance. Let us start with the Langevin equation for a particle in the presence of a conservative potential $U(x)$ (see Eq. (1.57)),

$$m\ddot{x}(t) = -U'(x) - \tilde{\gamma}\dot{x}(t) + \tilde{\eta}(t), \tag{1.248}$$

with $\langle \tilde{\eta}(t) \rangle = 0$ and $\langle \tilde{\eta}(t)\tilde{\eta}(t') \rangle = \tilde{\Gamma}\delta(t - t')$, as usual.

The most delicate point is the correct formulation of detailed balance when the microscopic state is labeled by a quantity, the momentum $p = m\dot{x}$, which changes sign with time reversal. In fact, until now we have tacitly assumed to make reference to spin or lattice gas systems, whose microscopic states are invariant under time reversal. For such systems, the formulation of detailed balance is simply $p_\alpha W_{\beta\alpha} = p_\beta W_{\alpha\beta}$ (see Eq. (1.129)), where $W_{\beta\alpha}$ is the transition rate between state α and state β and p_α is the probability to be in state α. Once we introduce the momentum, we should rather prove that

$$p_\alpha W_{\beta\alpha} = p_{\beta^*} W_{\alpha^*\beta^*}, \tag{1.249}$$

where the asterisk means the time-reversed state. In our case, labels read $\alpha = (x, p)$, $\beta = (x', p')$, $\alpha^* = (x, -p)$, and $\beta^* = (x', -p')$.

For a small time dt, Eq. (1.248) means

$$x(t + dt) = x(t) + \frac{p(t)}{m}dt \tag{1.250}$$

$$p(t + dt) = p(t) - U'(x(t))dt - \tilde{\gamma}\frac{p(t)}{m}dt + dW, \tag{1.251}$$

where dW is the infinitesimal increment of a Wiener process, as defined in Eq. (1.150). In practice, it is a Gaussian distributed, stochastic variable with zero average and variance $\langle (dW)^2 \rangle = \int_0^{dt} d\tau' \int_0^{dt} d\tau'' \langle \eta(\tau')\eta(\tau'') \rangle = \tilde{\Gamma}dt$. Therefore, we can write

$$W_{\beta\alpha} = \delta\left(x' - x - \frac{p}{m}dt\right) \frac{1}{\sqrt{2\pi\tilde{\Gamma}dt}} \exp\left\{ -\frac{\left(p' - p + U'(x)dt + \tilde{\gamma}\frac{p}{m}dt\right)^2}{2\tilde{\Gamma}dt} \right\}$$

$$W_{\alpha^*\beta^*} = \delta\left(x - x' + \frac{p'}{m}dt\right) \frac{1}{\sqrt{2\pi\tilde{\Gamma}dt}} \exp\left\{ -\frac{\left(p' - p + U'(x')dt - \tilde{\gamma}\frac{p'}{m}dt\right)^2}{2\tilde{\Gamma}dt} \right\}.$$

https://doi.org/10.1017/9781107278974.002 Published online by Cambridge University Press

From Eqs. (1.250) and (1.251) we find that $(x' - x)$ is of order dt and $(p' - p)$ is of order $\sqrt{dt}$, so we can evaluate the ratio

$$\frac{W_{\beta\alpha}}{W_{\alpha^*\beta^*}} = \exp\left\{\frac{\tilde{\gamma}}{\tilde{\Gamma}}\left(\frac{p^2 - p'^2}{m} - \frac{2p}{m}U'(x)dt\right) + o(dt)\right\} \tag{1.252}$$

$$= \exp\left\{\frac{2\tilde{\gamma}}{\tilde{\Gamma}}(E(\alpha) - E(\beta^*)) + o(dt)\right\}, \tag{1.253}$$

where

$$E(\alpha) \equiv E(x,p) = \frac{p^2}{2m} + U(x). \tag{1.254}$$

Eq. (1.253) can be written as the ratio p_{β^*}/p_α, therefore attesting to the validity of the detailed balance, if and only if $2\tilde{\gamma}/\tilde{\Gamma}$ is the inverse absolute temperature, i.e.,

$$\tilde{\Gamma} = 2\tilde{\gamma}T. \tag{1.255}$$

We might repeat the same procedure in the simpler case of overdamped Langevin equation for the Brownian particle, so that the inertial term $m\ddot{x}(t)$ is negligible. Rather than having the two coupled Eqs. (1.250) and (1.251), we simply have

$$x(t + dt) = x(t) - \frac{U'(x(t))}{\tilde{\gamma}} + dW, \tag{1.256}$$

and similar calculations lead to the same result (1.255).

1.7 Generalized Random Walks

In the previous sections we have described diffusive processes ruled by Gaussian probability distributions of observables in the mathematical language of discrete as well as of continuous stochastic processes.

On the other hand, the astonishing richness of natural phenomena encourages us to explore new territories of mathematics of stochastic processes. In fact, there is a broad class of phenomena that escape standard diffusive processes and exhibit quite a different statistical description. As physical examples we can mention turbulent fluid regimes, the dynamics of ions in an optical lattice, the diffusion of light in heterogeneous media, hopping processes of molecules along polymers, etc. Instances from other fields of science are the spreading of epidemics, the foraging behavior of bacteria and animals, the statistics of earthquakes and air traffic, and the evolution of the stock market. This incomplete list is enough to point out the importance of having suitable mathematical tools for describing the statistics of such a widespread class of phenomena that have in common an anomalous diffusive behavior.

The French mathematician Paul Lévy pioneered the statistical description of anomalous diffusive stochastic processes that are usually called Lévy processes; see Fig. 1.13 for an experimental example. Just to fix the ideas from the very beginning, a symmetric

https://doi.org/10.1017/9781107278974.002 Published online by Cambridge University Press

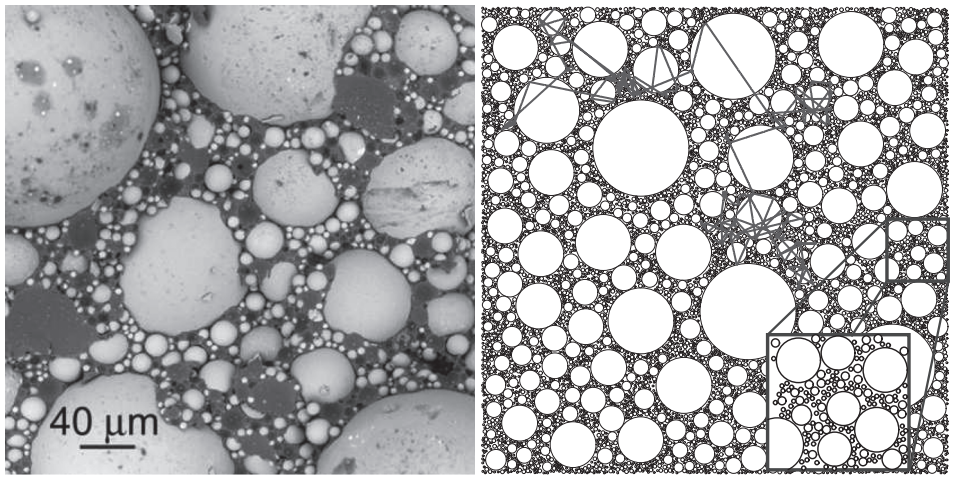

Fig. 1.13 Left: Electron micrograph of a Lévy glass. From M. Burresi et al., Weak localization of light in superdiffusive random systems, *Physical Review Letters*, **108** (2012) 110604. Right: Simulated photon walk in a two-dimensional Lévy glass, with the inset showing the scale invariance of the glass. From P. Barthelemy, J. Bertolotti and D. S. Wiersma, Lévy walk in an inhomogeneous medium, *Nature*, **453** (2008) 495–498.

stochastic process described by the variable $x(t)$ (without prejudice of generality one can take $\langle x(t)\rangle = 0$) is said to be anomalous if its variance obeys the relation

$$\langle x^2(t)\rangle \propto t^\alpha, \quad 0 < \alpha < 2. \tag{1.257}$$

For $\alpha = 1$ we recover the standard diffusive process (therefore, diffusion is not anomalous), while the cases $0 < \alpha < 1$ and $1 < \alpha < 2$ correspond to subdiffusive and superdiffusive stochastic processes, respectively. The case $\alpha = 0$ corresponds, for example, to the Ornstein–Uhlenbeck process (see page 45), while the case $\alpha = 2$ corresponds to ballistic motion.

1.7.1 Continuous Time Random Walk

The examples discussed in Section 1.5.2 describe fully discretized random walks: the walker performs a random sequence of equal-length space steps, each one in a unit time step, i.e. the walker's steps are performed at a velocity with constant modulus. This simple way of representing a random walk can be generalized by assuming that the space steps, x, are continuous, independent, identically distributed random variables, according to a probability distribution $\eta(x)$ (as we have already discussed in Section 1.2.2 for a Poisson distribution of step lengths). Similarly, we can assume that also the time intervals, t, needed to perform a step are continuous, independent, identically distributed random variables, according to a probability distribution $\omega(t)$. This implies that in this continuous time random walk (CTRW) model the walker performs elementary steps with different velocities. On the other hand, the very concept of velocity is still ill-defined, as well as for discrete random walk models, where the macroscopic motion is of diffusive type, despite

https://doi.org/10.1017/9781107278974.002 Published online by Cambridge University Press

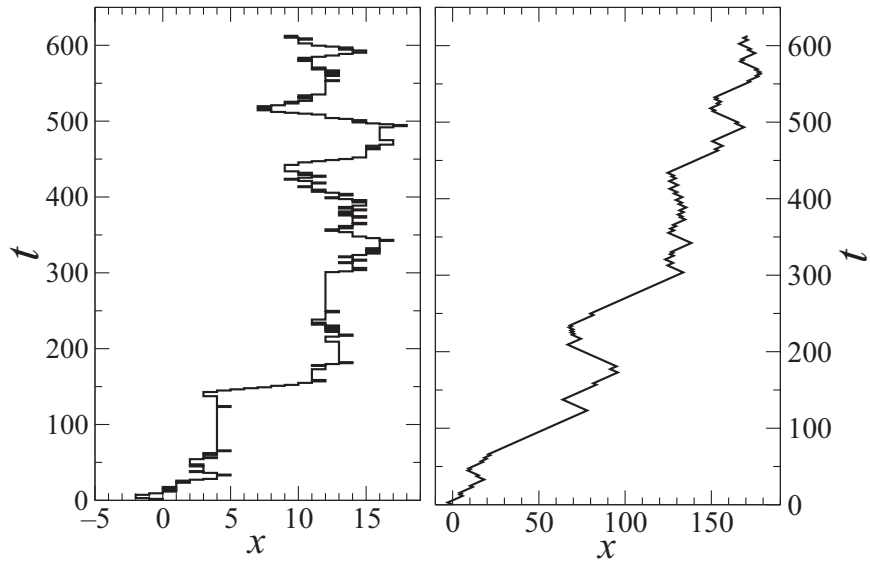

Fig. 1.14 Left: CTRW with step increment ± 1. Right: Lévy walk with velocity ± 1. Both evolutions are generated using the same time distribution $\omega(t) \sim 1/(t^{1+\alpha})$, with $\alpha = 1.5$ and the same distribution of $+/-$ signs for the direction of motion.

the unit step being performed at constant modulus of the velocity. In this perspective, it is more appropriate to interpret the continuous random variable t as a residence time of the walker at a given position in space, before performing a step in a vanishing time; see Fig. 1.14.

Let us now translate the CTRW model into a suitable mathematical language. For the sake of simplicity, we discuss here the one-dimensional case, whose extension to higher dimension does not present additional difficulties. If we assume that the space where the random walker moves is isotropic, we have to assume that a jump of length x is equally probable to a jump of length $-x$; i.e., we have to take into account a normalized symmetric distribution, $\eta(x) = \eta(-x)$ and $\int_{-\infty}^{+\infty} dx \eta(x) = 1$. Notice that, as for the standard Brownian motion, we have

$$\langle x \rangle = \int_{-\infty}^{+\infty} x \eta(x) \mathrm{d}x = 0, \tag{1.258}$$

$$\langle x^2 \rangle = \int_{-\infty}^{+\infty} x^2 \eta(x) \mathrm{d}x \neq 0. \tag{1.259}$$

The normalized probability distribution for the residence time $\omega(t)$ is defined for $0 \leq t \leq \infty$, with $\int_0^\infty \omega(t) dt = 1$.[25]

[25] When dealing with probabilities depending on values of physical quantities, such as space lengths or times as in this case, one has to remember that all of these values are pure numbers. They have to be multiplied by suitable physical units to recover their physical meaning. For instance, in this way the normalization conditions for $\eta(x)$ and $\omega(t)$ do not demand that these probability distributions have the dimension of an inverse length and of an inverse time, respectively.

https://doi.org/10.1017/9781107278974.002 Published online by Cambridge University Press

It is useful to introduce the quantity

$$\psi(t) = \int_t^{+\infty} \omega(\tau)\mathrm{d}\tau = 1 - \int_0^t \omega(\tau)\mathrm{d}\tau, \tag{1.260}$$

which represents the probability of the walker not to make a step until time t (and called survival probability). Now we want to define the random dynamics of the walker by introducing the probability $p(x,t)$ that a continuous-time random walker arrives at position x at time t, starting from the origin $x = 0$ at time $t = 0$. This probability can be written as[26]

$$p(x,t) = \int_{-\infty}^{+\infty} \mathrm{d}y \int_0^t \mathrm{d}\tau\, \eta(y)\, \omega(\tau)\, p(x-y, t-\tau) + \delta(x)\,\delta(t), \tag{1.261}$$

which means that the position x can be reached at time t by any walker that arrived at any other position $x - y$ at time $t - \tau$, by making a jump of length y ($-\infty < y < +\infty$) after a waiting time τ ($0 \le \tau \le t$). Notice that causality is preserved, meaning that the evolution of the walker at time t depends only on what happened at previous times. We can use $p(x,t)$ and the survival probability $\psi(t)$ to obtain the probability that a continuous-time random walker is at position x at time t,

$$\rho(x,t) = \int_0^t \mathrm{d}\tau\, p(x, t-\tau)\psi(\tau). \tag{1.262}$$

In fact, $\rho(x,t)$ is the sum over the time intervals τ during which a walker, which arrived at x at any previous time $t - \tau$, survived until time t.

We can use the Fourier–Laplace transformation to solve the integral equations (1.261) and (1.262). Let us denote with k and s the dual variables of x and t, respectively. Eq. (1.261) is a convolution product for both variables x and t; by Fourier-transforming on x and by Laplace-transforming on t, this equation takes the simple form[27]

$$p(k,s) = \eta(k)\,\omega(s)\,p(k,s) + 1, \tag{1.263}$$

thus yielding

$$p(k,s) = \frac{1}{1 - \eta(k)\,\omega(s)}, \tag{1.264}$$

where

$$\eta(k) = \int_{-\infty}^{+\infty} dx\, e^{-ikx}\eta(x), \tag{1.265}$$

$$\omega(s) = \int_0^{+\infty} dt\, e^{-st}\omega(t). \tag{1.266}$$

Before performing the Fourier–Laplace transform of $\rho(x,t)$, it is useful to evaluate the Laplace transform of the survival probability, $\psi(t)$. Since its derivative is minus the residence time probability (see Eq. (1.260)), we can write

[26] This equation is a generalization of the case of continuous time steps of Eq. (D.36) in Appendix D.4.

[27] In the exact definition of Laplace transform the lower limit of time integration is 0^-, so the integration of $\delta(t)$ gives 1 and not 1/2. Additionally, $p(x,t)$ can be extended to $t < 0$ by assuming $p(x,t) \equiv 0$ at negative times.

https://doi.org/10.1017/9781107278974.002 Published online by Cambridge University Press

$$\omega(s) = \int_0^\infty dt e^{-st} \omega(t) \tag{1.267}$$

$$= -\int_0^\infty dt e^{-st} \psi'(t) \tag{1.268}$$

$$= 1 - s\psi(s), \tag{1.269}$$

where we have integrated by parts and used the condition $\psi(0) = 1$. We can use this result in the form

$$\psi(s) = \frac{1 - \omega(s)}{s} \tag{1.270}$$

to perform the Fourier–Laplace transform of $\rho(x,t)$, thus obtaining (see Eq. (1.262))

$$\rho(k,s) = \frac{1}{1 - \eta(k)\,\omega(s)} \frac{1 - \omega(s)}{s}. \tag{1.271}$$

This is quite a useful equation, since the Fourier–Laplace transform of the density of continuous time random walkers is found to depend in a simple way on $\eta(k)$ and $\omega(s)$. Moreover, the relation between the momenta of the distribution of $\eta(x)$ and the derivatives of its Fourier transform $\eta(k)$, which is reported in Appendix D.4, applies also in this case, where discrete time steps have been substituted with continous time steps, ruled by the probability density function $\omega(t)$. More precisely, by replacing $\eta(k)$ with $\rho(k,t)$, Eq. (D.43) becomes

$$\langle x^m(t) \rangle = i^m \frac{d^m}{d\,k^m} \rho(k,t) \Big|_{k=0}. \tag{1.272}$$

For the Laplace-transformed momenta we can write

$$\langle x^m(s) \rangle = \int_0^{+\infty} dt \langle x^m(t) \rangle\, e^{-s\,t} = i^m \frac{d^m}{d\,k^m} \rho(k,s) \Big|_{k=0}. \tag{1.273}$$

In particular, for $m = 2$, we have

$$\begin{aligned} \langle x^2(s) \rangle &= -\frac{d^2}{d\,k^2} \rho(k,s) \Big|_{k=0} \\ &= -\frac{1 - \omega(s)}{s} \frac{d^2}{d\,k^2} \frac{1}{1 - \eta(k)\,\omega(s)} \Big|_{k=0} \\ &= -\frac{\omega(s)}{s\,[1 - \omega(s)]} \eta''(k=0) - \frac{2\omega^2(s)}{s\,[1 - \omega(s)]^2} (\eta'(k=0))^2, \end{aligned} \tag{1.274}$$

where we have used that $\eta(k=0) = \int_{-\infty}^{+\infty} dx \eta(x) = 1$. The two derivatives appearing in Eq. (1.274) encapsulate relevant information about the distribution $\eta(x)$. More generally, it is interesting to investigate the limits $k \to 0$ and $s \to 0$, which provide information on the asymptotic behaviour of $\rho(x,t)$ at large distances and for long times. In fact, in these limits we can use the approximate expressions

$$\eta(k) \approx \int_{-\infty}^{+\infty} dx \left(1 - ikx - \frac{k^2}{2} x^2 + \cdots\right) \eta(x) = 1 - ik\langle x \rangle - \frac{k^2}{2} \langle x^2 \rangle + \cdots \tag{1.275}$$

$$\omega(s) \approx \int_0^{+\infty} dt\, (1 - st + \cdots) \omega(t) = 1 - \theta\, s + \cdots, \tag{1.276}$$

https://doi.org/10.1017/9781107278974.002 Published online by Cambridge University Press

where we have considered the leading terms of the Taylor series expansion of the exponentials in Eqs. (1.265) and (1.266) and have introduced the quantities

$$\langle x \rangle = \int_{-\infty}^{+\infty} x\eta(x)dx \tag{1.277}$$

$$\langle x^2 \rangle = \int_{-\infty}^{+\infty} x^2\eta(x)dx \tag{1.278}$$

$$\theta = \langle t \rangle = \int_{0}^{+\infty} t\,\omega(t)\,dt. \tag{1.279}$$

For a symmetric distribution of spatial jumps, $\langle x \rangle = 0$ and $\sigma^2 = \langle x^2 \rangle$. Coming back to Eq. (1.274), we can replace $\eta'(k=0) = -i\langle x \rangle = 0$ and $\eta''(k=0) = -\langle x^2 \rangle = -\sigma^2$, obtaining

$$\langle x^2(s) \rangle = \frac{\omega(s)}{s\,[1-\omega(s)]}\sigma^2. \tag{1.280}$$

If we also replace the $s \to 0$ expansion of $\omega(s)$, we find

$$\langle x^2(s) \rangle = \frac{\sigma^2}{\theta s^2}, \tag{1.281}$$

whose Laplace antitransform is

$$\langle x^2(t) \rangle = \frac{\sigma^2}{\theta}t. \tag{1.282}$$

The last equation is proved observing that $\int_0^\infty dte^{-st}t = 1/s^2$.

In conclusion, if θ and σ^2 are finite, we obtain standard diffusion, with

$$D = \frac{\sigma^2}{2\theta}. \tag{1.283}$$

This result can also be found by using expansions (1.275) and (1.276) in Eq. (1.271). Then, at leading order we can write $\big(1-\eta(k)\omega(s)\big) \simeq \sigma^2k^2/2 + \theta s$ and $\big(1-\omega(s)\big) \simeq \theta s$, so that

$$\rho(k,s) \simeq \frac{1}{\left(\theta s + \frac{\sigma^2}{2}k^2\right)}\frac{\theta s}{s} \simeq \left(s + \frac{\sigma^2}{2\theta}k^2\right)^{-1}. \tag{1.284}$$

As shown in Appendix D.3, this is the Fourier–Laplace transformation of the solution of the diffusive equation

$$\frac{\partial \rho(x,t)}{\partial t} = D\,\frac{\partial^2 \rho(x,t)}{\partial x^2}. \tag{1.285}$$

The CTRW model provides also interesting examples of anomalous diffusion. For instance, let us consider the case where σ^2 is finite, while θ diverges. This means that $\omega(t)$ is dominated by a power law for $t \to +\infty$,

$$\omega(t) \sim \frac{\theta^\alpha}{t^{\alpha+1}}, \tag{1.286}$$

with $0 < \alpha < 1$ and θ^α having the dimension of a fractional time, $[t]^\alpha$, in such a way that ω has the correct dimension of inverse time.

https://doi.org/10.1017/9781107278974.002 Published online by Cambridge University Press

Notice that, even in the presence of such an asymptotic behavior, $\omega(t)$ is a well-defined probability distribution provided it can be normalized, namely if one assumes that $\omega(t)$ is a smooth function that does not diverge for $t \to 0^+$.[28] On the other hand, the average residence time (see Eq. (1.279)) is found to diverge as $t^{1-\alpha}$ in the limit $t \to +\infty$ and the CTRW describes a much slower evolution than standard diffusion, i.e. a subdiffusive process . In this case the Fourier–Laplace transform of $\omega(t)$ reads

$$\omega(s) = 1 - \theta^\alpha s^\alpha. \tag{1.287}$$

This result is a generalization of Eq. (1.276), which is not applicable here because $\alpha < 1$ and $\langle t \rangle$ diverges. However, $\omega(t)$ is still normalizable, so it must be

$$\lim_{s \to 0} \omega(s) = 1, \tag{1.288}$$

which allows us to write

$$\begin{aligned} \omega(s) - \omega(s=0) &= \int_0^\infty dt(e^{-st} - 1)\omega(t) \\ &= \left(\int_0^{1/s} + \int_{1/s}^\infty \right) dt(e^{-st} - 1)\omega(t) \\ &\approx -\int_{1/s}^\infty dt\omega(t) + O(s) \\ &= -\theta^\alpha s^\alpha \int_1^\infty \frac{d\tau}{\tau^{1+\alpha}}, \end{aligned} \tag{1.289}$$

where we have approximated $(e^{-st} - 1)$ with 0 in the first integral ($t < 1/s$) and with -1 in the second integral ($t > 1/s$). In conclusion, since the definite integral in the last line is a number and θ^α has just a dimensional role, we obtain Eq. (1.287).

We can now follow the same procedure adopted for the standard diffusive case, so in the limits $k \to 0$ and $s \to 0$ we can use the approximations $1 - \eta(k)\omega(s) \simeq \sigma^2 k^2/2 + (\theta\, s)^\alpha$ and $1 - \omega(s) \simeq (\theta\, s)^\alpha$, thus obtaining at leading order the expression

$$\rho(k,s) \simeq \frac{\theta^\alpha s^{\alpha-1}}{\frac{\sigma^2 k^2}{2} + (\theta s)^\alpha} \simeq \left(s + \frac{\sigma^2 k^2}{2\theta^\alpha} s^{1-\alpha} \right)^{-1}, \tag{1.290}$$

which reduces to Eq. (1.284) for $\alpha = 1$. Using this expression of $\rho(k,s)$ we can first compute the probability of finding the walker at the origin. This is given by the expression

$$\rho(x=0,s) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} d\,k\ \rho(k,s) \simeq \frac{\theta^\alpha\, s^{\alpha-1}}{2\,\pi} \int_{-\infty}^{+\infty} d\,k\ \left(k^2 \frac{\sigma^2}{2} + \theta^\alpha\, s^\alpha \right)^{-1}. \tag{1.291}$$

With the change of variable $y = k\sqrt{\sigma^2/(2\,\theta^\alpha\, s^\alpha)}$, we can write

$$\rho(x=0,s) \simeq \frac{1}{\pi\sqrt{2}} \frac{\theta^{\alpha/2}}{\sigma\, s^{1-\alpha/2}} \int_{-\infty}^{+\infty} dy\ \frac{1}{1+y^2} \simeq \frac{1}{\sqrt{2}} \frac{\theta^{\alpha/2}}{\sigma\, s^{1-\alpha/2}}. \tag{1.292}$$

[28] Heuristic arguments suggest that physically interesting processes demand a vanishing probability distribution for a zero residence time of the CTRW.

https://doi.org/10.1017/9781107278974.002 Published online by Cambridge University Press

Since $\rho(x = 0, s)$ does not have a finite limit for $s \to 0$, the line of reasoning to obtain (1.289) cannot be used and a more rigorous approach, making use of the so-called Tauberian theorems, should be adopted.[29] Using such theorems, the result for the probability of finding the walker at the origin is given by

$$\rho(0,t) = \frac{1}{\sqrt{2}} \frac{\theta^{\alpha/2}}{\sigma\,\Gamma(1-\alpha/2)}\, t^{-\alpha/2}. \tag{1.293}$$

We can therefore conclude that this probability decays algebraically in time with an exponent $\alpha/2$, which for normal diffusion is $1/2$. It is important to stress that this expression provides just the asymptotic (i.e., large t) behavior of the probability density function of finding the walker at the origin.

From (1.290) we can compute all momenta of the distribution. Here we just compute the time-dependent average square displacement, using the general formula (1.274), with $\langle x \rangle = 0$, and $\omega(s)$ given by (1.287), thus yielding up to leading order in k

$$\langle x^2(s) \rangle = -\frac{\omega(s)}{s\,[1-\omega(s)]}\sigma^2 \approx \frac{1}{\theta^\alpha\, s^{\alpha+1}}\,\sigma^2. \tag{1.294}$$

By applying the results of the Tauberian theorems we obtain

$$\langle x^2(t) \rangle = \frac{1}{\Gamma(1+\alpha)} \frac{\sigma^2}{\theta^\alpha} t^\alpha = 2\,D_\alpha\, t^\alpha, \tag{1.295}$$

where D_α is a generalized diffusion coefficient.

For the sake of completeness we want to mention that (1.290) can be read as the Fourier–Laplace transform of the following, fractional partial differential equation

$$\frac{\partial^\alpha \rho(x,t)}{\partial t^\alpha} = D_\alpha \frac{\partial^2 \rho(x,t)}{\partial x^2}. \tag{1.296}$$

Although the important topic of fractional partial differential equations cannot be addressed here, we can nevertheless observe on an intuitive ground that (1.295) follows by a simple dimensional analysis of (1.296). In conclusion, Eq. (1.295) is the basic physical relation describing a subdiffusive random walk (i.e., $0 < \alpha < 1$) in continuous space and time.

Another interesting example contained in the CTRW model is the case where θ is finite, while $\eta(x)$ is a normalized symmetric probability density function, whose asymptotic behavior obeys the power law

$$\eta(x) \sim |x|^{-(1+\alpha)}, \tag{1.297}$$

with $0 < \alpha < 2$. These processes are known as Lévy flights. For symmetry reasons the first momentum of $\eta(x)$ is zero, while the second momentum is found to diverge as

[29] The Tauberian theorems allow us to determine the Laplace transform of a function $f(t)$ whose behavior for large t is of the form

$$f(t) \simeq t^{\gamma-1}\,\Phi(t) \quad \text{with} \quad 0 < \gamma < +\infty,$$

where $\Phi(t)$ is a slowly varying function of t (e.g., $\lim_{t\to+\infty} \Phi(t) = c$, where c is a constant, or even $\phi(t) = [\log(t)]^n$). The Tauberian theorems tell us that the Laplace transform of $f(t)$ is

$$f(s) = \Gamma(\gamma)\, s^{-\gamma}\, \Phi(1/s).$$

https://doi.org/10.1017/9781107278974.002 Published online by Cambridge University Press

$$\langle x^2 \rangle = \lim_{R\to\infty} \int^R dx\, |x|^{1-\alpha} \approx \lim_{R\to\infty} R^{2-\alpha}. \tag{1.298}$$

The Fourier transform of $\eta(x)$ can be evaluated with the spirit of Eq. (1.289), using the fact that the distribution probability is normalizable, so $\eta(k=0) = 1$:

$$\eta(k=0) - \eta(k) = \int_{-\infty}^{+\infty} dx \left(1 - e^{-ikx}\right) \eta(x) \tag{1.299}$$

$$\approx 2 \int_{1/|k|}^{\infty} dx \frac{1 - \cos(|k|x)}{x^{1+\alpha}} \tag{1.300}$$

$$= 2|k|^\alpha \int_1^\infty dy \frac{1 - \cos y}{y^{1+\alpha}}. \tag{1.301}$$

Finally, we find

$$\eta(k) \sim 1 - (\sigma\, |k|)^\alpha, \tag{1.302}$$

where σ is a suitable physical length scale.

In the limits $k \to 0$ and $s \to 0$, from Eq. (1.271) we obtain the approximate expression

$$\rho(k,s) \simeq \left(s + \frac{\sigma^\alpha}{\theta}|k|^\alpha\right)^{-1}. \tag{1.303}$$

By introducing the quantity $\mathcal{D}_\alpha = \sigma^\alpha/\theta$ we can write

$$\rho(k,t) \simeq e^{-\mathcal{D}_\alpha\, t\, |k|^\alpha}, \tag{1.304}$$

where $\rho(k,t)$ is the inverse Laplace transform of $\rho(k,s)$ with respect to s, because $\int_0^\infty dt e^{-st} e^{-\mathcal{D}_\alpha t|k|^\alpha} = 1/(s + \mathcal{D}_\alpha |k|^\alpha)$. The probability distribution $\rho(x,t)$ can be obtained by antitransforming $\rho(k,t)$ in the Fourier space of k, but this is not an easy task, because the exponent α in (1.304) is not an integer. Nonetheless, it is possible to derive the fractional momenta of order q of $\rho(x,t)$, as

$$\langle |x(t)|^q \rangle = \int_{-\infty}^{+\infty} \rho(x,t)|x(t)|^q dx \tag{1.305}$$

$$= \frac{1}{2\pi} \int_{-\infty}^{+\infty} dx |x|^q \int_{-\infty}^{+\infty} dk e^{ikx} e^{-\mathcal{D}_\alpha\, t\, |k|^\alpha} \tag{1.306}$$

$$= (\mathcal{D}_\alpha t)^{q/\alpha} \frac{1}{2\pi} \int_{-\infty}^{+\infty} dy |y|^q \int_{-\infty}^{+\infty} dp e^{ipy} e^{-|p|^\alpha} \tag{1.307}$$

$$\sim (\mathcal{D}_\alpha t)^{q/\alpha}, \tag{1.308}$$

where we have made two changes of variable, $p = (\mathcal{D}_\alpha t)^{1/\alpha} k$ and $y = x/(\mathcal{D}_\alpha t)^{1/\alpha}$, and the final integral on the right is just a number. In particular, for $q = 2$ we have

$$\langle x^2(t) \rangle \sim (\mathcal{D}_\alpha t)^{2/\alpha}. \tag{1.309}$$

As expected, for $\alpha = 2$ we recover the standard diffusive behavior, while for $\alpha < 2$ the mean squared displacement grows in time more than linearly. Notice that $\alpha = 1$ corresponds to a ballistic propagation. As we are going to discuss in the following section, a suitable description of superdiffusion in the CTRW requires consideration of finite velocities by the introduction of the Lévy walk process.

https://doi.org/10.1017/9781107278974.002 Published online by Cambridge University Press

1.7.2 Lévy Walks

Lévy processes are a general class of stochastic Markov processes with independent stationary time and space increments. For example, above we studied Lévy flights, characterized by a narrow distribution of time steps and a distribution of space increments given by Eq. (1.297). The main feature of the latter distribution is that not all of its momenta are finite. As a consequence, it cannot be fully characterized by its first two momenta, i.e. the mean $\langle x \rangle$ and the variance $\langle x^2 \rangle - \langle x \rangle^2$, as is the case for Gaussian probability distributions.

The main advantage of studying Lévy flights is that they are quite simple mathematical models of anomalous diffusion. Their main drawback is that we are implicitly assuming that each step lasts over a fixed time interval, independent of the step length. Since this can be arbitrarily large, the random walker seems to move with an arbitrarily large velocity, a manifest drawback to giving a physical interpretation to such processes. Such physical interpretation can be restored by introducing Lévy walks (LWs), where the distribution of lengths is still given by

$$\eta(x) \sim \frac{1}{|x|^{1+\alpha}}, \tag{1.310}$$

but the path between the starting and the end points of a jump is assumed to be run by the walker at constant velocity $\mathbf{v}$; i.e., the time t spent in a jump is proportional to its length x, with $t = x/|\mathbf{v}|$ (see Fig. 1.14). It can be argued that a LW can be equally defined as a process where a walker moves at constant velocity in a time step t, whose asymptotic distribution is $\omega(t) \sim t^{-(1+\alpha)}$. If $\alpha < 2$ the variance of $\omega(t)$ diverges and the distances of the jumps run by the walker obey the probability distribution (1.310). In the context of LW, one can describe superdiffusive behavior.

In general terms, we can formulate the LW process as a suitable modification of the CTRW model discussed in Section 1.7.1. The probability $p(x,t)$ that a LW arrives at position x at time t, starting from the origin $x = 0$ at time $t = 0$, still obeys an equation like (1.261), namely,

$$p(x,t) = \int_{-\infty}^{+\infty} \mathrm{d}y \int_0^t \mathrm{d}\tau\, \Omega(y,\tau)\, p(x-y, t-\tau) + \delta(x)\,\delta(t), \tag{1.311}$$

where $\Omega(y,\tau)$ is the joint probability distribution of performing a step of length y in a time τ. Since a LW runs at constant velocity v, this quantity can be expressed as

$$\Omega(x,t) = \delta(|x| - vt)\,\omega(t), \tag{1.312}$$

where

$$\omega(t) \sim t^{-(1+\alpha)} \tag{1.313}$$

for large values of t. Notice that this expression admits the possibility that a walker has not yet completed a LW at time t. By Fourier-transforming on x and by Laplace-transforming on t, Eq. (1.311) yields

$$p(k,s) = \frac{1}{1 - \Omega(k,s)}. \tag{1.314}$$

https://doi.org/10.1017/9781107278974.002 Published online by Cambridge University Press

If $\Omega(x,t) = \eta(x)\omega(t)$ as in Eq. (1.261), we have $\Omega(k,s) = \eta(k)\omega(s)$ and we recover Eq. (1.264). In the present case instead, the Fourier–Laplace transform of $\Omega(x,t)$ (see Eq. (1.312)) is

$$\begin{aligned}\Omega(k,s) &= \frac{1}{2}\int_{-\infty}^{+\infty} dx \int_0^{+\infty} dt e^{-ikx-st}\Big(\delta(-x-vt)+\delta(x-vt)\Big)\,\omega(t)\\ &= \frac{1}{2}\left[\int_0^{+\infty} dt e^{-(s-ivk)t}\,\omega(t) + \int_0^{+\infty} dt e^{-(s+ivk)t}\,\omega(t)\right]\\ &= \frac{1}{2}\big[\omega(s-ivk)+\omega(s+ivk)\big] \equiv \mathrm{Re}\,\omega(s+ivk). \end{aligned} \tag{1.315}$$

For $\alpha < 1$, we can use the result $\omega(s) = 1 - \theta^\alpha s^\alpha$, obtained for the CTRW model (see Eq. (1.287)), to write

$$\Omega(k,s) \simeq 1 - \frac{1}{2}\theta^\alpha[(s-ivk)^\alpha + (s+ivk)^\alpha]. \tag{1.316}$$

This equation must be handled with care, because we have noninteger powers of binomials. We must first take the limit of large distances ($k \to 0$), then the limit of long times ($s \to 0$); otherwise, asymptotic dynamics are limited by the artificially imposed finite size of the system. Therefore, we obtain

$$\Omega(k,s) \approx 1 - \theta^\alpha s^\alpha - \frac{\theta^\alpha v^2}{2}\alpha(1-\alpha)k^2 s^{\alpha-2}. \tag{1.317}$$

In analogy with definition (1.262), we want to use these quantities to obtain an expression for the probability that a LW is at position x at time t,

$$\rho(x,t) = \int_{-\infty}^{+\infty} \mathrm{d}x' \int \mathrm{d}\tau p(x-x', t-\tau)\Psi(x',\tau), \tag{1.318}$$

where now $\Psi(x',\tau)$ is the probability that a LW moves exactly by a distance x' in a time τ. Thus

$$\Psi(x,t) = \delta(|x| - vt)\int_t^{+\infty} \omega(t')\mathrm{d}t', \tag{1.319}$$

which corresponds to the motion of a walker that proceeds at constant speed v under the condition that no scattering event occurs before time t. The same kind of calculation performed to obtain $\Omega(k,s)$ yields the expression of the Fourier–Laplace transform $\Psi(k,s)$,

$$\Psi(k,s) = \mathrm{Re}\psi(s+ivk), \tag{1.320}$$

where we recall that $\psi(s)$ is the Laplace transform of $\psi(t) = \int_t^\infty \omega(t')dt'$ and it is given by Eq. (1.270), which was obtained for CTRW, but still holds for LW.

By combining these results, the expression of the Fourier–Laplace transform of (1.318) is

$$\rho(k,s) = p(k,s)\,\Psi(k,s) = \frac{\Psi(k,s)}{1-\Omega(k,s)}, \tag{1.321}$$

https://doi.org/10.1017/9781107278974.002 Published online by Cambridge University Press

where the denominator can be found by Eq. (1.317). Making use of (1.320) and (1.270) the numerator can be evaluated in a similar manner, giving

$$\Psi(k,s) = \theta^\alpha s^{\alpha-1} + \frac{\theta^\alpha v^2}{2}(\alpha-1)(2-\alpha)k^2\, s^{\alpha-3}. \tag{1.322}$$

We finally obtain

$$\rho(k,s) \simeq \frac{1}{s}\,\frac{s^\alpha - c^2\,k^2\,s^{\alpha-2}}{s^\alpha + b^2\,k^2\,s^{\alpha-2}}\,, \tag{1.323}$$

where $b^2 = v^2\alpha(1-\alpha)/2$ and $c^2 = v^2(1-\alpha)(2-\alpha)/2$.

By applying Eq. (1.273) to this result, after some lengthy algebra and then by Laplace antitransforming in the s variable, we finally obtain

$$\langle x^2(t)\rangle = V^2 t^2, \tag{1.324}$$

i.e. a ballistic behavior, where the characteristic spread velocity is given by $V = v\sqrt{(1-\alpha)} < v$. In the special case $\alpha = 1$, V vanishes and logarithmic corrections of the form $t^2/\ln(t)$ are expected for the mean square displacement.

For $1 < \alpha < 2$ we face quite a different situation because $\omega(t) \sim \theta^\alpha/t^{\alpha+1}$ has a finite first moment; i.e., we are dealing with a finite characteristic time,

$$\tau = \langle t\rangle = \int_0^{+\infty} dt\, t\,\omega(t)t < \infty, \tag{1.325}$$

and we can obtain the expression of the Laplace transform of $\omega(t)$ as

$$\begin{aligned}\omega(s) - \omega(s=0) &= \int_0^\infty dt(e^{-st}-1)\omega(t)\\ &= \int_0^\infty dt(-st + \frac{1}{2}s^2t^2 + \cdots)\omega(t)\\ &= -s\tau + \int_0^{+\infty} dt\left(\frac{1}{2}s^2t^2 + \cdots\right)\omega(t),\end{aligned} \tag{1.326}$$

where the last integral still converges in $t = 0$ if we replace the large t expression, $\omega(t) \sim \theta^\alpha/t^{\alpha+1}$. So we obtain

$$\omega(s) \simeq 1 - \tau s + \theta^\alpha \int_0^{+\infty} dt \frac{\left(\frac{1}{2}s^2t^2 + \cdots\right)}{t^{\alpha+1}} \tag{1.327}$$

$$\simeq 1 - \tau s + \theta^\alpha s^\alpha \int_0^{+\infty} d\tau \frac{\left(\frac{1}{2}\tau^2 + \cdots\right)}{\tau^{\alpha+1}} \tag{1.328}$$

$$\simeq 1 - \tau s + As^\alpha, \tag{1.329}$$

where the quantity A is proportional to θ^α.

https://doi.org/10.1017/9781107278974.002 Published online by Cambridge University Press

Substituting this expression into Eq. (1.315) we obtain

$$\begin{aligned}\Omega(k,s) &\simeq 1-\tau s+\frac{1}{2}A[(s-ivk)^{\alpha}+(s+ivk)^{\alpha}]\\ &\simeq 1-\tau s+A\,s^{\alpha}-\frac{A\,v^2}{2}\alpha(\alpha-1)k^2s^{\alpha-2}\\ &\simeq 1-\tau s+\frac{A\,v^2}{2}\alpha(\alpha-1)k^2s^{\alpha-2},\end{aligned} \tag{1.330}$$

where the term s^{α} has been neglected with respect to the linear term s. As for the passage from the first to the second line, we first consider (as before) the limit $k \to 0$, then $s \to 0$. Making use again of (1.320) and (1.270), we can also compute

$$\begin{aligned}\Psi(k,s) &\simeq \tau-\frac{1}{2}A\,[(s-ivk)^{\alpha-1}+(s+ivk)^{\alpha-1}]\\ &\simeq \tau-A\,s^{\alpha-1}+\frac{A\,v^2}{2}(\alpha-1)(\alpha-2)k^2s^{\alpha-3}\\ &\simeq \tau+\frac{A\,v^2}{2}(\alpha-1)(\alpha-2)k^2s^{\alpha-3},\end{aligned} \tag{1.331}$$

where $s^{\alpha-1}$ has been neglected with respect to the constant term. By substituting (1.330) and (1.331) into (1.321) we eventually obtain

$$\rho(k,s) \simeq \frac{1}{s}\,\frac{\tau-c_1^2\,k^2\,s^{\alpha-3}}{\tau+b_1^2\,k^2\,s^{\alpha-3}}\,, \tag{1.332}$$

where $c_1^2 = Av^2(\alpha-1)(2-\alpha)/2$ and $b_1^2 = Av^2\alpha(\alpha-1)/2$. By applying Eq. (1.273) to this result, after (again) some lengthy algebra and then by Laplace antitransforming in the s variable (using the Tauberian theorems; see note 29 above), we obtain an expression for the mean square displacement,

$$\langle x^2(t)\rangle \simeq Ct^{3-\alpha}, \tag{1.333}$$

where $C \propto A(\alpha-1)/\tau$. In this case we find a superdiffusive behavior that interpolates between the ballistic case ($\alpha = 1$) and standard diffusion ($\alpha = 2$). Also for $\alpha = 2$, logarithmic corrections of the form $t\ln(t)$ have to be expected.

For $\alpha > 2$ also terms of order $s^{\alpha-2}$ become negligible and $\rho(k,s)$, at leading order, recovers the form (1.284); i.e., the LW boils down to a standard diffusive process. In fact, for $\alpha > 2$ the probability distribution $\omega(t)$ has finite average and variance and the asymptotic behavior of the corresponding LW can be argued to be equivalent to a standard Gaussian process equipped with the same average and variance. Summarizing, we have found that LW can be classified according to the following scheme:

$$\langle x^2(t)\rangle \propto \begin{cases} t, & \text{if } \alpha > 2 \quad \text{diffusive}\\ t^{3-\alpha}, & \text{if } 1 < \alpha < 2 \quad \text{superdiffusive}\\ t^2, & \text{if } 0 < \alpha < 1 \quad \text{ballistic.}\end{cases} \tag{1.334}$$

Notice that, at variance with the CTRW, the LW model cannot include subdiffusive propagation, because such a regime is incompatible with a finite velocity of the walker.

https://doi.org/10.1017/9781107278974.002 Published online by Cambridge University Press

In this perspective the subdiffusive regime described by the CTRW model is shown to be unsuitable for phenomenological applications.

1.8 Bibliographic Notes

A pedagogical introduction to the Boltzmann transport equation and related topics can be found in the book by K. Huang, *Statistical Mechanics*, 2nd ed. (Wiley, 1987). A more detailed account of the many aspects related to this fundamental equation is contained in the book by C. Cercignani, *The Boltzmann Equation and Its Applications* (Springer, 1988).

A historical perspective about kinetic theory and statistical mechanics is contained in the book by S. G. Brush, *The Kind of Motion We Call Heat* (North Holland, 1976).

Textbooks providing a basic and modern approach to equilibrium and nonequilibrium statistical mechanics are D. Chandler, *Introduction to Modern Statistical Mechanics* (Oxford University Press, 1987), and L. Peliti, *Statistical Mechanics in a Nutshell* (Princeton University Press, 2011).

The book by A. Einstein, *Investigations on the Theory of the Brownian Movement* (Dover, 1956), collects the five papers written by the great scientist between 1905 and 1908 about this central topic for the future developments of statistical mechanics. The book contains also very interesting notes by the editor, R. Fürth.

An interesting book in support of the atomic theory of matter was published in French by J. B. Perrin in 1913, and it is now available in several editions, with the English title *Atoms*.

An interesting article about the theory of Brownian motion is the original paper by G. E. Uhlenbeck and L. S. Ornstein, On the theory of the Brownian motion, *Physical Review*, **36** (1930) 823–841.

A mathematical approach to the theory of Brownian motion is contained in the book by E. Nelson, *Dynamical Theories of Brownian Motion*, 2nd ed. (Princeton University Press, 2001).

A useful reference for many questions faced in this chapter and beyond are the lecture notes by B. Derrida, *Fluctuations et grandes déviations autour du Seconde Principe* (Collège de France, 2015–2016).

A mathematical approach to the perspective of probability theory of random walk and Markov chains and processes can be found in W. Feller, *An Introduction to Probability Theory and Its Applications*, vol. 1, 3rd ed. (John Wiley & Sons, 1968).

A detailed and complete account about Monte Carlo methods and their applications can be found in K. Binder and D. W. Heermann, *Monte Carlo Simulation in Statistical Physics*, 3rd ed. (Springer, 1997).

A clear, introductory textbook on stochastic processes is D. S. Lemons, *An Introduction to Stochastic Processes in Physics* (John Hopkins University Press, 2002).

A general and complete book about stochastic processes and their applications is C. W. Gardiner, *Handbook of Stochastic Methods*, 3rd ed. (Springer, 2004).

The Fokker–Planck equation in most of its facets and applications is discussed in H. Risken, *The Fokker–Planck Equation: Methods of Solution and Applications* (Springer-Verlag, 1984).

An extended overview about recent achievements in the theory of random walks and Lévy processes is contained in J. Klafter and I. M. Sokolov, *First Steps in Random Walks: From Tools to Applications* (Oxford University Press, 2011). This book includes additional information about fractional partial differential equations.

A reference textbook of functional analysis is W. Rudin, *Functional Analysis* (McGraw-Hill, 1991), where details about Tauberian theorems can be found.

https://doi.org/10.1017/9781107278974.002 Published online by Cambridge University Press