

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 continuous 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.

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.

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,

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,

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

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 velocity² 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.

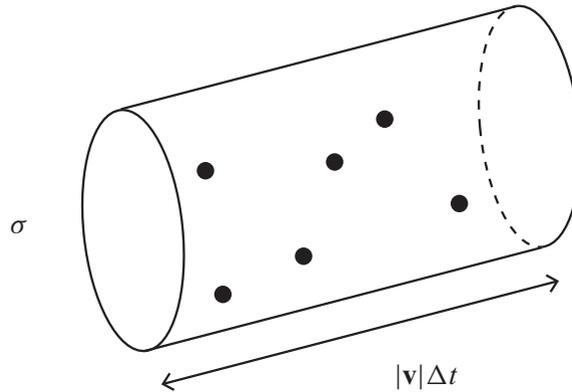


Fig. 1.1 Illustration of the concept of cross section. The black dots in the cylinder spanned by the cross section σ represent the centers of molecules hit in the time interval Δt by a molecule moving at speed v .

pairwise, keeping their total momentum and energy constant (elastic collisions³). 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.

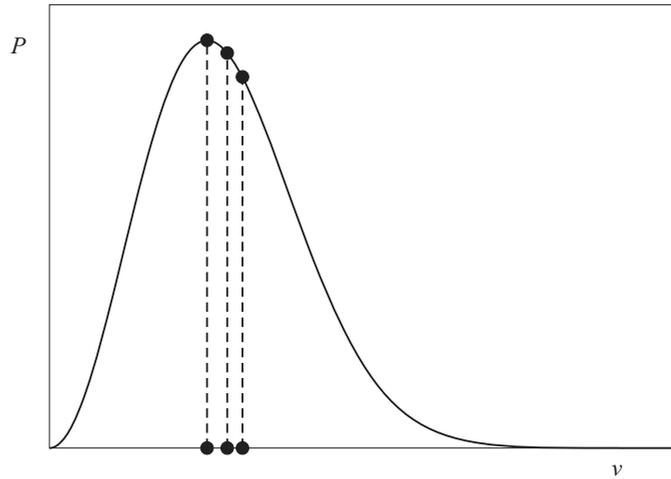


Fig. 1.2 The Maxwell distribution, Eq. (1.5). We indicate, from left to right, the most likely velocity v_{\max} , the average velocity $\langle v \rangle$, and the square root of the average square velocity, $\langle v^2 \rangle^{1/2}$, whose expressions are given in Eq. (1.6).

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

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

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

$$\tau = \frac{1}{n\sigma \langle v \rangle} \quad (1.7)$$

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

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

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

$$\tau = \frac{1}{n\sigma \langle v_r \rangle}, \quad (1.9)$$

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

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

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.

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

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

and obtain

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

from which we can evaluate the mean free path,

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

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

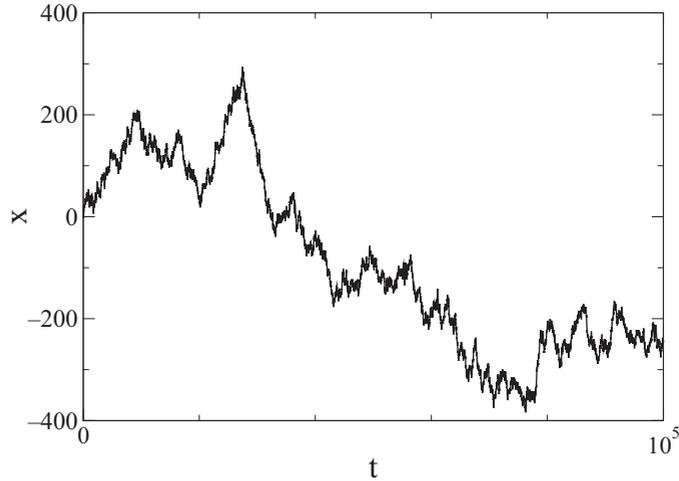


Fig. 1.3 Plot of a one-dimensional random walk, corresponding to $\ell = 1$.

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 i th segment run by the particle after the i th 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 $N \rightarrow \infty$ the average value $\mathbf{X}/N \rightarrow 0$. 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.

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

$$\langle X^2 \rangle = \ell^2 N, \quad (1.18)$$

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

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

and

$$\langle X^2 \rangle = \langle \ell^2 \rangle N. \quad (1.20)$$

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

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

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

$$\langle X^2 \rangle = 2\lambda^2 N. \quad (1.22)$$

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

$$\langle X^2 \rangle = 2\lambda \langle v \rangle t. \quad (1.23)$$

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 $T = 20^\circ\text{C}$ we have $\langle v \rangle \sim 450$ m/s, while $\lambda \sim 0.06$ μm . 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).

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, 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

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

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