Chapter 6
Classical Distribution Function and Transport Equation

6.1 Introduction

When a system made of a large number of molecules is considered, the description of the dynamics of each individual member of the system is practically impossible, and it is necessary to resort to the methods of Statistical Mechanics. The chapter introduces the concept of distribution function in the phase space and provides the definition of statistical average (over the phase space and momentum space) of a dynamical variable. The derivation of the equilibrium distribution in the classical case follows, leading to the Maxwell–Boltzmann distribution. The analysis proceeds with the derivation of the continuity equation in the phase space: the collisionless case is treated first, followed by the more general case where the collisions are present, this leading to the Boltzmann Transport Equation. In the Complements, after a discussion about the condition of a vanishing total momentum and angular momentum in the equilibrium case, and the derivation of statistical averages based on the Maxwell-Boltzmann distribution, the Boltzmann H-theorem is introduced. This is followed by an illustration of the apparent paradoxes brought about by Boltzmann's Transport Equation and H-theorem: the violation of the symmetry of the laws of mechanics with respect to time reversal, and the violation of Poincaré's time recurrence. The illustration is carried out basing on Kac's ring model. The chapter is completed by the derivation of the equilibrium limit of the Boltzmann Transport Equation.

6.2 Distribution Function

Consider a system made of N identical particles with no constraints. For the sake of simplicity, point-like particles are assumed, so that the total number of degrees of freedom is $3\,N$. The dynamics of the jth particle is described by the canonical

M. Rudan, *Physics of Semiconductor Devices*,
DOI 10.1007/978-1-4939-1151-6_6

coordinates q_{1j}, q_{2j}, q_{3j}, p_{1j}, p_{2j}, p_{3j}, that belong to the 6-dimensional μ-space introduced in Sect. 1.9.

If the number of particles is large, the description of the dynamics of each individual belonging to the system is in fact impossible. For instance, the number density of air at 20°C and 1 atm is about 2.5×10^{19} cm^{-3}. Even if the measurement of the initial conditions were possible, it would not be feasible in practice ([66], Sect. 1). This problem is present also when the number of particles is much lower than in the example above.

The problem is faced by adopting the viewpoint of *statistical mechanics*, whose object is not the description of the dynamical behavior of the individual particles but, rather, that of the distribution of the dynamical properties over the phase space. To this purpose one identifies each point of the μ-space with the pair of vectors $\mathbf{q} = (q_1, q_2, q_3)$, $\mathbf{p} = (p_1, p_2, p_3)$ pertaining to it, and considers the elementary volume $\mathrm{d}\omega = \mathrm{d}^3 q\ \mathrm{d}^3 p$ of the μ-space centered at $(\mathbf{q}, \mathbf{p})$, with $\mathrm{d}^3 q = \mathrm{d}q_1\,\mathrm{d}q_2\,\mathrm{d}q_3$ and the like for $\mathrm{d}^3 p$. Then, the number of particles $\mathrm{d}N$ that at time t belong to $\mathrm{d}\omega$ is given by

$$\mathrm{d}N = f_\mu(\mathbf{q}, \mathbf{p}, t)\,\mathrm{d}\omega, \tag{6.1}$$

with f_μ the concentration in the μ-space. The procedure here is similar to that carried out in Sect. 23.2, the difference being that the space considered here is the phase space instead of the configuration space of Sect. 23.2.[1] In both cases, the motion of the particles is described as that of a continuous fluid: in fact, index j is dropped from the canonical coordinates, which do not indicate any more a specific particle, but the center of the elementary cell of the phase space where the concentration f_μ is considered. As in Sect. 23.2, this procedure is legitimate if the cells of volume $\mathrm{d}\omega$ into which the phase space is partitioned can be treated as infinitesimal quantities in the scale of the problem that is being investigated, and the number of particles within each cell is large enough to make their average properties significant. The concentration f_μ is also called *distribution function*. By definition it fulfills the normalization condition

$$\int f_\mu(\mathbf{q}, \mathbf{p}, t)\,\mathrm{d}\omega = N, \tag{6.2}$$

where the integral is 6-dimensional and extends over the whole μ-space. As the order of integration is immaterial, the calculation can be split into two steps, namely,

$$n(\mathbf{q}, t) = \iiint_{-\infty}^{+\infty} f_\mu(\mathbf{q}, \mathbf{p}, t)\,\mathrm{d}^3 p, \qquad N = \iiint_{-\infty}^{+\infty} n(\mathbf{q}, t)\,\mathrm{d}^3 q. \tag{6.3}$$

The function $n(\mathbf{q}, t)$ defined by the first of (6.3) is the concentration of the particles in the configuration space.

---

[1] Note that here the symbol N indicates the number of particles; instead, in Sect. 23.2 the number of particles is indicated with $\mathcal{N}$, whereas N indicates the concentration.

Basing on the distribution function it is possible to define the average of a dynamical function. For the sake of generality the dynamical function is considered to be a vector that depends on all canonical coordinates and time, say, $\mathbf{a} = \mathbf{a}(\mathbf{q}, \mathbf{p}, t)$. Due to the smallness of the cell size one assumes that the dynamical function takes the same value for all particles within the same cell. As a consequence, the product $\mathbf{a}\, f_\mu\, \mathrm{d}\omega$ is the cell value of the function weighed by the number of particles belonging to the cell. The *statistical average* of $\mathbf{a}$ over the phase space is then

$$\mathrm{Av}\,[\mathbf{a}](t) = \frac{1}{N} \int \mathbf{a}(\mathbf{q}, \mathbf{p}, t)\, f_\mu(\mathbf{q}, \mathbf{p}, t)\, \mathrm{d}\omega, \tag{6.4}$$

where the integral is 6-dimensional, while the average over the momentum space is

$$\overline{\mathbf{a}}(\mathbf{q}, t) = \frac{1}{n(\mathbf{q}, t)} \iiint_{-\infty}^{+\infty} \mathbf{a}(\mathbf{q}, \mathbf{p}, t)\, f_\mu(\mathbf{q}, \mathbf{p}, t)\, \mathrm{d}^3 p. \tag{6.5}$$

Using the expression of N given by (6.2), and that of n given by the first of (6.3), shows that the definitions (6.4, 6.5) indeed provide the weighed averages of interest. By way of example, the dynamical function may be identified with the particle velocity $\mathbf{u}$: using the Cartesian coordinates one finds for the average velocity $\mathbf{v}$ in the configuration space the expression

$$\mathbf{v}(\mathbf{r}, t) = \frac{1}{n(\mathbf{r}, t)} \iiint_{-\infty}^{+\infty} \mathbf{u}(\mathbf{r}, \mathbf{p}, t)\, f_\mu(\mathbf{r}, \mathbf{p}, t)\, \mathrm{d}^3 p. \tag{6.6}$$

Similarly, the average Hamiltonian function in the configuration space reads

$$\overline{H}(\mathbf{r}, t) = \frac{1}{n(\mathbf{r}, t)} \iiint_{-\infty}^{+\infty} H(\mathbf{r}, \mathbf{p}, t)\, f_\mu(\mathbf{r}, \mathbf{p}, t)\, \mathrm{d}^3 p. \tag{6.7}$$

6.3 Statistical Equilibrium

This section deals with the properties of a system of particles in a condition of macroscopic equilibrium. Considering that in general the systems that are considered are composed of a large number of particles or molecules, the statistical concepts introduced in Sect. 6.2 will be used. Generally speaking, the condition of statistical equilibrium is fulfilled if the distribution function is independent of time. This condition may be achieved in different ways: for instance, $f_\mu = \mathrm{const}$ fulfills the required condition. A more general definition of a distribution function fulfilling the condition of statistical equilibrium is $f_\mu = f_\mu(c)$, where c is any constant of motion of the system. In case of a conservative system, energy is the most natural constant of motion to be used.

To proceed, consider a conservative system having a total energy E_S, enclosed in a stationary container of volume Ω. Let the walls of the container be such that no energy flows across them. Also, the container is assumed to be sufficiently massive so that it can be regarded as stationary despite the transfer of kinetic energy due to the molecules' collisions with the walls. If any external force acts on the molecules, it is assumed to be independent of time and conservative ([110], Sect. 26). Finally, the total momentum and angular momentum of the system are assumed to vanish; this condition is by no means obvious and requires some reasoning, as detailed in Sect. 6.6.1.

So far it has not been explicitly indicated whether the molecules that form the system are identical to each other or not; in the practical cases it is to be expected that the system under consideration be made of a mixture of different atoms or molecules. As the extension to a mixture is straightforward ([110], Sect. 30), the analysis is limited here to a system made of identical molecules. It should be noted that the molecules are identical, but distinguishable from each other: from the point of view of Classical Mechanics a continuous observation of the trajectory of each molecule is in fact possible, without disturbing its motion. As a consequence, systems that differ by the exchange of two molecules are to be considered as different from each other.

To proceed one assumes that the number of molecules forming the system is N, and that each molecule has R degrees of freedom. The canonical coordinates that describe the motion of a single molecule are then $q_1, \dots, q_R$, $p_1, \dots, p_R$, so that the number of dimensions of the μ-space is $2R$. As anticipated in Sect. 6.2, the description of the precise state of each molecule is impossible in practice; the state of the system will then be specified in a somewhat less precise manner, as detailed below. First, each q axis of the μ-space is divided into equal intervals of size $\Delta q_1, \dots, \Delta q_R$ and, similarly, each p axis is divided into equal intervals of size $\Delta p_1, \dots, \Delta p_R$. As a consequence, the μ-space is partitioned into elements, called *cells*, whose volume and units are, respectively,

$$\Delta M = (\Delta q_1 \, \Delta p_1) \dots (\Delta q_R \, \Delta p_R), \qquad [\Delta M] = (\mathrm{J\,s})^R. \tag{6.8}$$

The partitioning of the μ-space into cells has the advantage, first, that the set of cells is countable. Besides that, the partitioning into cells of finite size has a deeper meaning, that becomes apparent when the theory outlined in this section is extended to the quantum-mechanical case. In fact, due to the Heisenberg uncertainty relation (Sect. 10.6), the precision by which two conjugate variables can simultaneously be known is limited, so that each product $\Delta q_i \, \Delta p_i$ in (6.8) is bounded below, the bound being of the order of the Planck constant.

After the partitioning is accomplished, the state of the system is assigned by specifying the numbers $N_1, N_2, \dots \geq 0$ of molecules that belong, respectively, to the cell labeled $1, 2, \dots$; such numbers are subjected to the constraint $N_1 + N_2 + \dots = N$ which, in view of the calculations that follow, is more conveniently expressed as

$$F_N(N_1, N_2, \dots) = 0, \qquad F_N = N - \sum_i N_i. \tag{6.9}$$

The sum in (6.9) may be thought of as extending to all cells, due to the fact that only in a finite number of cases the cell population N_i differs from zero. Clearly, the description using the numbers $N_1, N_2, \dots$ is less precise than that given by the molecules' trajectories, and provides a partial specification of the state of each molecule within the limits of the size of ΔM. This means, among other things, that identical molecules belonging to different cells are distinguishable from each other, whereas identical molecules belonging to the same cell are not distinguishable.

As mentioned above, the total energy of the system is E_S, which provides a second constraint to be vested with mathematical form. It is provisionally assumed that the system is *dilute*, namely, that the energy of the interaction among the molecules is negligible. It follows that one can approximately assign to each molecule a total energy that corresponds to its position, momentum, and internal configuration, in other terms, to the cell where the molecule belongs. Letting the energy corresponding to the ith cell be E_i, the constraint on energy reads $N_1\, E_1 + N_2\, E_2 + \dots = E_S$, namely,

$$F_E(N_1, N_2, \dots) = 0, \qquad F_E = E_S - \sum_i N_i\, E_i. \tag{6.10}$$

The above reasoning does not apply to concentrated systems, where the interaction energy is strong. However, it can be shown that (6.10) still holds, albeit with a different interpretation of E_i ([110], Sect. 29). Another observation is that, given the constraints to which the system is subjected, the set of numbers $N_1, N_2, \dots$ may not be unique. It is therefore necessary to extract the set, that actually describes the system, from a larger number of sets made of all possible sets that are compatible with the constraints.

To proceed, let $N_1, N_2, \dots$ be a set that provides a molecules' distribution compatible with the constraints; such a set is called *accessible state*. It is postulated that no accessible state is privileged with respect to any other; this is in fact the fundamental hypothesis of equal *a priori* probability of the accessible states, upon which Statistical Mechanics is based ([110], Sect. 23). Remembering that the particles are identical to each other, any system obtained from the original distribution by exchanging two molecules is also compatible with the constraints. However, the system resulting from the exchange of molecules belonging to different cells is different from the original one because such molecules are distinguishable; in contrast, an exchange within the same cell does not produce a different system. As the total number of possible exchanges of N molecules of the system as a whole is $N!$, and the number of possible exchanges within the ith cell is $N_i!$, the total number of different systems corresponding to a set $N_1, N_2, \dots$ is

$$W(N_1, N_2, \dots) = \frac{N!}{N_1!\, N_2! \dots} \tag{6.11}$$

As time evolves, the interactions among the molecules makes the numbers $N_1, N_2, \dots$ to change, so that, if the system is inspected at successive times, it may be found to belong to different accessible states (as the principles of Classical Mechanics apply here, such an inspection does not perturb the state of the system); in principle, given enough time, the system will go through all accessible states. Now, the dependence of W on $N_1, N_2, \dots$ is quite strong, so that some accessible states correspond to a large number W of systems, others to a much smaller number.[2] As no accessible state is privileged with respect to any other, the majority of inspections carried out onto the system will provide the accessible state that maximizes W. Such an accessible state corresponds, by definition, to the condition of *statistical equilibrium* of the system under consideration.

6.4 Maxwell-Boltzmann Distribution

The analysis carried out in Sect. 6.3 led to the conclusion that the condition of statistical equilibrium of the system is found by maximizing the expression of W given by (6.11), under the constraints (6.9) and (6.10). The calculation is based upon the Lagrange method, that determines the free maximum of an auxiliary function F embedding the constraints. It is convenient to maximize $\log W$, which is a monotonic function of W, instead of W itself, so that the auxiliary function reads

$$F(N_1, N_2, \dots, \alpha, \beta) = \log W + \alpha\, F_N + \beta\, F_E, \tag{6.12}$$

where α, β are the *Lagrange multipliers*, respectively related to the total number of molecules and total energy of the system. In a typical system the total number of molecules and the populations of the majority of non empty cells are very large,[3] so that the Stirling approximation (C.97) is applicable; it follows, after neglecting terms of the form $\log\sqrt{2\pi\, N}$ and $\log\sqrt{2\pi\, N_i}$,

$$\log W = \log(N!) - \sum_i \log(N_i!) \simeq N\,\log(N) - N - \sum_i N_i\,\log(N_i) + \sum_i N_i, \tag{6.13}$$

where $-N$ and $\sum_i N_i$ cancel each other due to (6.9). The function to maximize then becomes $F = N\,\log(N) - \sum_i N_i\,\log(N_i) + \alpha\, F_N + \beta\, F_E$. Here the property of N_i of being very large is again of help, because, on account of the fact that a change of N_i by one unit is negligibly small with respect to N_i itself, in calculating the maximum one treats the integers N_i as continuous variables. Taking the derivative of F with respect to, say, N_r, and equating it to zero, yields $\log N_r + 1 = -\alpha - \beta\, E_r$. Neglecting the unity at the left hand side eventually yields the *Maxwell-Boltzmann distribution law*

$$N_r = \exp(-\alpha - \beta\, E_r). \tag{6.14}$$

---

[2] This is apparent even if the numbers $N_1, N_2, \dots$ are much smaller than in realistic systems. Let for instance $N = 8$: the combination $N_1 = 8$, $N_2 = N_3 = \dots = 0$ yields $W = 1$, whereas the combination $N_2 = N_3 = N_4 = N_5 = 2$, $N_1 = N_6 = N_7 = \dots = 0$ yields $W = 2.520$. It is implied that $8\, E_1 = 2\,(E_2 + E_3 + E_4 + E_5) = E_S$.

[3] The hypothesis that the populations are large is not essential. A more complicate calculation, in which such a hypothesis is not present, leads to the same result [22].

The Lagrange multipliers are then determined from (6.9, 6.10); the first one yields

$$\sum_r \exp(-\alpha - \beta\, E_r) = N, \qquad N\, \exp(\alpha) = \sum_r \exp(-\beta\, E_r) = Z, \tag{6.15}$$

where Z denotes the *partition function*.[4]. Extracting $\exp(\alpha)$ from (6.15) and replacing it into (6.10) provides the relation

$$\frac{E_S}{N} = \frac{1}{Z}\sum_r E_r\, \exp(-\beta\, E_r) = -\frac{\partial}{\partial\beta}\log Z, \tag{6.16}$$

with β the only unknown. This procedure is able to express β and, consequently, α, in terms of the dynamical properties of the molecules. A different method to determine the parameters, that relates them with macroscopic quantities typical of Thermodynamics, like pressure and temperature, is shown in the following. First, one introduces a constant C such that

$$C N\, \Delta M = \exp(-\alpha), \tag{6.17}$$

where ΔM is given by (6.8). After replacing (6.17) in (6.14), the cell's volume ΔM is made smaller and smaller so that, after dropping index r, (6.14) is recast in differential form as

$$\mathrm{d}N = C\, N\, \exp(-\beta E)\, \mathrm{d}M, \qquad \frac{1}{C} = \int \exp(-\beta E)\, \mathrm{d}M, \tag{6.18}$$

where the integral is extended over the µ-space and energy E is expressed in terms of the $2\,R$ coordinates $q_1, \dots, q_R, p_1, \dots, p_R$. The nature of the system is now specified as that of a monatomic gas of mass m, described by the Cartesian coordinates x_i and conjugate momenta p_i. Integrating (6.18) over the container's volume Ω yields the number $\mathrm{d}N_p$ of atoms belonging to the elementary volume of the momentum space. As the gas is assumed to be dilute, the energy in (6.18) is substantially of the kinetic type,

$$E \simeq \frac{1}{2m}\left(p_1^2 + p_2^2 + p_3^2\right), \tag{6.19}$$

so that the integrand is independent of the spatial coordinates. It follows

$$\mathrm{d}N_p = C\, N\, \Omega\, \exp(-\beta\, E)\, \mathrm{d}p_1\, \mathrm{d}p_2\, \mathrm{d}p_3. \tag{6.20}$$

The integral of (6.20) over the momentum space yields N, whence

$$\frac{1}{C\,\Omega} = \iiint_{-\infty}^{+\infty} \exp(-\beta\, E)\, \mathrm{d}p_1\, \mathrm{d}p_2\, \mathrm{d}p_3. \tag{6.21}$$

---

[4] Symbol Z comes from the German term *Zustandssumme* ("sum over states") ([95], Chap. II).

The pressure P exerted by the gas is uniform over all the container's walls, hence it can be calculated with reference to any surface element $\mathrm{d}\Sigma$ belonging to them. One can then choose a surface element normal to the x_1 axis, placed in such a way that the atoms impinging on it travel in the positive direction. Let p_1 be the component of momentum of one of such atoms before hitting the wall; after the atom is reflected by the wall, the component transforms into $-p_1$, so that the variation in the momentum component along the x_1 direction is $2p_1$. The p_2, p_3 components, instead, are left unchanged. The product of $2\,p_1$ by the number of atoms hitting $\mathrm{d}\Sigma$ in the unit time provides the force $\mathrm{d}F$ exerted by the gas on the surface element, whence the pressure is obtained as $P = \mathrm{d}F/\mathrm{d}\Sigma$. Now, consider an elementary cylinder whose base and height are $\mathrm{d}\Sigma$ and $\mathrm{d}x_1$ respectively; the number $\mathrm{d}\tilde{N}_1$ of atoms that belong to such a cylinder and whose momentum p_1 belongs to $\mathrm{d}p_1$ is obtained as the product of the atoms' concentration $\mathrm{d}N_p/\Omega$ times the cylinder's volume $\mathrm{d}\Sigma\,\mathrm{d}x_1 = \mathrm{d}\Sigma\,\mathrm{d}t\,\mathrm{d}x_1/\mathrm{d}t = \mathrm{d}\Sigma\,(p_1/m)\,\mathrm{d}t$, integrated over the other two momenta p_2, p_3:

$$\mathrm{d}\tilde{N}_1 = \mathrm{d}\Sigma\,\frac{p_1}{m}\,\mathrm{d}t\,\mathrm{d}p_1\,C\,N \iint_{-\infty}^{+\infty} \exp(-\beta\,E)\,\mathrm{d}p_2\,\mathrm{d}p_3. \tag{6.22}$$

As each atom in (6.22) undergoes a change $2p_1$ in the momentum component due to the reflexion at the wall, the force $\mathrm{d}F = P\mathrm{d}\Sigma$ is obtained by multiplying (6.22) by $2p_1$, dividing it by $\mathrm{d}t$, and integrating over p_1 between 0 and $+\infty$; in fact, only the atoms that move towards the wall must be accounted for. Eliminating $\mathrm{d}\Sigma$ yields

$$\frac{P}{C\,N} = \frac{2}{m}\int_0^{+\infty}\iint_{-\infty}^{+\infty} p_1^2\,\exp(-\beta\,E)\,\mathrm{d}p_1\,\mathrm{d}p_2\,\mathrm{d}p_3. \tag{6.23}$$

Due to the form of (6.19), the integrals in (6.21) and (6.23) are products of one-dimensional integrals over p_1, p_2, and p_3. As a consequence, in the ratio between (6.23) and (6.21) the integrals over p_2 and p_3 cancel each other, to yield

$$\frac{P\Omega}{N} = \frac{2}{m}\,\frac{\int_0^{+\infty} p_1^2\,\exp[-\beta\,p_1^2/(2\,m)]\,\mathrm{d}p_1}{\int_{-\infty}^{+\infty}\exp[-\beta\,p_1^2/(2\,m)]\,\mathrm{d}p_1}. \tag{6.24}$$

Letting Y indicate the integral at the denominator of (6.24), the integral at the numerator is found to be $-m\,\mathrm{d}Y/\mathrm{d}\beta$. Using (C.27) one finds

$$P\Omega = -2\,\frac{\mathrm{d}\sqrt{2\,\pi\,m/\beta}/\mathrm{d}\beta}{\sqrt{2\,\pi\,m/\beta}} = \frac{N}{\beta}. \tag{6.25}$$

The assumption that the gas used in the derivation of (6.25) is dilute makes it possible to consider it as a perfect gas, for which the phenomenological relation $P\Omega = N k_B T$ holds, with $k_B = 1.38\times 10^{-23}$ J/K the *Boltzmann constant* and T the gas temperature. Comparing with (6.25) yields

$$\beta = \frac{1}{k_B\,T}. \tag{6.26}$$

It can be shown that the validity of (6.26) is not limited to the case where the simple derivation shown here applies. Actually, (6.26) is found to hold for any system that follows the Maxwell-Boltzmann distribution law ([110], Sect. 32) and also, as shown in Sects. 15.9.1, 15.9.5, for quantum systems in equilibrium.

6.5 Boltzmann Transport Equation

The expressions worked out in Sect. 6.4 show the important role of the distribution function. It is then necessary to determine the equation fulfilled by it when the system is not in an equilibrium condition. The derivation is made of two steps; in the first one the interactions between molecules are neglected, in the second one they are accounted for. To start with the first step one observes that, due to the neglect of collisions, the only force acting on the molecules is that of an external field. To denote the position in the phase space it is convenient to use the symbol **s** introduced in Sect. 1.8. Here the symbol has a slightly different meaning, because the space is 6-dimensional instead of being $2n$-dimensional. However, the relations (1.57, 1.58, 1.59) still hold. Applying to the μ-space the same reasoning used in Sect. 23.2 to find the continuity equation (23.3), and considering the case where no particles are generated or destroyed, yields

$$\frac{\partial f_\mu}{\partial t} + \mathrm{div}_{\mathbf{s}} \left(\dot{\mathbf{s}}\, f_\mu \right) = 0. \tag{6.27}$$

From (1.58, 1.59) it follows $\mathrm{div}_{\mathbf{s}}(\dot{\mathbf{s}}\, f_\mu) = \dot{\mathbf{s}} \cdot \mathrm{grad}_{\mathbf{s}} f_\mu$. Replacing the latter into (6.27) and using the Cartesian coordinates yields the *Boltzmann collisionless equation*[5]

$$\frac{\partial f_\mu}{\partial t} + \dot{\mathbf{r}} \cdot \mathrm{grad}_{\mathbf{r}} f_\mu + \dot{\mathbf{p}} \cdot \mathrm{grad}_{\mathbf{p}} f_\mu = 0. \tag{6.28}$$

From the meaning of a continuity equation it follows that $-\mathrm{div}_{\mathbf{s}} \left(\dot{\mathbf{s}}\, f_\mu \right)$ is the time variation of f_μ per unit volume $\mathrm{d}\omega$ of the μ-space. As f_μ depends on **r**, **p**, and t, (6.28) is recast in compact form as $\mathrm{d}f_\mu/\mathrm{d}t = 0$.

To accomplish the second step, namely, adding the effects of collisions, one observes that the latter produce a further time change in f_μ. In principle, one might incorporate such effects into $-\mathrm{div}_{\mathbf{s}} \left(\dot{\mathbf{s}}\, f_\mu \right)$; however, it is more convenient to keep the effects of collisions separate from those of the external field. In fact, assuming as before that the system under consideration is dilute, each molecule spends a relatively large fraction of time far enough from the other molecules not to suffer any interaction; in other terms, the time during which a molecule is subjected to the external field is much longer than that involved in a collision. For this reason it is preferable to write the continuity equation, when the collisions are accounted for, as

$$\frac{\partial f_\mu}{\partial t} + \dot{\mathbf{r}} \cdot \mathrm{grad}_{\mathbf{r}} f_\mu + \dot{\mathbf{p}} \cdot \mathrm{grad}_{\mathbf{p}} f_\mu = C, \tag{6.29}$$

---

[5] In plasma physics, (6.28) is also called *Vlasov equation* ([85], Sect. 13.2).

called *Boltzmann Transport Equation*. In (6.29), term C indicates the time variation of f_μ per unit volume $\mathrm{d}\omega$ due to collisions, whereas $-\mathrm{div}_{\mathbf{s}}\left(\dot{\mathbf{s}}\, f_\mu\right)$ keeps the meaning of variation due to the external field. The compact form of (6.29) reads in this case

$$\frac{\mathrm{d}f_\mu}{\mathrm{d}t} = C, \qquad \int C\,\mathrm{d}\omega = 0. \tag{6.30}$$

where the second relation is due to the normalization condition (6.2). In the equilibrium condition the distribution function has no explicit dependence on time ($\partial f_\mu/\partial t = 0$) and depends on constants of motion only, so that $C = 0$. The condition $C = 0$ does not prevent collisions from happening; in fact, in the equilibrium condition the change in the state of two colliding particles is balanced by simultaneous state changes of other particles, that occur in the same elementary volume $\mathrm{d}\omega$, in such a way that the distribution function is left unchanged (*principle of detailed balance*).

In the calculation of C it is assumed that collisions are of the *binary* type, namely, that they involve only two particles at the time because the probability of a simultaneous interaction of more than two particles is negligibly small. This hypothesis, along with the assumption of a short duration of the interactions, greatly simplifies the calculation of C. This issue will not be pursued further here, because it will be developed directly in the quantum case (Sect. 19.3.1). It is worth observing that in a general non-equilibrium condition it is $C \neq 0$; the second relation in (6.30) then indicates that the form of C must be such, that in the integration over the μ-space every elementary contribution to the integral is balanced by an opposite contribution.

When the system under consideration is made of charged particles, the external field that matters is the electromagnetic one; if the particles are identical to each other, (6.29) takes the form

$$\frac{\partial f_\mu}{\partial t} + \mathbf{u}\cdot\mathrm{grad}_{\mathbf{r}} f_\mu + e\,(\mathbf{E} + \mathbf{u}\wedge\mathbf{B})\cdot\mathrm{grad}_{\mathbf{p}} f_\mu = C, \tag{6.31}$$

with $\mathbf{E}(\mathbf{r}, t)$ the electric field, $\mathbf{B}(\mathbf{r}, t)$ the magnetic induction, e the common value of the particles' charge, and $\mathbf{u}$ the velocity (Sect. 1.3.2).

6.6 Complements

6.6.1 Momentum and Angular Momentum at Equilibrium

In the introductory discussion about statistical equilibrium, carried out in Sect. 6.3, it has been assumed that the total momentum and angular momentum of the system vanish. To discuss this issue, consider the box-shaped container whose cross section is shown in Fig. 6.1, filled with molecules identical to each other, having initial positions near the right wall of the container. If the initial velocities are normal to the wall and equal to each other, as schematically indicated by the arrows, the total momentum $\mathbf{P}$ of the particles at $t = 0$ is different from zero. If the left and right walls are perfectly reflecting and parallel to each other, the particles keep bouncing

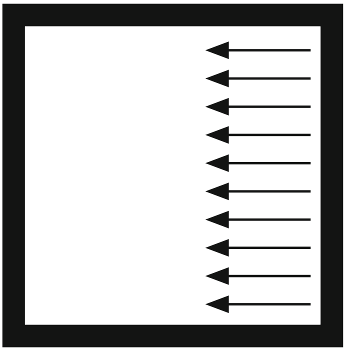


Fig. 6.1 Schematic picture used for discussing the issue of the total momentum of identical molecules within a container

back and forth between the two walls, and the total momentum alternates between **P** and $-\mathbf{P}$. As the container's mass is large, absorbing the momentum $2\,\mathbf{P}$ leaves the stationary condition of the container unaltered. On the other hand, as remarked in Sect. 1.3, this picture should not be regarded as describing a "system of particles", because the latter have no mutual interaction. To establish the interaction one must assume that the initial velocities are slightly misaligned, or the walls are not exactly parallel, or both; in this way the molecules will start colliding with each other and, after some time, their velocities will not be parallel any more. If each collision, and reflection at the walls, is energy conserving, the total energy of the system does not change; in contrast, opposite velocity components of different molecules compensate each other in the calculation of the total momentum, so that the latter will vanish after a sufficiently large number of collisions.[6] A similar argument is applicable to the case of the cylindrical container whose cross section is shown in Fig. 6.2, where the initial positions and velocities of the molecules are such that all molecules would move along the square described by the arrows. In this case the total angular momentum with respect to the cylinder's axis is different from zero. A slight misalignment of the initial conditions, or a deviation of the container's wall from the perfectly-cylindrical form, or both, will eventually make the molecules to collide with each other.

6.6.2 *Averages Based on the Maxwell-Boltzmann Distribution*

In a system made of classical particles in equilibrium at temperature T, each having R degrees of freedom, the average occupation number at energy E_r is given by (6.14). In general, the number of energy levels is large and their separation small, so that one

---

[6] If the walls are perfectly reflecting, and the collisions are elastic (Sect. 3.5), the molecular motions are reversible so that, in both examples of this section, the initial condition is recovered by reversing the direction of time. More comments about this are made in Sect. 6.6.4.

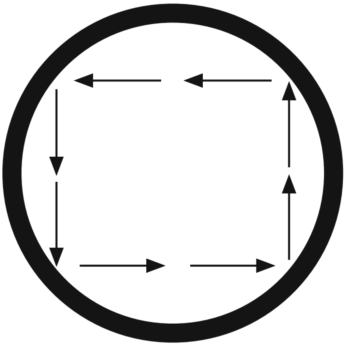


Fig. 6.2 Schematic picture used for discussing the issue of the total angular momentum of identical molecules within a container

disposes of the index and considers the number of particles belonging to the infinitesimal interval $\mathrm{d}q_1\,\mathrm{d}p_1 \ldots \mathrm{d}q_R\,\mathrm{d}p_R$ centered at $(q_1, p_1, \ldots, q_R, p_R)$. After dropping the index, the energy becomes a function of the position $(q_1, p_1, \ldots, q_R, p_R)$ of the interval's center, $E = E(q_1, p_1, \ldots, q_R, p_R)$; in turn, the Maxwell–Boltzmann distribution (6.14) takes the form $\exp(-\alpha - \beta E)$. Given these premises, and extending to the case of R degrees of freedom the definitions of Sect. 6.2, the statistical average over the Maxwell–Boltzmann distribution of a function $\zeta(q_1, p_1, \ldots, q_R, p_R)$ is

$$\overline{\zeta} = \frac{\int \ldots \int \zeta\ \exp(-\beta E)\,\mathrm{d}q_1\,\mathrm{d}p_1 \ldots \mathrm{d}q_R\,\mathrm{d}p_R}{\int \ldots \int \exp(-\beta E)\,\mathrm{d}q_1\,\mathrm{d}p_1 \ldots \mathrm{d}q_R\,\mathrm{d}p_R}, \tag{6.32}$$

where the factor $\exp(-\alpha)$ has been canceled out. An interesting case occurs when ζ depends on the generalized coordinates through the total energy only ($\zeta = \zeta(E)$) and, in turn, the energy is a positive-definite quadratic form of the coordinates,

$$E = a_1\,q_1^2 + b_1\,p_1^2 + \ldots + a_R\,q_R^2 + b_R\,p_R^2, \qquad a_i, b_i > 0. \tag{6.33}$$

Letting $n = 2\,R$ and using the Herring–Vogt transformation (17.66) yields

$$E = \eta/\beta, \qquad \eta = u_1^2 + \ldots + u_n^2, \tag{6.34}$$

where

$$u_1 = \sqrt{\beta\,a_1}\,q_1,\ u_2 = \sqrt{\beta\,b_1}\,p_1, \ldots, u_{n-1} = \sqrt{\beta\,a_R}\,q_R,\ u_n = \sqrt{\beta\,b_R}\,p_R. \tag{6.35}$$

In turn it is $\mathrm{d}u_1 \ldots \mathrm{d}u_n = c\,\mathrm{d}q_1\,\mathrm{d}p_1 \ldots \mathrm{d}q_R\,\mathrm{d}p_R$, with $c = \beta^R \sqrt{a_1\,b_1 \ldots a_R\,b_R}$. Using the procedure involving the density of states illustrated in Sect. B.5 yields

$$\overline{\zeta} = \frac{\int_0^{+\infty} \zeta(\eta)\ \exp(-\eta)\,b(\eta)\,\mathrm{d}\eta}{\int_0^{+\infty} \exp(-\eta)\,b(\eta)\,\mathrm{d}\eta} = \frac{\int_0^{+\infty} \zeta(\eta)\ \exp(-\eta)\,\eta^{n/2-1}\,\mathrm{d}\eta}{\int_0^{+\infty} \exp(-\eta)\,\eta^{n/2-1}\,\mathrm{d}\eta}, \tag{6.36}$$

where the last expression is obtained after canceling out the numerical factors appearing in (B.40). An important case of (6.36) is $\zeta = E = \eta/\beta$, which yields the average energy of the particles. Remembering (C.88, C.89), and using (6.26), one finds

$$\overline{E} = k_B\, T\, \frac{\Gamma(n/2+1)}{\Gamma(n/2)} = \frac{n}{2}\, k_B\, T = R\, k_B\, T. \tag{6.37}$$

The physical systems where the energy is a quadratic form of the type (6.33) are made of linear-harmonic oscillators, like those deriving from the diagonalization of the Hamiltonian function of a system of particles near the mechanical-equilibrium point (Sect. 3.10), or from the expression of the energy of the electromagnetic field *in vacuo* in terms of modes (Sect. 5.6). For systems of this type the average energy of the particles equals $k_B\, T$ times the number R of degrees of freedom of each particle.

Another important system is the dilute one, where the energy is essentially kinetic. In this case the form of the latter is found by letting $a_i \to 0$ in (6.33), so that the energy is made of a sum of R terms instead of $n = 2\, R$. Thus, the average energy of the particles is given by an expression of the form (6.37) where $n/2$ is replaced with $R/2$:

$$\overline{E} = k_B\, T\, \frac{\Gamma(R/2+1)}{\Gamma(R/2)} = R\, \frac{k_B\, T}{2}. \tag{6.38}$$

The above shows that in a dilute system the average energy of the particles equals $k_B\, T/2$ times the number R of degrees of freedom of each particle. The result expressed by (6.37) or (6.38) is called *principle of equipartition of energy*.

6.6.3 Boltzmann's H-Theorem

A measure of the extent by which the condition of a system departs from the equilibrium case is given by Boltzmann's H_B quantity, whose definition is that of statistical average of $\log f_\mu$. The case considered here refers to a spatially-homogeneous system in the absence of external forces, so that $\mathrm{d} f_\mu/\mathrm{d}t = \partial f_\mu/\partial t$. Remembering (6.4) one finds

$$H_B(t) = \mathrm{Av}\,[\,\log f_\mu] = \frac{1}{N} \int f_\mu\, \log f_\mu\, \mathrm{d}\omega, \tag{6.39}$$

whose time derivative, using (6.30), reads

$$\frac{\mathrm{d}H_B}{\mathrm{d}t} = \frac{1}{N} \int \left(\frac{\partial f_\mu}{\partial t}\, \log f_\mu + \frac{\partial f_\mu}{\partial t} \right) \mathrm{d}\omega = \int C\, \log f_\mu\, \mathrm{d}\omega. \tag{6.40}$$

As indicated in Sect. 6.5, the collision term will be worked out directly in the quantum case. However, it is worth anticipating that the analysis of the collision term C leads, both in the classical and quantum cases, to an important conclusion: the time derivative $\mathrm{d}H_B/\mathrm{d}t$ is negative for any distribution function f_μ different from the equilibrium one, while it vanishes in the equilibrium case. This result is the *Boltzmann*

H-theorem. It implies that if a uniform system is initially set in a condition described by a non-equilibrium distribution function, and the external forces are removed, then the initial distribution can not be stationary: an equilibration process occurs, that brings the distribution to the equilibrium one, and whose duration is dictated by the time constants typical of C. The decrease of H_B with respect to time while the system reaches the equilibrium condition reminds one of the behavior of entropy. In fact, it can be shown that H_B is the entropy apart from a negative multiplicative factor and an additive constant ([113], Sect. 18.3).[7]

6.6.4 Paradoxes — Kac-Ring Model

It is known that the Boltzmann Transport Equation (6.29), and the H-theorem derived from it, bring about two apparent paradoxes: the first one is that the equation contains irreversibility, because any initial distribution function, different from the equilibrium one, evolves towards the equilibrium distribution when the external forces are removed, whereas an opposite evolution never occurs. This outcome is in contrast with the symmetry of the laws of mechanics with respect to time reversal (Sect. 2.6.1). The second paradox is the violation of Poincaré's time recurrence, which states that every finite mechanical system returns to a state arbitrarily close to the initial one after a sufficiently long time (called *Poincaré cycle*); this is forbidden by the H-theorem, that prevents entropy from decreasing back to the initial value.

A thorough discussion about the mathematical origin of the paradoxes can be found, e.g., in ([113], Sect. 18.4); a qualitative insight into the question is given by a simple model, called *Kac's ring model*, also reported in [113] and taken from [60]. In the model, N objects are uniformly distributed over a circle, so that at time $t = 0$ each object is ascribed to a specific arc. The objects have two possible states, say, either "0" or "1". The time variable is discrete so that, when time evolves from $k\,\Delta t$ to $(k+1)\,\Delta t$, $k = 0, 1, 2, \ldots$, each object moves clockwise from the arc it occupied at time $k\,\Delta t$ to the next arc. A number $n < N$ of markers is present along the circle: specifically, the markers' positions are at the junctions between two neighboring arcs. The objects that cross the position of a marker change the state from "0" to "1" or vice versa; those that do not cross the position of a marker keep their state.

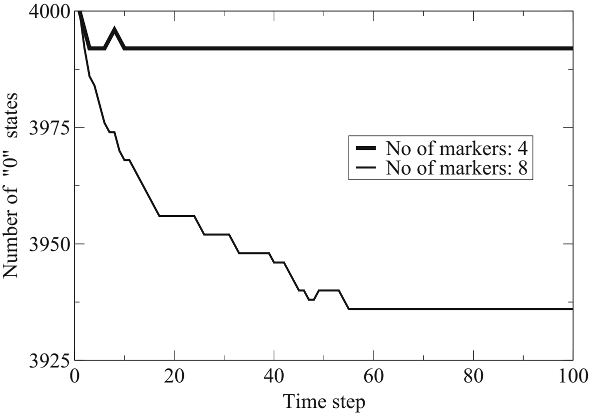
Given the number of objects and markers, the initial state of each object, and the markers' positions along the circle, one wants to investigate the time evolution of the states. Such an evolution is obviously time reversible and fulfills Poincaré's time recurrence; in fact, the set of objects goes back into the initial condition after N time steps if n is even, and after $2\,N$ time steps if n is odd.

Providing the time evolution of the individual object's state is in fact a microscopic description of the system; as remarked in Sect. 6.2, such a description

---

[7] Compare with the definition of entropy given in Sect. 15.9.1 which, at first sight, looks different. The equivalence between the two definitions is justified in Sects. 47 and 102 of [110].

Fig. 6.3 Kac-ring model: computer calculation of the time evolution of the number of "0" states in two samples made of $N = 4{,}000$ objects, which at time $t = 0$ were all set to "0". The markers of the two samples are $n = 4$ and $n = 8$, respectively, and the number of time steps is much smaller than N



becomes impossible when the number of objects in the system is large. A less detailed, macroscopic description of the Kac ring consists, for instance, in providing the time evolution of the number of "0" states. However, the outcome of the latter analysis seems to indicate that an irreversible process takes place; for instance, Fig. 6.3 shows a computer calculation of the time evolution of the number of "0" states in two samples made of $N = 4000$ objects, which at time $t = 0$ were all set to "0". The markers of the two samples are $n = 4$ and $n = 8$, respectively, and the number of time steps is much smaller than N. Both curves tend to decrease and, after some fluctuations (that depend on the markers' positions), stabilize to a constant value; the same behavior occurs at a larger numbers of markers, although the number of time steps necessary to reach a constant value increases (curve $n = 16$ in Fig. 6.4). A further increase in the number of markers makes the fluctuations more pronounced (curve $n = 32$ in the same figure).

On the other hand, a similar calculation using a number of time steps larger than the number of objects shows that the stabilization at or around a constant value is eventually lost: the system fulfills Poincaré's time recurrence and recovers the initial condition (Fig. 6.5). Such an outcome is not detectable in real many-body systems, because the Poincaré cycle is enormously long with respect to the typical time scales of experiments.[8]

6.6.5 Equilibrium Limit of the Boltzmann Transport Equation

As remarked in Sect. 6.5, in the equilibrium condition the distribution function has no explicit dependence on time ($\partial f_\mu / \partial t = 0$) and depends on constants of motion

---

[8] A crude estimate of the Poincaré cycle yields $\sim \exp(N)$, with N the total number of molecules in the system ([50], Sect. 4.5). In typical situations such a time is longer than the age of the universe.

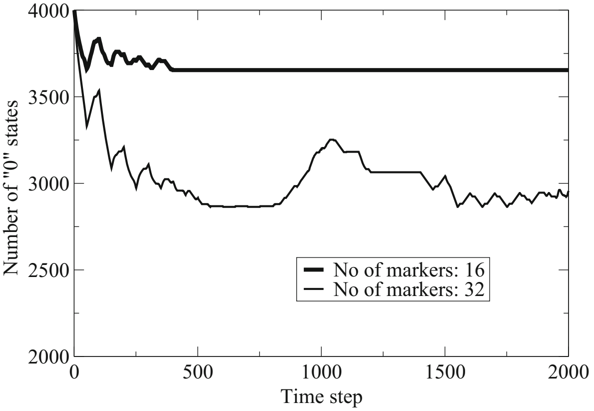


Fig. 6.4 Kac-ring model: computer calculation of the time evolution of the number of “0” states in two samples made of $N = 4000$ objects, which at time $t = 0$ were all set to “0”. The markers of the two samples are $n = 16$ and $n = 32$, respectively, and the number of time steps is much smaller than N

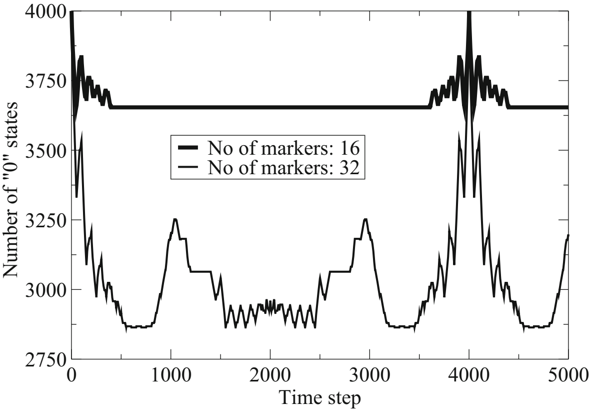


Fig. 6.5 Kac-ring model: computer calculation of the time evolution of the number of “0” states in two samples made of $N = 4000$ objects, which at time $t = 0$ were all set to “0”. The markers of the two samples are $n = 16$ and $n = 32$, respectively, and the number of time steps is larger than N

only, so that $C = 0$. From (6.29) it then follows $\dot{\mathbf{r}} \cdot \mathrm{grad}_{\mathbf{r}} f_\mu + \dot{\mathbf{p}} \cdot \mathrm{grad}_{\mathbf{p}} f_\mu = 0$. In case of a conservative system, energy is the most natural constant of motion to be used; in fact it is $H(\mathbf{r}, \mathbf{p}) = E = \text{const}$, with H the Hamiltonian function (Sect. 1.5). From $f_\mu = f_\mu(H)$ one derives $\mathrm{grad}_{\mathbf{r}} f_\mu = (\mathrm{d} f_\mu / \mathrm{d} H)\, \mathrm{grad}_{\mathbf{r}} H$ and $\mathrm{grad}_{\mathbf{p}} f_\mu = (\mathrm{d} f_\mu / \mathrm{d} H)\, \mathrm{grad}_{\mathbf{p}} H$, so that the equilibrium limit of the Boltzmann Transport Equation reads

$$\frac{\mathrm{d} f_\mu}{\mathrm{d} H} \left(\dot{\mathbf{r}} \cdot \mathrm{grad}_{\mathbf{r}} H + \dot{\mathbf{p}} \cdot \mathrm{grad}_{\mathbf{p}} H \right) = 0, \tag{6.41}$$

Apart from the trivial case $f_\mu = \text{const}$, it is $\mathrm{d} f_\mu / \mathrm{d} H \neq 0$. On the other hand, recasting the Hamilton equations (1.42) as $\dot{\mathbf{r}} = \mathrm{grad}_{\mathbf{p}} H$, $\dot{\mathbf{p}} = -\mathrm{grad}_{\mathbf{r}} H$, and replacing them in (6.41), reduces the term in parentheses to $-\dot{\mathbf{r}} \cdot \dot{\mathbf{p}} + \dot{\mathbf{p}} \cdot \dot{\mathbf{r}}$, this showing that the equation is fulfilled identically regardless of the explicit form of $f_\mu(H)$. This result implies that the equilibrium distribution (6.14) can not be extracted solely from the equilibrium limit of (6.29); its derivation requires also the maximization procedure described in Sects. 6.3 and 6.4.

The equilibrium limit of the Boltzmann Transport Equation described in this section applies in general; as a consequence, it includes also cases like that of (6.31), where a magnetic force is involved. This seems to contradict one of the hypotheses used in Sect. 6.3 to derive the equilibrium distribution, namely, that the system is acted upon by conservative forces. To clarify the issue one uses the concepts illustrated in Sects. 1.5,1.6; first, one notes that the equilibrium limit is achieved by making the external force independent of time, so that the scalar and vector potentials whence **E** and **B** derive are independent of time as well: $\varphi = \varphi(\mathbf{r})$, $\mathbf{A} = \mathbf{A}(\mathbf{r})$ (the dependence on space is kept to account for the possibility that the system under consideration is not uniform). It follows that the Hamiltonian function, which in this case is given by (1.35), is still a constant of motion; as a consequence, the procedure leading to (6.41) is applicable. One notes in passing that each summand in the resulting identity $-\dot{\mathbf{r}} \cdot \dot{\mathbf{p}} + \dot{\mathbf{p}} \cdot \dot{\mathbf{r}}$ becomes in this case

$$\dot{\mathbf{p}} \cdot \mathbf{u} = e\,(\mathbf{E} + \mathbf{u} \wedge \mathbf{B}) \cdot \mathbf{u} = e\,\mathbf{E} \cdot \mathbf{u}, \tag{6.42}$$

where the mixed product vanishes due to (A.32).

Problems

6.1 Calculate the average energy like in (6.37) assuming that energy, instead of being a continuous variable, has the discrete form $E_n = nh\nu$, $n = 0, 1, 2, \ldots$, $h\nu = \text{const}$. This is the hypothesis from which Planck deduced the black-body's spectral energy density (Sect. 7.4.1).