CHAPTER

3

THE PROBLEM OF KINETIC THEORY

3.1 FORMULATION OF THE PROBLEM

The system under consideration in the classical kinetic theory of gases is a dilute gas of N molecules enclosed in a box of volume V. The temperature is sufficiently high and the density is sufficiently low for the molecules to be localized wave packets whose extensions are small compared to the average intermolecular distance. For this to be realized the average de Broglie wavelength of a molecule must be much smaller than the average interparticle separation:

$$\frac{\hbar}{\sqrt{2mkT}}\left(\frac{N}{V}\right)^{1/3} \ll 1 \tag{3.1}$$

Under such conditions each molecule may be considered a classical particle with a rather well-defined position and momentum. Furthermore, two molecules may be considered to be distinguishable from each other. The molecules interact with each other through collisions whose nature is specified through a given differential scattering cross section σ. Throughout our discussion of kinetic theory only the special case of a system of one kind of molecule will be considered.

An important simplification of the problem is made by ignoring the atomic structure of the walls containing the gas under consideration. That is, the physical walls of the container are replaced by idealized surfaces which act on an impinging gas molecule in a simple way, e.g., reflecting it elastically.

We are not interested in the motion of each molecule in detail. Rather, we are interested in the distribution function $f(\mathbf{r}, \mathbf{p}, t)$, so defined that

$$f(\mathbf{r}, \mathbf{p}, t)\, d^3r\, d^3p \tag{3.2}$$

is the number of molecules which, at time t, have positions lying within a volume element d^3r about $\mathbf{r}$ and momenta lying within a momentum-space element d^3p about p. The volume elements d^3r and d^3p are not to be taken literally as

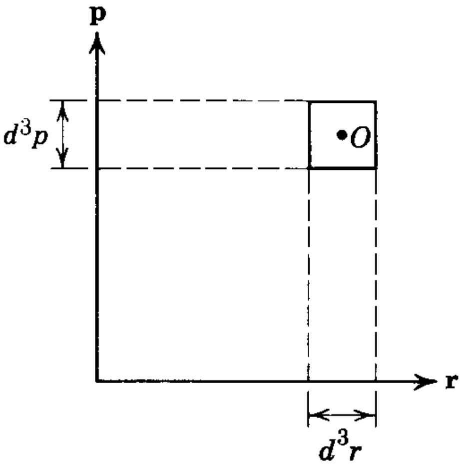


Fig. 3.1 The six-dimensional μ space of a molecule.

mathematically infinitesimal quantities. They are finite volume elements which are large enough to contain a very large number of molecules and yet small enough so that compared to macroscopic dimensions they are essentially points. That such a choice is possible can be seen by an example. Under standard conditions there are about 3×10^{19} molecules/cm^3 in a gas. If we choose $d^3r \sim 10^{-10}$ cm^3, which to us is small enough to be called a point, there are still on the order of 3×10^9 molecules in d^3r.

To make the definition of $f(\mathbf{r}, \mathbf{p}, t)$ more precise, let us consider the six-dimensional space, called the μ space, spanned by the coordinates* $(\mathbf{r}, \mathbf{p})$ of a molecule. The μ space is schematically represented in Fig. 3.1. A point in this space represents a state of a molecule. At any instant of time, the state of the entire system of N molecules is represented by N points in μ space. Let a volume element $d^3r\,d^3p$ be constructed about each point in μ space, such as that shown about the point O in Fig. 3.1. If we count the number of points in this volume element, the result is by definition $f(\mathbf{r}, \mathbf{p}, t)\,d^3r\,d^3p$. If the sizes of these volume elements are chosen so that each of them contains a very large number of points, such as 10^9, and if the density of these points does not vary rapidly from one element to a neighboring element, then $f(\mathbf{r}, \mathbf{p}, t)$ may be regarded as a continuous function of its arguments. If we cover the entire μ space with such volume elements, we can make the approximation

$$\sum f(\mathbf{r}, \mathbf{p}, t)\,d^3p\,d^3r \approx \int f(\mathbf{r}, \mathbf{p}, t)\,d^3p\,d^3r \tag{3.3}$$

where the sum on the left extends over all the centers of the volume elements, and the integral on the right side is taken in the sense of calculus. Such an approximation will always be understood.

Having defined the distribution function, we can express the information that there are N molecules in the volume V through the normalization condition

$$\int f(\mathbf{r}, \mathbf{p}, t)\,d^3r\,d^3p = N \tag{3.4}$$

---

*For brevity, the collection of spatial and momentum coordinates $(\mathbf{r}, \mathbf{p})$ is referred to as the coordinates of a molecule.

If the molecules are uniformly distributed in space, so that f is independent of $\mathbf{r}$, then

$$\int f(\mathbf{r}, \mathbf{p}, t)\, d^3p = \frac{N}{V} \tag{3.5}$$

The aim of kinetic theory is to find the distribution function $f(\mathbf{r}, \mathbf{p}, t)$ for a given form of molecular interaction. The limiting form of $f(\mathbf{r}, \mathbf{p}, t)$ as $t \to \infty$ would then contain all the equilibrium properties of the system. The aim of kinetic theory therefore includes the derivation of the thermodynamics of a dilute gas.

To fulfill this aim, our first task is to find the equation of motion for the distribution function. The distribution function changes with time, because molecules constantly enter and leave a given volume element in μ space. Suppose there were no molecular collisions (i.e., $\sigma = 0$). Then a molecule with the coordinates $(\mathbf{r}, \mathbf{p})$ at the instant t will have the coordinates $(\mathbf{r} + \mathbf{v}\,\delta t, \mathbf{p} + \mathbf{F}\,\delta t)$ at the instant $t + \delta t$, where $\mathbf{F}$ is the external force acting on a molecule, and $\mathbf{v} = \mathbf{p}/m$ is the velocity. We may take δt to be a truly infinitesimal quantity. Thus all the molecules contained in a μ-space element $d^3r\, d^3p$, at $(\mathbf{r}, \mathbf{p})$, at the instant t, will all be found in an element $d^3r'\, d^3p'$, at $(\mathbf{r} + \mathbf{v}\,\delta t, \mathbf{p} + \mathbf{F}\,\delta t)$, at the instant $t + \delta t$. Hence in the absence of collisions we have the equality

$$f(\mathbf{r} + \mathbf{v}\,\delta t, \mathbf{p} + \mathbf{F}\,\delta t, t + \delta t)\, d^3r'\, d^3v' = f(\mathbf{r}, \mathbf{v}, t)\, d^3r\, d^3v$$

which reduces to

$$f(\mathbf{r} + \mathbf{v}\,\delta t, \mathbf{p} + \mathbf{F}\,\delta t, t + \delta t) = f(\mathbf{r}, \mathbf{p}, t) \tag{3.6}$$

because $d^3r\, d^3p = d^3r'\, d^3p'$. The last fact is easily established if we assume that the external force $\mathbf{F}$ depends on position only. At any instant t, we may choose $d^3r\, d^3p$ to be a six-dimensional cube. It is sufficient to show that the area of any projection of this cube, say, $dx\, dp_x$, does not change. A simple calculation will show that this projection, originally a square, becomes a parallelogram of the same area in the time δt, as illustrated in Fig. 3.2. This invariance is valid as long as $(\mathbf{r}, \mathbf{p})$ are canonically conjugate generalized coordinates.

When there are collisions (i.e., $\sigma > 0$), equality (3.6) must be modified. We write

$$f(\mathbf{r} + \mathbf{v}\,\delta t, \mathbf{p} + \mathbf{F}\,\delta t, t + \delta t) = f(\mathbf{r}, \mathbf{p}, t) + \left(\frac{\partial f}{\partial t}\right)_{\text{coll}} \delta t \tag{3.7}$$

which *defines* $(\partial f/\partial t)_{\text{coll}}$. Expanding the left-hand side to the first order in δt, we obtain the equation of motion for the distribution function as we let $\delta t \to 0$:

$$\left(\frac{\partial}{\partial t} + \frac{\mathbf{p}}{m} \cdot \nabla_{\mathbf{r}} + \mathbf{F} \cdot \nabla_{\mathbf{p}}\right) f(\mathbf{r}, \mathbf{p}, t) = \left(\frac{\partial f}{\partial t}\right)_{\text{coll}} \tag{3.8}$$

where $\nabla_{\mathbf{r}}, \nabla_{\mathbf{p}}$ are, respectively, the gradient operators with respect to $\mathbf{r}$ and $\mathbf{p}$. This equation is not meaningful until we explicitly specify $(\partial f/\partial t)_{\text{coll}}$. It is in specifying this term that the assumption that the system is a dilute gas becomes relevant.

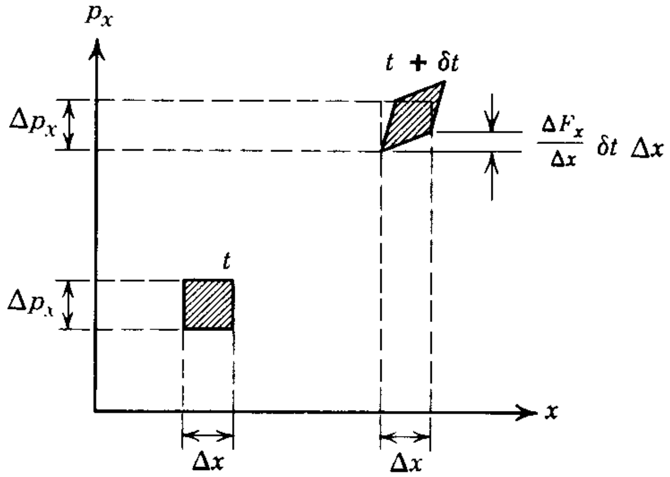


Fig. 3.2 The invariance of the volume element in μ space under dynamical evolution in time.

An explicit form for $(\partial f/\partial t)_{\text{coll}}$ can be obtained by going back to its definition (3.7). Consider Fig. 3.3, where the square labeled A represents the μ-space volume element at $\{\mathbf{r}, \mathbf{p}, t\}$ and the one labeled B represents that at $\{\mathbf{r} + \mathbf{v}\,\delta t, \mathbf{p} + \mathbf{F}\,\delta t, t + \delta t\}$, where δt eventually tends to zero. During the time interval δt, some molecules in A will be removed from A by collision. We regard A as so small that *any* collision that a molecule in A suffers will knock it out of A. Such a molecule with not reach B. On the other hand, there are molecules outside A which, through collisions, will get into A during the time interval δt. These will be in B. Therefore the number of molecules in B at $t + \delta t$, as $\delta t \to 0$, equals the original number of molecules in A at time t plus the *net* gain of molecules in A due to collisions during the time interval δt. This statement is the content of (3.7), and may be expressed in the form

$$\left(\frac{\partial f}{\partial t}\right)_{\text{coll}} \delta t = (\bar{R} - R)\,\delta t \tag{3.9}$$

where $R\,\delta t\,d^3r\,d^3p$ = no. of collisions occurring during the time between t and $t + \delta t$ in which one of the *initial* molecules is in $d^3r\,d^3p$ about $(\mathbf{r}, \mathbf{p})$ (3.10)

$\bar{R}\,\delta t\,d^3r\,d^3p$ = no. of collisions occurring during the time between t and $t + \delta t$, in which one of the *final* molecules is in $d^3r\,d^3p$ about $(\mathbf{r}, \mathbf{p})$ (3.11)

Strictly speaking, we make a small error here. For example, in (3.10), we are implicitly assuming that if a molecule qualifies under the description, none of its

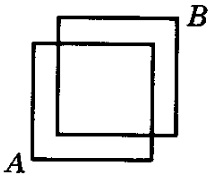


Fig. 3.3 A volume element in μ space at the times t and $t + \delta t$.

partners in collision qualifies. This error is negligible because of the smallness of d^3p.

To proceed further, we assume that the gas is extremely dilute, so that we may consider only binary collisions and ignore the possibility that three or more molecules may collide simultaneously. This considerably simplifies the evaluation of R and $\overline{R}$. It is thus natural to study the nature of binary collisions next.

3.2 BINARY COLLISIONS

We consider an elastic collision in free space between two spinless molecules of respective masses m_1 and m_2. The momenta of the molecules in the initial state are denoted by $\mathbf{p}_1$ and $\mathbf{p}_2$, respectively, and the energies by ϵ_1 and ϵ_2, with $\epsilon_i = p_i^2/2m_i$. The corresponding quantities in the final state are indicated by a prime. Momentum and energy conservation require that

$$\begin{aligned}\mathbf{p}_1 + \mathbf{p}_2 &= \mathbf{p}_1' + \mathbf{p}_2' \\ E = \epsilon_1 + \epsilon_2 &= \epsilon_1' + \epsilon_2' \end{aligned} \tag{3.12}$$

where E is the total energy. We define the total mass M and reduced mass μ by

$$M \equiv m_1 + m_2, \qquad \mu \equiv \frac{m_1 m_2}{m_1 + m_2} \tag{3.13}$$

and the total momentum $\mathbf{P}$ and relative momentum $\mathbf{p}$ by

$$\begin{aligned}\mathbf{P} &\equiv \mathbf{p}_1 + \mathbf{p}_2 \\ \mathbf{p} &\equiv \frac{m_2\mathbf{p}_1 - m_1\mathbf{p}_2}{m_1 + m_2} = \mu(\mathbf{v}_1 - \mathbf{v}_2)\end{aligned} \tag{3.14}$$

where $\mathbf{v}_i = \mathbf{p}_i/m_i$ is the velocity. Solving these for $\mathbf{p}_1$ and $\mathbf{p}_2$ gives

$$\mathbf{p}_1 = \frac{m_1}{M}\mathbf{P} - \mathbf{p}, \qquad \mathbf{p}_2 = \frac{m_2}{M}\mathbf{P} + \mathbf{p} \tag{3.15}$$

It can be easily verified that the total energy is given by

$$E = \frac{P^2}{2M} + \frac{p^2}{2\mu} \tag{3.16}$$

The conditions for momentum-energy conservation become simply

$$\mathbf{P} = \mathbf{P}', \qquad |\mathbf{p}| = |\mathbf{p}'| \tag{3.17}$$

That is, the collision merely rotates the relative momentum without changing its magnitude. Let the angle between $\mathbf{p}'$ and $\mathbf{p}$ be θ, and the azimuthal angle of $\mathbf{p}'$ about $\mathbf{p}$ be ϕ. These angles completely specify the kinematics of the collision. They are collectively denoted by Ω, and are called the scattering angles. We depict the kinematics geometrically in Fig. 3.4. If the potential responsible for the scattering is a central potential (i.e., dependent only on the magnitude of the distance between the molecules), then the scattering is independent of ϕ.

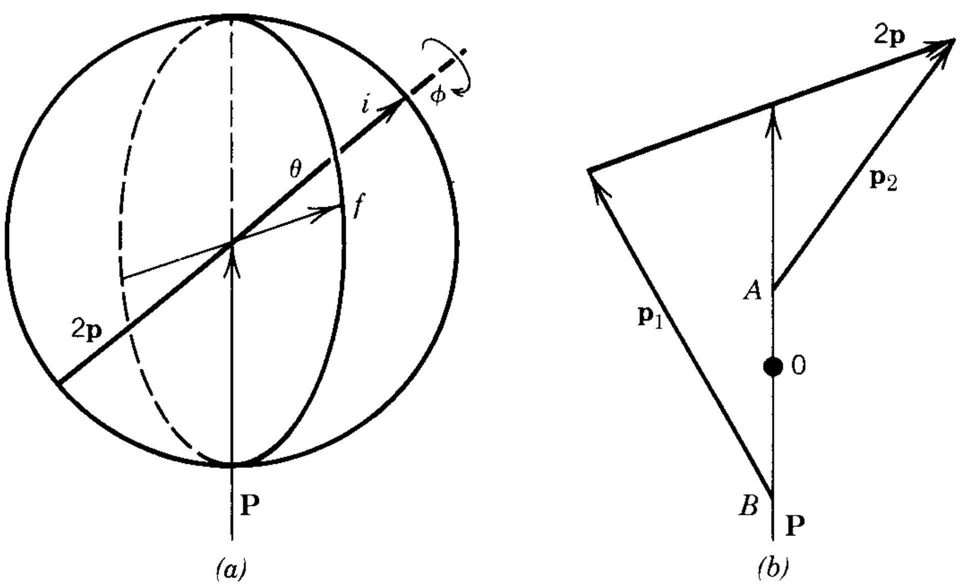


Fig. 3.4 Geometrical representation of the kinematics of a binary collision. (*a*) The total momentum **P** is unchanged (momentum conservation). The relative momentum is rotated from the initial direction *i* to the final direction *f*, without change in magnitude (energy conservation). The scattering angles are the angles θ, ϕ of *f* relative to *i*. (*b*) The individual momenta of the colliding partners p_1 and p_2 may be constructed from **P** and 2**p**, as shown. The points *A* and *B* will coincide at *O*, the midpoint of **P**, if the colliding molecules have equal mass.

The dynamical aspects of the collision are contained in the differential cross section $d\sigma/d\Omega$, which is defined experimentally as follows. Consider a beam of particle 2 incident on particle 1, regarded as the target. The incident flux I is defined as the number of incident particles crossing unit area per second, from the viewpoint of the target:

$$I = n|\mathbf{v}_1 - \mathbf{v}_2| \tag{3.18}$$

where n is the density of particles in the incident beam. The differential cross section $d\sigma/d\Omega$ is defined by the statement

$$I(d\sigma/d\Omega)\,d\Omega \equiv \text{Number of incident molecules scattered per second into the solid angle element } d\Omega \text{ about the direction } \Omega \tag{3.19}$$

The differential cross section has the dimension of area. The number of molecules scattered into $d\Omega$ per second is equal to the number of molecules in the incident beam crossing an area $d\sigma/d\Omega$ per second. The total cross section is the number of molecules scattered per second, regardless of scattering angle:

$$\sigma_{\text{tot}} = \int d\Omega\,\frac{d\sigma}{d\Omega} \tag{3.20}$$

In classical mechanics the differential cross section can be calculated from the intermolecular potential as follows. First we transform the coordinate system to the center-of-mass system, in which the total momentum is zero. Since we are considering only the nonrelativistic domain, this involves a trivial translation of

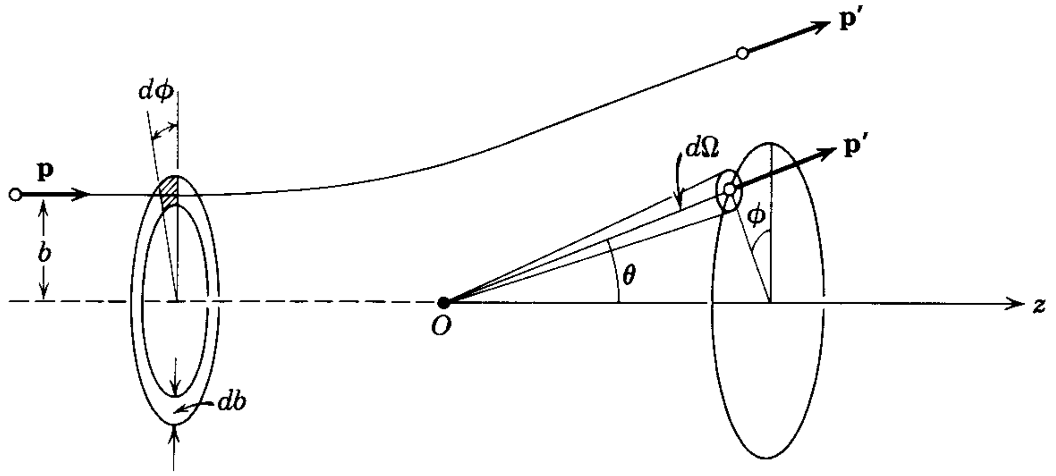


Fig. 3.5 Classical scattering of a molecule by a fixed center of force O.

all velocities by a constant amount. We only need to follow the trajectory of one of the particles, which will move along an orbit, as if it were scattered by a fixed center of force O, as illustrated in Fig. 3.5. It approaches O with momentum $\mathbf{p}$, the relative momentum, and will recede from O with momentum $\mathbf{p}'$, the rotated relative momentum. The normal distance between the line of approach and O is called the impact parameter b. By conservation of angular momentum, it is also the normal distance between the line of recession and O. This is indicated in Fig. 3.5, together with the scattering angles. From the geometry it is clear that

$$I\frac{d\sigma}{d\Omega}\,d\Omega = Ib\,db\,d\phi \tag{3.21}$$

We can find the relation between b and the scattering angles from the classical orbit equation, thereby obtaining $d\sigma/d\Omega$ as a function of the scattering angles.*

The use of classical mechanics to calculate the differential cross section in this problem is an old tradition (started by Maxwell) predating quantum mechanics. To be correct, however, we must use quantum mechanics, notwithstanding the fact that between collisions we regard the molecules as classical particles. The reason is that when the molecules collide their wave functions necessarily overlap, and they see each other as plane waves of definite momenta rather than wave packets of well-defined positions. Furthermore, formulating the scattering problem quantum mechanically makes the kinematics and symmetries of the problem more obvious.

In quantum mechanics the basic quantity in a scattering problem is the transition matrix (T matrix), whose elements are the matrix elements of a certain operator $T(E)$ between the initial (i) and final (f) state:

$$\begin{aligned} T_{\mathrm{fi}} &= \langle 1', 2' | T(E) | 1, 2 \rangle \\ T(E) &= \mathscr{H}' + \mathscr{H}'(E - \mathscr{H}_0 + i\epsilon)^{-1}\mathscr{H}' + \cdots \end{aligned} \tag{3.22}$$

where $\mathscr{H}_0$ is the unperturbed Hamiltonian, $\mathscr{H}'$ the potential, and $\epsilon \to 0^+$. A

---

*See any book on classical mechanics, for example, L. D. Landau and E. M. Lifshitz, *Mechanics* (Pergamon, Oxford, 1960), Chapter IV.

collision is a transition from the initial state to a set of final states. For final states in the infinitesimal momentum-space element $d^3p_1'\,d^3p_2'$, the rate is*

$$\begin{aligned} dP_{12\to 1'2'} &= I\,d\sigma = d^3p_1'\,d^3p_2'\,\delta^4(P_\mathrm{f}-P_\mathrm{i})|T_\mathrm{fi}|^2 \\ \delta^4(P_\mathrm{f}-P_\mathrm{i}) &\equiv \delta^3(\mathbf{P}-\mathbf{P}')\delta(E-E') \end{aligned} \tag{3.23}$$

The transition rate into any region of momentum space can be obtained by integrating the above over the appropriate region. To obtain the differential cross section, integrate over the recoil momentum $\mathbf{p}_1$ (which is fixed by momentum conservation) and the magnitude p_2' (which is determined by energy conservation) to obtain

$$I\frac{d\sigma}{d\Omega} = \int dp_2'\,p_2'^2 \int d^3p_1'\,\delta^4(P_\mathrm{f}-P_\mathrm{i})|T_\mathrm{fi}|^2 \tag{3.24}$$

The integrations are trivial, and yield a factor representing the density of final states because of the δ functions that enforce momentum-energy conservation. For the formal manipulations that we are going to do, however, it is best to leave the integrations undone.

The T matrix is invariant under spatial rotations and reflections and under time reversal, because all molecular interactions originate in the electromagnetic interaction, which have these invariance properties. Explicitly, we have

$$\begin{aligned} \langle \mathbf{p}_2',\mathbf{p}_1'|T|\mathbf{p}_1,\mathbf{p}_2\rangle &= \langle R\mathbf{p}_2', R\mathbf{p}_1'|T|R\mathbf{p}_1, R\mathbf{p}_2\rangle \\ \langle \mathbf{p}_2',\mathbf{p}_1'|T|\mathbf{p}_1,\mathbf{p}_2\rangle &= \langle -\mathbf{p}_2, -\mathbf{p}_1|T|-\mathbf{p}_1', -\mathbf{p}_2'\rangle \end{aligned} \tag{3.25}$$

where $R\mathbf{p}$ is the vector obtained from $\mathbf{p}$ after performing a spatial rotation about an axis, and/or a reflection with respect to a plane. For elastic scattering the density of states are the same in the initial and final states. Thus invariances of the T matrix directly implies corresponding invariances of the differential cross section.

If the molecules have spin, (3.25) remains valid provided we interpret $\mathbf{p}$ to include the spin coordinate. A rotation rotates both the momentum and the spin, but a reflection does not affect the spin. Under time reversal both the momentum and the spin change sign.

From (3.25) we can deduce that the inverse collision, defined as the collision with initial and final states interchanged, has the same T matrix (and hence the same cross section):

$$\langle \mathbf{p}_2',\mathbf{p}_1'|T|\mathbf{p}_1,\mathbf{p}_2\rangle = \langle \mathbf{p}_2,\mathbf{p}_1|T|\mathbf{p}_1',\mathbf{p}_2'\rangle \qquad \text{or} \qquad T_\mathrm{fi} = T_\mathrm{if} \tag{3.26}$$

To show this we represent the collision by the schematic drawing A of Fig. 3.6, which is self-explanatory. The diagram A' beneath it has the same meaning as

---

*A word about normalization. With the definition (3.22), and with single-particle states normalized to one particle per unit volume, there should be a factor $(2\pi\hbar)^3$ multiplying the three-dimensional δ function for momentum conservation, a factor $2\pi/\hbar$ multiplying the δ function for energy conservation, and a factor $(2\pi\hbar)^{-3}$ multiplying each volume element d^3p of momentum space. Since we are not going to calculate cross sections, these factors are a nuisance to write out. We redefine the T matrix appropriately to absorb these factors.

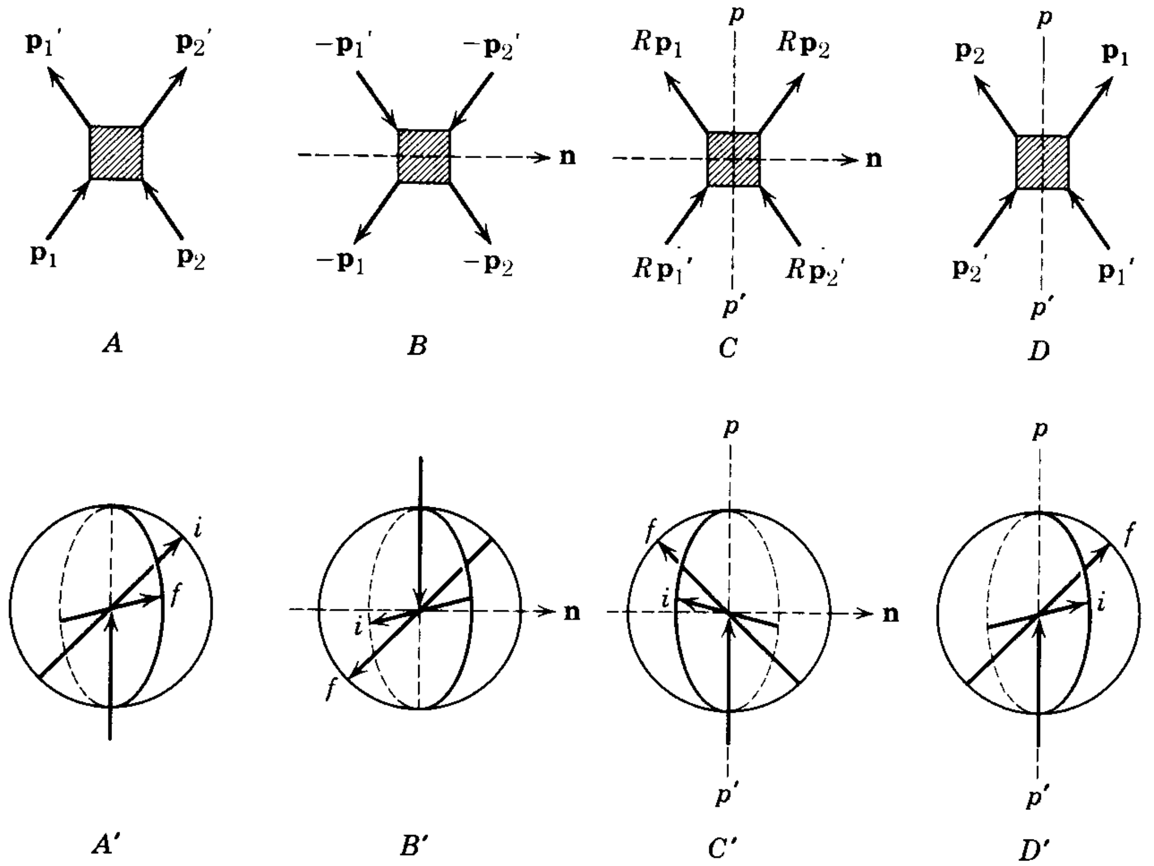


Fig. 3.6 Symmetry operations that take a collision to the inverse collision. In A', B', C', and D', i and f, respectively, denote initial and final relative momenta.

Fig. 3.4. The T matrix for this collision is the same as the time-reversed collision represented by B. Now rotate the coordinate system through 180° about a suitable axis $\mathbf{n}$ perpendicular to the total momentum, and then reflect with respect to a plane pp' perpendicular to $\mathbf{n}$. As a result we obtain the collision D, which is the inverse of the original collision, and which has the same T matrix because of (3.25).

If collisions were treated classically, the inverse collision could be very different from the original collision. As a concrete example consider the classical collision between a sphere and a wedge.* A glance at Fig. 3.7 proves the point. But this is irrelevant for molecular scattering, because molecules are not describable as "wedges" or the like. A nonspherically symmetric molecule is one with nonzero spin, and exists in an eigenstate of the spin. The angular orientation, being conjugate to the angular momentum, is completely uncertain. The symmetry between collision and inverse collision remains valid when the spin is taken into account, as stated previously.

3.3 THE BOLTZMANN TRANSPORT EQUATION

To derive an explicit formula for $(\partial f/\partial t)_{\mathrm{coll}}$, we assume that the gas is sufficiently dilute that only binary collisions need be taken into account. The effect of

---

*Both made of concrete.

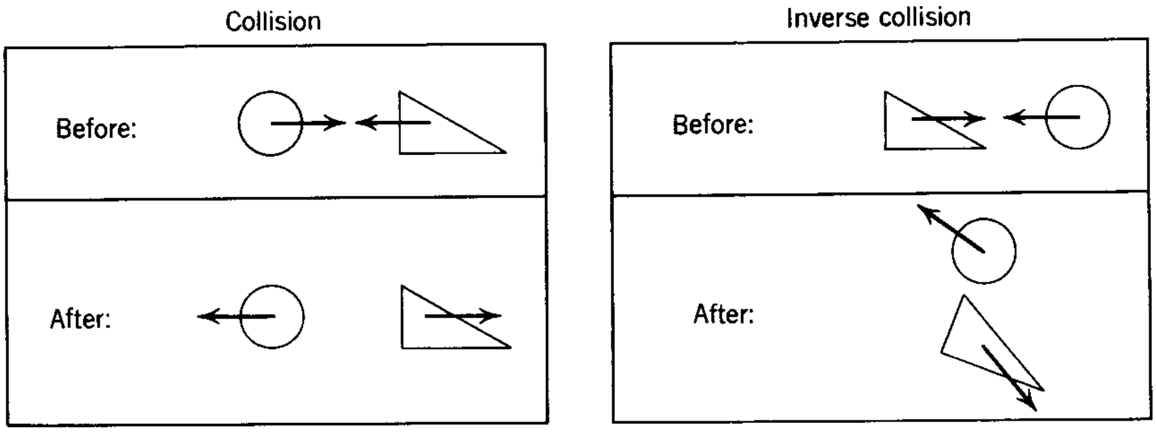


Fig. 3.7 Classical collision between macroscopic objects.

external forces on collisions are ignored on the assumption that these forces, if present, would vary little over the range of the intermolecular potential.

The number of transitions $12 \to 1'2'$ in a volume element d^3r at $\mathbf{r}$, owing to collisions during the time interval δt is

$$dN_{12}\, dP_{12 \to 1'2'}\, \delta t$$

where dN_{12} is the initial number of colliding pairs $(\mathbf{p}_1, \mathbf{p}_2)$. We introduce the two-particle correlation function F by

$$dN_{12} = F(\mathbf{r}, \mathbf{p}_1, \mathbf{p}_2, t)\, d^3r\, d^3p_1\, d^3p_2 \tag{3.27}$$

Thus, in the notation of (3.10), we have

$$R\, \delta t\, d^3r\, d^3p_1 = \delta t\, d^3r\, d^3p_1 \int d^3p_2\, dP_{12 \to 1'2'}\, F(\mathbf{r}, \mathbf{p}_1, \mathbf{p}_2, t) \tag{3.28}$$

Using (3.23), we obtain

$$R = \int d^3p_2\, d^3p_1'\, d^3p_2'\, \delta^4(P_\mathrm{f} - P_\mathrm{i})|T_\mathrm{fi}|^2 F(\mathbf{r}, \mathbf{p}_1, \mathbf{p}_2, t) \tag{3.29}$$

Similarly, we find

$$\bar{R} = \int d^3p_2\, d^3p_1'\, d^3p_2'\, \delta^4(P_\mathrm{i} - P_\mathrm{f})|T_\mathrm{if}|^2 F(\mathbf{r}, \mathbf{p}_1', \mathbf{p}_2', t) \tag{3.30}$$

The δ functions in (3.29) and (3.30) are identical, and $T_\mathrm{fi} = T_\mathrm{if}$ by (3.25). Hence

$$\left(\frac{\partial f}{\partial t}\right)_\mathrm{coll} = \bar{R} - R = \int d^3p_2\, d^3p_1'\, d^3p_2'\, \delta^4(P_\mathrm{f} - P_\mathrm{i})|T_\mathrm{fi}|^2 (F_{1'2'} - F_{12}) \tag{3.31}$$

where $F_{12} = F(\mathbf{r}, \mathbf{p}_1, \mathbf{p}_2, t)$. Note that we can integrate over the vector $\mathbf{p}_1'$ and the magnitude p_2, so that the differential cross section appears in the integrand of (3.30):

$$\left(\frac{\partial f}{\partial t}\right)_\mathrm{coll} = \int d^3p_2\, d\Omega\, |\mathbf{v}_1 - \mathbf{v}_2| (d\sigma/d\Omega)(F_{1'2'} - F_{12}) \tag{3.32}$$

For formal manipulations, however, it is more convenient to leave it in the form (3.31).

The expression we obtained is exact for a sufficiently dilute gas. But it contains the unknown correlation function F.

We now introduce the crucial assumption

$$F(\mathbf{r}, \mathbf{p}_1, \mathbf{p}_2, t) \approx f(\mathbf{r}, \mathbf{p}_1, t) f(\mathbf{r}, \mathbf{p}_2, t) \tag{3.33}$$

This says that the momenta of two particles in the volume element d^3r are uncorrelated, so that the probability of finding them simultaneously is the product of the probability of finding each alone. This is known as the "assumption of molecular chaos." It is necessary to obtain a closed equation for the distribution function, but there is otherwise no justification for it at this point. We shall come back to analyze its meaning later.

With the assumption of molecular chaos, we have

$$\left(\frac{\partial f_1}{\partial t}\right)_{\text{coll}} = \int d^3p_2\, d\Omega |\mathbf{v}_1 - \mathbf{v}_2| (d\sigma/d\Omega)(f_1' f_2' - f_1 f_2) \tag{3.34}$$

where the following abbreviations have been used:

$$\begin{aligned} f_1 &\equiv f(\mathbf{r}, \mathbf{p}_1, t) \\ f_2 &\equiv f(\mathbf{r}, \mathbf{p}_2, t) \\ f_1' &\equiv f(\mathbf{r}, \mathbf{p}_1', t) \\ f_2' &\equiv f(\mathbf{r}, \mathbf{p}_2', t) \end{aligned} \tag{3.35}$$

Substituting (3.34) into (3.8) we obtain the Boltzmann transport equation

$$\left(\frac{\partial}{\partial t} + \frac{\mathbf{p}_1}{m} \cdot \nabla_r + \mathbf{F} \cdot \nabla_{p_1}\right) f_1 = \int d^3p_2\, d^3p_1'\, d^3p_2'\, \delta^4(P_{\text{f}} - P_{\text{i}}) |T_{\text{fi}}|^2 (f_2' f_1' - f_2 f_1) \tag{3.36}$$

which is a nonlinear integro-differential equation for the distribution function.

We have considered only the case of a single species of spinless molecules. If we consider different types of molecules, then we have to introduce a separate distribution function for each type, and the collision term will couple them if different types of molecules can scatter each other. If the molecules have spin, or if we consider excitation of the molecules through scattering, then the different spin states or excited states should be considered as different species of molecules.

3.4 THE GIBBSIAN ENSEMBLE

Gibbs introduced the idea of a statistical ensemble to describe a macroscopic system, which has proved to be a very important concept. We shall use it here to present another approach to the Boltzmann transport equation.

A state of the gas under consideration can be specified by the $3N$ canonical coordinates $q_1, \ldots, q_{3N}$ and their conjugate momenta $p_1, \ldots, p_{3N}$. The $6N$-dimensional space spanned by $\{p_i, q_i\}$ is called the Γ space, or phase space, of the system. A point in Γ space represents a state of the entire N-particle system, and is referred to as the *representative point*. This is in contrast to the μ space introduced earlier, which refers to only one particle.

It is obvious that a very large (in fact, infinite) number of states of the gas corresponds to a given macroscopic condition of the gas. For example, the condition that the gas is contained in a box of volume 1 cm^3 is consistent with an infinite number of ways to distribute the molecules in space. Through macroscopic measurements we would not be able to distinguish between two gases existing in different states (thus corresponding to two distinct representative points) but satisfying the same macroscopic conditions. Thus when we speak of a gas under certain macroscopic conditions, we are in fact referring not to a single state, but to an infinite number of states. In other words, we refer not to a single system, but to a collection of systems, identical in composition and macroscopic condition but existing in different states. With Gibbs, we call such a collection of systems an *ensemble*, which is geometrically represented by a distribution of representative points in Γ space, usually a continuous distribution. It may be conveniently described by a density function $\rho(p, q, t)$, where (p, q) is an abbreviation for $(p_1, \ldots, p_{3N}; q_1, \ldots, q_{3N})$, so defined that

$$\rho(p, q, t)\, d^{3N}p\, d^{3N}q \tag{3.37}$$

is the number of representative points that at time t are contained in the infinitesimal volume element $d^{3N}p\, d^{3N}q$ of Γ space centered about the point (p, q). An ensemble is completely specified by $\rho(p, q, t)$. It is to be emphasized that members of an ensemble are mental copies of a system and do not interact with one another.

Given $\rho(p, q, t)$ at any time t, its subsequent values are determined by the dynamics of molecular motion. Let the Hamiltonian of a system in the ensemble be $\mathscr{H}(p_1, \ldots, p_{3N}; q_1, \ldots, q_{3N})$. The equations of motion for a system are given by

$$\begin{aligned} \dot{p}_i &= -\frac{\partial \mathscr{H}}{\partial q_i} \qquad (i = 1, \ldots, 3N) \\ \dot{q}_i &= \frac{\partial \mathscr{H}}{\partial p_i} \qquad (i = 1, \ldots, 3N) \end{aligned} \tag{3.38}$$

These will tell us how a representative point moves in Γ space as time evolves. We assume that the Hamiltonian does not depend on any time derivative of p and q. It is then clear that (3.38) is invariant under time reversal and that (3.38) uniquely determines the motion of a representative point for all times, when the position of the representative point is given at any time. It follows immediately from these observations that the locus of a representative point is either a simple closed curve or a curve that never intersects itself. Furthermore, the loci of two distinct representative points never intersect.

We now prove the following theorem.

LIOUVILLE'S THEOREM

$$\frac{\partial\rho}{\partial t}+\sum_{i=1}^{3N}\left(\frac{\partial\rho}{\partial p_i}\dot{p}_i+\frac{\partial\rho}{\partial q_i}\dot{q}_i\right)=0 \tag{3.39}$$

Proof Since the total number of systems in an ensemble is conserved, the number of representative points leaving any volume in Γ space per second must be equal to the rate of decrease of the number of representative points in the same volume. Let ω be an arbitrary volume in Γ space and let S be its surface. If we denote by **v** the $6N$-dimensional vector whose components are

$$\mathbf{v}\equiv(\dot{p}_1,\dot{p}_2,\ldots,\dot{p}_{3N};\dot{q}_1,\dot{q}_2,\ldots,\dot{q}_{3N})$$

and **n** the vector locally normal to the surface S, then

$$-\frac{d}{dt}\int_\omega d\omega\,\rho=\int_S dS\,\mathbf{n}\cdot\mathbf{v}\rho$$

With the help of the divergence theorem in $6N$-dimensional space, we convert this to the equation

$$\int_\omega d\omega\left[\frac{\partial\rho}{\partial t}+\nabla\cdot(\mathbf{v}\rho)\right]=0 \tag{3.40}$$

where ∇ is the $6N$-dimensional gradient operator:

$$\nabla\equiv\left(\frac{\partial}{\partial p_1},\frac{\partial}{\partial p_2},\ldots,\frac{\partial}{\partial p_{3N}};\frac{\partial}{\partial q_1},\frac{\partial}{\partial q_2},\ldots,\frac{\partial}{\partial q_{3N}}\right)$$

Since ω is an arbitrary volume the integrand of (3.40) must identically vanish. Hence

$$\begin{aligned}-\frac{\partial\rho}{\partial t}=\nabla\cdot(\mathbf{v}\rho)&=\sum_{i=1}^{3N}\left[\frac{\partial}{\partial p_i}(\dot{p}_i\rho)+\frac{\partial}{\partial q_i}(\dot{q}_i\rho)\right]\\&=\sum_{i=1}^{3N}\left(\frac{\partial\rho}{\partial p_i}\dot{p}_i+\frac{\partial\rho}{\partial q_i}\dot{q}_i\right)+\sum_{i=1}^{3N}\rho\left(\frac{\partial\dot{p}_i}{\partial p_i}+\frac{\partial\dot{q}_i}{\partial q_i}\right)\end{aligned}$$

By the equations of motion (3.38) we have

$$\frac{\partial\dot{p}_i}{\partial p_i}+\frac{\partial\dot{q}_i}{\partial q_i}=0\qquad(i=1,\ldots,3N)$$

Therefore

$$-\frac{\partial\rho}{\partial t}=\sum_{i=1}^{3N}\left(\frac{\partial\rho}{\partial p_i}\dot{p}_i+\frac{\partial\rho}{\partial q_i}\dot{q}_i\right)$$ ∎

Liouville's theorem is equivalent to the statement

$$\frac{d\rho}{dt}=0 \tag{3.41}$$

since by virtue of the equations of motion p_i and q_i are functions of the time. Its geometrical interpretation is as follows. If we follow the motion of a representative point in Γ space, we find that the density of representative points in its neighborhood is constant. Hence the distribution of representative points moves in Γ space like an incompressible fluid.

The observed value of a dynamical quantity O of the system, which is generally a function of the coordinates and conjugate momenta, is supposed to be its averaged value taken over a suitably chosen ensemble:

$$\langle O \rangle = \frac{\int d^{3N}p\, d^{3N}q\, O(p,q)\rho(p,q,t)}{\int d^{3N}p\, d^{3N}q\, \rho(p,q,t)} \tag{3.42}$$

This is called the ensemble average of O. Its time dependence comes from that of ρ, which is governed by Liouville's theorem. In principle, then, this tells us how a quantity approaches equilibrium—the central question of kinetic theory. In the next section we shall derive the Boltzmann transport equation using this approach.

Under certain conditions one can prove an *ergodic theorem*, which says that if one waits a sufficiently long time, the locus of the representative point of a system will cover the entire accessible phase space. More precisely, it says that the representative point comes arbitrarily close to any point in the accessible phase space. This would indicate that the ensemble corresponding to thermodynamic equilibrium is one for which ρ is constant over the accessible phase space. This is actually what we shall assume.*

3.5 THE BBGKY HIERARCHY

One can define correlation functions f_s, which give the probability of finding s particles having specified positions and momenta, in the systems forming an ensemble. The function f_1 is the familiar distribution function. The exact equations of motion for f_s in classical mechanics can be written down. They show that to find f_1 we need to know f_2, which in turns depends on a knowledge of f_3, and so on till we come the full N-body correlation function f_N. This system of equations is known as the BBGKY† hierarchy. We shall derive it and show how the chain of equations can be truncated to yield the Boltzmann transport equation. The "derivation" will not be any more rigorous than the one already given, but it will give new insight into the nature of the approximations.

Consider an ensemble of systems, each being a gas of N molecules enclosed in volume V, with Hamiltonian $\mathscr{H}$. Instead of the general notation $\{p_i, q_i\}$

---

*See the remarks about the relevance of the ergodic theorem in Section 4.5.

†BBGKY stands for Bogoliubov-Born-Green-Kirkwood-Yvon. For a detailed discussion and references see N. N. Bogoliubov in *Studies in Statistical Mechanics*, J. de Boer and G. E. Uhlenbeck, Eds., Vol. I (North-Holland, Amsterdam, 1962).

($i = 1, \ldots, 3N$), we shall denote the coordinates by the Cartesian vectors $\{\mathbf{p}_i, \mathbf{r}_i\}$ ($i = 1, \ldots, N$), for which we use the abbreviation

$$z_i = (\mathbf{p}_i, \mathbf{r}_i), \qquad \int dz_i = \int d^3p_i \, d^3r_i \tag{3.43}$$

The density function characterizing the ensemble is denoted by $\rho(1, \ldots, N, t)$, and assumed to be symmetric in $z_1, \ldots, z_N$. Its integral over all phase space is a constant by Liouville's theorem; hence we can normalized it to unity:

$$\int dz_1 \cdots dz_N \, \rho(1, \ldots, N, t) = 1 \tag{3.44}$$

Thus the ensemble average of any function $O(1, \ldots, N)$ of molecular coordinates can be written as

$$\langle O \rangle \equiv \int dz_1 \cdots dz_N \, \rho(1, \ldots, N, t) O(1, \ldots, N) \tag{3.45}$$

Using the Hamiltonian equations of motion (3.38), we rewrite Liouville's theorem in the form

$$\frac{\partial \rho}{\partial t} = \sum_{i=1}^{N} \left(\nabla_{p_i} \rho \cdot \nabla_{r_i} \mathscr{H} - \nabla_{r_i} \rho \cdot \nabla_{p_i} \mathscr{H} \right) \tag{3.46}$$

Assume that the Hamiltonian is of the form

$$\begin{aligned} \mathscr{H} &= \sum_{i=1}^{N} \frac{\mathbf{p}^2}{2m} + \sum_{i=1}^{N} U_i + \sum_{i<j} v_{ij} \\ U_i &= U(\mathbf{r}_i) \\ v_{ij} &= v_{ji} = v(|\mathbf{r}_i - \mathbf{r}_j|) \end{aligned} \tag{3.47}$$

Then

$$\begin{aligned} \nabla_{p_i} \mathscr{H} &= \frac{\mathbf{p}_i}{m} \\ \nabla_{r_i} \mathscr{H} &= -\mathbf{F}_i - \sum_{\substack{j=1 \\ (j \neq i)}}^{N} \mathbf{K}_{ij} \end{aligned} \tag{3.48}$$

where

$$\begin{aligned} \mathbf{F}_i &= -\nabla_{r_i} U(\mathbf{r}_i) \\ \mathbf{K}_{ij} &= -\nabla_{r_i} v(|\mathbf{r}_i - \mathbf{r}_j|) \end{aligned} \tag{3.49}$$

Liouville's theorem can now be cast in the form

$$\left[\frac{\partial}{\partial t} + h_N(1, \ldots, N) \right] \rho(1, \ldots, N) = 0 \tag{3.50}$$

where

$$h_N(1,\dots,N) = \sum_{i=1}^{N} S_i + \frac{1}{2}\sum_{\substack{i,j=1\\(i\neq j)}}^{N} P_{ij}$$

$$S_i \equiv \frac{\mathbf{p}_i}{m}\cdot\nabla_{r_i} + \mathbf{F}_i\cdot\nabla_{p_i} \tag{3.51}$$

$$P_{ij} \equiv \mathbf{K}_{ij}\cdot\nabla_{p_i} + \mathbf{K}_{ji}\cdot\nabla_{p_j} = \mathbf{K}_{ij}\cdot\left(\nabla_{p_i} - \nabla_{p_j}\right)$$

The single-particle distribution function is defined by

$$f_1(\mathbf{p},\mathbf{r},t) \equiv \left\langle \sum_{i=1}^{N} \delta^3(\mathbf{p}-\mathbf{p}_i)\delta^3(\mathbf{r}-\mathbf{r}_i)\right\rangle = N\int dz_2 \cdots dz_N\, \rho(1,\dots,N,t) \tag{3.52}$$

The factor N in the last form comes from the fact that all terms in the sum in the preceding term have the same value, owing to the fact that ρ is symmetric in $z_1,\dots,z_N$. Integrating f_1 over z_1 yields the correct normalization N, by virtue of (3.44).

The general s-particle distribution function, or correlation function, is defined by

$$f_s(1,\dots,z,t) \equiv \frac{N!}{(N-s)!}\int dz_{s+1}\cdots dz_N\, \rho(1,\dots,N,t) \qquad (s=1,\dots,N) \tag{3.53}$$

The combinatorial factor in front comes from the fact that we do not care which particle is at z_1, which is at z_2, etc. The equation of motion is

$$\frac{\partial}{\partial t} f_s = \frac{N!}{(N-s)!}\int dz_{s+1}\cdots dz_N \frac{\partial}{\partial t}\rho = -\frac{N!}{(N-s)!}\int dz_{s+1}\cdots dz_N\, h_N\rho \tag{3.54}$$

We isolate those terms in h_N involving only the coordinates $z_1,\dots,z_s$:

$$\begin{aligned} h_N(1,\dots,N) &= \sum_{i=1}^{s} S_i + \sum_{s+1}^{N} S_i + \frac{1}{2}\sum_{\substack{i,j=1\\(i\neq j)}}^{s} P_{ij} + \frac{1}{2}\sum_{\substack{i,j=s+1\\(i\neq j)}}^{N} P_{ij} + \sum_{i=1}^{s}\sum_{j=s+1}^{N} P_{ij} \\ &= h_s(1,\dots,s) + h_{N-s}(s+1,\dots,N) + \sum_{i=1}^{s}\sum_{j=s+1}^{N} P_{ij} \end{aligned} \tag{3.55}$$

Note that

$$\int dz_{s+1}\cdots dz_N\, h_{N-s}(s+1,\dots,N)\rho(1,\dots,N) = 0 \tag{3.56}$$

because h_{N-s} consists of gradient terms in $\mathbf{p}$ with $\mathbf{p}$-independent coefficients, and a gradient term in $\mathbf{r}$ with an $\mathbf{r}$-independent coefficient. Thus the integral evaluates ρ on the boundary of phase space, where we assume ρ to vanish. Substituting (3.55) into (3.54), we obtain

$$\begin{aligned}\left(\frac{\partial}{\partial t}+h_s\right)f_s &= -\frac{N!}{(N-s)!}\int dz_{s+1}\cdots dz_N\sum_{i=1}^{s}\sum_{j=s+1}^{N}P_{ij}\rho(1,\ldots,N)\\ &= -\sum_{i=1}^{s}\int dz_{s+1}P_{i,s+1}\frac{N!}{(N-s+1)!}\int dz_{s+2}\cdots dz_N\,\rho(1,\ldots,N)\\ &= -\sum_{i=1}^{s}\int dz_{s+1}P_{i,s+1}f_{s+1}(1,\ldots,s+1)\end{aligned}\tag{3.57}$$

In passing from the first to the second equation we have used the fact that the sum over j gives $N-s$ identical terms. Now substitute P_{ij} from (3.51), and note that the second term there does not contribute, because it leads to a vanishing surface term. We then arrive at

$$\left(\frac{\partial}{\partial t}+h_s\right)f_s(1,\ldots,s) = -\sum_{i=1}^{s}\int dz_{s+1}\mathbf{K}_{i,s+1}\cdot\nabla_{p_i}f_{s+1}(1,\ldots,s+1)$$

$$(s=1,\ldots,N)\tag{3.58}$$

which is the BBGKY hierarchy. The left side of each of the equations above is a "streaming term," involving only the s particles under consideration. For $s>1$ it includes the effect of intermolecular scattering among the s particles. The right-hand side is the "collision integral," which describes the effect of scattering between the particles under consideration with an "outsider," thus coupling f_s to f_{s+1}.

The first two equations in the hierarchy read

$$\left(\frac{\partial}{\partial t}+\frac{\mathbf{p}_1}{m}\cdot\nabla_{r_1}+\mathbf{F}_1\cdot\nabla_{p_1}\right)f_1(z_1,t) = -\int dz_2\,\mathbf{K}_{12}\cdot\nabla_{p_1}f_2(z_1,z_2,t)\tag{3.59}$$

$$\begin{aligned}&\left[\frac{\partial}{\partial t}+\frac{\mathbf{p}_1}{m}\cdot\nabla_{r_1}+\frac{\mathbf{p}_2}{m}\cdot\nabla_{r_2}+\mathbf{F}_1\cdot\nabla_{p_1}+\mathbf{F}_2\cdot\nabla_{p_2}+\tfrac{1}{2}\mathbf{K}_{12}\cdot\left(\nabla_{p_1}-\nabla_{p_2}\right)\right]\\ &\times f_2(z_1,z_2,t)\\ &= -\int dz_3\left(\mathbf{K}_{13}\cdot\nabla_{p_1}+\mathbf{K}_{23}\cdot\nabla_{p_2}\right)f_3(z_1,z_2,z_3,t)\end{aligned}\tag{3.60}$$

The terms in the equations above have dimensions f_s/time, and different time scales are involved:

$$\begin{aligned}\mathbf{K}\cdot\nabla_p &\sim \frac{1}{\tau_c}\\ \mathbf{F}\cdot\nabla_p &\sim \frac{1}{\tau_e}\\ \frac{\mathbf{p}}{m}\cdot\nabla_r &\sim \frac{1}{\tau_s}\end{aligned}\tag{3.61}$$

where τ_c is the duration of a collision, τ_e is the time for a molecule to traverse a characteristic distance over which the external potential varies significantly, and τ_s is the time for a molecule to traverse a characteristic distance over which the correlation function varies significantly. The time τ_c is the shortest, and τ_e the longest.

The equation for f_1 is unique in the hierarchy, in that "streaming" sets a rather slow time scale, for it does not involve intermolecular scattering, (there being only one particle present.) The collision integral, which has more rapid variations, sets the time scale of f_1. This is why the equilibrium condition is determined by the vanishing of the collision integral.

In contrast, the equation for f_2 (and higher ones as well) contains a collision term of the order $1/\tau_c$ on the left side. The collision integral on the right side is smaller by a factor of the order nr_0^3 (where n is the density, and r_0 the range of the intermolecular potential) because the integration of $\mathbf{r}_3$ extends only over a volume of radius r_0. Now $r_0 \approx 10^{-8}$ cm and $n \approx 10^{19}$ cm^{-3} under standard conditions. Hence $nr_0^3 \approx 10^{-5}$. Thus for f_2 (and higher equations too) the time scale is set by the streaming terms instead of the collision integral, which we shall neglect.

With neglect of the right side of (3.60), the hierarchy is truncated at f_2, and we have only two coupled equations for f_1 and f_2:

$$\left(\frac{\partial}{\partial t} + \frac{\mathbf{p}_1}{m} \cdot \nabla_{r_1}\right) f_1(z_1, t) = -\int_{r_0} dz_2\, \mathbf{K}_{12} \cdot \nabla_{p_1} f_2(z_1, z_2, t) \equiv \left(\frac{\partial f_1}{\partial t}\right)_{\text{coll}} \tag{3.62}$$

$$\left[\frac{\partial}{\partial t} + \frac{\mathbf{p}_1}{m} \cdot \nabla_{r_1} + \frac{\mathbf{p}_2}{m} \cdot \nabla_{r_2} + \tfrac{1}{2}\mathbf{K}_{12} \cdot \left(\nabla_{p_1} - \nabla_{p_2}\right)\right] f_2(z_1, z_2, t) = 0 \tag{3.63}$$

where we have set all external forces to zero, for simplicity. We shall also assume for simplicity that the force $\mathbf{K}$ vanishes outside a range r_0. To remind us of this fact, we put the subscript r_0 on the integral in the first equation, indicating that the spatial part of the integral is subject to $|\mathbf{r}_1 - \mathbf{r}_2| < r_0$.

The salient qualitative features of (3.62) and (3.63) are that f_2 varies in time with characteristic period τ_c, and in space with characteristic distance r_0, while f_1 varies much less rapidly, by a factor nr_0^3. Thus f_1 measures space and time with much coarser scales than f_2.

The correlations in f_2 are produced by collisions between particles 1 and 2. When their positions are so far separated as to be out of molecular interaction range, we expect that there will be no correlation between 1 and 2, and f_2 will assume a product form (neglecting, of course, possible correlations produced by collisions with a third particle):

$$f_2(z_1, z_2, t) \xrightarrow[|\mathbf{r}_1 - \mathbf{r}_2| \gg r_0]{} f_1(z, t) f_1(z_2, t) \tag{3.64}$$

To evaluate $(\partial f_1/\partial t)_{\text{coll}}$, however, we need f_2 not in the uncorrelated region, but in the region where the two particles are colliding. To look at this

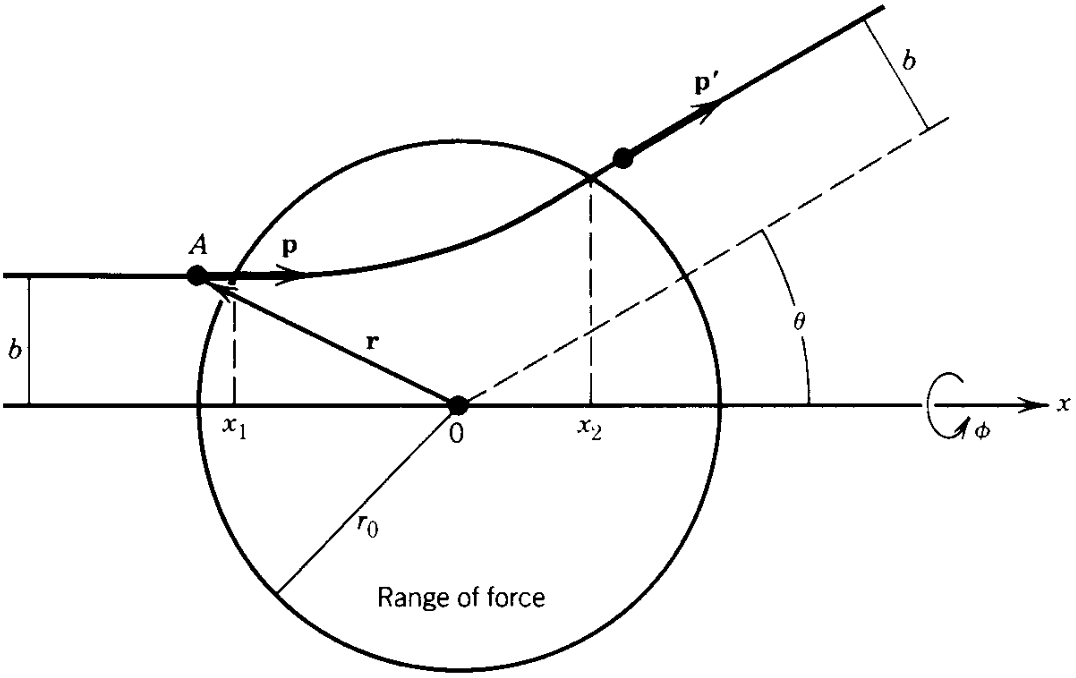


Fig. 3.8 Illustration of behavior of two-particle correlation function. The separation between the two particles is **r**, and the relative momentum **p**. The two particles are correlated only inside the range of the intermolecular force, indicated by the sphere of radius r_0. Outside the sphere, the correlation function is a product of two one-particle distribution functions. In equilibrium there is a steady scattering of beams of particles of all momenta, from all directions, at all impact parameters.

region it is convenient to use total and relative coordinates, defined as follows:

$$\begin{aligned} \mathbf{P} &= \mathbf{p}_2 + \mathbf{p}_1 \qquad & \mathbf{p} &= \frac{\mathbf{p}_2 - \mathbf{p}_1}{2} \\ \mathbf{R} &= \frac{\mathbf{r}_2 + \mathbf{r}_1}{2} \qquad & \mathbf{r} &= \mathbf{r}_2 - \mathbf{r}_1 \end{aligned} \tag{3.65}$$

Then (3.63) becomes

$$\begin{gathered} \left(\frac{\partial}{\partial t} + \frac{\mathbf{p}}{m} \cdot \nabla_R + \frac{\mathbf{p}}{m} \cdot \nabla_r + \mathbf{K}(\mathbf{r}) \cdot \nabla_p\right) f_2(\mathbf{P}, \mathbf{R}, \mathbf{p}, \mathbf{r}, t) = 0 \\ \mathbf{K}(\mathbf{r}) \equiv -\nabla_r v(r) \end{gathered} \tag{3.66}$$

Transform to the center-of-mass system by putting $\mathbf{P} = 0$. The above can then be rewritten, to first order in dt, as the streaming condition (with **P** and **R** suppressed for clarity):

$$f_2\left(\mathbf{p} + \mathbf{K}(\mathbf{r})\, dt, \mathbf{r} + \frac{\mathbf{p}}{m}\, dt, t + dt\right) = f_2(\mathbf{p}, \mathbf{r}, t) \tag{3.67}$$

It traces the classical trajectories in the force field **K** centered at O, as illustrated in Fig. 3.8. If f_2 were peaked at point A initially, then (3.67) says that as time goes on the peak will move along the trajectory for that particular initial condition.

The equilibrium situation, for which $\partial f_2/\partial t = 0$, is a steady-state scattering, by the force field **K**, of a beam of particles consisting of all momenta, at all impact parameters. Referring to Fig. 3.8, we may describe the steady state as follows: Outside the sphere of interaction the uncorrelated factorized form (3.64) holds. However, boundary values of the momenta are correlated through the fact that momenta entering the sphere at a specific impact parameter must leave the sphere at the correct scattering angle, and vice versa.

To "derive" the Boltzmann transport equation, we assume that, since f_2 has a shorter time scale than f_1, it reaches equilibrium earlier than f_1. Thus we set $\partial f_2/\partial t = 0$, and assume f_2 has attained the equilibrium form described earlier. Similarly, we assume that the range of force r_0 is essentially zero from the point of view of f_1. Thus in the factorized form of f_2 just before and after a collision, we can put $\mathbf{r}_2$ and $\mathbf{r}_1$ both equal to the same value.

With this in mind, we substitute (3.63) into (3.62) to obtain

$$\begin{aligned}\left(\frac{\partial f_1}{\partial t}\right)_{\text{coll}} &\equiv -\int_{r_0} dz_2\, \mathbf{K}_{12} \cdot \nabla_{p_1} f_2(z_1, z_2, t) \\ &= -\int_{r_0} dz_2\, \mathbf{K}_{12} \cdot (\nabla_{p_1} - \nabla_{p_2}) f_2(z_1, z_2, t) \\ &= \frac{1}{m}\int_{r_0} dz_2\, (\mathbf{p}_1 \cdot \nabla_{r_1} + \mathbf{p}_2 \cdot \nabla_{r_2}) f_2(z_1, z_2, t) \qquad (3.68)\end{aligned}$$

Using the coordinates defined in (3.66), and neglecting the gradient with respect to **R**, we have

$$\begin{aligned}\left(\frac{\partial f_1}{\partial t}\right)_{\text{coll}} &= -\frac{1}{m}\int d^3p_2 \int_{r<r_0} d^3r\, (\mathbf{p}_1 - \mathbf{p}_2) \cdot \nabla_r f_2 \\ &= \frac{1}{m}\int d^3p_2\, |\mathbf{p}_1 - \mathbf{p}_2| \int d\phi\, b\, db \int_{x_1}^{x_2} dx\, \frac{\partial}{\partial x} f_2 \qquad (3.69)\end{aligned}$$

where the notation is indicated in Fig. 3.8. Now we set

$$\begin{aligned} f_2(x_1) &= f_1(\mathbf{p}_1) f_1(\mathbf{p}_2) \\ f_2(x_2) &= f_1(\mathbf{p}_1') f_1(\mathbf{p}_2') \end{aligned}$$

where $\mathbf{p}_1', \mathbf{p}_2'$ are the final momenta in the scattering process, when the initial momenta are $\mathbf{p}_1$, $\mathbf{p}_2$ and the impact parameter is b. Using the definition (3.21) of the classical cross section, we finally have

$$\left(\frac{\partial f_1}{\partial t}\right)_{\text{coll}} = \int d^3p_2\, d\Omega\, |\mathbf{v}_1 - \mathbf{v}_2| (d\sigma/d\Omega)(f_1' f_2' - f_1 f_2) \qquad (3.70)$$

which is the same as (3.34).

PROBLEMS

3.1 Give a few numerical examples to show that the condition (3.1) is fulfilled for physical gases at room temperatures.

3.2 Explain qualitatively why all molecular interactions are electromagnetic in origin.

3.3 For the collision between perfectly elastic spheres of diameter a,

(*a*) calculate the differential cross section with classical mechanics in the coordinate system in which one of the spheres is initially at rest;

(*b*) compare your answer with the quantum mechanical result. Consider both the low-energy and the high-energy limit. (See, e.g., L. I. Schiff, *Quantum Mechanics*, 2nd ed. (McGraw-Hill, New York, 1955), p. 110).

3.4 Consider a mixture of two gases whose molecules have masses m and M, respectively, and which are subjected to external forces $\mathbf{F}$ and $\mathbf{Q}$, respectively. Denote the respective distribution functions by f and g. Assuming that only binary collisions between molecules are important, derive the Boltzmann transport equation for the system.

3.5 This problem illustrates in a trivial case how the ensemble density tends to a uniform density over the accessible phase space. Consider an ensemble of systems, each of which consists of a single free particle in one dimension with momentum p and coordinate q. The particle is confined to a one-dimensional box with perfectly reflecting walls located at $q = -1$ and $q = 1$ (in arbitrary units.) Draw a square box of unit sides in the pq plane (the phase space). Draw a square of sides $1/2$ in the upper left corner of this box. Let the initial ensemble correspond to filling this corner box uniformly with representative points.

(*a*) What is the accessible part of the phase space? (i.e., the region that the representative points can reach through dynamical evolution from the initial condition.)

(*b*) Consider how the shape of the distribution of representative point changes at regular successive time intervals. How does the distribution look after a long time?

Suggestion: When a particle is being reflected at a wall, its momentum changes sign. Represent what happens in phase space by continuing the locus of the representative point to a fictitious adjacent box in pq space, as if the wall were absent. "Fold" the adjacent box onto the original box properly to get the actual trajectory of the representative point. After a long time, you need many such adjacent boxes. The "folding back" will then give you a picture of the distribution.

CHAPTER 4

THE EQUILIBRIUM STATE OF A DILUTE GAS

4.1 BOLTZMANN'S *H* THEOREM

We define the equilibrium distribution function as the solution of the Boltzmann transport equation that is independent of time. We shall see that it is also the limiting form of the distribution function as the time tends to infinity. Assume that there is no external force. It is then consistent to assume further that the distribution function is independent of $\mathbf{r}$ and hence can be denoted by $f(\mathbf{p}, t)$. The equilibrium distribution function, denoted by $f_0(\mathbf{p})$, is the solution to the equation $\partial f(\mathbf{p}, t)/\partial t = 0$. According to the Boltzmann transport equation (3.36), $f_0(\mathbf{p})$ satisfies the integral equation

$$0 = \int d^3p_2\, d^3p_1'\, d^3p_2'\, \delta^4(P_\mathrm{f} - P_\mathrm{i})|T_{\mathrm{fi}}|^2(f_2'f_1' - f_2f_1) \tag{4.1}$$

where p_1 is a given momentum.

A sufficient condition for $f_0(\mathbf{p})$ to solve (4.1) is

$$f_0(\mathbf{p}_2')f_0(\mathbf{p}_1') - f_0(\mathbf{p}_2)f_0(\mathbf{p}_1) = 0 \tag{4.2}$$

where $\{\mathbf{p}_1, \mathbf{p}_2\} \to \{\mathbf{p}_1', \mathbf{p}_2'\}$ is any possible collision (i.e., one with nonvanishing cross section). We show that this condition is also necessary, and we thus arrive at the interesting conclusion that $f_0(p)$ is independent of $d\sigma/d\Omega$, as long as the latter is nonzero.

To show the necessity of (4.2) we define with Boltzmann the functional

$$H(t) \equiv \int d^3v\, f(\mathbf{p}, t) \log f(\mathbf{p}, t) \tag{4.3}$$

where $f(\mathbf{p}, t)$ is the distribution function at time t, satisfying

$$\frac{\partial f(p_1, t)}{\partial t} = \int d^3p_2\, d^3p_1'\, d^3p_2'\, \delta^4(P_\mathrm{f} - P_\mathrm{i})|T_{\mathrm{fi}}|^2(f_2'f_1' - f_2f_1) \tag{4.4}$$

Differentiation of (4.3) yields

$$\frac{dH(t)}{dt} = \int d^3v \frac{\partial f(\mathbf{p}, t)}{\partial t}[1 + \log f(\mathbf{p}, t)] \tag{4.5}$$

Therefore $\partial f/\partial t = 0$ implies $dH/dt = 0$. This means that a necessary condition for $\partial f/\partial t = 0$ is $dH/dt = 0$. We now show that the statement

$$\frac{dH}{dt} = 0 \tag{4.6}$$

is the same as (4.2). It would then follow that (4.2) is also a necessary condition for the solution of (4.1). To this end we prove the following theorem.

BOLTZMANN'S *H* THEOREM

If f satisfies the Boltzmann transport equation, then

$$\frac{dH(t)}{dt} \le 0 \tag{4.7}$$

Proof Substituting (4.4) into the integrand of (4.5) we have*

$$\frac{dH}{dt} = \int d^3p_2\, d^3p_1'\, d^3p_2'\, \delta^4(P_\mathrm{f} - P_\mathrm{i})|T_\mathrm{fi}|^2(f_2'f_1' - f_2f_1)(1 + \log f_1) \tag{4.8}$$

Interchanging $\mathbf{p}_1$ and $\mathbf{p}_2$ in this integrand leaves the integral invariant because T_fi is invariant under such an interchange. Making this change of variables of integration and taking one-half of the sum of the new expression and (4.8), we obtain

$$\frac{dH}{dt} = \frac{1}{2}\int d^3p_2\, d^3p_1'\, d^3p_2'\, \delta^4(P_\mathrm{f} - P_\mathrm{i})|T_\mathrm{fi}|^2 \times(f_2'f_1' - f_2f_1)[2 + \log(f_1f_2)] \tag{4.9}$$

This integral is invariant under the interchange of $\{\mathbf{p}_1, \mathbf{p}_2\}$ and $\{\mathbf{p}_1', \mathbf{p}_2'\}$ because for every collision there is an inverse collision with the same T matrix. Hence

$$\frac{dH}{dt} = -\frac{1}{2}\int d^3p_2\, d^3p_1'\, d^3p_2'\, \delta^4(P_\mathrm{f} - P_\mathrm{i})|T_\mathrm{fi}|^2(f_2'f_1' - f_2f_1) \times[2 + \log(f_1'f_2')] \tag{4.10}$$

Taking half the sum of (4.9) and (4.10) we obtain

$$\frac{dH}{dt} = \frac{1}{4}\int d^3p_2\, d^3p_1'\, d^3p_2'\, \delta^4(P_\mathrm{f} - P_\mathrm{i})|T_\mathrm{fi}|^2(f_2'f_1' - f_2f_1) \times[\log(f_1f_2) - \log(f_1'f_2')] \tag{4.11}$$

The integrand of the integral in (4.11) is never positive. ■

---

*Note that the use of (4.4) presupposes that the state of the system under consideration satisfies the assumption of molecular chaos.

As a by-product of the proof, we deduce from (4.11) that $dH/dt = 0$ if and only if the integrand of (4.11) identically vanishes. This proves that the statement (4.6) is identical with (4.2). It also shows that under an arbitrary initial condition $f(\mathbf{p}, t) \underset{t\to\infty}{\to} f_0(\mathbf{p})$.

4.2 THE MAXWELL-BOLTZMANN DISTRIBUTION

It has been shown that the equilibrium distribution function $f_0(\mathbf{p})$ is a solution of (4.2). It will be called the Maxwell-Boltzmann distribution. To find it, let us take the logarithm of both sides of (4.2):

$$\log f_0(\mathbf{p}_1) + \log f_0(\mathbf{p}_2) = \log f_0(\mathbf{p}_1') + \log f_0(\mathbf{p}_2') \tag{4.12}$$

Since $\{\mathbf{p}_1, \mathbf{p}_2\}$ and $\{\mathbf{p}_1', \mathbf{p}_2'\}$ are, respectively, the initial and final velocities of *any* possible collision, (4.12) has the form of a conservation law. If $\chi(\mathbf{p})$ is any quantity associated with a molecule of velocity $\mathbf{p}$, such that $\chi(\mathbf{p}_1) + \chi(\mathbf{p}_2)$ is conserved in a collision between molecules $\mathbf{p}_1$ and $\mathbf{p}_2$, a solution of (4.12) is

$$\log f_0(\mathbf{p}) = \chi(\mathbf{p})$$

The most general solution of (4.12) is

$$\log f_0(\mathbf{p}) = \chi_1(\mathbf{p}) + \chi_2(\mathbf{p}) + \cdots$$

where the list $\chi_1, \chi_2, \ldots$ exhausts all independently conserved quantities. For spinless molecules, these are the energy and the momentum of a molecule, and, of course, a constant. Hence $\log f$ is a linear combination of $\mathbf{p}^2$ and the three components of $\mathbf{p}$ plus an arbitrary constant:

$$\log f_0(\mathbf{p}) = -A(\mathbf{p} - \mathbf{p}_0)^2 + \log C$$

or

$$f_0(\mathbf{p}) = Ce^{-A(\mathbf{p}-\mathbf{p}_0)^2} \tag{4.13}$$

where C, A, and the three components of $\mathbf{p}_0$ are five arbitrary constants. We can determine these constants in terms of observed properties of the system.

Applying the condition (3.5), and denoting the particle density N/V by n we have

$$n = C\int d^3p\, e^{-A(\mathbf{p}-\mathbf{p}_0)^2} = C\int d^3p\, e^{-Ap^2} = C\left(\frac{\pi}{A}\right)^{3/2}$$

from which we conclude that $A > 0$ and

$$C = \left(\frac{A}{\pi}\right)^{3/2} n \tag{4.14}$$

Let the average momentum $\langle\mathbf{p}\rangle$ of a gas molecule be defined by

$$\langle\mathbf{p}\rangle \equiv \frac{\int d^3p\, \mathbf{p} f_0(\mathbf{p})}{\int d^3p\, f_0(\mathbf{p})} \tag{4.15}$$

Then

$$\langle \mathbf{p} \rangle = \frac{C}{n} \int d^3p \, \mathbf{p} e^{-A(\mathbf{p}-\mathbf{p}_0)^2} = \frac{C}{n} \int d^3p \, (\mathbf{p} + \mathbf{p}_0) e^{-Ap^2} = \mathbf{p}_0 \tag{4.16}$$

Thus we must take $\mathbf{p}_0 = 0$, if the gas has no translational motion as a whole.

Next we calculate the average energy ϵ of a molecule, defined by

$$\epsilon \equiv \frac{\int d^3p \, (p^2/2m) f_0(\mathbf{p})}{\int d^3p f_0(\mathbf{p})} \tag{4.17}$$

We have, setting $\mathbf{p}_0 = 0$,

$$\epsilon = \frac{C}{2nm} \int d^3p \, \frac{p^2}{2m} e^{-Ap^2} = \frac{2\pi C}{nm} \int_0^\infty dp \, p^4 e^{-Ap^2} = \frac{3}{4Am}$$

The constant A is therefore related to the average energy by

$$A = \frac{3}{4\epsilon m} \tag{4.18}$$

Substituting this into (4.14) we obtain for the constant C the expression

$$C = n \left(\frac{3}{4\pi\epsilon m} \right)^{3/2} \tag{4.19}$$

To relate the average energy ϵ to a directly measurable quantity, let us find the equation of state corresponding to the equilibrium distribution function. We do this by calculating the pressure, which is defined as the average force per unit area exerted by the gas on one face of a perfectly reflecting plane exposed to the gas. Let the disk shown in Fig. 4.1 represent such a unit area, and let us call the axis normal to it the x axis. A molecule can hit this disk only if the x component of its momentum p_x is positive. Then it loses an amount of momentum $2p_x$ upon reflection from this disk. The number of molecules reflected by the disk per second is the number of molecules contained in the cylinder shown in Fig. 4.1 with $v_x > 0$. This number is $v_x f_0(\mathbf{p}) \, d^3p$, with $v_x > 0$. Therefore the pressure is,

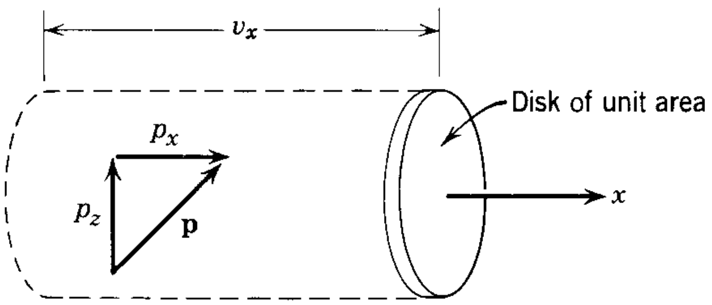


Fig. 4.1 Illustration for the calculation of the pressure.

for a gas with zero average velocity

$$\begin{aligned} P &= \int_{v_x>0} d^3p\, 2p_x v_x f_0(\mathbf{p}) = \int d^3p\, p_x v_x f_0(\mathbf{p}) \\ &= \frac{C}{m}\int d^3p\, p_x^2 e^{-Ap^2} = \frac{C}{3m}\int d^3p\, p^2 e^{-Ap^2} \end{aligned} \tag{4.20}$$

where the last step comes about because $f_0(\mathbf{p})$ depends only on $|\mathbf{p}|$ so that the average values of p_x^2, p_y^2, and p_2^2 are all equal to one-third of the average of $p^2 = p_x^2 + p_y^2 + p_z^2$. Finally we notice that

$$P = \tfrac{2}{3}C\int d^3p\, \frac{p^2}{2m} e^{-Ap^2} = \tfrac{2}{3}n\epsilon \tag{4.21}$$

This is the equation of state. Experimentally we define the temperature T by $P = nkT$, where k is Boltzmann's constant. Hence

$$\epsilon = \tfrac{3}{2}kT \tag{4.22}$$

In terms of the temperature T, the average momentum $\mathbf{p}_0$, and the particle density n the equilibrium distribution function for a dilute gas in the absence of external force is

$$f_0(\mathbf{p}) = \frac{n}{(2\pi mkT)^{3/2}} e^{-(\mathbf{p}-\mathbf{p}_0)^2/2mkT} \tag{4.23}$$

This is the Maxwell-Boltzmann distribution, the probability of finding a molecule with momentum $\mathbf{p}$ in the gas, under equilibrium conditions.*

If a perfectly reflecting wall is introduced into the gas, $f_0(\mathbf{p})$ will remain unchanged because $f_0(\mathbf{p})$ depends only on the magnitude of $\mathbf{p}$, which is unchanged by reflection from the wall.

For a gas with $\mathbf{p}_0 = 0$ it is customary to define the most probable speed $\bar{v}$ of a molecule by the value of v at which $4\pi p^2 f(\mathbf{p})$ attains a maximum. We easily find $\bar{p} = \sqrt{2mkT}$. The most probable speed is therefore

$$\bar{v} = \sqrt{\frac{2kT}{m}} \tag{4.24}$$

The root mean square speed v_{rms} is defined by

$$v_{\text{rms}} \equiv \left[\frac{\int d^3p\, v^2 f_0(\mathbf{p})}{\int d^3p\, f_0(\mathbf{p})}\right]^{1/2} = \sqrt{\frac{3kT}{m}} \tag{4.25}$$

At room temperatures these speeds for an O_2 gas are of the order of magnitude of 10^5 cm/s.

---

*We have assumed, in accordance with experimental facts, that the temperature T is independent of the average momentum $\mathbf{p}_0$.

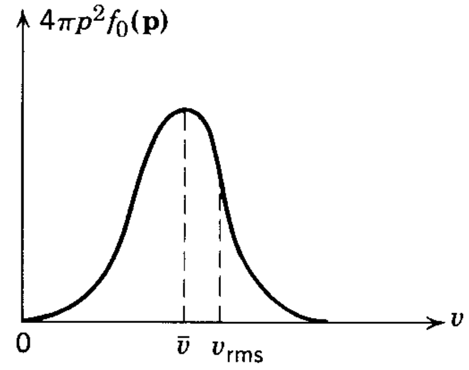


Fig. 4.2 The Maxwell-Boltzmann distribution.

A plot of $4\pi p^2 f_0(\mathbf{p})$ against $v = p/m$ is shown in Fig. 4.2. We notice that $f_0(\mathbf{p})$ does not vanish, as it should, when v exceeds the velocity of light c. This is because we have used Newtonian dynamics for the molecules instead of the more correct relativistic dynamics. The error is negligible at room temperatures, because $\bar{v} \ll c$. The temperature above which relativistic dynamics must be used can be roughly estimated by putting $\bar{v} = c$, from which we obtain $kT \approx mc^2$. Hence $T \approx 10^{13}$ K for H_2.

Let us now consider the equilibrium distribution for a dilute gas in the presence of an external conservative force field given by

$$\mathbf{F} = -\nabla\phi(\mathbf{r}) \tag{4.26}$$

We assert that the equilibrium distribution function is now

$$f(\mathbf{r},\mathbf{p}) = f_0(\mathbf{p})e^{-\phi(\mathbf{r})/kT} \tag{4.27}$$

where $f_0(\mathbf{p})$ is given by (4.23). To prove this we show that (4.27) satisfies Boltzmann's equation. We see immediately that $\partial f/\partial t = 0$ because (4.27) is independent of the time. Furthermore $(\partial f/\partial t)_{\text{coll}} = 0$ because $\phi(\mathbf{r})$ is independent of $\mathbf{p}$:

$$\left(\frac{\partial f}{\partial t}\right)_{\text{coll}} = e^{-2\phi(\mathbf{r})/kT}\int d^3p_2\, d^3p_1'\, d^3p_2'\, \delta^4(P_{\text{f}} - P_{\text{i}})|T_{\text{fi}}|^2(f_2'f_1' - f_2 f_1) = 0$$

Hence it is only necessary to verify that

$$\left(\frac{\mathbf{p}}{m}\cdot\nabla_{\mathbf{r}} + \mathbf{F}\cdot\nabla_{\mathbf{p}}\right)f(\mathbf{r},\mathbf{p}) = 0$$

and this is trivial. We may absorb the factor $\exp(-\phi/kT)$ in (4.27) into the density n and write

$$f(\mathbf{r},\mathbf{p}) = \frac{n(\mathbf{r})}{(2\pi mkT)^{3/2}}e^{-(\mathbf{p}-\mathbf{p}_0)^2/2mkT} \tag{4.28}$$

where

$$n(\mathbf{r}) = \int d^3p\, f(\mathbf{r},\mathbf{p}) = n_0 e^{-\phi(\mathbf{r})/kT} \tag{4.29}$$

Finally we derive the thermodynamics of a dilute gas. We have defined the temperature by (4.22) and we have obtained the equation of state. By the very definition of the pressure, the work done by the gas when its volume increases by dV is $P\,dV$. The internal energy is defined by

$$U(T) = N\epsilon = \tfrac{3}{2}NkT \tag{4.30}$$

which is obviously a function of the temperature alone.

The analog of the first law of thermodynamics now takes the form of a definition for the heat absorbed by the system:

$$dQ = dU + P\,dV \tag{4.31}$$

It tells us that heat added to the system goes into the mechanical work $P\,dV$ and the energy of molecular motion dU. From (4.31) and (4.30) we obtain for the heat capacity at constant volume

$$C_V = \tfrac{3}{2}Nk \tag{4.32}$$

The analog of the second law of thermodynamics is Boltzmann's H theorem, where we identify H with the negative of the entropy per unit volume divided by Boltzmann's constant:

$$H = -\frac{S}{Vk} \tag{4.33}$$

Thus the H theorem states that for a fixed volume (i.e., for an isolated gas) the entropy never decreases, which is a statement of the second law.

To justify (4.33) we calculate H in equilibrium:

$$H_0 = \int d^3p\, f_0 \log f_0 = n\left\{\log\left[n\left(\frac{1}{2\pi mkT}\right)^{3/2}\right] - \frac{3}{2}\right\}$$

Using the equation of state we can rewrite this as

$$-kVH_0 = \tfrac{3}{2}Nk\log(PV^{5/3}) + \text{constant} \tag{4.34}$$

We recognize that the right-hand side is the entropy of an ideal gas in thermodynamics. It follows from (4.34), (4.33), and (4.31) that $dS = dQ/T$.

Thus we have derived all of classical thermodynamics for a dilute gas; and moreover, we were able to calculate the equation of state and the specific heat. The third law of thermodynamics cannot be derived here because we have used classical mechanics and thus are obliged to confine our considerations to high temperatures.

4.3 THE METHOD OF THE MOST PROBABLE DISTRIBUTION

We have noted the interesting fact that the Maxwell-Boltzmann distribution is independent of the detailed form of molecular interactions, as long as they exist. This fact endows the Maxwell-Boltzmann distribution with universality. We might suspect that as long as we are interested only in the *equilibrium* behavior of

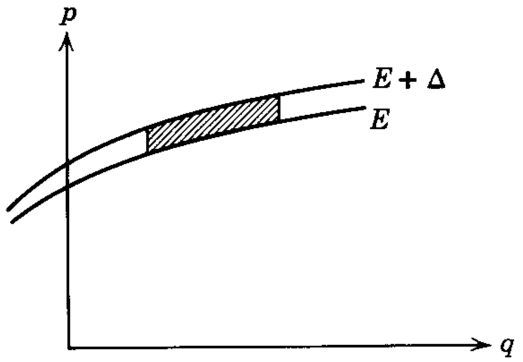


Fig. 4.3 The microcanonical ensemble corresponding to a gas contained in a finite volume with energy between E and $E + \Delta$.

a gas there is a way to derive the Maxwell-Boltzmann distribution without explicitly mentioning molecular interactions. Such a derivation is now supplied. Through it we shall understand better the meaning of the Maxwell-Boltzmann distribution. The conclusion we reach will be the following. If we choose a state of the gas at random from among all its possible states consistent with certain macroscopic conditions, the probability that we shall choose a Maxwell-Boltzmann distribution is overwhelmingly greater than that for any other distribution.

We shall use the approach of the Gibbsian ensemble described in Sec. 3.4. We assume that in equilibrium the system is equally likely to be found in any state consistent with the macroscopic conditions. That is, the density function is a constant over the accessible portion of Γ space.

Specifically we consider a gas of N molecules enclosed in a box of volume V with perfectly reflecting walls. Let the energy of the gas lie between E and $E + \Delta$, with $\Delta \ll E$. The ensemble then consists of a uniform distribution of points in a region of Γ space bounded by the energy surfaces of energies E and $E + \Delta$, and the surfaces corresponding to the boundaries of the containing box, as illustrated schematically in Fig. 4.3. Since the walls are perfectly reflecting, energy is conserved, and a representative point never leaves this region. By Liouville's theorem the distribution of representative points moves likes an incompressible fluid, and hence maintains a constant density at all times. This ensemble is called a *microcanonical ensemble*.

Next consider an arbitrary distribution function of a gas. A molecule in the gas is confined to a finite region of μ space because the values of p and q are restricted by the macroscopic conditions. Cover this finite region of μ space with volume elements of volume $\omega = d^3p\, d^3q$, and number them from 1 to K, where K is a very large number which eventually will be made to approach infinity. We refer to these volume elements as cells. An arbitrary distribution function is defined if we specify the number of molecules n_i found in the ith cell. These are called occupation numbers, and they satisfy the conditions

$$\sum_{i=1}^{K} n_i = N \tag{4.35}$$

$$\sum_{i=1}^{K} \epsilon_i n_i = E \tag{4.36}$$

where ϵ_i is the energy of a molecule in the ith cell:

$$\epsilon_i = \frac{p_i^2}{2m}$$

where $\mathbf{p}_i$ is the momentum of the ith cell. It is in (4.36), and only in (4.36), that the assumption of a dilute gas enters. An arbitrary set of integers $\{n_i\}$ satisfying (4.35) and (4.36) defines an arbitrary distribution function. The value of the distribution function in the ith cell, denoted by f_i, is

$$f_i = \frac{n_i}{\omega} \tag{4.37}$$

This is the distribution function for one member in the ensemble. The equilibrium distribution function is the above averaged over the microcanonical ensemble, which assigns equal weight to all systems satisfying (4.35) and (4.36):

$$\bar{f}_i = \frac{\langle n_i \rangle}{\omega}$$

This is the same definition as (3.52) except that we have replaced the infinitesimal element $d^3r\,d^3p$ by a finite cell of volume ω. Unfortunately this ensemble average is difficult to calculate. So we shall adopt a somewhat different approach, which will yield the same result for a sufficiently large system.

It is clear that if the state of the gas is given, then f is uniquely determined; but if f is given, the state of the gas is not uniquely determined. For example, interchanging the positions of two molecules in the gas leads to a new state of the gas, and hence moves the representative in Γ space; but that does not change the distribution function. Thus a given distribution function f corresponds not to a point, but to a volume in Γ space, which we call the *volume occupied by* f. We shall assume that the equilibrium distribution function is the *most probable distribution function*, i.e., that which occupies the largest volume in Γ space.

The procedure is then as follows:

(*a*) Choose an arbitrary distribution function by choosing an arbitrary set of allowed occupations numbers. Calculate the volumes it occupies by counting the number of systems in the ensemble that have these occupation numbers.

(*b*) Vary the distribution function to maximize the volume.

Let us denote by $\Omega\{n_i\}$ the volume in Γ space occupied by the distribution function corresponding to the occupation numbers $\{n_i\}$. It is proportional to the number of ways of distributing N distinguishable molecules among K cells so that there are n_i of them in the ith cell ($i = 1, 2, \ldots, K$). Therefore

$$\Omega\{n_i\} \propto \frac{N!}{n_1!n_2!n_3! \cdots n_K!} g_1^{n_1} g_2^{n_2} \cdots g_K^{n_K} \tag{4.38}$$

where g_i is a number that we will put equal to unity at the end of the calculation but that is introduced here for mathematical convenience. Taking the logarithm of (4.38) we obtain

$$\log \Omega\{n_i\} = \log N! - \sum_{i=1}^{K} \log n_i! + \sum_{i=1}^{K} n_i \log g_i + \text{constant}$$

Now assume that each n_i is a very large integer, so we can use Stirling's approximation, $\log n_i! \approx n_i \log n_i - 1$. We then have

$$\log \Omega\{n_i\} = N \log N - \sum_{i=1}^{K} n_i \log n_i! + \sum_{i=1}^{K} n_i \log g_i + \text{constant} \qquad (4.39)$$

To find the equilibrium distribution we vary the set of integers $\{n_i\}$ subject to the conditions (4.35) and (4.36) until $\log \Omega$ attains a maximum. Let $\{\bar{n}_i\}$ denote the set of occupation numbers that maximizes $\log \Omega$. By the well-known method Lagrange multipliers we have

$$\delta[\log \Omega\{n_i\}] - \delta\left(\alpha \sum_{i=1}^{K} n_i + \beta \sum_{i=1}^{K} \epsilon_i n_i\right) = 0 \qquad (n_i = \bar{n}_i) \qquad (4.40)$$

where α, β are Lagrange's multipliers. Now the n_i can be considered independent of one another. Substituting (4.39) into (4.40) we obtain

$$\sum_{i=1}^{K} [-(\log n_i + 1) + \log g_i - \alpha - \beta\epsilon_i]\, \delta n_i = 0 \qquad (n_i = \bar{n}_i)$$

Since δn_i are independent variations, we obtain the equilibrium condition by setting the summand equal to zero:

$$\begin{aligned} \log \bar{n}_i &= -1 + \log g_i - \alpha - \beta\epsilon_i \\ \bar{n}_i &= g_i e^{-\alpha - \beta\epsilon_i - 1} \end{aligned} \qquad (4.41)$$

The most probable distribution function is, by (4.37) and (4.41),

$$\bar{f}_i = C e^{-\beta\epsilon_i} \qquad (4.42)$$

where C is a constant. The determination of the constants C and β proceeds in the same way as for (4.13). Writing $\bar{f}_i \equiv f(\mathbf{p}_i)$, we see that $f(\mathbf{p})$ is the Maxwell-Boltzmann distribution (4.23) for $\mathbf{p}_0 = 0$. To show that (4.41) actually corresponds to a maximum of $\log \Omega\{n_i\}$ we calculate the second variation. It is easily shown that the second variation of the quantity on the left side of (4.40), for $n_i = \bar{n}_i$, is

$$-\sum_{i=1}^{K} \frac{1}{n_i}(\delta n_i)^2 < 0$$

We have obtained the Maxwell-Boltzmann distribution as the most probable distribution, in the sense that among all the systems satisfying the macroscopic conditions the Maxwell-Boltzmann distribution is the distribution common to the largest number of them. The question remains: What fraction of these systems have the Maxwell-Boltzmann distributions? In other words, how probable is the most probable distribution? The probability for the occurrence of any set of occupation numbers $\{n_i\}$ is given by

$$P\{n_i\} = \frac{\Omega\{n_i\}}{\sum\limits_{\{n'_j\}} \Omega\{n'_j\}} \tag{4.43}$$

where the sum in the denominator extends over all possible sets of integers $\{n'_j\}$ satisfying (4.35) and (4.36). The probability for finding the system in the Maxwell-Boltzmann distribution is therefore $P\{\bar{n}_i\}$. A direct calculation of $P\{\bar{n}_i\}$ is not easy. We shall be satisfied with an estimate, which, however, becomes an exact evaluation if this probability approaches unity.

The ensemble average of n_i is given by

$$\langle n_i \rangle = \frac{\sum\limits_{\{n_j\}} n_i \Omega\{n_j\}}{\sum\limits_{\{n_j\}} \Omega\{n_j\}} \tag{4.44}$$

It is obvious from (4.38) that

$$\langle n_i \rangle = g_i \frac{\partial}{\partial g_i} \log \left[\sum_{\{n_j\}} \Omega\{n_j\} \right] \tag{4.45}$$

if we let $g_i \to 1$. The deviations from the average value can be estimated by calculating the mean square fluctuation $\langle n_i^2 \rangle - \langle n_i \rangle^2$. We can express $\langle n_i^2 \rangle$ in terms of $\langle n_i \rangle$ as follows:

$$\langle n_i^2 \rangle \equiv \frac{\sum n_i^2 \Omega}{\sum \Omega} = \frac{g_i \dfrac{\partial}{\partial g_i} \left(g_i \dfrac{\partial}{\partial g_i} \sum \Omega \right)}{\sum \Omega} \tag{4.46}$$

where the sum Σ extends over all allowed $\{n_i\}$. Through the series of steps given next we obtain the desired results:

$$\begin{aligned} \langle n_i^2 \rangle &= g_i \frac{\partial}{\partial g_i} \left(\frac{1}{\sum \Omega} g_i \frac{\partial}{\partial g_i} \sum \Omega \right) - g_i \left(\frac{\partial}{\partial g_i} \frac{1}{\sum \Omega} \right) g_i \frac{\partial}{\partial g_i} \sum \Omega \\ &= g_i \frac{\partial}{\partial g_i} \left(g_i \frac{\partial}{\partial g_i} \log \sum \Omega \right) + \left(\frac{1}{\sum \Omega} g_i \frac{\partial}{\partial g_i} \sum \Omega \right)^2 \\ &= g_i \frac{\partial}{\partial g_i} \langle n_i \rangle + \langle n_i \rangle^2 \end{aligned} \tag{4.47}$$

Therefore the mean square fluctuation is

$$\langle n_i^2 \rangle - \langle n_i \rangle^2 = g_i \frac{\partial}{\partial g_i} \langle n_i \rangle \tag{4.48}$$

where we must let $g_i \to 1$ at the end of the calculation.

If the mean square fluctuation is large compared to $\langle n_i \rangle^2$, then $\langle n_i \rangle$ may differ considerably from $\bar{n}_i$; but then neither of them will be physically meaningful. If the mean square fluctuation is small compared to $\langle n_i \rangle^2$, we may expect $\langle n_i \rangle$ and $\bar{n}_i$ to be almost equal. We assume the latter is so, and we shall see that this is a consistent assumption.* Putting

$$\langle n_i \rangle \approx \bar{n}_i$$

we find from (4.41) and (4.48) that

$$\langle n_i^2 \rangle - \langle n_i \rangle^2 \approx \bar{n}_i$$

or

$$\sqrt{\left\langle \left(\frac{n_i}{N} \right)^2 \right\rangle - \left\langle \frac{n_i}{N} \right\rangle^2} \approx \frac{\sqrt{\bar{n}_i/N}}{\sqrt{N}} \tag{4.49}$$

Since $\bar{n}_i/N$ is less than one, the right side of (4.49) becomes vanishingly small if N is the number of molecules in 1 mol of gas, namely $N \approx 10^{23}$. This result implies that the probability $P\{n_i\}$ defined by (4.43) has an extremely sharp peak at $\{n_i\} = \{\bar{n}_i\}$. The width of the peak is such that $P\{n_i\}$ is essentially reduced to zero when any n_i/N differ from $\bar{n}_i/N$ by a number of the order of $1/\sqrt{N}$. A schematic plot of $P\{n_i\}$ is shown in Fig. 4.4. We shall call the distributions lying within the peak "essentially Maxwell-Boltzmann" distributions. They are physically indistinguishable from the strict Maxwell-Boltzmann distribution. From these considerations we conclude that in a physical gas any state picked out at random from among all those satisfying the given macroscopic conditions will almost certainly have a distribution function that is Maxwell-Boltzmann.

The meaning of the Maxwell-Boltzmann distribution is therefore as follows. If a dilute gas is prepared in an arbitrary initial state, and if there exist interactions to enable the gas to go into states other than the initial state, the gas will in time almost certainly become Maxwell-Boltzmann, because among all possible states of the gas satisfying the macroscopic conditions (which are conserved by the interactions), almost all of them have the Maxwell-Boltzmann distribution. This, however, does not tell us how long it will take for the gas to reach the equilibrium situation. Nor does it rule out the possibility that the gas may never reach the equilibrium situation, nor that of leaving the equilibrium situation after attaining it. From this point of view, we see that the laws of thermodynamics are not rigorously true but only overwhelmingly probable.

To illustrate these ideas, consider a gas enclosed in a cubical box with perfectly reflecting walls. Suppose initially the gas molecules are distributed in an arbitrary way within the box, and all have exactly the same velocity parallel to one edge of the box. If there are no interactions, this distribution will be maintained indefinitely, and the system never becomes Maxwell-Boltzmann. For such a gas thermodynamics is invalid. If there is molecular interaction, *no matter how small*, the initial distribution will, through collisions, change with time. Since almost any state of the gas will have a Maxwell-Boltzmann distribution, it is

---

*This assumption can be proved by the method described in Chapter 9. The desired result is essentially stated in (9.29).

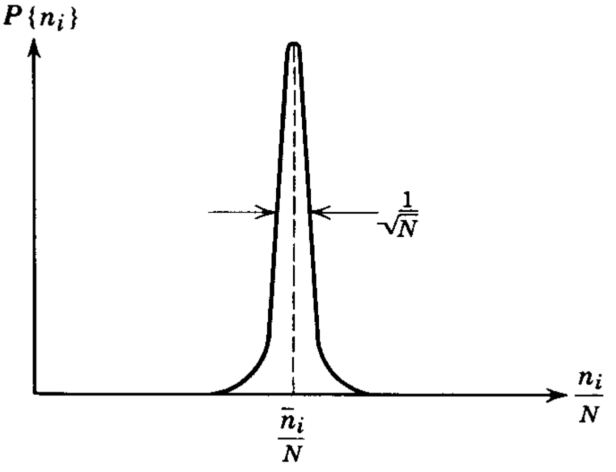


Fig. 4.4 Probability of a gas having the occupation numbers $\{n_i\}$. The most probable occupation numbers $\{\bar{n}_i\}$ correspond to the Maxwell-Boltzmann distribution. Occupation numbers $\{n_i\}$ for which $P\{n_i\}$ lies within the peak are called "essentially Maxwell-Boltzmann" distributions.

reasonable that the distribution after a *sufficiently long time*, depending on the collision cross section, will become Maxwell-Boltzmann. The considerations we have made cannot tell us how long this time must be. They only tell us what the equilibrium situation is, if it is reached.

The derivation of the Maxwell-Boltzmann distribution presented here is independent of the earlier derivation based on the Boltzmann transport equation. Neither of these derivations is rigorous. In the present one there are assumptions that we did not justify, and in the previous one there was the assumption of molecular chaos, which remains unproved and is not related to the assumptions made here. The present method seems to be more satisfactory as a derivation of the Maxwell-Boltzmann distribution because it reveals more clearly the statistical nature of the Maxwell-Boltzmann distribution. The method of the most probable distribution, however, does not furnish information about a gas not in equilibrium, whereas the Boltzmann transport equation does. Hence the main value of the Boltzmann equation lies in its application to nonequilibrium phenomena.

4.4 ANALYSIS OF THE *H* THEOREM

We now discuss the physical implication of Boltzmann's H theorem. For a given distribution function $f(\mathbf{p}, t)$, H is defined by

$$H = \int d^3p\, f(\mathbf{p}, t) \log f(\mathbf{p}, t) \tag{4.50}$$

The time evolution of H is determined by the time evolution of $f(\mathbf{p}, t)$, which does not in general satisfy the Boltzmann transport equation. It satisfies the Boltzmann transport equation only at the instant when the assumption of molecular chaos happens to be valid.

The H theorem states that *if at a given instant t the state of the gas satisfies the assumption of molecular chaos*, then at the instant $t + \epsilon(\epsilon \to 0)$,

(*a*) $\dfrac{dH}{dt} \leq 0$

(*b*) $\dfrac{dH}{dt} = 0$ if and only if $f(\mathbf{v}, t)$ is the Maxwell-Boltzmann distribution.

The proof of the theorem given earlier is rigorous in the limiting case of an infinitely dilute gas. Therefore an inquiry into the validity of the H theorem can only be an inquiry into the validity of the assumption of molecular chaos.

We recall that the assumption of molecular chaos states the following: If $f(\mathbf{p}, t)$ is the probability of finding a molecule with velocity $\mathbf{p}$ at time t, the probability of simultaneously finding a molecule with velocity $\mathbf{p}$ and a molecule with velocity $\mathbf{p}'$ at time t is $f(\mathbf{p}, t)f(\mathbf{p}', t)$. This assumption concerns the correlation between two molecules and has nothing to say about a form of the distribution function. Thus a state of the gas possessing a given distribution function may or may not satisfy the assumption of molecular chaos. For brevity we call a state of the gas a state of "molecular chaos" if it satisfies the assumption of molecular chaos.

We now show that when the gas is in a state of "molecular chaos" H is at a local peak. Consider a dilute gas, in the absence of external force, prepared with an initial condition that is invariant under time reversal.* Under these conditions, the distribution function depends on the magnitude but not the direction of $\mathbf{v}$. Let the gas be in a state of "molecular chaos" and be non-Maxwell-Boltzmann at time $t = 0$. According to the H theorem $dH/dt \leq 0$ at $t = 0^+$. Now consider another gas, which at $t = 0$ is precisely the same as the original one except that all molecular velocities are reversed in direction. This gas must have the same H and must also be in a state of "molecular chaos." Therefore for this new gas $dH/dt \leq 0$ at $t = 0^+$. On the other hand, according to the invariance of the equations of motion under time reversal, the future of the new gas is the past of the original gas. Therefore for the original gas we must have

$$\frac{dH}{dt} \leq 0 \qquad \text{at } t = 0^+$$

$$\frac{dH}{dt} \geq 0 \qquad \text{at } t = 0^-$$

Thus H is at a local peak,[†] as illustrated in Fig. 4.5.

When H is not at a local peak, such as at the points a and c in Fig. 4.5, the state of the gas is not a state of "molecular chaos." Hence molecular collisions, which are responsible for the change of H with time, can create "molecular chaos" when there is none and destroy "molecular chaos" once established.

It is important to note that dH/dt is not necessarily a continuous function of time; it can be changed abruptly by molecular collisions. Overlooking this fact might lead us to conclude, erroneously, that the H theorem is inconsistent with the invariance under time reversal. A statement of the H theorem that is manifestly invariant under time reversal is the following. If there is "molecular chaos" now, then $dH/dt \leq 0$ in the next instant. If there will be "molecular chaos" in the next instant, then $dH/dt \geq 0$ now.

*These simplifying features are introduced to avoid the irrelevant complications arising from the time reversal properties of the external force and the agent preparing the system.

[†] The foregoing argument is due to F. E. Low (unpublished).

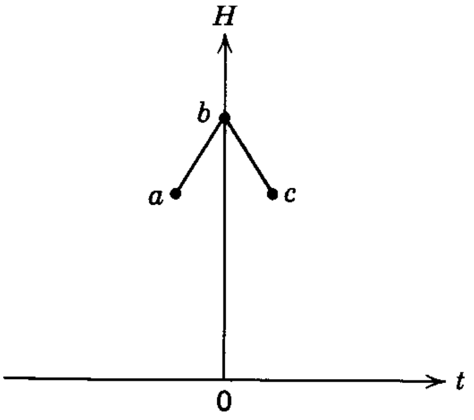


Fig. 4.5 H is at a local peak when the gas is in a state of "molecular chaos."

We now discuss the general behavior of H as a function of time. Our discussion rests on the following premises.

(*a*) H is at its smallest possible value when the distribution function is strictly Maxwell-Boltzmann. This easily follows from (4.50), and it is independent of the assumption of molecular chaos.*

(*b*) Molecular collisions happen at random, i.e., the time sequence of the states of a gas is a sequence of states chosen at random from those that satisfy the macroscopic conditions. This assumption is plausible but unproved.

From these premises it follows that the distribution function of the gas is almost always essentially Maxwell-Boltzmann, i.e., a distribution function contained within the peak shown in Fig. 4.4. The curve of H as a function of time consists mostly of microscopic fluctuations above the minimum value. Between two points at which H is at the minimum value there is likely to be a small peak.

If at any instant the gas has a distribution function appreciably different from the Maxwell-Boltzmann distribution, then H is appreciably larger than the minimum value. Since collisions are assumed to happen at random, it is overwhelmingly probable that after the next collision the distribution will become essentially Maxwell-Boltzmann and H will decrease to essentially the minimum value. By time reversal invariance it is overwhelmingly probable that before the last collision H was at essentially the minimum value. Thus H is overwhelmingly likely to be at a sharp peak when the gas is in an improbable state. The more improbable the state, the sharper the peak.

A very crude model of the curve of H as a function of time is shown in Fig. 4.6. The duration of a fluctuation, large or small, should be of the order of the time between two successive collisions of a molecule, i.e., 10^{-11} sec for a gas under ordinary conditions. The large fluctuations, such as that labeled a in Fig. 4.6, almost never occur spontaneously.† We can, of course, prepare a gas in an improbable state, e.g., by suddenly removing a wall of the container of the gas, so that H is initially at a peak. But it is overwhelmingly probable that within a few collision times the distribution would be reduced to an essentially Maxwell-Boltzmann distribution.

---

*See Problem 4.9.

†See Problems 4.5 and 4.6.

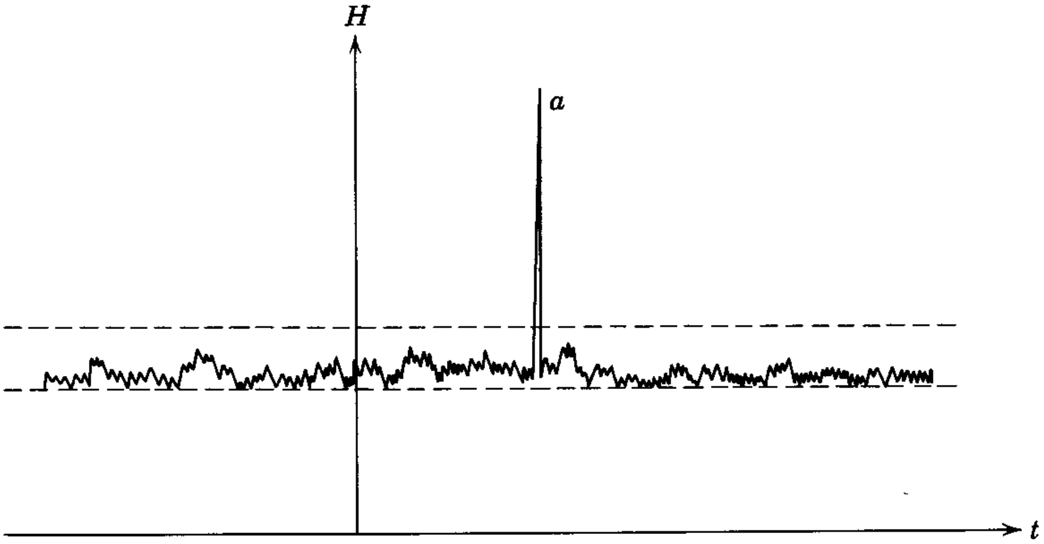


Fig. 4.6 H as a function of time. The range of values of H lying between the two horizontal dashed lines is called the "noise range."

Most of the time the value of H fluctuates within a small range above the minimum value. This range, shown enclosed by dashed lines in Fig. 4.6, corresponds to states of the gas with distribution functions that are essentially Maxwell-Boltzmann, i.e., distribution functions contained within the peak of Fig. 4.4. We call this range the "noise range." These features of the curve of H have been deduced only through plausibility arguments, but they are in accord with experience. We can summarize them as follows.

(a) For all practical purposes H never fluctuates spontaneously above the noise range. This corresponds to the observed fact that a system in thermodynamic equilibrium never spontaneously goes out of equilibrium.

(b) If at an instant H has a value above the noise range, then, for all practical purposes, H always decreases after that instant. In a few collision times its value will be within the noise range. This corresponds to the observed fact that if a system is initially not in equilibrium (the initial state being brought about by external agents), it always tends to equilibrium. In a few collision times it will be in equilibrium. This feature, together with (a), constitutes the second law of thermodynamics.

(c) Most of the time the value of H fluctuates in the noise range, in which dH/dt is as frequently positive as negative. (This is not a contradiction to the H theorem, because the H theorem merely requires that when the system is in a state of "molecular chaos," then $dH/dt \leq 0$ in the next instant.) These small fluctuations produce no observable change in the equation of state and other thermodynamic quantities. When H is in the noise range, the system is, for all practical purposes, in thermodynamic equilibrium. These fluctuations, however, do lead to observable effects,

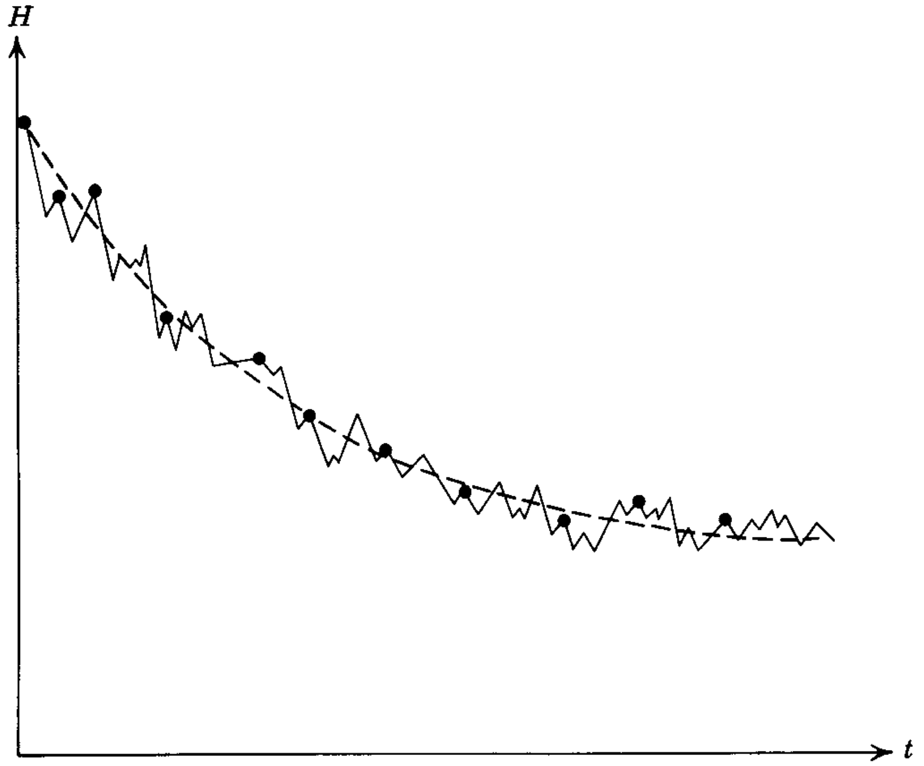


Fig. 4.7 The solid curve is H as a function of time for a gas initially in an improbable state. The dots are the points at which there is "molecular chaos." The dashed curve is that predicted by the Boltzmann transport equation.

e.g., the fluctuation scattering of light. We witness it in the blue of the sky.

We have argued that whenever the distribution function is not strictly Maxwell-Boltzmann, H is likely to be at a local peak. On the other hand, it was shown earlier that in a state of "molecular chaos" H *is* at a local peak. Thus we may regard a state of "molecular chaos" as a convenient mathematical model for a state that does not have a strictly Maxwell-Boltzmann distribution function. Hence the Boltzmann transport equation may be regarded as valid in a statistical sense. To illustrate this, let us imagine that a gas is prepared in an improbable initial state. The curve of H as a function of time might look like the solid curve in Fig. 4.7. Let us mark with a dot a point on this curve at which the gas is in a state of "molecular chaos." All these dots must be at a local peak of H (but not all local peaks are marked with a dot). By assumption of the randomness of the time sequence of states, they are likely to be evenly distributed in time. The distribution of dots might look like that illustrated in Fig. 4.7.

A solution to the Boltzmann transport equation would yield a smooth curve of negative slope that tries to fit these dots, as shown by the dashed curve in Fig. 4.7. It is in this sense that the Boltzmann transport equation provides a description of the approach to equilibrium.

These arguments make it only plausible that the Boltzmann transport equation is useful for the description of the approach to equilibrium. The final test lies in the comparison of results with experiments.

4.5 THE POINCARÉ CYCLE

When Boltzmann announced the H theorem a century ago, objections were raised against it on the ground that it led to "paradoxes." These are the so-called "reversal paradox" and "recurrence paradox," both based on the erroneous statement of the H theorem that $dH/dt \leq 0$ at all times. The correct statement of the H theorem, as given in the last section, is free from such objections. We mention these "paradoxes" purely for historical interest.

The "reversal paradox" is as follows: *The H theorem singles out a preferred direction of time. It is therefore inconsistent with time reversal invariance.* This is not a paradox, because the statement of the alleged paradox is false. We have seen in the last section that time reversal invariance is consistent with the H theorem, because dH/dt need not be a continuous function of time. In fact, we have made use of time reversal invariance to deduce interesting properties of the curve of H.

The "recurrence paradox" is based on the following true theorem.

POINCARÉ'S THEOREM

A system having a finite energy and confined to a finite volume will, after a sufficiently long time, return to an arbitrarily small neighborhood of almost any given initial state.

By "almost any state" is meant any state of the system, except for a set of measure zero (i.e., a set that has no volume, e.g., a discrete point set). A neighborhood of a state has an obvious definition in terms of the Γ space of the system.

A proof of Poincaré's theorem is given at the end of this section. This theorem implies that H is an almost periodic function of time. The "recurrence paradox" arises in an obvious way, if we take the statement of the H theorem to be $dH/dt \leq 0$ at all times. Since this is not the statement of the H theorem, there is no paradox. In fact, Poincaré's theorem furnishes further information concerning the curve of H.

Most of the time H lies in the noise range. Poincaré's theorem implies that the small fluctuations in the noise range repeat themselves. This is only to be expected.

For the rare spontaneous fluctuations above the noise range, Poincaré's theorem requires that if one such fluctuation occurs another one must occur after a sufficiently long time. The time interval between two large fluctuations is called a *Poincaré cycle*. A crude estimate (see Problem 4.7) shows that a Poincaré cycle is of the order of e^N, where N is the total number of molecules in the system. Since $N \approx 10^{23}$, a Poincaré cycle is extremely long. In fact, it is essentially the same number, be it $10^{10^{23}}$ s or $10^{10^{23}}$ ages of the universe, (the age of the universe being a mere 10^{10} years.) Thus it has nothing to do with physics.

We mentioned the ergodic theorem in Section 3.4, but did not use it as a basis for the microcanonical ensemble, even though, on the surface, it seems to be the justification we need. The reason is that existing proofs of the theorem all share a characteristic of the proof the Poincaré theorem given below, i.e., an avoidance of dynamics. For this reason, they cannot provide the true relaxation time for a system to reach local equilibrium, (typically about 10^{-15} s for real systems,) but have a characteristic time scale of the order of the Poincaré cycle. For this reason, the ergodic theorem has so far been an interesting mathematical exercise irrelevant to physics.

Proof of Poincaré's Theorem Let a state of the system be represented by a point in Γ space. As time goes on, any point in Γ space traces out a locus that is uniquely determined by any given point on the locus. Let g_0 be an arbitrary volume element in Γ space of volume ω_0. After time t all the points in g_0 will be in another volume element g_t, of volume ω_t, which is uniquely determined by g_0. By Liouville's theorem, $\omega_t = \omega_0$.

Let Γ_0 denote the subspace that is the union of all g_t for $0 \leq t < \infty$. Let its volume be Ω_0. Similarly, let Γ_τ denote the subspace that is the union of all g_t for $\tau \leq t < \infty$. Let its volume be Ω_τ. The numbers Ω_0 and Ω_τ are finite because, since the energy of the system and the spatial extension of the system are finite, a representative point is confined to a finite region of Γ space. The definitions immediately imply that Γ_0 contains Γ_τ.

We may think of Γ_0 and Γ_τ in a different way. Imagine the region Γ_0 to be filled uniformly with representative points. As time goes on, Γ_0 will evolve into some other regions that are uniquely determined. It is clear, from the definitions, that after a time τ, Γ_0 will become Γ_τ. Hence, by Liouville's theorem,

$$\Omega_0 = \Omega_\tau$$

We recall that Γ_0 contains all the future destinations of the points in g_0, and Γ_τ contains all the future destinations of the points in g_τ, which in turn is evolved from g_0 after the time τ. It has been shown that Γ_0 has the same volume as Γ_τ. Therefore Γ_0 and Γ_τ must contain the same set of points except for a set of measure zero.

In particular, Γ_τ contains all of g_0 except for a set of measure zero. But, by definition, all points in Γ_τ are future destinations of the points in g_0. Therefore all points in g_0, except for a set of measure zero, must return to g_0 after a sufficiently long time. Since g_0 can be made as small as we wish, Poincaré's theorem follows. ■

PROBLEMS

4.1 Describe an experimental method for the verification of the Maxwell-Boltzmann distribution.

4.2 A cylindrical column of gas of given temperature rotates about a fixed axis with constant angular velocity. Find the equilibrium distribution function.

4.3 (*a*) What fraction of the H_2 gas at sea level and at a temperature of 300 K can escape from the earth's gravitational field?

(*b*) Why do we still have H_2 gas in the atmosphere at sea level?

4.4 Using relativistic dynamics for gas molecules find, for a dilute gas of zero total momentum,

(*a*) the equilibrium distribution function;

(*b*) the equation of state.

Answer. PV is independent of the volume. Hence it is NkT by definition of T.

4.5 (*a*) *Estimate* the probability that a stamp (mass = 0.1 g) resting on a desk top at room temperature (300 K) will spontaneously fly up to a height of 10^{-8} cm above the desk top.

Hint. Think not of one stamp but of an infinite number of noninteracting stamps placed side by side. Formulate an argument showing that these stamps obey the Maxwell-Boltzmann distribution.

Answer. Let m = mass of stamps, h = height, g = acceleration of gravity. Probability $\approx e^{-mgh/kT}$

4.6 A room of volume $3 \times 3 \times 3\ \text{m}^3$ is under standard conditions (atmospheric pressure and 300 K).

(*a*) *Estimate* the probability that at any instant of time a 1-cm^3 volume anywhere within this room becomes totally devoid of air because of spontaneous statistical fluctuations.

(*b*) estimate the same for a 1-Å^3 volume.

Answer. Let N = total number of air molecules, V = volume of room, v = the volume devoid of air. Probability $\approx e^{-N(v/V)}$

4.7 Suppose the situation referred to in Problem 4.6*a* has occurred. Describe qualitatively the behavior of the distribution function thereafter. Estimate the time it takes for such a situation to occur again, under the assumption that molecular collisions are such that the time sequence of the state of the system is a random sequence of states.

4.8 (*a*) Explain why in (4.42) we arrived at the formula for the Maxwell-Boltzmann distribution for a gas with no average momentum ($\mathbf{p}_0 = 0$), although average momentum was not specified as a macroscopic condition in (4.35) and (4.36).

(*b*) Derive the Maxwell-Boltzmann distribution for a gas with average velocity $\mathbf{v}_0$, using the method of the most probable distribution.

4.9 Let

$$H = \int d^3p\, f(\mathbf{p}, t) \log f(\mathbf{p}, t)$$

where $f(\mathbf{v}, t)$ is arbitrary except for the conditions

$$\int d^3p\, f(\mathbf{p}, t) = n$$

$$\frac{1}{n}\int d^3p\, \frac{p^2}{2m} f(\mathbf{p}, t) = \epsilon$$

Show that H is minimum when f is the Maxwell-Boltzmann distribution.

CHAPTER

5

TRANSPORT PHENOMENA

5.1 THE MEAN FREE PATH

To begin our discussion on the approach to equilibrium of a gas initially not in equilibrium, we introduce the qualitative concept of the mean free path and related quantities.

A gas is not in equilibrium when the distribution function is different from the Maxwell-Boltzmann distribution. The most common case of a nonequilibrium situation is that in which the temperature, density, and average velocity are not constant throughout the gas. To approach equilibrium, these nonuniformities have to be ironed out through the transport of energy, mass, and momentum from one part of the gas to another. The mechanism of transport is molecular collision, and the average distance over which molecular properties can be transported in one collision is the *mean free path*. It is the average distance traveled by a molecule between successive collisions. We give an estimate of its order of magnitude.

The number of collisions happening per second per unit volume at the point $\mathbf{r}$ in a gas is given by

$$Z = \int d^3p_1\, d^3p_2\, d^3p_1'\, d^3p_2'\, \delta^4(P_f - P_i)|T_{fi}|^2 f(\mathbf{r},\mathbf{p}_1,t) f(\mathbf{r},\mathbf{p}_2,t)$$

where $f(\mathbf{r},\mathbf{p},t)$ is the distribution function. The integration over $\mathbf{p}_1'$ and $\mathbf{p}_2'$ can be immediately effected to yield

$$Z = \int d^3p_1 \int d^3p_2\, \sigma_{\text{tot}}\, |\mathbf{v}_1 - \mathbf{v}_2| f(\mathbf{r},\mathbf{p}_1,t) f(\mathbf{r},\mathbf{p}_2,t) \tag{5.1}$$

A free path is defined as the distance traveled by a molecule between two successive collisions. Since it takes two molecules to make a collision, every collision terminates two free paths. The total number of free paths occurring per second per unit volume is therefore $2Z$. Since there are n molecules per unit volume, the average number of free paths traveled by a molecule per second is

$2Z/n$. The *mean free path*, which is the average length of a free path, is given by

$$\lambda = \frac{n}{2Z}\bar{v} \tag{5.2}$$

where $\bar{v} = \sqrt{2kT/m}$ is the most probable speed of a molecule. The average duration of a free path is called the *collision time* and is given by

$$\tau = \frac{\lambda}{\bar{v}} \tag{5.3}$$

For a gas in equilibrium, $f(\mathbf{r}, \mathbf{p}, t)$ is the Maxwell-Boltzmann distribution. Assume for an order-of-magnitude estimate that σ_{tot} is insensitive to the energy of the colliding molecules and may be replaced by a constant of the order of πa^2 where a is the molecular diameter. Then we have

$$\begin{aligned} Z &= \frac{\sigma_{\text{tot}}}{m}\int d^3p_1 \int d^3p_2 |\mathbf{p}_1 - \mathbf{p}_2| f(\mathbf{p}_1) f(\mathbf{p}_2) \\ &= \frac{\sigma_{\text{tot}}\, n^2}{m(2\pi mkT)^3}\int d^3p_1 \int d^3p_2 |\mathbf{p}_1 - \mathbf{p}_2| \exp\left[-\frac{p_1^2 + p_2^2}{2mkT}\right] \\ &= \frac{\sigma_{\text{tot}} n^2}{m(2\pi mkT)^3}\int d^3P \int d^3p |\mathbf{p}| \exp\left[-\frac{1}{kT}\left(\frac{P^2}{4m} + \frac{p^2}{m}\right)\right] \end{aligned}$$

where $\mathbf{P} = \mathbf{p}_1 + \mathbf{p}_2$, $\mathbf{p} = \frac{1}{2}(\mathbf{p}_2 - \mathbf{p}_1)$. The integrations are elementary and give

$$Z = 2n^2\sigma_{\text{tot}}\sqrt{\frac{kT}{\pi m}} = \sqrt{\frac{2}{\pi}}\, n^2 \sigma_{\text{tot}}\, \bar{v} \tag{5.4}$$

Therefore

$$\begin{aligned} \lambda &= \frac{\sqrt{\pi/8}}{n\sigma_{\text{tot}}} \\ \tau &= \frac{\sqrt{\pi/8}}{n\sigma_{\text{tot}}\,\bar{v}} \end{aligned} \tag{5.5}$$

We see that the mean free path is independent of the temperature and is inversely proportional to the density times the total cross section.

The following are some numerical estimates. For H_2 gas at its critical point,

$$\lambda \approx 10^{-7}\ \text{cm}$$

$$\tau \approx 10^{-11}\ \text{s}$$

For H_2 gas in interstellar space, where the density is about 1 molecule/cm^3,

$$\lambda \approx 10^{15}\ \text{cm}$$

The diameter of H_2 has been taken to be about 1 Å.

From these qualitative estimates, it is expected that in H_2 gas under normal conditions, for example, any nonuniformity in density or temperature over distances of order 10^{-7} cm will be ironed out in the order of 10^{-11} s. Variations in density or temperature over macroscopic distances may persist for a long time.

5.2 EFFUSION

An important quantity governing the behavior of a gas is the ratio of the mean free path to some other characteristic length, such as

The size of the box containing the gas.
The diameter of a hole through which gas molecules may pass.
The wavelength of density fluctuations.

When the mean free path is large compared to any other length in the problem, the gas is said to be in the *collisionless regime*. A practical example is the process of *effusion*, whereby a gas leaks through a very small hole of diameter much smaller than the mean free path—a phenomenon of great interest to all experimentalists who maintain vacuum systems.

In effusion the gas molecules do not collide as they go through the hole. Therefore the flux I through the hole, defined as the number of molecules crossing the hole per second per unit area of the hole, is just the flux of molecules impinging on the surface area of the hole. The contribution to the flux from molecules of velocity $\mathbf{v}$ is given by

$$dI = d^3p\, v_x f(\mathbf{p})$$

where the x axis is chosen normal to the hole. The total flux is therefore

$$I = \int_{v_x > 0} d^3p\, v_x f(\mathbf{p})$$

Assuming the Maxwell-Boltzmann distribution, we have

$$\begin{aligned} I &= \frac{nm^3}{(2\pi mkT)^{3/2}} \int_0^\infty dv_x\, v_x e^{-mv_x^2/2kT} \int_{-\infty}^\infty dv_y\, e^{-mv_y^2/2kT} \int_{-\infty}^\infty dv_z\, e^{-mv_z^2/2kT} \\ &= n\sqrt{\frac{kT}{2\pi m}} = \frac{n\bar{v}}{2\sqrt{\pi}} \end{aligned}$$

Eliminating n through $P = nkT$, we obtain

$$I = \frac{P}{\sqrt{2\pi mkT}} \tag{5.6}$$

The inverse proportionality to $\sqrt{m}$ makes the process useful as a means of separating isotopes.

The opposite of the collisionless regime is one in which the mean free path is much smaller than the other characteristic lengths of the problem, exemplified by the flow of a gas through a very large hole. In this case the gas molecules will undergo many collisions as they pass through the hole, and will "thermalize" locally. The prevailing condition is known as the *hydrodynamic regime*, and will be the subject of the rest of this chapter.

5.3 THE CONSERVATION LAWS

To investigate nonequilibrium phenomena, we must solve the Boltzmann transport equation, with given initial conditions, to obtain the distribution function as a function of time. Some rigorous properties of any solution to the Boltzmann equation may be obtained from the fact that in any molecular collision there are dynamical quantities that are rigorously conserved.

Let $\chi(\mathbf{r}, \mathbf{p})$ be any quantity associated with a molecule of velocity $\mathbf{p}$ located at $\mathbf{r}$, such that in any collision $\{\mathbf{p}_1, \mathbf{p}_2\} \to \{\mathbf{p}_1', \mathbf{p}_2'\}$ taking place at $\mathbf{r}$, we have

$$\chi_1 + \chi_2 = \chi_1' + \chi_2' \tag{5.7}$$

where $\chi_1 = \chi(\mathbf{r}_1, \mathbf{p}_1)$, etc. We call χ a conserved property. The following theorem holds.

THEOREM

$$\int d^3p\, \chi(\mathbf{r}, \mathbf{p}) \left[\frac{\partial f(\mathbf{r}, \mathbf{p}, t)}{\partial t} \right]_{\text{coll}} = 0 \tag{5.8}$$

where $(\partial f/\partial t)_{\text{coll}}$ is the right side of (3.36).*

Proof By definition of $(\partial f/\partial t)_{\text{coll}}$ we have

$$\int d^3p\, \chi \left(\frac{\partial f}{\partial t} \right)_{\text{coll}} = \int d^3p_1\, d^3p_2\, d^3p_1'\, d^3p_2'\, \delta^4(P_{\text{f}} - P_{\text{i}}) |T_{\text{fi}}|^2 (f_2' f_1' - f_2 f_1) \chi_1 \tag{5.9}$$

Making use of the properties of T_{fi} discussed in Section 3.2, and proceeding in a manner similar to the proof of the H theorem, we make each of the following interchanges of integration variables.

First: $\mathbf{p}_1 \rightleftarrows \mathbf{p}_2$
Next: $\mathbf{p}_1 \rightleftarrows \mathbf{p}_1'$ and $\mathbf{p}_2 \rightleftarrows \mathbf{p}_2'$
Next: $\mathbf{p}_1 \rightleftarrows \mathbf{p}_2'$ and $\mathbf{p}_2 \rightleftarrows \mathbf{p}_1'$

For each case we obtain a different form for the same integral. Adding the three

---

*Note that it is not required that f be a solution of the Boltzmann transport equation.

new formulas so obtained to (5.9) and dividing the result by 4 we get

$$\int d^3p\,\chi\left(\frac{\partial f}{\partial t}\right)_{\text{coll}} = \frac{1}{4}\int d^3p_1\,d^3p_2\,d^3p_1'\,d^3p_2'\,\delta^4(P_f - P_i)|T_{fi}|^2$$
$$\times(f_2'f_1' - f_2f_1)(\chi_1 + \chi_2 - \chi_1' - \chi_2') \equiv 0 \qquad \blacksquare$$

The conservation theorem relevant to the Boltzmann transport equation is obtained by multiplying the Boltzmann transport equation on both sides by χ and then integrating over **p**. The collision term vanishes by virtue of (5.8), and we have*

$$\int d^3p\,\chi(\mathbf{r},\mathbf{p})\left(\frac{\partial}{\partial t} + \frac{p_i}{m}\frac{\partial}{\partial x_i} + F_i\frac{\partial}{\partial v_i}\right)f(\mathbf{r},\mathbf{p},t) = 0 \tag{5.10}$$

We may rewrite (5.10) in the form

$$\frac{\partial}{\partial t}\int d^3p\,\chi f + \frac{\partial}{\partial x_i}\int d^3p\,\chi\frac{p_i}{m}f - \int d^3p\,\frac{\partial\chi}{\partial x_i}\frac{p_i}{m}f + \int d^3p\,\frac{\partial}{\partial p_i}(\chi F_i f)$$
$$-\int d^3p\,\frac{\partial\chi}{\partial p_i}F_i f - \int d^3p\,\chi\frac{\partial F_i}{\partial p_i}f = 0 \tag{5.11}$$

The fourth term vanishes if $f(\mathbf{r},\mathbf{p},t)$ is assumed to vanish when $|\mathbf{p}| \to \infty$. This conservation theorem is most useful in hydrodynamics, where the velocity $\mathbf{v} = \mathbf{p}/m$ rather than the momentum **p** is a directly measurable quantity. Accordingly, we shall reexpress **p** in terms of **v**, where convenient. We also define the average value $\langle A\rangle$ by

$$\langle A\rangle \equiv \frac{\int d^3p\,Af}{\int d^3p\,f} = \frac{1}{n}\int d^3p\,Af \tag{5.12}$$

where

$$n(\mathbf{r},t) \equiv \int d^3p\,f(\mathbf{r},\mathbf{p},t) \tag{5.13}$$

We obtain finally the desired theorem:

CONSERVATION THEOREM

$$\frac{\partial}{\partial t}\langle n\chi\rangle + \frac{\partial}{\partial x_i}\langle nv_i\chi\rangle - n\left\langle v_i\frac{\partial\chi}{\partial x_i}\right\rangle - \frac{n}{m}\left\langle F_i\frac{\partial\chi}{\partial v_i}\right\rangle - \frac{n}{m}\left\langle\frac{\partial F_i}{\partial v_i}\chi\right\rangle = 0 \tag{5.14}$$

where χ is any conserved property. Note that $\langle nA\rangle = n\langle A\rangle$ because n is

---

*The summation convention, whereby a repeated vector index is understood to be summed from 1 to 3, is used.

independent of **v**. From now on we restrict our attention to velocity-independent external forces so that the last term of (5.14) may be dropped.

For simple molecules the independent conserved properties are mass, momentum, and energy. For charged molecules we also include the charge, but this extension is trivial. Accordingly we set successively

$$\chi = m \qquad \text{(mass)}$$

$$\chi = mv_i \quad (i = 1, 2, 3) \qquad \text{(momentum)}$$

$$\chi = \tfrac{1}{2}m|\mathbf{v} - \mathbf{u}(\mathbf{r}, t)|^2 \qquad \text{(thermal energy)}$$

where

$$\mathbf{u}(\mathbf{r}, t) \equiv \langle \mathbf{v} \rangle$$

We should then have three independent conservation theorems.

For $\chi = m$ we have immediately

$$\frac{\partial}{\partial t}(mn) + \frac{\partial}{\partial x_i}\langle mnv_i \rangle = 0$$

or, introducing the mass density

$$\rho(\mathbf{r}, t) \equiv mn(\mathbf{r}, t)$$

we obtain

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{u}) = 0 \tag{5.15}$$

Next we put $\chi = mv_i$, obtaining

$$\frac{\partial}{\partial t}\langle \rho v_i \rangle + \frac{\partial}{\partial x_j}\langle \rho v_i v_j \rangle - \frac{1}{m}\rho F_i = 0 \tag{5.16}$$

To reduce this further let us write

$$\begin{aligned}\langle v_i v_j \rangle &= \big\langle (v_i - u_i)(v_j - u_j) \big\rangle + \langle v_i \rangle u_j + u_i \langle v_j \rangle - u_i u_j \\ &= \big\langle (v_i - u_i)(v_j - u_j) \big\rangle + u_i u_j\end{aligned}$$

Substituting this into (5.16) we obtain

$$\rho\left(\frac{\partial u_i}{\partial t} + u_j \frac{\partial u_i}{\partial x_j}\right) = \frac{1}{m}\rho F_i - \frac{\partial}{\partial x_j}\big\langle \rho(v_i - u_i)(v_j - u_j) \big\rangle \tag{5.17}$$

Introducing the abbreviation

$$P_{ij} \equiv \rho\big\langle (v_i - u_i)(v_j - u_j) \big\rangle$$

which is called the *pressure tensor*, we finally have

$$\left(\frac{\partial}{\partial t} + u_j \frac{\partial}{\partial x_j}\right) u_i = \frac{1}{m} F_i - \frac{1}{\rho}\frac{\partial}{\partial x_j} P_{ij} \tag{5.18}$$

Finally we set $\chi = \frac{1}{2}m|\mathbf{v} - \mathbf{u}|^2$. Then

$$\frac{1}{2}\frac{\partial}{\partial t}\langle \rho|\mathbf{v} - \mathbf{u}|^2\rangle + \frac{1}{2}\frac{\partial}{\partial x_i}\langle \rho v_i|\mathbf{v} - \mathbf{u}|^2\rangle - \frac{1}{2}\rho\left\langle v_i\frac{\partial}{\partial x_i}|\mathbf{v} - \mathbf{u}|^2\right\rangle = 0 \quad (5.19)$$

We define the *temperature* by

$$kT \equiv \theta \equiv \tfrac{1}{3}m\langle |\mathbf{v} - \mathbf{u}|^2\rangle$$

and the *heat flux* by

$$\mathbf{q} = \tfrac{1}{2}\rho\big\langle(\mathbf{v} - \mathbf{u})|\mathbf{v} - \mathbf{u}|^2\big\rangle$$

We then have

$$\begin{aligned}\tfrac{1}{2}\rho\langle v_i|\mathbf{v} - \mathbf{u}|^2\rangle &= \tfrac{1}{2}m\rho\big\langle(v_i - u_i)|\mathbf{v} - \mathbf{u}|^2\big\rangle + \tfrac{1}{2}\rho u_i\langle|\mathbf{v} - \mathbf{u}|^2\rangle \\ &= q_i + \tfrac{3}{2}\rho\theta u_i\end{aligned}$$

and

$$\rho\big\langle v_i(v_j - u_j)\big\rangle = \rho\big\langle(v_i - u_i)(v_j - u_j)\big\rangle + \rho u_i\langle v_j - u_j\rangle = P_{ij}$$

Thus (5.19) can be written

$$\frac{3}{2}\frac{\partial}{\partial t}(\rho\theta) + \frac{\partial q_i}{\partial x_i} + \frac{3}{2}\frac{\partial}{\partial x_i}(\rho\theta u_i) + mP_{ij}\frac{\partial u_j}{\partial x_i} = 0$$

Since $P_{ij} = P_{ji}$

$$mP_{ij}\frac{\partial u_j}{\partial x_i} = P_{ij}\frac{m}{2}\left(\frac{\partial u_j}{\partial x_i} + \frac{\partial u_i}{\partial x_j}\right) \equiv P_{ij}\Lambda_{ij}$$

The final form is then obtained after a few straightforward steps:

$$\rho\left(\frac{\partial}{\partial t} + u_i\frac{\partial}{\partial x_i}\right)\theta + \frac{2}{3}\frac{\partial}{\partial x_i}q_i = -\frac{2}{3}\Lambda_{ij}P_{ij} \quad (5.20)$$

The three conservation theorems are summarized in (5.21), (5.22), and (5.23).

$$\frac{\partial \rho}{\partial t} + \nabla\cdot(\rho\mathbf{u}) = 0 \qquad \text{(conservation of mass)} \quad (5.21)$$

$$\rho\left(\frac{\partial}{\partial t} + \mathbf{u}\cdot\nabla\right)\mathbf{u} = \frac{\rho}{m}\mathbf{F} - \nabla\cdot\overleftrightarrow{P} \qquad \text{(conservation of momentum)} \quad (5.22)$$

$$\rho\left(\frac{\partial}{\partial t} + \mathbf{u}\cdot\nabla\right)\theta = -\tfrac{2}{3}\nabla\cdot\mathbf{q} - \tfrac{2}{3}\overleftrightarrow{P}\cdot\overleftrightarrow{\Lambda} \qquad \text{(conservation of energy)} \quad (5.23)$$

where $\overleftrightarrow{P}$ is a dyadic whose components are P_{ij}, $\nabla\cdot\overleftrightarrow{P}$ is a vector whose ith component is $\partial P_{ij}/\partial x_j$, and $\overleftrightarrow{P}\cdot\overleftrightarrow{\Lambda}$ is a scalar $P_{ij}\Lambda_{ij}$. The auxiliary quantities are defined as follows.

$$\rho(\mathbf{r}, t) \equiv m\int d^3v f(\mathbf{r}, \mathbf{v}, t) \qquad \text{(mass density)} \tag{5.24}$$

$$\mathbf{u}(\mathbf{r}, t) \equiv \langle \mathbf{v} \rangle \qquad \text{(average velocity)} \tag{5.25}$$

$$\theta(\mathbf{r}, t) \equiv \tfrac{1}{3}m\langle |\mathbf{v} - \mathbf{u}|^2 \rangle \qquad \text{(temperature)} \tag{5.26}$$

$$\mathbf{q}(\mathbf{r}, t) \equiv \tfrac{1}{2}m\rho\big\langle (\mathbf{v} - \mathbf{u})|\mathbf{v} - \mathbf{u}|^2 \big\rangle \qquad \text{(heat flux vector)} \tag{5.27}$$

$$P_{ij} \equiv \rho\big\langle (v_i - u_i)(v_j - u_j) \big\rangle \qquad \text{(pressure tensor)} \tag{5.28}$$

$$\Lambda_{ij} \equiv \tfrac{1}{2}m\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right) \tag{5.29}$$

Although the conservation theorems are exact, they have no practical value unless we can actually solve the Boltzmann transport equation and use the distribution function so obtained to evaluate the quantities (5.24)–(5.29). Despite the fact that these quantities have been given rather suggestive names, their physical meaning, if any, can only be ascertained after the distribution function is known. We shall see that when it is known these conservation theorems become the physically meaningful equations of hydrodynamics.

5.4 THE ZERO-ORDER APPROXIMATION

From now on we shall work in the hydrodynamic regime, where the mean free path is small compared to other characteristic lengths. This means that gas molecules make a large number of collisions within a small space. Consequently they come to local equilibrium rapidly. In the lowest-order approximation it is natural to assume that the gas has a local Maxwell-Boltzmann distribution, with slowly varying temperature, density, and average velocity:

$$f(\mathbf{r}, \mathbf{p}, t) \approx f^{(0)}(\mathbf{r}, \mathbf{p}, t) \tag{5.30}$$

where

$$f^{(0)}(\mathbf{r}, \mathbf{p}, t) = \frac{n}{(2\pi m\theta)^{3/2}} \exp\left[-\frac{m}{2\theta}(\mathbf{v} - \mathbf{u})^2 \right] \tag{5.31}$$

where $n, \theta, \mathbf{u}$ are all slowly varying functions of $\mathbf{r}$ and t. It is obvious that (5.30) cannot be an exact solution of the Boltzmann transport equation. On the one hand we have

$$\left(\frac{\partial f^{(0)}}{\partial t} \right)_{\text{coll}} = 0 \tag{5.32}$$

because $n, \theta, \mathbf{u}$ do not depend on $\mathbf{v}$. On the other hand it is clear that in general

$$\left(\frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla_{\mathbf{r}} + \frac{\mathbf{F}}{m} \cdot \nabla_{\mathbf{v}} \right) f^{(0)}(\mathbf{r}, \mathbf{p}, t) \neq 0 \tag{5.33}$$

We postpone the discussion of the accuracy of the approximation (5.30). For the moment let us assume that it is a good approximation and discuss the physical consequences.

If (5.30) is a good approximation, the left side of (5.33) must be approximately equal to zero. This in turn would mean that $n, \boldsymbol{\theta}, \mathbf{u}$ are such that the conservation theorems (5.21)–(5.23) are approximately satisfied. The conservation theorems then become the equations restricting the behaviour of $n, \boldsymbol{\theta}, \mathbf{u}$. To see what they are, we must calculate $\mathbf{q}$ and P_{ij} to the lowest order. The results are denoted respectively by $\mathbf{q}^{(0)}$ and $P_{ij}^{(0)}$. Let $C(\mathbf{r}, t) = n(m/2\pi\theta)^{3/2}$ and $A(\mathbf{r}, t) = m/2\boldsymbol{\theta}$. We easily obtain

$$\begin{aligned} \mathbf{q}^{(0)} &= \frac{1}{2}\frac{\rho}{n}\int d^3v\,(\mathbf{v}-\mathbf{u})|\mathbf{v}-\mathbf{u}|^2 C(\mathbf{r}, t)\,e^{-A(\mathbf{r}, t)|\mathbf{v}-\mathbf{u}|^2} \\ &= \tfrac{1}{2}m^2 C(\mathbf{r}, t)\int d^3U\,\mathbf{U}U^2 e^{-A(\mathbf{r}, t)U^2} = 0 \end{aligned} \tag{5.34}$$

$$\begin{aligned} P_{ij}^{(0)} &= \frac{\rho}{n}C(\mathbf{r}, t)\int d^3v\,(v_i - u_i)(v_j - u_j)\,e^{-A(\mathbf{r}, t)|\mathbf{v}-\mathbf{u}|^2} \\ &= mC(\mathbf{r}, t)\int d^3U\,U_iU_j e^{-A(\mathbf{r}, t)U^2} = \delta_{ij}P \end{aligned} \tag{5.35}$$

where

$$P = \tfrac{1}{3}\rho\left(\frac{m}{2\pi\theta}\right)^{3/2}\int d^3\mathbf{U}\,U^2 e^{-A(\mathbf{r}, t)U^2} = n\theta \tag{5.36}$$

which is the local hydrostatic pressure.

Substituting these into (5.21) and (5.23), and noting that

$$\nabla\cdot\overleftrightarrow{P}^{(0)} = \nabla P$$

$$\overleftrightarrow{P}^{(0)}\cdot\overleftrightarrow{\Lambda} = P\sum_{i=1}^{3}\Lambda_{ii} = mP\nabla\cdot\mathbf{u}$$

We obtain the equations

$$\frac{\partial\rho}{\partial t} + \nabla\cdot(\rho\mathbf{u}) = 0 \qquad \text{(continuity equation)} \tag{5.37}$$

$$\left(\frac{\partial}{\partial t} + \mathbf{u}\cdot\nabla\right)\mathbf{u} + \frac{1}{\rho}\nabla P = \frac{\mathbf{F}}{m} \qquad \text{(Euler's equation)} \tag{5.38}$$

$$\left(\frac{\partial}{\partial t} + \mathbf{u}\cdot\nabla\right)\theta + \frac{1}{c_V}(\nabla\cdot\mathbf{u})\theta = 0 \tag{5.39}$$

where $c_V = \frac{3}{2}$. These are the hydrodynamic equations for the nonviscous flow of a gas. They possess solutions describing flow patterns that persist indefinitely. Thus, in this approximation, the local Maxwell-Boltzmann distribution never decays to the true Maxwell-Boltzmann distribution. This is in rough accord with experience, for we know that a hydrodynamic flow, left to itself, takes a long time to die out.

Although derived for dilute gases, (5.37)–(5.39) are also used for liquids because these equations can also be derived through heuristic arguments which indicate that they are of a more general validity.

We shall now briefly point out some of the consequences of (5.37)–(5.39) that are of practical interest.

The quantity $(\partial/\partial t + \mathbf{u}\cdot\nabla)\,X$ is known as the "material derivative of X," because it is the time rate of change of X to an observer moving with the local average velocity $\mathbf{u}$. Such an observer is said to be moving along a streamline. We now show that in the zero-order approximation a dilute gas undergoes only adiabatic transformations to an observer moving along a streamline. Equations (5.37) and (5.39) may be rewritten as

$$\left(\frac{\partial}{\partial t} + \mathbf{u}\cdot\nabla\right)\rho = -\rho\nabla\cdot\mathbf{u}$$

$$-\frac{3}{2}\frac{\rho}{\theta}\left(\frac{\theta}{\partial t} + \mathbf{u}\cdot\nabla\right)\theta = \rho\nabla\cdot\mathbf{u}$$

Adding these two equations we obtain

$$\left(\frac{\partial}{\partial t} + \mathbf{u}\cdot\nabla\right)\rho - \frac{3}{2}\frac{\rho}{\theta}\left(\frac{\partial}{\partial t} + \mathbf{u}\cdot\nabla\right)\theta = 0$$

or

$$\left(\frac{\partial}{\partial t} + \mathbf{u}\cdot\nabla\right)\left(\rho\theta^{-3/2}\right) = 0 \tag{5.40}$$

Using the equation of state $P = \rho\theta/m$ we can convert (5.40) to the condition

$$P\rho^{-5/3} = \text{constant} \quad \text{(along a streamline)} \tag{5.41}$$

This is the condition for adiabatic transformation for an ideal gas, since $c_P/c_V = \frac{5}{3}$.

Next we derive the linear equation for a sound wave. Let us restrict ourselves to the case in which $\mathbf{u}$ and all the space and time derivatives of $\mathbf{u}$, ρ, and θ are small quantities of the first order. For $\mathbf{F} = 0$, (5.37) and (5.38) may be replaced by

$$\frac{\partial\rho}{\partial t} + \rho\nabla\cdot\mathbf{u} = 0 \tag{5.42}$$

$$\rho\frac{\partial\mathbf{u}}{\partial t} + \nabla P = 0 \tag{5.43}$$

$$\tfrac{3}{2}\rho\frac{\partial\theta}{\partial t} - \theta\frac{\partial\rho}{\partial t} = 0 \tag{5.44}$$

where quantities smaller than first-order ones are neglected. Note that (5.44) is none other than (5.40) or (5.41). Taking the divergence of (5.43) and the time derivative of (5.42), and subtracting one resulting equation from the other, we obtain

$$\nabla^2 P - \frac{\partial^2 \rho}{\partial t^2} = 0 \tag{5.45}$$

in which higher-order quantities are again neglected. Now P is a function of ρ and θ, but the latter are not independent quantities, being related to each other through the condition of adiabatic transformation (5.44). Hence we may regard P as a function of ρ alone, and write

$$\nabla^2 P = \nabla \cdot \left[\left(\frac{\partial P}{\partial \rho} \right)_S \nabla \rho \right] \approx \left(\frac{\partial P}{\partial \rho} \right)_S \nabla^2 \rho$$

where $(\partial P / \partial \rho)_S$ is the adiabatic derivative, related to the adiabatic compressibility κ_S by

$$\kappa_S = \frac{1}{\rho} \left(\frac{\partial \rho}{\partial P} \right)_S = \frac{3}{5} \frac{m}{\rho \theta} \tag{5.46}$$

Thus (5.45) can be written in the form

$$\nabla^2 \rho - \rho \kappa_S \frac{\partial^2 \rho}{\partial t^2} = 0 \tag{5.47}$$

which is a wave equation for ρ, describing a sound wave with a velocity of propagation c given by

$$c = \frac{1}{\sqrt{\rho \kappa_S}} = \sqrt{\frac{5}{3} \frac{\theta}{m}} = \sqrt{\frac{5}{6}}\, \bar{v} \tag{5.48}$$

It is hardly surprising that the adiabatic compressibility enters here, because in the present approximation there can be no heat conduction in the gas, as (5.34) indicates.

Finally consider the case of steady flow under the influence of a conservative external force field, i.e., under the conditions

$$\begin{aligned} \mathbf{F} &= -\nabla \phi \\ \frac{\partial \mathbf{u}}{\partial t} &= 0 \end{aligned} \tag{5.49}$$

Using the vector identity

$$(\mathbf{u} \cdot \nabla) \mathbf{u} = \tfrac{1}{2} \nabla (u^2) - \mathbf{u} \times (\nabla \times \mathbf{u}) \tag{5.50}$$

we can rewrite (5.38) as follows

$$\nabla \left(\tfrac{1}{2} u^2 + \frac{1}{\rho} P + \frac{1}{m} \phi \right) = \mathbf{u} \times (\nabla \times \mathbf{u}) - \frac{\theta}{m} \frac{\nabla \rho}{\rho} \tag{5.51}$$

Two further specializations are of interest. First, in the case of uniform density and irrotational flow, namely, $\nabla \rho = 0$ and $\nabla \times \mathbf{u} = 0$, we have

$$\nabla\left(\tfrac{1}{2}u^2 + \frac{1}{\rho}P + \frac{1}{m}\phi\right) = 0 \tag{5.52}$$

which is *Bernoulli's equation*. Second, in the case of uniform temperature and irrotational flow, namely, $\nabla \theta = 0$ and $\nabla \times \mathbf{u} = 0$, we have

$$\nabla\left(\tfrac{1}{2}u^2 + \frac{1}{m}\phi\right) = -\frac{\theta}{m}\nabla(\log \rho)$$

which may be immediately integrated to yield

$$\rho = \rho_0 \exp\left[-\frac{1}{\theta}\left(\tfrac{1}{2}mu^2 + \phi\right)\right] \tag{5.53}$$

where ρ_0 is an arbitrary constant.

5.5 THE FIRST-ORDER APPROXIMATION

We now give an estimate of the error incurred in the zero-order approximation (5.30). Let $f(\mathbf{r}, \mathbf{p}, t)$ be the exact distribution function, and let

$$g(\mathbf{r}, \mathbf{p}, t) \equiv f(\mathbf{r}, \mathbf{p}, t) - f^{(0)}(\mathbf{r}, \mathbf{p}, t) \tag{5.54}$$

We are interested in the magnitude of g as compared to $f^{(0)}$. First let us estimate the order of magnitude of $(\partial f/\partial t)_{\text{coll}}$. We have, by definition,

$$\begin{aligned}\left(\frac{\partial f}{\partial t}\right)_{\text{coll}} &= \int d^3p_2\, d^3p_1'\, d^3p_2'\, \delta^4(P_{\text{f}} - P_{\text{i}})|T_{\text{fi}}|^2(f_2' f_1' - f_2 f_1)\\ &\approx \int d^3p_2\, d^3p_1'\, d^3p_2'\, \delta^4(P_{\text{f}} - P_{\text{i}})|T_{\text{fi}}|^2\\ &\quad\times\left(f_2^{(0)\prime} g_1' - f_2^{(0)} g_1 + g_2' f_1^{(0)\prime} - g_2 f_1^{(0)}\right)\end{aligned} \tag{5.55}$$

where we have used (5.54), the fact that $(\partial f^{(0)}/\partial t)_{\text{coll}} = 0$, and the assumption that g is a small quantity whose square can be neglected. An order-of-magnitude estimate of (5.55) may be obtained by calculating the second term of the right side of (5.55), which is

$$-g(\mathbf{r}, \mathbf{p}_1, t)\int d^3p_2\, \sigma_{\text{tot}}|\mathbf{v}_2 - \mathbf{v}_1| f_2^{(0)} = -\frac{g(\mathbf{r}, \mathbf{p}_1, t)}{\tau} \tag{5.56}$$

where τ is a number of the order of magnitude of the collision time. Thus if we put

$$\left(\frac{\partial f}{\partial t}\right)_{\text{coll}} \approx -\frac{f - f^{(0)}}{\tau} \tag{5.57}$$

we obtain results that are qualitatively correct.* With (5.57) the Boltzmann transport equation becomes

$$\left(\frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla_r = \frac{\mathbf{F}}{m} \cdot \nabla_v\right)\left(f^{(0)} + g\right) \approx -\frac{g}{\tau} \tag{5.58}$$

Assuming $g \ll f^{(0)}$, we can neglect g on the left side of (5.58). Assume further that $f^{(0)}$ varies by a significant amount (i.e., of the order of itself) only when $|\mathbf{r}|$ varies by a distance L. Then (5.58) furnishes the estimate

$$\bar{v}\frac{f^{(0)}}{L} \approx -\frac{g}{\tau}$$

or

$$\frac{g}{f^{(0)}} \approx -\frac{\lambda}{L} \tag{5.59}$$

where λ is a length of the order of the mean free path. From these considerations we conclude that $f^{(0)}$ is a good approximation if the local density, temperature, and velocity have characteristic wavelengths L much larger than the mean free path λ. The corrections to $f^{(0)}$ would be of the order of λ/L.

A systematic expansion of f in powers of λ/L is furnished by the Chapman-Enskog expansion, which is somewhat complicated. In order not to lose sight of the physical aspects of the problem, we give a qualitative discussion of the first-order approximation based on the approximate equation (5.58). The precise value of τ cannot be ascertained. For the present we have to be content with the knowledge that τ is of the order of the collision time. Thus we put

$$f = f^{(0)} + g \tag{5.60}$$

where, with (5.58), we take

$$g = -\tau\left(\frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla_r + \frac{\mathbf{F}}{m} \cdot \nabla_v\right)f^{(0)} \tag{5.61}$$

To calculate g, note that $f^{(0)}$ depends on $\mathbf{r}$ and t only through the functions ρ, θ, and $\mathbf{u}$. Thus we need the derivatives

$$\begin{aligned}
\frac{\partial f^{(0)}}{\partial \rho} &= \frac{f^{(0)}}{\rho} \\
\frac{\partial f^{(0)}}{\partial \theta} &= \frac{1}{\theta}\left(\frac{m}{2\theta}U^2 - \frac{3}{2}\right)f^{(0)} \\
\frac{\partial f^{(0)}}{\partial u_i} &= \frac{m}{\theta}U_i f^{(0)} \\
\frac{\partial f^{(0)}}{\partial v_i} &= -\frac{m}{\theta}U_i f^{(0)}
\end{aligned} \tag{5.62}$$

---

*Techniques useful for solving the Boltzmann transport equation, together with results for a few simple intermolecular potentials, may be found in S. Chapman and T. G. Cowling, *The Mathematical Theory of Non-Uniform Gases*, 2nd ed. (Cambridge University Press, Cambridge, 1952).

where

$$\mathbf{U} \equiv \mathbf{v} - \mathbf{u}(\mathbf{r}, t) \tag{5.63}$$

Hence

$$\begin{aligned} g &= -\tau\left(\frac{\partial}{\partial t} + v_i\frac{\partial}{\partial x_i} + \frac{F_i}{m}\frac{\partial}{\partial v_i}\right)f^{(0)} \\ &= -\tau f^{(0)}\left[\frac{1}{\rho}D(\rho) + \frac{1}{\theta}\left(\frac{m}{2\theta}U^2 - \frac{3}{2}\right)D(\theta) + \frac{m}{\theta}U_jD(u_j) - \frac{1}{\theta}\mathbf{F}\cdot\mathbf{U}\right] \end{aligned} \tag{5.64}$$

where

$$D(X) \equiv \left(\frac{\partial}{\partial t} + v_i\frac{\partial}{\partial x_i}\right)X \tag{5.65}$$

Using the zero-order hydrodynamic equations (5.37)–(5.39), we can show that

$$\begin{aligned} D(\rho) &= -\rho(\nabla\cdot\mathbf{u}) + \mathbf{U}\cdot\nabla\rho \\ D(\theta) &= -\tfrac{2}{3}\theta\nabla\cdot\mathbf{u} + \mathbf{U}\cdot\nabla\theta \\ D(u_j) &= -\frac{1}{\rho}\frac{\partial P}{\partial x_j} + \frac{F_j}{m} + U_i\frac{\partial u_j}{\partial x_i} \end{aligned} \tag{5.66}$$

where $P = \rho\theta/m$. Substituting these into (5.64) we obtain

$$\begin{aligned} g = -\tau f^{(0)}\Bigg[&-(\nabla\cdot\mathbf{u}) + \mathbf{U}\cdot\frac{\nabla\rho}{\rho} + \frac{1}{\theta}\left(\frac{m}{2\theta}U^2 - \frac{3}{2}\right)\left(-\frac{2}{3}\theta\nabla\cdot\mathbf{u} + \mathbf{U}\cdot\nabla\theta\right) \\ &+ \frac{m}{\theta}\left(-\mathbf{U}\cdot\frac{\nabla P}{\rho} + \mathbf{U}\cdot\frac{\mathbf{F}}{m} + U_iU_j\frac{\partial u_j}{\partial x_i}\right) - \frac{1}{\theta}\mathbf{F}\cdot\mathbf{U}\Bigg] \end{aligned}$$

which, after some rearrangement and cancellation of terms, becomes

$$g = -\tau\left[\frac{1}{\theta}\frac{\partial\theta}{\partial x_i}U_i\left(\frac{m}{2\theta}U^2 - \frac{5}{2}\right) + \frac{1}{\theta}\Lambda_{ij}\left(U_iU_j - \tfrac{1}{3}\delta_{ij}U^2\right)\right]f^{(0)} \tag{5.67}$$

where Λ_{ij} is defined by (5.29).

It is now necessary to calculate $\mathbf{q}$ and P_{ij} with the help of (5.60) to obtain the equations of hydrodynamics to the first order. We have

$$\mathbf{q} = \frac{m\rho}{2n}\int d^3p\,(\mathbf{v} - \mathbf{u})|\mathbf{v} - \mathbf{u}|^2g$$

Noting that the second term of (5.67) does not contribute to this integral, we obtain

$$\mathbf{q} = -\frac{\tau m^5}{2}\int d^3U\,\mathbf{U}U^2\left(\frac{m}{2\theta}U^2 - \frac{5}{2}\right)\frac{1}{\theta}U_i\frac{\partial\theta}{\partial x_i}f^{(0)}$$

or

$$\mathbf{q} = -K\,\nabla\theta \tag{5.68}$$

where

$$K = \frac{m^5\tau}{6\theta}\int d^3U\,U^4\left(\frac{m}{2\theta}U^2 - \frac{5}{2}\right)f^{(0)} = \tfrac{5}{2}\tau\theta n \tag{5.69}$$

It is clear from (5.68) that K is to be identified as the coefficient of thermal conductivity. It is also clear that $|\mathbf{q}|$ is a small quantity of the first order, being of the order of λ/L.

For the pressure tensor P_{ij}, only the second term of (5.67) contributes:

$$P_{ij} = \frac{\rho}{n}\int d^3p\,(v_i - u_i)(v_j - u_j)\left(f^{(0)} + g\right) = \delta_{ij}P + P'_{ij} \tag{5.70}$$

where $P = \rho\theta/m$ and

$$P'_{ij} = -\frac{\tau\rho m^3}{\theta n}\Lambda_{kl}\int d^3U\,U_iU_j\left(U_kU_l - \tfrac{1}{3}\delta_{kl}U^2\right)f^{(0)} \tag{5.71}$$

To evaluate this, note that P'_{ij} is a symmetric tensor of zero trace (i.e., $\sum_{i=1}^{3} P'_{ii} = 0$), and it depends linearly on the symmetric tensor Λ_{ij}. Therefore P'_{ij} must have the form

$$P'_{ij} = -\frac{2\mu}{m}\left(\Lambda_{ij} - \frac{m}{3}\delta_{ij}\nabla\cdot\mathbf{u}\right) \tag{5.72}$$

where $m\nabla\cdot\mathbf{u}$ is none other than the trace of Λ_{ij}:

$$\sum_{i=1}^{3}\Lambda_{ii} = m\sum_{i=1}^{3}\frac{\partial u_i}{\partial x_i} = m\nabla\cdot\mathbf{u} \tag{5.73}$$

and μ is a constant. It remains to calculate μ. For this purpose it suffices to calculate any component of P'_{ij}, e.g., P'_{12}. From (5.71) we have

$$\begin{aligned} P'_{12} &= -\frac{\tau m^4}{\theta}\Lambda_{kl}\int d^3U\,U_1U_2\left(U_kU_l - \tfrac{1}{3}\delta_{kl}U^2\right)f^{(0)} \\ &= -2\frac{\tau m^4}{\theta}\Lambda_{12}\int d^3U\,U_1^2U_2^2f^{(0)} \end{aligned}$$

Therefore

$$\mu = \frac{\tau m^5}{\theta}\int d^3U\,U_1^2U_2^2f^{(0)} = \tau n\theta \tag{5.74}$$

With this we have

$$P_{ij} = \delta_{ij}P - \frac{2\mu}{m}\left(\Lambda_{ij} - \frac{m}{3}\delta_{ij}\nabla\cdot\mathbf{u}\right) \tag{5.75}$$

The second term is of the order of λ/L. The coefficient μ turns out to be the coefficient of viscosity, as we show shortly.

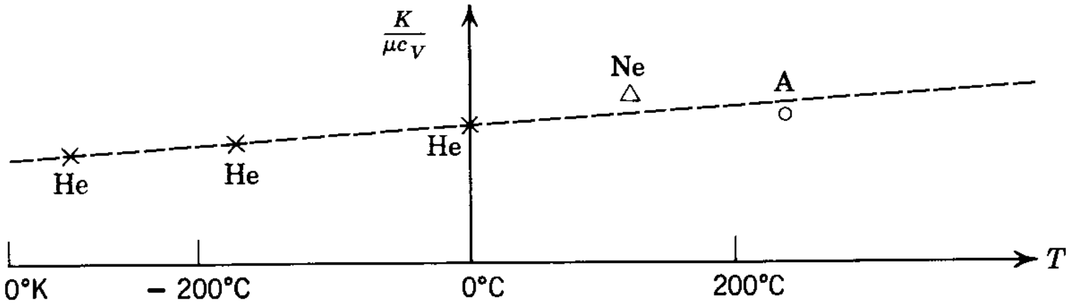


Fig. 5.1 Ratio of thermal conductivity to the product of viscosity and specific heat for different dilute gases.

A comparison of (5.74) with (5.69) shows that

$$\frac{K}{\mu} = \tfrac{5}{2} = \tfrac{5}{3}c_V \tag{5.76}$$

Since the unknown collision time τ drops out in this relation, we might expect (5.76) to be of quantitative significance. A plot of some experimental data for different dilute gases in Fig. 5.1 shows that it is indeed so.

Let us put, with (5.6),

$$\tau \approx \sqrt{\frac{m}{kT}}\,\frac{1}{na^2} \tag{5.77}$$

where a is the molecular diameter. Then we find that

$$\mu \approx K \approx \frac{\sqrt{mkT}}{a^2} \tag{5.78}$$

5.6 VISCOSITY

To show that (5.74) is the coefficient of viscosity, we independently calculate the coefficient of viscosity using its experimental definition. Consider a gas of uniform and constant density and temperature, with an average velocity given by

$$\begin{aligned} u_x &= A + By \\ u_y &= u_z = 0 \end{aligned} \tag{5.79}$$

where A and B are constants. The gas may be thought of as being composed of different layers sliding over each other, as shown in Fig. 5.2. Draw any plane perpendicular to the y axis, as shown by the dotted line in Fig. 5.2. Let F' be the frictional force experienced by the gas above this plane, per unit area of the plane. Then the coefficient of viscosity μ is experimentally defined by the relation

$$F' = -\mu\frac{\partial u_x}{\partial y} \tag{5.80}$$

The gas above the plane experiences a frictional force by virtue of the fact that it

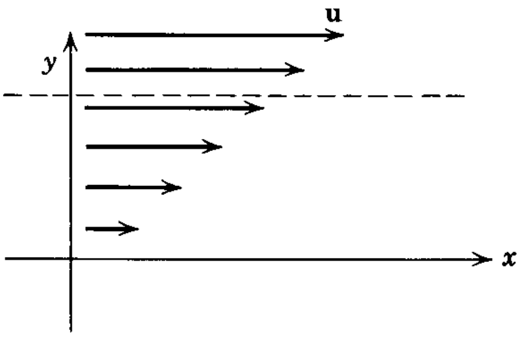


Fig. 5.2 Horizontal flow of a gas with average velocity increasing linearly with height.

suffers a net loss of "x component of momentum" to the gas below. Thus

$$F' \equiv \textit{net} \text{ amount of "}x\text{ component of momentum" transported per sec across unit area in the } y \text{ direction} \tag{5.81}$$

The quantity being transported is $m(v_x - u_x)$, whereas the flux effective in the transport is $n(v_y - u_y)$. Hence we have

$$F' = mn\langle (v_x - u_x)(v_y - u_y)\rangle = m^4 \int d^3v\, (v_x - u_x)(v_y - u_y)(f^{(0)} + g) \tag{5.82}$$

We easily see that the term $f^{(0)}$ does not contribute to the integral in (5.82). The first correction g may be obtained directly from the approximate Boltzmann transport equation

$$\begin{gathered} \mathbf{v} \cdot \nabla f^{(0)} = -\frac{g}{\tau} \\ g = -\frac{\tau m}{\theta} v_y (v_x - u_x) B f^{(0)} = -\frac{\tau m}{\theta} U_y U_x \frac{\partial u_x}{\partial y} f^{(0)} \end{gathered} \tag{5.83}$$

where $\mathbf{U} \equiv \mathbf{v} - \mathbf{u}$. Thus

$$F' = -\frac{\partial u_x}{\partial y} \frac{\tau m^5}{\theta} \int d^3U\, U_x^2 U_y^2 f^{(0)} \tag{5.84}$$

A comparison between this and (5.80) yields

$$\mu = \frac{\tau m^5}{\theta} \int d^3U\, U_x^2 U_y^2 f^{(0)} \tag{5.85}$$

which is identical with (5.74).

From the nature of this derivation it is possible to understand physically why μ has the order of magnitude given by (5.78). Across the imaginary plane mentioned previously, a net transport of momentum exists, because molecules constantly cross this plane in both directions. The flux is the same in both directions, being of the order of $n\sqrt{kT/m}$. On the average, however, those that cross from above to below carry more "x component of momentum" than the opposite ones, because the average velocity u_x is greater above than below. Since most molecules that cross the plane from above originated within a mean free path λ above the plane, their u_x is in excess of the local u_x below the plane by the amount $\lambda(\partial u_x/\partial y)$. Hence the net amount of "x component of momentum" transported per second from above to below, per unit area of the plane, is

$$\lambda n m \sqrt{\frac{kT}{m}} \frac{\partial u_x}{\partial y} = \frac{\sqrt{mkT}}{a^2} \frac{\partial u_x}{\partial y} \tag{5.86}$$

Therefore

$$\mu \approx \frac{\sqrt{mkT}}{a^2} \tag{5.87}$$

It is interesting to note that according to (5.87) μ is independent of the density for a given temperature. When Maxwell first derived this fact, he was so surprised that he put it to experimental test by observing the rate of damping of a pendulum suspended in gases of different densities. To his satisfaction, it was verified.

According to (5.87) the coefficient of viscosity increases as the molecular diameter decreases, everything else being constant. This is physically easy to understand because the mean free path λ increases with decreasing molecular diameter. For a given gradient $\partial u_x/\partial y$, the momentum transported across any plane normal to the y axis obviously increases as λ increases. When λ becomes so large that it is comparable to the size of the container of the gas, the whole method adopted here breaks down, and the coefficient of viscosity ceases to be a meaningful concept.

As a topic related to the concept of viscosity we consider the boundary condition for a gas flowing past a wall. A gas, unlike a liquid, does not stick to the wall of its container. Rather, it slips by with an average velocity u_0. To determine u_0, it is necessary to know how the gas molecules interact with the wall. We make the simplifying assumption that a fraction $1 - \alpha$ of the molecules striking the wall is reflected elastically while the remaining fraction α is absorbed by the wall, only to return to the gas later with thermal velocity. The number α is called the *coefficient of accommodation*. Suppose the wall is the xy plane, as shown in Fig. 5.3. Then the downward flux of particles is given by

$$m^3 \int_{-\infty}^{\infty} dv_x \int_{-\infty}^{\infty} dv_y \int_0^{\infty} dv_z \, n v_z f^{(0)} = n \sqrt{\frac{\theta}{2\pi m}} \tag{5.88}$$

The particles that reach the wall came from a mean free path λ above the wall.

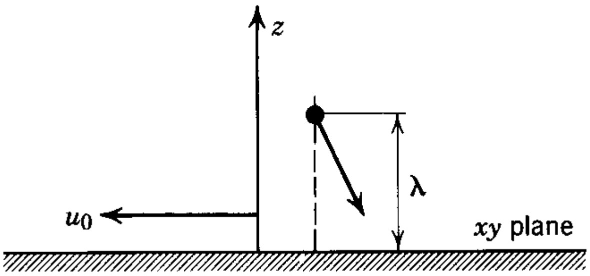


Fig. 5.3 A gas slipping past a wall.

Thus the gas loses to a unit area of the wall an amount of momentum per second equal to

$$F' = -\alpha nm\sqrt{\frac{\theta}{2\pi m}}\left[u_0 + \lambda\left(\frac{\partial u}{\partial z}\right)_0\right] \tag{5.89}$$

where $(\partial u/\partial z)_0$ is the normal gradient of u at the wall. This is the force of friction per unit area that the wall exerts on the gas, and must equal $-\mu(\partial u/\partial z)_0$. Hence the boundary condition at the wall is

$$\alpha nm\sqrt{\frac{\theta}{2\pi m}}\left[u_0 + \lambda\left(\frac{\partial u}{\partial z}\right)_0\right] = \mu\left(\frac{\partial u}{\partial z}\right)_0$$

or

$$u_0 = \left(\sqrt{\frac{2\pi}{m\theta}}\,\frac{\mu}{n\alpha} - \lambda\right)\left(\frac{\partial u}{\partial z}\right)_0 \tag{5.90}$$

Using $\mu = \tau n\theta$ and $\lambda = \beta\tau\sqrt{2\pi\theta/m}$, where β is a constant of the order of unity, we obtain the boundary condition

$$u_0 = s\lambda\left(\frac{\partial u}{\partial z}\right)_0 \tag{5.91}$$

where

$$s = \frac{1-\alpha\beta}{\alpha\beta}$$

is an empirical constant which may be called the "slipping coefficient." When $s = 0$ there is no slipping at the wall. In general the velocity of slip is equal to the velocity in the gas at a distance of s mean free paths from the wall. Usually $s\lambda$ is a few mean free paths.

5.7 VISCOUS HYDRODYNAMICS

The equations of hydrodynamics in the first-order approximation can be obtained by substituting $\mathbf{q}$ and P_{ij}, given respectively in (5.68) and (5.75), into the conservation theorems (5.21)–(5.23). We first evaluate a few relevant quantities.

$$\nabla\cdot\mathbf{q} = -\nabla(K\nabla\theta) = -K\nabla^2\theta - \nabla K\cdot\nabla\theta \tag{5.92}$$

$$\frac{\partial P_{ij}}{\partial x_j} = \frac{\partial P}{\partial x_i} - \mu\left[\nabla^2 u_i + \frac{1}{3}\frac{\partial}{\partial x_i}(\nabla\cdot\mathbf{u})\right] - \frac{2}{m}\frac{\partial\mu}{\partial x_j}\left(\Lambda_{ij} - \frac{m}{3}\delta_{ij}\nabla\cdot\mathbf{u}\right) \tag{5.93}$$

$$P_{ij}\Lambda_{ij} = mP(\nabla\cdot\mathbf{u}) - \frac{2\mu}{m}\Lambda_{ij}\Lambda_{ij} + \tfrac{2}{3}\mu m(\nabla\cdot\mathbf{u})^2 \tag{5.94}$$

The quantity $\Lambda_{ij}\Lambda_{ij}$ can be further reduced:

$$\Lambda_{ij}\Lambda_{ij} = \frac{m^2}{4}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right)\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right) = \frac{m^2}{2}\frac{\partial u_i}{\partial x_j}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right)$$

Now we reduce the two terms above separately:

$$\frac{\partial u_i}{\partial x_j}\frac{\partial u_i}{\partial x_j} = \frac{\partial}{\partial x_j}\left(u_i\frac{\partial u_i}{\partial x_j}\right) - u_i\frac{\partial^2 u_i}{\partial x_j\,\partial x_j} = \tfrac{1}{2}\nabla^2(u^2) - \mathbf{u}\cdot\nabla^2\mathbf{u}$$

$$\frac{\partial u_i}{\partial x_j}\frac{\partial u_j}{\partial x_i} = \left(\frac{\partial u_i}{\partial x_j} - \frac{\partial u_j}{\partial x_i}\right)\left(\frac{\partial u_j}{\partial x_i} - \frac{\partial u_i}{\partial x_j}\right) + \frac{\partial u_i}{\partial x_j}\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\frac{\partial u_j}{\partial x_i} - \frac{\partial u_j}{\partial x_i}\frac{\partial u_i}{\partial x_j}$$

$$= -2(\nabla\times\mathbf{u})^2 + 2\frac{\partial u_i}{\partial x_j}\frac{\partial u_i}{\partial x_j} - \frac{\partial u_j}{\partial x_i}\frac{\partial u_i}{\partial x_j}$$

Hence

$$\frac{\partial u_i}{\partial x_j}\frac{\partial u_j}{\partial x_i} = -(\nabla\times\mathbf{u})^2 + \frac{\partial u_i}{\partial x_j}\frac{\partial u_i}{\partial x_j}$$

and finally

$$\Lambda_{ij}\Lambda_{ij} = \frac{m^2}{2}\left[\nabla^2(u^2) - 2\mathbf{u}\cdot\nabla^2\mathbf{u} - |\nabla\times\mathbf{u}|^2\right] \tag{5.95}$$

Substituting (5.92)–(5.94) into (5.21)–(5.23) we obtain

$$\frac{\partial\rho}{\partial t} + \nabla\cdot(\rho\mathbf{u}) = 0 \tag{5.96}$$

$$\rho\left(\frac{\partial}{\partial t} + \mathbf{u}\cdot\nabla\right)\mathbf{u} = \frac{\mathbf{F}}{m} - \nabla\left(P - \frac{\mu}{3}\nabla\cdot\mathbf{u}\right) + \mu\nabla^2\mathbf{u} + \mathbf{R} \tag{5.97}$$

$$\rho\left(\frac{\partial}{\partial t} + \mathbf{u}\cdot\nabla\right)\theta = \frac{K}{c_V}\nabla^2\theta + \frac{1}{c_V}\nabla K\cdot\nabla\theta - \frac{1}{c_V} \times\left[m\rho(\nabla\cdot\mathbf{u}) + \tfrac{2}{3}\mu m(\nabla\cdot\mathbf{u})^2 - \mu m\left\{\nabla^2(u^2) - 2\mathbf{u}\cdot\nabla^2\mathbf{u} - |\nabla\times\mathbf{u}|^2\right\}\right] \tag{5.98}$$

where $c_V = \frac{3}{2}$ and $\mathbf{R}$ is a vector whose components are given by

$$R_i = \frac{2}{m}\frac{\partial\mu}{\partial x_j}\left(\Lambda_{ij} - \frac{m}{3}\delta_{ij}\nabla\cdot\mathbf{u}\right) \tag{5.99}$$

In these equations the quantities of first-order smallness are μ, K, $\mathbf{u}$, and the derivatives of ρ, θ, and $\mathbf{u}$. Keeping only quantities of first-order smallness, we can neglect all terms involving derivatives of μ and K and the last four terms on the right side of (5.98). We then have the equations of hydrodynamics to the first order:

$$\frac{\partial\rho}{\partial t} + \nabla \cdot (\rho\mathbf{u}) = 0 \qquad \text{(continuity equation)} \tag{5.100}$$

$$\left(\frac{\partial}{\partial t} + \mathbf{u}\cdot\nabla\right)\mathbf{u} = \frac{\mathbf{F}}{m} - \frac{1}{\rho}\nabla\left(P - \frac{\mu}{3}\nabla\cdot\mathbf{u}\right) + \frac{\mu}{\rho}\nabla^2\mathbf{u}$$

$$\text{(Navier-Stokes equation)} \tag{5.101}$$

$$\left(\frac{\partial}{\partial t} + \mathbf{u}\cdot\nabla\right)\theta = -\frac{1}{c_V}(\nabla\cdot\mathbf{u})\theta + \frac{K}{\rho c_V}\nabla^2\theta \qquad \text{(heat conduction equation)} \tag{5.102}$$

where $c_V = \frac{3}{2}$. The boundary condition to be used when a wall is present is the slip boundary condition (5.91).

If $\mathbf{u} = 0$, (5.102) reduces to

$$\rho c_V \frac{\partial\theta}{\partial t} - K\nabla^2\theta = 0 \tag{5.103}$$

which is the familiar diffusion equation governing heat conduction. This equation can be derived intuitively from the fact that $\mathbf{q} = -K\nabla\theta$. Although we have proved this fact only for a dilute gas, it is experimentally correct for liquids and solids as well. For this reason (5.103) is often applied to systems other than a dilute gas.

The Navier-Stokes equation can also be derived on an intuitive basis provided we take the meaning of viscosity from experiments. We discuss this derivation in the next section.

5.8 THE NAVIER-STOKES EQUATION

We give a phenomenological derivation of the Navier-Stokes equation to show why it is expected to be valid even for liquids. Some examples of its use are then discussed.

Consider a small element of fluid whose volume is $dx_1\,dx_2\,dx_3$ and whose velocity is $\mathbf{u}(\mathbf{r}, t)$. According to Newton's second law the equation of motion of this element of fluid is

$$m\frac{d\mathbf{u}}{dt} = \mathscr{F}$$

where m is the mass of the fluid element and $\mathscr{F}$ is the total force acting on the fluid element. Let the mass density of the fluid be ρ and let there be two forces acting on any element of fluid: A force due to agents external to the fluid, and a force due to neighboring fluid elements. These forces *per unit volume* will be

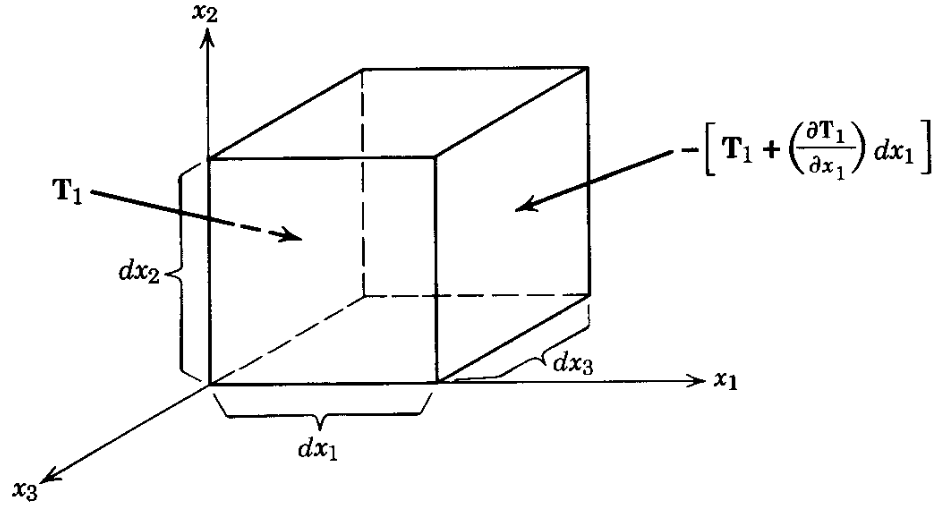


Fig. 5.4 Forces acting on an element of fluid.

respectively denoted by $\mathbf{F}_1$ and $\mathbf{G}$. Thus we can write

$$m = \rho\, dx_1\, dx_2\, dx_3$$

$$\mathscr{X} = (\mathbf{F}_1 + \mathbf{G})\, dx_1\, dx_2\, dx_3$$

Therefore Newton's second law for a fluid element takes the form

$$\rho\left(\frac{\partial}{\partial t} + \mathbf{u} \cdot \nabla\right)\mathbf{u} = \mathbf{F}_1 + \mathbf{G} \tag{5.104}$$

Thus the derivation of the Navier-Stokes equation reduces to the derivation of a definite expression for $\mathbf{G}$.

Let us choose a coordinate system such that the fluid element under consideration is a cube with edges along the three coordinate axes, as shown in Fig. 5.4. The six faces of this cube are subjected to forces exerted by neighboring fluid elements. The force on each face is such that its direction is determined by the direction of the normal vector to the face. That is, its direction depends on which side of the face is considered the "outside." This is physically obvious if we remind ourselves that this force arises from hydrostatic pressure and viscous drag. Let $\mathbf{T}_i$ be the force per unit area acting on the face whose normal lies along the x_i axis. Then the forces *per unit area* acting on the two faces normal to the x_i axis are, respectively (see Fig. 5.4),

$$\mathbf{T}_i, \quad -\left(\mathbf{T}_i + \frac{\partial \mathbf{T}_i}{\partial x_i} dx_i\right) \qquad (i = 1, 2, 3) \tag{5.105}$$

The total force acting on the cube by neighboring fluid elements is then given by

$$\mathbf{G}\, dx_1\, dx_2\, dx_3 = -\left(\frac{\partial \mathbf{T}_1}{\partial x_1} + \frac{\partial \mathbf{T}_2}{\partial x_2} + \frac{\partial \mathbf{T}_3}{\partial x_3}\right) dx_1\, dx_2\, dx_3 \tag{5.106}$$

We denote the components of the vectors $\mathbf{T}_1, \mathbf{T}_2, \mathbf{T}_3$ as follows:

$$\begin{aligned} \mathbf{T}_1 &= (P_{11}, P_{12}, P_{13}) \\ \mathbf{T}_2 &= (P_{21}, P_{22}, P_{23}) \\ \mathbf{T}_3 &= (P_{31}, P_{32}, P_{33}) \end{aligned} \tag{5.107}$$

Then

$$G_i = -\frac{\partial P_{ji}}{\partial x_j} \tag{5.108}$$

or

$$\mathbf{G} = -\nabla \cdot \overleftrightarrow{P} \tag{5.109}$$

With this, (5.104) becomes

$$\rho\left(\frac{\partial}{\partial t} + \mathbf{u} \cdot \nabla\right)\mathbf{u} = \mathbf{F}_1 - \nabla \cdot \overleftrightarrow{P} \tag{5.110}$$

which is of the same form as (5.22) if we set $\mathbf{F}_1 = \rho \mathbf{F}/m$, where $\mathbf{F}$ is the external force per molecule and m is the mass of a molecule. To derive the Navier-Stokes equation, we only have to deduce a more explicit form for P_{ij}. We postulate that (5.110) is valid, whatever the coordinate system we choose. It follows that P_{ij} is a tensor.

We assume the fluid under consideration to be isotropic, so that there can be no intrinsic distinction among the axes x_1, x_2, x_3. Accordingly we must have

$$P_{11} = P_{22} = P_{33} \equiv P \tag{5.111}$$

where P is by definition the hydrostatic pressure. Thus P_{ij} can be written in the form

$$P_{ij} = \delta_{ij} P + P'_{ij} \tag{5.112}$$

where P'_{ij} is a traceless tensor, namely,

$$\sum_{i=1}^{3} P'_{ii} = 0 \tag{5.113}$$

This follows from the fact that (5.113) is true in one coordinate system and that the trace of a tensor is independent of the coordinate system.

Next we make the physically reasonable assumption that the fluid element under consideration, which is really a point in the fluid, has no intrinsic angular momentum. This assumption implies that P_{ij}, and hence P'_{ij}, is a symmetric tensor:

$$P'_{ij} = P'_{ji} \tag{5.114}$$

To see this we need only remind ourselves of the meaning of, for example, P'_{12}. A glance at Fig. 5.5*a* makes (5.114) obvious.

Finally we incorporate into P_{ij} the empirical connection between the shear force applied to a fluid element and the rate of deformation of the same fluid element. A shear force F' per unit area acting parallel to a face of a cube of fluid

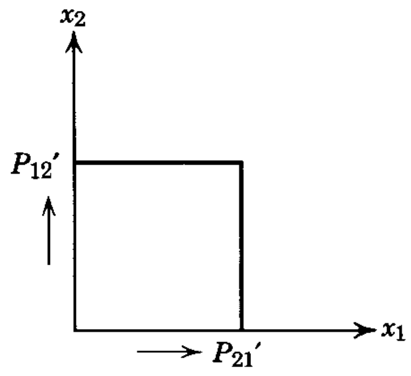


Fig. 5.5a Nonrotation of fluid element implies $P_{12}' = P_{21}'$.

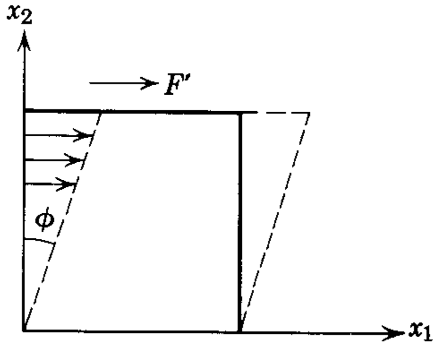


Fig. 5.5b Deformation of fluid element due to shear force.

tends to stretch the cube into a parallelopiped at a rate given by $R' = \mu(d\phi/dt)$, where μ is the coefficient of viscosity and ϕ is the angle shown in Fig. 5.5*b*.

Consider now the effect of P_{12}' on one fluid element. It can be seen from Fig. 5.5*c*, where P_{12}' is indicated in its positive sense in accordance with (5.105), that

$$P_{21}' = -\mu\left(\frac{d\phi_1}{dt} + \frac{d\phi_2}{dt}\right) = -\mu\left(\frac{\partial u_2}{\partial x_1} + \frac{\partial u_1}{\partial x_2}\right) \tag{5.115}$$

In general we have

$$P_{ij}' = -\mu\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right) \qquad (i \neq j) \tag{5.116}$$

To make P_{ij}' traceless we must take

$$P_{ij}' = -\mu\left[\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right) - \frac{2}{3}\delta_{ij}\nabla \cdot \mathbf{u}\right] \tag{5.117}$$

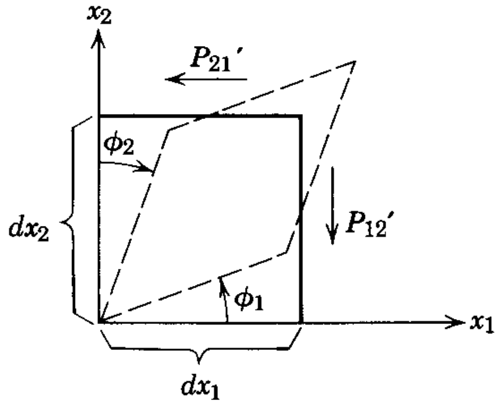


Fig. 5.5c P_{12}' as shear force.

Therefore

$$P_{ij} = \delta_{ij}P - \mu\left[\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right) - \frac{2}{3}\delta_{ij}\nabla \cdot \mathbf{u}\right] \tag{5.118}$$

which is identical in form to (5.75). This completes the phenomenological derivation, which makes it plausible that the Navier-Stokes equation is valid for dilute gas and dense liquid alike.

5.9 EXAMPLES IN HYDRODYNAMICS

To illustrate the mathematical techniques of dealing with the equations of hydrodynamics (5.110)–(5.102), we consider two examples of the application of the Navier-Stokes equation to a liquid.

Incompressible Flow

We consider the following problem: A sphere of radius r is moving with instantaneous velocity $\mathbf{u}_0$ in an infinite, nonviscous, incompressible fluid of constant density in the absence of external force. The Navier-Stokes equation reduces to Euler's equation:

$$\rho\left(\frac{\partial}{\partial t} + \mathbf{u} \cdot \nabla\right)\mathbf{u} = -\nabla P \tag{5.119}$$

where $\mathbf{u}$ is the velocity field of the liquid and P the pressure as given by the equation of state of the fluid. Let us choose the center of the sphere to be the origin of the coordinate system and label any point in space by either the rectangular coordinates (x, y, z) or the spherical coordinates (r, θ, ϕ). The boundary conditions shall be such that the normal component of $\mathbf{u}$ vanishes on the surface of the sphere and that the liquid is at rest at infinity:

$$\begin{gathered}[\mathbf{r} \cdot \mathbf{u}(\mathbf{r})]_{r=a} - (\mathbf{r} \cdot \mathbf{u}_0)_{r=a} = 0 \\ \mathbf{u}(\mathbf{r}) \underset{r=\infty}{\rightarrow} 0\end{gathered} \tag{5.120}$$

Note that incompressibility means $\partial V/\partial P = 0$, or that the density is independent of P. Therefore ∇P is arbitrary, and adjusts itself to whatever the boundary condition demands. Since there is no source for the fluid, we must have everywhere

$$\nabla \cdot \mathbf{u} = 0 \tag{5.121}$$

Taking the curl of both sides of (5.119), remembering that ρ is a constant, and neglecting terms of the form $(\partial \mathbf{u}/\partial x_i)(\partial \mathbf{u}/dx_j)$, we find that

$$\left(\frac{\partial}{\partial t} + \mathbf{u} \cdot \nabla\right)(\nabla \times \mathbf{u}) = 0 \tag{5.122}$$

i.e., that $\nabla \times \mathbf{u}$ is constant along a streamline. Note that P drops out because $\nabla \times (\nabla \times P) \equiv 0$. Since very far from the sphere we have $\nabla \times \mathbf{u} = 0$, it follows that everywhere

$$\nabla \times \mathbf{u} = 0 \tag{5.123}$$

This means that $\mathbf{u}$ is the gradient of some function:

$$\mathbf{u} = \nabla \Phi \tag{5.124}$$

where Φ is called the velocity potential. By (5.120) and (5.121) the equation and boundary conditions for Φ are

$$\begin{gathered} \nabla^2 \Phi(\mathbf{r}) = 0 \\ \left(\frac{\partial \Phi}{\partial r}\right)_{r=a} = u_0 \cos\theta \\ \Phi(\mathbf{r}) \underset{r\to\infty}{\to} 0 \end{gathered} \tag{5.125}$$

where θ is the angle between u_0 and $\mathbf{r}$, as shown in Fig. 5.6.

The most general solution to $\nabla^2\Phi = 0$ is a superposition of solid harmonics.* Since the boundary condition involves $\cos\theta$, we try the solution

$$\Phi(\mathbf{r}) = A \frac{\cos\theta}{r^2} \qquad (r \geq a) \tag{5.126}$$

which is a solid harmonic of order 1 and is the potential that would be set up if a dipole source were placed at the center of the sphere. Choosing $A = -\frac{1}{2}u_0 a^3$ satisfies the boundary conditions. Therefore

$$\Phi(r) = -\tfrac{1}{2}u_0 a^3 \frac{\cos\theta}{r^2} \qquad (r \geq a) \tag{5.127}$$

This is the only solution of (5.125), by the well-known uniqueness theorem of the Laplace equation. The velocity field of the fluid is then given by

$$\mathbf{u}(\mathbf{r}) = -\tfrac{1}{2}u_0 a^3 \nabla \frac{\cos\theta}{r^2} \qquad (r \geq a) \tag{5.128}$$

The streamlines can be sketched immediately, and they look like the electric field due to a dipole, as shown in Fig. 5.6.

Let us calculate the kinetic energy of the fluid. It is given by the integral

$$\begin{aligned} \text{K.E.} = \int d^3r \tfrac{1}{2}\rho|\mathbf{u}|^2 &= \frac{\rho}{2}\left(\frac{u_0 a^3}{2}\right)^2 \int_{r\geq a} d^3r \nabla \frac{\cos\theta}{r^2} \cdot \nabla \frac{\cos\theta}{r^2} \\ &= \frac{\rho}{2}\left(\frac{u_0 a^3}{2}\right)^2 \int_{r\geq a} d^3r \nabla \cdot \left[\frac{\cos\theta}{r^2} \nabla \frac{\cos\theta}{r^2}\right] \\ &= -\frac{\rho}{2}\left(\frac{u_0 a^3}{2}\right)^2 \int_{r=a} d\mathbf{S} \cdot \left[\frac{\cos\theta}{r^2} \nabla \frac{\cos\theta}{r^2}\right] \\ &= -\frac{\rho}{2}\left(\frac{u_0 a^3}{2}\right)^2 a^2 \int_0^{2\pi} d\phi \int_{-1}^{+1} d(\cos\theta) \left(\frac{\cos\theta}{r^2} \frac{\partial}{\partial r} \frac{\cos\theta}{r^2}\right)_{r=a} = \tfrac{1}{2}m' u_0^2 \end{aligned} \tag{5.129}$$

---

*A solid harmonic is $r^l Y_{lm}$ or $r^{-l-1}Y_{lm}$, where Y_{lm} is a spherical harmonic.

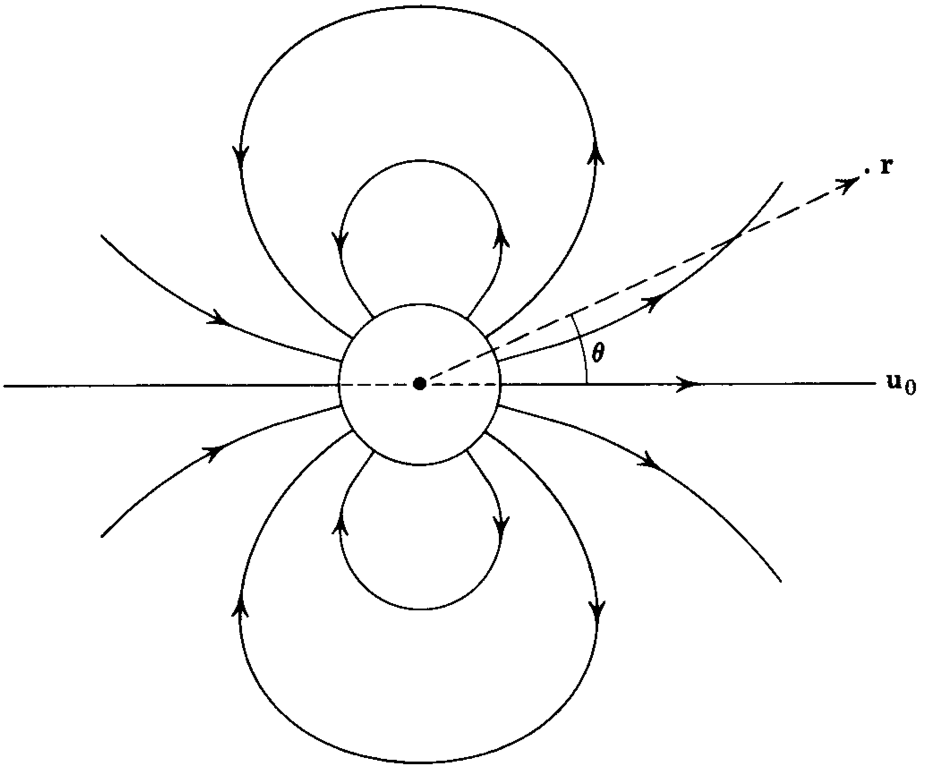


Fig. 5.6 Streamlines in a nonviscous liquid in the presence of a moving sphere.

where m' is half the mass of the displaced fluid:

$$m' = \tfrac{1}{2}\left(\tfrac{4}{3}\pi a^3 \rho\right) \tag{5.130}$$

If the sphere has a mechanical mass m, the total kinetic energy of the system of liquid plus sphere is

$$E = \tfrac{1}{2}(m_0 + m')u_0^2 \tag{5.131}$$

The mass $m_0 + m'$ may be interpreted to be the effective mass of the sphere, since (5.131) is the total energy that has to be supplied for the sphere to move with velocity u_0.

Stokes' Law

We proceed to solve the same problem when the fluid has a nonvanishing coefficient of viscosity μ. The Navier-Stokes equation will be taken to be

$$0 = -\nabla\left(P - \frac{\mu}{3}\nabla \cdot \mathbf{u}\right) + \mu\nabla^2\mathbf{u} \tag{5.132}$$

on the assumption that the material derivative of **u**, which gives rise to the effective mass, is small compared to the viscous terms. We return to examine the validity of this approximation later. Since there is no source for the fluid, we still require $\nabla \cdot \mathbf{u} = 0$, and (5.132) becomes the simultaneous equations

$$\begin{aligned} \nabla^2\mathbf{u} &= \frac{1}{\mu}\nabla P \\ \nabla \cdot \mathbf{u} &= 0 \end{aligned} \tag{5.133}$$

with the boundary condition that the fluid sticks to the sphere. Let us translate the coordinate system so that the sphere is at rest at the origin while the fluid at infinity flows with uniform constant velocity $\mathbf{u}_0$. The equations (5.133) remain invariant under the translation, whereas the boundary conditions become

$$\begin{gathered}[\mathbf{u}(\mathbf{r})]_{r=a} = 0 \\ \mathbf{u}(\mathbf{r}) \underset{r\to\infty}{\to} \mathbf{u}_0\end{gathered} \tag{5.134}$$

Taking the divergence of both sides of the first equation of (5.133), we obtain

$$\nabla^2 P = 0 \tag{5.135}$$

Thus the pressure, whatever it is, must be a linear superposition of solid harmonics. A systematic way to proceed would be to write P as the most general superposition of solid harmonics and to determine the coefficient by requiring that (5.133) be satisfied. We take a short cut, however, and *guess* that P is, apart from an additive constant, a pure solid harmonic of order 1:

$$P = P_0 + \mu P_1 \frac{\cos\theta}{r^2} \tag{5.136}$$

where P_0 and P_1 are constants to be determined later. With this, the problem reduces to solving the inhomogeneous Laplace equation

$$\nabla^2 \mathbf{u} = P_1 \nabla \frac{\cos\theta}{r^2} \tag{5.137}$$

subject to the conditions

$$\begin{gathered}\nabla\cdot\mathbf{u} = 0 \\ [\mathbf{u}(\mathbf{r})]_{r=a} = 0 \\ \mathbf{u}(\mathbf{r}) \underset{r\to\infty}{\to} \mathbf{u}_0\end{gathered} \tag{5.138}$$

A particular solution of (5.137) is

$$\mathbf{u}_1 = -\frac{P_1}{6} r^2 \nabla \frac{\cos\theta}{r^2} = -\frac{P_1}{6}\left(\frac{\hat{z}}{r} - 3\mathbf{r}\frac{z}{r^3}\right) \tag{5.139}$$

where $\hat{z}$ denotes the unit vector along the z axis, which lies along $\mathbf{u}_0$. It is easily verified that (5.139) solves (5.137), if we note that $1/r$ and z/r^3 are both solid harmonics. Thus,

$$\nabla^2 \mathbf{u}_1 = -\frac{P_1}{6}\left[-3\nabla^2\left(\frac{\mathbf{r}z}{r^3}\right)\right] = P_1 \nabla\left(\frac{z}{r^3}\right) = P_1 \nabla \frac{\cos\theta}{r^2} \tag{5.140}$$

The complete solution is obtained by adding an appropriate homogeneous solution to (5.139) to satisfy (5.138). By inspection we see that the complete solution is

$$\mathbf{u} = \mathbf{u}_0\left(1 - \frac{a}{r}\right) + \frac{1}{4} u_0 a (r^2 - a^2) \nabla \frac{\cos\theta}{r^2} \tag{5.141}$$

where we have set

$$P_1 = -\tfrac{3}{2}u_0 a \tag{5.142}$$

to have $\nabla \cdot \mathbf{u} = 0$.

We now calculate the force acting on the sphere by the fluid. By definition the force per unit area acting on a surface whose normal point along the x_j axis is $-\mathbf{T}_j$ of (5.107). It follows that the force per unit area acting on a surface element of the sphere is

$$\mathbf{f} = -\left(\frac{x}{r}\mathbf{T}_1 + \frac{y}{r}\mathbf{T}_2 + \frac{z}{r}\mathbf{T}_3\right) = -\hat{r} \cdot \overleftrightarrow{P} \tag{5.143}$$

where $\hat{r}$ is the unit vector in the radial direction and $\overleftrightarrow{P}$ is given by (5.118). The total force experienced by the sphere is

$$\mathbf{F}' = \int dS\,\mathbf{f} \tag{5.144}$$

where dS is a surface element of the sphere and the integral extends over the entire surface of the sphere. Thus it is sufficient to calculate $\mathbf{f}$ for $r = a$.

The vector $\hat{r} \cdot \overleftrightarrow{P}$ has the components

$$\begin{aligned}(\hat{r} \cdot \overleftrightarrow{P})_i &= \frac{1}{r}x_j P_{ji} = \frac{1}{r}x_j\left[\delta_{ji}P - \mu\left(\frac{\partial u_j}{\partial x_i} + \frac{\partial u_i}{\partial x_j}\right)\right] \\ &= \frac{x_i}{r}P - \frac{\mu}{r}\left[\frac{\partial}{\partial x_i}(x_j u_j) - u_i + x_j\frac{\partial}{\partial x_j}u_i\right]\end{aligned}$$

Hence

$$\mathbf{f} = -\hat{r}P + \frac{\mu}{r}[\nabla(\mathbf{r} \cdot \mathbf{u}) - \mathbf{u} + (\mathbf{r} \cdot \nabla)\mathbf{u}] \tag{5.145}$$

where P is given by (5.136) and (5.142), and $\mathbf{u}$ is given by (5.141). Since $\mathbf{u} = 0$ when $r = a$, we only need to consider the first and the last terms in the bracket. The first term is zero at $r = a$ by a straightforward calculation. At $r = a$ the second term is found to be

$$\frac{1}{r}[(\mathbf{r} \cdot \nabla)\mathbf{u}]_{r=a} = \left(\frac{\partial \mathbf{u}}{\partial r}\right)_{r=a} = \frac{3}{2}\frac{\mathbf{u}_0}{a} - \frac{3}{2}\hat{r}u_0\frac{\cos\theta}{a} \tag{5.146}$$

When this is substituted into (5.145), the second term exactly cancels the dipole part of $\hat{r}P$, and we obtain

$$(\mathbf{f})_{r=a} = -\hat{r}P_0 + \frac{3}{2}\frac{\mu}{a}\mathbf{u}_0$$

The constant P_0 is unknown, but it does not contribute to the force on the sphere. From (5.144) we obtain

$$\mathbf{F}' = 6\pi\mu a\mathbf{u}_0 \tag{5.147}$$

which is Stokes' law.

The validity of (5.141) depends on the smallness of the material derivative of $\mathbf{u}$ as compared to $\mu \nabla^2 \mathbf{u}$. Both these quantities can be computed from (5.141). It is then clear that we must require

$$\frac{\rho u_0 a}{\mu} \ll 1 \tag{5.148}$$

Thus Stokes' law holds only for small velocities and small radii of the sphere. A more elaborate treatment shows that a more accurate formula for $\mathbf{F}'$ is

$$\mathbf{F}' = 6\pi\mu a \mathbf{u}_0 \left(1 + \frac{3}{8}\frac{\rho u_0 a}{\mu} + \cdots\right) \tag{5.149}$$

The pure number $\rho u_0 a/\mu$ is called the Reynolds number. When the Reynolds number becomes large, turbulence sets in and streamline motion completely breaks down.

PROBLEMS

5.1 Make order-of-magnitude estimates for the mean free path and the collision time for

(*a*) H_2 molecules in a hydrogen gas in standard condition (diameter of $H_2 = 2.9$ Å);

(*b*) protons in a plasma (gas of totally ionized H_2) at $T = 3 \times 10^5$ K, $n = 10^{15}$ protons/cm^3, $\sigma = \pi r^2$, where $r = e^2/kT$;

(*c*) protons in a plasma at the same density as (*b*) but at $T = 10^7$ K, where thermonuclear reactions occur;

(*d*) protons in the sun's corona, which is a plasma at $T = 10^6$ K, $n = 10^6$ protons/cc;

(*e*) slow neutrons of energy 0.5 MeV in ^{238}U ($\sigma \approx \pi r^2$, $r \approx 10^{-13}$ cm).

5.2 A box made of perfectly reflecting walls is divided by a perfectly reflecting partition into compartments 1 and 2. Initially a gas at temperature T_1 was confined in compartment 1, and compartment 2 was empty. A small hole of dimension much less than the mean free path of the gas is opened in the partition for a short time to allow a small fraction of the gas to escape into compartment 2. The hole is then sealed off and the new gas in compartment 2 comes to equilibrium.

(*a*) During the time when the hole was open, what was the flux dI of molecules crossing into compartment 2 with speed between v and $v + dv$?

(*b*) During the same time, what was the average energy per particle ϵ of the molecules crossing into compartment 2?

(*c*) After final equilibrium has been established, what is the temperature T_2 in compartment 2?

Answer. $T_2 = \frac{4}{3}T_1$.

5.3 (*a*) Explain why it is meaningless to speak of a sound wave in a gas of strictly noninteracting molecules.

(*b*) In view of (*a*), explain the meaning of a sound wave in an ideal gas.

5.4 Show that the velocity of sound in a real substance is to a good approximation given by $c = 1/\sqrt{\rho\kappa_S}$, where ρ is the mass density and κ_S the adiabatic compressibility, by the following steps.

(*a*) Show that in a sound wave the density oscillates adiabatically if

$$K \ll c\lambda\rho c_V$$

where K = coefficient of thermal conductivity
λ = wavelength of sound wave
ρ = mass density
c_V = specific heat
c = velocity of sound

(*b*) Show by numerical examples, that the criterion stated in (*a*) is well satisfied in most practical situations.

5.5 A flat disk of unit area is placed in a dilute gas at rest with initial temperature T. Face A of the disk is at temperature T, and face B is at temperature $T_1 > T$ (see sketch). Molecules striking face A reflect elastically. Molecules striking face B are absorbed by the disk, only to re-emerge from the same face with a Maxwellian distribution of temperature T_1.

(*a*) Assume that the mean free path in the gas is much smaller than the dimension of the disk. Present an argument to show that after a few collision times the gas can be described by the hydrodynamic equations, with face B replaced by a boundary condition for the temperature.

(*b*) Write down the first-order hydrodynamic equations for (*a*), neglecting the flow of the gas. Show that there is no net force acting on the disk.

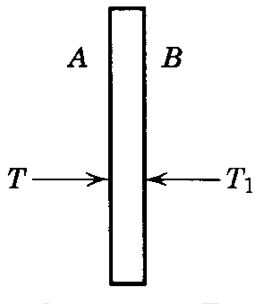
(*c*) Assume that the mean free path is much larger than the dimensions of the disk. Find the net force acting on the disk.

5.6 A square vane, of area 1 cm^2, painted white on one side, black on the other, is attached to a vertical axis and can rotate freely about it (see the sketch). Suppose the arrangement is placed in He gas at room temperature and sunlight is allowed to shine on the vane. Explain qualitatively why

(*a*) at high density of the gas the vane does not move;

(*b*) at extremely small densities the vane rotates;

(*c*) at some intermediate density the vane rotates in a sense opposite to that in (*b*). *Estimate* this intermediate density and the corresponding pressure.



Gas at temp. T **Fig. P5.5**

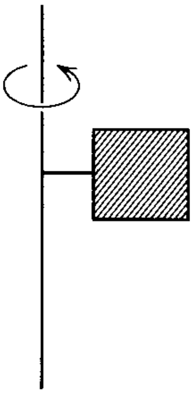


Fig. P5.6

5.7 A dilute gas, infinite in extension and composed of charged molecules, each of charge e and mass m, comes to equilibrium in an infinite lattice of fixed ions. In the absence of an external electric field the equilibrium distribution function is

$$f^{(0)}(\mathbf{p}) = n(2\pi mkT)^{-3/2} e^{-p^2/2mkT}$$

where n and T are constants. A weak uniform electric field E is then turned on, leading to a new equilibrium distribution function. Assume that a collision term of the form

$$\left(\frac{\partial f}{\partial t}\right)_{\text{coll}} = -\frac{f - f^{(0)}}{\tau}$$

where τ is a collision time, adequately takes into account the effect of collisions among molecules and between molecules and lattice. Calculate

(*a*) the new equilibrium distribution function f, to the first order;

(*b*) the electrical conductivity σ, defined by the relation

$$ne\langle \mathbf{v} \rangle = \sigma \mathbf{E}$$