

# KINETIC THEORY OF GASES AND THE BOLTZMANN EQUATION

In Section 7-2 we introduced the concept of phase space and distribution functions in phase space. We also derived the Liouville equation, which is the equation of motion that the phase space distribution function must satisfy. Since we were interested only in equilibrium statistical mechanics at that time, we did not consider the Liouville equation in any detail. In this chapter we shall review the concept of phase space and derive the Liouville equation again. We shall then introduce reduced distribution functions and derive the Bogoliubov, Born, Green, Kirkwood, Yvon (BBGKY) hierarchy. This hierarchy is the nonequilibrium generalization of the Kirkwood integral equation hierarchy for the fluid distribution functions,  $g^{(n)}(\mathbf{r}_1, \dots, \mathbf{r}_n)$ , of Chapter 13. Nobody has yet devised a successful way to uncouple the BBGKY hierarchy, and so in Section 18-4 we shall derive a physical, yet approximate, equation for the distribution function for gases. This equation, called the Boltzmann equation, is the central equation of the rigorous kinetic theory of gases. In Section 18-5, we shall derive some of the general consequences of the Boltzmann equation that can be determined without actually solving it completely. We shall discuss its solution in Chapter 19. The standard reference for most of this chapter is Hirschfelder, Curtiss, and Bird. Mazo (in "Additional Reading") also discusses these topics well.

## 18-1 PHASE SPACE AND THE LIOUVILLE EQUATION

Consider a system of  $N$  point particles. The classical dynamical state of this system is specified by the  $3N$  momentum components  $p_1, p_2, \dots, p_{3N}$  and the  $3N$  spatial coordinates  $q_1, \dots, q_{3N}$ . We can construct a  $6N$ -dimensional space whose coordinates are  $q_1, q_2, \dots, p_1, \dots, p_{3N}$ . One point in this *phase space* completely specifies the microscopic dynamical state of our  $N$ -particle system. As the system evolves in time, this *phase point* moves through phase space in a manner completely dictated by the equations of motion of the system. Actually, one never knows (nor really cares to know) the  $6N$  coordinates of a macroscopic system. Rather, one knows just a few

macroscopic mechanical properties of the system, such as the energy, volume, velocity, etc. Clearly there are a great number of points in phase space that are compatible with the few variables that we know about the system. The set of all such phase points constitutes an *ensemble* of systems. The number of systems in an ensemble approaches infinity, and so the set of phase points that could possibly represent our system becomes quite dense. This allows us to define a *density of phase points* or *distribution function* as the fraction of phase points contained in the volume  $dq_1 dq_2 \dots dp_{3N}$ . We shall denote the phase space distribution function by  $f_N(q_1, q_2, \dots, p_{3N}, t)$ , or more conveniently by  $f_N(p, q, t)$ . We shall often use this abbreviated notation. Similarly, we shall often denote  $dq_1, dq_2 \dots dp_{3N}$  by  $dp dq$ . The density  $f_N(p, q, t)$  is normalized such that

$$\int f_N(p, q, t) dp dq = 1$$

Since each phase point moves in time according to the equations of motion of the system it describes,  $f_N$  itself must obey some sort of equation of motion. The equation that  $f_N(p, q, t)$  satisfies can be readily determined by using the methods of the previous chapter, particularly, the argument associated with Eqs. (17-1) to (17-5). The number of phase points within some arbitrary volume  $v$  is

$$n = \mathcal{N} \int_v f_N(p, q, t) dp dq$$

where we are using the condensed notation of letting  $p$  and  $q$  denote all the spatial coordinates and momenta necessary to specify a system in the ensemble. The rate of change of the number of phase points within  $v$  is

$$\frac{dn}{dt} = \mathcal{N} \int_v \frac{\partial f_N}{\partial t} dp dq \quad (18-1)$$

Since phase points are neither created nor destroyed, the rate of change of  $n$  must be given by the rate at which phase points flow through the surface enclosing  $v$ . The rate of flow of phase points is  $\mathcal{N} f_N \mathbf{u}$ , where  $\mathbf{u}$  is not just the  $3N$ -dimensional vector  $(\dot{q}_1, \dot{q}_2, \dots, \dot{q}_{3N})$ , but the  $6N$ -dimensional vector  $(\dot{q}_1, \dots, \dot{p}_1, \dots, \dot{p}_{3N})$  since the spatial coordinates and momenta play an equivalent role in phase space. We integrate this flow over the surface to get

$$\frac{dn}{dt} = -\mathcal{N} \int_S f_N \mathbf{u} \cdot d\mathbf{S}$$

The negative sign here indicates that an outflow of phase points yields a negative value for  $dn/dt$  since  $\mathbf{u} \cdot d\mathbf{S}$  is positive if  $\mathbf{u}$  is directed outward from  $v$  and negative if  $\mathbf{u}$  is directed inward.

The surface integral can be transformed to a volume integral by using Gauss' theorem to get

$$\frac{dn}{dt} = -\mathcal{N} \int_v \nabla \cdot (f_N \mathbf{u}) dp dq \quad (18-2)$$

If we subtract Eq. (18-1) from Eq. (18-2) and realize that this equation is valid for any choice of  $v$ , we have the equation for the conservation of phase points

$$\frac{\partial f_N}{\partial t} + \nabla \cdot (f_N \mathbf{u}) = 0 \quad (18-3)$$

in which it should be clear that since we are dealing with phase space

$$\mathbf{u} = (\dot{q}_1, \dots, \dot{q}_{3N}, \dot{p}_1, \dots, \dot{p}_{3N})$$

and

$$\begin{aligned} \nabla \cdot f_N \mathbf{u} &= \sum_{j=1}^{3N} \frac{\partial}{\partial q_j} (f_N \dot{q}_j) + \sum_{j=1}^{3N} \frac{\partial}{\partial p_j} (f_N \dot{p}_j) \\ &= \sum_{j=1}^{3N} \left\{ \frac{\partial f_N}{\partial q_j} \dot{q}_j + \frac{\partial f_N}{\partial p_j} \dot{p}_j \right\} + \sum_{j=1}^{3N} \left\{ \frac{\partial \dot{q}_j}{\partial q_j} + \frac{\partial \dot{p}_j}{\partial p_j} \right\} f_N \end{aligned}$$

But Eq. (7-27) shows that the summand of the second summation here is zero, and so Eq. (18-3) becomes

$$\frac{\partial f_N}{\partial t} + \sum_{j=1}^{3N} \frac{\partial f_N}{\partial q_j} \dot{q}_j + \sum_{j=1}^{3N} \frac{\partial f_N}{\partial p_j} \dot{p}_j = 0 \quad (18-4)$$

Using Hamilton's equations of motion,

$$\dot{p}_i = -\frac{\partial H}{\partial q_i} \quad \text{and} \quad \dot{q}_i = \frac{\partial H}{\partial p_i}$$

Eq. (18-4) can be written

$$\frac{\partial f_N}{\partial t} + \sum_{j=1}^{3N} \left( \frac{\partial H}{\partial p_j} \frac{\partial f_N}{\partial q_j} - \frac{\partial H}{\partial q_j} \frac{\partial f_N}{\partial p_j} \right) = 0 \quad (18-5)$$

The summation here is called a Poisson bracket and is commonly denoted by  $\{H, f_N\}$ ; so Eq. (18-5) is often written as

$$\frac{\partial f_N}{\partial t} + \{H, f_N\} = 0 \quad (18-6)$$

This is the Liouville equation, the most fundamental equation of statistical mechanics. In fact, it can be shown that the Liouville equation is equivalent to the  $6N$  Hamilton equations of motion of the  $N$ -body system.\*

In Cartesian coordinates, the Liouville equation reads

$$\frac{\partial f_N}{\partial t} + \sum_{j=1}^N \frac{\mathbf{p}_j}{m_j} \cdot \nabla_{\mathbf{r}_j} f_N + \sum_{j=1}^N \mathbf{F}_j \cdot \nabla_{\mathbf{p}_j} f_N = 0 \quad (18-6')$$

In this equation  $\nabla_{\mathbf{r}_j}$  denotes the gradient with respect to the spatial variables in  $f_N$ ;  $\nabla_{\mathbf{p}_j}$  denotes the gradient with respect to the momentum variables in  $f_N$ ; and  $\mathbf{F}_j$  is the total force on the  $j$ th particle.

One often sees the Liouville equation written as

$$i \frac{\partial f_N}{\partial t} = L f_N \quad (18-7)$$

where  $L$  is the Liouville operator,

$$L = -i \left( \sum_{j=1}^N \frac{\mathbf{p}_j}{m_j} \cdot \nabla_{\mathbf{r}_j} + \sum_{j=1}^N \mathbf{F}_j \cdot \nabla_{\mathbf{p}_j} \right) \quad (18-8)$$

\* Mazo, "Additional Reading," p. 23; M. Beran, *Amer. J. Phys.*, 35, p. 242, 1967.

The Liouville operator has been defined in such a way as to bring the Liouville equation into the form of the Schrödinger equation. A formal, and sometimes useful, solution to Eq. (18-7) is

$$f_N(\mathbf{p}, \mathbf{r}, t) = e^{-iLt} f_N(\mathbf{p}, \mathbf{r}, 0) \quad (18-9)$$

Note that the operator  $\exp(-iLt)$  displaces  $f_N$  ahead a distance  $t$  in time. This operator is called the *time displacement operator* of the system.

## 18-2 REDUCED DISTRIBUTION FUNCTIONS

Once we have the distribution function  $f_N(p, q, t)$ , we may compute the ensemble average of any dynamical variable,  $A(p, q, t)$ , from the equation

$$\langle A(t) \rangle = \int A(p, q, t) f_N(p, q, t) dp dq \quad (18-10)$$

It turns out that the dynamical variables of interest are functions of either the coordinates and momenta of just a few particles or can be written as a sum over such functions. A familiar example of this is the total intermolecular potential of the system. To a good approximation, this can be written as a sum over pair-wise potentials, and so

$$\langle U \rangle = \sum_{i,j} \int \cdots \int u(\mathbf{r}_i, \mathbf{r}_j) f_N(\mathbf{r}_1, \dots, \mathbf{p}_N, t) d\mathbf{r}_1 \cdots d\mathbf{p}_N \quad (18-11)$$

We encountered similar integrands when we studied the equilibrium theory of liquids. There we integrated over the coordinates of all the particles except  $i$  and  $j$  and called the resulting function of  $\mathbf{r}_i$  and  $\mathbf{r}_j$  a radial distribution function. We do the same thing here. We define *reduced distribution functions*  $f_N^{(n)}(\mathbf{r}_1, \dots, \mathbf{r}_n, \mathbf{p}_1, \dots, \mathbf{p}_n, t)$  by

$$f_N^{(n)}(\mathbf{r}_1, \dots, \mathbf{r}_n, \mathbf{p}_1, \dots, \mathbf{p}_n, t) = \frac{N!}{(N-n)!} \int \cdots \int f_N(\mathbf{r}_1, \dots, \mathbf{p}_N, t) d\mathbf{r}_{n+1} \cdots d\mathbf{r}_N d\mathbf{p}_{n+1} \cdots d\mathbf{p}_N \quad (18-12)$$

We shall usually drop the  $N$  subscript and furthermore write this simply as  $f^{(n)}(\mathbf{r}^n, \mathbf{p}^n, t)$ . Usually only  $f^{(1)}$  and  $f^{(2)}$  are necessary, and therefore we want to derive an equation for  $f^{(1)}$  and  $f^{(2)}$ . To do this, write the force  $\mathbf{F}_j$  appearing in the Liouville equation as the sum of the forces due to the other molecules in the system  $\sum_i \mathbf{F}_{ij}$  and an external force  $\mathbf{X}_j$ . Then multiply through by  $N!/(N-n)!$  and integrate over  $d\mathbf{r}_{n+1} \cdots d\mathbf{r}_N d\mathbf{p}_{n+1} \cdots d\mathbf{p}_N$  to get (Problem 18-2)

$$\begin{aligned} \frac{\partial f^{(n)}}{\partial t} + \sum_{j=1}^n \frac{\mathbf{p}_j}{m_j} \cdot \nabla_{\mathbf{r}_j} f^{(n)} + \sum_{j=1}^n \mathbf{X}_j \cdot \nabla_{\mathbf{p}_j} f^{(n)} \\ + \frac{N!}{(N-n)!} \sum_{i,j=1}^n \int \cdots \int \mathbf{F}_{ij} \cdot \nabla_{\mathbf{p}_j} f d\mathbf{r}_{n+1} \cdots d\mathbf{r}_N d\mathbf{p}_{n+1} \cdots d\mathbf{p}_N = 0 \end{aligned} \quad (18-13)$$

We have used the fact that  $f$  vanishes outside the walls of the container and when  $\mathbf{p}_i = \pm \infty$ . The last term in Eq. (18-13) can be broken up into two parts:

$$\begin{aligned} \sum_{i,j=1}^n \mathbf{F}_{ij} \cdot \nabla_{\mathbf{p}_j} f^{(n)} \\ + \frac{N!}{(N-n)!} \sum_{j=1}^n \sum_{i=n+1}^N \int \cdots \int \mathbf{F}_{ij} \cdot \nabla_{\mathbf{p}_j} f d\mathbf{r}_{n+1} \cdots d\mathbf{r}_N d\mathbf{p}_{n+1} \cdots d\mathbf{p}_N \end{aligned}$$



The second term here can be written as

$$\sum_{j=1}^n \iint \mathbf{F}_{j,n+1} \cdot \nabla_{\mathbf{p}_j} f^{(n+1)} d\mathbf{r}_{n+1} d\mathbf{p}_{n+1}$$

Putting all this together finally gives an exact equation for  $f^{(n)}$ , namely, (Problem 18-3),

$$\begin{aligned} \frac{\partial f^{(n)}}{\partial t} + \sum_{j=1}^n \frac{\mathbf{p}_j}{m_j} \cdot \nabla_{\mathbf{r}_j} f^{(n)} + \sum_{j=1}^n \mathbf{X}_j \cdot \nabla_{\mathbf{p}_j} f^{(n)} \\ + \sum_{i,j=1}^n \mathbf{F}_{ij} \cdot \nabla_{\mathbf{p}_i} f^{(n)} + \sum_{j=1}^n \iint \mathbf{F}_{j,n+1} \cdot \nabla_{\mathbf{p}_j} f^{(n+1)} d\mathbf{r}_{n+1} d\mathbf{p}_{n+1} = 0 \end{aligned} \quad (18-14)$$

This is the so-called Bogoliubov, Born, Green, Kirkwood, Yvon (BBGKY) hierarchy. This is the time-dependent generalization of the hierarchy that we derived earlier in the equilibrium theory of fluids. In fact, if one assumes that

$$f^{(n)} = g^{(n)}(\mathbf{r}_1, \dots, \mathbf{r}_n) \exp\left\{-\frac{1}{2mkT} \sum_{j=1}^n p_j^2\right\}$$

multiplies Eq. (18-14) through by  $\mathbf{p}_i$ ,  $1 \leq i \leq n$ , and integrates over all momenta, one obtains the equilibrium hierarchy for  $g^{(n)}(\mathbf{r}_1, \dots, \mathbf{r}_n)$  (Problem 18-4). It would seem natural at this point to truncate this hierarchy by some sort of a superposition approximation, but so far this approach has not been successful.\* We shall end up deriving approximate equations for  $f^{(1)}$  and  $f^{(2)}$ .

Everything we have done up to now has been independent of density; i.e., it has been applicable to any density. Now we shall specialize to systems of dilute gases.

### 18-3 FLUXES IN DILUTE GASES

In a dilute gas, most of the molecules are not interacting with any other molecule and are just traveling along between collisions. Because of this, the macroscopic properties of a gas depend upon only the singlet distribution function  $f_j^{(1)}(\mathbf{r}, \mathbf{p}_j, t)$ . The subscript  $j$  here denotes the singlet distribution function of species  $j$ . This is the central distribution function of any theory of transport in dilute gases. In this section we shall define a number of averages over  $f_j^{(1)}$  and derive molecular expressions for the important flux quantities in terms of integrals over  $f_j^{(1)}$ . Since we shall be concerned only with gases in this and the following sections, we shall drop the superscript (1) from here on. We shall also write our equations in velocity space rather than momentum space, and so the distribution function of interest becomes  $f_j(\mathbf{r}, \mathbf{v}_j, t)$ . We shall renormalize  $f_j$  such that the integral of this distribution function over all velocities is the number density of  $j$  particles at the point  $\mathbf{r}$  at time  $t$ , i.e.,

$$\rho_j(\mathbf{r}, t) = \int f_j(\mathbf{r}, \mathbf{v}_j, t) d\mathbf{v}_j \quad (18-15)$$

Furthermore, if  $N_j$  is the total number of  $j$  molecules in our system, then

$$N_j = \iint f_j(\mathbf{r}, \mathbf{v}_j, t) d\mathbf{r} d\mathbf{v}_j \quad (18-16)$$

We shall now define a number of important average velocities.  $\mathbf{v}_j$  is the *linear*

\* See, for example, R. G. Mortimer, *J. Chem. Phys.*, 48, p. 1023, 1968.

*velocity* of a molecule of species  $j$ ; i.e., it is the velocity with respect to a coordinate system fixed in space. The average velocity is given by

$$\mathbf{v}_j(\mathbf{r}, t) = \frac{1}{\rho_j} \int \mathbf{v}_j f(\mathbf{r}, \mathbf{v}_j, t) d\mathbf{v}_j \quad (18-17)$$

and represents the macroscopic flow of species  $j$ . The *mass average velocity* is defined by

$$\mathbf{v}_0(\mathbf{r}, t) = \frac{\sum_j m_j \rho_j \mathbf{v}_j}{\sum_j m_j \rho_j} \quad (18-18)$$

Note that the denominator here is the mass density  $\rho_m(\mathbf{r}, t)$ . This velocity is often called the flow velocity or stream velocity. The momentum density of the gas is the same as if all the molecules were moving with velocity  $\mathbf{v}_0$ . The *peculiar velocity* is the velocity of a molecule relative to the flow velocity. The peculiar velocity  $\mathbf{V}_j$  is

$$\mathbf{V}_j = \mathbf{v}_j - \mathbf{v}_0 \quad (18-19)$$

The average of this peculiar velocity is the *diffusion velocity* (Problem 18-5). Clearly,

$$\bar{\mathbf{V}}_j = \frac{1}{\rho_j} \int (\mathbf{v}_j - \mathbf{v}_0) f_j(\mathbf{r}, \mathbf{v}_j, t) d\mathbf{v}_j \quad (18-20)$$

It is easy to show that (Problem 18-6)

$$\sum_j \rho_j m_j \bar{\mathbf{V}}_j = 0 \quad (18-21)$$

When we studied the elementary kinetic theory of gases, we saw that the various transport coefficients were related to molecular transport of mass, momentum, and kinetic energy. Let these molecular properties be designated collectively by  $\psi_j$ , where  $j$  refers to the particular species. We now derive expressions for the fluxes of these properties. Figure 18-1 shows a surface  $dS$  moving with velocity  $\mathbf{v}_0$ . The quantity  $\mathbf{n}$  is a unit vector normal to  $dS$ , and  $dS = \mathbf{n} dS$ . All the molecules that have velocity  $\mathbf{V}_j = \mathbf{v}_j - \mathbf{v}_0$  and that cross  $dS$  in the time interval  $(t, t + dt)$  must have been in a cylinder of length  $|\mathbf{V}_j| dt$  and base  $dS$ . This cylinder is shown in Fig. 18-1 and has a volume  $(\mathbf{n} \cdot \mathbf{V}_j) dS dt$ . Since there are  $f_j d\mathbf{v}_j$  molecules per unit volume with relative velocity  $\mathbf{V}_j$ , the number of  $j$  molecules that cross  $dS$  in  $dt$  is given by

$$(f_j d\mathbf{v}_j)(\mathbf{n} \cdot \mathbf{V}_j) dS dt$$

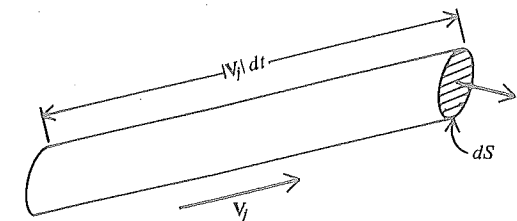


Figure 18-1. The cylinder containing all those molecules of species  $j$  with velocity  $\mathbf{V}_j$ , which cross the surface  $dS$  during the time interval  $dt$ . (From J. O. Hirschfelder, C. F. Curtiss, and R. B. Bird, *Molecular Theory of Gases and Liquids*. New York: Wiley, 1954.)