John Wiley & Sons
New York 1983
ISBN 0-471-09045-X

IMPORTANT FORMULAS

(For electrons; note how each formula scales with mass and charge.)

Plasma Frequency: $\omega_e(s^{-1}) = \left(\frac{4\pi ne^2}{m_e}\right)^{1/2} = 2\pi\ 9000\sqrt{n(cm^{-3})}$

Debye Length: $\lambda_e(cm) = \left(\frac{T_e}{4\pi ne^2}\right)^{1/2} = 740\sqrt{T_e(eV)/n(cm^{-3})}$

Gyrofrequency: $|\Omega_e|(s^{-1}) = \frac{eB}{m_e c} = 2 \times 10^7\ B(Gauss)$

Plasma Parameter: $\Lambda = n\lambda_e^3 = 4 \times 10^8\ T_e^{3/2}(eV)/n^{1/2}(cm^{-3})$

Speed: $v(cm/s) = (2E/m_e)^{1/2} = 6 \times 10^7\ E^{1/2}(eV)$

Thermal Speed: $v_e(cm/s) = (T_e/m_e)^{1/2} = 4 \times 10^7\ T_e^{1/2}(eV)$

Gyroradius: $r_g(cm) = v_\perp/\Omega = 3E_\perp^{1/2}(eV)/B(Gauss)$

D.R. Nicholson
Introduction to Plasma Theory

Preface

The purpose of this book is to teach the basic theoretical principles of plasma physics. It is not intended to be an encyclopedia of results and techniques. Nor is it intended to be used primarily as a reference book. It is intended to develop the basic techniques of plasma physics from the beginning, namely, from Maxwell's equations and Newton's law of motion. Absolutely no previous knowledge of plasma physics is assumed. Although the book is primarily intended for a one year course at the first or second year graduate level, it can also be used for a one or two semester course at the junior or senior undergraduate level. Such an undergraduate course would make use of that half of the book which assumes a knowledge only of undergraduate electricity and magnetism. The other half of the book, suitable for the graduate level, requires familiarity with complex variables, Fourier transformation, and the Dirac delta function.

The book is organized in a logical fashion. Although this is not the standard organization of an introductory course in plasma physics, I have found that students at the graduate level respond well to this organization. After the introductory material of Chapters 1 and 2 (single particle motion), the exact theories of Chapters 3 to 5 (Klimontovich and Liouville equations), which are equivalent to Maxwell's equations plus Newton's law of motion, are replaced via approximations by the Vlasov equation of Chapter 6. Further approximations lead to the fluid theory (Chapter 7) and magnetohydrodynamic theory (Chapter 8). The book concludes with two chapters on discrete particle effects (Chapter 9) and weak turbulence theory (Chapter 10). Chapter 6, and Chapters 7 and 8, are meant to be self-contained, so that the book can easily be used by instructors who wish the standard organization. Thus, the introductory material of Chapters 1 and 2 can be immediately followed by Chapters 7 and 8. This would be enough material for a

one semester undergraduate course, while the first half of a two semester graduate course could continue with Chapter 6 on Vlasov theory, followed in the second semester by Chapters 3 to 5 on kinetic theory and then by Chapters 9 and 10.

It is a pleasure to acknowledge the help of many individuals in writing this book. My views on plasma physics have been shaped over the years by dozens of plasma physicists, especially Allan N. Kaufman and Martin V. Goldman. The students in graduate plasma physics courses at the University of Colorado and the University of Iowa have contributed many useful suggestions (Sun Guo-Zheng deserves special mention). The manuscript was professionally typed and edited by Alice Conwell Shank, Gail Maxwell, Susan D. Imhoff, and Janet R. Kephart. The figures were skillfully drafted by John R. Birkbeck, Jr. and Jeana K. Wonderlich. The preparation of this book was supported by the University of Colorado, the University of Iowa, the United States Department of Energy, the United States National Aeronautics and Space Administration, and the United States National Science Foundation.

Dwight R. Nicholson

Contents

CHAPTER 1

Introduction

1.1 INTRODUCTION

A *plasma* is a gas of charged particles, in which the potential energy of a typical particle due to its nearest neighbor is much smaller than its kinetic energy. The *plasma state* is the fourth state of matter: heating a solid makes a liquid, heating a liquid makes a gas, heating a gas makes a plasma. (Compare the ancient Greeks' earth, water, air, and fire.) The word plasma comes from the Greek plásma, meaning "something formed or molded." It was introduced to describe ionized gases by Tonks and Langmuir [1]. More than 99% of the known universe is in the plasma state. (Note that our definition excludes certain configurations such as the electron gas in a metal and so-called "strongly coupled" plasmas which are found, for example, near the surface of the sun. These need to be treated by techniques other than those found in this book.)

In this book, we shall always consider plasma having roughly equal numbers of singly charged ions ($+e$) and electrons ($-e$), each with average density n_0 (particles per cubic centimeter). In nature many plasmas have more than two species of charged particles, and many ions have more than one electron missing. It is easy to generalize the results of this book to such plasmas.

EXERCISE Name a well-known proposed source of energy that involves plasma with more than one species of ion.

1.2 DEBYE SHIELDING

In a plasma we have many charged particles flying around at high speeds. Consider a special test particle of charge $q_T > 0$ and infinite mass, located at the origin of a

three-dimensional coordinate system containing an infinite, uniform plasma. The test charge repels all other ions, and attracts all electrons. Thus, around our test charge the electron density n_e increases and the ion density decreases. The test ion gathers a *shielding cloud* that tends to cancel its own charge (Fig. 1.1).

Consider Poisson's equation relating the electric potential φ to the charge density ρ due to electrons, ions, and test charge,

$$\nabla^2\varphi = -4\pi\rho = 4\pi e(n_e - n_i) - 4\pi q_T\,\delta(\mathbf{r}) \tag{1.1}$$

where $\delta(\mathbf{r}) \equiv \delta(x)\delta(y)\delta(z)$ is the product of three Dirac delta functions. After the introduction of the test charge, we wait for a long enough time that the electrons with temperature T_e have come to thermal equilibrium with themselves, and the ions with temperature T_i have come to thermal equilibrium with themselves, but not so long that the electrons and ions have come to thermal equilibrium with each other at the same temperature (see Section 1.6). Then equilibrium statistical mechanics predicts that

$$n_e = n_0 \exp\left(\frac{e\varphi}{T_e}\right), \qquad n_i = n_0 \exp\left(\frac{-e\varphi}{T_i}\right) \tag{1.2}$$

where each density becomes n_0 at large distances from the test charge where the potential vanishes. Boltzmann's constant is absorbed into the temperatures T_e and T_i, which have units of energy and are measured in units of electron-volts (eV).

Assuming that $e\varphi/T_e << 1$ and $e\varphi/T_i << 1$, we expand the exponents in (1.2) and write (1.1) away from $\mathbf{r} = 0$ as

$$\nabla^2\varphi = \frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{d\varphi}{dr}\right) = 4\pi n_0 e^2\left(\frac{1}{T_e} + \frac{1}{T_i}\right)\varphi \tag{1.3}$$

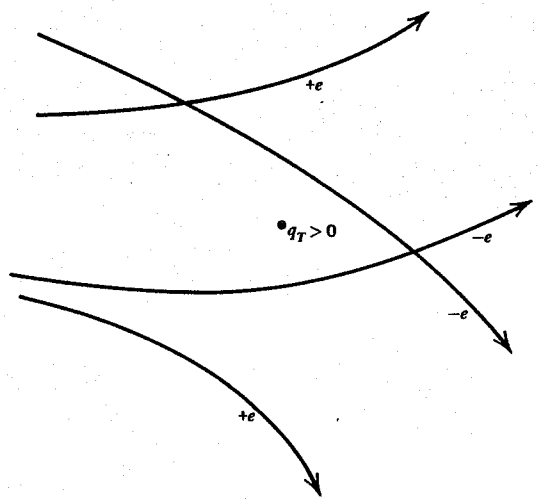

Fig. 1.1 A test charge in a plasma attracts particles of opposite sign and repels particles of like sign, thus forming a shielding cloud that tends to cancel its charge.

If we define the electron and ion *Debye lengths*

$$\lambda_{e,i} \equiv \left(\frac{T_{e,i}}{4\pi n_0 e^2}\right)^{1/2} \tag{1.4}$$

and the *total Debye length*

$$\lambda_D^{-2} = \lambda_e^{-2} + \lambda_i^{-2} \tag{1.5}$$

Eq. (1.3) then becomes

$$\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{d\varphi}{dr}\right) = \lambda_D^{-2}\varphi \tag{1.6}$$

Trying a solution of the form $\varphi = \tilde{\varphi}/r$, we find

$$\frac{d^2\tilde{\varphi}}{dr^2} = \lambda_D^{-2}\tilde{\varphi} \tag{1.7}$$

The solution that falls off properly at large distances is $\tilde{\varphi} \propto \exp(-r/\lambda_D)$. From elementary electricity and magnetism we know that the solution to (1.1) at locations very close to $\mathbf{r} = 0$ is $\varphi = q_T/r$; thus, the desired solution to (1.1) at all distances is

$$\varphi = \frac{q_T}{r}\exp\left(\frac{-r}{\lambda_D}\right) \tag{1.8}$$

The potential due to a test charge in a plasma falls off much faster than in vacuum. This phenomenon is known as *Debye shielding*, and is our first example of plasma *collective behavior*. For distances $r >>$ the Debye length λ_D, the shielding cloud effectively cancels the test charge q_T. Numerically, the Debye length of species s with temperature T_s is roughly $\lambda_s \approx 740[T_s(\text{eV})/n(\text{cm}^{-3})]^{1/2}$ in units of cm.

EXERCISE Prove that the net charge in the shielding cloud exactly cancels the test charge q_T.

It is not necessary that q_T be a special particle. In fact, each particle in a plasma tries to gather its own shielding cloud. However, since the particles are moving, they are not completely successful. In an equal temperature plasma ($T_e = T_i$), a typical slowly moving ion has the full electron component of its shielding cloud and a part of the ion component, while a typical rapidly moving electron has a part of the electron component of its shielding cloud and almost none of the ion component.

1.3 PLASMA PARAMETER

In a plasma where each species has density n_0, the distance between a particle and its nearest neighbor is roughly $n_0^{-1/3}$. The average potential energy Φ of a particle due to its nearest neighbor is, in absolute value,

$$|\Phi| \sim \frac{e^2}{r} \sim n_0^{1/3}\,e^2 \tag{1.9}$$

Our definition of a plasma requires that this potential energy be much less than the typical particle's kinetic energy

$$\frac{1}{2} m_s\langle v^2\rangle = \frac{3}{2} T_s = \frac{3}{2} m_s v_s^2 \tag{1.10}$$

where m_s is the mass of species s, $\langle\ \rangle$ means an average over all particle velocities at a given point in space, and we have defined the *thermal speed* v_s of species s by

$$v_s \equiv \left(\frac{T_s}{m_s}\right)^{1/2} \tag{1.11}$$

For electrons, $v_e \approx 4 \times 10^7\ T_e^{1/2}$ (eV) in units of cm/s. Our definition of a plasma requires

$$n_0^{1/3}\, e^2 << T_s \tag{1.12}$$

or

$$n_0^{2/3} \left(\frac{T_s}{n_0\, e^2}\right) >> 1 \tag{1.13}$$

Raising each side of (1.13) to the 3/2 power, and recalling the definition (1.4) of the Debye length, we have (dropping factors of 4π, etc.)

$$\boxed{\Lambda_s \equiv n_0 \lambda_s^3 >> 1} \tag{1.14}$$

where Λ_s is called the *plasma parameter of species s*. (*Note:* Some authors call Λ_s^{-1} the plasma parameter.) The plasma parameter is just the number of particles of species s in a box each side of which has length the Debye length (a Debye cube). Equation (1.14) tells us that, by definition, a plasma is an ionized gas that has many particles in a Debye cube. Numerically, $\Lambda_s \approx 4 \times 10^8\, T_s^{3/2}(\text{eV})/n_0^{1/2}(\text{cm}^{-3})$. We will often substitute the total Debye length λ_D in (1.14), and define the result $\Lambda \equiv n_0 \lambda_D^3$ to be the *plasma parameter*.

EXERCISE Evaluate the electron thermal speed, electron Debye length, and electron plasma parameter for the following plasmas.

(a) A tokamak or mirror machine with $T_e \approx 1$ keV, $n_0 \approx 10^{13}$ cm^{-3}.
(b) The solar wind near the earth with $T_e \approx 10$ eV, $n_0 \approx 10$ cm^{-3}.
(c) The ionosphere at 300 km above the earth's surface with $T_e \approx 0.1$ eV, $n_0 \approx 10^6$ cm^{-3}.
(d) A laser fusion, electron beam fusion, or ion beam fusion plasma with $T_e \approx 1$ keV, $n_0 \approx 10^{20}$ cm^{-3}.
(e) The sun's center with $T_e \approx 1$ keV, $n_0 \approx 10^{23}$ cm^{-3}.

It is fairly easy to see why many ionized gases found in nature are indeed plasmas. If the potential energy of a particle due to its nearest neighbor were greater than its kinetic energy, then there would be a strong tendency for electrons and ions to bind together into atoms, thus destroying the plasma. The need to keep ions and electrons from forming bound states means that most plasmas have temperatures in excess of one electron-volt.

EXERCISE The temperature of intergalactic plasma is currently unknown, but it could well be much lower than 1 eV. How could the plasma maintain itself at such a low temperature? (*Hint:* $n_0 \approx 10^{-5}$ cm^{-3}).

Of course, it is possible to find situations where a plasma exists jointly with another state. For example, in the lower ionosphere there are regions where 99% of the atoms are neutral and only 1% are ionized. In this *partially ionized plasma*, the ionized component can be a legitimate plasma according to (1.14), where Λ_s should be calculated using only the parameters of the ionized component. Typically, there will be a continuous exchange of particles between the unionized gas and the ionized plasma, through the processes of atomic recombination and ionization.

We can now evaluate the validity of the assumption made before (1.3), that $e\varphi/T_s << 1$. This assumption is most severe for the nearest neighbor to the test charge (which we now take to have charge $q_T = +e$). Using the unshielded form of the potential, we require

$$\frac{e}{T_s}\left(\frac{e}{r}\right) \approx \frac{e}{T_s}\left(\frac{e}{n_0^{-1/3}}\right) << 1 \tag{1.15}$$

or

$$n_0^{1/3}\, e^2 << T_s \tag{1.16}$$

which is just the condition (1.12) required by the definition of a plasma. Thus, our derivation of Debye shielding is correct for any ionized gas that is indeed a plasma.

1.4 PLASMA FREQUENCY

Consider a hypothetical slab of plasma of thickness L, where for the present we consider the ions to have infinite mass, but equal density n_0 and opposite charge to the electrons while the electrons are held rigidly in place with respect to each other, but can move freely through the ions. Suppose the electron slab is displaced a distance δ to the right of the ion slab and then allowed to move freely (Fig. 1.2). What happens?

An electric field will be set up, causing the electron slab to be pulled back toward the ions. When the electrons exactly overlap the ions, the net force is zero, but the electron slab has substantial speed to the left. Thus, the electron slab overshoots, and the net result is harmonic oscillation. The frequency of the oscillation is called the *electron plasma frequency*. It depends only on the electron density, the electron charge, and the electron mass. Let's calculate it.

Poisson's equation in one dimension is ($\partial_x \equiv \partial/\partial x$)

$$\partial_x E = 4\pi\rho \tag{1.17}$$

where E is the electric field. Referring to Fig. 1.3, we take the boundary condition $E(x = 0) = 0$, and assume throughout that $\delta << L$. From (1.17) the electric field over most of the slab is $4\pi n_0 e\delta$, and the force per unit area on the electron slab is (electric field) $\times$ (charge per unit area) or $-4\pi n_0^2 e^2 \delta L$. Newton's second law is

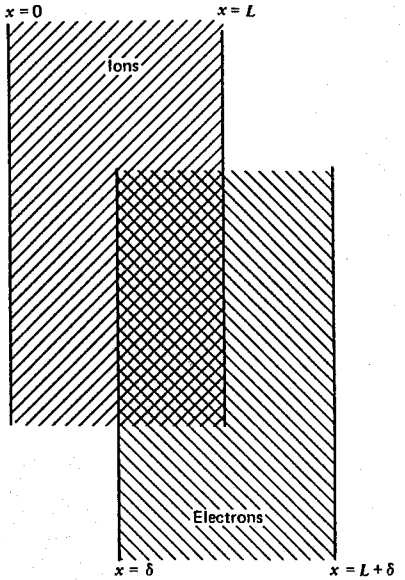

Fig. 1.2 Plasma slab model used to calculate the plasma frequency.

(force per unit area) = (mass per unit area) × (acceleration), or

$$(-4\pi n_0^2 e^2 \delta L) = (n_0 m_e L)(\ddot{\delta}) \tag{1.18}$$

where an overdot is a time derivative. Equation (1.18) is in the standard form of a harmonic oscillator equation,

$$\ddot{\delta} + \left(\frac{4\pi n_0 e^2}{m_e}\right)\delta = 0 \tag{1.19}$$

with characteristic frequency

$$\boxed{\omega_e \equiv \left(\frac{4\pi n_0 e^2}{m_e}\right)^{1/2}} \tag{1.20}$$

which is called the *electron plasma frequency*. Numerically, $\omega_e = 2\pi \times 9000\, n_e^{1/2}$ (cm^{-3}) in units of s^{-1}.

EXERCISE Calculate the electron plasma frequency ω_e and $\omega_e/2\pi$ (e.g., in MHz and kHz) for the five plasmas in the exercise below (1.14).

By analogy with the electron plasma frequency (1.20) we define the ion plasma frequency ω_i for a general ion species with density n_i and ion charge Ze as

$$\omega_i \equiv \left(\frac{4\pi n_i Z^2 e^2}{m_i}\right)^{1/2} \tag{1.21}$$

The total plasma frequency ω_p for a two-component plasma is defined as

$$\omega_p^2 \equiv \omega_e^2 + \omega_i^2 \tag{1.22}$$

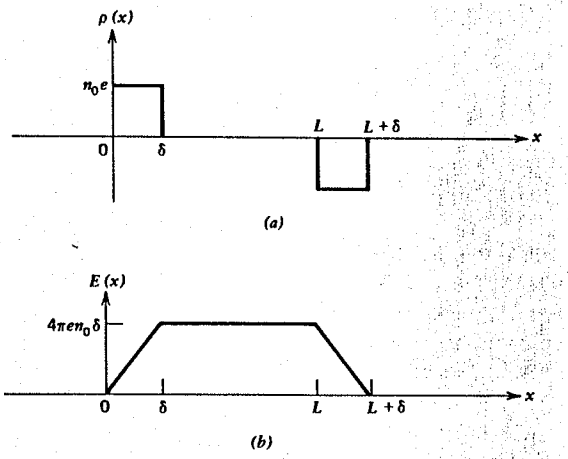

Fig. 1.3 Calculation of the electron plasma frequency. (*a*) Charge density. (*b*) Electric field.

(See Problem 1.3.) For most plasmas in nature $\omega_e >> \omega_i$, so $\omega_p^2 \approx \omega_e^2$. We will see in a later chapter that the general response of an unmagnetized plasma to a perturbation in the electron density is a set of oscillations with frequencies very close to the electron plasma frequency ω_e.

The relation among the Debye length λ_s, the plasma frequency ω_s, and the thermal speed v_s, for the species s, is

$$\boxed{\lambda_s = v_s/\omega_s} \tag{1.23}$$

EXERCISE Demonstrate (1.23).

1.5 OTHER PARAMETERS

Many of the plasmas in nature and in the laboratory occur in the presence of magnetic fields. Thus, it is important to consider the motion of an individual charged particle in a magnetic field. The Lorentz force equation for a particle of charge q_s and mass m_s moving in a constant magnetic field $\mathbf{B} = B_0\hat{z}$ is

$$m_s\ddot{\mathbf{r}} = \frac{q_s}{c}(\dot{\mathbf{r}} \times B_0\hat{z}) \tag{1.24}$$

For initial conditions $\mathbf{r}(t = 0) = (x_0, y_0, z_0)$ and $\dot{\mathbf{r}}(t = 0) = (0, v_\perp, v_z)$ the solution of (1.24) is

$$\begin{aligned} x(t) &= x_0 + \frac{v_\perp}{\Omega_s}(1 - \cos\Omega_s t) \\ y(t) &= y_0 + \frac{v_\perp}{\Omega_s}\sin\Omega_s t \\ z(t) &= z_0 + v_z t \end{aligned} \tag{1.25}$$

where we have defined the *gyrofrequency*

$$\boxed{\Omega_s \equiv \frac{q_s B_0}{m_s c}} \tag{1.26}$$

EXERCISE Verify that (1.25) is the solution of (1.24) with the desired initial conditions.

Numerically, $\Omega_e = -2 \times 10^7 B_0$ (gauss, abbreviated G) in units of s^{-1}, and $\Omega_i = 10^4 B_0$ (gauss) in units of s^{-1} if the ions are protons.

The nature of the motion (1.25) is a constant velocity in the $\hat{z}$-direction, and a circular gyration in the x-y plane with angular frequency $|\Omega_s|$ and center at the *guiding center* position $\mathbf{r}_{gc}$ given by

$$\mathbf{r}_{gc} = (x_0 + v_\perp/\Omega_s, y_0, z_0 + v_z t) \tag{1.27}$$

The radius of the circle in the x-y plane is the *gyroradius* $v_\perp/|\Omega_s|$. The *mean gyroradius* r_s of species s is defined by setting $v_\perp$ equal to the thermal speed, so

$$\boxed{r_s \equiv v_s/|\Omega_s|} \tag{1.28}$$

EXERCISE In the exercise below (1.14), calculate and order the frequencies ω_e, ω_i, $|\Omega_e|$, Ω_i; also calculate the gyroradii r_e and r_i; take $T_i = T_e$ and use the following parameters.

(a) Protons, $B_0 = 10$ kG.
(b) Protons, $B_0 = 10^{-5}$ G.
(c) O^+ ions, $B_0 = 0.5$ G.
(d) Deuterons, $B_0 = 0$ and $B_0 = 10^6$ G.
(e) Protons, $B_0 = 100$ G.

At this point, let us briefly me .ion relativistic and quantum effects. For simplicity, we shall always treat nonrelativistic plasmas. In principle, there is no difficulty in generalizing any of the results of this course to include special relativistic effects; these are discussed at length in the book by Clemmow and Dougherty [2].

EXERCISE To what regime of electron temperature are we limited by the nonrelativistic assumption? How about ion temperature if the ions are protons?

There are, of course, many plasmas in which special relativistic effects do become important. For example, cosmic rays may be thought of as a component of the interstellar and intergalactic plasma with relativistic temperature.

We shall also neglect quantum mechanical effects. For most of the laboratory and astrophysical plasmas in which we might be interested, this is a good assumption. There are, of course, plasmas in which quantum effects are very important. An example would be solid state plasmas. As a rough criterion for the neglect of quantum effects, one might require that the typical de Broglie length $h/m_s v_s$ be much less than the average distance between particles $n_0^{-1/3}$.

EXERCISE What is the maximum density allowed by this criterion for electrons with temperature

(a) 10 eV?
(b) 1 keV?
(c) 100 keV?

In other applications, such as collisions (see next section), one might require the de Broglie length to be much smaller than the distance of closest approach of the colliding particles.

In addition to these assumptions, we shall also neglect the magnetic field in many of the sections of this book. This neglect is made for simplicity, in order that the basic physical phenomena can be elucidated without the complications of a magnetic field. In practice, the magnetic field can usually be ignored when the typical frequency (inverse time scale) of a phenomenon is much larger than the gyrofrequencies of both plasma species.

1.6 COLLISIONS

A typical charged particle in a plasma is at any instant interacting electrostatically (see Problem 1.5) with many other charged particles. If we did not know about Debye shielding, we might think that a typical particle is simultaneously having Coulomb collisions with all of the other particles in the plasma. However, the field of our typical particle is greatly reduced from its vacuum field at distances greater than a Debye length, so that the particle is really not colliding with particles at large distances. Thus, we may roughly think of each particle as undergoing Λ simultaneous Coulomb collisions.

From our definition of a plasma, we know that the potential energy of interaction of each particle with its nearest neighbor is small. Since the potential energy is a measure of the effect of a collision, this means that the strongest one of its Λ simultaneous collisions (the one with its nearest neighbor) is relatively weak. Thus, a typical charged particle in a plasma is simultaneously undergoing Λ weak collisions. We shall soon see that even though Λ is a large number for a plasma, the total effect of all the simultaneous collisions is still weak. Of course, a weak effect can still be a very important effect. In the magnetic bottles like tokamaks and mirror machines currently being used to study controlled thermonuclear fusion plasmas, ion-ion collisions are one of the most important loss mechanisms.

Mathematically, the importance of collisions is contained in an expression called the *collision frequency*, which is the inverse of the time it takes for a particle to suffer a collision. Exactly what is meant by a collision of a charged particle depends upon the definition, and we will consider two different definitions with different physical content. Our mathematical derivation of the collision frequency is an approximate one, intended to be simple but yet to yield the correct results within factors of two or so. A more rigorous development can be found in the book by Spitzer [3]. (See Problem 1.6.)

Consider the situation shown in Fig. 1.4. A particle of charge q, mass m is incident on another particle of charge q_0 and infinite mass with incident speed v_0.

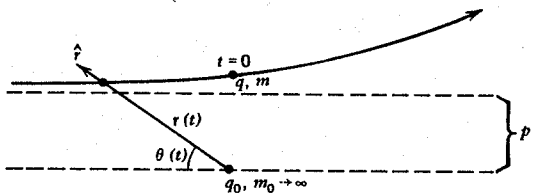

Fig. 1.4 Parameters used in the discussion of the collision frequency in Section 1.6.

If the incident particle were undeflected, it would have position $x = v_0 t$ along the upper dashed line in Fig. 1.4, being at $x = 0$ directly above the scattering charge q_0 at $t = 0$. The separation p of the two dashed lines is the *impact parameter*. If the scattering angle is small, the final parallel speed (parallel to the dashed lines) will be quite close to v_0. The perpendicular speed $v_\perp$ can be obtained by calculating the total perpendicular impulse

$$mv_\perp = \int_{-\infty}^{\infty} dt\, F_\perp(t) \tag{1.29}$$

where $F_\perp$ is the perpendicular force that the particle experiences in its orbit. Since the scattering angle $v_\perp/v_0$ is small, we can to a good approximation use the *unperturbed orbit* $x = v_0 t$ to evaluate the right side of (1.29). This approximation is a very useful one in plasma physics. In Fig. 1.4, Newton's second law with the Coulomb force law is

$$m\ddot{\mathbf{r}} = \frac{qq_0}{r^2}\,\hat{r} \tag{1.30}$$

where $\hat{r}$ is a unit vector in the r-direction. Then

$$F_\perp = \frac{qq_0}{r^2}\sin\theta = \frac{qq_0 \sin\theta}{(p/\sin\theta)^2} = \frac{qq_0}{p^2}\sin^3\theta \tag{1.31}$$

where we have used $p = r\sin\theta$ since the particle is assumed to be traveling along the upper dashed line. Equation (1.29) then reads

$$v_\perp = \frac{qq_0}{mp^2}\int_{-\infty}^{\infty} dt \sin^3\theta(t) \tag{1.32}$$

The relation between θ and t is obtained from

$$x = -r\cos\theta = \frac{-p\cos\theta}{\sin\theta} = v_0 t \tag{1.33}$$

so that

$$dt = \frac{p}{v_0}\,\frac{d\theta}{\sin^2\theta} \tag{1.34}$$

EXERCISE Verify (1.34).

Using (1.34) in (1.32), we find

$$v_\perp = \frac{qq_0}{mv_0 p}\int_0^{\pi} d\theta \sin\theta = \frac{2qq_0}{mv_0 p} \tag{1.35}$$

Defining the quantity

$$p_0 \equiv \frac{2qq_0}{mv_0^2} \tag{1.36}$$

we have

$$\frac{v_\perp}{v_0} = \frac{p_0}{p} \tag{1.37}$$

which is strictly valid only when $v_\perp << v_0, p >> p_0$. In some books, the parameter p_0 is called the *Landau length*.

EXERCISE Show that if $qq_0 > 0$, then p_0 is the distance of closest possible approach for a particle of initial speed v_0.

Although (1.37) is not valid for large angle collisions, let us use it to get a rough idea of the impact parameter p which yields a large angle collision; we do this by setting $v_\perp$ equal to v_0 in (1.37) to obtain $p = p_0$. Thus, any impact parameter $p \leq p_0$ will yield a large angle collision. Suppose the incident particle is an electron, and the (almost) stationary scatterer is an ion. (Although Fig. 1.4 shows a repulsive collision, our development is equally valid for attractive collisions.) The *cross section* for scattering through a large angle by one ion is πp_0^2. Consider an electron that enters a gas of ions. It will have a large angle collision after a time given roughly by setting (the total cross section of the ions in a tube of unit cross-sectional area, and length equal to the distance traveled) equal to (the unit area), or (time) × (velocity) × (number per unit volume) × (cross section) = 1. The inverse of this time gives us the collision frequency ν_L for large angle collisions; thus

$$\nu_L = \pi n_0 v_0 p_0^2 = \frac{4\pi n_0 q^2 q_0^2}{m^2 v_0^3} = \frac{4\pi n_0 e^4}{m_e^2 v_0^3} \tag{1.38}$$

Note that ν_L is proportional to the inverse third power of the particle speed.

Recall that a typical charged particle in a plasma is simultaneously undergoing Λ collisions. Only a very few of these are of the large angle type that lead to (1.38), since a large angle collision involves a potential energy of interaction comparable to the kinetic energy of the incident particle and, by the definition of a plasma, the potential energy of a particle due to its nearest neighbor is small compared to its kinetic energy. Thus, a particle undergoes many more small angle collisions than large angle collisions. It turns out that the cumulative effect of these small angle collisions is substantially larger than the effect of the large angle collisions, as we shall now show.

Unlike the large angle collisions, the many small angle collisions can produce a large effect only after many of them occur. But these small angle collisions produce velocity changes in random directions, some up, some down, some left, some right. We need to know how to measure the cumulative effect of many small random events.

Consider a variable Δx that is the sum of many small random variables Δx_i, $i = 1, 2, \ldots, N$,

$$\Delta x = \Delta x_1 + \Delta x_2 + \ldots + \Delta x_N \tag{1.39}$$

Suppose $\langle \Delta x_i \rangle = 0$ for each i and $\langle (\Delta x_i)^n \rangle$ is the same for each i, where $\langle\ \rangle$ indicates ensemble average [4]. Furthermore, suppose $\langle \Delta x_i\, \Delta x_j \rangle = 0$ if $i \neq j$, so that Δx_i is uncorrelated with Δx_j, $i \neq j$. Then by (1.39) we have $\langle \Delta x \rangle = 0$, and

$$\begin{aligned} \langle (\Delta x)^2 \rangle &= \left\langle \left(\sum_{i=1}^{N} \Delta x_i \right)^2 \right\rangle \\ &= \sum_{i=1}^{N} \langle (\Delta x_i)^2 \rangle \\ &= N \langle (\Delta x_1)^2 \rangle \end{aligned} \tag{1.40}$$

Consider a typical particle moving in the z-direction through a gas of scattering centers. As it moves, it suffers many small angle collisions given by $v_\perp$ which can be decomposed into random variables Δv_x and Δv_y. These latter have just the properties of our random variable Δx_i above. For one collision, with a given impact parameter p (Fig. 1.5), we have from (1.37)

$$\langle v_\perp{}^2 \rangle = \langle (\Delta v_x)^2 \rangle + \langle (\Delta v_y)^2 \rangle = \frac{v_0{}^2 p_0{}^2}{p^2} \tag{1.41}$$

Since Δv_x must have the same statistical properties as Δv_y, we must have

$$\langle (\Delta v_x)^2 \rangle = \langle (\Delta v_y)^2 \rangle = \frac{1}{2} \frac{v_0{}^2 p_0{}^2}{p^2} \tag{1.42}$$

Then by (1.40) we have, for the total x velocity $\Delta v_x{}^{\text{tot}}$,

$$\langle (\Delta v_x{}^{\text{tot}})^2 \rangle = N \langle (\Delta v_x)^2 \rangle = \frac{N}{2} \frac{v_0{}^2 p_0{}^2}{p^2} \tag{1.43}$$

Since we are considering a particle moving through a gas of scattering centers, it is more useful for our purposes to have the time derivative of (1.43), where on the right we shall have $dN/dt = 2\pi p\, dp\, n_0 v_0$ as the number of scattering centers, with impact parameter between p and $p + dp$, which our incident particle encounters per unit time. The time derivative of (1.43) is then

$$\frac{d}{dt} \langle (\Delta v_x{}^{\text{tot}})^2 \rangle = \pi n_0 v_0{}^3 p_0{}^2 \frac{dp}{p} \tag{1.44}$$

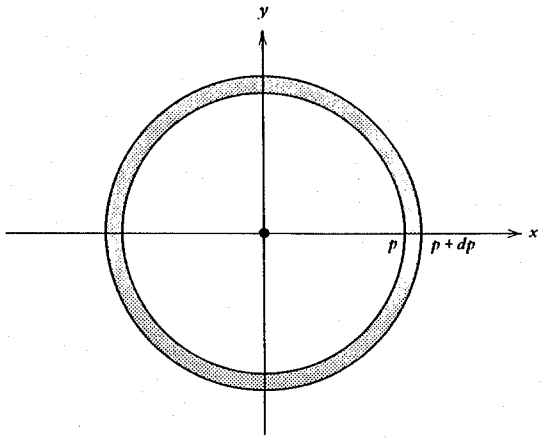

Fig. 1.5 The incident particle is located at the origin and is traveling into the paper. It makes simultaneous small angle collisions with all of the scattering centers randomly distributed with impact parameters between p and $p + dp$.

We have calculated (1.44) for only one set of impact parameters between p and $p + dp$. The same logic that led to (1.40) also allows us to sum (integrate) the right side of (1.44) over all impact parameters to obtain a total change in mean square velocity in the $\hat{x}$-direction. Likewise, we can add the total $\hat{x}$-direction and the total $\hat{y}$-direction mean square velocities to obtain a total mean square perpendicular velocity $\langle (\Delta v_\perp{}^{\text{tot}})^2 \rangle$. With this final factor of two we have

$$\frac{d}{dt} \langle (\Delta v_\perp{}^{\text{tot}})^2 \rangle = 2\pi n_0 v_0{}^3 p_0{}^2 \int_{p_{\min}}^{p_{\max}} \frac{dp}{p} \tag{1.45}$$

What should we use for $p_{\max}$ and $p_{\min}$? Recall that our derivation of the scattering angle $v_\perp / v_0$ in (1.37) uses the Coulomb force law. However, we know from Section 1.2 that the true force law is modified by Debye shielding and is essentially negligible at distances (impact parameters) much greater than a Debye length. Thus, it is consistent with the approximate nature of the present calculation to replace $p_{\max}$ with λ_D. In the case of $p_{\min}$, we use the fact that our scattering formula (1.37) is not valid for impact parameters $p < |p_0|$ to replace $p_{\min}$ by $|p_0|$. Equation (1.45) is then

$$\frac{d}{dt} \langle (\Delta v_\perp{}^{\text{tot}})^2 \rangle = 2\pi n_0 v_0{}^3 p_0{}^2 \ln \left(\frac{\lambda_D}{|p_0|} \right) \tag{1.46}$$

Since the logarithm is such a slowly varying function of its argument, it will suffice to make a very rough evaluation of λ_D / p_0. In the definition of p_0 in (1.36) we take $q = -e$, $q_0 = +e$, $m = m_e$, and for this rough calculation replace v_0 by the electron thermal speed v_e to obtain

$$\frac{\lambda_D}{|p_0|} \approx \frac{\lambda_D m_e v_e{}^2}{2e^2} \approx \frac{m_e \lambda_D{}^3 \omega_e{}^2}{2e^2} \approx 2\pi n_0 \lambda_D{}^3 = 2\pi\Lambda \tag{1.47}$$

where we have ignored the difference between λ_D and λ_e. Dropping the small factor 2π compared to the large plasma parameter Λ, and using the definition (1.36) of p_0, we find that (1.46) becomes

$$\frac{d}{dt} \langle (\Delta v_\perp{}^{\text{tot}})^2 \rangle = \frac{8\pi n_0 e^4}{m_e{}^2 v_0} \ln \Lambda \tag{1.48}$$

A reasonable definition for the scattering time due to small angle collisions is the time it takes $\langle (\Delta v_\perp{}^{\text{tot}})^2 \rangle$ to equal $v_0{}^2$ according to (1.48); the inverse of this time is the *collision frequency* ν_c due to small-angle collisions:

$$\boxed{\nu_c = \frac{8\pi n_0 e^4 \ln \Lambda}{m_e{}^2 v_0{}^3}} \tag{1.49}$$

Note again the inverse cube dependence on the velocity v_0. One important aspect of ν_c is that it is a factor $2 \ln \Lambda$ larger than the collision frequency ν_L for large

angle collisions given by (1.38). This is a substantial factor in a plasma ($\ln \Lambda = 14$ if $\Lambda = 10^6$). Thus, the deflection of a charged particle in a plasma is predominantly due to the many random small angle collisions that it suffers, rather than the rare large angle collisions.

Throughout one's study of plasma physics, it is useful to identify each phenomenon as a collective effect or as a single particle effect. The oscillation of the plasma slab in Section 1.4, characterized by the plasma frequency ω_e, is a collective effect involving many particles acting simultaneously to produce a large electric field. The collisional deflection of a particle, represented by the collision frequency ν_c in (1.49), is a single particle effect caused by many collisions with individual particles that do not act cooperatively.

EXERCISE Is the Debye shielding described in Section 1.2 a collective effect or a single particle effect?

It is instructive to calculate the ratio of ν_c to ω_e, which is, taking a typical speed $v_0 = v_e$ in (1.49),

$$\frac{\nu_c}{\omega_e} \approx \frac{8\pi n_0 e^4 \ln \Lambda}{m_e^2 v_e^3 \omega_e} = \frac{\ln \Lambda}{2\pi n_0 \lambda_e^3} = \frac{\ln \Lambda}{2\pi \Lambda_e} \tag{1.50}$$

By crudely dropping the factor $\ln \Lambda / 2\pi$ and replacing Λ_e by Λ, we have the easily remembered but very approximate expression

$$\boxed{\frac{\nu_c}{\omega_e} \approx \frac{1}{\Lambda}} \tag{1.51}$$

Thus, the collision frequency in a plasma is very much smaller than the plasma frequency. In this respect, single particle effects are less important than collective effects. A wave with frequency near ω_e will oscillate many times before being substantially damped because of collisions.

EXERCISE What is the ratio of the collisional mean free path, for a typical electron, to the electron Debye length?

The collision frequency ν_c that we calculated in (1.49) is the one appropriate to the collisions of electrons with ions, ν_{ei}. The collision frequency ν_{ee} of electrons with electrons could be calculated in the same way, by moving to the center-of-mass frame rather than taking the scattering center to have infinite mass. This procedure would only introduce factors of two or so, so that within such factors we have $\nu_{ee} \approx \nu_{ei}$. Next, consider ion-ion collisions between ions having the same temperature as the electrons that have collision frequency ν_{ee}. Equation (1.49) yields, with m_e replaced by m_i and $v_i = (m_e/m_i)^{1/2} v_e$ instead of v_0, $\nu_{ii} \approx (m_e/m_i)^{1/2} \nu_{ee}$. Finally, consider ions scattered by electrons (or Mack trucks scattered by pedestrians). This calculation in the center-of-mass frame would introduce another factor of $(m_e/m_i)^{1/2}$, so that $\nu_{ie} \approx (m_e/m_i)\nu_{ee}$.

Suppose an electron–proton plasma is prepared in such a way that the electrons and protons have arbitrary velocity distributions, and comparable but not equal temperatures. On the time scale $\nu_{ee}^{-1} \approx \nu_{ei}^{-1} \approx \Lambda\, \omega_e^{-1}$, the electrons will thermalize via electron-electron and electron-ion collisions and obtain a Maxwellian distribution. On a time scale 43 times longer, the ions will thermalize and obtain a Maxwellian at the ion temperature via ion–ion collisions. Finally, on a time scale 43 times longer still, the electrons and ions will come to the same temperature via ion–electron collisions.

This completes our brief introduction to the important basic concepts of plasma physics. In the next chapter, we shall consider the motion of single charged particles in electric and magnetic fields.

REFERENCES

[1] L. Tonks and I. Langmuir, *Phys. Rev.*, *33*, 195 (1929).

[2] P. C. Clemmow and J. P. Dougherty, *Electrodynamics of Particles and Plasmas*, Addison-Wesley, Reading, Mass., 1969.

[3] L. Spitzer, Jr., *Physics of Fully Ionized Gases*, 2nd ed., Wiley-Interscience, New York, 1962.

[4] F. Reif, *Fundamentals of Statistical and Thermal Physics*, McGraw-Hill, New York, 1965.

PROBLEMS

1.1 Debye Shielding

In the discussion of Debye shielding in Section 1.2, suppose that the ions are infinitely massive and thus cannot respond to the introduction of the test charge. How does the answer change?

1.2 Potential Energy (Birdsall's Problem)

A sphere of plasma has equal uniform densities n_0 of electrons and infinitely massive ions. The electrons are moved to the surface of the sphere, which they cover uniformly. What is the potential energy in the system? Sketch the electric field and electric potential as a function of radius. If the electrons initially had temperature T_e, and it is found that the potential energy is equal to the total initial electron kinetic energy, what is the radius of the sphere in terms of the electron Debye length?

1.3 Total Plasma Frequency

In the discussion of the plasma frequency in Section 1.4, suppose the ions are not infinitely massive but have mass m_i. Modify the discussion to show that the slabs oscillate with the total plasma frequency defined in (1.22).

1.4 Plasma in a Gravitational Field

Consider an electron–proton plasma with equal temperatures $T = T_e = T_i$, no magnetic field, and a gravitational acceleration g in the $-\hat{z}$-direction. We desire the

densities $n_e(z)$ and $n_i(z)$, where $z = 0$ can be thought of as the surface of a planet. If the electrons and ions were neutral, their densities would be given by the Boltzmann law $n_{e,i} \propto \exp(-m_{e,i}\, gz/T)$. Then the scale height $T/m_{e,i}\, g$ would be quite different for electrons and ions. However, this would give rise to huge electric fields that would tend to move ions up and electrons down. Taking into account the electric field, use the Boltzmann law and the initial guess that $n_e(z) \approx n_i(z)$, to be checked at the end of the calculation, to find self-consistent electron and ion density distributions.

1.5 Electrostatic Interaction

Show that in nonrelativistic plasma, the Coulomb force between two typical particles is much more important than the magnetic field part of the Lorentz force.

1.6 Collisions

Read Sections 5.1, 5.2, and 5.3 of Spitzer [3] and compare his treatment of collisions to our Section 1.6. Watch out for differences in notation, and explain all apparent differences of factors of two.

CHAPTER 2

Single Particle Motion

2.1 INTRODUCTION

A plasma consists of many charged particles moving in self-consistent electric and magnetic fields. The fields affect the particle orbits, and the particle orbits affect the fields. The general solution of any problem in plasma physics can be quite complicated. In this chapter, we consider the motion of a single charged particle moving in prescribed fields. After studying this part of the problem in isolation, we can proceed in following chapters to include these particle orbits in the self-consistent determination of the fields.

2.2 E × B DRIFTS

Consider a particle with $v_z = 0$ gyrating in a magnetic field $\mathbf{B}_0$ in the $\hat{z}$-direction, with an electric field $\mathbf{E}_0$ in the $-\hat{y}$-direction perpendicular to the magnetic field as in Fig. 2.1. (The symbol $\hat{a}$ always means a unit vector in the a-direction.) The electric field $\mathbf{E}_0$ cannot accelerate the particle indefinitely, because the magnetic field will turn the particle. (The component of electric field E_z, which we ignore here, can accelerate particles indefinitely. In a plasma, the resulting current usually acts to cancel the charge that caused the electric field in the first place. There are, however, important cases where this cancellation is hindered; for example, the earth's aurora, and tokamak runaway electrons.) What does happen? When the charge q_s is positive, the ion is accelerated on the way down. This gives it a larger local gyroradius at the bottom of its orbit than at the top; recall that the gyroradius is $r_s = v_\perp/\Omega_s$. Thus, the motion will be a spiral in the x-y plane as shown in Fig. 2.2, where we have used the symmetry of the situation to draw the upward part of each orbit. We see that the orbit does not connect to itself, but has jumped a

accelerated drift to a force (being careful with signs), plug in the $\mathbf{F} \times \mathbf{B}$ formula (2.8), and compare the result to (2.41).

2.4 Mirror Machines

(a) A mirror machine has mirror ratio 2. A Maxwellian group of electrons is released at the center of the machine. In the absence of collisions, what fraction of these electrons is confined?

(b) Suppose the mirror machine has initially equal densities $n \approx 10^{13}$ cm^{-3} of electrons and protons, each Maxwellian with a temperature 1 keV $\approx 10^7$ °C. The machine is roughly one meter in size in both directions. Recalling our discussion of collisions from Chapter 1, estimate very roughly the time for
   (1) loss of the unconfined electrons;
   (2) loss of the unconfined ions;
   (3) loss of many of the initially confined electrons (due primarily to which kind of collision?); why do not all of the electrons leave?;
   (4) loss of the initially confined ions (due primarily to which kind of collision?).

   For fusion purposes (supposing the protons were replaced by deuterium or tritium) which of these numbers is the most relevant?

2.5 Drift Energy

A particle of mass m and charge q in a uniform magnetic field $\mathbf{B} = B_0\hat{z}$ is set into motion in the $\hat{x}$-direction by an electric field $E(t)\hat{y}$ that varies slowly from zero to a final value E_0. Thus, at the final time the particle has an $\mathbf{E} \times \mathbf{B}$ drift $\mathbf{v}_0$.

(a) Use energy arguments to show that the particle's guiding center must have been displaced a distance v_0/Ω in the direction of the electric field.

(b) Integrate the polarization drift velocity from time zero to time infinity to obtain a displacement. Does the answer agree with (a)?

CHAPTER 3

Plasma Kinetic Theory I: Klimontovich Equation

3.1 INTRODUCTION

In this chapter, we begin a study of the basic equations of plasma physics. The word "kinetic" means "pertaining to motion," so that plasma kinetic theory is the theory of plasma taking into account the motions of all of the particles. This can be done in an exact way, using the *Klimontovich equation* of the present chapter or the *Liouville equation* of the next chapter. However, we are usually not interested in the exact motion of all of the particles in a plasma, but rather in certain average or approximate characteristics. Thus, the greatest usefulness of the exact Klimontovich and Liouville equations is as starting points for the derivation of approximate equations that describe the average properties of a plasma.

In classical plasma physics, we think of the particles as point particles, each with a given charge and mass. Suppose we have a gas consisting of only one particle. This particle has an orbit $\mathbf{X}_1(t)$ in three-dimensional configuration space $\mathbf{x}$. The orbit $\mathbf{X}_1(t)$ is the set of positions $\mathbf{x}$ occupied by the particle at successive times t. Likewise, the particle has an orbit $\mathbf{V}_1(t)$ in three-dimensional velocity space $\mathbf{v}$. We combine three-dimensional configuration space $\mathbf{x}$ and three-dimensional velocity space $\mathbf{v}$ into six-dimensional phase space $(\mathbf{x},\mathbf{v})$. The density of one particle in this phase space is

$$N(\mathbf{x},\mathbf{v},t) = \delta[\mathbf{x} - \mathbf{X}_1(t)]\delta[\mathbf{v} - \mathbf{V}_1(t)] \tag{3.1}$$

where $\delta[\mathbf{x} - \mathbf{X}_1] \equiv \delta(x - X_1)\delta(y - Y_1)\delta(z - Z_1)$, etc. (The properties of the Dirac delta function are reviewed in Ref. [1], p. 29, and in Ref. [2], pp. 53–54.) Note that $\mathbf{X}_1$, $\mathbf{V}_1$ are the Lagrangian coordinates of the particle itself, whereas $\mathbf{x}$, $\mathbf{v}$ are the Eulerian coordinates of the phase space.

EXERCISE At any time t, the density of particles integrated over all phase space must yield the total number of particles in the system. Verify this for the density (3.1).

Next, suppose we have a system with two point particles, with respective orbits $[\mathbf{X}_1(t), \mathbf{V}_1(t)]$ and $[\mathbf{X}_2(t), \mathbf{V}_2(t)]$ in phase space $(\mathbf{x},\mathbf{v})$. By analogy to (3.1), the particle density is

$$N(\mathbf{x},\mathbf{v},t) = \sum_{i=1}^{2} \delta[\mathbf{x} - \mathbf{X}_i(t)]\delta[\mathbf{v} - \mathbf{V}_i(t)] \tag{3.2}$$

EXERCISE Repeat the previous exercise for (3.2).

Now suppose that a system contains two species of particles, electrons and ions, and each species has N_0 particles. Then the density N_s of species s is

$$N_s(\mathbf{x},\mathbf{v},t) = \sum_{i=1}^{N_0} \delta[\mathbf{x} - \mathbf{X}_i(t)]\delta[\mathbf{v} - \mathbf{V}_i(t)] \tag{3.3}$$

and the total density N is

$$N(\mathbf{x},\mathbf{v},t) = \sum_{s,i} N_s(\mathbf{x},\mathbf{v},t) \tag{3.4}$$

EXERCISE Repeat the previous exercise for (3.4).

If we know the exact positions and velocities of the particles at one time, then we know them at all later times. This can be seen as follows. The position $\mathbf{X}_i(t)$ of particle i satisfies the equation

$$\dot{\mathbf{X}}_i(t) = \mathbf{V}_i(t) \tag{3.5}$$

where an overdot means a time derivative. Likewise, the velocity $\mathbf{V}_i(t)$ of particle i satisfies the Lorentz force equation

$$m_s\dot{\mathbf{V}}_i(t) = q_s\mathbf{E}^m[\mathbf{X}_i(t),t] + \frac{q_s}{c}\,\mathbf{V}_i(t) \times \mathbf{B}^m[\mathbf{X}_i(t),t] \tag{3.6}$$

where the superscript m indicates that the electric and magnetic fields are the *m*icroscopic fields self-consistently produced by the point particles themselves, together with externally applied fields. [On the right of (3.6), the portion of $\mathbf{E}^m$ and $\mathbf{B}^m$ produced by particle i itself is deleted.] The microscopic fields satisfy Maxwell's equations

$$\nabla \cdot \mathbf{E}^m(\mathbf{x},t) = 4\pi\rho^m(\mathbf{x},t) \tag{3.7}$$

$$\nabla \cdot \mathbf{B}^m(\mathbf{x},t) = 0 \tag{3.8}$$

$$\nabla \times \mathbf{E}^m(\mathbf{x},t) = -\frac{1}{c}\,\frac{\partial \mathbf{B}^m(\mathbf{x},t)}{\partial t} \tag{3.9}$$

and

$$\nabla \times \mathbf{B}^m(\mathbf{x},t) = \frac{4\pi}{c}\,\mathbf{J}^m(\mathbf{x},t) + \frac{1}{c}\,\frac{\partial \mathbf{E}^m(\mathbf{x},t)}{\partial t} \tag{3.10}$$

The microscopic charge density is

$$\rho^m(\mathbf{x},t) = \sum_{s,i} q_s \int d\mathbf{v}\, N_s(\mathbf{x},\mathbf{v},t) \tag{3.11}$$

while the microscopic current is

$$\mathbf{J}^m(\mathbf{x},t) = \sum_{s,i} q_s \int d\mathbf{v}\mathbf{v} N_s(\mathbf{x},\mathbf{v},t) \tag{3.12}$$

EXERCISE Convince yourself that (3.11) and (3.12) yield the correct charge density and current.

Equations 3.7 to 3.12 determine the exact fields in terms of the exact particle orbits, while (3.5) and (3.6) determine the exact particle orbits in terms of the exact fields. The entire set of equations is closed, so that if the positions and velocities of all particles, and the fields, are known exactly at one time, then they are known exactly at all later times.

3.2 KLIMONTOVICH EQUATION

An exact equation for the evolution of a plasma is obtained by taking the time derivative of the density N_s. From (3.3), this is

$$\frac{\partial N_s(\mathbf{x},\mathbf{v},t)}{\partial t} = -\sum_{i=1}^{N_0} \dot{\mathbf{X}}_i \cdot \nabla_{\mathbf{x}}\, \delta[\mathbf{x} - \mathbf{X}_i(t)]\delta[\mathbf{v} - \mathbf{V}_i(t)]$$
$$- \sum_{i=1}^{N_0} \dot{\mathbf{V}}_i \cdot \nabla_{\mathbf{v}}\, \delta[\mathbf{x} - \mathbf{X}_i(t)]\delta[\mathbf{v} - \mathbf{V}_i(t)] \tag{3.13}$$

where we have used the relations

$$\frac{\partial}{\partial a} f(a - b) = -\frac{\partial}{\partial b} f(a - b)$$

and

$$\frac{d}{dt} f[g(t)] = \frac{df}{dg}\,\dot{g}$$

and where $\nabla_{\mathbf{x}} \equiv (\partial_x,\partial_y,\partial_z)$ and $\nabla_{\mathbf{v}} \equiv (\partial_{v_x},\partial_{v_y},\partial_{v_z})$. Using (3.5) and (3.6), we can write $\dot{\mathbf{X}}_i$ and $\dot{\mathbf{V}}_i$ in terms of $\mathbf{V}_i$ and the fields $\mathbf{E}^m$ and $\mathbf{B}^m$, whereupon (3.13) becomes

$$\frac{\partial N_s(\mathbf{x},\mathbf{v},t)}{\partial t} = -\sum_{i=1}^{N_0} \mathbf{V}_i \cdot \nabla_{\mathbf{x}}\, \delta[\mathbf{x} - \mathbf{X}_i]\delta[\mathbf{v} - \mathbf{V}_i]$$
$$- \sum_{i=1}^{N_0} \left\{ \frac{q_s}{m_s}\,\mathbf{E}^m[\mathbf{X}_i(t),t] + \frac{q_s}{m_s c}\,\mathbf{V}_i \times \mathbf{B}^m[\mathbf{X}_i(t),t] \right\}$$
$$\cdot \nabla_{\mathbf{v}}\, \delta[\mathbf{x} - \mathbf{X}_i]\delta[\mathbf{v} - \mathbf{V}_i] \tag{3.14}$$

An important property of the Dirac delta function is

$$a\delta(a - b) = b\delta(a - b)$$

EXERCISE How would one prove this relation?

This relation allows us to replace $\mathbf{V}_i(t)$ with $\mathbf{v}$, and $\mathbf{X}_i(t)$ with $\mathbf{x}$, on the right of (3.14) (but not in the arguments of the delta functions) so that (3.14) becomes

$$\frac{\partial N_s(\mathbf{x},\mathbf{v},t)}{\partial t} = -\mathbf{v}\cdot\nabla_{\mathbf{x}}\sum_{i=1}^{N_0}\delta[\mathbf{x}-\mathbf{X}_i]\delta[\mathbf{v}-\mathbf{V}_i] - \left[\frac{q_s}{m_s}\mathbf{E}^m(\mathbf{x},t)+\frac{q_s}{m_s c}\mathbf{v}\times\mathbf{B}^m(\mathbf{x},t)\right]\cdot\nabla_{\mathbf{v}}\sum_{i=1}^{N_0}\delta[\mathbf{x}-\mathbf{X}_i]\delta[\mathbf{v}-\mathbf{V}_i] \tag{3.15}$$

But the two summations on the right of (3.15) are just the density (3.3); therefore

$$\boxed{\frac{\partial N_s(\mathbf{x},\mathbf{v},t)}{\partial t}+\mathbf{v}\cdot\nabla_{\mathbf{x}}N_s+\frac{q_s}{m_s}\left(\mathbf{E}^m+\frac{\mathbf{v}}{c}\times\mathbf{B}^m\right)\cdot\nabla_{\mathbf{v}}N_s=0} \tag{3.16}$$

This is the exact *Klimontovich equation* (Klimontovich [3]; Dupree [4]).

The Klimontovich equation, together with Maxwell's equations, constitute an exact description of a plasma. Given the initial positions and velocities of the particles, the initial densities $N_e(\mathbf{x},\mathbf{v},t=0)$ and $N_i(\mathbf{x},\mathbf{v},t=0)$ are given exactly by (3.3). The initial fields are then chosen to be consistent with Maxwell's equations (3.7) to (3.12). With these initial conditions the problem is completely deterministic, and the densities and fields are exactly determined for all time.

In practice, we never carry out this procedure. The Klimontovich equation contains every one of the exact single particle orbits. This is far more information than we want or need. What we really want is information about certain average properties of the plasma. We do not really care about all of the individual electromagnetic fields contributed by the individual charges. What we do care about is the average long-range electric field, which might exist over many thousands or millions of interparticle spacings. The usefulness of the Klimontovich equation comes from its role as a starting point in the derivation of equations that describe the average properties of a plasma.

The Klimontovich equation can be thought of as expressing the incompressibility of the "substance" $N_s(\mathbf{x},\mathbf{v},t)$ as it moves about in the $(\mathbf{x},\mathbf{v})$ phase space. (Is it any wonder that a point particle is incompressible?) This can be seen as follows. Imagine a hypothetical particle with charge q_s, mass m_s, which at time t finds itself at the position $(\mathbf{x},\mathbf{v})$. This hypothetical particle has an orbit in phase space determined by the fields in the system. Imagine taking a time derivative of any quantity along this orbit (such a time derivative is called a *convective derivative*). This derivative must include the time variation produced by the changing position in $(\mathbf{x},\mathbf{v})$ space as well as the explicit time variation of the quantity. Thus, it must be given by

$$\frac{D}{Dt}\equiv\frac{\partial}{\partial t}+\left.\frac{d\mathbf{x}}{dt}\right|_{\text{orbit}}\cdot\nabla_{\mathbf{x}}+\left.\frac{d\mathbf{v}}{dt}\right|_{\text{orbit}}\cdot\nabla_{\mathbf{v}} \tag{3.17}$$

where by $d\mathbf{x}/dt|_{\text{orbit}}$ we mean the change in position $\mathbf{x}$ of the hypothetical particle with time; likewise for $d\mathbf{v}/dt|_{\text{orbit}}$. But for our hypothetical particle at position $(\mathbf{x},\mathbf{v})$ in phase space we know that

$$\left.\frac{d\mathbf{x}}{dt}\right|_{\text{orbit}}=\mathbf{v} \tag{3.18}$$

and

$$\left.\frac{d\mathbf{v}}{dt}\right|_{\text{orbit}}=\frac{q_s}{m_s}\left[\mathbf{E}^m(\mathbf{x},t)+\frac{\mathbf{v}}{c}\times\mathbf{B}^m(\mathbf{x},t)\right] \tag{3.19}$$

Thus,

$$\frac{D}{Dt}=\frac{\partial}{\partial t}+\mathbf{v}\cdot\nabla_{\mathbf{x}}+\frac{q_s}{m_s}\left[\mathbf{E}^m(\mathbf{x},t)+\frac{\mathbf{v}}{c}\times\mathbf{B}^m(\mathbf{x},t)\right]\cdot\nabla_{\mathbf{v}} \tag{3.20}$$

and the Klimontovich equation (3.16) simply says

$$\boxed{\frac{D}{Dt}N_s(\mathbf{x},\mathbf{v},t)=0} \tag{3.21}$$

The density of particles of species s is a constant in time, as measured along the orbit of a hypothetical particle of species s. This is true whether we are moving along the orbit of an actual particle, in which case the density is infinite, or whether we are moving along a hypothetical orbit that is not occupied by an actual particle, in which case the density is zero. Note that the density is only constant as measured along orbits of hypothetical particles; in $(\mathbf{x},\mathbf{v})$ space at a given time it is not constant but is zero or infinite.

There is yet a third way to think of the Klimontovich equation. Any fluid in which the fluid density $f(\mathbf{r},t)$ is neither created nor destroyed satisfies a continuity equation

$$\partial_t f(\mathbf{r},t)+\nabla_{\mathbf{r}}\cdot(f\mathbf{V})=0 \tag{3.22}$$

where $\nabla_{\mathbf{r}}$ is the divergence vector in the phase space under consideration, and $\mathbf{V}$ is a vector that gives the time rate of change of a fluid element at a point in phase space. (See, for example, Symon [5], p. 317.) In the present case, $\nabla_{\mathbf{r}}=(\nabla_{\mathbf{x}},\nabla_{\mathbf{v}})$ and $\mathbf{V}=(d\mathbf{x}/dt|_{\text{orbit}},d\mathbf{v}/dt|_{\text{orbit}})$. Since the particle density is neither created nor destroyed, it must satisfy a continuity equation of the form

$$\partial_t N_s(\mathbf{x},\mathbf{v},t)+\nabla_{\mathbf{x}}\cdot(\mathbf{v}N_s)+\nabla_{\mathbf{v}}\cdot\left\{\frac{q_s}{m_s}\left[\mathbf{E}^m+\frac{\mathbf{v}}{c}\times\mathbf{B}^m\right]N_s\right\}=0 \tag{3.23}$$

It is left as a problem to demonstrate that the continuity equation (3.23) is equivalent to the Klimontovich equation (3.16).

3.3 PLASMA KINETIC EQUATION

Although the Klimontovich equation is exact, we are really not interested in exact solutions of it. These would contain all of the particle orbits, and would thus be far too detailed for any practical purpose. What we really would like to know are the average properties of a plasma. The Klimontovich equation tells us whether or not a particle with infinite density is to be found at a given point $(\mathbf{x},\mathbf{v})$ in phase space.

What we really want to know is how many particles are likely to be found in a small volume $\Delta\mathbf{x}\,\Delta\mathbf{v}$ of phase space whose center is at $(\mathbf{x},\mathbf{v})$. Thus, we really are not interested in the spikey function $N_s(\mathbf{x},\mathbf{v},t)$, but rather in the smooth function

$$f_s(\mathbf{x},\mathbf{v},t) \equiv \langle N_s(\mathbf{x},\mathbf{v},t)\rangle \tag{3.24}$$

The most rigorous way to interpret $\langle\ \rangle$ is as an ensemble average [6] over an infinite number of realizations of the plasma, prepared according to some prescription. For example, we could prepare an ensemble of equal temperature plasmas, each in thermal equilibrium, and each with a test charge q_T at the origin of configuration space. The resulting f_e and f_i would then be consistent with the discussion of Debye shielding in Section 1.2.

There is another useful interpretation of the *distribution function* $f_s(\mathbf{x},\mathbf{v},t)$, the number of particles of species s per unit configuration space per unit velocity space. Suppose we are interested in long range electric and magnetic fields that extend over distances much larger than a Debye length. Then we can imagine a box, centered around the point $\mathbf{x}$ in configuration space, of a size much greater than a mean interparticle spacing, but much smaller than a Debye length (this is easy to do in a plasma; why?) We can now count the number of particles of species s in the box at time t with velocities in the range $\mathbf{v}$ to $\mathbf{v} + \Delta\mathbf{v}$, divide by (the size of the box multiplied by $\Delta v_x\,\Delta v_y\,\Delta v_z$), and call the result $f_s(\mathbf{x},\mathbf{v},t)$. This number will of course fluctuate with time but, if there are very many particles in the box, the fluctuations will be tiny and the $f_s(\mathbf{x},\mathbf{v},t)$ obtained in this manner will agree very well with that obtained in the more rigorous ensemble averaging procedure.

An equation for the time evolution of the distribution function $f_s(\mathbf{x},\mathbf{v},t)$ can be obtained from the Klimontovich equation (3.16) by ensemble averaging. We define δN_s, $\delta\mathbf{E}$, and $\delta\mathbf{B}$ by

$$\begin{aligned} N_s(\mathbf{x},\mathbf{v},t) &= f_s(\mathbf{x},\mathbf{v},t) + \delta N_s(\mathbf{x},\mathbf{v},t) \\ \mathbf{E}^m(\mathbf{x},\mathbf{v},t) &= \mathbf{E}(\mathbf{x},\mathbf{v},t) + \delta\mathbf{E}(\mathbf{x},\mathbf{v},t) \end{aligned} \tag{3.25}$$

and

$$\mathbf{B}^m(\mathbf{x},\mathbf{v},t) = \mathbf{B}(\mathbf{x},\mathbf{v},t) + \delta\mathbf{B}(\mathbf{x},\mathbf{v},t)$$

where $\mathbf{B} \equiv \langle\mathbf{B}^m\rangle$, $\mathbf{E} \equiv \langle\mathbf{E}^m\rangle$, and $\langle\delta N_s\rangle = \langle\delta\mathbf{E}\rangle = \langle\delta\mathbf{B}\rangle = 0$. Inserting these definitions into (3.16) and ensemble averaging, we obtain

$$\boxed{\begin{aligned} &\frac{\partial f_s(\mathbf{x},\mathbf{v},t)}{\partial t} + \mathbf{v}\cdot\nabla_{\mathbf{x}} f_s + \frac{q_s}{m_s}\left(\mathbf{E} + \frac{\mathbf{v}}{c}\times\mathbf{B}\right)\cdot\nabla_{\mathbf{v}} f_s \\ &\qquad = -\frac{q_s}{m_s}\langle(\delta\mathbf{E} + \frac{\mathbf{v}}{c}\times\delta\mathbf{B})\cdot\nabla_{\mathbf{v}}\,\delta N_s\rangle \end{aligned}} \tag{3.26}$$

Equation (3.26) is the exact form of the *plasma kinetic equation*. We shall meet other forms of this equation in the next chapter.

The left side of (3.26) consists only of terms that vary smoothly in $(\mathbf{x},\mathbf{v})$ space. The right side is the ensemble average of the products of very spikey quantities like $\delta\mathbf{E} = \mathbf{E}^m - \langle\mathbf{E}^m\rangle$ and δN_s. Thus, the left side of (3.26) contains terms that are insensitive to the discrete-particle nature of the plasma, while the right side of (3.26) is very sensitive to the discrete-particle nature of the plasma. But the discrete-particle nature of a plasma is what gives rise to collisional effects, so that

the left side of (3.26) contains smoothly varying functions representing collective effects, while the right side represents the collisional effects. We have seen in Section 1.6 that the ratio of the importance of collisional effects to the importance of collective effects is sometimes given by $1/\Lambda$, which is a very small number. We might guess that for many phenomena in a plasma, the right side of (3.26) has a size $1/\Lambda$ compared to each of the terms on the left side; thus the right side can be neglected for the study of such phenomena. This indeed is the case, as shown in the next two chapters.

This important point can be illustrated by a hypothetical exercise. Imagine that we break each electron into an infinite number of pieces, so that $n_0 \to \infty$, $m_e \to 0$, and $e \to 0$, while $n_0 e$ = constant, e/m_e = constant, and v_e = constant.

EXERCISE Show that in this hypothetical exercise, ω_e = constant, λ_e = constant, but $T_e \to 0$ and $\Lambda_e \to \infty$.

Then any volume, no matter how small, would contain an infinite number of point particles, each represented by a delta function with infinitesimal charge. Statistical mechanics tells us that the relative fluctuations in such a plasma would vanish, since the fluctuations in the number of particles N_0 in a certain volume is proportional to the square root of that number. Thus, on the right side of (3.26) we have $\delta N_s \sim N_0^{1/2} \sim \Lambda_e^{1/2}$, and $\delta\mathbf{E}$ and $\delta\mathbf{B}$, which are produced by δN_s behaving like (from Poisson's equation) $\sim e\delta N_s \sim N_0^{-1}N_0^{1/2} \sim N_0^{-1/2} \sim \Lambda_e^{-1/2}$, so that the right side becomes constant. On the left, however, each term becomes infinite as $f_s \to \infty$. Thus, the relative importance of the right side vanishes $\sim N_0^{-1} \sim \Lambda_e^{-1}$, and we have

$$\boxed{\frac{\partial f_s(\mathbf{x},\mathbf{v},t)}{\partial t} + \mathbf{v}\cdot\nabla f_s + \frac{q_s}{m_s}\left(\mathbf{E} + \frac{\mathbf{v}}{c}\times\mathbf{B}\right)\cdot\nabla_{\mathbf{v}} f_s = 0} \tag{3.27}$$

which is the *Vlasov [7] equation* (sometimes referred to as the collisionless Boltzmann equation). This approximate equation, which neglects collisional effects, is often called the most important equation in plasma physics. Its properties will be explored in detail in Chapter 6.

The fields $\mathbf{E}$ and $\mathbf{B}$ of (3.27) are the ensemble averaged fields of (3.25). They must satisfy the ensemble averaged versions of Maxwell's equations (3.7) to (3.12), which are

$$\begin{aligned} \nabla\cdot\mathbf{E}(\mathbf{x},t) &= 4\pi\rho \\ \nabla\cdot\mathbf{B}(\mathbf{x},t) &= 0 \\ \nabla\times\mathbf{E}(\mathbf{x},t) &= -\frac{1}{c}\frac{\partial\mathbf{B}}{\partial t} \\ \nabla\times\mathbf{B}(\mathbf{x},t) &= \frac{4\pi}{c}\mathbf{J} + \frac{1}{c}\frac{\partial\mathbf{E}}{\partial t} \\ \rho(\mathbf{x},t) \equiv \langle\rho^m\rangle &= \sum_{e,i} q_s\int d\mathbf{v}\, f_s(\mathbf{x},\mathbf{v},t) \end{aligned}$$

and

$$\mathbf{J}(\mathbf{x},t) \equiv \langle\mathbf{J}^m\rangle = \sum_{e,i} q_s\int d\mathbf{v}\,\mathbf{v} f_s(\mathbf{x},\mathbf{v},t) \tag{3.28}$$

In the next two chapters we shall approach the plasma kinetic equation (3.26) from another direction, and shall use approximate methods to evaluate the collisional right side. In Chapter 6 we shall take up the study of the Vlasov equation (3.27).

REFERENCES

[1] J. D. Jackson, *Classical Electrodynamics*, 2nd ed., Wiley, New York, 1975.

[2] K. Gottfried, *Quantum Mechanics*, Benjamin, New York, 1966.

[3] Yu. L. Klimontovich, *The Statistical Theory of Non-equilibrium Processes in a Plasma*, M.I.T. Press, Cambridge, Mass., 1967.

[4] T. H. Dupree, *Phys. Fluids*, *6*, 1714 (1963).

[5] K. R. Symon, *Mechanics*, 3rd ed., Addison-Wesley, Reading, Mass., 1971.

[6] F. Reif, *Fundamentals of Statistical and Thermal Physics*, McGraw-Hill, New York, 1965.

[7] A. A. Vlasov, *J. Phys. (U.S.S.R.)*, *9*, 25 (1945).

PROBLEM

3.1 Klimontovich as Continuity

Prove that the continuity equation (3.23) is equivalent to the Klimontovich equation (3.16).

CHAPTER 4

Plasma Kinetic Theory II: Liouville Equation

4.1 INTRODUCTION

In addition to the Klimontovich equation, there is another equation, the *Liouville equation*, which provides an exact description of a plasma. Like the Klimontovich equation, the Liouville equation is of no direct use, but provides a starting point for the construction of approximate statistical theories. One of the most useful practical results of this approach is to provide us with an approximate form for the right side of the plasma kinetic equation (3.26), which tells us how the distribution function changes in time due to collisions.

The Klimontovich equation describes the behavior of individual particles. By contrast, the Liouville equation describes the behavior of *systems*. Consider first a "system" consisting of one charged particle. Suppose we measure this particle's position in a coordinate system $\mathbf{x}_1$; then the orbit of the particle $\mathbf{X}_1(t)$ is the set of positions $\mathbf{x}_1$ occupied by the particle at consecutive times t. Likewise, in velocity space we denote the orbit of the particle $\mathbf{V}_1(t)$; this is the set of velocities taken by the particle at consecutive times t; these velocities are measured in a coordinate system $\mathbf{v}_1$. We thus have a phase space $(\mathbf{x}_1,\mathbf{v}_1) = (x_1,y_1,z_1,v_{x_1},v_{y_1},v_{z_1})$. In this six-dimensional phase space there is one "system" consisting of one particle. The *density of systems* in this phase space is

$$N(\mathbf{x}_1,\mathbf{v}_1,t) = \delta[\mathbf{x}_1 - \mathbf{X}_1(t)]\delta[\mathbf{v}_1 - \mathbf{V}_1(t)] \tag{4.1}$$

Next, consider a system of two particles. We introduce a set of coordinate axes for each particle. Particle 1 has $(\mathbf{x}_1,\mathbf{v}_1)$ coordinate axes as before. Particle 2 has $(\mathbf{x}_2,\mathbf{v}_2)$ coordinate axes that lay right on top of the $(\mathbf{x}_1,\mathbf{v}_1)$ coordinate axes. The orbit $\mathbf{X}_1(t)$, $\mathbf{V}_1(t)$ of particle 1 is measured with respect to the $(\mathbf{x}_1,\mathbf{v}_1)$ coordinate axes, while the orbit $\mathbf{X}_2(t)$, $\mathbf{V}_2(t)$ of particle 2 is measured with respect to the $(\mathbf{x}_2,\mathbf{v}_2)$ coordinate

axes. We now introduce an entirely new phase space, having twelve dimensions. The phase space is

$$(\mathbf{x}_1,\mathbf{v}_1,\mathbf{x}_2,\mathbf{v}_2) = (x_1,y_1,z_1,v_{x_1},v_{y_1},v_{z_1},x_2,y_2,z_2,v_{x_2},v_{y_2},v_{z_2}) \tag{4.2}$$

In this twelve-dimensional phase space, there is one system that is occupying the point $[\mathbf{x}_1 = \mathbf{X}_1(t), \mathbf{v}_1 = \mathbf{V}_1(t), \mathbf{x}_2 = \mathbf{X}_2(t), \mathbf{v}_2 = \mathbf{V}_2(t)]$ at time t. The density of systems in this phase space is

$$N(\mathbf{x}_1,\mathbf{v}_1,\mathbf{x}_2,\mathbf{v}_2,t) = \delta[\mathbf{x}_1 - \mathbf{X}_1(t)]\delta[\mathbf{v}_1 - \mathbf{V}_1(t)]\delta[\mathbf{x}_2 - \mathbf{X}_2(t)]\delta[\mathbf{v}_2 - \mathbf{V}_2(t)] \tag{4.3}$$

EXERCISE Show that there is indeed one system in the phase space by integrating the density (4.3) over all phase space.

Note that the density N in (4.3) is completely different from the density N_s used in the previous chapter in the discussion of the Klimontovich equation. The density N_s in Ch. 3 is the density of particles in six-dimensional phase space. The density N in (4.3) is the density of systems (each having two particles) in twelve-dimensional phase space.

Finally, suppose that we have a system of N_0 particles. With each particle i, $i = 1,2,\ldots N_0$, we associate a six-dimensional coordinate system $(\mathbf{x}_i,\mathbf{v}_i)$. Using these $6N_0$ coordinate axes, we construct a $6N_0$-dimensional phase space, analogous to the twelve-dimensional phase space in (4.2). There is one system in $6N_0$-dimensional phase space; therefore the density of systems, by analogy with the density of systems (4.3), is

$$N(\mathbf{x}_1,\mathbf{v}_1,\mathbf{x}_2,\mathbf{v}_2\ldots\mathbf{x}_{N_0},\mathbf{v}_{N_0},t) = \prod_{i=1}^{N_0}\delta[\mathbf{x}_i - \mathbf{X}_i(t)]\delta[\mathbf{v}_i - \mathbf{V}_i(t)] \tag{4.4}$$

where $\Pi_{j=1}^{n} f_j \equiv f_1 f_2 \ldots f_n$.

EXERCISE Use (4.4) to prove that there is one system in all of phase space.

4.2 LIOUVILLE EQUATION

As with the Klimontovich equation in Chapter 3, the Liouville equation is obtained by taking the time derivative of the appropriate density. In this case, we take the time derivative of the density of systems (4.4). Because the density of systems (4.4) is the product of $6N_0$ terms, its time derivative involves the sum of $6N_0$ terms. Using the relation

$$\frac{\partial}{\partial t}\delta[\mathbf{x}_i - \mathbf{X}_i(t)] = -\frac{\partial \mathbf{X}_i}{\partial t}\cdot\nabla_{\mathbf{x}_i}\delta[\mathbf{x}_i - \mathbf{X}_i(t)] \tag{4.5}$$

and similar relations encountered in the previous chapter, the time derivative of (4.4) is

$$\frac{\partial N}{\partial t} + \sum_{i=1}^{N_0}\mathbf{V}_i(t)\cdot\nabla_{\mathbf{x}_i}\prod_{j=1}^{N_0}\delta(\mathbf{x}_j - \mathbf{X}_j)\delta(\mathbf{v}_j - \mathbf{V}_j)$$

$$+ \sum_{i=1}^{N_0}\dot{\mathbf{V}}_i\cdot\nabla_{\mathbf{v}_i}\prod_{j=1}^{N_0}\delta(\mathbf{x}_j - \mathbf{X}_j)\delta(\mathbf{v}_j - \mathbf{V}_j) = 0 \tag{4.6}$$

Using $a\delta(a - b) = b\delta(a - b)$ to replace $\mathbf{V}_i$ by $\mathbf{v}_i$, and similarly for $\dot{\mathbf{V}}_i$ so that for the remainder of this chapter

$$\dot{\mathbf{V}}_i(t) = \frac{q_{s_i}}{m_{s_i}}\left[\mathbf{E}^m(\mathbf{x}_i,t) + \frac{\mathbf{v}_i}{c}\times\mathbf{B}^m(\mathbf{x}_i,t)\right] \tag{4.7}$$

and noting that the products are just the density of systems N, (4.6) becomes

$$\boxed{\frac{\partial N}{\partial t} + \sum_{i=1}^{N_0}\mathbf{v}_i\cdot\nabla_{\mathbf{x}_i}N + \sum_{i=1}^{N_0}\dot{\mathbf{V}}_i(t)\cdot\nabla_{\mathbf{v}_i}N = 0} \tag{4.8}$$

which is the *Liouville equation*. When combined with Maxwell's equations and the Lorentz force equation, the Liouville equation is an exact description of a plasma. For a two-component plasma with $N_0/2$ electrons and $N_0/2$ ions, the expression for $\mathbf{V}_i(t)$ will depend upon whether the ith particle is an electron or a proton. The Liouville equation has all of the advantages and all of the disadvantages of the Klimontovich equation. Because it contains all of the exact six-dimensional orbits of the individual particles in a single system orbit in $6N_0$-dimensional space, it contains far more information than we want or need. Its usefulness is as a starting point in deriving a reduced statistical description, which with appropriate approximations can yield practical information.

Equation (4.8) has the form of a convective time derivative in the $6N_0$-dimensional phase space,

$$\frac{D}{Dt}N(\mathbf{x}_1,\mathbf{v}_1,\mathbf{x}_2,\mathbf{v}_2,\ldots,\mathbf{x}_{N_0},\mathbf{v}_{N_0},t) = 0 \tag{4.9}$$

where

$$\frac{D}{Dt} \equiv \frac{\partial}{\partial t} + \sum_{i=1}^{N_0}\mathbf{v}_i\cdot\nabla_{\mathbf{x}_i} + \sum_{i=1}^{N_0}\dot{\mathbf{V}}_i(t)\cdot\nabla_{\mathbf{v}_i} \tag{4.10}$$

Here, $\dot{\mathbf{V}}_i(t)$ is expressed in terms of the position $(\mathbf{x}_1,\mathbf{v}_1,\mathbf{x}_2,\mathbf{v}_2,\ldots,\mathbf{x}_{N_0},\mathbf{v}_{N_0})$ of the system in $6N_0$-dimensional phase space, since that position determines the positions of the particles in six-dimensional space and thus the fields at all points in six-dimensional space through Maxwell's equations. Thus, the convective time derivative, taken along the system orbit in $6N_0$-dimensional phase space, is zero. The density of systems is incompressible.

The Liouville equation (4.8) can also be put in the form of a continuity equation. Recall the vector identity $\nabla\cdot(a\mathbf{b}) = \mathbf{b}\cdot\nabla a + a\nabla\cdot\mathbf{b}$. Then

$$\mathbf{v}_i\cdot\nabla_{\mathbf{x}_i}N = \nabla_{\mathbf{x}_i}\cdot(\mathbf{v}_i N) \tag{4.11}$$

since $\mathbf{v}_i$ and $\mathbf{x}_i$ are independent variables. Similarly,

$$\dot{\mathbf{V}}_i\cdot\nabla_{\mathbf{v}_i}N = \nabla_{\mathbf{v}_i}\cdot(\dot{\mathbf{V}}_i N) \tag{4.12}$$

since

$$\nabla_{\mathbf{v}_i}\cdot\dot{\mathbf{V}}_i = \nabla_{\mathbf{v}_i}\cdot\left\{\frac{q_{s_i}}{m_{s_i}}\left[\mathbf{E}^m(\mathbf{x}_i,t) + \frac{\mathbf{v}_i}{c}\times\mathbf{B}^m(\mathbf{x}_i,t)\right]\right\} = 0 \tag{4.13}$$

EXERCISE Prove (4.13).

Then the Liouville equation (4.8) becomes

$$\frac{\partial N}{\partial t} + \sum_{i=1}^{N_0} \nabla_{\mathbf{x}_i} \cdot (\mathbf{v}_i N) + \sum_{i=1}^{N_0} \nabla_{\mathbf{v}_i} \cdot (\dot{\mathbf{V}}_i N) = 0 \tag{4.14}$$

In the form of a continuity equation, the Liouville equation expresses the conservation of systems in $6N_0$-dimensional phase space.

As we have introduced it, the Liouville equation describes the exact orbit of a single point in $6N_0$-dimensional phase space. An example is shown in Fig. 4.1, which is a projection of the orbit onto three of the $6N_0$ dimensions. As the individual particles of the system move about in six-dimensional space, the system itself moves along a continuous orbit in $6N_0$-dimensional phase space.

Suppose that we have an ensemble of such systems, prepared at time t_0. At any later time $t \geq t_0$, we define

$$f_{N_0}(\mathbf{x}_1,\mathbf{v}_1,\mathbf{x}_2,\mathbf{v}_2, \ldots, \mathbf{x}_{N_0},\mathbf{v}_{N_0},t)d\mathbf{x}_1 d\mathbf{v}_1 d\mathbf{x}_2 d\mathbf{v}_2 \ldots d\mathbf{x}_{N_0} d\mathbf{v}_{N_0}$$

to be the probability that a particular system is at the point $(\mathbf{x}_1,\mathbf{v}_1, \ldots, \mathbf{x}_{N_0},\mathbf{v}_{N_0})$ in $6N_0$-dimensional phase space, that is, the probability that $\mathbf{X}_1(t)$ lies between $\mathbf{x}_1$ and $\mathbf{x}_1 + d\mathbf{x}_1$, *and* $\mathbf{V}_1(t)$ lies between $\mathbf{v}_1$ and $\mathbf{v}_1 + d\mathbf{v}_1$ *and* $\mathbf{X}_2(t)$ lies between $\mathbf{x}_2$ and $\mathbf{x}_2 + d\mathbf{x}_2$, *and* etc. Since f_{N_0} is a probability density, its integral over all $6N_0$ dimensions must be unity.

Each system in the ensemble moves along an orbit like that shown in Fig. 4.1. We can think of this orbit as carrying its "piece" of probability along with it. A large probability for point A in Fig. 4.1 at time t_0 implies a large probability for point B at time t. In other words, we can think of the probability density as a fluid moving in the $6N_0$-dimensional phase space. Each element in the probability fluid moves along an exact orbit as given by the solution of the Liouville equation (4.8). Since each element of probability fluid moves along a continuous orbit, and since probability is neither created nor destroyed, the probability fluid must satisfy a

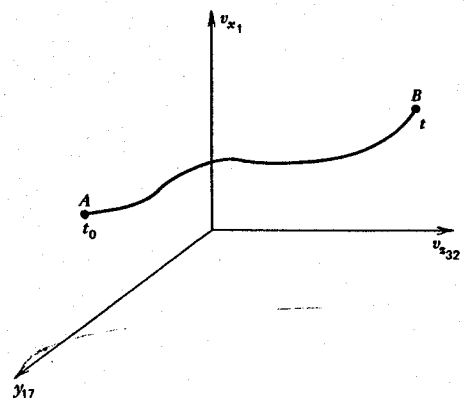

Fig. 4.1 A projection onto three dimensions of a typical system orbit in $6N_0$-dimensional phase space.

continuity equation in $6N_0$-dimensional phase space of the form (4.14). Thus, f_{N_0} must satisfy

$$\frac{\partial f_{N_0}}{\partial t} + \sum_{i=1}^{N_0} \nabla_{\mathbf{x}_i} \cdot (\mathbf{v}_i f_{N_0}) + \sum_{i=1}^{N_0} \nabla_{\mathbf{v}_i} \cdot (\dot{\mathbf{V}}_i f_{N_0}) = 0 \tag{4.15}$$

where $\dot{\mathbf{V}}_i(t)$ is, as usual, calculated from the Lorentz force equation (4.7) and the fields $\mathbf{E}^m$ and $\mathbf{B}^m$ are the exact fields appropriate to the system that occupies this particular point in $6N_0$-dimensional phase space.

We shall only be concerned with smooth functions f_{N_0}. Thus, we might think of a drop of ink placed in a glass of water. The initial drop contains all those systems that have a finite probability of being represented in the ensemble of systems at time t_0. Ignoring diffusion, the drop may lengthen, contract, distort, squeeze, break into pieces, deform, etc., as time progresses. However, the total volume of ink is always constant; the total probability is always unity. The convection of the probability ink is expressed mathematically by reversing the steps that led from the Liouville equation (4.8) to the continuity equation (4.14). (See Problem 4.1.) Equation (4.15) becomes

$$\boxed{\frac{\partial f_{N_0}}{\partial t} + \sum_{i=1}^{N_0} \mathbf{v}_i \cdot \nabla_{\mathbf{x}_i} f_{N_0} + \sum_{i=1}^{N_0} \dot{\mathbf{V}}_i \cdot \nabla_{\mathbf{v}_i} f_{N_0} = 0} \tag{4.16}$$

which by (4.10) is

$$\frac{Df_{N_0}}{Dt} = 0 \tag{4.17}$$

Equation (4.16) is the Liouville equation for the probability density f_{N_0}. Thus, the density of the probability ink is a constant provided that we move with the ink. The probability density f_{N_0} is incompressible.

4.3 BBGKY HIERARCHY

As discussed above, the density f_{N_0} represents the joint probability density that particle 1 has coordinates between $(\mathbf{x}_1,\mathbf{v}_1)$ and $(\mathbf{x}_1 + d\mathbf{x}_1,\mathbf{v}_1 + d\mathbf{v}_1)$ *and* particle 2 has coordinates between $(\mathbf{x}_2,\mathbf{v}_2)$ and $(\mathbf{x}_2 + d\mathbf{x}_2,\mathbf{v}_2 + d\mathbf{v}_2)$, *and* etc. We may also consider reduced probability distributions

$$f_k(\mathbf{x}_1,\mathbf{v}_1,\mathbf{x}_2,\mathbf{v}_2, \ldots, \mathbf{x}_k,\mathbf{v}_k,t) \equiv V^k \int d\mathbf{x}_{k+1} d\mathbf{v}_{k+1} \ldots d\mathbf{x}_{N_0} d\mathbf{v}_{N_0} f_{N_0} \tag{4.18}$$

which give the joint probability of particles 1 through k having the coordinates $(\mathbf{x}_1,\mathbf{v}_1)$ to $(\mathbf{x}_1 + d\mathbf{x}_1,\mathbf{v}_1 + d\mathbf{v}_1)$ *and* . . . *and* $(\mathbf{x}_k,\mathbf{v}_k)$ to $(\mathbf{x}_k + d\mathbf{x}_k,\mathbf{v}_k + d\mathbf{v}_k)$, irrespective of the coordinates of particles $k + 1, k + 2, \ldots, N_0$. The factor V^k on the right of (4.18) is a normalization factor, where V is the finite spatial volume in which f_{N_0} is nonzero for all $\mathbf{x}_1,\mathbf{x}_2, \ldots, \mathbf{x}_{N_0}$ (Fig. 4.2). At the end of our theoretical development, we will take the limit $N_0 \to \infty$, $V \to \infty$, in such a way that $n_0 = N_0/V$ is a constant giving the average number of particles per unit real space. For the present, we assume that $f_{N_0} \to 0$ as $x_i \to \pm\infty$ or $y_i \to \pm\infty$ or $z_i \to \pm\infty$ for any

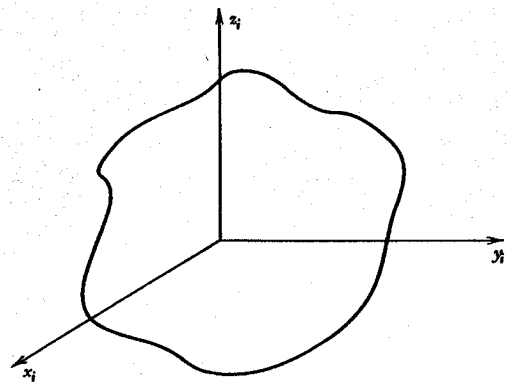

Fig. 4.2 Finite spatial volume V in which f_{N_0} is nonzero for any $\mathbf{x}_i$, $i = 1, \ldots, N_0$.

i. Likewise, because there are no particles with infinite speed, $f_{N_0} \to 0$ as $v_{x_i} \to \pm\infty$ or $v_{y_i} \to \pm\infty$ or $v_{z_i} \to \pm\infty$ for any i.

In this development, we do not care which one of the N_0 particles is called particle number 1, etc. Thus, we always choose probability densities f_{N_0} that are completely symmetric with respect to the particle labels. For example,

$$f_{N_0}(\ldots z_7 = 2 \text{ cm} \ldots z_{13} = 5 \text{ cm} \ldots t) = f_{N_0}(\ldots z_7 = 5 \text{ cm} \ldots z_{13} = 2 \text{ cm} \ldots t) \tag{4.19}$$

provided all of the other independent variables are the same. Here, we must interchange all of the $i = 7$ variables with all of the $i = 13$ variables. This means that when we set $k = 1$ in (4.18), the function $f_1(\mathbf{x}_1, \mathbf{v}_1, t)$ is (to within a normalization constant) the number of particles per unit real space per unit velocity space. Thus, this function $f_1(\mathbf{x}_1, \mathbf{v}_1, t)$ has the same meaning (to within a normalization constant) as the function $f_s(\mathbf{x}, \mathbf{v}, t)$ introduced in the previous chapter in connection with the plasma kinetic equation.

To keep the theory as simple as possible, we shall ignore any external electric and magnetic fields. We shall deal with only one species of N_0 particles; it is easy enough to generalize the results to a plasma with two species of $N_0/2$ particles each at the end of the development. For some purposes, such as calculating electron-electron collisional effects, the second species can be introduced as a smeared-out ion background of density n_0, which simply neutralizes the total electron charge. Finally, we adopt the *Coulomb model*, which ignores the magnetic fields produced by the charged particle motion. In this model, the acceleration

$$\dot{\mathbf{V}}_i(t) = \sum_{j=1}^{N_0} \mathbf{a}_{ij} \tag{4.20}$$

where

$$\mathbf{a}_{ij} = \frac{q_s^2}{m_s |\mathbf{x}_i - \mathbf{x}_j|^3} (\mathbf{x}_i - \mathbf{x}_j) \tag{4.21}$$

is the acceleration of particle i *due to* the Coulomb electric field of particle j. Since a particle exerts no force on itself, we use (4.21) only if $i \neq j$; if $i = j$, we use $\mathbf{a}_{ii} = 0$. Equation (4.21) replaces Maxwell's equations and the Lorentz force law. The Liouville equation (4.16) becomes

$$\frac{\partial f_{N_0}}{\partial t} + \sum_{i=1}^{N_0} \mathbf{v}_i \cdot \nabla_{\mathbf{x}_i} f_{N_0} + \sum_{i=1}^{N_0} \sum_{j=1}^{N_0} \mathbf{a}_{ij} \cdot \nabla_{\mathbf{v}_i} f_{N_0} = 0 \tag{4.22}$$

Equations for the reduced distributions f_k are obtained by integrating the Liouville equation (4.22) over all $\mathbf{x}_{k+1}, \mathbf{v}_{k+1}, \mathbf{x}_{k+2}, \mathbf{v}_{k+2}, \ldots \mathbf{x}_{N_0}, \mathbf{v}_{N_0}$. For example, to obtain the equation for f_{N_0-1} we integrate (4.22) over all $\mathbf{x}_{N_0}$ and $\mathbf{v}_{N_0}$, obtaining

$$\textcircled{1} \quad \int d\mathbf{x}_{N_0} d\mathbf{v}_{N_0} \frac{\partial f_{N_0}}{\partial t} + \textcircled{2} \quad \int d\mathbf{x}_{N_0} d\mathbf{v}_{N_0} \sum_{i=1}^{N_0} \mathbf{v}_i \cdot \nabla_{\mathbf{x}_i} f_{N_0} + \textcircled{3} \quad \int d\mathbf{x}_{N_0} d\mathbf{v}_{N_0} \sum_{i=1}^{N_0} \sum_{j=1}^{N_0} \mathbf{a}_{ij} \cdot \nabla_{\mathbf{v}_i} f_{N_0} = 0 \tag{4.23}$$

Term ① is easy, since we can move the time derivative outside the integral to obtain

$$\textcircled{1} = \frac{\partial}{\partial t} \int d\mathbf{x}_{N_0} d\mathbf{v}_{N_0} f_{N_0} = V^{1-N_0} \frac{\partial}{\partial t} f_{N_0-1} \tag{4.24}$$

where the definition (4.18) has been used. Term ② is also easy. In the first $N_0 - 1$ terms in the sum, the integration variables are independent of the operator $\mathbf{v}_i \cdot \nabla_{\mathbf{x}_i}$; this operator can then be moved outside the integration and we again obtain a term proportional to f_{N_0-1}. The last term in the sum, with $i = N_0$, is

$$\begin{aligned} &\int d\mathbf{x}_{N_0} d\mathbf{v}_{N_0} (v_{x_{N_0}} \partial_{x_{N_0}} + v_{y_{N_0}} \partial_{y_{N_0}} + v_{z_{N_0}} \partial_{z_{N_0}}) f_{N_0} \\ &= \int d\mathbf{v}_{N_0} dy_{N_0} dz_{N_0} v_{x_{N_0}} f_{N_0} \Big|_{x_{N_0}=-\infty}^{x_{N_0}=+\infty} + 2 \text{ similar terms} \\ &= 0 \end{aligned} \tag{4.25}$$

since f_{N_0} vanishes at the boundaries of the system that have been placed at $x_{N_0} = \pm\infty$, etc. Thus,

$$\textcircled{2} = V^{1-N_0} \sum_{i=1}^{N_0-1} \mathbf{v}_i \cdot \nabla_{\mathbf{x}_i} f_{N_0-1} \tag{4.26}$$

Term ③ is not much harder. Splitting the double sum

$$\sum_{i=1}^{N_0} \sum_{j=1}^{N_0} g_{ij} = \sum_{i=1}^{N_0-1} \sum_{j=1}^{N_0-1} g_{ij} + \sum_{j=1}^{N_0-1} g_{N_0 j} + \sum_{i=1}^{N_0-1} g_{iN_0} + g_{N_0 N_0}$$

we get

$$③ = V^{1-N_0} \sum_{i=1}^{N_0-1} \sum_{j=1}^{N_0-1} \mathbf{a}_{ij} \cdot \nabla_{\mathbf{v}_i} f_{N_0-1}$$
$$+ \int d\mathbf{x}_{N_0}\, d\mathbf{v}_{N_0} \sum_{j=1}^{N_0-1} \mathbf{a}_{N_0 j} \cdot \nabla_{\mathbf{v}_{N_0}} f_{N_0}$$
$$+ \int d\mathbf{x}_{N_0}\, d\mathbf{v}_{N_0} \sum_{i=1}^{N_0-1} \mathbf{a}_{iN_0} \cdot \nabla_{\mathbf{v}_i} f_{N_0} \qquad (4.27)$$

where the $i = N_0, j = N_0$ term has been discarded because $\mathbf{a}_{N_0 N_0} = 0$. The second term on the right vanishes after direct integration with respect to $dv_{x_{N_0}}$ and evaluation at $v_{x_{N_0}} = \pm\infty$, etc. The remaining terms in ①, ②, and ③, after multiplication by V^{N_0-1}, are

$$① \qquad ② \qquad ③$$
$$\frac{\partial}{\partial t} f_{N_0-1} + \sum_{i=1}^{N_0-1} \mathbf{v}_i \cdot \nabla_{\mathbf{x}_i} f_{N_0-1} + \sum_{i=1}^{N_0-1} \sum_{j=1}^{N_0-1} \mathbf{a}_{ij} \cdot \nabla_{\mathbf{v}_i} f_{N_0-1}$$
$$④$$
$$+ V^{N_0-1} \sum_{i=1}^{N_0-1} \int d\mathbf{x}_{N_0}\, d\mathbf{v}_{N_0}\, \mathbf{a}_{iN_0} \cdot \nabla_{\mathbf{v}_i} f_{N_0} = 0 \qquad (4.28)$$

This is the desired equation for f_{N_0-1}. Notice that it does not depend only on f_{N_0-1}; the last term ④ depends on f_{N_0}. We have made no approximations in deriving (4.28); within the Coulomb model, it is exact.

Having succeeded in deriving the equation for f_{N_0-1}, let us proceed to derive the equation for f_{N_0-2}. To do this, we integrate (4.28) over all $\mathbf{x}_{N_0-1}$ and over all $\mathbf{v}_{N_0-1}$. As in (4.24), term ① yields $V\, \partial_t f_{N_0-2}$.

EXERCISE Use the definition (4.18) to explain the difference between the power of V encountered here and that encountered in (4.24).

As in (4.26), term ② yields one term that vanishes upon integration, leaving a sum from 1 to $N_0 - 2$. In term ③, we do as in (4.27); we split the double $(N_0 - 1)$ sum into a double $(N_0 - 2)$ sum plus two single $(N_0 - 2)$ sums, the $i = N_0 - 1$, $j = N_0 - 1$ term vanishing since $\mathbf{a}_{N_0-1, N_0-1} = 0$. Term ③ becomes

$$③ = V \sum_{i=1}^{N_0-2} \sum_{j=1}^{N_0-2} \mathbf{a}_{ij} \cdot \nabla_{\mathbf{v}_i} f_{N_0-2}$$
$$+ \int d\mathbf{x}_{N_0-1}\, d\mathbf{v}_{N_0-1} \sum_{i=1}^{N_0-2} \mathbf{a}_{i,N_0-1} \cdot \nabla_{\mathbf{v}_i} f_{N_0-1}$$
$$+ \int d\mathbf{x}_{N_0-1}\, d\mathbf{v}_{N_0-1} \sum_{j=1}^{N_0-2} \mathbf{a}_{N_0-1,j} \cdot \nabla_{\mathbf{v}_{N_0-1}} f_{N_0-1} \qquad (4.29)$$

The last term on the right vanishes upon direct integration with respect to $\mathbf{v}_{N_0-1}$.

For term ④ we have

$$④ = V^{N_0-1} \sum_{i=1}^{N_0-1} \int d\mathbf{x}_{N_0-1}\, d\mathbf{v}_{N_0-1}\, d\mathbf{x}_{N_0}\, d\mathbf{v}_{N_0}\, \mathbf{a}_{iN_0} \cdot \nabla_{\mathbf{v}_i} f_{N_0}$$
$$= V^{N_0-1} \sum_{i=1}^{N_0-2} \int d\mathbf{x}_{N_0}\, d\mathbf{v}_{N_0}\, \mathbf{a}_{iN_0} \cdot \nabla_{\mathbf{v}_i} \int d\mathbf{x}_{N_0-1}\, d\mathbf{v}_{N_0-1} f_{N_0} \qquad (4.30)$$

where the $N_0 - 1$ term in the sum vanishes upon doing the $d\mathbf{v}_{N_0-1}$ integration. The variables $(\mathbf{x}_{N_0}, \mathbf{v}_{N_0})$ and $(\mathbf{x}_{N_0-1}, \mathbf{v}_{N_0-1})$ are simply dummy variables of integration on the far right of (4.30). Therefore, we can switch the labels N_0 and $N_0 - 1$, so that $\mathbf{a}_{iN_0}$ becomes $\mathbf{a}_{i,N_0-1}$. The density f_{N_0} can stay the same, however, because it has the symmetry property (4.19). Equation (4.30) becomes

$$④ = V^{N_0-1} \sum_{i=1}^{N_0-2} \int d\mathbf{x}_{N_0-1}\, d\mathbf{v}_{N_0-1}\, \mathbf{a}_{i,N_0-1} \cdot \nabla_{\mathbf{v}_i} \underbrace{\int d\mathbf{x}_{N_0}\, d\mathbf{v}_{N_0} f_{N_0}}_{V^{1-N_0} f_{N_0-1}}$$
$$= \sum_{i=1}^{N_0-2} \int d\mathbf{x}_{N_0-1}\, d\mathbf{v}_{N_0-1}\, \mathbf{a}_{i,N_0-1} \cdot \nabla_{\mathbf{v}_i} f_{N_0-1} \qquad (4.31)$$

which is identical with the middle term on the right of term ③ in (4.29). Collecting all of the remaining terms in ①, ②, ③, and ④ and dividing by V, we obtain

$$\frac{\partial}{\partial t} f_{N_0-2} + \sum_{i=1}^{N_0-2} \mathbf{v}_i \cdot \nabla_{\mathbf{x}_i} f_{N_0-2} + \sum_{i=1}^{N_0-2} \sum_{j=1}^{N_0-2} \mathbf{a}_{ij} \cdot \nabla_{\mathbf{v}_i} f_{N_0-2}$$
$$+ \frac{2}{V} \sum_{i=1}^{N_0-2} \int d\mathbf{x}_{N_0-1}\, d\mathbf{v}_{N_0-1}\, \mathbf{a}_{i,N_0-1} \cdot \nabla_{\mathbf{v}_i} f_{N_0-1} = 0 \qquad (4.32)$$

This equation for f_{N_0-2} is quite similar in structure to (4.28) for f_{N_0-1}. Notice again that this equation does not involve only f_{N_0-2}, but also involves f_{N_0-1} in the last term on the left.

By comparing (4.28) and (4.32), we see a pattern emerging. Using the same manipulations that we have been using (see Problem 4.2), we can generate an equation similar to (4.28) and (4.32) for arbitrary k. This equation is

$$\boxed{\begin{aligned} &\frac{\partial}{\partial t} f_k + \sum_{i=1}^{k} \mathbf{v}_i \cdot \nabla_{\mathbf{x}_i} f_k + \sum_{i=1}^{k} \sum_{j=1}^{k} \mathbf{a}_{ij} \cdot \nabla_{\mathbf{v}_i} f_k \\ &+ \frac{(N_0 - k)}{V} \sum_{i=1}^{k} \int d\mathbf{x}_{k+1}\, d\mathbf{v}_{k+1}\, \mathbf{a}_{i,k+1} \cdot \nabla_{\mathbf{v}_i} f_{k+1} = 0 \end{aligned}} \qquad (4.33)$$

for $k = 1, 2, \ldots, N_0 - 2$. This is the ***BBGKY hierarchy*** (*B*ogoliubov [1]; *B*orn and *G*reen [2]; *K*irkwood [3, 4]; and *Y*von [5]). Each equation for f_k is coupled to the next higher equation through the f_{k+1} term.

EXERCISE Verify that (4.22) for f_{N_0} and (4.32) for f_{N_0-2} are in agreement with (4.33). Verify that (4.28) for f_{N_0-1} is in agreement with (4.33), provided that f_{N_0} is replaced by $V^{N_0} f_{N_0}$ in (4.33) [see (4.18)].

As it stands, the BBGKY hierarchy (4.33) is still exact (within the Coulomb model) and is just as hard to solve as the original Liouville equation (4.22). It consists of N_0 coupled integro-differential equations. Progress will come only when we take just the first few equations, for $k = 1$, $k = 2$, etc., and then use an approximation to close the set and cut off the dependence on higher equations.

From (4.33) the $k = 1$ equation is

$$\partial_t f_1(\mathbf{x}_1,\mathbf{v}_1,t) + \mathbf{v}_1 \cdot \nabla_{\mathbf{x}_1} f_1 + \frac{N_0 - 1}{V} \int d\mathbf{x}_2\, d\mathbf{v}_2\, \mathbf{a}_{12} \cdot \nabla_{\mathbf{v}_1} f_2(\mathbf{x}_1,\mathbf{v}_1,\mathbf{x}_2,\mathbf{v}_2,t) = 0 \tag{4.34}$$

This is coupled to the $k = 2$ equation through f_2. One way to proceed is to find some approximation for f_2 in terms of f_1. If we can do this, then (4.34) will be written entirely in terms of f_1, and we will have a complete description of the time evolution of $f_1(\mathbf{x}_1,\mathbf{v}_1,t)$ given the initial value $f_1(\mathbf{x}_1,\mathbf{v}_1,t = 0)$.

This is a good point at which to repeat our interpretation of the functions $f_1(\mathbf{x}_1,\mathbf{v}_1,t)$ and $f_2(\mathbf{x}_1,\mathbf{v}_1,\mathbf{x}_2,\mathbf{v}_2,t)$. We have said before that f_1 is equivalent to f_s in the plasma kinetic equation; when multiplied by $n_0 \equiv N_0/V$, it is the ensemble averaged number of particles per unit real space per unit velocity space at the point $(\mathbf{x}_1,\mathbf{v}_1)$ in six-dimensional phase space.

EXERCISE Use the definition to show that $\int d\mathbf{v}_1 f_1(\mathbf{x}_1,\mathbf{v}_1,t) = 1$ provided that none of the functions f_k, $k = 1,2,\ldots,N_0$ depend upon the positions $\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_{N_0}$.

We may also say that $f_1(\mathbf{x}_1,\mathbf{v}_1,t)d\mathbf{x}_1\, d\mathbf{v}_1$ is the probability that a given particle finds itself in the region of phase space between $(\mathbf{x}_1,\mathbf{v}_1)$ and $(\mathbf{x}_1 + d\mathbf{x}_1,\mathbf{v}_1 + d\mathbf{v}_1)$. The interpretation of f_2 is similar to the interpretation of f_1. The function f_2 is the ensemble averaged number of particles per unit $\mathbf{x}_1$ real space per unit $\mathbf{x}_2$ real space per unit $\mathbf{v}_1$ velocity space per unit $\mathbf{v}_2$ velocity space. We may also say that $f_2(\mathbf{x}_1,\mathbf{v}_1,\mathbf{x}_2,\mathbf{v}_2,t)$ is proportional to the joint probability that particle 1 finds itself at $(\mathbf{x}_1,\mathbf{v}_1)$ *and* particle 2 finds itself at $(\mathbf{x}_2,\mathbf{v}_2)$. Since in this discussion all particles are of the same species, we know that an exact expression for f_2 would include the fact that no two particles (electrons, for example) can occupy the same spatial location. Thus, an exact expression for f_2 must have the property that $f_2 \to 0$ as $\mathbf{x}_1 \to \mathbf{x}_2$, regardless of the values of $\mathbf{v}_1$ and $\mathbf{v}_2$. In developing an approximate expression for f_2, we could of course lose this property. Another property that f_2 should have is symmetry with respect to the particle labels: $f_2(\mathbf{x}_1,\mathbf{v}_1,\mathbf{x}_2,\mathbf{v}_2,t) = f_2(\mathbf{x}_2,\mathbf{v}_2,\mathbf{x}_1,\mathbf{v}_1,t)$. This symmetry occurs because the original f_{N_0} has such symmetry, by assumption.

It turns out that f_2 has an intimate relation to f_1, which can be seen by an elementary example from probability theory. Suppose we have two loaded dice, each of which always rolls a five. Then the probability distribution for the value of the throws of either die is

$$P_1(x) = \delta(x - 5) \tag{4.35}$$

The joint probability that the value of the first die will be x and the value of the second die will be y is

$$P_2(x,y) = \delta(x - 5)\delta(y - 5) \tag{4.36}$$

But by (4.35) this is just

$$P_2(x,y) = P_1(x)P_1(y) \tag{4.37}$$

This separation always occurs when two quantities are *statistically independent*; that is, the value of one quantity does not depend on the value of the other quantity. Thus, it is always useful in considering joint probability distributions to factor out the piece that would be there if the two quantities were uncorrelated. Thus, for the dice we have

$$P_2(x,y) = P_1(x)P_1(y) + \delta P(x,y) \tag{4.38}$$

where $\delta P(x,y) = 0$ by (4.37). For a plasma, we define the *correlation function* $g(\mathbf{x}_1,\mathbf{v}_1,\mathbf{x}_2,\mathbf{v}_2,t)$ by

$$f_2(\mathbf{x}_1,\mathbf{v}_1,\mathbf{x}_2,\mathbf{v}_2,t) = f_1(\mathbf{x}_1,\mathbf{v}_1,t)f_1(\mathbf{x}_2,\mathbf{v}_2,t) + g(\mathbf{x}_1,\mathbf{v}_1,\mathbf{x}_2,\mathbf{v}_2,t) \tag{4.39}$$

This is the first step in the *Mayer* [6] *cluster expansion.*

EXERCISE From the definitions of f_{N_0} and f_k, convince yourself that f_2 has the same units as $f_1 f_1$.

We are ready to insert the form (4.39) into the equation (4.34) for f_1, which becomes

$$\partial_t f_1(\mathbf{x}_1,\mathbf{v}_1,t) + \mathbf{v}_1 \cdot \nabla_{\mathbf{x}_1} f_1 + n_0 \int d\mathbf{x}_2\, d\mathbf{v}_2\, \mathbf{a}_{12} \cdot \nabla_{\mathbf{v}_1} [f_1(\mathbf{x}_1,\mathbf{v}_1,t)f_1(\mathbf{x}_2,\mathbf{v}_2,t) + g(\mathbf{x}_1,\mathbf{v}_1,\mathbf{x}_2,\mathbf{v}_2,t)] = 0 \tag{4.40}$$

where we have replaced $(N_0 - 1)/V$ by n_0 because we are interested only in systems with $N_0 >> 1$.

Suppose one assumes that the correlation function vanishes. That is, we assume that the particles in the plasma behave as if they were completely independent of the particular positions and velocities of the other particles. This assumption would be exactly valid if we performed the pulverization procedure discussed in the previous chapter, in which $n_0 \to \infty, e \to 0, m_e \to 0, \Lambda \to \infty; n_0 e =$ constant, $e/m_e =$ constant, $v_e =$ constant, $\omega_e =$ constant, and $\lambda_e =$ constant. Then each particle would have zero charge, and its presence would not affect any other particle. Collective effects could of course still happen, as these involve only f_1 and not g. When we set g equal to zero, (4.40) becomes

$$\partial_t f_1 + \mathbf{v}_1 \cdot \nabla_{\mathbf{x}_1} f_1 + [n_0 \int d\mathbf{x}_2\, d\mathbf{v}_2\, \mathbf{a}_{12} f_1(\mathbf{x}_2,\mathbf{v}_2,t)] \cdot \nabla_{\mathbf{v}_1} f_1(\mathbf{x}_1,\mathbf{v}_1,t) = 0 \tag{4.41}$$

But the quantity in brackets is just the acceleration $\mathbf{a}_{12}$ produced on particle 1 by particle 2, integrated over the probability distribution $f_1(\mathbf{x}_2,\mathbf{v}_2,t)$ of particle 2. This is the ensemble averaged acceleration experienced by particle 1 due to all other particles,

$$\mathbf{a}(\mathbf{x}_1,t) \equiv n_0 \int d\mathbf{x}_2\, d\mathbf{v}_2\, \mathbf{a}_{12} f_1(\mathbf{x}_2,\mathbf{v}_2,t) \tag{4.42}$$

EXERCISE Convince yourself that $\mathbf{a}$ is normalized correctly.

Then (4.41) becomes

$$\boxed{\partial_t f_1 + \mathbf{v}_1 \cdot \nabla_{\mathbf{x}_1} f_1 + \mathbf{a} \cdot \nabla_{\mathbf{v}_1} f_1 = 0} \tag{4.43}$$

which we recognize as our old friend the Vlasov equation.

The Vlasov equation is probably the most useful equation in plasma physics, and a large portion of this book is devoted to its study. For our present purposes, however, it is not enough. It does not include the collisional effects that are represented by the two-particle correlation function g. We would like to have at least an approximate equation that does include collisional effects and that, therefore, predicts the temporal evolution of f_1 due to collisions. We must therefore return to the exact $k = 1$ equation (4.40) and find some method to evaluate g.

Since g is defined through (4.39) as $g = f_2 - f_1 f_1$, we must go back to the $k = 2$ equation in the BBGKY hierarchy in order to obtain an equation for f_2 and, hence, for g. Setting $k = 2$ in (4.33) and using $(N_0 - 2)/V \approx n_0$, one has

$$\underset{①}{\partial_t f_2} + \underset{②}{(\mathbf{v}_1 \cdot \nabla_{\mathbf{x}_1} + \mathbf{v}_2 \cdot \nabla_{\mathbf{x}_2}) f_2} + \underset{③}{(\mathbf{a}_{12} \cdot \nabla_{\mathbf{v}_1} + \mathbf{a}_{21} \cdot \nabla_{\mathbf{v}_2}) f_2}$$

$$+ \underset{④}{n_0 \int d\mathbf{x}_3\, d\mathbf{v}_3 (\mathbf{a}_{13} \cdot \nabla_{\mathbf{v}_1} + \mathbf{a}_{23} \cdot \nabla_{\mathbf{v}_2}) f_3} = 0 \tag{4.44}$$

We have seen that it is useful to factor out the part $f_1 f_1$ of $f_2 = f_1 f_1 + g$, which exists when the particles are uncorrelated. Likewise, it is useful to factor from f_3 the part that would exist when the particles are uncorrelated, plus those parts that result from two-particle correlations. This leads to the next step in the Mayer cluster expansion, which is

$$f_3(123) = f_1(1) f_1(2) f_1(3) + f_1(1) g(23) + f_1(2) g(13) + f_1(3) g(12) + h(123) \tag{4.45}$$

where we have introduced a simplified notation: $(1) \equiv (\mathbf{x}_1, \mathbf{v}_1)$, $(2) \equiv (\mathbf{x}_2, \mathbf{v}_2)$, and $(3) \equiv (\mathbf{x}_3, \mathbf{v}_3)$. Equation (4.45) will be explored further in Problems 4.4 and 4.5.

Our procedure is to insert (4.45) into (4.44) and neglect $h(123)$. This means that we neglect three-particle correlations, or three-body collisions. It turns out that these correlations are of higher order in the plasma parameter Λ; therefore their neglect is quite well justified for many purposes. The resulting set of equations constitute two equations in two unknowns f_1 and g. Thus, we have truncated the BBGKY hierarchy while retaining the effects of collisions to a good approximation.

Inserting (4.45) for f_3 and $f_2 = f_1 f_1 + g$ into the $k = 2$ BBGKY equation (4.44), we find for the numbered terms:

$$① = \underset{ⓐ}{\dot{f}_1(1) f_1(2)} + \underset{ⓒ}{\dot{f}_1(2) f_1(1)} + \dot{g}(12)$$

$$② = \underset{ⓑ}{\mathbf{v}_1 \cdot \nabla_{\mathbf{x}_1} f_1(1) f_1(2)} + \mathbf{v}_1 \cdot \nabla_{\mathbf{x}_1} g(12) + \{1 \leftrightarrow 2\}$$

$$③ = \mathbf{a}_{12} \cdot \nabla_{\mathbf{v}_1} f_1(1) f_1(2) + \mathbf{a}_{12} \cdot \nabla_{\mathbf{v}_1} g(12) + \{1 \leftrightarrow 2\}$$

$$④ = n_0 \int d3\, \mathbf{a}_{13} \cdot \nabla_{\mathbf{v}_1} [\underset{ⓒ}{f_1(1) f_1(2)} f_1(3) + f_1(1) g(23) + \underset{ⓓ}{f_1(2) g(13)} + f_1(3) g(12)] + \{1 \leftrightarrow 2\} \tag{4.46}$$

where $d3 \equiv d\mathbf{x}_3\, d\mathbf{v}_3$ and $\{1 \leftrightarrow 2\}$ means that all of the preceding terms on the right side are repeated with the symbols 1 and 2 interchanged. Recall that $g(12) = g(21)$ by the symmetry of f_2. Many of the terms in (4.46) can be eliminated using the $k = 1$ BBGKY equation (4.40). For example,

$$ⓐ + ⓑ + ⓒ + ⓓ = \{\dot{f}_1(1) + \mathbf{v}_1 \cdot \nabla_{\mathbf{x}_1} f_1(1) + n_0 \int d3\, \mathbf{a}_{13} \cdot \nabla_{\mathbf{v}_1} [f_1(1) f_1(3) + g(13)]\} f_1(2)$$

$$= [\text{left side of (4.40)}] f_1(2) = 0 \tag{4.47}$$

Term ⓔ likewise combines with three of the $\{1 \leftrightarrow 2\}$ terms to vanish, leaving

$$\boxed{\begin{aligned} &\dot{g}(12) + (\mathbf{v}_1 \cdot \nabla_{\mathbf{x}_1} + \mathbf{v}_2 \cdot \nabla_{\mathbf{x}_2}) g(12) = \\ &- (\mathbf{a}_{12} \cdot \nabla_{\mathbf{v}_1} + \mathbf{a}_{21} \cdot \nabla_{\mathbf{v}_2}) [f_1(1) f_1(2) + g(12)] \\ &- \{n_0 \int d3\, \mathbf{a}_{13} \cdot \nabla_{\mathbf{v}_1} [f_1(1) g(23) + f_1(3) g(12)] + \{1 \leftrightarrow 2\}\} \end{aligned}} \tag{4.48}$$

Together with (4.40) which in the condensed notation reads

$$\boxed{\dot{f}_1(1) + \mathbf{v}_1 \cdot \nabla_{\mathbf{x}_1} f_1(1) + n_0 \int d2\, \mathbf{a}_{12} \cdot \nabla_{\mathbf{v}_1} [f_1(1) f_1(2) + g(12)] = 0} \tag{4.49}$$

we have two equations in the two unknowns f_1 and g. We have truncated the BBGKY hierarchy by ignoring three-particle correlations.

In practice, (4.48) and (4.49) are impossibly difficult to solve, either analytically or numerically. They are two coupled nonlinear integro-differential equations in a twelve-dimensional phase space. The present thrust of plasma kinetic theory consists in finding certain approximations to $g(12)$ that are then inserted in (4.49). Using the definition of the acceleration $\mathbf{a}$ in (4.42), we rewrite (4.49) as

$$\boxed{\dot{f}_1(1) + \mathbf{v}_1 \cdot \nabla_{\mathbf{x}_1} f_1 + \mathbf{a} \cdot \nabla_{\mathbf{v}_1} f_1 = - n_0 \int d2\, \mathbf{a}_{12} \cdot \nabla_{\mathbf{v}_1} g(12)} \tag{4.50}$$

which is in exactly the same form as the plasma kinetic equation (3.26).

Most of the discussion in this chapter has been exact, in particular, the derivation of the Liouville equation and the BBGKY hierarchy. Even the approximations that lead to (4.48) and (4.49) are extremely good ones, for example, $1 \ll N_0$ and the neglect of three-particle collisions. By contrast, the approximations needed to convert (4.48) and (4.49) into manageable form are sometimes quite drastic and

less justifiable, as will be seen in the next chapter. Further discussion of the Liouville equation and the BBGKY hierarchy can be found in the books of Montgomery and Tidman [7], Montgomery [8], Clemmow and Dougherty [9], Krall and Trivelpiece [10], and Klimontovich [11].

REFERENCES

[1] N. N. Bogoliubov, *Problems of a Dynamical Theory in Statistical Physics*, State Technical Press, Moscow, 1946.

[2] M. Born and H. S. Green, *A General Kinetic Theory of Liquids*, Cambridge University Press, Cambridge, England, 1949.

[3] J. G. Kirkwood, *J. Chem. Phys.*, *14*, 180 (1946).

[4] J. G. Kirkwood, *J. Chem. Phys.*, *15*, 72 (1947).

[5] J. Yvon, *La Theorie des Fluides et l'Equation d'Etat*, Hermann et Cie, Paris, 1935.

[6] J. E. Mayer and M. G. Mayer, *Statistical Mechanics*, Wiley, New York, 1940.

[7] D. C. Montgomery and D. A. Tidman, *Plasma Kinetic Theory*, McGraw-Hill, New York, 1964.

[8] D. C. Montgomery, *Theory of the Unmagnetized Plasma*, Gordon and Breach, New York, 1971.

[9] P. C. Clemmow and J. P. Dougherty, *Electrodynamics of Particles and Plasmas*, Addison-Wesley, Reading, Mass., 1969.

[10] N. A. Krall and A. W. Trivelpiece, *Principles of Plasma Physics*, McGraw-Hill, New York, 1973.

[11] Yu. L. Klimontovich, *The Statistical Theory of Non-equilibrium Processes in a Plasma*, M.I.T. Press, Cambridge, Mass., 1967.

PROBLEMS

4.1 Continuity vs. Convective

Demonstrate the equivalence between the convective derivative form of the Liouville equation (4.16) and the continuity equation (4.15).

4.2 BBGKY Hierarchy

Integrate (4.32) over all $\mathbf{x}_{N_0-2}$ and $\mathbf{v}_{N_0-2}$ to obtain the $k = N_0 - 3$ equation of the BBGKY hierarchy, and compare your result to (4.33).

4.3 Normalization

Explain in detail the normalization of (4.42).

4.4 Three-Point Correlations (Coins)

In (4.45) we define a three-point joint probability function f_3 in terms of the one-point probability f_1, the two-point correlation function g, and the three-point correlation function h. Suppose we apply this kind of thinking to the case of three coins, each of which can come up heads (+) or tails (−). What is the meaning of f_3 in this case? Write out f_3 in the form (4.45), and evaluate f_3, f_1, g, and h in each of the following cases.

(a) All three coins are "honest," that is, each coin is equally likely to come up heads or tails, and each coin is unaffected by any other coin.

(b) Because the coins are mysteriously locked together, in any one throw all three are heads or tails, the result changing randomly from throw to throw.

(c) All three coins always come up tails.

(d) The first two coins always come up heads, while the third is honest. Note that here the probability functions are not symmetric, so that, for example, $f_1(1)$ is not the same function as $f_1(3)$.

4.5 Three-Point Correlations (Dice)

In (4.45) we define a three-point joint probability function f_3 in terms of the one-point probability f_1, the two-point correlation function g, and the three-point correlation function h. Suppose we apply this kind of thinking to the case of three dice, each of which can take on integer values from one through six. What is the meaning of f_3 in this case? Write out f_3 in the form (4.45), and evaluate f_3, f_1, g, and h in each of the following cases.

(a) All three dice are "honest," that is, the value of each die is equally likely one through six and is independent of the value of any other die.

(b) Because the dice are mysteriously locked together, in one throw all three always show the same value, the value changing randomly from throw to throw with all six values equally likely.

(c) All of the dice always come up "five."

(d) The first two dice always come up "two"; the other one is "honest."

4.6 BBGKY Hierarchy

In this chapter, we derive the BBGKY hierarchy from the Liouville equation. This can be done in a completely different way [10], starting with the Klimontovich equation. Explain, by using words and writing equations only for illustration, how the $k = 1$ and $k = 2$ equations of the BBGKY hierarchy can be obtained from the Klimontovich equation.

CHAPTER 5

Plasma Kinetic Theory III: Lenard–Balescu Equation

5.1 BOGOLIUBOV'S HYPOTHESIS

In the preceding chapter, the BBGKY hierarchy is truncated by neglecting three-particle correlations (three-body collisions). For a good plasma, this is probably a very good approximation, although no rigorous proof exists. The spirit of the approximation is the same as that of Section 1.6, where the collision frequency is calculated as a series of two-body collisions, even though the particle is interacting with Λ particles simultaneously. Since the collision of particle A with particle B is usually a small angle collision, its effect on the orbit of particle A is small, thus making a negligible effect on the simultaneous collision of particle A with particle C.

The result of our truncation of the BBGKY hierarchy is the set of coupled equations (4.48) and (4.50) in the two unknowns $f_1(\mathbf{x}_1,\mathbf{v}_1,t)$ and $g(\mathbf{x}_1,\mathbf{v}_1,\mathbf{x}_2,\mathbf{v}_2,t)$. These equations are quite intractable in general. However, there is one set of simplifying assumptions that is both physically very important and allows the exact (almost) solution of (4.48) and (4.50).

Consider a spatially homogeneous ensemble of plasmas. This means that any function of one spatial variable must be independent of that variable; so $f_1(\mathbf{x}_1,\mathbf{v}_1,t) = f_1(\mathbf{v}_1,t)$ and $\mathbf{a}(\mathbf{x}_1,t) = \mathbf{a}(t) = 0$ by (4.21) and (4.42). Any ensemble averaged function of two spatial variables can only be a function of the difference between those variables; therefore we write $g = g(\mathbf{x}_1 - \mathbf{x}_2,\mathbf{v}_1,\mathbf{v}_2,t)$. With these assumptions, (4.50) simplifies considerably and becomes

$$\boxed{\partial_t f_1(\mathbf{v}_1,t) = -n_0 \int d\mathbf{x}_2\, d\mathbf{v}_2\, \mathbf{a}_{12} \cdot \nabla_{\mathbf{v}_1} g(\mathbf{x}_1 - \mathbf{x}_2,\mathbf{v}_1,\mathbf{v}_2,t)} \tag{5.1}$$

Equation 4.48 simplifies since two terms are of the form

$$[n_0 \int d3\ \mathbf{a}_{13} f_1(3)] \cdot \nabla_{\mathbf{v}_1} g(12) = \mathbf{a} \cdot \nabla_{\mathbf{v}_1} g(12) = 0 \tag{5.2}$$

leaving

$$\begin{aligned} \partial_t g(\mathbf{x}_1 - \mathbf{x}_2,\mathbf{v}_1,\mathbf{v}_2,t) + \mathbf{v}_1 \cdot \nabla_{\mathbf{x}_1} g(12) + \mathbf{v}_2 \cdot \nabla_{\mathbf{x}_2} g(12) \\ + (\mathbf{a}_{12} \cdot \nabla_{\mathbf{v}_1} + \mathbf{a}_{21} \cdot \nabla_{\mathbf{v}_2}) g(12) \\ + n_0 \int d3\ \mathbf{a}_{13} \cdot \nabla_{\mathbf{v}_1} f_1(1) g(23) + n_0 \int d3\ \mathbf{a}_{23} \cdot \nabla_{\mathbf{v}_2} f_1(2) g(13) \\ = -(\mathbf{a}_{12} \cdot \nabla_{\mathbf{v}_1} + \mathbf{a}_{21} \cdot \nabla_{\mathbf{v}_2}) f_1(1)\, f_1(2) \end{aligned} \tag{5.3}$$

We now wish to argue that the fourth term on the left is smaller than all the other terms and can be discarded. Recall the pulverization procedure of the previous chapter. By that argument, as well as the discussion of collisions in Section 1.6, we argue that the two-point correlation function g is higher order in the plasma parameter Λ than f_1; thus $g/f_1 \sim \Lambda^{-1}$. The acceleration $\mathbf{a}_{12} \sim e^2/m_e \sim \Lambda^{-1}$ since e/m_e is constant and $e \sim n_0^{-1} \sim \Lambda^{-1}$; here, we phrase our discussion in terms of electrons. Thus, all terms in (5.3) are $\sim \Lambda^{-1}$ except for the fourth term on the left, which is $\sim \Lambda^{-2}$. We discard this term, leaving

$$\boxed{\frac{\partial g(12)}{\partial t} + V_1 g + V_2 g = S} \tag{5.4}$$

where V_1 and V_2 are operators defined by

$$\begin{aligned} V_1 g(12) &= \mathbf{v}_1 \cdot \nabla_{\mathbf{x}_1} g(12) \\ &+ [n_0 \int d3\mathbf{a}_{13}\, g(23)] \cdot \nabla_{\mathbf{v}_1} f_1(1) \end{aligned} \tag{5.6}$$

$$\begin{aligned} V_2 g(12) &= \mathbf{v}_2 \cdot \nabla_{\mathbf{x}_2} g(12) \\ &+ [n_0 \int d3\mathbf{a}_{23}\, g(13)] \cdot \nabla_{\mathbf{v}_2} f_1(2) \end{aligned} \tag{5.7}$$

and the source function S is

$$S(\mathbf{x}_1 - \mathbf{x}_2,\mathbf{v}_1,\mathbf{v}_2) = -(\mathbf{a}_{12} \cdot \nabla_{\mathbf{v}_1} + \mathbf{a}_{21} \cdot \nabla_{\mathbf{v}_2}) f_1(1)\, f_1(2) \tag{5.8}$$

In this chapter we alternate between the notations (1) and $(\mathbf{x}_1,\mathbf{v}_1)$ depending on convenience. For simplicity, we suppose that we are dealing with an electron plasma. A neutralizing ion background can be thought to be present; it is considered to be smoothed out so that it does not contribute explicitly to the acceleration $\mathbf{a}_{ij}$, which by (4.21) is

$$\mathbf{a}_{ij} = \frac{e^2}{m_e |\mathbf{x}_i - \mathbf{x}_j|^3} (\mathbf{x}_i - \mathbf{x}_j) \tag{5.9}$$

The important physical situation to which this discussion applies is as follows. Imagine a beam of electrons incident on a Maxwellian electron plasma in the

$\hat{x}$-direction. Then the function

$$F(v_x) \equiv \int dv_y\, dv_z\, f_1(\mathbf{v}) \tag{5.10}$$

has the form shown in Fig. 5.1. Ignoring questions of stability (see Chapter 6), we recognize that the beam of electrons represented by the bump at large positive v_x will experience collisions that will eventually ($t \to \infty$) produce a new Maxwellian at a higher temperature. By the discussion of Section 1.6 we can predict the time scale for this process to be $\sim \nu_{ee} \sim \omega_e/\Lambda$. The solution of (5.1) and (5.4) which we are about to obtain should yield a very good theoretical description for this important process. This evolution is encountered in such applications as electron beam-pellet fusion and (when generalized to ions) ohmic heating of tokamaks.

The further assumption that allows us to solve the (still very complicated) set of equations (5.1) and (5.4) is *Bogoliubov's hypothesis*. The assumption is that the two-point correlation function g relaxes on a time scale very short compared to the time scale on which f_1 relaxes [1]. Imagine introducing a test electron into a plasma. The other electrons will adjust to the presence of the test electron in roughly the time it takes for them to have a collision with it. With a typical speed v_e and a typical length λ_e, the time for a collision is $\sim \lambda_e/v_e \sim \omega_e^{-1}$. By contrast, the time for f_1 to change because of collisions is $\sim \Lambda\omega_e^{-1}$; thus it is indeed quite reasonable to assume that g relaxes quickly compared to f_1. Mathematically, we incorporate this assumption by ignoring the time dependence of $f_1(\mathbf{v}_1,t)$ and $f_1(\mathbf{v}_2,t)$ in the source function S on the right of (5.4). Equation (5.4) is then a linear equation for g with a known, constant (in time) source function on the right. We can solve such a linear equation for $g(\mathbf{x}_1 - \mathbf{x}_2,\mathbf{v}_1,\mathbf{v}_2,t \to \infty)$ where $t \to \infty$ is understood to refer to the short time scale on which g relaxes. The solution for g will then depend on the factors $f_1(\mathbf{v}_1,t)$ and $f_1(\mathbf{v}_2,t)$ in the source function (5.8). When this solution for g is substituted into the right side of (5.1), there results a single nonlinear integro-differential equation in the one unknown function f_1. We

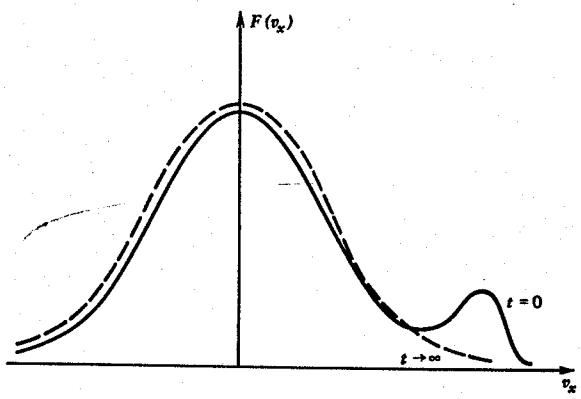

Fig. 5.1 Distribution $F(v_x)$ defined in (5.10) for an electron beam incident on a plasma.

have finally achieved our goal of truncating the BBGKY hierarchy and have expressed the entire plasma kinetic equation (5.1) in terms of the one unknown function $f_1(\mathbf{v}_1,t)$.

The implementation of this procedure is straightforward but complicated. In order to understand it, it is useful to have first studied the material in Chapter 6 on the Vlasov equation. Thus, we will not perform the derivation here; it is included in Appendix A. The reader who is studying plasma physics for the first time may wish to accept the results as given here, and proceed to read Appendix A after a thorough study of Chapter 6.

The solution of (5.1) and (5.4) uses the techniques of Fourier transformation in space, Laplace transformation in time, and their inverses. The conventions used in this book are as follows:

$$f(\mathbf{k}) = \int \frac{d\mathbf{x}}{(2\pi)^3}\, e^{-i\mathbf{k}\cdot\mathbf{x}} f(\mathbf{x}) \tag{5.11}$$

$$f(\mathbf{x}) = \int d\mathbf{k}\, e^{i\mathbf{k}\cdot\mathbf{x}} f(\mathbf{k}) \tag{5.12}$$

$$f(\omega) = \int_0^\infty dt\, e^{i\omega t} f(t) \tag{5.13}$$

$$f(t) = \int_L \frac{d\omega}{2\pi}\, e^{-i\omega t} f(\omega) \tag{5.14}$$

where the integrals over $\mathbf{x}$, $\mathbf{k}$, and t are usually along the real axes while the integral over ω is along the Laplace inversion contour to be discussed later.

Expressed in terms of the difference variable $\mathbf{x} = \mathbf{x}_1 - \mathbf{x}_2$, the acceleration $\mathbf{a}_{12}$ in (5.9) is

$$\mathbf{a}_{12}(\mathbf{x}) = \frac{e^2}{m_e|\mathbf{x}|^3}\,\mathbf{x} \tag{5.15}$$

with Fourier transform

$$\mathbf{a}_{12}(\mathbf{k}) = \frac{-i\mathbf{k}}{m_e}\,\varphi(k) \tag{5.16}$$

where

$$\varphi(k) = \frac{e^2}{2\pi^2 k^2} \tag{5.17}$$

is the Fourier transform of the Coulomb potential

$$\varphi(x) = \frac{e^2}{|\mathbf{x}|} \tag{5.18}$$

(See Problem 5.1.) Then, as shown in Appendix A, the solution of (5.1) and (5.4), under the Bogoliubov hypothesis, is

$$\boxed{\frac{\partial f(\mathbf{v},t)}{\partial t} = -\frac{8\pi^4 n_0}{m_e^2}\,\nabla_{\mathbf{v}}\cdot\int d\mathbf{k}\, d\mathbf{v}'\, \mathbf{k}\mathbf{k}\cdot\frac{\varphi^2(k)}{|\epsilon(\mathbf{k},\mathbf{k}\cdot\mathbf{v})|^2} \times \delta[\mathbf{k}\cdot(\mathbf{v}-\mathbf{v}')][f(\mathbf{v})\nabla_{\mathbf{v}'}f(\mathbf{v}') - f(\mathbf{v}')\nabla_{\mathbf{v}}f(\mathbf{v})]} \tag{5.19}$$

which is the *Lenard–Balescu equation* (Refs. [2] to [6]). In this equation, we have dropped the subscript 1 from $\mathbf{v}_1$, and the subscript 1 from f_1, and have used the *dielectric function*

$$\epsilon(\mathbf{k},\omega) = 1 + \frac{\omega_e^2}{k^2}\int d\mathbf{v}\,\frac{\mathbf{k}\cdot\nabla_{\mathbf{v}} f(\mathbf{v})}{\omega - \mathbf{k}\cdot\mathbf{v}} \tag{5.20}$$

which will be studied in detail in the next chapter. The velocity integral must be performed along the Landau contour, as discussed in the next chapter. The interpretation of the Lenard–Balescu equation (5.19), and several alternate forms, will be discussed in the next section.

5.2 LENARD-BALESCU EQUATION

The Lenard–Balescu equation (5.19) is obtained from the BBGKY hierarchy after several assumptions: three-particle correlations are negligible, the ensemble of plasmas is spatially homogeneous, and the two-particle correlation function g relaxes much faster than the one-particle distribution function f_1. Thus, the Lenard–Balescu equation is applicable to situations such as the collisional relaxation of a beam in a plasma, but is not applicable in general to the collisional damping of spatially inhomogeneous wave motion or any phenomena that involve high frequencies like ω_e.

The right side of (5.19) represents the physics of two-particle collisions, since the right side of (5.1) is proportional to the two-particle correlation function g. This is indicated by the factor $\varphi(k)/\epsilon(\mathbf{k},\mathbf{k}\cdot\mathbf{v})$, which appears squared. It will be shown in the next chapter that the dielectric function $\epsilon(\mathbf{k},\omega)$ represents the plasma shielding of the field of a test charge. Thus, this term in (5.19) represents the interaction of one particle (together with its shielding cloud) with the potential field of another particle (together with its shielding cloud); that is, the collision of two shielded particles.

There is a problem with the Lenard–Balescu equation (5.19) as it stands. If one converts the k integration into spherical coordinates, and takes into account the forms (5.17) of $\varphi(k)$ and (5.20) of $\epsilon(\mathbf{k},\omega)$, one finds that at large k the integral diverges like $\int dk/k \sim \ln k$. Thus, just as in the derivation of the collision frequency in Section 1.6, we find a logarithmic divergence at large k, or small distances. In Section 1.6 we cut off the spatial integral at the lower limit p_0, where p_0 is the impact parameter for large angle collisions. It is argued in Section 1.6 that the physical formulation is not valid for large angle collisions, thus producing an unphysical divergence at short distances. The same thing is going on here. The derivation of the Lenard–Balescu equation is based on the assumption that in the expression

$$f_2(12) = f_1(1)\,f_1(2) + g(12) \tag{5.21}$$

we have $|g| \ll |f_1 f_1|$. This assumption led us to discard a term in (5.3) to obtain (5.4). However, this assumption is not always valid. It is not possible for two electrons to get very close to each other; therefore, we must have $f_2 \to 0$ as $\mathbf{x}_1 \to \mathbf{x}_2$, which implies $g = -f_1 f_1$. Thus, for small values of $|\mathbf{x}_1 - \mathbf{x}_2|$ (large k), it is not correct to assume $|g| \ll |f_1 f_1|$. In practice, since the divergence is

logarithmic, we can simply cut off the integral in (5.19) at some upper limit wave number corresponding to some lower limit spatial scale. For this purpose, the impact parameter (Landau length) p_0 for large angle collisions (see Section 1.6) would be a reasonable choice.

The Lenard–Balescu equation (5.19) has several desirable features [4–5]. These are:

(a) If $f \geq 0$ at $t = 0$, $f \geq 0$ at all t.
(b) Particles are conserved: $d/dt \int d\mathbf{v}\, f(\mathbf{v},t) = 0$.
(c) Momentum is conserved: $d/dt \int d\mathbf{v}\mathbf{v}\, f(\mathbf{v},t) = 0$.
(d) Kinetic energy is conserved: $d/dt \int d\mathbf{v}\, v^2 f(\mathbf{v},t) = 0$.
(e) Any Maxwellian is a time-independent solution.
(f) As $t \to \infty$, any f satisfying (a) approaches a Maxwellian.

A simplified but fairly accurate form of the Lenard–Balescu equation (5.19) can be obtained as follows. We rewrite (5.19) in the form

$$\frac{\partial f(\mathbf{v},t)}{\partial t} = -\nabla_{\mathbf{v}}\cdot\int d\mathbf{v}'\,\overleftrightarrow{\mathbf{Q}}\,(\mathbf{v},\mathbf{v}')\cdot(\nabla_{\mathbf{v}} - \nabla_{\mathbf{v}'})f(\mathbf{v})f(\mathbf{v}') \tag{5.22}$$

with the tensor

$$\begin{aligned}\overleftrightarrow{\mathbf{Q}}\,(\mathbf{v},\mathbf{v}') &= -\frac{8\pi^4 n_0}{m_e^2}\int d\mathbf{k}\,\frac{\mathbf{k}\mathbf{k}\varphi^2(k)}{|\epsilon(\mathbf{k},\mathbf{k}\cdot\mathbf{v})|^2}\,\delta[\mathbf{k}\cdot(\mathbf{v}-\mathbf{v}')] \\ &= -\frac{2n_0 e^4}{m_e^2}\int d\mathbf{k}\,\frac{\mathbf{k}\mathbf{k}}{k^4}\,\frac{\delta[\mathbf{k}\cdot(\mathbf{v}-\mathbf{v}')]}{\left|1 + \dfrac{\psi}{k^2\lambda_e^2}\right|^2}\end{aligned} \tag{5.23}$$

where the definition (5.17) has been used, and where the dimensionless function ψ is found from (5.20) to be

$$\psi(\mathbf{k},\mathbf{k}\cdot\mathbf{v}) = v_e^2\int d\mathbf{v}'\,\frac{\mathbf{k}\cdot\nabla_{\mathbf{v}'} f(\mathbf{v}')}{\mathbf{k}\cdot(\mathbf{v}-\mathbf{v}')} \tag{5.24}$$

Again, the velocity integral must be performed along the Landau contour, as discussed in the next chapter. The wave number integral in (5.23) is performed as follows. When we orient the $\hat{k}_1$ axis in the $\mathbf{v} - \mathbf{v}'$ direction, the Q_{ij} component of the tensor $\overleftrightarrow{\mathbf{Q}}$ is

$$Q_{ij}(\mathbf{v},\mathbf{v}') = -\frac{2n_0 e^4}{m_e^2}\int dk_1\,dk_2\,dk_3\,\frac{k_i k_j}{k^4}\,\frac{1}{|\mathbf{v}-\mathbf{v}'|}\,\frac{\delta(k_1)}{|1 + (\psi/k^2\lambda_e^2)|^2} \tag{5.25}$$

The factor $\delta(k_1)$ implies $Q_{ij} = 0$ if either $i = 1$ or $j = 1$. The k_1 integration is trivially performed using this factor. In cylindrical coordinates with $k_2 = k\cos\theta$, $k_3 = k\sin\theta$, and cutting off the integration at an upper wave number $k_0 = p_0^{-1}$, we find, using Q_{33} as an example,

$$Q_{33}(\mathbf{v},\mathbf{v}') = -\frac{2n_0 e^4}{m_e^2|\mathbf{v}-\mathbf{v}'|}\int_0^{2\pi} d\theta\sin^2\theta\int_0^{k_0}\frac{dk}{k}\,\frac{1}{|1 + (\psi/k^2\lambda_e^2)|^2} \tag{5.26}$$

Since ψ is a function of θ but not of k [see (5.24)], the wave number integration can be performed (Problem 5.3). The result is

$$Q_{33}(\mathbf{v},\mathbf{v}') = -\frac{n_0 e^4}{m_e^2|\mathbf{v}-\mathbf{v}'|}\int_0^{2\pi} d\theta\,\frac{\mathrm{Im}[\psi \ln(1+k_0^2\lambda_e^2/\psi)]}{\mathrm{Im}(\psi)}\sin^2\theta \quad (5.27)$$

It turns out (as can be seen more clearly after a study of the following chapter) that the dimensionless function ψ is of order unity. In addition, we recognize the factor $k_0\lambda_e = \lambda_e/p_0$ from Section 1.6 to be (within factors of order unity) the plasma parameter Λ. Thus, we neglect unity compared to $k_0^2\lambda_e^2/\psi$, and $\ln(\psi)$ compared to $\ln(k_0^2\lambda_e^2) \approx \ln(\Lambda^2) = 2\ln\Lambda$, to obtain

$$Q_{33}(\mathbf{v},\mathbf{v}') = Q_{22}(\mathbf{v},\mathbf{v}') = -\frac{2\pi n_0 e^4}{m_e^2|\mathbf{v}-\mathbf{v}'|}\ln\Lambda . \quad (5.28)$$

Similar arguments yield $Q_{23} = Q_{32} = 0$. A tensor with only the Q_{22} and Q_{33} components nonzero can be conveniently expressed in terms of the unit tensor $\vec{\mathbf{I}} \equiv \hat{k}_1\hat{k}_1 + \hat{k}_2\hat{k}_2 + \hat{k}_3\hat{k}_3$; with $\mathbf{g} \equiv \mathbf{v} - \mathbf{v}'$ and recalling that $\hat{k}_1 = \hat{g}$, we have

$$\vec{\mathbf{Q}}(\mathbf{v},\mathbf{v}') = -\frac{2\pi n_0 e^4 \ln\Lambda}{m_e^2}\,\frac{g^2\vec{\mathbf{I}} - \mathbf{gg}}{g^3} \quad (5.29)$$

This expression is known as the *Landau form* for $\vec{\mathbf{Q}}$.

With $\vec{\mathbf{Q}}$ in the form (5.29), it is possible to put the Lenard–Balescu equation (5.22) in the form of a *Fokker–Planck equation*. The general Fokker–Planck equation is a very important equation in all aspects of statistical physics, and is derived from first principles in Appendix B. Following Montgomery and Tidman [5], we notice that

$$\nabla_{\mathbf{v}}\nabla_{\mathbf{v}} g = \frac{g^2\vec{\mathbf{I}} - \mathbf{gg}}{g^3} \quad (5.30)$$

so that with an integration by parts (5.22) becomes

$$\begin{aligned}\partial_t f(\mathbf{v},t) &= \frac{2\pi n_0 e^4 \ln\Lambda}{m_e^2}\,\nabla_{\mathbf{v}}\cdot[(\nabla_{\mathbf{v}} f)\cdot\nabla_{\mathbf{v}}\nabla_{\mathbf{v}}\int d\mathbf{v}'\, g f(\mathbf{v}') \\ &\quad - f(\mathbf{v})\int d\mathbf{v}'\,\nabla_{\mathbf{v}}(\nabla_{\mathbf{v}}\cdot\nabla_{\mathbf{v}})g\, f(\mathbf{v}')] \\ &= \frac{2\pi n_0 e^4 \ln\Lambda}{m_e^2}\{\nabla_{\mathbf{v}}\nabla_{\mathbf{v}} : [f(\mathbf{v})\nabla_{\mathbf{v}}\nabla_{\mathbf{v}}\int d\mathbf{v}'\, g f(\mathbf{v}')] \\ &\quad - 2\,\nabla_{\mathbf{v}}\cdot[f(\mathbf{v})\int d\mathbf{v}'\,\nabla_{\mathbf{v}}(\nabla_{\mathbf{v}}\cdot\nabla_{\mathbf{v}})g\, f(\mathbf{v}')]\} \\ &= \frac{2\pi n_0 e^4 \ln\Lambda}{m_e^2}\{-4\nabla_{\mathbf{v}}\cdot\left[f(\mathbf{v})\nabla_{\mathbf{v}}\int d\mathbf{v}'\,\frac{f(\mathbf{v}')}{g}\right] \\ &\quad + \nabla_{\mathbf{v}}\nabla_{\mathbf{v}} : [f(\mathbf{v})\nabla_{\mathbf{v}}\nabla_{\mathbf{v}}\int d\mathbf{v}'\, g f(\mathbf{v}')]\} \end{aligned} \quad (5.31)$$

where in the first step we have used $\nabla_{\mathbf{v}} g = -\nabla_{\mathbf{v}'} g$, and in the third step we have used $(\nabla_{\mathbf{v}}\cdot\nabla_{\mathbf{v}})g = 2/g$. This is in the standard form of a Fokker–Planck equation (see Appendix B),

$$\boxed{\frac{\partial f(\mathbf{v},t)}{\partial t} = -\nabla_{\mathbf{v}}\cdot[\mathbf{A} f(\mathbf{v})] + \frac{1}{2}\nabla_{\mathbf{v}}\nabla_{\mathbf{v}} : [\vec{\mathbf{B}} f(\mathbf{v})]} \quad (5.32)$$

where the *coefficient of dynamic friction*

$$\mathbf{A}(\mathbf{v},t) = \frac{8\pi n_0 e^4 \ln\Lambda}{m_e^2}\,\nabla_{\mathbf{v}}\int d\mathbf{v}'\,\frac{f(\mathbf{v}',t)}{|\mathbf{v}-\mathbf{v}'|} \quad (5.33)$$

and the *diffusion coefficient*

$$\vec{\mathbf{B}}(\mathbf{v},t) = \frac{4\pi n_0 e^4 \ln\Lambda}{m_e^2}\,\nabla_{\mathbf{v}}\nabla_{\mathbf{v}}\int d\mathbf{v}'\,|\mathbf{v}-\mathbf{v}'|\, f(\mathbf{v}',t) \quad (5.34)$$

With the coefficients (5.33) and (5.34), Eq. (5.32) is known as the *Landau form* of the Fokker–Planck equation.

The meaning of the terms in the Fokker–Planck equation is discussed in Appendix B. The coefficient of dynamic friction $\mathbf{A}$ represents the slowing down of a typical particle because of many small angle collisions. The diffusion coefficient represents the increase of a typical particle's velocity (in the direction perpendicular to its instantaneous velocity) because of many small angle collisions. Thus, the two terms on the right side of the Fokker–Planck equation (5.32) tend to balance each other. They are in perfect balance when f is a Maxwellian, as shown in Problem 5.5.

The Landau form of the Fokker–Planck equation (5.32) has been solved numerically by MacDonald et al. [7] (Fig. 5.2). The initial distribution function $f(\mathbf{v},t=0) = f(|\mathbf{v}|,t=0)$ is spherically symmetric in velocity space. Figure 5.2 shows the steady progression of the distribution, as time increases, toward a Maxwellian. At late times, there is an overshoot at low speeds, which indicates that it

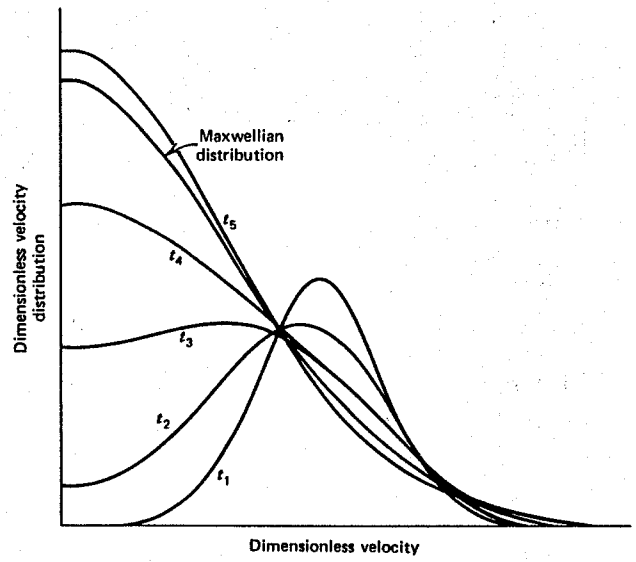

Fig. 5.2 Time evolution of a spherically symmetric electron distribution function as obtained from a numerical solution of the Landau form of the Fokker–Planck equation (5.32) by MacDonald et al. [7].

takes a long time to populate the high speed tail of the Maxwellian. (Remember that Coulomb collisions become quite weak for fast particles.)

There exist even simpler forms of the Fokker–Planck equation [4] but these are not too accurate and are used only to get a rough idea of collisional effects. One is

$$\frac{\partial f}{\partial t} = \nu \nabla_{\mathbf{v}} \cdot [(\mathbf{v} - \mathbf{v}_0) f + v_e^2 \nabla_{\mathbf{v}} f] \tag{5.35}$$

where ν is a collision frequency, and $\mathbf{v}_0$ is a constant velocity. An even cruder model, which is not related to the development of the present chapter, is the *Krook model*,

$$\frac{\partial f}{\partial t} = -\nu(f - f_0) \tag{5.36}$$

where f_0 is the appropriate Maxwellian distribution. Equation (5.36) is also called the BGK equation, after *B*hatnagar, *G*ross, and *K*rook [8].

This brings us to the end of our study of plasma kinetic theory including the effects of two-body collisions. The material in this chapter can be truly appreciated only after a careful study of Appendices A and B. However, Appendix A itself can best be understood after one has mastered the treatment of the Vlasov equation, to which we turn our attention in the next chapter.

REFERENCES

[1] N. N. Bogoliubov, *Problems of a Dynamical Theory in Statistical Physics*, State Technical Press, Moscow, 1946.

[2] A. Lenard, *Ann. Phys. (New York)*, *10*, 390 (1960).

[3] R. Balescu, *Phys. Fluids*, *3*, 52 (1960).

[4] P. C. Clemmow and J. P. Dougherty, *Electrodynamics of Particles and Plasmas*, Addison-Wesley, Reading, Mass., 1969.

[5] D. C. Montgomery and D. A. Tidman, *Plasma Kinetic Theory*, McGraw-Hill, New York, 1964.

[6] D. C. Montgomery, *Theory of the Unmagnetized Plasma*, Gordon and Breach, New York, 1971.

[7] W. M. MacDonald, M. N. Rosenbluth, and W. Chuck, *Phys. Rev.*, *107*, 350 (1957).

[8] P. L. Bhatnagar, E. P. Gross, and M. Krook, *Phys. Rev.*, *94*, 511 (1954).

[9] M. Rosenbluth, W. M. MacDonald, and D. L. Judd, *Phys. Rev.*, *107*, 1 (1957).

PROBLEMS

5.1 Fourier Transforms

Find the Fourier transforms (5.16) and (5.17). (*Hint:* Use spherical polar coordinates with $\mathbf{k} \cdot \mathbf{x} = kr \cos \theta$.)

5.2 Lenard–Balescu Equation

After referring to Clemmow and Dougherty [4], and Montgomery and Tidman [5], sketch the proofs of properties (a) to (f) of the Lenard–Balescu equation as listed below (5.21).

5.3 An Integral

With the help of a table of integrals, perform the integration in (5.26).

5.4 Simpler Derivation of the Landau Form

The development of the Landau form for $\vec{\mathbf{Q}}$, from (5.23) to (5.28), is the standard one. However, a simpler one exists. In (5.23), replace ϵ by unity, and cut off the wave number integration at a lower wave number λ_e^{-1} as well as at the upper wave number p_0^{-1}. Show that (5.28) results. The replacement of ϵ by unity is equivalent to ignoring the shielding, as can be seen in (5.20).

5.5 Maxwellian

Show that a Maxwellian is an exact time-independent solution of both the Lenard–Balescu equation (5.19) and the Landau form of the Fokker–Planck equation (5.32).

5.6 Two-Point Correlation Function

Discuss the meaning of $f_2 = f_1 f_1 + g$. Why should g depend on f_1? In particular, how would g change as we turn up the temperature of a Maxwellian?

5.7 Plasmas and Brownian Motion

Discuss the analogy between collisional effects on a particle in a plasma and Brownian motion. Explain why the collisional effects can be described by a Fokker–Planck equation. Thus, using only words, explain how we could use the results of Section 1.6 on collisions to obtain the Fokker–Planck equation directly, without starting from Liouville → BBGKY → Lenard–Balescu → Fokker–Planck. This is actually the technique used by Rosenbluth et al. [9].

5.8 Units

Check all of the units in (5.19) to (5.36). Using crude dimensional arguments, derive the model (5.35) from the Fokker–Planck equation (5.32) and the coefficients (5.33) and (5.34).

APPENDIX A

Derivation of the Lenard–Balescu Equation

In this appendix, we complete the derivation of the Lenard–Balescu equation (5.19) starting from Eqs. (5.1) and (5.4), which in turn are obtained from the BBGKY hierarchy by discarding three-particle correlations (Refs. [1] to [5]). From (5.1) and (5.4) to (5.8), we have

$$\partial_t f_1(\mathbf{v}_1,t) = -n_0 \int d\mathbf{x}_2\, d\mathbf{v}_2 \mathbf{a}_{12} \cdot \nabla_{\mathbf{v}_1} g(\mathbf{x}_1 - \mathbf{x}_2,\mathbf{v}_1,\mathbf{v}_2,t) \tag{A.1}$$

$$\frac{\partial}{\partial \tilde{t}}\, g(\mathbf{x}_1 - \mathbf{x}_2,\mathbf{v}_1,\mathbf{v}_2,\tilde{t}) + V_1 g + V_2 g = S(\mathbf{x}_1 - \mathbf{x}_2,\mathbf{v}_1,\mathbf{v}_2) \tag{A.2}$$

$$V_1 g(12) = \mathbf{v}_1 \cdot \nabla_{\mathbf{x}_1} g(12) + [n_0 \int d3\ \mathbf{a}_{13}\, g(32)] \cdot \nabla_{\mathbf{v}_1} f_1(\mathbf{v}_1) \tag{A.3}$$

$$V_2 g(12) = \mathbf{v}_2 \cdot \nabla_{\mathbf{x}_2} g(12) + [n_0 \int d3\ \mathbf{a}_{23}\, g(13)] \cdot \nabla_{\mathbf{v}_2} f_1(\mathbf{v}_2) \tag{A.4}$$

$$S(\mathbf{x}_1 - \mathbf{x}_2,\mathbf{v}_1,\mathbf{v}_2) = -(\mathbf{a}_{12} \cdot \nabla_{\mathbf{v}_1} + \mathbf{a}_{21} \cdot \nabla_{\mathbf{v}_2}) f_1(\mathbf{v}_1) f_1(\mathbf{v}_2) \tag{A.5}$$

where we have used $g(32) = g(23)$, we alternate between the notations (1) and $(\mathbf{x}_1,\mathbf{v}_1)$ depending on convenience, and we recall from Chapter 5 that we wish to solve for $g(\tilde{t} \to \infty)$ where $\tilde{t}$ is the fast time scale on which g relaxes. On this fast time scale, the functions f_1 and thus S are considered to be constants. We shall also need, from (5.9),

$$\mathbf{a}_{ij} = \frac{e^2}{m_e |\mathbf{x}_i - \mathbf{x}_j|^3} (\mathbf{x}_i - \mathbf{x}_j) \tag{A.6}$$

Using the Fourier transform conventions in Chapter 5, we spatially Fourier transform these equations with respect to $\mathbf{x}_1$ and $\mathbf{x}_2$. Because of the appearance of the combination $(\mathbf{x}_1 - \mathbf{x}_2)$, we obtain the factor $\delta(\mathbf{k}_1 + \mathbf{k}_2)$ in several places and, thus, can replace $\mathbf{k}_2$ by $-\mathbf{k}_1$.

EXERCISE For any function $f(\mathbf{x}_1 - \mathbf{x}_2)$, show that the double Fourier transform with respect to $\mathbf{x}_1$ and $\mathbf{x}_2$ is $\delta(\mathbf{k}_1 + \mathbf{k}_2) f(\mathbf{k}_1)$ where $f(\mathbf{k})$ is the Fourier transform of $f(\mathbf{x})$ with respect to $\mathbf{x}$.

EXERCISE Show that $\int d\mathbf{x} f_1(\mathbf{x}) f_2(\mathbf{x}) = (2\pi)^3 \int d\mathbf{k} f_1(-\mathbf{k}) f_2(\mathbf{k})$ for any functions f_1 and f_2; here, $f_1(\mathbf{k})$ is the Fourier transform of $f_1(\mathbf{x})$, etc., as usual.

EXERCISE Show that the double Fourier transform of $\int d\mathbf{x}_3 f_1(\mathbf{x}_1 - \mathbf{x}_3) \times f_2(\mathbf{x}_2 - \mathbf{x}_3)$ is $(2\pi)^3 \delta(\mathbf{k}_1 + \mathbf{k}_2) f_1(\mathbf{k}_1) f_2(-\mathbf{k}_1)$.

With the results of these exercises, and Eq. (5.16) for $\mathbf{a}_{12}(\mathbf{k})$, the Fourier transformed version of (A.1) to (A.6) is

$$\partial_t f_1(\mathbf{v}_1, t) = -\frac{i n_0 (2\pi)^3}{m_e} \nabla_{\mathbf{v}_1} \cdot \int d\mathbf{v}_2 \, d\mathbf{k}_1 \, \mathbf{k}_1 \, \varphi(k_1) g(\mathbf{k}_1, \mathbf{v}_1, \mathbf{v}_2, \tilde{t} = \infty) \tag{A.7}$$

$$\frac{\partial}{\partial \tilde{t}} g(\mathbf{k}_1, \mathbf{v}_1, \mathbf{v}_2, \tilde{t}) + V_1 g + V_2 g = S(\mathbf{k}_1, \mathbf{v}_1, \mathbf{v}_2) \tag{A.8}$$

$$V_1 g(12) = i\mathbf{k}_1 \cdot \mathbf{v}_1 g(12)$$

$$- \frac{n_0 (2\pi)^3}{m_e} i\mathbf{k}_1 \cdot \nabla_{\mathbf{v}_1} f_1(\mathbf{v}_1) \varphi(k_1) \int d\mathbf{v}_3 \, g(\mathbf{k}_1, \mathbf{v}_3, \mathbf{v}_2, \tilde{t}) \tag{A.9}$$

$$V_2 g(12) = -i\mathbf{k}_1 \cdot \mathbf{v}_2 \, g(12)$$

$$+ \frac{n_0 (2\pi)^3}{m_e} i\mathbf{k}_1 \cdot \nabla_{\mathbf{v}_2} f_1(\mathbf{v}_2) \varphi(k_1) \int d\mathbf{v}_3 g(\mathbf{k}_1, \mathbf{v}_1, \mathbf{v}_3, \tilde{t}) \tag{A.10}$$

$$S(\mathbf{k}_1, \mathbf{v}_1, \mathbf{v}_2) = \frac{\varphi(k_1)}{m_e} i\mathbf{k}_1 \cdot (\nabla_{\mathbf{v}_1} - \nabla_{\mathbf{v}_2}) f_1(\mathbf{v}_1) f_1(\mathbf{v}_2) \tag{A.11}$$

Our goal is to express the right side of (A.7) in terms of f_1 by solving (A.8) for g. With (A.7) in its present form, the remainder of the calculation can be performed in wave number space; because of the factor i on the right of (A.7) and the fact that the right of (A.7) must be real, we need only calculate the imaginary part of $g(\mathbf{k}_1, \mathbf{v}_1, \mathbf{v}_2, \tilde{t} = \infty)$.

The solution of (A.8) for $g(\mathbf{k}_1, \mathbf{v}_1, \mathbf{v}_2, \tilde{t} = \infty)$ is accomplished by Laplace transforming with respect to the fast time $\tilde{t}$.

EXERCISE For any function $g(t)$, show that the Laplace transform of dg/dt is $-g(t = 0) - i\omega \, g(\omega)$.

With the result of this exercise, the Laplace transform of (A.8) is

$$-g(\mathbf{k}_1, \mathbf{v}_1, \mathbf{v}_2, \tilde{t} = 0) - i\omega \, g(\mathbf{k}_1, \mathbf{v}_1, \mathbf{v}_2, \omega) + V_1 \, g(12\omega)$$

$$+ V_2 g(12\omega) = -\frac{1}{i\omega} S(\mathbf{k}_1, \mathbf{v}_1, \mathbf{v}_2) \tag{A.12}$$

where $g(\omega)$ is defined only for $\omega_i \equiv \mathrm{Im}(\omega)$ sufficiently large, and where the operators V_1 and V_2 can be regarded as numbers since they have no time dependence in (A.9) and (A.10). Solving (A.12) for $g(\omega)$ we find

$$g(\omega) = \frac{g(\tilde{t} = 0) - (S/i\omega)}{-i\omega + V_1 + V_2} \tag{A.13}$$

We require $g(\tilde{t} = \infty)$: This can be obtained from the inverse Laplace transform of (A.13). It turns out that distribution functions $f_1(\mathbf{v})$ that are stable in the Vlasov sense (Chapter 6) are such that the zeros of $-i\omega + V_1 + V_2$ always occur in the lower half ω-plane. We consider only such stable distribution functions $f_1(\mathbf{v})$. Thus, the inverse Laplace transform

$$g(\tilde{t}) = \int_L \frac{d\omega}{2\pi} \frac{g(\tilde{t} = 0) - S/i\omega}{-i\omega + V_1 + V_2} e^{-i\omega t} \tag{A.14}$$

can be performed by deforming the Laplace contour as shown in Fig. A.1. Since poles in the lower half plane contribute only damped functions of time, $\sim \exp(\omega_i t)$, the only pole that contributes to $g(\tilde{t} = \infty)$ is the one at $\omega = 0$; therefore,

$$g(\tilde{t} = \infty) = \lim_{\omega \to 0} \frac{S}{-i\omega + V_1 + V_2} \tag{A.15}$$

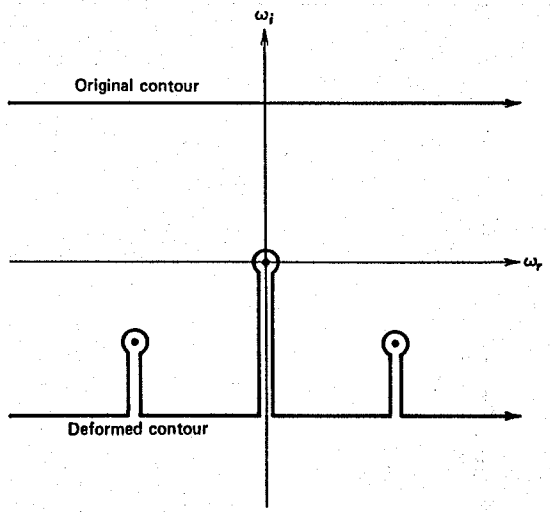

Fig. A.1 Inverse Laplace contour for calculating $g(\tilde{t} = \infty)$.

where we retain the $\lim_{\omega\to 0}$ to help us interpret other contour integrations that occur in the calculation.

At this point, we introduce a trick that allows us to treat the operators V_1 and V_2 separately, rather than in the combination $V_1 + V_2$. Consider

$$\frac{1}{-i\omega + V_1 + V_2} = \int_0^\infty dt\, e^{(i\omega - V_1 - V_2)t}$$

$$= \int_0^\infty dt\, e^{i\omega t} \int_{C_1} \frac{d\omega_1}{2\pi} \frac{e^{-i\omega_1 t}}{-i\omega_1 + V_1} \int_{C_2} \frac{d\omega_2}{2\pi} \frac{e^{-i\omega_2 t}}{-i\omega_2 + V_2}$$

$$= \int_{C_1} \frac{d\omega_1}{2\pi} \int_{C_2} \frac{d\omega_2}{2\pi} \frac{1}{-i\omega_1 + V_1} \frac{1}{-i\omega_2 + V_2} \frac{1}{-i(\omega - \omega_1 - \omega_2)} \tag{A.16}$$

where the contours C_1 and C_2 must be chosen so that $\omega_i > \omega_{1i} + \omega_{2i}$. Then (A.15) becomes

$$g(\mathbf{k}_1,\mathbf{v}_1,\mathbf{v}_2,\tilde{t} = \infty)$$

$$= \lim_{\omega\to 0} \int_{C_1} \frac{d\omega_1}{2\pi} \int_{C_2} \frac{d\omega_2}{2\pi} \frac{1}{-i\omega_2 + V_2} \frac{1}{-i\omega_1 + V_1} \frac{S(\mathbf{k}_1,\mathbf{v}_1,\mathbf{v}_2)}{-i(\omega - \omega_1 - \omega_2)} \tag{A.17}$$

In expressions (A.13) to (A.17), we interpret the meaning of an inverse operator $(-i\omega_1 + V_1)^{-1}F$ acting on a function F to be that function G such that $F = (-i\omega_1 + V_1)G$.

We first need

$$\alpha(\mathbf{k}_1,\mathbf{v}_1,\mathbf{v}_2) = \frac{1}{-i\omega_1 + V_1} S(\mathbf{k}_1,\mathbf{v}_1,\mathbf{v}_2) \tag{A.18}$$

such that

$$S(\mathbf{k}_1,\mathbf{v}_1,\mathbf{v}_2) = (-i\omega_1 + V_1)\alpha(\mathbf{k}_1,\mathbf{v}_1,\mathbf{v}_2)$$

$$= (-i\omega_1 + i\mathbf{k}_1\cdot\mathbf{v}_1)\alpha(\mathbf{k}_1,\mathbf{v}_1,\mathbf{v}_2) - \frac{i(2\pi)^3 n_0}{m_e} \mathbf{k}_1\cdot\nabla_{\mathbf{v}_1}$$

$$\times f_1(\mathbf{v}_1)\varphi(k_1)\int d\mathbf{v}_3\, \alpha(\mathbf{k}_1,\mathbf{v}_3,\mathbf{v}_2) \tag{A.19}$$

In order to solve this for α we must first eliminate $\int d\mathbf{v}_3\, \alpha$; we express (A.19) as

$$\alpha(\mathbf{k}_1,\mathbf{v}_1,\mathbf{v}_2) = \frac{1}{-i\omega_1 + i\mathbf{k}_1\cdot\mathbf{v}_1} [S(\mathbf{k}_1,\mathbf{v}_1,\mathbf{v}_2)$$

$$+ \frac{i(2\pi)^3 n_0}{m_e} \mathbf{k}_1\cdot\nabla_{\mathbf{v}_1} f_1(\mathbf{v}_1)\varphi(k_1)\int d\mathbf{v}_3\, \alpha(\mathbf{k}_1,\mathbf{v}_3,\mathbf{v}_2)] \tag{A.20}$$

and integrate over all $\mathbf{v}_1$ to find

$$\int d\mathbf{v}_1\, \alpha(\mathbf{k}_1,\mathbf{v}_1,\mathbf{v}_2)$$

$$= \int d\mathbf{v}_1 \frac{S(\mathbf{k}_1,\mathbf{v}_1,\mathbf{v}_2)}{-i\omega_1 + i\mathbf{k}_1\cdot\mathbf{v}_1} + \left[\int d\mathbf{v}_3\, \alpha(\mathbf{k}_1,\mathbf{v}_3,\mathbf{v}_2)\right]$$

$$\times \frac{i(2\pi)^3 n_0}{m_e} \varphi(k_1)\int d\mathbf{v}_1 \frac{\mathbf{k}_1\cdot\nabla_{\mathbf{v}_1} f_1(\mathbf{v}_1)}{-i\omega_1 + i\mathbf{k}_1\cdot\mathbf{v}_1} \tag{A.21}$$

Realizing that $\mathbf{v}_3$ on the right is merely a dummy variable of integration, we find

$$\int d\mathbf{v}_3\, \alpha(\mathbf{k}_1,\mathbf{v}_3,\mathbf{v}_2) = \frac{1}{\epsilon(\mathbf{k}_1,\omega_1)} \int d\mathbf{v}_3 \frac{S(\mathbf{k}_1,\mathbf{v}_3,\mathbf{v}_2)}{-i\omega_1 + i\mathbf{k}_1\cdot\mathbf{v}_3} \tag{A.22}$$

where

$$\epsilon(\mathbf{k}_1,\omega_1) = 1 + \frac{\omega_e^2}{k_1^2}\int d\mathbf{v}_1 \frac{\mathbf{k}_1\cdot\nabla_{\mathbf{v}_1} f_1(\mathbf{v}_1)}{\omega_1 - \mathbf{k}_1\cdot\mathbf{v}_1} \tag{A.23}$$

is the dielectric function encountered in Chapter 6. Thus, (A.20) becomes

$$\alpha(\mathbf{k}_1,\mathbf{v}_1,\mathbf{v}_2) = \frac{1}{-i\omega_1 + i\mathbf{k}_1\cdot\mathbf{v}_1} \left[S(\mathbf{k}_1,\mathbf{v}_1,\mathbf{v}_2)\right.$$

$$\left. + \frac{i(2\pi^3) n_0 \mathbf{k}_1\cdot\nabla_{\mathbf{v}_1} f_1(\mathbf{v}_1)\varphi(k_1)/m_e}{\epsilon(\mathbf{k}_1,\omega_1)} \int d\mathbf{v}_3 \frac{S(\mathbf{k}_1,\mathbf{v}_3,\mathbf{v}_2)}{-i\omega_1 + i\mathbf{k}_1\cdot\mathbf{v}_3}\right] \tag{A.24}$$

which completes the inversion of the operator $(-i\omega_1 + V_1)^{-1}$

Next, we need

$$\int d\mathbf{v}_2\, \beta(\mathbf{k}_1,\mathbf{v}_1,\mathbf{v}_2) = \int d\mathbf{v}_2 \frac{1}{-i\omega_2 + V_2} \alpha(\mathbf{k}_1,\mathbf{v}_1,\mathbf{v}_2) \tag{A.25}$$

where we have noted from (A.7) that we need $\int d\mathbf{v}_2\, g$ rather than g, allowing us to use the compact analogue of (A.22). Noting that V_2 is the same as V_1 if the sign of $\mathbf{k}_1$ is changed and if $\mathbf{v}_1$ and $\mathbf{v}_2$ are interchanged appropriately, we find

$$\int d\mathbf{v}_2\, \beta(\mathbf{k}_1,\mathbf{v}_1,\mathbf{v}_2) = \frac{1}{\epsilon(-\mathbf{k}_1,\omega_2)} \int d\mathbf{v}_2 \frac{\alpha(\mathbf{k}_1,\mathbf{v}_1,\mathbf{v}_2)}{-i\omega_2 - i\mathbf{k}_1\cdot\mathbf{v}_2} \tag{A.26}$$

With the result (A.26) we have from (A.17)

$$\int d\mathbf{v}_2\, g(\mathbf{k}_1,\mathbf{v}_1,\mathbf{v}_2,\tilde{t} = \infty)$$

$$= \lim_{\omega\to 0} \int_{C_1} \frac{d\omega_1}{2\pi} \int_{C_2} \frac{d\omega_2}{2\pi} \frac{1}{\epsilon(-\mathbf{k}_1,\omega_2)} \frac{1}{-i(\omega - \omega_1 - \omega_2)}$$

$$\times \int d\mathbf{v}_2 \frac{1}{-i\omega_2 - i\mathbf{k}_1\cdot\mathbf{v}_2} \frac{1}{-i\omega_1 + V_1} S(\mathbf{k}_1,\mathbf{v}_1,\mathbf{v}_2) \tag{A.27}$$

We perform the ω_2 integration first, along the contour C_2 shown in Fig. (A.2). Since the integrand behaves like ω_2^{-2} for large ω_2, we can close the contour upward

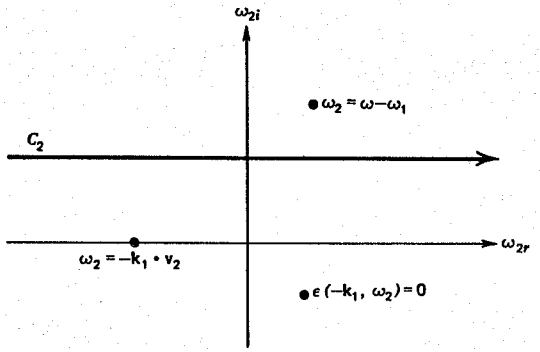

Fig. A.2 Contour C_2 used in evaluating (A.27).

and pick up only the pole at $\omega_2 = \omega - \omega_1$, yielding

$$\int d\mathbf{v}_2\, g(\mathbf{k}_1,\mathbf{v}_1,\mathbf{v}_2,\tilde{t} = \infty) = \lim_{\omega\to 0} \int d\mathbf{v}_2 \int_{C_1} \frac{d\omega_1}{2\pi}$$

$$\times \frac{1}{\epsilon(-\mathbf{k}_1,\omega - \omega_1)} \frac{1}{-i(\omega - \omega_1) - i\mathbf{k}_1 \cdot \mathbf{v}_2} \frac{1}{-i\omega_1 + V_1} S(\mathbf{k}_1,\mathbf{v}_1,\mathbf{v}_2) \quad \text{(A.28)}$$

Inserting the results (A.11) and (A.24) we have

$$\int d\mathbf{v}_2\, g(\mathbf{k}_1,\mathbf{v}_1,\mathbf{v}_2,\tilde{t} = \infty) = \lim_{\omega\to 0} \int d\mathbf{v}_2 \int_{C_1} \frac{d\omega_1}{2\pi}$$

$$\times \frac{1}{\epsilon(-\mathbf{k}_1,\omega - \omega_1)} \frac{1}{-i(\omega - \omega_1) - i\mathbf{k}_1 \cdot \mathbf{v}_2} \frac{1}{-i\omega_1 + i\mathbf{k}_1 \cdot \mathbf{v}_1}$$

$$\times \left[\frac{i\mathbf{k}_1}{m_e} \cdot \varphi(k_1)(\overset{①}{\nabla_{\mathbf{v}_1}} - \overset{②}{\nabla_{\mathbf{v}_2}}) f_1(\mathbf{v}_1) f_1(\mathbf{v}_2)\right.$$

$$- \frac{i(2\pi)^3 n_0 \mathbf{k}_1 \cdot \nabla_{\mathbf{v}_1} f_1(\mathbf{v}_1) \varphi^2(k_1)/m_e^2}{\epsilon(\mathbf{k}_1,\omega_1)}$$

$$\left.\times \int d\mathbf{v}_3 \frac{\mathbf{k}_1 \cdot (\overset{③}{\nabla_{\mathbf{v}_3}} - \overset{④}{\nabla_{\mathbf{v}_2}})}{\omega_1 - \mathbf{k}_1 \cdot \mathbf{v}_3} f_1(\mathbf{v}_3) f_1(\mathbf{v}_2)\right] \quad \text{(A.29)}$$

There are four numbered terms in the square brackets. Including the $\mathbf{v}_2$ integration and the denominator containing $\mathbf{v}_2$, we have

$$② = \frac{-i\mathbf{k}_1 \cdot \varphi(k_1)}{m_e} f_1(\mathbf{v}_1) \int d\mathbf{v}_2 \frac{\nabla_{\mathbf{v}_2} f_1(\mathbf{v}_2)}{-i(\omega - \omega_1) - i\mathbf{k}_1 \cdot \mathbf{v}_2}$$

$$= \frac{-f_1(\mathbf{v}_1)}{n_0(2\pi)^3} [\epsilon(-\mathbf{k}_1,\omega - \omega_1) - 1] \quad \text{(A.30)}$$

where (A.23) has been used. Similarly,

$$③ = \frac{-i(2\pi)^3 n_0}{\epsilon(\mathbf{k}_1,\omega_1)} \left[\int d\mathbf{v}_2 \frac{f_1(\mathbf{v}_2)}{-i(\omega - \omega_1) - i\mathbf{k}_1 \cdot \mathbf{v}_2}\right]$$

$$\times \mathbf{k}_1 \cdot \nabla_{\mathbf{v}_1} f_1(\mathbf{v}_1) \varphi(k_1)/m_e \left[\frac{\varphi(k_1)}{m_e} \int d\mathbf{v}_3 \frac{\mathbf{k}_1 \cdot \nabla_{\mathbf{v}_3} f_1(\mathbf{v}_3)}{\omega_1 - \mathbf{k}_1 \cdot \mathbf{v}_3}\right]$$

$$= \left[1 - \frac{1}{\epsilon(\mathbf{k}_1,\omega_1)}\right] (-i)\mathbf{k}_1 \cdot \nabla_{\mathbf{v}_1} f_1(\mathbf{v}_1) \frac{\varphi(k_1)}{m_e}$$

$$\times \int d\mathbf{v}_2 \frac{f_1(\mathbf{v}_2)}{-i(\omega - \omega_1) - i\mathbf{k}_1 \cdot \mathbf{v}_2} \quad \text{(A.31)}$$

where (A.23) has been used again. Likewise,

$$④ = \left[i\mathbf{k}_1 \cdot \frac{\varphi(k_1)}{m_e} (2\pi)^3 n_0 \int d\mathbf{v}_2 \frac{\nabla_{\mathbf{v}_2} f_1(\mathbf{v}_2)}{-i(\omega - \omega_1) - i\mathbf{k}_1 \cdot \mathbf{v}_2}\right]$$

$$\times \frac{\mathbf{k}_1 \cdot \nabla_{\mathbf{v}_1} f_1(\mathbf{v}_1) \varphi(k_1)/m_e}{\epsilon(\mathbf{k}_1,\omega_1)} \int d\mathbf{v}_3 \frac{f_1(\mathbf{v}_3)}{\omega_1 - \mathbf{k}_1 \cdot \mathbf{v}_3}$$

$$= \frac{[\epsilon(-\mathbf{k}_1,\omega - \omega_1) - 1]}{\epsilon(\mathbf{k}_1,\omega_1)} \mathbf{k}_1 \cdot \nabla_{\mathbf{v}_1} f_1(\mathbf{v}_1) \frac{\varphi(k_1)}{m_e}$$

$$\times \int d\mathbf{v}_3 \frac{f_1(\mathbf{v}_3)}{\omega_1 - \mathbf{k}_1 \cdot \mathbf{v}_3} \quad \text{(A.32)}$$

Cancelling term ① with one of the terms in term ③, we combine the remaining terms to obtain

$$\int d\mathbf{v}_2 g(\mathbf{k}_1,\mathbf{v}_1,\mathbf{v}_2,\tilde{t} = \infty) = \lim_{\omega\to 0} \int_{C_1} \frac{d\omega_1}{2\pi} \frac{1}{-i\omega_1 + i\mathbf{k}_1 \cdot \mathbf{v}_1}$$

$$\times \left\{\left[1 - \frac{1}{\epsilon(-\mathbf{k}_1,\omega - \omega_1)}\right]\left[-\frac{f_1(\mathbf{v}_1)}{n_0(2\pi)^3} + \frac{\mathbf{k}_1 \cdot \nabla_{\mathbf{v}_1} f_1(\mathbf{v}_1) \varphi(k_1)/m_e}{\epsilon(\mathbf{k}_1,\omega_1)}\right.\right.$$

$$\left.\left.\times \int d\mathbf{v}_2 \frac{f_1(\mathbf{v}_2)}{\omega_1 - \mathbf{k}_1 \cdot \mathbf{v}_2}\right] + \frac{i\mathbf{k}_1 \cdot \nabla_{\mathbf{v}_1} f_1(\mathbf{v}_1) \varphi(k_1)/m_e}{\epsilon(\mathbf{k}_1,\omega_1)\epsilon(-\mathbf{k}_1,\omega - \omega_1)}\right.$$

$$\left.\times \int d\mathbf{v}_2 \frac{f_1(\mathbf{v}_2)}{-i(\omega - \omega_1) - i\mathbf{k}_1 \cdot \mathbf{v}_2}\right\} \quad \text{(A.33)}$$

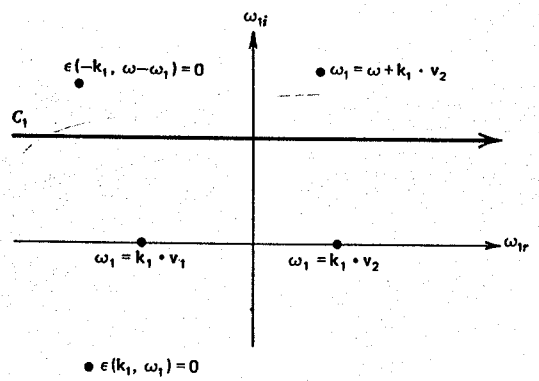

Fig. A.3 Contour C_1 used in evaluating (A.33).

where a change of dummy variable has occurred in term ⓓ. Recalling from below (A.16) that we still have $\omega_i > \omega_{1i}$ along the C_1 contour, and recalling that stable distributions $f_1(v)$ imply that the zeroes of $\epsilon(k,\omega)$ occur only for $\omega_i < 0$, the C_1 contour and the poles of the integrand on the right of (A.33) are as shown in Fig. A.3. Note that not all of the poles occur for each of the terms in (A.33).

Term ⓒ is evaluated by closing the contour C_1 downward, yielding a contribution only from the pole at $\omega_1 = k_1 \cdot v_1$, which gives

$$ⓒ = -\frac{f_1(v_1)}{n_0(2\pi)^3}\left[1 - \frac{1}{\epsilon(-k_1,\omega - k_1\cdot v_1)}\right] \tag{A.34}$$

EXERCISE Convince yourself that the integrand falls off fast enough at large ω_1 to allow the contour to be closed downward.

Term ⓐ × ⓓ vanishes when the contour is closed upward. Finally, we consider the remaining two pieces together; these are

$$\begin{aligned} ⓑ \times ⓓ + ⓒ = \lim_{\omega\to 0} \int_{C_1} \frac{d\omega_1}{2\pi} \frac{1}{-i\omega_1 + ik_1\cdot v_1} \frac{1}{\epsilon(-k_1,\omega-\omega_1)} \\ \times \frac{1}{\epsilon(k_1,\omega_1)} k_1\cdot\nabla_{v_1} f_1(v_1) \frac{\varphi(k_1)}{m_e} \int dv_2 f_1(v_2) \\ \times \left[\frac{-1}{\omega_1 - k_1\cdot v_2} + \frac{1}{\omega_1 - \omega - k_1\cdot v_2}\right] \end{aligned} \tag{A.35}$$

At this point it is convenient to use the fact that in (A.7) we only need the imaginary part of $\int dv_2\, g$. As we move the contour in Fig. A.3 down to the real axis (let $\omega \to 0$), the integrand of (A.35) appears to vanish. However, we must be careful at the pole $\omega_1 = k_1 \cdot v_1$ and at the two poles that pinch the contour at $\omega_1 = k_1 \cdot v_2$ and $\omega_1 = \omega + k_1 \cdot v_2$. Recall the Plemelj formulas,

$$\lim_{\eta\to 0} \frac{1}{\omega - a \pm i\eta} = P\left(\frac{1}{\omega - a}\right) \mp i\pi\delta(\omega - a) \tag{A.36}$$

where the upper sign is used when a contour passes above a pole, and the lower sign is used when a contour passes below a pole. Then

$$\begin{aligned} \mathrm{Re}\lim_{\omega\to 0} \frac{1}{\omega_1 - k_1\cdot v_1}\left[\frac{-1}{\omega_1 - k_1\cdot v_2} + \frac{1}{\omega_1 - \omega - k_1\cdot v_2}\right] \\ = \mathrm{Re}\left[P\left(\frac{1}{\omega_1 - k_1\cdot v_1}\right) - i\pi\delta(\omega_1 - k_1\cdot v_1)\right] \\ \times\left[-P\left(\frac{1}{\omega_1 - k_1\cdot v_2}\right) + i\pi\delta(\omega_1 - k_1\cdot v_2)\right. \\ \left. + P\left(\frac{1}{\omega_1 - k_1\cdot v_2}\right) + i\pi\delta(\omega_1 - k_1\cdot v_2)\right] \\ = 2\pi^2\delta(\omega_1 - k_1\cdot v_1)\delta(\omega_1 - k_1\cdot v_2) \end{aligned} \tag{A.37}$$

where Re indicates the real part. If we use one of the δ-functions to perform the ω_1 integration, (A.35) yields

$$\begin{aligned} \mathrm{Im}[ⓑ \times ⓓ + ⓒ] = i\pi k_1\cdot\nabla_{v_1} f_1(v_1) \frac{\varphi(k_1)}{m_e} \\ \times \int dv_2 \frac{\delta[k_1\cdot(v_1 - v_2)] f_1(v_2)}{|\epsilon(k_1, k_1\cdot v_1)|^2} \end{aligned} \tag{A.38}$$

where we have used the fact that $\epsilon(-k,-\omega) = \epsilon^*(k,\omega)$ when ω is real.

EXERCISE Demonstrate this fact from the definition (A.23) of $\epsilon(k,\omega)$. Show that for ω real, $\mathrm{Im}[\epsilon(k,\omega)] = -i\pi(\omega_e^2/k^2)\int dv\,[k\cdot\nabla_v f_1(v)]\delta(\omega - k\cdot v)$.

Similarly, if one uses the results of the exercise,

$$\begin{aligned} \mathrm{Im}[ⓒ] = \frac{f_1(v_1)}{n_0(2\pi)^3} \frac{\mathrm{Im}[\epsilon(k_1, k_1\cdot v_1)]}{|\epsilon(k_1, k_1\cdot v_1)|^2} = \frac{-i\pi f_1(v_1)\varphi(k_1)/m_e}{|\epsilon(k_1, k_1\cdot v_1)|^2} \\ \times \int dv_2\,[k_1\cdot\nabla_{v_2} f_1(v_2)]\delta[k_1\cdot(v_1 - v_2)] \end{aligned} \tag{A.39}$$

Finally, inserting (A.38) and (A.39) into (A.7), one obtains

$$\begin{aligned} \partial_t f_1(v_1, t) = -\frac{8\pi^4 n_0}{m_e^2} \nabla_{v_1}\cdot\int dk_1\, dv_2 \frac{k_1 k_1\cdot\varphi^2(k_1)}{|\epsilon(k_1, k_1\cdot v_1)|^2} \\ \times \delta[k_1\cdot(v_1 - v_2)][f_1(v_1)\nabla_{v_2} f_1(v_2) - f_1(v_2)\nabla_{v_1} f_1(v_1)] \end{aligned} \tag{A.40}$$

which with appropriate changes of variables is the Lenard–Balescu equation (5.19).

REFERENCES

[1] D. C. Montgomery and D. A. Tidman, *Plasma Kinetic Theory*, McGraw-Hill, New York, 1964.

[2] P. C. Clemmow and J. P. Dougherty, *Electrodynamics of Particles and Plasmas*, Addison-Wesley, Reading, Mass., 1969.

[3] D. C. Montgomery, *Theory of the Unmagnetized Plasma*, Gordon and Breach, New York, 1971.

[4] A. Lenard, *Ann. Phys. (New York)*, *10*, 390 (1960).

[5] R. Balescu, *Phys. Fluids*, *3*, 52 (1960).

APPENDIX B

Langevin Equation, Fluctuation–Dissipation Theorem, Markov Process, and Fokker–Planck Equation

B.1 LANGEVIN EQUATION AND FLUCTUATION-DISSIPATION THEOREM

The discussion of plasma kinetic theory, including collisions, in Chapters 3, 4, and 5, led to the Fokker–Planck form of the plasma kinetic equation in (5.31). This is not a coincidence. In this appendix, it is shown that the Fokker–Planck equation arises naturally whenever a probability distribution [i.e., the one particle distribution function $f_s(\mathbf{v},t)$] changes slowly in time because of huge numbers of small changes (i.e., small angle collisions).

In order to motivate the Fokker–Planck equation, we use a physical example that is simpler than a plasma; namely, the case of Brownian motion. This will lead us to the related topics of the *Langevin equation*, the *fluctuation–dissipation theorem*, and *Markov processes*. As we study the example of Brownian motion, ask yourself how each step corresponds to its analogue in the plasma case.

The Langevin equation arises whenever a variable experiences a slow time variation as a result of a rapidly varying force. The best known example of this is the case of Brownian motion. A large particle (mass $\sim 10^{-12}$ gram) exhibits Brownian motion when bombarded by the molecules in air (mass $\sim 10^{-22}$ gram). The path of the particle may look as shown in Fig. B.1. The human eye, looking through a microscope, cannot see the fine structure on the curve shown, and so instead [1, 2] sees the curve in Fig. B.2. The wandering motion is, essentially, a random walk due to the large number of collisions that the particle suffers per unit time with the gas molecules. Picking out one of the dimensions of the motion, we can write Newton's force law in one spatial dimension,

$$\frac{dv(t)}{dt} = F(t) \tag{B.1}$$

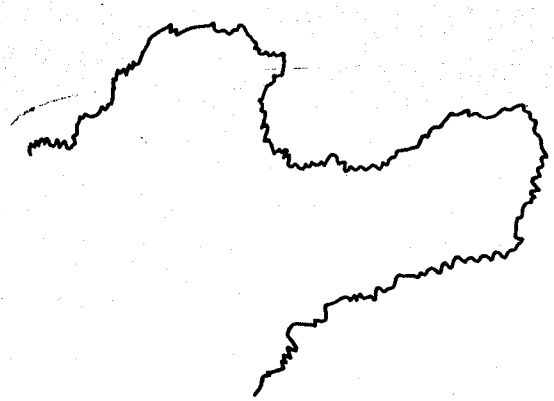

Fig. B.1 Path of a Brownian particle.

where $F(t)$ is the force per unit mass on the Brownian particle. Thus, $F(t)$ contains the sum of many collisions, each lasting an extremely short time.

To study the physics of (B.1), we can consider an ensemble of realizations, each having the same initial speed $v(t = 0) = v_0$ but different random functions $F(t)$. Our intuition tells us that the overall effect of the many collisions will be to slow the Brownian particle, so that $\langle v(t)\rangle = 0$ as $t \to \infty$.

Microscopically, the Brownian particle slows because it collides with more particles in the direction of motion than in the opposite direction. It thus gives up net kinetic energy to the gas molecules, which leave the collision with a net gain in right-going momentum.

This discussion leads to the conclusion that the ensemble average of the force on the right of (B.1) must contain a term that tends to slow the Brownian particle.

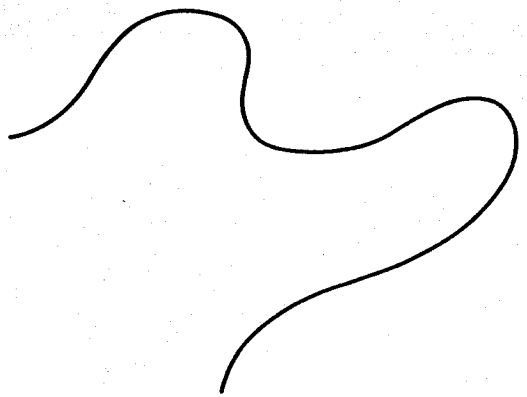

Fig. B.2 Path of a Brownian particle as seen by the human eye.

Thus, we split the force $F(t)$ into two terms,

$$F(t) = \langle F(t)\rangle + \delta F(t) \tag{B.2}$$

so that $\langle \delta F(t)\rangle = 0$. The ensemble averaged part of $F(t)$ will depend on the properties of the gas, and on the speed v of the Brownian particle. Suppose we Taylor expand this quantity in terms of the particle speed v:

$$\langle F(t)\rangle = c_1 + c_2 v + c_3 v^2 + \ldots \tag{B.3}$$

When $v = 0$, we want $\langle F\rangle = 0$, since there is then no preferred direction; thus, $c_1 = 0$. Let us then keep only the next term in (B.3). Because we expect this term to slow the particle, we introduce the minus sign explicitly through the introduction of a new constant ν such that $c_2 = -\nu$; our force equation (B.1) now reads

$$\boxed{\frac{dv(t)}{dt} = -\nu v(t) + \delta F(t)} \tag{B.4}$$

which is the famous *Langevin equation* (Refs. [3] to [7]).

The constant ν in (B.4) represents dissipation. This can be seen by taking the ensemble average of (B.4)

$$\frac{d}{dt}\langle v(t)\rangle = -\nu\langle v(t)\rangle \tag{B.5}$$

so that

$$\langle v(t)\rangle = v_0 e^{-\nu t} \tag{B.6}$$

(Recall that each realization of the ensemble has initial speed v_0). Thus, the characteristic slowing down time is ν^{-1}, and since the slowing down means a decrease in kinetic energy, ν represents dissipation.

Let us next investigate some of the statistical properties of (B.4). This equation is a linear inhomogeneous first order ordinary differential equation and thus is easy to solve. We have

$$\frac{dv(t)}{dt} + \nu v = \delta F(t) \tag{B.7}$$

Multiplying each side by $e^{\nu t}$ we have

$$\frac{d}{dt}\left[v(t)e^{\nu t}\right] = e^{\nu t}\,\delta F(t) \tag{B.8}$$

Thus

$$v(t)e^{\nu t} = v_0 + \int_0^t dt'\,\delta F(t')e^{\nu t'} \tag{B.9}$$

or

$$\boxed{v(t) = v_0\, e^{-\nu t} + e^{-\nu t}\int_0^t dt'\,\delta F(t')e^{\nu t'}} \tag{B.10}$$

The ensemble average of this equation reproduces (B.6),

$$\langle v(t)\rangle = v_0 e^{-\nu t} \tag{B.6}$$

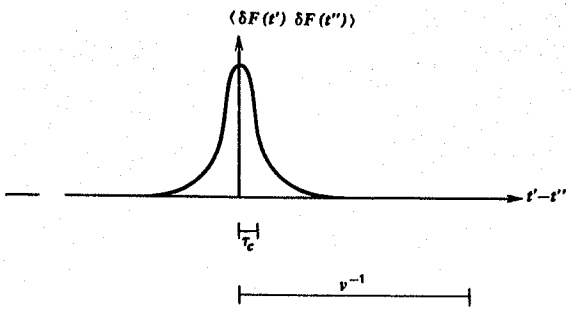

Fig. B.3 Autocorrelation function of the fluctuating force $\delta F(t)$, indicating the relative time scales τ_c and ν^{-1}.

Next, we square the velocity and ensemble average. Using (B.10), we have

$$\langle v^2(t)\rangle = \langle [v_0\, e^{-\nu t} + e^{-\nu t}\int_0^t dt'\, \delta F(t') e^{\nu t'}]$$

$$\times\, [v_0 e^{-\nu t} + e^{-\nu t}\int_0^t dt''\, \delta F(t'') e^{\nu t''}]\rangle$$

$$= v_0^2 e^{-2\nu t} + e^{-2\nu t}\int_0^t dt'\, e^{\nu t'}\int_0^t dt''\, \langle \delta F(t')\delta F(t'')\rangle e^{\nu t''} \qquad \text{(B.11)}$$

where two terms have disappeared in the ensemble average.

We now make the important assumption that δF is only correlated with itself over a time τ_c extremely short compared to the characteristic dissipation time ν^{-1} (Fig. B.3). We furthermore assume that δF is a stationary process, so that $\langle \delta F(t')\delta F(t'')\rangle$ is only a function of the time difference $t' - t''$. The correlation time τ_c is roughly the time of one molecular collision.

We are interested in the integral

$$I \equiv e^{-2\nu t}\int_0^t dt'\, e^{\nu t'}\int_0^t dt''\, \langle \delta F(t')\delta F(t'')\rangle e^{\nu t''} \qquad \text{(B.12)}$$

The above arguments indicate that the integrand is only important (nonzero) for $t' \approx t''$, as shown in Fig. B.4. With the change of variable $y = t' - t''$, $dy = -dt''$, (B.12) becomes

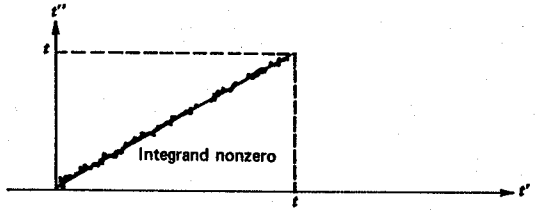

Fig. B.4 Region of the t'-t'' plane that contains a substantial contribution to the integral in (B.12).

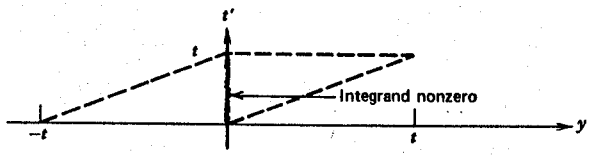

Fig. B.5 Region of the t'-y plane that contains a substantial contribution to the integral in (B.13).

$$I = e^{-2\nu t}\int_0^t dt'\, e^{\nu t'}\int_{t'-t}^{t'} dy\, e^{\nu t' - \nu y}\langle \delta F(t')\delta F(t' - y)\rangle \qquad \text{(B.13)}$$

By stationarity, we can write

$$\langle \delta F(t')\delta F(t' - y)\rangle = \langle \delta F(0)\delta F(-y)\rangle$$

$$= \langle \delta F(0)\delta F(y)\rangle \qquad \text{(B.14)}$$

where the last equality is due to the evenness of the correlation function. The integral in (B.13) is now substantial in the region shown in Fig. B.5, where $y = t' - t'' \approx 0$. Since the integrand is only important near $y \approx 0$, we can replace the upper limit of y-integration by $+\infty$ and the lower limit of y-integration by $-\infty$. Then (B.13) becomes

$$I = e^{-2\nu t}\int_0^t dt'\, e^{2\nu t'}\int_{-\infty}^{\infty} dy\, \langle \delta F(0)\delta F(y)\rangle \qquad \text{(B.15)}$$

where we have discarded the factor $e^{-\nu y}$ that is unity when $y \approx 0$ where the integrand is important. The t' integration can now be performed,

$$I = \frac{1}{2\nu}(1 - e^{-2\nu t})\int_{-\infty}^{\infty} dy\, \langle \delta F(0)\delta F(y)\rangle \qquad \text{(B.16)}$$

so that the full equation (B.11) now reads

$$\boxed{\langle v^2(t)\rangle = v_0^2 e^{-2\nu t} + \frac{1}{2\nu}(1 - e^{-2\nu t})\int_{-\infty}^{\infty} dy\, \langle \delta F(0)\delta F(y)\rangle} \qquad \text{(B.17)}$$

If we allow the time to become very large compared to the dissipation time ν^{-1}, then we obtain an expression for the thermal fluctuations of v^2,

$$\boxed{\langle v^2(t)\rangle \xrightarrow{t\to\infty} \frac{1}{2\nu}\int_{-\infty}^{\infty} dy\, \langle \delta F(0)\delta F(y)\rangle} \qquad \text{(B.18)}$$

However, we know from elementary thermodynamics that in thermal equilibrium, the Brownian particle will have $\frac{1}{2}T$ of kinetic energy per degree of freedom (Boltzmann's constant is as usual absorbed into the temperature T). Thus, elementary thermodynamics predicts

$$\tfrac{1}{2}M\langle v^2(t)\rangle = \tfrac{1}{2}T \qquad \text{(B.19)}$$

or

$$\langle v^2(t)\rangle = \frac{T}{M} \tag{B.20}$$

Equating (B.18) and (B.20) we have

$$\frac{T}{M} = \frac{1}{2\nu}\int_{-\infty}^{\infty} dy\, \langle \delta F(0)\delta F(y)\rangle \tag{B.21}$$

or

$$\boxed{\nu = \frac{M}{2T}\int_{-\infty}^{\infty} dy\, \langle \delta F(0)\delta F(y)\rangle} \tag{B.22}$$

which is the *fluctuation–dissipation* theorem.

Equation (B.22) expresses the amazing fact that the dissipation of a Brownian particle is directly related to the correlation function $\langle \delta F(0)\delta F(y)\rangle$ of the fluctuating force $F(t) = \langle F(t)\rangle + \delta F(t)$ whose ensemble average $\langle F\rangle$ produces the dissipation. This is a fundamental result of physics that applies in many situations; in the theory of electric circuits it is known as *Nyquist's theorem*.

This concludes our discussion of the Langevin equation and the fluctuation–dissipation theorem. In the next section, we shall consider the related topic of *Markov processes* and derive the *Fokker–Planck equation*.

B.2 MARKOV PROCESSES AND FOKKER–PLANCK EQUATION

In the previous section, we considered the behavior of a Brownian particle and derived the Langevin equation together with a fluctuation–dissipation theorem. In this section, we show how the behavior of a Brownian particle can be described by a Fokker–Planck equation. The Fokker–Planck equation is a very general equation in physics; it describes not only Brownian particles, but any phenomenon that in some approximate sense can be thought of as a *Markov process*.

A *Markov process* is one whose value at the next measuring time depends only on its value at the present measuring time, and not on any previous measuring time. Thus, if $x(t)$ is the random process, and $x_n \equiv x(t_n)$, with $t_n > t_{n-1} > \ldots > t_1 > t_0$, a Markov process has a probability density such that

$$\rho(x_n|x_{n-1}\, x_{n-2}\ldots x_1 x_0) = \rho(x_n|x_{n-1}) \tag{B.23}$$

where the notation $\rho(a|b)$ means "the probability density of a given that b was true." Thus, for a Markov process, the probability that $x_n = 5$ depends only on what the value of x_{n-1} was; it does not depend on what the values of x_{n-2}, x_{n-3}, etc. were.

There are both *discrete* and *continuous* Markov processes. An example of a discrete Markov process is given by flipping a coin. A trivial example comes if we give each toss a value $x(t_n) \equiv x_n = +1$ for a toss of "heads" and a value $x_n = -1$ for a toss of "tails." Then x is clearly a Markov process, since $\rho(x_n) = \frac{1}{2}\delta(x_n - 1) + \frac{1}{2}\delta(x_n + 1)$ does not depend on x_{n-1}, much less on x_{n-2}, x_{n-3}, etc.

A better example of a discrete Markov process is given by defining the random variable

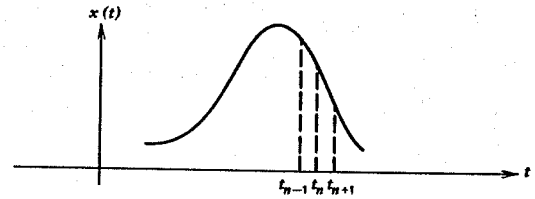

Fig. B.6 Any function in nature can be drawn as a smooth curve as shown.

$$X(t_n) \equiv X_n \equiv \sum_{i=1}^{n} x_i \tag{B.24}$$

where the x_i are given by the coin tosses of the previous paragraph. Now X is clearly a Markov process, whose probability density at t_n very definitely depends on the value of X_{n-1}, but on no previous value.

EXERCISE Calculate $\rho(X_n|X_{n-1})$ for this example.

To give an example of a continuous Markov process is more difficult, because a continuous Markov process cannot exist in nature. To see this, consider any random function that we can draw as a smooth curve, as in Fig. B.6. Now, on the time scale shown, it appears that x_{n+1} not only depends on x_n, but also on x_{n-1}. That is, x_{n+1} not only depends on x_n, but also on the derivative of the function $dx(t)/dt|_{t=t_n}$, which can be written

$$\left.\frac{dx(t)}{dt}\right|_{t=t_n} = \frac{x_n - x_{n-1}}{\Delta t} \tag{B.25}$$

Thus, this function is not a Markov process. In fact, no function that is a continuous curve and, therefore no physical function, can be a Markov process.

This does not mean that Markov processes cannot be a good approximation to a physical process. Consider the velocity function of the Brownian particle in the previous section (Fig. B.7). We have seen that the velocity consists of a rapid fluctuation due to each molecular collision, together with a slowing down or net friction force. Thus, on the time scale of molecular collisions, the process is not Markovian. However, on the much longer time scale of many collision times, the

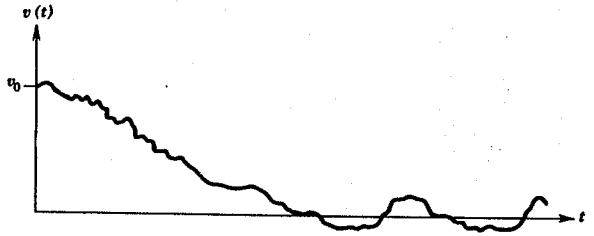

Fig. B.7 One realization of the velocity of a Brownian particle in a particular direction.

situation is very nearly Markovian. The Brownian particle is performing a random walk in velocity space, and soon forgets the details of its orbit near $t = 0$; it does, however, remember its velocity v_0 at $t = 0$.

Thus, we consider the process to have three time scales (Fig. B.8): the collision time τ_c, which is the autocorrelation time of the force $\delta F(t)$ in the Langevin equation; the time Δt after which we may assume to good approximation that the process is Markovian; and the dissipation time ν^{-1}. We must have $\Delta t >> \tau_c$; we shall further assume in this section that $\Delta t << \nu^{-1}$.

Let us develop some of the mathematical properties of Markov processes. This development will lead us to the Fokker–Planck equation.

Consider the probability of a sequence of values of the random function $x(t)$. This is

$$\rho(x_n,x_{n-1},\ldots,x_2,x_1,x_0) \equiv \{\text{probability that, at time } t_0, \text{ the process } x(t) \text{ has the value } x_0 \textit{ and} \text{ at time } t_1, x(t) \text{ has the value } x_1, \textit{ and} \ldots \textit{ and} \text{ at time } t_n, x(t) \text{ has the value } x_n\} \text{ where } t_n > t_{n-1} > t_{n-2} \ldots > t_1 > t_0 \tag{B.26}$$

By the definition (B.23) we can write

$$\rho(x_n,x_{n-1},\ldots,x_0) = \rho(x_n|x_{n-1},x_{n-2},\ldots,x_0)$$

$$\times \rho(x_{n-1},x_{n-2},\ldots,x_0) = \rho(x_n|x_{n-1})\rho(x_{n-1},x_{n-2},\ldots,x_0) \tag{B.27}$$

The same procedure can now be applied to the last factor on the right of (B.27), so that

$$\rho(x_{n-1},x_{n-2},\ldots,x_0) = \rho(x_{n-1}|x_{n-2})\rho(x_{n-2},\ldots,x_0) \tag{B.28}$$

and so on until we have finally, for a Markov process,

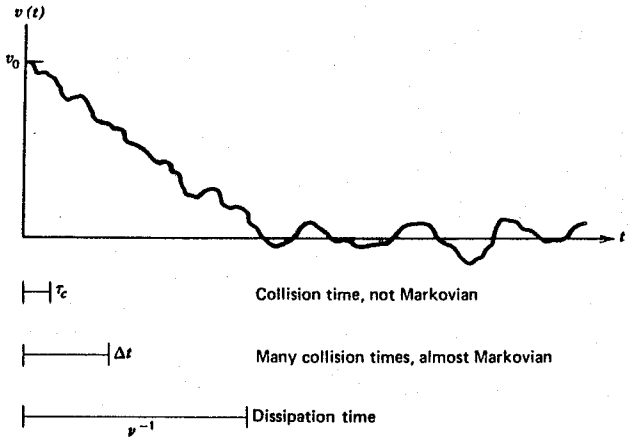

Fig. B.8 Three time scales of Brownian motion.

$$\boxed{\rho(x_n,x_{n-1},x_{n-2},\ldots,x_0) = \rho(x_n|x_{n-1})\rho(x_{n-1}|x_{n-2}) \ldots \rho(x_2|x_1)\rho(x_1|x_0)\rho(x_0)} \tag{B.29}$$

By elementary considerations it must also be true that

$$\rho(x_n,x_{n-1},x_{n-2},\ldots,x_1,x_0) = \rho(x_n,x_{n-1},\ldots,x_1|x_0)\rho(x_0) \tag{B.30}$$

Comparing (B.30) and (B.29) we find

$$\rho(x_n,x_{n-1},\ldots,x_1|x_0)$$

$$= \rho(x_n|x_{n-1})\rho(x_{n-1}|x_{n-2}) \ldots \rho(x_1|x_0) \tag{B.31}$$

In particular, we can choose $n = 2$ to obtain

$$\rho(x_2,x_1|x_0) = \rho(x_2|x_1)\rho(x_1|x_0) \tag{B.32}$$

Let us now integrate this expression over all possible x_1 to obtain

$$\rho(x_2|x_0) = \int dx_1\, \rho(x_2,x_1|x_0) \tag{B.33}$$

or

$$\boxed{\rho(x_2|x_0) = \int dx_1\, \rho(x_2|x_1)\rho(x_1|x_0)} \tag{B.34}$$

which is the *Chapman–Kolmogorov equation*, or *Smoluchowsky equation* [8].

Suppose we identify x_1 with time t and x_2 as $x(t + \Delta t)$. Suppose we further assume that

$$\rho(x_0) \equiv \rho(x,t = t_0) = \delta(x - x_0) \tag{B.35}$$

Then we can drop the references to x_0 in (B.34), and write

$$\rho(x_2|x_0) = \rho(x,t + \Delta t) \tag{B.36}$$

that is, x_2 is now denoted by x, and

$$\rho(x_1|x_0) = \rho(x_1,t) \tag{B.37}$$

We can also change the notation of $\rho(x_2|x_1)$; with the definition

$$\Delta x \equiv x - x_1 \tag{B.38}$$

we can write

$$\rho(x_2|x_1) = \rho(x,t + \Delta t|x - \Delta x,t)$$

$$= \psi(\Delta x,t + \Delta t|x - \Delta x,t) \tag{B.39}$$

where the transition probability ψ is defined by (B.39); ψ gives the probability that at time $t + \Delta t$, the random process has made a jump of Δx from its previous value $x - \Delta x$ at time t.

With these notational changes, we can rewrite (B.34) as

$$\rho(x,t + \Delta t) = \int d(\Delta x)\psi(\Delta x,t + \Delta t|x - \Delta x,t)\rho(x - \Delta x,t) \tag{B.40}$$

The value x appears on the right of (B.40) only in the combination $x - \Delta x$. Thus, if we assume that *all of the important physics happens for small* Δx, then we can make a Taylor series expansion on the right of (B.40), obtaining

$$\rho(x,t + \Delta t) = \int d(\Delta x) \sum_{l=0}^{\infty} \frac{(-\Delta x)^l}{l!} \times \left\{ \frac{\partial^l}{\partial x^l} \left[\psi(\Delta x, t + \Delta t | x - \Delta x, t) \rho(x - \Delta x, t) \right]_{x-\Delta x=x} \right\}$$

or

$$\rho(x,t + \Delta t) = \int d(\Delta x) \sum_{l=0}^{\infty} \frac{(-\Delta x)^l}{l!} \frac{\partial^l}{\partial x^l} \times [\psi(\Delta x, t + \Delta t | x,t) \rho(x,t)] \tag{B.41}$$

If the infinite sum converges, and if we can interchange the summation and integration, then we can write

$$\rho(x,t + \Delta t) = \sum_{l=0}^{\infty} \frac{(-1)^l}{l!} \frac{\partial^l}{\partial x^l} \times [\rho(x,t) \int d(\Delta x)(\Delta x)^l \psi(\Delta x, t + \Delta t | x,t)] \tag{B.42}$$

The quantity given by the Δx integration is just the expectation value or ensemble average of $(\Delta x)^l$,

$$\langle (\Delta x)^l \rangle \equiv \int d(\Delta x)(\Delta x)^l \psi(\Delta x, t + \Delta t | x,t) \tag{B.43}$$

which is itself a function of x,t through ψ. Equation (B.42) becomes

$$\rho(x,t + \Delta t) = \sum_{l=0}^{\infty} \frac{(-1)^l}{l!} \frac{\partial^l}{\partial x^l} [\rho(x,t) \langle (\Delta x)^l \rangle (x,t)] \tag{B.44}$$

Moving the $l = 0$ term to the left side, and dividing by Δt, we have

$$\frac{\rho(x,t + \Delta t) - \rho(x,t)}{\Delta t} = \sum_{l=1}^{\infty} \frac{(-1)^l}{l! \Delta t} \frac{\partial^l}{\partial x^l} [\rho(x,t) \langle (\Delta x)^l \rangle (x,t)] \tag{B.45}$$

We next take the limit as $\Delta t \to$ "0". This means that we let Δt become very small, much smaller than any macroscopic time scale (e.g., ν^{-1}). However, Δt cannot really go to zero, because this development has assumed that Δt is large enough to justify the Markovian assumption. Thus, the left side of (B.45) becomes

$$\lim_{\Delta t \to \text{"0"}} \frac{\rho(x,t + \Delta t) - \rho(x,t)}{\Delta t} = \frac{\partial \rho(x,t)}{\partial t} \tag{B.46}$$

where the time derivative refers to macroscopic time. Equation (B.45) becomes

$$\frac{\partial \rho}{\partial t} = \sum_{l=1}^{\infty} (-1)^l \frac{\partial^l}{\partial x^l} \left[\lim_{\Delta t \to \text{"0"}} \frac{\langle (\Delta x)^l \rangle}{l! \Delta t} \rho(x,t) \right] \tag{B.47}$$

Defining the diffusion coefficients

$$D^{(l)}(x,t) \equiv \lim_{\Delta t \to \text{"0"}} \frac{\langle (\Delta x)^l \rangle}{l! \Delta t} \tag{B.48}$$

Equation (B.47) is

$$\frac{\partial \rho(x,t)}{\partial t} = \sum_{l=1}^{\infty} (-1)^l \frac{\partial^l}{\partial x^l} [D^{(l)}(x,t) \rho(x,t)] \tag{B.49}$$

If we keep only the first two terms on the right of (B.49), we have

$$\boxed{\frac{\partial \rho(x,t)}{\partial t} = -\frac{\partial}{\partial x} [D^{(1)}(x,t) \rho(x,t)] + \frac{\partial^2}{\partial x^2} [D^{(2)}(x,t) \rho(x,t)]} \tag{B.50}$$

which is the well-known *Fokker–Planck equation* [9].

For Brownian motion, the random variable x is replaced by the particle velocity $v(t)$. We shall leave it as an exercise to determine the diffusion coefficients $D^{(1)}(v,t)$ and $D^{(2)}(v,t)$.

EXERCISE Use the results of the previous section to evaluate the coefficients in Eq. (B.50) for Brownian motion. Show that $D^{(1)}(v,t) = -\nu v$, and $D^{(2)}(v,t) = \nu T/M$, so that the Fokker–Planck equation associated with the Langevin equation of Brownian motion is

$$\boxed{\frac{\partial \rho(v,t)}{\partial t} = \nu \frac{\partial}{\partial v} (v\rho) + \frac{\nu T}{M} \frac{\partial^2}{\partial v^2} \rho} \tag{B.51}$$

EXERCISE Use the results of the previous section to show that $D^{(3)}(v,t) \sim \Delta t$ and, thus, vanishes as $\Delta t \to$ "0".

We can now understand why we are able to write the Lenard–Balescu equation in the form of a Fokker–Planck equation,

$$\frac{\partial f(\mathbf{v}_1,t)}{\partial t} = -\nabla_{\mathbf{v}_1} \cdot (\mathbf{A} f) + \frac{1}{2} \nabla_{\mathbf{v}_1} \nabla_{\mathbf{v}_1} : (\overleftrightarrow{\mathbf{B}} f) \tag{B.52}$$

Because the derivation of Lenard–Balescu assumed $g(1,2) \ll f_1(1) f_1(2)$, we have effectively limited ourselves to small angle two-body collisions. The quantity $f(\mathbf{v}_1,t)$ may be thought of as the probability density of particles in velocity space. Thus, $f(\mathbf{v}_1,t)$ is changing slowly on the time scale for a two-body collision. All of these features are precisely those assumed in the derivation of the Fokker–Planck equation. It should come as no surprise to us that the Lenard–Balescu equation can be written in the form of the Fokker–Planck equation. The coefficient A in (B.52) is called the coefficient of *dynamic friction*, and plays the same role as νv in the Fokker–Planck equation (B.51) for Brownian motion. It represents the slowing