Contents

Part II

STATISTICAL PHYSICS

Statistical Physics

Version 1203.1.K, 23 September 2012

In this second Part of the book, we shall study aspects of classical statistical physics that every physicist should know but that are not usually treated in elementary thermodynamics courses. Our study will lay the microphysical (particle-scale) foundations for the continuum physics of Parts III—VII, and it will elucidate the intimate connections between relativistic statistical physics and the Newtonian theory, and between quantum statistical physics and the classical theory. Our treatment will be both Newtonian and relativistic. Readers who prefer a fully Newtonian treatment can skip the (rather few) relativistic sections. Throughout, we presume that readers are familiar with elementary thermodynamics, but not with other aspects of statistical physics.

In Chap. 3, we will study *kinetic theory* — the simplest of all formalisms for analyzing systems of huge numbers of particles (e.g., molecules of air, or neutrons diffusing through a nuclear reactor, or photons produced in the big-bang origin of the Universe). In kinetic theory, the key concept is the "distribution function" or "number density of particles in phase space", $\mathcal{N}$; i.e., the number of particles of some species (e.g. electrons) per unit 3-dimensional volume of ordinary space and per unit 3-dimensional volume of momentum space. In special relativity, despite first appearances, this $\mathcal{N}$ turns out to be a geometric, reference-frame-independent entity (a scalar field in phase space). This $\mathcal{N}$ and the frame-independent laws it obeys provide us with a means for computing, from microphysics, the macroscopic quantities of continuum physics: mass density, thermal energy density, pressure, equations of state, thermal and electrical conductivities, viscosities, diffusion coefficients, ...

In Chap. 4, we will develop the foundations of *statistical mechanics*. Here our statistical study will be more sophisticated than in kinetic theory: we shall deal with "ensembles" of physical systems. Each ensemble is a (conceptual) collection of a huge number of physical systems that are identical in the sense that they all have the same degrees of freedom, but different in that their degrees of freedom may be in different physical states. For example, the systems in an ensemble might be balloons that are each filled with 10^{23} air molecules so each is describable by 3×10^{23} spatial coordinates (the x, y, z of all the molecules) and 3×10^{23} momentum coordinates (the p_x, p_y, p_z of all the molecules). The state of one of the balloons is fully described, then, by 6×10^{23} numbers. We introduce a distribution function $\mathcal{N}$ which is a function of these 6×10^{23} different coordinates, i.e., it is defined in a phase space with 6×10^{23} dimensions. This distribution function tells us how many systems (balloons) in our ensemble lie in a unit volume of that phase space. Using this distribution function we will study such issues as the statistical meaning of entropy, the relationship between entropy and information, the statistical origin of the second law of thermodynamics, the statistical meaning of "thermal equilibrium", and the evolution of ensembles into thermal equilibrium. Our applications will include derivations of the Fermi-Dirac distribution for fermions in thermal equilibrium and the Bose-Einstein distribution for bosons, a study of Bose-Einstein condensation in a dilute gas, and explorations of the meaning and role of entropy in gases, in black holes and in the universe as a whole.

In Chap. 5, we will use the tools of statistical mechanics to study *statistical thermodynamics*, i.e. ensembles of systems that are in or near thermal equilibrium (also called statistical equilibrium). Using statistical mechanics, we shall derive the laws of thermodynamics, and we shall learn how to use thermodynamic and statistical mechanical tools, hand in hand, to study not only equilibria, but also the probabilities for random, spontaneous fluctuations away from equilibrium. Among the applications we shall study are: (i) chemical and particle reactions such as ionization equilibrium in a hot gas, and electron-positron pair formation in a still hotter gas; and (ii) phase transitions, such as the freezing, melting, vaporization and condensation of water. We shall focus special attention on a Ferromagnetic phase transition in which the magnetic moments of atoms spontaneously align with each other as iron is cooled, using it to illustrate two elegant and powerful techniques of statistical physics: the renormalization group, and Monte Carlo methods.

In Chap. 6, we will develop the theory of *random processes* (a modern, mathematical aspect of which is the theory of stochastic differential equations). Here we shall study the dynamical evolution of processes that are influenced by a huge number of factors over which we have little control and little knowledge, except their statistical properties. One example is the Brownian motion of a dust particle being buffeted by air molecules; another is the motion of a pendulum used, e.g., in a gravitational-wave interferometer, when one monitors that motion so accurately that one can see the influences of seismic vibrations and of fluctuating "thermal" ("Nyquist") forces in the pendulum's suspension wire. The position of such a dust particle or pendulum cannot be predicted as a function of time, but one can compute the probability that it will evolve in a given manner. The theory of random processes is a theory of the evolution of the position's probability distribution (and the probability distribution for any other entity driven by random, fluctuating influences). Among the random-process concepts we shall study are spectral densities, correlation functions, the Fokker-Planck equation which governs the evolution of probability distributions, and the fluctuation-dissipation theorem which says that, associated with any kind of friction there must be fluctuating forces whose statistical properties are determined by the strength of the friction and the temperature of the entities that produce the friction.

The theory of random processes, as treated in Chap. 6, also includes the theory of signals and noise. At first sight this undeniably important topic, which lies at the heart of experimental and observational science, might seem outside the scope of this book. However, we shall discover that it is intimately connected to statistical physics and that similar principles to those used to describe, say, Brownian motion are appropriate when thinking about, for example, how to detect the electronic signal of a rare particle event against a strong and random background. We shall study, for example, techniques for extracting weak signals from noisy data by filtering the data, and the limits that noise places on the accuracies of physics experiments and on the reliability of communications channels.

Chapter 3

Kinetic Theory

Version 1203.1.K, 23 September 2012

Please send comments, suggestions, and errata via email to kip@caltech.edu, or on paper to Kip Thorne, 350-17 Caltech, Pasadena CA 91125

Box 3.1
Reader's Guide

- This chapter develops nonrelativistic (Newtonian) kinetic theory and also the relativistic theory. Sections and exercises labeled [N] are Newtonian and those labeled [R] are relativistic. The N material can be read without the R material, but the R material requires the N material as a foundation. The R material is all Track Two, **T2**.
- This chapter relies on the geometric viewpoint about physics developed in Chap. 1 for Newtonian physics and in Chap. 2 for relativistic physics. It especially relies on
  - Secs. 1.4 and 1.5.2 on Newtonian particle kinetics,
  - Secs. 2.4.1 and 2.6 on relativistic particle kinetics,
  - Sec. 1.8 for the Newtonian conservation laws for particles,
  - Sec. 2.6 for the relativistic number-flux 4-vector and its conservation law,
  - Sec. 1.9 on the Newtonian stress tensor and its role in conservation laws for momentum,
  - Sec. 2.13 on the relativistic stress-energy tensor and its role in conservation laws for 4-momentum,
  - and aspects of relativity theory that Newtonian readers will need, in Sec. 1.10.
- The Newtonian parts of this chapter are a crucial foundation for the remainder of Part II of this book (Statistical Physics), for small portions of Part V (Fluid Mechanics; especially equations of state, the origin of viscosity, and the diffusion of heat in fluids), and for half of Part VI (Plasma Physics: Chaps. 22 and 23).

3.1 Overview

In this chapter, we shall study kinetic theory, the simplest of all branches of statistical physics. Kinetic theory deals with the statistical distribution of a "gas" made from a huge number of "particles" that travel freely, without collisions, for distances (*mean free paths*) long compared to their sizes.

Examples of particles (*italicized*) and phenomena that can be studied via kinetic theory are these: (i) How *galaxies*, formed in the early universe, congregate into clusters as the universe expands. (ii) How spiral structure develops in the distribution of a galaxy's *stars.* (iii) How, deep inside a white-dwarf star, relativistic degeneracy influences the equation of state of the star's *electrons and protons.* (iv) How a supernova explosion affects the evolution of the density and temperature of *interstellar molecules.* (v) How anisotropies in the expansion of the universe affect the temperature distribution of the *cosmic microwave photons*—the remnants of the big-bang. (vi) How changes of a metal's temperature affect its thermal and electrical conductivity (with the heat and current carried by *electrons*). (vii) Whether *neutrons* in a nuclear reactor can survive long enough to maintain a nuclear chain reaction and keep the reactor hot.

Most of these applications involve particle speeds small compared to light and so can be studied with Newtonian theory, but some involve speeds near or at the speed of light and require relativity. Accordingly, we shall develop both versions of the theory, Newtonian and relativistic, and shall demonstrate that the Newtonian theory is the low-speed limit of the relativistic theory. As is discussed in the Reader's Guide (Box 3.1), the relativistic material is all Track Two and can be skipped by readers who are focusing on the (nonrelativistic) Newtonian theory.

We begin in Sec. 3.2 by introducing the concepts of momentum space, phase space (the union of physical space and momentum space), and the distribution function (number density of particles in phase space). We meet several different versions of the distribution function, all equivalent, but each designed to optimize conceptual thinking or computations in a particular arena (e.g., photons, plasma physics, and the interface with quantum theory). In Sec. 3.3, we study the distribution functions that characterize systems of particles in thermal equilibrium. There are three such equilibrium distributions: one for quantum mechanical particles with half-integral spin (fermions), another for quantum particles with integral spin (bosons), and a third for classical particles. As special applications, we derive the Maxwell velocity distribution for low-speed, classical particles (Ex. 3.4) and its high-speed relativistic analog (Ex. 3.5 and Fig. 3.6) and we compute the effects of observers' motions on their measurements of the cosmic microwave radiation created in our universe's big-bang origin (Ex. 3.6). In Sec. 3.4, we learn how to compute macroscopic, physical-space quantities (particle density and flux, energy density, stress tensor, stress-energy tensor, ...) by integrating over the momentum portion of phase space. In Sec. 3.5, we show that, if the distribution function is isotropic in momentum space, in some reference frame, then on macroscopic scales the particles constitute a perfect fluid; we use our momentum-space integrals to evaluate the equations of state of various kinds of perfect fluids: a nonrelativistic, hydrogen gas in both the classical, nondegenerate regime and the regime of electron degeneracy (Sec. 3.5.2), a relativistically degenerate gas (Sec. 3.5.4), and a photon gas (Sec. 3.5.5); and we use our results to discuss the physical nature of matter as a function of density and temperature (Fig. 3.7).

In Sec. 3.6, we study the evolution of the distribution function, as described by Liouville's theorem and by the associated collisionless Boltzmann equation when collisions between particles are unimportant, and by the Boltzmann transport equation when collisions are significant; and we use a simple variant of these evolution laws to study the heating of the Earth by the Sun, and the key role played by the Greenhouse effect (Ex. 3.14). Finally, in Sec. 3.7, we learn how to use the Boltzmann transport equation to compute the transport coefficients (diffusion coefficient, electrical conductivity, thermal conductivity, and viscosity) which describe the diffusive transport of particles, charge, energy, and momentum through a gas of particles that collide frequently; and we use the Boltzmann transport equation to study a chain reaction in a nuclear reactor (Ex. 3.20).

Readers who feel overwhelmed by the enormous amount and variety of applications in this chapter (and throughout this book) should remember the authors' goals: We want readers to learn the fundamental concepts of kinetic theory (and other topics in this book), and want them to meet a variety of applications so they will understand how the fundamental concepts are used. However, we do not expect readers to become expert in or even remember all these applications. To do so would require much more time and effort than most readers can afford or should expend.

3.2 Phase Space and Distribution Function

3.2.1 [N] Newtonian Number density in phase space, $\mathcal{N}$

In Newtonian, 3-dimensional space (*physical space*), consider a particle with rest mass m that moves along a path $\mathbf{x}(t)$ as universal time t passes (Fig. 3.1a). The particle's time-varying velocity and momentum are $\mathbf{v}(t) = d\mathbf{x}/dt$ and $\mathbf{p}(t) = m\mathbf{v}$. The path $\mathbf{x}(t)$ is a curve in the physical space, and the momentum $\mathbf{p}(t)$ is a time-varying, coordinate-independent vector in the physical space.

It is useful to introduce an auxiliary 3-dimensional space, called *momentum space*, in which we place the tail of $\mathbf{p}(t)$ at the origin. As time passes, the tip of $\mathbf{p}(t)$ sweeps out a curve in momentum space (Fig. 3.1b). This momentum space is "secondary" in the sense that it relies for its existence on the physical space of Fig. 3.1a. Any Cartesian coordinate system of physical space, in which the location $\mathbf{x}(t)$ of the particle has coordinates (x, y, z), induces in momentum space a corresponding coordinate system (p_x, p_y, p_z). The 3-dimensional physical space and 3-dimensional momentum space together constitute a 6-dimensional *phase space*, with coordinates (x, y, z, p_x, p_y, p_z).

In this chapter, we study a collection of a very large number of identical particles (all with the same rest mass m)[1]. As tools for this study, consider a tiny 3-dimensional volume $d\mathcal{V}_x$ centered on some location $\mathbf{x}$ in physical space, and a tiny 3-dimensional volume $d\mathcal{V}_p$ centered on location $\mathbf{p}$ in momentum space. Together these make up a tiny 6-dimensional

[1] In Ex. 3.2 and Box 3.2 we shall extend kinetic theory to particles with a range of rest masses.

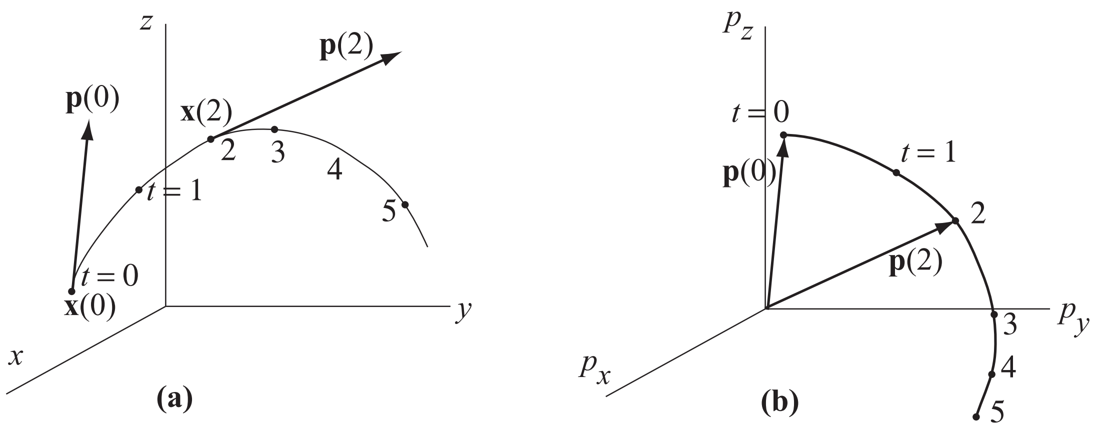

Fig. 3.1: (a) Euclidean physical space, in which a particle moves along a curve $\mathbf{x}(t)$ that is parameterized by universal time t, and in which the particle's momentum $\mathbf{p}(t)$ is a vector tangent to the curve. (b) Momentum space in which the particle's momentum vector $\mathbf{p}$ is placed, unchanged, with its tail at the origin. As time passes, the momentum's tip sweeps out the indicated curve $\mathbf{p}(t)$.

volume

$$d^2\mathcal{V} \equiv d\mathcal{V}_x d\mathcal{V}_p \ . \tag{3.1}$$

In any Cartesian coordinate system, we can think of $d\mathcal{V}_x$ as being a tiny cube located at (x, y, z) and having edge lengths dx, dy, dz; and similarly for $d\mathcal{V}_p$. Then, as computed in this coordinate system, these tiny volumes are

$$d\mathcal{V}_x = dx\, dy\, dz \ , \quad d\mathcal{V}_p = dp_x\, dp_y\, dp_z \ , \quad d^2\mathcal{V} = dx\, dy\, dz\, dp_x\, dp_y\, dp_z \ . \tag{3.2}$$

Denote by dN the number of particles (all with rest mass m) that reside inside $d^2\mathcal{V}$ in phase space (at some moment of time t). Stated more fully: dN is the number of particles that, at time t, are located in the 3-volume $d\mathcal{V}_x$ centered on the location $\mathbf{x}$ in physical space, and that also have momentum vectors whose tips at time t lie in the 3-volume $d\mathcal{V}_p$ centered on location $\mathbf{p}$ in momentum space. Denote by

$$\boxed{\mathcal{N}(\mathbf{x}, \mathbf{p}, t) \equiv \frac{dN}{d^2\mathcal{V}} = \frac{dN}{d\mathcal{V}_x\, d\mathcal{V}_p}} \tag{3.3}$$

the *number density of particles at location* $(\mathbf{x}, \mathbf{p})$ *in phase space at time* t. This is also called the *distribution function.*

This distribution function is kinetic theory's principal tool for describing any collection of a large number of identical particles.

In Newtonian theory, the volumes $d\mathcal{V}_x$ and $d\mathcal{V}_p$ occupied by our collection of dN particles are independent of the reference frame that we use to view them. Not so in relativity theory: $d\mathcal{V}_x$ undergoes a Lorentz contraction when one views it from a moving frame, and $d\mathcal{V}_p$ also changes; but (as we shall see in Sec. 3.2.2), their product $d^2\mathcal{V} = d\mathcal{V}_x d\mathcal{V}_p$ is the same in all frames. Therefore, in both Newtonian theory and relativity theory, the distribution function $\mathcal{N} = dN/d^2\mathcal{V}$ is independent of reference frame, and also, of course, independent of any choice of coordinates. It is a coordinate-independent scalar in phase space.

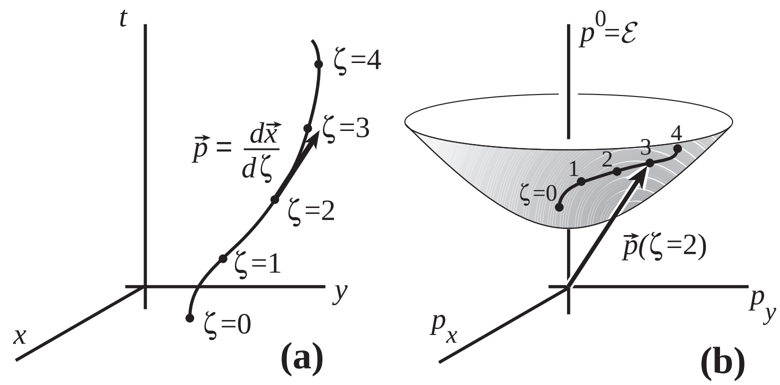

Fig. 3.2: (a) The world line $\vec{x}(\zeta)$ of a particle in spacetime (with one spatial coordinate, z, suppressed), parameterized by a parameter ζ that is related to the particle's 4-momentum by $\vec{p} = d\vec{x}/d\zeta$. (b) The trajectory of the particle in momentum space. The particle's momentum is confined to the mass hyperboloid, $\vec{p}^2 = -m^2$ (also known as the mass shell).

3.2.2 T2 [R] Relativistic Number Density in Phase Space, $\mathcal{N}$

Phase Space

We shall define the special relativistic distribution function in precisely the same way as the non-relativistic one, $\mathcal{N}(\mathbf{x}, \mathbf{p}, t) \equiv dN/d^2\mathcal{V} = dN/d\mathcal{V}_x d\mathcal{V}_p$, except that now $\mathbf{p}$ is the relativistic momentum, ($\mathbf{p} = m\mathbf{v}/\sqrt{1-v^2}$ if the particle has nonzero rest mass m). This definition of $\mathcal{N}$ appears, at first sight, to be frame-dependent, since the physical 3-volume $d\mathcal{V}_x$ and momentum 3-volume $d\mathcal{V}_p$ do not even exist until we have selected a specific reference frame. In other words, this definition appears to violate our insistence that relativistic physical quantities be described by frame-independent geometric objects that live in 4-dimensional spacetime. In fact, the distribution function defined in this way *is* frame-independent, though it does not look so. In order to elucidate this, we shall develop carefully and somewhat slowly the 4-dimensional spacetime ideas that underlie this relativistic distribution function:

Consider, as shown in Fig. 3.2a, a classical particle with rest mass m, moving through spacetime along a world line $\mathcal{P}(\zeta)$, or equivalently $\vec{x}(\zeta)$, where ζ is an affine parameter related to the particle's 4-momentum by

$$\vec{p} = d\vec{x}/d\zeta \tag{3.4a}$$

[Eq. (2.10)]. If the particle has non-zero rest mass, then its 4-velocity $\vec{u}$ and proper time τ are related to its 4-momentum and affine parameter by

$$\vec{p} = m\vec{u}\,, \quad \zeta = \tau/m \tag{3.4b}$$

[Eqs. (2.10) and (2.11)], and we can parameterize the world line by either τ or ζ. If the particle has zero rest mass, then its world line is null and τ does not change along it, so we have no choice but to use ζ as the world line's parameter.

The particle can be thought of not only as living in four-dimensional spacetime (Fig. 3.2a), but also as living in a four-dimensional momentum space (Fig. 3.2b). Momentum space, like spacetime, is a geometric, coordinate-independent concept: each point in momentum space corresponds to a specific 4-momentum $\vec{p}$. The tail of the vector $\vec{p}$ sits at the origin of momentum space and its tip sits at the point representing $\vec{p}$. The momentum-space diagram drawn in Fig. 3.2b has as its coordinate axes the components (p^0, $p^1 = p_1 \equiv p_x$, $p^2 = p_2 \equiv p_y$, $p^3 = p_3 \equiv p_z$) of the 4-momentum as measured in some *arbitrary* inertial frame. Because the squared length of the 4-momentum is always $-m^2$,

$$\vec{p} \cdot \vec{p} = -(p^0)^2 + (p_x)^2 + (p_y)^2 + (p_z)^2 = -m^2 \ , \tag{3.4c}$$

the particle's 4-momentum (the tip of the 4-vector $\vec{p}$) is confined to a hyperboloid in momentum space. This *mass hyperboloid* requires no coordinates for its existence; it is the frame-independent set of points, in momentum space, for which $\vec{p} \cdot \vec{p} = -m^2$. If the particle has zero rest mass, then $\vec{p}$ is null and the mass hyperboloid is a cone with vertex at the origin of momentum space. As in Chap. 2, we shall often denote the particle's energy p^0 by

$$\mathcal{E} \equiv p^0 \tag{3.4d}$$

(with the $\mathcal{E}$ in script font to distinguish it from the energy $E = \mathcal{E} - m$ with rest mass removed and its nonrelativistic limit $E = \frac{1}{2}mv^2$), and we shall embody the particle's spatial momentum in the 3-vector $\mathbf{p} = p_x\mathbf{e}_x + p_y\mathbf{e}_y + p_z\mathbf{e}_z$, and therefore shall rewrite the mass-hyperboloid relation (3.4c) as

$$\mathcal{E}^2 = m^2 + |\mathbf{p}|^2 \ . \tag{3.4e}$$

If no forces act on the particle, then its momentum is conserved and its location in momentum space remains fixed. A force (e.g., due to an electromagnetic field) will push the particle's 4-momentum along some curve in momentum space that lies on the mass hyperboloid. If we parameterize that curve by the same parameter ζ as we use in spacetime, then the particle's trajectory in momentum space can be written abstractly as $\vec{p}(\zeta)$. Such a trajectory is shown in Fig. 3.2b.

Because the mass hyperboloid is three dimensional, we can characterize the particle's location on it by just three coordinates rather than four. We shall typically use as those coordinates the spatial components of the particle's 4-momentum, (p_x, p_y, p_z) or the spatial momentum vector $\mathbf{p}$ as measured in some specific (but usually arbitrary) inertial frame.

Momentum space and spacetime, taken together, constitute the relativistic *phase space*. We can regard phase space as eight dimensional (four spacetime dimensions plus four momentum space dimensions). Alternatively, if we think of the 4-momentum as confined to the three-dimensional mass hyperboloid, then we can regard phase space as seven dimensional. This 7 or 8 dimensional phase space, by contrast with the non-relativistic 6-dimensional phase space, is frame-independent. No coordinates or reference frame are actually needed to define spacetime and explore its properties, and none are needed to define and explore 4-momentum space or the mass hyperboloid — though inertial (Lorentz) coordinates are often helpful in practical situations.

Volumes in Phase Space and Distribution Function

Turn attention, now, from an individual particle to a collection of a huge number of identical particles, each with the same rest mass m, and allow m to be finite or zero, it does not matter which. Examine those particles that pass close to a specific event $\mathcal{P}$ (also denoted $\vec{x}$) in spacetime; and *examine them from the viewpoint of a specific observer, who lives in a*

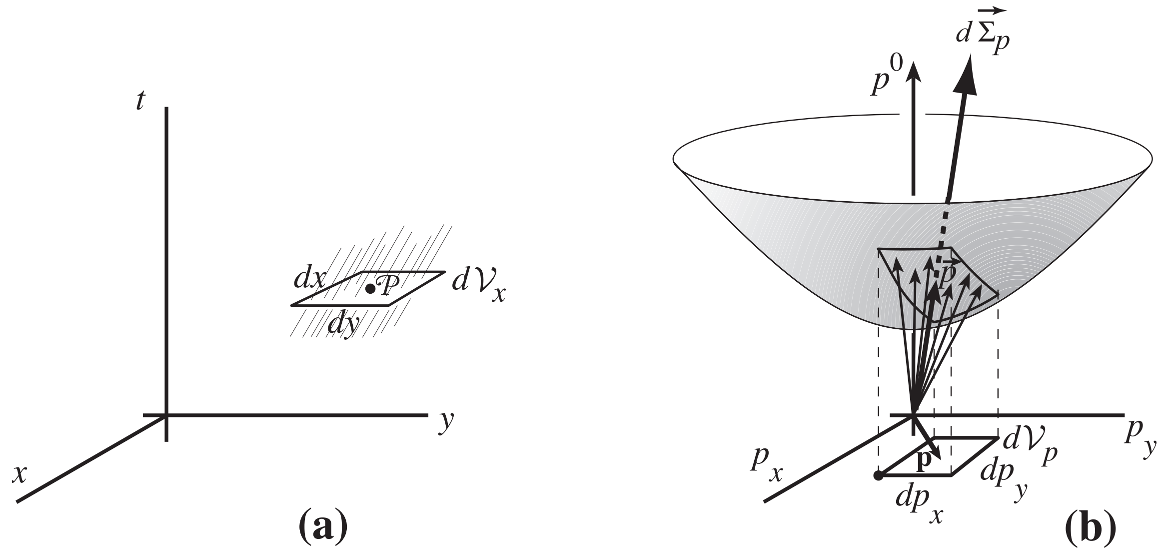

Fig. 3.3: Definition of the distribution function from the viewpoint of a specific observer in a specific inertial reference frame, whose coordinate axes are used in these drawings: At the event $\mathcal{P}$ [denoted by the dot in drawing a], the observer selects a 3-volume $d\mathcal{V}_x$, and she focuses on the set $\mathcal{S}$ of particles that lie in $d\mathcal{V}_x$ and have momenta lying in a region of the mass hyperboloid that is centered on $\vec{p}$ and has 3-momentum volume $d\mathcal{V}_p$ [drawing (b)]. If dN is the number of particles in that set $\mathcal{S}$, then $\mathcal{N}(\mathcal{P}, \vec{p}) \equiv dN/d\mathcal{V}_x d\mathcal{V}_p$.

specific inertial reference frame. Figure 3.3a is a spacetime diagram drawn in that observer's frame. As seen in that frame, the event $\mathcal{P}$ occurs at time t and at spatial location (x, y, z).

We ask the observer, at the time t of the chosen event, to define the distribution function $\mathcal{N}$ in identically the same way as in Newtonian theory, except that $\mathbf{p}$ is the relativistic spatial momentum $\mathbf{p} = m\mathbf{v}/\sqrt{1 - v^2}$ instead of the nonrelativistic $\mathbf{p} = m\mathbf{v}$. Specifically, the observer, in her inertial frame, chooses a tiny 3-volume

$$d\mathcal{V}_x = dxdydz \tag{3.5a}$$

centered on location $\mathbf{x}$ (little horizontal rectangle shown in drawing 3.3a) and a tiny 3-volume

$$d\mathcal{V}_p = dp_x dp_y dp_z \tag{3.5b}$$

centered on $\mathbf{p}$ in momentum space (little rectangle in the p_x-p_y plane in drawing 3.3b). Ask the observer to focus on that set $\mathcal{S}$ of particles which lie in $d\mathcal{V}_x$ and have spatial momenta in $d\mathcal{V}_p$ (Fig. 3.3). If there are dN particles in this set $\mathcal{S}$, then the observer will identify

$$\mathcal{N} \equiv \frac{dN}{d\mathcal{V}_x d\mathcal{V}_p} \equiv \frac{dN}{d^2\mathcal{V}} \tag{3.6}$$

as the *number density of particles in phase space.*

Notice in drawing 3.3b that the *4-momenta* of the particles in $\mathcal{S}$ have their tails at the origin of momentum space (as by definition do all 4-momenta), and have their tips in a tiny rectangular box on the mass hyperboloid — a box centered on the 4-momentum

$\vec{p}$ whose spatial part is $\mathbf{p}$ and temporal part is $p^0 = \mathcal{E} = \sqrt{m^2 + \mathbf{p}^2}$. The momentum volume element $d\mathcal{V}_p$ is the projection of that mass-hyperboloid box onto the horizontal (p_x, p_y, p_z) plane in momentum space. [The mass-hyperboloid box itself can be thought of as a (frame-independent) vectorial 3-volume $d\vec{\Sigma}_p$ — the momentum-space version of the vectorial 3-volume introduced in Sec. 2.12.1; see below.]

The number density $\mathcal{N}$ depends on the location $\mathcal{P}$ in spacetime of the 3-volume $d\mathcal{V}_x$ and on the 4-momentum $\vec{p}$ about which the momentum volume *on the mass hyperboloid* is centered: $\mathcal{N} = \mathcal{N}(\mathcal{P}, \vec{p})$. From the chosen observer's viewpoint, it can be regarded as a function of time t and spatial location $\mathbf{x}$ (the coordinates of $\mathcal{P}$) and of spatial momentum $\mathbf{p}$.

At first sight, one might expect $\mathcal{N}$ to depend also on the inertial reference frame used in its definition, i.e., on the 4-velocity of the observer. If this were the case, i.e., if $\mathcal{N}$ at fixed $\mathcal{P}$ and $\vec{p}$ were different when computed by the above prescription using different inertial frames, then we would feel compelled to seek some other object—one that is frame-independent—to serve as our foundation for kinetic theory. This is because the principle of relativity insists that all fundamental physical laws should be expressible in frame-independent language.

Fortunately, the distribution function (3.6) is frame-independent by itself, i.e. it is a frame-independent scalar field in phase space, so we need seek no further.

Proof of Frame Independence of $\mathcal{N} = dN/d^2\mathcal{V}$:

To prove the frame independence of $\mathcal{N}$, we shall consider, first, the frame dependence of the spatial 3-volume $d\mathcal{V}_x$, then the frame dependence of the momentum 3-volume $d\mathcal{V}_p$, and finally the frame dependence of their product $d^2\mathcal{V} = d\mathcal{V}_x d\mathcal{V}_p$ and thence of the distribution function $\mathcal{N} = dN/d^2\mathcal{V}$.

The thing that identifies the 3-volume $d\mathcal{V}_x$ and 3-momentum $d\mathcal{V}_p$ is the set of particles $\mathcal{S}$. We shall select that set once and for all and hold it fixed, and correspondingly the number of particles dN in the set will be fixed. Moreover, we shall assume that the particles' rest mass m is nonzero and shall deal with the zero-rest-mass case at the end by taking the limit $m \to 0$. Then there is a preferred frame in which to observe the particles $\mathcal{S}$: their own rest frame, which we shall identify by a prime.

In their rest frame and at a chosen event $\mathcal{P}$, the particles $\mathcal{S}$ occupy the interior of some box with imaginary walls that has some 3-volume $d\mathcal{V}_{x'}$. As seen in some other "laboratory" frame, their box has a Lorentz-contracted volume $d\mathcal{V}_x = \sqrt{1 - v^2}\, d\mathcal{V}_{x'}$. Here v is their speed as seen in the laboratory frame. The Lorentz-contraction factor is related to the particles' energy, as measured in the laboratory frame, by $\sqrt{1 - v^2} = m/\mathcal{E}$, and therefore $\mathcal{E} d\mathcal{V}_x = m d\mathcal{V}_{x'}$. The right-hand side is a frame-independent constant m times a well-defined number that everyone can agree on: the particles' rest-frame volume $d\mathcal{V}_{x'}$; i.e.

$$\boxed{\mathcal{E} d\mathcal{V}_x = (\text{a frame-independent quantity})} \; . \tag{3.7a}$$

Thus, the spatial volume $d\mathcal{V}_x$ occupied by the particles is frame-dependent, and their energy $\mathcal{E}$ is frame-dependent, but the product of the two is independent of reference frame.

Turn now to the frame dependence of the particles' 3-volume $d\mathcal{V}_p$. As one sees from Fig. 3.3b, $d\mathcal{V}_p$ is the projection of the frame-independent mass-hyperboloid region $d\vec{\Sigma}_p$ onto the laboratory's xyz 3-space. Equivalently, it is the time component $d\Sigma^0_p$ of $d\vec{\Sigma}_p$. Now, the 4-vector $d\vec{\Sigma}_p$, like the 4-momentum $\vec{p}$, is orthogonal to the mass hyperboloid at the common point where they intersect it, and therefore $d\vec{\Sigma}_p$ is parallel to $\vec{p}$. This means that, when one goes from one reference frame to another, the time components of these two vectors will grow or shrink in the same manner, i.e., $d\vec{\Sigma}_p^0 = d\mathcal{V}_p$ is proportional to $p^0 = \mathcal{E}$. This means that their ratio is frame-independent:

$$\boxed{\frac{d\mathcal{V}_p}{\mathcal{E}} = \text{(a frame-independent quantity)}} \,. \tag{3.7b}$$

(If this sophisticated argument seems too slippery to you, then you can develop an alternative, more elementary proof using simpler two-dimensional spacetime diagrams: Exercise 3.1.)

By taking the product of Eqs. (3.7a) and (3.7b) we see that for our chosen set of particles $\mathcal{S}$,

$$d\mathcal{V}_x d\mathcal{V}_p = d^2\mathcal{V} = \text{(a frame-independent quantity)} \;; \tag{3.7c}$$

and since the number of particles in the set, dN, is obviously frame-independent, we conclude that

$$\boxed{\mathcal{N} = \frac{dN}{d\mathcal{V}_x d\mathcal{V}_p} \equiv \frac{dN}{d^2\mathcal{V}} = \text{(a frame-independent quantity)}} \,. \tag{3.8}$$

Although we assumed nonzero rest mass, $m \neq 0$, in our derivation, the conclusions that $\mathcal{E} d\mathcal{V}_x$ and $d\mathcal{V}_p/\mathcal{E}$ are frame-independent continue to hold if we take the limit as $m \to 0$ and the 4-momenta become null. Correspondingly, all of Eqs. (3.7a) – (3.8) are valid for particles with zero rest mass as well as nonzero.

****

EXERCISES

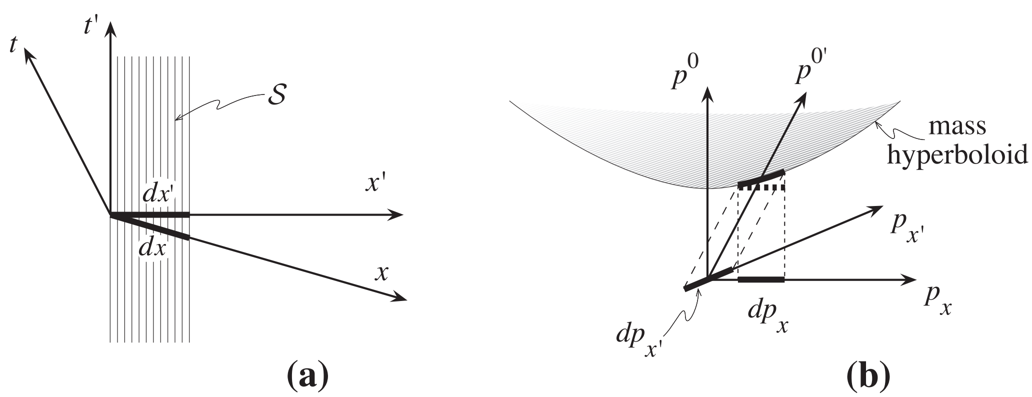

Fig. 3.4: (a) Spacetime diagram drawn from viewpoint of the (primed) rest frame of the particles $\mathcal{S}$ for the special case where the laboratory frame moves in the $-x'$ direction with respect to them. (b) Momentum space diagram drawn from viewpoint of the unprimed observer.

Exercise 3.1 **T2** [R] *Derivation and Practice: Frame-Dependences of $d\mathcal{V}_x$ and $d\mathcal{V}_p$*
Use the two-dimensional spacetime diagrams of Fig. 3.4 to show that $\mathcal{E}d\mathcal{V}_x$ and $d\mathcal{V}_p/\mathcal{E}$ are frame-independent [Eqs. (3.7a) and (3.7b)].

Exercise 3.2 ***Example:* **T2** [R] *Distribution function for Particles with a Range of Rest Masses*
A galaxy such as our Milky Way contains $\sim 10^{12}$ stars—easily enough to permit a kinetic-theory description of their distribution; and each star contains so many atoms ($\sim 10^{56}$) that the masses of the stars can be regarded as continuously distributed, not discrete. Almost everywhere in a galaxy, the stars move with speeds small compared to light, but deep in the cores of some galaxies there occasionally may develop a cluster of stars and black holes with relativistic speeds. In this exercise we shall explore the foundations for treating such a system: "particles" with continuously distributed rest masses and relativistic speeds.

(a) For a subset $\mathcal{S}$ of particles like that of Fig. 3.3 and associated discussion, but with a range of rest masses dm centered on some value m, introduce the phase-space volume $d^2\mathcal{V} \equiv d\mathcal{V}_x d\mathcal{V}_p dm$ that the particles $\mathcal{S}$ occupy. Explain why this occupied volume is frame-invariant.

b Show that this invariant occupied volume can be rewritten as $d^2\mathcal{V} = (d\mathcal{V}_x\,\mathcal{E}/m)(d\mathcal{V}_p d\mathcal{E}) = (d\mathcal{V}_x\,\mathcal{E}/m)(dp^0 dp^x dp^y dp^z)$. Explain the physical meaning of each of the terms in parentheses, and show that each is frame-invariant.

If the number of particles in the set $\mathcal{S}$ is dN, then we define the frame-invariant distribution function by

$$\mathcal{N} \equiv \frac{dN}{d^2\mathcal{V}} = \frac{dN}{d\mathcal{V}_x d\mathcal{V}_p dm} \,. \tag{3.9}$$

This is a function of location $\mathcal{P}$ in 4-dimensional spacetime and location $\vec{p}$ in 4-dimensional momentum space (not confined to the mass hyperboloid), i.e. a function of location in 8-dimensional phase space. We will explore the evolution of this distribution function in Box 3.2 below.

****

3.2.3 [N] Distribution function $f(\mathbf{x}, \mathbf{v}, t)$ for Particles in a Plasma.

The normalization that one uses for the distribution function is arbitrary; renormalize $\mathcal{N}$ by multiplying with any constant, and $\mathcal{N}$ will still be a geometric, coordinate-independent and frame-independent quantity and will still contain the same information as before. In this book, we shall use several renormalized versions of $\mathcal{N}$, depending on the situation. We shall now introduce them, beginning with the version used in plasma physics:

In Part VI, when dealing with nonrelativistic plasmas (collections of electrons and ions that have speeds small compared to light), we will regard the distribution function as depending on time t, location $\mathbf{x}$ in Euclidean space, and velocity $\mathbf{v}$ (instead of momentum

$\mathbf{p} = m\mathbf{v}$), and we will denote it by

$$f(t, \mathbf{x}, \mathbf{v}) \equiv \frac{dN}{d\mathcal{V}_x \, d\mathcal{V}_v} = \frac{dN}{dxdydz \, dv_x dv_y dv_z} = m^3 \mathcal{N} \; . \tag{3.10}$$

(This change of viewpoint and notation when transitioning to plasma physics is typical of the textbook you are reading. When presenting any subfield of physics, we shall usually adopt the conventions, notation, and also the system of units that are generally used in that subfield.)

3.2.4 [N & R] Distribution function I_ν/ν^3 for Photons.

[*Note to those readers who are restricting themselves to the Newtonian portions of this book:* Please read Sec. 1.10, which lists a few items of special relativity that you will need. As described there, you can deal with photons fairly easily by simply remembering that a photon has zero rest mass and has an energy $\mathcal{E} = h\nu$ and momentum $\mathbf{p} = (h\nu/c)\mathbf{n}$, where ν is its frequency and $\mathbf{n}$ is a unit vector pointing in its spatial direction.]

When dealing with photons or other zero-rest-mass particles, one often expresses $\mathcal{N}$ in terms of the *spectral intensity*, I_ν. This quantity (also sometimes called the specific intensity) is defined as follows; see Fig. 3.5: An observer places a CCD (or other measuring device) perpendicular to the photons' propagation direction $\mathbf{n}$—perpendicular as measured in her reference frame. The region of the CCD that the photons hit has surface area dA as measured by her, and because the photons move at the speed of light c, the product of that surface area with c times the time dt that they take to all go through the CCD is equal to the volume they occupy at a specific moment of time:

$$d\mathcal{V}_x = dA \, cdt \; . \tag{3.11a}$$

Focus attention on a set $\mathcal{S}$ of photons that all have nearly the same frequency ν and propagation direction $\mathbf{n}$ as measured by the observer. Their energies $\mathcal{E}$ and momenta $\mathbf{p}$ are related to ν and $\mathbf{n}$ by

$$\mathcal{E} = h\nu \; , \quad \mathbf{p} = (h\nu/c)\mathbf{n} \; , \tag{3.11b}$$

where h is Planck's constant. Their frequencies lie in a range $d\nu$ centered on ν, and they come from a small solid angle $d\Omega$ centered on $-\mathbf{n}$; and the volume they occupy in momentum space is related to these by

$$d\mathcal{V}_p = |\mathbf{p}|^2 d\Omega d|\mathbf{p}| = (h\nu/c)^2 d\Omega (h d\nu/c) = (h/c)^3 \nu^2 d\Omega d\nu \; . \tag{3.11c}$$

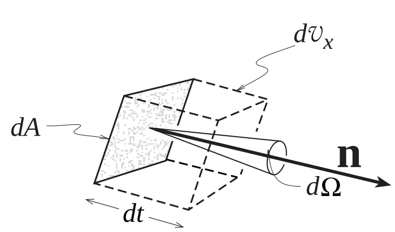

Fig. 3.5: Geometric construction used in defining the "spectral intensity" I_ν.

The photons' spectral intensity, as measured by the observer, is defined to be the total energy,

$$d\mathcal{E} = h\nu dN \tag{3.11d}$$

(where dN is the number of photons) that crosses the CCD per unit area dA, per unit time dt, per unit frequency $d\nu$, and per unit solid angle $d\Omega$ (i.e., "per unit everything"):

$$I_\nu \equiv \frac{d\mathcal{E}}{dAdtd\nu d\Omega} \,. \tag{3.12}$$

(This I_ν is sometimes denoted $I_{\nu\Omega}$.) From Eqs. (3.8), (3.11) and (3.12) we readily deduce the following relationship between this spectral intensity and the distribution function:

$$\boxed{\mathcal{N} = \frac{c^2}{h^4}\frac{I_\nu}{\nu^3}} \,. \tag{3.13}$$

This relation shows that, with an appropriate renormalization, I_ν/ν^3 *is the photons' distribution function.*

Astronomers and opticians regard the spectral intensity (or equally well I_ν/ν^3) as a function of the photon propagation direction $\mathbf{n}$, photon frequency ν, location $\mathbf{x}$ in space, and time t. By contrast, nonrelativistic physicists regard the distribution function $\mathcal{N}$ as a function of the photon momentum $\mathbf{p}$, location in space, and time; and relativistic physicists regard it as a function of the photon 4-momentum $\vec{p}$ (on the photons' mass hyperboloid, which is the light cone) and of location $\mathcal{P}$ in spacetime. Clearly, the information contained in these three sets of variables, the astronomers' set and the two physicists' sets, is the same.

If two different astronomers in two different reference frames at the same event in spacetime examine the same set of photons, they will measure the photons to have different frequencies ν (because of the Doppler shift between their two frames); and they will measure different spectral intensities I_ν (because of Doppler shifts of frequencies, Doppler shifts of energies, dilation of times, Lorentz contraction of areas of CCD's, and aberrations of photon propagation directions and thence distortions of solid angles). However, if each astronomer computes the ratio of the spectral intensity that she measures to the cube of the frequency she measures, that ratio, according to Eq. (3.13), will be the same as computed by the other astronomer; i.e., the distribution function I_ν/ν^3 *will be frame-independent.*

3.2.5 [N & R] Mean Occupation Number, η

Although this book is about classical physics, we cannot avoid making frequent contact with quantum theory. The reason is that classical physics is quantum mechanical in origin. Classical physics is an approximation to quantum physics, and not conversely. Classical physics is derivable from quantum physics, and not conversely.

In statistical physics, the classical theory cannot fully shake itself free from its quantum roots; it must rely on them in crucial ways that we shall meet in this chapter and the next. Therefore, rather than try to free it from its roots, we shall expose the roots and profit from them by introducing a quantum mechanically based normalization for the distribution function: the "mean occupation number" η.

As an aid in defining the mean occupation number, we introduce the concept of the *density of states*: Consider a particle of mass m, described quantum mechanically. Suppose that the particle is known to be located in a volume $d\mathcal{V}_x$ (as observed in a specific inertial reference frame) and to have a spatial momentum in the region $d\mathcal{V}_p$ centered on $\mathbf{p}$. Suppose, further, that *the particle does not interact with any other particles or fields*; for example, ignore Coulomb interactions. (In portions of Chaps. 4 and 5 we will include interactions.) Then how many single-particle quantum mechanical states[2] are available to the free particle? This question is answered most easily by constructing (in some arbitrary inertial frame), a complete set of wave functions for the particle's spatial degrees of freedom, with the wave functions (i) confined to be eigenfunctions of the momentum operator, and (ii) confined to satisfy the standard periodic boundary conditions on the walls of a box with volume $d\mathcal{V}_x$. For simplicity, let the box have edge length L along each of the three spatial axes of the Cartesian spatial coordinates, so $d\mathcal{V}_x = L^3$. (This L is arbitrary and will drop out of our analysis shortly.) Then a complete set of wave functions satisfying (i) and (ii) is the set $\{\psi_{j,k,l}\}$ with

$$\psi_{j,k,l}(x,y,z) = \frac{1}{L^{3/2}} e^{i(2\pi/L)(jx+ky+lz)} e^{-i\omega t} \tag{3.14a}$$

[cf., e.g., pp. 1440–1442 of Cohen-Tannoudji, Diu and Laloe (1977), especially the Comment at the end of this section]. Here the demand that the wave function take on the same values at the left and right faces of the box ($x = -L/2$ and $x = +L/2$), and at the front and back faces, and at the top and bottom faces (the demand for periodic boundary conditions) dictates that the quantum numbers j, k, and l be integers. The basis states (3.14a) are eigenfunctions of the momentum operator $(\hbar/i)\boldsymbol{\nabla}$ with momentum eigenvalues

$$p_x = \frac{2\pi\hbar}{L} j \; , \quad p_y = \frac{2\pi\hbar}{L} k \; , \quad p_z = \frac{2\pi\hbar}{L} l \; ; \tag{3.14b}$$

and correspondingly, the wave function's frequency ω has the following values in Newtonian theory [N] and relativity [R]:

$$\text{[N]}\ \hbar\omega = E = \frac{\mathbf{p}^2}{2m} = \frac{1}{2m}\left(\frac{2\pi\hbar}{L}\right)^2 (j^2+k^2+l^2) \; ; \tag{3.14c}$$

$$\text{[R]}\ \hbar\omega = \mathcal{E} = \sqrt{m^2+\mathbf{p}^2} \to m + E \text{ in the Newtonian limit.} \tag{3.14d}$$

Equations (3.14b) tell us that the allowed values of the momentum are confined to "lattice sites" in 3-momentum space with one site in each cube of side $2\pi\hbar/L$. Correspondingly, the total number of states in the region $d\mathcal{V}_x d\mathcal{V}_p$ of phase space is the number of cubes of side $2\pi\hbar/L$ in the region $d\mathcal{V}_p$ of momentum space:

$$dN_{\text{states}} = \frac{d\mathcal{V}_p}{(2\pi\hbar/L)^3} = \frac{L^3 d\mathcal{V}_p}{(2\pi\hbar)^3} = \frac{d\mathcal{V}_x d\mathcal{V}_p}{h^3}. \tag{3.15}$$

This is true no matter how relativistic or nonrelativistic the particle may be.

[2] A quantum mechanical state for a single particle is called an "orbital" in the chemistry literature and in the classic thermal physics textbook by Kittel and Kroemer (1980); we shall use physicists' more conventional but cumbersome phrase "single-particle quantum state", and also, sometimes, the word "mode".

Thus far we have considered only the particle's spatial degrees of freedom. Particles can also have an internal degree of freedom called "spin". For a particle with spin s, the number of independent spin states is

$$g_s = \begin{cases} 2s+1 & \text{if } m \neq 0\text{; e.g., an electron or proton or atomic nucleus,} \\ 2 & \text{if } m = 0 \text{ \& } s > 0\text{; e.g., a photon } (s=1) \text{ or graviton } (s=2), \\ 1 & \text{if } m = 0 \text{ \& } s = 0\text{; i.e., a hypothetical massless scalar particle.} \end{cases} \tag{3.16}$$

A notable exception is each species of neutrino or antineutrino, which has nonzero rest mass and spin 1/2, but $g_s = 1$ rather than $g_s = 2s+1 = 2$.[3]) We shall call this number of internal spin states g_s the particle's *multiplicity*. [It will turn out to play a crucial role in computing the *entropy* of a system of particles (Chap. 4); i.e., it places the imprint of quantum theory on the entropy of even a highly classical system.]

Taking account of both the particle's spatial degrees of freedom and its spin degree of freedom, we conclude that the total number of independent quantum states available in the region $d\mathcal{V}_x d\mathcal{V}_p \equiv d^2\mathcal{V}$ of phase space is $dN_{\rm states} = (g_s/h^3)d^2\mathcal{V}$, and correspondingly the *number density of states in phase space* is

$$\mathcal{N}_{\rm states} \equiv \frac{dN_{\rm states}}{d^2\mathcal{V}} = \frac{g_s}{h^3} \ . \tag{3.17}$$

[Relativistic remark: Note that, although we derived this number density of states using a specific inertial frame, it is a frame-independent quantity, with a numerical value depending only on Planck's constant and (through g_s) the particle's rest mass m and spin s.]

The ratio of the number density of particles to the number density of quantum states is obviously the number of particles in each state (the state's *occupation number*), averaged over many neighboring states—but few enough that the averaging region is small by macroscopic standards. In other words, this ratio is the quantum states' *mean occupation number* η:

$$\eta = \frac{\mathcal{N}}{\mathcal{N}_{\rm states}} = \frac{h^3}{g_s}\mathcal{N} \ ; \quad \text{i.e.,} \ \boxed{\mathcal{N} = \mathcal{N}_{\rm states}\eta = \frac{g_s}{h^3}\eta} \ . \tag{3.18}$$

The mean occupation number η plays an important role in quantum statistical mechanics, and its quantum roots have a profound impact on classical statistical physics:

From quantum theory we learn that the allowed values of the occupation number for a quantum state depend on whether the state is that of a *fermion* (a particle with spin 1/2, 3/2, 5/2, ...) or that of a *boson* (a particle with spin 0, 1, 2, ...). For fermions, no two particles can occupy the same quantum state, so the occupation number can only take on the eigenvalues 0 and 1. For bosons, one can shove any number of particles one wishes into the same quantum state, so the occupation number can take on the eigenvalues 0, 1, 2, 3, ... Correspondingly, the mean occupation numbers must lie in the ranges

$$0 \le \eta \le 1 \text{ for fermions} \ , \quad 0 \le \eta < \infty \text{ for bosons} \ . \tag{3.19}$$

Quantum theory also teaches us that, when $\eta \ll 1$, the particles, whether fermions or bosons, behave like *classical, discrete, distinguishable particles*; and when $\eta \gg 1$ (possible only for

[3] The reason is the particle's fixed chirality: -1 for neutrino and +1 for antineutrino; in order to have $g_s = 2$, a spin 1/2 particle must admit both chiralities.

bosons), the particles behave like a *classical wave* — if the particles are photons ($s = 1$), like a classical electromagnetic wave; and if they are gravitons ($s = 2$), like a classical gravitational wave. This role of η in revealing the particles' physical behavior will motivate us frequently to use η as our distribution function instead of $\mathcal{N}$.

Of course η, like $\mathcal{N}$, is a function of location in phase space, $\eta(\mathcal{P}, \vec{p})$ in relativity with no inertial frame chosen; or $\eta(t, \mathbf{x}, \mathbf{p})$ in both relativity and Newtonian theory when an inertial frame is in use.

****

EXERCISES

Exercise 3.3 ***Practice & Example:* [N & R] *Regimes of Particulate and Wave-like Behavior*

(a) Cygnus X-1 is a source of X-rays that has been studied extensively by astronomers. The observations (X-ray, optical, and radio) show that it is a distance $r \sim 6000$ lyr from Earth, and it consists of a very hot disk of X-ray-emitting gas that surrounds a black hole with mass $15 M_\odot$, and the hole in turn is in a binary orbit with a heavy companion star. Most of the X-ray photons have energies $\mathcal{E} \sim 3$ keV, their energy flux arriving at Earth is $F \sim 10^{-23} \mathrm{Wm}^{-2}\mathrm{s}^{-1}$, and the portion of the disk that emits most of them has radius roughly 10 times that of the black hole, i.e., $R \sim 500$ km. Make a rough estimate of the mean occupation number of the X-rays' photon states. Your answer should be in the region $\eta \ll 1$, so the photons behave like classical, distinguishable particles. Will the occupation number change as the photons propagate from the source to Earth?

(b) A highly nonspherical supernova in the Virgo cluster of galaxies (40 million light years from Earth) emits a burst of gravitational radiation with frequencies spread over the band 500 Hz to 2000 Hz, as measured at Earth. The burst comes out in a time of about 10 milliseconds, so it lasts only a few cycles, and it carries a total energy of roughly $10^{-3} M_\odot c^2$, where $M_\odot = 2 \times 10^{33}$ g is the mass of the sun. The emitting region is about the size of the newly forming neutron-star core (10 km), which is small compared to the wavelength of the waves; so if one were to try to resolve the source spatially by imaging the gravitational waves with a gravitational lens, one would see only a blur of spatial size one wavelength rather than seeing the neutron star. What is the mean occupation number of the burst's graviton states? Your answer should be in the region $\eta \gg 1$, so the gravitons behave like a classical gravitational wave.

****

3.3 [N & R] Thermal-Equilibrium Distribution Functions

In Chap. 4 we will introduce with care, and explore in detail, the concept of "statistical equilibrium"—also called "thermal equilibrium". That exploration will lead to a set of distribution functions for particles that are in statistical equilibrium. In this section we will summarize those equilibrium distribution functions, so as to be able to use them for examples and applications of kinetic theory.

If a collection of many identical particles is in thermal equilibrium in the neighborhood of an event $\mathcal{P}$ then, as we shall see in Chap. 4, there is a special inertial reference frame (the *mean rest frame* of the particles near $\mathcal{P}$) in which there are equal numbers of particles of any given speed going in all directions, i.e. the mean occupation number η is a function only of the magnitude $|\mathbf{p}|$ of the particle momentum and does not depend on the momentum's direction. Equivalently, η is a function of the particle's energy. In the relativistic regime, we use two different energies, one denoted $\mathcal{E}$ that includes the contribution of the particle's rest mass and the other denoted E that omits the rest mass and thus represents kinetic energy (cf. Sec. 1.10):

$$\boxed{E \equiv \mathcal{E} - m = \sqrt{m^2 + \mathbf{p}^2} - m \to \frac{\mathbf{p}^2}{2m} \text{ in the low-velocity, Newtonian limit.}} \tag{3.20}$$

In the nonrelativistic, Newtonian regime we shall use only $E = \mathbf{p}^2/2m$.

Most readers will already know that the details of the thermal equilibrium are fixed by two quantities: the mean density of particles and the mean energy per particle, or equivalently (as we shall see) by the *chemical potential* μ and the *temperature* T. By analogy with our treatment of relativistic energy, we shall use two different chemical potentials: one, $\tilde{\mu}$, that includes rest mass and the other,

$$\mu \equiv \tilde{\mu} - m \,, \tag{3.21}$$

that does not. In the Newtonian regime we shall use only μ.

As we shall prove by an elegant argument in Chap. 4, in thermal equilibrium the mean occupation number has the following form at all energies, relativistic or nonrelativistic:

$$\boxed{\eta = \frac{1}{e^{(E-\mu)/k_BT} + 1} \quad \text{for fermions}} \,, \tag{3.22a}$$

$$\boxed{\eta = \frac{1}{e^{(E-\mu)/k_BT} - 1} \quad \text{for bosons}} \,. \tag{3.22b}$$

Here $k_B = 1.381 \times 10^{-16}\text{erg K}^{-1} = 1.381 \times 10^{-23}\text{J K}^{-1}$ is Boltzmann's constant. Equation (3.22a) for fermions is called the *Fermi-Dirac distribution*; Eq. (3.22b) for bosons is called the *Bose-Einstein* distribution. In the relativistic regime, we can also write these distribution functions in terms of the energy $\mathcal{E}$ that includes the rest mass as

$$[\mathrm{R}] \quad \eta = \frac{1}{e^{(E-\mu)/k_BT} \pm 1} = \frac{1}{e^{(\mathcal{E}-\tilde{\mu})/k_BT} \pm 1} \,. \tag{3.22c}$$

Notice that the equilibrium mean occupation number (3.22a) for fermions lies in the range 0 to 1 as required, while that (3.22b) for bosons lies in the range 0 to ∞. In the regime $\mu \ll -k_BT$, the mean occupation number is small compared to unity for all particle energies E (since E is never negative, i.e. $\mathcal{E}$ is never less than m). This is the domain of distinguishable, classical particles, and in it both the Fermi-Dirac and Bose-Einstein distributions become

$$\boxed{\eta \simeq e^{-(E-\mu)/k_BT} = e^{-(\mathcal{E}-\tilde{\mu})/k_BT} \quad \text{when } \mu \equiv \tilde{\mu} - m \ll -k_BT \quad \text{(classical particles)}} \,. \tag{3.22d}$$

This limiting distribution is called the *Boltzmann distribution*.[4]

By scrutinizing the distribution functions (3.22), one can deduce that the larger the temperature T at fixed μ, the larger will be the typical energies of the particles, and the larger the chemical potential μ at fixed T, the larger will be the total density of particles; see Ex. 3.4 and Eqs. (3.37). For bosons, μ must always be negative or zero, i.e. $\tilde{\mu}$ cannot exceed the particle rest mass m; otherwise η would be negative at low energies, which is physically impossible. For bosons with μ extremely close to zero, there is a huge number of very low energy particles, leading quantum mechanically to a *boson condensate*; we shall study boson condensates in Sec. 4.9.

In the special case that the particles of interest can be created and destroyed completely freely, with creation and destruction constrained only by the laws of 4-momentum conservation, the particles quickly achieve a thermal equilibrium in which the relativistic chemical potential vanishes, $\tilde{\mu} = 0$ (as we shall see in Sec. 5.5). For example, inside a box whose walls are perfectly emitting and absorbing and have temperature T, the photons acquire the mean occupation number (3.22b) with zero chemical potential, leading to the standard *black-body (Planck)* form

$$\eta = \frac{1}{e^{h\nu/k_BT} - 1} \; , \quad \mathcal{N} = \frac{2}{h^3}\frac{1}{e^{h\nu/k_BT} - 1} \; , \quad I_\nu = \frac{(2h/c^2)\nu^3}{e^{h\nu/k_BT} - 1} \; . \tag{3.23}$$

(Here we have set $E = h\nu$ where ν is the photon frequency as measured in the box's rest frame, and in the third expression we have inserted the factor c^{-2} so that I_ν will be in ordinary units.)

On the other hand, if one places a fixed number of photons inside a box whose walls cannot emit or absorb them but can scatter them, exchanging energy with them in the process, then the photons will acquire the Bose-Einstein distribution (3.22b) with temperature T equal to that of the walls and with nonzero chemical potential μ fixed by the number of photons present; the more photons there are, the larger will be the chemical potential.

****

EXERCISES

Exercise 3.4 ***Example:* [N] *Maxwell Velocity Distribution*
Consider a collection of thermalized, classical particles with nonzero rest mass, so they have the Boltzmann distribution (3.22d). Assume that the temperature is low enough, $k_BT \ll mc^2$ that they are nonrelativistic.

(a) Explain why the total number density of particles n in physical space (as measured in the particles' mean rest frame) is given by the integral $n = \int \mathcal{N} d\mathcal{V}_p$. Show that $n \propto e^{\mu/k_BT}$, and derive the proportionality constant. [Hint: use spherical coordinates in momentum space so $d\mathcal{V}_p = 4\pi p^2 dp$ with $p \equiv |\mathbf{p}|$.] Your answer should be Eq. (3.37a) below.

[4] Lynden-Bell (1967) identifies a fourth type of thermal distribution which occurs in the theory of violent relaxation of star clusters. It corresponds to individually distinguishable, classical particles (in his case stars with a range of masses) that obey the same kind of exclusion principle as fermions.

(b) Explain why the mean energy per particle is given by $\bar{E} = \int (p^2/2m) d\mathcal{V}_p$. Show that $\bar{E} = \frac{3}{2} k_B T$.

(c) Show that $P(v)dv \equiv$(probability that a randomly chosen particle will have speed $v \equiv |\mathbf{v}|$ in the range dv) is given by

$$\boxed{P(v) = \frac{4}{\sqrt{\pi}} \frac{v^2}{v_o^3} e^{-v^2/v_o^3} \,, \quad \text{where } v_o = \sqrt{\frac{2k_B T}{m}}} \,. \tag{3.24}$$

This is called the *Maxwell velocity distribution*; it is graphed in Fig. 3.6a. Notice that the peak of the distribution is at speed v_o.

[Side Remark: In the normalization of probability distributions such as this one, you will often encounter integrals of the form $\int_0^\infty x^{2n} e^{-x^2} dx$. You can evaluate this quickly via integration by parts, if you have memorized that $\int_0^\infty e^{-x^2} dx = \sqrt{\pi}/2$.]

(d) Consider particles confined to move in a plane, or in one dimension (on a line). What is their speed distribution $P(v)$ and at what speed does it peak?

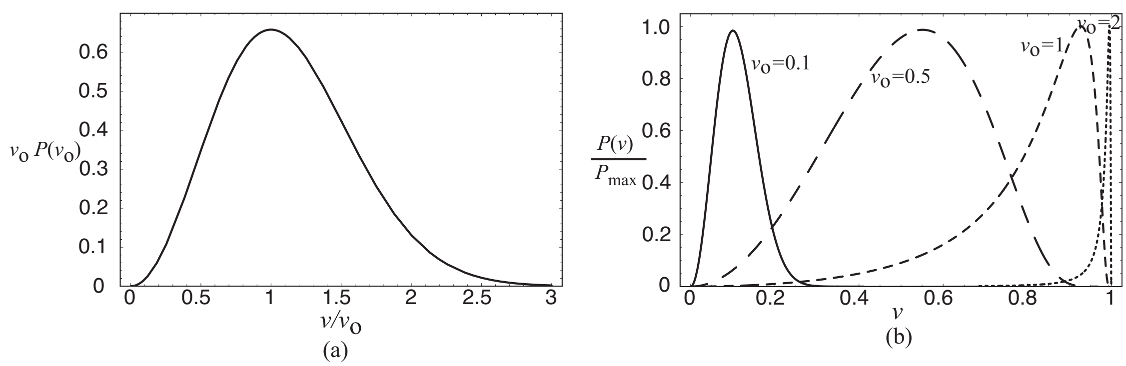

Fig. 3.6: (a) Maxwell velocity distribution for thermalized, classical, nonrelativistic particles. (b) Extension of the Maxwell velocity distribution into the relativistic domain. In both plots $v_o = \sqrt{2k_B T/m}$

Exercise 3.5 *Problem:* **T2** [R] *Velocity Distribution for Thermalized, Classical, Relativistic Particles*

Show that for thermalized, classical relativistic particles the probability distribution for the speed [relativistic version of the Maxwell distribution (3.24)] is

$$P(v) = \text{constant} \frac{v^2}{(1-v^2)^{5/2}} \exp\left[-\frac{2/v_o^2}{\sqrt{1-v^2}} \right] \,, \quad \text{where } v_o = \sqrt{\frac{2k_B T}{m}} \,. \tag{3.25}$$

This is plotted in Fig. 3.6b for a sequence of four temperatures ranging from the nonrelativistic regime $k_B T \ll m$ toward the ultrarelativistic regime $k_B T \gg m$. In the ultrarelativistic regime the particles are (almost) all moving at very close to the speed of light, $v = 1$.

Exercise 3.6 ***Example:* **[*T2*]** [R] *Observations of Cosmic Microwave Radiation from Earth*

The universe is filled with cosmic microwave radiation left over from the big bang. At each event in spacetime the microwave radiation has a mean rest frame; and as seen in that mean rest frame the radiation's distribution function η is almost precisely isotropic and thermal with zero chemical potential:

$$\eta = \frac{1}{e^{h\nu/k_B T_o} - 1} , \quad \text{with } T_o = 2.725\,\mathrm{K} . \tag{3.26}$$

Here ν is the frequency of a photon as measured in the mean rest frame.

(a) Show that the spectral intensity of the radiation as measured in its mean rest frame has the *Planck spectrum*

$$I_\nu = \frac{(2h/c^2)\nu^3}{e^{h\nu/k_B T_o} - 1} . \tag{3.27}$$

Plot this spectral intensity as a function of wavelength and from your plot determine the wavelength of the intensity peak.

(b) Show that η can be rewritten in the frame-independent form

$$\eta = \frac{1}{e^{-\vec{p}\cdot\vec{u}_o/k_B T_o} - 1} , \tag{3.28}$$

where $\vec{p}$ is the photon 4-momentum and $\vec{u}_o$ is the 4-velocity of the mean rest frame. [Hint: See Sec. 2.6 and especially Eq. (2.29).]

(c) In actuality, the Earth moves relative to the mean rest frame of the microwave background with a speed v of roughly 400 km/sec toward the Hydra-Centaurus region of the sky. An observer on Earth points his microwave receiver in a direction that makes an angle θ with the direction of that motion, as measured in the Earth's frame. Show that the spectral intensity of the radiation received is precisely Planckian in form [Eq. (3.23)], but with a direction-dependent *Doppler-shifted temperature*

$$T = T_o \left(\frac{\sqrt{1 - v^2}}{1 - v\cos\theta} \right) . \tag{3.29}$$

Note that this Doppler shift of T is precisely the same as the Doppler shift of the frequency of any specific photon. Note also that the θ dependence corresponds to an anisotropy of the microwave radiation as seen from Earth. Show that because the Earth's velocity is small compared to the speed of light, the anisotropy is very nearly dipolar in form. Measurements by the WMAP satellite give $T_o = 2.725$ K and (averaged over a year) an amplitude of 3.346×10^{-3} K for the dipolar temperature variations (Bennet et. al. 2003). What, precisely, is the value of the Earth's year-averaged speed v?

****

3.4 Macroscopic Properties of Matter as Integrals over Momentum Space

3.4.1 [N] Particle Density n, Flux $\mathbf{S}$, and Stress Tensor $\boldsymbol{\mathsf{T}}$

If one knows the Newtonian distribution function $\mathcal{N} = (g_s/h^3)\eta$ as a function of momentum $\mathbf{p}$ at some location $(\mathbf{x}, t)$ in space and time, one can use it to compute various macroscopic properties of the particles. Specifically:

From the definition $\mathcal{N} \equiv dN/d\mathcal{V}_x d\mathcal{V}_p$ of the distribution function, it is clear that the number density of particles $n(\mathbf{x}, t)$ in physical space is given by the integral

$$\boxed{n = \frac{dN}{d\mathcal{V}_x} = \int \frac{dN}{d\mathcal{V}_x d\mathcal{V}_p} d\mathcal{V}_p = \int \mathcal{N} d\mathcal{V}_p} \, . \tag{3.30a}$$

Similarly, the number of particles crossing a unit surface in the y-z plane per unit time, i.e. the x component of the flux of particles, is

$$S_x = \frac{dN}{dydzdt} = \int \frac{dN}{dxdydzd\mathcal{V}_p} \frac{dx}{dt} d\mathcal{V}_p = \int \mathcal{N} \frac{p_x}{m} d\mathcal{V}_p \, ,$$

where $dx/dt = p_x/m$ is the x component of the particle velocity. This and the analogous equations for S_y and S_z can be combined into the single geometric, coordinate-independent integral

$$\boxed{\mathbf{S} = \int \mathcal{N} \, \mathbf{p} \, \frac{d\mathcal{V}_p}{m}} \, . \tag{3.30b}$$

Notice that, if we multiply this $\mathbf{S}$ by the particles' mass m, the integral becomes the momentum density:

$$\boxed{\boldsymbol{\mathcal{G}} = m\mathbf{S} = \int \mathcal{N} \mathbf{p} \, d\mathcal{V}_p} \, . \tag{3.30c}$$

Finally, since the stress tensor $\boldsymbol{\mathsf{T}}$ is the flux of momentum [Eq. (1.33)], its j-x component (j component of momentum crossing a unit area in the y-z plane per unit time) must be

$$T_{jx} = \int \frac{dN}{dydzdtd\mathcal{V}_p} p_j \, d\mathcal{V}_p = \int \frac{dN}{dxdydzd\mathcal{V}_p} \frac{dx}{dt} p_j d\mathcal{V}_p = \int \mathcal{N} p_j \frac{p_x}{m} d\mathcal{V}_p \, .$$

This and the corresponding equations for T_{jy} and T_{jz} can be collected together into the single geometric, coordinate-independent integral

$$T_{jk} = \int \mathcal{N} p_j p_k \frac{d\mathcal{V}_p}{m} \, , \quad \text{i.e.,} \; \boxed{\boldsymbol{\mathsf{T}} = \int \mathcal{N} \mathbf{p} \otimes \mathbf{p} \frac{d\mathcal{V}_p}{m}} \, . \tag{3.30d}$$

Notice that the number density n is the zero'th moment of the distribution function in momentum space [Eq. (3.30a)], and aside from factors $1/m$ the particle flux vector is the first moment [Eq. (3.30b)], and the stress tensor is the second moment [Eq, (3.30d)]. All three moments are geometric, coordinate-independent quantities, and they are the simplest such quantities that one can construct by integrating the distribution function over momentum space.

3.4.2 T2 [R] Relativistic Number-Flux 4-Vector $\vec{S}$ and Stress-Energy Tensor $\boldsymbol{T}$

When we switch from Newtonian theory to special relativity's 4-dimensional spacetime viewpoint, we require that all physical quantities be described by geometric, frame-independent objects (scalars, vectors, tensors, ...) in 4-dimensional spacetime. We can construct such objects as momentum-space integrals over the frame-independent, relativistic distribution function $\mathcal{N}(\mathcal{P},\vec{p}) = (g_s/h^3)\eta$. The frame-independent quantities that can appear in these integrals are (i) $\mathcal{N}$ itself, (ii) the particle 4-momentum $\vec{p}$, and (iii) the frame-independent integration element $d\mathcal{V}_p/\mathcal{E}$ [Eq. (3.7b)], which takes the form $dp_x dp_y dp_z/\sqrt{m^2+\mathbf{p}^2}$ in any inertial reference frame. By analogy with the Newtonian regime, the most interesting such integrals are the lowest three moments of the distribution function:

$$R \equiv \int \mathcal{N}\,\frac{d\mathcal{V}_p}{\mathcal{E}}\ ; \tag{3.31a}$$

$$\boxed{\vec{S} \equiv \int \mathcal{N}\vec{p}\,\frac{d\mathcal{V}_p}{\mathcal{E}}}\ , \quad \text{i.e. } S^\mu \equiv \int \mathcal{N}p^\mu\,\frac{d\mathcal{V}_p}{\mathcal{E}}\ ; \tag{3.31b}$$

$$\boxed{\boldsymbol{T} \equiv \int \mathcal{N}\vec{p}\otimes\vec{p}\,\frac{d\mathcal{V}_p}{\mathcal{E}}}\ , \quad \text{i.e. } T^{\mu\nu} \equiv \int \mathcal{N}p^\mu p^\nu\,\frac{d\mathcal{V}_p}{\mathcal{E}}\ . \tag{3.31c}$$

Here and throughout this chapter, relativistic momentum-space integrals unless otherwise specified are taken over the entire mass hyperboloid.

We can learn the physical meanings of each of the momentum-space integrals (3.31) by introducing a specific but arbitrary inertial reference frame, and using it to perform a 3+1 split of spacetime into space plus time [cf. the paragraph containing Eq. (2.28)]. When we do this and rewrite $\mathcal{N}$ as $dN/d\mathcal{V}_x d\mathcal{V}_p$, the scalar field R of Eq. (3.31a) takes the form

$$R = \int \frac{dN}{d\mathcal{V}_x d\mathcal{V}_p}\,\frac{1}{\mathcal{E}}\,d\mathcal{V}_p \tag{3.32}$$

(where of course $d\mathcal{V}_x = dxdydz$ and $d\mathcal{V}_p = dp_x dp_y dp_z$). This is the sum, over all particles in a unit 3-volume, of the inverse energy. Although it is intriguing that this quantity is a frame-independent scalar, it is not a quantity that appears in any important way in the laws of physics.

By contrast, the 4-vector field $\vec{S}$ of Eq. (3.31b) plays a very important role in physics. Its time component in our chosen frame is

$$S^0 = \int \frac{dN}{d\mathcal{V}_x d\mathcal{V}_p}\frac{p^0}{\mathcal{E}}d\mathcal{V}_p = \int \frac{dN}{d\mathcal{V}_x d\mathcal{V}_p}d\mathcal{V}_p \tag{3.33a}$$

(since p^0 and $\mathcal{E}$ are just different notations for the same thing, the relativistic energy $\sqrt{m^2+\mathbf{p}^2}$ of a particle). Obviously, this S^0 is the number of particles per unit spatial volume as measured in our chosen inertial frame:

$$S^0 = n = (\text{number density of particles})\ . \tag{3.33b}$$

The x component of $\vec{S}$ is

$$S^x = \int \frac{dN}{d\mathcal{V}_x d\mathcal{V}_p}\frac{p^x}{\mathcal{E}}\, d\mathcal{V}_p = \int \frac{dN}{dxdydz\, d\mathcal{V}_p}\frac{dx}{dt} d\mathcal{V}_p = \int \frac{dN}{dtdydz\, d\mathcal{V}_p} d\mathcal{V}_p \;, \tag{3.33c}$$

which is the number of particles crossing a unit area in the y-z plane per unit time, i.e. the x-component of the particle flux; and similarly for other directions j,

$$S^j = (j\text{-component of the particle flux vector } \mathbf{S}). \tag{3.33d}$$

[In Eq. (3.33c), the second equality follows from

$$\frac{p^j}{\mathcal{E}} = \frac{p^j}{p^0} = \frac{dx^j/d\zeta}{dt/d\zeta} = \frac{dx^j}{dt} = \; (j\text{-component of velocity}) \;, \tag{3.33e}$$

where ζ is the "affine parameter" such that $\vec{p} = d\vec{x}/d\zeta$.] Since S^0 is the particle number density and S^j is the particle flux, $\vec{S}$ *must be the number-flux 4-vector* introduced and studied in Sec. 2.12.3. Notice that in the Newtonian limit, where $p^0 = \mathcal{E} \to m$, the temporal and spatial parts of the formula $\vec{S} = \int \mathcal{N}\vec{p}(d\mathcal{V}_p/\mathcal{E})$ reduce to $S^0 = \int \mathcal{N} d\mathcal{V}_p$ and $\mathbf{S} = \int \mathcal{N}\mathbf{p}(d\mathcal{V}_p/m)$, which are the coordinate-independent expressions (3.30a) and (3.30b) for the Newtonian number density of particles and flux of particles.

Turn to the quantity $\boldsymbol{T}$ defined by the integral (3.31c). When we perform a 3+1 split of it in our chosen inertial frame, we find the following for its various parts:

$$T^{\mu 0} = \int \frac{dN}{d\mathcal{V}_x d\mathcal{V}_p} p^\mu p^0 \frac{d\mathcal{V}_p}{p^0} = \int \frac{dN}{d\mathcal{V}_x d\mathcal{V}_p} p^\mu d\mathcal{V}_p \;, \tag{3.34a}$$

is the μ-component of 4-momentum per unit volume (i.e., T^{00} is the energy density and T^{j0} is the momentum density). Also,

$$T^{\mu x} = \int \frac{dN}{d\mathcal{V}_x d\mathcal{V}_p} p^\mu p^j \frac{d\mathcal{V}_p}{p^0} = \int \frac{dN}{dxdydz d\mathcal{V}_p}\frac{dx}{dt} p^\mu d\mathcal{V}_p = \int \frac{dN}{dtdydz d\mathcal{V}_p} p^\mu d\mathcal{V}_p \tag{3.34b}$$

is the amount of μ-component of 4-momentum that crosses a unit area in the y-z plane per unit time; i.e., it is the x-component of flux of μ-component of 4-momentum. More specifically, T^{0x} is the x-component of energy flux (which is the same as the momentum density T^{x0}) and T^{jx} is the x component of spatial-momentum flux—or, equivalently, the jx component of the stress tensor. These and the analogous expressions and interpretations of $T^{\mu y}$ and $T^{\mu z}$ can be summarized by

$$T^{00} = (\text{energy density}) \;, \quad T^{j0} = (\text{momentum density}) = T^{0j} = (\text{energy flux}) \;, \quad T^{jk} = (\text{stress tensor}) \;. \tag{3.34c}$$

Therefore [cf. Eq. (2.67f)], $\boldsymbol{T}$ *must be the stress-energy tensor* introduced and studied in Sec. 2.13. Notice that in the Newtonian limit, where $\mathcal{E} \to m$, the coordinate-independent equation (3.31c) for the spatial part of the stress-energy tensor (the *stress*) becomes $\int \mathcal{N}\mathbf{p} \otimes \mathbf{p}\, d\mathcal{V}_p/m$, which is the same as our coordinate-independent equation (3.30d) for the stress tensor.

3.5 Isotropic Distribution Functions and Equations of State

3.5.1 [N] Newtonian Density, Pressure, Energy Density and Equation of State

Let us return, for a while, to Newtonian theory:

If the Newtonian distribution function is isotropic in momentum space, i.e. is a function only of the magnitude $p \equiv |\mathbf{p}| = \sqrt{p_x^2 + p_y^2 + p_z^2}$ of the momentum (as is the case, for example, when the particle distribution is thermalized), then the particle flux $\mathbf{S}$ vanishes (equal numbers of particles travel in all directions), and the stress tensor is isotropic, $\mathbf{T} = P\mathbf{g}$, i.e. $T_{jk} = P\delta_{jk}$; i.e. it is that of a perfect fluid. [Here P is the isotropic pressure and $\mathbf{g}$ is the metric tensor of Euclidian 3-space, with Cartesian components equal to the Kronecker delta; Eq. (1.9f).] In this isotropic case, the pressure can be computed most easily as 1/3 the trace of the stress tensor (3.30d):

$$P = \frac{1}{3}T_{jj} = \frac{1}{3}\int \mathcal{N}(p_x^2 + p_y^2 + p_z^2)\frac{d\mathcal{V}_p}{m} = \frac{1}{3}\int_0^\infty \mathcal{N}p^2\frac{4\pi p^2 dp}{m} = \frac{4\pi}{3m}\int_0^\infty \mathcal{N}p^4 dp\ . \tag{3.35a}$$

Here in the third step we have written the momentum-volume element in spherical polar coordinates as $d\mathcal{V}_p = p^2 \sin\theta d\theta d\phi dp$ and have integrated over angles to get $4\pi p^2 dp$. Similarly, we can reexpress the number density of particles (3.30a) and the corresponding mass density as

$$n = 4\pi\int_0^\infty \mathcal{N}p^2 dp\ , \quad \rho \equiv mn = 4\pi m\int_0^\infty \mathcal{N}p^2 dp\ . \tag{3.35b}$$

Finally, because each particle carries an energy $E = p^2/2m$, the energy density in this isotropic case (which we shall denote by U) is 3/2 the pressure:

$$U = \int \frac{p^2}{2m}\mathcal{N}d\mathcal{V}_p = \frac{4\pi}{2m}\int_0^\infty \mathcal{N}p^4 dp = \frac{3}{2}P\ ; \tag{3.35c}$$

cf. Eq. (3.35a).

If we know the distribution function for an isotropic collection of particles, Eqs. (3.35) give us a straightforward way of computing the collection's number density of particles n, mass density $\rho = nm$, perfect-fluid energy density U, and perfect-fluid pressure P as measured in the particles' mean rest frame. For a thermalized gas, the distribution functions (3.22a), (3.22b), and (3.22d) [with $\mathcal{N} = (g_s/h^3)\eta$] depend on two parameters: the temperature T and chemical potential μ, so this calculation gives n, U and P in terms of μ and T. One can then invert $n(\mu, T)$ to give $\mu(n, T)$ and insert into the expressions for U and P to obtain *equations of state* for thermalized, nonrelativistic particles

$$U = U(\rho, T)\ , \quad P = P(\rho, T)\ . \tag{3.36}$$

For a gas of nonrelativistic, classical particles, the distribution function is Boltzmann [Eq. (3.22d)], $\mathcal{N} = (g_s/h^3)e^{(\mu-E)/k_BT}$ with $E = p^2/2m$, and this procedure gives, quite easily (Ex. 3.7):

$$\boxed{n = \frac{g_s e^{\mu/k_BT}}{\lambda_{T\,\mathrm{dB}}^3} = \frac{g_s}{h^3}(2\pi m k_B T)^{3/2}e^{\mu/k_BT}}\ , \tag{3.37a}$$

$$\boxed{U = \frac{3}{2} n k_B T \ , \quad P = n k_B T} \ . \tag{3.37b}$$

Notice that the mean energy per particle is

$$\bar{E} = \frac{3}{2} k_B T \ . \tag{3.37c}$$

In Eq. (3.37a), $\lambda_{T\,\mathrm{dB}} \equiv h/\sqrt{2\pi m k_B T}$ is the particles' *thermal de Broglie wavelength*, i.e. the wavelength of Schrödinger wave-function oscillations for a particle with the thermal kinetic energy $E = \pi k_B T$. Note that the classical regime $\eta \ll 1$ (i.e. $\mu/k_B T \ll -1$), in which our computation is being performed, corresponds to a mean number of particles in a thermal de Broglie wavelength small compared to one, $n\lambda_{T\,\mathrm{dB}}^3 \ll 1$, which should not be surprising.

3.5.2 [N] Equations of State for a Nonrelativistic Hydrogen Gas

As an application, consider ordinary matter. Figure 3.7 shows its physical nature as a function of density and temperature, near and above "room temperature", 300 K. We shall study solids (lower right) in Part IV, fluids (lower middle) in Part V, and plasmas (middle) in Part VI.

Our kinetic theory tools are well suited to any situation where the particles have mean free paths large compared to their sizes. This is generally true in plasmas and sometimes

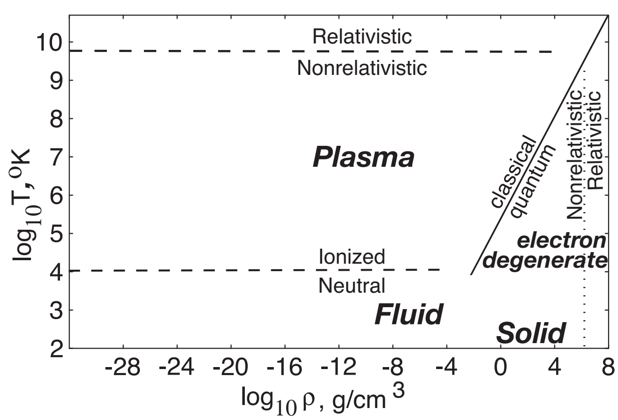

Fig. 3.7: Physical nature of hydrogen at various densities and temperatures. The plasma regime is discussed in great detail in Part VI, and the equation of state there is Eq. (3.38). The region of electron degeneracy (to the right of the slanted solid line) is analyzed in Sec. 3.5.4, and for the nonrelativistic regime (between slanted solid line and vertical dotted line) in the second half of Sec. 3.5.2. The boundary between the plasma regime and the electron-degenerate regime (slanted solid line) is Eq. (3.39); that between nonrelativistic degeneracy and relativistic degeneracy (vertical dotted line) is Eq. (3.44). The relativistic/nonrelativistic boundary (upper dashed curve) is governed by electron-positron pair production (Ex. 5.8 and Fig. 5.7). The ionized-neutral boundary (lower dashed curve) is governed by the Saha equation (Ex. 5.9 and Fig. 20.1). For more details on the Plasma regime and its boundaries, see Fig. 20.1.

in fluids (e.g. air and other gases, but not water and other liquids), and even sometimes in solids (e.g. electrons in a metal). Here we shall focus on a nonrelativistic plasma, i.e. the region of Fig. 3.7 that is bounded by the two dashed lines and the slanted solid line. For concreteness and simplicity, we shall regard the plasma as made solely of hydrogen. (This is a good approximation in most astrophysical situations; the modest amounts of helium and traces of other elements usually do not play a major role in equations of state. By contrast, for a laboratory plasma it can be a poor approximation; for quantitative analyses one must pay attention to the plasma's chemical composition.)

A nonrelativistic hydrogen plasma consists of a mixture of two fluids (gases): free electrons and free protons, in equal numbers. Each fluid has a particle number density $n = \rho/m_p$, where ρ is the total mass density and m_p is the proton mass. (The electrons are so light that they do not contribute significantly to ρ.) Correspondingly, the energy density and pressure include equal contributions from the electrons and protons and are given by [cf. Eqs. (3.37b)]

$$U = 3(k_B/m_p)\rho T \ , \quad P = 2(k_B/m_p)\rho T \ . \tag{3.38}$$

In "zero'th approximation", the high-temperature boundary of validity for this equation of state is the temperature $T_{\rm rel} = m_e c^2/k_B = 6 \times 10^9$ K at which the electrons become strongly relativistic (top dashed line in Fig. 3.7). In Ex. 5.8, we shall compute the thermal production of electron-positron pairs in the hot plasma and thereby shall discover that the upper boundary is actually somewhat lower than this (Figs. 5.7 and 20.1). The bottom dashed line in Fig. 3.7 is the temperature $T_{\rm rec} \sim$ (ionization energy of hydrogen)/(a few k_B) $\sim 10^4$ K at which electrons and protons begin to recombine and form neutral hydrogen. In Ex. 5.9 on the Saha equation, we shall analyze the conditions for ionization-recombination equilibrium and thereby shall refine this boundary (Fig. 20.1). The solid right boundary is the point at which the electrons cease to behave like classical particles, because their mean occupation number η_e ceases to be $\ll 1$. As one can see from the Fermi-Dirac distribution (3.22a), for typical electrons (which have energies $E \sim k_B T$), the regime of classical behavior ($\eta_e \ll 1$; left side of solid line) is $\mu_e \ll -k_B T$ and the regime of strong quantum behavior ($\eta_e \simeq 1$; *electron degeneracy*; right side of solid line) is $\mu_e \gg +k_B T$. The slanted solid boundary in Fig. 3.7 is thus the location $\mu_e = 0$, which translates via Eq. (3.37a) to

$$\rho = \rho_{\rm deg} \equiv 2m_p/\lambda_{\rm T\,dB}^3 = (2m_p/h^3)(2\pi m_e k_B T)^{3/2} = 0.01(T/10^4{\rm K})^{3/2}{\rm g/cm}^3 \ . \tag{3.39}$$

Although the hydrogen gas is *degenerate* to the right of this boundary, we can still compute its equation of state using our kinetic-theory equations (3.47b) and (3.47c), so long as we use the quantum mechanically correct distribution function for the electrons—the Fermi-Dirac distribution (3.22a).[5] In this electron-degenerate region, $\mu_e \gg k_B T$, the electron mean occupation number $\eta_e = 1/(e^{(E-\mu_e)/k_B T} + 1)$ has the form shown in Fig. 3.8 and thus can be well approximated by $\eta_e = 1$ for $E = p^2/2m_e < \mu_e$ and $\eta_e = 0$ for $E > \mu_e$; or, equivalently by

$$\eta_e = 1 \ \text{ for } p < p_F \equiv \sqrt{2m_e\mu_e} \ , \quad \eta_e = 0 \ \text{ for } p > p_F \ . \tag{3.40}$$

[5]Our kinetic-theory analysis and the Fermi-Dirac distribution ignore Coulomb interactions between the electrons, and the electrons and the ions. They thereby miss so-called *Coulomb corrections* to the equation of state (Sec. 22.6.3) and other phenomena that are often important in condensed matter physics, but rarely important in astrophysics.

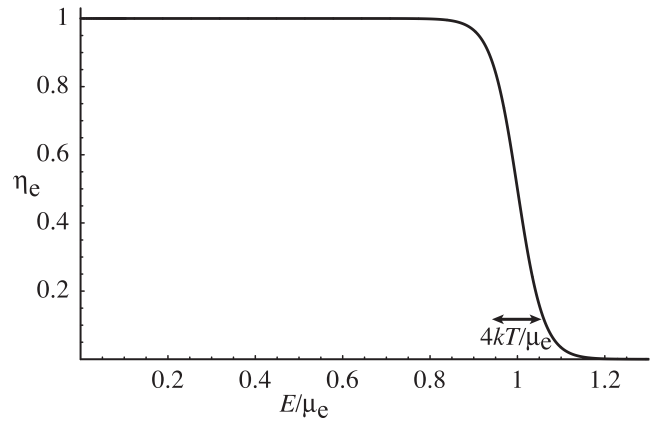

Fig. 3.8: The Fermi-Dirac distribution function for electrons in the nonrelativistic, degenerate regime $k_BT \ll \mu_e \ll m_e$, with temperature such that $k_BT/\mu_e = 0.03$. Note that η_e drops from near one to near zero over the range $\mu_e - 2k_BT \lesssim E \lesssim \mu_e + 2k_BT$. See Ex. 3.10b.

Here p_F is called the *Fermi momentum.* (The word "degenerate" refers to the fact that almost all the quantum states are fully occupied or are empty; i.e., η_e is everywhere nearly one or zero.) By inserting this degenerate distribution function [or, more precisely, $\mathcal{N}_e = (2/h^3)\eta_e$] into Eqs. (3.35) and integrating, we obtain $n_e \propto {p_F}^3$ and $P_e \propto {p_F}^5$. By then setting $n_e = n_p = \rho/m_p$ and solving for $p_F \propto n_e^{1/3} \propto \rho^{1/3}$ and inserting into the expression for P_e and evaluating the constants, we obtain (Ex. 3.8) the following equation of state for the electron pressure

$$P_e = \frac{1}{20}\left(\frac{3}{\pi}\right)^{2/3} \frac{m_e c^2}{\lambda_c^3}\left(\frac{\rho}{m_p/\lambda_c^3}\right)^{5/3} . \tag{3.41}$$

Here

$$\lambda_c = h/m_e c = 2.426 \times 10^{-10}\text{cm} \tag{3.42}$$

is the electron Compton wavelength.

The rapid growth $P_e \propto \rho^{5/3}$ of the electron pressure with increasing density is due to the degenerate electrons' being being confined by the Pauli Exclusion Principle to regions of ever shrinking size, causing their zero-point motions and associated pressure to grow. By contrast, the protons, with their far larger rest masses, remain nondegenerate [until their density becomes $(m_p/m_e)^{3/2} \sim 10^5$ times higher than Eq. (3.39)], and so their pressure is negligible compared to that of the electrons: the total pressure is

$$P = P_e = \text{Eq. (3.41) in the regime of nonrelativistic electron degeneracy.} \tag{3.43}$$

This is the equation of state for the interior of a low-mass white-dwarf star, and for the outer layers of a high-mass white dwarf—aside from tiny corrections due to Coulomb interactions. We shall use it in Sec. 13.3.2 to explore the structures of white dwarfs. It is also the equation of state for a neutron star, with m_e replaced by the rest mass of a neutron m_n (since neutron degeneracy pressure dominates over that due to the star's tiny number of electrons and protons) — except that for neutron stars there are large corrections due to the strong nuclear force; see, e.g., Shapiro and Teukolsky (1983).

When the density of hydrogen, in this degenerate regime, is pushed on upward to

$$\rho_{\rm rel\ deg} = \frac{8\pi m_p}{3\lambda_c^3} \simeq 1.0 \times 10^6 {\rm g/cm}^3 \tag{3.44}$$

(dotted vertical line in Fig. 3.7), the electrons' zero-point motions become relativistically fast (the electron chemical potential μ_e becomes of order $m_e c^2$), so the non-relativistic, Newtonian analysis fails and the matter enters a domain of "relativistic degeneracy" (Sec. 3.5.4 below). Both domains, nonrelativistic degeneracy ($\mu_e \ll m_e c^2$) and relativistic degeneracy ($\mu_e \gtrsim m_e c^2$), occur for matter inside a massive white-dwarf star—the type of star that the Sun will become when it dies; see Shapiro and Teukolsky (1983). In Sec. 26.3.5, we shall see how general relativity (spacetime curvature) modifies a star's structure and helps force sufficiently massive white dwarfs to collapse.

The (almost) degenerate Fermi-Dirac distribution function shown in Fig. 3.8 has a thermal tail whose width is $4kT/\mu_e$. As the temperature T is increased, the number of electrons in this tail increases, thereby increasing the electrons' total energy $E_{\rm tot}$. This increase is responsible for the electrons' *specific heat* (Ex. 3.10)—a quantity of importance for both the electrons in a metal (e.g. a copper wire) and the electrons in a white dwarf star. The electrons dominate the specific heat when the temperature is sufficiently low; but at higher temperatures it is dominated by the energies of sound waves (see Ex. 3.10, where we use the kinetic theory of *phonons* to compute the sound waves' specific heat).

3.5.3 T2 [R] Relativistic Density, Pressure, Energy Density and Equation of State

We turn, now, to the relativistic domain of kinetic theory, initially for a single species of particle with rest mass m and then (in the next subsection) for matter composed of electrons and protons.

The relativistic *mean rest frame* of the particles, at some event $\mathcal{P}$ in spacetime, is that frame in which the particle flux $\mathbf{S}$ vanishes. We shall denote by $\vec{u}_{\rm rf}$ the 4-velocity of this mean rest frame. As in Newtonian theory (above), we are especially interested in distribution functions $\mathcal{N}$ that are *isotropic* in the mean rest frame, i.e., distribution functions that depend on the magnitude $|\mathbf{p}| \equiv p$ of the spatial momentum of a particle, but not on its direction—or equivalently that depend solely on the particles' energy

$$\begin{aligned} \mathcal{E} &= -\vec{u}_{\rm rf} \cdot \vec{p} \quad \text{expressed in frame-independent form [Eq. (2.29)]}, \\ \mathcal{E} &= p^0 = \sqrt{m^2 + p^2} \quad \text{in mean rest frame} . \end{aligned} \tag{3.45}$$

Such isotropy is readily produced by particle collisions (discussed later in this chapter).

Notice that isotropy in the mean rest frame, i.e., $\mathcal{N} = \mathcal{N}(\mathcal{P}, \mathcal{E})$ does not imply isotropy in any other inertial frame. As seen in some other ("primed") frame, $\vec{u}_{\rm rf}$ will have a time component $u^{0'}_{\rm rf} = \gamma$ and a space component $\mathbf{u'}_{\rm rf} = \gamma \mathbf{V}$ [where $\mathbf{V}$ is the mean rest frame's velocity relative to the primed frame and $\gamma = (1 - \mathbf{V}^2)^{-1/2}$]; and correspondingly, in the primed frame $\mathcal{N}$ will be a function of

$$\mathcal{E} = -\vec{u}_{\rm rf} \cdot \vec{p} = \gamma[(m^2 + \mathbf{p'}^2)^{\frac{1}{2}} - \mathbf{V} \cdot \mathbf{p'}] \ , \tag{3.46}$$

which is anisotropic: it depends on the direction of the spatial momentum $\mathbf{p}'$ relative to the velocity $\mathbf{V}$ of the particle's mean rest frame. An example is the cosmic microwave radiation as viewed from Earth, Ex. 3.6 above.

As in Newtonian theory, isotropy greatly simplifies the momentum-space integrals (3.31) that we use to compute macroscopic properties of the particles: (i) The integrands of the expressions $S^j = \int \mathcal{N} p^j (d\mathcal{V}_p/\mathcal{E})$ and $T^{j0} = T^{0j} = \int \mathcal{N} p^j p^0 (d\mathcal{V}_p/\mathcal{E})$ for the particle flux, energy flux and momentum density are all odd in the momentum-space coordinate p^j and therefore give vanishing integrals: $S^j = T^{j0} = T^{0j} = 0$. (ii) The integral $T^{jk} = \int \mathcal{N} p^j p^k d\mathcal{V}_p/\mathcal{E}$ produces an isotropic stress tensor, $T^{jk} = P g^{jk} = P\delta^{jk}$, whose pressure is most easily computed from its trace, $P = \frac{1}{3}T^{jj}$. Using this and the relations $|\mathbf{p}| \equiv p$ for the magnitude of the momentum, $d\mathcal{V}_p = 4\pi p^2 dp$ for the momentum-space volume element, and $\mathcal{E} = p^0 = \sqrt{m^2 + p^2}$ for the particle energy, we can easily evaluate Eqs. (3.31) for the particle number density $n = S^0$, the total density of mass-energy T^{00} (which we shall denote ρ—the same notation as we use for mass density in Newtonian theory), and the pressure P. The results are:

$$n \equiv S^0 = \int \mathcal{N} d\mathcal{V}_p = 4\pi \int_0^\infty \mathcal{N} p^2 dp \; , \tag{3.47a}$$

$$\rho \equiv T^{00} = \int \mathcal{N}\mathcal{E} d\mathcal{V}_p = 4\pi \int_0^\infty \mathcal{N}\mathcal{E} p^2 dp \; , \tag{3.47b}$$

$$P = \frac{1}{3}\int \mathcal{N} p^2 \frac{d\mathcal{V}_p}{\mathcal{E}} = \frac{4\pi}{3}\int_0^\infty \mathcal{N} \frac{p^4 dp}{\sqrt{m^2+p^2}} \; . \tag{3.47c}$$

3.5.4 T2 [R] Equation of State for a Relativistic Degenerate Hydrogen Gas

Return to the hydrogen gas whose nonrelativistic equations of state were computed in Sec. 3.5.1. As we deduced there, at densities $\rho \gtrsim 10^5 \mathrm{g/cm}^3$ (near and to the right of the vertical dotted line in Fig. 3.7) the electrons are squeezed into such tiny volumes that their zero-point energies are $\gtrsim m_e c^2$, forcing us to treat them relativistically.

We can do so with the aid of the following approximation for the relativistic Fermi-Dirac mean occupation number $\eta_e = 1/[e^{(\mathcal{E}-\tilde{\mu}_e/k_B T)} + 1]$:

$$\eta_e \;\simeq\; 1 \;\text{ for } \mathcal{E} < \tilde{\mu}_e \equiv \mathcal{E}_F\text{; i.e., for } p < p_F = \sqrt{\mathcal{E}_F^2 - m^2} \; , \tag{3.48}$$

$$\eta_e \;\simeq\; 0 \;\text{ for } \mathcal{E} > \mathcal{E}_F\text{; i.e., for } p > p_F \; . \tag{3.49}$$

Here $\mathcal{E}_F$ is called the relativistic *Fermi energy* and p_F the relativistic *Fermi momentum.* By inserting this η_e along with $\mathcal{N}_e = (2/h^3)\eta_e$ into the integrals (3.47) for the electrons' number density n_e, total density of mass-energy ρ_e and pressure, P_e, and performing the integrals (Ex. 3.9), we obtain results that are expressed most simply in terms of a parameter t (not to be confused with time) defined by

$$\mathcal{E}_F \equiv \tilde{\mu}_e \equiv m_e \cosh(t/4) \; , \quad p_F \equiv \sqrt{\mathcal{E}_F^2 - m_e^2} \equiv m_e \sinh(t/4) \; . \tag{3.50a}$$

The results are:

$$n_e = \frac{8\pi}{3\lambda_c^3}\left(\frac{p_F}{m_e}\right)^3 = \frac{8\pi}{3\lambda_c^3}\sinh^3(t/4) \;, \tag{3.50b}$$

$$\rho_e = \frac{8\pi\, m_e}{\lambda_c^3}\int_0^{p_F/m_e} x^2\sqrt{1+x^2}\,dx = \frac{\pi\, m_e}{4\lambda_c^3}[\sinh(t) - t] \;, \tag{3.50c}$$

$$P_e = \frac{8\pi\, m_e}{\lambda_c^3}\int_0^{p_F/m_e} \frac{x^4}{\sqrt{1+x^2}}dx = \frac{\pi\, m_e}{12\lambda_c^3}[\sinh(t) - 8\sinh(t/2) + 3t] \;. \tag{3.50d}$$

These parametric relationships for ρ_e and P_e as functions of the electron number density n_e are sometimes called the Anderson-Stoner equation of state, because they were first derived by Wilhelm Anderson and Edmund Stoner in 1930 [see the history on pp. 153–154 of Thorne (1994)]. They are valid throughout the full range of electron degeneracy, from nonrelativistic up to ultra-relativistic.

In a white-dwarf star, the protons, with their high rest mass, are nondegenerate, the total density of mass-energy is dominated by the proton rest-mass density, and since there is one proton for each electron in the hydrogen gas, that total is

$$\rho \simeq m_p n_e = \frac{8\pi m_p}{3\lambda_c^3}\sinh^3(t/4) \;. \tag{3.51a}$$

By contrast (as in the nonrelativistic regime), the pressure is dominated by the electrons (because of their huge zero-point motions), not the protons; and so the total pressure is

$$P = P_e = \frac{\pi\, m_e}{12\lambda_c^3}[\sinh(t) - 8\sinh(t/2) + 3t] \;. \tag{3.51b}$$

In the low-density limit, where $t \ll 1$ so $p_F \ll m_e = m_e c$, we can solve the relativistic equation (3.50b) for t as a function of $n_e = \rho/m_p$ and insert the result into the relativistic expression (3.51b); the result is the nonrelativistic equation of state (3.41).

The dividing line $\rho = \rho_{\rm rel\ deg} = 8\pi m_p/3\lambda_c^3 \simeq 1.0 \times 10^6 \text{g/cm}^3$ between nonrelativistic and relativistic degeneracy is the point where the electron Fermi momentum is equal to the electron rest mass, i.e. $\sinh(t/4) = 1$. The equation of state (3.51a), (3.51b) implies

$$\begin{aligned} P_e &\propto \rho^{5/3} \quad \text{in the nonrelativistic regime, } \rho \ll \rho_{\rm rel\ deg} \;, \\ P_e &\propto \rho^{4/3} \quad \text{in the relativistic regime, } \rho \gg \rho_{\rm rel\ deg} \;. \end{aligned} \tag{3.51c}$$

These asymptotic equations of state turn out to play a crucial role in the structure and stability of white-dwarf stars [Secs. 13.3.2 and 26.3.5; Chap. 4 of Thorne (1994)].

3.5.5 [N & R] Equation of State for Radiation

As was discussed at the end of Sec. 3.3, for a gas of thermalized photons in an environment where photons are readily created and absorbed, the distribution function has the black-body (Planck) form $\eta = 1/(e^{\mathcal{E}/k_B T} - 1)$, which we can rewrite as $1/(e^{p/k_B T} - 1)$ since the energy

$\mathcal{E}$ of a photon is the same as the magnitude p of its momentum. In this case, the relativistic integrals (3.47) give (see Ex. 3.12)

$$\boxed{n = bT^3 \; , \quad \rho = aT^4 \; , \quad P = \frac{1}{3}\rho} \; , \tag{3.52a}$$

where

$$b = 16\pi\zeta(3)\frac{k_B^3}{h^3c^3} = 20.28\,\mathrm{cm}^{-3}\mathrm{K}^{-3} \; , \tag{3.52b}$$

$$a = \frac{8\pi^5}{15}\frac{k_B^4}{h^3c^3} = 7.56\times 10^{-15}\mathrm{erg\,cm}^{-3}\,\mathrm{K}^{-4} = 7.56\times 10^{-16}\mathrm{J\,m}^{-3}\,\mathrm{K}^{-4} \tag{3.52c}$$

are *radiation constants.* Here $\zeta(3) = \sum_{n=1}^{\infty} n^{-3} = 1.2020569...$ is the Riemann Zeta function.

As we shall see in Sec. 28.4, when the Universe was younger than about 100,000 years, its energy density and pressure were predominantly due to thermalized photons, plus neutrinos which contributed approximately the same as the photons, so its equation of state was given by Eq. (3.52a) with the coefficient changed by a factor of order unity. Einstein's general relativistic field equations (Part VII) relate the energy density ρ of these photons and neutrinos to the age of the universe τ as measured in the photons' and neutrinos' mean rest frame:

$$\frac{3}{32\pi G\tau^2} = \rho \simeq aT^4 \tag{3.53a}$$

[Eq. (28.45)]. Here G is Newton's gravitation constant. Putting in numbers, we find that

$$\rho = \frac{4.9\times 10^{-12}\mathrm{g/cm}^3}{(\tau/1\mathrm{yr})^2} \; , \quad T \simeq \frac{10^6\mathrm{K}}{\sqrt{\tau/1\mathrm{yr}}} \; . \tag{3.53b}$$

This implies that, when the universe was one minute old, its radiation density and temperature were about 1 $\mathrm{g/cm}^3$ and 6×10^8 K. These conditions were well suited for burning hydrogen to helium; and, indeed, about 1/4 of all the mass of the universe did get burned to helium at this early epoch. We shall examine this in further detail in Sec. 28.4.

****

EXERCISES

Exercise 3.7 *Derivation & Practice:* [N] *Equation of State for Nonrelativistic, Classical Gas*

Consider a collection of identical, classical (i.e., with $\eta \ll 1$) particles with a distribution function $\mathcal{N}$ which is thermalized at a temperature T such that $k_BT \ll mc^2$ (nonrelativistic temperature).

(a) Show that the distribution function, expressed in terms of the particles' momentum or velocity in their mean rest frame, is

$$\mathcal{N} = \frac{g_s}{h^3} e^{\mu/k_BT} e^{-p^2/2mk_BT} \; , \quad \text{where } p = |\mathbf{p}| = mv \; , \tag{3.54}$$

with v being the speed of a particle.

(b) Show that the number density of particles in the mean rest frame is given by Eq. (3.37a).

(c) Show that this gas satisfies the equations of state (3.37b).

Note: The following integrals, for nonnegative integral values of q, will be useful:

$$\int_0^\infty x^{2q} e^{-x^2} dx = \frac{(2q-1)!!}{2^{q+1}} \sqrt{\pi} \ , \tag{3.55}$$

where $n!! \equiv n(n-2)(n-4)\ldots(2 \text{ or } 1)$; and

$$\int_0^\infty x^{2q+1} e^{-x^2} dx = \frac{1}{2} q! \ . \tag{3.56}$$

Exercise 3.8 *Derivation and Practice:* [N] *Equation of State for Nonrelativistic, Electron-Degenerate Hydrogen*
Derive Eq. (3.41) for the electron pressure in a nonrelativistic, electron-degenerate hydrogen gas.

Exercise 3.9 *Derivation and Practice:* **T2** [R] *Equation of State for Relativistic, Electron-Degenerate Hydrogen*
Derive the equation of state (3.50) for an electron-degenerate hydrogen gas. (Note: It might be easiest to compute the integrals with the help of symbolic manipulation software such as Mathematica or Maple.)

Exercise 3.10 *Example:* [N] *Specific Heat for Nonrelativistic, Degenerate Electrons in White Dwarfs and in Metals*
Consider a nonrelativistically degenerate electron gas at finite but small temperature.

(a) Show that the inequalities $k_B T \ll \mu_e \ll m_e$ are equivalent to the words "nonrelativistically degenerate".

(b) Show that the electrons' mean occupation number $\eta_e(E)$ has the form depicted in Fig. 3.8: It is near unity out to (nonrelativistic) energy $E \simeq \mu_e - 2k_B T$, and it then drops to nearly zero over a range of energies $\Delta E \sim 4k_B T$.

(c) If the electrons were nonrelativistic but *non*degenerate, their thermal energy density would be $U = \frac{3}{2} n k_B T$, so the total electron energy (excluding rest mass) in a volume V containing $N = nV$ electrons would be $E_{\rm tot} = \frac{3}{2} N k_B T$, and the electron specific heat, at fixed volume, would be

$$C_V \equiv \left(\frac{\partial E_{\rm tot}}{\partial T}\right)_V = \frac{3}{2} N k_B \quad \text{(nondegenerate, nonrelativistic).} \tag{3.57}$$

Using the semiquantitative form of η_e depicted in Fig. 3.8, show that to within a factor of order unity the specific heat of degenerate electrons is smaller than in the nondegenerate case by a factor $\sim k_B T/\mu_e$:

$$C_V \equiv \left(\frac{\partial E_{\rm tot}}{\partial T}\right)_V \sim \left(\frac{k_B T}{\mu_e}\right) N k_B \quad \text{(degenerate, nonrelativistic).} \tag{3.58}$$

(d) Compute the multiplicative factor in this equation for C_V. More specifically, show that, to first order in k_BT/μ_e,

$$C_V = \frac{\pi^2}{2}\left(\frac{k_BT}{\mu_e}\right) N k_B \ . \tag{3.59}$$

(e) As an application, consider hydrogen inside a white dwarf star with density $\rho = 10^5$ g cm^{-3} and temperature $T = 10^6$ K. (These are typical values for a white dwarf interior). What are the numerical values of μ_e/m_e and k_BT/μ_e for the electrons? What is the numerical value of the dimensionless factor $(\pi^2/2)(k_BT/\mu_e)$ by which degeneracy reduces the electron specific heat?

(f) As a second application, consider the electrons inside a copper wire in a laboratory on Earth at room temperature. Each copper atom donates about one electron to a "gas" of freely traveling (conducting) electrons, and keeps the rest of its electrons bound to itself. (We neglect interaction of this electron gas with the ions, thereby missing important condensed-matter complexities such as conduction bands and what distinguishes conducting materials from insulators.)

What are the numerical values of μ_e/m_e and k_BT/μ_e for the conducting electron "gas"? Verify that these are in the range corresponding to nonrelativistic degeneracy. What is the value of the factor $(\pi^2/2)(k_BT/\mu_e)$ by which degeneracy reduces the electron specific heat? At room temperature, this electron contribution to the specific heat is far smaller than the contribution from thermal vibrations of the copper atoms (i.e., thermal sound waves, i.e. thermal *phonons*), but at very low temperatures the electron contribution dominates, as we shall see in the next exercise.

Exercise 3.11 *Example:* [N] *Specific Heat for Phonons in an Isotropic Solid*

In Sec. 12.2 we will study classical sound waves propagating through an isotropic, elastic solid. As we shall see, there are two types of sound waves: *longitudinal* with frequency-independent speed C_L, and *transverse* with a somewhat smaller frequency-independent speed C_T. For each type of wave, $s = L$ or T, the material of the solid undergoes an elastic displacement $\boldsymbol{\xi} = A\mathbf{f}_s \exp(i\mathbf{k}\cdot\mathbf{x} - \omega t)$, where A is the wave amplitude, $\mathbf{f}_s$ is a unit vector (polarization vector) pointing in the direction of the displacement, $\mathbf{k}$ is the wave vector, and ω is the wave frequency. The wave speed is $C_s = \omega/|\mathbf{k}|$ $(= C_L$ or $C_T)$. Associated with these waves are quanta called *phonons*. As for any wave, each phonon has a momentum related to its wave vector by $\mathbf{p} = \hbar\mathbf{k}$, and an energy related to its frequency by $E = \hbar\omega$. Combining these relations we learn that the relationship between a phonon's energy and the magnitude $p = |\mathbf{p}|$ of its momentum is $E = C_s p$. This is the same relationship as for photons, but with the speed of light replaced by the speed of sound! For longitudinal waves, $\mathbf{f}_L$ is in the propagation direction $\mathbf{k}$ so there is just one polarization, $g_L = 1$. For transverse waves, $\mathbf{f}_T$ is orthogonal to $\mathbf{k}$, so there are two orthogonal polarizations (e.g. $\mathbf{f}_T = \mathbf{e}_x$ and $\mathbf{f}_T = \mathbf{e}_y$ when $\mathbf{k}$ points in the $\mathbf{e}_z$ direction); i.e., $g_T = 2$.

(a) Phonons of both types, longitudinal and transverse, are bosons. Why? [Hint: each normal mode of an elastic body can be described, mathematically, as a harmonic oscillator.]

(b) Phonons are fairly easily created, absorbed, scattered and thermalized. A general argument regarding chemical reactions (Sec. 5.5) can be applied to phonon creation and absorption to deduce that, once they reach complete thermal equilibrium with their environment, the phonons will have vanishing chemical potential $\mu = 0$. What, then, will be their distribution functions η and $\mathcal{N}$?

(c) Ignoring the fact that the sound waves' wavelengths $\lambda = 2\pi/|\mathbf{k}|$ cannot be smaller than about twice the spacing between the atoms of the solid, show that the total phonon energy (wave energy) in a volume V of the solid is identical to that for black-body photons in a volume V, but with the speed of light c replaced by the speed of sound C_s, and with the photons' number of spin states, 2, replaced by g_s (2 for transverse waves, 1 for longitudinal): $E_{\rm tot} = a_s T^4 V$, with $a_s = g_s(4\pi^5/15)(k_B^4/h^3 C_s^3)$; cf. Eqs. (3.52).

(d) Show that the specific heat of the phonon gas (the sound waves) is $C_V = 4a_s T^3 V$. This scales as T^3, whereas in a metal the specific heat of the degenerate electrons scales as T [previous exercise], so at sufficiently low temperatures the electron specific heat will dominate over that of the phonons.

(e) Show that in the phonon gas, only phonon modes with wavelengths longer than $\sim \lambda_T = C_s h/k_B T$ are excited; i.e., for $\lambda \ll \lambda_T$ the mean occupation number is $\eta \ll 1$; for $\lambda \sim \lambda_T$, $\eta \sim 1$; and for $\lambda \gg \lambda_T$, $\eta \gg 1$. As T is increased, λ_T gets reduced. Ultimately it becomes of order the interatomic spacing, and our computation fails because most of the modes that our calculation assumes are thermalized actually don't exist. What is the critical temperature (*Debye temperature*) at which our computation fails and the T^3 law for C_V changes? Show by a $\sim$ one-line argument that above the Debye temperature C_V is independent of temperature.

Exercise 3.12 *Derivation & Practice:* [N, R] *Equation of State for a Photon Gas*

(a) Consider a collection of photons with a distribution function $\mathcal{N}$ which, in the mean rest frame of the photons, is isotropic. Show, using Eqs. (3.47b) and (3.47c), that this photon gas obeys the equation of state $P = \frac{1}{3}\rho$.

(b) Suppose the photons are thermalized with zero chemical potential, i.e., they are isotropic with a black-body spectrum. Show that $\rho = aT^4$, where a is the radiation constant of Eq. (3.52c). *Note*: Do not hesitate to use Mathematica or Maple or other computer programs to evaluate integrals!

(c) Show that for this isotropic, black-body photon gas the number density of photons is $n = bT^3$ where b is given by Eq. (3.52b), and that the mean energy of a photon in the gas is

$$\bar{\mathcal{E}}_\gamma = \frac{\pi^4}{30\zeta(3)} k_B T = 2.7011780...\, k_B T \; . \tag{3.60}$$

****

3.6 [N & R] Evolution of the Distribution Function: Liouville's Theorem, the Collisionless Boltzmann Equation, and the Boltzmann Transport Equation

We now turn to the issue of how the distribution function $\eta(\mathcal{P}, \vec{p})$, or equivalently $\mathcal{N} = (g_s/h^3)\eta$, evolves from point to point in phase space. We shall explore the evolution under the simple assumption that between their very brief collisions, the particles all move freely, uninfluenced by any forces. It is straightforward to generalize to a situation where the particles interact with electromagnetic or gravitational or other fields as they move, and we shall do so in Box 3.2, Sec. 4.3, and Box 28.4. However, in the body of this chapter we shall restrict attention to the very common situation of free motion between collisions.

Initially we shall even rule out collisions; only at the end of this section will we restore them, by inserting them as an additional term in our collision-free evolution equation for η.

The foundation for our collision-free evolution law will be *Liouville's Theorem*: Consider a set $\mathcal{S}$ of particles which are initially all near some location in phase space and initially occupy an infinitesimal (frame-independent) phase-space volume $d^2\mathcal{V} = d\mathcal{V}_x d\mathcal{V}_p$ there. Pick a particle at the center of the set $\mathcal{S}$ and call it the "fiducial particle". Since all the particles in $\mathcal{S}$ have nearly the same initial position and velocity, they subsequently all move along nearly the same trajectory (world line); i.e., they all remain congregated around the fiducial particle. Liouville's theorem says that the phase-space volume occupied by the particles $\mathcal{S}$ is conserved along the trajectory of the fiducial particle:

$$\frac{d}{d\ell}(d\mathcal{V}_x d\mathcal{V}_p) = 0 \; ; . \tag{3.61}$$

Here ℓ is an arbitrary parameter along the trajectory; for example, in Newtonian theory it could be universal time t or distance l travelled, and in relativity it could be proper time τ as measured by the fiducial particle (if its rest mass is nonzero) or the affine parameter ζ that is related to the fiducial particle's 4-momentum by $\vec{p} = d\vec{x}/d\zeta$.

We shall prove Liouville's theorem with the aid of the diagrams in Fig. 3.9. Assume, for simplicity, that the particles have nonzero rest mass. Consider the region in phase space occupied by the particles, as seen in the inertial reference frame (rest frame) of the fiducial

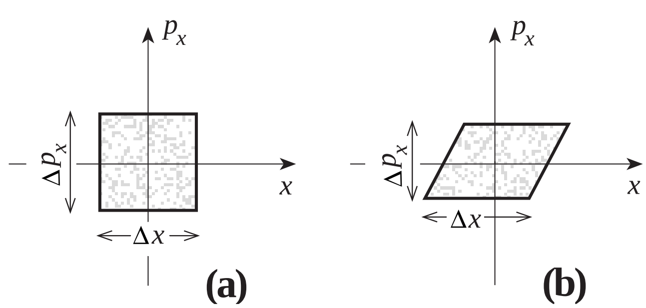

Fig. 3.9: The phase space region (x-p_x part) occupied by a set $\mathcal{S}$ of particles with finite rest mass, as seen in the inertial frame of the central, fiducial particle. (a) The initial region. (b) The region after a short time.

particle, and choose for ℓ the time t of that inertial frame (or in Newtonian theory the universal time t). Choose the particles' region $d\mathcal{V}_x d\mathcal{V}_p$ at $t = 0$ to be a rectangular box centered on the fiducial particle, i.e. on the origin $x^j = 0$ of its inertial frame (Fig. 3.9a). Examine the evolution with time t of the 2-dimensional slice $y = p_y = z = p_z = 0$ through the occupied region. The evolution of other slices will be similar. Then, as t passes, the particle at location (x, p_x) moves with velocity $dx/dt = p_x/m$, (where the nonrelativistic approximation to the velocity is used because all the particles are very nearly at rest in the fiducial particle's inertial frame). Because the particles move freely, each has a conserved p_x, and their motion $dx/dt = p_x/m$ (larger speed higher in the diagram) deforms the particles' phase space region into a skewed parallelogram as shown in Fig. 3.9b. Obviously, the area of the occupied region, $\Delta x \Delta p_x$, is conserved.

This same argument shows that the x-p_x area is conserved at *all* values of y, z, p_y, p_z; and similarly for the areas in the y-p_y planes and the areas in the z-p_z planes. As a consequence, the total volume in phase space, $d\mathcal{V}_x d\mathcal{V}_p = \Delta x \Delta p_x \Delta y \Delta p_y \Delta z \Delta p_z$ is conserved.

Although this proof of Liouville's theorem relied on the assumption that the particles have nonzero rest mass, the theorem is also true for particles with zero rest mass—as one can deduce by taking the relativistic limit as the rest mass goes to zero and the particles' 4-momenta become null.

Since, in the absence of collisions or other nongravitational interactions, the number dN of particles in the set $\mathcal{S}$ is conserved, Liouville's theorem immediately implies also the conservation of the number density in phase space, $\mathcal{N} = dN/d\mathcal{V}_x d\mathcal{V}_p$:

$$\boxed{\frac{d\mathcal{N}}{d\ell} = 0 \quad \text{along the trajectory of a fiducial particle}} \; . \tag{3.62}$$

This conservation law is called the *collisionless Boltzmann equation*; and in the context of plasma physics (Part VI) it is sometimes called the *Vlasov equation.* Note that it says that *not only is the distribution function $\mathcal{N}$ frame-independent; $\mathcal{N}$ also is constant along the phase-space trajectory of any freely moving particle.*

The collisionless Boltzmann equation is actually far more general than is suggested by the above derivation; see Box 3.2, which is best read after finishing this section.

The collisionless Boltzmann equation is most nicely expressed in the frame-independent form (3.62). For some purposes, however, it is helpful to express the equation in a form that relies on a specific but arbitrary choice of inertial reference frame. Then $\mathcal{N}$ can be regarded as a function of the reference frame's seven phase-space coordinates, $\mathcal{N} = \mathcal{N}(t, x^j, p_k)$, and the collisionless Boltzmann equation (3.62) takes the coordinate-dependent form

$$\frac{d\mathcal{N}}{d\ell} = \frac{dt}{d\ell}\frac{\partial \mathcal{N}}{\partial t} + \frac{dx^j}{d\ell}\frac{\partial \mathcal{N}}{\partial x^j} + \frac{dp_j}{d\ell}\frac{\partial \mathcal{N}}{\partial p_j} = \frac{dt}{d\ell}\left(\frac{\partial \mathcal{N}}{\partial t} + v_j \frac{\partial \mathcal{N}}{\partial x_j}\right) = 0 \; . \tag{3.63}$$

Here we have used the equation of straight-line motion $dp_j/dt = 0$ for the particles and have set dx^j/dt equal to the particle velocity $v^j = v_j$.

Since our derivation of the collisionless Boltzmann equation relied on the assumption that no particles are created or destroyed as time passes, the collisionless Boltzmann equation in turn should guarantee conservation of the number of particles, $\partial n/\partial t + \boldsymbol{\nabla} \cdot \mathbf{S} = 0$ in

Newtonian theory (Sec. 1.8), and $\vec{\nabla} \cdot \vec{S} = 0$ relativistically (Sec. 2.12.3). Indeed, this is so; see Ex. 3.13. Similarly, since the collisionless Boltzmann equation is based on the law of momentum (or 4-momentum) conservation for all the individual particles, it is reasonable to expect that the collisionless Boltzmann equation will guarantee the conservation of their total Newtonian momentum [$\partial \boldsymbol{\mathcal{G}}/\partial t + \boldsymbol{\nabla} \cdot \boldsymbol{T} = 0$, Eq. (1.36)] and their relativistic 4-momentum [$\vec{\nabla} \cdot \boldsymbol{T} = 0$, Eq. (2.73a)]. Indeed, these conservation laws do follow from the collisionless Boltzmann equation; see Ex. 3.13.

Thus far we have assumed that the particles move freely through phase space with no collisions. If collisions occur, they will produce some nonconservation of $\mathcal{N}$ along the trajectory of a freely moving, noncolliding fiducial particle, and correspondingly, the collisionless Boltzmann equation will get modified to read

$$\boxed{\frac{d\mathcal{N}}{d\ell} = \left(\frac{d\mathcal{N}}{d\ell}\right)_{\text{collisions}}}\,, \tag{3.64}$$

where the right-hand side represents the effects of collisions. This equation, with collision terms present, is called the *Boltzmann transport equation*. The actual form of the collision terms depends, of course, on the details of the collisions. We shall meet some specific examples in the next section [Eqs. (3.77), (3.84a), (3.85), and Ex. 3.20] and in our study of plasmas (Chaps. 22 and 23).

Whenever one applies the collisionless Boltzmann equation or Boltzmann transport equation to a given situation, it is helpful to simplify one's thinking in two ways: (i) Adjust the normalization of the distribution function so it is naturally tuned to the situation. For example, when dealing with photons, I_ν/ν^3 is typically best, and if—as is usually the case—the photons do not change their frequencies as they move and only a single reference frame is of any importance, then I_ν alone may do; see Ex. 3.14. (ii) Adjust the differentiation parameter ℓ so it is also naturally tuned to the situation.

****

EXERCISES

Exercise 3.13 [N & R] *Derivation and Problem: Collisionless Boltzmann Equation Implies Conservation of Particles and of 4-Momentum*
Consider a collection of freely moving, noncolliding particles, which satisfy the collisionless Boltzmann equation $d\mathcal{N}/d\ell = 0$.

(a) Show that this equation guarantees that the Newtonian particle conservation law $\partial n/\partial t + \boldsymbol{\nabla} \cdot \mathbf{S} = 0$ and momentum conservation law $\partial \boldsymbol{\mathcal{G}}/\partial t + \boldsymbol{\nabla} \cdot \boldsymbol{T} = 0$ are satisfied, where n, $\mathbf{S}$, $\boldsymbol{\mathcal{G}}$ and $\boldsymbol{T}$ are expressed in terms of the distribution function $\mathcal{N}$ by the Newtonian momentum-space integrals (3.30).

(b) Show that the relativistic Boltzmann equation guarantees the relativistic conservation laws $\vec{\nabla} \cdot \vec{S} = 0$ and $\vec{\nabla} \cdot \boldsymbol{T} = 0$, where the number-flux 4-vector $\vec{S}$ and the stress-energy tensor $\boldsymbol{T}$ are expressed in terms of $\mathcal{N}$ by the momentum-space integrals (3.31).

Box 3.2

T2 Sophisticated Derivation of Relativistic Collisionless Boltzmann Equation

Denote by $\vec{X} \equiv \{\mathcal{P}, \vec{p}\}$ a point in 8-dimensional phase space. In an inertial frame the coordinates of $\vec{X}$ are $\{x^0, x^1, x^2, x^3, p_0, p_1, p_2, p_3\}$. [We use up indices ("contravariant" indices) on x and down indices ("covariant" indices) on p because this is the form required in Hamilton's equations below; i.e., it is p_α not p^α that is canonically conjugate to x^α.] Regard $\mathcal{N}$ as a function of location $\vec{X}$ in 8-dimensional phase space. The fact that our particles all have the same rest mass so $\mathcal{N}$ is nonzero only on the mass hyperboloid means that as a function of $\vec{X}$, $\mathcal{N}$ entails a delta function. For the following derivation, that delta function is irrelevant; the derivation is valid also for distributions of non-identical particles, as treated in Ex. 3.2.

A particle in our distribution $\mathcal{N}$ at location $\vec{X}$ moves through phase space along a world line with tangent vector $d\vec{X}/d\zeta$, where ζ is its affine parameter. The product $\mathcal{N} d\vec{X}/d\zeta$ represents the number-flux 8-vector of particles through spacetime, as one can see by an argument analogous to Eq. (3.33c). We presume that, as the particles move through phase space, none are created or destroyed. The law of particle conservation in phase space, by analogy with $\vec{\nabla} \cdot \vec{S} = 0$ in spacetime, takes the form $\vec{\nabla} \cdot (\mathcal{N} d\vec{X}/d\zeta) = 0$. In terms of coordinates in an inertial frame, this conservation law says

$$\frac{\partial}{\partial x^\alpha}\left(\mathcal{N}\frac{dx^\alpha}{d\zeta}\right) + \frac{\partial}{\partial p^\alpha}\left(\mathcal{N}\frac{dp^\alpha}{d\zeta}\right) = 0 \; . \tag{1}$$

The motions of individual particles in phase space are governed by Hamilton's equations

$$\frac{dx^\alpha}{d\zeta} = \frac{\partial \mathcal{H}}{\partial p^\alpha} \; , \quad \frac{dp^\alpha}{d\zeta} = -\frac{\partial \mathcal{H}}{\partial x^\alpha} \; . \tag{2}$$

For the freely moving particles of this chapter, the relativistic Hamiltonian is [cf. Sec. 8.4 of Goldstein, Poole and Safko (2002) or p. 489 of Misner, Thorne and Wheeler (1973)]

$$\mathcal{H} = \frac{1}{2}(p_\alpha p_\beta g^{\alpha\beta} - m^2) \; . \tag{3}$$

Our derivation of the collisionless Boltzmann equation does not depend on this specific form of the Hamiltonian; it is valid for any Hamiltonian and thus, e.g., for particles interacting with an electromagnetic field or even a relativistic gravitational field (spacetime curvature; Part VII). By inserting Hamilton's equations (2) into the 8-dimensional law of particle conservation (1), we obtain

$$\frac{\partial}{\partial x^\alpha}\left(\mathcal{N}\frac{\partial \mathcal{H}}{\partial p^\alpha}\right) - \frac{\partial}{\partial p^\alpha}\left(\mathcal{N}\frac{\partial \mathcal{H}}{\partial x^\alpha}\right) = 0 \; . \tag{4}$$

Using the rule for differentiating products, and noting that the terms involving two

Box 3.2, Continued

derivatives of $\mathcal{H}$ cancel, we bring this into the form

$$0 = \frac{\partial \mathcal{N}}{\partial x^\alpha}\frac{\partial \mathcal{H}}{\partial p^\alpha} - \frac{\partial \mathcal{N}}{\partial p^\alpha}\frac{\partial \mathcal{H}}{\partial x^\alpha} = \frac{\partial \mathcal{N}}{\partial x^\alpha}\frac{dx^\alpha}{d\zeta} - \frac{\partial \mathcal{N}}{\partial p^\alpha}\frac{dp^\alpha}{d\zeta} = \frac{d\mathcal{N}}{d\zeta} \;, \tag{5}$$

which is the collisionless Boltzmann equation. [To get the second expression we have used Hamilton's equations, and the third follows directly from the formulas of differential calculus.] *Thus, the collisionless Boltzmann equation is a consequence of just two assumptions: conservation of particles and Hamilton's equations for the motion of each particle. This implies it has very great generality.* We shall extend and explore this generality in the next chapter.

Exercise 3.14 ***Example:* [N] *Solar Heating of the Earth: The Greenhouse Effect*
In this example we shall study the heating of the Earth by the Sun. Along the way, we shall derive some important relations for black-body radiation.

Since we will study photon propagation from the Sun to the Earth with Doppler shifts playing a negligible role, there is a preferred inertial reference frame: that of the Sun and Earth with their relative motion neglected. We shall carry out our analysis in that frame. Since we are dealing with thermalized photons, the natural choice for the distribution function is I_ν/ν^3; and since we use just one unique reference frame, each photon has a fixed frequency ν, so we can forget about the ν^3 and use I_ν.

(a) Assume, as is very nearly true, that each spot on the sun emits black-body radiation in all outward directions with a temperature $T_\odot = 5800$ K. Show, by integrating over the black-body I_ν, that the total energy flux (i.e. power per unit surface area) F emitted by the Sun is

$$F \equiv \frac{d\mathcal{E}}{dtdA} = \sigma T_\odot^4 \;, \quad \text{where } \sigma = \frac{ac}{4} = \frac{2\pi^5}{15}\frac{k_B^4}{h^3c^2} = 5.67\times 10^{-5}\frac{\text{erg}}{\text{cm}^2\text{sK}^4} \;. \tag{3.65}$$

b Since the distribution function I_ν is conserved along each photon's trajectory, observers on Earth, looking at the sun, see identically the same black-body spectral intensity I_ν as they would if they were on the Sun's surface. (No wonder our eyes hurt if we look directly at the Sun!) By integrating over I_ν at the Earth [and not by the simpler method of using Eq. (3.65) for the flux leaving the Sun], show that the energy flux arriving at Earth is $F = \sigma T_\odot^4 (R_\odot/r)^2$, where $R_\odot = 696,000$ km is the Sun's radius and $r = 1.496\times 10^8$ km is the distance from the Sun to Earth.

(c) Our goal is to compute the temperature $T_\oplus$ of the Earth's surface. As a first attempt, assume that all the Sun's flux arriving at Earth is absorbed by the Earth's surface, heating it to the temperature $T_\oplus$, and then is reradiated into space as black-body radiation at temperature $T_\oplus$. Show that this leads to a surface temperature of

$$T_\oplus = T_\odot \left(\frac{R_\odot}{2r}\right)^{1/2} = 280\,\text{K} = 7\,\text{C} \;. \tag{3.66}$$

This is a bit cooler than the correct mean surface temperature (287 K = 14 C).

(d) Actually, the Earth has an "albedo" of $A \simeq 0.30$, which means that 30 per cent of the sunlight that falls onto it gets reflected back into space with an essentially unchanged spectrum, rather than being absorbed. Show that with only a fraction $1 - A = 0.70$ of the solar radiation being absorbed, the above estimate of the Earth's temperature becomes

$$T_\oplus = T_\odot \left(\frac{\sqrt{1-A}\, R_\odot}{2r} \right)^{1/2} = 256\,\mathrm{K} = -17\,\mathrm{C}\ . \tag{3.67}$$

This is even farther from the correct answer.

(e) The missing piece of physics, which raises the temperature from −17C to something much nearer the correct 14C, is the *Greenhouse Effect*: The absorbed solar radiation has most of its energy at wavelengths $\sim 0.5\mu$m (in the visual band), which pass rather easily through the Earth's atmosphere. By contrast, the black-body radiation that the Earth's surface wants to radiate back into space, with its temperature $\sim$ 300K, is concentrated in the infrared range from $\sim 8\mu$m to $\sim 30\mu$m. Water molecules and carbon dioxide in the Earth's atmosphere absorb about 40% of the energy that the Earth tries to reradiate at these energies,[6] causing the reradiated energy to be about 60% that of a black body at the Earth's surface temperature. Show that with this (oversimplified!) "Greenhouse" correction, $T_\oplus$ becomes about 290K = +17C, which is within a few degrees of the true mean temperature. Of course, the worry is that human activity is increasing the amount of carbon dioxide in the atmosphere by enough to raise the Earth's temperature significantly further and disrupt our comfortable lives.

Exercise 3.15 ***Challenge:* [N] *Olbers' Paradox and Solar Furnace*
Consider a universe (not ours!) in which spacetime is flat and infinite in size and is populated throughout by stars that cluster into galaxies like our own and our neighbors, with interstellar and intergalactic distances similar to those in our neighborhood. Assume that the galaxies are *not* moving apart, i.e., there is no universal expansion. Using the collisionless Boltzmann equation for photons, show that the Earth's temperature in this universe would be about the same as the surface temperatures of the universe's hotter stars, $\sim 10,000$ K, so we would all be fried. This is called Olber's Paradox. What features of our universe protect us from this fate?

Motivated by this model universe, describe a design for a furnace that relies on sunlight for its heat and achieves a temperature nearly equal to that of the sun's surface, 5800 K.

3.7 [N] Transport Coefficients

In this section we turn to a practical application of kinetic theory: the computation of *transport coefficients*. Our primary objective is to illustrate the use of kinetic theory, but

[6] See, e.g., Sec. 11.22 of Allen (2000).

the transport coefficients themselves are also of interest: they will play important roles in Parts V and VI of this book (Fluid Mechanics and Plasma Physics).

What are transport coefficients? An example is *electrical conductivity* κ_e. When an electric field $\mathbf{E}$ is imposed on a sample of matter, Ohm's law tells us that the matter responds by developing a current density

$$\boxed{\mathbf{j} = \kappa_e \mathbf{E}} \,. \tag{3.68a}$$

The electrical conductivity is high if electrons can move through the material with ease; it is low if electrons have difficulty moving. The impediment to electron motion is scattering off other particles—off ions, other electrons, phonons (sound waves), plasmons (plasma waves), ... Ohm's law is valid when (as almost always) the electrons scatter many times, so they *diffuse* (*random-walk their way*) through the material. In order to compute the electrical conductivity, one must analyze, statistically, the effects of the many scatterings on the electrons' motions. The foundation for an accurate analysis is the Boltzmann transport equation.

Another example of a transport coefficient is *thermal conductivity* κ, which appears in the law of heat conduction

$$\boxed{\mathbf{F} = -\kappa \boldsymbol{\nabla} T} \,. \tag{3.68b}$$

Here $\mathbf{F}$ is the diffusive energy flux from regions of high temperature T to low. The impediment to heat flow is scattering of the conducting particles; and, correspondingly, the foundation for accurately computing κ is the Boltzmann transport equation.

Other examples of transport coefficients are (i) the *coefficient of shear viscosity* $\eta_{\rm shear}$ which determines the stress T_{ij} (diffusive flux of momentum) that arises in a shearing fluid

$$\boxed{T_{ij} = -2\eta_{\rm shear}\sigma_{ij}} \,, \tag{3.68c}$$

where σ_{ij} is the fluid's "rate of shear", Ex. 3.18; and (ii) the *diffusion coefficient* D, which determines the diffusive flux of particles $\mathbf{S}$ from regions of high particle density n to low,

$$\boxed{\mathbf{S} = -D \boldsymbol{\nabla} n} \,. \tag{3.68d}$$

There is a *diffusion equation* associated with each of these transport coefficients. For example, the differential law of particle conservation $\partial n/\partial t + \boldsymbol{\nabla} \cdot \mathbf{S} = 0$ [Eq. (1.30)], when applied to material in which the particles scatter many times so $\mathbf{S} = -D\boldsymbol{\nabla} n$, gives the following diffusion equation for the particle number density:

$$\boxed{\frac{\partial n}{\partial t} = D\nabla^2 n} \,, \tag{3.69}$$

where we have assumed that D is spatially constant. In Ex. 3.16, by exploring solutions to this equation, we shall see that the rms (root means square) distance $\bar{l}$ the particles travel is proportional to the square root of their travel time, $\bar{l} = \sqrt{4Dt}$, a behavior characteristic of diffusive random walks.[7]

[7] Einstein derived this diffusion law $\bar{l} = \sqrt{4Dt}$, and he used it and his formula for the diffusion coefficient D, along with observational data about the diffusion of sugar molecules in water, to demonstrate the physical reality of molecules, determine their sizes, and deduce the numerical value of Avogadro's number; see the historical discussion in Chap. 5 of Pais (1982).

Similarly, the law of energy conservation, when applied to diffusive heat flow $\mathbf{F} = -\kappa \boldsymbol{\nabla} T$, leads to a diffusion equation for the thermal energy density U and thence for temperature [Ex. 3.17 and Eq. (18.4)]; Maxwell's equations in a magnetized fluid, when combined with Ohm's law $\mathbf{j} = \kappa_e \mathbf{E}$, lead to a diffusion equation (19.6) for magnetic field lines; and the law of angular momentum conservation, when applied to a shearing fluid with $T_{ij} = -2\eta_{\rm shear}\sigma_{ij}$, leads to a diffusion equation (14.6) for vorticity.

These diffusion equations, and all other physical laws involving transport coefficients, are approximations to the real world—approximations that are valid if and only if (*i*) many particles are involved in the transport of the quantity of interest (charge, heat, momentum, particles) and (*ii*) on average each particle undergoes many scatterings in moving over the length scale of the macroscopic inhomogeneities that drive the transport. This second requirement can be expressed quantitatively in terms of the *mean free path* λ between scatterings (i.e., the mean distance a particle travels between scatterings, as measured in the mean rest frame of the matter) and the *macroscopic inhomogeneity scale* $\mathcal{L}$ for the quantity that drives the transport (for example, in heat transport that scale is $\mathcal{L} \sim T/|\boldsymbol{\nabla} T|$, i.e., it is the scale on which the temperature changes by an amount of order itself). In terms of these quantities, the second criterion of validity is $\lambda \ll \mathcal{L}$. These two criteria (many particles and $\lambda \ll \mathcal{L}$) together are called *diffusion criteria*, since they guarantee that the quantity being transported (charge, heat, momentum, particles) will diffuse through the matter. If either of the two diffusion criteria fails, then the standard transport law (Ohm's law, the law of heat conduction, the Navier-Stokes equation, or the particle diffusion equation) breaks down and the corresponding transport coefficient becomes irrelevant and meaningless.

The accuracy with which one can compute a transport coefficient using the Boltzmann transport equation depends on the accuracy of one's description of the scattering. If one uses a high-accuracy collision term $(d\mathcal{N}/d\ell)_{\rm collisions}$ in the Boltzmann equation, one can derive a highly accurate transport coefficient. If one uses a very crude approximation for the collision term, one's resulting transport coefficient might be accurate only to within an order of magnitude—in which case, it was probably not worth the effort to use the Boltzmann equation; a simple order-of-magnitude argument would have done just as well. If the interaction between the diffusing particles and the scatterers is electrostatic or gravitational ("long-range" $1/r^2$ interaction forces), then the particles cannot be idealized as moving freely between collisions, and an accurate computation of transport coefficients requires a more sophisticated analysis: the Fokker-Planck equation developed in Sec. 6.9 and discussed, for plasmas, in Secs. 20.4.3 and 20.5.

In the following three subsections, we shall compute the coefficient of thermal conductivity κ for hot gas inside a star (where short-range collisions hold sway and the Boltzmann equation is highly accurate). We shall do so first by an order-of-magnitude argument, and then by the Boltzmann equation with an accurate collision term. In Exs. 3.18 and 3.19 readers will have the opportunity to compute the coefficient of viscosity and the diffusion coefficient for particles using moderately accurate collision terms, and in Ex. 3.20 (neutron diffusion in a nuclear reactor), we will meet diffusion in momentum space, by contrast with diffusion in physical space.

****

EXERCISES

Exercise 3.16 ***Example: Solution of Diffusion Equation in an Infinite, Homogeneous Medium*

(a) Show that the following is a solution to the diffusion equation (3.69) for particles in a homogeneous, infinite medium:

$$\boxed{n = \frac{N}{(4\pi Dt)^{3/2}} e^{-r^2/4Dt}}\,, \tag{3.70}$$

(where $r \equiv \sqrt{x^2+y^2+z^2}$ is radius), and that it satisfies $\int n\, d\mathcal{V}_x = N$, so N is the total number of particles. Note that this is a Gaussian distribution with width $\sigma = \sqrt{4Dt}$. Plot this solution for several values of σ. In the limit as $t \to 0$, the particles are all localized at the origin. As time passes, they random-walk (diffuse) away from the origin, traveling a mean distance $\sigma = \sqrt{4Dt}$ after time t. We will meet this square-root-of-time evolution in other random-walk situations elsewhere in this book.

(b) Suppose that the particles have an arbitrary initial distribution $n_o(\mathbf{x})$ at time $t = 0$. Show that their distribution at a later time t is given by the following "Greens function" integral:

$$n(\mathbf{x}, t) = \int \frac{n_o(\mathbf{x}')}{(4\pi Dt)^{3/2}} e^{-|\mathbf{x}-\mathbf{x}'|^2/4Dt} d\mathcal{V}_{x'}\,. \tag{3.71}$$

Exercise 3.17 ***Problem: Diffusion Equation for Temperature*

Use the law of energy conservation to show that, when heat diffuses through a homogeneous medium whose pressure is being kept fixed, the evolution of the temperature perturbation $\delta T \equiv T -$ (average temperature) is governed by the diffusion equation

$$\frac{\partial T}{\partial t} = \chi \nabla^2 T \quad \text{where} \quad \chi = \kappa/\rho c_P \tag{3.72}$$

is called the *thermal diffusivity*. Here c_P is the specific heat per unit mass at fixed pressure, and ρc_P is that specific heat per unit volume. For the extension of this to heat flow in a moving fluid, see Eq. (18.4).

****

3.7.1 Problem to be Analyzed: Diffusive Heat Conduction Inside a Star

The specific transport-coefficient problem we shall treat here in the text is for heat transport through hot gas deep inside a young, massive star. We shall confine attention to that portion of the star in which the temperature is $10^7\,\mathrm{K} \lesssim T \lesssim 10^9\mathrm{K}$, the mass density is

$\rho \lesssim 10\,\mathrm{g/cm}^3(T/10^7\mathrm{K})^2$, and *heat is carried primarily by diffusing photons* rather than by diffusing electrons or ions or by convection. (We shall study convection in Chap. 18.) In this regime the primary impediment to the photons' flow is collisions with electrons. The lower limit on temperature, 10^7K, guarantees that the gas is almost fully ionized, so there is a plethora of electrons to do the scattering. The upper limit on density, $\rho \sim 10\,\mathrm{g/cm}^3(T/10^7\mathrm{K})^2$ guarantees that (i) the inelastic scattering, absorption, and emission of photons by electrons accelerating in the coulomb fields of ions ("bremsstrahlung" processes) are unimportant as impediments to heat flow compared to *scattering off free electrons*; and (ii) the scattering electrons are nondegenerate, i.e., they have mean occupation numbers η small compared to unity and thus behave like classical, free, charged particles. The upper limit on temperature, $T \sim 10^9$K, guarantees that (i) the electrons which do the scattering are moving thermally at much less than the speed of light (the mean thermal energy $\frac{3}{2}k_BT$ of an electron is much less than its rest mass-energy m_ec^2); and (ii) the scattering is nearly elastic, with negligible energy exchange between photon and electron, and is describable with good accuracy by the *Thomson scattering cross section*:

In the rest frame of the electron, which to good accuracy will be the same as the mean rest frame of the gas since the electron's speed relative to the mean rest frame is $\ll c$, the differential cross section $d\sigma$ for a photon to scatter from its initial propagation direction $\mathbf{n}'$ into a unit solid angle $d\Omega$ centered on a new propagation direction $\mathbf{n}$ is

$$\frac{d\sigma(\mathbf{n}' \to \mathbf{n})}{d\Omega} = \frac{3}{16\pi}\sigma_{\rm T}[1 + (\mathbf{n}\cdot\mathbf{n}')^2] \ . \tag{3.73a}$$

Here $\sigma_{\rm T}$ is the total Thomson cross section [the integral of the differential cross section (3.73a) over solid angle]:

$$\sigma_{\rm T} = \int \frac{d\sigma(\mathbf{n}' \to \mathbf{n})}{d\Omega} d\Omega = \frac{8\pi}{3} r_o^2 = 0.665 \times 10^{-24}\,\mathrm{cm}^2 \ , \tag{3.73b}$$

where $r_o = e^2/m_ec^2$ is the classical electron radius. For a derivation and discussion of the Thomson cross sections (3.73) see, e.g., Sec. 14.8 of Jackson (1999).

3.7.2 Order-of-Magnitude Analysis

Before embarking on any complicated calculation, it is *always* helpful to do a rough, order-of-magnitude analysis, thereby identifying the key physics and the approximate answer. The first step of a rough analysis of our heat transport problem is to identify the magnitudes of the relevant lengthscales. The inhomogeneity scale $\mathcal{L}$ for the temperature, which drives the heat flow, is the size of the hot stellar core, a moderate fraction of the Sun's radius: $\mathcal{L} \sim 10^5$ km. The mean free path of a photon can be estimated by noting that, since each electron presents a cross section $\sigma_{\rm T}$ to the photon and there are n_e electrons per unit volume, the probability of a photon being scattered when it travels a distance l through the gas is of order $n_e\sigma_{\rm T}l$; and therefore to build up to unit probability for scattering, the photon must travel a distance

$$\lambda \sim \frac{1}{n_e\sigma_T} \sim \frac{m_p}{\rho\sigma_{\rm T}} \sim 3\,\mathrm{cm}\left(\frac{1\mathrm{g/cm}^3}{\rho}\right) \sim 3\,\mathrm{cm} \ . \tag{3.74}$$

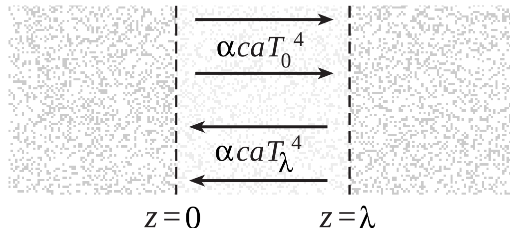

Fig. 3.10: Heat exchange between two layers of gas separated by a distance of one photon mean free path in the direction of the gas's temperature gradient.

Here m_p is the proton rest mass, $\rho \sim 1$ g/cm^3 is the mass density in the core of a young, massive star, and we have used the fact that stellar gas is mostly hydrogen to infer that there is approximately one nucleon per electron in the gas and hence that $n_e \sim \rho/m_p$. Note that $\mathcal{L} \sim 10^5$ km is 3×10^4 times larger than $\lambda \sim 3$ cm, and the number of electrons and photons inside a cube of side $\mathcal{L}$ is enormous, so the diffusion description of heat transport is quite accurate.

In the diffusion description, the heat flux $\mathbf{F}$ as measured in the gas's rest frame is related to the temperature gradient $\boldsymbol{\nabla} T$ by the law of diffusive heat conduction $\mathbf{F} = -\kappa \boldsymbol{\nabla} T$. To estimate the thermal conductivity κ, orient the coordinates so the temperature gradient is in the z direction, and consider the rate of heat exchange between a gas layer located near $z = 0$ and a layer one photon-mean-free-path away, at $z = \lambda$ (Fig. 3.10). The heat exchange is carried by photons that are emitted from one layer, propagate nearly unimpeded to the other, and then scatter. Although the individual scatterings are nearly elastic (and we thus are ignoring changes of photon frequency in the Boltzmann equation), tiny changes of photon energy add up over many scatterings to keep the photons nearly in local thermal equilibrium with the gas. Thus, we shall approximate the photons and gas in the layer at $z = 0$ to have a common temperature T_0 and those in the layer at $z = \lambda$ to have a common temperature $T_\lambda = T_0 + \lambda dT/dz$. Then the photons propagating from the layer at $z = 0$ to that at $z = \lambda$ carry an energy flux

$$F_{0\to\lambda} = \alpha c a (T_0)^4 \ , \tag{3.75a}$$

where a is the radiation constant of Eq. (3.52c), $a(T_0)^4$ is the photon energy density at $z = 0$, and α is a dimensionless constant of order $1/4$ that accounts for what fraction of the photons at $z = 0$ are moving rightward rather than leftward, and at what mean angle to the z direction. (Throughout this section, by contrast with early sections of this chapter, we shall use non-geometrized units, with the speed of light c present explicitly). Similarly, the flux of energy from the layer at $z = \lambda$ to the layer at $z = 0$ is

$$F_{\lambda\to 0} = -\alpha c a (T_\lambda)^4 \ ; \tag{3.75b}$$

and the net rightward flux, the sum of (3.75a) and (3.75b), is

$$F = \alpha c a [(T_0)^4 - (T_\lambda)^4] = -4\alpha c a T^3 \lambda \frac{dT}{dz}. \tag{3.75c}$$

Noting that 4α is approximately one, inserting expression (3.74) for the photon mean free path, and comparing with the law of diffusive heat flow $\mathbf{F} = -\kappa \boldsymbol{\nabla} T$, we conclude that the thermal conductivity is

$$\kappa \sim aT^3 c\lambda = \frac{acT^3}{\sigma_{\rm T} n_e} \; . \tag{3.76}$$

3.7.3 Analysis Via the Boltzmann Transport Equation

With these physical insights and rough answer in hand, we turn to a Boltzmann transport analysis of the heat transfer. Our first step is to formulate the Boltzmann transport equation for the photons (including effects of Thomson scattering off the electrons) in the rest frame of the gas. Our second step will be to solve that equation to determine the influence of the heat flow on the distribution function $\mathcal{N}$, and our third step will be to compute the thermal conductivity κ by an integral over that $\mathcal{N}$.

To simplify the analysis we use, as the parameter ℓ in the Boltzmann transport equation $d\mathcal{N}/d\ell = (d\mathcal{N}/d\ell)_{\rm collisions}$, the distance l that a fiducial photon travels, and we regard the distribution function $\mathcal{N}$ as a function of location $\mathbf{x}$ in space, the photon propagation direction (unit vector) $\mathbf{n}$, and the photon frequency ν: $\mathcal{N}(\mathbf{x}, \mathbf{n}, \nu)$. Because the photon frequency does not change during free propagation nor in the Thomson scattering, it can be treated as a constant when solving the Boltzmann equation.

Along the trajectory of a fiducial photon, $\mathcal{N}(\mathbf{x}, \mathbf{n}, \nu)$ will change as a result of two things: (*i*) the scattering of photons out of the $\mathbf{n}$ direction and into other directions, and (*ii*) the scattering of photons from other directions $\mathbf{n}'$ into the $\mathbf{n}$ direction. These effects produce the following two collision terms in the Boltzmann transport equation (3.64):

$$\frac{d\mathcal{N}(\mathbf{x}, \mathbf{n}, \nu)}{dl} = -\sigma_{\rm T} n_e \mathcal{N}(\mathbf{x}, \mathbf{n}, \nu) + \int \frac{d\sigma(\mathbf{n}' \to \mathbf{n})}{d\Omega} n_e \mathcal{N}(\mathbf{x}, \mathbf{n}', \nu) d\Omega' \; . \tag{3.77}$$

Because the mean free path $\lambda = 1/\sigma_T n_e \sim 3$cm is so short compared to the length scale $\mathcal{L} \sim 10^5$km of the temperature gradient, the heat flow will show up as a tiny correction to an otherwise isotropic, perfectly thermal distribution function. Thus, we can write the photon distribution function as the sum of an unperturbed, perfectly isotropic and thermalized piece $\mathcal{N}_0$ and a tiny, anisotropic perturbation $\mathcal{N}_1$:

$$\mathcal{N} = \mathcal{N}_0 + \mathcal{N}_1 \; , \quad \text{where} \;\; \mathcal{N}_0 = \frac{2}{h^3} \frac{1}{e^{h\nu/k_B T} - 1} \; . \tag{3.78a}$$

Here the perfectly thermal piece $\mathcal{N}_0(\mathbf{x}, \mathbf{n}, \nu)$ has the standard black-body form (3.23); it is independent of $\mathbf{n}$ and it depends on $\mathbf{x}$ only through the temperature $T = T(\mathbf{x})$. If the photon mean free path were vanishingly small, there would be no way for photons at different locations $\mathbf{x}$ to discover that the temperature is inhomogeneous; and, correspondingly, $\mathcal{N}_1$ would be vanishingly small. The finiteness of the mean free path permits $\mathcal{N}_1$ to be finite, and so it is reasonable to expect (and turns out to be true) that the magnitude of $\mathcal{N}_1$ is

$$\mathcal{N}_1 \sim \frac{\lambda}{\mathcal{L}} \mathcal{N}_0 \; . \tag{3.78b}$$

Thus, $\mathcal{N}_0$ is the leading-order term, and $\mathcal{N}_1$ is the first-order correction in an expansion of the distribution function $\mathcal{N}$ in powers of $\lambda/\mathcal{L}$. This is called a *two-lengthscale expansion*; see Box 3.3.

Box 3.3
Two-Lengthscale Expansions

Equation (3.78b) is indicative of the mathematical technique that underlies Boltzmann-transport computations: a perturbative expansion in the dimensionless ratio of two lengthscales, the tiny mean free path λ of the transporter particles and the far larger macroscopic scale $\mathcal{L}$ of the inhomogeneities that drive the transport. Expansions in lengthscale ratios $\lambda/\mathcal{L}$ are called *two-lengthscale expansions*, and are widely used in physics and engineering. Most readers will previously have met such an expansion in quantum mechanics: the WKB approximation, where λ is the lengthscale on which the wave function changes and $\mathcal{L}$ is the scale of changes in the potential $V(x)$ that drives the wave function. Kinetic theory itself is the result of a two-lengthscale expansion: It follows from the more sophisticated statistical-mechanics formalism in Chap. 4, in the limit where the particle sizes are small compared to their mean free paths. In this book we shall use two-lengthscale expansions frequently—e.g., in the geometric optics approximation to wave propagation (Chap. 7), in the study of boundary layers in fluid mechanics (Secs. 14.4 and 15.5), in the quasi-linear formalism for plasma physics (Chap. 23), and in the definition of a gravitational wave (Sec. 27.4).

Inserting $\mathcal{N} = \mathcal{N}_0 + \mathcal{N}_1$ into our Boltzmann transport equation (3.77) and using $d/dl = \mathbf{n}\cdot\boldsymbol{\nabla}$ for the derivative with respect to distance along the fiducial photon trajectory, we obtain

$$\begin{aligned} n_j \frac{\partial \mathcal{N}_0}{\partial x_j} + n_j \frac{\partial \mathcal{N}_1}{\partial x_j} &= \left[-\sigma_{\rm T} n_e \mathcal{N}_0 + \int \frac{d\sigma(\mathbf{n}' \to \mathbf{n})}{d\Omega} n_e c \mathcal{N}_0 d\Omega' \right] \\ &+ \left[-\sigma_{\rm T} n_e c \mathcal{N}_1(\mathbf{n},\nu) + \int \frac{d\sigma(\mathbf{n}' \to \mathbf{n})}{d\Omega} n_e c \mathcal{N}_1(\mathbf{n}',\nu) d\Omega' \right] \ . \end{aligned} \tag{3.78c}$$

Because $\mathcal{N}_0$ is isotropic, i.e., is independent of photon direction $\mathbf{n}'$, it can be pulled out of the integral over $\mathbf{n}'$ in the first square bracket on the right side; and when this is done, the first and second terms in that square bracket cancel each other. Thus, the unperturbed part of the distribution, $\mathcal{N}_0$, completely drops out of the right side of Eq. (3.78c). On the left side the term involving the perturbation $\mathcal{N}_1$ is tiny compared to that involving the unperturbed distribution $\mathcal{N}_0$, so we shall drop it; and because the spatial dependence of $\mathcal{N}_0$ is entirely due to the temperature gradient, we can bring the first term and the whole transport equation into the form

$$n_j \frac{\partial T}{\partial x_j} \frac{\partial \mathcal{N}_0}{\partial T} = -\sigma_{\rm T} n_e \mathcal{N}_1(\mathbf{n},\nu) + \int \frac{d\sigma(\mathbf{n}' \to \mathbf{n})}{d\Omega} n_e \mathcal{N}_1(\mathbf{n}',\nu) d\Omega' \ . \tag{3.78d}$$

The left side of this equation is the amount by which the temperature gradient causes $\mathcal{N}_0$ to fail to satisfy the Boltzmann equation, and the right side is the manner in which the perturbation $\mathcal{N}_1$ steps into the breach and enables the Boltzmann equation to be satisfied.

Because the left side is linear in the photon propagation direction n_j (i.e., it has a $\cos\theta$ dependence in coordinates where $\boldsymbol{\nabla} T$ is in the z-direction; i.e., it has a "dipolar", $l = 1$

angular dependence), $\mathcal{N}_1$ must also be linear in n_j, i.e. dipolar, in order to fulfill Eq. (3.78d). Thus, we shall write $\mathcal{N}_1$ in the dipolar form

$$\mathcal{N}_1 = K_j(\mathbf{x}, \nu) n_j \ , \tag{3.78e}$$

and we shall solve the transport equation (3.78d) for the function K_j.

[*Important side remark*: This is a special case of a general situation: When solving the Boltzmann transport equation in diffusion situations, one is performing a power series expansion in $\lambda/\mathcal{L}$; see Box 3.3. The lowest-order term in the expansion, $\mathcal{N}_0$, is isotropic, i.e., it is monopolar in its dependence on the direction of motion of the diffusing particles. The first-order correction, $\mathcal{N}_1$, is down in magnitude by $\lambda/\mathcal{L}$ from $\mathcal{N}_0$ and is dipolar (or sometimes quadrupolar; see Ex. 3.18) in its dependence on the particles' direction of motion. The second-order correction, $\mathcal{N}_2$, is down in magnitude by $(\lambda/\mathcal{L})^2$ from $\mathcal{N}_0$ and its multipolar order is one higher than $\mathcal{N}_1$ (quadrupolar here; octupolar in Ex. 3.18). And so it continues on up to higher and higher orders.[8]]

When we insert the dipolar expression (3.78e) into the angular integral on the right side of the transport equation (3.78d) and notice that the differential scattering cross section (3.73a) is unchanged under $\mathbf{n}' \to -\mathbf{n}'$, but $K_j n'_j$ changes sign, we find that the integral vanishes. As a result the transport equation (3.78d) takes the simplified form

$$n_j \frac{\partial T}{\partial x_j} \frac{\partial \mathcal{N}_0}{\partial T} = -\sigma_{\rm T} n_e K_j n_j \ , \tag{3.78f}$$

from which we can read off the function K_j and thence $\mathcal{N}_1 = K_j n_j$:

$$\mathcal{N}_1 = -\frac{\partial \mathcal{N}_0/\partial T}{\sigma_T n_e} \frac{\partial T}{\partial x_j} n_j \ . \tag{3.78g}$$

Notice that, as claimed above, the perturbation has a magnitude

$$\frac{\mathcal{N}_1}{\mathcal{N}_0} \sim \frac{1}{\sigma_T n_e} \frac{1}{T} |\boldsymbol{\nabla} T| \sim \frac{\lambda}{\mathcal{L}}. \tag{3.78h}$$

Having solved the Boltzmann transport equation to obtain $\mathcal{N}_1$, we can now evaluate the energy flux F_i carried by the diffusing photons. Relativity physicists will recognize that flux as the T^{0i} part of the stress-energy tensor and will therefore evaluate it as

$$F_i = T^{0i} = c^2 \int \mathcal{N} p^0 p^i \frac{d\mathcal{V}_p}{p^0} = c^2 \int \mathcal{N} p_i d\mathcal{V}_p \tag{3.79}$$

[cf. Eq. (3.31c) with the factors of c restored]. Newtonian physicists can deduce this formula by noticing that photons with momentum $\mathbf{p}$ in $d\mathcal{V}_p$ carry energy $E = |\mathbf{p}|c$ and move with velocity $\mathbf{v} = c\mathbf{p}/|\mathbf{p}|$, so their energy flux is $\mathcal{N} E \mathbf{v}\, d\mathcal{V}_p = c^2 \mathcal{N} \mathbf{p}\, d\mathcal{V}_p$; integrating this over momentum space gives Eq. (3.79). Inserting $\mathcal{N} = \mathcal{N}_0 + \mathcal{N}_1$ into this equation and noting that the integral over $\mathcal{N}_0$ vanishes, and inserting Eq. (3.78g) for $\mathcal{N}_1$, we obtain

$$F_i = c^2 \int \mathcal{N}_1 p_i d\mathcal{V}_p = -\frac{c}{\sigma_{\rm T} n_e} \frac{\partial T}{\partial x_j} \frac{\partial}{\partial T} \int \mathcal{N}_0 c n_j p_i d\mathcal{V}_p \ . \tag{3.80a}$$

[8] For full details in nonrelativistic situations see, e.g., Grad (1957); and for full relativistic details see, e.g., Thorne (1981).

The relativity physicist will identify the integral as Eq. (3.31c) for the photons' stress tensor T_{ij} (since $n_j = p_j/p_0 = p_j/\mathcal{E}$). The Newtonian physicist, with a little thought, will recognize the integral in Eq. (3.80a) as the j-component of the flux of i-component of momentum, which is precisely the stress tensor. Since this stress tensor is being computed with the isotropic, thermalized part of $\mathcal{N}$, it is isotropic, $T_{ji} = P\delta_{ji}$, and its pressure has the standard black-body-radiation form $P = \frac{1}{3}aT^4$ [Eqs. (3.52a)]. Replacing the integral in Eq. (3.80a) by this black-body stress tensor, we obtain our final answer for the photons' energy flux:

$$F_i = -\frac{c}{\sigma_{\rm T} n_e}\frac{\partial T}{\partial x_j}\frac{d}{dT}\left(\frac{1}{3}aT^4\delta_{ji}\right) = -\frac{c}{\sigma_{\rm T} n_e}\frac{4}{3}aT^3\frac{\partial T}{\partial x_i} \ . \tag{3.80b}$$

Thus, *from the Boltzmann transport equation we have simultaneously derived the law of diffusive heat conduction* $\mathbf{q} = -\kappa\boldsymbol{\nabla}T$ *and evaluated the coefficient of heat conductivity*

$$\kappa = \frac{4}{3}\frac{acT^3}{\sigma_{\rm T} n_e} \ . \tag{3.81}$$

Notice that this heat conductivity is 4/3 times our crude, order-of-magnitude estimate (3.76).

The above calculation, while somewhat complicated in its details, is conceptually fairly simple. The reader is encouraged to go back through the calculation and identify the main conceptual steps (expansion of distribution function in powers of $\lambda/\mathcal{L}$, insertion of zero-order plus first-order parts into the Boltzmann equation, multipolar decomposition of the zero and first-order parts with zero-order being monopolar and first-order being dipolar, neglect of terms in the Boltzmann equation that are smaller than the leading ones by factors $\lambda/\mathcal{L}$, solution for the coefficient of the multipolar decomposition of the first-order part, reconstruction of the first-order part from that coefficient, and insertion into a momentum-space integral to get the flux of the quantity being transported). Precisely these same steps are used to evaluate all other transport coefficients that are governed by classical physics. For examples of other such calculations see, e.g., Shkarofsky, Johnston, and Bachynski (1966).

As an application of the thermal conductivity (3.81), consider a young (main-sequence) 7 solar mass ($7M_\odot$) star as modeled, e.g., on page 480 of Clayton (1968). Just outside the star's convective core, at radius $r \simeq 0.8R_\odot \simeq 6\times10^5$km (where $R_\odot$ is the Sun's radius), the density and temperature are $\rho \simeq 5{\rm g/cm}^3$ and $T \simeq 1.6\times10^7$K, so the number density of electrons is $n_e \simeq \rho/m_p \simeq 3\times10^{24}{\rm cm}^{-3}$. For these parameters, Eq. (3.81) gives a thermal conductivity $\kappa \simeq 7\times10^{17}{\rm erg\ s^{-1}cm^{-2}K^{-1}}$. The lengthscale $\mathcal{L}$ on which the temperature is changing is approximately the same as the radius, so the temperature gradient is $|\boldsymbol{\nabla}T| \sim T/r \sim 3\times10^{-4}$K/cm. The law of diffusive heat transfer then predicts a heat flux $F = \kappa|\boldsymbol{\nabla}T| \sim 2\times10^{14}{\rm erg\ s^{-1}cm^{-2}}$, and thus a total luminosity $L = 4\pi r^2 q \sim 8\times10^{36}$erg/s $\simeq 2000L_\odot$ (2000 solar luminosities). What a difference the mass makes! The heavier a star, the hotter its core, the faster it burns, and the higher its luminosity. Increasing the mass by a factor 7 drives the luminosity up by 2000.

****

EXERCISES

Exercise 3.18 [N] ***Example: Viscosity of a Monatomic Gas*
Consider a nonrelativistic fluid that, in the neighborhood of the origin, has fluid velocity

$$v_i = \sigma_{ij} x_j \ , \tag{3.82}$$

with σ_{ij} symmetric and trace free. As we shall see in Sec. 13.7.1, this represents a purely shearing flow, with no rotation or volume changes of fluid elements; σ_{ij} is called the fluid's *rate of shear*. Just as a gradient of temperature produces a diffusive flow of heat, so the gradient of velocity embodied in σ_{ij} produces a diffusive flow of momentum, i.e. a stress. In this exercise we shall use kinetic theory to show that, for a monatomic gas with isotropic scattering of atoms off each other, this stress is

$$\boxed{T_{ij} = -2\eta_{\rm shear}\sigma_{ij}} \ , \tag{3.83a}$$

with the coefficient of shear viscosity

$$\boxed{\eta_{\rm shear} \simeq \frac{1}{3}\rho\lambda v_{th}} \ , \tag{3.83b}$$

where ρ is the gas density, λ is the atoms' mean free path between collisions, and $v_{\rm th} = \sqrt{3kT_B/m}$ is the atoms' rms speed. Our analysis will follow the same route as the analysis of heat conduction in Secs. 3.7.2 and 3.7.3.

(a) Derive Eq. (3.83b) for the shear viscosity, to within a factor of order unity, by an order of magnitude analysis like that in Sec. 3.7.2.

(b) Regard the atoms' distribution function $\mathcal{N}$ as being a function of the magnitude p and direction $\mathbf{n}$ of an atom's momentum, and of location $\mathbf{x}$ in space. Show that, if the scattering is isotropic with cross section σ_s and the number density of atoms is n, then the Boltzmann transport equation can be written as

$$\frac{d\mathcal{N}}{dl} = \mathbf{n}\cdot\boldsymbol{\nabla}\mathcal{N} = -\frac{1}{\lambda}\mathcal{N} + \int \frac{1}{4\pi\lambda}\mathcal{N}(p,\mathbf{n}',\mathbf{x})d\Omega' \ , \tag{3.84a}$$

where $\lambda = 1/n\sigma_s$ is the atomic mean free path (mean distance traveled between scatterings) and l is distance traveled by a fiducial atom.

(b) Explain why, in the limit of vanishingly small mean free path, the distribution function has the following form:

$$\mathcal{N}_0 = \frac{n}{(2\pi m k_B T)^{3/2}} \exp[-(\mathbf{p} - m\boldsymbol{\sigma}\cdot\mathbf{x})^2/2mk_BT] \ . \tag{3.84b}$$

(c) Solve the Boltzmann transport equation (3.84a) to obtain the leading-order correction $\mathcal{N}_1$ to the distribution function at $\mathbf{x} = 0$. [Answer: $\mathcal{N}_1 = -(\lambda\, p/k_BT)\sigma_{ab}n_a n_b \mathcal{N}_0$.]

(d) Compute the stress via a momentum-space integral. Your answer should be Eq. (3.83a) with $\eta_{\rm shear}$ given by Eq. (3.83b) to within a few tens of per cent accuracy. [Hint: Along the way you will need the following angular integral:

$$\int n_a n_b n_i n_j d\Omega = \frac{4\pi}{15}(\delta_{ab}\delta_{ij} + \delta_{ai}\delta_{bj} + \delta_{aj}\delta_{bi}) \ . \tag{3.84c}$$

Derive this by arguing that the integral must have the above delta-function structure, and by then computing the multiplicative constant by performing the integral for $a = b = i = j = z$.]

Exercise 3.19 [N] *Example: Diffusion Coefficient in the "Collision-Time" Approximation*
Consider a collection of identical "test particles" with rest mass $m \neq 0$ that diffuse through a collection of thermalized "scattering centers". (The test particles might be molecules of one species, and the scattering centers might be molecules of a much more numerous species.) The scattering centers have a temperature T such that $k_B T \ll mc^2$, so if the test particles acquire this same temperature they will have thermal speeds small compared to the speed of light, as measured in the mean rest frame of the scattering centers. We shall study the effects of scattering on the test particles using the following "collision-time" approximation for the collision terms in the Boltzmann equation, which we write in the mean rest frame of the scattering centers:

$$\left(\frac{d\mathcal{N}}{dt}\right)_{\rm collision} = (\mathcal{N}_0 - \mathcal{N})\frac{1}{\hat\tau} \ , \quad \text{where } \mathcal{N}_0 \ \equiv \ \frac{e^{-p^2/2mk_BT}}{(2\pi m k_B T)^{3/2}} n \ . \tag{3.85}$$

Here the time derivative d/dt is taken moving with a fiducial test particle along its unscattered trajectory, $p = |\mathbf{p}|$ is the magnitude of the test particles' spatial momentum, $n = \int \mathcal{N} d\mathcal{V}_p$ is the number density of test particles, and $\hat\tau$ is a constant to be discussed below.

(a) Show that this collision term preserves test particles in the sense that

$$\left(\frac{dn}{dt}\right)_{\rm collision} \equiv \int \left(\frac{d\mathcal{N}}{dt}\right)_{\rm collision} dp_x dp_y dp_z = 0 \ . \tag{3.86}$$

(b) Explain why this collision term corresponds to the following physical picture: Each test particle has a probability per unit time $1/\hat\tau$ of scattering; and when it scatters, its direction of motion is randomized and its energy is thermalized at the scattering centers' temperature.

(c) Suppose that the temperature T is homogeneous (spatially constant), but the test particles are distributed inhomogeneously, $n = n(\mathbf{x}) \neq$const. Let $\mathcal{L}$ be the length scale on which their number density n varies. What condition must $\mathcal{L}$, $\hat\tau$, T, and m satisfy in order that the diffusion approximation be reasonably accurate? Assume that this condition is satisfied.

(d) Compute, in order of magnitude, the particle flux $\mathbf{S} = -D\boldsymbol{\nabla} n$ produced by the gradient of the number density n, and thereby evaluate the diffusion coefficient D.

(e) Show that the Boltzmann transport equation takes the form

$$\frac{\partial \mathcal{N}}{\partial t} + \frac{p_j}{m}\frac{\partial \mathcal{N}}{\partial x_j} = \frac{1}{\hat{\tau}}(\mathcal{N}_0 - \mathcal{N}) \; . \tag{3.87a}$$

(f) Show that to first order in a small diffusion-approximation parameter, the solution of this equation is $\mathcal{N} = \mathcal{N}_0 + \mathcal{N}_1$, where $\mathcal{N}_0$ is as defined in Eq. (3.85) above, and

$$\mathcal{N}_1 = -\frac{p_j \hat{\tau}}{m}\frac{\partial n}{\partial x_j}\frac{e^{-p^2/2mk_BT}}{(2\pi m k_B T)^{3/2}} \; . \tag{3.87b}$$

Note that $\mathcal{N}_0$ is monopolar (independent of the direction of $\mathbf{p}$), while $\mathcal{N}_1$ is dipolar (linear in $\mathbf{p}$).

(g) Show that the perturbation $\mathcal{N}_1$ gives rise to a particle flux given by Eq. (3.68d), with the diffusion coefficient

$$D = \frac{k_B T}{m}\hat{\tau} \; . \tag{3.88}$$

How does this compare with your order of magnitude estimate in part (d)?

Exercise 3.20 [N] *****Example: Neutron Diffusion in a Nuclear Reactor***

Here are some salient, oversimplified facts about nuclear reactors (see, e.g., Stephenson 1954, especially Chap. 4): A reactor's core is made of a mixture of natural uranium (0.72 percent ^{235}U and 99.28 percent ^{238}U) and a solid or liquid material such as carbon (graphite) or water, made of low-atomic-number atoms and called the *moderator*. For concreteness we shall assume that the moderator is graphite. Slow (thermalized) neutrons, with kinetic energies ~ 0.1 eV, get captured by the ^{235}U nuclei and trigger them to fission, releasing ~ 170 MeV of kinetic energy per fission (which ultimately goes into heat and then electric power), and also releasing an average of about 2 fast neutrons (kinetic energies ~ 1 MeV). The fast neutrons must be slowed to thermal speeds so as to capture onto ^{235}U atoms and induce further fissions. The slowing is achieved by scattering off the moderator atoms—a scattering in which the crucial effect, energy loss, occurs in *momentum space*. The momentum-space scattering is elastic and isotropic in the center-of-mass frame, with total cross section (to scatter off one of the moderator's carbon atoms) $\sigma_s \simeq 4.8 \times 10^{-24}\text{cm}^2 \equiv 4.8$ barns. Using the fact that in the moderator's rest frame, the incoming neutron has a much higher kinetic energy than the moderator carbon atoms, and using energy and momentum conservation and the isotropy of the scattering, one can show that in the moderator's rest frame, the logarithm of the neutron's energy is reduced in each scattering by an average amount ξ that is independent of energy and is given by:

$$\xi \equiv -\overline{\Delta \ln E} = 1 + \frac{(A-1)^2}{2A}\ln\left(\frac{A-1}{A+1}\right) \simeq 0.158 \; . \tag{3.89}$$

Here $A \simeq 12$ is the ratio of the mass of the scattering atom to the mass of the scattered neutron.

There is a dangerous hurdle that the diffusing neutrons must overcome during their slowdown: as the neutrons pass through a critical energy region of about 7 to 6 eV, the ^{238}U atoms can absorb them. The absorption cross section has a huge resonance there, with width about 1 eV and resonance integral $\int \sigma_a d\ln E \simeq 240$ barns. For simplicity, we shall approximate the cross section in this absorption resonance by $\sigma_a \simeq 1600$ barns at $6\text{eV} < E < 7\text{eV}$, and zero outside this range. To achieve a viable fission chain reaction and keep the reactor hot, it is necessary that about half of the neutrons (one per original ^{235}U fission) slow down through this resonant energy without getting absorbed. Those that make it through will thermalize and trigger new ^{235}U fissions (about one per original fission), maintaining the chain reaction.

We shall idealize the uranium and moderator atoms as homogeneously mixed on length-scales small compared to the neutron mean free path, $\lambda_s = 1/(\sigma_s n_s) \simeq 2$ cm, where n_s is the number density of moderator (carbon) atoms. Then the neutrons' distribution function $\mathcal{N}$, as they slow down, will be isotropic in direction and independent of position; and in our steady state situation, it will be independent of time. It therefore will depend only on the magnitude p of the neutron momentum, or equivalently on the neutron kinetic energy $E = p^2/2m$: $\mathcal{N} = \mathcal{N}(E)$.

Use the Boltzmann transport equation or other considerations to develop the theory of the slowing down of the neutrons in momentum space, and of their struggle to pass through the ^{238}U resonance region without getting absorbed. More specifically:

(a) Use as the distribution function not $\mathcal{N}(E)$ but rather $n_E(E) \equiv dN/d\mathcal{V}_x dE =$ (number of neutrons per unit volume and per unit kinetic energy), and denote by $q(E)$ the number of neutrons per unit volume that slow down through energy E per unit time. Show that outside the resonant absorption region these two quantities are related by

$$q = \sigma_s n_s \xi E \, n_E v \text{ , where } \quad v = \sqrt{2mE} \tag{3.90}$$

is the neutron speed, so q contains the same information as the distribution function n_E. Explain why the steady-state operation of the nuclear reactor requires q to be independent of energy in this non-absorption region, and infer that $n_E \propto E^{-3/2}$.

(b) Show further that inside the resonant absorption region, $6\text{eV} < E < 7\text{eV}$, the relationship between q and E is modified:

$$q = (\sigma_s n_s + \sigma_a n_a)\xi E \, n_E v \ . \tag{3.91}$$

Here n_s is the number density of scattering (carbon) atoms and n_a is the number density of absorbing (^{238}U) atoms. [Hint: require that the rate at which neutrons scatter into a tiny interval of energy $\delta E \ll \xi E$ is equal to the rate at which they leave that tiny interval.] Then show that the absorption causes q to vary with energy according to the following differential equation:

$$\frac{d\ln q}{d\ln E} = \frac{\sigma_a n_a}{(\sigma_s n_s + \sigma_a n_a)\xi} \ . \tag{3.92}$$

(c) By solving this differential equation in our idealization of constant σ_a over the range 7 to 6 eV, show that the condition to maintain the chain reaction is

$$\frac{n_s}{n_a} \simeq \frac{\sigma_a}{\sigma_s}\left(\frac{\ln(7/6)}{\xi \ln 2} - 1\right) \simeq 0.41\frac{\sigma_a}{\sigma_s} \simeq 140\ . \tag{3.93}$$

Thus, to maintain the reaction in the presence of the huge ^{238}U absorption resonance for neutrons, it is necessary that more than 99 per cent of the reactor volume be taken up by moderator atoms and less than 1 per cent by uranium atoms.

We reiterate that this is a rather idealized version of what happens inside a nuclear reactor, but it provides insight into some of the important processes and the magnitudes of various relevant quantities. For a graphic example of an additional complexity, see the description of "Xenon poisoning" of the chain reaction in the first production-scale nuclear reactor (built during World War II to make plutonium for the first American atomic bombs), in John Archibald Wheeler's autobiography (Wheeler 1998).

****

Bibliographic Note

Newtonian kinetic theory is treated in many textbooks on statistical physics. At an elementary level, Chap. 14 of Kittel and Kroemer (1980) is rather good. At a more advanced level, see, e.g., Chap. 3 of Kardar (2007), Chap. 11 of Reichl (2009), and Secs. 7.9–7.13 and Chaps. 12, 13, and 14 of Reiff (1965). For a very advanced treatment with extensive applications to electrons and ions in plasmas, and electrons, phonons and quasi-particles in liquids and solids, see Lifshitz and Pitaevskii (1981).

Relativistic kinetic theory is rarely touched on in statistical-physics textbooks but should be. It is well known to astrophysicists. The treatment in this chapter is easily lifted into general relativity theory; see, e.g., Sec. 22.6 of Misner, Thorne and Wheeler (1973) and for a bit of the flavor, Box 28.4 of this book.

Bibliography

Allen, C. W. 2000. *Allen's Astrophysical Quantitites*, fourth edition, Ed. A. N. Cox, Berlin: Springer-Verlag.

Bennett, C. L. et. al. "First-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Preliminary Maps and Basic Results," *Astrophysical Supplement Series*, **148**, 1, 2003.

Cohen-Tannoudji, Claude, Diu, Bernard, and Laloe, Franck. 1977. *Quantum Mechanics*, Volume 2. New York: Wiley.

Box 3.4
Important Concepts in Chapter 3

- **Foundational Concepts**
  - Momentum space and phase space, Secs. 3.2.1 and 3.2.2.
  - Distribution function, three important variants: $\mathcal{N} = dN/d\mathcal{V}_x d\mathcal{V}_p$ (Secs. 3.2.1 and 3.2.2), mean occupation number η (Sec. 3.2.5), and for photons the spectral intensity I_ν (Sec. 3.2.4).
  - **T2** Frame invariance of $\mathcal{N}$, η and I_ν/ν^3 in relativity theory, Sec. 3.2.2.
  - Liouville's theorem and collisionless Boltzmann equation (constancy of distribution function along a particle trajectory in phase space): Sec. 3.6.
  - Boltzmann transport equation with collisions: Eq. (3.64).
  - Thermal equilibrium: Fermi-Dirac, Bose-Einstein and Boltzmann distribution functions: Sec. 3.3.
  - Electron degeneracy: Sec. 3.5.2.
  - Macroscopic properties of matter expressed as momentum-space integrals—density, pressure, stress tensor, stress-energy tensor: Sec. 3.4.
  - Two-lengthscale expansion: Box 3.3.
- **Equations of State Computed Via Kinetic Theory:**
  - Computed using momentum-space integrals: Eq. (3.36) and preceding discussion.
  - Important cases: nonrelativistic, classical gas, Sec. 3.5.1; radiation, Sec. 3.5.5, **T2** degenerate Fermi gas, Eq. (3.41) and Sec. 3.5.4;
  - Density-temperature plane for hydrogen: Sec. 3.5.2 and Fig. 3.7.
- **Transport Coefficients**
  - Defined: beginning of Sec. 3.7.
  - Electrical conductivity, thermal conductivity, shear viscosity and diffusion coefficient [Eqs. (3.68)]
  - Order-of-magnitude computations of: Sec. 3.7.2
  - Computations using Boltzmann transport equation: Sec. 3.7.3.

Clayton, Donald D. 1968. *Principles of Stellar Evolution and Nucleosynthesis* New York:McGraw-Hill

Grad, Harold. 1957. *Principles of the Kinetic Theory of Gases in Handbuch der Physik*, Berlin: Springer Verlag, 205–294.

Goldstein, Herbert, Poole, Charles, and Safko, John. 2002. *Classical Mechanics*, Third Edition, New York: Addison-Wesley.

Jackson, John, David. 1999. *Classical Electrodynamics*, Third Edition, New York: Wiley.

Kardar, M. 2007. *Statistical Physics of Particles*, Cambridge: Cambridge University Press.

Kittel, C., & Kroemer, H. 1980. *Thermal Physics*, San Francisco: Freeman.

Lifshitz, E.M. & Pitaevskii, L.P. 1981. *Physical Kinetics*, Volume 10 of Landau & Lifshitz *Course of Theoretical Physics*.

Lynden-Bell, D. 1967. *Monthly Notices of the Royal Astronomical Society*, **136**, 101.

Misner, Charles W., Thorne, Kip S., & Wheeler, John A. 1973. *Gravitation*, San Francisco: Freeman.

Reichl, L. E. 2009. *A Modern Course in Statistical Physics*, third revised edition, London: Arnold.

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

Shapiro, S.L. and Teukolsky, S.A. 1983. *Black Holes, White Dwarfs and Neutron Stars*, New York: Wiley.

Shkarofsky, I. P., Johnston, T. W., & Bachynski, M. P. 1966. *The Particle Kinetics of Plasmas*, New York: Addison-Wesley.

Stephenson, Richard. 1954. *Introduction to Nuclear Engineering*, New York: McGraw-Hill.

Thorne, Kip S. 1981. “Relativistic radiative transfer: Moment formalisms”, *Monthly Notices of the Royal Astronomical Society*, 194, 439–473.

Thorne, Kip S. 1994. *Black Holes and Time Warps: Einstein's Outrageous Legacy*, New York: W.W. Norton.

Wheeler, John Archibald. 1998. *Geons, Black Holes and Quantum Foam: A Life in Physics*, New York: W.W. Norton, pp. 54–59.