

Part I
STATISTICAL PHYSICS

Statistical Physics

Version 0402.2.K, 12 October 2004

In this first part of the book we shall study aspects of classical statistical physics that every physicist should know, but are not usually treated in elementary thermodynamics courses. Our study will lay the microphysical (particle-scale) foundations for the continuum physics of Parts II—VI. As a central feature of our approach, we shall emphasize the intimate connections between the relativistic formulation of statistical physics and its nonrelativistic limit, and between quantum statistical physics and the classical theory. Throughout, we shall presume that the reader is familiar with elementary thermodynamics, but not with other aspects of statistical physics.

In Chap. 2 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 per unit 3-dimensional volume of ordinary space and per unit 3-dimensional volume of momentum space. Despite first appearances, \mathcal{N} turns out to be a geometric, frame-independent entity. This \mathcal{N} and the frame-independent laws it obeys provide us with a means for computing, from microphysics, macroscopic quantities that continuum physics: mass density, thermal energy density, pressure, equations of state, thermal and electrical conductivities, viscosities, diffusion coefficients,

In Chap. 3 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 have the same degrees of freedom, but different in that their degrees of freedom may be in different 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} coordinates (the x , y , z of all the molecules) and 3×10^{23} momenta (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 and momenta, 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 “thermodynamic equilibrium”, and the evolution of ensembles into

thermodynamic equilibrium. Among our applications will be a derivation of the Fermi-Dirac distribution for Fermions in thermodynamic equilibrium and the Bose-Einstein distribution for Bosons, a development of the theory of Bose-Einstein condensation in a dilute gas, an exploration of the meaning and role of entropy in gases, in black holes and in the universe as a whole.

In Chap. 4 we will use the tools of statistical mechanics to study *statistical thermodynamics*, i.e. ensembles of systems that are in or near thermodynamic 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 of these tools that we shall study are: (i) a 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. 5 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 Chapter 5, 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 is that it is intimately connected to statistical physics and that similar principles to those originally 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 techniques for extracting weak signals from noisy data by filtering of the data, and the limits that noise places on the accuracies of physics experiments and on the reliability of communications channels.

Chapter 2

Kinetic Theory

Version 0402.2.K, 12 October 2004

[Changes from 0402.1.K: (i) Some rewording of parts (a) and (b) of Ex. 2.14 on neutron diffusion in a nuclear reactor. (ii) Minor typos fixed.

2.1 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. The particles might be, for example, galaxies; and our goal might be to study how galaxies, born randomly in the early universe, congregate into clusters as the universe expands. Or, the particles might be stars that reside in our own galaxy, and our goal might be to study how spiral structure develops in the distribution of our galaxy’s stars. Or, the particles might be molecules in interstellar space, and our goal might be to study how a supernova explosion affects the evolution of the interstellar gas’s density and temperature distribution. Or, the particles might be photons in the cosmic microwave radiation, and our goal might be to study how anisotropies in the expansion of the universe affect the photons’ temperature distribution. Or, the particles might be electrons in a metal, and our goal might be to study how changes of the metal’s temperature affect its thermal and electrical conductivity. Or, the particles might be neutrons in a nuclear reactor, and our goal might be to study whether they can survive long enough to maintain a nuclear chain reaction and keep the reactor hot.

For all such problems, a powerful conceptual tool is the “distribution function” or “number density of particles in phase space”. In order to develop this concept in a manner valid both for photons, which move with the speed of light, and for galaxies, which move slowly compared to light, we shall use the tools of special relativity. We begin in Sec. 2.2 by introducing the concepts of momentum space, phase space, and the distribution function. In Sec. 2.3 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 mechanical particles with integral spin (bosons), and a third for classical particles. As special applications we derive

the Maxwell-Boltzmann velocity distribution for low-speed, classical particles (Ex. 2.3), and we compute the effects of an observer’s motion on his measurement of the cosmic microwave radiation created in the big-bang origin of the universe (Ex. 2.4). In Sec. 2.4 we meet the number-flux four-vector and the stress-energy tensor associated with a collection of particles, and we learn how to compute these quantities by integrating over the momentum portion of phase space. We show that, if the momentum distribution is isotropic in some reference frame, then on macroscopic scales the particles constitute a perfect fluid, and we use our momentum-space integrals to evaluate the equations of state of various kinds of perfect fluids: a photon gas (Ex. 2.5), a nonrelativistic, classical gas (Ex. 2.6), and an electron-degenerate hydrogen gas (Ex. 2.7), and we use our results to discuss the physical nature of matter as a function of density and temperature. In Sec. 2.5 we study the evolution of the distribution function, as described by Liouville’s theorem and 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. 2.10). Finally, in Sec. 2.6 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 also use the Boltzmann transport equation to study a chain reaction in a nuclear reactor (Ex. 2.14).

2.2 Phase Space and Distribution Function

2.2.1 Number density in phase space

Consider 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 \vec{p} by

$$\vec{p} = d\vec{x}/d\zeta . \quad (2.1a)$$

[Eq. (1.28)]. If the particle has non-zero rest mass, then (as we saw in Chap. 1) the particle’s 4-velocity \vec{u} and the proper time τ along its world line are related to its 4-momentum and affine parameter by

$$\vec{p} = m\vec{u} , \quad \zeta = \tau/m . \quad (2.1b)$$

The particle can be thought of not only as living in four-dimensional spacetime [Fig. 2.1(a)], but also as living in a four-dimensional momentum space [Fig. 2.1(b)]. 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 head sits at the point representing \vec{p} . The momentum-space diagram drawn in Fig. 2.1(b) has as its coordinate axes the components (p^0, p^1, p^2, p^3) of the 4-momentum as measured in some arbitrary Lorentz frame. Because the squared length of the 4-momentum is always $-m^2$,

$$\vec{p} \cdot \vec{p} = -(p^0)^2 + (p_x)^2 + (p_y)^2 + (p_z)^2 = -m^2 , \quad (2.1c)$$

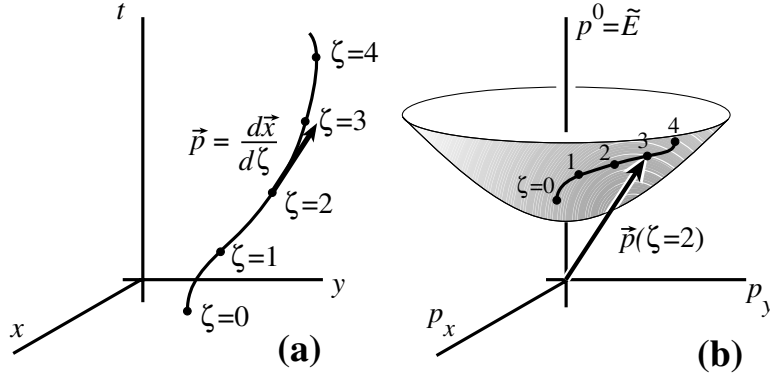


Fig. 2.1: (a) The world line $\vec{x}(\zeta)$ of a particle in spacetime (with one spatial coordinate, z , suppressed), parametrized by a parameter ζ that is related to the particle's 4-momentum by $\vec{p} = d\vec{x}/d\zeta$. (b) The trajectory of the particle in momentum space. The particle's momentum is confined to the mass hyperboloid, $\vec{p}^2 = -m^2$.

the particle's 4-momentum is confined to a hyperboloid in momentum space. This hyperboloid is described mathematically by Eq. (2.1c) and is called the *mass hyperboloid*. We shall often denote the particle's energy p^0 by

$$\mathcal{E} \equiv p^0 \quad (2.1d)$$

(with the \mathcal{E} in script font to distinguish it from the nonrelativistic energy of a particle, $E = \frac{1}{2}mv^2$), and we shall embody its 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 (2.1c) as

$$\mathcal{E}^2 = m^2 + |\mathbf{p}|^2 \quad (2.1e)$$

If no forces act on the particle, then its momentum is conserved and its location in momentum space remains fixed. However, forces (e.g., due to an electromagnetic field) can push the particle's 4-momentum along some curve in momentum space that lies on the mass hyperboloid. If we parametrize 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. 2.1(b). 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 inertial frame.

Momentum space and spacetime, taken together, are called *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.

Turn attention, now, from an individual particle to a collection of a huge number of particles. Assume, for simplicity, that all the particles have the same rest mass m (they are identical particles); and allow m to be finite or zero, it does not matter. Examine those particles that pass close to a specific event \mathcal{P} in spacetime; and examine them from the

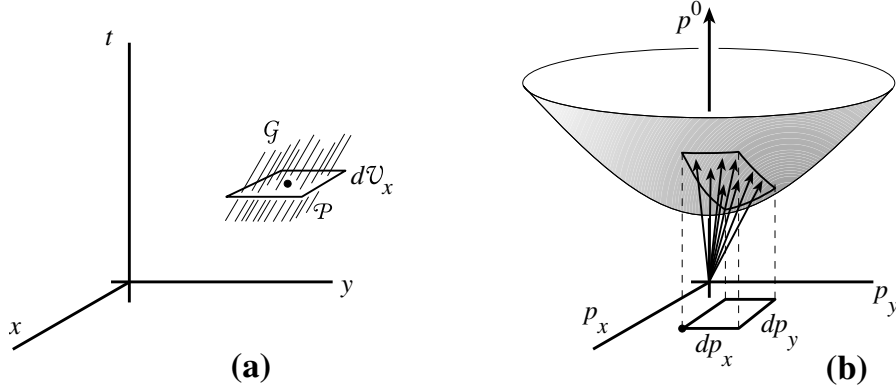


Fig. 2.2: Definition of the distribution function from the viewpoint of a specific observer in a specific reference frame: At the event \mathcal{P} the observer selects a 3-volume $d\mathcal{V}_x$, and she focuses on the set \mathcal{G} of particles that lie in $d\mathcal{V}_x$ and have momenta lying in a region of the mass hyperboloid which is centered on \vec{p} and has 3-momentum volume $d\mathcal{V}_p$. If dN is the number of particles in that set \mathcal{G} , then $\mathcal{N}(\mathcal{P}, \vec{p}) \equiv dN/d\mathcal{V}_x d\mathcal{V}_p$.

viewpoint of a specific observer, who lives in a specific inertial reference frame. Fig. 2.1(a) 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 select a specific, small, 3-dimensional box with edge lengths dx , dy , dz , with the event \mathcal{P} at its center, and with imaginary (i.e., not physical) walls. She measures this box to have an ordinary, three-dimensional volume

$$d\mathcal{V}_x = dx dy dz . \quad (2.2a)$$

Focus attention on a specific box-like region of the mass hyperboloid, centered on a specific 4-momentum \vec{p} in momentum space and with spatial momenta in the range $(p_x \pm \frac{1}{2} dp_x, p_y \pm \frac{1}{2} dp_y, p_z \pm \frac{1}{2} dp_z)$; Fig. 2.2(b). Our chosen observer attributes to that region a 3-dimensional *momentum volume*

$$d\mathcal{V}_p = dp_x dp_y dp_z . \quad (2.2b)$$

Ask the observer to focus on that collection \mathcal{G} of particles which lie in the spatial 3-volume $d\mathcal{V}_x$ [Fig. 2.2(a)] and have spatial momenta in the 3-volume $d\mathcal{V}_p$ [Fig. 2.2(b)]. If there are dN particles in this collection \mathcal{G} , then the observer will identify

$$\mathcal{N}(\mathcal{P}, \vec{p}) \equiv \frac{dN}{d\mathcal{V}_x d\mathcal{V}_p} \quad (2.3)$$

as the *number density of particles in phase space*. This number density 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 $d\mathcal{V}_p$ is centered. Regarded as a function of these quantities, $\mathcal{N}(\mathcal{P}, \vec{p})$ is called the particles' *distribution function*. It is the fundamental concept of kinetic theory.

At first sight one might expect \mathcal{N} to depend also on the inertial reference frame used in its definition, i.e., on the velocity of the observer. If this were the case, i.e., if \mathcal{N} at fixed \mathcal{P} and \mathbf{p} were different when computed by the above prescription using different inertial frames,

then we would feel compelled to seek some other frame-independent geometric object 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 (2.3) is frame-independent by itself, i.e. it is a frame-independent scalar field in phase space; so we need seek no further for a geometric foundation for kinetic theory. 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 $d\mathcal{V}_x d\mathcal{V}_p$ and of the distribution function $\mathcal{N} = dN/d\mathcal{V}_x d\mathcal{V}_p$. In studying this frame dependence the thing that identifies the 3-volume $d\mathcal{V}_x$ and 3-momentum $d\mathcal{V}_p$ is the set of particles \mathcal{G} . 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 later by taking the limit $m \rightarrow 0$. Then there is a preferred frame from which to observe the particles: their own rest frame. We shall denote by $x^{\alpha'}$ the Lorentz coordinates of that rest frame, and shall retain our previous notation x^α for the Lorentz coordinates of our previous frame. We shall orient the coordinate axes of the two frames so they are related by a pure boost in the x direction with speed v and with $\gamma = 1/\sqrt{1-v^2}$; see Fig. 2.3.

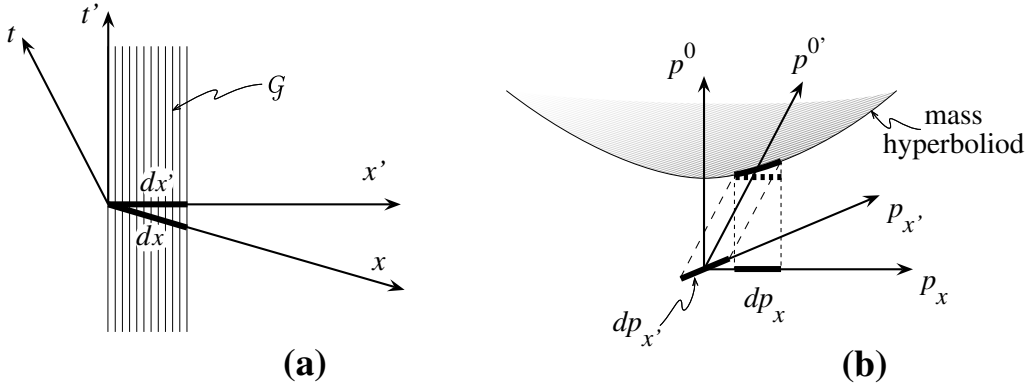


Fig. 2.3: (a) Spacetime diagram drawn from viewpoint of the (primed) rest frame of the particles \mathcal{G} . The length of the box occupied by the particles, dx , as measured in the moving (unprimed) frame, is Lorentz contracted, so $dx = \gamma^{-1}dx'$ and $d\mathcal{V}_x = \gamma^{-1}d\mathcal{V}'_x$. (b) Momentum space diagram drawn from viewpoint of the unprimed observer. The ranges of x -components of momenta occupied by the particles \mathcal{G} are related by $dp_{x'} = \gamma^{-1}dp_x$, so that $d\mathcal{V}'_p = \gamma^{-1}d\mathcal{V}_p$.

Figure 2.3(a) is a spacetime diagram that shows the world lines of a few of the particles in the set \mathcal{G} . These world lines go straight upward because the diagram's chosen reference frame is the particles' (primed) rest frame. The leftmost and rightmost world lines demark the walls of the particles' box. The box's length dx' as measured in the particles' rest frame is the square root of the spacetime interval along the thick horizontal line in the figure. Its length dx as measured in the unprimed reference frame is the square root of the spacetime interval along the thick slanted line. The slanted line dx is the hypotenuse of a right triangle in spacetime with spatial leg dx' and temporal leg vdx' . By the spacetime

“Pythagorean theorem” (equation for the interval) applied to this triangle, we have $dx^2 = dx'^2 - (vdx')^2 = (1 - v^2)dx'^2 = \gamma^{-2}dx'^2$, so $dx = \gamma^{-1}dx'$, which is the standard formula for the *Lorentz contraction of a moving box along its direction of motion*. Since lengths are preserved along the perpendicular directions, $dy = dy'$, $dz = dz'$, we conclude that $d\mathcal{V}_x \equiv dx dy dz = \gamma^{-1}dx' dy' dz' = \gamma^{-1}d\mathcal{V}_{x'}$. Since the dN particles all have nearly the same energy, $\mathcal{E} = m\gamma$, as seen in the unprimed frame, this Lorentz-contraction relation can be rewritten as $\mathcal{E}d\mathcal{V}_x = md\mathcal{V}_{x'}$, which says that

$$\mathcal{E}d\mathcal{V}_x = (\text{a frame-independent quantity}) . \quad (2.4a)$$

Thus, the spatial volume occupied by the particles is not frame-independent, but the product of their volume and their energy is.

Fig. 2.3(b) shows the mass hyperboloid, the region on that hyperboloid occupied by the momenta of the particles \mathcal{G} , and the projections dp_x and $dp_{x'}$ of that box on the two frames' axes. We have chosen to draw this diagram in the unprimed frame, rather than the particles' primed frame, because we thereby obtain a convenient right triangle to use in our computation. The hypotenuse of the right triangle is $dp_{x'}$ (thick slanted curve), the horizontal leg is dp_x (thick dashed curve), and the vertical leg is vdp_x ; so the spacetime “Pythagorean theorem” says $(dp_{x'})^2 = (dp_x)^2 - (vdp_x)^2 = \gamma^{-2}(dp_x)^2$. Therefore, $dp_{x'} = \gamma^{-1}dp_x$, which is a reversal from the relation $dx = \gamma^{-1}dx'$. Correspondingly, there is a reversal from Eq. (2.4a):

$$\frac{d\mathcal{V}_p}{\mathcal{E}} = (\text{a frame-independent quantity}) . \quad (2.4b)$$

This result can also be understood by examining the little box on the mass hyperboloid in Fig. 2.2(b). It is a small element of 3-volume $d\vec{\Sigma}_p$ in 4-momentum space [cf. Eq. (1.121) and associated discussion]. Regarded as a 4-vector in momentum space, this 3-volume points perpendicular to the little box in the diagram, i.e. it points in the same direction as the 4-momentum \vec{p} of a particle at the center of the 3-volume; i.e., $d\vec{\Sigma}_p$ and \vec{p} are parallel to each other. The time component of \vec{p} is the particle energy \mathcal{E} , and the time component of the 3-volume $d\vec{\Sigma}_p$ is $d\Sigma_p^0 = d\mathcal{V}_p$ (the projection of the box into the horizontal plane). Because $d\vec{\Sigma}_p$ and \vec{p} are parallel, when we transform from one reference frame to another, their time components must grow or shrink in unison, i.e. the ratio $d\Sigma_p^0/p^0 = d\mathcal{V}_p/\mathcal{E}$ must not change; it is frame invariant.

By taking the product of Eqs. (2.4a) and (2.4b) we see that for our chosen set of particles \mathcal{G} ,

$$d\mathcal{V}_x d\mathcal{V}_p = (\text{a frame-independent quantity}) ; \quad (2.4c)$$

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

$$\mathcal{N} = \frac{dN}{d\mathcal{V}_x d\mathcal{V}_p} \equiv \frac{dN}{d\mathcal{V}^2} = (\text{a frame-independent quantity}) . \quad (2.5)$$

Here we have introduced the notation $d\mathcal{V}^2 \equiv d\mathcal{V}_x d\mathcal{V}_p$ for the frame-independent six-dimensional phase-space volume occupied by the particles, a notation that suppresses any reference to the individually frame-dependent $d\mathcal{V}_x$ and $d\mathcal{V}_p$.

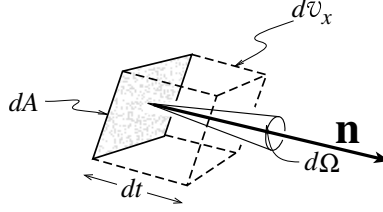


Fig. 2.4: Geometric construction used in defining the “specific intensity” I_ν .

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 $m \rightarrow 0$ and the 4-momenta become null. Correspondingly, all of Eqs. (2.4a) – (2.5) are valid for particles with zero rest mass as well as nonzero.

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, 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:

2.2.2 The distribution function $f(t, \mathbf{x}, \mathbf{v})$ for particles in a plasma.

In Part V, when dealing with nonrelativistic plasmas (collections of electrons and ions that have speeds small compared to light), we shall regard the distribution function as a function of time t , location \mathbf{x} in Euclidean space, and velocity \mathbf{v} (instead of momentum), and we shall use the notation and normalization

$$f(t, \mathbf{x}, \mathbf{v}) = \frac{dN}{dx dy dz dv_x dv_y dv_z} = m^3 \mathcal{N}. \quad (2.6)$$

However, we shall not use this normalization in the present chapter.

2.2.3 The distribution function I_ν/ν^3 for photons.

When dealing with photons or other zero-rest-mass particles, one often reexpresses \mathcal{N} in terms of the *specific intensity*, I_ν . This quantity is defined as follows (cf. Fig. 2.4): An observer places a CCD (or other measuring device) perpendicular to the spatial direction \mathbf{n} of propagation of the photons—perpendicular as measured in her Lorentz 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 $c = 1$, the product of that surface area with 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

$$d\mathcal{V}_x = dA dt. \quad (2.7a)$$

The photons all have nearly the same frequency ν as measured by the observer, and their 4-momenta as measured by her have components related to that frequency and to their

propagation direction \mathbf{n} by

$$\mathcal{E} = p^0 = h\nu, \quad \mathbf{p} = h\nu\mathbf{n}, \quad (2.7b)$$

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

$$d\mathcal{V}_p = |\mathbf{p}|^2 d\Omega d|\mathbf{p}| = h^3 \nu^2 d\Omega d\nu. \quad (2.7c)$$

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

$$d\mathcal{E} = h\nu dN, \quad (2.7d)$$

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"):

$$I_\nu \equiv \frac{d\mathcal{E}}{dA dt d\nu d\Omega}. \quad (2.8)$$

(This I_ν is sometimes denoted $I_{\nu\Omega}$.) From Eqs. (2.5), (2.7) and (2.8) we readily deduce the following relationship between this specific intensity and the distribution function:

$$\mathcal{N} = \frac{c^2 I_\nu}{h^4 \nu^3}. \quad (2.9)$$

Here the factor c^2 has been inserted so that I_ν is expressed in ordinary (cgs or mks) units in accord with astronomers' conventions. This relation shows that, with an appropriate renormalization, I_ν/ν^3 is the photons' distribution function.

Astronomers and opticians regard the specific 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, physicists regard the distribution function \mathcal{N} as a function of the photon 4-momentum \vec{p} (on the mass hyperboloid) and of location \mathcal{P} in spacetime. Clearly, the information contained in these two sets of variables, the astronomers' set and the physicists' set, is the same.

If two different astronomers in two different Lorentz 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 specific 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 specific intensity that she measures to the cube of the frequency she measures, that ratio, according to Eq. (2.9), will be the same as computed by the other astronomer; i.e., the distribution function I_ν/ν^3 will be frame-independent.

2.2.4 The mean occupation number η

Although this book is about classical physics, we cannot avoid making occasional 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 relativistic 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 Lorentz frame) and to have a spatial momentum in the region $d\mathcal{V}_p$ centered on \vec{p} . Suppose, further, that the particle does not interact with any other particles or fields. How many single-particle quantum mechanical states¹ are available to the free particle? This question is answered most easily by constructing, in the particle’s Lorentz 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 Lorentz 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

$$\psi_{j,k,l}(x, y, z) = \frac{1}{L^{3/2}} e^{i(2\pi/L)(jx+ky+lz)} e^{-i\omega t} , \quad (2.10a)$$

where

$$\omega = \sqrt{(m/\hbar)^2 + (2\pi/L)^2(j^2 + k^2 + l^2)} . \quad (2.10b)$$

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. Note that ω is a relativistic frequency, i.e. it includes the influence of the particle’s rest mass, i.e. it is related to the particle’s energy by $\mathcal{E} = \hbar\omega$; and correspondingly, Eq. (2.10b) is the relation $\mathcal{E} = \sqrt{m^2 + |\mathbf{p}|^2}$ [Eq.(2.1e)] in disguise. The basis states (2.10a) are eigenfunctions of the momentum operator $(\hbar/i)\nabla$ with momentum eigenvalues

$$p_x = \frac{2\pi\hbar}{L} , \quad p_y = \frac{2\pi\hbar}{L} , \quad p_z = \frac{2\pi\hbar}{L} . \quad (2.10c)$$

Thus, 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:

$$dN_{\text{states}} = \frac{d\mathcal{V}_p}{(2\pi\hbar/L)^3} = \frac{L^3 d\mathcal{V}_p}{(2\pi\hbar)^3} = \frac{d\mathcal{V}_x d\mathcal{V}_p}{h^3} . \quad (2.10d)$$

¹A 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”.

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

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

$$g_s = \begin{cases} 2s + 1 & \text{if } m \neq 0; \text{ e.g., an electron or proton or atomic nucleus,} \\ 2 & \text{if } m = 0 \text{ \& } s > 0; \text{ e.g., a photon } (s = 1) \text{ or graviton } (s = 2), \\ 1 & \text{if } m = 0 \text{ \& } s = 0; \text{ i.e., a hypothetical massless scalar particle.} \end{cases} \quad (2.11)$$

We shall call this number of internal spin states the particle's *multiplicity*.

Taking account both of 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 $d\mathcal{V}_x d\mathcal{V}_p \equiv d\mathcal{V}^2$ of phase space is $dN_{\text{states}} = (g_s/h^3)d\mathcal{V}^2$, and correspondingly the *number density of states in phase space* is

$$\mathcal{N}_{\text{states}} \equiv \frac{dN_{\text{states}}}{d\mathcal{V}^2} = \frac{g_s}{h^3}. \quad (2.12)$$

Note that, although we derived this number density of states using a specific Lorentz reference 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 .

In quantum statistical mechanics one often focuses attention on the number of particles that reside in a given quantum state. That number is called the state's *occupation number*. Select a specific event \mathcal{P} in spacetime and a specific 4-momentum \vec{p} for particles of a specific type. Then the ratio of the number density of particles at (\mathcal{P}, \vec{p}) to the number density of states is the *mean occupation number* η of the states near (\mathcal{P}, \vec{p}) :

$$\eta(\mathcal{P}, \vec{p}) \equiv \frac{\mathcal{N}(\mathcal{P}, \vec{p})}{\mathcal{N}_{\text{states}}} = \frac{h^3}{g_s} \mathcal{N}. \quad (2.13)$$

Because η is our most fundamental version of the distribution function, we shall generally use this relation in the rearranged form

$$\mathcal{N} \equiv \frac{dN}{d\mathcal{V}^2} = \frac{g_s}{h^3} \eta. \quad (2.14)$$

From quantum field 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 $1/2, 3/2, 5/2, \dots$) or that of a *boson* (a particle with spin $0, 1, 2, \dots$). 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 $0, 1, 2, 3, \dots$. Correspondingly, the mean occupation numbers must lie in the ranges

$$0 \leq \eta \leq 1 \text{ for fermions, } \quad 0 \leq \eta \leq \infty \text{ for bosons.} \quad (2.15)$$

Quantum field 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 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].

EXERCISES

Exercise 2.1 *Example: Distribution function for Particles with a Range of Rest Masses*
Consider a collection of particles (e.g. stars) that are not all identical, but instead have different rest masses.

- (a) For a subset \mathcal{G} of particles like that of Fig. 2.2 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{G} 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{G} is dN , then we define the frame-invariant distribution function by

$$\mathcal{N} \equiv \frac{dN}{d^2\mathcal{V}} = \frac{dN}{d\mathcal{V}_x d\mathcal{V}_p dm} . \quad (2.16)$$

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 2.1 below.

Exercise 2.2 *Example: Regimes of Particulate and Wave-like Behavior*

- (a) A gamma-ray burster is an astrophysical object (probably a fireball of hot gas exploding outward from the vicinity of a newborn black hole or colliding neutron stars or black holes) at a cosmological distance from earth ($\sim 10^{10}$ light years). The fireball emits gamma rays with individual photon energies as measured at earth $\mathcal{E} \sim 100$ keV. These photons arrive at Earth in a burst whose total energy per unit area is roughly 10^{-6} ergs/cm² and that lasts about one second. Assume the diameter of the emitting surface as seen from earth is ~ 1000 km and there is no absorption along the route to earth. Make a rough estimate of the mean occupation number of the burst's 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 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.

2.3 Thermal Equilibrium

In Chap. 3 we will introduce with care, and explore in detail, the concept of “statistical equilibrium”—also called “thermal equilibrium”. For now, we rely on the reader's prior experience with this concept. 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. 3, there is a special Lorentz frame (the *mean rest frame* of the particles near \mathcal{P}) in which the mean occupation number takes on the following form:

$$\eta = \frac{1}{e^{(\mathcal{E}-\mu_R)/kT} + 1} \quad \text{for fermions ,} \quad (2.17a)$$

$$\eta = \frac{1}{e^{(\mathcal{E}-\mu_R)/kT} - 1} \quad \text{for bosons .} \quad (2.17b)$$

Here $\mathcal{E} \equiv p^0$ is the total energy of an individual particle as measured in that mean rest frame. It is expressible in terms of the particle 4-momentum \vec{p} and the rest-frame 4-velocity \vec{u}_{rf} by

$$\mathcal{E} \equiv -\vec{p} \cdot \vec{u}_{\text{rf}} \quad (2.18a)$$

[Eq. (1.69)]; in terms of the spatial momentum \mathbf{p} as measured in the mean rest frame, it is

$$\mathcal{E} = (m^2 + \mathbf{p}^2)^{\frac{1}{2}} . \quad (2.18b)$$

The quantity k in Eqs. (2.17) is Boltzmann's constant, $k = 1.381 \times 10^{-16}$ erg K⁻¹ = 1.381×10^{-23} J K⁻¹; T is the *temperature* of the collection of particles and μ_R is the particles' relativistic *chemical potential* (the chemical potential including rest mass). When we remove the rest mass we get the nonrelativistic chemical potential

$$\mu \equiv \mu_R - m . \quad (2.19)$$

The chemical potential μ_R , the temperature T , and the rest-frame 4-velocity \vec{u}_{rf} can depend on location in spacetime \mathcal{P} . Thus, the dependence of the equilibrium η on \mathcal{P} is through $\mu_R(\mathcal{P})$, $T(\mathcal{P})$, and $\vec{u}_{\text{rf}}(\mathcal{P})$; while its dependence on the particle 4-momentum is only through $\mathcal{E} = -\vec{p} \cdot \vec{u}_{\text{rf}}$.

Notice that the equilibrium mean occupation number (2.17a) for fermions lies in the range 0 to 1 as required, while that (2.17b) for bosons lies in the range 0 to ∞ . The equilibrium fermion distribution (2.17a) is called the *Fermi-Dirac* distribution, while the equilibrium boson distribution (2.17b) is called the *Bose-Einstein* distribution. In the regime $(m - \mu_R)/kT \gg 1$ (i.e. $\mu \ll -kT$), the mean occupation number is small compared to unity for all particle energies $\mathcal{E} \geq m$, and is the same for fermions and bosons:

$$\eta = e^{-(\mathcal{E}-\mu_R)/kT} \quad \text{when } \mu \equiv \mu_R - m \ll -kT . \quad (2.20)$$

This is the regime of *classical*, distinguishable-particle behavior, and this limiting distribution is called the *Boltzmann distribution*.

The temperature T should already be familiar to the reader as a measure of the average energy per particle in thermal equilibrium. The chemical potential μ_R might not be so familiar. We shall learn its nature by exploring the consequences of the thermal distributions (2.17a), (2.17b), (2.20). As we shall see, μ_R is a measure of how many particles are present. It can be positive or negative (though for bosons it cannot exceed the rest mass m ; otherwise η would be negative at low energies, $\mathcal{E} \simeq m$);² and the more positive μ_R is, the larger is the number of particles.

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 thermodynamic equilibrium in which the chemical potential vanishes, $\mu_R = 0$ (as we shall see in Chap. 3). For example, photons inside a box with perfectly emitting and absorbing walls that have temperature T acquire the mean occupation number (2.17b) with zero chemical potential; expressed in terms of photon frequency ν as measured in the box's rest frame, the corresponding distribution function in its various normalizations takes on the standard *black-body (Planck)* form

$$\eta = \frac{1}{e^{h\nu/kT} - 1} , \quad \mathcal{N} = \frac{2}{h^3} \frac{1}{e^{h\nu/kT} - 1} , \quad I_\nu = \frac{(2h/c^2)\nu^3}{e^{h\nu/kT} - 1} . \quad (2.21)$$

(Here, 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 (2.17b) with temperature T equal to that of the walls and with nonzero chemical potential μ_R fixed by the number of photons present; the more photons there are, the larger will be the chemical potential.

When the particles have finite rest mass m and kT is small compared to m , then almost all of them will have speeds small compared to light. In this nonrelativistic regime, we shall

²For bosons with μ_r extremely close to m , i.e. $\mu \simeq 0$, a *boson condensate* can occur; we shall study boson condensates in Sec. 3.4.4.

replace the particle energy $\mathcal{E} = m/\sqrt{1-v^2} = \sqrt{m^2 + |\mathbf{p}|^2}$ by its nonrelativistic limit

$$E = \mathcal{E} - m = \frac{1}{2}mv^2 = \frac{|\mathbf{p}|^2}{2m}, \quad (2.22)$$

from which the rest mass has been removed, and we shall use the nonrelativistic chemical potential $\mu = \mu_R - m$. The thermal distributions then take on nonrelativistic forms that are identical to their fully relativistic forms (2.17a), (2.17b), (2.20), but with μ_R replaced by μ and \mathcal{E} replaced E .³

$$\eta = \frac{1}{e^{(E-\mu)/kT} + 1} \quad \text{for fermions (Fermi-Dirac distribution),} \quad (2.23a)$$

$$\eta = \frac{1}{e^{(E-\mu)/kT} - 1} \quad \text{for bosons (Bose-Einstein distribution),} \quad (2.23b)$$

$$\eta = e^{-(E-\mu)/kT} \quad \text{for classical regime } \mu \ll -kT \text{ (Boltzmann distribution).} \quad (2.23c)$$

EXERCISES

Exercise 2.3 Example: Velocity Distribution for Thermalized, Classical Particles

Consider a collection of thermalized, classical particles with zero spin and nonzero rest mass, so they have the relativistic Boltzmann distribution (2.20). Denote by t, x, y, z the Lorentz coordinates of their mean rest frame.

- (a) Show that the chemical potential determines the total number density of particles n in physical space, as measured in the mean rest frame; more specifically, show that $n \equiv dN/dxdydz \propto e^{\mu_R/kT}$.
- (b) Show that the distribution of particle energies as measured in the mean rest frame is

$$\frac{dN}{dxdydzd\mathcal{E}} = \left(\frac{4\pi}{h^3} e^{\mu_R/kT} \right) \mathcal{E} \sqrt{\mathcal{E}^2 - m^2} e^{-\mathcal{E}/kT}. \quad (2.24)$$

- (c) What is the distribution of particle speeds $v = |\mathbf{v}|$, $dN/dxdydzdv$? Show that in the limit $v \ll 1$ your answer reduces to the *Maxwell velocity distribution*

$$\frac{dN}{dxdydzdv} = \left(\frac{4\pi m^3}{h^3} e^{\mu/kT} \right) v^2 \exp\left(\frac{-\frac{1}{2}mv^2}{kT} \right). \quad (2.25)$$

Draw a graph of this distribution.

³Lynden-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.

Exercise 2.4 *Example: 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:

$$\eta = \frac{1}{e^{h\nu/kT_o} - 1}, \quad \text{with } T_o = 2.73 \text{ K} . \quad (2.26)$$

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

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

$$I_\nu = \frac{(2h/c^2)\nu^3}{e^{h\nu/kT_o} - 1} . \quad (2.27)$$

Plot this specific 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

$$\eta = \frac{1}{e^{-\vec{p} \cdot \vec{u}_o/kT_o} - 1}, \quad (2.28)$$

where \vec{p} is the photon 4-momentum and \vec{u}_o is the 4-velocity of the mean rest frame.

- (c) In actuality the earth moves relative to the mean rest frame of the microwave background with a speed v of about 600 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 specific intensity of the radiation received is precisely Planckian in form [Eq. (2.21)], but with a direction-dependent *Doppler-shifted temperature*

$$T = T_o \left(\frac{\sqrt{1 - v^2}}{1 - v \cos \theta} \right) . \quad (2.29)$$

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. What is the magnitude $\Delta T/T$ of the variations between maximum and minimum microwave temperature on the sky? It was by measuring these variations that astronomers⁴ discovered the motion of the earth relative to the mean rest frame of the cosmic microwave radiation.

⁴Corey and Wilkinson (1976), and Smoot, Gorenstein, and Muller (1977).

2.4 Macroscopic Properties of Matter as Integrals over Momentum Space

2.4.1 Number-Flux Vector and Stress-Energy Tensor

The constraint that the laws of physics be frame-independent relationships between frame-independent objects is a powerful one. We shall see an example of that power in this section.

Our objective is to study macroscopic properties of a distribution of particles. Guided by the Principle of Relativity (Chap. 1), we shall describe those properties as geometric objects (scalar, vector, and tensor fields) that live in spacetime, and we shall try to express those geometric objects as frame-independent integrals over momentum space. The integrals must satisfy a number of properties:

- i.* Their integrands must involve the distribution function $\mathcal{N} = (g_s/h^3)\eta$, since \mathcal{N} is the only frame-independent object that we have to describe the state of the particles.
- ii.* The most interesting of the integrals (and the only ones we shall study) will be linear in \mathcal{N} , so that if we double the number of particles present, we double the amount of the quantity being computed.
- iii.* Besides \mathcal{N} , the only other frame-independent quantity we know that lives in momentum space is the 4-momentum \vec{p} ; and, thus, the integrand can involve \mathcal{N} , \vec{p} , and no other quantities.
- iv.* The integration element must be frame-independent, so it must be $d\mathcal{V}_p/\mathcal{E} = dp_x dp_y dp_z/\mathcal{E}$ when expressed in terms of the momentum components in any Lorentz frame.

The only integrals satisfying these four properties (aside from normalization) are

$$R \equiv \int \mathcal{N} \frac{d\mathcal{V}_p}{\mathcal{E}}, \quad (2.30a)$$

$$S^\mu \equiv \int \mathcal{N} p^\mu \frac{d\mathcal{V}_p}{\mathcal{E}}, \quad (2.30b)$$

$$T^{\mu\nu} \equiv \int \mathcal{N} p^\mu p^\nu \frac{d\mathcal{V}_p}{\mathcal{E}}, \quad (2.30c)$$

and higher-order (cubic, quartic, ...) “moments” of \mathcal{N} . Here and throughout this chapter, momentum space integrals unless otherwise specified are taken over the entire mass hyperboloid. We shall explore each of the geometric, frame-independent entities (2.30) by computing its components in a Lorentz frame.

In any Lorentz frame the scalar field R of Eq. (2.30a) is

$$R = \int \mathcal{N} \frac{1}{\mathcal{E}} dp_x dp_y dp_z. \quad (2.31)$$

This is the sum, over all particles in a unit 3-volume, of the inverse energy. Although it is intriguing that this quantity is frame-independent, 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. (2.30b) plays a very important role in physics. Its time component in a Lorentz frame is

$$S^0 = \int \mathcal{N} dp_x dp_y dp_z = \int \frac{dN}{dx dy dz dp_x dp_y dp_z} dp_x dp_y dp_z . \quad (2.32a)$$

Obviously, this is the number of particles per unit three-dimensional volume n , i.e., the *particle number density*, as measured in the Lorentz frame:

$$S^0 = n . \quad (2.32b)$$

The x -component of \vec{S} , similarly, is

$$\begin{aligned} S_x &= \int \mathcal{N} p_x \frac{dp_x dp_y dp_z}{\mathcal{E}} = \int \frac{dN}{dx dy dz dp_x dp_y dp_z} \frac{dx}{dt} dp_x dp_y dp_z \\ &= \int \frac{dN}{dt dy dz dp_x dp_y dp_z} dp_x dp_y dp_z . \end{aligned} \quad (2.32c)$$

Here the second equality follows from

$$\frac{p_x}{\mathcal{E}} = \frac{p_x}{p^0} = \frac{dx/d\zeta}{dt/d\zeta} = \frac{dx}{dt} = (\textit{x-component of velocity}) . \quad (2.32d)$$

It should be obvious from Eq. (2.32c) that S_x is the flux of particles in the x -direction, i.e., the number of particles that cross a unit area orthogonal to the x -direction per unit time, as measured in the Lorentz frame. Similarly, S_y and S_z are the fluxes of particles in the y and z directions; and \mathbf{S} (denoted S_j in index notation) is the *particle flux vector*—a 3-dimensional vector residing in the three-dimensional space of the chosen Lorentz frame. The full, 4-dimensional quantity \vec{S} (S^μ in index notation) is the *number-flux 4-vector* introduced in Sec. 1.11.2.

If the particles are conserved, then \vec{S} satisfies the conservation law $S^\alpha{}_{;\alpha} = 0$; i.e., $\vec{\nabla} \cdot \vec{S} = 0$; or in a Lorentz frame $\partial n / \partial t + \nabla \cdot \mathbf{S} = 0$ [Eq. (1.130)]. Notice that $\partial n / \partial t$ has the interpretation of the increase in the number of particles in a unit volume per unit time, i.e. the increase in the number of particles per unit 4-dimensional volume of spacetime; and the term $\nabla \cdot \mathbf{S}$ accounts for this increase by the flow of particles across the spatial volume's walls. Correspondingly, if particles are actually being created or destroyed, the frame-invariant evolution law for \vec{S} must take the form

$$\vec{\nabla} \cdot \vec{S} = (\textit{number of particles created minus number annihilated per unit spacetime volume}). \quad (2.33)$$

Turn to the quantity $T^{\mu\nu}$ defined by the integral (2.30c). By evaluating its components in a chosen Lorentz frame, we can easily verify that this is the particles' *stress-energy tensor*, which we studied in Sec. 1.12. Specifically, the component

$$T^{\mu 0} = \int \mathcal{N} p^\mu p^0 \frac{dp_x dp_y dp_z}{p^0} = \int \frac{dN}{dx dy dz dp_x dp_y dp_z} p^\mu dp_x dp_y dp_z \quad (2.34a)$$

is obviously the μ -component of 4-momentum per unit volume; i.e., T^{00} is the energy per unit volume and T^{j0} is the j -component of momentum per unit volume. Similarly, the component

$$\begin{aligned} T^{\mu x} &= \int \mathcal{N} p^\mu p_x \frac{dp_x dp_y dp_z}{p^0} = \int \frac{dN}{dx dy dz dp_x dp_y dp_z} \frac{dx}{dt} p^\mu dp_x dp_y dp_z \\ &= \int \frac{dN}{dt dy dz dp_x dp_y dp_z} p^\mu dp_x dp_y dp_z \end{aligned} \quad (2.34b)$$

is the amount of μ -component of 4-momentum per unit time that crosses a unit area orthogonal to the x -direction; 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 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 the familiar stress-energy-tensor relations

$$\begin{aligned} T^{00} &= (\text{energy density}) , & T^{j0} &= (\text{momentum density}) = T^{0j} = (\text{energy flux}) , \\ T^{jk} &= (\text{stress tensor}) . \end{aligned} \quad (2.34c)$$

Notice that the energy density includes the contribution of rest mass, since $\mathcal{E} = p^0$ includes rest mass. Later we shall find quite useful the following specialization of Eq. (2.30c) to the stress tensor:

$$\begin{aligned} T_{jk} &= \int \mathcal{N} p_j p_k \frac{dp_x dp_y dp_z}{\mathcal{E}} \\ &\simeq \int \mathcal{N} p_j p_k \frac{dp_x dp_y dp_z}{m} \quad \text{in the nonrelativistic (low-velocity) limit.} \end{aligned} \quad (2.34d)$$

The *mean rest frame* of the particles, at some event \mathcal{P} , is that frame in which the particle flux S_j vanishes. We shall denote by \vec{u}_{rf} the 4-velocity of this mean rest frame. Especially interesting are 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. Since the energy of a particle as measured in the mean rest frame is also a function of $|\mathbf{p}| = p$ and not of the momentum direction ($\mathcal{E} = \sqrt{m^2 + p^2}$), we can equally well say that isotropic distributions are those for which \mathcal{N} depends on the 4-momentum only through the energy as measured in the mean rest frame, i.e., only through

$$\begin{aligned} \mathcal{E} &= -\vec{u}_{\text{rf}} \cdot \vec{p} \quad \text{expressed in frame-independent form,} \\ \mathcal{E} &= p^0 = (m^2 + p^2)^{\frac{1}{2}} \quad \text{in mean rest frame .} \end{aligned} \quad (2.35)$$

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 Lorentz frame. As seen in some other (“primed”) frame, \vec{u}_{rf} will have a time component $u_{\text{rf}}^0 = \gamma$ and a space component $\mathbf{u}'_{\text{rf}} = \gamma \mathbf{V}$ [where \mathbf{V} is the mean rest frame’s velocity relative to the primed frame and $\gamma = (1 - \mathbf{V}^2)^{\frac{1}{2}}$]; and correspondingly, \mathcal{N} will be a function of

$$\mathcal{E} = -\vec{u}_{\text{rf}} \cdot \vec{p} = \gamma[(m^2 + \mathbf{p}'^2)^{\frac{1}{2}} - \mathbf{V} \cdot \mathbf{p}'] , \quad (2.36)$$

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. 2.4 above.

2.4.2 Perfect Fluids and Equations of State

Consider a collection of particles whose distribution function is isotropic in their mean rest frame. Such isotropy is readily produced by collisions between the particles. As measured in the mean rest frame, the components of the number-flux 4-vector (2.30b) are

$$n \equiv S^0 = \int \mathcal{N} dp_x dp_y dp_z = 4\pi \int_0^\infty \mathcal{N} p^2 dp = \text{number density} , \quad (2.37a)$$

$$S_j = \int \mathcal{N} \frac{p_j}{\mathcal{E}} dp_x dp_y dp_z = 0 , \quad \mathcal{E} = \sqrt{m^2 + p^2} . \quad (2.37b)$$

Here in the third equality of Eq. (2.37a) we have switched from Cartesian to spherical coordinates in momentum space. The integral (2.37b) vanishes because the integrand $\mathcal{N} p_j / \mathcal{E}$ is odd in the integration variable p_j , and the integration being over the entire mass hyperboloid runs from $p_j = -\infty$ to $p_j = +\infty$. Similarly, the time-time component of the stress-energy tensor (2.30c), i.e., the energy density as measured in the rest frame, is

$$\rho \equiv T^{00} = \int \mathcal{N} \mathcal{E} dp_x dp_y dp_z = 4\pi \int_0^\infty \mathcal{N} \mathcal{E} p^2 dp , \quad (2.38a)$$

and the energy flux (same as momentum density) is

$$T^{0j} = \int \mathcal{N} p_j dp_x dp_y dp_z = 0 , \quad (2.38b)$$

which vanishes because the integrand is odd in p_j . Finally, because the distribution function is isotropic in the momentum, the stress tensor T_{jk} is isotropic, i.e., it is a multiple of the metric $g_{jk} = \delta_{jk}$ of the rest frame's 3-dimensional space, so

$$T_{jk} = P \delta_{jk} . \quad (2.38c)$$

Here P is a scalar that is defined only in this isotropic situation.

Notice that the stress-energy tensor (2.38a)–(2.38c) is that of a *perfect fluid* [Eq. (1.154)]. The mean rest frame of the particles, in which we are computing its components, is the rest frame of the fluid, and ρ and P are the fluid's density and pressure as measured in this mean rest frame.

The most computationally useful expression for the fluid's pressure P is that obtained by taking one-third the trace of the stress tensor [Eq. (2.34d)]:

$$P = \frac{1}{3}(T_{xx} + T_{yy} + T_{zz}) = \frac{1}{3} \int \mathcal{N} |\mathbf{p}|^2 \frac{dp_x dp_y dp_z}{\mathcal{E}} = \frac{4\pi}{3} \int_0^\infty \mathcal{N} \frac{p^4 dp}{\sqrt{m^2 + p^2}} , \quad (2.38d)$$

where $p = |\mathbf{p}|$ is the magnitude of the momentum.

If we know the distribution function for an isotropic collection of relativistic particles, Eqs. (2.37a), (2.38a), and (2.38d) give us a straightforward way of computing the collection's number density of particles n , its perfect-fluid energy density ρ , and its perfect-fluid pressure P as measured in the mean rest frame. For the thermalized distribution functions (2.17a),

(2.17b), and (2.20) [with $\mathcal{N} = (g_s/h^3)\eta$], that calculation gives n , ρ and P in terms of the particles' temperature T and chemical potential μ_R . One can then invert the expression for $n(\mu_R, T)$ to give $\mu_R(n, T)$ and then insert into the expressions for ρ and P to obtain *equations of state* for thermalized particles

$$\rho = \rho(n, T) , \quad P = P(n, T) . \quad (2.39)$$

The details of this calculation are carried out in Ex. 2.5 for a gas of thermalized photons with vanishing chemical potential (i.e., with an isotropic, black-body distribution); the result is

$$\rho = aT^4 , \quad P = \frac{1}{3}\rho , \quad (2.40a)$$

where

$$a = \frac{8\pi^5}{15} \frac{k^4}{h^3 c^3} = 7.56 \times 10^{-15} \text{erg cm}^{-3} \text{K}^{-4} = 7.56 \times 10^{-16} \text{J m}^{-3} \text{K}^{-4} \quad (2.40b)$$

is the ‘‘radiation constant.’’

As we shall see in Sec. 27.4, 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 roughly the same as the photons), so its equation of state was given by Eq. (2.40a) with the coefficient changed by a factor of order unity. Einstein's general relativistic field equations (Part VI) required that

$$\frac{3}{32\pi G\tau^2} = \rho \simeq aT^4 \quad (2.41a)$$

[Eq. (27.46)], where G is Newton's gravitation constant and τ was the age of the universe as measured in the mean rest frame of the photons. Putting in numbers, we find that

$$\rho = \frac{4.9 \times 10^{-12} \text{g/cm}^3}{(\tau/1\text{yr})^2} , \quad T \simeq \frac{10^6 \text{K}}{\sqrt{\tau/1\text{yr}}} . \quad (2.41b)$$

This implies that when the universe was one minute old, its radiation density and temperature were about 1 g/cm^3 and $6 \times 10^8 \text{ K}$. 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. 27.4.

Returning to our use of kinetic theory to compute equations of state: In the nonrelativistic limit, the energy \mathcal{E} of a particle becomes $m + E$ where $E = p^2/2m$, the relativistic energy density ρ becomes $\rho = mn + \epsilon$, where ϵ is the energy density *excluding* rest mass, and Eqs. (2.37a), (2.38a), and (2.38d) reduce to the following nonrelativistic expressions:

$$n = \int_0^\infty \mathcal{N} 4\pi p^2 dp , \quad \epsilon = \frac{3}{2}P = \int_0^\infty \mathcal{N} \frac{p^2}{2m} 4\pi p^2 dp . \quad (2.42)$$

In Ex. 2.6, we use these expressions to derive the equations of state for a thermalized, nonrelativistic ($kT \ll mc^2$) classical ($\eta \ll 1$, i.e. $\mu/kT \ll -1$) gas; the result is

$$n = \frac{g_s e^{\mu/kT}}{\lambda_{T\text{dB}}^3} = \frac{g_s}{h^3} (2\pi m kT)^{3/2} e^{\mu/kT} , \quad (2.43a)$$

$$\epsilon = \frac{3}{2}nkT, \quad P = nkT, \quad (2.43b)$$

so the mean energy per particle is

$$\bar{E} = \frac{3}{2}kT. \quad (2.43c)$$

Here $\lambda_{T\text{dB}} \equiv (2\pi mkT)^{3/2}/h^3$ is the particles' *thermal deBroglie wavelength*, i.e. the wavelength of Schroedinger wave-function oscillations for a particle with the thermal kinetic energy $E = \pi kT$. Note that the classical regime $\eta \ll 1$ (i.e. $\mu/kT \ll -1$) corresponds to a mean number of particles in a thermal deBroglie wavelength small compared to one, $n\lambda_{T\text{dB}}^3 \ll 1$, which should not be surprising.

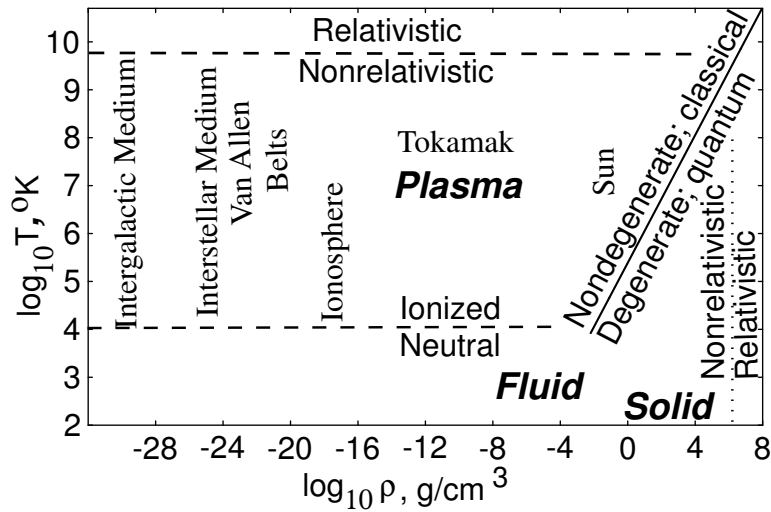


Fig. 2.5: Physical nature of matter at various densities and temperatures.

As an application, consider ordinary matter. Fig. 2.5 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 III, fluids (lower middle) in Part IV, and plasmas (middle) in Part V. Our kinetic theory tools are well suited any situation where the particles have mean free paths large compared to their sizes; this is generally true in plasmas, and sometimes in fluids (e.g. air and other gases but not water) and even sometimes in solids (e.g. for electrons in a metal). Here we shall focus on a nonrelativistic plasma, i.e. the region of Fig. 2.5 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.)

Our nonrelativistic Hydrogen plasma consists of a mixture of two gases (or “fluids”): 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 , and because the plasma is nonrelativistic, the thermal contributions to ρ are negligible.) Correspondingly, the equation of state includes equal contributions from the electrons and protons and is given by [cf. Eq. 2.43b]

$$\epsilon = 3(k/m_p)\rho T, \quad P = 2(k/m_p)\rho T. \quad (2.44)$$

The upper dashed boundary (Fig. 2.5) on the region of validity for this equation of state is the temperature $T_{\text{rel}} = m_e c^2/k = 6 \times 10^9$ K at which the electrons become strongly relativistic. In Ex. 4.5 we shall compute the thermal production of electron-positron pairs in the hot plasma and thereby shall discover that the upper boundary of validity is actually somewhat lower than this (Fig. 4.6). The lower dashed boundary is the temperature $T_{\text{rec}} \sim (\text{ionization energy of hydrogen})/(\text{a few } k) \sim 10^4$ K at which electrons and protons begin to recombine and form neutral hydrogen. In Ex. 4.6 we shall analyze the conditions for ionization-recombination equilibrium and thereby shall refine this boundary. The solid right boundary is the point at which the electrons cease to behave like classical particles, because their mean occupation number η ceases to be small compared to unity. As one can see from the Fermi-Dirac distribution (2.23a), for typical electrons (which have energies $E \sim kT$), the regime of classical behavior ($\eta \ll 1$; left side of solid line) is $\mu \ll -kT$ and the regime of strong quantum behavior ($\eta \sim 1$; electron degeneracy; right side of solid line) is $\mu \gg +kT$. The slanted solid boundary in Fig. 2.5 is thus the location $\mu = 0$, which translates via Eq. (2.43a) to

$$\rho = \rho_{\text{deg}} \equiv (2m_p/h^3)(2\pi m_e kT)^{3/2} = 0.01(T/10^4\text{K})^{3/2} \text{g/cm}^3. \quad (2.45)$$

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 (2.38a) and (2.38d), so long as we use the quantum mechanically correct distribution function for the electrons—the Fermi-Dirac distribution (2.23a). In this electron-degenerate region, the proton pressure retains its classical value, $P \propto \rho$, while the electron pressure grows much more rapidly with density. Heuristically, this is because the electrons are being confined by the Pauli exclusion Principle to regions of ever shrinking size, causing their zero-point motions and associated pressure to grow. As a result, the proton pressure becomes negligible and the electron pressure dominates. Moreover, when the density gets sufficiently high, the zero-point motions become relativistically fast (the electron chemical potential $\mu_e = \mu_{Re} - m_e$ becomes of order m_e), so the non-relativistic, Newtonian analysis fails and the matter enters a domain of “relativistic degeneracy”. Both domains, nonrelativistic degeneracy ($\mu_e \ll m_e$) and relativistic degeneracy ($\mu_e \gtrsim m_e$), occur for matter inside a massive white-dwarf star—the type of star that the Sun will become when it dies. We shall study the structures of such stars in the Fluid-Mechanics part of this book (Chap. 12); and in Chap. 25 we shall see how general relativity (spacetime curvature) modifies that structure and helps force sufficiently massive white dwarfs to collapse.

The equation of state for electron-degenerate hydrogen will be a key foundation for our study of these stars. That equation of state, due to Wilhelm Anderson and Edmund Stoner in 1930 [see the history on pp. 153–154 of Thorne (1994)] and derived in Ex. 2.7, is the following:

The electrons’ chemical potential including rest mass, μ_{Re} , is called their *Fermi energy* (including rest mass) and denoted $\mathcal{E}_F \equiv \mu_{Re}$, the corresponding momentum $p_F \equiv \sqrt{\mathcal{E}_F^2 - m_e^2}$

is called their *Fermi momentum*, and we introduce a dimensionless parameter t related to these Fermi energy and momentum by

$$\mathcal{E}_F \equiv \mu_{Re} \equiv m_e \cosh(t/4), \quad p_F \equiv \sqrt{\mathcal{E}_F^2 - m_e^2} \equiv m_e \sinh(t/4). \quad (2.46a)$$

Then the number density of electrons n_e , their total density of mass-energy ρ_e , and their pressure P are (Ex. 2.7)

$$\begin{aligned} n_e &= \frac{8\pi}{(h/m_e)^3} \left(\frac{p_F}{m_e}\right)^3 = \frac{8\pi}{(h/m_e)^3} \sinh^3(t/4), \\ \rho_e &= \frac{8\pi m_e}{(h/m_e)^3} \int_0^{p_F/m_e} x^2 \sqrt{1+x^2} dx = \frac{\pi}{(h/m_e)^3} [\sinh(t) - t], \\ P_e &= \frac{8\pi}{(h/m_e)^3} \int_0^{p_F/m_e} \frac{x^4}{\sqrt{1+x^2}} dx = \frac{\pi}{12(h/m_e)^3} [\sinh(t) - 8 \sinh(t/2) + 3t]. \end{aligned} \quad (2.46b)$$

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

$$\rho \simeq m_p n_e = \frac{8\pi m_p}{(h/m_e)^3} \sinh^3(t/4). \quad (2.46c)$$

The dividing line between nonrelativistic and relativistic degeneracy is the point where the electron Fermi momentum is equal to the rest mass, i.e. $\sinh t = 1$, which corresponds to a total mass density

$$\rho_{\text{rel deg}} = \frac{8\pi m_p}{3(h/m_e)^3} \simeq 1.0 \times 10^6 \text{ g/cm}^3. \quad (2.46d)$$

The equation of state (2.46b) implies

$$\begin{aligned} P_e &\propto \rho^{5/3} && \text{in nonrelativistic regime, } \rho \ll \rho_{\text{rel deg}}, \\ P_e &\propto \rho^{4/3} && \text{in relativistic regime, } \rho \gg \rho_{\text{rel deg}}. \end{aligned} \quad (2.46e)$$

These asymptotic equations of state turn out to play a crucial role in the structure and stability of white-dwarf stars [Chaps. 12 and 25; Chap. 4 of Thorne(1994)].

EXERCISES

Exercise 2.5 *Example and Derivation: 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. (2.38a) and (2.38d), that this photon gas obeys the equation of state $P = \frac{1}{3}\rho$.

n	1	2	3	4	5
$\zeta(n)$	∞	$\frac{\pi^2}{6}$	1.20206	$\frac{\pi^4}{90}$	1.03693
B_{2n}	$\frac{1}{6}$	$-\frac{1}{30}$	$\frac{1}{42}$	$-\frac{1}{30}$	$\frac{5}{66}$

Table 2.1: Zeta function and Bernoulli Numbers

(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. (2.40b). *Note:* Equations in the note at the end of this problem may be helpful.

(c) Show that for this isotropic, black-body photon gas the number density of photons is

$$n = bT^3, \quad \text{where } b = 16\pi\zeta(3) \left(\frac{k}{hc}\right)^3, \quad (2.47a)$$

and that the mean energy of a photon in the gas is

$$\bar{\mathcal{E}}_\gamma = \frac{\pi^4}{30\zeta(3)}kT = 2.7011780\dots kT. \quad (2.47b)$$

Here $\zeta(q)$ is the Riemann Zeta function.

Note: The Riemann Zeta function $\zeta(q)$, for $q > 1$, is given by

$$\zeta(q) = \sum_{n=1}^{\infty} \frac{1}{n^q}, \quad (2.48a)$$

and it appears in the following integral, which arises in this problem:

$$\int_0^{\infty} \frac{x^{q-1}}{e^x - 1} dx = \Gamma(q)\zeta(q). \quad (2.48b)$$

Here $\Gamma(q)$ is the gamma function, which is equal to $(q-1)!$ if q is an integer. For the special case where q is an even integer and is ≥ 2 , so $q-1$ is odd and ≥ 1 , the integral (2.48b) takes on the value

$$\int_0^{\infty} \frac{x^{2q-1}}{e^x - 1} dx = (-1)^{q-1} (2\pi)^{2q} \frac{B_{2q}}{4q}, \quad (2.48c)$$

where B_n is the Bernoulli number. Some values of the Riemann Zeta function and Bernoulli number are given in Table 2.1.

Exercise 2.6 *Example and Derivation: Equation of State for a 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 $kT \ll mc^2$ (nonrelativistic temperature).

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

$$\mathcal{N} = \frac{g_s}{h^3} e^{\mu/kT} e^{-p^2/2mkT}, \quad \text{where } p = |\mathbf{p}| = mv, \quad (2.49)$$

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. (2.43a).
 (c) Show that this gas satisfies the equations of state (2.43b).

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

$$\int_0^\infty x^{2q} e^{-x^2} dx = \frac{(2q-1)!!}{2^{q+1}} \sqrt{\pi}, \quad (2.50)$$

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

$$\int_0^\infty x^{2q+1} e^{-x^2} dx = \frac{1}{2} q!. \quad (2.51)$$

Exercise 2.7 *Problem and Derivation: Equation of State for Electron-Degenerate Hydrogen*
 Derive the equation of state (2.46b) for an electron-degenerate hydrogen gas. In your derivation, approximate the temperature T as zero, explain why this approximation is justified, and deduce that the electrons' mean occupation number is 1 for $\mathcal{E} < \mathcal{E}_F$ and 0 for $\mathcal{E} > \mathcal{E}_F$. (Note: It might be easiest to compute the integrals with the help of symbolic manipulation software such as Mathematica or Maple.)

Exercise 2.8 *Example: 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 $kT \ll \mu_e \ll m_e$ are equivalent to the words “nonrelativistically degenerate”. Here $\mu_e \equiv \mu_{Re} - m_e$ is the chemical potential (Fermi energy) excluding rest mass.
 (b) Show that the electrons' mean occupation number $\eta(E)$ has the form depicted in Fig. 2.6: It is near unity out to (nonrelativistic) energy $E \simeq \mu_e - 2kT$, and it then drops to nearly zero over a range of energies $\Delta E \sim 4kT$.
 (c) If the electrons were nonrelativistic but *nondegenerate*, their thermal energy density would be $\epsilon = \frac{3}{2}nkT$, so the total electron energy (excluding rest mass) in a volume V containing $N = nV$ electrons would be $E = \frac{3}{2}NkT$, and the electron specific heat, at fixed volume, would be

$$C_V \equiv \left(\frac{\partial E}{\partial T} \right)_V = \frac{3}{2} Nk \quad (\text{nondegenerate, nonrelativistic}). \quad (2.52)$$

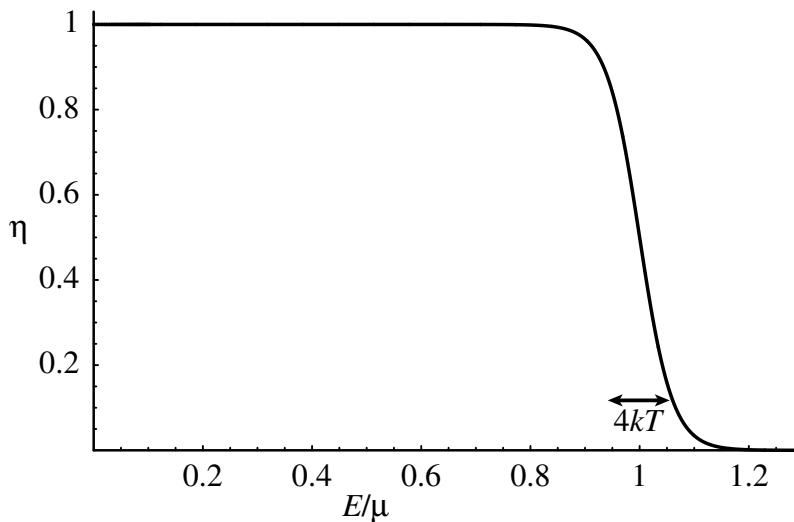


Fig. 2.6: The Fermi-Dirac distribution function for electrons in the nonrelativistic, degenerate regime $kT \ll \mu_e \ll m_e$, with temperature such that $kT/\mu = 0.03$. Note that η drops from near one to near zero over the range $\mu - 2kT \lesssim E \lesssim \mu + 2kT$.

Using the semiquantitative form of η depicted in Fig. 2.6, 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 $\sim kT/\mu_e$:

$$C_V \equiv \left(\frac{\partial E}{\partial T} \right)_V \sim \left(\frac{kT}{\mu_e} \right) Nk \quad (\text{degenerate, nonrelativistic}). \quad (2.53)$$

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

$$C_V = \frac{\pi^2}{2} \left(\frac{kT}{\mu_e} \right) Nk. \quad (2.54)$$

- (e) As a specific application, consider hydrogen inside a white dwarf star with density $\rho = 10^5 \text{ g cm}^{-3}$ and temperature $T = 10^6 \text{ K}$. (These are typical values for a white dwarf interior). What are the numerical values of μ_e/m_e and kT/μ_e for the electrons? What is the numerical value of the dimensionless factor $(\pi^2/2)(kT/\mu_e)$ 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. What are the numerical values of μ_e/m_e and kT/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 $(\pi^2/2)(kT/\mu_e)$ 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 2.9 *Example: Specific Heat for Phonons in an Isotropic Solid*

In Sec. 11.3 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 $c_s = \omega/|\mathbf{k}|$. 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 $\mathcal{E} = \hbar\omega$. Combining these relations we learn that the relationship between a phonon's energy and the magnitude $p = |\mathbf{p}|$ of its momentum is $\mathcal{E} = c_s p$. 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?
- (b) Phonons are fairly easily created, absorbed, scattered and thermalized. A general argument regarding chemical reactions (Sec. 4.4) 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 $\mu_R = 0$ some temperature T . 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 larger 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): $\mathcal{E}_{\text{Total}} = a_s T^4 V$, with $a_s = g_s(4\pi^5/15)(k^4/h^3 c_s^3)$; cf. Eqs. (2.40).
- (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.
- (d) Show that in the phonon gas, only phonon modes with wavelengths longer than $\sim \lambda_T = c_s h/kT$ 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.

2.5 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 2.1, Sec. 3.3, and Chap. 27. 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 collection \mathcal{G} of particles which are initially all near some location $(\mathcal{P}_o, \vec{p}_o)$ in phase space and initially occupy an infinitesimal (frame-independent) phase-space volume $d\mathcal{V}_x d\mathcal{V}_p$ there. Pick a particle at the center of the collection \mathcal{G} and call it the "fiducial particle". Since all the particles in \mathcal{G} have nearly the same initial position \mathcal{P}_o and 4-velocity $\vec{u}_o = \vec{p}_o/m$, they subsequently all move along nearly the same world line through spacetime; i.e., they all remain congregated around the fiducial particle. We thus can regard their volume $d\mathcal{V}_x d\mathcal{V}_p$ in phase space as a function of the fiducial particle's parameter ζ (so normalized that the particle's straight world line is $\mathcal{P}(\zeta)$ and its conserved 4-momentum is $\vec{p}_o = d\mathcal{P}/d\zeta$). Liouville's theorem says that the phase-space volume occupied by the particles \mathcal{G} is conserved,

$$\frac{d}{d\zeta}(d\mathcal{V}_x d\mathcal{V}_p) = 0 ; \quad (2.55)$$

i.e., it is a constant along the world line of the fiducial particle.

We can prove Liouville's theorem with the aid of the diagrams in Fig. 2.7. Assume, for simplicity, that the particles have nonzero rest mass, and switch from ζ to the fiducial particle's proper time $\tau = m\zeta$ as the parameter with respect to which changes are monitored. Consider the region in phase space occupied by the particles, as seen in the Lorentz frame of the fiducial particle (so Lorentz coordinate time t is equal to τ). Choose that region to be a rectangular box at the initial event \mathcal{P}_o [Fig. 2.7(a)], and set $t = \tau = 0$ at that initial event. 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 $t = \tau$ passes,

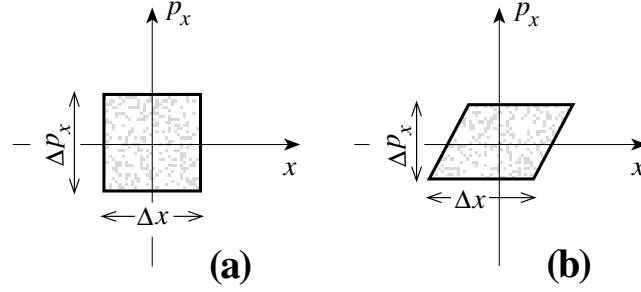


Fig. 2.7: The phase space region (x - p_x part) occupied by a collection of particles with finite rest mass, as seen in the Lorentz frame of the central, fiducial particle. (a) The initial region. (b) The region after a short time.

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 Lorentz frame. Because the particles move freely, each one 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 as shown in Fig. 2.7(b). 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 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 collection \mathcal{G} will be conserved, Liouville's theorem immediately implies also the conservation of the number density in phase space, $\mathcal{N} = dN/d\mathcal{V}_x d\mathcal{V}_p$:

$$\frac{d\mathcal{N}}{d\zeta} = 0 \quad \text{along the trajectory of a fiducial particle.} \quad (2.56)$$

This conservation law is called the *collisionless Boltzmann equation*; and in the context of plasma physics 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 2.1, which is best read after finishing this section.

The collisionless Boltzmann equation is most nicely expressed in the frame-independent form (2.56). For some purposes, however, it is helpful to express the equation in a form that relies on a specific but arbitrary choice of Lorentz reference frame. Then $\mathcal{N}(\mathcal{P}, \vec{p})$ can be regarded as a function of the seven phase-space coordinates, $\mathcal{N} = \mathcal{N}(x^\mu, p_j)$, and the collisionless Boltzmann equation (2.56) then takes the coordinate-dependent form

$$\frac{d\mathcal{N}}{d\zeta} = \frac{dx^\mu}{d\zeta} \frac{\partial \mathcal{N}}{\partial x^\mu} + \frac{dp_j}{d\zeta} \frac{\partial \mathcal{N}}{\partial p_j} = \sqrt{m^2 + |\mathbf{p}|^2} \left(\frac{\partial \mathcal{N}}{\partial t} + v_j \frac{\partial \mathcal{N}}{\partial x_j} \right) = 0. \quad (2.57)$$

Here we have used the equations of motion for the fiducial particle: $dx^\mu/d\zeta = p^\mu$, $p^0 = \mathcal{E} = \sqrt{m^2 + |\mathbf{p}|^2}$, $p_j = p^0 v_j$, and $dp_j/dt = 0$.

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, $\vec{\nabla} \cdot \vec{S} = 0$ [cf. Eq. (2.33)]. Indeed, this is so; see Ex. 2.10. Similarly, since the collisionless Boltzmann equation is based on the law of 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 4-momentum, i.e. will guarantee that $\vec{\nabla} \cdot \mathbf{T} = 0$. Indeed, this conservation law does follow from the collisionless Boltzmann equation; see Ex. 2.10.

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

$$\frac{d\mathcal{N}}{d\zeta} = \left(\frac{d\mathcal{N}}{d\zeta} \right)_{\text{collisions}}, \quad (2.58)$$

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. (2.67), (2.72a), and Ex. 2.14] and in our study of plasmas (Chaps. 21 and 22).

Whenever one applies the collisionless Boltzmann 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. 2.11. (ii) Adjust the differentiation parameter (ζ above) so it is also naturally tuned to the situation. For example, instead of ζ , one might use the distance l or time t traveled in some preferred inertial reference frame.

EXERCISES

Exercise 2.10 *Derivation and Problem: Collisionless Boltzmann Implies Conservation of Particles and of 4-Momentum*

Consider a collection of freely moving, noncolliding particles, which satisfy the collisionless Boltzmann equation $d\mathcal{N}/d\zeta = 0$. Show that this equation guarantees that the conservation laws $\vec{\nabla} \cdot \vec{S} = 0$ and $\vec{\nabla} \cdot \mathbf{T} = 0$ are satisfied, where the number-flux vector \vec{S} and the stress-energy tensor \mathbf{T} are expressed in terms of \mathcal{N} by the momentum-space integrals (2.30b) and (2.30c).

Exercise 2.11 *Problem: Solar Heating of the Earth: The Greenhouse Effect*

Box 2.1

Sophisticated Derivation of collisionless Boltzmann Equation

Denote by $\vec{X} \equiv \{\mathcal{P}, \vec{p}\}$ a point in 8-dimensional phase space. In a Lorentz frame the coordinates of \vec{X} can be taken to be $\{x^0, x^1, x^2, x^3, p_0, p_1, p_2, p_3\}$. [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. 2.1.

A particle in our distribution 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 Eqs. (2.32). 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 a Lorentz frame, this conservation law says

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

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

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

For the freely moving particles of this chapter, the relativistic Hamiltonian is [cf. Eq. (8.62) of Goldstein (1980) and pp. 488–489 of Misner, Thorne and Wheeler (1973)]

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

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 VI). By inserting Hamilton’s equations (2) into the 8-dimensional law of particle conservation (1), we obtain

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

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

Box 2.1, Continued

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

$$0 = \frac{\partial \mathcal{N}}{\partial x^\alpha} \frac{\partial \mathcal{H}}{\partial p^\alpha} - \frac{\partial \mathcal{N}}{\partial p^\alpha} \frac{\partial \mathcal{H}}{\partial x^\alpha} = \frac{\partial \mathcal{N}}{\partial x^\alpha} \frac{dx^\alpha}{d\zeta} - \frac{\partial \mathcal{N}}{\partial p^\alpha} \frac{dp^\alpha}{d\zeta} = \frac{d\mathcal{N}}{d\zeta}, \quad (1)$$

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; cf. Eq. (2.57).] *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, which implies it has very great generality.* We shall extend and explore this generality in the next chapter.

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

$$F \equiv \frac{d\mathcal{E}}{dt dA} = \sigma T_\odot^4, \quad \text{where } \sigma = \frac{ac}{4} = \frac{2\pi^5}{15} \frac{k^4}{h^3 c^2} = 5.67 \times 10^{-5} \frac{\text{erg}}{\text{cm}^2 \text{sK}^4}. \quad (2.59)$$

- (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 specific 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. (2.59) 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

$$T_\oplus = T_\odot \left(\frac{R_\odot}{2r} \right)^{1/2} = 280 \text{ K} = 7 \text{ C}. \quad (2.60)$$

This is a bit cooler than the correct mean surface temperature ($287\text{ K} = 14\text{ C}$).

- (d) Actually, the Earth has an “albedo” of $A = 0.30$, which means that 40 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

$$T_{\oplus} = T_{\odot} \left(\frac{\sqrt{1 - AR_{\odot}}}{2r} \right)^{1/2} = 255\text{ K} = -18\text{ C} . \quad (2.61)$$

This is even farther from the correct answer.

- (e) The missing piece of physics, which raises the temperature from -27C to something much nearer the correct 14C , is the *Greenhouse Effect*: The absorbed solar radiation has most of its energy at wavelengths $\sim 0.5\mu\text{m}$ (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 $\sim 300\text{K}$, is concentrated in the infrared range from $\sim 8\mu\text{m}$ to $\sim 30\mu\text{m}$. Water molecules and carbon dioxide in the Earth’s atmosphere absorb about half of the energy that the Earth tries to reradiate at these energies,⁵ causing the reradiated energy to be about half that of a black body at the Earth’s surface temperature. Show that with this “Greenhouse” correction, T_{\oplus} becomes about $293\text{K} = +20\text{C}$. Of course, the worry is that human activity will increase the amount of carbon dioxide in the atmosphere by enough to raise the Earth’s temperature still further, thereby disrupting our comfortable lives.

Exercise 2.12 Challenge: Olbers’ Paradox and Solar Furnace

Consider a universe in which spacetime is flat 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, $\sim 10,000\text{ K}$, so we would all be fried. 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, 5770 K .

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

⁵See, e.g., the section and figures on “Absorption of Atmospheric Gases” in Allen (2000)

What are transport coefficients? An example is the *electrical conductivity* σ of a sample of matter. Ohm's law says that, as measured in the sample's proper reference frame, the current density \mathbf{J} is directly proportional to the electric field \mathbf{E} , with proportionality constant σ ; i.e., $\mathbf{J} = \sigma \mathbf{E}$. 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 the electron motion is scattering off other particles—off ions, other electrons, phonons (sound waves), plasmons (plasma waves), In order to compute the electrical conductivity, one must analyze the effects of those scatterings on the motion of the electrons; and since huge numbers of electrons and scatterers are involved, the analysis must be statistical. The foundation for an accurate analysis is the Boltzmann transport equation.

Another example of a transport coefficient is the *thermal conductivity* κ , which enters into the law of heat conduction, $\mathbf{q} = -\kappa \nabla T$ where \mathbf{q} is the energy flux and T is the temperature as measured in the mean rest frame of the matter. The impediment to heat conduction is the scattering of the conducting particles off other particles; and, correspondingly, the foundation for accurately computing the thermal conductivity κ is the Boltzmann transport equation.

Other examples of transport coefficients are the *coefficient of shear viscosity* η_{shear} which enters into the Navier Stokes equation of fluid mechanics (Part IV of this book, where it is denoted η), and which describes the ability of a fluid to transport momentum in directions transverse to its flow; and the *diffusion coefficient* D which enters into the diffusion equation (Ex. 2.13 below) and describes the ability of particles to diffuse through matter.

The physical laws into which these transport coefficients enter (Ohm's law, the law of heat conduction, the Navier-Stokes equation, the diffusion equation) 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 for validity 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 $\mathcal{L} \sim T/|\nabla T|$, 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 that $\lambda \ll \mathcal{L}$. 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 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 $(d\mathcal{N}/d\zeta)_{\text{collisions}}$ 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.

In this section we shall compute the coefficient of thermal conductivity κ first by an order-of-magnitude argument, and then by the Boltzmann equation with a highly accurate collision term. In Ex. 2.13 readers will have the opportunity to compute the diffusion coefficient using a moderately accurate collision term.

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

The specific problem we shall treat here in the text is 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 $10^7 \text{ K} \lesssim T \lesssim 10^9 \text{ K}$, the mass density is $\rho \lesssim 10 \text{ g/cm}^3 (T/10^7 \text{ K})^2$, and *heat is carried primarily by diffusing photons* rather than by diffusing electrons or ions or by convection. (We shall study convection in Chap. 17.) In this regime the primary impediment to the photons' flow is collisions with electrons. The lower limit on temperature, $10^7 \text{ K} \sim T$, guarantees that the gas is almost fully ionized, so there is a plethora of electrons to do the scattering. The upper limit on density, $\rho \sim 10 \text{ g/cm}^3 (T/10^7 \text{ K})^2$ guarantees that (i) 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; and (ii) the inelastic scattering, absorption, and emission of photons by electrons that are accelerating in the coulomb fields of ions ("bremsstrahlung" processes) are unimportant as impediments to heat flow compared to *scattering off free electrons*. The upper limit on temperature, $T \sim 10^9 \text{ K}$, guarantees that (i) the electrons which do the scattering are moving thermally at much less than the speed of light (the mean thermal energy $1.5kT$ of an electron is much less than its rest mass-energy $m_e c^2$); 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

$$\frac{d\sigma(\mathbf{n}' \rightarrow \mathbf{n})}{d\Omega} = \frac{3}{16\pi} \sigma_T [1 + (\mathbf{n} \cdot \mathbf{n}')^2] . \quad (2.62a)$$

Here σ_T is the total Thomson cross section [the integral of the differential cross section (2.62a) over solid angle]

$$\sigma_T = \int \frac{d\sigma(\mathbf{n}' \rightarrow \mathbf{n})}{d\Omega} d\Omega = \frac{8\pi}{3} r_o^2 = 0.665 \times 10^{-24} \text{ cm}^2 , \quad (2.62b)$$

where $r_o = e^2/m_e c^2$ is the classical electron radius. For a derivation and discussion of the differential Thomson cross section (2.62a) and the total Thomson cross section (2.62b) see, e.g., Sec. 14.8 of Jackson (1999).

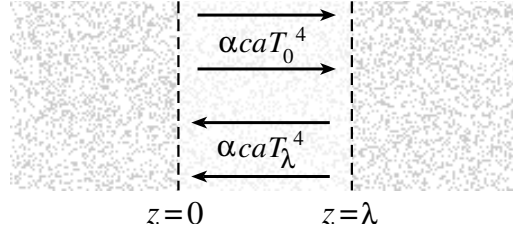


Fig. 2.8: Heat exchange between two layers of gas separated by a distance of one photon mean free path in the direction of the gas’s temperature gradient.

2.6.2 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 σ_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 $n_e \sigma_T l$; and therefore to build up to unit probability for scattering, the photon must travel a distance

$$\lambda \sim \frac{1}{n_e \sigma_T} \sim \frac{m_p}{\rho \sigma_T} \sim 3 \text{ cm} \left(\frac{1 \text{ g/cm}^3}{\rho} \right) \sim 3 \text{ cm} . \quad (2.63)$$

Here m_p is the proton rest mass, $\rho \sim 1 \text{ g/cm}^3$ 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 $n_e \simeq \rho/m_p$. Note that $\mathcal{L} \sim 10^5$ km is roughly 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{q} as measured in the gas’s rest frame is related to the temperature gradient ∇T by

$$\mathbf{q} = -\kappa \nabla T , \quad (2.64)$$

where κ is the thermal conductivity. To estimate 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. 2.8). 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

$$q_{\lambda \rightarrow 0} = \alpha ca(T_0)^4, \quad (2.65a)$$

where a is the radiation constant of Eq. (2.40b), $a(T_0)^4$ 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 cgs units with c present explicitly). Similarly, the flux of energy from the layer at $z = \lambda$ to the layer at $z = 0$ is

$$q_{\lambda \rightarrow 0} = -\alpha ca(T_\lambda)^4; \quad (2.65b)$$

and the net rightward flux, the sum of (2.65a) and (2.65b), is

$$q = \alpha ca[(T_0)^4 - (T_\lambda)^4] = -4\alpha caT^3\lambda \frac{dT}{dz}. \quad (2.65c)$$

Noting that 4α is of order unity, inserting expression (2.63) for the photon mean free path, and comparing with the law of diffusive heat flow $\mathbf{q} = -\kappa \nabla T$, we conclude that the thermal conductivity is

$$\kappa \sim aT^3 c\lambda = \frac{acT^3}{\sigma_T n_e}. \quad (2.66)$$

2.6.3 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.

The effects of electron scattering are most easily written down in terms of the specific intensity I_ν of the photons, so we initially shall use I_ν as our distribution function. (We can use I_ν rather than $I_\nu/\nu^3 = (h^4/c^2)\mathcal{N}$ because each photon's frequency ν is conserved in our chosen Lorentz frame as it propagates.) It is easiest to study the scattering as a function of the Lorentz coordinate time t the photons have been traveling, or equally well the distance ct traveled, so we shall use t as our independent variable.

Consider those photons that propagate very nearly in the \mathbf{n} direction as seen in the rest frame of the gas, and that have frequencies very nearly equal to ν . The specific intensity [Eq. (2.8)] describes the energy per unit time per unit area per unit solid angle per unit frequency ($d\mathcal{E}/dtdAd\Omega d\nu$; “energy per unit everything”) carried by those photons. If observers in the rest frame of the gas watch the evolution of I_ν in the neighborhood of a fiducial photon (i.e., moving with that photon at the speed of light in the \mathbf{n} direction), they will see I_ν 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. Since these scatterings leave the frequency of the scattered photon unchanged (a key property of Thomson scattering), we can treat the frequency ν and the frequency interval $d\nu$ as constant in our description of the scattering. Scattering out of and into the

\mathbf{n} direction are embodied quantitatively in the two terms that contribute to the following expression for the rate of change of I_ν :

$$\frac{dI_\nu(\mathbf{n}, \nu)}{dt} = -\sigma_T n_e c I_\nu(\mathbf{n}, \nu) + \int \frac{d\sigma(\mathbf{n}' \rightarrow \mathbf{n})}{d\Omega} n_e c I_\nu(\mathbf{n}', \nu) d\Omega' . \quad (2.67)$$

Here t is proper time (Lorentz coordinate time) in the gas's rest frame, n_e is the number density of electrons in that frame, and all other quantities have previously been defined. The first term represents scattering out of the \mathbf{n} direction into other directions; the second, scattering into the \mathbf{n} direction from other directions \mathbf{n}' . That these terms have the forms shown should be obvious from the physical meanings of the scattering cross sections, together with the fact that the photons are moving with the speed of light c .

From the relation $\mathcal{N} = \text{constant} \times I_\nu / \nu^3$ [Eq. (2.9)], and the fact that the photon frequency ν (as measured in the gas rest frame) is a constant along the trajectory of our fiducial photon, we infer that the distribution function \mathcal{N} must satisfy the same transport equation (2.67) as the specific intensity:

$$\frac{d\mathcal{N}(t, \mathbf{x}, \mathbf{n}, \nu)}{dt} = -\sigma_T n_e c \mathcal{N}(t, \mathbf{x}, \mathbf{n}, \nu) + \int \frac{d\sigma(\mathbf{n}' \rightarrow \mathbf{n})}{d\Omega} n_e c \mathcal{N}(t, \mathbf{x}, \mathbf{n}', \nu) d\Omega' . \quad (2.68a)$$

Here our notation shows explicitly that \mathcal{N} depends not only on the photon propagation direction and frequency (as measured in the mean rest frame of the gas), but also on time t and spatial location \mathbf{x} . To make the notation fully explicit, we must also note that the time derivative d/dt on the left side is taken moving with the fiducial photon, i.e.,

$$\frac{d}{dt} = \frac{\partial}{\partial t} + c n_j \frac{\partial}{\partial x_j} . \quad (2.68b)$$

(There are no contributions $(d\nu/dt)\partial/\partial\nu$ or $(dn_j/dt)\partial/\partial n_j$ in (2.68b) because photon frequency ν and direction of propagation n_j are constants along the trajectory of a fiducial photon in the gas's mean rest frame; i.e., the tiny effects of gravitational redshift and of gravitational light deflection are being neglected.)

Equation (2.68a), with d/dt given by (2.68b), is the Boltzmann transport equation for the photons, written using the gas's proper time t as the parameter along the fiducial photon trajectory. The terms on the right side of (2.68a) are the explicit realizations, for our star's gas, of the Boltzmann equation's collision terms.

Because the mean free path λ is so short compared to the length scale \mathcal{L} 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 :

$$\mathcal{N} = \mathcal{N}_0 + \mathcal{N}_1 , \quad \text{where } \mathcal{N}_0 = \frac{2}{h^3} \frac{1}{e^{h\nu/kT} - 1} . \quad (2.69a)$$

Here the perfectly thermal piece \mathcal{N}_0 has the standard black-body form (2.21); it depends on the photon 4-momentum only through the frequency ν as measured in the mean rest

Box 2.2
Two-Lengthscale Expansions

Equation (2.69b) 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 lengthscales ratios λ/\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 lengthscales 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. 3, 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. 6), in the study of boundary layers in fluid mechanics (Secs. 13.4 and 14.4), and in the definition of a gravitational wave (Chap. 26).

frame of the gas, and it depends on location in spacetime only through the temperature T , which we assume is time independent in the star but is a function of spatial location, $T = T(\mathbf{x})$. 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 it is reasonable to expect (and turns out to be true) that the magnitude of \mathcal{N}_1 is

$$\mathcal{N}_1 \sim \frac{\lambda}{\mathcal{L}_0} \mathcal{N}_0. \quad (2.69b)$$

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 λ/\mathcal{L} . This is called a *two-lengthscale expansion*; see Box 2.2.

Because the temperature and density inside the star are independent of time t , the distribution function $\mathcal{N} = \mathcal{N}_0 + \mathcal{N}_1$ is independent of time; and the Boltzmann transport equation (2.68), for $\mathcal{N} = \mathcal{N}_0 + \mathcal{N}_1$ takes the form

$$\begin{aligned} cn_j \frac{\partial \mathcal{N}_0}{\partial x_j} + cn_j \frac{\partial \mathcal{N}_1}{\partial x_j} &= \left[-\sigma_T n_e c \mathcal{N}_0 + \int \frac{d\sigma(\mathbf{n}' \rightarrow \mathbf{n})}{d\Omega} n_e c \mathcal{N}_0 d\Omega' \right] \\ &+ \left[-\sigma_T n_e c \mathcal{N}_1(\mathbf{n}, \nu) + \int \frac{d\sigma(\mathbf{n}' \rightarrow \mathbf{n})}{d\Omega} n_e c \mathcal{N}_1(\mathbf{n}', \nu) d\Omega' \right]. \end{aligned} \quad (2.69c)$$

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 (2.69c). 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

$$n_j \frac{\partial T}{\partial x_j} \frac{\partial \mathcal{N}_0}{\partial T} = -\sigma_T n_e \mathcal{N}_1(\mathbf{n}, \nu) + \int \frac{d\sigma(\mathbf{n}' \rightarrow \mathbf{n})}{d\Omega} n_e \mathcal{N}_1(\mathbf{n}', \nu) d\Omega' . \quad (2.69d)$$

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 ∇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. (2.69d). Thus, we shall write \mathcal{N}_1 in the dipolar form

$$\mathcal{N}_1 = K_j(\mathbf{x}, \nu) n_j , \quad (2.69e)$$

and we shall solve the transport equation (2.69d) for the function K_j .

[*Important side remark:* This is a special case of a general situation: When solving the Boltzmann equation in diffusion situations, one is performing a power series expansion in λ/\mathcal{L} ; see Box 2.2. 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 λ/\mathcal{L} from \mathcal{N}_0 and is dipolar 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 is quadrupolar in its direction dependence. And so it continues on up to higher and higher orders.^{6]}

When we insert the dipolar expression (2.69e) into the angular integral on the right side of the transport equation (2.69d) and notice that the differential scattering cross section (2.62a) is unchanged under $\mathbf{n}' \rightarrow -\mathbf{n}'$, but $K_j n'_j$ changes sign, we find that the integral vanishes. As a result the transport equation (2.69d) takes the simplified form

$$n_j \frac{\partial T}{\partial x_j} \frac{\partial \mathcal{N}_0}{\partial T} = -\sigma_T n_e K_j n_j , \quad (2.69f)$$

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

$$\mathcal{N}_1 = -\frac{\partial \mathcal{N}_0 / \partial T}{\sigma_T n_e} \frac{\partial T}{\partial x_j} n_j . \quad (2.69g)$$

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

$$\frac{\mathcal{N}_1}{\mathcal{N}_0} \sim \frac{1}{\sigma_T n_e T} |\nabla T| \sim \frac{\lambda}{\mathcal{L}} . \quad (2.69h)$$

⁶For full details in nonrelativistic situations see, e.g., Grad (1957); and for full details in the fully relativistic case see, e.g., Thorne (1981).

We can now evaluate the energy flux carried by the diffusing photons. That flux q_i is the T^{0i} part of the stress-energy tensor. By inserting $\mathcal{N} = \mathcal{N}_0 + \mathcal{N}_1$ into the momentum-space integral (2.30c) for the stress-energy tensor and noting that because \mathcal{N}_0 is isotropic its contribution to the integral vanishes, we obtain

$$q_i \equiv T^{0i} = \int \mathcal{N}_1 p^0 p_i \frac{d\mathcal{V}_p}{p^0} = -\frac{1}{\sigma_{\text{T}} n_e} \frac{\partial T}{\partial x_j} \int \frac{\partial \mathcal{N}_0}{\partial T} n_j p^0 p_i \frac{d\mathcal{V}_p}{p^0}. \quad (2.70a)$$

By then noting that for photons $n_j p^0 = p_j$ and pulling the derivative with respect to temperature out from under the integral sign and noting that the integral then becomes that for the photons' stress tensor $T^{ji} = P\delta^{ji}$, and by using $P = \frac{1}{3}aT^4$ [Eq. (2.40a)] for the photons' pressure P , we obtain

$$\begin{aligned} q_i &= -\frac{c}{\sigma_{\text{T}} n_e} \frac{\partial T}{\partial x_j} \frac{d}{dT} \int \mathcal{N}_0 p_j p_i \frac{d\mathcal{V}_p}{p^0} = -\frac{c}{\sigma_{\text{T}} n_e} \frac{\partial T}{\partial x_j} \frac{d}{dT} \left(\frac{1}{3} a T^4 \delta^{ji} \right) \\ &= -\frac{c}{\sigma_{\text{T}} n_e} \frac{4}{3} a T^3 \frac{\partial T}{\partial x_i}. \end{aligned} \quad (2.70b)$$

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

$$\kappa = \frac{4}{3} \frac{acT^3}{\sigma_{\text{T}} n_e}. \quad (2.71)$$

Notice that this heat conductivity is 4/3 times our crude, order-of-magnitude estimate (2.66).

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 λ/\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 λ/\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 (2.71), 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 $r \simeq 0.8R_{\odot} \simeq 6 \times 10^5 \text{ km}$ (where R_{\odot} is the Sun's radius), the density and temperature are $\rho \simeq 5 \text{ g/cm}^3$ and $T \simeq 1.6 \times 10^7 \text{ K}$, so the number density of electrons is $n_e \simeq \rho/m_p \simeq 3 \times 10^{24} \text{ cm}^{-3}$. For these parameters, Eq. (2.71) gives a thermal conductivity $\kappa \simeq 7 \times 10^{17} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ K}^{-1}$. The lengthscale \mathcal{L} on which the temperature is changing is approximately the same as the radius, so the temperature gradient is $|\nabla T| \sim T/r \sim 3 \times 10^{-4} \text{ K/cm}$. The law of diffusive heat transfer then predicts a heat flux $q = \kappa |\nabla T| \sim 2 \times 10^{14} \text{ erg s}^{-1} \text{ cm}^{-2}$, and thus a total luminosity $L = 4\pi r^2 q \sim 8 \times 10^{36} \text{ erg/s} \simeq 2000 L_{\odot}$ (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 2.13 *Example: Diffusion Coefficient Computed 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 $kT \ll mc^2$, 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:

$$\left(\frac{d\mathcal{N}}{dt}\right)_{\text{collision}} = (\mathcal{N}_0 - \mathcal{N})\frac{1}{\hat{\tau}}, \quad \text{where } \mathcal{N}_0 \equiv \frac{e^{-p^2/2mkT}}{(2\pi mkT)^{3/2}}n. \quad (2.72a)$$

All frame-dependent quantities appearing here are evaluated in the mean rest frame of the scattering centers; in particular, t is Lorentz time, T is the temperature of the scattering centers, $p = |\mathbf{p}|$ is the magnitude of the test particles’ spatial momentum, n is the number density of test particles

$$n = \int \mathcal{N} dp_x dp_y dp_z, \quad (2.72b)$$

and $\hat{\tau}$ is a constant to be discussed below.

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

$$\left(\frac{dn}{dt}\right)_{\text{collision}} \equiv \int \left(\frac{d\mathcal{N}}{dt}\right)_{\text{collision}} dp_x dp_y dp_z = 0. \quad (2.73)$$

- (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 that the test particles are distributed inhomogeneously, $n = n(\mathbf{x}) \neq \text{const}$. 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) The inhomogeneity of n induces a net flux of particles \mathbf{S} , from high-density regions toward low-density regions, given by

$$\mathbf{S} = -D\nabla n . \quad (2.74a)$$

This is the analog of Ohm's law $\mathbf{J} = \sigma\mathbf{E}$ for the flow of electric charge and of $\mathbf{q} = -\kappa\nabla T$ for the flow of heat, and D is the *diffusion coefficient*. Give an order-of-magnitude derivation of the diffusion coefficient D .

- (e) Show that the law of particle conservation reduces, in this situation, to the following *diffusion equation* for $n(t, \mathbf{x})$:

$$\frac{\partial n}{\partial t} + \nabla \cdot (-D\nabla n) = 0 . \quad (2.74b)$$

- (f) Show that for this nonrelativistic problem the Boltzmann transport equation takes the form

$$\frac{\partial \mathcal{N}}{\partial t} + \frac{p_j}{m} \frac{\partial \mathcal{N}}{\partial x_j} = \frac{1}{\hat{\tau}} (\mathcal{N}_0 - \mathcal{N}) . \quad (2.75a)$$

- (g) 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. (2.72a) above, and

$$\mathcal{N}_1 = -\frac{p_j \hat{\tau}}{m} \frac{\partial n}{\partial x_j} \frac{e^{-p^2/2mkT}}{(2\pi mkT)^{3/2}} . \quad (2.75b)$$

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

- (h) Show that the perturbation \mathcal{N}_1 gives rise to a particle flux given by Eq. (2.74a), with the diffusion coefficient

$$D = \frac{kT}{m} \hat{\tau} . \quad (2.76)$$

Exercise 2.14 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 of ~ 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 in order for them to capture onto ^{235}U atoms and induce further fissions. The slowing is achieved by scattering off the moderator atoms. The scattering is elastic and is isotropic in the center-of-mass frame with total cross section $\sigma_s \simeq 4.8 \times 10^{-24} \text{cm}^2 \equiv 4.8$ barns. Using the fact that 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 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:

$$\xi \equiv -\overline{\Delta \ln E} = 1 + \frac{(A-1)^2}{2A} \ln \left(\frac{A-1}{A+1} \right) \simeq 0.158 . \quad (2.77)$$

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, $\lambda_s = 1/(\sigma_s n_s) \simeq 2.5$ 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 neutron kinetic energy E , $\mathcal{N} = \mathcal{N}(E)$. Use the Boltzmann transport equation or other considerations to develop the theory of the slowing down of the neutrons, and of their struggle to pass through the ^{238}U resonance region without getting absorbed. More specifically:

- (a) Use as the distribution function $n_E(E) \equiv dN/d^3x dE =$ (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

$$q = \sigma_s n_s \xi E n_E v , \text{ where } v = \sqrt{2mE} \quad (2.78)$$

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:

$$q = (\sigma_s n_s + \sigma_E n_E) \xi E n_E v . \quad (2.79)$$

Here n_a is the number density of absorbing (^{238}U) atoms and n_s is the number density of scattering (carbon) 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:

$$\frac{d \ln q}{d \ln E} = \frac{\sigma_a n_a}{(\sigma_s n_s + \sigma_a n_a) \xi} . \quad (2.80)$$

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

$$\frac{n_s}{n_a} \simeq \frac{\sigma_a}{\sigma_s} \left(\frac{\ln(7/6)}{\xi \ln 2} - 1 \right) \simeq 0.41 \frac{\sigma_a}{\sigma_s} \simeq 140 . \quad (2.81)$$

Thus, to maintain the reaction in the presence of the huge ^{238}U absorption resonance for neutrons, it is necessary that less than 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).

Bibliography

Allen, C. W. 2000. *Astrophysical Quantities*, fourth edition, London: The Athelone Press.

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

Corey, B. E. & Wilkinson, David T. 1976. *Bulletin of the American Astronomical Society*, **8**, 351.

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

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

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

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.
- Shkarofsky, I. P., Johnston, T. W., & Bachynski, M. P. 1966. *The Particle Kinetics of Plasmas*, New York: Addison-Wesley.
- Smoot, George F., Gorenstein, Mark V., & Muller, Richard A. 1977. Determination of anisotropy in the cosmic black body radiation”, *Physical Review Letters*, 39, 898–901.
- 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.