# Chapter 3

# Statistical Mechanics

## 3.1 Overview

While kinetic theory, as we discussed in the preceding chapter, gives a powerful description of some statistical features of matter, other features are outside its realm and must be treated using the more sophisticated tools of statistical mechanics. Examples are:
(i) *Correlations*: Kinetic theory's distribution function $\mathcal{N}$ tells us, on average, how many particles will occupy a given phase-space volume, but it says nothing about whether the particles like to clump together, or avoid each other. It is therefore inadequate to describe the distribution of galaxies, which aggregate under their mutual gravitational attraction, or that of electrons (Chap. 21), which are mutually repulsive and thus are spatially anti-correlated.
(ii) *Fluctuations*: The "test masses" in an interferometric detector for gravitational waves hang from wires as pendula and move in response to the waves. Molecules of gas hitting a pendular test mass also make it move. Kinetic theory predicts how many molecules will hit in one millisecond, on average, and how much resulting pressure squeezes the mass from all sides. But kinetic theory's distribution function $\mathcal{N}$ cannot tell us the probability that in one millisecond more molecules will hit one side than the others and thereby cause pendular motion that might be confused with a gravitational wave. The probability distribution for fluctuations is an essential tool for analyzing the noise in this and any other physical experiment, and it falls in the domain of statistical mechanics, not kinetic theory.
(iii) *Strongly interacting particles*: As should be familiar, the thermal motions of an ionic crystal are best described not by focusing attention on individual atoms (as in the "Einstein theory"), but instead by decomposing the motion into normal modes (phonons; "Debye theory"). Statistical mechanics must be invoked to describe the degree of excitation of the phonons.
(iv) *Microscopic origin of thermodynamic laws*: The laws of classical thermodynamics can be (and often are) derived from a few elementary, macroscopic postulates without any reference to the microscopic, atomic nature of matter. Kinetic theory provides a microscopic foundation for some of thermodynamics' abstract macroscopic ideas (e.g. the first law of thermodynamics) and permits the computation of equations of state. However a full appreciation of the second law and of behavior at phase transitions requires the machinery of

statistical mechanics.

In this chapter we shall develop the conceptual foundations for classical statistical mechanics, restricting our attention mainly to systems where the quantum mechanical occupation number of the constituent particles, averaged over some finite volume of phase space, is small, so that it does not matter if we are dealing with fermions or bosons. In Section 3.2 we shall introduce the concepts of systems, ensembles of systems, and the distribution function for an ensemble. In Section 3.3 we shall use Hamiltonian dynamics to study the evolution of an ensemble's distribution function and derive Liouville's theorem. In Section 3.4, we shall develop the concept of statistical equilibrium and shall derive the general forms of distribution functions for ensembles of systems that have reached statistical equilibrium. This is applied to a particular type of ensemble, called (for historical reasons) *microcanonical*, in Section 3.5. In Section 3.6 we shall introduce the concept of the entropy of an ensemble of systems and shall show that an ensemble of identical systems that are isolated from the external universe maximizes its entropy by evolving into statistical equilibrium. Having laid all these foundations, we shall develop illustrative applications of them in Sections 3.7, 3.8 and a number of exercises. Our examples will include a simple monatomic gas in both the nonrelativistic and ultrarelativistic domains, an ionized-hydrogen plasma, the mean occupation numbers of boson and fermion states, stars, galaxies, black holes, and the universe as a whole. For galaxies, black holes and the universe we will have to confront the role of gravity in statistical mechanics (Section 3.8).

## 3.2 Systems, Ensembles, and Distribution Functions

### 3.2.1 Systems

*Systems* play in statistical mechanics the same role as is played by particles in kinetic theory. A system is any physical entity that we can *imagine* replicating many times. Examples are a galaxy, the sun, a sapphire crystal, the fundamental mode of vibration of that crystal, an aluminum atom in that crystal, an electron from that aluminum atom, ....

Statistical mechanics focuses special attention on systems that couple only weakly to the rest of the universe. Stated more precisely, we are interested in systems whose relevant "internal" evolution times, $\tau_{\rm int}$, are short compared with the "external" times, $\tau_{\rm ext}$, with which they exchange energy, entropy, particles, etc with their surroundings. Such systems are said to be *semiclosed*, and in the idealized limit where one completely ignores their external interactions, they are said to be *closed*. The statistical-mechanics formalism for dealing with them is a variant of a two-lengthscale expansion (cf. Box 2.2). Let us look at our examples.

For a galaxy of, say, $10^{10}$ stars, $\tau_{\rm int}$ is the time it takes a star to cross the galaxy, $\tau_{\rm int} \sim 10^8$ yr. The external time can be thought of as the time since the galaxy's birth, or its last interaction with a neighboring galaxy, and this is $\tau_{\rm ext} \sim 10^{10}$ yr, so the galaxy is semiclosed. For a small volume of gas inside the sun, say a 1 m cube, $\tau_{\rm int}$ is the timescale for the constituent electrons, ions and photons to interact through collisions, typically $\tau_{\rm int} \sim 10^{-11}$ s, which is much smaller than the time for external heat to diffuse from the cube's surface to its center, $\tau_{\rm ext} \sim 10^{-5}$ s, so the cube is semiclosed. By contrast, an individual atom in a crystal is so strongly coupled to its neighboring atoms by electrostatic forces that $\tau_{\rm int} \sim \tau_{\rm ext}$, which means that the atom is not semiclosed. However, if instead we consider a particular, vibrational mode of the crystal as our system, then we can regard $\tau_{\rm int}$ as the mode's period, and $\tau_{\rm ext}$ as the time to couple to other modes and thereby damp the chosen mode's vibrations; and in general the damping time is long, $\tau_{\rm ext} \gg \tau_{\rm int}$, so the mode is semiclosed. (For a highly-polished, cold sapphire crystal, $\tau_{\rm ext}$ can be $\sim 10^9 \tau_{\rm int}$.) It therefore makes sense to treat the crystal's normal modes, and not its atoms, as the physical systems. Finally, for a free electron, the internal timescale is that of its vacuum fluctuations, $\tau_{\rm int} \sim \hbar/mc^2 \sim 10^{-21}$ s which is much shorter than any interaction time $\tau_{\rm ext}$ for which classical physics is appropriate, so the electron is semiclosed.

When a semiclosed classical system is idealized as closed, so its interactions with the external universe are ignored, then its evolution can be described using Hamiltonian dynamics.[1] The system's classical state is described by *generalized coordinates* $\mathbf{q} \equiv \{q_j\}$ and *generalized momenta* $\mathbf{p} \equiv \{p_j\}$, where the index $j$ runs from 1 to $W =$ (the number of degrees of freedom). The evolution of $\mathbf{q}, \mathbf{p}$ is governed by *Hamilton's equations*

$$\frac{dq_j}{dt} = \frac{\partial H}{\partial p_j} , \quad \frac{dp_j}{dt} = -\frac{\partial H}{\partial q_j} , \tag{3.1}$$

where $H(\mathbf{q}, \mathbf{p})$ is the *Hamiltonian* and each equation is really $W$ separate equations. Note that, because the system is idealized as closed, there is no explicit time dependence in the Hamiltonian. Of course, not all physical systems (e.g. not those with strong internal dissipation) are Hamiltonian, although in principle this restriction can usually be circumvented by increasing the number of degrees of freedom to include the cause of the dissipation. Let us return to our examples.

For an individual star inside a galaxy, there are three degrees of freedom ($W = 3$) which we might choose to be the motion along three mutually orthogonal directions with Cartesian basis vectors $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ so $q_1 = x_1$, $q_2 = x_2$, $q_3 = x_3$. Because the star's speed is small compared to light, its momenta $p_1, p_2, p_3$ are just the product of its mass and its velocity components along these axes, and so the Hamiltonian is

$$H(\mathbf{q}, \mathbf{p}) = \frac{1}{2m}({p_1}^2 + {p_2}^2 + {p_3}^2) + m\Phi(q_1, q_2, q_3) , \tag{3.2}$$

where $m$ is the stellar mass and $\Phi(q_1, q_2, q_3)$ is the gravitational potential (whose sign we take to be negative). Now make a *canonical transformation* [2] of these coordinates (in this case the same as a coordinate transformation and an associated change of basis; Exercise 3.1) to new $Q_1 = r$, $Q_2 = \theta$, $Q_3 = \phi$, where $(r, \theta, \phi)$ are the star's spherical polar coordinates with $r$ measured from the center of the galaxy. The corresponding, *canonically conjugate* momenta turn out to be $P_1 = p_r$, $P_2 = rp_\theta$, $P_3 = r \sin\theta p_\phi$ where $p_r$, $p_\theta$ and $p_\phi$ are the components of the star's momentum along unit vectors that point in the $r$, $\theta$, and $\phi$ directions. In terms of these new coordinates, the Hamiltonian Eq. (3.2) takes the form

$$H = \frac{1}{2m}\left({P_1}^2 + \frac{{P_2}^2}{r^2} + \frac{{P_3}^2}{r^2 \sin^2\theta}\right) + m\Phi(r, \theta, \phi) . \tag{3.3}$$

[1]See, for example, Goldstein (1980).
[2]Goldstein [3]bid

Now consider not just one star, but $K \sim 10^{10}$ of them in a galaxy. There are now $W = 3K$ degrees of freedom and the Hamiltonian is simply the sum of the Hamiltonians for each individual star as long as we ignore interactions between stars.

If our system is the fundamental mode of a sapphire crystal, then the number of degrees of freedom is only $W = 1$ and we can take the single generalized coordinate $q$ to be the displacement of one end of the crystal from equilibrium; there will be an "effective mass" $M$ for the mode (approximately equal to the actual mass of the crystal) such that the generalized momentum is $p = M dq/dt$. The Hamiltonian will be the standard one for a harmonic oscillator,

$$H(p,q) = \frac{p^2}{2M} + \frac{1}{2}M\omega^2 q^2 \ . \tag{3.4}$$

If we want to describe a whole crystal with $K \sim 10^{30}$ atoms, then we sum $W = 3K$ oscillator Hamiltonians for the crystal's $W$ normal modes and add an interaction potential $H_{\rm int}$ that accounts for the weak interactions between modes.

$$H = \sum_{j=1}^{W} \left\{ \frac{p_j^2}{2M_j} + \frac{1}{2} M_j {\omega_j}^2 (q_j)^2 \right\} + H_{\rm int}(q_1, \ldots, q_W, p_1, \ldots, p_W) \ . \tag{3.5}$$

Here $M_j$ is the effective mass of mode $j$ and $\omega_j$ is the mode's angular frequency.

### 3.2.2 Ensembles

In kinetic theory we study, statistically, a collection of a huge number of particles, which we can regard as a *system*. Similarly, in statistical mechanics, we study, statistically, a collection or *ensemble* of a huge number of systems. Of course this is only a theoretical device, like a thought experiment, which enables us to make statistical arguments. As we shall see, there are many different ways that we can imagine forming an ensemble and this freedom can be put to use to solve many different types of problems.

In some applications, we require that all the systems in the ensemble be *identical* in the sense that they all have identically the same number of degrees of freedom $W$, and be governed by Hamiltonians with identically the same functional forms $H(\mathbf{q}, \mathbf{p})$, and have identically the same volume $V$ and total internal energy $\tilde{E}$ — but the values of the generalized coordinates and momenta at a specific time $t$, ($\mathbf{q}(t)$ and $\mathbf{p}(t)$) need not be the same; i.e., the systems need not be in the same *state* at time $t$. If such a conceptual ensemble of identical systems (first studied by Boltzmann) evolves until it reaches "statistical equilibrium" (Sec. 3.4), it then is called *microcanonical*; see Table 3.1.

Sometimes we will deal with an ensemble of systems that can exchange energy (heat) with their surroundings so that the internal energy of each system can fluctuate. If we regard the surroundings (sometimes called *heat baths*) as having far greater heat capacity than the individual systems, and if statistical equilibrium has been reached, then we call this sort of ensemble (introduced by Gibbs) *canonical*.

At the next level of freedom, the systems can also expand, i.e. they can exchange volume as well as energy with their surroundings. This was also studied by Gibbs and in equilibrium is known as the *Gibbs* ensemble. The fourth ensemble in common use is Pauli's *grand*

| **Ensemble** | **Quantities Exchanged with Surroundings** |
|---|---|
| Microcanonical | nothing |
| Canonical | $\tilde{E}$ |
| Gibbs | $\tilde{E},\ V$ |
| Grand Canonical | $\tilde{E},\ N_I$ |

**Table 3.1:** Different types of ensemble used in this chapter. The first column is the name given to the ensemble if it is in statistical equilibrium (Sec. 3.4). The column on the right lists the physical quantities that can vary from system to system as a result of weak interaction with their surroundings.

*canonical* ensemble in which each system can exchange energy and particles (but not volume) with its surroundings.

### 3.2.3 Distribution Functions

In kinetic theory we described the statistical properties of a collection of particles by a distribution function $\mathcal{N}(t; \mathbf{x}, \mathbf{p})$ = (number of particles per unit volume of 7-dimensional phase space, near the event $\{t, \mathbf{x}\}$ and near the 4-momentum $\{\tilde{E}, \mathbf{p}\}$ on the mass hyperboloid). (Henceforth, in keeping with the usual Newtonian-motivated viewpoint of Hamiltonian dynamics, we shall separate off time and regard the phase space as 6-dimensional, even in relativistic situations.) Similarly, in statistical mechanics we shall describe the statistical properties of an ensemble of systems by a distribution function equal to the number of systems per unit volume in a $2W$-dimensional phase space that is analogous to kinetic theory's 6-dimensional one:

$$\mathcal{N}(t; W; q_1, q_2, \ldots, q_W; p_1, p_2, \ldots, p_W) \equiv \frac{dN}{d^W q d^W p} . \tag{3.6}$$

Here

$$d^W q \equiv dq_1 dq_2 \cdots dq_W , \quad d^W p \equiv dp_1 dp_2 \cdots dp_W \tag{3.7}$$

are the volumes in generalized coordinate space and in generalized momentum space respectively, and $dN$ is the total number of systems that, at time $t$, have $W$ degrees of freedom, have their generalized coordinates lying in the region $d^W q$ centered on the values $(q^1, q^2, \ldots, q^W)$ of Eq. (3.6), and have their generalized momenta lying in the region $d^W p$ centered on the $(p_1, p_2, \ldots, p_W)$ of Eq. (3.6).

Just as the event $\{\mathbf{x}, t\}$ and 4-momentum $\{\mathbf{p}, \tilde{E}\}$ in kinetic theory are geometric, frame-independent objects, similarly location in phase space in statistical mechanics is a geometric, coordinate-independent concept (though our notation does not emphasize it). The quantities $\{q_j, p_k\}$ appearing in Eq. (3.6) are the coordinates in phase space of that location. When one makes a canonical transformation from one set of generalized coordinates to another, the $\{q_j, p_k\}$ change but the geometric location in phase space does not. Moreover, just as the individual spatial and momentum volumes $d\mathcal{V}_x$ and $d\mathcal{V}_p$ in kinetic theory are frame-dependent but their product $d\mathcal{V}_x d\mathcal{V}_p$ is frame-independent [c.f. Eqs. (2.8)–(2.10)], so also in statistical mechanics $d^W q$ and $d^W p$ are dependent on the choice of canonical coordinates and change under a canonical transformation, but the product $d^W q d^W p$ is independent of the choice of canonical coordinates and is unchanged by a canonical transformation. Correspondingly, the statistical mechanical distribution function $\mathcal{N}(t; W; \mathbf{q}; \mathbf{p})$, like its kinetic-theory counterpart, is a geometric, coordinate-independent quantity; i.e., it is unchanged by a canonical transformation. See Exercise 3.1.

The development of classical thermodynamics was one of the crowning achievements of nineteenth century science. However, the theory was inevitably incomplete and had to remain so until the development of quantum theory. A major difficulty, that we have already confronted in the last chapter, was how to count the number of quantum states available to a system. As we saw, the number density of quantum mechanical states in the 6-dimensional, single-particle phase space of kinetic theory is (ignoring particle spin) $\mathcal{N}_{\rm states} = 1/h^3$, where $h$ is Planck's constant. Generalising to the $2W$-dimensional phase space of statistical mechanics, the number density of states is $1/h^W$, (one factor of $1/h$ for each of the canonical pairs $(q_1, p_1)$, $(q_2, p_2)$, $\cdots$, $(q_W, p_W)$.) Formally, this follows from the canonical quantization procedure of elementary quantum mechanics.

There was a second problem in nineteenth century classical thermodynamics, that of *distinguishability*. (If we swap two similar atoms in phase space do we have a new state or not? If we mix two containers of the same gas at the same temperature and pressure, does the entropy increase?) This problem was recognized classically, but had not been resolved in a completely satisfactory manner. When the laws of quantum mechanics were developed, it became clear that all identical particles are indistinguishable, so having particle 1 at location $\mathcal{A}$ in phase space and an identical particle 2 at location $\mathcal{B}$ must be counted as the same state as particle 1 at $\mathcal{B}$ and particle 2 at $\mathcal{A}$. Correspondingly, if we attribute half the quantum state to the classical phase space location $\{1 \text{ at } \mathcal{A}, 2 \text{ at } \mathcal{B}\}$ and the other half to $\{1 \text{ at } \mathcal{B}, 2 \text{ at } \mathcal{A}\}$, then the classical number density of states per unit volume of phase space must be reduced by a factor 2—and more generally by some *multiplicity factor* $\mathcal{M}$. In general, therefore, we can write the actual number density of states in phase space $dN_{\rm states}/d^W q d^W p \equiv \mathcal{N}_{\rm states}$ as

$$\mathcal{N}_{\rm states} = \frac{1}{\mathcal{M} h^W} \tag{3.8}$$

(ignoring spins). For $K$ identical particles, it is not hard to see that $\mathcal{M} = K!$

In addition to resolving the problems of counting of states and distinguishability, quantum mechanics adds two more features to statistical physics. **First**, we must distinguish between *bosons*, of which an unlimited number can occupy a given state and *fermions*, for which the occupation number is either 0 or 1. Classical statistical mechanics is only appropriate when the phase-space particle density is so low that most states are unoccupied, and double or higher occupancy, even if allowed, is rare and thus irrelevant. **Second**, even if we start with a system whose wave function $\psi$ is in a *pure* eigenstate (ordinary, everyday type of eigenstate), the system may evolve into a *mixed state* as a result of (i) interaction with the rest of the universe and (ii) our choice not to keep track of correlations between the universe and the system; Box 3.1 and Sec. 3.6 below. The system's initial, pure state can be described in geometric, basis-independent quantum language by a state vector ("ket") $|\psi\rangle$; but its final, mixed state requires a different kind of quantum description: a *density operator* $\hat{\rho}$.

In the classical limit, the quantum mechanical density operator $\hat{\rho}$ becomes the classical (i.e. ordinary, non-quantum) probability $\rho(t, W, \mathbf{q}, \mathbf{p})$ that the system at time $t$ will be found to have dimensionality $W$ and to be in a pure quantum state at the location $\{\mathbf{q}, \mathbf{p}\}$ in the classical phase space; see Box 3.1. (For example, if the system is an atom with zero spin, the pure quantum state at $\{\mathbf{q}, \mathbf{p}\}$ can be regarded as a plane wave state with momentum $\mathbf{p}$, inside a box of arbitrary size $L$ whose center is at $\mathbf{q}$; c.f. Sec. 2.3.) The combination $(W, \mathbf{q}, \mathbf{p})$, which we regard as identifying the pure quantum state, we shall label with the integer $n$, and correspondingly we sometimes shall denote $\rho$ by $\rho_n$. Clearly, the sum of the probability $\rho_n$ over all quantum states must be unity:

$$\sum_n \rho_n \equiv \sum_W \int \rho \frac{d^W q d^W p}{\mathcal{M} h^W} \equiv \sum_W \int \rho \mathcal{N}_{\text{states}} d\Gamma_W = 1, \tag{3.9}$$

where we have denoted the volume element in phase space $d^W q d^W p$ by $d\Gamma_W$:

$$d\Gamma_W \equiv d^W q d^W p \,. \tag{3.10}$$

For the microcanonical, canonical, and Gibbs ensembles, only one value of $W$ is allowed and we just write $d\Gamma$; but for the grand canonical ensemble with its variable number of particles $N_I$, $W$ is variable and we must use $d\Gamma_W$ and sum over $W$. *Note that Eq.* (3.9) *gives the classical meaning of the sum over quantum states* $\sum_n$. We shall use the symbol $\sum_n$ extensively.

If $N$ is the total number of systems in the ensemble, then the classical distribution function (i.e., the number of systems per unit volume of phase space) is related to the probability $\rho$ by

$$\mathcal{N} = N \mathcal{N}_{\text{states}} \rho = \frac{N\rho}{\mathcal{M} h^W} \,. \tag{3.11}$$

This relation tells us that $\rho$ is nothing but a renormalized variant of the distribution function $\mathcal{N}$; we therefore shall sometimes refer to it as the *probabilistic distribution function* or just the *distribution function*.

Because the ensemble is almost always a purely conceptual thing, and does not exist in reality, the total number of systems $N$ in the ensemble is rarely of any interest, and correspondingly, the classical distribution function $\mathcal{N}$, which depends on $N$, carries around excess, uninteresting baggage. For this reason, we shall not use $\mathcal{N}$ at all in this book; we shall focus attention solely on the $N$-independent distribution function $\rho$.

Each system in an ensemble is endowed with an energy that is equal to its Hamiltonian, $\tilde{E} = H(\mathbf{q}, \mathbf{p})$. Because different systems reside at different locations $(\mathbf{q}, \mathbf{p})$ in phase space, they typically will have different energies. A quantity of much interest is the *ensemble-averaged energy*, which is the average value of $\tilde{E}$ over all systems in the ensemble

$$\langle \tilde{E} \rangle = \sum_n \rho_n \tilde{E}_n = \sum_W \int \rho \tilde{E} \mathcal{N}_{\text{states}} d\Gamma_W. \tag{3.12}$$

For any other function $A(\mathbf{q}, \mathbf{p})$ defined on the phase space of a system, for example the linear momentum or the angular momentum, one can compute an ensemble average by the obvious analog of Eq. (3.12):

$$\langle A \rangle = \sum_n \rho_n A_n \,. \tag{3.13}$$

**Box 3.1 Density Operator and Quantum Statistical Mechanic**

Consider a single quantum mechanical system that is in a *pure* state $|\psi\rangle$. One can formulate the theory of such a pure state equally well in terms of $|\psi\rangle$ or the *density operator* $\hat{\varrho} \equiv |\psi\rangle\langle\psi|$; c.f. Feynman (1972). For example, the expectation value of some observable, described by a Hermitian operator $\hat{A}$, can be expressed equally well as $\langle A\rangle = \langle\psi|\hat{A}|\psi\rangle$ or as $\langle A\rangle = \mathrm{Trace}(\hat{\varrho}\hat{A})$. (In any basis $|\phi_i\rangle$, "Trace" is just the trace of the matrix product $\sum_j \varrho_{ij}A_{jk}$, where $\varrho_{ij} = \langle\phi_i|\hat{\varrho}|\phi_j\rangle$ is called the *density matrix* in that basis, and similarly for $A_{jk}$.) If our chosen system interacts with the external universe and we have no knowledge of the correlations that the interaction creates between the system and the universe, then the interaction drives the system into a mixed state, which is describable by a density operator $\hat{\varrho}$ but not by a ket vector $|\psi\rangle$. This $\hat{\varrho}$ can be regarded as a classical-type average of $|\psi\rangle\langle\psi|$ over an ensemble of systems, each of which has interacted with the external universe and then has been driven into a pure state $|\psi\rangle$ by a measurement of the universe (or, equivalently, by "a trace over the universe's degrees of freedom"). If the systems in the ensemble behave nearly classically, then it turns out that in the basis $|\phi_n\rangle$, whose states are labeled by the classical variables $n = \{W, \mathbf{q}, \mathbf{p}\}$, the density matrix $\varrho_{nm} \equiv \langle\phi_n|\hat{\varrho}|\phi_m\rangle$ is nearly diagonal. The classical probability $\rho_n$ of classical statistical mechanics (and of this book when dealing with classical or quantum systems) is then equal to the diagonal value of this density matrix: $\rho_n = \varrho_{nn}$. It can be demonstrated that the equation of motion for the density operator $\hat{\varrho}$, when the systems in the quantum mechanical ensemble are all evolving freely (no significant interactions with the external universe), is

$$\frac{\partial\hat{\varrho}}{\partial t} = -\frac{1}{i\hbar}[\hat{\varrho}, \hat{H}] \, . \tag{1}$$

This is the quantum statistical analog of Liouville's equation (Sec. 3.3 below), and the quantum mechanical commutator $[\hat{\varrho}, \hat{H}]$ appearing here is the quantum mechanical analog of the Poisson bracket $[\rho, H]_{\mathbf{q},\mathbf{p}}$, which appears in Liouville's equation. If the quantum systems are in eigenstates of their Hamiltonians, then $\hat{\varrho}$ commutes with $\hat{H}$ so the density matrix is constant in time and there will be no transitions. This is the quantum analog of the classical $\rho$ being constant in time and thus a constant of the motion; Sec. 3.3 below.

****************************

## EXERCISES

**Exercise 3.1** *Derivation and Example: Canonical Transformation*

Let $(q_j, p_k)$ be one set of generalized coordinates and momenta for a given system, and let $(Q_j, P_k)$ be another set. Then (except in degenerate cases, for which a different generating function must be used) there is a generating function $F(q_j, P_k)$, which depends on the "old" coordinates $q_j$ and "new" momenta $P_k$, such that

$$p_j = \frac{\partial F}{\partial q_j} \;, \quad Q_j = \frac{\partial F}{\partial P_j} \;. \tag{3.14}$$

(a) As an example, what are the new coordinates and momenta in terms of the old that result from

$$F = \sum_{i=1}^{W} f_i(q_j) P_i \;, \tag{3.15}$$

where $f_i$ are arbitrary functions of the old coordinates?

(b) The canonical transformation generated by Eq. (3.14) for arbitrary $F(q_j, P_k)$ leaves unchanged the value, but not the functional form, of the Hamiltonian at each point in phase space, i.e., $H$ is a geometric, coordinate-independent function (scalar field) of location in phase space. Show, for the special case of a system with one degree of freedom (one $q$, one $p$, one $Q$, and one $P$) that, if Hamilton's equations (3.1) are satisfied in the old variables $(q, p)$, then they will be satisfied in the new variables $(Q, P)$.

(c) Show, for a system with one degree of freedom, that although $dq \neq dQ$ and $dp \neq dP$, the volume in phase space is unaffected by the canonical transformation: $dpdq = dPdQ$.

(d) Hence show that for any closed path in phase space, $\oint pdq = \oint PdQ$.

(e) As a higher-dimensional example ($W = 3$), consider a star moving in a galaxy [Eqs. (3.2), (3.3) and associated discussion]. Show that $d^3q = dx_1 dx_2 dx_3 \neq d^3Q = dr d\theta d\phi$, and $d^3p \neq d^3P$, but $d^3q d^3p = d^3Q d^3P$.

****************************

## 3.3 Liouville's Theorem and the Evolution of the Distribution Function

In kinetic theory the distribution function $\mathcal{N}$ was not only a frame-independent entity; it was also a constant along the trajectory of any freely moving particle, so long as collisions between particles were negligible. Similarly, in statistical mechanics the probability $\rho$ is not only coordinate-independent (unaffected by canonical transformations); *$\rho$ is also a constant along the phase-space trajectory of any freely evolving system*, so long as the systems in the ensemble are not interacting significantly with the external universe, i.e. so long as they can be idealized as closed. This is the statistical mechanical version of *Liouville's theorem*, and its proof is a simple exercise in Hamiltonian mechanics, analogous to the "sophisticated" proof of the Vlasov equation in Box 2.1:

Since the ensemble's systems are closed, no system changes its dimensionality $W$ during its evolution. This permits us to fix $W$ in the proof. Since no systems are created or destroyed during the evolution, the probability (which is proportional to the distribution function $\mathcal{N}$), must satisfy the conservation law

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial q_j}\left(\rho \frac{dq_j}{dt}\right) + \frac{\partial}{\partial p_j}\left(\rho \frac{dp_j}{dt}\right) = 0, \tag{3.16}$$

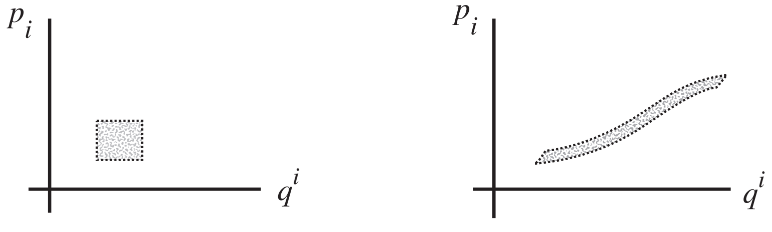


**Fig. 3.1:** Liouville's theorem: (a) The region in the $q_i - p_i$ part of phase space (with $i$ fixed) occupied by a set $\mathcal{G}$ of identical, closed systems at time $t = 0$. (b) The region occupied by the same set of systems a short time later, $t > 0$. The Hamiltonian-generated evolution of the individual systems has moved them in such a manner as to skew the region they occupy, but the volume $\Delta q_i \Delta p_i$ is unchanged.

where the implicit sum is now from 1 to $W$ rather than 1 to 3 as for the Vlasov equation. (This conservation law is the analog of the one for particle conservation, $\partial n/\partial t + \partial S_j/\partial x_j = 0$. In the particle conservation law $n$ is the number of particles per unit volume of physical space; in Eq. (3.16), $\rho$ is, aside from a multiplicative constant, the number of systems per unit volume of phase space. In the particle conservation law, $S_j$ is the flux of particles in the $j$ direction of physical space; in Eq. (3.16), $\rho dq_j/dt$ is, aside from a constant, the flux of systems in the $j$ direction of physical space and $\rho dp_j/dt$, the flux of systems in the $j$ direction of momentum space.) We can use Hamilton's Eq. (3.1) to rewrite the conservation law for systems (3.16) in the form

$$\frac{d\rho}{dt} \equiv \frac{\partial \rho}{\partial t} + \left(\frac{\partial \rho}{\partial q_j}\right)\left(\frac{\partial H}{\partial p_j}\right) - \left(\frac{\partial \rho}{\partial p_j}\right)\left(\frac{\partial H}{\partial q_j}\right) = \frac{\partial \rho}{\partial t} + [\rho, H]_{\mathbf{q},\mathbf{p}} = 0, \tag{3.17}$$

where $[\rho, H]_{\mathbf{q},\mathbf{p}}$ is the *Poisson bracket.*[4] As with the Vlasov equation (2.95), $d\rho/dt$ as defined in Eq. (3.17) is just the rate of change of the probability following a system in its motion through phase space. Therefore, the probability $\rho$ is constant along the system's phase space trajectory, as was to be proved.

We shall call Eq. (3.17), which embodies this Liouville theorem, the statistical mechanical *Liouville equation* or *Vlasov equation* or *collisionless Boltzmann equation.*

As a simple, qualitative example, consider a system consisting of hot gas expanding adiabatically so that its large random kinetic energy is converted into ordered radial motion. If we examine a set $\mathcal{G}$ of such systems very close to each other in phase space, then it is apparent that, as the expansion proceeds, the size of $\mathcal{G}$'s physical-space volume $d^W q$ increases as the size of its momentum-space volume $d^W p$ diminishes, so that the product $d^W q d^W p$ remains constant (Fig. 3.1), and correspondingly $\rho \propto \mathcal{N} = dN/d^W q d^W p$ is constant.

What happens if the systems being studied interact weakly with their surroundings? We must then include an interaction term on the right hand side of Eq. (3.17), thereby converting

[4]Goldstein [5]bid.

it into the statistical mechanical version of the Boltzmann transport equation:

$$\left(\frac{d\rho}{dt}\right)_{\text{moving with a fiducial system}} = \left(\frac{d\rho}{dt}\right)_{\text{interactions}} . \tag{3.18}$$

Here the time derivative on the left is taken moving with a "fiducial" system that does *not* interact with the external universe.

## 3.4 Statistical Equilibrium

Consider an ensemble of identical systems, all of which have the same huge number of degrees of freedom (dimensionality $W \gg 1$). Put all the systems initially in identically the same state, and then let them exchange heat (but not particles or volume or anything else) with an external *thermal bath* that has a huge heat capacity and is in thermodynamic equilibrium at some temperature $T$. (For example, the systems might be impermeable cubes of gas 1000 kilometers on a side near the center of the sun, and the thermal bath might be all the surrounding gas near the sun's center; or the systems might be identical sapphire crystals inside a huge cryostat, and the thermal bath might be the cryostat's huge store of liquid Helium.) After a sufficiently long time, $t \gg \tau_{\text{ext}}$, the ensemble will settle down into equilibrium with the bath, i.e. it will become the *canonical ensemble* mentioned in Table **??** above. In this final, canonical equilibrium state, the probability $\rho(t, \mathbf{q}, \mathbf{p})$ will be independent of time $t$, and it no longer will be affected by interactions with the external environment; i.e., the interaction terms in the evolution equation (3.18) will have ceased to have any net effect: on average, for each interaction event that feeds energy into a system there will be an interaction event that takes away an equal amount of energy. The distribution function, therefore, will satisfy the interaction-free, collisionless Boltzmann equation (3.17) with the time derivative $\partial\rho/\partial t$ removed:

$$[\rho, H]_{\mathbf{q},\mathbf{p}} \equiv \frac{\partial \rho}{\partial q_j}\frac{\partial H}{\partial p_j} - \frac{\partial \rho}{\partial p_j}\frac{\partial H}{\partial q_j} = 0 \tag{3.19}$$

We shall use the phrase *statistical equilibrium* to refer to any ensemble whose distribution function has attained such a state and thus satisfies Eq. (3.19).

Equation (3.19) is a well-known equation in Hamiltonian mechanics; it says that $\rho$ is a function solely of constants of the individual systems' Hamiltonian-induced motions; i.e., $\rho$ can depend on location $(\mathbf{q}, \mathbf{p})$ in phase space only through those constants of the motion. Among the constants of motion, in typical situations (for typical Hamiltonians), are the system's energy $\tilde{E}$, its linear momentum $\mathbf{P}$, its angular momentum $\mathbf{J}$, its number $N_I$ of conserved particles of various types $I$ (e.g. electrons, protons, ...), and its volume $V$. Note that these constants of motion are all additive: if we double the size of a system, they each double. We shall call such additive constants of the Hamiltonian-induced motion *extensive variables* (a term borrowed from thermodynamics) and we shall denote them by an enumerated list $K_1, K_2, \ldots$.

Now, the systems that we are studying have exchanged energy $\tilde{E}$ with their environment (the thermal bath) and thereby have acquired some range of $\tilde{E}$'s; therefore, $\rho$ can depend on $\tilde{E}$. However, the systems have not exchanged anything else with their environment, and they all thus have retained their original (identical) values of the other extensive variables $K_A$; therefore, $\rho$ must be a delta function in the others. We shall write

$$\rho = \rho(\tilde{E}) \ , \tag{3.20}$$

and shall not write down the delta functions explicitly.

As an aid in discovering the form of the function $\rho(\tilde{E})$, let us decompose each system in the ensemble into a huge number of subsystems. For example, each system might be a cube 1000 kilometers on a side inside the sun and its subsystems might be the $10^9$ 1-meter cubes into which one can divide it, or the systems might be identical sapphire crystals each containing $10^{27}$ atoms, and the subsystems might be the crystals' $3 \times 10^{27}$ normal modes of vibration. We shall label the subsystems of each system by an integer $a$ in such a way that subsystem $a$ in one system has the same Hamiltonian as subsystem $a$ in any other system. (For the sapphire crystals, $a = 1$ could be the fundamental mode of vibration, $a = 2$ the first harmonic, $a = 3$ the second harmonic, etc.) The subsystems with fixed $a$ make up a *subensemble* because of their relationship to the original ensemble.

Because the full ensemble is in statistical equilibrium, the subensembles will also be in equilibrium; and therefore their probabilities must be functions of those extensive variables $\tilde{E}, K_A$ that they can exchange with each other:

$$\rho_a = \rho_a(\tilde{E}_a, K_{1\,a}, K_{2\,a}, ...) \ . \tag{3.21}$$

(Although each system can exchange only energy $\tilde{E}$ with its heat bath, the subsystems may be able to exchange other quantities with each other; for example, if subsystem $a$ is a 1-meter cube inside the sun, then it can exchange energy $\tilde{E}_a$ and particles of all species $I$, so $K_{I\,a} = N_{I\,a}$.)

Since there is such a huge number of subsystems in each system, it is reasonable to expect that in statistical equilibrium there will be no significant correlations at any given time between the actual state of subsystem $a$ and the state of any other subsystem. In other words, the probability $\rho_a(W_a, \mathbf{q}_a, \mathbf{p}_a)$ that subsystem $a$ is in a quantum state with $W_a$ degrees of freedom and with its generalized coordinates and momenta near the values $(\mathbf{q}_a, \mathbf{p}_a)$ is independent of the state of any other subsystem. This lack of correlations, which has the mathematical statement

$$\rho(\tilde{E}) = \prod_a \rho_a \ , \tag{3.22}$$

is called *statistical independence*. (Statistical independence is actually a consequence of a "2-lengthscale approximation" [Box 2.2]. The size of each subsystem is far smaller than that of the full system, and precise statistical independence arises in the limit as the ratio of these sizes goes to zero.)

Statistical independence places a severe constraint on the functional forms of $\rho$ and $\rho_a$, as the following argument shows. By taking the logarithm of Eq. (3.22), we obtain

$$\ln \rho(\tilde{E}) = \sum_a \ln \rho_a(\tilde{E}_a, K_{1\,a}, ...) \ . \tag{3.23}$$

We also know, since energy is a linearly additive quantity, that

$$\tilde{E} = \sum_a \tilde{E}_a. \tag{3.24}$$

Now, we have not stipulated the way in which the systems are decomposed into subsystems. For our solar example, the subsystems might have been 2-m cubes or 7-m cubes rather than 1-m cubes. Exploiting this freedom, one can deduce that Eqs. (3.23) and (3.24) can be satisfied simultaneously if and only if $\ln\rho$ and $\ln\rho_a$ depend linearly on the energies $\tilde{E}$ and $\tilde{E}_a$, with the same proportionality constant $\beta$:

$$\begin{aligned} \ln\rho_a &= \beta\tilde{E}_a + (\text{some function of } K_{1\,a}, K_{2\,a}, ...) && (3.25)\\ \ln\rho &= \beta\tilde{E} + \text{constant}. && (3.26)\\ & && (3.27)\end{aligned}$$

Since the parameter $\beta$ is the same for all the systems and their subsystems, it must be some characteristic of the thermal bath with which the systems and subsystems equilibrated. By exponentiating the second of Eqs. (3.27) and noting that it has the same functional form as the Boltzmann distribution (2.32) of kinetic theory, we infer that $\beta$ must be $-1/kT$, where $T$ is the temperature of the thermal bath.

To summarize, an ensemble of identical systems with many degrees of freedom $W \gg 1$ which have reached statistical equilibrium by exchanging energy but nothing else with a huge thermal bath has the following *canonical* distribution function:

$$\rho_{\text{canonical}} = C\exp(-\tilde{E}/kT)\ . \tag{3.28}$$

Here $\tilde{E}(\mathbf{q},\mathbf{p})$ is the energy of a system at location $\{\mathbf{q},\mathbf{p}\}$ in phase space, $k$ is Boltzmann's constant, $T$ is the temperature of the heat bath, and $C$ is whatever normalization constant is required to guarantee that $\sum_n \rho_n = 1$. Actually, we have proved more than this. Not only must the ensemble of huge systems ($W \gg 1$) have the energy dependence $\rho \propto \exp(-\tilde{E}/kT)$, so must each subensemble of smaller systems, $\rho_a \propto \exp(-\tilde{E}_a/kT)$, even if (for example) the subensemble's identical subsystems have only one degree of freedom $W_a = 1$. Thus, if the subsystems exchanged only heat with their parent systems, then they must have the same canonical distribution (3.28) as the parents. This shows that the canonical distribution is the equilibrium state independently of the number of degrees of freedom $W$.

We can easily generalize this canonical distribution to an ensemble of systems that exchange other additive conserved quantitities (extensive variables) $K_1, K_2, ...$, in addition to energy $\tilde{E}$, with a huge, thermalized bath. By an obvious generalization of the above argument, the resulting statistical equilibrium distribution function must have the form

$$\rho = C\exp\left(\beta\tilde{E} + \sum_A \beta_A K_A\right)\ . \tag{3.29}$$

When the extensive variables $K_A$ that are exchanged with the bath, and that thus appear explicitly in the distribution function $\rho$, are energy $\tilde{E}$, and/or momentum $\mathbf{P}$, and/or angular momentum $\mathbf{J}$, and/or the number $N_I$ of the species $I$ of conserved particles, and/or volume $V$, it is conventional to rename the multiplicative factors $\beta$ and $\beta_A$ so that $\rho$ takes on the following form

$$\rho = C \exp\left[\frac{-\tilde{E} + \mathbf{U}\cdot\mathbf{P} + \mathbf{\Omega}\cdot\mathbf{J} + \sum_I \tilde{\mu}_I N_I - PV}{kT}\right] . \tag{3.30}$$

Here $k$ is Boltzmann's constant; $T$, $\mathbf{U}$, $\mathbf{\Omega}$, $\tilde{\mu}_I$, and $P$ are constants that are the same for all systems and subsystems, i.e., that characterize the full ensemble and all its subensembles and that therefore must have been acquired from the bath; and any extensive variable that is not exchanged with the bath must be omitted from the exponential and be replaced by an implicit delta function.

As we have seen, $T$ is the *temperature* that the ensemble and subensembles acquired from the bath; i.e., it is the bath temperature. From the Lorentz transformation law for energy and momentum $[\tilde{E}' = \gamma(\tilde{E} - \mathbf{U}\cdot\mathbf{P})]$ we see that, if we were to transform to a reference frame that moves with velocity $\mathbf{U}$ with respect to our original frame, then the $\exp(\mathbf{U}\cdot\mathbf{P}/kT)$ term in $\rho$ would disappear, and the distribution function would be isotropic in $\mathbf{P}$. This tells us that $\mathbf{U}$ is the *velocity* of the bath with respect to our chosen reference frame. By a similar argument, $\mathbf{\Omega}$ is the bath's *angular velocity* with respect to an inertial frame. By comparison with Eq. (2.32) we see that $\tilde{\mu}_I$ is the *chemical potential* of the conserved species $I$. Finally, experience with elementary thermodynamics suggests (and it turns out to be true) that $P$ is the bath's *pressure*. The bath-induced quantitites $T$, $\mathbf{U}$, $\mathbf{\Omega}$, $\tilde{\mu}_I$, and $P$ are called *intensive* variables. Note that, by contrast with the corresponding extensive variables $\tilde{E}$, $\mathbf{P}$, $\mathbf{J}$, $N_I$, and $V$, these intensive variables do not double when the size of a system is doubled, i.e. they are not additive; rather, they are properties of the ensemble as a whole and thus are independent of the systems' sizes.

Henceforth (except in Sec. 3.8, when discussing black-hole atmospheres), we shall restrict our baths always to be at rest in our chosen reference frame and to be non-rotating with respect to inertial frames, so that $\mathbf{U} = \mathbf{\Omega} = 0$. The distribution function $\rho$ can then either be a delta function in the system momentum $\mathbf{P}$ and angular momentum $\mathbf{J}$ (if momentum and angular momentum are not exchanged with the bath), or it can involve no explicit dependence whatsoever on $\mathbf{P}$ and $\mathbf{J}$ (if momentum and angular momentum *are* exchanged with the bath; c.f. Eq. (3.30) with $\mathbf{U} = \mathbf{\Omega} = 0$). In either case, if energy is the only other quantity exchanged with the bath, then the distribution function is said to be canonical and has the form (3.28). If, in addition to energy, volume can also be exchanged with the bath (e.g., if the systems are floppy bags of gas whose volumes can change and through which heat can flow), then the equilibrium is the *Gibbs ensemble*, which has the distribution function

$$\rho_{\rm Gibbs} = C \exp[(-\tilde{E} - PV)/kT] \tag{3.31}$$

(and an implicit delta function in $N_I$ and possibly in $\mathbf{J}$ and $\mathbf{P}$). If the exchanged quantities are energy and particles but not volume (e.g., if the systems are 1-m cubes inside the sun with totally imaginary walls through which particles and heat can flow), then the equilibrium is the *grand canonical ensemble*, with

$$\rho_{\rm grand\ canonical} = C \exp\left[\left(-\tilde{E} + \sum_I \tilde{\mu}_I N_I\right)/kT\right] \tag{3.32}$$

(and an implicit delta function in $V$ and perhaps in $\mathbf{J}$ and $\mathbf{P}$).

We mention, as a preview of an issue to be addressed in Chap. 4, that an individual system, picked randomly from the ensemble and then viewed as a bath for its own tiny subsystems, will *not* have *identically* the same temperature $T$, and/or chemical potential $\tilde{\mu}_I$, and/or pressure $P$ as the huge bath with which the ensemble has equilibrated; rather, the individual system's $T$, $\tilde{\mu}_I$, and/or $P$ can fluctuate a tiny bit about the huge bath's values (about the values that appear in the above probabilities), just as its $\tilde{E}$, $N_I$, and/or $V$ fluctuate. We shall study these fluctuations in Chap. 4.

Most of our applications of statistical mechanics (later in this chapter and in Chap. 4) will deal with systems which have such low temperatures ($kT \ll m_I c^2$ for all relevant species $I$) that almost all their particles move with speeds small compared to light. In this nonrelativistic regime, it is convenenient to remove the rest-mass contributions from the total energy $\tilde{E}$ of a system and from the chemical potentials $\tilde{\mu}_I$ of each species; i.e., it is convenient to introduce the nonrelativistic energy and chemical potentials

$$E \equiv \tilde{E} - \sum_I N_I m_I \; , \quad \mu_I \equiv \tilde{\mu}_I - m_I \; . \tag{3.33}$$

By a redefinition of the normalization constants $C$, the canonical, Gibbs, and grand canonical distributions then take on identically the same forms as above [Eqs. (3.28), (3.31), and (3.32)] but with the tildes removed:

$$\rho_{\text{canonical}} = C \exp(-E/kT) \; , \tag{3.34}$$

$$\rho_{\text{Gibbs}} = C \exp[(-E - PV)/kT] \; . \tag{3.35}$$

$$\rho_{\text{grand canonical}} = C \exp\left[\left(-E + \sum_I \mu_I N_I\right)/kT\right] \tag{3.36}$$

The concepts and results that we are developing in this chapter have enormous generality. They are valid (when handled with sufficient care) for quantum systems as well as classical, and they are valid for semiclosed or closed systems of any type whatsoever. The systems need not resemble the examples we have met in the text. They can be radically different, but so long as they are closed or semiclosed, our concepts and results will apply. Exercise 3.2 is a very important example in which each system is a single-particle quantum state of a field, rather than being a collection of particles or normal modes of a crystal; and the result of that exercise is of great importance: it is the Fermi-Dirac or Bose-Einstein distribution for the mean number of quanta in the single-particle quantum state.

****************************

**EXERCISES**

**Exercise 3.2** *Derivation and Example: Derivation of the Bose-Einstein and Fermi-Dirac Distributions*

The goal of this exercise is to show that the Fermi-Dirac and Bose-Einstein distributions

$$\bar{\eta} = \frac{1}{e^{(\tilde{E}-\tilde{\mu})/kT}+1} , \quad \bar{\eta} = \frac{1}{e^{(\tilde{E}-\tilde{\mu})/kT}-1} , \tag{3.37}$$

which we used in our kinetic-theory studies in the last chapter [Eqs. (2.28), (2.29)], are both consequences of the grand canonical ensemble. Our proof will illustrate the closeness of classical statistical mechanics and quantum statistical mechanics: the proof is fundamentally quantum mechanical because the regime $\bar{\eta} \sim 1$ is quantum mechanical (it violates the classical condition $\bar{\eta} \ll 1$); nevertheless, the proof makes use of precisely the same concepts and techniques as we have developed for our classical studies.

Consider an ensemble of complex systems in statistical equilibrium. Each system can be regarded as made up of a large number of fermions (electrons, protons, neutrons, ...) and/or bosons (photons, gravitons, alpha particles, phonons, ...). We shall analyze each system by identifying a complete set of single-particle quantum states (which we shall call *modes*) into which the particles can be inserted. (For photons, these "modes" are the normal modes of the classical electromagnetic field; for phonons in a crystal, they are the normal modes of the crystal's vibrations; for nonrelativistic electrons or protons or alpha particles, they are energy eigenstates of the nonrelativistic Schroedinger equation; for relativistic electrons, they are energy eigenstates of the Dirac equation.) A complete enumeration of modes is the starting point for the *second quantization* formulation of quantum field theory, and also the starting point for this exercise.

Choose one specific mode $\mathcal{S}$ [e.g., a nonrelativistic electron plane-wave mode in a box of side $L$ with spin up and $\mathbf{p} = (5, 3, 17)h/L$)]. There is one such mode $\mathcal{S}$ in each of the systems in our ensemble, and these modes (all identical in their properties) form a subensemble of our original ensemble. Because each of these modes $\mathcal{S}$ can exchange energy and particles with all the other modes in its system, the subensemble is grand canonically distributed.

The quantum states allowed to the mode $\mathcal{S}$ are states in which $\mathcal{S}$ contains a finite number of particles (quanta), $n$. Denote by $\tilde{E}_{\mathcal{S}}$ the energy of one particle residing in the mode $\mathcal{S}$. Then the mode's total energy when it is in the state $|n\rangle$ (when it contains $n$ quanta) is $\tilde{E}_n = n\tilde{E}_{\mathcal{S}}$. [Note: for a freely traveling, relativistic electron mode, $\tilde{E}_{\mathcal{S}} = (m^2 + \mathbf{p}^2)^{\frac{1}{2}}$ where $\mathbf{p}$ is the mode's momentum $p_x = jh/L$ etc.; for a phonon mode with eigenfrequency of vibration $\omega$, $\tilde{E}_{\mathcal{S}} = \hbar\omega$; etc.] Since the distribution of the ensemble's modes among the allowed quantum states is grand canonical, the probability $\rho_n$ of being in state $|n\rangle$ is

$$\rho_n = \text{const} \times \exp\left(\frac{\tilde{\mu}n - \tilde{E}_n}{kT}\right) , \tag{3.38}$$

where $\tilde{\mu}$ and $T$ are the chemical potential and temperature of the bath of other modes with which the mode $\mathcal{S}$ interacts.

(a) Suppose that $\mathcal{S}$ is a fermion mode (i.e., its particles have half-integral spin). Then the Pauli exclusion principle dictates that $\mathcal{S}$ cannot contain more than one particle; i.e., $n$ can take on only the values 0 and 1. Show that the normalization of $\rho_n$ ($\sum_n \rho_n = 1$) implies that

$$\rho_0 = \frac{1}{1+\exp[(\tilde{\mu}-\tilde{E}_{\mathcal{S}})/kT]} , \quad \rho_1 = \frac{\exp[(\tilde{\mu}-\tilde{E}_{\mathcal{S}})/kT]}{1+\exp[(\tilde{\mu}-\tilde{E}_{\mathcal{S}})/kT]} . \tag{3.39}$$

This is the grand canonical distribution for a fermion mode. Show, further, that this grand canonical distribution has a mean occupation number

$$\bar{\eta} \equiv \langle n \rangle = \frac{1}{\exp[(\tilde{E}_{\mathcal{S}} - \tilde{\mu})/kT] + 1} , \tag{3.40}$$

in accord with Eq. (2.28).

(b) Suppose that $\mathcal{S}$ is a boson mode (i.e., its particles have integral spin). Then $\mathcal{S}$ can contain any nonnegative number of quanta; i.e., $n$ can assume the values 0, 1, 2, 3, ... . Show that normalization of $\rho_n$ implies

$$\rho_n = \left[1 - \exp\left(\frac{\tilde{\mu} - \tilde{E}_{\mathcal{S}}}{kT}\right)\right] \exp\left(\frac{n(\tilde{\mu} - \tilde{E}_{\mathcal{S}})}{kT}\right) . \tag{3.41}$$

This is the grand canonical distribution for bosons. Show that the mean number of bosons in $\mathcal{S}$ is

$$\bar{\eta} \equiv \langle n \rangle = \frac{1}{\exp[(\tilde{E}_{\mathcal{S}} - \tilde{\mu})/kT] - 1} \tag{3.42}$$

in accord with Eq. (2.29).

****************************

## 3.5 The Microcanonical Ensemble and the Ergodic Hypothesis

Turn attention, now, from ensembles of systems that interact with an external, thermal bath, to an ensemble of identical, precisely closed systems, i.e. systems that have no interactions whatsoever with the external universe. By "identical" we mean that every system in the ensemble has (i) precisely the same set of degrees of freedom, and thus (ii) precisely the same number of degrees of freedom $W$, (iii) precisely the same Hamiltonian, and (iv) precisely the same values $\tilde{E}, K_a$ for all the additive constants of motion (except perhaps momentum $\mathbf{P}$ and angular momentum $\mathbf{J}$).

Suppose that these systems begin with values of $(\mathbf{q}, \mathbf{p})$ that are not all precisely the same; i.e. that are spread out in some (arbitrary) manner over a *hypersurface* in phase space that has $H(\mathbf{q}, \mathbf{p})$ equal to the common value of energy $\tilde{E}$. Of course, we cannot choose systems whose energy is *precisely* equal to $\tilde{E}$. For most $\tilde{E}$ this would be a set of measure zero. Instead we let the systems occupy a tiny range of energy between $\tilde{E}$ and $\tilde{E} + \delta\tilde{E}$ and then discover (in Exercise 3.8) that our results are highly insensitive to $\delta\tilde{E}$ as long as it is extremely small compared with $\tilde{E}$.

It seems reasonable to expect that this ensemble, after evolving for a time much longer than its longest internal dynamical time scale $t \gg \tau_{\rm int}$, will achieve statistical equilibrium, i.e. will evolve into a state with

$$\frac{\partial \rho}{\partial t} = 0 . \tag{3.43}$$

(In the next section we will justify this expectation.) The distribution function $\rho$ will then satisfy the collisionless Boltzmann equation (3.19) and therefore will be a function only of the Hamiltonian's additive constants of the motion $\tilde{E}, K_A$. However, we already know that $\rho$ is a delta function in $K_A$ and a delta function with a tiny but finite spread in $\tilde{E}$; and the fact that it cannot depend on any other phase space quantities then implies that *$\rho$ is a constant over the hypersurface in phase space that has the prescribed values of $K_A$ and $\tilde{E}$, and is zero everywhere else in phase space.* This equilibrium ensemble is called *microcanonical.*

There is a subtle aspect of this microcanonical ensemble that deserves discussion. Suppose that we split each system in the ensemble up into a huge number of subsystems that can exchange energy (but for concreteness nothing else) with each other. We thereby obtain a huge number of subensembles, in the manner of Sec. 3.4. The original systems can be regarded as a thermal bath for the subsystems, and correspondingly it seems reasonable to expect (and indeed is true) that the subensembles will have canonical distribution functions, $\rho_a = Ce^{-\tilde{E}_a/kT}$. One might also expect the subensembles to be statistically independent, so that $\rho = \prod_a \rho_a$. However, such independence is not possible, since together with additivity of energy $\tilde{E} = \sum_a \tilde{E}_a$, it would imply that $\rho = Ce^{-\tilde{E}/kT}$, i.e. that the full ensemble is canonically distributed rather than microcanonical. What is wrong here?

The answer is that there in fact is a tiny correlation between the subensembles: If, at some moment of time, subsystem $a = 1$ happens to have an unusually large energy, then the other subsystems must correspondingly have a little less energy than usual; and this very slightly invalidates the statistical-independence relation $\rho = \prod_a \rho_a$, thereby enabling the full ensemble to be microcanonical even though all its subensembles are canonical. In the language of two-lengthscale expansions, where one expands in the dimensionless ratio (size of subsystems)/(size of full system) [Box 2.2], this correlation is a higher-order correction to statistical independence.

We are now in a position to understand more deeply the nature of the thermalized bath that we have invoked to drive ensembles into statistical equilibrium. That bath can be any huge system which contains the systems we are studying as subsystems; and the bath's *thermal equilibrium* can be either a microcanonical statistical equilibrium, or a statistical equilibrium involving exponentials of its extensive variables.

### 3.5.1 The Ergodic Hypothesis

The ensembles that we have been studying are almost always just conceptual ones that do not exist in the real universe. We have introduced them and paid so much attention to them not for their own sake, but because they can be powerful tools for studying the properties of a single, individual system that really does exist in the universe or in our laboratory.

This power comes about because a sequence of "snapshots" of the single system, taken at times separated by sufficiently large intervals $\Delta t$, has a probability distribution $\rho$ for the snapshots' instantaneous locations $\{\mathbf{q}, \mathbf{p}\}$ in phase space that is the same as the distribution function $\rho$ of some conceptual ensemble. If the single system is closed, so its evolution is driven solely by its own Hamiltonian, then the time between snapshots should be $\Delta t \gg \tau_{\rm int}$ and its snapshots will be microcanonically distributed. If the single system exchanges energy, and only energy, with a thermal bath, then the time between snapshots should be $\Delta t \gg \tau_{\rm ext}$

and its snapshots will be canonically distributed; and similarly for the other types of bath interactions. This property of snapshots is equivalent to the statement that for the individual system, *the long-term time average* $\bar{A}$ *of any function of the system's location in phase space* $A(\mathbf{q}(t), \mathbf{p}(t))$ *is equal to the ensemble average* $\langle A \rangle = \sum_n A_n \rho_n$ *of* $A(\mathbf{q}, \mathbf{p}) = A_n$.

This property of the snapshots comes about because of *ergodicity*: the individual system, as it evolves, visits all accessible quantum states for a fraction of the time that is equal to the ensemble's probability $\rho_n$.

At first sight, ergodicity may seem obvious. However, it is not a universal property of all systems: one can easily devise idealized examples of non-ergodic behavior (e.g., an elastic billiard ball bouncing around a square billiard table). On the other hand, generic systems, whose properties and parameters are not carefully fine tuned, do typically behave ergodically—but to prove so is one of the most difficult problems in statistical mechanics. We shall assume throughout this book's discussion of statistical physics that all the systems we study are indeed ergodic, and we shall call this the *ergodic hypothesis.*

## 3.6 Entropy and the Evolution into Statistical Equilibrium

For *any* ensemble of systems, whether it is in statistical equilibrium or not, and also whether it is quantum mechanical or not, the ensemble's *entropy* $S$ is defined, in words, by the following awful sentence: $S$ is the mean value (ensemble average) of the logarithm of the probability that a random system in the ensemble occupies a given quantum state, averaged over states and multiplied by $-k$. More specifically, denoting the probability that a system is in state $n$ by $\rho_n$, the ensemble's entropy $S$ is the following sum over quantum states (or equivalently integral over phase space):

$$S = -k \sum_n \rho_n \ln \rho_n \ . \tag{3.44}$$

Entropy is a precise measure of our lack of information about the state of any system chosen at random from an ensemble. In this sense, *the entropy can be regarded as a property of a random individual system in the ensemble*, as well as of the ensemble itself. If one increases the number of systems in the ensemble, one does not increase the entropy.

If all the systems are in the same quantum state, e.g. in the state $n = 17$, then $\rho_n = \delta_{n,17}$ so we know precisely the state of any system pulled at random from the ensemble, and Eq. (3.44) dictates that the entropy vanish. Vanishing entropy thus corresponds to a perfect knowledge of the system's quantum state; it corresponds to the quantum state being pure.

By contrast, consider a system in microcanonical statistical equilibrium. In this case, all states are equally likely ($\rho$ is constant), so if there are $N_{\rm states}$ states all together, then $\rho_n = 1/N_{\rm states}$ and the entropy (3.44) takes the form[6]

$$S = k \ln N_{\rm states} \ . \tag{3.45}$$

[6]This formula, with slightly different notation, can be found on Boltzmann's tomb.

The entropy, so defined, has some important properties. One property is that whenever the ensemble can be broken up into statistically independent subensembles of subsystems (as is generally the case for big systems in statistical equilibrium), so $\rho = \prod_a \rho_a$, then the entropy is additive, $S = \sum_a S_a$; see Exercise 3.4. This permits us to regard the entropy, like the systems' additive constants of the motion, as an extensive variable.

A second, very important property is the fact that, as an ensemble of systems evolves with time, its entropy cannot decrease and it generally tends to increase. This is the statistical mechanical version of the *second law of thermodynamics*.

As an example of this second law, consider two different gases (e.g., nitrogen and oxygen) in a container, separated by a thin membrane. One set of gas molecules is constrained to lie on one side of the membrane; the other set lies on the opposite side. The total number of available states $N_{\rm states}$ is less than if the membrane is ruptured and the two gases are allowed to mix. The mixed state is accessible from the partitioned state and not vice-versa. When the membrane is removed, the entropy begins to increase in accord with the second law of thermodynamics.

Since any ensemble of identical, closed systems will ultimately, after a time $t \gg \tau_{\rm int}$, evolve into microcanonical statistical equilibrium, it must be that the microcanonical distribution function $\rho =$ constant has a larger entropy than any other distribution function that the ensemble could acquire. That this, indeed, is so, can be demonstrated formally as follows: Consider the class of all distribution functions $\rho$ that: (i) vanish unless the constants of motion have the prescribed values $\tilde{E}$ (in the tiny range $\delta\tilde{E}$) and $K_A$, (ii) can be non-zero anywhere in the region of phase space, $\mathcal{S}_o$, where the prescribed values $\tilde{E}$, $K_A$ are taken on, and (iii) are correctly normalized so that

$$\int_{\mathcal{S}_o} \rho \mathcal{N}_{\rm states} d\Gamma = 1 \; . \tag{3.46}$$

We ask which $\rho$ in this class gives the largest entropy $S$. The requirement that the entropy be extremal (stationary) under variations $\delta\rho$ of $\rho$ that preserve the normalization (3.46) is embodied in the variational principle[7]

$$\delta \int_{\mathcal{S}_o} (-k\rho \ln \rho - \Lambda\rho) \mathcal{N}_{\rm states} d\Gamma = 0 \; . \tag{3.47}$$

Here $\Lambda$ is a Lagrange multiplier that enforces the normalization. Performing the variation, we find that

$$\int_{\mathcal{S}_o} (-k \ln \rho - k - \Lambda) \delta\rho \mathcal{N}_{\rm states} d\Gamma = 0 \; , \tag{3.48}$$

which is satisfied if and only if $\rho$ is a constant, $\rho = e^{-1-\Lambda/k}$, independent of location in the allowed region $\mathcal{S}_o$ of phase space; i.e., if and only if $\rho$ is that of the microcanonical ensemble. This calculation actually only shows that the microcanonical ensemble has stationary entropy. To show it is a maximum, one must perform the second variation; i.e., one must compute the second-order contribution of $\delta\rho$ to $\delta S$. When one is varying around the

[7]See, e.g., Mathews & Walker (1964), Chap. 12.

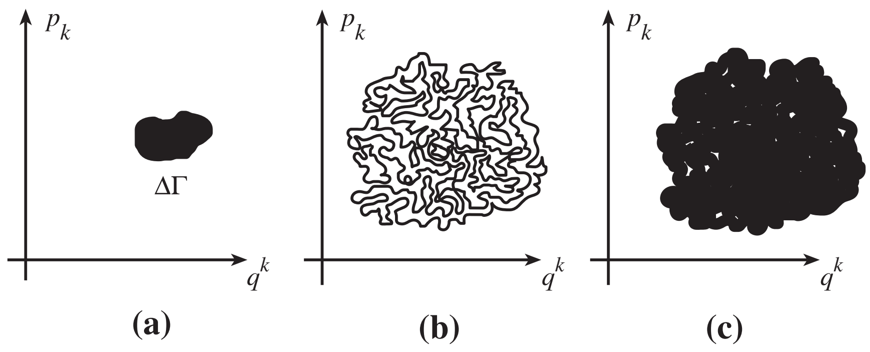


**Fig. 3.2:** Evolution of a classical ensemble into statistical equilibrium by means of phase mixing followed by coarse-graining of one's viewpoint.

microcanonical distribution, that second-order contribution is easily seen to be

$$\delta^2 S = \int_{\mathcal{S}_o} \Big(-k\frac{(\delta\rho)^2}{\rho}\Big)\mathcal{N}_{\rm states}d\Gamma < 0 \;. \tag{3.49}$$

Thus, the microcanonical distribution does *maximize* the entropy, as claimed.

### 3.6.1 The Approach to Equilibrium

There is an apparent paradox that lies at the heart of statistical mechanics and, at various stages in the development of the subject has led to confusion and even despair.[8] It still causes controversy.[9] Its simplest and most direct expression is to ask how can the time-reversible, microscopic laws, encoded in a time-independent Hamiltonian, lead to the remorseless increase of the entropy of the universe?

Consider, first, a classical, microcanonical ensemble of precisely closed systems (no interaction at all with the external universe). Assume, for simplicity, that at time $t = 0$ all the systems are concentrated in a small but finite region of phase space with volume $\Delta\Gamma$, as shown in Fig. 3.2(a), with

$$\rho = \frac{1}{\mathcal{N}_{\rm states}\Delta\Gamma} \quad \text{in the occupied region,} \tag{3.50}$$

$$\rho = 0 \qquad \text{everywhere else .} \tag{3.51}$$

$$\tag{3.52}$$

As time passes each system evolves under the action of the systems' common Hamiltonian. As we saw in Fig. 3.2, this evolution distorts the occupied region of phase space; but Liouville's theorem dictates that the occupied region's volume, $\Delta\Gamma$, remain unchanged; and, correspondingly, that the ensemble's entropy

$$S = -k\int(\rho\ln\rho)\mathcal{N}_{\rm states}d\Gamma = k\ln(\mathcal{N}_{\rm states}\Delta\Gamma) \tag{3.53}$$

[8]Boltzmann committed suicide.
[9]Hawking (1989), Penrose (1989).

remain unchanged.

How can this be so? The ensemble is supposed to evolve into statistical equilibrium, with its distribution function uniformly spread out over that entire portion of phase space allowed by the Hamiltonian's constants of motion—a portion of phase space far far larger than $\Delta\Gamma$—and in the process the entropy is supposed to increase.

Fig. 3.2 (b,c) resolves the paradox. As time passes, the occupied region becomes more and more distorted. It retains its phase-space volume, but gets strung out into a winding, contorted surface which (by virtue of the ergodic hypothesis) ultimately passes arbitrarily close to any given point in the region allowed by the constants of motion. This ergodic wandering is called *phase mixing*. Ultimately the physicist gets tired of keeping track of all these contortions of the occupied region and chooses instead to take a *coarse-grained* viewpoint which averages over scales larger than the distance between adjacent portions of the occupied surface and thereby regards the ensemble as having become spread over the entire allowed region. Or, more typically, the physicist will perform a coarse-grained smearing out on some given, constant scale at all times; and once the transverse scale of the ensemble's lengthening and narrowing phase-space region drops below the smearing scale, its smeared volume and its entropy start to increase. Thus, for an ensemble of closed systems it is the physicist's choice to perform coarse-grain averaging that causes entropy to increase and causes the ensemble to evolve into statistical equilibrium.

The situation is a bit more subtle for an ensemble of systems interacting with a thermal bath. The evolution into statistical equilibrium is driven by the interactions. Thus, it would appear at first sight that the physicist is not, this time, to blame for the entropy increase and the achievement of statistical equilibrium. A deeper examination, however, reveals the physicist's ultimate culpability. If the physicist were willing to keep track of all those dynamical degrees of freedom of the bath which are influenced by and influence the systems in the ensemble, then the physicist could incorporate these degrees of freedom into the description of the systems and define a phase space volume that obeys Liouville's theorem and thus does not increase, and an entropy that correspondingly remains constant. However, physicists instead generally choose to ignore the microscopic details of the bath, and that choice forces them to attribute a growing entropy to the ensemble and regard the ensemble as approaching statistical equilibrium.

When one reexamines these issues in quantum mechanical language, one discovers that the entropy increase is caused by the physicists' discarding the quantum mechanical correlations, essentially the off-diagonal terms in the density matrix (c.f. Box 3.1) that get built up through the systems' interaction with the rest of the universe. This discarding of correlations is accomplished through a *reduction* or *collapse* of the quantum state of the ensemble (which, if the state was originally pure, makes it mixed). From this viewpoint, then, it is the choice of the physicist to reduce the state of the ensemble by discarding correlations with the external universe that causes the entropy increase and the evolution into statistical equilibrium. Heuristically, we can say that the entropy does not increase *until* the physicist actually (or figuratively) tries to measure it.

This then raises a most intriguing question. What if we regard the universe as the ultimate microcanonical system? In this case, we might expect that the entropy of the universe will remain identically zero for all time, unless physicists (or other intelligent beings) perform some sort of coarse graining or discard some sort of correlations. However, such coarse graining or discarding are made deeply subtle by the fact that the physicists (or intelligent beings) are themselves part of the system being studied. Further discussion of these questions introduces fascinating, though ill-understood, quantum mechanical and cosmological considerations to which we shall briefly return in Part VI.

****************************

## EXERCISES

**Exercise 3.3** *Practice and Problem: Estimating Entropy*

Make rough estimates of the entropy of the following systems, assuming they are in statistical equilibrium:

(a) An electron in a hydrogen atom at room temperature
(b) A glass of wine
(c) The Pacific ocean
(d) An ice cube
(e) The universe (This is mostly contained in the 3 K microwave background radiation.)

**Exercise 3.4** *Derivation: Additivity of Entropy for Statistically Independent Systems*

Consider an ensemble of classical systems with each system made up of a large number of statistically independent subsystems, so $\rho = \prod_a \rho_a$. Show that the entropy of the full ensemble is equal to the sum of the entropies of the subensembles $a$: $S = \sum_a S_a$.

**Exercise 3.5** *Practice: Entropy of Thermalized Mode of a Field*

Consider a mode $\mathcal{S}$ of a fermionic or bosonic field, as discussed in Exercise 3.2 above. Suppose that an ensemble of identical such modes is in statistical equilibrium with a heat and particle bath and thus is grand-canonically distributed.

(a) Show that if $\mathcal{S}$ is fermionic, then the ensemble's entropy is

$$\begin{aligned} S_{\mathcal{S}} &= -k[\bar\eta \ln \bar\eta + (1-\bar\eta)\ln(1-\bar\eta)] && (3.54) \\ &\simeq -k\bar\eta(\ln\bar\eta - 1) \quad \text{in the classical regime } \bar\eta \ll 1 \; , && (3.55) \end{aligned}$$

where $\bar\eta$ is the mode's fermionic mean occupation number (3.40).

(b) Show that if the mode is bosonic, then the entropy is

$$\begin{aligned} S_{\mathcal{S}} &= k[(\bar\eta+1)\ln(\bar\eta+1) - \bar\eta\ln\bar\eta] && (3.56) \\ &\simeq -k\bar\eta(\ln\bar\eta - 1) \quad \text{in the classical regime } \bar\eta \ll 1 \; , && (3.57) \end{aligned}$$

where $\bar\eta$ is the bosonic mean occupation number (3.42). Note that in the classical regime, $\bar\eta \simeq e^{-(\tilde E - \tilde\mu)/kT} \ll 1$, the entropy is insensitive to whether the mode is bosonic or fermionic.

(c) Show that the entropy per particle in units of Boltzmann's constant (which we denote by $\sigma$) is $\sigma = S_{\mathcal{S}}/\bar\eta k$. Plot $\sigma$ as a function of $\bar\eta$ for fermions and for bosons. Show analytically that for degenerate fermions ($\bar\eta \simeq 1$) and for the bosons' classical-wave regime ($\bar\eta \gg 1$) the entropy per particle is small compared to unity.

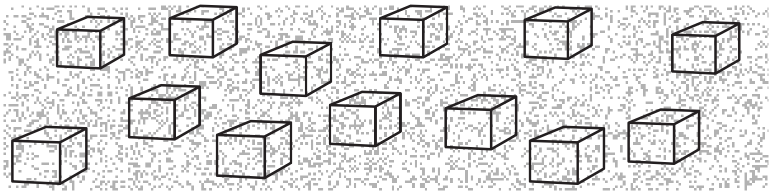

**Fig. 3.3:** An ensemble of gas cells, each with volume $V$, inside a heat and particle bath.

****************************

## 3.7 Statistical Mechanics of an Ideal Monatomic Gas

We now turn to the first of several examples that illustrate the formalism we have developed: An ideal, relativistic monatomic gas. In this section and Exercise 3.6 we shall compute the gas's statistical properties in the context of the grand canonical ensemble; and then in Exercise 3.8 we shall compute its entropy in the microcanonical ensemble. In Chap. 4 we shall return to the gas and explore in greater detail its properties and the relationships between its various ensembles.

We consider the ensemble of systems illustrated in Fig. 3.3. Each system in the ensemble is a cell of fixed volume $V$, with imaginary walls, inside a huge thermal bath of a monatomic gas (e.g. helium or atomic hydrogen or neutrons or photons or ...). Since the cells' walls are imaginary, the cells can and do exchange energy and atoms with the bath. We presume that the gas particles do not interact with each other (no potential energies in the Hamiltonian) so the gas is *ideal*. The bath is characterized by a chemical potential $\tilde{\mu}$ for these particles and by a temperature $T$, and each particle has rest mass $m$. In most textbooks on statistical mechanics one does not include the particles' rest mass in the chemical potential; however, we wish to allow the particles to be relativistic (e.g. they might be photons, which move at the speed of light), so we will be careful to include the rest mass in both the chemical potential $\tilde{\mu}$ and in the energy $\tilde{E} = (m^2 + |\mathbf{p}|^2)^{1/2} = (m^2c^4 + |\mathbf{p}|^2c^2)^{1/2}$ of each particle. We assume that the chemical potential is sufficiently small (sufficiently few particles) that the mean occupation number of the particles' quantum states, $\bar{\eta}$, is small compared to unity, so they behave classically. According to Eqs. (3.40) and (3.42) with $\tilde{E} \simeq mc^2$ (which gives the largest $\bar{\eta}$), this means that

$$\mu \equiv \tilde{\mu} - mc^2 \ll -kT \ . \tag{3.58}$$

Note that the quantity $\mu$ is the chemical potential with the particle's rest mass subtracted off; i.e., it is the nonrelativistic chemical potential of Eq. (3.33). (However, we do not require for now that $kT$ be $\ll mc^2$, i.e., that the particles be nonrelativistic.)

We presume that our ensemble of cells has reached statistical equilibrium with the bath, so its probabilistic distribution function has the grand canonical form (3.32):

$$\rho_n = \frac{1}{Z}\exp\left(\frac{-\tilde{E}_n + \tilde{\mu}N_n}{kT}\right) = \exp\left(\frac{\Omega - \tilde{E}_n + \tilde{\mu}N_n}{kT}\right) \ . \tag{3.59}$$

Here $\tilde{E}_n$ is the energy of a system that is in the many-particle quantum state $|n\rangle$, $N_n$ is the number of particles in that quantum state, and $1/Z \equiv e^{\Omega/kT}$ is the normalization constant that guarantees $\sum_n \rho_n = 1$; i.e.,

$$Z \equiv \exp\left(\frac{-\Omega}{kT}\right) \equiv \sum_n \exp\left(\frac{-\tilde{E}_n + \tilde{\mu} N_n}{kT}\right) . \tag{3.60}$$

This normalization constant, whether embodied in $Z$ or in $\Omega$, is a function of the bath's temperature $T$ and chemical potential $\tilde{\mu}$, and also of the cells' common volume $V$ (which enters into the listing of the available states $|n\rangle$). When regarded as a function of $T$, $\tilde{\mu}$, and $V$, the quantity $Z(T, \tilde{\mu}, V)$ is called the gas's *grand partition function*, and $\Omega(T, \tilde{\mu}, V)$ is called its *grand potential*. The following general argument shows that, once one has computed the explicit functional form for the grand potential (or equally well for the grand partition function), one can then derive from it all the thermodynamic properties of the thermally equilibrated gas. The argument is so general that it applies to *every* grand canonical ensemble of systems, not just to our chosen, monatomic gas; but for simplicity, we shall restrict ourselves to systems made of a single type of particle (which need not for now be monatomic or free of potential-energy interactions).

We introduce, as key quantities in the argument, the mean energy and mean number of particles in the ensemble's systems (cells):

$$\overline{\tilde{E}} \equiv \sum_n \rho_n \tilde{E}_n \ , \quad \text{and } \overline{N} \equiv \sum_n \rho_n N_n \ . \tag{3.61}$$

We now ask how the grand potential will change if the temperature $T$ and chemical potential $\tilde{\mu}$ of the bath and therefore of the ensemble are slowly altered with the volumes $V$ of the cells held fixed. The answer for the change $d\Omega$ produced by changes $dT$ and $d\tilde{\mu}$ can be derived from the normalization equation (3.60), which we rewrite as

$$1 = \sum_n \rho_n = \sum_n \exp\left(\frac{\Omega - \tilde{E}_n + \tilde{\mu} N_n}{kT}\right) . \tag{3.62}$$

Since the normalization condition must continue to hold as $T$ and $\tilde{\mu}$ change, the sum in Eq. (3.62) must be left unchanged, which means that

$$0 = \sum_n \left(\frac{d\Omega + N_n d\tilde{\mu} - (\Omega - \tilde{E}_n + \tilde{\mu} N_n) T^{-1} dT}{kT}\right) \rho_n \ . \tag{3.63}$$

Using expressions (3.61) for the mean energy and the mean number of particles, and rearranging terms, we obtain

$$d\Omega = -\overline{N} d\tilde{\mu} + (\Omega - \overline{\tilde{E}} + \tilde{\mu}\overline{N}) T^{-1} dT \ . \tag{3.64}$$

This change can be reexpressed in a more useful form by introducing the ensemble's entropy. We evaluate the entropy $S = -k \sum_n \rho_n \ln \rho_n$ using expression (3.59) for $\rho_n$, which says that

$\ln \rho_n = (\Omega - \tilde{E}_n + \tilde{\mu} N_n)/kT$, so

$$S = -k \sum_n \rho_n \ln \rho_n = -k \sum_n \rho_n \left( \frac{\Omega - \tilde{E}_n + \tilde{\mu} N_n}{kT} \right) = -\frac{\Omega - \overline{\tilde{E}} + \tilde{\mu} \overline{N}}{T}; \tag{3.65}$$

or, equivalently

$$\Omega = \overline{\tilde{E}} - TS - \tilde{\mu} \overline{N} \; . \tag{3.66}$$

By inserting expression (3.66) into Eq. (3.64), we obtain

$$d\Omega = -\overline{N} d\tilde{\mu} - S dT \; . \tag{3.67}$$

Equation (3.67) has several important consequences. The first consequence is the fact that it is actually the *First Law of Thermodynamics* in disguise. To see this, insert expression (3.66) for $\Omega$ into (3.67), thereby bringing it into the form

$$d\overline{\tilde{E}} = \tilde{\mu} d\overline{N} + T dS, \tag{3.68}$$

which is the familiar form of the first law of thermodynamics, but with the "$-PdV$" work, associated with a change in a cell's volume, omitted because the cells have fixed volume $V$. If we (momentarily) pass from our original grand canonical ensemble, all of whose cells have the same $V$, $\tilde{\mu}$, and $T$, to another grand canonical ensemble whose cells have the same $\tilde{\mu}$ and $T$ as before, but have slightly larger volumes, $V + dV$, then according to Eq. (3.66) with $\tilde{\mu}$ and $T$ fixed, $\Omega$ will change by $d\Omega = d\overline{\tilde{E}} - TdS - \tilde{\mu} d\overline{N}$ (where $dS$ and $d\overline{N}$ are the changes of entropy and mean number of particles induced by the volume change); and by the elementary first law of thermodynamics, $d\overline{\tilde{E}} = -PdV + \tilde{\mu} d\overline{N} + TdS$, this change of $\Omega$ at fixed $\tilde{\mu}$ and $T$ is simply $-PdV$. Combining with Eq. (3.67), this gives for the change of $\Omega$ when all of $\tilde{\mu}$, $T$, and $V$ change:

$$d\Omega = -PdV - \overline{N} d\tilde{\mu} - S dT \; . \tag{3.69}$$

A second consequence of Eq. (3.67), in the generalized form (3.69), is the fact that it tells us how to compute the mean number of particles, the entropy, and the pressure in terms of $\tilde{\mu}$, $T$, and $V$:

$$\overline{N} = -\left( \frac{\partial \Omega}{\partial \tilde{\mu}} \right)_{V,T} , \quad S = -\left( \frac{\partial \Omega}{\partial T} \right)_{V,\tilde{\mu}} , \quad P = -\left( \frac{\partial \Omega}{\partial V} \right)_{\tilde{\mu},T} . \tag{3.70}$$

A third consequence is the fact that Eqs. (3.70) and (3.66) determine the cell's full set of standard thermodynamic parameters, $\{\overline{\tilde{E}}, T, \overline{N}, \tilde{\mu}, V, P, S\}$ in terms of the ensemble's three independent quantities $T$, $\tilde{\mu}$, $V$. Thus, as soon as the functional form of $\Omega(T, \tilde{\mu}, V)$ is known, from it we can generate all the thermodynamic properties of our equilibrated cells of gas. As we shall see in Chap. 4 and shall illustrate in Exercise 3.8, the resulting thermodynamic properties are independent of the equilibrium ensemble used to compute them: grand canonical, canonical, Gibbs, or microcanonical. This is completely general: *When confronted by a new* *type of system, one can compute its thermodynamic properties using whatever equilibrium ensemble is most convenient; the results will be ensemble-independent.*

Returning to the specific case of a monatomic, relativistic gas, we can compute $\Omega(T, \tilde{\mu}, V)$ by carrying out explicitly the sum over states of Eq. (3.60). We first fix (temporarily) the number of particles $N$ in an individual system and its corresponding number of degrees of freedom $W = 3N$. We then perform the integral over phase space. Then we sum over all $N$ from 1 to $\infty$. The details are spelled out in Exercise 3.6 below. In the non-relativistic limit, $kT \ll mc^2$, this calculation yields

$$\Omega(T, \mu, V) = -kTV \frac{(2\pi mkT)^{3/2}}{h^3} e^{\mu/kT} , \quad \text{where } \mu \equiv \tilde{\mu} - mc^2 , \tag{3.71}$$

and we have used the "classical particles" condition (3.58). In the ultra-relativistic limit, $kT \gg mc^2$ (e.g. for photons or neutrinos or an ultrarelativistic gas of electrons), the calculation gives

$$\Omega(T, \tilde{\mu}, V) = -\frac{8\pi V (kT)^4}{h^3} e^{\tilde{\mu}/kT} . \tag{3.72}$$

In the non-relativistic limit, differentiation of the grand potential (3.71) gives [with the aid of Eqs. (3.70) and (3.66)] the following thermodynamic relations

$$\overline{N} = \tfrac{(2\pi mkT)^{3/2}}{h^3} e^{\mu/kT} V , \quad S = \left(\tfrac{5}{2} - \tfrac{\mu}{kT}\right) k\overline{N} , \tag{3.73}$$

$$P = kT \tfrac{(2\pi mkT)^{3/2}}{h^3} e^{\mu/kT} = \tfrac{\overline{N}}{V} kT , \quad \overline{E} = \tfrac{3}{2} kT\overline{N} . \tag{3.74}$$

The corresponding ultrarelativistic expressions are derived in Exercise 3.6. Note that the first of the nonrelativistic expressions (3.74) is the same number density of particles, $\overline{N}/V$, as we derived from kinetic theory in the last chapter [Eq. (2.88)], but with $g_s = 1$ since we have ignored spin degrees of freedom in the above analysis. The second expression says that each particle in a nonrelativistic, thermalized, ideal, classical gas carries an entropy of $5/2 - \mu/kT$ in units of Boltzmann's constant. The third and fourth expressions are the standard, nonrelativistic ideal-gas equations of state, $P = (\overline{N}/V)kT$ and $\epsilon \equiv \overline{E}/V = \frac{3}{2}nkT$ [Eq. (2.89)].

We note in passing that the entropy per particle in units of Boltzman's constant,

$$\sigma \equiv S/Nk , \tag{3.75}$$

is a very useful concept in both quantitative analyses and order-of-magnitude analyses [see, e.g., Exercise 3.9; also the discussion of entropy in the expanding universe in Sec. 3.8]. One reason is the second law of thermodynamics. Another is that in the real universe $\sigma$ generally lies somewhere between 0 and 100 and thus is a natural quantity in terms of which to think and remember. For example, for ionized hydrogen gas in the nonrelativistic, classical domain, Eqs. (3.74) [with $\mu/kT$ from the second equation inserted into the first] say that the protons' entropy per proton is a factor $\simeq 10$ higher than that of the electrons, $\sigma_p - \sigma_e = \frac{3}{2}\ln(m_p/m_e) = 11.27 \simeq 10$, and thus for an ionized hydrogen gas most of the entropy

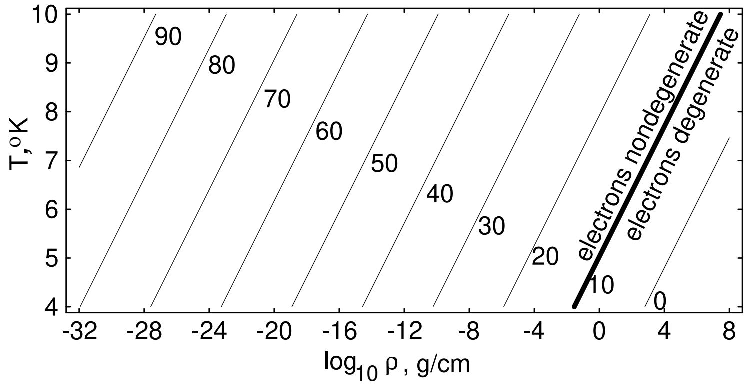


**Fig. 3.4:** The proton entropy per proton $\sigma_p$ for an ionized hydrogen gas. The electron entropy per electron, $\sigma_e$, is a factor $\simeq 10$ smaller; the electrons are degenerate when $\sigma_e \simeq \sigma_p - 10 \lesssim 1$. The protons are degenerate when $\sigma_p \lesssim 1$.

is in the protons. The proton entropy per proton grows logarithmically with decreasing density $\rho$,

$$\sigma_p = \frac{5}{2} - \frac{\mu_p}{kT} = \frac{5}{2} + \ln\left[\frac{2m_p}{\rho}\left(\frac{2\pi m_p kT}{h^2}\right)^{3/2}\right] ; \tag{3.76}$$

[Eqs. (3.74) and (3.74)]. This entropy per proton is plotted as a function of density and temperature in Fig. 3.4. It ranges from $\sim 1$ near the onset of degeneracy (the failure of the classical approximation), to about 100 at the lowest density that occurs in the universe, $\sim 10^{-29}$ g/cm$^3$. This range is an example of the fact that the logarithms of almost all dimensionless numbers that occur in nature lie between $\sim -100$ and $+100$. In Ex. 3.5 above it was shown that for Fermions (including, of course, protons and electrons), the entropy per particle in the degenerate region (lower right of Fig. 3.4) is small compared to one.

The statistical mechanical tools developed in this chapter not only can reproduce all the thermodynamic results that kinetic theory gives us; they can do much more. For example, as we shall see in the following two chapters (and we learn as a preview in Exercise 3.7), these tools can be used to study statistical fluctuations of the physical quanties, whose mean values we have computed.

****************************

## EXERCISES

**Exercise 3.6** *Derivation and Example: Grand Canonical Ensemble for a Monatomic Gas*

Consider the cells of ideal, classical, monatomic gas with volume $V$ that reside in the heat and particle bath of Fig. 3.3. Assume initially that the temperature $T$ of the bath has an arbitrary magnitude relative to the rest mass-energy $mc^2$ of the particles, but require $kT \ll -\mu$ so all the particles behave classically. Ignore the particles' spin degrees of freedom, if any.

(a) The number of particles in a given system can be anything from $N = 0$ to $N = \infty$. Restrict attention, for the moment, to a situation in which the cell contains a precise number of particles, $N$. Show that the multiplicity is

$$\mathcal{M} = N! \tag{3.77}$$

so long as the probability $\rho$ is always small and the system can be regarded classically.

(b) Still holding fixed the number of particles in the cell, show that the number of degrees of freedom $W$, the number density of states in phase space $\mathcal{N}_{\rm states}$ and the energy $\tilde{E}_N$ in the cell are

$$W = 3N\ , \quad \mathcal{N}_{\rm states} = \frac{1}{N!h^{3N}}\ , \quad \tilde{E}_N = \sum_{A=1}^{N}({\bf p}_A{}^2 + m^2)^{\frac{1}{2}}\ , \tag{3.78}$$

where ${\bf p}_A$ is the momentum of classical particle number $A$.

(c) Show that the sum over states (3.60) for the grand partition function becomes

$$Z = e^{-\Omega/kT} = \sum_{N=0}^{\infty}\frac{V^N}{N!h^{3N}}e^{\tilde{\mu}N/kT}\left[\int_0^\infty \exp\left(-\frac{(p^2+m^2)^{\frac{1}{2}}}{kT}\right)4\pi p^2 dp\right]^N\ . \tag{3.79}$$

(d) Show that in the nonrelativistic limit this gives Eq. (3.71), and in the extreme relativistic limit it gives Eq. (3.72).

(e) For the extreme relativistic limit use your result (3.72) for the grand potential $\Omega(V, T, \tilde{\mu})$ to derive the mean number of particles $\overline{N}$, the pressure $P$, the entropy $S$, and the mean energy $\overline{\tilde{E}}$ as functions of $V$, $T$, and $\tilde{\mu}$. Note that for a photon gas, because of the spin degree of freedom, the correct values of $\overline{N}$, $\overline{\tilde{E}}$ and $S$ will be twice as large as you obtain in this calculation. Show that $\overline{\tilde{E}}/V = 3P$ (a relation valid for any ultrarelativistic gas); and that $\overline{\tilde{E}}/\overline{N} = 3kT$ (which is higher than the $2.70 \ldots kT$ for black-body radiation, as derived in Exercise 2.5, because in the classical regime of $\bar{\eta} \ll 1$ photons don't cluster in the same state at low frequency, thereby lowering the mean photon energy.)

**Exercise 3.7** *Example: Probability Distribution for the Number of Particles in a Cell*

Suppose that we make a large number of measurements of the number of atoms in one of the systems of the ensemble of Fig. 3.3, (and Ex. 3.6), and that from those measurements we compute a probability $p_N$ for that cell to contain $N$ particles.

(a) How widely spaced in time must the measurements be to guarantee that the measured probability distribution is the same as that which one computes from the ensemble of cells at a specific moment of time?

(b) Assume that the measurements are widely enough separated for your criterion [part (a)] to be satisfied. Use the grand canonical distribution to show that the probability $p_N$ is given by the Poisson distribution

$$p_N = e^{-\overline{N}}(\overline{N}^N/N!)\ . \tag{3.80}$$

(c) Show that the mean number of particles in a single system, as predicted by this distribution, is $\langle N\rangle = \overline{N}$, and the root mean square deviation from this is

$$\Delta N \equiv \langle (N - \overline{N})^2\rangle^{\frac{1}{2}} = \overline{N}^{\frac{1}{2}}\ . \tag{3.81}$$

**Exercise 3.8** *Example: Entropy of a Monatomic Gas in the Microcanonical Ensemble*

Consider a microcanonical ensemble of closed cells with volume $V$, each containing precisely $N$ particles of a monatomic, nonrelativistic, ideal, classical gas, and each containing a nonrelativistic energy $E \equiv \tilde{E} - Nmc^2$. For the moment (by contrast with the text's discussion of the microcanonical ensemble), assume that $E$ is precisely fixed instead of being spread over some tiny but finite range.

(a) Explain why the region $\mathcal{S}_o$ of phase space accessible to each system is

$$|x_A^j| < L/2 \ , \quad \sum_{A=1}^{N} \frac{1}{2m} |\mathbf{p}_A|^2 = E \ , \tag{3.82}$$

where $A$ labels the particles.

(b) In order to compute the entropy of the microcanonical ensemble, we compute the volume in phase space $\Delta\Gamma$ that it occupies, then multiply by the number density of states in phase space (which is independent of location in phase space), and then take the logarithm. Explain why

$$\Delta\Gamma \equiv \prod_{A=1}^{N} \int_{\mathcal{S}_o} dx_A dy_A dz_A dp_A^x dp_A^y dp_A^z \tag{3.83}$$

vanishes. This illustrates the "set of measure zero" statement in the text (second paragraph of Sec. 3.5), which we used to assert that we must allow the systems' energies to be spread over some tiny but finite range.

(c) Now permit the energies of our ensemble's cells to lie in the tiny but finite range $E_o - \delta E_o < E < E_o$. Show that

$$\Delta\Gamma = V^N [\mathcal{V}_\nu(a) - \mathcal{V}_\nu(a - \delta a)] \ , \tag{3.84}$$

where $\mathcal{V}_\nu(a)$ is the volume of a sphere of radius a in a Euclidean space with $\nu \gg 1$ dimensions, and where

$$a \equiv \sqrt{2mE_o} \ , \quad \frac{\delta a}{a} \equiv \frac{1}{2}\frac{\delta E_o}{E_o} \ , \quad \nu \equiv 3N \ . \tag{3.85}$$

It can be shown (and you might want to try) that

$$\mathcal{V}_\nu(a) = \frac{\pi^{\nu/2}}{(\nu/2)!} a^\nu \quad \text{for } \nu \gg 1 \ . \tag{3.86}$$

(d) Show that, so long as $1 \gg \delta E_o / E_o \gg 1/N$ (where $N$ in practice is an exceedingly huge number),

$$\mathcal{V}_\nu(a) - \mathcal{V}_\nu(a - \delta a) \simeq \mathcal{V}_\nu(a)[1 - e^{-\nu\delta a/a}] \ \simeq \ \mathcal{V}_\nu(a) \ , \tag{3.87}$$

which is independent of $\delta E_o$ and thus will produce a value for $\Delta\Gamma$ and thence $N_{\rm states}$ and $S$ independent of $\delta E_o$, as desired. From this and with the aid of Stirling's approximation, $n! \simeq 2\pi n^{1/2}(n/e)^n$ for large $n$, and taking account of the multiplicity $\mathcal{M} = N!$, show that the entropy of the microcanonically distributed cells is given by

$$S(V, E, N) = Nk \ln \left[ \frac{V}{N} \left(\frac{E}{N}\right)^{3/2} \left(\frac{4\pi m}{3h^2}\right)^{3/2} e^{5/2} \right] \ . \tag{3.88}$$

This is known as the *Sackur-Tetrode* equation. [Note that the entropy per particle differs from that for the protons in ionized hydrogen, Eq. (3.76), by a factor 2 inside the logarithm because of the proton's two spin states.]

(e) Returning to the grand canonical ensemble of the text and Exercise 3.6, show that its entropy (3.74), when expressed in terms of $V/\overline{N}$ and $\overline{E}/\overline{N}$, takes precisely the same form as this Sackur-Tetrode equation. This illustrates the fact that the thermodynamic properties of a thermally equilibrated system are independent of the nature of its statistical equilibrium, i.e. independent of the type of bath (if any) that has brought it to equilibrium.

**Exercise 3.9** *Problem: Primordial Element Formation*

When the expanding universe was $t \sim 10^{-4}$ second old, it contained equal numbers of protons and neutrons, plus far more photons, electrons, and positrons, all in statistical equilibrium at a temperature $T \sim 10^{12}$ K. As the universe continued to expand, its temperature fell as $T \propto 1/\sqrt{t}$. Nuclear reactions in this expanding plasma were continually trying to make alpha particles (helium nuclei): $2n + 2p \rightarrow \alpha + 7$ Mev, with the 7 Mev of energy able to go into thermalized photons. Using an order-of-magnitude computation based on entropy per baryon, show that the second law of thermodynamics prevented this helium formation from occuring until the temperature had fallen to some critical value $T_{\rm crit}$, and thereafter the second law encouraged helium formation to occur. Compute $T_{\rm crit}$ and the corresponding time $t$ at which helium could start to form. The fraction of the universe's protons and neutrons that actually became $\alpha$'s (about 25%) was controlled by the actual rates for the relevant nuclear reactions and by the complete shutoff of these reactions via beta decay of the neutrons at $t \sim$ (neutron half life) = 13 minutes.

****************************

## 3.8 Statistical Mechanics in the Presence of Gravity

Systems with significant gravity behave quite differently, statistical mechanically, than systems without gravity. This has led to much controversy as to whether statistical mechanics can really be applied to gravitating systems. Despite that controversy, statistical mechanics *has* been applied in the presence of gravity in a variety of ways, with great success, and with important, fundamental conclusions. In this section we sketch some of those applications: to galaxies, black holes, the universe as a whole, and the formation of structure in the universe. Our discussion is intended to give just the flavor of these subjects and not full details, so we shall state a number of things without derivation. This is necessary in part because many of the phenomena we shall describe rely for their justification on general relativity (Part VI) and/or quantum field theory in curved spacetime (not treated in this book).

### 3.8.1

Galaxies We shall idealize a galaxy as a spherically symmetric cluster of $N \sim 10^{11}$ stars each with mass $m$ comparable to that of the sun, $m \sim 1 M_\odot$.[10] The stars are bound under

[10] $1\mathrm{M}_\odot \equiv 2 \times 10^{30}$ kg.

their collective self-gravity. (In fact we know that there is also dark matter present, but we will ignore its effects. In addition, most galaxies including our own are not spherical, a fact we shall also ignore.) Now the characteristic size of a galaxy is $R \sim 10$ kiloparsecs (kpc).[11] Therefore the magnitude of its Newtonian gravitational potential in units of $c^2$ is $|\Phi/c^2| \sim GNm/Rc^2 \sim 10^{-6} \ll 1$ (where $G$ is Newton's gravitational constant) and the characteristic speed of the stars in the galaxy is $v \sim (GNm/R)^{1/2} \sim 200$ km s$^{-1} \ll c$, which means the gravity and stars can be analyzed in the nonrelativistic, Newtonian framework (cf. Fig. 1.1 and Part VI).

The time it takes stars to cross the galaxy is $\tau_{\rm int} \sim 2R/v \sim 10^8$ yr.[12] This time is short compared with the age of a galaxy, $\sim 10^{10}$ yr. Galaxies have distant encounters with their neighbors on timescales that can be smaller than their ages but still much longer than $\tau_{\rm int}$; in this sense, they can be thought of as semiclosed systems, weakly coupled to their environments. However, we shall idealize our chosen galaxy as fully closed (no interaction with its environment). Direct collisions between stars are exceedingly rare, and strong two-star gravitational encounters, which happen when the impact parameter is smaller than $\sim Gm/v^2 \sim R/N$, are also neglibibly rare. We can therefore regard each of our galaxy's stars as moving under a gravitational potential determined by the smoothed-out mass of all the other stars, and can use Hamiltonian dynamics to describe their motions.

We imagine that we have an ensemble of such galaxies, all with the same number of stars $N$, the same mass $M = Nm$, and the same energy $E$ (in a tiny range $\delta E$), and we begin our study of that ensemble by making an order of magnitude estimate of the probability $\rho$ of finding a chosen galaxy from the ensemble in some chosen (single-particle) quantum state. We compute that probability from the corresponding probabilities for its subsystems, individual stars: The phase-space volume available to each star in the galaxy is $\sim R^3(mv)^3$, the density of states in each star's phase space is $1/h^3$, the number of available states is the product of these $\sim (Rmv/h)^3$, and the probability of the star occupying the chosen state, or any other state, is the reciprocal of this, $\sim (h/Rmv)^3$. The probability of the galaxy occupying a state in its phase space is the product of the probabilities for each of its $N$ stars [Eq. (3.22)]:

$$\rho \sim \left(\frac{h}{Rmv}\right)^{3N} \sim 10^{-1000}. \tag{3.89}$$

This very small number suggests that it is somewhat silly of us to use quantum mechanics to normalize the distribution function (i.e., silly to use the probabilistic distribution function $\rho$), when dealing with a system as classical as a whole galaxy. Silly, perhaps; but dangerous, no. The key point is that, so far as classical statistical mechanics is concerned, the only important feature of $\rho$ is that it is proportional to the classical distribution function $\mathcal{N}$; its absolute normalization is usually not important, classically. It was this fact that permitted so much progress to be made in statistical mechanics prior to the advent of quantum mechanics.

Are real galaxies in statistical equilibrium? To gain insight into this question, we shall estimate the entropy of a galaxy in our ensemble and shall then ask whether that entropy has any chance of being the maximum value allowed to the galaxy's stars (as it must be if

[11] 1kpc$\equiv 3.1 \times 10^{16}$ km.

[12] 1 yr$\sim \pi \times 10^7$ s.

the galaxy is in statistical equilibrium).

Obviously, the stars (by contrast with electrons) are distinguishable, so we can assume multiplicity $\mathcal{M} = 1$ when estimating the galaxy's entropy. Ignoring the (negligible) correlations between stars, the entropy computed by integating $\rho \ln \rho$ over the galaxy's full $6N$ dimensional phase space is just $N$ times the entropy associated with a single star, which is approximately $S \sim Nk \ln(\Delta\Gamma/h^3)$ [Eqs. (3.53) and (3.8)], where $\Delta\Gamma$ is the phase space volume over which the star wanders in its ergodic, Hamiltonian-induced motion, i.e. the phase space volume available to the star. We shall express this entropy in terms of the galaxy's total mass $M = Nm$ and its total nonrelativistic energy $E \equiv \tilde{E} - Mc^2 \sim -GM^2/2R$. Since the characteristic stellar speed is $v \sim (GM/R)^{1/2}$, the volume of phase space over which the star wanders is $\Delta\Gamma \sim (mv)^3 R^3 \sim (GMm^2R)^{3/2} \sim (-G^2M^3m^2/2E)^{3/2}$, and the entropy therefore is

$$S_{\rm Galaxy} \sim (M/m)k \ln(\Delta\Gamma/h^3) \sim (3M/2m)k \ln(-G^2M^3m^2/2Eh^2)\ . \tag{3.90}$$

Is this the maximum possible entropy available to the galaxy, given the constraints that its mass be $M$ and its nonrelativistic energy be $E$? No. Its entropy can be made larger by removing a single star from the galaxy to radius $r \gg R$, where the star's energy is negligible. The entropy of the remaining stars will decrease slightly since the mass $M$ diminishes by $m$ at constant $E$. However, the entropy associated with the removed star, $\sim (3/2)\ln(GMm^2r/h^2)$ can be made arbitrarily large by making its radius $r$ arbitrarily large. Therefore, by this thought experiment we discover that galaxies cannot be in a state of maximum entropy at fixed $\tilde{E}$ and $M$, and they therefore cannot be in a true statistical equilibrium state.[13] (One might wonder whether there is entropy associated with the galaxy's gravitational field, and whether that entropy invalidates our analysis. The answer is no. The gravitational field has no randomness , beyond that of the stars themselves, and thus no entropy; its structure is uniquely determined, via Newton's field equation, by the stars' spatial distribution.)

In a real galaxy or other star cluster, rare near-encounters between stars in the cluster core cause individual stars to be ejected from the core into distant orbits or be ejected from the cluster altogether. These ejections increase the cluster's entropy in just the manner of our thought experiment. The core of the galaxy shrinks, a diffuse halo grows, and the total number of stars in the galaxy gradually decreases. This evolution to ever larger entropy is demanded by the laws of statistical mechanics, but by contrast with systems without gravity, it does not bring the cluster to statistical equilibrium. The long-range influence of gravity prevents a true equilibrium from being reached. Ultimately, the cluster's core may collapse to form a black hole.

### 3.8.2 Black Holes

Quantum field theory predicts that, near the horizon of a black hole, the vacuum fluctuations of quantized fields behave thermally, as seen by stationary (non-infalling) observers. More

[13] A true equilibrium *can* be achieved if the galaxy is enclosed in an idealized spherical box whose walls prevent stars from escaping, or if the galaxy lives in an infinite thermalized bath of stars so that, on average, whenever one star is ejected into a distant orbit in the bath, another gets injected into the galaxy; see, e.g., Ogorodnikov (1965) and Lynden-Bell (1967). However, in the real universe galaxies are not surrounded by walls or by thermalized star baths.

specifically, such observers see the horizon to be surrounded by an atmosphere that is in statistical equilibrium (a "thermalized" atmosphere) and that rotates with the same angular velocity $\mathbf{\Omega}_H$ as the hole's horizon.[14] The atmosphere contains all types of particles that can exist in nature. Very few of the particles manage to escape from the hole's gravitational pull; most fly up to some maximum height, then fall back down toward its horizon; only if they start out moving almost vertically upward (i.e., with near zero angular momentum) do they have any hope to escape. The few that do escape make up a tiny trickle of "Hawking radiation" that will ultimately cause the black hole to evaporate, unless it grows more rapidly due to infall of material from the external universe.

In discussing the distribution function for the hole's thermalized, rotating atmosphere, one must take account of the fact (Part VI) that the locally measured energy of a particle decreases as it climbs out of the hole's gravitational field; one does so by attributing to the particle the energy that it would ultimately have if it were to escape from the hole's gravitational grip. This is called the particle's "redshifted" energy and is denoted by $\tilde{E}_\infty$. This $\tilde{E}_\infty$ is conserved along the particle's world line, as is the projection $\mathbf{j}\cdot\hat{\mathbf{\Omega}}_H$ of the particle's orbital angular momentum $\mathbf{j}$ along the hole's spin axis (unit direction $\hat{\mathbf{\Omega}}_H$).

The hole's atmosphere is thermalized with vanishing chemical potentials, with a redshifted temperature $T_H$ equal to the temperature of the Hawking radiation that escapes, and with the rotational angular velocity $\mathbf{\Omega}_H$. Correspondingly, the probability that some randomly selected particle will be in an atmospheric quantum state with energy $E_\infty$ and angular momentum $\mathbf{j}$ is given by the following equilibrium distribution function [Eq. (3.30)]

$$\rho = C \exp\left[\frac{-\tilde{E}_\infty + \mathbf{\Omega}_H \cdot \mathbf{j}}{kT_H}\right] , \tag{3.91}$$

and similarly for the entire atmosphere. The temperature $T_H$ and angular velocity $\mathbf{\Omega}_H$, like all properties of a black hole, are determined completely by the hole's spin angular momentum $\mathbf{J}_H$ and its mass $M_H$; and, to within factors of order unity, they have magnitudes

$$T_H \sim \frac{\hbar}{8\pi G M_H/c^3} \simeq \frac{6\times 10^{-8}\mathrm{K}}{M_H/M_\odot} , \quad \Omega_H \sim \frac{J_H}{M_H(2GM_H/c^2)^2} . \tag{3.92}$$

Notice how small the hole's temperature is, if its mass is greater than or of order $M_\odot$. For such holes the thermal atmosphere is of no practical interest, though it has deep implications for fundamental physics. Only for tiny black holes (that might, e.g., have been formed in the big bang) is $T_H$ high enough to be physically interesting.

Suppose that the black hole evolves much more rapidly by accreting matter than by emitting Hawking radiation. Then the evolution of its entropy can be deduced from the first law of thermodynamics for its atmosphere. By techniques analogous to those in Sec. 3.8 above, one can show that the atmosphere's equilibrium distribution BHDistribution implies the following form for the first law (where we set $c = 1$):

$$dM = T_H dS_H + \mathbf{\Omega}_H \cdot d\mathbf{J}_H , \tag{3.93}$$

[14]This remarkable conclusion, due to Stephen Hawking, William Unruh, and Paul Davies, is discussed pedagogically in the book by Thorne, Price and Macdonald (1986).

where $dM_H$ is the change of the black-hole's mass due to the accretion (with each infalling particle contributing its $E_\infty$ to $dM_H$), $d\mathbf{J}_H$ is the change of the hole's spin angular momentum due to the accretion (with each infalling particle contributing its $\mathbf{j}$), and $dS_H$ is the increase of the black hole's entropy.

Because this first law can be derived (as described above) via the techniques of statistical mechanics, we are guaranteed that the hole's entropy increase has the standard statistical mechanical origin and interpretation: If $N_{\rm states}$ is the total number of quantum states that the infalling material could have been in (subject only to the requirement that the total infalling mass-energy be $dM_H$ and total infalling angular momentum be $d\mathbf{J}_H$), then $dS_H = k \log N_{\rm states}$ [cf. Eq. (3.45)]. In other words, the hole's entropy increases by $k$ times the logarithm of the number of quantum mechanically different ways that we could have produced its changes of mass and angular momentum, $dM_H$ and $d\mathbf{J}_H$. Correspondingly, we can regard the hole's total entropy as $k$ times the logarithm of the number of ways in which it could have been made. That number of ways is enormous, and correspondingly the hole's entropy is enormous: The above analysis, when carried out in full detail, reveals that the entropy is

$$S_H = k\frac{A_H}{4{l_P}^2} \sim 1 \times 10^{77} k \left(\frac{M}{M_\odot}\right)^2 \tag{3.94}$$

where $A_H \sim 4\pi(2GM_H/c^2)$ is the surface area of the hole's horizon and $l_P = \sqrt{G\hbar/c^3} = 1.616 \times 10^{-33}$ cm is the Planck length.

What is it about a black hole that leads to this peculiar thermal behavior and enormous entropy? Why is a hole so different from a star or galaxy? The answer lies in the black-hole horizon and the fact that things which fall inward through the horizon cannot get back out. In quantum field theory it is the horizon that produces the thermal behavior. In statistical mechanics it is the horizon that produces the loss of information about how the black-hole was made and the corresponding entropy increase. In this sense, the horizon for a black hole plays a role analogous to coarse graining in conventional classical statistical mechanics. [15]

The above statistical mechanical description of a hole's atmosphere and thermal behavior is based on the laws of quantum field theory in curved spacetime—laws in which the atmosphere's fields (electromagnetic, neutrino, etc.) are quantized but the hole itself is not governed by the laws of quantum mechanics. As of 1999, a much deeper understanding is arising from string theory—the most promising approach to quantum gravity and to quantization of black holes. Indeed, the thermal properties of black holes, and most especially their entropy, is a powerful testing ground for candidate theories of quantum gravity.

[15]It is not completely clear in 1999 whether the information about what fell into the hole gets completely lost, in principle as well as in practice, or whether—as with coarse-graining—the loss of information is the physicist's fault. It might be that the information is somehow retained in the black-hole atmosphere and we have just coarse-grained it away by our faulty understanding of the relevant black-hole quantum mechanics; cf. the Hawking-Thorne-Preskill bet on the wall in Bridge Annex.

### 3.8.3 The Universe

[16] Observations and theory agree that the universe began in a highly thermalized state, with all types of particles (except gravitons) in mutual statistical equilibrium. As the universe expanded, the energies of the particles and the equilibrium temperature all got cosmologically redshifted by the same amount, thereby maintaining the equilibrium; see Part VI. In the absence of gravity, this expanding equilibrium state would have had maximum entropy, and the second law of thermodynamics would have forbidden the development of galaxies, stars, and people.

Fortunately, gravity exists and through its long-range force it produced condensations with an attendant *increase* of entropy. The enormous power of gravity to do this is epitomized by the following:

After an early epoch of "inflation" (Part VI), the universe's expansion settled into a more leisurely form in which, when its age was $t$, a region with radius $R \sim t$ was able to communicate with itself. Here and below we set $c = 1$. As we shall show in Part VI, the temperature of this region, at time $t$, was given by $kT \sim E_P\sqrt{t_P/t}$, where $E_P = \sqrt{\hbar c^5/G} \sim 10^{19}$ Gev and $t_P = \sqrt{\hbar G/c^5} \sim 10^{-43}$ s are the Planck energy and Planck time.[17] As we shall also show in Part VI, the number of particles inside the communicable region was $N \sim (t/t_P)^{3/2}$ ($\sim 10^{91}$ today). Since each particle, in order of magnitude, carried an entropy $k$ [cf. Eq. (3.75) and associated discussion], the total entropy in the communicable region at time $t$ was $S \sim k(t/t_P)^{3/2}$. This actual entropy should be compared with the maximum entropy that could be produced, in the communicable region, with the aid of gravity. That maximum entropy would arise if we were to collapse the entire mass $M \sim NkT \sim E_P(t/t_P)$ of the communicable region into a black hole, thereby giving up all information about the content of the communicable region except its mass and (approximately vanishing) angular momentum: $S_{\rm max} \sim k(GM)^2/{l_P}^2 \sim k(t/t_P)^2$.

Notice that near the Planck time (when space and time as we know them presumably were coming into existence), the actual entropy in the communicable region and the maximum achievable entropy were of the same order, $S \sim S_{\rm max} \sim k$, so it was difficult to produce gravitational condensations. However, as the universe expanded, the actual entropy $S \sim k(t/t_P)^{3/2}$ lagged further and further behind the maximum achievable entropy $S_{\rm max} \sim k(t/t_P)^2$, so gravitational condensations became more and more favorable, statistically. Ultimately, those condensations did occur under the inexorable driving force of gravity, producing (with aid from the other fundamental forces) galaxies, stars and people.

[16] This subsection is based on Frautschi (1982), which contains many fascinating ideas and facts about entropy in the expanding universe.

[17] This formula predicts a temperature today for the cosmic microwave radiation $T_o \sim 100$ K that is too high by a factor $\sim 30$ because the details of the expansion changed, altering this formula a bit, when the universe was about a million years old and nonrelativistic matter became the dominant source of gravity rather than radiation; see Part VI.

### 3.8.4 Structure Formation in the Expanding Universe: Violent Relaxation and Phase Mixing

The formation of stars and galaxies ("structure") by gravitational condensation provides a nice illustration of the phase mixing and coarse graining that underly the second law of thermodynamics (Sec. 3.6 above):

It is believed that a galaxy's stars formed out of the universe's almost uniform, expanding gas, when slight overdensities (presumably irregular in shape) stopped expanding and began to contract under their mutual gravitational attraction. This gas had little internal motion, so the stars were formed with very small relative velocities. Correspondingly, although the physical volume $\mathcal{V}_x$ occupied by the galaxy's $N$ stars was initially somewhat larger than its volume today, its stars' kinetic-theory momentum-space volume $\mathcal{V}_p$ was far smaller than today. Translated into the language of an ensemble of such galaxies, the initial coordinate-space volume $\int d^{3N}x \sim {\mathcal{V}_x}^N$ occupied by the ensemble's galaxies was moderately larger than today, while its momentum-space volume $\int d^{3N}p \sim {\mathcal{V}_p}^N$ was far smaller. The phase-space volume must therefore have increased considerably during the galaxy formation—with the increase due to a big increase in the relative momenta of neighboring stars. For this to occur, it was necessary that the stars changed their relative energies during the collapse, and this required a time-dependent Hamiltonian. In other words the gravitational potential $\Phi$ felt by the stars must have varied rapidly so that the individual stellar energies would vary according to

$$\frac{dE}{dt} = \frac{\partial H}{\partial t} = m\frac{\partial \Phi}{\partial t} \ . \tag{3.95}$$

The largest changes of energy occurred when the galaxy was collapsing dynamically, so that the potential changed significantly on the time it took the stars to cross the galaxy, $\tau_{\rm int} \sim R/v$. Numerical simulations show that this energy transfer was highly efficient. This process is known as *violent relaxation*. Although violent relaxation could create the observed stellar distribution functions, it was not by itself a means of diluting the phase space density, since Liouville's theorem still applied.

The mechanism that changed the phase space density was *phase mixing* and *coarse-graining* (Sec. 3.6 above). During the initial collapse, when the stars are newly formed, they can be thought of as following highly perturbed radial orbits. The orbits of nearby stars are somewhat similar, though not identical. This means that small elements of phase space became highly contorted as the stars moved along their paths in phase space. Let us make a simple model of this process by assuming that the individual stars initially populated a fraction $\sim f$ of the final volume of the occupied available phase space. After one dynamical timescale $\tau_{int} \sim R/v$, we expect that this small volume will be deformed into a convoluted surface that folds back upon itself once or twice while still occupying the same total volume of phase space. After $n$ such dynamical times, there will be $\sim 2^n$ such folds (cf. Fig. 3.2 (b) above). After $n \sim -\log_2 f$ dynamical timescales the spacing between folds becomes comparable with the characteristic thickness of this convoluted surface and it is no longer practical to distinguish the original distribution function. Coarse-graining has been accomplished for all practical purposes; only a pathological physicist would resist it and insist on trying to continue keeping track of which contorted phase space regions have the original high density and which do not. For a galaxy we might expect that $f \sim 10^{-3}$ and so this natural coarse-graining can occur in a time $\sim 10\,\tau_{int} \sim 10^9$ yr., which is 10 times shorter than the present age of galaxies. It need therefore not be a surprise that the galaxy we know best, our own Milky Way, exhibits no obvious vestigial trace of its initial high-density (low phase-space-volume) distribution function.

## Bibliographic Note

Statistical Mechanics has inspired a variety of readable and innovative texts. The classic treatment is Tolman (1938). Amongst more modern approaches that deal in much greater depth with the topics covered by this chapter are Lifshitz & Pitaevski (1980), Pathria (1972), Reichl (1980), Riedi (1988) and Reif (1965). A highly individual and advanced treatment, emphasizing quantum statistical mechanics is Feynman (1972). A particularly readable account in which statistical mechanics is used heavily to describe the properties of solids, liquids and gases is Goodstein (1985).

## Bibliography


Feynman, R. P. 1972. *Statistical Mechanics*,Reading: Benjamin-Cummings

Frautschi, Steven, 1982. *Entropy in an expanding universe*, Science 217, 593–599.

Goldstein, Herbert, 1980. *Classical Mechanics. (Second Edition)*, New York: Addison Wesley

Goodstein, D. L., 1985. *States of Matter*, New York: Dover Publications

Hawking, Stephen W. 1976. *Particle creation by black holes*, Communications in Mathematical Physics, 43, 199–220.

Hawking, S. W. 1989. *Entropy decrease in contracting universe*

Kittel, Charles 1958. *Elementary Statistical Physics*, New York: Wiley.

Lifshitz, E. M. & Pitaevskii, L. P. 1980. *Statistical Physics. Part 1 (Third Edition)*,Oxford: Pergamon.

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

Mathews, Jon & Walker, Robert L. 1964. *Mathematical Methods of Physics*, New York: Benjamin.

Ogorodnikov, K. F. 1965. *Dynamics of Stellar Systems*, Oxford: Pergamon.

Pathria, R. K. 1972. *Statistical Mechanics*, Oxford: Pergamon.

Penrose, R. 1989. *The Emperor's New Mind*, Oxford: Oxford University Press.

Reichl, L. E. 1980. *A Modern Course in Statistical Physics*, London: Arnold.

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

Riedi, P.C. 1988. *Thermal Physics*Oxford: Oxford Science Publications.

Thorne, Kip S., Price, Richard H., & Macdonald, Douglas M., 1986. *Black Holes: The Membrane Paradigm*, New Haven: Yale University Press.

Tolman, R. C. 1938. *The Principles of Statistical Mechanics*, Oxford: Oxford University Press.