

Chapter 2. Principles of Physical Statistics

This chapter is the keystone of this course. It starts with a brief discussion of such basic notions of statistical physics as statistical ensembles, probability, and ergodicity. Then the so-called microcanonical distribution postulate is formulated, simultaneously with the statistical definition of entropy. This basis enables a ready derivation of the famous Gibbs (“canonical”) distribution – the most frequently used tool of statistical physics. Then we will discuss one more, “grand canonical” distribution, which is more convenient for some tasks. In particular, it is immediately used for the derivation of the most important Boltzmann, Fermi-Dirac, and Bose-Einstein statistics of independent particles, which will be repeatedly utilized in the following chapters.

2.1. Statistical ensembles and probability

As was already discussed in Sec. 1.1, statistical physics deals with situations when either unknown initial conditions, or system’s complexity, or the laws of its motion (as in the case of quantum mechanics) do not allow a definite prediction of measurement results. The main formalism for the analysis of such systems is the probability theory, so let me start with a very brief review of its basic concepts, using an informal “physical” language – less rigorous but (hopefully) more transparent than standard mathematical treatments,¹ and quite sufficient for our purposes.

Consider $N \gg 1$ independent similar experiments carried out with *apparently* similar systems (i.e. systems with identical *macroscopic* parameters such as volume, pressure, etc.), but still giving, by any of the reasons listed above, different results of measurements. Such a collection of experiments, together with a fixed method of result processing, is a good example of a *statistical ensemble*. Let us start from the case when each experiment may have M different *discrete* outcomes, and the number of experiments giving these outcomes is N_1, N_2, \dots, N_M , so

$$\sum_{m=1}^M N_m = N. \quad (2.1)$$

The *probability* of each outcome, for the given statistical ensemble, is then defined as

$$W_m \equiv \lim_{N \rightarrow \infty} \frac{N_m}{N}. \quad (2.2)$$

Probability:
definition

Though this definition is so close to our everyday experience that it is almost self-evident, a few remarks may still be relevant.

First, the probabilities W_m depend on the exact statistical ensemble they are defined for, notably including the method of result processing. As the simplest example, consider throwing the standard cubic-shaped dice many times. For the ensemble of all thrown and counted dice, the probability of each outcome (say, “1”) is 1/6. However, nothing prevents us from defining another statistical ensemble of dice-throwing experiments in which all outcomes “1” are discounted. Evidently, the probability of

¹ For the reader interested in a more rigorous approach, I can recommend, for example, Chapter 18 of the famous handbook by G. Korn and T. Korn – see MA Sec. 16(ii).

finding the outcome “1” in this modified (but legitimate) ensemble is 0, while for all other five outcomes (“2” to “6”), it is 1/5 rather than 1/6.

Second, a statistical ensemble does not necessarily require N similar physical systems, e.g., N distinct dice. It is intuitively clear that tossing the same die N times constitutes an ensemble with similar statistical properties. More generally, a set of N experiments with the same system gives a statistical ensemble equivalent to the set of experiments with N different systems, provided that the experiments are kept independent, i.e. that outcomes of past experiments do not affect the experiments to follow. Moreover, for many physical systems of interest, no special preparation of each new experiment is necessary, and N experiments separated by sufficiently long time intervals, form a “good” statistical ensemble – the property called *ergodicity*.²

Third, the reference to infinite N in Eq. (2) does not strip the notion of probability of its practical relevance. Indeed, it is easy to prove (see Chapter 5) that, at very general conditions, at finite but sufficiently large N , the numbers N_m are approaching their *average* (or *expectation*) values³

$$\langle N_m \rangle \equiv W_m N, \quad (2.3)$$

with the relative deviations decreasing as $\sim 1/\langle N_m \rangle^{1/2}$, i.e. as $1/N^{1/2}$.

Now let me list those properties of probabilities that we will immediately need. First, dividing both sides of Eq. (1) by N and following the limit $N \rightarrow \infty$, we get the well-known *normalization condition*

$$\sum_{m=1}^M W_m = 1; \quad (2.4)$$

just remember that it is true only if each experiment definitely yields one of the outcomes N_1, N_2, \dots, N_M .

Second, if we have an additive function of the results,

$$f = \frac{1}{N} \sum_{m=1}^M N_m f_m, \quad (2.5)$$

where f_m are some definite (deterministic) coefficients, the *statistical average* (also called the *expectation value*) of this function is naturally defined as

² The most popular counter-examples are provided by some energy-conserving systems. Consider, for example, a system of particles placed in a potential that is a quadratic-polynomial function of its coordinates. The theory of oscillations tells us (see, e.g., CM Sec. 6.2) that this system is equivalent to a set of non-interacting harmonic oscillators. Each of these oscillators conserves its own initial energy E_j forever, so the statistics of N measurements of one such system may differ from that of N different systems with a random distribution of E_j , even if the total energy of the system, $E = \sum_j E_j$, is the same. Such non-ergodicity, however, is a rather feeble phenomenon and is readily destroyed by any of many mechanisms, such as weak interaction with the environment (leading, in particular, to oscillation damping), potential anharmonicity (see, e.g., CM Chapter 5), and chaos (CM Chapter 9), all of them growing fast with the number of particles in the system, i.e. the number of its degrees of freedom. This is why an overwhelming part of real-life systems are ergodic; for readers interested in non-ergodic exotics, I can recommend the monograph by V. Arnold and A. Avez, *Ergodic Problems of Classical Mechanics*, Addison-Wesley, 1989.

³ Here (and everywhere in this series) angle brackets $\langle \dots \rangle$ mean averaging over a statistical ensemble, which is generally different from averaging over time – as it will be the case in quite a few examples considered below.

$$\langle f \rangle \equiv \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{m=1}^M \langle N_m \rangle f_m, \quad (2.6)$$

so using Eq. (3) we get

Expectation
value via
probabilities

$$\langle f \rangle = \sum_{m=1}^M W_m f_m. \quad (2.7)$$

Note that Eq. (3) may be considered as the particular form of this general result, when all $f_m = 1$. Eq. (5) with these f_m defines what is sometimes called the *counting function*.

Next, the spectrum of possible experimental outcomes is frequently continuous for all practical purposes. (Think, for example, about the set of positions of the marks left by bullets fired into a target from afar.) The above formulas may be readily generalized to this case; let us start from the simplest situation when all different outcomes may be described by just one continuous scalar variable q – which replaces the discrete index m in Eqs. (1)-(7). The basic relation for this case is the self-evident fact that the probability dW of having an outcome within a very small interval dq surrounding some point q is proportional to the magnitude of that interval:

$$dW = w(q)dq, \quad (2.8)$$

where $w(q)$ is some function of q , which does not depend on dq . This function is called *probability density*. Now all the above formulas may be recast by replacing the probabilities W_m with the products (8), and the summation over m , with the integration over q . In particular, instead of Eq. (4) the normalization condition now becomes

$$\int w(q)dq = 1, \quad (2.9)$$

where the integration should be extended over the whole range of possible values of q . Similarly, instead of the discrete values f_m participating in Eq. (5), it is natural to consider a function $f(q)$. Then instead of Eq. (7), the expectation value of the function may be calculated as

Expectation
value via
probability
density

$$\langle f \rangle = \int w(q)f(q)dq. \quad (2.10)$$

It is also straightforward to generalize these formulas to the case of more variables. For example, the state of a classical particle with three degrees of freedom may be fully described by the probability density w defined in the 6D space of its generalized radius-vector \mathbf{q} and momentum \mathbf{p} . As a result, the expectation value of a function of these variables may be expressed as a 6D integral

$$\langle f \rangle = \int w(\mathbf{q}, \mathbf{p}) f(\mathbf{q}, \mathbf{p}) d^3 q d^3 p. \quad (2.11)$$

Some systems considered in this course consist of components whose quantum properties cannot be ignored, so let us discuss how $\langle f \rangle$ should be calculated in this case. If by f_m we mean measurement results, then Eq. (7) (and its generalizations) remains valid, but since these numbers themselves may be affected by the intrinsic quantum-mechanical uncertainty, it may make sense to have a bit deeper look into this situation. Quantum mechanics tells us⁴ that the most general expression for the expectation value of an observable f in a certain ensemble of macroscopically similar systems is

⁴ See, e.g., QM Sec. 7.1.

$$\langle f \rangle = \sum_{m,m'} W_{mm'} f_{m'm} \equiv \text{Tr}(Wf). \quad (2.12)$$

Here $f_{mm'}$ are the matrix elements of the quantum-mechanical operator \hat{f} corresponding to the observable f , in a full basis of orthonormal states m ,

$$f_{mm'} = \langle m | \hat{f} | m' \rangle, \quad (2.13)$$

while the coefficients $W_{mm'}$ are the elements of the so-called *density matrix* W , which represents, in the same basis, the *density operator* \hat{W} describing properties of this ensemble. Eq. (12) is evidently more general than Eq. (7), and is reduced to it only if the density matrix is diagonal:

$$W_{mm'} = W_m \delta_{mm'} \quad (2.14)$$

(where $\delta_{mm'}$ is the Kronecker symbol), when the diagonal elements W_m play the role of probabilities of the corresponding states.

Thus formally, the largest difference between the quantum and classical description is the presence, in Eq. (12), of the off-diagonal elements of the density matrix. They have the largest values in a *pure* (also called “coherent”) ensemble, in which the state of the system may be described with *state vectors*, e.g., the *ket-vector*

$$|\alpha\rangle = \sum_m \alpha_m |m\rangle, \quad (2.15)$$

where α_m are some (generally, complex) coefficients. In this case, the density matrix elements are merely

$$W_{mm'} = \alpha_m^* \alpha_{m'}, \quad (2.16)$$

so the off-diagonal elements are of the same order as the diagonal elements. For example, in the very important particular case of a *two-level system*, the pure-state density matrix is

$$W = \begin{pmatrix} \alpha_1^* \alpha_1 & \alpha_1^* \alpha_2 \\ \alpha_2^* \alpha_1 & \alpha_2^* \alpha_2 \end{pmatrix}, \quad (2.17)$$

so the product of its off-diagonal components is as large as that of the diagonal components.

In the most important basis of stationary states, i.e. the eigenstates of the system’s time-independent Hamiltonian, the coefficients α_m oscillate in time as⁵

$$\alpha_m(t) = \alpha_m(0) \exp\left\{-i \frac{E_m}{\hbar} t\right\} \equiv |\alpha_m| \exp\left\{-i \frac{E_m}{\hbar} t + i\varphi_m\right\}, \quad (2.18)$$

where E_m are the corresponding eigenenergies, φ_m are constant phase shifts, and \hbar is the Planck constant. This means that while the diagonal terms of the density matrix (16) remain constant, its off-diagonal components are oscillating functions of time:

⁵ Here I use the Schrödinger picture of quantum dynamics, in which the matrix elements $f_{mm'}$ representing quantum-mechanical operators, do not evolve in time. The final results of this discussion do not depend on the particular picture – see, e.g., QM Sec. 4.6.

$$W_{mm'} = \alpha_{m'}^* \alpha_m = |\alpha_m \cdot \alpha_{m'}| \exp\left\{i \frac{E_m - E_{m'}}{\hbar} t\right\} \exp\{i(\varphi_{m'} - \varphi_m)\}. \quad (2.19)$$

Due to the extreme smallness of \hbar on the human scale of things), minuscule random perturbations of eigenenergies are equivalent to substantial random changes of the phase multipliers, so the time average of any off-diagonal matrix element tends to zero. Moreover, even if our statistical ensemble consists of systems with exactly the same E_m , but different values φ_m (which are typically hard to control at the initial preparation of the system), the average values of all $W_{mm'}$ (with $m \neq m'$) vanish again.

This is why, besides some very special cases, typical statistical ensembles of quantum particles are far from being pure, and in most cases (certainly including the thermodynamic equilibrium), a good approximation for their description is given by the opposite limit of the so-called *classical mixture*, in which all off-diagonal matrix elements of the density matrix equal zero, and its diagonal elements W_{mm} are merely the probabilities W_m of the corresponding eigenstates. In this case, for the observables compatible with energy, Eq. (12) is reduced to Eq. (7), with f_m being the eigenvalues of the variable f , so we may base our further discussion on this key relation and its continuous extensions (10)-(11).

2.2. Microcanonical ensemble and distribution

Now we may move to the now-standard approach to statistical mechanics, based on the three statistical ensembles introduced in the 1870s by Josiah Willard Gibbs.⁶ The most basic of them is the so-called *microcanonical statistical ensemble*⁷ defined as a set of macroscopically similar closed (isolated) systems with *virtually* the same total energy E . Since in quantum mechanics the energy of a closed system is quantized, in order to make the forthcoming discussion suitable for quantum systems as well, it is convenient to include in the ensemble all systems with energies E_m within a relatively narrow interval $\Delta E \ll E$ (see Fig. 1) that is nevertheless much larger than the average distance δE between the energy levels, so that the number M of different quantum states within the interval ΔE is large, $M \gg 1$. Such choice of ΔE is only possible if $\delta E \ll E$; however, the reader should not worry too much about this condition, because the most important applications of the microcanonical ensemble are for very large systems (and/or very high energies) when the energy spectrum is very dense.⁸

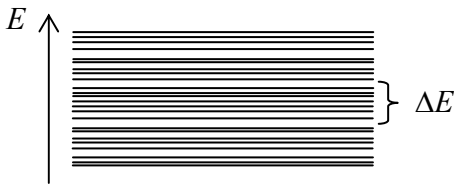


Fig. 2.1. A *very* schematic image of the microcanonical ensemble. (Actually, the ensemble deals with quantum states rather than energy levels. An energy level may be degenerate, i.e. correspond to several states.)

⁶ Personally, I believe that the genius of J. Gibbs, praised by Albert Einstein as the “greatest mind in American history”, is still insufficiently appreciated, and agree with R. Millikan that Gibbs “did for statistical mechanics and thermodynamics what [...] Maxwell did for electrodynamics”.

⁷ The terms “microcanonical”, as well as “canonical” (see Sec. 4 below) are apparently due to Gibbs and I was unable to find out his motivation for the former name. (“Canonical” in the sense of “standard” or “common” is quite appropriate, but why “micro”? Perhaps to reflect the smallness of ΔE ?)

⁸ Formally, the main result of this section, Eq. (20), is valid for any M (including $M = 1$); it is just less informative for small M (and trivial for $M = 1$).

This ensemble serves as the basis for the formulation of the postulate which is most frequently called the *microcanonical distribution* (or, more adequately, “the main statistical postulate” or “the main statistical hypothesis”): *in the thermodynamic equilibrium of a microcanonical ensemble, all its states have equal probabilities,*

$$W_m = \frac{1}{M} = \text{const.} \quad (2.20)$$

Micro-canonical distribution

Though in some constructs of statistical mechanics, this equality is derived from other axioms, which look more plausible to their authors, I believe that Eq. (20) may be taken as the starting point of the statistical physics, supported “just” by the compliance of all its corollaries with experimental observations.

Note that the postulate (20) is closely related to the macroscopic irreversibility of the systems that are microscopically virtually reversible (closed): if such a system was initially in a certain state, its time evolution with even minuscule interactions with the environment (which is necessary for reaching the thermodynamic equilibrium) eventually leads to the uniform distribution of its probability among all states with essentially the same energy. Each of these states is not “better” than the initial one; rather, in a macroscopic system, there are just so many of these states that the chance to find the system in the initial state is practically nil – again, think about the ink drop diffusion into a glass of water.⁹

Now let us find a suitable definition of entropy S of a microcanonical ensemble’s member – for now, in the thermodynamic equilibrium only. This was done in 1877 by another giant of statistical physics, Ludwig Eduard Boltzmann – on the basis of the prior work by James Clerk Maxwell on the kinetic theory of gases – see Sec. 3.1 below. In present-day terminology, since S is a measure of *disorder*, it should be related to the amount of information¹⁰ *lost* when the system went irreversibly from the full order to its current state – in equilibrium, to the full disorder, i.e. from one definite state to the microcanonical distribution (20). In an even more convenient formulation, this is the amount of information necessary to find the exact state of a certain system in a microcanonical ensemble.

In the information theory, the amount of information necessary to make a definite choice between two options with equal probabilities (Fig. 2a) is defined as

$$I(2) \equiv \log_2 2 = 1. \quad (2.21)$$

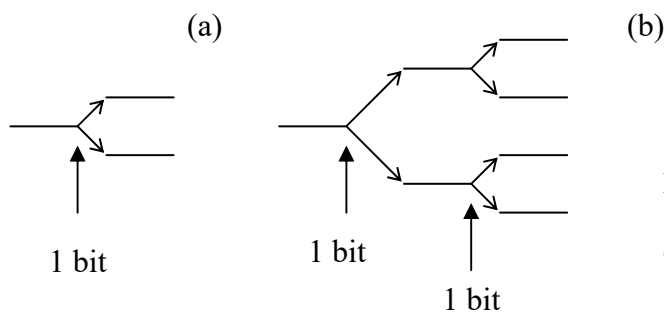


Fig. 2.2. “Logarithmic trees” of binary decisions for choosing between (a) $M = 2$, and (b) $M = 4$ opportunities with equal probabilities.

⁹ Though I have to move on, let me note that the microcanonical distribution (20) is a very nontrivial postulate, and my advice to the reader is to find some time to give additional thought to this keystone of the whole building of statistical mechanics.

¹⁰ I will rely on the reader’s common sense and intuitive understanding of what information is, because even in the formal information theory, this notion is essentially postulated – see, e.g., the wonderfully clear short textbook by J. Pierce, *An Introduction to Information Theory*, Dover, 1980.

This unit of information is called a *bit*. Now, if we need to make a choice between four equally probable opportunities, it can be made in two similar steps (Fig. 2b), each requiring one bit of information, so the total amount of information necessary for the choice is

$$I(4) = 2I(2) = 2 \equiv \log_2 4. \quad (2.22)$$

An obvious extension of this process to the choice between $M = 2^m$ states gives

$$I(M) = mI(2) = m \equiv \log_2 M. \quad (2.23)$$

This measure, if extended naturally to any integer M , is quite suitable for the definition of entropy at equilibrium, with the only difference that, following tradition, the binary logarithm is replaced with the natural one:¹¹

$$S \equiv \ln M. \quad (2.24a)$$

Using Eq. (20), we may recast this definition in its most frequently used form

Entropy
in
equilibrium

$$S = \ln \frac{1}{W_m} \equiv -\ln W_m. \quad (2.24b)$$

(Again, please note that Eqs. (24) are valid in thermodynamic equilibrium only!)

Note that Eq. (24) satisfies the major properties of the entropy discussed in thermodynamics. First, it is a unique characteristic of the disorder. Indeed, according to Eq. (20), M (at fixed ΔE) is the only possible measure characterizing the microcanonical distribution, and so is its unique function $\ln M$. This function also satisfies another thermodynamic requirement to the entropy, of being an extensive variable. Indeed, for several independent systems, the joint probability of a certain state is just a product of the partial probabilities, and hence, according to Eq. (24), their entropies just add up.

Now let us see whether Eqs. (20) and (24) are compatible with the 2nd law of thermodynamics. For that, we need to generalize Eq. (24) for S to an arbitrary state of the system (generally, out of thermodynamic equilibrium), with an arbitrary set of state probabilities W_m . Let us first recognize that M in Eq. (24) is just the number of possible ways to commit a particular system to a certain state m ($m = 1, 2, \dots, M$), in a statistical ensemble where each state is equally probable. Now let us consider a more general ensemble, still consisting of a large number $N \gg 1$ of similar systems, but with a certain number $N_m = W_m N \gg 1$ of systems in each of M states, with the factors W_m not necessarily equal. In this case, the evident generalization of Eq. (24) is that the entropy S_N of the whole ensemble is

$$S_N \equiv \ln M(N_1, N_2, \dots), \quad (2.25)$$

where $M(N_1, N_2, \dots)$ is the number of ways to commit a particular system to a certain state m while keeping all numbers N_m fixed. This number $M(N_1, N_2, \dots)$ is clearly equal to the number of ways to distribute N distinct balls between M different boxes, with the fixed number N_m of balls in each box, but

¹¹ This is of course just the change of a constant factor: $S(M) = \ln M = \ln 2 \times \log_2 M = \ln 2 \times I(M) \approx 0.693 I(M)$. A review of Chapter 1 shows that nothing in thermodynamics prevents us from choosing such a constant coefficient arbitrarily, with the corresponding change of the temperature scale – see Eq. (1.9). In particular, in the SI units, where Eq. (24b) becomes $S = -k_B \ln W_m$, one bit of information corresponds to the entropy change $\Delta S = k_B \ln 2 \approx 0.693 k_B \approx 0.965 \times 10^{-23}$ J/K. (The formula “ $S = k \log W$ ” is engraved on L. Boltzmann’s tombstone in Vienna.)

in no particular order within it. Comparing this description with the definition of the so-called *multinomial coefficients*,¹² we get

$$M(N_1, N_2, \dots) = {}^N C_{N_1, N_2, \dots, N_M} \equiv \frac{N!}{N_1! N_2! \dots N_M!}, \quad \text{with } N = \sum_{m=1}^M N_m. \quad (2.26)$$

To simplify the resulting expression for S_N , we can use the famous *Stirling formula*, in its crudest, de Moivre's form,¹³ whose accuracy is suitable for most purposes of statistical physics:

$$\ln(N!)_{N \rightarrow \infty} \rightarrow N(\ln N - 1). \quad (2.27)$$

When applied to our current problem, this formula gives the following average entropy per system,¹⁴

$$\begin{aligned} S &\equiv \frac{S_N}{N} = \frac{1}{N} \left[\ln(N!) - \sum_{m=1}^M \ln(N_m!) \right]_{N_m \rightarrow \infty} \rightarrow \frac{1}{N} \left[N(\ln N - 1) - \sum_{m=1}^M N_m (\ln N_m - 1) \right] \\ &\equiv - \sum_{m=1}^M \frac{N_m}{N} \ln \frac{N_m}{N}, \end{aligned} \quad (2.28)$$

and since this result is only valid in the limit $N_m \rightarrow \infty$ anyway, we may use Eq. (2) to represent it as

$$S = - \sum_{m=1}^M W_m \ln W_m = \sum_{m=1}^M W_m \ln \frac{1}{W_m}. \quad (2.29)$$

Entropy
out of
equilibrium

This extremely important result¹⁵ may be interpreted as the average of the entropy values given by Eq. (24), weighed with specific probabilities W_m per the general formula (7).¹⁶

Now let us find what distribution of probabilities W_m provides the largest value of the entropy (29). The answer is almost evident from a good glance at Eq. (29). For example, if for a subgroup of $M' \leq M$ states, the coefficients W_m are constant and equal to $1/M'$, so $W_m = 0$ for all other states, all M' non-zero terms in the sum (29) are equal to each other, so

$$S = M' \frac{1}{M'} \ln M' \equiv \ln M', \quad (2.30)$$

and the closer M' to its maximum value M the larger S . Hence, the maximum of S is reached at the uniform distribution given by Eq. (24).

¹² See, e.g., MA Eq. (2.3). Despite the intimidating name, Eq. (26) may be very simply derived. Indeed, $N!$ is just the number of all possible permutations of N balls, i.e. of the ways to place them in *certain* positions – say, inside M boxes. Now take into account that the particular order of the balls in each box is not important, that number should be divided by all numbers $N_m!$ of possible permutations of balls within each box – that's it!

¹³ See, e.g., MA Eq. (2.10).

¹⁴ Strictly speaking, I should use the notation $\langle S \rangle$ here. However, following the style accepted in thermodynamics, I will drop the averaging signs until we will really need them to avoid confusion. Again, this shorthand is not too bad because the relative fluctuations of entropy (as those of any macroscopic variable) are very small at $N \gg 1$.

¹⁵ With the replacement of $\ln W_m$ with $\log_2 W_m$ (i.e. division of both sides by $\ln 2$), Eq. (29) becomes the famous *Shannon* (or “Boltzmann-Shannon”) *formula* for the average information I per symbol in a long communication string using M different symbols, with probability W_m each.

¹⁶ In some textbooks, this interpretation is even accepted as the derivation of Eq. (29); however, it is evidently less rigorous than the one outlined above.

In order to prove this important fact more strictly, let us find the maximum of the function given by Eq. (29). If its arguments W_1, W_2, \dots, W_M were completely independent, this could be done by finding the point (in the M -dimensional space of the coefficients W_m) where all partial derivatives $\partial S/\partial W_m$ equal zero. However, since the probabilities are constrained by condition (4), the differentiation has to be carried out more carefully, taking into account this interdependence:

$$\left[\frac{\partial}{\partial W_m} S(W_1, W_2, \dots) \right]_{\text{cond}} = \frac{\partial S}{\partial W_m} + \sum_{m' \neq m} \frac{\partial S}{\partial W_{m'}} \frac{\partial W_{m'}}{\partial W_m}. \quad (2.31)$$

At the maximum of the function S , such expressions should be equal to zero for all m . This condition yields $\partial S/\partial W_m = \lambda$, where the so-called *Lagrange multiplier* λ is independent of m . Indeed, at such point Eq. (31) becomes

$$\left[\frac{\partial}{\partial W_m} S(W_1, W_2, \dots) \right]_{\text{cond}} = \lambda + \sum_{m' \neq m} \lambda \frac{\partial W_{m'}}{\partial W_m} \equiv \lambda \left(\frac{\partial W_m}{\partial W_m} + \sum_{m' \neq m} \frac{\partial W_{m'}}{\partial W_m} \right) \equiv \lambda \frac{\partial}{\partial W_m} (1) = 0. \quad (2.32)$$

For our particular expression (29), the condition $\partial S/\partial W_m = \lambda$ yields

$$\frac{\partial S}{\partial W_m} \equiv \frac{d}{dW_m} [-W_m \ln W_m] \equiv -\ln W_m - 1 = \lambda. \quad (2.33)$$

The last equality holds for all m (and hence the entropy reaches its maximum value) only if W_m is independent of m . Thus the entropy (29) indeed reaches its maximum value (24) at equilibrium.

To summarize, we see that the statistical definition (24) of entropy does fit all the requirements imposed on this variable by thermodynamics. In particular, we have been able to *prove* the 2nd law of thermodynamics using that definition together with the fundamental postulate (20).

Now let me discuss one possible point of discomfort with that definition: the values of M , and hence W_m , depend on the accepted energy interval ΔE of the microcanonical ensemble, for whose choice no exact guidance is offered. However, if the interval ΔE contains many states, $M \gg 1$, as was assumed before, then with a very small relative error (vanishing in the limit $M \rightarrow \infty$), M may be represented as

$$M = g(E)\Delta E, \quad (2.34)$$

where $g(E)$ is the *density of states* of the system:

$$g(E) \equiv \frac{d\Sigma(E)}{dE}, \quad (2.35)$$

$\Sigma(E)$ being the total number of states with energies *below* E . (Note that the average interval δE between energy levels, mentioned at the beginning of this section, is just $\Delta E/M = 1/g(E)$.) Plugging Eq. (34) into Eq. (24), we get

$$S = \ln M = \ln g(E) + \ln \Delta E, \quad (2.36)$$

so the only effect of a particular choice of ΔE is an offset of the entropy by a constant, and in Chapter 1 we have seen that such a constant offset does not affect any measurable quantity. Of course, Eq. (34), and hence Eq. (36) are only precise in the limit when the density of states $g(E)$ is so large that the range available for the appropriate choice of ΔE ,

$$g^{-1}(E) \ll \Delta E \ll E, \quad (2.37)$$

is sufficiently broad: $g(E)E = E/\delta E \gg 1$.

In order to get some gut feeling of the functions $g(E)$ and $S(E)$ and the feasibility of the condition (37), and also to see whether the microcanonical distribution may be directly used for calculations of thermodynamic variables in particular systems, let us apply it to a microcanonical ensemble of many sets of $N \gg 1$ independent, similar harmonic oscillators with frequency ω . (Please note that the requirement of a virtually fixed energy is applied, in this case, to the *total* energy E_N of each set of oscillators, rather to energy E of a *single* oscillator – which may be virtually arbitrary though certainly much less than $E_N \sim NE \gg E$.) Basic quantum mechanics tells us¹⁷ that the eigenenergies of such an oscillator form a discrete, equidistant spectrum:

$$E_m = \hbar\omega \left(m + \frac{1}{2} \right), \quad \text{where } m = 0, 1, 2, \dots \quad (2.38)$$

If ω is kept constant, the ground-state energy $\hbar\omega/2$ does not contribute to any thermodynamic properties of the system,¹⁸ so for the sake of simplicity we may take that point as the energy origin, and replace Eq. (38) with $E_m = m\hbar\omega$. Let us carry out an approximate analysis of the system for the case when its average energy per oscillator,

$$E = \frac{E_N}{N}, \quad (2.39)$$

is much larger than the energy quantum $\hbar\omega$.

For one oscillator, the number of states with energy ε_1 below a certain value $E_1 \gg \hbar\omega$ is evidently $\Sigma(E_1) \approx E_1/\hbar\omega \equiv (E_1/\hbar\omega)/1!$ (Fig. 3a). For two oscillators, all possible values of the total energy ($\varepsilon_1 + \varepsilon_2$) below some level E_2 correspond to the points of a 2D square grid within the right triangle shown in Fig. 3b, giving $\Sigma(E_2) \approx (1/2)(E_2/\hbar\omega)^2 \equiv (E_2/\hbar\omega)^2/2!$. For three oscillators, the possible values of the total energy ($\varepsilon_1 + \varepsilon_2 + \varepsilon_3$) correspond to those points of the 3D cubic grid, that fit inside the right pyramid shown in Fig. 3c, giving $\Sigma(E_3) \approx (1/3)[(1/2)(E_3/\hbar\omega)^3] \equiv (E_3/\hbar\omega)^3/3!$, etc.

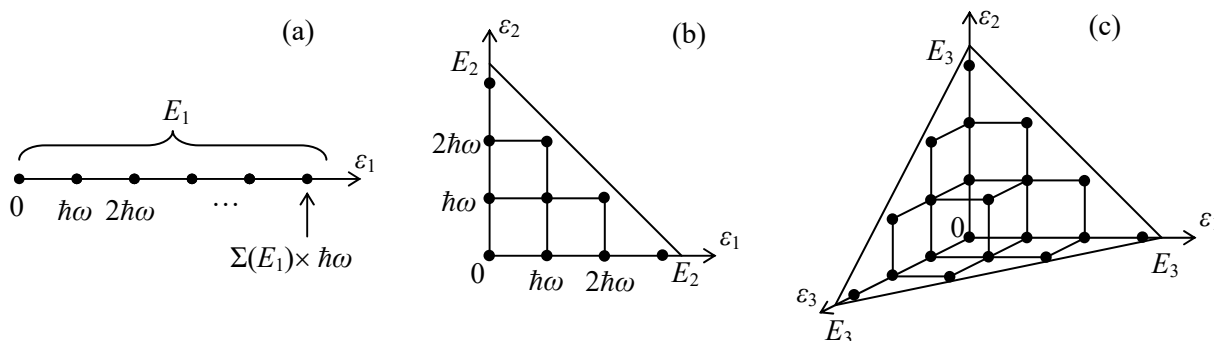


Fig. 2.3. Calculating functions $\Sigma(E_N)$ for systems of (a) one, (b) two, and (c) three harmonic oscillators.

¹⁷ See, e.g., QM Secs. 2.9 and 5.4.

¹⁸ Let me hope that the reader knows that the ground-state energy *is* experimentally measurable – for example, using the famous *Casimir effect* – see, e.g., QM Sec. 9.1. (In Sec. 5.5 below I will briefly discuss another method of experimental observation of that energy.)

An evident generalization of these formulas to arbitrary N gives the number of states¹⁹

$$\Sigma(E_N) \approx \frac{1}{N!} \left(\frac{E_N}{\hbar\omega} \right)^N. \quad (2.40)$$

Differentiating this expression over the energy, we get

$$g(E_N) \equiv \frac{d\Sigma(E_N)}{dE_N} = \frac{1}{(N-1)!} \frac{E_N^{N-1}}{(\hbar\omega)^N}, \quad (2.41)$$

so

$$S_N(E_N) = \ln g(E_N) + \text{const} = -\ln[(N-1)!] + (N-1) \ln E_N - N \ln(\hbar\omega) + \text{const}. \quad (2.42)$$

For $N \gg 1$, we may ignore the difference between N and $(N-1)$ in both instances, and use the Stirling formula (27) to simplify this result as

$$S_N(E) - \text{const} \approx N \left(\ln \frac{E_N}{N\hbar\omega} + 1 \right) \approx N \left(\ln \frac{E}{\hbar\omega} \right) \equiv \ln \left[\left(\frac{E}{\hbar\omega} \right)^N \right]. \quad (2.43)$$

(The second step is only valid at very high $E/\hbar\omega$ ratios when the logarithm in Eq. (43) is substantially larger than 1.) Returning for a second to the density of states, we see that in the limit $N \rightarrow \infty$, it is exponentially large:

$$g(E_N) = e^{S_N} \approx \left(\frac{E}{\hbar\omega} \right)^N, \quad (2.44)$$

so the conditions (37) may be indeed satisfied within a very broad range of ΔE .

Now we can use Eq. (43) to find all thermodynamic properties of the system, though only in the limit $E \gg \hbar\omega$. Indeed, according to thermodynamics, if the system's volume and the number of particles in it are fixed, the derivative dS/dE is nothing else than the reciprocal temperature in thermal equilibrium – see Eq. (1.9). In our current case, we imply that the harmonic oscillators are distinct, for example by their spatial positions. Hence, even if we can speak of some volume of the system, it is certainly fixed.²⁰ Differentiating Eq. (43) over energy E , we get

$$\frac{1}{T} \equiv \frac{dS_N}{dE_N} = \frac{N}{E_N} = \frac{1}{E}. \quad (2.45)$$

Classical
oscillator:
average
energy

Reading this result backward, we see that the average energy E of a harmonic oscillator equals T (i.e. $k_B T_K$ is SI units). At this point, the first-time student of thermodynamics should be very much relieved to see that the counter-intuitive thermodynamic definition (1.9) of temperature does indeed correspond to what we all have known about this notion from our kindergarten physics courses.

The result (45) may be readily generalized. Indeed, in quantum mechanics, a harmonic oscillator with eigenfrequency ω may be described by the Hamiltonian operator

¹⁹ The coefficient $1/N!$ in this formula has the geometrical meaning of the (hyper)volume of the N -dimensional right pyramid with unit sides.

²⁰ For the same reason, the notion of pressure P in such a system is not clearly defined, and neither are any thermodynamic potentials but E and F .

$$\hat{H} = \frac{\hat{p}^2}{2m} + \frac{\kappa\hat{q}^2}{2}, \quad (2.46)$$

where q is some generalized coordinate, p is the corresponding generalized momentum, m is the oscillator's mass,²¹ and κ is its spring constant, so $\omega = (\kappa/m)^{1/2}$. Since in the thermodynamic equilibrium the density matrix is always diagonal in the basis of stationary states m (see Sec. 1 above), the quantum-mechanical averages of the kinetic and potential energies may be found from Eq. (7):

$$\left\langle \frac{p^2}{2m} \right\rangle = \sum_{m=0}^{\infty} W_m \langle m | \frac{\hat{p}^2}{2m} | m \rangle, \quad \left\langle \frac{\kappa q^2}{2} \right\rangle = \sum_{m=0}^{\infty} W_m \langle m | \frac{\kappa\hat{q}^2}{2} | m \rangle, \quad (2.47)$$

where W_m is the probability to occupy the m^{th} energy level, while bra- and ket-vectors describe the stationary state corresponding to that level.²² However, both classical and quantum mechanics teach us that for any m , the bra-ket expressions under the sums in Eqs. (47), which represent the average kinetic and mechanical energies of the oscillator on its m^{th} energy level, are equal to each other, and hence each of them is equal to $E_m/2$. Hence, even though we do not know the exact probability distribution W_m yet (it will be calculated in Sec. 5 below), we may conclude that in the “classical limit” $T \gg \hbar\omega$,

$$\boxed{\left\langle \frac{p^2}{2m} \right\rangle = \left\langle \frac{\kappa q^2}{2} \right\rangle = \frac{T}{2}}. \quad (2.48) \quad \text{Equipartition theorem}$$

Now let us consider a system with an arbitrary number of degrees of freedom, described by a more general Hamiltonian:²³

$$\hat{H} = \sum_j \hat{H}_j, \quad \text{with } \hat{H}_j = \frac{\hat{p}_j^2}{2m_j} + \frac{\kappa_j \hat{q}_j^2}{2}, \quad (2.49)$$

with (generally, different) frequencies $\omega_j = (\kappa_j/m_j)^{1/2}$. Since the “modes” (effective harmonic oscillators) contributing to this Hamiltonian, are independent, the result (48) is valid for each of the modes. This is the famous *equipartition theorem*: at thermal equilibrium at temperature $T \gg \hbar\omega_j$, the average energy of each so-called *half-degree of freedom* (which is defined as any variable, either p_j or q_j , giving a quadratic contribution to the system's Hamiltonian), is equal to $T/2$.²⁴ In particular, for each of three Cartesian component contributions to the kinetic energy of a free-moving particle, this theorem is valid for *any* temperature, because such components may be considered as 1D harmonic oscillators with vanishing potential energy, i.e. $\omega_j = 0$, so condition $T \gg \hbar\omega_j$ is fulfilled at any temperature.

²¹ I am using this fancy font for the mass to avoid any chance of its confusion with the state number.

²² Note again that while we have committed the energy E_N of N oscillators to be fixed (to apply Eq. (36), valid only for a microcanonical ensemble at thermodynamic equilibrium), the single oscillator's energy E in our analysis may be arbitrary – within the very broad limits $\hbar\omega \ll E \leq E_N \sim NT$.

²³ As a reminder, the Hamiltonian of any system whose classical Lagrangian function is an arbitrary quadratic form of its generalized coordinates and the corresponding generalized velocities may be brought to the form (49) by an appropriate choice of “normal coordinates” q_j which are certain linear combinations of the original coordinates – see, e.g., CM Sec. 6.2.

²⁴ This also means that in the classical limit, the heat capacity of a system is equal to one-half of the number of its half-degrees of freedom (in the SI units, multiplied by k_B).

I believe that this case study of harmonic oscillator systems was a fair illustration of both the strengths and the weaknesses of the microcanonical ensemble approach.²⁵ On one hand, we could readily calculate virtually everything we wanted in the classical limit $T \gg \hbar\omega$, but calculations for an arbitrary $T \sim \hbar\omega$, though possible, would be rather unpleasant because for that, all vertical steps of the function $\Sigma(E_N)$ have to be carefully counted. In Sec. 4 below, we will see that other statistical ensembles are much more convenient for such calculations.

Let me conclude this section with a short notice on deterministic classical systems with just a few degrees of freedom (and even simpler mathematical objects called “maps”) that may exhibit essentially disordered behavior, called *deterministic chaos*.²⁶ Such chaotic system may be approximately characterized by an entropy defined similarly to Eq. (29), where W_m are the probabilities to find it in different small regions of phase space, at well-separated small time intervals. On the other hand, one can use an expression slightly more general than Eq. (29) to define the so-called *Kolmogorov* (or “Kolmogorov-Sinai”) entropy K that characterizes the speed of loss of the information about the initial state of the system, and hence what is called the “chaos depth”. In the definition of K , the sum over m is replaced with the summation over all possible permutations $\{m\} = m_0, m_1, \dots, m_{N-1}$ of small space regions, and W_m is replaced with $W_{\{m\}}$, the probability of finding the system in the corresponding regions m at time moment t_m , with $t_m = m\tau$, in the limit $\tau \rightarrow 0$, with $N\tau = \text{const}$. For chaos in the simplest objects, 1D maps, K is equal to the Lyapunov exponent $\lambda > 0$.²⁷ For systems of higher dimensionality, which are characterized by several Lyapunov exponents λ , the Kolmogorov entropy is equal to the phase-space average of the sum of all positive λ . These facts provide a much more practicable way of (typically, numerical) calculation of the Kolmogorov entropy than the direct use of its definition.²⁸

2.3. Maxwell’s Demon, information, and computation

Before proceeding to other statistical distributions, I would like to make a detour to address one more popular concern about Eq. (24) – the direct relation between entropy and information. Some physicists are still uneasy with entropy being nothing else than the (deficit of) information, though, to the best of my knowledge, nobody has yet been able to suggest any experimentally verifiable difference between these two notions. Let me give one example of their direct relationship.²⁹ Consider a cylinder containing just one molecule (considered as a point particle), and separated into two halves by a movable partition with a door that may be opened and closed at will, at no energy cost – see Fig. 4a. If the door is open and the system is in thermodynamic equilibrium, we do not know on which side of the partition the molecule is. Here the disorder, i.e. the entropy has the largest value, and there is no way to get, from a large ensemble of such systems in equilibrium, any useful mechanical energy.

²⁵ The reader is strongly urged to solve Problem 2, whose task is to do a similar calculation for another key (“two-level”) physical system, and compare the results.

²⁶ See, e.g., CM Chapter 9 and the literature therein.

²⁷ For the definition of λ , see, e.g., CM Eq. (9.9).

²⁸ For more discussion, see, e.g., either Sec. 6.2 of the monograph H. G. Schuster and W. Just, *Deterministic Chaos*, 4th ed., Wiley-VHS, 2005, or the monograph by Arnold and Avez, cited in Sec. 1.

²⁹ This system is frequently called the *Szilard engine*, after L. Szilard who published its detailed theoretical discussion in 1929, but is essentially a straightforward extension of the thought experiment suggested by J. Maxwell as early as 1867.

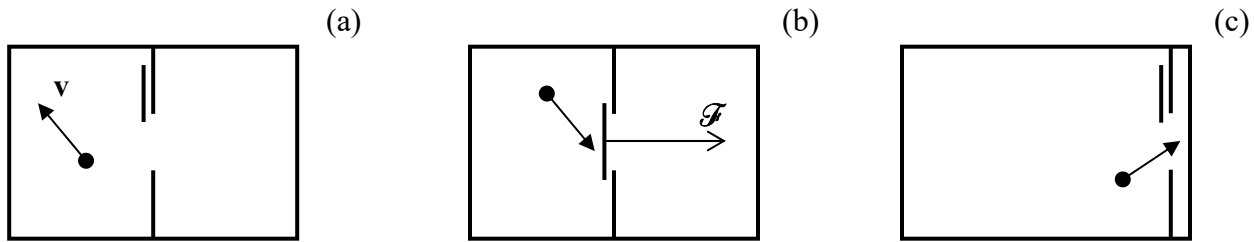


Fig. 2.4. The Szilard engine: a cylinder with a single molecule and a movable partition: (a) before and (b) after closing the door, and (c) after opening the door at the end of the expansion stage.

Now, let us consider that we know (as instructed by, in Lord Kelvin’s formulation, an omniscient *Maxwell’s Demon*) on which side of the partition the molecule is currently located. Then we may close the door, trapping the molecule, so its repeated impacts on the partition create, on average, a pressure force \mathcal{F} directed toward the empty part of the volume (in Fig. 4b, the right one). Now we can get from the molecule some mechanical work, say by allowing the force \mathcal{F} to move the partition to the right, and picking up the resulting mechanical energy by some deterministic (zero-entropy) external mechanism. After the partition has been moved to the right end of the volume, we can open the door again (Fig. 4c), equalizing the molecule’s average pressure on both sides of the partition, and then slowly move the partition back to the middle of the volume – without its resistance, i.e. without doing any substantial work. With the continuing help from Maxwell’s Demon, we can repeat the cycle again and again, and hence make the system perform unlimited mechanical work, fed “only” by the molecule’s thermal motion, and the information about its position – thus implementing the perpetual motion machine of the 2nd kind – see Sec. 1.6. The fact that such heat engines do not exist means that getting any *new* information, at a non-zero temperature (i.e. at a substantial thermal agitation of particles) has a non-zero energy cost.

In order to evaluate this cost, let us calculate the maximum work per cycle that can be made by the Szilard engine (Fig. 4), assuming that it is constantly in the thermal equilibrium with a heat bath of temperature T . Formula (21) tells us that the information supplied by the demon (on what exactly half of the volume contains the molecule) is exactly one bit, $I(2) = 1$. According to Eq. (24), this means that by getting this information we are changing the entropy of our system by

$$\Delta S_I = -\ln 2. \quad (2.50)$$

Now, it would be a mistake to plug this (negative) entropy change into Eq. (1.19). First, that relation is only valid for slow, reversible processes. Moreover (and more importantly), this equation, as well as its irreversible version (1.41), is only valid for a fixed statistical ensemble. The change ΔS_I does not belong to this category and may be formally described by the change of the statistical ensemble – from the one consisting of all similar systems (experiments) with an unknown location of the molecule to a new ensemble consisting of the systems with the molecule in its certain (in Fig. 4, left) half.³⁰

Now let us consider a slow expansion of the “gas” after the door had been closed. At this stage, we do not need the Demon’s help any longer (i.e. the statistical ensemble may be fixed), and can indeed use the relation (1.19). At the assumed isothermal conditions ($T = \text{const}$), this relation may be integrated

³⁰ This procedure of the *statistical ensemble re-definition* is the central point of the connection between physics and information theory, and is crucial in particular for any (or rather any meaningful :-)) discussion of measurements in quantum mechanics – see, e.g., QM Secs. 2.5 and 10.1.

over the whole expansion process, getting $\Delta Q = T\Delta S$. At the final position shown in Fig. 4c, the system's entropy should be the same as initially, i.e. before the door had been opened, because we again do not know where in the volume the molecule is. This means that the entropy was replenished, during the reversible expansion, from the heat bath, by $\Delta S = -\Delta S_I = +\ln 2$, so $\Delta Q = T\Delta S = T\ln 2$. Since by the end of the whole cycle, the internal energy E of the system is the same as before, all this heat could have gone into the mechanical energy obtained during the expansion. Thus the maximum obtained work per cycle (i.e. for each obtained information bit) is $T\ln 2$ ($k_B T_K \ln 2$ in the SI units), about 4×10^{-21} Joule at room temperature. This is exactly the energy cost of getting one bit of *new* information about a system at temperature T . The smallness of that amount on the everyday human scale has left the Szilard engine an academic theoretical exercise for almost a century. However, recently several such devices, of various physical nature, were implemented experimentally (with the Demon's role played by an instrument measuring the position of the particle without a substantial effect on its motion), and the relation $\Delta Q = T\ln 2$ was proved, with a gradually increasing precision.³¹

Actually, discussion of another issue closely related to Maxwell's Demon, namely energy consumption at numerical calculations, was started earlier, in the 1960s. It was motivated by the exponential (*Moore's-law*) progress of the digital integrated circuits, which has led in particular, to a fast reduction of the energy ΔE "spent" (turned into heat) per one binary logic operation. In the recent generations of semiconductor digital integrated circuits, the typical ΔE is still above 10^{-17} J, i.e. still exceeds the room-temperature value of $T\ln 2 \approx 4 \times 10^{-21}$ J by several orders of magnitude.³² Still, some engineers believe that thermodynamics imposes this important lower limit on ΔE and hence presents an insurmountable obstacle to the future progress of computation. Unfortunately, in the 2000s this delusion resulted in a substantial and unjustified shift of electron device research resources toward using "non-charge degrees of freedom" such as spin (as if they do not obey the general laws of statistical physics!), so the issue deserves at least a brief discussion.

Let me believe that the reader of these notes understands that, in contrast to naïve popular talk, computers *do not create* any new information; all they can do is reshape ("process") the input information, *losing* most of it on the go. Indeed, any digital computation algorithm may be decomposed into simple, binary logical operations, each of them performed by a circuit called the *logic gate*. Some of these gates (e.g., the logical NOT performed by inverters, as well as memory READ and WRITE operations) do not change the amount of information in the computer. On the other hand, such *information-irreversible* logic gates as two-input NAND (or NOR, or XOR, etc.) erase one bit at each operation, because they turn two input bits into one output bit – see Fig. 5a.

In 1961, Rolf Landauer argued that each logic operation should turn into heat at least energy

$$\Delta E_{\min} = T \ln 2 \equiv k_B T_K \ln 2. \quad (2.51)$$

This result may be illustrated with the Szilard engine (Fig. 4), operated in a reversed cycle. At the first stage, with the partition's door closed, it uses external mechanical work $\Delta E = T\ln 2$ to reduce the volume in that the molecule is confined, from V to $V/2$, pumping heat $\Delta Q = \Delta E$ into the heat bath. To model a logically irreversible logic gate, let us now open the door in the partition, and thus lose one bit of

Irreversible
computation:
energy cost

³¹ See, for example, A. Bérut *et al.*, *Nature* **483**, 187 (2012); J. Koski *et al.*, *PNAS USA* **111**, 13786 (2014); Y. Jun *et al.*, *Phys. Rev. Lett.* **113**, 190601 (2014); J. Peterson *et al.*, *Proc. Roy. Soc. A* **472**, 20150813 (2016).

³² In practical computers, the effective ΔE is even much higher (currently, above $\sim 10^{-15}$ J) due to the high energy cost of moving data across a multi-component system, in particular between its logic and memory chips.

information about the molecule's position. Then we will never get the work $T \ln 2$ back, because moving the partition back to the right, with the door open, takes place at zero average pressure. Hence, Eq. (51) gives a fundamental limit for energy loss (per bit) at the logically irreversible computation.

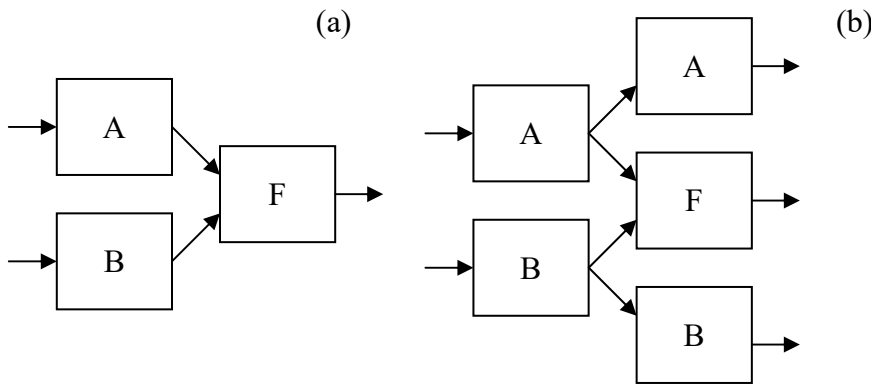


Fig. 2.5. Simple examples of (a) irreversible and (b) potentially reversible logic circuits. Each rectangle denotes a circuit storing one bit of information.

However, in 1973 Charles Bennett came up with convincing arguments that it is possible to avoid such energy loss by using only operations that are *reversible* not only physically, but also logically.³³ For that, one has to avoid any loss of information, i.e. any erasure of intermediate results, for example in the way shown in Fig. 5b.³⁴ At the end of all calculations, after the result has been copied into memory, the intermediate results may be “rolled back” through reversible gates to be eventually merged into a copy of input data, again without erasing a single bit. The minimal energy dissipation at such reversible calculation tends to zero as the operation speed is decreased, so the average energy loss per bit may be less than the perceived “fundamental thermodynamic limit” (51). The price to pay for this ultralow dissipation is a very high complexity of the hardware necessary for the storage of all intermediate results. However, using irreversible gates sparsely, it may be possible to reduce the complexity dramatically, so in the future such *mostly* reversible computation may be able to reduce energy consumption in practical digital electronics.³⁵

Before we leave Maxwell's Demon behind, let me use it to revisit, for one more time, the relation between the reversibility of the classical and quantum mechanics of Hamiltonian systems and the irreversibility possible in thermodynamics and statistical physics. In the thought experiment shown in Fig. 4, the laws of mechanics governing the motion of the molecule are reversible at all times. Still, at partition's motion to the right, driven by molecular impacts, the entropy grows, because the molecule picks up the heat $\Delta Q > 0$, and hence the entropy $\Delta S = \Delta Q/T > 0$, from the heat bath. The physical mechanism of this irreversible entropy (read: disorder) growth is the interaction of the molecule with uncontrollable components of the heat bath and the resulting loss of information about its motion. Philosophically, such emergence of irreversibility in large systems is a strong argument against *reductionism* – a naïve belief that by knowing the exact laws of Nature at the lowest, most fundamental

³³ C. Bennett, *IBM J. Res. Devel.* **17**, 525 (1973); see also C. Bennett, *Int. J. Theor. Phys.* **21**, 905 (1982).

³⁴ For that, all gates have to be *physically reversible*, with no static power consumption. Such logic devices do exist, though they are still not very practicable – see, e.g., K. Likharev, *Int. J. Theor. Phys.* **21**, 311 (1982). (Another reason why I am citing, rather reluctantly, my own paper is that it also gave a constructive proof that the reversible computation may also beat the perceived “fundamental quantum limit”, $\Delta E \Delta t > \hbar$, where Δt is the time of the binary logic operation.)

³⁵ Many currently explored schemes of *quantum computing* are also reversible – see, e.g., QM Sec. 8.5 and references therein.

level of its complexity, we can readily understand all phenomena on the higher levels of its organization. In reality, the macroscopic irreversibility of large systems is a good example³⁶ of a new law (in this case, the 2nd law of thermodynamics) that becomes relevant on a substantially new, higher level of complexity – without defying the lower-level laws. Without such new laws, very little of the higher-level organization of Nature may be understood.

2.4. Canonical ensemble and the Gibbs distribution

As was shown in Sec. 2, the microcanonical distribution may be directly used for solving some important problems. However, its further development, also due to J. Gibbs, turns out to be much more convenient for calculations.

Let us consider a statistical ensemble of macroscopically similar systems, each in thermal equilibrium with a heat bath of the same temperature T (Fig. 6a). Such an ensemble is called *canonical*.

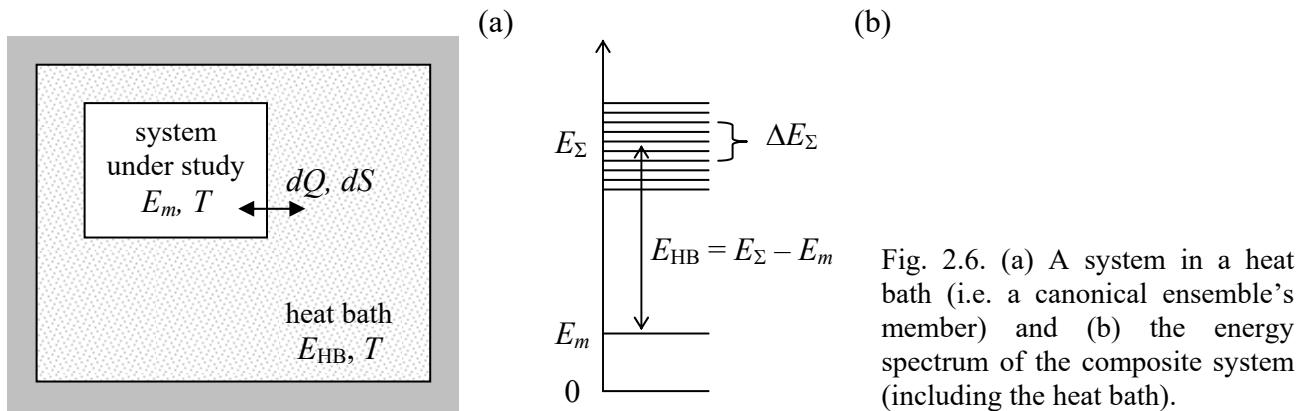


Fig. 2.6. (a) A system in a heat bath (i.e. a canonical ensemble's member) and (b) the energy spectrum of the composite system (including the heat bath).

It is intuitively evident that if the heat bath is sufficiently large, any thermodynamic variables characterizing the system under study should not depend on the *heat bath's* environment. In particular, we may assume that the heat bath is thermally insulated, so the total energy E_Σ of the *composite* system, consisting of the system of our interest plus the heat bath, does not change in time. For example, if the system under study is in a certain (say, m^{th}) quantum state, then the sum

$$E_\Sigma = E_m + E_{\text{HB}} \quad (2.52)$$

is time-independent. Now let us partition the considered *canonical* ensemble of such systems into much smaller sub-ensembles, each being a *microcanonical* ensemble of composite systems whose total, time-independent energies E_Σ are the same – as was discussed in Sec. 2, within a certain small energy interval $\Delta E_\Sigma \ll E_\Sigma$ – see Fig. 6b. Due to the very large size of each heat bath in comparison with that of the system under study, the heat bath's density of states g_{HB} is very high, and ΔE_Σ may be selected so

$$\frac{1}{g_{\text{HB}}} \ll \Delta E_\Sigma \ll |E_m - E_{m'}| \ll E_{\text{HB}}, \quad (2.53)$$

where m and m' are any states of the system of our interest.

³⁶ Another famous example is Charles Darwin's theory of biological evolution.

According to the microcanonical distribution, within each of these microcanonical sub-ensembles, the probabilities to find the composite system in any state are equal. Still, the heat bath energies $E_{\text{HB}} = E_{\Sigma} - E_m$ (Fig. 6b) of the members of this sub-ensemble may be different – due to the difference in E_m . The probability $W(E_m)$ to find the system of our interest (within the selected sub-ensemble) in a state with energy E_m is proportional to the number ΔM of the corresponding heat baths in the sub-ensemble. As Fig. 6b shows, in this case we may write $\Delta M = g_{\text{HB}}(E_{\text{HB}})\Delta E_{\Sigma}$. As a result, within the microcanonical sub-ensemble with the total energy E_{Σ} ,

$$W_m \propto \Delta M = g_{\text{HB}}(E_{\text{HB}})\Delta E_{\Sigma} = g_{\text{HB}}(E_{\Sigma} - E_m)\Delta E_{\Sigma}. \quad (2.54)$$

Let us simplify this expression further, using the Taylor expansion with respect to relatively small $E_m \ll E_{\Sigma}$. However, here we should be careful. As we have seen in Sec. 2, the density of states of a large system is a nearly exponential function of energy, so if we applied the Taylor expansion directly to Eq. (54), the Taylor series would only converge for very small E_m . A much broader applicability range may be obtained by taking the logarithms of both parts of Eq. (54) first:

$$\ln W_m = \text{const} + \ln[g_{\text{HB}}(E_{\Sigma} - E_m)] + \ln(\Delta E_{\Sigma}) = \text{const} + S_{\text{HB}}(E_{\Sigma} - E_m), \quad (2.55)$$

where the last step used Eq. (36) for the heat bath, and incorporated $\ln(\Delta E_{\Sigma})$ into the (inconsequential) constant. Now, we can Taylor-expand the (much more smooth) function of energy on the right-hand side of Eq. (55), and limit ourselves to the two leading terms of the series:

$$\ln W_m \approx \text{const} + S_{\text{HB}} \Big|_{E_m=0} - \frac{dS_{\text{HB}}}{dE_{\text{HB}}} \Big|_{E_m=0} E_m. \quad (2.56)$$

But according to Eq. (1.9), the derivative participating in this expression is nothing other than the reciprocal temperature of the heat bath, which (due to the large bath size) does not depend on whether E_m is equal to zero or not. Since our system of interest is in the thermal equilibrium with the bath, this is also the temperature T of the system – see Eq. (1.8). Hence Eq. (56) is merely

$$\ln W_m = \text{const} - \frac{E_m}{T}. \quad (2.57)$$

This equality describes a substantial decrease of W_m as E_m is increased by $\sim T$, and hence our linear approximation (56) is virtually exact as soon as E_{HB} is much larger than T – the condition that is rather easy to satisfy, because as we have seen in Sec. 2, the average energy per one degree of freedom of the system of the heat bath is also of the order of T , so its total energy is much larger than T because of its much larger size.

Now we should be careful again because so far, Eq. (57) was only derived for a sub-ensemble with a certain fixed E_{Σ} . However, since the second term on the right-hand side of Eq. (57) includes only E_m and T , which are independent of E_{Σ} , this relation, perhaps with different constant terms, is valid for all sub-ensembles of the canonical ensemble, and hence for that ensemble as the whole. Hence for the total probability of finding our system of interest in a state with energy E_m , in the canonical ensemble with temperature T , we can write

$$W_m = \text{const} \times \exp\left\{-\frac{E_m}{T}\right\} \equiv \frac{1}{Z} \exp\left\{-\frac{E_m}{T}\right\}. \quad (2.58)$$

Gibbs
distribution

This is the famous *Gibbs distribution*,³⁷ sometimes called the “canonical distribution”, which is arguably the summit of statistical physics,³⁸ because it may be used for a straightforward (or at least conceptually straightforward :-) calculation of all statistical and thermodynamic variables of a vast range of systems. Its physical sense is very clear: the interaction with the heat bath “punishes” the system states (by the reduction of their probability) for having higher energies – on the scale T of its thermal agitation.

Now let us calculate the coefficient Z participating in Eq. (58). Requiring, per Eq. (4), the sum of all W_m to be equal to 1, we get

$$Z = \sum_m \exp\left\{-\frac{E_m}{T}\right\}, \quad (2.59)$$

Statistical
sum

where the summation is formally extended to all quantum states of the system, though in practical calculations, the sum may be truncated to include only the states that are noticeably occupied. The apparently humble normalization coefficient Z turns out to be so important for applications that it has a special name – or actually, two names: either the *statistical sum* or the *partition function* of the system. To appreciate the importance of Z , let us use the general expression (29) for entropy to calculate it for the particular case of the canonical ensemble, i.e. the Gibbs distribution (58) of the probabilities W_n :

$$S = -\sum_m W_m \ln W_m = \frac{\ln Z}{Z} \sum_m \exp\left\{-\frac{E_m}{T}\right\} + \frac{1}{ZT} \sum_m E_m \exp\left\{-\frac{E_m}{T}\right\}. \quad (2.60)$$

On the other hand, according to the general rule (7), the thermodynamic (i.e. ensemble-averaged) value E of the internal energy of the system is

$$E = \sum_m W_m E_m = \frac{1}{Z} \sum_m E_m \exp\left\{-\frac{E_m}{T}\right\}, \quad (2.61a)$$

so the second term on the right-hand side of Eq. (60) is just E/T , while the first term equals $\ln Z$, due to Eq. (59). (By the way, using the notion of *reciprocal temperature* $\beta \equiv 1/T$, with the account of Eq. (59), Eq. (61a) may be also rewritten as

$$E = -\frac{\partial(\ln Z)}{\partial\beta}. \quad (2.61b)$$

E from Z

This formula is very convenient for calculations if our prime interest is the average internal energy E rather than F or W_n .) With these substitutions, Eq. (60) yields a very simple relation between the statistical sum and the entropy of the system:

$$S = \frac{E}{T} + \ln Z. \quad (2.62)$$

³⁷ The temperature dependence of the type $\exp\{-\text{const}/T\}$, especially when showing up in rates of certain events, e.g., chemical reactions, is also frequently called the *Arrhenius law* – after chemist S. Arrhenius who has noticed this law in numerous experimental data. In all cases I am aware of, the Gibbs distribution is the underlying reason for the Arrhenius law. (We will see several examples of that later in this course.)

³⁸ This is the opinion of many physicists, including Richard Feynman – who climbs on this “summit” already on the first page of his brilliant book *Statistical Mechanics*, CRC Press, 1998. (This is a collection of lectures on a few diverse, mostly advanced topics of statistical physics, rather than its systematic course, so it can hardly be used as the first textbook on the subject. However, I can highly recommend its first chapter to all my readers.)

Now using Eq. (1.33), we see that Eq. (62) gives a straightforward way to calculate the free energy F of the system from nothing other than its statistical sum (and temperature):

$$F \equiv E - TS = -T \ln Z. \quad (2.63) \quad F \text{ from } Z$$

The relations (61b) and (63) play a key role in the connection of statistics to thermodynamics, because they enable the calculation, from Z alone, of the thermodynamic potentials of the system in equilibrium, and hence of all other variables of interest, using the general thermodynamic relations – see especially the circular diagram shown in Fig. 1.6, and its discussion in Sec. 1.4. Let me only note that to calculate pressure P from the second of Eqs. (1.35), we would need to know the explicit dependence of F , and hence of the statistical sum Z on the system's volume V . This would require the calculation, by appropriate methods of either classical or quantum mechanics, of the dependence of the eigenenergies E_m on the volume. Numerous examples of such calculations will be given later in the course.

Before proceeding to such examples, let us notice that Eqs. (59) and (63) may be readily combined to give an elegant equality,

$$\exp\left\{-\frac{F}{T}\right\} = \sum_m \exp\left\{-\frac{E_m}{T}\right\}. \quad (2.64)$$

This formula, together with Eq. (59), enables us to rewrite the Gibbs distribution (58) in another form:

$$W_m = \exp\left\{\frac{F - E_m}{T}\right\}, \quad (2.65)$$

more convenient for some applications. In particular, this expression shows that since all probabilities W_m are below 1, F is always *lower* than the lowest energy level. Also, Eq. (65) clearly shows that the probabilities W_m do not depend on the energy reference, i. e. on an arbitrary constant added to all E_m – and hence to E and F .

2.5. Harmonic oscillator statistics

The last property may be immediately used in our first example of the Gibbs distribution application to a particular but very important system – the harmonic oscillator, for a much more general case than was done in Sec. 2, namely for an arbitrary relation between T and $\hbar\omega$.³⁹ Let us consider a canonical ensemble of similar oscillators, each in contact with a heat bath of temperature T . Selecting the ground-state energy $\hbar\omega/2$ for the origin of E , the oscillator eigenenergies (38) become $E_m = m\hbar\omega$ (with $m = 0, 1, \dots$), so the Gibbs distribution (58) for probabilities of these states is

$$W_m = \frac{1}{Z} \exp\left\{-\frac{E_m}{T}\right\} = \frac{1}{Z} \exp\left\{-\frac{m\hbar\omega}{T}\right\}, \quad (2.66)$$

with the following statistical sum:

$$Z = \sum_{m=0}^{\infty} \exp\left\{-\frac{m\hbar\omega}{T}\right\} \equiv \sum_{m=0}^{\infty} \lambda^m, \quad \text{where } \lambda \equiv \exp\left\{-\frac{\hbar\omega}{T}\right\} \leq 1. \quad (2.67)$$

³⁹ The task of making a very similar (and even simpler) calculation for another key class of quantum-mechanical objects, two-level systems, is left for the reader's exercise.

This is just the well-known infinite geometric progression (the “geometric series”),⁴⁰ with the sum

$$Z = \frac{1}{1 - \lambda} \equiv \frac{1}{1 - e^{-\hbar\omega/T}}, \quad (2.68)$$

so Eq. (66) yields

$$W_m = (1 - e^{-\hbar\omega/T}) e^{-m\hbar\omega/T}. \quad (2.69)$$

Figure 7a shows W_m for several lower energy levels, as functions of temperature, or rather of the $T/\hbar\omega$ ratio. The plots show that the probability of finding the oscillator in each particular state (except for the ground one, with $m = 0$) vanishes in both low- and high-temperature limits, and reaches its maximum value $W_m \sim 0.3/m$ at $T \sim m\hbar\omega$, so the contribution $m\hbar\omega W_m$ of each excited level to the average oscillator energy E is always smaller than $\hbar\omega$.

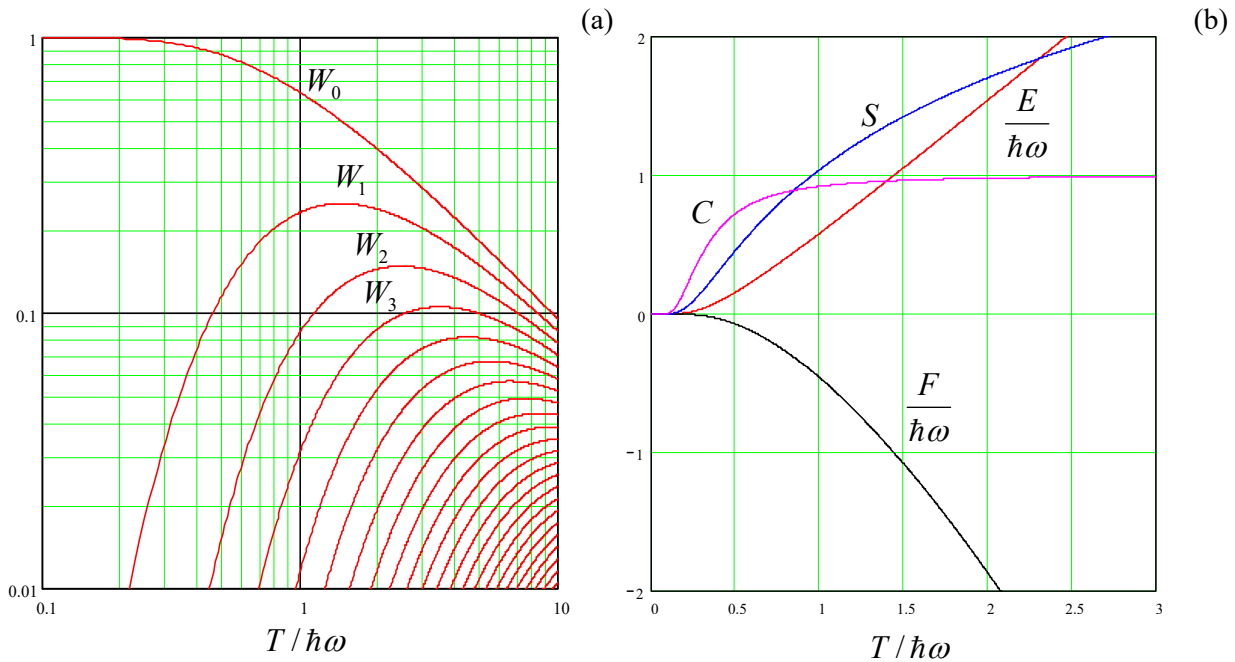


Fig. 2.7. Statistical and thermodynamic parameters of a harmonic oscillator, as functions of temperature.

This average energy may be calculated in either of two ways: either using Eq. (61a) directly:

$$E = \sum_{m=0}^{\infty} E_m W_m = (1 - e^{-\hbar\omega/T}) \sum_{m=0}^{\infty} m\hbar\omega e^{-m\hbar\omega/T}, \quad (2.70)$$

or (simpler) using Eq. (61b), as

$$E = -\frac{\partial}{\partial \beta} \ln Z = \frac{\partial}{\partial \beta} \ln(1 - \exp\{-\beta\hbar\omega\}), \quad \text{where } \beta \equiv \frac{1}{T}. \quad (2.71)$$

Both methods give (of course) the same result,⁴¹

⁴⁰ See, e.g., MA Eq. (2.8b).

$$E = E(\omega, T) = \hbar\omega \frac{1}{e^{\hbar\omega/T} - 1}, \quad (2.72)$$

Harmonic oscillator: average energy

which is valid for arbitrary temperature and plays a key role in many fundamental problems of physics. The red line in Fig. 7b shows this result as a function of the normalized temperature. At relatively low temperatures, $T \ll \hbar\omega$, the oscillator is predominantly in its lowest (ground) state, and its energy (on top of the constant zero-point energy $\hbar\omega/2$, which was used in our calculation as the reference) is exponentially small: $E \approx \hbar\omega \exp\{-\hbar\omega/T\} \ll T, \hbar\omega$. On the other hand, in the high-temperature limit, the energy tends to T . This is exactly the result (a particular case of the equipartition theorem) that was obtained in Sec. 2 from the microcanonical distribution. Please note how much simpler is the calculation using the Gibbs distribution, even for an arbitrary ratio $T/\hbar\omega$.

To complete the discussion of the thermodynamic properties of the harmonic oscillator, we can calculate its free energy using Eq. (63):

$$F = T \ln \frac{1}{Z} = T \ln(1 - e^{-\hbar\omega/T}). \quad (2.73)$$

Now the entropy may be found from thermodynamics: either from the first of Eqs. (1.35), $S = -(\partial F/\partial T)_V$, or (even more easily) from Eq. (1.33): $S = (E - F)/T$. Both relations give, of course, the same result:

$$S = \frac{\hbar\omega}{T} \frac{1}{e^{\hbar\omega/T} - 1} - \ln(1 - e^{-\hbar\omega/T}). \quad (2.74)$$

Finally, since in the general case, the dependence of the oscillator properties (essentially, of ω) on volume V is not specified, such variables as P , μ , G , W , and Ω are not defined, and what remains is to calculate the average heat capacity C per one oscillator:

$$C = \frac{\partial E}{\partial T} = \left(\frac{\hbar\omega}{T}\right)^2 \frac{e^{\hbar\omega/T}}{(e^{\hbar\omega/T} - 1)^2} \equiv \left[\frac{\hbar\omega/2T}{\sinh(\hbar\omega/2T)}\right]^2. \quad (2.75)$$

The calculated thermodynamic variables are plotted in Fig. 7b as functions of temperature. In the low-temperature limit ($T \ll \hbar\omega$), they all tend to zero. On the other hand, in the high-temperature limit ($T \gg \hbar\omega$), $F \rightarrow -T \ln(T/\hbar\omega) \rightarrow -\infty$, $S \rightarrow \ln(T/\hbar\omega) \rightarrow +\infty$, and $C \rightarrow 1$ (in the SI units, $C \rightarrow k_B$). Note that the last limit is the direct corollary of the equipartition theorem: each of the two “half-degrees of freedom” of the oscillator gives, in the classical limit, the same contribution $C = 1/2$ into its heat capacity.

Now let us use Eq. (69) to discuss the statistics of the quantum oscillator described by the Hamiltonian (46), in the coordinate representation. Again using the density matrix’s diagonality in thermodynamic equilibrium, we may use a relation similar to Eqs. (47) to calculate the probability density to find the oscillator at coordinate q :

$$w(q) = \sum_{m=0}^{\infty} W_m w_m(q) = \sum_{m=0}^{\infty} W_m |\psi_m(q)|^2 = (1 - e^{-\hbar\omega/T}) \sum_{m=0}^{\infty} e^{-m\hbar\omega/T} |\psi_m(q)|^2, \quad (2.76)$$

⁴¹ It was first obtained in 1924 by S. Bose and is sometimes called the *Bose distribution* – a particular case of the *Bose-Einstein distribution* to be discussed in Sec. 8 below.

where $\psi_m(q)$ is the normalized eigenfunction of the m^{th} stationary state of the oscillator. Since each $\psi_m(q)$ is proportional to the Hermite polynomial⁴² that requires at least m elementary functions for its representation, working out the sum in Eq. (76) is a bit tricky,⁴³ but the final result is rather simple: $w(q)$ is just a normalized *Gaussian distribution* (the “bell curve”),

$$w(q) = \frac{1}{(2\pi)^{1/2} \delta q} \exp\left\{-\frac{q^2}{2(\delta q)^2}\right\}, \quad (2.77)$$

with $\langle q \rangle = 0$, and

$$\langle q^2 \rangle = (\delta q)^2 = \frac{\hbar}{2m\omega} \coth \frac{\hbar\omega}{2T}. \quad (2.78)$$

Since the function $\coth \xi$ tends to 1 at $\xi \rightarrow \infty$, and diverges as $1/\xi$ at $\xi \rightarrow 0$, Eq. (78) shows that the width δq of the coordinate distribution is nearly constant (and equal to that, $(\hbar/2m\omega)^{1/2}$, of the ground-state wavefunction ψ_0) at $T \ll \hbar\omega$, and grows as $(T/m\omega^2)^{1/2} \equiv (T/\kappa)^{1/2}$ at $T/\hbar\omega \rightarrow \infty$.

As a sanity check, we may use Eq. (78) to write the following expression,

$$U \equiv \left\langle \frac{\kappa q^2}{2} \right\rangle = \frac{\hbar\omega}{4} \coth \frac{\hbar\omega}{2T} \rightarrow \begin{cases} \hbar\omega/4, & \text{for } T \ll \hbar\omega, \\ T/2, & \text{for } \hbar\omega \ll T, \end{cases} \quad (2.79)$$

for the average potential energy of the oscillator. To comprehend this result, let us recall that Eq. (72) for the average full energy E was obtained by counting it from the ground state energy $\hbar\omega/2$ of the oscillator. If we add this reference energy to that result, we get

$$E = \frac{\hbar\omega}{e^{\hbar\omega/T} - 1} + \frac{\hbar\omega}{2} \equiv \frac{\hbar\omega}{2} \coth \frac{\hbar\omega}{2T}. \quad (2.80)$$

Harmonic
oscillator:
E

We see that for arbitrary temperature, $U = E/2$, as was already discussed in Sec. 2. This means that the average kinetic energy, equal to $E - U$, is also the same:⁴⁴

$$\left\langle \frac{p^2}{2m} \right\rangle = \left\langle \frac{\kappa q^2}{2} \right\rangle = \frac{E}{2} = \frac{\hbar\omega}{4} \coth \frac{\hbar\omega}{2T}. \quad (2.81)$$

In the classical limit $T \gg \hbar\omega$, each of the energies tends to $T/2$, reproducing the equipartition theorem (48).

2.6. Two important applications

The results of the previous section, especially Eq. (72), have innumerable applications in physics and related disciplines, and here I have time for a brief discussion of only two of them.

(i) Blackbody radiation. Let us consider a free-space volume V limited by non-absorbing (i.e. ideally reflecting) walls. Electrodynamics tells us⁴⁵ that the electromagnetic field in such a “cavity” may be represented as a sum of modes with a time evolution similar to that of the usual harmonic oscillator.

⁴² See, e.g., QM Sec. 2.10.

⁴³ The calculation may be found, e.g., in QM Sec. 7.2.

⁴⁴ As a reminder: the equality of these two averages, for arbitrary temperatures, was proved already in Sec. 2.

If the volume V is large enough,⁴⁶ the number of these modes within a small range dk of the wave vector magnitude k is

$$dN = \frac{gV}{(2\pi)^3} d^3k = \frac{gV}{(2\pi)^3} 4\pi k^2 dk, \quad (2.82)$$

where for electromagnetic waves, the degeneracy factor g is equal to 2, due to their two different independent (e.g., linear) polarizations of waves with the same wave vector \mathbf{k} . With the linear, isotropic dispersion relation for waves in vacuum, $k = \omega/c$, Eq. (82) yields

$$dN = \frac{2V}{(2\pi)^3} 4\pi \frac{\omega^2 d\omega}{c^3} \equiv V \frac{\omega^2}{\pi^2 c^3} d\omega. \quad (2.83)$$

On the other hand, quantum mechanics says⁴⁷ that the energy of such a “field oscillator” is quantized per Eq. (38), so at thermal equilibrium its average energy is described by Eq. (72). Plugging that result into Eq. (83), we see that the spectral density of the electromagnetic field’s energy, per unit volume, is

$$u(\omega) \equiv \frac{E}{V} \frac{dN}{d\omega} = \frac{\hbar\omega^3}{\pi^2 c^3} \frac{1}{e^{\hbar\omega/T} - 1}. \quad (2.84)$$

Planck's
radiation
law

This is the famous *Planck’s blackbody radiation law*.⁴⁸ To understand why its common name mentions radiation, let us consider a small planar part, of area dA , of a surface that completely absorbs electromagnetic waves incident from any direction. (Such “perfect black body” approximation may be closely approached using special experimental structures, especially in limited frequency intervals.) Figure 8 shows that if the arriving wave was planar, with the incidence angle θ , then the power $d\mathcal{P}_\theta(\omega)$ absorbed by the surface of small area dA , within a small frequency interval $d\omega$, i.e. the energy incident at that area in unit time, would be equal to the radiation energy within the same frequency interval, contained inside an imaginary cylinder (shaded in Fig. 8) of height c , base area $dA \cos\theta$, and hence volume $dV = c dA \cos\theta$:

$$d\mathcal{P}_\theta(\omega) = u(\omega) d\omega dV = u(\omega) d\omega c dA \cos\theta. \quad (2.85)$$

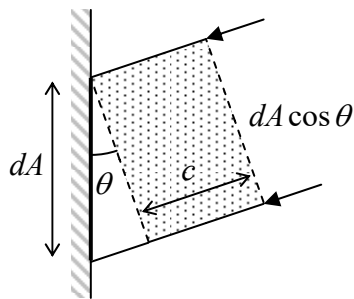


Fig. 2.8. Calculating the relation between $d\mathcal{P}_\theta(\omega)$ and $u(\omega)d\omega$.

⁴⁵ See, e.g., EM Sec. 7.8.

⁴⁶ In our current context, the volume should be much larger than $(c\hbar/T)^3$, where $c \approx 3 \times 10^8$ m/s is the speed of light. For the room temperature ($T \approx k_B \times 300\text{K} \approx 4 \times 10^{-21}$ J), this lower bound is of the order of 10^{-16} m³.

⁴⁷ See, e.g., QM Sec. 9.1.

⁴⁸ Let me hope the reader knows that this law was first suggested in 1900 by Max Planck as an empirical fit for the experimental data on blackbody radiation, and this was the historic point at which the Planck constant \hbar (or rather $h \equiv 2\pi\hbar$) was introduced – see, e.g., QM Sec. 1.1.

Since the thermally-induced field is isotropic, i.e. propagates equally in all directions, this result should be averaged over all solid angles within the polar angle interval $0 \leq \theta \leq \pi/2$:

$$\frac{d\mathcal{P}(\omega)}{dAd\omega} = \frac{1}{4\pi} \int \frac{d\mathcal{P}_\theta(\omega)}{dAd\omega} d\Omega = cu(\omega) \frac{1}{4\pi} \int_0^{\pi/2} \sin\theta d\theta \int_0^{2\pi} d\varphi \cos\theta = \frac{c}{4} u(\omega). \quad (2.86)$$

Hence Planck's expression (84), multiplied by $c/4$, gives the power absorbed by such a "blackbody" surface. But at thermal equilibrium, this absorption has to be exactly balanced by the surface's own radiation, due to its non-zero temperature T .

I hope the reader is familiar with the main features of the Planck law (84), including its general shape (Fig. 9), with the low-frequency asymptote $u(\omega) \propto \omega^2$ (due to its historic significance, bearing the special name of the *Rayleigh-Jeans law*), the exponential drop at high frequencies (the *Wien law*), and the resulting maximum of the function $u(\omega)$, reached at the frequency ω_{\max} with

$$\hbar\omega_{\max} \approx 2.82T, \quad (2.87)$$

i.e. at the wavelength $\lambda_{\max} = 2\pi/k_{\max} = 2\pi c/\omega_{\max} \approx 2.22 c\hbar/T$.

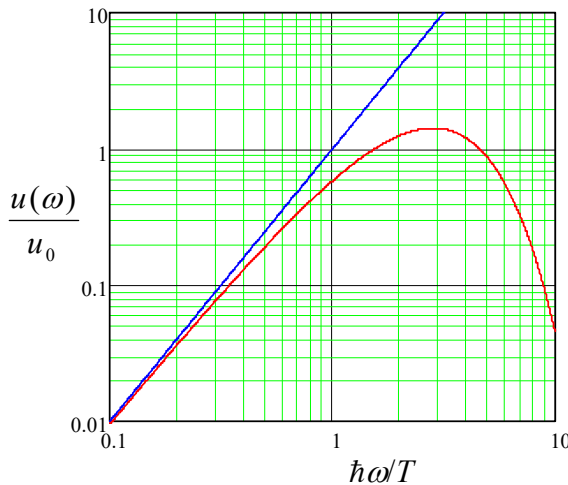


Fig. 2.9. The frequency dependence of the blackbody radiation density, normalized by $u_0 \equiv T^3/\pi^2\hbar^2c^3$, according to the Planck law (red line) and the Rayleigh-Jeans law (blue line).

Still, I cannot help mentioning a few important particular values: one corresponding to the visible light ($\lambda_{\max} \sim 500$ nm) for the Sun's effective surface temperature $T_K \approx 6,000$ K, and another one corresponding to the mid-infrared range ($\lambda_{\max} \sim 10$ μm) for the Earth's surface temperature $T_K \approx 300$ K. The balance of these two radiations, absorbed and emitted by the Earth, determines its surface temperature and hence has the key importance for all life on our planet. This is why it is at the front and center of the current climate change discussions. As one more example, the cosmic microwave background (CMB) radiation, closely following the Planck law with $T_K = 2.725$ K (and hence having the maximum density at $\lambda_{\max} \approx 1.9$ mm), and in particular its (very small) anisotropy, is a major source of data for modern cosmology.

Now let us calculate the total energy E of the blackbody radiation inside some volume V . It may be found from Eq. (84) by its integration over all frequencies:^{49:50}

⁴⁹ The last step in Eq. (88) uses a table integral, equal to $\Gamma(4)\zeta(4) = (3!)(\pi^4/90) = \pi^4/15$ – see, e.g., MA Eq. (6.8b), with $s = 4$, and then MA Eqs. (6.7e), and (2.7b).

$$E = V \int_0^{\infty} u(\omega) d\omega = V \int_0^{\infty} \frac{\hbar \omega^3}{\pi^2 c^3} \frac{d\omega}{e^{\hbar\omega/T} - 1} = \frac{VT^4}{\pi^2 \hbar^3 c^3} \int_0^{\infty} \frac{\xi^3 d\xi}{e^{\xi} - 1} = V \frac{\pi^2}{15 \hbar^3 c^3} T^4. \quad (2.88)$$

Using Eq. (86) to recast Eq. (88) into the total power radiated by a blackbody surface, we get the well-known *Stefan* (or “Stefan-Boltzmann”) law⁵¹

$$\frac{d\mathcal{P}}{dA} = \frac{\pi^2}{60 \hbar^3 c^2} T^4 \equiv \sigma T_K^4, \quad (2.89a) \quad \text{Stefan law}$$

where σ is the *Stefan-Boltzmann constant*

$$\sigma \equiv \frac{\pi^2}{60 \hbar^3 c^2} k_B^4 \approx 5.67 \times 10^{-8} \frac{\text{W}}{\text{m}^2 \text{K}^4}. \quad (2.89b) \quad \text{Stefan-Boltzmann constant}$$

By this point, the thoughtful reader should have an important concern ready: Eq. (84) and hence Eq. (88) are based on Eq. (72) for the average energy of each oscillator, referred to its ground-state energy $\hbar\omega/2$. However, the radiation power should not depend on the energy origin; why have not we included the ground energy of each oscillator into the integration (88), as we have done in Eq. (80)? The answer is that usual radiation detectors only measure the *difference* between the power \mathcal{P}_{in} of the incident radiation (say, that of a blackbody surface with temperature T) and their own back-radiation power \mathcal{P}_{out} , corresponding to some effective temperature T_d of the detector – see Fig. 10. But however low T_d is, the temperature-independent contribution $\hbar\omega/2$ of the ground-state energy to the back radiation is always there. Hence, the term $\hbar\omega/2$ drops out from the balance, and cannot be detected – at least in this simple way. This is the reason why we had the right to ignore this contribution in Eq. (88) – very fortunately, because it would lead to the integral’s divergence at its upper limit. However, let me repeat that the ground-state energy of the electromagnetic field oscillators is physically real and may be important – see Sec. 5.5 below.

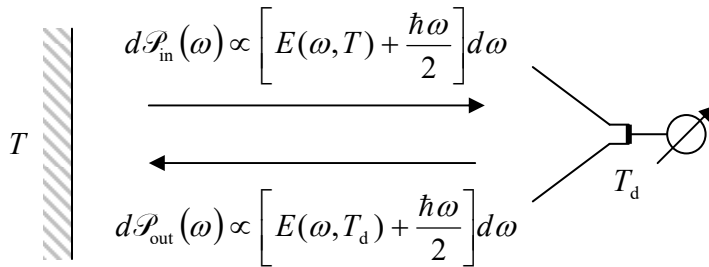


Fig. 2.10. The power balance at the electromagnetic radiation power measurement.

One more interesting result may be deduced from the free energy F of the electromagnetic radiation, which may be calculated by integration of Eq. (73) over all the modes, with the appropriate weight (83):

⁵⁰ Note that the heat capacity $C_V \equiv (\partial E / \partial T)_V$, following from Eq. (88), is proportional to T^3 at any temperature, and hence does not obey the trend $C_V \rightarrow \text{const}$ at $T \rightarrow \infty$. This is the result of the unlimited growth, with temperature, of the number of thermally-excited field oscillators with frequencies ω below T/\hbar .

⁵¹ Its functional part ($E \propto T^4$) was deduced in 1879 by Joseph Stefan from earlier experiments by John Tyndall. Theoretically, it was proved in 1884 by L. Boltzmann, using a result derived earlier by Adolfo Bartoli from the Maxwell equations for the electromagnetic field – all well before Max Planck’s work.

$$F = \sum_{\omega} T \ln(1 - e^{-\hbar\omega/T}) \rightarrow \int_0^{\infty} T \ln(1 - e^{-\hbar\omega/T}) \frac{dN}{d\omega} d\omega = \int_0^{\infty} T \ln(1 - e^{-\hbar\omega/T}) \left(V \frac{\omega^2}{\pi^2 c^3} \right) d\omega. \quad (2.90)$$

Representing $\omega^2 d\omega$ as $d(\omega^3)/3$, we can readily work out this integral by parts, reducing it to a table integral similar to that in Eq. (88), and getting a surprisingly simple result:

$$F = -V \frac{\pi^2}{45\hbar^3 c^3} T^4 \equiv -\frac{E}{3}. \quad (2.91)$$

Now we can use the second of the general thermodynamic relations (1.35) to calculate the pressure exerted by the radiation on the walls of the containing volume V :⁵²

$$P = -\left(\frac{\partial F}{\partial V} \right)_T = \frac{\pi^2}{45\hbar^3 c^3} T^4 = \frac{E}{3V}. \quad (2.92a)$$

Rewritten in the form,

$$PV = \frac{E}{3}, \quad (2.92b)$$

Photon gas:
PV vs. E

this result may be considered as the equation of state of the electromagnetic field, i.e. from the quantum-mechanical point of view, of the *photon gas*. Note that the equation of state (1.44) of the ideal classical gas may be represented in a similar form, but with a coefficient generally different from Eq. (92). Indeed, according to the equipartition theorem, for an ideal gas of non-relativistic particles whose internal degrees of freedom are in a fixed (say, ground) state, the temperature-dependent energy is that of the three translational “half-degrees of freedom”, $E = 3N(T/2)$. Expressing from here the product $NT = (2E/3)$, and plugging it into Eq. (1.44), we get a relation similar to Eq. (92), but with a twice larger factor before E . On the other hand, a *relativistic* treatment of the classical gas shows that Eq. (92) is valid for any gas in the ultra-relativistic limit, $T \gg mc^2$, where m is the rest mass of the gas’ particle. Evidently, photons (i.e. particles with $m = 0$) satisfy this condition at any energy.⁵³

Finally, let me note that Eq. (92) allows for the following interesting interpretation. The last of Eqs. (1.60), being applied to Eq. (92), shows that in this particular case, the grand thermodynamic potential Ω equals $(-E/3)$, so according to Eq. (91), it is equal to F . But according to the definition of Ω , i.e. the first of Eqs. (1.60), this means that the chemical potential of the electromagnetic field excitations (photons) vanishes:

$$\mu = \frac{F - \Omega}{N} = 0. \quad (2.93)$$

In Sec. 8 below, we will see that the same result follows from the comparison of Eq. (72) and the general Bose-Einstein distribution for arbitrary bosons. So, from the statistical point of view, photons may be considered bosons with zero chemical potential.

(ii) Specific heat of solids. The heat capacity of solids is readily measurable, and in the early 1900s, its experimentally observed temperature dependence served as an important test for the then-

⁵² This formula may be also derived from the expression for the forces exerted by the electromagnetic radiation on the walls (see, e.g. EM Sec. 9.8), but the above calculation is much simpler.

⁵³ Note that according to Eqs. (1.44), (88), and (92), the difference between the equations of state of the photon gas and an ideal gas of non-relativistic particles, expressed in the more usual form $P = P(V, T)$, is much more dramatic: $P \propto T^4 V^0$ vs. $P \propto T^1 V^{-1}$.

emerging quantum theories. However, the theoretical calculation of C_V is not simple⁵⁴ – even for insulators, whose specific heat at realistic temperatures is due to thermally-induced vibrations of their crystal lattice alone.⁵⁵ Indeed, at relatively low frequencies, a solid may be treated as an elastic continuum. This continuum supports three different modes of mechanical waves with the same frequency ω , that all obey linear dispersion laws, $\omega = vk$, but the velocity $v = v_l$ for one of these modes (the *longitudinal sound*) is higher than that (v_t) of two other modes (the *transverse sound*).⁵⁶ At such frequencies, the wave mode density may be described by an evident generalization of Eq. (83):

$$dN = V \frac{1}{(2\pi)^3} \left(\frac{1}{v_l^3} + \frac{2}{v_t^3} \right) 4\pi\omega^2 d\omega. \quad (2.94a)$$

For what follows, it is convenient to rewrite this relation in a form similar to Eq. (83):

$$dN = \frac{3V}{(2\pi)^3} 4\pi \frac{\omega^2 d\omega}{v^3}, \quad \text{with } v \equiv \left[\frac{1}{3} \left(\frac{1}{v_l^3} + \frac{2}{v_t^3} \right) \right]^{-1/3}. \quad (2.94b)$$

However, the basic wave theory shows⁵⁷ that as the frequency ω of a sound wave in a periodic structure is increased so much that its half-wavelength π/k approaches the crystal period d , the dispersion law $\omega(k)$ becomes nonlinear before the frequency reaches its maximum at $k = \pi/d$. To make things even more complex, 3D crystals are generally anisotropic, so the dispersion law is different in different directions of the wave propagation. As a result, the exact statistics of thermally excited sound waves, and hence the heat capacity of crystals, is rather complicated and specific for each particular crystal type.

In 1912, P. Debye suggested an approximate theory of the specific heat's temperature dependence, which is in surprisingly good agreement with experiment for many insulators, including polycrystalline and amorphous materials. In his model, the linear (*acoustic*) dispersion law $\omega = vk$, with the effective sound velocity v defined by the second of Eqs. (94b), is assumed to be exact all the way up to some cutoff frequency ω_D , the same for all three wave modes. This *Debye frequency* may be defined by the requirement that the total number of acoustic modes, calculated within this model from Eq. (94b),

$$N = V \frac{1}{(2\pi)^3} \frac{3}{v^3} \int_0^{\omega_D} 4\pi\omega^2 d\omega = \frac{V\omega_D^3}{2\pi^2 v^3}, \quad (2.95)$$

is equal to the universal number $N = 3nV$ of the degrees of freedom (and hence of independent oscillation modes) in a 3D system of nV elastically coupled particles, where n is the atomic density of the crystal, i.e. the number of atoms per unit volume.⁵⁸ For this model, Eq. (72) immediately yields the following expression for the average energy and specific heat (in thermal equilibrium at temperature T):

$$E = V \frac{1}{(2\pi)^3} \frac{3}{v^3} \int_0^{\omega_D} \frac{\hbar\omega}{e^{\hbar\omega/T} - 1} 4\pi\omega^2 d\omega \equiv 3nVT D(x)_{x=T_D/T}, \quad (2.96)$$

⁵⁴ Due to a rather low-temperature expansion of solids, the difference between their C_V and C_P is small.

⁵⁵ In good conductors (e.g., metals), specific heat is contributed (and at low temperatures, dominated) by free electrons – see Sec. 3.3 below.

⁵⁶ See, e.g., CM Sec. 7.7.

⁵⁷ See, e.g., CM Sec. 6.3, in particular, Fig. 6.5 and its discussion.

⁵⁸ See, e.g., CM Sec. 6.2.

Debye
law

$$c_V \equiv \frac{C_V}{nV} = \frac{1}{nV} \left(\frac{\partial E}{\partial T} \right)_V = 3 \left[D(x) - x \frac{dD(x)}{dx} \right]_{x=T_D/T}, \quad (2.97)$$

where $T_D \equiv \hbar\omega_D$ is called the *Debye temperature*,⁵⁹ and

$$D(x) \equiv \frac{3}{x^3} \int_0^x \frac{\xi^3 d\xi}{e^\xi - 1} \rightarrow \begin{cases} 1, & \text{for } x \rightarrow 0, \\ \pi^4/5x^3, & \text{for } x \rightarrow \infty, \end{cases} \quad (2.98)$$

is the *Debye function*. The red lines in Fig. 11 show the temperature dependence of the specific heat c_V (per particle) within the Debye model. At high temperatures, it approaches a constant value of three, corresponding to the energy $E = 3nVT$, in agreement with the equipartition theorem for each of three degrees of freedom (i.e. six half-degrees of freedom) of each mode. (This value of c_V is known as the *Dulong-Petit law*.) In the opposite limit of low temperatures, the specific heat is much smaller:

$$c_V \approx \frac{12\pi^4}{5} \left(\frac{T}{T_D} \right)^3 \ll 1, \quad (2.99)$$

reflecting the reduction of the number of excited phonons with $\hbar\omega < T$ as the temperature is decreased.

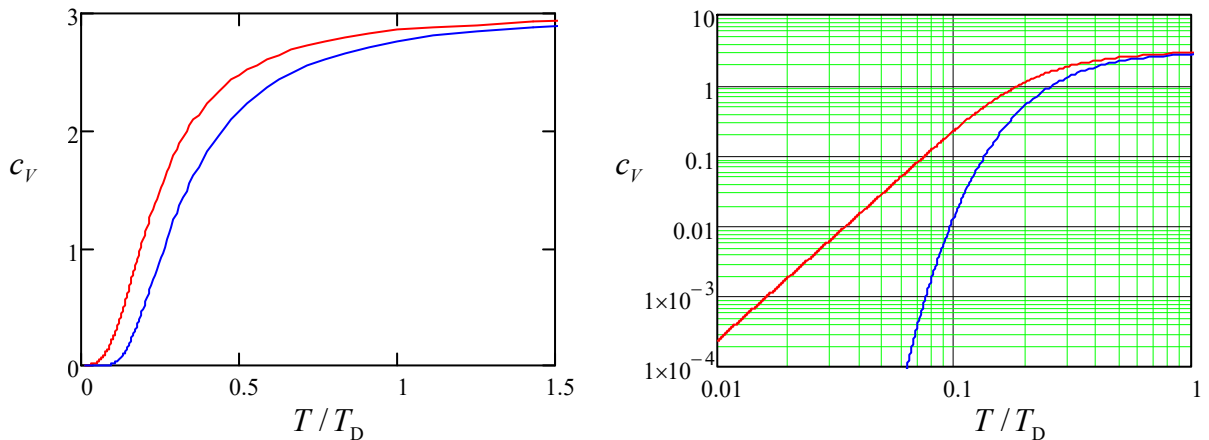


Fig. 2.11. The specific heat as a function of temperature in the Debye (red lines) and Einstein (blue lines) models.

As a historic curiosity, P. Debye's work followed one by A. Einstein, who had suggested (in 1907) a simpler model of crystal vibrations. In his model, all $3nV$ independent oscillatory modes of nV atoms of the crystal have approximately the same frequency, say ω_E , and Eq. (72) immediately yields

$$E = 3nV \frac{\hbar\omega_E}{e^{\hbar\omega_E/T} - 1}, \quad (2.100)$$

so the specific heat is functionally similar to Eq. (75):

⁵⁹ In the SI units, the Debye temperature T_D is of the order of a few hundred K for most simple solids (e.g., ~ 430 K for aluminum and ~ 340 K for copper), with somewhat lower values for crystals with heavy atoms (~ 105 K for lead), and reaches its highest value ~ 2200 K for diamond, with its relatively light atoms and very stiff lattice.

$$c_V \equiv \frac{1}{nV} \left(\frac{\partial E}{\partial T} \right)_V = 3 \left[\frac{\hbar \omega_E / 2T}{\sinh(\hbar \omega_E / 2T)} \right]^2. \quad (2.101)$$

This dependence $c_V(T)$ is shown with blue lines in Fig. 11 (assuming, for the sake of simplicity, that $\hbar \omega_E = T_D$). At high temperatures, this result does satisfy the universal Dulong-Petit law ($c_V = 3$), but for $T \ll T_D$, Einstein's model predicts a much faster (exponential) drop of the specific heat as the temperature is reduced. (The difference between the Debye and Einstein models is not too spectacular on the linear scale, but in the log-log plot shown on the right panel of Fig. 11, it is rather dramatic.⁶⁰) The Debye model is in much better agreement with experimental data for simple, monoatomic crystals, thus confirming the conceptual correctness of his wave-based approach.

Note, however, that when a genius such as Albert Einstein makes an error, there is usually some deep and important background under it. Indeed, crystals with the basic cell consisting of atoms of two or more types (such as NaCl, etc.), feature two or more separate branches of the dispersion law $\omega(k)$ – see, e.g., Fig. 12. While the lower, “acoustic” branch is virtually similar to those for monoatomic crystals and may be approximated by the Debye model, $\omega = vk$, reasonably well, the upper (“optical”⁶¹) branch does not approach $\omega = 0$ at any k . Moreover, for large values of the atomic mass ratio r , the optical branches are almost flat, with virtually k -independent frequencies ω_0 , which correspond to simple oscillations of each light atom between its heavy neighbors. For thermal excitations of such oscillations, and their contribution to the specific heat, Einstein's model (with $\omega_E = \omega_0$) gives a very good approximation, so for such solids, the specific heat may be well described by a sum of the Debye and Einstein laws (97) and (101), with appropriate weights.

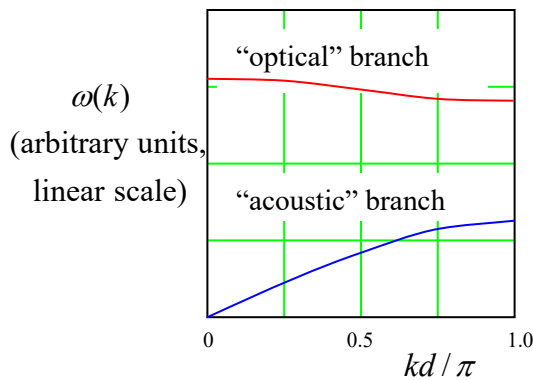


Fig. 2.12. The dispersion relation for mechanical waves in a simple 1D model of a solid, with similar interparticle distances d , but alternating particle masses, plotted for a particular mass ratio $r = 5$ – see CM Chapter 6.

2.7. Grand canonical ensemble and distribution

As we have seen, the Gibbs distribution is a very convenient way to calculate the statistical and thermodynamic properties of systems with a fixed number N of particles. However, for systems in which N may vary, another distribution is preferable for applications. Several examples of such situations (as

⁶⁰ This is why there is the following general “rule of thumb” in quantitative sciences: if you plot your data on a linear rather than log scale, you better have a good excuse ready. (An example of a valid excuse: the variable you are plotting changes its sign within the range you want to exhibit.)

⁶¹ This term stems from the fact that at $k \rightarrow 0$, the mechanical waves corresponding to these branches have phase velocities $v_{\text{ph}} \equiv \omega(k)/k$ that are much higher than that of the acoustic waves, and may approach the speed of light. As a result, these waves can strongly interact with electromagnetic (practically, optical) waves of the same frequency, while acoustic waves cannot.

well as the basic thermodynamics of such systems) have already been discussed in Sec. 1.5. Perhaps even more importantly, statistical distributions for systems with variable N are also applicable to some ensembles of independent particles in certain single-particle states even if the number of the particles is fixed – see the next section.

With this motivation, let us consider what is called the *grand canonical ensemble* (Fig. 13). It is similar to the canonical ensemble discussed in Sec. 4 (see Fig. 6) in all aspects, besides that now the system under study and the heat bath (in this case, more often called the *environment*) may exchange not only heat but also particles. In this ensemble, all environments are in both the thermal and chemical equilibrium, with their temperatures T and chemical potentials μ the same for all members.

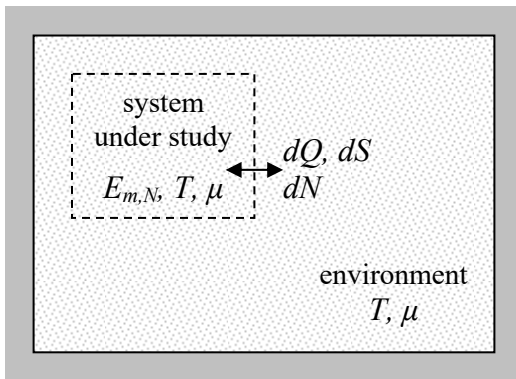


Fig. 2.13. A member of the grand canonical ensemble.

Let us assume that the system of interest is also in chemical and thermal equilibrium with its environment. Then using exactly the same arguments as in Sec. 4 (including the specification of microcanonical sub-ensembles with fixed E_Σ and N_Σ), we may generalize Eq. (55), taking into account that the entropy S_{env} of the environment is now a function of not only its energy $E_{\text{env}} = E_\Sigma - E_{m,N}$,⁶² but also of the number of particles $N_{\text{env}} = N_\Sigma - N$, with E_Σ and N_Σ fixed:

$$\begin{aligned} \ln W_{m,N} &\propto \ln M = \ln g_{\text{env}}(E_\Sigma - E_{m,N}, N_\Sigma - N) + \ln(\Delta E_\Sigma) = S_{\text{env}}(E_\Sigma - E_{m,N}, N_\Sigma - N) + \text{const} \\ &\approx S_{\text{env}} \Big|_{E_\Sigma, N_\Sigma} - \frac{\partial S_{\text{env}}}{\partial E_{\text{env}}} \Big|_{E_\Sigma, N_\Sigma} E_{m,N} - \frac{\partial S_{\text{env}}}{\partial N_{\text{env}}} \Big|_{E_\Sigma, N_\Sigma} N + \text{const}. \end{aligned} \quad (2.102)$$

To simplify this relation, let us rewrite Eq. (1.52) in the following equivalent form:

$$dS = \frac{1}{T} dE + \frac{P}{T} dV - \frac{\mu}{T} dN. \quad (2.103)$$

Hence, if the entropy S of a system is expressed as a function of E , V , and N , then

$$\left(\frac{\partial S}{\partial E} \right)_{V,N} = \frac{1}{T}, \quad \left(\frac{\partial S}{\partial V} \right)_{E,N} = \frac{P}{T}, \quad \left(\frac{\partial S}{\partial N} \right)_{E,V} = -\frac{\mu}{T}. \quad (2.104)$$

Applying the first one and the last one of these relations to the last form of Eq. (102), and using the equality of the temperatures T and the chemical potentials μ in the system under study and its environment, at equilibrium (as was discussed in Sec. 1.5), we get

⁶² The additional index in the new notation $E_{m,N}$ for the energy of the system of interest reflects the fact that its spectrum is generally dependent on the number N of particles in it.

$$\ln W_{m,N} = S_{\text{env}}(E_{\Sigma}, N_{\Sigma}) - \frac{1}{T} E_{m,N} + \frac{\mu}{T} N + \text{const}. \quad (2.105)$$

Again, exactly as at the derivation of the Gibbs distribution in Sec. 4, we may argue that since $E_{m,N}$, T , and μ do not depend on the choice of environment's size, i.e. on E_{Σ} and N_{Σ} , the probability $W_{m,N}$ for a system to have N particles and be in m^{th} quantum state in the whole grand canonical ensemble should also obey Eq. (105). As a result, we get the so-called *grand canonical distribution*:

$$W_{m,N} = \frac{1}{Z_G} \exp\left\{\frac{\mu N - E_{m,N}}{T}\right\} \equiv \frac{1}{Z_G} \exp\left\{\frac{\mu N}{T}\right\} \exp\left\{-\frac{E_{m,N}}{T}\right\}, \quad (2.106) \quad \text{Grand canonical distribution}$$

where, just as in the case of the Gibbs distribution (2.58), the constant Z_G (most often called the *grand statistical sum*, but sometimes the “grand partition function”) should be determined from the probability normalization condition. However, now the summation of the probabilities $W_{m,N}$ should be over all possible values of both m and N :

$$Z_G = \sum_{m,N} \exp\left\{\frac{\mu N - E_{m,N}}{T}\right\}. \quad (2.107) \quad \text{Grand canonical sum}$$

The last multiplier in the last form of Eq. (106) is the same as in the Gibbs distribution, and its physical interpretation is similar: states are “punished” by lower probability for their excessively higher energy. The handwaving interpretation of the first multiplier, with its opposite sign, is different: in the absence of the energy-related penalty $E_{m,N}$, the environment with an average particle energy $\mu > 0$ “wants” to flood the system with more particles.

Now let us see how the grand canonical distribution may be used for calculations of measurable variables. First, using the general Eq. (29) to calculate the entropy from Eq. (106) (exactly like we did it for the canonical ensemble), we get the following expression,

$$S = -\sum_{m,N} W_{m,N} \ln W_{m,N} = \frac{E}{T} - \frac{\mu \langle N \rangle}{T} + \ln Z_G, \quad (2.108)$$

which is evidently a generalization of Eq. (62).⁶³ We see that now the grand thermodynamic potential Ω (rather than the free energy F) may be expressed directly via the normalization coefficient Z_G :

$$\Omega \equiv F - \mu \langle N \rangle = E - TS - \mu \langle N \rangle = T \ln \frac{1}{Z_G} = -T \ln \sum_{m,N} \exp\left\{\frac{\mu N - E_{m,N}}{T}\right\}. \quad (2.109) \quad \Omega \text{ from } Z_G$$

Finally, solving the last equality for Z_G , and plugging the result back into Eq. (106), we can rewrite the grand canonical distribution in the form

$$W_{m,N} = \exp\left\{\frac{\Omega + \mu N - E_{m,N}}{T}\right\}, \quad (2.110)$$

similar to Eq. (65) for the Gibbs distribution. Indeed, in the particular case when the number N of particles is fixed, $N = \langle N \rangle$, so $\Omega + \mu N = \Omega + \mu \langle N \rangle \equiv F$, Eq. (110) is reduced to Eq. (65).

⁶³ The average number of particles $\langle N \rangle$ is exactly what was called N in thermodynamics (see Chapter 1), but I keep this explicit notation here to make a clear distinction between this average value of the variable, and its particular values participating in Eqs. (102)-(110).

2.8. Systems of independent particles

Now let us apply the general statistical distributions discussed above to a simple but very important case when the system we are considering consists of many similar particles whose direct interactions are negligible. As a result, each particular energy value $E_{m,N}$ of such a system may be represented as a sum of energies ε_k of the particles, where the index k numbers single-particle states – rather than those of the whole system as the index m does.

Let us start with the classical limit. In classical mechanics, the energy quantization effects are negligible, i.e. there is a formally infinite number of quantum states k within each finite energy interval. However, it is convenient to keep, for the time being, the discrete-state language, with the understanding that the average number $\langle N_k \rangle$ of particles in each of these states, usually called the *state occupancy*, is very small. In this case, we may apply the Gibbs distribution to the canonical ensemble of *single particles*, and hence use it with the substitution $E_m \rightarrow \varepsilon_k$, so Eq. (58) becomes

Boltzmann
distribution

$$\langle N_k \rangle = c \exp\left\{-\frac{\varepsilon_k}{T}\right\} \ll 1, \quad (2.111)$$

where the constant c should be found from the normalization condition:

$$\sum_k \langle N_k \rangle = 1. \quad (2.112)$$

This is the famous *Boltzmann distribution*.⁶⁴ Despite its formal similarity to the Gibbs distribution (58), let me emphasize the conceptual difference between these two important formulas. The Gibbs distribution describes the probability to find the *whole system* on one of its states with energy E_m , and it is *always* valid – more exactly, for any canonical ensemble of systems in thermodynamic equilibrium. On the other hand, the Boltzmann distribution describes the occupancy of an energy level of a *single particle*, and, as we will see in just a minute, is valid for quantum particles *only in the classical limit* $\langle N_k \rangle \ll 1$, even if the particles do not interact directly.

The last fact may be surprising, because it may seem that as soon as particles of the system are independent, nothing prevents us from using the Gibbs distribution to derive Eq. (111), regardless of the value of $\langle N_k \rangle$. This is indeed true if the particles are *distinguishable*, i.e. may be distinguished from each other – say by their definitely different spatial positions, or by the states of certain internal degrees of freedom (say, spin), or by any other “pencil mark”. However, it is an experimental fact that elementary particles of each particular type (say, electrons) are *identical* to each other, i.e. cannot be “pencil-marked”.⁶⁵ For such particles we have to be more careful: even if they do not interact *directly*,

⁶⁴ The distribution was first suggested in 1877 by L. Boltzmann. For the particular case when ε is the kinetic energy of a free classical particle (and hence has a continuous spectrum), it is reduced to the *Maxwell distribution* (see Sec. 3.1 below), which was derived earlier – in 1860.

⁶⁵ This fact invites a natural question: what particles are “elementary enough” for their identity? For example, protons and neutrons have an internal structure, in some sense consisting of quarks and gluons; can they be considered elementary? Next, if protons and neutrons are elementary, are atoms? molecules? What about *really* large molecules (such as proteins)? viruses? The general answer to these questions, given by quantum mechanics (or rather experiment :-), is that any particles/systems, no matter how large and complex they are, are identical if they not only have the same internal structure but also are exactly in the same internal quantum state – for example, in the ground state of all their internal degrees of freedom. Evidently, the more complex are the particles/systems, the harder it is to enforce this situation in experiment.

there is still some *indirect* dependence in their behavior, which is especially evident for the so-called *fermions* (elementary particles with semi-integer spin): they obey the *Pauli exclusion principle* that forbids two identical particles to be in the same quantum state, even if they do not interact directly.⁶⁶

Note that the term “the same quantum state” carries a heavy meaning load here. For example, if two particles are confined to stay at different spatial positions (say, reliably locked in different boxes), they are *distinguishable* even if they are internally *identical*. Thus the Pauli principle, as well as other particle identity effects such as the Bose-Einstein condensation to be discussed in the next chapter, are important only when identical particles may move in the same spatial region. To emphasize this fact, it is common to use, instead of “identical”, a more precise (though grammatically rather unpleasant) adjective *indistinguishable*.

In order to take these effects into account, let us examine the statistical properties of a system of many non-interacting but indistinguishable particles (at the first stage of calculation, either fermions or bosons) in equilibrium, applying the grand canonical distribution (109) to a very unusual grand canonical ensemble: a subset of particles in the same quantum state k (Fig. 14).

single-particle energy levels:

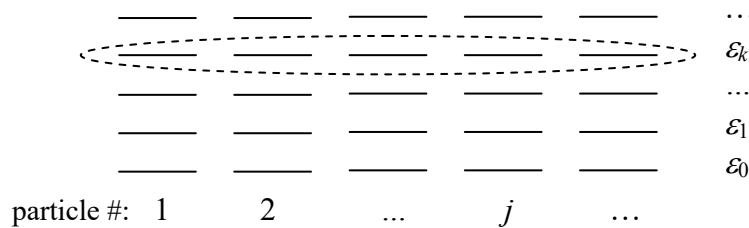


Fig. 2.14. The grand canonical ensemble of particles in the same quantum state with energy ε_k – schematically.

In this ensemble, the role of the environment may be played just by the set of particles in all other states $k' \neq k$, because due to infinitesimal interactions, the particles may gradually change their states. In the resulting equilibrium, the chemical potential μ and temperature T of the system should not depend on the state number k , though the grand thermodynamic potential Ω_k of the chosen particle subset may. Replacing N with N_k – the particular (not average!) number of particles in the selected k^{th} state, and the particular energy value $E_{m,N}$ with $\varepsilon_k N_k$, we reduce the final form of Eq. (109) to

$$\Omega_k = -T \ln \left[\sum_{N_k} \exp \left\{ \frac{\mu N_k - \varepsilon_k N_k}{T} \right\} \right] \equiv -T \ln \left[\sum_{N_k} \left(\exp \left\{ \frac{\mu - \varepsilon_k}{T} \right\} \right)^{N_k} \right], \quad (2.113)$$

where the summation should be carried out over all possible values of N_k . For the final calculation of this sum, the elementary particle type is essential.

On one hand, for *fermions*, obeying the Pauli principle, the numbers N_k in Eq. (113) may take only two values, either 0 (the state k is unoccupied) or 1 (the state is occupied), and the summation gives

$$\Omega_k = -T \ln \left[\sum_{N_k=0,1} \left(\exp \left\{ \frac{\mu - \varepsilon_k}{T} \right\} \right)^{N_k} \right] \equiv -T \ln \left(1 + \exp \left\{ \frac{\mu - \varepsilon_k}{T} \right\} \right). \quad (2.114)$$

⁶⁶ For a more detailed discussion of this issue, see, e.g., QM Sec. 8.1.

Now the state occupancy may be calculated from the last of Eqs. (1.62) – in this case, with the (average) N replaced with $\langle N_k \rangle$:

Fermi-Dirac
distribution

$$\langle N_k \rangle = - \left(\frac{\partial \Omega_k}{\partial \mu} \right)_{T,V} = \frac{1}{e^{(\varepsilon_k - \mu)/T} + 1}. \quad (2.115)$$

This is the famous *Fermi-Dirac distribution*, derived in 1926 independently by Enrico Fermi and Paul Dirac.

On the other hand, *bosons* do not obey the Pauli principle, and for them the numbers N_k can take any non-negative integer values. In this case, Eq. (1.13) turns into the following equality:

$$\Omega_k = -T \ln \left[\sum_{N_k=0}^{\infty} \left(\exp \left\{ \frac{\mu - \varepsilon_k}{T} \right\} \right)^{N_k} \right] \equiv -T \ln \sum_{N_k=0}^{\infty} \lambda^{N_k}, \quad \text{with } \lambda \equiv \exp \left\{ \frac{\mu - \varepsilon_k}{T} \right\}. \quad (2.116)$$

This sum is just the usual geometric series, which converges if $\lambda < 1$, giving

$$\Omega_k = -T \ln \frac{1}{1 - \lambda} \equiv T \ln \left(1 - \exp \left\{ \frac{\mu - \varepsilon_k}{T} \right\} \right), \quad \text{for } \mu < \varepsilon_k. \quad (2.117)$$

In this case, the average occupancy, again calculated using Eq. (1.62) with N replaced with $\langle N_k \rangle$, obeys the *Bose-Einstein distribution*,

Bose-Einstein
distribution

$$\langle N_k \rangle = - \left(\frac{\partial \Omega_k}{\partial \mu} \right)_{T,V} = \frac{1}{e^{(\varepsilon_k - \mu)/T} - 1}, \quad \text{for } \mu < \varepsilon_k, \quad (2.118)$$

which was derived in 1924 by Satyendra Nath Bose (for the particular case $\mu = 0$) and generalized in 1925 by Albert Einstein to the case of arbitrary chemical potential. In particular, comparing Eq. (118) with Eq. (72), we see that harmonic oscillator's *excitations*,⁶⁷ each with energy $\hbar\omega$, may be considered as bosons, with the chemical potential equal to zero. As a reminder, we have already obtained this equality ($\mu = 0$) in a different way – see Eq. (93). Its physical interpretation is that the oscillator excitations may be created inside the system, so there is no energy cost μ of moving them into the system under consideration from its environment.

The simple form of Eqs. (115) and (118), and their similarity (besides “only” the difference of the signs before the unity in their denominators), is one of the most beautiful results of physics. This similarity, however, should not disguise the fact that the energy dependences of the occupancies $\langle N_k \rangle$ given by these two formulas are very much different – see their linear and semi-log plots in Fig. 15.

In the Fermi-Dirac statistics, the level occupancy is not only finite, but is below 1 at any energy, while in the Bose-Einstein it may be above 1, and diverges at $\varepsilon_k \rightarrow \mu$. However, as the temperature is increased, it eventually becomes much larger than the difference ($\varepsilon_k - \mu$). In this limit, $\langle N_k \rangle \ll 1$, so both quantum distributions coincide with each other, as well as with the classical Boltzmann distribution (111) with $c = \exp\{\mu/T\}$:

⁶⁷ As the reader certainly knows, for electromagnetic field oscillators, such excitations are called *photons*; for mechanical oscillation modes, *phonons*, etc. It is important, however, not to confuse such mode *excitations* with the oscillators as such, and be very careful in prescribing to them certain spatial locations – see, e.g., QM Sec. 9.1.

$$\langle N_k \rangle \rightarrow \exp\left\{\frac{\mu - \varepsilon_k}{T}\right\}, \quad \text{for } \langle N_k \rangle \rightarrow 0. \quad (2.119)$$

Boltzmann
distribution:
identical
particles

This distribution (also shown in Fig. 15) may be, therefore, understood also as the high-temperature limit for indistinguishable particles of both sorts.

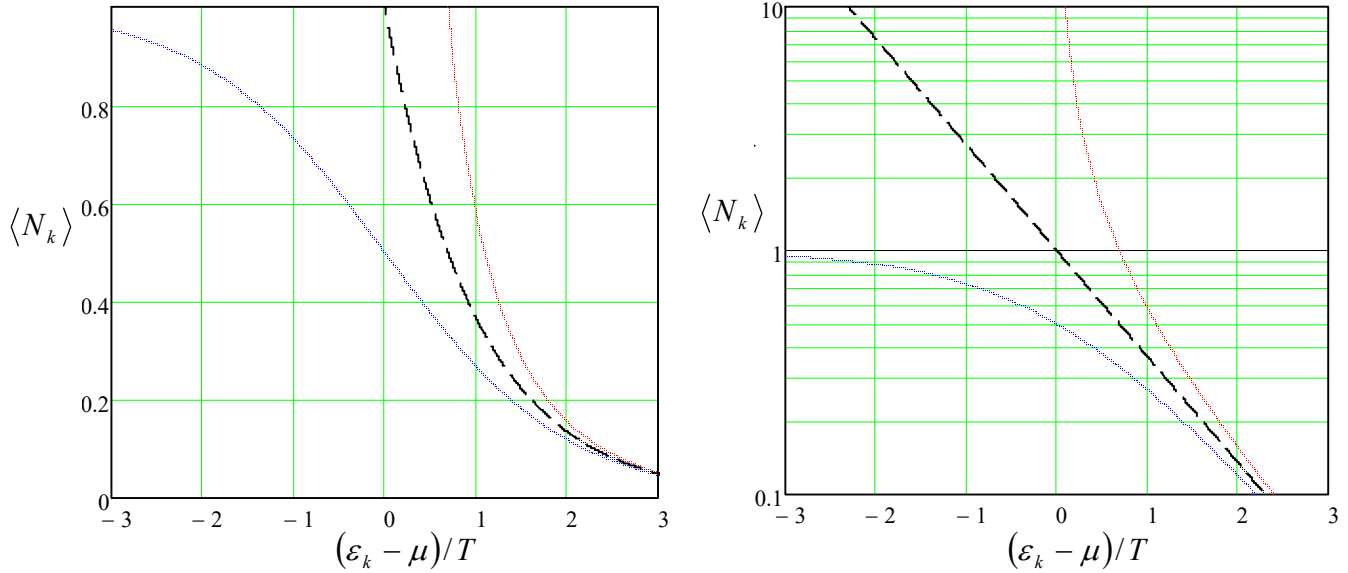


Fig. 2.15. The Fermi-Dirac (blue line), Bose-Einstein (red line), and Boltzmann (dashed line) distributions for indistinguishable quantum particles. (The last distribution is valid only asymptotically, at $\langle N_k \rangle \ll 1$.)

A natural question now is how to find the chemical potential μ participating in Eqs. (115), (118), and (119). In the grand canonical ensemble as such (Fig. 13), with the number of particles variable, the value of μ is imposed by the system's environment. However, both the Fermi-Dirac and Bose-Einstein distributions are also *approximately* applicable (in thermal equilibrium) to systems with a *fixed but very large* number N of particles. In these conditions, the role of the environment for some subset of $N' \ll N$ particles is essentially played by the remaining $N - N'$ particles. In this case, μ may be found by the calculation of $\langle N \rangle$ from the corresponding probability distribution and then requiring the result to be equal to the genuine number of particles in the system. In the next section, we will perform such calculations for several particular systems.

For that and other applications, it will be convenient for us to have ready formulas for the entropy S of a general (i.e. not necessarily equilibrium) state of systems of independent Fermi and Bose particles, expressed not as a function of W_m of the whole system, as in Eq. (29), but via the occupancy numbers $\langle N_k \rangle$. For that, let us consider an ensemble of *composite systems*, each consisting of $M \gg 1$ similar but distinct *component systems*, numbered by index $m = 1, 2, \dots, M$, with independent (i.e. not directly interacting) particles – see Fig. 16. Let us assume that though in each of the M component systems, the number $N_k^{(m)}$ of particles in their k^{th} quantum state may be different, their total number $N_k^{(\Sigma)}$ in the composite system is fixed. As a result, the total energy of the composite system is fixed as well,

$$\sum_{m=1}^M N_k^{(m)} = N_k^{(\Sigma)} = \text{const}, \quad E_k = \sum_{m=1}^M N_k^{(m)} \varepsilon_k = N_k^{(\Sigma)} \varepsilon_k = \text{const}, \quad (2.120)$$

so an ensemble of many such composite systems (with the same k), in equilibrium, is microcanonical. According to Eq. (24a), the average entropy S_k per component system in this microcanonical ensemble may be calculated as

$$S_k = \lim_{M \rightarrow \infty} \frac{\ln M_k}{M}, \quad (2.121)$$

where M_k is the number of possible different ways such a composite system (with fixed $N_k^{(\Sigma)}$) may be implemented.

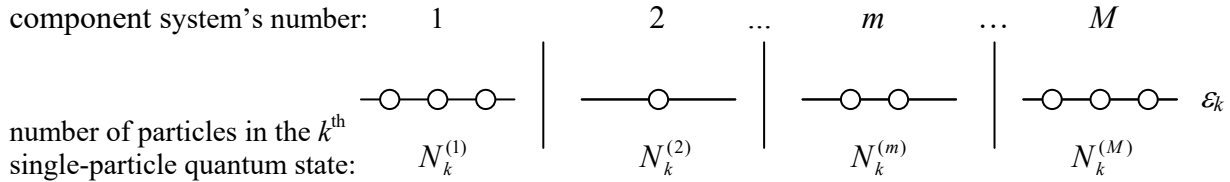


Fig. 2.16. An example of a composite system of $N_k^{(\Sigma)}$ particles in the k^{th} quantum state, distributed between M component systems.

Let us start with the calculation of M_k for Fermi particles – for which the Pauli principle is valid. Here the level occupancies $N_k^{(m)}$ may be only equal to either 0 or 1, so the distribution problem is solvable only if $N_k^{(\Sigma)} \leq M$, and evidently equivalent to the choice of $N_k^{(\Sigma)}$ balls (in *arbitrary* order) from the total number of M *distinct* balls. Comparing this formulation with the definition of the binomial coefficient,⁶⁸ we immediately get

$$M_k = {}^M C_{N_k^{(\Sigma)}} = \frac{M!}{(M - N_k^{(\Sigma)})! N_k^{(\Sigma)}!}. \quad (2.122)$$

From here, using the Stirling formula (again, in its simplest form (27)), we get

Fermions:
entropy

$$S_k = -\langle N_k \rangle \ln \langle N_k \rangle - (1 - \langle N_k \rangle) \ln(1 - \langle N_k \rangle), \quad (2.123)$$

where

$$\langle N_k \rangle \equiv \lim_{M \rightarrow \infty} \frac{N_k^{(\Sigma)}}{M} \quad (2.124)$$

is exactly the average occupancy of the k^{th} single-particle state in each system, which was discussed earlier in this section. Since for a Fermi system, $\langle N_k \rangle$ is always between 0 and 1, its entropy (123) cannot be negative – see the blue line in Fig. 17.

In the Bose case, where the Pauli principle is not valid, the number $N_k^{(m)}$ of particles in the k^{th} state in each of the systems is an arbitrary (non-negative) integer. Let us consider $N_k^{(\Sigma)}$ particles and $(M - 1)$ partitions (shown by vertical lines in Fig. 16) between M systems as $(M - 1 + N_k^{(\Sigma)})$ mathematical objects ordered along one axis. Each specific location of the partitions evidently fixes all $N_k^{(m)}$. Hence M_k may be calculated as the number of possible ways to distribute the $(M - 1)$ indistinguishable partitions among these $(M - 1 + N_k^{(\Sigma)})$ ordered objects, i.e. as the following binomial coefficient:⁶⁹

⁶⁸ See, e.g., MA Eq. (2.2).

⁶⁹ See also MA Eq. (2.4).

$$M_k = {}^{M+N_k-1}C_{M-1} = \frac{(M-1+N_k^{(\Sigma)})!}{(M-1)!N_k^{(\Sigma)}!}. \quad (2.125)$$

Applying the Stirling formula (27) again, we get the following result,

$$S_k = -\langle N_k \rangle \ln \langle N_k \rangle + (1 + \langle N_k \rangle) \ln(1 + \langle N_k \rangle), \quad (2.126) \text{ Bosons: entropy}$$

which again differs from the Fermi case (123) “only” by the signs in the second term, and is valid for any positive $\langle N_k \rangle$ - see the red line in Fig. 17.

In the classical limit when the average occupancies $\langle N_k \rangle$ of the state is much smaller than 1, the Fermi and Bose expressions for S_k tend to the same *Boltzmann limit*:

$$S_k \rightarrow -\langle N_k \rangle (\ln \langle N_k \rangle - 1) \rightarrow -\langle N_k \rangle \ln \langle N_k \rangle, \quad \text{for } \langle N_k \rangle \rightarrow 0. \quad (2.127)$$

(The last expression may be also obtained from the functionally similar Eq. (29), by considering an ensemble of systems consisting of just one classical particle each, so $E_m \rightarrow \varepsilon_k$ and $W_m \rightarrow \langle N_k \rangle$.)

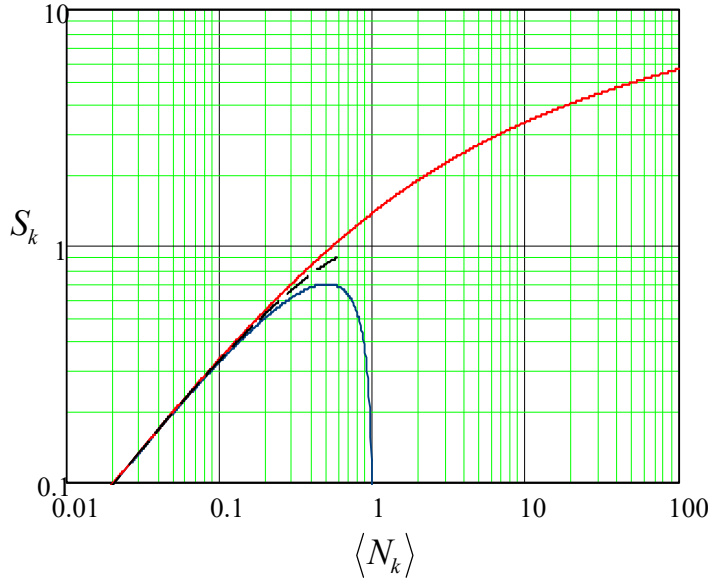


Fig. 2.17. Entropy of particles in a quantum state as a function of its average occupancy $\langle N_k \rangle$, for fermions (blue line) and bosons (red line). The dashed line shows their common asymptote at $\langle N_k \rangle \rightarrow 0$, given by the first of Eqs. (127).

Expressions (123) and (126) are valid for an arbitrary (possibly, non-equilibrium) case; they may be also used for an alternative derivation of the Fermi-Dirac (115) and Bose-Einstein (118) distributions, which are valid only in equilibrium. For that, we may use the method of Lagrange multipliers, requiring (just like it was done in Sec. 2) the total entropy of a system of N independent, similar particles,

$$S = \sum_k S_k, \quad (2.128)$$

considered as a function of state occupancies $\langle N_k \rangle$, to attain its maximum, under the conditions of the fixed total number of particles N and total energy E :

$$\sum_k \langle N_k \rangle = N = \text{const}, \quad \sum_k \langle N_k \rangle \varepsilon_k = E = \text{const}. \quad (2.129)$$

The completion of this calculation is left for the reader’s exercise.

2.9. Exercise problems

2.1. A famous example of macroscopic irreversibility was suggested in 1907 by P. Ehrenfest. Two dogs share $2N \gg 1$ fleas. Each flea may jump onto another dog, and the rate Γ of such events (i.e. the probability of jumping per unit time) does not depend either on time or on the location of other fleas. Find the time evolution of the average number of fleas on a dog, and of the flea-related part of the total dogs' entropy (at arbitrary initial conditions), and prove that the entropy can only grow.⁷⁰

2.2. Use the microcanonical distribution to calculate thermodynamic properties (including the entropy, all relevant thermodynamic potentials, and the heat capacity) of a two-level system in thermodynamic equilibrium with its environment, at a temperature T that is comparable with the energy gap Δ . For each variable, sketch its temperature dependence, and find its asymptotic values (or trends) in the low-temperature and high-temperature limits.

Hint: The two-level system is any quantum system with just two different stationary states, whose energies (say, E_0 and E_1) are separated by a gap $\Delta \equiv E_1 - E_0$. Its popular (but by no means the only!) example is the spin- $1/2$ of a particle, e.g., an electron, in an external magnetic field.⁷¹

2.3. Solve the previous problem using the Gibbs distribution. Also, calculate the probabilities of the energy level occupation, and give physical interpretations of your results, in both temperature limits.

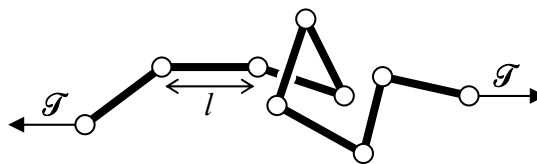
2.4. A quantum spin- $1/2$ particle with a gyromagnetic ratio γ is placed into an external magnetic field $\mathcal{H} = \mathcal{H} \mathbf{n}_z$. Neglecting the possible orbital motion of the particle, calculate its average magnetization $\langle m_z \rangle$ as a function \mathcal{H} , and in particular its low-field magnetic susceptibility χ , in thermal equilibrium at temperature T . Calculate the same characteristics for a classical magnetic dipole \mathbf{m} of a fixed magnitude m_0 , free to change its direction in space, and compare the results.

Hint: The low-field magnetic susceptibility of a single particle is defined⁷² as

$$\chi = \left. \frac{\partial \langle m_z \rangle}{\partial \mathcal{H}} \right|_{\mathcal{H} \rightarrow 0}.$$

2.5.* Calculate the weak-field magnetic susceptibility of a hydrogen atom, at room temperature. Is this response to the field paramagnetic or diamagnetic? Compare the result with the estimated susceptibility of a hydrogen *molecule* H_2 .

2.6. N similar stiff rods of length l are connected with the joints that allow for free 3D rotation, to form a chain – see the figure on the right. The chain, in thermal



⁷⁰ This is essentially a simpler (and funnier :-)) version of the particle scattering model used by L. Boltzmann to prove his famous *H-theorem* (1872). Besides the historical significance of that theorem, the model used in it (see Sec. 6.2 below) is as cartoonish, and not more general.

⁷¹ See, e.g., QM Secs. 4.6 and 5.1, in particular, Eq. (4.167).

⁷² This “atomic” (or “molecular”) susceptibility χ should not be confused with the “volumic” susceptibility $\chi_m \equiv \partial \mathcal{M}_z / \partial \mathcal{H}$, where \mathcal{M} is the magnetization, i.e. the magnetic moment of a unit volume of a system – see, e.g., EM Eq. (5.111). For a uniform medium with $n \equiv N/V$ non-interacting dipoles per unit volume, $\chi_m = n\chi$.

equilibrium at temperature T , is stretched by a fixed force \mathcal{F} . Calculate the spring constant κ of the chain in the elastic limit $\mathcal{F} \rightarrow 0$.

2.7. Calculate the low-field magnetic susceptibility of a particle with an arbitrary (either integer or semi-integer) spin s , neglecting its orbital motion. Compare the result with the solution of Problem 4.

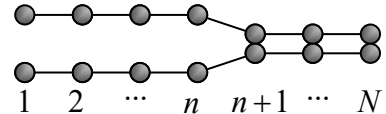
Hint: Quantum mechanics⁷³ tells us that the Cartesian component m_z of the magnetic moment of such a particle, in the direction of the applied field, has $(2s + 1)$ stationary values:

$$m_z = \gamma \hbar m_s, \quad \text{with } m_s = -s, -s + 1, \dots, s - 1, s,$$

where γ is the gyromagnetic ratio of the particle, and \hbar is Planck's constant.

2.8. Analyze the possibility of using a system of non-interacting spin- $\frac{1}{2}$ particles placed into a controllable external magnetic field, for refrigeration.

2.9. A rudimentary “zipper” model of DNA replication is a chain of N links that may be either open or closed – see the figure on the right. Opening a link increases the system's energy by $\Delta > 0$; a link may change its state (either open or closed) only if all links to the left of it are open, while all those on the right of it, are closed. Calculate the average number of open links in thermal equilibrium, and analyze its temperature dependence, especially for the case $N \gg 1$.



2.10. Use the microcanonical distribution to calculate the average entropy, energy, and pressure of a classical 3D particle of mass m , with no internal degrees of freedom, free to move in volume V , at temperature T .

Hint: Try to make a more accurate calculation than has been done in Sec. 2.2 for the system of N harmonic oscillators. For that, you would need to know the volume V_d of a d -dimensional hypersphere of a unit radius. To avoid being too cruel, I am providing it:

$$V_d = \pi^{d/2} / \Gamma\left(\frac{d}{2} + 1\right),$$

where $\Gamma(\xi)$ is the gamma function.⁷⁴

2.11. Solve the previous problem using the Gibbs distribution.

2.12. Calculate the average energy, entropy, free energy, and the equation of state of a classical 2D particle (without internal degrees of freedom), free to move within area A , at temperature T , starting from:

- (i) the microcanonical distribution, and
- (ii) the Gibbs distribution.

Hint: For the equation of state, make the appropriate modification of the notion of pressure.

⁷³ See, e.g., QM Sec. 5.7, in particular Eq. (5.169).

⁷⁴ For its definition and main properties, see, e.g., MA Eqs. (6.6)-(6.9).

2.13. A quantum particle of mass m is confined to free motion along a 1D segment of length a . Using any approach you like, calculate the average force the particle exerts on the “walls” (ends) of such a “1D potential well” in thermal equilibrium, and analyze its temperature dependence, focusing on the low-temperature and high-temperature limits.

Hint: You may consider the series $\Theta(\xi) \equiv \sum_{n=1}^{\infty} \exp\{-\xi n^2\}$ a known function of ξ .⁷⁵

2.14. Rotational properties of diatomic molecules (such as N_2 , CO , etc.) may be reasonably well described by the so-called *dumbbell model*: two point particles, of masses m_1 and m_2 , with a fixed distance d between them. Ignoring the translational motion of the molecule as a whole, use this model to calculate its heat capacity, and spell out the result in the limits of low and high temperatures. Is your solution valid for the so-called *homonuclear* molecules consisting of two similar atoms, such as H_2 , O_2 , N_2 , etc.?

2.15.* Modify the solution of the previous problem for homonuclear molecules. Specifically, consider the cases of molecules H_2 and N_2 . For the first of them, compute the equilibrium ratio of the number of the ortho- and parahydrogen molecules as a function of temperature.

Hint: Use the value of d that gives the experimentally observed difference of 1.455 kJ/mol between the ground state energies of these two hydrogen species (“spin isomers”).

2.16. Calculate the heat capacity of a heteronuclear diatomic molecule by using the simple model described in Problem 14, but now assuming that the rotation is confined to one plane.⁷⁶

2.17. A classical, rigid, strongly elongated body (such as a thin needle) is free to rotate about its center of mass and is in thermal equilibrium with its environment. Are the angular velocity vector $\boldsymbol{\omega}$ and the angular momentum vector \mathbf{L} , on average, directed along the elongation axis of the body, or normal to it?

2.18. Two similar classical electric dipoles, of a fixed magnitude d , are separated by a fixed distance r . Assuming that each dipole moment vector \mathbf{d} may point in any direction and that the system is in thermal equilibrium, write general expressions for its statistical sum Z , average interaction energy E , heat capacity C , and entropy S , and calculate them explicitly in the high-temperature limit.

2.19. A classical 1D particle of mass m , residing in the potential well

$$U(x) = \alpha|x|^\gamma, \quad \text{with } \gamma > 0,$$

is in thermal equilibrium with its environment, at temperature T . Calculate the average values of its potential energy U and the full energy E :

- (i) directly from the Gibbs distribution, and
- (ii) by using the virial theorem of classical mechanics⁷⁷ and the equipartition theorem.

⁷⁵ It may be reduced to the so-called elliptic theta-function $\theta_3(z, \tau)$ for a particular case $z = 0$ – see, e.g., Sec. 16.27 in the Abramowitz-Stegun handbook cited in MA Sec. 16(ii). However, you do not need that (or any other) handbook to solve this problem.

⁷⁶ This is a reasonable model of the constraints imposed on small atomic groups (e.g., ligands) by their atomic environment inside some large molecules.

2.20. For a slightly anharmonic classical 1D oscillator with mass m and potential energy

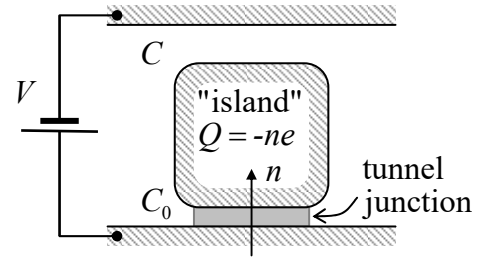
$$U(x) = \frac{\kappa}{2}x^2 + \alpha x^3$$

with a small coefficient α , in thermal equilibrium with its environment, calculate:

- (i) the statistical average of the coordinate x , and
- (ii) the deviation of the heat capacity from its basic value $C=1$,

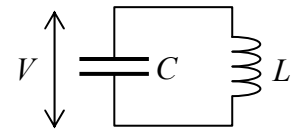
in the first (linear) approximation in low temperature T .

2.21. A small conductor (in this context, usually called the *single-electron island*) is placed between two conducting electrodes, with voltage V applied between them. The gap between one of the electrodes and the island is so narrow that electrons may tunnel quantum-mechanically through this gap (the “weak tunnel junction”) – see the figure on the right. Calculate the average charge of the island as a function of V at temperature T .



Hint: The quantum-mechanical tunneling of an electron through a weak junction⁷⁸ between two macroscopic conductors and its subsequent energy relaxation, may be treated as a single inelastic (energy-dissipating) event, so the only energy essential for the thermal equilibrium of the system is its electrostatic potential energy.

2.22. A lumped LC circuit (see the figure on the right) is in thermodynamic equilibrium with its environment. Calculate the r.m.s. fluctuation $\delta V \equiv \langle V^2 \rangle^{1/2}$ of the voltage across it, for an arbitrary ratio $T/\hbar\omega$, where $\omega = (LC)^{-1/2}$ is the resonance frequency of this “tank circuit”.



2.23. Derive Eq. (92) from simplistic arguments, by representing the blackbody radiation as an ideal gas of photons treated as classical ultra-relativistic particles. What do similar arguments give for an ideal gas of classical but non-relativistic particles?

2.24. Calculate the enthalpy, the entropy, and the Gibbs energy of blackbody electromagnetic radiation in thermal equilibrium with temperature T inside volume V , and then use these results to find the law of temperature and pressure drop at an adiabatic expansion.

2.25. As was mentioned in Sec. 6(i), the relation between the temperature T_{\odot} of the visible Sun’s surface and that (T_{\oplus}) of the Earth’s surface follows from the balance of the thermal radiation they emit. Prove that the experimentally observed relation indeed follows, with good precision, from a simple model in which the surfaces radiate as perfect black bodies with constant temperatures.

Hint: You may pick up the experimental values you need from any reliable source.

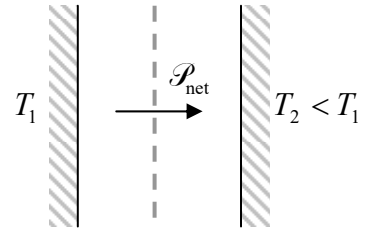
⁷⁷ See, e.g., CM Problem 1.12.

⁷⁸ In this particular context, the adjective “weak” denotes a junction with a tunneling transparency so low that the tunneling electron’s wavefunction loses its quantum-mechanical coherence before the electron has a chance to tunnel back. In a typical junction of a macroscopic area, this condition is fulfilled if its effective resistance is much higher than the quantum unit of resistance (see, e.g., QM Sec. 3.2), $R_Q \equiv \pi\hbar/2e^2 \approx 6.5 \text{ k}\Omega$.

2.26. If the surface is not perfectly radiation-absorbing (“black”), the electromagnetic power of its thermal radiation differs from the Planck radiation law by a frequency-dependent factor $\varepsilon < 1$, called *emissivity*. Prove that such a surface reflects the $(1 - \varepsilon)$ fraction of the incident radiation.

2.27. If two black surfaces, facing each other, have different temperatures (see the figure on the right), then according to the Stefan radiation law (89), there is a net flow of thermal radiation, from the hotter surface to the colder one:

$$\frac{\mathcal{P}_{\text{net}}}{A} = \sigma(T_1^4 - T_2^4).$$



For many applications, notably including most low-temperature experiments, this flow is detrimental. One way to suppress it is to reduce the emissivity ε (for its definition, see the previous problem) of both surfaces – say by covering them with shiny metallic films. An alternative way toward the same goal is to place, between the surfaces, a thin layer (usually called the *thermal shield*), with a low emissivity of both surfaces – see the dashed line in Fig. above. Assuming that the emissivity is the same in both cases and neglecting its possible dependence on the angle and frequency, find out which way is more efficient.

2.28. Two perfectly reflecting parallel plates of area A are separated by a free-space gap of a constant thickness $t \ll A^{1/2}$. Calculate the energy of the thermally-induced electromagnetic field inside the gap in thermal equilibrium, with temperature T in the range

$$\frac{\hbar c}{A^{1/2}} \ll T \ll \frac{\hbar c}{t}.$$

Does the field push the plates apart?

2.29. Use the Debye theory to estimate the specific heat of aluminum at room temperature (say, 300 K) and express the result in the following popular units:

- (i) eV/K per atom,
- (ii) J/K per mole, and
- (iii) J/K per gram.

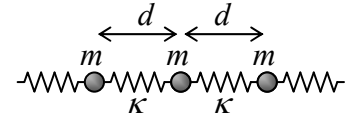
Compare the last number with the experimental value (from a reliable source).

2.30. Low-temperature specific heat of some solids has a considerable contribution from the thermal excitation of spin waves, whose dispersion law at $\omega \rightarrow 0$ scales as $\omega \propto k^2$.⁷⁹ Neglecting anisotropy, calculate the temperature dependence of this contribution to C_V at low temperatures, and discuss conditions of its experimental observation.

Hint: Just as the photons and phonons discussed in section 2.6, the quantum excitations of spin waves (called *magnons*) may be considered non-interacting bosonic quasiparticles with zero chemical potential, whose statistics obeys Eq. (2.72).

⁷⁹ Note that the same dispersion law is typical for bending waves in thin elastic rods – see, e.g., CM Sec. 7.8.

2.31. Derive a general expression for the specific heat of a very long straight chain of similar particles of mass m , confined to move only in the direction of the chain and elastically interacting with effective spring constants κ – see the figure on the right. Spell out the result in the limits of very low and very high temperatures. Would using the Debye approximation change these results?



Hint: You may like to use the following integral:⁸⁰ $\int_0^{+\infty} \frac{\xi^2 d\xi}{\sinh^2 \xi} = \frac{\pi^2}{6}$.

2.32. Use the Debye approximation to obtain a general expression for the longitudinal phonon contribution to the specific heat of a stand-alone monatomic layer of an elastic material (such as graphene). Find its explicit temperature dependence at $T \rightarrow 0$.

2.33. Calculate the r.m.s. thermal fluctuation of an arbitrary point of a uniform guitar string of length l , stretched by force \mathcal{F} , at temperature T . Evaluate your result for $l = 0.7$ m, $\mathcal{F} = 10^3$ N, and room temperature.

Hint: You may like to use the following series: $\sum_{n=1}^{\infty} \frac{\sin^2(\pi n \xi)}{(\pi n)^2} = \frac{\xi(1-\xi)}{2}$, for $0 \leq \xi \leq 1$.

2.34. Use the general Eq. (123) to re-derive the Fermi-Dirac distribution (115) for a system in equilibrium.

2.35. Each of two identical particles, not interacting directly, may be in any of two quantum states, with the single-particle energies ε equal to 0 and Δ . Write down the statistical sum Z of the system, and use it to calculate its average total energy E at temperature T , for the cases when the particles are:

- (i) distinguishable (say, by their spatial positions);
- (ii) indistinguishable fermions;
- (iii) indistinguishable bosons.

Analyze and interpret the temperature dependence of $\langle E \rangle$ for each case, assuming that $\Delta > 0$.

2.36. Each of $N \gg 1$ indistinguishable fermions has two non-degenerate energy levels separated by gap Δ . Calculate the chemical potential of their system in thermal equilibrium at temperature T , if the direct interaction of the particles is negligible.

⁸⁰ It may be reduced, via integration by parts, to the table integral MA Eq. (6.8d) with $n = 1$.