

Ensembles

from Statistical Physics using Mathematica

© James J. Kelly, 1996-2002

Two methods for constructing canonical probability distributions are presented. The first is based upon thermal interaction between a sample and a much larger reservoir of heat that determines the temperature of the sample. The second maximizes the entropy of an ensemble subject to constraints upon its mean energy and perhaps other variables. Application of these methods to several prototypical systems illustrates important general features of thermostatistics.

Microcanonical ensemble

■ Definition

The simplest and most fundamental ensemble is constructed as the set of isolated systems, each prepared in a unique stationary quantum state, such that each and every state within the energy range $\{E, E + \delta E\}$ is represented exactly once. The stipulations that each member of the ensemble is prepared in a stationary state and that the systems are isolated from all external interactions imply that the state of each system is constant and the ensemble is stationary. Since each microstate within the specified energy range is represented once and only once, the statistical weight assigned to each member of the ensemble is simply the inverse of the total number of states available. Therefore, the probability P_i of selecting a particular state i at random from a microcanonical ensemble containing $\Gamma[E, \delta E]$ states is simply

$$P_i = \frac{1}{\Gamma[E, \delta E]} \quad \text{for } E \leq E_i \leq E + \delta E$$
$$P_i = 0 \quad \text{otherwise}$$

Ensemble averages for property \hat{x} are then computed as unweighted averages over the values x_i for each of the Γ members of the ensemble, according to

$$\langle \hat{x} \rangle = \sum_{i \in \Gamma} P_i x_i = \frac{1}{\Gamma} \sum_{i \in \Gamma} x_i$$

The thermodynamic internal energy U is the ensemble average of \hat{H} , such that

$$U = \langle \hat{H} \rangle = \sum_{i \in \Gamma} P_i E_i = \frac{1}{\Gamma} \sum_{i \in \Gamma} E_i$$

Similarly, if the volume or particle number vary within the ensemble, the thermodynamic volume and particle number are also interpreted as ensemble averages.

Entropy in a microcanonical ensemble is obtained directly from the multiplicity function $\Gamma[E, \delta E] = g[E] \delta E$ according to

$$S = k_B \ln \Gamma \approx k_B \ln g$$

where the energy interval is usually sufficiently small that for systems with large multiplicity there is no practical (logarithmic) difference between the total number of states Γ within the interval and the density of states $g[E]$ at the central energy of the interval. Similarly, if this interval is small, there is no practical difference between E and U , so that $g[E] \rightarrow g[U]$. Thus, the fundamental thermodynamic relation for the microcanonical ensemble specifies the dependence of entropy upon internal energy and any other thermodynamic variables needed to specify the ensemble. All other thermodynamic properties of the system can be obtained from this fundamental relation. For example, the temperature can be obtained using

$$\frac{1}{T} = \frac{\partial S}{\partial U} \implies \beta = \frac{1}{k_B T} = \frac{\partial \ln g[U]}{\partial U}$$

where other variables are held constant. In the microcanonical ensemble temperature measures the energy dependence of the multiplicity function for isolated systems. For isolated systems, you specify the mean energy and then the internal dynamics decide the temperature.

The next few sections provide examples of the application of the microcanonical ensemble to prototypical systems which illustrate some important features of thermostatistics. Here we summarize and discuss the results, leaving technical details and more thorough development to separate notebooks.

■ Example: binary systems

Consider a system of N weakly interacting elements for which each can occupy only two states. For example, if we neglect kinetic energies and mutual interactions, a system of N spin $\frac{1}{2}$ atoms or nuclei subject to an external magnetic field can be approximated as an ideal two-level system. Suppose that N_1 elements are found in the state with energy $\varepsilon_1 = -\varepsilon$ and N_2 in the state with energy $\varepsilon_2 = +\varepsilon$. It is convenient to define an order parameter $-1 \leq x \leq 1$, or alignment variable, to represent the excess population of the lower level

$$\begin{aligned} N_1 &= \frac{N}{2} (1 + x) & N &= N_1 + N_2 \\ N_2 &= \frac{N}{2} (1 - x) & x &= (N_1 - N_2)/N \end{aligned}$$

such that the energy becomes

$$U = N_1 \varepsilon_1 + N_2 \varepsilon_2 = -N x \varepsilon$$

The number of states with N_1 elements in level ε_1 and $N - N_1$ elements in level ε_2 is given by the binomial coefficient

$$g[x] = \binom{N}{N_1} = \frac{N!}{N_1! N_2!} = \frac{N!}{(\frac{N}{2} (1+x))! (\frac{N}{2} (1-x))!}$$

The entropy $S[x]$ is proportional to the logarithm of the multiplicity, which with the aid of Stirling's formula and straightforward algebra becomes

$$S = N k_B \left(\frac{1+x}{2} \text{Log} \left[\frac{2}{1+x} \right] + \frac{1-x}{2} \text{Log} \left[\frac{2}{1-x} \right] \right)$$

The temperature is determined by the thermodynamic relation

$$\frac{1}{T} = \frac{\partial S}{\partial U} = -\frac{1}{N \varepsilon} \frac{\partial S}{\partial x} \implies \frac{\varepsilon}{k_B T} = \frac{1}{2} \text{Log} \left[\frac{1+x}{1-x} \right]$$

Solving for the order parameter, we find

$$x = \text{Tanh}[\beta \varepsilon] \implies U = -N \varepsilon \text{Tanh}[\beta \varepsilon]$$

where $\beta \equiv 1/k_B T$. After some tedious algebra, we can express entropy and free energy in terms of temperature as

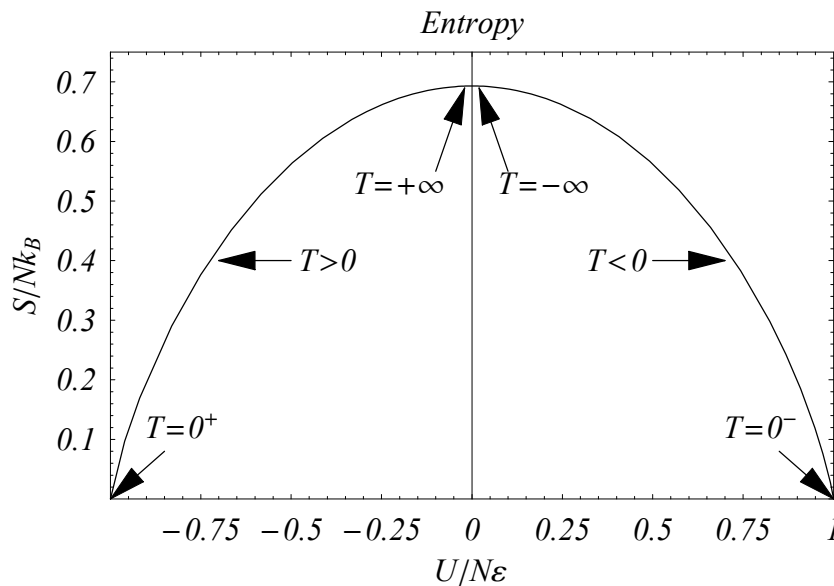
$$S = N k_B (\text{Log}[2 \text{Cosh}[\beta \varepsilon]] - \beta \varepsilon \text{Tanh}[\beta \varepsilon]) \implies F = -N k_B T \text{Log}[2 \text{Cosh}[\beta \varepsilon]]$$

Finally, the heat capacity becomes

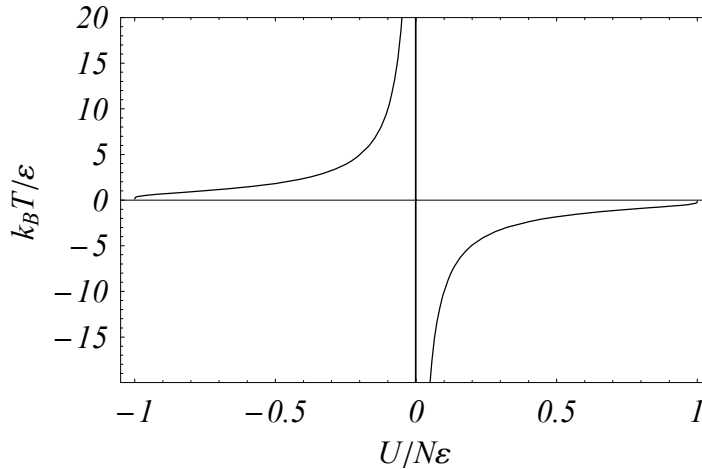
$$C = \frac{\partial U}{\partial T} = N k_B (\beta \varepsilon \text{Sech}[\beta \varepsilon])^2$$

where we assume that any external parameters that govern ε , such as magnetic field, are held fixed.

The figure below shows the dependence of entropy upon the order parameter. When $x \rightarrow +1$ all elements are found in the lower energy level and the system has its lowest possible energy. The entropy vanishes when the system is completely ordered and the temperature approaches zero, $T \rightarrow 0^+$. More states become available as the energy and temperature of the system increases. As $T \rightarrow \infty$ both upper and lower states are populated equally and the entropy approaches its maximum possible value, $S_{\text{max}} = N k_B \ln 2$. The temperature becomes infinite at $U = 0$ because the slope of $S[U]$ vanishes, but is positive for slightly negative U and appears to be negative for slightly positive U . As the energy increases further an increasing fraction of the elements are found in the upper level and the entropy decreases as the system become increasingly ordered. In this region the temperature appears to be negative because the slope of $S[U]$ is negative. Finally, the entropy vanishes as the energy approaches its maximum possible value, corresponding to $T \rightarrow 0^-$.



The concept of negative temperature can be disconcerting, at first, but it is a characteristic feature of isolated systems with bounded ranges of energy and entropy. Any system for which the entropy is a convex function of energy, as shown above, can find itself with entropy decreasing as the upper limit of the energy range is approached. If the slope of $S[U]$ is negative, so is the temperature. However, one would naturally describe increasing energy as heating, such that states with negative temperature are actually hotter than states with positive temperature. The figure below shows the dependence of temperature on internal energy: $T = 0^+$ is cold, T increases as the system is heated, and $T \rightarrow +\infty$ as both states are populated equally. As the system is heated further the temperature *increases* from $+\infty$ to $-\infty$, which are infinitesimally close in energy, and then becomes less negative until the hottest temperature is reached at $T = 0^-$.

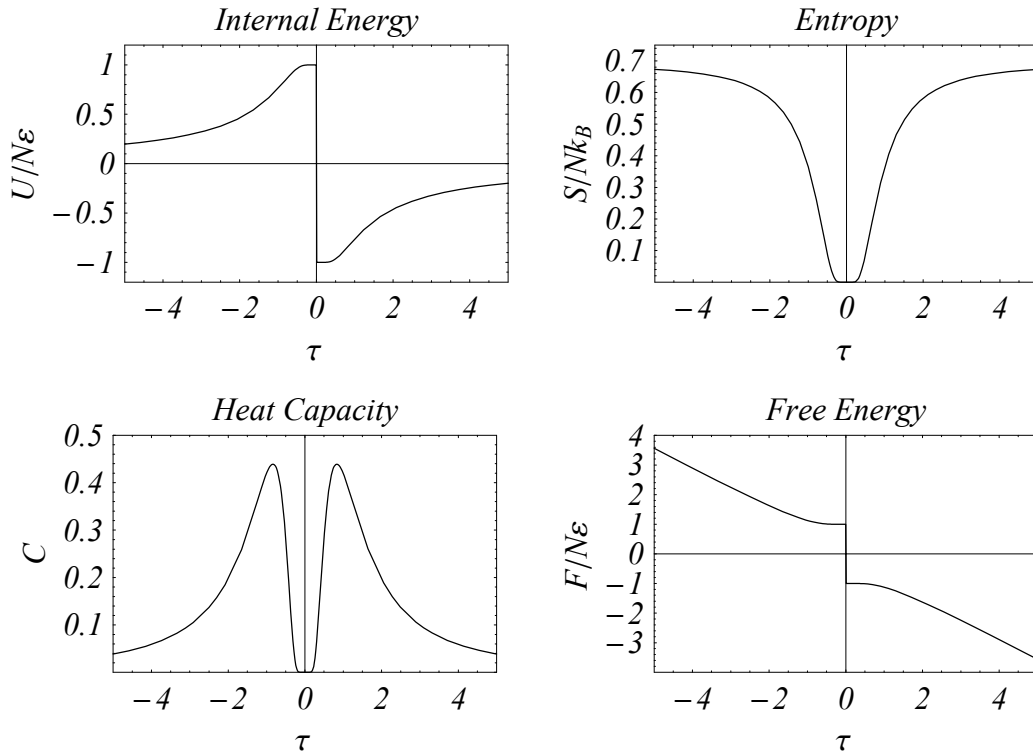


The classic example of negative temperature is provided by nuclear spin systems which interact so weakly with the surrounding crystal lattice that macrostates with nuclear magnetization opposed to the external magnetic field can be maintained in constrained or partial equilibrium for a long time. This phenomenon was first demonstrated by Purcell and Pound. The nuclear spins were aligned parallel to a strong magnetic field. The field was then reversed so rapidly that the spins remained in their original orientation and found themselves aligned antiparallel with the magnetic field. Hence, the nuclear spin system is characterized by negative temperature. The spin system then gradually gives up its excess internal energy and returns to a state aligned with the external field with a relaxation time of the order of 5 minutes. Although the macrostate changes and is not truly in equilibrium at negative temperature, the relaxation time is so long compared with atomic time scales that it can be considered to be a quasistatic process that is always sufficiently near equilibrium to apply the concepts and methods of equilibrium statistical mechanics.

A bounded system with greater population of the upper energy level is described as having a *population inversion*. Such systems are not normally in a true equilibrium because it is not possible in practice to isolate a system from its environment. The temperature of the environment will always be positive because the energy of any statistically normal system with kinetic contributions is unbounded. Recognizing that the any system with negative temperature is hotter than any system with positive temperature, heat will flow from the system with negative temperature until it eventually reaches thermal equilibrium at positive temperature. Nevertheless, if the relaxation time is sufficiently long, such as some nuclear spin systems, the system can be maintained in *partial* or *constrained equilibrium* long enough to observe negative temperature. Other systems, such as lasers, require active pumping to maintain the population inversion.

For a system whose energy is limited to a finite range, the entropy will generally be small at the extreme values of the internal energy and maximal somewhere in between, as seen above for the two-level system. The temperature, $T = (\frac{\partial S}{\partial U})^{-1}$, will then be positive for U below the entropy peak, negative above, and infinite at the peak where $\frac{\partial S}{\partial U} = 0$. The figure below illustrates the dependence of temperature upon internal energy and shows that there is a singularity at $U = 0$. The difference in energy between states with very large negative temperature and states with very large positive temperature is actually very small. Therefore, the heat capacity must vanish at the entropy peak because an infinitesimal change in energy produces an infinite change in temperature.

The principal thermodynamic functions for a binary system are plotted below as functions of a dimensionless *reduced temperature* variable $\tau \equiv k_B T / \varepsilon$ defined as the ratio between the thermal energy $k_B T$ and the characteristic excitation energy ε . The discontinuity in internal and free energies at $\tau = 0$ is associated with the zero entropy for fully ordered systems in which all elements are found either in the lower state for $\tau \rightarrow 0^+$ or the upper state for $\tau \rightarrow 0^-$. The heat capacity vanishes at $\tau = 0$ because a fully ordered system cannot absorb heat without a finite change of temperature. Similarly, the capacity also vanishes at $T = \pm \infty$ because a system with maximum entropy has $\frac{\partial S}{\partial T} = 0 \implies C = 0$. Therefore, the heat capacity must exhibit a maximum for intermediate temperatures. This characteristic feature of systems with bounded energy is called a *Schottky anomaly*; it is anomalous because the heat capacity for normal systems with unlimited capacity to absorb energy increases monotonically with temperature.



Finally, it is useful to examine the limiting behaviors of these functions. For high temperatures we find

$$\begin{aligned}\frac{S}{N k_B} &\sim \ln 2 - \frac{y^2}{2} + \frac{y^4}{4} + \dots \\ \frac{U}{N \epsilon} &\sim -y + \frac{y^2}{3} + \dots \\ \frac{C}{N k_B} &\sim y^2 - y^4\end{aligned}$$

while at low temperatures

$$\begin{aligned}\frac{S}{N k_B} &\approx e^{-2y} (1 + 2y) \\ \frac{U}{N \epsilon} &\approx -(1 - 2e^{-2y}) \\ \frac{C}{N k_B} &\approx 4y^2 e^{-2y}\end{aligned}$$

where $y = \tau^{-1}$ is a dimensionless reduced energy variable defined as the reciprocal of the reduced temperature. The exponential factors obtained for low temperatures are characteristic of systems with a gap in their excitation spectrum. At very low temperature most elements are found in the lowest energy level and the population of the upper energy level is

$$\frac{N_2}{N_1} \approx e^{-2y} = \text{Exp}\left[-\frac{\Delta}{k_B T}\right]$$

where $\Delta = \epsilon_2 - \epsilon_1 = 2\epsilon$ is the energy gap between the two levels. Hence, the heat capacity is small because the entropy and the change in entropy with temperature are small. Systems with a continuous excitation spectrum normally exhibit power-law behavior at low temperature, but the finite gap for the binary system produces exponential suppression of the low-temperature heat capacity.

Algebraic details and more thorough presentations of the thermodynamics of binary systems can be found in *spin-half.nb* using the microcanonical ensemble and *thermo2.nb* using the canonical ensemble.

■ Example: harmonic oscillators

Consider a system of harmonic oscillators and assume that interactions between these oscillators can be neglected. For example, electromagnetic radiation in a cavity, vibrational excitations of molecules, or the vibrations of atoms about lattice sites in a crystal may be analyzed in terms of independent normal modes of oscillation. Each individual oscillator has an energy spectrum $\varepsilon_n = \hbar\omega(n + \frac{1}{2})$ consisting of an infinite sequence of equally spaced levels where n is the total number of quanta, $\hbar\omega$ is the fundamental quantum of energy for the system, and the zero-point energy is $\hbar\omega/2$.

For example, consider a single particle in a 3-dimensional harmonic well. The total number of quanta is then $n = n_x + n_y + n_z$ where each of the three spatial dimensions can be treated as an independent oscillator. The degeneracy of each single-particle energy level is equal to the total number of ways that n quanta can be distributed among the three independent modes. There are $n + 1$ possible values of n_x between 0 and n . For each of these values, we can choose n_y to be anywhere between 0 and $n - n_x$. The number of quanta along the z -axis is then determined. Hence, the degeneracy of a single-particle level with n quanta is given by the sum

$$g[n] = \sum_{n_x=0}^n (n - n_x + 1) = \sum_{k=1}^{n+1} k = \frac{1}{2} (n + 1)(n + 2)$$

whose value can be demonstrated by induction.

The degeneracy for a system of many independent oscillators can now be obtained by extending this argument to f independent degrees of freedom, where f is the total number of oscillator modes rather than simply the total number of particles. Let n represent the total number of quanta, so that the internal energy of the system is $U = \hbar\omega(n + \frac{f}{2})$. The degeneracy is simply the number of distinct ways that n indistinguishable objects (quanta) can be distributed among f distinguishable boxes (vibrational modes). Suppose that the boxes are represented by vertical lines marking their boundaries and that the objects are represented by circles, as sketched below for a particular configuration.

•• | •••• | | ••• | ••••• | | •

The outermost walls need not be included. The number of combinations of n objects and $f - 1$ interior partitions is simply $(f + n - 1)!$. However, because the $n!$ permutations of the balls or the $(f - 1)!$ permutations of the partitions among themselves leave the system invariant, the degeneracy is simply the binomial coefficient

$$g[n, f] = \binom{f + n - 1}{n} = \frac{(f + n - 1)!}{n! (f - 1)!}$$

The earlier result for a single three-dimensional oscillator is recovered by choosing $f \rightarrow 3$.

In the limit of large n and large f , we can employ Stirling's approximation to obtain

$$\ln g \approx (f + n) \ln[f + n] - n \ln n = f((1 + x) \ln[1 + x] - x \ln x)$$

where $x = n/f$ is the average excitation per oscillator. The entropy can now be expressed as

$$S = k_B f ((1 + x) \ln[1 + x] - x \ln x)$$

where $x + \frac{1}{2} = U/f\hbar\omega$. Evaluating the temperature, we find

$$\frac{1}{k_B T} = \frac{\partial \ln g}{\partial U} \implies f y = \frac{\partial \ln g}{\partial x} = f \ln \frac{1 + x}{x} \implies x = \frac{1}{e^y - 1}$$

where $y = \frac{\hbar\omega}{k_B T}$ is a dimensionless variable that represents the ratio between oscillator and thermal energies. Therefore, we obtain the thermal equation of state

$$U = f \hbar \omega \left(\frac{1}{e^{\hbar\omega/k_B T} - 1} + \frac{1}{2} \right) = \frac{f \hbar \omega}{2} \text{Coth}[y]$$

and heat capacity

$$\frac{C_V}{f k_B} = \frac{y^2 e^y}{(e^y - 1)^2} \quad \text{where} \quad y = \frac{\hbar \omega}{k_B T}$$

It is useful to define a dimensionless reduced temperature as $\tau = y^{-1} = k_B T / \hbar \omega$. At high temperature, we can expand these functions in powers of the small variable y , to find

$$U \approx f k_B T \left(1 + \frac{1}{12} \left(\frac{\hbar \omega}{k_B T} \right)^2 + \dots \right)$$

$$C_V \approx f k_B \left(1 - \frac{1}{12} \left(\frac{\hbar \omega}{k_B T} \right)^2 + \dots \right)$$

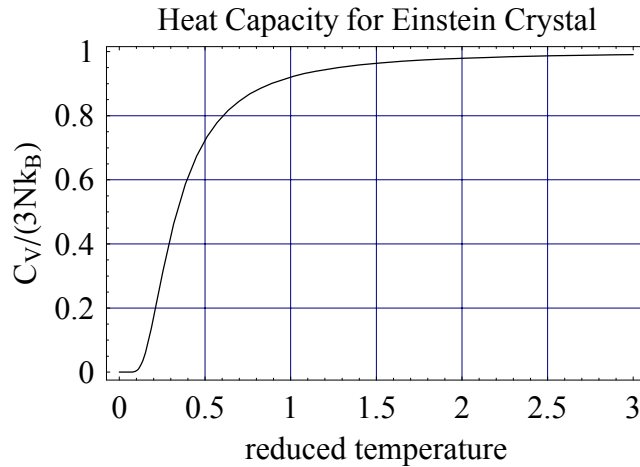
Thus, at high temperature the heat capacity per oscillator approaches the limiting value k_B . At low temperature, where y is large, we isolate the exponential factors to obtain

$$U \approx \frac{f \hbar \omega}{2} (1 + e^{-y})$$

$$C_V \approx f k_B y^2 e^{-y}$$

Notice that at low temperature the heat capacity resembles that of the binary system because only the first excited state matters when the thermal energy is small compared with the excitation energy. Under those circumstances the thermodynamics is dominated by the gap in the excitation spectrum. However, unlike the binary system, the internal energy is proportional to temperature and the heat capacity approaches a nonzero constant at high temperature because the energy spectrum is not bounded. Thus, the oscillator system is considered statistically normal while the binary system is not.

Einstein's model of the heat capacity of a crystal containing N atoms in a lattice treated the $3N$ normal modes of vibration as independent harmonic oscillators with a single average frequency ω_E . If we define a characteristic temperature $T_E = \hbar\omega_E / k_B$, the reduced temperature becomes $\tau = T / T_E$. The figure below displays the dependence of the heat capacity upon reduced temperature. At high temperatures, $\tau \gg 1$, one finds the classical heat capacity $C_V \rightarrow 3N k_B$ expected for equipartition among $3N$ coordinates and $3N$ momenta. However, at low temperatures, $\tau < 1$, the heat capacity is much smaller quantum mechanically than classically because the finite energy required to excite a quantized oscillator *freezes out* states that would participate in a classical theory that permits arbitrarily small excitations. Although this model is simplistic, the explanation of reduced heat capacity at low temperature in terms of quantization of energy provided important support for the early quantum theory.

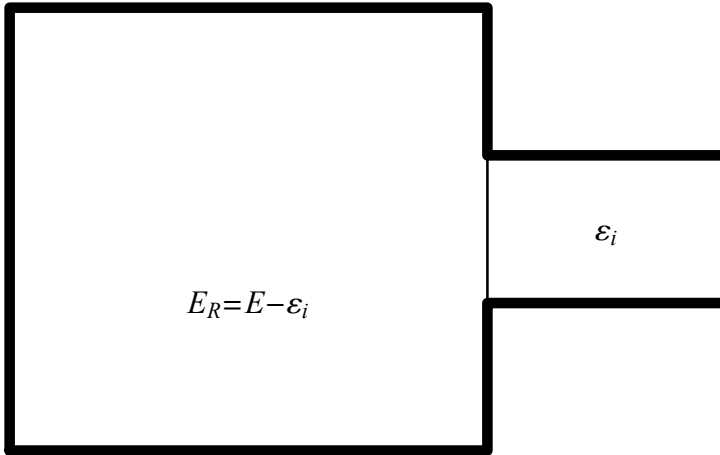


We will return to this and more refined models shortly. Further information on these systems can be found in the notebooks *hotherm.nb*, *debye.nb*, and *planck.nb*.

Canonical ensemble

■ Thermal interaction with a reservoir

Suppose that a large isolated system is partitioned asymmetrically into two parts, one of which is very much larger than the other, separated by a rigid, impermeable wall which nevertheless conducts heat. The total energy E of the composite system is conserved. We seek the probability P_i that, in thermal equilibrium, the small system is found in a particular quantum state i with energy ε_i . The large system serves as a heat reservoir with total energy $E_R = E - \varepsilon_i$ that is much greater than the energy $\varepsilon_i \ll E_R$ of the system of interest. Henceforth, we refer to the large system as the "reservoir" and to the smaller subsystem as simply the "system".



Energy sharing between a reservoir with energy $E_R = E - \varepsilon_i$ and a small system with energy ε_i in state i . Thick walls isolate the composite system while the two subsystems exchange heat through the thin rigid, impermeable, diathermal wall.

Although we have specified the state of the system precisely, the reservoir may occupy any of its states which satisfy the constraint $E_R + \varepsilon_i = E$ upon the total energy. According to the postulate of equal *a priori* probabilities, each of these states is found with equal likelihood. Therefore, P_i is proportional to the multiplicity Ω_R for the reservoir, such that

$$P_i = C \Omega_R[E - \varepsilon_i]$$

where C is a normalization constant. Expanding $\ln \Omega_R$ for $\varepsilon \ll E$,

$$\ln \Omega_R[E - \varepsilon] \approx \ln \Omega_R[E] - \varepsilon \frac{\partial \ln \Omega_R}{\partial E}$$

and identifying the entropy and temperature of the reservoir as

$$S_R = k_B \ln \Omega_R \quad \frac{1}{T} = \frac{\partial S_R}{\partial E}$$

we find

$$\ln P_i = -\frac{\varepsilon_i}{k_B T} + C'$$

where C' is some other constant. Therefore, the relative probability that the system occupies a particular quantum state is given by the Boltzmann factor

$$P_i \propto e^{-\beta \varepsilon_i} \quad \beta \equiv \frac{1}{k_B T}$$

To interpret this result, consider two states labelled 1 and 2 with energies ε_1 and ε_2 , respectively. The ratio of probabilities for finding the system in either of these states, in thermal equilibrium with a reservoir at temperature T , is then

$$\frac{P_1}{P_2} = \text{Exp}\left[-\frac{\Delta \varepsilon}{k_B T}\right]$$

where $\Delta \varepsilon = \varepsilon_1 - \varepsilon_2$ is the energy difference between the two states. The state with higher energy occurs with smaller probability because the number of states available to the reservoir, and hence the entropy of the reservoir, is smaller if less energy is available to it. This effect is relatively small if the temperature is large enough so that the typical thermal energy $k_B T$ is much larger than the energy differences among states of the smaller system. Under these circumstances all states of the system are populated with essentially equal frequency,

$$k_B T \gg \Delta\varepsilon \implies P_1 \approx P_2$$

corresponding to a highly disordered state of high entropy. On the other hand, if the thermal energy is comparable to or smaller than the spacing between energy levels of the system, the probability that a level with energy ε is populated decreases exponentially with energy. Under these circumstances, states of higher energy are effectively frozen out:

$$k_B T \lesssim \Delta\varepsilon \implies P_1 \ll P_2$$

With fewer quantum states populated with appreciable frequency, the system exhibits a higher degree of order and possesses lower entropy.

Therefore, the canonical probability distribution

$$P_i = Z^{-1} e^{-\beta\varepsilon_i} \quad \beta = \frac{1}{k_B T} \quad Z = \sum_i e^{-\beta\varepsilon_i}$$

describes the population of states for a small system in thermal equilibrium with a much larger reservoir of heat. The reservoir must be large enough so that its macroscopic properties are not affected by fluctuations in the energy drawn by the smaller system. The reservoir then defines the temperature of the smaller system precisely even if its energy fluctuations are relatively large (compared with its mean energy). The relative probability for each quantum state of the system is described by the Boltzmann factor $e^{-\beta\varepsilon}$ and the absolute probability is obtained by normalizing the distribution to unity. The normalization factor is determined by the partition function Z , where the sum extends over all possible states of the system. The Boltzmann factor plays a central role throughout the theory of statistical mechanics and, hence, is of fundamental significance.

It is important to remember that P_i is the probability for a *unique* quantum state. The probability $P[\varepsilon_i]$ for finding any state with energy ε_i and degeneracy $g[\varepsilon_i]$ is then $P_i g[\varepsilon_i]$, such that

$$\frac{P[\varepsilon_1]}{P[\varepsilon_2]} = \frac{g[\varepsilon_1]}{g[\varepsilon_2]} \text{Exp}\left[-\frac{\Delta\varepsilon}{k_B T}\right]$$

is the ratio of probabilities for two energy levels. Similarly, it is important to remember that the partition function is defined by a sum over microstates rather than over energy levels. If an energy level ε_i has degeneracy g_i , it must be counted not once but g_i times. If in a particular application it is more convenient to sum over energy levels instead of states, then each level must be assigned a statistical weight equal to its degeneracy. The definition of the partition function then becomes

$$Z = \sum_i g_i e^{-\beta\varepsilon_i}$$

where the summation index now enumerates levels. Finally, if the spacing between energy levels becomes sufficiently small and the average number of states within a suitable interval $d\varepsilon$ is a smooth function of energy, we can replace the sum by an integral, such that

$$Z = \int d\varepsilon g[\varepsilon] e^{-\beta\varepsilon}$$

where $g[\varepsilon]$ is the density of states defined as the total number of states between ε and $\varepsilon + d\varepsilon$ according to

$$g[\varepsilon] d\varepsilon = \sum_{\varepsilon \leq \varepsilon_i \leq \varepsilon + d\varepsilon} g_i$$

However, we must remember that if a single state or small group of states becomes important, it may be necessary to return to explicit summation because integration with respect to a smooth density of states may miss important effects arising from details of the spectrum of quantum states. The phenomenon of Bose condensation is the most famous example of this type.

It is also important to recognize that the derivation of the canonical ensemble does not require the system to be of macroscopic size. In fact, we can apply the Boltzmann factor to analyze the relative probabilities of the states of a single atom which interacts with a reservoir. The reservoir, of course, must be large enough to define a statistically significant temperature. However, it need not be a material system — it could, for example, be a thermal distribution of electromagnetic radiation which bathes an atom, taken to be the system. We then imagine that there exists a statistically large ensemble of macroscopically identical systems interacting with similar environments. The ensemble then represents a thermodynamic system subject to statistical analysis. The properties of such an ensemble may exhibit large dispersions, but can be analyzed using the canonical probability distribution. The members of the ensemble need not be large enough themselves to possess narrow distributions of properties. Thus, statistical mechanics is more general than thermodynamics, which relies on extensive potentials. Those potentials need not be strictly extensive for small systems.

■ Maximizing disorder given fixed mean energy

We wish to construct an ensemble consisting of many macroscopically identical isolated systems for which the mean energy of the ensemble is specified. The probability distribution $\{P_i\}$ should be chosen to maximize disorder subject to the constraints

$$U = \sum_i P_i E_i, \quad \sum_i P_i = 1$$

Thus, using

$$\begin{aligned} dS &= -k_B \sum_i \ln P_i dP_i = 0 \\ dU &= \sum_i E_i dP_i = 0 \\ \sum_i dP_i &= 0 \end{aligned}$$

and introducing Lagrange multipliers with convenient factors, we find

$$\ln P_i + \lambda_1 E_i + \lambda_2 = 0 \implies P_i = \text{Exp}[-\lambda_1 E_i - \lambda_2]$$

Applying the constraints, we find

$$\begin{aligned} \lambda_2 &= \ln \sum_i e^{-\lambda_1 E_i} \\ U &= -\frac{\partial \lambda_2}{\partial \lambda_1} \\ S &= k_B (\lambda_1 U + \lambda_2) \end{aligned}$$

The Lagrange multipliers can be connected to conventional thermodynamic quantities using

$$\frac{1}{T} = \frac{\partial S}{\partial U} \implies \lambda_1 = \frac{1}{k_B T} \equiv \beta$$

Notice that we are holding external parameters, such as volume, fixed so that for this partial derivative the energy levels E_i and the normalization λ_2 are constant. Thus, we find that the normalization factor is related to the Helmholtz free energy by

$$\lambda_2 = \frac{TS - U}{k_B T} = -\beta F$$

Once again defining the canonical partition function Z as the normalization of the probability distribution, the canonical probability distribution which maximizes the disorder within an ensemble of specified mean energy now becomes

$$P_i = \frac{e^{-\beta E_i}}{Z} = e^{-\beta(E_i - F)}$$

$$Z = \sum_i e^{-\beta E_i} = e^{-\beta F}$$

Therefore, we have derived the same canonical probability distribution from two rather different perspectives. The method based upon the thermal interaction between a system (sample) and a much larger reservoir of heat (environment) specifies the temperature through the variation of the entropy of the reservoir with the energy drawn by the sample. The entropy of the composite system is maximized when all states of the reservoir with the energy determined by the state of the sample are equally likely. As the sample and the reservoir interact, the state of each changes rapidly. Hence, the probability for a microstate of the sample represents the fraction of the time a particular sample can be found in that state and is governed by the temperature of the reservoir. Conversely, the method based upon the entropy of the sample specifies the mean energy for an ensemble of isolated systems and derives the temperature of the sample from the energy dependence of the disorder within that ensemble. Here the probability for a microstate represents its frequency within an ensemble of stationary states. This method is closely related to the microcanonical ensemble, but does not require the internal energy to be constrained to an infinitesimal interval. Both methods give equivalent results because the ergodic hypothesis requires objective and *a priori* probabilities to be equal. Both of these methods can be generalized to include other interactions between the sample and its environment. For example, if we allow the sample and the reservoir to exchange volume through motion of the wall between them, the thermal-interaction method would include the dependence of the multiplicity function for the reservoir upon volume as well as energy. Alternatively, the information-theory method would need another Lagrange multiplier to constrain the mean volume of the sample. Both approaches will produce the same probability distribution for this (T, p) ensemble; your preference is an aesthetic choice.

■ Thermodynamics of canonical ensembles

Canonical ensembles are represented by probability distributions of the form

$$P_i = Z^{-1} e^{-\beta E_i} \iff P[E] = Z^{-1} g[E] e^{-\beta E}$$

$$Z = \sum_i e^{-\beta E_i} \iff Z = \int dE g[E] e^{-\beta E}$$

The ensemble average for any variable x is then

$$\langle \hat{x} \rangle = \sum_i P_i x_i = Z^{-1} \sum_i e^{-\beta E_i} x_i \iff \langle \hat{x} \rangle = Z^{-1} \int dE g[E] e^{-\beta E} x[E]$$

For example, the average internal energy becomes

$$U = \langle \hat{H} \rangle = \frac{\sum_i e^{-\beta E_i} E_i}{\sum_i e^{-\beta E_i}} = -\frac{1}{Z} \left(\frac{\partial Z}{\partial \beta} \right) = -\frac{\partial}{\partial \beta} \ln Z$$

Therefore, the partition function determines the thermodynamic properties of the system completely.

It is useful to relate the partition function to the Helmholtz free energy, which is the thermodynamic potential whose natural variables are those specified in the definition of the canonical ensemble, namely temperature and volume. Recognizing that $S = -\left(\frac{\partial F}{\partial T}\right)_V$, we identify

$$F = U - TS = U + T \left(\frac{\partial F}{\partial T} \right)_V$$

Using

$$T \frac{\partial}{\partial T} = -\beta \frac{\partial}{\partial \beta}$$

we can write

$$U = F + \beta \left(\frac{\partial F}{\partial \beta} \right)_V = \frac{\partial}{\partial \beta} (\beta F) = -T^2 \frac{\partial}{\partial T} \left(\frac{F}{T} \right)_V$$

Thus, comparing expressions for U , we find

$$\frac{\partial}{\partial \beta} (\beta F) = -\frac{\partial}{\partial \beta} \ln Z \implies F = -k_B T \ln Z - T S_0$$

where S_0 is a constant of integration.

The constant of integration can be determined by examining the behavior of the entropy for temperatures so low that the partition function is dominated by the quantum state of lowest energy, E_0 . We then find that

$$T \rightarrow 0 \implies F \approx E_0 - T S_0 \implies S \approx S_0$$

However, the third law of thermodynamics stipulates the entropy approaches zero as the temperature approaches absolute zero. Microscopically, if we assume that the ground state is nondegenerate, or at least has macroscopically negligible degeneracy, then the entropy must vanish when the temperature is so low that only the ground state is available to the system. Thus, we conclude that $S_0 = 0$.

Therefore, the thermodynamics of any system can be derived from knowledge of its partition function and the general relationship

$$F = -k_B T \ln Z \implies Z = e^{-\beta F}$$

We can then compute U by differentiating $\ln Z$ with respect to β or by evaluating S from F and using $U = F + T S$. Finally, the equation of state $p = p[T, V]$ is most readily determined from the free energy $F[T, V]$ using $p = -\left(\frac{\partial F}{\partial V}\right)_T$. These relationships are summarized below.

$$\begin{aligned} Z = e^{-\beta F} &\implies F = -k_B T \ln Z \\ F = U - T S &\implies dF = -S dT - p dV \text{ (simple compressible)} \\ S &= -\left(\frac{\partial F}{\partial T}\right)_V \\ p &= -\left(\frac{\partial F}{\partial V}\right)_T \\ U &= -\frac{\partial}{\partial \beta} \ln Z = \frac{\partial}{\partial \beta} (\beta F) = -T^2 \frac{\partial}{\partial T} \left(\frac{F}{T} \right)_V = F + T S \end{aligned}$$

■ Properties of the partition function

The statistical analysis of many systems will be found to benefit from several general properties of the partition function. First, suppose that the energy of each state is shifted by the same constant δ , such that

$$E_i \rightarrow E_i + \delta \implies Z \rightarrow e^{-\beta \delta} Z$$

We then find that the energy and free energy are shifted by the same amount, whereas the entropy and pressure are unaffected.

$$E_i \rightarrow E_i + \delta \implies F \rightarrow F + \delta, \quad U \rightarrow U + \delta, \quad S \rightarrow S, \quad p \rightarrow p$$

Second, suppose that two parts of a system interact weakly so that their energies are additive. We then find that the partition function reduces to the product of two separate partition functions, one for each part, and that the free energy reduces to the sum of two separate free energies. The state of a *separable* (or *factorizable*) system can be described by two

indices representing the separate states of the two independent systems, such that $E_{i,j} = E_i^{(1)} + E_j^{(2)}$ is the total energy of the product state (i, j) composed of contributions $E_i^{(1)}$ from subsystem (1) and $E_j^{(2)}$ from subsystem (2). We then find

$$E_{i,j} = E_i^{(1)} + E_j^{(2)} \implies Z = Z_1 Z_2 \implies F = F_1 + F_2$$

where

$$\begin{aligned} Z_1 &= \sum_i \text{Exp}[-\beta E_i^{(1)}] & Z_2 &= \sum_j \text{Exp}[-\beta E_j^{(2)}] \\ F_1 &= -k_B T \text{Log}[Z_1] & F_2 &= -k_B T \text{Log}[Z_2] \end{aligned}$$

Hence, F , U , and S are extensive variables for weakly interacting systems. This argument can be extended to a system of N identical noninteracting constituents with the important result

$$\text{noninteracting} \implies Z_N = (Z_1)^N \implies F = -N k_B T \text{Log}[Z_1]$$

where Z_1 is the single-particle partition function.

■ Example: binary system

Suppose that each member of a system of N noninteracting constituents can occupy two states with energies ε and $\varepsilon + \Delta$, where ε is the ground-state energy and Δ is the energy gap. The single-particle partition function is then

$$Z_1 = \text{Exp}\left[-\frac{\varepsilon}{k_B T}\right] + \text{Exp}\left[-\frac{\varepsilon + \Delta}{k_B T}\right] = e^{-\beta\varepsilon} (1 + e^{-y})$$

where $y = \beta\Delta = \Delta/k_B T$ is the reduced excitation energy defined as the ratio between excitation and thermal energies. The principal thermodynamic functions are

$$\begin{aligned} F &= N \left(\varepsilon - \frac{\Delta}{y} \text{Log}[1 + e^{-y}] \right) \\ U &= N \left(\varepsilon + \frac{\Delta}{1 + e^y} \right) \\ S &= N k_B \left(\frac{y}{1 + e^y} + \text{Log}[1 + e^{-y}] \right) \end{aligned}$$

where the energy offset $N\varepsilon$ that appears in both F and U has no thermodynamical significance because it is not coupled to temperature. Thus, the heat capacity

$$C_V = N k_B \frac{y^2 e^y}{(1 + e^y)^2}$$

depends only upon the reduced excitation energy and is independent of the ground-state energy. With a little algebra one can demonstrate that these results are consistent with those obtained earlier using the microcanonical ensemble, but you will probably agree that the derivation is easier using the canonical ensemble. The microcanonical ensemble is useful for deriving general relationships, such as the form of the canonical probability distribution, but applications are usually easier to make with the canonical ensemble.

■ Example: ideal paramagnetism

In this example we compare and contrast the quantum mechanical and classical theories of ideal paramagnetism. The orientational energy for a permanent magnetic dipole moment $\vec{\mu}$ in an external magnetic field \vec{B} is $\varepsilon = -\vec{\mu} \cdot \vec{B}$. For atomic physics the temperature scale for paramagnetic behavior is determined by the Bohr magneton

$$\mu_0 \rightarrow \mu_B = \frac{e \hbar}{2 m_e c} \implies \frac{\mu_B}{k_B} = 0.672 \text{ Kelvin/Tesla}$$

while for nuclear physics the nuclear magneton and corresponding temperature scale is smaller by the ratio of electron to proton masses:

$$\mu_0 \rightarrow \mu_N = \frac{e \hbar}{2 m_p c} \implies \frac{\mu_N}{k_B} = 0.000366 \text{ Kelvin/Tesla}$$

In the quantum mechanical theory, the magnetic moment can be expressed as $\vec{\mu} = g \mu_0 \vec{j}$ where μ_0 is the unit of magnetic moment (Bohr or nuclear magneton, as appropriate), \vec{j} is the total angular momentum, usually containing both spin and orbital contributions, and g is a gyromagnetic ratio describing the internal structure of the particle or atom. Recognizing that the component of angular momentum along the magnetic field is quantized, the discrete spectrum of orientational energy levels becomes $\varepsilon_m = -\frac{m}{j} \mu B$ where μ is the magnetic moment in the aligned state with $m = j_z = j$. Thus, it is useful to define $y = \mu B / k_B T$ as the ratio between orientational and thermal energies such that the single-particle partition function becomes

$$Z_1 = \sum_{m=-j}^j \text{Exp}[m y / j] = \frac{\text{Sinh}\left[\frac{2j+1}{2j} y\right]}{\text{Sinh}\left[\frac{y}{2j}\right]}$$

The principal thermodynamic functions can now be evaluated straightforwardly. Some of these functions are summarized below and the gory details can be found in *paramag.nb*.

$$\begin{aligned} F &= -N k_B T \text{Log}\left[\frac{\text{Sinh}\left[\frac{2j+1}{2j} y\right]}{\text{Sinh}\left[\frac{y}{2j}\right]}\right] \\ M &= -\left(\frac{\partial F}{\partial B}\right)_T = N \mu \left(\frac{(2j+1)}{2j} \text{Coth}\left[\frac{(2j+1)y}{2j}\right] - \frac{1}{2j} \text{Coth}\left[\frac{y}{2j}\right]\right) \\ U &= -M B \\ C_B &= \left(\frac{\partial U}{\partial T}\right)_B = N k_B \left(\left(\frac{y}{2j} \text{Csch}\left[\frac{y}{2j}\right]\right)^2 - \left(\frac{(2j+1)y}{2j} \text{Csch}\left[\frac{(2j+1)y}{2j}\right]\right)^2\right) \end{aligned}$$

One can easily demonstrate that our previous results for spin $\frac{1}{2}$ are recovered when $j \rightarrow \frac{1}{2}$.

It is instructive to compare these results with a classical theory for which the orientation of the magnetic dipole is a continuous function of angle, such that $\varepsilon = -\mu B \cos\theta$. The summation over states becomes an integral over the unit sphere, such that

$$Z_1 = 2\pi \int_{-1}^1 dx e^{xy} = 4\pi \frac{\text{Sinh}[y]}{y}$$

where the alignment variable $x = \cos\theta$ is uniformly distributed between $\{-1, 1\}$. The corresponding thermodynamic functions constructed in like manner are:

$$\begin{aligned} F &= -N k_B T \text{Log}\left[4\pi \frac{\text{Sinh}[y]}{y}\right] \\ M &= -\left(\frac{\partial F}{\partial B}\right)_T = N \mu \left(\text{Coth}[y] - \frac{1}{y}\right) \\ U &= -M B \\ C_B &= \left(\frac{\partial U}{\partial T}\right)_B = N k_B \left(1 - \left(\frac{y}{\text{Sinh}[y]}\right)^2\right) \end{aligned}$$

Notice that for fixed μ the quantum mechanical and classical formulas for M and C_B are identical in the limit $j \rightarrow \infty$. We must hold μ fixed because in the classical theory there μ is independent of j . For the quantum mechanical theory one can imagine that this limit is taken with $\{g \rightarrow 0, j \rightarrow \infty\}$ such that $\mu = g \mu_0 j$ remains constant.

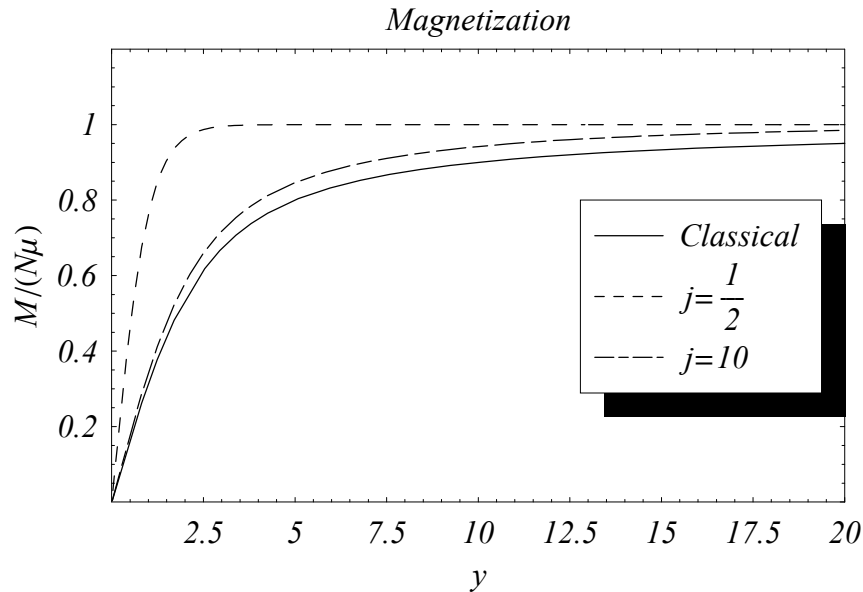
The figure below compares the classical magnetization with quantum mechanical results for small and large spins. For all cases saturation is observed for large y , corresponding to large B or small T , but the quantum mechanical theory shows that saturation is approached more quickly for small j than predicted by the classical theory. For both theories one finds that the magnetic susceptibility for large temperatures or small fields is described by Curie's law

$$\chi = -\left(\frac{\partial M}{\partial B}\right)_T \approx \frac{\chi_0}{T}$$

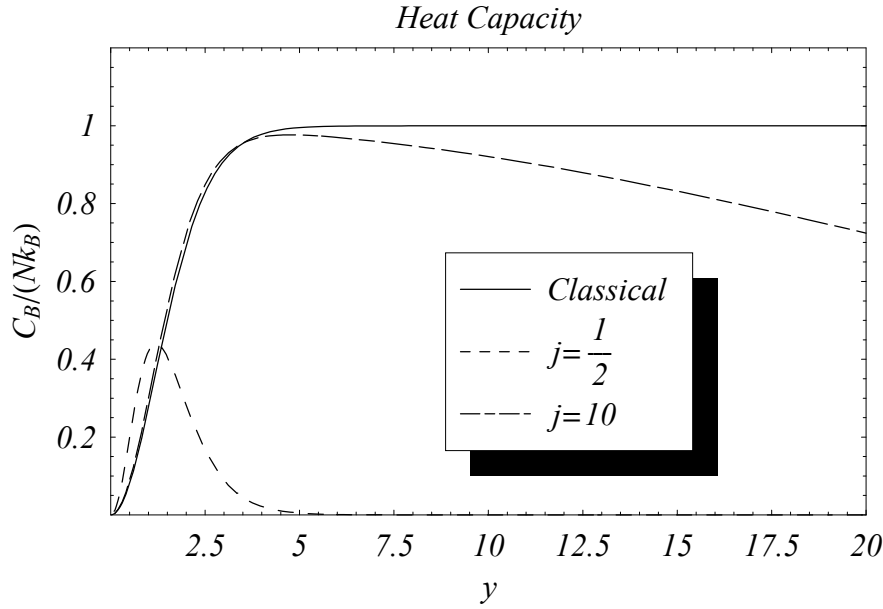
but the classical and quantum mechanical Curie constants

$$\begin{aligned} \text{classical} &\quad \Rightarrow \chi_0 = \frac{N \mu^2}{3 k_B} \\ \text{quantum mechanical} &\quad \Rightarrow \chi_0 = \frac{N g \mu_0^2 j(j+1)}{3 k_B} = \frac{N \mu^2}{3 k_B} \frac{j+1}{j} \end{aligned}$$

differ by a factor of $(j+1)/j$, which is 3 for spin $\frac{1}{2}$. The spin-dependent factor in χ_0 is simply the quantum mechanical expectation of $\langle j^2 \rangle = j(j+1)$. Although these models of the magnetization are qualitatively consistent, the quantization of orientation of the magnetic dipole moment strongly affects the Curie constant for small spins.



The figure below compares the magnetic heat capacities for the classical and quantum mechanical theories. Classically one finds that the heat capacity saturates for large y , but quantum mechanically one finds that the heat capacity approaches zero for large y . The difference between these theories is especially great for small spins. As the spin increases the quantum mechanical heat capacity stays near the classical limit longer, but eventually it peaks and decreases, failing to maintain saturation. Thus, the Schottky heat-capacity anomaly we discussed for binary systems is present for higher spins also, differing in shape and scale but not in character. Why does C_B differ from the classical limit more strongly than M ? A classical system can absorb infinitesimally small amounts of heat because its energy spectrum is continuous, but because the quantum mechanical energy spectrum is quantized energy can be absorbed only in multiples of the energy spacing $\Delta\varepsilon = g \mu_0 B$. Therefore, the gap between the lowest two energy levels inhibits heat transfer when $k_B T$ is appreciably smaller than $\Delta\varepsilon$. When practically all atoms are found in the lowest energy level, the gap to the next level is more important than the spectrum of higher levels. If there is a finite gap, the thermodynamics for very low temperatures resembles that of the two-state system.



■ Example: lattice vibrations

A crystal consists of N atoms arranged upon a lattice. The atoms may be considered distinguishable by their lattice positions. Let $x_i[t]$ denote one of the $3N$ coordinates specifying the atomic positions and let

$$H = \frac{1}{2} m \sum_{i=1}^{3N} \dot{x}_i^2 + V[x_1, \dots, x_{3N}]$$

represent the total kinetic and potential energies of the crystal. Assuming that atomic vibrations are limited to small displacements from the equilibrium positions \bar{x}_i , the potential energy can be expanded to second order according to

$$V[x_1, \dots, x_{3N}] \approx \bar{V}[\bar{x}_1, \dots, \bar{x}_{3N}] + \frac{1}{2} \sum_{i,j=1}^{3N} \left(\frac{\partial^2 V}{\partial x_i \partial x_j} \right)_{\bar{x}_1, \dots, \bar{x}_{3N}} (x_i - \bar{x}_i)(x_j - \bar{x}_j) + \dots$$

We assume that the constant term $\bar{V} = -N\phi$ is proportional to the number of atoms, where ϕ is the average binding energy per atom, and define a curvature matrix

$$\alpha_{i,j} = \left(\frac{\partial^2 V}{\partial x_i \partial x_j} \right)_{\bar{x}_1, \dots, \bar{x}_{3N}}$$

such that

$$V[x_1, \dots, x_{3N}] \approx -N\phi + \frac{1}{2} \sum_{i,j=1}^{3N} \alpha_{i,j} \xi_i \xi_j + \dots \implies H \approx -N\phi + \frac{1}{2} m \sum_{i=1}^{3N} \dot{\xi}_i^2 + \frac{1}{2} \sum_{i,j=1}^{3N} \alpha_{i,j} \xi_i \xi_j$$

where displacements from equilibrium are denoted $\xi_i = x_i - \bar{x}_i$. It is useful to treat ξ_i as a component of a vector $\boldsymbol{\xi}$, such that the model Hamiltonian can be expressed in the form

$$H = -N\phi + \frac{1}{2} (m \dot{\boldsymbol{\xi}} \cdot \dot{\boldsymbol{\xi}} + \boldsymbol{\xi} \cdot \boldsymbol{\alpha} \cdot \boldsymbol{\xi})$$

The matrix α is clearly symmetric and must also be positive-definite because the potential energy was expanded around the stable equilibrium configuration. Therefore, the potential energy can be diagonalized by an orthogonal transformation \mathbf{S} , such that

$$\tilde{\mathbf{S}} \alpha \mathbf{S} = \lambda$$

where λ is a diagonal matrix with positive eigenvalues $\lambda_{i,j} = \lambda_i \delta_{i,j}$ and where $\tilde{\mathbf{S}}$ is the transpose of \mathbf{S} such that $\tilde{\mathbf{S}} \mathbf{S} = \mathbf{1}$. Defining normal coordinates

$$\xi = \mathbf{S} \cdot \mathbf{q} \implies \xi \cdot \alpha \cdot \xi = \mathbf{q} \cdot \lambda \cdot \mathbf{q} = \sum_{i=1}^{3N} \lambda_i q_i^2$$

and momenta

$$\mathbf{p} = m \dot{\mathbf{q}} \implies m \dot{\xi} \cdot \dot{\xi} = \frac{\mathbf{p} \cdot \mathbf{p}}{2m}$$

the Hamiltonian reduces to

$$H = -N\phi + \sum_{i=1}^{3N} \left(\frac{p_i^2}{2m} + \frac{1}{2} m \omega_i^2 q_i^2 \right)$$

where the frequencies are identified by $\lambda_i = m \omega_i^2$. Thus, expressing the displacements in terms of *normal modes* of vibration reduces the Hamiltonian to a sum of independent terms, each of which has the form of a harmonic oscillator. There are $3N$ normal modes with a spectrum of vibrational frequencies $\{\omega_i, i = 1, 3N\}$ that depends in a complicated manner upon the interatomic potential that binds the crystal together. Although it may be difficult or impossible to compute these frequencies from first principles, the excitation spectrum is readily measured by neutron scattering or other means.

The analysis has been essentially classical to this point, but quantization of the Hamiltonian for a harmonic oscillator should be quite familiar. We simply replace the classical Hamiltonian function by the quantum mechanical operator

$$\hat{H} = -N\phi + \sum_{\alpha=1}^{3N} \hbar \omega_{\alpha} \left(\hat{n}_{\alpha} + \frac{1}{2} \right)$$

where $\hat{n} = \hat{a}_{\alpha}^{\dagger} \hat{a}_{\alpha}$ is the number operator and where $\hat{a}_{\alpha}^{\dagger}$ and \hat{a}_{α} are creation and destruction operators for quantized excitations of mode α . Vibrational quanta are known as *phonons*, named by analogy with electromagnetic quanta known as photons. The thermodynamic behavior of a wide variety of physical systems can be analyzed in terms of quantized excitations of various types. If one can approximate the Hamiltonian by a sum of independent normal modes, the subsequent thermodynamic analysis is typically very similar to that for lattice vibrations. Hence, this example represents a very important prototype.

The canonical partition function for this separable Hamiltonian of this type

$$Z = \text{Tr}[\text{Exp}[-\beta \hat{H}]] = e^{\beta N \phi} \prod_{\alpha=1}^{3N} Z_{\alpha}$$

factorizes into single-mode partition functions

$$Z_{\alpha} = \sum_{n_{\alpha}=0}^{\infty} \text{Exp}\left[-\beta \hbar \omega_{\alpha} \left(n_{\alpha} + \frac{1}{2} \right)\right] = \frac{\text{Exp}[\beta \hbar \omega_{\alpha} / 2]}{\text{Exp}[\beta \hbar \omega_{\alpha}] - 1}$$

where each mode can carry an arbitrarily large number of phonons. The internal energy

$$U = -\frac{\partial \ln Z}{\partial \beta} = -N \phi + \sum_{\alpha=1}^{3N} \hbar \omega_{\alpha} \left(\bar{n}_{\alpha} + \frac{1}{2} \right)$$

separates into a temperature-independent zero-point and binding energy plus the sum of the excitation energies for each mode carrying an average number of phonons given by

$$\bar{n}_{\alpha} = \frac{1}{\text{Exp}[\beta \hbar \omega_{\alpha}] - 1}$$

To complete this model, we must determine the spectrum of normal modes, $\{\omega_{\alpha}, \alpha = 1, 3N\}$. If we assume that the spectrum is essentially continuous, we can replace the discrete summation by an integral of the form

$$U = -N \phi + \int d\omega g[\omega] \frac{\hbar \omega}{\text{Exp}[\beta \hbar \omega] - 1}$$

where $g[\omega]$ is the density of states such that $g[\omega] d\omega$ is the number of states between ω and $\omega + d\omega$. The density of states must be normalized according to

$$\int d\omega g[\omega] = 3N$$

such that the total number of states represents three degrees of freedom per atom.

The simplest version of this model, first proposed by Einstein, assumes that all frequencies can be approximated by a single average value ω_E . It is then convenient to define a reduced temperature $\tau = T/T_E$ where the Einstein temperature is given by $\hbar \omega_E = k_B T_E$. Evaluating the internal energy

$$g[\omega] = 3N \delta[\omega - \omega_E] \implies U = 3N \hbar \omega_E \left(\frac{1}{\text{Exp}[\frac{T_E}{T}] - 1} + \frac{1}{2} \right) - N \phi$$

we find that the heat capacity takes the form

$$C_V = \frac{3}{4} N k_B \left(\frac{T_E}{T} \right)^2 \text{Csch} \left[\frac{T_E}{2T} \right]^2$$

With a little algebra one can show that

$$\begin{aligned} T \rightarrow \infty &\implies C_V \rightarrow 3N k_B \\ T \rightarrow 0 &\implies C_V \approx 3N k_B \left(\frac{T_E}{T} \right)^2 \text{Exp} \left[-\frac{T_E}{T} \right] \end{aligned}$$

Thus, at high temperature one obtains the classical equipartition theorem but at low temperature the heat capacity is exponentially suppressed.

Debye proposed a more realistic model of the energy spectrum in which $g[\omega] \propto \omega^2$ for $\omega \leq \omega_D$. The constant of proportionality is determined easily using the normalization condition

$$a \int_0^{\omega_D} \omega^2 d\omega = 3N \implies a = \frac{9N}{\omega_D^3} \implies g[\omega] = 9N \frac{\omega^2}{\omega_D^3} \Theta[\omega_D - \omega]$$

where the Heaviside truth function $\Theta[x]$ is unity for $x \geq 0$ and zero otherwise. For this model the internal energy can be expressed in the form

$$U + N \phi = \frac{9N \hbar}{\omega_D^3} \int_0^{\omega_D} d\omega \frac{\omega^3}{\text{Exp}[\beta \hbar \omega] - 1} = 9N \hbar \omega_D \left(\frac{T}{T_D} \right)^4 \int_0^{T_D/T} dy \frac{y^3}{e^y - 1}$$

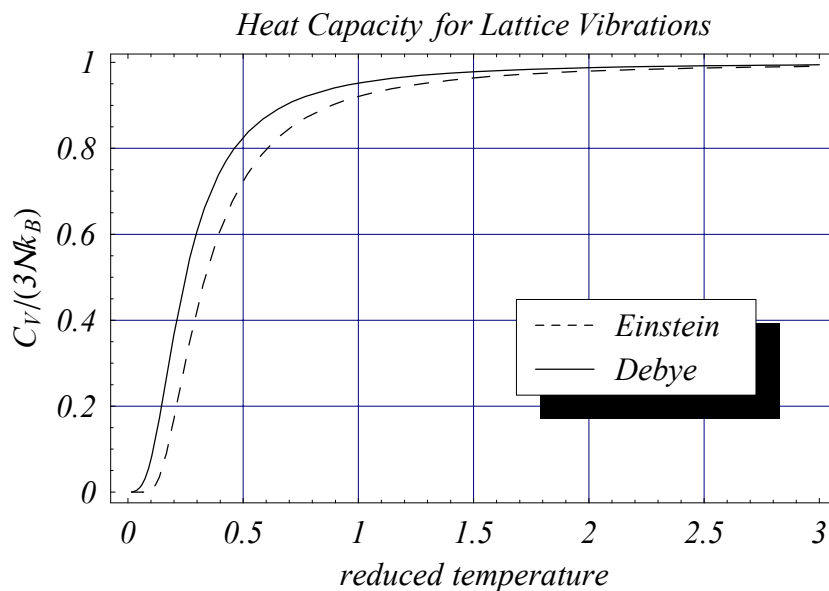
where the Debye temperature, T_D , is related to the Debye cut-off frequency, ω_D , by $k_B T_D = \hbar \omega_D$. Although this integral can be evaluated in terms of named functions, we refer the reader to the notebook *debye.nb* because those functions will not be familiar to most readers. For the present purposes it is sufficient to note that the low-temperature behavior can be obtained by extending the range of integration to ∞ and using the definite integral

$$\int_0^{\infty} dy \frac{y^3}{e^y - 1} = \frac{\pi^4}{15}$$

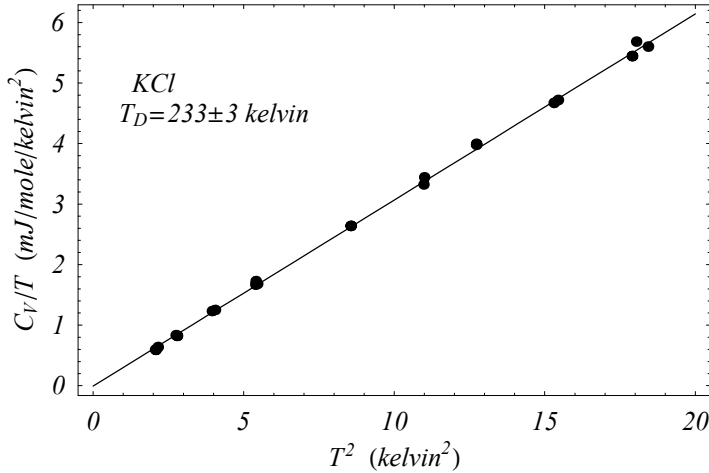
to obtain

$$T \ll T_D \Rightarrow U + N\phi \approx \frac{3\pi^4}{5} N \hbar \omega_D \left(\frac{T}{T_D}\right)^4 \Rightarrow C_V \approx \frac{12\pi^4}{5} N k_B \left(\frac{T}{T_D}\right)^3$$

The Debye model predicts the low-temperature heat capacity is governed by a power law rather than by an exponential function because its continuous energy spectrum does not feature an energy gap.



The low frequency or long wavelength excitations represent collective modes in which the displacements of neighboring atoms are similar — these are acoustical waves with correlated motions. Excitation quanta are smaller for low frequency than for high frequency, but the number of collective modes decreases as the correlation between atomic motions increases. Therefore, although the excitation spectrum is continuous and the crystal can accept arbitrarily small excitation energy (in the limit $N \rightarrow \infty$), the heat capacity is suppressed at low temperature because the fraction of the available modes that are thermodynamically active is proportional to $(T/T_D)^3$. We obtain a power-law instead of exponential suppression because there is no gap in the excitation spectrum for collective modes of oscillation. The figure below compares some of the earliest low-temperature data for the capacity in an insulating crystal with a fit based upon the predicted $(T/T_D)^3$ behavior. The data are fit very well by this power law and the fitted value of T_D is consistent with calculations based upon either elastic moduli or upon the acoustical excitation spectrum. Conversely, high-frequency modes represent essentially independent oscillations of individual atoms about their equilibrium positions with a limiting frequency that is inversely proportional to the lattice spacing. Therefore, at high temperature the thermodynamics of the system are well approximated by a system of independent oscillators.



Ref.: P. H. Keesom and N. Pearlman, Phys. Rev. 91, 1354 (1953)

It is also of interest to compare these results with those of the corresponding classical theory. A classical oscillator can accept infinitesimal excitation energy, such that the single-mode partition function is represented by an integral of the form

$$Z_\alpha = \int_0^\infty e^{-\beta \varepsilon} d\varepsilon = k_B T$$

instead of a discrete sum over quanta. Thus, the average energy per mode

$$\bar{\varepsilon}_\alpha = -\frac{\partial \text{Log}[Z_\alpha]}{\partial \beta} = k_B T$$

is independent of frequency and the internal energy

$$U + N \phi = \int_0^\infty d\omega g[\omega] \bar{\varepsilon}[\omega] = 3 N k_B T \implies C_V = 3 N k_B$$

is simply the total number of modes times the average energy per mode. Therefore, the classical heat capacity for lattice vibrations is independent of temperature. This result, known as the Dulong-Petit law, is an example of the classical equipartition theorem. However, the low-temperature data for insulating solids follow the T^3 behavior of the Debye theory and disagree with the classical prediction. The essential difference between the two theories is the discretization of the single-mode energy spectrum in the quantum theory.

Unfortunately, there are several serious inconsistencies in the classical theory. First, because the partition function should be dimensionless we should have included a density of states factor in our integral. In this case, the density of states would simply be constant with units of inverse energy, but classical physics offers no natural energy scale by which to determine its value. Consequently, the classical entropy is undetermined because it depends upon this scale parameter. More importantly, the number of modes for a classical system is infinite. Although we assumed that the number of modes is limited to the $3 N$ vibrational modes of N atoms, a classical mass can always be decomposed into smaller constituents bound together by internal forces. If an object appears to be rigid, that just means that its internal binding forces are strong. However, no matter how strong those binding forces are, it should always be possible to induce small-amplitude oscillations against those forces. To be more specific, we know that atoms consist of electrons bound to a central nucleus by electromagnetic forces. Classically we would expect small-amplitude vibrations about the equilibrium configuration of those constituents. Furthermore, the nucleus itself consists of protons and neutrons which should also exhibit vibrational modes. The nucleons themselves are composite objects which should possess a spectrum of normal modes. Therefore, the number of normal modes is much larger than $3 N$ and, if one takes classical physics seriously, should probably be infinite! In order to obtain finite results for heat capacities, classical theories must limit themselves to a finite set of active modes even though the theory offers no dynamical method for suppressing stiff internal modes. There is no natural energy scale

in the theory and all modes are capable of accepting energy even when their vibrational amplitudes are infinitesimal. Quantum mechanics, on the other hand, offers a natural method to decide whether a mode is active or passive. Due to the quantization of energies, oscillators whose quanta are larger than the typical thermal energy, $k_B T$, are inactive. Thus, all modes with $\hbar \omega \gg k_B T$ are *frozen out* and may be safely ignored. Consequently, the calculation of heat capacity need only include those modes that are potentially active.

Quantization is essential to the success of statistical mechanics. Although many textbooks begin with classical statistical mechanics, that approach is fundamentally flawed. Classical physics cannot compute entropy or heat capacity without appealing to quantum mechanics, or *ad hoc* prescriptions, to count the number of states and to limit the number of active modes. Therefore, we have chosen to develop statistical mechanics from a quantum mechanical perspective from the outset. Nevertheless, we often find that *semiclassical* models can be useful and provide valuable insight when derived as a suitable limit from quantum mechanics. We use quantum mechanics to evaluate the density of states and to decide which modes are active. We then use classical methods to simplify the calculations, when appropriate. This approach will be developed in *semiclassical.nb*.

Relationship between canonical and microcanonical ensembles

In the canonical ensemble, the probability that a system will be found with energy E between E and $E + dE$ is

$$P[E] dE = Z^{-1} e^{-\beta E} g[E] dE \qquad Z = \int dE g[E] e^{-\beta E}$$

where $g[E]$ is the density of states. Normally $g[E]$ is a rapidly increasing function of energy, resembling E^N , whereas the Boltzmann factor $e^{-\beta E}$ is a rapidly decreasing function. Thus, the product $g[E] e^{-\beta E}$ for a large system is usually sharply peaked about the most probable energy \tilde{E} . The most probable energy is determined simply by setting the derivative wrt energy to zero

$$\left(\frac{\partial \ln P[E]}{\partial E} \right)_{E=\tilde{E}} = 0 \implies \beta = \left(\frac{\partial \ln g[E]}{\partial E} \right)_{E=\tilde{E}}$$

with both Z and β treated as constants. The parameter $\beta = (k_B T_R)^{-1}$ is fixed by the temperature of the reservoir, T_R , which is unaffected by energy exchange with the system of interest, whose heat capacity is much smaller. If we interpret $\frac{1}{T} = k_B \frac{\partial \ln g}{\partial E}$ as the inverse temperature of the system, we find that maximizing the probability with respect to the internal energy of the system is equivalent to the equilibrium condition $T = T_R$. Therefore, the fixed energy in the microcanonical ensemble is equivalent to the most probable energy in the canonical ensemble. Similarly, the width of the peak in the canonical energy distribution is equivalent to the interval used by the microcanonical ensemble to bound the energy.

We can also compare the mean energy U with the most probable energy \tilde{E} .

$$U = \int dE P[E] E = Z^{-1} \int dE g[E] e^{-\beta E} E = - \frac{\partial}{\partial \beta} \ln Z$$

Near \tilde{E} , we expand

$$\text{Log}[g[E] e^{-\beta E}] \approx \text{Log}[g[\tilde{E}] e^{-\beta \tilde{E}}] + \frac{1}{2} (E - \tilde{E})^2 \left(\frac{\partial^2}{\partial E^2} \text{Log}[g[E] e^{-\beta E}] \right)_{E=\tilde{E}} + \dots$$

and define

$$\sigma^{-2} = -\left(\frac{\partial^2}{\partial E^2} \text{Log}[g[E] e^{-\beta E}]\right)_{E=\tilde{E}}$$

so that

$$P[E] \approx P[\tilde{E}] \text{Exp}\left[-\frac{(E - \tilde{E})^2}{2\sigma^2}\right]$$

displays a Gaussian shape in the vicinity of the peak. Expressing the density of states in terms of entropy, the variance in energy becomes

$$\frac{\partial^2}{\partial E^2} \text{Log}[g[E] e^{-\beta E}] = \frac{1}{k_B} \left(\frac{\partial^2 S}{\partial E^2}\right) \implies \sigma^{-2} = -\frac{1}{k_B} \left(\frac{\partial^2 S}{\partial E^2}\right)_{E=\tilde{E}}$$

If the width σ is sufficiently narrow that higher-order terms may be neglected, then $\tilde{E} \approx U$. Furthermore, the width of the energy distribution is related to thermodynamic properties by

$$\sigma^2 = -k_B \left(\frac{\partial^2 S}{\partial E^2}\right)_{E=\tilde{E}}^{-1} = -k_B \left(\frac{\partial}{\partial U} \frac{1}{T}\right)^{-1} = k_B T^2 C_V$$

where $C_V = \frac{\partial U}{\partial T}$ is the isochoric heat capacity. Since $C_V \propto N$ and $U \propto N$, normally, we expect that $\sigma \propto N^{1/2} \implies \frac{\sigma}{U} \propto N^{-1/2}$ so that for large systems the energy distribution becomes very sharp, practically a δ -function, except near a phase transition or other divergence in C_V .

Finally, we note that

$$g[U] e^{-\beta U} = \text{Exp}[\ln g] e^{-\beta U} = \text{Exp}[-\beta(U - TS)] = e^{-\beta F} \implies P[\tilde{E}] \approx e^{-\beta F}$$

where $F = U - TS$ is the Helmholtz free energy. Therefore, the canonical probability distribution becomes

$$P[E] \approx e^{-\beta F} \text{Exp}\left[-\frac{(E - U)^2}{2\sigma^2}\right] \quad \sigma^2 = k_B T^2 C_V$$

for E near $\tilde{E} = U$. Furthermore, the partition function becomes

$$Z = \int dE g[E] e^{-\beta E} \approx \int dE e^{-\beta F} \text{Exp}\left[-\frac{(E - U)^2}{2\sigma^2}\right] = e^{-\beta F} \sqrt{2\pi} \sigma$$

so that

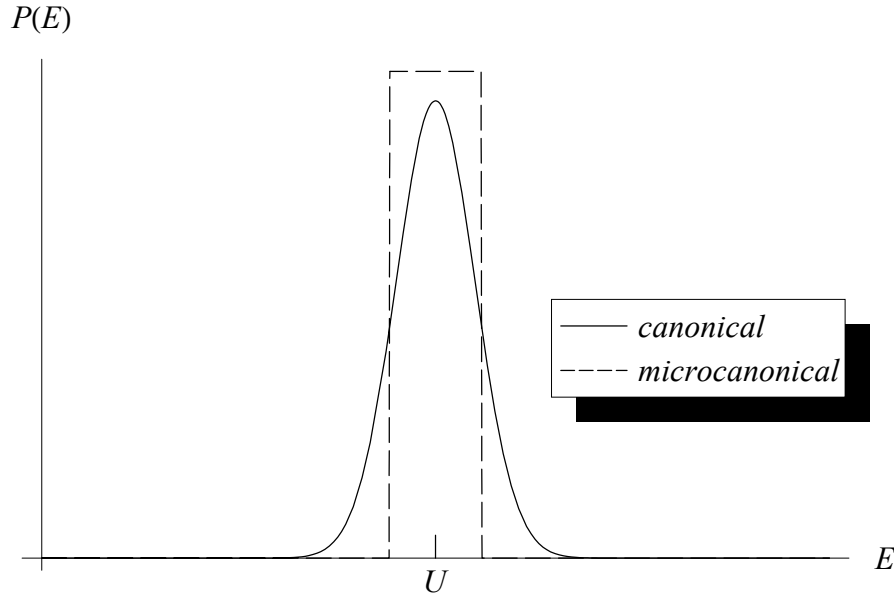
$$\ln Z \approx -\beta F + \frac{1}{2} \text{Log}[2\pi k_B T^2 C_V]$$

where the second term originates from the normalization of the probability distribution. However, since $F \propto N$ the second term, which is proportional to $\ln N$, may be neglected for large N . Hence, we conclude that $Z \approx e^{-\beta F}$ and that

$$P[E] \approx Z^{-1} \text{Exp}\left[-\frac{(E - U)^2}{2\sigma^2}\right]$$

near the peak.

Therefore, the canonical and microcanonical ensembles provide essentially equivalent statistical description of the thermodynamics of macroscopic systems. The two probability distributions compared below are equivalent for all practical purposes. For large systems the widths of these distributions become extremely small compared with U , so that the difference between Gaussian and square distributions becomes negligible.



Grand canonical ensemble

■ Derivation

Consider a small system in thermal and diffusive contact with a much larger reservoir with which it may exchange both heat and particles. The combined system is isolated, so that the total energy E and total particle number N is conserved. These are shared according to

$$\begin{aligned} E_R + E_i &= E_T & E_i &\ll E_R \\ N_R + N_i &= N_T & N_i &\ll N_T \end{aligned}$$

The probability that the sample is found in a particular state i is proportional to the total number of states available to the reservoir, such that

$$P_i = C \Omega_R[E_T - E_i, N_T - N_i]$$

Expanding $\ln \Omega_R$ for small (E_i, N_i) , we find

$$\ln \Omega_R[E_T - E_i, N_T - N_i] \approx \ln \Omega_R[E_T, N_T] - \beta(E_i - \mu N_i)$$

where the expansion coefficients

$$\beta = \left(\frac{\partial \ln \Omega_R}{\partial E_R} \right)_{E_R=E_T} \quad \beta \mu = - \left(\frac{\partial \ln \Omega_R}{\partial N_R} \right)_{N_R=N_T}$$

are defined by recognizing that the fundamental relation for $S = k_B \ln \Omega_R$ takes the form

$$T dS = dU + p dV - \mu dN \implies \left(\frac{\partial S}{\partial U} \right)_{V,N} = \frac{1}{T} \quad \left(\frac{\partial S}{\partial N} \right)_{U,V} = -\frac{\mu}{T}$$

It is not necessary to include subscripts on T or μ because thermal equilibrium requires those intensive parameters to be the same for both the sample and the reservoir. Hence, the reservoir multiplicity function takes the form

$$\Omega_R[E_T - E_i, N_T - N_i] \approx \Omega_R[E_T, N_T] \text{Exp}[-\beta(E_i - \mu N_i)]$$

Therefore, the population of states of a system in diffusive contact with a reservoir of energy and particles is governed by the *grand canonical distribution*

$$P_i = \mathcal{Z}^{-1} \text{Exp}[-\beta(E_i - \mu N_i)] \quad \mathcal{Z} = \sum_i \text{Exp}[-\beta(E_i - \mu N_i)]$$

where \mathcal{Z} is known as the grand partition function. The mean number of particles in the system is determined by the ensemble average

$$N = \mathcal{Z}^{-1} \sum_i \text{Exp}[-\beta(E_i - \mu N_i)] N_i = k_B T \left(\frac{\partial \ln \mathcal{Z}}{\partial \mu} \right)_{T,V}$$

The bridge to thermodynamics is most conveniently developed through the *grand potential*

$$\mathcal{G}[T, V, \mu] = F - \mu N = U - TS - \mu N$$

The grand potential can be simplified using the Euler relation

$$U = TS - pV - \mu N \implies \mathcal{G} = -pV$$

In terms of its natural independent variables, the fundamental relation for \mathcal{G} now takes the form

$$d\mathcal{G} = -S dT - p dV - N d\mu$$

from which we identify

$$S = -\left(\frac{\partial \mathcal{G}}{\partial T} \right)_{V,\mu} \quad p = -\left(\frac{\partial \mathcal{G}}{\partial V} \right)_{T,\mu} \quad N = -\left(\frac{\partial \mathcal{G}}{\partial \mu} \right)_{T,V}$$

The relationship for N permits the identifications

$$\mathcal{G} = -k_B T \ln \mathcal{Z} \implies \frac{pV}{k_B T} = \ln \mathcal{Z}$$

The validity of this identification depends upon recovering the thermodynamic expressions for S and p from the statistical expressions for the grand partition function. Differentiating \mathcal{G} with respect to T , we find

$$\frac{\partial \mathcal{G}}{\partial T} = \frac{\mathcal{G}}{T} - k_B T \frac{\partial \ln \mathcal{Z}}{\partial T} = \frac{1}{T} \left(\mathcal{G} + \frac{\partial \ln \mathcal{Z}}{\partial \beta} \right)$$

From the definition of \mathcal{Z} , we find

$$\frac{\partial \ln \mathcal{Z}}{\partial \beta} = -(U - \mu N) \implies \frac{\partial \mathcal{G}}{\partial T} = \frac{1}{T} (\mathcal{G} - U + \mu N) = -S$$

in agreement with the thermodynamic relationship. Similarly, differentiating \mathcal{G} with respect to V , we find

$$\frac{\partial \mathcal{G}}{\partial V} = -k_B T \frac{\partial \ln \mathcal{Z}}{\partial V}$$

However, $\ln \mathcal{Z}$ depends upon V only through its dependence on the energy levels E_i , whereby

$$\left(\frac{\partial \ln \mathcal{Z}}{\partial V}\right)_{T,\mu} = -\frac{\beta}{\mathcal{Z}} \sum_i \text{Exp}[-\beta(E_i - \mu N_i)] \frac{\partial E_i}{\partial V} = \beta \left\langle -\frac{\partial E}{\partial V} \right\rangle = \beta p$$

Thus, the thermodynamic relationship for pressure,

$$p = k_B T \left(\frac{\partial \ln \mathcal{Z}}{\partial V}\right)_{T,\mu} = -\left(\frac{\partial \mathcal{G}}{\partial V}\right)_{T,\mu}$$

is also recovered from the grand partition function. Finally, we note that

$$U - \mu N = \frac{\partial \ln \mathcal{Z}}{\partial \beta} \implies U = \mu N + \frac{\partial}{\partial \beta} (\beta \mathcal{G}) = \mu N + \mathcal{G} + T S$$

reproduces the Euler relation. Therefore, the standard thermodynamic relationships can be extracted from the grand partition function through the identification of the grand potential as

$$\mathcal{G}[T, V, \mu] = U - T S - \mu N = -p V = -k_B T \ln \mathcal{Z} \implies \frac{p V}{k_B T} = \ln \mathcal{Z}$$

Alternatively, the grand canonical probability distribution can be derived by maximizing the disorder within an ensemble of isolated systems with specified mean energy and particle number.

$$\begin{aligned} dS = 0 &\implies \sum_i (1 + \ln P_i) dP_i = 0 \\ U = \sum_i P_i E_i &\implies \sum_i E_i dP_i = 0 \\ N = \sum_i P_i N_i &\implies \sum_i N_i dP_i = 0 \\ 1 = \sum_i P_i &\implies \sum_i dP_i = 0 \end{aligned}$$

Using Lagrange multipliers, we find

$$\ln P_i + \beta E_i - \beta \mu N_i + \ln \mathcal{Z} = 0 \implies P_i = \mathcal{Z}^{-1} \text{Exp}[-\beta(E_i - \mu N_i)]$$

where, blessed with hindsight, we have chosen to associate Lagrange multipliers β , $\beta \mu$, and $\ln \mathcal{Z}$ with the energy, particle number, and normalization constraints from the outset.

■ Relationship to canonical ensemble

The microcanonical ensemble stipulates that each state whose energy lies within a specified narrow energy interval occurs with equal probability; the multiplicity $\Gamma[E] = g[E] \delta E$ is used to normalize the probability distribution and hence serves as a microcanonical partition function. The canonical probability distribution $P[E] \propto e^{-\beta E} g[E] dE$ describes an ensemble composed of a collection of microcanonical ensembles with various energies where each subensemble is weighted by the Boltzmann factor $e^{-\beta E}$. The canonical partition function which normalizes the probability distribution for this ensemble of ensembles is then a weighted sum of microcanonical partition functions, namely $Z = \int dE g[E] e^{-\beta E}$. Similarly, the grand canonical ensemble consists of a canonical ensemble with different particle numbers weighted by the relative probability $e^{\beta \mu N}$. It is convenient to define the fugacity parameter as $z = e^{\beta \mu}$, whereby the grand partition function becomes

$$\mathcal{Z}[T, V, \mu] = \sum_N z^N Z_N[T, V] = \sum_N e^{\beta \mu N} Z_N[T, V]$$

where $Z_N[T, V]$ is the canonical partition function for a system of N particles. Thus, the factor z^N plays a role similar to that of the Boltzmann factor. The grand partition function is now seen to be a weighted sum of canonical partition functions for subensembles with varying numbers of particles.

The relationship between the canonical and grand canonical ensembles can be clarified further by recognizing that for large N , $\ln \mathcal{Z}$ will be dominated by its single largest term, such that

$$0 = \frac{\partial}{\partial N} \text{Log}[e^{\beta\mu N} Z_N] = \beta\mu + \frac{\partial}{\partial N} \text{Log}[Z_N] \quad \text{when } N \rightarrow \tilde{N}$$

However, extensivity of thermodynamic potentials requires $\text{Log}[Z_N] \propto N$ so that the maximization condition becomes

$$0 = \beta\mu + \frac{1}{\tilde{N}} \text{Log}[Z_{\tilde{N}}] \implies \mu \tilde{N} = -k_B T \text{Log}[Z_{\tilde{N}}]$$

Now, using $z = e^{\beta\mu} \implies \mu = k_B T \text{Log}[z]$, we find that the fugacity satisfies

$$\tilde{N} \text{Log}[z] = -\text{Log}[Z_{\tilde{N}}] \implies Z_{\tilde{N}} = z^{-\tilde{N}}$$

at the peak of the N distribution. Thus, the most probable particle number is determined by the condition

$$\frac{\mu \tilde{N}}{k_B T} = -\text{Log}[Z_{\tilde{N}}] \implies Z_{\tilde{N}}^{1/\tilde{N}} = \frac{\mu}{k_B T}$$

Although it may appear that the grand canonical ensemble is more complicated than the other ensembles, sometimes unnecessarily so, it nevertheless plays an important role in a wide variety of applications. It is often inconvenient to enforce a strict constraint upon particle number or an approximation method may violate particle number conservation, in which case the grand canonical ensemble may provide a more convenient method to adjust the particle number through μ . Similarly, the grand canonical ensemble provides an easier method for analyzing the effects of indistinguishability among identical particles than does the canonical ensemble. Furthermore, in many systems the particle number really is indeterminate, such as for photons or phonons. For other systems, particles may be exchanged between several components of the system, with the equilibrium division of particles governed by the chemical potential. For example, when a gas is in equilibrium with its own liquid phase, with these two component separated by gravity, the particle numbers in the two phases are statistical variables. Since the number in the condensed phase is usually much larger, the condensed phase serves as a particle reservoir for the vapor. The vapor pressure is then governed by the chemical potential established by the reservoir.

Density operators for canonical ensembles

In the energy representation the microcanonical ensemble is represented by the density matrix

$$\rho_n = \Gamma^{-1} \Theta[\delta E - 2 | E - E_n]]$$

where the Heaviside or unit step function is defined by $\Theta[x] = 1$ for $x > 0$, $\Theta[x] = \frac{1}{2}$ for $x = 0$, and $\Theta[x] = 0$ for $x < 0$. Thus, in an arbitrary representation we could represent the microcanonical density operator in the form

$$\hat{\rho} = \Gamma^{-1} \Theta[\delta E - 2 | E - \hat{H}]] \quad \Gamma = \text{Tr} \Theta[\delta E - 2 | E - \hat{H}]]$$

However, this ugly operator is not particularly well suited to calculation, for which one should always return to the energy representation. Nevertheless, it does provide a formal definition of the ensemble.

The canonical ensemble is represented in the energy representation by the density matrix elements

$$\rho_n = Z^{-1} e^{-\beta E_n} \quad Z = \sum_n e^{-\beta E_n}$$

where the summation is with respect to states (not levels). Thus, in an arbitrary basis we can employ the canonical density operator

$$\hat{\rho} = Z^{-1} e^{-\beta \hat{H}} \quad Z = \text{Tr} e^{-\beta \hat{H}}$$

where the exponential is interpreted as a power series. The ensemble average for any observable is then

$$\langle \hat{f} \rangle = \text{Tr} \hat{\rho} \hat{f} = \frac{\text{Tr} e^{-\beta \hat{H}} \hat{f}}{\text{Tr} e^{-\beta \hat{H}}}$$

In particular, the internal energy is given by

$$U = \langle \hat{H} \rangle = -\frac{\partial}{\partial \beta} \text{Log}[\text{Tr} e^{-\beta \hat{H}}] = -\frac{\partial}{\partial \beta} \text{Log}[Z]$$

Similarly, the entropy is

$$S = \langle -k_B \text{Log}[\hat{\rho}] \rangle = -k_B \text{Tr} \hat{\rho} \text{Log}[\hat{\rho}] = -k_B Z^{-1} \text{Tr}[e^{-\beta \hat{H}} (-\beta \hat{H} - \text{Log}[Z])]$$

or

$$S = \frac{1}{T} (U + k_B T \text{Log}[Z])$$

Therefore, we find

$$F = U - TS = -k_B T \text{Log}[Z] \implies Z = e^{-\beta F}$$

in agreement with our previous results. Furthermore, we can define a Helmholtz operators as

$$\hat{F} = \hat{H} - T \hat{S} = \hat{H} + k_B T \text{Log}[\hat{\rho}]$$

so that

$$F = \text{Tr} e^{-\beta \hat{H}} \hat{F} = \text{Tr} \hat{\rho} \hat{H} - T \text{Tr} \hat{\rho} \hat{S} = U - TS$$

Thus, we can complete the circle by solving for $\hat{\rho}$, with the result

$$\hat{\rho} = e^{\beta(\hat{F} - \hat{H})} = e^{-\hat{S}/k_B}$$

Following this pattern, the grand canonical density matrix becomes

$$\rho_n = \mathcal{Z}^{-1} e^{-\beta(E_n - \mu N_n)} \quad \mathcal{Z} = \sum_n e^{-\beta(E_n - \mu N_n)}$$

in the energy representation. However, since this ensemble does not have fixed particle number, its states and operators do not live in a unique Hilbert space but must be constructed in a *Fock space* consisting of the union of all Hilbert spaces of similar nature but different particle numbers. Hence, in the Fock space there must exist a number operator \hat{N} such that $\hat{N} |n\rangle = N_n |n\rangle$ for state n where N_n is the eigenvalue of the number operator in a representation where both \hat{H} and \hat{N} are

diagonal. The state index n subsumes the entire set of quantum numbers, include particle number, needed to specify a state in Fock space. The corresponding statistical operators are then

$$\hat{\rho} = \mathcal{Z}^{-1} e^{-\beta(\hat{E} - \mu \hat{N})} \quad \mathcal{Z} = \text{Tr} e^{-\beta(\hat{E} - \mu \hat{N})}$$

such that

$$S = -k_B \text{Tr} \hat{\rho} \text{Log}[\hat{\rho}] = -k_B \text{Tr}[\hat{\rho}(-\beta \hat{H} + \beta \mu \hat{N} - \text{Log}[\mathcal{Z}])]]$$

or

$$TS = U - \mu N + k_B T \text{Log}[\mathcal{Z}]$$

where

$$U = \text{Tr} \hat{\rho} \hat{H} \quad N = \text{Tr} \hat{\rho} \hat{N}$$

with the traces taken with respect to Fock space, which includes a summation over particle number. Thus, we find

$$\mathcal{Z} = \text{Exp}[-\beta(U - TS - \mu N)] = e^{-\beta \mathcal{G}} \implies \mathcal{G} = -k_B T \text{Log}[\mathcal{Z}]$$

where $\mathcal{G} = U - TS - \mu N = -pV$ is the grand potential, as before. Therefore, the canonical density operator can also be expressed as

$$\hat{\rho} = e^{\beta(\mathcal{G} - \hat{H} + \mu \hat{N})}$$

Finally, if we define an operator for \mathcal{G} , we find that

$$\hat{\mathcal{G}} = \hat{H} - T \hat{S} - \mu \hat{N} \implies \hat{\rho} = e^{-\hat{\mathcal{G}}/k_B}$$

provides a closed loop of operator definitions.

Therefore, the fundamental definition

$$\hat{\rho} = e^{-\hat{\mathcal{G}}/k_B} \implies \hat{S} = -k_B \text{Log}[\hat{\rho}]$$

can be regarded as the primary template for the density operators corresponding to various ensembles. By expressing S in terms of the appropriate thermodynamic variables, replacing those that are variable by operators and treating those that are constrained as numbers, one can construct the desired canonical density matrix practically by inspection.

$$\begin{aligned} \text{variable } E; \text{ fixed } T, V, N &\implies \hat{S} = \frac{1}{T} (\hat{H} - F) &\implies \hat{\rho} = \text{Exp}[\beta(F - \hat{H})] \\ \text{variable } E, N; \text{ fixed } T, V, \mu &\implies \hat{S} = \frac{1}{T} (\hat{H} - \mu \hat{N} - \mathcal{G}) &\implies \hat{\rho} = \text{Exp}[\beta(\mathcal{G} - \hat{H} + \mu \hat{N})] \\ \text{variable } E, V; \text{ fixed } T, p, N &\implies \hat{S} = \frac{1}{T} (\hat{H} + p \hat{V} - G) &\implies \hat{\rho} = \text{Exp}[\beta(G - \hat{H} - p \hat{V})] \end{aligned}$$

The thermodynamic potential whose natural variables are those specified by the external constraints provides the normalization of the density operator.

Problems

▼ Maximum disorder in banded system

Suppose that the states available to a particular system can be divided into a sequence of bands where each band contains g_i equally likely states with energy ε_i . Thus, for this problem the index i enumerates energy levels rather than quantum states. If g_i is large, there are many different states all with the same energy ε_i . Let p_i be the probability that the system is found in some state within band i . The system can exchange energy with its environment, but its average internal energy, U , is known.

- Find a general expression for the entropy of this system in terms of g_i and p_i . It is crucial to remember that disorder sums over states, not levels, so that the degeneracy of each level must be considered carefully.
- Use the method of Lagrange multipliers to determine the probabilities p_i that maximize entropy subject to the constraint that the internal energy is fixed. Express the Lagrange multipliers in terms of the probability normalization and the temperature of the system.
- Suppose that there are just two levels with energies $\varepsilon = \bar{\varepsilon} \pm \frac{\Delta}{2}$. Express the average internal energy in terms of temperature, the band gap Δ , and the ratio $r = g_2/g_1$ where $\varepsilon_2 > \varepsilon_1$. Explain both the low and high temperature limits.
- Evaluate the heat capacity for the two-level system and produce sketches for several values of r . How does the Schottky peak change as r increases?

▼ A three-level system with degeneracy

Consider a system of \mathcal{N} noninteracting molecules which possess three possible configurations, two of which are degenerate. It is convenient to choose an energy scale which places the ground state at zero and the other two levels at ε . Let $\{\mathcal{N}_i, i = 0, 1, 2\}$ represent the numbers of molecules in each level.

- Evaluate the multiplicity function $g[\mathcal{N}_0, \mathcal{N}_1, \mathcal{N}_2]$.
- Express the entropy in terms of internal energy and total particle number assuming that all \mathcal{N}_i are large.
- Evaluate and sketch the temperature dependence of the heat capacity for this system.

▼ Canonical partition function for Einstein crystal

In the text we evaluated the canonical partition function for an Einstein crystal by treating the normal modes of vibration as distinguishable subsystems with energy $\varepsilon_n = (n + \frac{1}{2}) \hbar \omega$. Alternatively, we can compute the partition function by summation over the energy levels for the entire system of N particles if we properly account for the degeneracy of each energy level. The energy levels can be expressed as

$$E_n = n \hbar \omega + N_s \left(\frac{3}{2} \hbar \omega - \phi \right)$$

where n is the total number of quantum to be distributed among $3N$ degrees of freedom and ϕ is the binding energy per particle. Demonstrate that this method gives the same result.

▼ Simple model of polymer

Suppose that a polymer consists of a chain with N links of length ℓ and that each link is equally likely to be directed either toward the right or toward the left. We assume that the links can turn freely. This is the simplest model of rubber elasticity.

- Deduce the total number of configurations $g[n, N]$ which result in a total length $L = 2n\ell$.
- Find a simple expression for the entropy in the limit $n \ll N$ and explain why this condition represents the high temperature limit.
- More generally, determine the tension on the polymer as a function of its temperature and length. Sketch the result. Under what conditions is Hooke's law valid? [Hint: the fundamental thermodynamic relation for a simple extensible system has the form $dU = T dS + \tau dL$ where τ is the tension and where U is constant for this trivial model.]

▼ Frenkel defects

A perfect crystal consists of N atoms arranged on a regular lattice. Suppose that n atoms, where $1 \ll n \ll N_1$, are moved from lattice sites to interstitial sites and that there are N_2 possible interstitial sites, where $N_1 \approx N_2 \approx N$. Thus, the crystal becomes an imperfect crystal with n defects of the Frenkel type. If the energy required to move a single atom from a lattice site to an interstitial site is ϵ , determine the equilibrium number of defects at temperature T . Sketch and discuss $S[T]$, $U[T]$, and $C_V[T]$.

▼ One-dimensional Ising model

The one-dimensional Ising model of a magnetic system consists of a linear chain of N spins with two states each, either \uparrow or \downarrow . The energy contains two contributions.

- The nearest-neighbor coupling energy is $+J$ for antiparallel neighbors ($\uparrow\downarrow$ or $\downarrow\uparrow$) or $-J$ for parallel neighbors ($\uparrow\uparrow$ or $\downarrow\downarrow$).
- The interaction with the external magnetic field B , defined to be positive when parallel to the \uparrow direction, is $-\mu B$ for spin-up or $+\mu B$ for spin-down.

The energy of the system is determined by the number of up spins N_\uparrow , the number of down spins N_\downarrow , and the number of spin-flips N_\ddagger . For example, the following arrangement of spins

$$\uparrow \uparrow \downarrow \downarrow \downarrow \uparrow \downarrow \downarrow \downarrow \uparrow \uparrow \downarrow$$

has $N_\uparrow = 5$, $N_\downarrow = 6$, and $N_\ddagger = 5$ where the spin-flips are represented by the symbol \ddagger .

- Show that the total energy of a particular configuration of spins is given by

$$E = -J(N - 2N_{\uparrow}) - \mu B(N_{\uparrow} - N_{\downarrow})$$

b) Demonstrate that the multiplicity function for this system is approximated well by

$$g[N_{\uparrow}, N_{\downarrow}, N_{\uparrow\downarrow}] \approx 2 \frac{(N_{\uparrow}! N_{\downarrow}!)}{\left((N_{\uparrow} - \frac{N_{\uparrow\downarrow}}{2})! (N_{\downarrow} - \frac{N_{\uparrow\downarrow}}{2})! \left((\frac{N_{\uparrow\downarrow}}{2})! \right)^2 \right)}$$

Then find an expression for the total number of states, $\Omega[E, N, M]$, with energy E , particle number N , and magnetization M .

c) Use Stirling's approximation and the method of Lagrange multipliers to determine the values of N_{\uparrow} , N_{\downarrow} , and $N_{\uparrow\downarrow}$ which maximize the disorder in this system subject to the constraints that the total energy and particle number are fixed.

d) The general case is difficult, but the entropy and energy can be obtained analytically when $B = 0$. Show that

$$U = -NJ \tanh[\beta J] \quad \beta = \left(\frac{\partial \ln \Omega}{\partial E} \right)_N$$

for $B = 0$. What is the net magnetization under these conditions?

e) **Extra credit:** More ambitious students might try to solve the general case ($B \neq 0$) using *Mathematica*. This can be done either numerically or symbolically. Make plots which show the dependence of U upon B for several choices of β , μ , and J .

▼ Two-level system with degeneracies using canonical ensemble

Suppose that the states available to a particular system can be divided into two groups consisting of: 1) g_1 equally likely states with energy ε_1 ; and 2) g_2 equally likely states with energy $\varepsilon_2 > \varepsilon_1$. Let p_i be the probability that the system is found in a state within group i , such that $p_1 + p_2 = 1$. For example, the lower level might represent the valence band and the upper level the conduction band of a semiconductor, assuming that the gap between bands is large compared with the band widths.

a) Find a general expression for the entropy of this system in terms of g_i and p_i .

b) Evaluate the entropy assuming that the p_i follow the canonical probability distribution for temperature T .

c) Determine the entropy in both the low-temperature and high-temperature limits and discuss the physical interpretation of these limiting cases.

d) Express the heat capacity in terms of $r = g_2/g_1$ and $y = \Delta/k_B T$ where $\Delta = \varepsilon_2 - \varepsilon_1$ and produce sketches for several values of r . How does the Schottky peak change as r increases?

▼ A 4-state problem using density matrix

A system consisting of \mathcal{N} independent components with 4 internal states described by a single-particle hamiltonian of the form

$$H = \begin{pmatrix} \varepsilon_0 - \varepsilon_1 & 0 & 0 & -2\varepsilon_1 \\ 0 & \varepsilon_0 + \varepsilon_1 & 0 & 0 \\ 0 & 0 & \varepsilon_0 + \varepsilon_1 & 0 \\ -2\varepsilon_1 & 0 & 0 & \varepsilon_0 - \varepsilon_1 \end{pmatrix}$$

is in thermal equilibrium with a reservoir at temperature T .

- Construct the canonical single-particle density matrix $\rho = e^{-\beta H}$.
- Evaluate the mean energy per particle using the trace technique.
- Evaluate the mean entropy and free energy per particle.
- Verify that the same results are obtained by the usual partition-function method.

▼ Fluctuation of spin $\frac{1}{2}$ magnetization

Evaluate the root-mean-square fluctuation $\langle (\hat{M} - \bar{M})^2 \rangle$ of the magnetization of a system of N independent spin $\frac{1}{2}$ atoms in equilibrium at temperature T within an external magnetic field B . Compare with the mean magnetization, \bar{M} , itself.

▼ Polar gas

Suppose that a classical ideal gas consists of polar molecules with static electric dipole moment p_0 . In the presence of an electric field the energy of orientation is then $\varepsilon = -\vec{p} \cdot \vec{E}$. We seek to evaluate the electric susceptibility, $\chi_e = \frac{\partial P}{\partial E}$, where P is the average dipole moment per unit volume.

- Argue that the sum over orientations is equivalent to an integral with respect to $d\theta$, with $x = \cos \theta$.
- Evaluate the partition function for molecular orientation and compute the average dipole moment.
- Compute the susceptibility as a function of temperature and evaluate both low and high temperature limits. Sketch $\chi_e[T]$ and briefly discuss its characteristics.
- Estimate the characteristic orientational temperature for a typical molecular dipole moment in a field of 10^4 V/m.

▼ Ising ring

A more elegant solution for the one-dimensional Ising model uses the canonical ensemble to analyze a ring of spins. Assume that N spin $\frac{1}{2}$ atoms with magnetic moment μ are equally spaced around a circle. The energy of a particular arrangement of spins within a magnetic field B that is perpendicular to the ring can be expressed as

$$E = \sum_{i=1}^N (-\mu B \sigma_i - J \sigma_i \sigma_{i+1})$$

where $\sigma_i = \pm 1$ is the direction of spin of atom i , J represents the spin-spin interaction between nearest neighbors, and periodic boundary conditions $\sigma_{i+N} = \sigma_i$ are used on the ring.

- Show that the partition function can be represented in the form

$$Z = \sum_{\{\sigma_i\}} \prod_{i=1}^N z_{i,i+1}$$

where the sum includes all possible arrangements of the spins $\{\sigma_i = \pm 1\}$ and

$$z_{i,i+1} = \text{Exp}\left[\beta\left(\frac{\mu B}{2}(\sigma_i + \sigma_{i+1}) + J\sigma_i\sigma_{i+1}\right)\right]$$

is the partition function for a pair of spins. Then, recognizing that $z_{i,i+1}$ is a 2×2 matrix, show that $Z = \text{Tr } z^N$.

b) Using the fact that a real symmetric matrix can be diagonalized by an orthogonal transformation, obtain a simple but general expression for Z in terms of the eigenvalues of z . Thus, demonstrate that for large N we may use $F \approx -N k_B T \ln \lambda_1$ where λ_1 is the larger eigenvalue of z .

c) Now find explicit expressions for these eigenvalues and compute the magnetization and heat capacity of the system.

▼ Vibrating chain

Consider a vibrating chain consisting of N equal masses coupled by elastic bonds. This system can provide an improved model of one-dimensional lattice vibrations or a model of the transverse modes of very long polymers. The energy of the chain is $E = \sum_{k=1}^N n_k \hbar \omega_k$ where n_k is the number of vibrational quanta in a normal mode with frequency ω_k . [Note that we omit the zero-point energy throughout this problem.]

a) Show that the partition function can be expressed in the form

$$Z = \prod_{k=1}^N Z_k \quad Z_k = \frac{1}{1 - e^{-\beta \varepsilon_k}}$$

where $\varepsilon_k = \hbar \omega_k$ is the energy quantum for a mode with frequency ω_k and Z_k is the corresponding single-mode partition function.

b) Find an expression for the Helmholtz free energy, $F[T, N]$, and deduce the internal energy and the entropy directly from the free energy.

c) Recall that the frequency spectrum for a large linear chain has the form $\omega_k = \omega_N \text{Sin}\left[\frac{k\pi}{2N}\right]$. Using

$$\sum_{k=1}^N \dots \rightarrow N \int_0^1 \dots dx \quad \text{with } x = \frac{k}{N}$$

obtain an integral expression for the internal energy.

d) Evaluate the heat capacity in the high-temperature limit, $k_B T \gg \varepsilon_N$.

e) Evaluate the heat capacity in the low-temperature limit, $k_B T \ll \varepsilon_N$. However, since we are interested in low but not ultralow temperature, assume that the chain is large enough so that there remain many modes for which $k_B T \gg \varepsilon_k$.

f) How is the number of modes related to the number of atoms when this model is used to describe: i) longitudinal vibrations in one dimension, or ii) transverse vibrations of a linear chain? How should your results be modified for these situations.

▼ Einstein crystal with density dependence

The Einstein model of lattice vibrations can be improved by including a dependence of the vibration frequency upon the density of the crystal. We assume that the atoms are bound to their lattice sites by a harmonic potential such that the vibrational energy for mode n becomes

$$\varepsilon_n = (n_x + n_y + n_z) \hbar\omega - \varepsilon_B + \frac{\kappa}{2} \left(\frac{V - \sigma_0 N}{N} \right)^2$$

where ε_B is the mean binding energy, κ is an elasticity parameter, and σ_0 is the mean volume per atom.

a) Construct the partition function $Z[T, V, N]$ and calculate the Helmholtz free energy, F . Use F to determine the pressure and then evaluate the isothermal compressibility, $\kappa_T = -\frac{1}{V} \left(\frac{\partial V}{\partial p} \right)_T$, and isobaric expansivity, $\alpha = \frac{1}{V} \left(\frac{\partial V}{\partial T} \right)_p$. Interpret these results.

b) Deduce the chemical potential from F . Then calculate the Gibbs free enthalpy, $G[T, p, N] = F + pV$, as a function of temperature, pressure, and particle number. Show that $G[T, p, N] = N g[T, p]$ and compare the Gibbs free energy per particle with the chemical potential.

c) Use G to determine the equilibrium density.

d) Use F to express entropy as a function of T , V , and N . Then calculate the isochoric and isobaric heat capacities,

$$C_V = T \left(\frac{\partial S}{\partial T} \right)_{V,N} = \left(\frac{\partial U}{\partial T} \right)_{V,N} \quad C_p = T \left(\frac{\partial S}{\partial T} \right)_{p,N} = \left(\frac{\partial H}{\partial T} \right)_{p,N}$$

[Remember: enthalpy is defined as $H = U + pV$.] You should find that $C_p = C_V$ for this model. What further refinements of the model could produce a difference between the principal heat capacities?

▼ Sensitivity of low-temperature C_V to excitation spectrum

Consider a system for which the excitation spectrum satisfies a dispersion relation of the form $\omega \propto k^s$. Show that such excitations give a contribution to the heat capacity that at low temperatures is proportional to $T^{3/s}$. Compare graphically the temperature dependence of the heat capacities for $s = 2$, which applies to spin waves propagating in a ferromagnetic system, to those for $s = 1$, which applies to elastic waves in the lattice.

▼ Heat capacity of layered crystal

Graphite exhibits a layered crystalline structure for which the restoring forces parallel to each layer are much larger than those perpendicular to the layers. Hence, lattice vibrations can be separated into modes in which the atoms vibrate within the layer, characterized by Debye frequency ω_{\parallel} , and modes in which atoms vibrate perpendicular to the layers, characterized by Debye frequency ω_{\perp} , where $\omega_{\parallel} \gg \omega_{\perp}$. Vibrations within the plane are two-dimensional, whereas vibrations perpendicular to planes are one-dimensional.

a) Show that the heat capacity for lattice vibrations in an n -dimensional Debye crystal can be expressed in the form $C_V[T] \propto \tau^n f_n[\tau]$ where τ is an appropriate reduced temperature and $f_n[\tau]$ is a definite integral with τ as one of its limits.

b) Obtain explicit expressions for the temperature dependence of the heat capacity of graphite in each of the following regimes:

i) $k_B T \ll \hbar \omega_{\perp}$;

ii) $\hbar \omega_{\perp} \ll k_B T \ll \hbar \omega_{\parallel}$;

iii) $k_B T \gg \hbar \omega_{\parallel}$.

c) Sketch $C_V[T]$ assuming that $\omega_{\parallel}/\omega_{\perp} \sim 10$.

▼ (T,p) ensemble

Suppose that a small system (sample) is in thermal contact with a very much larger reservoir of heat. Further, suppose that these systems are separated by a movable wall. Therefore, in equilibrium the reservoir maintains constant temperature and pressure but permits fluctuations of the volume and energy of the sample. We seek to develop a statistical ensemble with temperature and pressure as its thermodynamic variables.

a) Construct the probability distribution and exhibit the partition function, $Z = Z[T, p, N]$, for this ensemble. This derivation should be performed using both methods we employed for the canonical ensemble; namely i) Taylor expansion of the multiplicity function for the reservoir, or ii) maximizing disorder subject to constraints on average energy and volume. How are these methods related?

b) Express the mean volume occupied by the sample in terms of the partition function.

c) Which thermodynamic potential provides the most natural connection between the statistical and thermodynamic descriptions of this ensemble? Express this potential in terms of the partition function and demonstrate that its derivatives are consistent with the appropriate form of the fundamental relation of thermodynamics.

d) Express the mean square volume fluctuation $\langle(\Delta V)^2\rangle$ in terms of the isothermal compressibility and evaluate the relative volume fluctuation $\langle(\Delta V)^2\rangle^{1/2}/\langle V\rangle$ for an ideal gas.

▼ Density fluctuations

a) Use the grand canonical ensemble to derive a general relationship for the variance $\sigma_{\rho}^2 = \langle \rho^2 \rangle - \langle \rho \rangle^2$ in particle density $\rho = N/V$ assuming fixed volume but variable particle number.

- b) Use thermodynamic methods to relate σ_ρ^2 to the isothermal compressibility, $\kappa_T = -\frac{1}{V} \left(\frac{\partial V}{\partial p} \right)_{T,N}$, assuming that the system is large enough to use extensivity. [Hints: use the Gibbs-Duhem relation and the fact that μ depends only on temperature and pressure. Be careful to distinguish between density, ρ , and pressure, p .]
- c) Explain the behavior of σ_ρ^2 at the liquid-vapor critical point.

▼ Correlation between energy and density

- a) Use the grand canonical ensemble to demonstrate that the correlation $\langle \Delta E \Delta N \rangle$ between fluctuations in energy and particle number is given by

$$\langle \Delta E \Delta N \rangle = \left(\frac{\partial U}{\partial N} \right)_\beta \langle (\Delta N)^2 \rangle$$

where

$$\langle (\Delta N)^2 \rangle = \frac{N^2 k_B T}{V} \kappa_T$$

Why are these variables correlated? What happens near the critical point for a fluid?

- b) Evaluate $\langle \Delta E \Delta \rho \rangle$ for an ideal gas, where $\rho = N/V$ is the particle density.

▼ Fluctuation in $E - \mu N$

Evaluate the variance in the quantity $J = E - \mu N$ for the grand canonical ensemble.