# IV. Classical Statistical Mechanics

## IV.A  General Definitions

• **Statistical Mechanics** *is a probabilistic approach to equilibrium macroscopic proper-
ties of large numbers of degrees of freedom.*

As discussed in chapter I, equilibrium properties of macroscopic bodies are phe-
nomenologically described by the laws of thermodynamics. The macro-state $M$, depends
on a relatively small number of thermodynamic coordinates. To provide a more funda-
mental derivation of these properties, we can examine the dynamics of the many degrees
of freedom $N$, comprising a macroscopic body. Description of each micro-state $\mu$, requires
an enormous amount of information, and the corresponding time evolution, governed by
the Hamiltonian equations (discussed in chapter II), is usually quite complicated. Rather
than following the evolution of an individual (pure) micro-state, statistical mechanics ex-
amines an ensemble of micro-states corresponding to a given (mixed) macro-state. It aims
to provide the probabilities $p_M(\mu)$, for the equilibrium ensemble. Liouville's theorem jus-
tifies the assumption that all accessible micro-states are equally likely in an equilibrium
ensemble. As discussed in chapter III, such assignment of probabilities is subjective. In
this chapter we shall provide unbiased estimates of $p_M(\mu)$ for a number of different equi-
librium ensembles. A central conclusion is that in the thermodynamic limit of large $N$ all
these ensembles are in fact equivalent. In contrast to kinetic theory, equilibrium statistical
mechanics leaves out the question of how various systems achieve equilibrium.

## IV.B  The Microcanonical Ensemble

Our starting point in thermodynamics is a mechanically and adiabatically isolated
system. In the absence of heat or work input to the system, the internal energy $E$, and
the generalized coordinates $\mathbf{x}$, are fixed, specifying a macro-state $M \equiv (E, \mathbf{x})$. The corre-
sponding set of mixed micro-states form the *microcanonical ensemble.* In classical statisti-
cal mechanics, these microstates are defined by points in phase space, their time evolution
governed by a Hamiltonian $\mathcal{H}(\mu)$, as discussed in Chapter II. Since the Hamiltonian equa-
tions (II.1) conserve the total energy of a given system, all micro-states are confined to the
surface $\mathcal{H}(\mu) = E$ in phase space. Assume that there are no other conserved quantities, so

that all points on this surface are mutually accessible. The central postulate of statistical mechanics is that the equilibrium probability distribution is given by

$$p_{(E,\mathbf{x})}(\mu) = \frac{1}{\Omega(E, \mathbf{x})} \cdot \begin{cases} 1 & \text{for } \mathcal{H}(\mu) = E \\ 0 & \text{otherwise} \end{cases}. \tag{IV.1}$$

Some remarks and clarification on the above postulate are in order:

(1) *Boltzmann's assumption of equal a priori equilibrium probabilities* refers to the above postulate, which is in fact the unbiased probability estimate in phase space subject to the constraint of constant energy. This assignment is consistent with, but not required by, Liouville's theorem. Note that the phase space, specifying the microstates $\mu$, must be composed of *canonically conjugate pairs*. Under a *canonical* change of variables, $\mu \to \mu'$, volumes in phase space are left invariant. The Jacobian of such transformations is unity, and the transformed probability, $p(\mu') = p(\mu) |\partial\mu/\partial\mu'|$, is again uniform on the surface of constant energy.

(2) *The normalization factor* $\Omega(E, \mathbf{x})$, is the area of the surface of constant energy $E$ in phase space. To avoid subtleties associated with densities that are non-zero only a surface, is sometimes more convenient to define the microcanonical ensemble by requiring $E - \Delta \leq \mathcal{H}(\mu) \leq E + \Delta$, i.e. assigning the energy of the ensemble up to an uncertainty of $\Delta$. In this case, the accessible phase space forms a shell of thickness $\Delta$ around the surface of energy $E$. The normalization is now the volume of the shell, $\Omega' \approx 2\Delta\Omega$. Since $\Omega$ typically depends exponentially on $E$, as long as $\Delta \sim \mathcal{O}(E^0)$ (or even $\mathcal{O}(E^1)$), the difference between the surface and volume of the shell is negligible in the $E \propto N \to \infty$ limit, and we shall use $\Omega$ and $\Omega'$ interchangeably.

(3) *The entropy* of this uniform probability distribution is given by

$$S(E, \mathbf{x}) = k_B \ln \Omega(E, \mathbf{x}). \tag{IV.2}$$

An additional factor of $k_B$ is introduced compared to the definition of eq.(III.68), so that the entropy has the correct dimensions of energy per degrees Kelvin, used in thermodynamics. $\Omega$ and $S$ are not changed by a canonical change of coordinates in phase space. For a collection of independent systems, the overall allowed phase space is the product of individual ones, i.e. $\Omega_{\text{Total}} = \prod_i \Omega_i$. The resulting entropy is thus additive, as expected for an extensive quantity.

Various results in thermodynamics now follow from eq.(IV.1), provided that we consider macroscopic systems with many degrees of freedom.

- *The Zeroth law:* Equilibrium properties are discussed in thermodynamics by placing two previously isolated systems in contact, and allowing them to exchange heat. We can similarly bring together two microcanonical systems, and allowing them to exchange energy, but not work. If the original systems have energies $E_1$ and $E_2$ respectively, the combined system has energy $E = E_1 + E_2$. Assuming that interactions between the two parts are small, each micro-state of the joint system corresponds to a pair of micro-states of the two components, i.e. $\mu = \mu_1 \otimes \mu_2$, and $\mathcal{H}(\mu_1 \otimes \mu_2) = \mathcal{H}_1(\mu_1) + \mathcal{H}_2(\mu_2)$. As the joint system is in a microcanonical ensemble of energy $E = E_1 + E_2$, in equilibrium

$$p_E(\mu_1 \otimes \mu_2) = \frac{1}{\Omega(E)} \cdot \begin{cases} 1 & \text{for } \mathcal{H}_1(\mu_1) + \mathcal{H}_2(\mu_2) = E \\ 0 & \text{otherwise} \end{cases}. \qquad \text{(IV.3)}$$

Since only the overall energy is fixed, the total allowed phase space is computed from

$$\Omega(E) = \int dE_1 \Omega_1(E_1) \Omega_2(E - E_1) = \int dE_1 \exp \left[ \frac{S_1(E_1) + S_2(E - E_1)}{k_B} \right]. \qquad \text{(IV.4)}$$

The properties of the two systems in the new joint equilibrium state are implicit in eq.(IV.3). We can make them explicit by examining the entropy that follows from eq.(IV.4). Extensivity of entropies suggests that $S_1$ and $S_2$, are proportional to the number of particles in the systems, making the integrand in eq.(IV.4) an exponentially large quantity. Hence the integral can be equated by the saddle point method to the maximum value of the integrand, obtained for energies $E_1^*$ and $E_2^* = E - E_1^*$, i.e.

$$S(E) = k_B \ln \Omega(E) \approx S_1(E_1^*) + S_2(E_2^*). \qquad \text{(IV.5)}$$

The position of the maximum is obtained by extremizing the exponent in eq.(IV.4) with respect to $E_1$, resulting in the condition,

$$\left. \frac{\partial S_1}{\partial E_1} \right|_{\mathbf{x_1}} = \left. \frac{\partial S_2}{\partial E_2} \right|_{\mathbf{x_2}}. \qquad \text{(IV.6)}$$

Although all joint micro-states are equally likely, the above results indicate that there are an exponentially larger number of states in the vicinity of $(E_1^*, E_2^*)$. Originally, the joint system starts in the vicinity of the point $(E_1, E_2)$. After the exchange of energy takes place, the combined system explores a whole set of new micro-states. The probabilistic arguments provide no information on the dynamics of evolution amongst these micro-states, or on the amount of time needed to establish equilibrium. However, once sufficient

time has elapsed so that the assumption of equal a priori probabilities is again valid, the system is overwhelmingly likely to be at a state with internal energies $(E_1^*, E_2^*)$. At this *equilibrium* point, condition (IV.6) is satisfied, specifying a relation between two functions of state. These state functions are thus equivalent to empirical temperatures, and indeed, consistent with the fundamental result of thermodynamics, we have

$$\left. \frac{\partial S}{\partial E} \right|_{\mathbf{x}} = \frac{1}{T}. \tag{IV.7}$$

- *The first law:* We next inquire about the variations of $S(E, \mathbf{x})$ with $\mathbf{x}$, by changing the coordinates by $\delta \mathbf{x}$. This results in doing work on the system by an amount $dW = \mathbf{J} \cdot \delta \mathbf{x}$, and changes the internal energy to $E + \mathbf{J} \cdot \delta \mathbf{x}$. The first order change in entropy is given by

$$\delta S = S(E + \mathbf{J} \cdot \delta \mathbf{x}, \mathbf{x} + \delta \mathbf{x}) - S(E, \mathbf{x}) = \left( \left. \frac{\partial S}{\partial E} \right|_{\mathbf{x}} \mathbf{J} + \left. \frac{\partial S}{\partial \mathbf{x}} \right|_E \right) \cdot \delta \mathbf{x} \quad . \tag{IV.8}$$

This change will occur spontaneously, taking the system into a more probable state, unless the quantity in brackets in zero. Using eq.(IV.7), this allows us to identify the derivatives

$$\left. \frac{\partial S}{\partial x_i} \right|_{E, x_{j \neq i}} = -\frac{J_i}{T}. \tag{IV.9}$$

Having thus identified all variations of $S$, we have

$$dS(E, \mathbf{x}) = \frac{dE}{T} - \frac{\mathbf{J} \cdot d\mathbf{x}}{T}, \quad \implies \quad dE = TdS + \mathbf{J} \cdot d\mathbf{x}, \tag{IV.10}$$

allowing us to identify the heat input $dQ = TdS$.

- *The second law:* Clearly, the above statistical definition of equilibrium rests on the presence of many degrees of freedom $N \gg 1$, which make it exponentially unlikely in $N$, that the combined systems is found with component energies different from $(E_1^*, E_2^*)$. By this construction, the equilibrium point has a larger number of accessible states than the starting point, i.e.

$$\Omega_1(E_1^*, \mathbf{x}_1)\Omega_2(E_2^*, \mathbf{x}_2) \geq \Omega_1(E_1, \mathbf{x}_1)\Omega_2(E_2, \mathbf{x}_2). \tag{IV.11}$$

In the process of evolving to the more likely (and more densely populated) regions, there is an irreversible loss of information, accompanied by an increase in entropy,

$$\delta S = S_1(E_1^*) + S_2(E_2^*) - S_1(E_1) - S_2(E_2) \geq 0, \tag{IV.12}$$

as required by the second law of thermodynamics. When the two bodies are first brought into contact, the equality in eq.(IV.6) does not hold. The change in entropy is such that

$$\delta S = \left( \left.\frac{\partial S_1}{\partial E_1}\right|_{\mathbf{x_1}} - \left.\frac{\partial S_2}{\partial E_2}\right|_{\mathbf{x_2}} \right) \delta E_1 = \left( \frac{1}{T_1} - \frac{1}{T_2} \right) \delta E_1 \geq 0, \tag{IV.13}$$

i.e. heat (energy) flows from the hotter to the colder body, as in Clausius's statement of the second law.

• *Stability conditions:* Since the point $(E_1^*, E_2^*)$ is a maximum, the second derivative of $S_1(E_1) + S_2(E_2)$ must be negative at this point, i.e.

$$\left.\frac{\partial^2 S_1}{\partial E_1^2}\right|_{\mathbf{x_1}} + \left.\frac{\partial^2 S_2}{\partial E_2^2}\right|_{\mathbf{x_2}} \leq 0. \tag{IV.14}$$

Applying the above condition to two parts of the same system, the condition of thermal stability, $C_{\mathbf{x}} \geq 0$, as discussed in section (I.I), is regained. Similarly, the second order changes in eq.(IV.8) must be negative, requiring that the matrtix $\partial^2 S/\partial x_i \partial x_j \big|_E$ be positive definite.

## IV.C  Two–Level Systems

Consider $N$ impurity atoms trapped in a solid matrix. Each impurity can be in one of two states, with energies 0 and $\epsilon$ respectively. This example is somewhat different from the situations considered so far, in that the allowed micro-states are discrete. Liouville's theorem applies to Hamiltonian evolution in a continuous phase space. Although, there is less ambiguity in enumeration of discrete states, the dynamics that ensures that all allowed micro-states are equally accessed will remain unspecified for the moment. (An example from quantum mechanical evolution will be presented later on.)

The micro-states of the two level system are specified by the set of *occupation numbers* $\{n_i\}$, where $n_i = 0$ or 1 depending on whether the $i^{\text{th}}$ impurity is in its ground state or excited. The overall energy is

$$\mathcal{H}\left(\{n_i\}\right) = \epsilon \sum_{i=1}^{N} n_i \equiv \epsilon N_1, \tag{IV.15}$$

where $N_1$ is the total number of excited impurities. The macro-state is specified by the total energy $E$, and the number of impurities $N$. The microcanonical probability is thus

$$p\left(\{n_i\}\right) = \frac{1}{\Omega(E, N)} \delta_{\epsilon \sum_i n_i, E} \quad . \tag{IV.16}$$

As there are $N_1 = E/\epsilon$ excited impurities, the normalization $\Omega$ is the number of ways of choosing $N_1$ excited levels among the available $N$, and given by the binomial coefficient

$$\Omega(E, N) = \frac{N!}{N_1! \, (N - N_1)!} \, . \tag{IV.17}$$

The entropy

$$S(E, N) = k_B \ln \frac{N!}{N_1! \, (N - N_1)!} \, , \tag{IV.18}$$

can be simplified by Stirling's formula in the limit of $N_1, N \gg 1$ to

$$\begin{aligned} S(E, N) &\approx -N k_B \left[ \frac{N_1}{N} \ln \frac{N_1}{N} + \frac{N - N_1}{N} \ln \frac{N - N_1}{N} \right] \\ &= -N k_B \left[ \left( \frac{E}{N\epsilon} \right) \ln \left( \frac{E}{N\epsilon} \right) + \left( 1 - \frac{E}{N\epsilon} \right) \ln \left( 1 - \frac{E}{N\epsilon} \right) \right] . \end{aligned} \tag{IV.19}$$

The equilibrium temperature can now be calculated from eq.(IV.7) as

$$\frac{1}{T} = \left. \frac{\partial S}{\partial E} \right|_N = -\frac{k_B}{\epsilon} \ln \left( \frac{E}{N\epsilon - E} \right) . \tag{IV.20}$$

Alternatively, the internal energy at a temperature $T$, is given by

$$E(T) = \frac{N\epsilon}{\exp \left( \frac{\epsilon}{k_B T} \right) + 1} \, . \tag{IV.21}$$

The internal energy is a monotonic function of temperature, increasing from a minimum value of 0 at $T = 0$ to a maximum value of $N\epsilon/2$ at infinite temperature. It is, however, possible to start with energies larger than $N\epsilon/2$, which correspond to *negative temperatures* from eq.(IV.20). The origin of the negative temperature is the *decrease* in the number of microstates with increasing energy, the opposite of what happens in most systems. Two level systems have an upper bound on their energy, and very few microstates close to this maximal energy. Hence increased energy leads to more order in the system. However, once a negative temperature system is brought into contact with the rest of the universe (or any portion of it without an upper bound in energy), it loses its excess energy and comes to equilibrium at a positive temperature. The world of negative temperatures is quite unusual in that systems can be cooled by adding heat, and heated by removing it. There are physical examples of systems temporarily prepared at a metastable equilibrium of negative temperature in lasers, and for magnetic spins.

The heat capacity of the system, given by

$$C = \frac{dE}{dT} = N k_B \left(\frac{\epsilon}{k_B T}\right)^2 \exp\left(\frac{\epsilon}{k_B T}\right) \left(\exp\left(\frac{\epsilon}{k_B T}\right) + 1\right)^{-2}, \quad \text{(IV.22)}$$

vanishes at both low and high temperatures. The vanishing of $C$ as $\exp\left(-\epsilon/k_B T\right)$ at low temperatures is characteristic of all systems with an *energy gap* separating the ground state and lowest excited states. The vanishing of $C$ at high temperatures is a *saturation* effect, common to systems with a maximum in the number of states as a function of energy. In between, the heat capacity exhibits a peak at a characteristic temperature of $T_\epsilon \propto \epsilon/k_B$.

Statistical mechanics provides much more than just macroscopic quantities such as energy and heat capacity. Eq.(IV.16) is a complete joint probability distribution with considerable information on the micro-states. For example, the unconditional probability for exciting a particular impurity is obtained from

$$p(n_1) = \sum_{\{n_2, \cdots, n_N\}} p\left(\{n_i\}\right) = \frac{\Omega(E - n_1 \epsilon, N - 1)}{\Omega(E, N)}. \quad \text{(IV.23)}$$

The second equality is obtained by noting that once the energy taken by the first impurity is specified, the remaining energy must be distributed among the other $N - 1$ impurities. Using eq.(IV.17),

$$p(n_1 = 0) = \frac{\Omega(E, N - 1)}{\Omega(E, N)} = \frac{(N - 1)!}{N_1! \, (N - N_1 - 1)!} \cdot \frac{N_1! \, (N - N_1)!}{N!} = 1 - \frac{N_1}{N}, \quad \text{(IV.24)}$$

and $p(n_1 = 1) = 1 - p(n_1 = 0) = N_1/N$. Using $N_1 = E/\epsilon$, and eq.(IV.21), the occupation probabilities at a temperature $T$ are

$$p(0) = \frac{1}{1 + \exp\left(-\frac{\epsilon}{k_B T}\right)}, \quad \text{and} \quad p(1) = \frac{\exp\left(-\frac{\epsilon}{k_B T}\right)}{1 + \exp\left(-\frac{\epsilon}{k_B T}\right)}. \quad \text{(IV.25)}$$