STATISTICAL MECHANICS

1. Introduction

The study of statistical mechanics (in second year) concerns using energy levels of a single molecule to predict the bulk thermodynamic properties of a sample of large numbers of the same molecule. In classical thermodynamics, properties such as pressure, heat capacity, equilibrium constants etc, have to be measured. With statistical mechanics, these properties can be derived from an understanding of the underlying microscopic properties, in particular, the energy levels of the molecule. It is 'statistical' since probability theory is used to average a large number of microstates to give an averaged macroscopic property. The central idea of this study is the **partition function** of a system.

The energy levels of a molecule can usually be obtained using either quantum mechanics (from theory, with approximations) or spectroscopy (from experiments). Three statistical ensembles can be distinguished, each suited to the modelling of an appropriate system:

(1) A *microcanonical* or *NVE ensemble* has a fixed number of molecules, N, fixed volume V and fixed total energy of the system E. This is used to model an *isolated* system, one that does not allow the exchange of either matter or energy with the surrounding.

(2) A *canonical* or *NVT ensemble* has a fixed number of molecules, N, a fixed volume V and a fixed temperature T. The energy of the system is not known exactly. This is used to model a *closed* system, one that allows the exchange of energy, but not matter, with the surrounding.

(3) A *grand canonical* or *μVT ensemble* has a fixed chemical potential mu, a fixed volume V and a fixed temperature T. Neither the energy nor the number of particles is known exactly. This is used to model an *open* system, which allows the exchange of both energy and matter, with the surrounding.

We are primarily interested in the first two ensembles at this stage.

2. The Boltzmann distribution

Consider a fixed NVE ensemble of a sample of identical molecules, we have

$$\text{total energy} = E = N\bar{E} = N\sum_i \epsilon_i P_i \tag{1}$$

$$\text{total number of possible microstates} = \sum_n W(n) \tag{2}$$

where $\bar{E}$ is the average energy per molecule, P_i is the probability of any selected molecule in state i; $W(n)$ is the number of microstates in configuration n. For each configuration $n \equiv (n_1, n_2, ..., n_i, ...)$, where n_i is the number of molecules found in state i, assuming the molecules can be *localised*, the total number of possible microstates is given by

$$W(n) = \frac{N!}{\Pi_i n_i!} \tag{3}$$

where the division by each $n_i!$ accounts for the fact that n_i molecules in state i are identical.

The probability that any molecule is in state i is the probability that the molecule is found in state i for a given configuration n, $p_i(n)$, weighted by the probability that the microstate is found in that particular configuration, $p(n)$, summed over all possible configurations, viz,

$$P_i = \sum_n p_i(n)p_(n) = \sum_n p_i(n)\frac{W(n)}{\sum_n W(n)} \tag{4}$$

This probability is also given by the ratio of the number of particles in quantum state i divided by the total number of particles:

$$P_i = \frac{n_i}{N}. \tag{5}$$

The **ergodic hypothesis** states that, over a long period of time, all accessible microstates are equally probable. In other words, each microstate in each configuration is equally likely. However, we will postulate that there is *one* particular configuration having the largest number of microstates (given by Eq. (3)) such that the total number of microstates in all other configurations are negligible. We can visualise this in Figure 1.
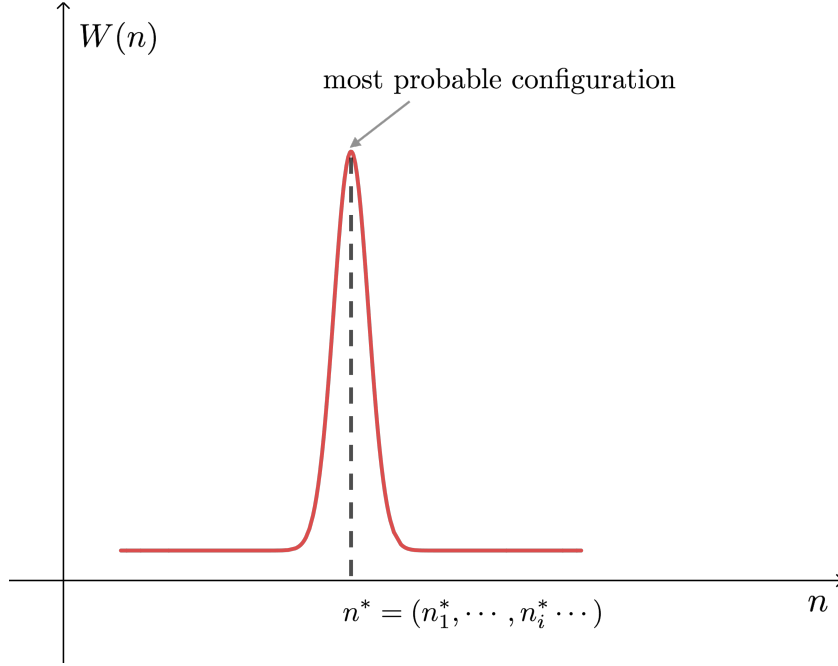
FIGURE 1. Configuration n^* has the highest number of accessible microstates out of all possible configurations. This configuration is called the most probable distribution and is considered to dominate over any other configurations.

With this assumption, we have

$$p(n) = \begin{cases} = 1 & \text{if } n = n^* \\ = 0 & \text{if } n \neq n^* \end{cases}$$

such that we have replaced the total number of microstates arising from all possible configurations by the total number of microstates arising from the most probable configuration:

$$\sum_n W(n) = W(n^*).$$

We now proceed to find this configuration $n^* = (n_1^*, \cdots, n_i^*, \cdots)$ for which the total number of microstates, $W(n^*)$, is a maximum, subject to the constraints that the total energy and total number of particles of the system (Eq.s (1) and (2)) remain the same. To maximise W, it is equally valid to maximise $\ln W$, since ln function is *monotonically increasing*. Applying the method of Lagrange

multipliers (see lectures), we arrive at the **Boltzmann distribution**:

$$P_i = \frac{n_i^*}{N} = \frac{e^{-\beta\epsilon_i}}{\sum_i e^{-\beta\epsilon_i}}$$

(6)

where $\beta = 1/kT$. The *molecular* partition function, q, is defined by

$$q = \sum_i e^{-\epsilon_i/kT} = \sum_j g_j e^{-\epsilon_j/kT}$$

(7)

where ϵ_i is the energy of the ith state and the summation index i goes over all states; alternatively, we can sum over all levels j, each having energy ϵ_j and degeneracy g_j. We see that the molecular partition function arises as a normalisation constant in the probability distribution function, from which all thermodynamic properties can be deduced. Note that this derivation of Maxwell-Boltzmann statistics is also valid for *indistinguishable* systems, provided that the temperature is large enough such that the number of accessible quantum states (measured by the partition function) are much larger than the total number of particles in the system.

For the **canonical ensemble**, where NVT are fixed, we can follow the same procedure as before, replacing the idea of 'particle' in a microcanonical ensemble, by the idea of a 'system' in a canonical ensemble, to arrive at

$$\mathbb{P}_i = \frac{\mathbb{n}_i^*}{\mathbb{N}} = \frac{e^{-\beta E_i}}{\sum_i e^{-\beta E_i}}$$

(8)

where now i refers to a selected system from the ensemble (instead of a selected molecule from an isolated system in the microcanonical ensemble), and E_i is the energy of the ith system in the ensemble. Eq. (8) is the *canonical distribution* and we can similarly define the canonical partition function

$$Q = \sum_i e^{-E_i/kT} = \sum_j g_j e^{-E_j/kT}$$

(9)

as before. The canonical ensemble is particularly important as many experimental set-ups have fixed volume (in a closed container) and fixed temperature (in a heat bath) and can be modelled using this ensemble.

4

3. Thermodynamic quantities from partition function

We consider the canonical ensemble here on. The *average* energy per system of ensemble is the same as the internal energy of the system and is given by

$$E = \frac{\mathbb{E}}{\mathbb{N}} = \sum_i E_i \mathbb{P}_i \tag{10}$$

where $\mathbb{E}$ is the total energy of the ensemble of all systems, $\mathbb{N}$ is the number of systems in the ensemble; the sum i runs over all states of the systems.

Aside

In statistics, the average/mean value of a random variable, X, is given by the sum of all the measured values, x_i, weighted by the probability distribution function, $p(X = x_i)$. For discrete variable, this is

$$\bar{x} = \langle x \rangle = \sum_i x_i \, p(X = x_i), \tag{11}$$

and for continuous variable, this is

$$\bar{x} = \langle x \rangle = \int x \, p(X = x) dx. \tag{12}$$

Here in statistical mechanics, each quantum state has a discrete energy level, thus the probability distribution function is discrete.

Now from probability theory, and substituting Eq. (8) into Eq. (10), we have

$$E = \sum_i E_i \mathbb{P}_i = \sum_i E_i \frac{e^{-\beta E_i}}{\sum_j e^{-\beta E_j}} = \frac{1}{Q} \sum_i E_i e^{-\beta E_i}$$

Notice that if we differentiate $e^{-\beta E_i}$ with respect to β (V is fixed as this is an NVT-ensemble), this gives us a copy of E_i, that is

$$\left(-\frac{\partial}{\partial \beta} e^{-\beta E_i} \right)_V = E_i e^{-\beta E_i}$$

so that

$$E = \frac{1}{Q} \sum_i \left(-\frac{\partial}{\partial \beta} e^{-\beta E_i} \right)_V = -\frac{1}{Q} \left(\frac{\partial}{\partial \beta} \right)_V \sum_i e^{-\beta E_i} = -\frac{1}{Q} \left(\frac{\partial Q}{\partial \beta} \right)_V = -\left(\frac{\partial \ln Q}{\partial \beta} \right)_V,$$

using chain rule, we have

$$E = -\left(\frac{\partial \ln Q}{\partial \beta} \right)_V = -\left(\frac{\partial T}{\partial \beta} \right)_V \left(\frac{\partial \ln Q}{\partial T} \right)_V = kT^2 \left(\frac{\partial \ln Q}{\partial T} \right)_V$$

where, in the last step, we have used $\beta = 1/kT$ in the following:

$$\frac{\partial T}{\partial \beta} = \frac{1}{\partial \beta / \partial T} = \frac{1}{-1/kT^2} = -kT^2$$

Thus, we have the energy expression completely specified in terms of the partition function:

$$E = kT^2 \left(\frac{\partial \ln Q}{\partial T} \right)_V$$

$$(13)$$

Knowing the partition function of a given system allows us to determine its internal energy via Eq. (13) above. This then allows us to find the heat capacity, again at fixed volume, C_V, via

$$C_V = \left(\frac{\partial E}{\partial T} \right)_V \tag{14}$$

The easiest way forward is to introduce the 'bridge equation' for the canonical ensemble:

$$A = -kT \ln Q$$

$$(15)$$

Note that you can of course follow the lectures and derive entropy S first and obtain this same expression via thermodynamic relation

$$A = E - TS, \tag{16}$$

but this 'bridge equation' is the easiest to proceed here. It is a 'bridge' between the macroscopic thermodynamic quantity A and the partition function, Q, arising from the microscopic microstates.

By rearranging Eq. (16) and substituting the bridge equation, we arrive at the expression of entropy S in terms of the partition function:

$$S = \frac{E}{T} - \frac{A}{T} = \frac{E}{T} + k \ln Q = kT \left(\frac{\partial \ln Q}{\partial T} \right)_V + k \ln Q \tag{17}$$

Here again, the entropy of the system can now be found by knowing the partition function of the system.

All other thermodynamic properties can now be found via the *fundamental thermodynamic relations*. For example, we have

$$dA = -PdV - SdT + \sum_i \mu_i dn_i$$

$$(18)$$

From this, the pressure of the system, P, is given by

$$P = -\left(\frac{\partial A}{\partial V}\right)_{T,n_i's} = -\left(\frac{\partial(-kT\ln Q)}{\partial V}\right)_{T,n_i} = kT\left(\frac{\partial \ln Q}{\partial V}\right)_{T,n_i} \tag{19}$$

again, you can now find thermodynamic pressure by simply knowing the partition function (a microscopic property) of a system.

Similarly, the entropy, S, is given by

$$S = -\left(\frac{\partial A}{\partial T}\right)_{V,n_i} = -\left(\frac{\partial(-kT\ln Q)}{\partial T}\right)_{V,n_i} = k\left(\frac{\partial(T\ln Q)}{\partial T}\right)_{V,n_i} \tag{20}$$

which gives you the same result as in Eq. (17) when you do product rule differentiation with respect to T in the last step above.

The chemical potential is similarly given by

$$\mu_i = \left(\frac{\partial A}{\partial n_i}\right)_{T,V,n_j \neq n_i} = \left(\frac{\partial(-kT\ln Q)}{\partial n_i}\right)_{T,V,n_j'} = -kT\left(\frac{\partial \ln Q}{\partial n_i}\right)_{T,V,n_j'} \tag{21}$$

All other thermodynamic energies can now be found:

$$H = E + PV = kT^2\left(\frac{\partial \ln Q}{\partial T}\right)_V + VkT\left(\frac{\partial \ln Q}{\partial V}\right)_{T,n_i's} \tag{22}$$

$$G = H - TS = A + PV = -kT\ln Q + VkT\left(\frac{\partial \ln Q}{\partial V}\right)_{T,n_i's} \tag{23}$$

The only thing now remains is to find the system-specific canonical partition function Q for a given system and substituting this into appropriate thermodynamic quantities to obtain those values. Two kinds of independent systems, each with either distinguishable molecules (as in solids) or indistinguishable molecules (as in gases), are commonly considered. The canonical partition function is given, since the particles are identical, by

$$Q = \begin{cases} q^N/N! & \text{indistinguishable molecules} \\ q^N & \text{distinguishable molecules} \end{cases} \tag{24}$$

Using these expressions, we are able to find all thermodynamic quantities in terms of the molecular partition function q.

Everything remains now is to find the molecular partition function q for a given system. The molecular partition function can be decomposed into individual components:

$$q = q_{trans}\, q_{rot}\, q_{vib}\, q_{elec} \tag{25}$$

since the total energy of a molecule is approximately $\epsilon = \epsilon_{trans} + \epsilon_{rot} + \epsilon_{vib} + \epsilon_{elec}$ and q is given by Eq. (7). To find the partition function due to each contribution,

we now need to find the energy expressions of each contribution, using quantum mechanics and spectroscopy.

4. Energy Levels of each contribution

4.1. **Translation.** The quantum mechanical model for translation is the **particle-in-a-box** model. The Schrödinger equation for this system, in 1D, is

$$\hat{H}\psi = \left(\frac{\hat{p}^2}{2m} + \hat{V}\right)\psi = -\frac{\hbar^2}{2m}\frac{d}{dx}\psi + \hat{V}\psi = E\psi \tag{26}$$

where the potential V is given by

$$V = \begin{cases} 0 & \text{if } 0 < x < L \\ \infty & \text{otherwise} \end{cases}$$

where L is the box length. Within the box, we can easily solve the following second-order linear ordinary differential equation

$$-\frac{\hbar^2}{2m}\frac{d}{dx}\psi = E\psi.$$

Applying boundary conditions that the wavefunction vanishes at both ends of the box, we arrive at energy quantization with the following energy expression:

$$\boxed{E_n = \frac{n^2 h^2}{8mL^2}} \tag{27}$$

The one-dimensional translational partition function is then given by

$$q_{tr,1D} = \sum_{n=1}^{\infty} e^{-\beta E_n}$$

We *assume* that the number of accessible translational energy levels is very large at temperature T (since thermal energy kT is much greater than the translational energy level spacings), we are going to approximate the sum above by an integral. We then have

$$q_{trans,1D} = \sum_{n=1}^{\infty} e^{-\beta E_n} \simeq \int_0^{\infty} e^{-\beta E_n} dn = \int_0^{\infty} \exp\left(-\frac{n^2 h^2}{8mL^2 kT}\right) dn = \left(\frac{2\pi mkT}{h^2}\right)^{\frac{1}{2}} L$$

where approximating the lower $n = 1$ in the sum by $n = 0$ in the integral is not going to make the approximation any worse (only one more level is added to many), and in the last step, a standard Gaussian integration was carried out.

8

By extension, the 3D translational partition function is given by

$$q_{trans,3D} = \left(\frac{2\pi mkT}{h^2}\right)^{\frac{3}{2}} L^3 = \left(\frac{2\pi mkT}{h^2}\right)^{\frac{3}{2}} V$$

Defining **de Broglie thermal wavelength** (units of length of course!) as

$$\Lambda = \frac{h}{\sqrt{2\pi mkT}},$$

(28)

we can write the translational molecular partition function, in 3D, as

$$q_{trans} = \frac{V}{\Lambda^3}$$

(29)

Note that by approximating the sum by an integral, we are ignoring any contributions from quantum effects. As a result, if you use the translational partition function Eq. (29) above and substituting it into the energy expression $E = kT^2(\partial \ln Q/\partial T)_V$, you would obtain the energy contribution from classical equipartition theorem, i.e., you would find $E = 3/2kT$.

4.2. **Rotation.** The quantum mechanical model for rotation is the **rigid rotor** model. In 3D, the full Schrödinger equation for this system is

$$\hat{H}\psi = \left(\frac{\hat{J}^2}{2I} + \hat{V}\right)\psi = \frac{\hat{J}^2}{2I}\psi = E\psi$$

(30)

where the potential V is zero, I is the moment of inertia and $I = \mu r^2$ for a diatomic and $I = \sum_i m_i r_i^2$ for a polyatomic. We consider only diatomic and linear polyatomics here. $\hat{J}^2$ is angular momentum operator and is given, in spherical polar coordinates, by

$$\hat{J}^2 = -\hbar^2 \left(\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\sin\theta\frac{\partial}{\partial\theta} + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right)$$

Note that the eigenfunctions of the rigid rotor are also the *spherical harmonics.*

We can solve this this differential equation (see here for the maths if interested) to get the energy for the rigid rotor, but it is enough to know that

$$\hat{J}^2 Y_{JM_J} = \hbar^2 J(J+1)Y_{JM_J}$$

(31)

so that the energy E in the Schrödinger equation Eq. (30) is given by

$$E_J = \frac{\hbar^2}{2I}J(J+1) = BJ(J+1)$$

(32)

where $B = \hbar^2/(2I)$ is the **rotational constant**. The *degeneracy* of each rotational energy level J is given by $g_J = 2J + 1$ due to the number of possible M_J values, therefore, the rotational molecular partition function is given by

$$q_{rot} = \sum_{J=0}^{\infty}(2J+1)e^{-\beta E_J} = \sum_{J=0}^{\infty}(2J+1)e^{-\beta BJ(J+1)}$$

(33)

If at temperature T, the thermal energy kT is much greater than the rotational energy level spacings (some multiple of the rotational constant), we can *assume* that many rotational levels are assessible, in which case we can approximate the sum above by an integral. We then have

$$q_{rot} = \sum_{J=0}^{\infty}(2J+1)e^{-\beta BJ(J+1)} \simeq \int_0^{\infty}(2J+1)\exp\left(-\frac{BJ(J+1)}{kT}\right)dJ = \frac{kT}{B}$$

We can now define a **rotational temperature** (units of temperature of course!):

$$\theta_{rot} = \frac{B}{k}$$

(34)

so that the rotational partition function can now be written as

$$q_{rot} = \frac{T}{\sigma\theta_{rot}}$$

(35)

where σ is a symmetry parameter that arises from nuclear spin statistics (see later).

Again, by approximating the sum as an integral, we obtain the classical results. It is important to note that this approximation is valid only when the number of accessible states is much greater than 1. If the temperature T is very low, it is possible that not many rotational levels are thermally accessible, so that the rotational partition function Eq. (33) must be evaluated term by term.

10

> **Aside**
>
> *For polyatomics, the rotational partition function using integral approximation is more complex, and is given by*
>
> $$q_{rot} = \frac{\pi^{\frac{1}{2}}}{\sigma} \left(\frac{T^3}{\theta_{rot,a}\,\theta_{rot,b}\,\theta_{rot,c}} \right)$$
>
> *where each rotational temperature is defined via its associated rotational rotational constants and moments of inertia:*
>
> $$\theta_{rot,a} = \frac{\tilde{A}}{k} \quad where\ \tilde{A} = \frac{h}{8\pi^2\tilde{c}I_a}$$
>
> *and so on.*

4.3. **Vibration.** The quantum mechanical model for a vibrating bond is the **simple harmonic oscillator** (SHO). The energies are given by

$$E_v = \left(v + \frac{1}{2} \right) \hbar\omega = \left(v + \frac{1}{2} \right) h\nu \tag{36}$$

where v is the vibrational quantum number and ν is the vibrational frequency, ω is the angular frequency.

The vibrational partition function is given by

$$q_{vib} = \sum_{v=0}^{\infty} e^{-\beta(v+1/2)h\nu} = \frac{e^{-1/2\beta h\nu}}{1 - e^{-\beta h\nu}} = \frac{e^{-h\nu/(2kT)}}{1 - e^{-h\nu/(kT)}} = e^{-E_0/(kT)}\, q'_{vib} \tag{37}$$

where $E_0 = \frac{1}{2}h\nu$ is the *zero-point energy* and $q'_{vib} = (1 - e^{-\beta h\nu})^{-1}$ is the vibrational partition function taking the vibrational ground state as energy zero. Note that this expression is *exact* (we used the sum of an infinite geometric series), within the SHO model.

We can define a vibrational temperature as

$$\boxed{\theta_{vib} = \frac{h\nu}{k}} \tag{38}$$

so that the partition function can be written as

$$\boxed{q_{vib} = \frac{e^{-\theta_{vib}/(2T)}}{1 - e^{-\theta_{vib}/T}}} \tag{39}$$

4.4. **Electronic.** Except for the simple case of hydrogen atoms, there is no quantum mechanical model for the energies of electronic levels; these have to be determined experimentally, usually by spectroscopy. The electronic partition function is simply given by

$$q_{elec} = \sum_i g_i \, e^{-\beta \epsilon_{elec,i}} \tag{40}$$

Commonly, the electronic excited states lie much higher in energy than the ground state such that they are thermally inaccessible and does not contribute to the electronic partition function. This occurs when the energy spacing between the electronic states is much larger than kT. In those cases, only the ground state is accessible, resulting in $q_{elec} = g_0 \, e^{-\beta \epsilon_{elec,0}}$, where g_0 is the degeneracy of the ground state. If we take the electronic ground state as energy zero, then, we have $q'_{elec} = g_0$.