Notes for Lecture 6

Semi-classical statistical mechanics

In the last lecture, we gave some general considerations to many particle problems. We studied pretty important things – the Poisson bracket, the Poincaré recurrence, the Liouville's theorem, and some basic issues regarding the entropy and the second law of thermodynamics. Especially, we noted that the microscopic explanation of the second law of thermodynamics is an active subject of research, still.

Now, we are going into the most established part of the statistical mechanics, namely the *equilibrium* statistical mechanics, leaving out topics pertaining to non-equilibrium to later in the course.

In **equilibrium statistical mechanics**, we do not, by design, worry too much about how the equilibrium state is brought about. In particular, in this lecture, we will be concerned with *semi-classical* statistical mechanics. Right from the beginning, let us be reminded that the so-called classical statistical mechanics does not really exist. Even when we apply mostly classical mechanics, two quantum aspects – the identical particle nature of gas molecules and the quantum of the phase space volume – must be considered, as emphasized in the previous lecture, making what we do *semi-classical statistical mechanics* rather than classical statistical mechanics. Note, however, that the first aspect – the identical nature of particles – is not always important (even for quantum statistical mechanics) if somehow particles are really distinguishable: a problem involving molecules fixed in space (like in a crystal) would be an example where particles can, and best, be treated as distinguishable.

6.1 Basic assumption of statistical mechanics

From Eq. 5.6, we see that

$$\frac{d\langle O\rangle}{dt} = \int d\Gamma\, \frac{\partial\rho}{\partial t} O \tag{6.1}$$

where $O = O(\vec{\mathbf{p}}, \vec{\mathbf{q}})$ is any dynamical variable, and $\rho = \rho(\vec{\mathbf{p}}, \vec{\mathbf{q}}, t)$ is the PDF defined in the phase space by an ensemble of states, through Eq. 5.5.

At the outset, let us note that some of what we did in the last lecture (e.g., Section 5.5) applies even when the Hamiltonian[1] H is explicitly time-dependent. However, from now on, we rule out such possibility, unless explicitly stated otherwise in any problem: as a rule, H is always time-independent from now on. Also, note that any variable O, which we consider the ensemble average for, is, by definition, not explicitly time-dependent ($\frac{\partial O}{\partial t} = 0$, i.e., $O = O(\vec{\mathbf{p}}, \vec{\mathbf{q}})$ as we have been noting).

Our interest is in $O(\vec{\mathbf{p}}, \vec{\mathbf{q}})$, which is a *macroscopic* dynamical variable, such as a thermodynamic state variable (such as the energy) or a macroscopic current. As we defined the thermal equilibrium state to be a state in which any thermodynamic variable is constant, we see that the following condition

$$\frac{\partial\rho(\vec{\mathbf{p}}, \vec{\mathbf{q}}, t)}{\partial t} = 0 \qquad \text{equibrium condition} \tag{6.2}$$

is a sufficient condition. Indeed, this gives **the basic definition of the equilibrium state in statistical mechanics**.

Let us consider a bit what the above definition means.

First of all, it means that the PDF function ρ is stationary in the phase space. This does *not* mean that each state in the ensemble is unchanging as a function of time. That would be possible only if the dynamics is trivial, i.e., if each particle were truly free. Such a system, if prepared in a non-equilibrium state, can never relax to the equilibrium state, in the first place. What really happens is that each particle in the system is constantly interacting with one another (and the environment, if applicable), all the time. Therefore, each state in the ensemble always changes as a function of time. In an equilibrium state, it is just that those states evolve in time in such a way that the whole distribution given by ρ is constant since those states only swap between one another.

Second, one should not confuse the above condition with the more general Liou-

---

[1] You may have noticed that in the last lecture, H has been used for two things, the Hamiltonian and the Boltzmann H function. From here on, H will mean only the former.

ville's theorem (LN 5), $d\rho/dt = 0$, which is always valid[2]. In equilibrium, we have *both* $\frac{d\rho}{dt} = 0$ and $\frac{\partial\rho}{\partial t} = 0$, and so we must have

$$\{\rho, H\} = 0 \tag{6.3}$$

according to Eq. 5.13. In view of Eq. 5.11, we then might conjecture that $\rho = \rho(H)$. This condition can be generalized to include any dynamical variable $L(\vec{\mathbf{p}}, \vec{\mathbf{q}})$, which is conserved, i.e., $\{L, H\} = 0$: $\rho = \rho(H, L_1 L_2, ...)$, where L_1, L_2, etc. are such conserved quantities. This leads to the following basic assumption.

$$\rho = \rho(H, L_1, L_2, ...) \qquad \text{equilibrium state, fundamental assumption} \tag{6.4}$$

This means that, for a given set of conserved quantities, *all accessible/possible states are equally likely to occur*, or *they have equal a priori probabilities.*

What are the examples of conserved quantities, L_1, L_2, ...? For a classical Hamiltonian, the number of particles is a conserved quantity, and it is a simple trivial example. The total momentum or the total angular momentum would have been another example, except that it is not an appropriate one, since a finite total momentum or a finite total angular momentum would imply a macroscopic current, not allowed for an equilibrium state, by our definition (LN 1). For this reason, for a classical mechanical system, H is the only one that matters[3], and we have

$$\rho = \rho(H) \qquad \text{equilibrium state.} \tag{6.5}$$

6.2 Microcanonical ensemble

For a closed system, the energy (E) and the "volume" (x)[4] are conserved. The collection of states that we consider under this condition are called a *microcanonical ensemble.*

---

[2] However, note the comment on "having to define a coarse grained density function in a non-equilibrium case," as discussed in LN 5, top of page 15.

[3] This is due to the fact that $\rho = \rho_1\rho_2$, and thus, $\log\rho = \log\rho_1 + \log\rho_2$, where $\rho_1 = \rho_1(H_1, L_{1,1}, L_{2,1}, ...)$ and $\rho_2 = \rho_2(H_2, L_{1,2}, L_{2,2}, ...)$ are the PDFs for the weakly interacting subsystems 1 and 2, respectively. This means that $\log\rho_1$ and $\log\rho_2$ must be *additive* constants of motion. In classical mechanics, only E (energy), $\vec{P}$ (linear momentum), and $\vec{L}$ (angular momentum) are additive constants (cf., Landau, Mechanics), and so the general form for $\log\rho$ must be $\log\rho = AE + \vec{B}\cdot\vec{P} + \vec{C}\cdot\vec{L}$, where $A, \vec{B}, \vec{C}$ are constants.

[4] As in LN 1, we use x as a general displacement parameter, and J as a general force parameter. For a P-V system, $x = V$ and $J = -P$. For an H-M system, $x = M$ and $J = H$.

6.2.1 Probability

The central postulate of the statistical mechanics (Eq. 6.5 or Eq. 6.4), means that, for a microcanonical ensemble in equilibrium,

$$p_{E,x}(\mathrm{µ}) = \frac{1}{\Omega(E,x)}\,\delta_{H(\mathrm{µ}),E} \tag{6.6}$$

where δ here is the Kroncker delta symbol and Ω is the number of all accessible/possible microstates[5] $\{\mathrm{µ}\}$.

Here, µ represents one unique state in the phase space: it corresponds to a unique set of initial conditions for $\vec{\mathbf{p}}(0)$ and $\vec{\mathbf{q}}(0)$ and the unique phase space trajectory from it. Namely, µ is equivalent to $(\vec{\mathbf{p}}(t), \vec{\mathbf{q}}(t))$. Thus, µ is a $6N$ dimensional vector, as each bold vector is defined as a $3N$ dimensional vector. Since what we are doing is really semi-classical statistical mechanics, it is quite appropriate to keep in mind that a microstate in quantum mechanics is the total many body wave function for the N body system.

What is $\Omega(E,x)$? In accordance with what we discussed in the last lecture, it is given by

$$\Omega(E,x) = \frac{\text{volume in phase space}}{h^{3N}} \tag{6.7}$$

Now, you may notice that there is a slight problem here. If the energy is really precisely defined, then the volume is simply zero. However, keep in mind that the energy really precisely well-defined in classical mechanics simply means a finite energy uncertainty given by the energy step in quantum mechanics. So, as we are thinking about the semi-classical physics here, it is quite reasonable to invoke a small width in energy. We shall define this width as

$$\Delta E = 2\Delta. \tag{6.8}$$

We shall see that this width plays no essential role in the computation of thermodynamic quantities (see page T14 of the textbook, which has the same argument, in the example of free particles).

Another problem is that the above counting does not take into account the identical particle nature of the N particles. If they come within a small distance to interact with each other (like in gas), then they can no longer be considered as non-identical. We need to divide by $N!$ to account for the multiple counting in classical mechanics.

$$\Omega(E,x) = \frac{\text{volume in phase space}}{h^{3N}\,N!} \qquad \text{for identical particles} \tag{6.9}$$

---

[5] The symbol µ is use for microstate. Distinguish it from μ, which means chemical potential (see later).

6.2.2 Primary state function – entropy

The entropy of the system is defined as

$$S(E,x) \equiv k_B \log \Omega(E,x) \tag{6.10}$$

and plays the primary role in characterizing a microcanonical ensemble. Note that this entropy is just k_B times what we defined in Eq. 5.28 (which was for general state, not just for equilibrium state), since in the current case $p_i = 1/\Omega(E,x)$, while $i = 1, 2, ..., \Omega(E,x)$. An importance difference is that the current definition applies only to the equilibrium state.

Let us calculate the entropy for a monatomic ideal gas[6]. In this case, the total energy is given by $E = \sum_{i=1}^{N} \frac{p_i^2}{2m}$. So, if the energy uncertainty is 2Δ, then the volume in phase space is $S_{3N} \cdot (2mE)^{3N/2-1} \cdot V^N \cdot \sqrt{2m}\frac{\Delta}{\sqrt{E}}$. Here, S_d is the total *solid* angle in d-dimensions ($S_1 = 2$, $S_2 = 2\pi$ and $S_3 = 4\pi$). It can be obtained by $I_d \equiv \left(\int_{-\infty}^{\infty} dx e^{-x^2}\right)^d = \pi^{d/2}$. Since the integral expands to $S_d \int_0^\infty dr r^{d-1} \exp(-r^2) = S_d \frac{1}{2} \int_0^\infty dt t^{d/2-1} \exp(-t) = S_d \, (d/2-1)!/2$, we get $S_d = 2\pi^{d/2}/(d/2-1)! = 2\pi^{d/2}/\Gamma(d/2)$. Thus, from Eq. 6.9,

$$\Omega(E,V) = \frac{2\pi^{3N/2}}{(3N/2-1)!} \cdot (2m)^{(3N-1)/2} \cdot E^{(3N-3)/2} V^N \Delta \cdot \frac{1}{h^{3N} N!} \tag{6.11}$$

The entropy is (ignoring such terms as $\log 2$, $\log E$ or $\log \Delta$ in comparison to the dominant term[7]),

$$S(E,V) \approx k_B \left[\frac{3N}{2} \log(2m\pi E/h^2) - \log(3N/2-1)! + N \log V - \log N!\right] \tag{6.12}$$

Using the Stirling's formula (Eq. 5.10), $\log N! \approx N \log N - N + O(\log N)$, we get

$$S(E,V) = k_B N \left[\frac{3}{2} \log\left(\frac{4m\pi E}{3h^2 N}\right) + \log \frac{V}{N} + \frac{5}{2}\right] \tag{6.13}$$

This is the so-called *Sackur-Tetrode equation.* This is a beautiful expression in that (1) it is extensive (i.e. when E, N, V are scaled by λ, S also scales as λS) and (2) it does not suffer from any unknown constant of order N. The fact (1) owes it to the consideration of the identical particle (the $1/N!$ factor) in Eq. 6.9, while the fact (2) owes it to the identification of h as the quantum of action in the same equation. However, note that this entropy is clearly incorrect as T goes to zero, as it

---

[6] From this point on to the rest of this section, we use $x = V$.

[7] Note that the energy width parameter 2Δ is completely irrelevant in this problem for large N. As long as some width (due to quantum effect or instrumental resolution, etc.) exists, it does not affect our final result.

does not approach a constant. To correct *this* deficiency, the semi-classical physics is not enough and one needs to do the full quantum statistics.

By the way, the above derivation of the Sackur-Tetrode equation is not necessarily the easiest. We will revisit this expression with another (more convenient) derivation later on in Section 6.3.4.

In relation to point (1) just made, consider the following. Suppose there are two different gases at the same temperature. N_1, V_1 characterizes the first gas and N_2, V_2 characterizes the second gas. Imagine that they are initially separated by a thin hard wall, which is suddenly removed. When the two are mixed, the gases are then spread to the total volume $V = V_1 + V_2$, the final entropy increases. This increase in entropy can be calculated from Eq. 6.13 as (noting that E/N is a function of only T for an ideal gas)

$$\begin{aligned}\Delta S_{mix,inh} &= k_B N_1 \log\frac{V}{N_1} + k_B N_2 \log\frac{V}{N_2} - k_B N_1 \log\frac{V_1}{N_1} - k_B N_2 \log\frac{V_2}{N_2} \\ &= k_B N_1 \log\frac{V}{V_1} + k_B N_2 \log\frac{V}{V_2} \qquad (6.14)\end{aligned}$$

However, if the two gases are identical, then

$$\Delta S_{mix,hom} = k_B N \log\frac{V}{N} - k_B N_1 \log\frac{V_1}{N_1} - k_B N_2 \log\frac{V_2}{N_2} \qquad (6.15)$$

These are the mixing entropy formulae for ideal gas. In particular, consider the case when the initial densities are identical with the final density $N_1/V_1 = N_2/V_2 = N/V$. The homogeneous mixing entropy is zero, in this case, according to Eq. 6.15, while the inhomogeneous mixing entropy is always positive. This is of course a correct result.

Had we stuck to the purely classical statistics where the identical particle nature of gas molecules when they come into contact is ignored, then we would have gotten a finite mixing entropy even for the homogeneous gases of the same density. This is called the "Gibbs paradox" for a historical reason.

6.2.3 Laws of thermodynamics

Once the entropy concept is firmly established, one can go ahead and make connections with the thermodynamics. Keep in mind that there cannot be a complete derivation of the laws of thermodynamics from the equilibrium statistical mechanics alone (cf. last lecture). However, we can make certain non-trivial connections to the laws of thermodynamics, starting from the concept of the entropy in equilibrium thermodynamics alone.

Suppose two systems with energy[8] $\mathcal{E}_1$ and $\mathcal{E}_2$ constitute a total system, which is a closed system. Assuming that they interact weakly, we have the total energy $E = \mathcal{E}_1 + \mathcal{E}_2$, and also the total volume $x = \chi_1 + \chi_2$. The total number of states is given by

$$\Omega(E) = \sum_{\mathcal{E}_1+\mathcal{E}_2=E,\ \chi_1+\chi_2=x} \Omega_1(\mathcal{E}_1,\chi_1)\Omega_2(\mathcal{E}_2,\chi_2). \tag{6.16}$$

So, the energy exchange between the systems is not through any mechanical work. Here, the sum can be considered as the sum over $\mathcal{E}_1$, which has *discrete* values with an energy step given by 2Δ as we defined in Section 6.2.1.

Note that both Ω_1 and Ω_2 are exponentially large numbers. Let us say that the product is maximized at a certain value of $\mathcal{E}_1 = E_1$ and the corresponding value of $\mathcal{E}_2 = E_2 = E - \mathcal{E}_1$. And similarly, for a certain value of $\chi_1 = x_1$ and the corresponding value of $\chi_2 = x_2 = x - x_1$. Then, by the rule of large numbers, we conclude that

$$\log\Omega(E,x) = \log\Omega_1(E_1,x_1) + \log\Omega_2(E_2,x_2). \tag{6.17}$$

Multiplying by k_B, we conclude, nicely, that the entropy is extensive

$$S_{tot}(E,x) = S_1(E_1,x_1) + S_2(E_2,x_2). \tag{6.18}$$

The zeroth law

The maximization condition for $S_1(\mathcal{E}_1,\chi_1) + S_2(\mathcal{E}_2,\chi_2)$, under the constraints $\mathcal{E}_1 + \mathcal{E}_2 = E$ and $\chi_1 + \chi_2 = x$, leads to the following conditions (cf., page 7 of LN 2).

$$\begin{aligned}\left(\frac{\partial S_1(E_1,x_1)}{\partial E_1}\right)_{x_1} &= \left(\frac{\partial S_2(E_2,x_2)}{\partial E_2}\right)_{x_2}\\ \left(\frac{\partial S_1(E_1,x_1)}{\partial x_1}\right)_{E_1} &= \left(\frac{\partial S_2(E_2,x_2)}{\partial x_2}\right)_{E_2}\end{aligned}$$

By defining

$$\frac{1}{T} \equiv \left(\frac{\partial S}{\partial E}\right)_x \qquad \text{statistical mechanical definition of temperature} \tag{6.19}$$

the first condition amounts to

$$T_1 = T_2 \qquad \text{in equilibrium.} \tag{6.20}$$

---

[8] From now on, we will use the convention that symbols written in Roman characters, such as E or x, generally correspond to the average values, or equivalently, most probable values, while calligraphic/Greek characters such as $\mathcal{E}$ or χ correspond to random variables. So, for a macroscopic system, only E and x are experimentally accessible values, while $\mathcal{E}$ and χ are not.

This is the basic content of the zeroth law of thermodynamics (LN 1), i.e., there is a parameter that is equalized by systems in equilibrium, when they exchange energy without doing work. Note that in the current statistical mechanism formalism, we are taking an inductive approach, where the entropy is the primary quantity, from which everything follows.

The second condition amounts to

$$J_1 = J_2 \qquad \text{in equilibrium} \tag{6.21}$$

if two systems exchange displacement x. We already knew this from lectures 1 and 2. Again, the difference now is that we deduce it from the maximum entropy condition, and by identifying the force J as

$$J \equiv -T\left(\frac{\partial S}{\partial x}\right)_E \tag{6.22}$$

The first law

We can justify the identification that we just made, easily, if we identify TdS as the definition of the heat $đQ$ in a reversible process. In particular, let us consider such a process with $đQ = 0$. Then, we have

$$\begin{aligned} 0 = đQ = TdS &= T\left(\frac{\partial S}{\partial E}\right)_x dE + T\left(\frac{\partial S}{\partial x}\right)_E dx \\ &= dE + T\left(\frac{\partial S}{\partial x}\right)_E dx \qquad \text{Eq. 6.19} \end{aligned}$$

Rearranging, and identifying J as in Eq. 6.22, we get

$$dE = TdS + Jdx \tag{6.23}$$

which is (a less general form of) the first law of thermodynamics.

The second law

Let us consider what the above consideration of two weakly interacting subsystems means in terms of the second law. Suppose that one starts with a configuration that does not maximize the summand in Eq.. 6.16. In such a case, $\Omega_1(\mathcal{E}_1, \chi_1)\Omega_2(\mathcal{E}_2, \chi_2) < \Omega_1(E_1, x_1)\Omega_2(E_2, x_2)$. That is, the number of states in the ensemble is less than the maximum entropy state. It is then clear that, such a configuration corresponds to a probability distribution that is *not* uniform across all accessible states of the maximum entropy ensemble. Thus the process of going from $\Omega_1(\mathcal{E}_1, \chi_1)\Omega_2(\mathcal{E}_2, \chi_2)$ to $\Omega_1(E_1, x_1)\Omega_2(E_2, x_2)$ can be thought of as **the process of maximizing the entropy of a closed system**, in accordance with the fundamental assumption, Eq. 6.5.

Of course, if the entropy is the maximum to start with, then it will stay constant.

6.3 Canonical ensemble

For a system that is in contact with a big heat reservoir kept at a constant temperature, then the system is constrained to have a constant temperature T. Assuming that no work is exchanged between the reservoir and the system, the "volume" of the system, x, is also constant.

The ensemble of states under the constant T and the constant x is called the *canonical ensemble.*

6.3.1 Probability

Let us consider two subsystems, one of which is very large. We will call the large system a reservoir, denoted by the subscript R, and the small system simply the system. First of all, the total system is a microcanonical ensemble, and the total number of states is given by

$$\Omega_{tot} = \sum_{\mathcal{E}} \Omega(\mathcal{E}, x)\Omega_R(\mathcal{E}_R, x_R) = \sum_{\mathcal{E}} \Omega(\mathcal{E}, x)\Omega_R(E_{tot} - \mathcal{E}, x_R) \tag{6.24}$$

where $\Omega(\mathcal{E}, x)$ is the number of states for the system at energy $\mathcal{E}$. As in the above, we use the calligraphic symbol $\mathcal{E}$ or $\mathcal{E}_R$ to denote the random variable, reserving roman symbols for the average/most-probable value of energy.

Now, the assumption of a large reservoir means that $E_{tot} \approx \mathcal{E}_R \gg \mathcal{E}$, which means that we can do[9]

$$\begin{aligned} k_B \log \Omega_R(\mathcal{E}_R, x_R) &= S_R(\mathcal{E}_R, x_R) = S_R(E_{tot} - \mathcal{E}, x_R) \\ &\approx S_R(E_{tot}, x_R) - \mathcal{E}\left(\frac{\partial S_R}{\partial E_R}\right)_{x_R} \\ &\approx S_R(E_{tot}, x_R) - \frac{\mathcal{E}}{T} \\ &\approx S_R(E_{tot}, x_R) - k_B \beta \mathcal{E} \qquad \beta \equiv 1/(k_B T) \end{aligned}$$

Therefore we can replace $\Omega_R(\mathcal{E}_R, x_R)$ above with $e^{-\beta\mathcal{E}} e^{S_R(E_{tot}, x_R)/k_B}$, getting

$$\Omega_{tot} = A \sum_{\mathcal{E}} \Omega(\mathcal{E}, x) e^{-\beta\mathcal{E}} \qquad (6.25)$$

where $A = e^{S_R(E_{tot}, x_R)/k_B}$ is a constant, independent of the parameters of the system.

Now, by Eq. 6.6, all Ω_{tot} microstates are all equally likely. But here, we must focus on one microstate µ of the system. By definition, this is one of the states $\Omega(\mathcal{E} = H(\mu), x)$, where $H(\mu)$ denotes the value of the Hamiltonian for the microstate µ. As there are $Ae^{-\beta\mathcal{E}}$ states possible for the reservoir for each microstate µ of the system with energy $\mathcal{E}$, we see that the probability distribution for µ must be such that

$$p(\mu) \propto \exp(-\beta\mathcal{E}_\mu) \qquad \mathcal{E}_\mu \equiv H(\mu)$$

Applying the probability sum rule, we then obtain

$$\begin{aligned} p(\mu) &= \frac{\exp(-\beta\mathcal{E}_\mu)}{Z} && \text{PDF, canonical ensemble} \qquad (6.26) \\ Z &= \sum_\mu \exp(-\beta\mathcal{E}_\mu) && \textbf{partition function} \qquad (6.27) \end{aligned}$$

where Z is the partition function, which is the normalization constant for the probability distribution for the canonical ensemble. It is an *extremely important function*! Note that the probability depends only on energy, while the sum is over microstates,

---

[9] Here, one *cannot* expand Ω_R directly for small $\mathcal{E}$. Why not? If one tries, then one gets, to the leading order of $\mathcal{E}$, $\Omega_R(E_{tot} - \mathcal{E}, x_R) \approx \Omega_R(E_{tot}, x_R)\,(1 - \beta\mathcal{E})$. Clearly this expansion is valid only for $|\beta\mathcal{E}| \ll 1$, a condition which *cannot* be satisfied for semi-classical statistical mechanics or quantum statistical mechanics for Fermions! Why did this happen? It is because Ω_R is such a fast rising function of $\mathcal{E}_R$.

some of which may have the same energy. So, we can re-express the probability with $\mathcal{E}$ as the random variable, instead of µ. For doing so, first note that

$$\begin{aligned} Z &= \sum_{\mu} \exp(-\beta\mathcal{E}_{\mu}) = \sum_{\mathcal{E}} \Omega(\mathcal{E}) \exp(-\beta\mathcal{E}) \\ &= \sum_{\mathcal{E}} \exp\left(\frac{\mathcal{S}(\mathcal{E})}{k_B}\right) \exp(-\beta\mathcal{E}) \end{aligned}$$

Therefore, by defining

$$\mathcal{F}(\mathcal{E}) \equiv \mathcal{E} - T\mathcal{S}(\mathcal{E}) \tag{6.28}$$

we can now express the PDF and the partition function in terms of $\mathcal{E}$ as the random variable

$$p(\mathcal{E}) = \frac{\exp(-\beta\mathcal{F}(\mathcal{E}))}{Z} \tag{6.29}$$

$$Z = \sum_{\mathcal{E}} \exp(-\beta\mathcal{F}(\mathcal{E})) \tag{6.30}$$

However, note that **Eq. 6.27 is the most fundamental and, often, most useful**.

6.3.2 Primary state function – Helmholtz free energy

The macroscopic Helmholtz free energy, F, is the value of $\mathcal{F}$ that maximizes the summand of Eq. 6.30 and it can be written as[10]

$$F \equiv E - TS = -k_B T \log Z \tag{6.31}$$

Note that F that maximizes the summand is the minimum of $\mathcal{F}$, in agreement with LN 2 (Section 2.3).

Another quantity that can easily be derived from Z is the energy and its cumulants. Here is a derivation. Note that Z can be formally thought of as a function of β. We form a rather artificial but extremely useful function

$$c(\beta, \gamma) \equiv \frac{Z(\beta)}{Z(\gamma)} \tag{6.32}$$

We can write

$$
\begin{aligned}
\langle \mathcal{E}^n \rangle &= \frac{\sum_\mu \mathcal{E}_\mu^n \exp(-\beta \mathcal{E}_\mu)}{Z} && (6.33)\\
&= \frac{1}{Z} \frac{\partial^n Z}{\partial(-\beta)^n} && (6.34)\\
&= \left. \frac{\partial^n c(\beta,\gamma)}{\partial(-\beta)^n} \right|_{\gamma=\beta} && (6.35)
\end{aligned}
$$

Recalling that $\langle x^n \rangle = \frac{\partial^n \tilde{p}(k)}{\partial(-ik)^n}$, we see that $c(\beta,\gamma) = \tilde{p}(k = -i\beta)$ as long as we consider γ another independent variable, which we put $\gamma = \beta$ only after all necessary differentiations with respect to β are completed. With this rule, we can regard c as the characteristic function with a complex $k = -i\beta$. Then, $\log c$ is the second characteristic function that generates cumulants. Namely, $\langle \mathcal{E}^n \rangle_c = \left. \frac{\partial^n c(\beta,\gamma)}{\partial(-\beta)^n} \right|_{\gamma=\beta}$, from which we arrive at a very useful relation

$$
\langle \mathcal{E}^n \rangle_c = \left(\frac{\partial^n \log Z}{\partial(-\beta)^n} \right)_x \qquad n = 1, 2, \ldots \qquad (6.36)
$$

So, effectively, $\log Z$ **is the second characteristic function for** $\mathcal{E}$ **with the substitution** $-ik \to -\beta$**.**

Let us examine the variance, in particular.

$$
\begin{aligned}
\langle \mathcal{E}^2 \rangle_c &= \left(\frac{\partial \langle \mathcal{E} \rangle_c}{\partial(-\beta)} \right)_x && (6.37)\\
&= \frac{dT}{d(-\beta)} \left(\frac{\partial E}{\partial T} \right)_x && (6.38)\\
&= k_B T^2 C_x && (6.39)
\end{aligned}
$$

Thus, the uncertainty in energy is given by

$$
\Delta E = \sqrt{\langle \mathcal{E}^2 \rangle_c} = T \sqrt{k_B C_x} \qquad (6.40)
$$

In particular, this means that ΔE is of order $\sqrt{N}$, since E and C_x are of order N. In **the thermodynamic limit** ($N \to \infty, V \to \infty, V/N \to$ finite), $\Delta E/E \to 0$ as $O(1/\sqrt{N})$, making the total energy value precisely defined.

6.3.3 Which ensemble to choose?

As we just found out, the canonical ensemble for which we allowed the energy value to adjust to keep temperature constant ends up attaining a unique energy value in the thermodynamic limit. Thus, for a given physical system of interest, we could either choose the microcanonical ensemble *or* the canonical ensemble. For computing the thermodynamic variables and their interrelations, it does not matter which one to choose. Or, any other ensemble discussed later on in this note will do fine too. It is the convenience of the calculation whereby a particular ensemble is chosen.

Having said this, you will find that calculations using the canonical ensemble is generally much more straightforward than those using the microcanonical ensemble!

6.3.4 Ideal gas

The partition function for a monatomic ideal gas is easily calculated as

$$Z = \int \frac{\prod_i d^3\vec{p}_i d^3\vec{q}_i}{N! h^{3N}} \exp\left(-\beta \sum_{i=1}^{N} \frac{\vec{p}^2}{2m}\right) \tag{6.41}$$

Here, we have incorporated the two key aspects of the semi-classical statistical mechanics, namely the phase space volume quantum h^{3N} (N is the number of particles) and the consideration of the particle exchange symmetry ($N!$).

Using $\int dx \exp(-\beta x^2/(2m)) = \sqrt{\pi}\sqrt{2mk_BT}$, we get

$$Z = \frac{V^N}{N!}\left(\sqrt{\frac{2\pi m k_B T}{h^2}}\right)^{3N} \tag{6.42}$$

Define

$$\lambda(T) \equiv \frac{h}{\sqrt{2\pi m k_B T}} \tag{6.43}$$

which may be called the **thermal de Broglie wavelength** as it is on the order of the de Broglie wavelength for the thermal energy $3k_BT/2$.

Note that the applicability of Newton's laws in semi-classical statistical mechanics means that $\lambda(T)^3 N/V \ll 1$. Using λ, and we get

$$\boxed{Z = \frac{1}{N!}\left(\frac{V}{\lambda^3}\right)^N \qquad (6.44)}$$

From this, F and E can be readily obtained using $F = -k_B T \log Z$ and $E = \left(\frac{\partial \log Z}{\partial \beta}\right)_V = \frac{3}{2} N k_B T$ (see the next section). And, then, the entropy S can be obtained, from $S = (E - F)/T$, with the identical result as in Section 6.2.2.

$$\begin{aligned} F &= -k_B T \left(N \log V - 3N \log \lambda - \log N!\right) \\ &\approx -k_B T \left(N \log V - 3N \log \lambda - N \log N + N\right) && \text{keep up to } O(N) \text{ for } \log N! \\ &= N k_B T \left(\log \frac{N}{V} - 1 + 3 \log \lambda\right) && (6.45) \\ S &= (E - F)/T \\ &= N k_B \left(\frac{5}{2} + \log \frac{V}{N} - 3 \log \lambda\right) && \text{Sackur-Tetrode; same as Eq. 6.13} \quad (6.46) \end{aligned}$$

In addition, the pressure can be obtained as (from the thermodynamic identity, Eq. 2.17),

$$P = -\left(\frac{\partial F}{\partial V}\right)_T = k_B T \left(\frac{\partial \log Z}{\partial V}\right)_T = \frac{N k_B T}{V} \qquad (6.47)$$

establishing the equation of state.

6.3.5 Equipartition theorem

Now, let us look at the energy more closely: $E = -\left(\frac{\partial \log Z}{\partial \beta}\right)_V$ applied to Eq. 6.42 gives

$$E = \frac{3N}{2} k_B T \qquad \text{monatomic ideal gas} \qquad (6.48)$$

since $Z \propto \beta^{-3N/2}$. The origin of this dependence on β is the fact that there are $3N$ degrees of freedom $(\vec{p}_1, ..., \vec{p}_N)$ that contribute to the quadratic energy of the form Ap^2 where A is positive. Now, suppose that our gas consists of O_2 molecules. The above partition function considers the monatomic gas, and thus considers only the center of mass motion of each oxygen molecule. However, there are internal motions per each molecule and they must be included in the partition function as well. This is simple to do within the Harmonic approximation, i.e. assuming that the inter-atomic motion within a molecule is governed by the Hooke's law. Then the inter-oxygen potential energy is of the form Ar^2 where r is the inter-oxygen distance minus the bond length. This r is another degree of freedom for the problem to be included in the partition function. In addition, there will be an internal kinetic energy term of the form Bp_r^2 where B involves the reduced mass. The involved integrals are the same integrals that we already performed above, and so this problem of including the internal degrees of freedom is straightforward. In particular, the result for the energy is very simple, and does not require any new calculation. This is because every degree of freedom that gives rise to the quadratic function in energy will contribute to one power of $\beta^{-1/2}$ to the partition function. This means that the total energy is given by

$$E = \frac{Q_d}{2} k_B T \qquad \textbf{equipartition theorem}\text{; any (semi-classical) ideal gas} \tag{6.49}$$

where Q_d is the total number of the quadratic degrees of freedom. This is the **equipartition theorem**. In the O_2 example, we get $Q_d = (3 + 2 + 2)N = 7N$, where the last two is from the rotational motion. However, at room temperature, the experiment does *not* show this equipartition energy. Why not? As we shall see later, quantum statistical mechanics is necessary to explain – the gist is that the vibrational modes are too high in energy (too large quantum energy level) so that they are not excited at room temperature. The discrepancy between the equipartition theorem and the experiment was a great impetus to the birth of quantum mechanics.