Notes for Lecture 7

Ensembles

In the previous lecture, we have established the micro-canonical ensemble and the canonical ensemble. For each different ensemble, it is important to know to ask, and answer, these following questions. First, what fundamental quantity we need to calculate first? (Ω and Z, respectively.) Second, what primary state function follows from it? ($S = k_B \log \Omega$ and $F = -k_B T \log Z$, respectively.) Third, how can we calculate other quantities of interest these first two quantities?

Note that the third question is a "how" question, not a "what" question, since *all* equilibrium thermodynamic functions can be calculated. A general answer to this how question is "**by using the fundamental thermodynamic identities (page 11 of LN 2) or by recognizing the log of the partition function as the cumulant generator**." We have used the latter fact to prove that $\left(\frac{\partial^n \log Z}{\partial(-\beta)^n}\right)_x = \langle \mathcal{E}^n \rangle_c$ ($n = 1, 2, \ldots$), which is an extremely general, nice, and useful result. For instance, by taking $n = 1$, we see that we can calculate the energy $E = \langle \mathcal{E} \rangle_c$ in a straightforward manner from the partition function Z. With $F = -k_B T \log Z$ and E at hand, $S = (E - F)/T$ and other quantities are derivable from the table given in page 11 of LN 2. However, this is not a unique approach. One can also note that $S = -\left(\frac{\partial F}{\partial T}\right)_x$ from the thermodynamic identity, and one can then derive other quantities from F and S, instead.

Later in this lecture, we will summarize these useful core results, including those for the two more common ensembles that we will discuss, now. Last, but not least, note that much of the formalism that we discussed in the previous lecture applies to the quantum statistical mechanics as well, despite the title of that lecture. Whether we are considering the semi-classical mechanics or the quantum mechanics is reflected upon the first quantity that we calculate (Ω or Z). The formal ways in which other thermodynamic functions follow from this first quantity is independent of the microscopic nature of mechanics (quantum or semi-classical).

7.1 Reservoir

Let us recall that the canonical ensemble arises naturally from the consideration of a system in contact with a heat reservoir. As any system that we are interested in tends to be in contact with the environment and the experimental equipment, the concept of a reservoir is a very practical one, and perhaps an even essential one, in a real sense, even though it is not necessary for the microcanonical ensemble. A **heat reservoir** constrains the equilibrium temperature, T, of the system, by virtue of its large size and its having a thermal contact with the system (exchange of S). Now, one can easily imagine a **work reservoir** as well: it constrains the equilibrium value of J of the system, by virtue of its large size and its having a mechanical contact with the system (exchange of x).

As discussed in Section 1.7, (J, x) pairs, i.e., pairs of "force–displacement" variables, include all pairs of variables that occur in a single term of the right hand side of the fundamental thermodynamic identity $dE = TdS + \sum_i J_i dx_i$, except for (T, S). So, (μ, N) can be such a pair, representing *chemical work*, performed through "the chemistry of exchanging particles." Other important examples are $(-P, V)$ and (H, M).

Suppose that such a work reservoir, constraining one "force" variable J, is in contact with our system.

Proceeding as in Section 6.3.1, we must look at the number of states for system *plus* reservoir, to establish the fundamental state function for the system. Note that what we need to do is actually quite simple, as before: we need to count the multiplicity of each microstate of the system, μ, by counting the number of reservoir states Ω_R consistent with the configuration of μ (as done just before Eq. 6.5).

Since the reservoir and the system exchange work, both the energy and the displacement of the reservoir fluctuate, and so we use calligraphic/Greek symbols to denote random variables, as in the last lecture: $\mathcal{E}_R$ and χ_R. The reservoir is large in energy ($\mathcal{E}_R$) and size (χ_R), and so, we can treat the system energy ($\mathcal{E}$) and the system size (χ) in $\mathcal{E}_R = E_{tot} - \mathcal{E}$ and $\chi_R = \chi_{tot} - \chi$, as small variables.

$$\begin{aligned}
k_B \log \Omega_R(\mathcal{E}_R, \chi_R) &= S_R(\mathcal{E}_R, \chi_R) \\
&= S_R(E_{tot} - \mathcal{E}, x_{tot} - \chi) \\
&\approx S_R(E_{tot}, x_{tot}) - \mathcal{E}\left(\frac{\partial S_R}{\partial E_R}\right)_{x_R} - \chi\left(\frac{\partial S_R}{\partial x_R}\right)_{E_R} \\
&\approx S_R(E_{tot}, x_{tot}) - \frac{\mathcal{E}}{T} + \frac{J\chi}{T} && \text{Eqs. 6.19, 6.22} \\
&\approx S_R(E_{tot}, x_{tot}) - k_B\beta(\mathcal{E} - J\chi) && \beta \equiv 1/(k_B T)
\end{aligned}$$

So, we get a similar result as before, and conclude that

$$p(\mu) \propto \exp\left[-\beta(\mathcal{E} - J\chi)\right] \tag{7.1}$$

For normalization, we can divide this exponential factor by $\sum_{\mu} \exp\left[-\beta(\mathcal{E} - J\chi)\right]$, which we can recognize as another kind of partition function. When the probability is expressed in terms of the energy of the system, we must take into account of the number of states of the system for a given energy, $\mathcal{E}$, as before (cf. Eqs. 6.28,6.29,6.30 and the discussion above them), and we get

$$p(\mathcal{E}) \propto \exp\left[-\beta(\mathcal{E} - T\mathcal{S} - J\chi)\right] \tag{7.2}$$

Most often, these results are applied to the situation where the reservoir constrains *both* T and J. Indeed, the important specific ensembles that we will consider below are for that type of situations. However, none of what we did so far in this section *requires* that T be constrained by the reservoir, while we definitely required that J be constrained. So, it is of some interest, in principle, to spend a bit of time to examine the nature of a pure work reservoir, which only constrains J but not T, as it does not exchange heat with the system. Notice that in the presence of a pure work reservoir the system's temperature *can and will* change[1]; it is just that in general (i.e., barring coincidence) it will not be the same as the temperature of the reservoir. However, the topic of a pure work reservoir will be treated as optional, since the topic is just a bit atypical.

Pure work reservoir (optional)

Here, by definition, there is no heat exchange between the system and the reservoir. This does *not* mean that the entropy of the reservoir is kept at a constant value.

If we do assume that the entropy of the reservoir remains constant, then we have $\Delta\mathcal{E}_R = J\Delta\chi_R$, and since $\Delta\mathcal{E}_R = -\mathcal{E}$ and $\Delta\chi_R = -\chi$, we get $\mathcal{E} - J\chi = 0$. So, the probability $p(\mu)$ is *independent of* $\mathcal{E}$ and χ! The system will settle in a state, in which the entropy of the system, S, is maximum. We are familiar with this type of behavior. A gas in a reversible adiabatic transformation is an example. If, in such a case, the pressure is constrained, then the volume of the system is determined so that[2] $dS = 0$, and the temperature is determined accordingly.

---

[1] For example, a volume change of an ideal gas under constant pressure would lead to a temperature change.

[2] The value of V that satisfies the adiabatic equation of state $f(P,V) =$ constant, given a value of P, will automatically satisfy $dS = 0$, since that is how the adiabatic equation of state, $f(P,V) = 0$ is determined. For example, $f(P,V) = PV^{\gamma}$ for an ideal gas, where $\gamma = (Q_d + 2N)/Q_d$, and Q_d is the total number of quadratic degrees of freedom (Eq. 6.49).

If, on the other hand, the entropy of the reservoir changes, then it can only increase, due to Eq. 2.8. In addition, let us suppose that the system's entropy does not change. In this case, we get, from Eq. 7.2, that the maximum entropy is equivalent to the minimization of the function, $\mathcal{E} - J\chi$, which is identified as **the enthalpy** for a P-V system. This conclusion is as discussed in pages 9,10 of LN 2. An example is a mechanical equilibrium in the presence of a potential field.

Lastly, we can still consider the case when both the entropy of the system S and the entropy of the reservoir $(\mathcal{E} - J\chi)/(-T_R)$ for a pure work reservoir increase. In this case, Eq. 7.2 shows that the potential function, $\mathcal{E} - T_R\mathcal{S} - J\chi$ is minimized, corresponding to the entropy maximization.

Note that in the last part, we use T_R to denote the temperature of the reservoir, not T, since the reservoir temperature is generally not the same as the system temperature for a pure work reservoir.

General reservoir

The last point just stated is a bit interesting, and seems worth emphasizing. Generalizing that point to any reservoir, which can constrain any combination of T or J_i's, one can note the point in a box given below.

If one looks at the derivation for Eq. 7.2 carefully, one recognizes that J and T are both the values for the reservoir. So, for the purpose of this paragraph, let us rewrite them as J_R and T_R, for the maximum clarity. Also, $\beta = \beta_R = 1/(k_B T_R)$.

Eq. 7.2 applies generally, for a system connected to any reservoir (including no reservoir). For a work reservoir, $J = J_R$, and, for a heat reservoir, $T = T_R$. However, this is valid only for the *final equilibrium state* of the system. During the transition to the equilibrium state, the state of the system may be described approximately as an equilibrium state, but $\mathcal{T} \neq T_R$ and $\mathcal{J} \neq J_R$ for such an intermediate state, in general. **None of $\mathcal{E}$, $\mathcal{S}$, and χ depend on T_R or J_R!** Of course, E, S, and x do depend on T and J.

For a system connected to a general reservoir (including no reservoir), the maximum entropy principle is equivalent to the minimization of

$$\beta_R \left(\mathcal{E} - T_R\mathcal{S} - J_R\chi\right). \tag{7.3}$$

T_R and J_R are the temperature and the force of the reservoir, which may *or may not* be equal to those of the system in the final equilibrium state, depending on the nature of (the connection to) the reservoir.

Note that this point was made before, basically: see page 15 of LN 2. Here, we are making this point in a more general language and in a broader light. Also, while it is not explicitly noted in the box, $J_R\chi$ can clearly be generalized to $\sum_i J_{R,i}\chi_i$ if there are many ways in which work can be performed. In any case, the above means that $d\mathcal{E} - T_R d\mathcal{S} - J_R d\chi = 0$ on equilibrium. Let us consider a pure work reservoir, for which we can[3] write $d\mathcal{E} = J d\chi$ and $d\mathcal{S} = 0$, from which it follows that $J = J_R$, in equilibrium. So, $J = J_R$ for a work reservoir, a proof that is rather independent of similar proofs that we already provided (cf. Section 2.2), since we are examining the particular case where $T \neq T_R$ ever.

7.2 Gibbs canonical ensemble

This is the case when J and x are mechanical variables (i.e. anything other than μ and N: $-P, V$ or H, M will do). So in a Gibbs canonical ensemble, T and J are fixed by the reservoir[4].

$$
\begin{aligned}
p(\mu) &= \frac{\exp\left(-\beta\left[\mathcal{E}(\mu) - J\chi\right]\right)}{X} && && (7.4)\\
X &= \sum_{\mu} \exp\left(-\beta\left[\mathcal{E}(\mu) - J\chi\right]\right) && \text{Gibbs partition function} && (7.5)\\
G &= E - TS - Jx \;=\; -k_B T \log X && \text{Gibbs free energy} && (7.6)\\
\langle \chi^n \rangle_c &= \frac{1}{\beta^n}\left(\frac{\partial^n \log X}{\partial J^n}\right)_T && n = 1, 2, \ldots && (7.7)\\
\langle \mathcal{H}^n \rangle_c &= \left(\frac{\partial^n \log X}{\partial(-\beta)^n}\right)_J && H = E - Jx \text{ (= enthalpy if } P\text{-}V) && (7.8)\\
G &= \mu N && \text{extensive system; Eq. 2.23} && (7.9)
\end{aligned}
$$

A warning on notation. Here H (or $\mathcal{H}$) is used for the enthalpy, not the Hamiltonian! Unfortunately, this is the standard notation for the enthalpy, and so there is room for

---

[3] This may not be obvious, but this is guaranteed by the fact that we can always pick a reversible path for given two points infinitesimally close in the phase space. One can also use a more general approach, as follows, and comes to a useful conclusion, which is consistent with this approach. $d\mathcal{E} = T d\mathcal{S} + J d\chi$, and so the condition becomes $(1 - T_R/T) d\mathcal{E} + (J T_R/T - J_R) d\chi = 0$. For a pure work reservoir, we do not have two independent variables, and so $d\mathcal{E}$ and $d\chi$ are not independent. So, we *cannot* conclude $T = T_R$ and $J = J_R$, as we would be able to do if the reservoir is both a heat reservoir and a work reservoir. However, we *can* pick a reversible path, from which we derive the necessary condition for the equilibrium: $J = J_R$. Plugging in this condition into $(1 - T_R/T) d\mathcal{E} + (J T_R/T - J_R) d\chi = 0$, and cancelling out $(1 - T_R/T)$, which is generally not zero, we get $d\mathcal{E} = J d\chi$. So, *any* state infinitesimally close to the final equilibrium state is reached by a reversible process. This is as expected, since $dS_{tot} = 0$ at the equilibrium point.

[4] Note that, except for the grand canonical ensemble, N for the system is considered a fixed number for all the ensembles considered so far.

confusion with the Hamiltonian. The good news is that we will *not* use it much. On the other hand, later on, we will encounter a few Hamiltonians, and use the symbol H for them.

7.3 Grand canonical ensemble

This applies case when $J = \mu$ and thus $x = N$. So, in a grand canonical ensemble T and μ are fixed while E, S and N are allowed to fluctuate. If any other "displacement variable" such as V or M exists, then it is considered fixed as in the canonical ensemble.

$$
\begin{aligned}
p(\mu) &= \frac{\exp\left(-\beta\left[\mathcal{E}(\mu) - \mu\mathcal{N}\right]\right)}{Q} && && (7.10)\\
Q &= \sum_{\mu} \exp\left(-\beta\left[\mathcal{E}(\mu) - \mu\mathcal{N}\right]\right) && \text{grand partition function} && (7.11)\\
\Phi &= E - TS - \mu N \;=\; -k_B T \log Q && \text{grand potential} && (7.12)\\
\langle \mathcal{N}^n \rangle_c &= \frac{1}{\beta^n}\left(\frac{\partial^n \log Q}{\partial \mu^n}\right)_T && n = 1, 2, \ldots && (7.13)\\
\langle (\mathcal{E} - \mu\mathcal{N})^n \rangle_c &= \left(\frac{\partial^n \log Q}{\partial(-\beta)^n}\right)_\mu && n = 1, 2, \ldots && (7.14)\\
\Phi &= Jx \quad (= -PV) && \text{extensive system; Eq. 2.21} && (7.15)
\end{aligned}
$$

The grand canonical ensemble comes handy when the calculation becomes much easier by allowing particle numbers unrestricted. This is often the case in quantum statistics. It follows, from Eq. 7.13, that $\Delta N \;=\; \sqrt{\langle \mathcal{N}^2 \rangle_c} = \sqrt{\frac{1}{\beta}\frac{\partial N}{\partial \mu}} = O(\sqrt{N})$ and so $\Delta N / N \to 0$ in the thermodynamic limit. This is much like what we found, using the canonical ensemble: $\Delta E / E = O(\sqrt{N}) \to 0$.

Note that we have, until now, not paying too much attention to N as a thermodynamic variable. Now that we just did, we must recognize that in a common situation N and another displacement parameter such as $x = V$ (or $x = M$) tend to occur at the same time. In such a case, it would be appropriate to take three independent variables, T, x, and N, for example, or any equivalent set, to describe the problem. This was the assumption for page 11 of LN 2. In the following grand summary section, we make the same assumption.

7.4 Summary of all ensembles

We are at a point, where it may be helpful to summarize most things about different ensembles and their characteristics in one place. The following table is such a summary. Note that first two rows summarize variables that form, when combined, the natural set of independent variables for the primary state function.

	microcanonical	canonical	Gibbs canonical	grand canonical
conserved (by law)	E, x, N	x, N	N	x
constrained (reservoir)	none	T	T, J	T, μ
what to calculate first	Ω = # of μ states	Z $= \sum_{\mu} e^{-\beta\mathcal{E}}$	X $= \sum_{\mu} e^{-\beta(\mathcal{E}-J\chi)}$ $= \sum_{\chi} e^{\beta J\chi} Z(\chi)$	Q $= \sum_{\mu} e^{-\beta(\mathcal{E}-\mu\mathcal{N})}$ $= \sum_{\mathcal{N}} e^{\beta\mu N} Z(\mathcal{N})$
primary state function (extensivity)	S $= k_B \log\Omega$	$F = E - TS$ $= -k_B T \log Z$	$G = F - Jx$ $= -k_B T \log X$ $= \mu N$	$\Phi = F - \mu N$ $= -k_B T \log Q$ $= Jx$
thermo identity	$dE = TdS$ $+Jdx + \mu dN$	$dF = -SdT$ $+Jdx + \mu dN$	$dG = -SdT$ $-xdJ + \mu dN$	$d\Phi = -SdT$ $+Jdx - Nd\mu$
cumulant ($n = 1, 2, \ldots$)		$\langle \mathcal{E}^n \rangle_c$ $= \left(\frac{\partial^n \log Z}{\partial(-\beta)^n} \right)_{x,N}$	$\langle (\mathcal{E} - J\chi)^n \rangle_c$ $= \left(\frac{\partial^n \log X}{\partial(-\beta)^n} \right)_{J,N}$	$\langle (\mathcal{E} - \mu\mathcal{N})^n \rangle_c$ $= \left(\frac{\partial^n \log Q}{\partial(-\beta)^n} \right)_{\mu,x}$
cumulant ($n = 1, 2, \ldots$)			$\langle \chi^n \rangle_c$ $= \left(\frac{\partial^n \log X}{\partial(\beta J)^n} \right)_{T,N}$	$\langle \mathcal{N}^n \rangle_c$ $= \left(\frac{\partial^n \log Q}{\partial(\beta\mu)^n} \right)_{T,x}$

How to read this SUPER-important table[5]

1. The first row contains conserved quantities, which are extensive.
2. The second row contains quantities that the reservoir regulates on the system (and they become the same between the system and the reservoir), and these quantities are intensive.

---

[5] This table is valid for the quantum statistical mechanics as well, with one caveat. In quantum mechanics, random variables such as $\mathcal{E}$, χ, and $\mathcal{N}$ must be considered as *operators*. If non-commuting operators are involved, then the cumulant formulas in the last row break down in general, while still remaining valid for $n = 1$. For instance, if $\mathcal{E}$ and χ do not commute (e.g., when $\chi = \mathcal{M}$ in a magnetic problem), then the cumulant formula for χ is not to be trusted for $n \geq 2$.

3. The quantities in the first two rows list quantities that are given initially when a problem starts. Namely, if a problem specifies temperature and pressure as given, then you must naturally consider a Gibbs canonical ensemble.

4. Then, calculate quantities in the third and fourth rows—these quantities are naturally a function of the variables given in the first two rows. When you do this calculation, the quantities in the first two rows must be constants; i.e., they should be independent of micro-states considered. If they are not, then the calculation is not valid.

5. Want any other quantities? Use the identity in the fifth row. For instance, to get the temperature in the microcanonical ensemble, note that it is given by $T = \left(\frac{\partial E}{\partial S}\right)_{x,N}$. Or, use those cumulant formulas in the last two rows, e.g., $V = \left(\frac{\partial \log X}{\partial(-\beta P)}\right)_{T,N}$.

6. A partition function is sort of (see Sec. 6.3.2) a first characteristic function (with β replacing ik), and therefore the cumulant formulas turn out the way that they do in the sixth and seventh rows (see Sec. 6.3.2 for the rigorous derivation, which is readily adaptable to all the cumulant formulas here).

7. In the thermodynamic limit ($N \to \infty$ and N/V = finite), all of the above ensembles are equivalent, since any thermodynamic variable has negligible fluctuation ($\Delta N/N \to 0$ (Sec. 7.3), $\Delta E/E \to 0$ (Sec. 6.3.2), etc.).

7.5 Probability theory point of view

For a given probability distribution, the quantity

$$s = -\sum_i p_i \log p_i \tag{7.16}$$

is the so-called Shannon entropy[6]. One can define the entropy of a thermodynamic system by

$$S = -k_B \sum_{\mu} p(\mu) \log p(\mu) \tag{7.17}$$

which is sometimes referred to as the Gibbs-Boltzmann entropy. Defined for an arbitrary PDF for microstates $p(\mu)$, it is a quantity defined not only for equilibrium states (Eq. 6.10), but also for any non-equilibrium state. We shall see that this

---

[6] This needs a bit of qualification. For the Shannon entropy, one must replace log with $\log_2$. So, the above function is the Shannon entropy up to a scale factor.

definition of the entropy connects well to another conventional entropy that is widely used in quantum statistics, the von Neumann entropy. In any case, note that the above Gibbs-Boltzmann entropy is k_B times the Shannon entropy.

From the above definition, one can easily re-derive all main results summarized in the table, from a slightly different approach than we used in previous sections and the previous lecture. This derivation provides an alternative way of thinking about various ensembles.

Microcanonical ensemble

If the entropy function above is to be maximized, under the single constraint, $\sum_\mu p(\mu) = 1$, then it is straightforward to show that the result is

$$p(\mu) = \frac{1}{\Omega} \qquad \Omega = \text{the total number of possible } \mu \text{ states} \tag{7.18}$$

Physically, this corresponds to the microcanonical ensemble, characterized by conservation laws of energy, x, and N, while there is no constraint imposed by the reservoir.

Canonical ensemble

If the entropy function above is to be maximized, under *two* constraints,

$$\sum_\mu p(\mu) = 1 \tag{7.19}$$

$$\sum_\mu \mathcal{E}(\mu) p(\mu) = E \tag{7.20}$$

then, we can proceed as follows. Using the undetermined Lagrange multiplier method, we define

$$f(p(\mu)) = -k_B \sum_\mu p(\mu) \log p(\mu) - \lambda_1 \left(\sum_\mu p(\mu) - 1 \right) - \lambda_2 \left(\sum_\mu \mathcal{E}(\mu) p(\mu) - E \right) \tag{7.21}$$

Having introduced the undetermined multipliers λ_1 and λ_2, we can treat each $p(\mu)$ as an independent variable, as long as we include λ_1 and λ_2 as unknown variables to be solved for with the help of the constraint equations themselves. Differentiating f with respect to $p(\mu)$'s, we get, for the maximum[7] of S,

$$\frac{\partial f}{\partial p(\mu)} = -k_B \log p(\mu) - k_B - \lambda_1 - \lambda_2 \mathcal{E}(\mu) = 0. \tag{7.22}$$

---

[7] Note that the second derivative is given by $\frac{\partial^2 f}{\partial p(\mu) \partial p(\mu')} = -k_B \frac{1}{p(\mu)} \delta_{p(\mu),p(\mu')} \le 0$, proving the maximum nature.

Thus, we get a result that

$$p(\mu) = \exp\left(-1 - \frac{\lambda_1}{k_B}\right)\exp\left(-\frac{\lambda_2}{k_B}\mathcal{E}\right)$$

Clearly, λ_1 is basically a normalization constant, which can be solved for, using the first constraint, $\sum_\mu p(\mu) = 1$. That is,

$$p(\mu) = \frac{\exp\left(-\frac{\lambda_2}{k_B}\mathcal{E}\right)}{Z}$$
$$Z \equiv \sum_\mu \exp\left(-\frac{\lambda_2}{k_B}\mathcal{E}\right)$$

The Z function can be turned into the $\mathcal{E}$ sum, which is,

$$Z = \sum_{\mathcal{E}} e^{\frac{\mathcal{S}}{k_B}} \exp\left(-\frac{\lambda_2}{k_B}\mathcal{E}\right) \approx \exp\left(\frac{S - \lambda_2 E}{k_B}\right)$$

where the rule of large numbers has been used for the $\mathcal{E}$ sum so that S and E correspond to the those values that maximize the summand. For the summand to be maximum, we must have $\delta\mathcal{S} - \lambda_2\delta\mathcal{E} = 0$, i.e., $\lambda_2 = \frac{\partial\mathcal{S}}{\partial\mathcal{E}}$. Thus, we identify λ_2 as $1/T$, and subsequently λ_2/k_B as $\beta = 1/(k_B T)$. Therefore, we get the same result as Eqs. 6.26 and 6.27, the standard results for the canonical ensemble PDF and the partition function. In particular, $-k_B T \log Z = E - TS$, and so $-k_B T \log Z$ is identified as the Helmholtz free energy F.

Now, the second constraint, $\sum_\mu \mathcal{E} p(\mu) = E$, clearly means that $E = \frac{\partial \log Z}{\partial(-\beta)}$. That this is consistent with the relation that we just obtained, $F = -k_B T \log Z$ is easy to demonstrate: $\frac{\partial \log Z}{\partial(-\beta)} = \frac{\partial \beta F}{\partial\beta} = F + \beta\frac{\partial F}{\partial\beta} = F + \beta\frac{\partial F}{\partial T}\frac{dT}{d\beta} = F + ST = E$. So, we reproduce all results of the canonical ensemble by maximizing the entropy, Eq. 7.17, under two constraints, Eqs. 7.19, 7.20. Along the way, note that we also proved that the most probable energy value for $\mathcal{E}$ is the same as the average value E. None of these results are new, but it is interesting to derive these results from a slightly different point of view, which is quite worth remembering.

Other ensembles

It should be clear from the work on the canonical ensemble that we can do just the same for Gibbs canonical ensemble or grand canonical ensemble. For the former, we can add one more constraint, $\sum_\mu \chi(\mu)p(\mu) = x$, while for the latter, we can add one more constraint, $\sum_\mu \mathcal{N}(\mu)p(\mu) = N$, to the two constraints, Eqs. 7.19, 7.20. Other than the fact that there is one more constraint, the mathematical steps to be used to re-produce all results in Sections 7.2 and 7.3 are very similar to the mathematical steps used in the canonical ensemble case. You are invited to work out these other ensembles from the above probability theory point of view.