

Unit I Review

Mohan Shankar

*Department of Chemistry, University of Virginia,
Charlottesville, Virginia*

E-mail: mohan-s@virginia.edu

ABSTRACT: This document attempts to provide an overview of the content that one might be tested on during the first midterm. You'll likely have a question(s) about entropy, but I didn't want to inundate you with you content, so I left it off. It is my hope that this document can provide a base-level recap of concepts rather than thoroughly reviewing all material.

Contents

1	Energy Differentials and Path (In-)Dependence	1
1.1	Differentials	2
1.2	Thermodynamic Definitions	3
1.3	Equations for Calculations	4
2	Probability and Partition Functions	5
2.1	Probability	5
2.2	Ensembles	6
2.3	Partition Functions	6
2.4	Molecular Partition Functions	8
2.4.1	Translation, Vibration, Rotation, and Electronic	9

1 Energy Differentials and Path (In-)Dependence

Although it's out of order compared to the textbook, I think it's appropriate to start with energy differentials. The most important one is the differential for energy (i.e. the 1st Law of Thermodynamics):

$$dU = \delta q_{rev} + \delta w_{rev} \begin{cases} \delta q_{rev} = TdS \\ \delta w_{rev} = -PdV \end{cases} \quad (1.1)$$

Here, dU represents the differential of a *path-independent* quantity while the δ for heat (q) and work (w) denotes the two quantities are *path-dependent*. I don't have a great reason why this is qualitatively the case, but the takeaway should be that the sum of two path-dependent quantities produces a path-independent quantity called a *state function*. You will see in the following subsection why this is the case, but looked [1.1](#), we can see the $U = U(S, V)$ meaning energy (U) is a function of entropy and volume. A visual analogy of [1.1](#) can be found below in Figure 1.

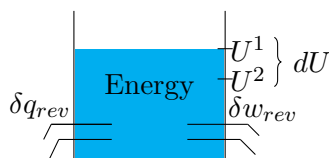


Figure 1. Visualization of path dependence.

The two taps correspond to a change in work and heat. Opening either tap leads to a change in energy that is, indeed, dependent on which path was taken. Recall $dU =$

$\delta q_{rev} + \delta w_{rev}$ while $\Delta U = q + w$. That is, **integrating** dU and $\delta q/\delta w$ produces ΔU and q/w respectively.

Some notes about convention and concepts: $\delta w_{rev} = -PdV$ is used so that when $\delta w_{rev} > 0$, we are doing work on the system by compressing it making $dV < 0$. Conversely, the system expanding means $dV > 0$ so $\delta w_{rev} < 0$ implying the system is doing work.

The reason this convention is useful is that energy must be *minimized* at equilibrium. As such:

Compression \longrightarrow Minimum Work

Compression means $dV < 0$ so 1.1 becomes $dU = TdS - (-) PdV$ so $dU = TdS + PdV$. For dU to be as small as possible, the work term ($+PdV$) should be minimized.

Expansion \longrightarrow Maximum Work

As mentioned above, compression means $dV < 0$ so 1.1 becomes $dU = TdS - (+) PdV$ thus $dU = TdS - PdV$. For dU to be as small as possible, the work term ($-PdV$) should be as large as possible.

Lastly, let's clear some points up. **Adiabatic** means no *heat* is transferred, but this doesn't mean temperature remains constant. Heat (q) is the transfer of energy through thermal contact, so it has units of energy— perhaps helping distinguish it from temperature. In fact, heat is a function of entropy rather than temperature. **Isothermal** means there's no change in temperature. For an *ideal gas* whose energy is *solely* dependent on temperature, an isothermal process means there is no change in energy (i.e. $\Delta U = 0$). There isn't a fancy word for **no work** done, but know that this condition is signaled by a constant volume.

1.1 Differentials

When considering a multi-variable function $f = f(x_1, x_2, \dots, x_n)$, a differential is an infinitesimal change in that multi-variable function. The general form for the differential of a multi-variable function of n variables is as follows:

$$df(x_1, x_2, \dots, x_n) = \left(\frac{\partial f}{\partial x_1} \right)_{i \neq j} dx_1 + \left(\frac{\partial f}{\partial x_2} \right)_{i \neq j} dx_2 + \dots + \left(\frac{\partial f}{\partial x_n} \right)_{i \neq j} dx_n \quad (1.2)$$

where x_i is used to reference the variables f is a function of. We see the general pattern is the partial derivative of the function with respect to one of its variables multiplied by the differential of the same variable. The subscript $i \neq j$ means that the partial derivative only cares about the variable indexed by i and is constant in all other variables. As an example:

$$f(x, y) = 6yx^2 \implies df = \left(\frac{\partial f}{\partial x} \right)_y dx + \left(\frac{\partial f}{\partial y} \right)_x dy = (12yx) dx + (6x^2) dy$$

Here, we leverage the nice feature of partial derivatives where we treat everything other than the variable in question as a constant. As such,

$$\left(\frac{\partial f}{\partial x}\right)_y = \left(\frac{\partial}{\partial x}\right)_y [6yx^2] = 12yx \implies \left(\frac{\partial f}{\partial y}\right)_x = \left(\frac{\partial}{\partial y}\right)_x [6yx^2] = 6x^2$$

It's important to note that the chain rule and product rule among others would still apply if relevant. We denote what f is constant with respect to when taking these partial derivatives by placing subscripts after the parenthesis. This choice is very important and should not be forgotten when writing out your partial derivatives. Lastly, the order of application for partial derivatives don't matter when differentiating something is an *exact differential*. $\frac{\partial^2}{\partial x \partial y} = \frac{\partial^2}{\partial y \partial x}$. This will be useful later when learning about Maxwell constructions.

To elaborate a bit on inexact differentials, let's consider $\delta w_{rev} = -PdV$. If we were to write dW where we guess $W = -PV$, the differential would look like

$$dW = \left(\frac{\partial W}{\partial V}\right)_P dV + \left(\frac{\partial W}{\partial P}\right)_V dP = -PdV - VdP$$

which is not the result we want. It should hopefully be clear that no combination of P and V when defining W can produce $dW = -PdV$ hence there is a need of the δ to denote an inexact differential.

Remark. You may see $\not{d}Q$ and $\not{d}W$ (or some other slash) rather than δ if you look online. This difference in notation has no effect on the actual meaning of the math.

1.2 Thermodynamic Definitions

As seen when comparing the energy differential in 1.1 and the general form in 1.2, we can express thermodynamic quantities as partial derivatives of energy. Perhaps the most famous is a quantity we use everyday: temperature. It's *defined* as follows:

$$T \equiv \left(\frac{\partial E}{\partial S}\right)_{\vec{x}} \quad (1.3)$$

Further,

$$\frac{1}{T} = \left(\frac{\partial S}{\partial E}\right)_{\vec{x}}$$

The other definitions that will be of use is that of specific heat. In its case, there are two forms depending on if volume or pressure is constant.

$$C_v \equiv \left(\frac{\partial U}{\partial T}\right)_V \quad (1.4)$$

$$C_P \equiv \left(\frac{\partial H}{\partial T}\right)_P \quad (1.5)$$

Qualitatively, we see that specific heat is defined to be the change in energy as temperature changes. This should hopefully coincide with how you understand it. Where the subtlety

lies is in distinguishing between 1.4 and 1.5 due to which “energy” we use.

In the case of 1.4, we use energy since it is a function of volume. As such, it is possible for us to hold that volume constant. For 1.5, we use enthalpy as that energetic quantity is a function of pressure. To summarize what the various energies are a function of:

$$U = U(S, V); H = H(S, P); A = A(T, V); G = G(T, P).$$

1.3 Equations for Calculations

The most common one that you’re likely to see will be the integral leveraging the fact that

$$C_V \equiv \left(\frac{\partial U}{\partial T} \right)_V \implies dU = C_V dT \implies \int_{U_1}^{U_2} dU = \underbrace{\int_{T_1}^{T_2} C_V dT}_{\text{use!}} \implies \Delta U = U_2 - U_1 = \bar{C}_V(T_2 - T_1)$$

thus we’ll use it when there’s a change in temperature at a constant volume. If volume were to change, we’d leverage

$$w = - \int_{V_1}^{V_2} P dV$$

where the *form* of P depends on the equation of state (i.e. Ideal Gas, van der Waals, etc.) If you get a question like this, make sure to carefully look at the units. Oftentimes, the value of R with its associated units will be changed to make the overall problem work out (consider PSET 2, Q4). The same form of the integral holds true for ΔH , only we integrate $C_P dT$.

The last, main integral to consider focusing on is the one for heat and entropy. If we consider that $dU = \underbrace{T dS}_{\delta q} - P dV$, we can re-arrange this to see that

$$dS = \frac{\delta q_{rev}}{T} \longrightarrow dS = \frac{dU}{T} + \frac{P dV}{T}$$

Additionally, we can re-arrange 1.1 to show $\delta q = dU - \delta w$ meaning $\delta q = dU + P dV$. Don’t get tripped up by the signs! Consequently, we can write

$$\delta q = C_V dT + P dV \implies q = \int_{T_1}^{T_2} C_V dT + \int_{V_1}^{V_2} P dV$$

Note. Suppose a process has two steps, one in which no work is done, and one in which work is done while both processes see a change in temperature. You can add the heat associated with each process q_1, q_2 to find q_{tot} .

Finally, you can find entropy using what we saw above:

$$dS = \frac{dU}{T} - \frac{\delta w}{T} = \frac{C_V dT}{T} + \frac{P dV}{T} \implies \Delta S = \int_{T_1}^{T_2} \frac{C_V}{T} dT + \int_{V_1}^{V_2} \frac{P}{T} dV$$

For questions like these, you’ll have to be able to parse key words/phrases like *adiabatic*, *constant volume*, *no work*, *isothermal*, etc.

2 Probability and Partition Functions

The chief object of interest for Statistical Mechanics is the partition function, denoted by Q in the textbook [?]. It has many uses, and in the case of the *Canonical Ensemble* takes the following form:

$$Q = \sum_j e^{-\beta E_j} \begin{cases} j = \text{index for distinct state} \\ \beta \equiv \frac{1}{k_B T} \end{cases} \quad (2.1)$$

As an aside, physicists typically denote the partition function using ‘ Z ’ or perhaps ‘ Ξ ’. This is purely a difference in notation — the meaning remains the same.

2.1 Probability

Before discussing the partition function further, I imagine it would be useful to discuss some formalism of probability. As such, we will start with a simple scenario: rolling a fair die.

Outcome	1	2	3	4	5	6
Probability (Weight)	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

Table 1. Outcome and associated probability of rolling a fair die.

If you were asked to give the expected value of a roll, what would your answer be? Intuitively, you might imagine it be 3.5 since that is the *average* result. That is, $\frac{1+2+3+4+5+6}{6} = \frac{21}{6} = 3.5$. This is a lucky break that works because the probability associated with each “state” is equal. Let’s now consider the scenario in Table 2.

Outcome	1	2	3	4	5	6
Probability (Weight)	$\frac{0}{6}$	$\frac{0}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{3}{6}$

Table 2. Outcome and associated probability of rolling an unfair die.

How would you find the expected value of a roll? As it turns out, you can multiply the probability associated with each state then add these values up.

$$\text{Average Roll} = \frac{0}{6} \cdot 1 + \frac{1}{6} \cdot 2 + \frac{1}{6} \cdot 3 + \frac{1}{6} \cdot 4 + \frac{1}{6} \cdot 5 + \frac{3}{6} \cdot 6 = 5$$

This result should make sense since we’ve eliminated the two lowest possible values to roll. Generalizing this idea in 2.2:

$$\langle \varphi \rangle = \sum_j p_j \varphi_j \quad (2.2)$$

Here, φ is an arbitrary observable quantity (energy, pressure, chemical potential, etc), ‘ $\langle \rangle$ ’ denotes an ensemble average at equilibrium, j is an index for the distinct states, and p_j the probability associated with each state.

In the event that the your probabilities are not normalized ($\sum_j p_j \neq 1$), you divide by the probabilities summed up, which is why we divide by Q for the partition function.

2.2 Ensembles

Physically, we can imagine the canonical ensemble using the following figure:

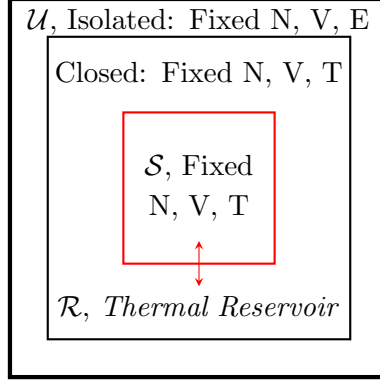


Figure 2. Visualization of a canonical ensemble as a subsystem of an isolated system.

Here, the outermost black box \mathcal{U} represents an *isolated* system. Isolated meaning the system *does not* exchange matter nor energy with its surroundings. \mathcal{S} and \mathcal{R} *exchange energy* indicated by \leftrightarrow but no energy from \mathcal{S}, \mathcal{R} is exchanged with \mathcal{U} . That is to say, \mathcal{S} and \mathcal{R} can exchange energy as heat, but both systems retain the particles present. In this case, these systems are characterized by the number of particles (N), their volume (V), and their temperature (T).

Justification: A glass of water sitting on a table has a well-defined average energy, but the energy is constantly fluctuating as it interacts with the environment. For such systems, it is more appropriate to think of them as sitting at fixed temperature T .

Remark. The term “reservoir” means that the energy of the subsystem is negligible compared with that of the reservoir. In particular, the subsystem can happily absorb or donate energy from or to the reservoir without changing the ambient temperature T .

2.3 Partition Functions

Returning back to the partition function (Q), it is formally a function that describes the statistical properties of a system in equilibrium. As the Simon and McQuarrie say, the probability of some state j for the canonical ensemble is proportional to $e^{-\beta E_j}$ where $\beta \equiv 1/k_B T$ is inverse temperature. Once you have these various energies, you can sum the terms up resulting in the partition function as shown in 2.1. Consequently, the probability associated with some state j is:

$$p_j = \frac{e^{-\beta E_j}}{Q}. \quad (2.3)$$

This formalism is invoked to ensure the probabilities sum up to 1 (i.e. they are normalized). Re-writing 2.3 in another manner:

$$p_j = \frac{e^{-\frac{E_j}{k_B T}}}{\sum_n e^{-\frac{E_n}{k_B T}}} \quad (2.4)$$

Here, one can explicitly see that the probability of some state j is proportional to $e^{-\frac{E_j}{k_B T}}$ and equal to this divided by the sum over all possible states. This is to say, j is a member of n possible states that is hand-picked.

You can consider the partition function to be the probability distribution for a particular system. Consequently, it is extremely useful as you can calculate many things using it including thermodynamic quantities. One such example is average energy of the system. Let's work it out by using the blueprint shown in 2.2.

$$\langle E \rangle = \sum_j p_j E_j = \sum_j E_j \frac{e^{-\beta E_j}}{\sum_j e^{-\beta E_j}} = \sum_j E_j \frac{e^{-\beta E_j}}{Q} \quad (2.5)$$

The question here is how can we “simplify” this process and get average energy $\langle E \rangle$ from Q ? Let's start by differentiating Q with respect to β :

$$Q = \sum_j e^{-\beta E_j} \implies \left(\frac{\partial Q}{\partial \beta} \right)_{N,V} = \sum_j \left(\frac{\partial}{\partial \beta} \right) e^{-\beta E_j} = \sum_j -E_j \cdot e^{-\beta E_j}$$

Comparing this expression to 2.5, we still need the factor of $1/Q$ to ensure our probabilities are normalized. That and we need to get rid of the negative sign that associated with E_j . We can do this by leveraging the nice property of the natural log's derivative, $\frac{d}{dx}[\ln(x)] = \frac{1}{x} \cdot \frac{d}{dx}[x]$ since we know that $\frac{d}{dx}$ is $\left(\frac{\partial Q}{\partial \beta} \right)_{N,V}$ and x is Q . Using this and adding the negative sign, we now have

$$\langle E \rangle = - \left(\frac{\partial \ln(Q)}{\partial \beta} \right)_{N,V} \quad (2.6)$$

To find the partial derivative needed to find $\langle E \rangle$ by differentiating with respect to temperature, we use the chain rule:

$$\langle E \rangle = - \left(\frac{\partial \ln(Q)}{\partial \beta} \right)_{N,V} \implies \left(\frac{\partial \ln(Q)}{\partial T} \right)_{N,V} = - \left(\frac{\partial \ln(Q)}{\partial \beta} \right)_{N,V} \cdot \left(\frac{\partial \beta}{\partial T} \right)_{N,V}$$

We know $\beta = 1/k_B T$, so $\left(\frac{\partial \beta}{\partial T} \right)_{N,V} = -\frac{1}{k_B T^2}$. Substituting this back in,

$$\left(\frac{\partial \ln(Q)}{\partial T} \right)_{N,V} = - \left(\frac{\partial \ln(Q)}{\partial \beta} \right)_{N,V} \cdot -\frac{1}{k_B T^2}$$

The negative signs cancel out, and we leverage the what we found in 2.6 to find:

$$\left(\frac{\partial \ln(Q)}{\partial T} \right)_{N,V} = \frac{\langle E \rangle}{k_B T^2}$$

rearrange and we see

$$k_B T^2 \left(\frac{\partial \ln(Q)}{\partial T} \right)_{N,V} = \langle E \rangle \quad (2.7)$$

It's worth pointing out that in the case of 2.6 and 2.7, our subscript for the partial derivative has N and V . This is because $Q = Q(N, V, T)$ thus we must specify what our derivatives are *independent* of when taking them.

To be precise, the notation of these partial derivatives reflect that we don't care to know how the number of particles N nor the volume of the system V affect the system in this case. This is not to say that changing the volume or number of particles would leave the energy of the system unaffected.

2.4 Molecular Partition Functions

The molecular partition function, denoted by q , requires ε_j values rather than E_j which can be determined using quantum mechanics. Recall that there are four main types of energy: translational, rotational, vibrational, and electronic modeled by particle in a box, rigid rotor, and harmonic oscillator respectively.

If the molecular partition functions are *independent*, you can multiply them together to find Q . The reasoning returns back to probability. Returning back to the analogy of the dice, we can imagine a single die as one system with associated states, rolling a value 1-6. We can add the probabilities associated with each state together which is why we get:

$$q = \sum_i e^{-\beta \varepsilon_i} = e^{-\beta \varepsilon_0} + e^{-\beta \varepsilon_1} + e^{-\beta \varepsilon_2} + \dots$$

of course, the actual probability associated with a state i is

$$\text{prob}(j) = \frac{e^{-\beta \varepsilon_j}}{q}$$

for the sake of normalization, but the work above is still valid.

$$Q = \begin{cases} \frac{q^N}{N!} & \text{if indistinguishable} \\ q^N & \text{if distinguishable} \end{cases} \quad (2.8)$$

The reason we multiply q 's to get Q is due to our ability to multiply *independent* probabilities together. If we suddenly care about the results of rolling two die, the probability of rolling a 1 and a 5 become $\frac{1}{6} \times \frac{1}{6}$ which should hopefully make sense. If we assume all N molecules have the same associated states, we can write this multiplication succinctly as q^N . If these molecules are indistinguishable, we add $1/N!$ as a necessary coefficient for combinatorics reasons—i.e. to avoid over-counting.

Finding average molecular energy, this time denoted as $\langle \varepsilon \rangle$ can be done using

$$\langle \varepsilon \rangle = - \left(\frac{\partial \ln(q)}{\partial \beta} \right)_{N,V} = k_B T^2 \left(\frac{\partial \ln(q)}{\partial T} \right)_{N,V} \quad (2.9)$$

2.4.1 Translation, Vibration, Rotation, and Electronic

I won't say much about these partition functions. You'll have to refer to quantum mechanics to learn where these things come from, but I promise it's a very fascinating story. Anyhoo, written out:

$$q_{\text{trans}}(T) = \left(\frac{2\pi m K_B T \ell^2}{h^2} \right)^{d/2} \quad (2.10)$$

d-dimensional translation. If $d = 3$, $q_{\text{trans}} = \left(\frac{2\pi m K_B T}{h^2} \right)^{3/2} V$ as ℓ^2 gets raised to the $3/2$ leaving $\ell^3 = V$.

For the vibrational partition function, you saw how the series converges to

$$q_{\text{vib}}(T) = \frac{e^{-\frac{\beta h \nu}{2}}}{1 - e^{-\beta h \nu}} \quad (2.11)$$

and if you ignore zero point energy,

$$q_{\text{vib}}(T) = \frac{1}{1 - e^{-\beta h \nu}} \quad (2.12)$$

Rotation is a bit of a pain, but we'll see it works out to be

$$q_{\text{rot}}(T) = \frac{k_B T}{\sigma h B} = \frac{1}{\beta \sigma h B} \quad (2.13)$$

where $B = h/8\pi I^2$ and σ is the symmetry number of the molecule.

Lastly, the electronic partition function is

$$q_{\text{el}}(T) = \sum_i g_{e_i} e^{-\beta \varepsilon_{e_i}} \quad (2.14)$$

where ε_{e_0} is always zero and ε_{e_1} is measured relative to that 0.