Notes for Lecture 10

Quantum statistical mechanics

Let us look back on what we did so far with semi-classical mechanics. We first defined a density function in the phase space. We discussed general properties such as the Liouville theorem (and the Poincaré recurrence, both in the classical regime or the quantum regime). We discussed some limited phenomenological theory of the time reversal symmetry breaking (i.e. how an irreversible process actually arises) – Boltzmann's H theorem. Then, we made a leap into the statistical mechanics, deciding not to worry too much about how the time reversal symmetry breaking occurs. Instead we asserted that the system settles for the most probable state and this maximum entropy principle was taken as an axiom. From this, we could reproduce all thermodynamics laws (except the third law, which cannot be derived from semi-classical physics).

Now, starting from quantum mechanics, we do pretty much the same. We are actually quite well prepared for this task, since much of what we did in semi-classical statistical mechanics applies also to the quantum statistical mechanics (e.g., all the thermodynamic relations and how we go from the partition function to the grand partition function, etc.). We just need to make sure to know what needs to be changed in quantum statistical mechanics.

10.1 Micro state

The picture of a micro state (μ) as a trajectory in phase space is a Newtonian concept and must be revised in quantum mechanics as the uncertainty principle (and accordingly, quantum fluctuations) becomes more and more important.

Instead, **a micro state in quantum mechanics means a wave function**. Please do not confuse this with a one particle wave function. When we talk about a wave function in statistical physics, it means a **many body (an N-body) wave function**, by default.

10.2 Density matrix – why?

In a standard course of quantum mechanics, the density matrix is often skipped. Instead, most of the time is spent on a state vector, $|\Psi\rangle$, or its wave function such as $\langle \vec{x}_1, \vec{x}_2, ..., \vec{x}_N \,|\, \Psi\rangle$. But, the density matrix is an essential topic of quantum mechanics, and it can be viewed as more general than a state vector. Why?

10.2.1 An example – light polarization

Let us consider a simple experiment involving the photon polarization. Any physics student knows that there is such a thing as "an unpolarized beam of light." How does one describe it? We shall see that the description is possible with the density matrix. We shall see that it is *not* possible to describe such a state with a state vector alone.

Let us describe experiments, first. A beam of light is propagating, and we define the propagation direction as the positive z axis. A polarizer is mounted perpendicular to the beam. Let us say that we have an "x" polarizer, whose output beam has x polarization. If we take the identical polarizer, and rotate it by 90 degrees around the z axis, then we get a "y" polarizer. Clearly, there is an arbitrariness in how we define the x axis, but what matters is that we define the xyz axes so that they form a right-handed Cartesian coordinate system. The polarization corresponding to the polarizer can be defined as a unit-length vector $\alpha\hat{x} + \beta\hat{y}$ (see below). If α and β are real numbers, then we have a linear polarizer, which is basically just the x polarizer rotated by an appropriate amount of angle. However, we can have a circular polarizer or an elliptical polarizer[1]: for such a polarizer, α or β *must* be complex (see below). **An unpolarized beam of light, propagating in the z direction, is defined as a beam of light, whose intensity is reduced by half by any polarizer mounted in the xy plane.** It is also found that when one uses linear polarizers in succession, then the second linear polarizer reduces the intensity further by $\cos^2\alpha$, where α is the angle between the polarization axes of the two polarizers.

Now, let us build a quantum mechanical description of this experiment[2]. Quan-

[1] These can be made of birefringent wave plates.

[2] You might ask why we do quantum mechanics, when most polarization experiments involve

tum mechanically, the polarizer carries out a "measurement." If we consider an x polarizer, then its measurement operator is given by $|h\rangle\langle h|$, where we defined $|h\rangle$ (h for "horizontal") as the state vector for the x polarized light. Namely, if the initial beam is given by $|\Psi\rangle$, then the final beam after the measurement is given by $|h\rangle\langle h\,|\,\Psi\rangle$, and similarly for other polarizations.

$$|\Psi\rangle \xRightarrow{measurement} |h\rangle\langle h\,|\,\Psi\rangle = C_x\,|h\rangle \qquad \text{by } x \text{ polarizer: } C_x = \langle h\,|\,\Psi\rangle \tag{10.1}$$

$$|\Psi\rangle \xRightarrow{measurement} |v\rangle\langle v\,|\,\Psi\rangle = C_y\,|v\rangle \qquad \text{by } y \text{ polarizer: } C_y = \langle v\,|\,\Psi\rangle \tag{10.2}$$

$$|\Psi\rangle \xRightarrow{measurement} |a\rangle\langle a\,|\,\Psi\rangle = C_a\,|a\rangle \qquad \text{by any polarizer: } C_a = \langle a\,|\,\Psi\rangle \tag{10.3}$$

The last one describes the most general polarization, where the polarization eigenstate $|a\rangle$ is defined as

$$|a\rangle = \alpha\,|h\rangle + \beta\,|v\rangle \qquad |\alpha|^2 + |\beta|^2 = 1 \text{ (normalization condition)}, \tag{10.4}$$

$$\alpha = \cos\left(\frac{\theta}{2}\right),\ \beta = e^{i\phi}\sin\left(\frac{\theta}{2}\right),\ \text{with real } \theta \text{ and } \phi. \tag{10.5}$$

In the last θ, ϕ form[3] for α and β, we adopted a convention to take α to be real, taking advantage of the fact that the overall phase factor of $|a\rangle$ is unimportant. If α, β are not "special values," then $|a\rangle$ describes an elliptically polarized light. Here are some special values of α, β, which describe linearly polarized light or circularly polarized light: $|a\rangle_{\alpha=1,\beta=0} = |h\rangle$, $|a\rangle_{\alpha=0,\beta=1} = |v\rangle$, $|a\rangle_{\alpha=\frac{1}{\sqrt{2}},\beta=\frac{i}{\sqrt{2}}} \equiv |R\rangle$ (right circularly polarized light), and $|a\rangle_{\alpha=\frac{1}{\sqrt{2}},\beta=\frac{-i}{\sqrt{2}}} \equiv |L\rangle$ (left circularly polarized light). In general, if α and β are both real, then $|a\rangle$ describes a linearly polarized light, with the polarization axis

classical light – light with lots of quanta. The answer is very simple. Photons hardly interact with one another (any interaction is indirectly through charged particles), and so you can study one photon or many photons, with the same result for virtually all topics such as polarization, transmission, reflection, energy, momentum, interference, diffraction, etc.

[3]Why the particular form of Eq. 10.5? Here, we shall use the spin ½ language, while keeping in mind the fact that the photon spin, the quantum mechanical meaning of photon polarization, is really 1. Using the spin ½ language, with mapping $|h\rangle \to |z\uparrow\rangle$ and $|v\rangle \to |z\downarrow\rangle$, consider the spin rotation operator $R(\theta,\phi) = \exp\left(-i\frac{\phi}{2}\sigma_z\right)\exp\left(-i\frac{\theta}{2}\sigma_y\right)$. $R(\theta,\phi)\,|z\uparrow\rangle$ gives the up eigenstate for the spin component pointing to the direction defined by polar angle θ and azimuthal angle ϕ. Noting that $\exp\left(-i\frac{\theta}{2}\sigma_y\right) = \cos\left(\frac{\theta}{2}\right) - i\sigma_y\sin\left(\frac{\theta}{2}\right) \doteq \begin{pmatrix}\cos(\theta/2) & -\sin(\theta/2)\\ \sin(\theta/2) & \cos(\theta/2)\end{pmatrix}$, we see that $R(\theta,\phi)\,|z\uparrow\rangle = e^{-i\frac{\phi}{2}}\left[\cos\left(\frac{\theta}{2}\right)|z\uparrow\rangle + e^{i\phi}\sin\left(\frac{\theta}{2}\right)|z\downarrow\rangle\right]$, which is equal to Eq. 10.5 except for the overall phase factor. So, this is why we used the form of Eq. 10.5, even though there are other equally general ways to write a unit-length 2-dimensional complex vector. From this discussion, it is worth pointing out that *any* general spin ½ state is expressible as $R(\theta,\phi)\,|z\uparrow\rangle$, which is by definition also a spin up eigenstate as mentioned above. As θ and ϕ changes, the direction pointed to by $R(\theta,\phi)\,|z\uparrow\rangle$ can be seen as tracing points on a unit sphere in the real 3-dimensional space, e.g., $|z\uparrow\rangle \to (0,0,1)$ and $R(\pi,0)\,|z\uparrow\rangle = |z\downarrow\rangle \to (0,0,-1)$, and in such visualization of a general spin ½ state as a point on a unit sphere in the real space, the unit sphere is called the "Bloch sphere."

defined by $\alpha\hat{x} + \beta\hat{y}$, and, if $\frac{\beta}{\alpha} = \pm i$, then it describes a circularly polarized light, for which we must conclude that $|\alpha| = |\beta| = 1/\sqrt{2}$.

Now, let us prove that an unpolarized beam of light cannot be described by a wave function. PROOF ✎ `Suppose that there` *is* `a state vector` $|\Psi\rangle$ `such that` $|\langle a|\Psi\rangle| = \frac{1}{\sqrt{2}}$ `for any polarization eigenstate` $|a\rangle$. `Taking` $|a\rangle$ `as` $|h\rangle$ `or` $|v\rangle$, `we must conclude that` $|\Psi\rangle = \frac{1}{\sqrt{2}}|h\rangle + \frac{e^{i\phi}}{\sqrt{2}}|v\rangle$, `where` ϕ `is an arbitrary, but fixed, real number. Clearly,` $|a\rangle \equiv \frac{1}{\sqrt{2}}|h\rangle + \frac{e^{i\phi}}{\sqrt{2}}|v\rangle$ `is a polarization eigenstate! Since,` $\langle a|\Psi\rangle = 1$, `we have a contradiction.` QED. ✐

How do we represent such a state in quantum mechanics? The answer is the density matrix. We may write

$$\rho_{xy} \equiv \frac{1}{2}|h\rangle\langle h| + \frac{1}{2}|v\rangle\langle v| \tag{10.6}$$

for the density matrix. Each term in the density matrix is a product of the probability (1/2 for either term in this case) and the projection operator that corresponds to the corresponding measurement. This is not the only way to represent a beam of unpolarized light as a density matrix. One can also choose to measure $|R\rangle$ and $|L\rangle$, and may write

$$\rho_{LR} \equiv \frac{1}{2}|R\rangle\langle R| + \frac{1}{2}|L\rangle\langle L| \tag{10.7}$$

Of course, a single density matrix must be able to describe all possible measurements, not just those measurements that one may use to construct a density matrix. For example, with ρ_{LR} in hand, one should be able to answer the question: "what is the probability to measure x polarization?" How do we calculate such a probability? Here is the answer.

> For a given density matrix, ρ, the probability that a measurement of an operator O will result in an eigenvalue o, with the corresponding eigenstate $|o\rangle$, is given by
>
> $$P_o = \langle o|\rho|o\rangle \tag{10.8}$$

It is left for your exercise to show that (1) $P_a = \langle a|\rho_{xy}|a\rangle = \frac{1}{2}$ for any $|a\rangle$ (Eq. 10.4), (2) $P_a = \langle a|\rho_{RL}|a\rangle = \frac{1}{2}$ for any $|a\rangle$ (Eq. 10.4), and (3) $\rho_{xy} = \rho_{RL}$. Of course, showing (1) and (3) are sufficient to show (2). The condition (3) can be generalized to $\rho_{xy} = \rho_{ab}$ where $|a\rangle = \alpha|h\rangle + \beta|v\rangle$, normalized, and $|b\rangle = \beta^*|h\rangle - \alpha^*|v\rangle$ (so that $\langle a|b\rangle = 0$ and $\langle a|a\rangle = \langle b|b\rangle = 1$).

So, now, we know how to write down the representation of an unpolarized light in quantum mechanics!

$$\rho \doteq \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{pmatrix} \qquad \text{density matrix for the polarization of unpolarized light} \tag{10.9}$$

where $\doteq$ means "is represented by" in the sense of the quantum mechanical representation using a certain set of basis.

By the property (3) above, it turns out that, in this particular case, it does not matter what basis we use to represent the density matrix; we always get a diagonal matrix with the two diagonal elements being equal to $\frac{1}{2}$. This is because the unpolarized light is a special case!

It is important to note that just by measuring the polarization in one direction does not get you enough information for obtaining the density matrix. Let us consider an example that illustrates this. Suppose one has a state of photons, and, by measurement, discovers that the light is 60 % x polarized and 40 % y polarized. One may wonder if the matrix, $\left(\begin{smallmatrix} 0.6 & 0 \\ 0 & 0.4 \end{smallmatrix}\right)$, would represent this state in the xy basis. The answer is no, in general. How do we know this? It is easy to see that when one changes the basis to the RL basis, the matrix becomes $\left(\begin{smallmatrix} 0.5 & 0.2 \\ 0.2 & 0.5 \end{smallmatrix}\right)$. So, this state corresponds to an equal mixture of right circularly polarized light and left circularly polarized light. So, why can't we write the matrix as $\left(\begin{smallmatrix} 0.5 & 0 \\ 0 & 0.5 \end{smallmatrix}\right)$ in the RL basis? Clearly, there is an ambiguity here, and we need some more information. As we shall see shortly, the density matrix is a Hermitian matrix with a unit trace and 0 or positive diagonal elements. Thus, for a 2×2 problem like the current one, we need to know three numbers. They can be taken to be the average values of the x polarization, the x' polarization, and the circular polarization. For an unpolarized beam of light, all these numbers are zero, and this is the reason why the diagonal matrix above, $\left(\begin{smallmatrix} 0.5 & 0 \\ 0 & 0.5 \end{smallmatrix}\right)$, is the correct representation for the unpolarized state. In general, the density matrix in the two dimensional Hilbert space can be written as $\rho = \frac{1+c_1\sigma_1+c_2\sigma_2+c_3\sigma_3}{2}$, where 1 means the 2×2 unit matrix, c_i's are real numbers and σ_i's are Pauli matrices[4].

That there are off-diagonal elements means that "things are not quite random" or "things are not completely incoherent."

[4] The properties of coefficients c_i can be further narrowed down: $\rho = \frac{1}{2}(1 + \vec{c}\cdot\vec{\sigma}) = \frac{1}{2}(1 + |\vec{c}|\,\hat{c}\cdot\vec{\sigma})$. From this, it can be shown that $|\vec{c}| \le 1$ (hint: begin by showing that the ensemble average value of $\hat{c}\cdot\vec{\sigma}$ is $|\vec{c}|$). $|\vec{c}| = 1$ defines the "Bloch sphere" (see footnote 3). $|\vec{c}| = 1$ if and only if the state is pure (i.e., $\rho = |a\rangle\langle a|$ for a polarization eigenstate $|a\rangle$; hint: begin by showing that for a pure state (Eq. 10.5) $\vec{c} = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$).

10.2.2 Is the Universe in a pure state?

A pure state in quantum mechanics means a state that is described by a single wave function. We do not know the answer to this question. But, some people may conjecture that the answer is yes. Let us assume so, mostly for the fun of it, in this section.

Let us consider ourselves doing experiments in this Universe. As a rule, we are investigating only a small part of the Universe. We can control parameters that pertain to that small part. Our system under investigation must be reasonably disconnected from the rest of the Universe; otherwise, we cannot be in control of our experiments, and our results will not be reproducible! However, clearly we cannot completely separate our system from the rest of the Universe. There is a weak interaction between them. So, the total wave function of the Universe can be written down as[5]

$$|\Psi_U\rangle = \sum_{i,j} C_{i,j} |i\rangle |R, j\rangle \tag{10.10}$$

where $|i\rangle$ is the wave function of the system (ortho-normalized), $|R, j\rangle$ is the state of the rest of the universe (ortho-normalized), and i, j are the quantum numbers that describe the system and the rest, respectively.

So, we are measuring a dynamical variable, O, of our system. By definition, this variable is dependent only on the system, not on the rest of the Universe. What is the expectation value for this measurement? Since we have only one overall wave function, the answer is easy.

$$\langle O\rangle = \langle \Psi_U | O | \Psi_U\rangle \tag{10.11}$$

$$= \sum_{i,i',j} C^*_{i',j} C_{i,j} \langle i' | O | i\rangle \tag{10.12}$$

where in the second step, the ortho-normality condition, $\langle \Psi_{R,j} | \Psi_{R,j'}\rangle = \delta_{j,j'}$ has been used along with the fact that O is j independent, both by assumptions.

Note that, our expression for $\langle O\rangle$ is somewhat complicated. If we assumed that the system is completely disconnected from the rest of the Universe, then we would have gotten no j sum. This motivates us to define a matrix that absorbs the j sum:

$$\rho_{i,i'} \equiv \sum_j C^*_{i',j} C_{i,j} \tag{10.13}$$

[5]If the system and the rest of the Universe are completely disconnected, then the wave function can be written as a product: $|\Psi_U\rangle = |\Psi\rangle |\Psi_R\rangle$. If the system and the rest of the Universe are strongly interacting (i.e., they are "entangled"), then $|i\rangle |R, j\rangle$ must be written as $|i, j\rangle$.

This matrix is now dependent only on the (generally composite) index i for the Hilbert space of the system only. We can define this matrix the density matrix of the system. So, we now have

$$\langle O \rangle = \sum_{i,i'} \rho_{i,i'} \langle i' | O | i \rangle = \mathrm{tr}\{\rho O\} \tag{10.14}$$

which says that, if we know the density matrix, then we can compute the expectation value for any operator by taking the trace of the matrix product, ρO.

In this view, the density matrix represents the "traced-out" representation of the wave function of the Universe, where "tracing out" means the sum over j (Eq. 10.13), the (generally composite) index for the Hilbert space of the rest of the Universe. The necessity to invoke the density matrix arises from the fact that our system is not isolated.

For the last point, we can note the following. If the system were isolated completely, then the usual quantum mechanics applies. We can then take $i = o$ as the eigenvalue of O, since the eigenstates of any observable form a complete set of states to span the Hilbert space. Then, we would have

$$\langle O \rangle = \sum_{o} |C_o|^2 o \tag{10.15}$$

which is the basic common wisdom of quantum mechanics.

10.3 Quantum ensemble, Density matrix

Let us recall (Section 5.4) that the ensemble is a device by which we assert that there is a many-to-one mapping from micro states (that we cannot fully measure or control) to a macro state (that we observe consistently, nonetheless).

In classical mechanics, we could distribute microstates as dots in the phase space. In non-equilibrium states, dots have different probabilities even if they have the same total energy value (and same other conserved quantities), while they must, by a fundamental postulate of statistical mechanics (Section 6.1, Eq. 6.5 or 6.4), have the same probabilities of occurrence in equilibrium.

How do we represent an ensemble in quantum mechanics? Clearly a single wave function is not the right answer. A state represented by a single wave function is a state that we know completely, by definition. For an ensemble of states, all we know is that a given state may have a probability to be measured to be in this or that wave function (micro state). So, this is the clue. What represents a measurement? It is the projection operator: $|\Psi_\alpha\rangle\langle\Psi_\alpha|$, where we assume that $|\Psi_\alpha\rangle$'s are all possible members of the ensemble, i.e., all possible micro states. This leads to the definition of the *density matrix*, the quantum equivalent to the probability density function in semi-classical statistical mechanics[6]:

$$\rho(t) \equiv \sum_\alpha p_\alpha |\Psi_\alpha\rangle\langle\Psi_\alpha| \tag{10.16}$$

where p_α is the probability for the system to be in state Ψ_α and thus

$$p_\alpha \geq 0, \qquad \sum_\alpha p_\alpha = 1 \tag{10.17}$$

By definition, p_α's are time-independent. However, the states, $|\Psi_\alpha\rangle$'s, are generally time-dependent, satisfying the Schrödinger equation, $i\hbar\frac{\partial}{\partial t}|\Psi_\alpha\rangle = H|\Psi_\alpha\rangle$, where H is the Hamiltonian of the system.

10.4 Properties of the density matrix

1. The density matrix is a Hermitian operator. This is clear because p_α's are real numbers and each projection operator is a Hermitian: $(|\Psi\rangle\langle\Psi|)^\dagger = \langle\Psi|^\dagger|\Psi\rangle^\dagger = |\Psi\rangle\langle\Psi|$.

2. $|\Psi_\alpha\rangle$'s represent observable states, and so they are usually taken as eigenstates of an Hermitian operator such as Hamiltonian. Then, they must form a complete and orthonormal set.

$$\sum_\alpha |\Psi_\alpha\rangle\langle\Psi_\alpha| = 1 \tag{10.18}$$

$$\langle\Psi_\alpha|\Psi_{\alpha'}\rangle = \delta_{\alpha,\alpha'} \tag{10.19}$$

However, note that it is *not* required in general that in the definition of the density matrix, Eq. 10.16, $|\Psi_\alpha\rangle$'s are complete and ortho-normal. In general, they can be over-complete or non-orthogonal. As an example, this is a perfectly fine way of defining a mixed state for photons: 30 % of a pure x-polarization

[6] The following general points are worth keeping in mind, regarding this definition of the density matrix. (1) Eq. 10.16 represents an *incoherent sum* of pure states, each of which is given by $|\Psi_\alpha\rangle\langle\Psi_\alpha|$. One must keep in mind that not any state, which has a probability p_α to be measured in state Ψ_α, can be written in the form of Eq. 10.16. In general, non-diagonal elements may exist, if the system has coherence (cf. the last paragraph in page 5). (2) $|\Psi\rangle_\alpha$'s are taken to be normalized, but they *do not* need to be mutually orthogonal. Also, they can be over-complete. See property 2 in the next section.

state plus 50 % of a pure right circularly polarized state plus 20 % of a pure left circularly polarized state, i.e., $\rho = 0.3\,|h\rangle\langle h| + 0.5\,|R\rangle\langle R| + 0.2\,|L\rangle\langle L|$, in the notation of Section 10.2.1. Even in such a case, it is always possible to diagonalize ρ with a complete orthonormal basis set, since ρ itself is Hermitian! So, we can take the above two conditions to be true, always.

3. Just as in the semi-classical case, a **pure state** is defined as an ensemble consisting of just one micro state. A **mixed state** is an ensemble consisting of more than one possible micro states. Note that $\rho^2 = \rho$ if and only if a pure state describes the density matrix (proof left for your exercise).

4. For an observable O, the ensemble average is given by

$$\overline{\langle O\rangle} = \mathrm{tr}\{\rho O\} \tag{10.20}$$

Note that for quantum statistical mechanics, we use the symbol for ensemble average as $\overline{\langle O\rangle}$, to avoid misunderstanding as the mere expectation in quantum mechanics $\langle O\rangle$.

Proof ✎ `By definition,`

$$\overline{\langle O\rangle} = \sum_\gamma p_\gamma \langle \Psi_\gamma\, | O\, | \Psi_\gamma\rangle$$

`We need to show that this is identical with tr`$\{\rho O\}$. `This can be accomplished by inserting the resolution of the identify above (Eq. 10.18).`

$$\begin{aligned}
\sum_\gamma p_\gamma \langle \Psi_\gamma\, | O\, | \Psi_\gamma\rangle &= \sum_{\gamma,\alpha} p_\gamma \langle \Psi_\gamma | \Psi_\alpha\rangle \langle \Psi_\alpha | O\, | \Psi_\gamma\rangle \\
&= \sum_{\gamma,\alpha} p_\alpha \langle \Psi_\gamma | \Psi_\alpha\rangle \langle \Psi_\alpha | O\, | \Psi_\gamma\rangle && \texttt{using Eq. 10.19} \\
&= \sum_\gamma \langle \Psi_\gamma\, | \rho O\, | \Psi_\gamma\rangle && \texttt{Eq. 10.16} \\
&= \mathtt{tr}\{\rho O\}
\end{aligned}$$

QED. ✎

5. In the Heisenberg picture of quantum mechanics, matrices evolve in time, and state vectors do not. In that picture,

$$\frac{d\rho}{dt} = 0 \tag{10.21}$$

This is the direct equivalent to Eq. 5.15. This can be called the **quantum Liouville theorem**. This comes from the fact that p_α's are by definition time-independent, and that states are time-independent in the Heisenberg picture (and so is the projection operator for *any* state). Note that for any operator[7] O, the following **Heisenberg equation of motion** holds[8]

$$\frac{dO}{dt} = \frac{i}{\hbar}[H,O] + \frac{\partial O}{\partial t} \tag{10.22}$$

where $\frac{\partial O}{\partial t}$, defined as the Heisenberg picture form of the time derivative of O in the Schrödinger picture, is non-zero only if O has an explicit time dependence. Applying the Liouville theorem above, we get[9]

$$\frac{\partial \rho}{\partial t} = -\frac{i}{\hbar}[H,\rho] \tag{10.23}$$

[7]And even when H is time dependent, while a time dependent Hamiltonian is not of our main interest.

[8]In case some of you are not familiar with the Heisenberg picture, here is a reminder. In the Schrödinger picture, $U(t)\,|\Psi(t=0)\rangle = |\Psi(t)\rangle$, where $U(t)$ is the time evolution operator determined by the Schrödinger equation. Operators do not get affected by $U(t)$, although they can have explicit time dependence. In the Heisenberg picture, we regard states as fixed in time, i.e., as $|\Psi(t=0)\rangle$, while any operator $O_s(t)$ evolves in time as $O_h(t) = U^\dagger(t)O_s(t)U(t)$, where h means the Heisenberg picture and s means the Schrödinger picture. These two pictures are equivalent, since the matrix element $\langle\Psi_a\,|O|\,\Psi_b\rangle$ is independent of which picture we use. Here is the derivation of the Heisenberg equation of motion. <u>Derivation</u> ✎

$$\begin{aligned}
U(t+dt)\,|\Psi(t=0)\rangle &= |\Psi(t+dt)\rangle \approx |\Psi(t)\rangle + dt\frac{d}{dt}|\Psi(t)\rangle \\
&= \left(1-\frac{i}{\hbar}H\,dt\right)|\Psi(t)\rangle \\
&= \left(1-\frac{i}{\hbar}H\,dt\right)U(t)\,|\Psi(t=0)\rangle \qquad \texttt{for } \textit{any} \texttt{ state } \Psi(t=0) \\
\therefore\quad U(t+dt) &= \left(1-\frac{i}{\hbar}H\,dt\right)U(t) \\
\frac{dU(t)}{dt} &= -\frac{i}{\hbar}HU(t)
\end{aligned}$$

`where` $U(t)$ `is defined as the time evolution operator. This result is very general, valid even when` H *`is`* `time dependent. From` $O_h(t) = U^\dagger(t)O_s(t)U(t)$`,`

$$\begin{aligned}
\frac{dO_h}{dt} &= \frac{dU^\dagger}{dt}O_sU + U^\dagger\frac{dO_s}{dt}U + U^\dagger O_s\frac{dU}{dt} \\
&= \frac{i}{\hbar}U^\dagger[H,O_s]U + U^\dagger\frac{dO_s}{dt}U
\end{aligned}$$

`If` H`'s at different times commute, as is commonly the case,` U `and` H `commute, and the first term becomes` $\frac{i}{\hbar}[H,O_h]$`. If` H`'s at different times do not commute, then` $\frac{i}{\hbar}[H,O]$ `can be computed in the Schrödinger picture and then transformed to the Heisenberg picture. The second term is, by definition,` $\frac{\partial O_h}{\partial t}$`.` <u>QED.</u> ✐

[9]You may find the following statement in some books: $\frac{d\rho}{dt} = -\frac{i}{\hbar}[H,\rho]$. Such a statement is made within the Schrödinger picture. Here, we use the Heisenberg picture, as it compares well with the semi-classical statistical mechanics picture.

For any observable O, which does not depend explicitly on time, its time derivative is given by

$$\frac{d\overline{\langle O\rangle}}{dt} = \mathrm{tr}\left\{\frac{\partial\rho}{\partial t}O\right\} \tag{10.24}$$

assuming that $|\Psi_\alpha\rangle$'s are taken as eigenstates of the Hamiltonian, which we now assume to be not explicitly dependent on time. So, this suggests that

$$\frac{\partial\rho}{\partial t} = -\frac{i}{\hbar}[H,\rho] = 0 \tag{10.25}$$

for an equilibrium state. And, we take this as the fundamental postulate here, just as we did in semi-classical statistical mechanics (Section 6.1).

6. For a micro canonical ensemble in equilibrium,

$$p_{\mathrm{\mu}} = \frac{1}{\Omega(\mathcal{E},V,N)} \tag{10.26}$$

where $\Omega(\mathcal{E},V,N)$ is the number of micro states, i.e. the number of many body states (i.e., wave functions) $|\Psi_{\mathrm{\mu}}\rangle$. The entropy is given by[10]

$$S(E,V,N) = k_B \log\Omega(E,V,N) \tag{10.27}$$

For a canonical ensemble, we have

$$Z(T,V,N) = \sum_{\mathrm{\mu}}\exp(-\beta\mathcal{E}_{\mathrm{\mu}}) \qquad H|\Psi_\alpha\rangle = \mathcal{E}_{\mathrm{\mu}}|\Psi_{\mathrm{\mu}}\rangle \tag{10.28}$$

$$p_{\mathrm{\mu}} = \frac{\exp(-\beta\mathcal{E}_{\mathrm{\mu}})}{Z(T,V,N)} \tag{10.29}$$

For a grand canonical ensemble, we have

$$\Phi(T,V,\mu) = \sum_{\alpha}\exp(-\beta(\mathcal{E}_\alpha-\mathcal{N}\mu)) \tag{10.30}$$

$$p_\alpha = \frac{\exp(-\beta(\mathcal{E}_\alpha-\mathcal{N}\mu))}{\Phi(T,V,\mu)} \tag{10.31}$$

Here, we use the symbol α or $\mathrm{\mu}$ interchangeably, to mean the same thing – the index for microstates. **It represents the collection of quantum numbers that characterize the energy eigenstate $|\Psi_\alpha(\mathcal{N})\rangle$ for $\mathcal{N}$ particles.** We shall find that the grand canonical ensemble is of particular importance for quantum statistics. For this reason, α may be preferred to $\mathrm{\mu}$, in order to avoid confusion between symbols ($\mathrm{\mu}$ for microstate and μ for chemical potential).

The case of the Gibbs canonical ensemble is left for your exercise.

[10]Note the similarity of this definition to the Boltzmann entropy for the semi-classical case, Eq. 6.10. This definition can be extended to non-equilibrium states: we can define the von Neumann entropy, $S = -k_B\,\mathrm{tr}\{\rho\log\rho\}$. So, the von Neumann entropy is clearly the quantum version of the Gibbs-Boltzmann entropy, Eq. 7.17. Note that due to the Liouville theorem, both Eq. 7.17 and the von Neumann entropy are *time-independent*.

10.5 The partition function

As part of the discussion in the previous section, we have just now discussed the various ensembles and the associated partition functions.

In particular, the partition function is already given in Eq. 10.28. Since this is the first time to discuss it in quantum statistical mechanics, let us have a closer look at the way it arises from the density matrix formalism. From the consideration of the canonical ensemble in contact with a reservoir, we know that the probability associated with each microstate in a canonical ensemble is proportional to $\exp(-\beta\mathcal{E}_\alpha)$. So, the density matrix is given by

$$\rho = \sum_\alpha \frac{\exp(-\beta\mathcal{E}_\alpha)}{Z} |\Psi_\alpha\rangle \langle\Psi_\alpha| \tag{10.32}$$

where Z is defined as the probability normalization constant, as it should be. By taking the trace of the density matrix with the energy eigenstate basis $|\Psi_{\alpha'}\rangle$, we get

$$\begin{aligned} \mathrm{tr}\{\rho\} &= \frac{\sum_{\alpha,\alpha'} \exp(-\beta\mathcal{E}_\alpha) \langle \Psi_{\alpha'} | \Psi_\alpha\rangle \langle \Psi_\alpha | \Psi_{\alpha'}\rangle}{Z} && (10.33) \\ &= \frac{\sum_{\alpha,\alpha'} \exp(-\beta\mathcal{E}_\alpha)(\delta_{\alpha,\alpha'})^2}{Z} \qquad \text{by Eq. 10.19} && (10.34) \\ &= \frac{\sum_\alpha \exp(-\beta\mathcal{E}_\alpha)}{Z} && (10.35) \end{aligned}$$

Note that $\mathrm{tr}\{\rho\} = 1$, by definition, due to the probability sum rule. Or, if you like, by Eq. 10.20, $\mathrm{tr}\{\rho\}$ is the ensemble average of the identity operator, and it must be 1. Therefore, we get

$$Z(T,V,N) = \sum_\alpha \exp(-\beta\mathcal{E}_\alpha) \tag{10.36}$$

While this derivation is trivial, it does show the density matrix formalism at work. It also shows the density matrix is naturally a diagonal matrix in the energy basis. But, note that if we use other basis, say position basis, then it will no longer be diagonal.

10.6 An example – ideal gas, semi-classical view of quantum statistics

While the quantum mechanical problem of non-interacting particles (ideal gas) is most conveniently described in the grand canonical ensemble, studying the canonical ensemble for the same problem has some pedagogical content. So, let us spend a little time doing so.

Here, we do *not* ask what happens to the ideal gas at very low temperature or at very high density[11]. Instead, we ask what is the leading order correction to the semi-classical statistical mechanics due to the quantum statistics?

Let us recall that for calculating the partition function within the semi-classical statistical mechanics (Sections 6.3.4, 8.1), the key step was to consider the partition function of only one particle, as the result for the N particle case is equal to the N-th power of one particle partition function divided by $N!$ (Section 8.1). Then, it might not come as a surprise for you that to answer the above question that we posed, it is sufficient to consider a two particle density matrix and a two particle partition function.

$$\rho_2 = \frac{\sum_\alpha e^{-\beta\mathcal{E}_\alpha} |\mathcal{E}_\alpha\rangle \langle\mathcal{E}_\alpha|}{Z_2} \tag{10.37}$$

$$Z_2 = \sum_\alpha e^{-\beta\mathcal{E}_\alpha} \tag{10.38}$$

We consider two fermions or two bosons. The quantum number, α, for two free particles is given by $\vec{k}_1, \vec{k}_2, s_1, s_2$, where $\vec{k}_1$, $\vec{k}_2$, are wave vectors, and s_1, s_2 are spin quantum numbers. We will assume that $s_1 = s_2$ throughout this section. In this spin subspace, we do not need to consider spin quantum numbers. Then, we have

$$\mathcal{E}_\alpha = \frac{\hbar(k_1^2 + k_2^2)}{2m} \tag{10.39}$$

$$|\mathcal{E}_\alpha\rangle = \frac{1}{\sqrt{2}}\left(\left|\vec{k}_1, \vec{k}_2\right\rangle + \eta\left|\vec{k}_2, \vec{k}_1\right\rangle\right) \qquad \vec{k}_1 \neq \vec{k}_2,\ \eta = 1 \text{ for boson, } -1 \text{ for fermion} \tag{10.40}$$

$$= \left|\vec{k}_1, \vec{k}_2\right\rangle \qquad \vec{k}_1 = \vec{k}_2,\ \text{boson} \tag{10.41}$$

$$= 0 \text{ (no such state)} \qquad \vec{k}_1 = \vec{k}_2,\ \text{fermion} \tag{10.42}$$

where the exchange symmetry of the two particle case is explicitly taken care of in the last three equations.

Notice that, if we simply ignore the case $\vec{k}_1 = \vec{k}_2$, then the calculation of the partition function becomes very simple, and becomes identical with the classical result. In general, however, we have quite a complicated problem, here, due to the exchange

[11] Asking such a question means considering the so-called **quantum degenerate regime**, i.e., the regime of temperature/density where the quantum statistics becomes singularly important. We have already encountered such a regime in Section 9.1 (and Section 8.5), where the Bose-Einstein statistics has the full effect. The quantum degenerate regime is a topic that we will pay our full attention to in later lectures.

symmetry. To deal with the exchange symmetry, it is convenient to introduce the **occupation number**, $n_{\vec{k}}$. Then we get the following results, which are valid for general values of the number of particles, N, while we are primarily interested in the $N=2$ case, in this section.

$$\begin{aligned}
\alpha &\equiv \{n_{\vec{k}}\} \qquad \sum_{\vec{k}} n_{\vec{k}} = N \qquad n_{\vec{k}} = 0,1 \text{ (fermion)},\ 0,1,2,\ldots \text{ (boson)} && (10.43)\\
\mathcal{E}_\alpha &= \sum_{\vec{k}} \frac{\hbar k^2}{2m} n_{\vec{k}} && (10.44)\\
|\mathcal{E}_\alpha\rangle &= \frac{1}{\sqrt{N!\prod_{\vec{k}} n_{\vec{k}}!}} \sum_P \eta^{s(P)} \left|P(\vec{\mathbf{k}})\right\rangle && (10.45)
\end{aligned}$$

Here, $\vec{\mathbf{k}}$ is the set of N wave vectors, some repeating $n_{\vec{k}}$ times if $n_{\vec{k}} > 1$, for which $n_{\vec{k}} > 0$. P is the permutation of those wave vectors. The function $s(P)$ is defined as the "sign of the permutation": it gives 1 if the permutation involves an even number of exchanges, and -1 if the permutation involves an odd number of exchanges. $\eta = 1$ for boson, and -1 for fermion, as already introduced above. Eq. 10.45 is written for general N (proof left for your exercise), while it is easy to see that our explicit wave functions given above (Eqs. 10.40,10.41,10.42) are consistent with it.

While the expression for the partition function seems simple, $Z = \sum_\alpha e^{-\beta\mathcal{E}_\alpha}$, it is not simple to evaluate it, since the α index is *not* simple, being highly constrained by the exchange symmetry.

Let us look at the density matrix ρ in the space representation, to figure out how we might achieve the goal of calculating the partition function. We shall carry out this calculation for the $N = 2$ case, while the extension to the general N case will become clear.

$$\begin{aligned}
\langle \vec{r}_1, \vec{r}_2 \,|\rho|\, \vec{r}\,'_1, \vec{r}\,'_2\rangle &= \sum_{\{\vec{k}_1,\vec{k}_2\}} \frac{e^{-\beta\mathcal{E}_\alpha}}{Z_2} \langle \vec{r}_1, \vec{r}_2 \,|\, \mathcal{E}_\alpha\rangle \langle \mathcal{E}_\alpha \,|\, \vec{r}\,'_1, \vec{r}\,'_2\rangle\\
&= \sum_{\vec{k}_1}\sum_{\vec{k}_2} \frac{n_{\vec{k}_1}! n_{\vec{k}_2}!}{2!} \frac{e^{-\beta\mathcal{E}_\alpha}}{Z_2} \langle \vec{r}_1, \vec{r}_2 \,|\, \mathcal{E}_\alpha\rangle \langle \mathcal{E}_\alpha \,|\, \vec{r}\,'_1, \vec{r}\,'_2\rangle\\
&= \sum_{\vec{k}_1}\sum_{\vec{k}_2}\sum_{P,P'} \frac{\eta^{s(P)+s(P')}}{(2!)^2} \frac{e^{-\beta\mathcal{E}_\alpha}}{Z_2} \left\langle \vec{r}_1, \vec{r}_2 \,|\, P(\vec{k}_1, \vec{k}_2)\right\rangle \left\langle P'(\vec{k}_1, \vec{k}_2) \,\middle|\, \vec{r}\,'_1, \vec{r}\,'_2\right\rangle
\end{aligned}$$

Here in the first line, the sum over $\{\vec{k}_1, \vec{k}_2\}$ is restricted to a unique set of two (possibly the same, for bosons) wave vectors, while from the second line on the sums over $\vec{k}_1$ and $\vec{k}_2$ are unrestricted. The over-counting introduced by this change in sum is taken care of by the division by the multinomial coefficient, $N!/\prod_{\vec{k}} n_{\vec{k}}$. To get the last equation, Eq. 10.45 (for $N = 2$) has been used.

It is easy to consider all permutations of two wave vectors.

$$\begin{aligned}\langle\vec{r}_1,\vec{r}_2\,|\rho|\,\vec{r}_1',\vec{r}_2'\rangle &= \frac{1}{(2!)^2}\sum_{\vec{k}_1}\sum_{\vec{k}_2}\frac{e^{-\beta\mathcal{E}_\alpha}}{Z_2}\Big[\left\langle\vec{r}_1,\vec{r}_2\,|\,\vec{k}_1,\vec{k}_2\right\rangle\left\langle\vec{k}_1,\vec{k}_2\,\middle|\,\vec{r}_1',\vec{r}_2'\right\rangle\\ &+\left\langle\vec{r}_1,\vec{r}_2\,|\,\vec{k}_2,\vec{k}_1\right\rangle\left\langle\vec{k}_2,\vec{k}_1\,\middle|\,\vec{r}_1',\vec{r}_2'\right\rangle\\ &+\eta\left\langle\vec{r}_1,\vec{r}_2\,|\,\vec{k}_2,\vec{k}_1\right\rangle\left\langle\vec{k}_1,\vec{k}_2\,\middle|\,\vec{r}_1',\vec{r}_2'\right\rangle\\ &+\eta\left\langle\vec{r}_1,\vec{r}_2\,|\,\vec{k}_1,\vec{k}_2\right\rangle\left\langle\vec{k}_2,\vec{k}_1\,\middle|\,\vec{r}_1',\vec{r}_2'\right\rangle\Big]\end{aligned}$$

Note that

$$\left\langle\vec{r}_1,\vec{r}_2\,|\,\vec{k}_1,\vec{k}_2\right\rangle = \frac{e^{i\vec{k}_1\cdot\vec{r}_1}e^{i\vec{k}_2\cdot\vec{r}_2}}{V} \tag{10.46}$$

where the one body wave function is normalized within the volume of V, with the wave function normalization factor $\frac{1}{\sqrt{V}}$. Inserting this wave function to the above expression for ρ, and also noting that $\vec{k}_1$ and $\vec{k}_2$ are dummy indices that can be swapped, we get

$$\langle\vec{r}_1,\vec{r}_2\,|\rho|\,\vec{r}_1',\vec{r}_2'\rangle = \frac{2}{(2!V)^2}\sum_{\vec{k}_1}\sum_{\vec{k}_2}\frac{e^{-\beta\mathcal{E}_\alpha}}{Z_2}\Big[e^{i\vec{k}_1\cdot(\vec{r}_1-\vec{r}_1')}e^{i\vec{k}_2\cdot(\vec{r}_2-\vec{r}_2')}+\eta e^{i\vec{k}_1\cdot(\vec{r}_1-\vec{r}_2')}e^{i\vec{k}_2\cdot(\vec{r}_2-\vec{r}_1')}\Big]$$

Since $\mathcal{E}_\alpha = \frac{\hbar(k_1^2+k_2^2)}{2m}$, this expression requires us to evaluate one summation

$$I = \sum_{\vec{k}}\exp\left(-\beta\frac{\hbar^2k^2}{2m}+i\vec{k}\cdot\vec{q}\right)$$

Note that, using a periodic boundary condition for a cubic volume, V, we get

$$\vec{k} = \frac{2\pi}{L}m_i \qquad\qquad m_i = \text{integer};\ i = x,y,z \tag{10.47}$$

Thus,

$$\sum_{\vec{k}} = \sum_{k_x,k_y,k_z} = \frac{V}{(2\pi)^3}\int d^3\vec{k} \tag{10.48}$$

where the volume of the system is given by $V = L^3$. Thus, the above summation becomes

$$\begin{aligned}I &= \frac{V}{(2\pi)^3}\int dk_xdk_ydk_z\ \exp\left(-\beta\frac{\hbar^2k^2}{2m}+i\vec{k}\cdot\vec{q}\right)\\ &= \frac{V}{(2\pi)^3}\left(\sqrt{\frac{2m\pi k_BT}{\hbar^2}}\right)^3\ \exp\left(-\frac{mk_BT}{2\hbar^2}q^2\right)\\ &= V\lambda^{-3}\exp\left(-\frac{\pi}{\lambda^2}q^2\right)\end{aligned}$$

where λ is the thermal de Broglie wave length defined in Eq. 6.43. Using this, we get

$$\begin{aligned}\langle \vec{r}_1, \vec{r}_2 \,|\rho|\, \vec{r}\,'_1, \vec{r}\,'_2\rangle &= \frac{1}{2! Z_2 \lambda^6}\left\{\exp\left(-\frac{\pi}{\lambda^2}\left[(\vec{r}_1 - \vec{r}\,'_1)^2 + (\vec{r}_2 - \vec{r}\,'_2)^2\right]\right)\right.\\ &\left. + \eta \exp\left(-\frac{\pi}{\lambda^2}\left[(\vec{r}_1 - \vec{r}\,'_2)^2 + (\vec{r}_2 - \vec{r}\,'_1)^2\right]\right)\right\} \qquad (10.49)\end{aligned}$$

Note that the density matrix is not diagonal in the position basis. From $\mathrm{tr}\{\rho\} = 1$, we get

$$Z_2 = \frac{1}{2! \lambda^6} \int d^3\vec{r}_1 d^3\vec{r}_2 \left(1 + \eta \exp\left[-\frac{2\pi}{\lambda^2}(\vec{r}_1 - \vec{r}_2)^2\right]\right) \qquad (10.50)$$

These two results are easily generalized to the arbitrary N case as

$$\langle \vec{r}_1, ..., \vec{r}_N \,|\rho|\, \vec{r}\,'_1, ..., \vec{r}\,'_N\rangle = \frac{1}{Z\, N! \lambda^{3N}} \sum_P \eta^{s(P)} \exp\left(-\frac{\pi}{\lambda^2}\sum_{i=1}^N (\vec{r}_i - \vec{r}\,'_{P(i)})^2\right) \qquad (10.51)$$

$$Z = \frac{1}{N! \lambda^{3N}} \int \prod_i d^3\vec{r}_i \sum_P \eta^{s(P)} \exp\left(-\frac{\pi}{\lambda^2}\sum_{i=1}^N (\vec{r}_i - \vec{r}_{P(i)})^2\right) \qquad (10.52)$$

Let us look at Z_2 in more detail.

$$Z_2 = \frac{1}{2! \lambda^6}\left[V^2 + V\eta\left(\frac{\lambda}{\sqrt{2}}\right)^3\right] = \frac{V^2}{2!\lambda^6}\left[1 + \eta \frac{\lambda^3}{V 2\sqrt{2}}\right] \qquad (10.53)$$

The first term corresponds to the semi-classical result $Z = \frac{V^N}{N!\lambda^{3N}} = \left(\frac{e}{n\lambda^3}\right)^N$, which results from the case when the permuation is the identity transformation (Eq. 10.52). As the indices are permuted, some $V \to \lambda^3$. Therefore, generally,

$$\begin{aligned} Z &= \left(\frac{e}{n\lambda^3}\right)^N \left[1 + a_1(n\lambda^3) + a_2(n\lambda^3)^2 + ... + a_{N-1}(n\lambda^3)^{N-1}\right] \qquad (10.54)\\ &= \left(\frac{e}{n\lambda^3}\right)^N \left[1 + \eta \frac{N-1}{2^{5/2}} n\lambda^3 + \ ...\ \right] \qquad (10.55)\end{aligned}$$

While what we have done so far is valid for any temperature or any density, clearly, they are most useful when $n\lambda^3 \ll 1$. In fact, we have just proven that in the limit of $n\lambda^3 \to 1$, we recover the semi-classical result!

Let us look at the pressure (for large N)

$$\begin{aligned} P &= -\left(\frac{\partial F}{\partial V}\right)_T = k_B T\left(\frac{\partial \log Z}{\partial V}\right)_T = k_B T n - \eta \frac{n^2}{2^{5/2}}\lambda^3 + ...\\ &= n k_B T\left[1 - \eta\frac{n\lambda^3}{2^{5/2}} + \ ...\ \right] \qquad (10.56)\end{aligned}$$

This implies that the first quantum correction to pressure is positive for fermions and negative for bosons. Thus, one might say that fermions tend to repel one another, while bosons tend to attract one another, *just due to statistics*.

This can be verified by examining the $N = 2$ density matrix, Eq. 10.49. The diagonal terms are given by

$$\langle \vec{r}_1, \vec{r}_2 \,|\rho|\, \vec{r}_1, \vec{r}_2 \rangle = \frac{1}{2! Z_2 \lambda^6} \left\{ 1 + \eta \exp\left(-\frac{2\pi r^2}{\lambda^2} \right) \right\} \qquad r \equiv |\vec{r}_1 - \vec{r}_2| \tag{10.57}$$

This shows that the probability to find two particles at distance r is (1) uniform for semi-classical ideal gas, (2) enhanced for bosons ($\eta = 1$), and (3) reduced for fermions ($\eta < -1$). The effect is, as expected, significant only for distance on the order of λ or smaller.

More quantitatively, one can interpret the above density matrix as occurring due to an effective two-body potential, $U_{eff}(r)$, such that

$$\langle \vec{r}_1, \vec{r}_2 \,|\rho|\, \vec{r}_1, \vec{r}_2 \rangle = \frac{1}{Z_2 2! \lambda^6} \exp\left(-\beta U_{eff}(r)\right) \tag{10.58}$$

$$U_{eff}(r) \equiv -k_B T \log \left\{ 1 + \eta \exp\left(-\frac{2\pi r^2}{\lambda^2} \right) \right\} \tag{10.59}$$

See Figure T5.1 for the plot of this effective potential, which nicely summarizes the quantum statistics as effective interaction: attractive for bosons, and repulsive for fermions.

10.7 Rotational mode

A quantum rigid body has the Hamiltonian

$$H_{rot} = \frac{\vec{L}^2}{2I} \tag{10.60}$$

where I is the rotational inertia, and $\vec{L}$ is the angular momentum.

The analysis of this problem (see below) with the quantized angular momentum proceeds much like for the single oscillator, except that the partition function cannot be summed exactly. Nevertheless, the (constant volume) heat capacity shows the same qualitative behavior – it is exponentially suppressed at low T, while it goes to the classical value k_B at high T.

The characteristic temperature scale for the quantum rotation can be defined as

$$\theta_{rot} = \frac{\hbar^2}{2Ik_B} \tag{10.61}$$

Using k_B = 25.85 meV / 300 K, $\hbar c$ = 1973 eV Å, and using typical values for Ic^2 (keeping in mind that $I = Ma^2$, where M is the atomic mass and a is the atomic dimension), one can estimate that $\theta_{rot} \sim 1$ K. So, the crossover from the quantum regime to the classical regime occurs at very low temperature in this case, quite different from the vibrational case, in which the crossover to the classical regime happens at room temperature or higher.

$$Z = \sum_{l=0}^{\infty}(2l+1)\exp\left(-\beta\frac{l(l+1)\hbar^2}{2I}\right) = \sum_{l=0}^{\infty}(2l+1)\exp\left(-\frac{\theta_{rot}}{T}l(l+1)\right) \tag{10.62}$$

In the classical regime, the l sum can be turned into an integral, with the continuous variable $x \equiv l(l+1)$. Then, we get

$$Z \approx \int_0^{\infty} dx \exp\left(-\frac{\theta_{rot}}{T}x\right) = \frac{T}{\theta_{rot}} \qquad T \gg \theta_{rot} \tag{10.63}$$

$$E \approx -\frac{\partial \log Z}{\partial \beta} = k_B T \qquad T \gg \theta_{rot} \tag{10.64}$$

$$C_V \approx k_B \qquad T \gg \theta_{rot} \tag{10.65}$$

The latter two results are as expected from the equipartition theorem. The rotational kinetic energy is given by $K\frac{I}{2}\dot{\theta}^2 + \frac{I}{2}\sin^2\theta\,\dot{\phi}^2 = L$, from which the Hamiltonian can be written as $H_{rot} = \frac{p_\theta^2}{2I} + \frac{p_\phi^2}{2I\sin^2\theta}$, with $p_\theta = \frac{\partial L}{\partial\dot{\theta}} = I\dot{\theta}$ and $p_\phi = \frac{\partial L}{\partial\dot{\phi}} = I\sin^2\theta\,\dot{\phi}$. Thus, the Hamiltonian consists of two quadratic terms, giving rise to the equipartition energy $\frac{k_BT}{2} \times 2$.

In the quantum regime, the partition function is can be approximated with the first few terms.

$$Z \approx 1 + 3\exp\left(-\frac{2\theta_{rot}}{T}\right) \qquad T \ll \theta_{rot} \tag{10.66}$$

$$E \approx 6k_B\theta_{rot}\exp\left(-\frac{2\theta_{rot}}{T}\right) \qquad T \ll \theta_{rot} \tag{10.67}$$

$$C_V \approx 3k_B\left(\frac{2\theta_{rot}}{T}\right)^2\exp\left(-\frac{2\theta_{rot}}{T}\right) \qquad T \ll \theta_{rot} \tag{10.68}$$