G25.2651: Statistical Mechanics

Notes for Lecture 11

I. PRINCIPLES OF QUANTUM STATISTICAL MECHANICS

The problem of quantum statistical mechanics is the quantum mechanical treatment of an N-particle system. Suppose the corresponding N-particle classical system has Cartesian coordinates

$$q_1, ..., q_{3N}$$

and momenta

$$p_1, ..., p_{3N}$$

and Hamiltonian

$$H = \sum_{i=1}^{3N} \frac{p_i^2}{2m_i} + U(q_1, ..., q_{3N})$$

Then, as we have seen, the quantum mechanical problem consists of determining the state vector $|\Psi(t)\rangle$ from the Schrödinger equation

$$H|\Psi(t)\rangle = i\hbar \frac{\partial}{\partial t}|\Psi(t)\rangle$$

Denoting the corresponding operators, $Q_1, ..., Q_{3N}$ and $P_1, ..., P_{3N}$, we note that these operators satisfy the commutation relations:

$$[Q_i, Q_j] = [P_i, P_j] = 0$$
$$[Q_i, P_j] = i\hbar I \delta_{ij}$$

and the many-particle coordinate eigenstate $|q_1...q_{3N}\rangle$ is a tensor product of the individual eigenstate $|q_1\rangle, ..., |q_{3N}\rangle$:

$$|q_1...q_{3N}\rangle = |q_1\rangle \cdots |q_{3N}\rangle$$

The Schrödinger equation can be cast as a partial differential equation by multiplying both sides by $\langle q_1...q_{3N}|$:

$$\langle q_1...q_{3N}|H|\Psi(t)\rangle = i\hbar \frac{\partial}{\partial t} \langle q_1...q_{3N}|\Psi(t)\rangle$$

$$\left[-\sum_{i=1}^{3N} \frac{\hbar^2}{2m_i} \frac{\partial^2}{\partial q_i^2} + U(q_1, ..., q_{3N}) \right] \Psi(q_1, ..., q_{3N}, t) = i\hbar \frac{\partial}{\partial t} \Psi(q_1, ..., q_{3N}, t)$$

where the many-particle wave function is $\Psi(q_1, ..., q_{3N}, t) = \langle q_1...q_{3N}|\Psi(t)\rangle$. Similarly, the expectation value of an operator $A = A(Q_1, ..., Q_{3N}, P_1, ..., P_{3N})$ is given by

$$\langle A \rangle = \int dq_1 \cdots dq_{3N} \Psi^*(q_1, ..., q_{3N}) A\left(q_1, ..., q_{3N}, \frac{\hbar}{i} \frac{\partial}{\partial q_1}, ..., \frac{\hbar}{i} \frac{\partial}{\partial q_{3N}}\right) \Psi(q_1, ..., q_{3N})$$

A. The density matrix and density operator

In general, the many-body wave function $\Psi(q_1, ..., q_{3N}, t)$ is far too large to calculate for a macroscopic system. If we wish to represent it on a grid with just 10 points along each coordinate direction, then for $N = 10^{23}$, we would need

$$10^{10^{23}}$$

total points, which is clearly enormous.

We wish, therefore, to use the concept of ensembles in order to express expectation values of observables $\langle A \rangle$ without requiring direct computation of the wavefunction. Let us, therefore, introduce an ensemble of systems, with a total of Z members, and each having a state vector $|\Psi^{(\alpha)}\rangle$, $\alpha = 1, ..., Z$. Furthermore, introduce an orthonormal set of vectors $|\phi_k\rangle$ ($\langle \phi_k | \phi_j \rangle = \delta_{ij}$) and expand the state vector for each member of the ensemble in this orthonormal set:

$$|\Psi^{(\alpha)}\rangle = \sum_k C_k^{(\alpha)} |\phi_k\rangle$$

The expectation value of an observable, averaged over the ensemble of systems is given by the average of the expectation value of the observable computed with respect to each member of the ensemble:

$$\langle A \rangle = \frac{1}{Z} \sum_{\alpha=1}^{Z} \langle \Psi^{(\alpha)} | A | \Psi^{(\alpha)} \rangle$$

Substituting in the expansion for $|\Psi^{(\alpha)}\rangle$, we obtain

$$\langle A \rangle = \frac{1}{Z} \sum_{k,l} C_k^{(\alpha)^*} C_l^{(\alpha)} \langle \phi_k | A | \phi_l \rangle$$

$$= \sum_{k,l} \left(\frac{1}{Z} \sum_{\alpha=1}^{Z} C_l^{(\alpha)} C_k^{(\alpha)^*} \right) \langle \phi_k | A | \phi_l \rangle$$

Let us define a matrix

$$\rho_{lk} = \sum_{\alpha=1}^{Z} C_l^{(\alpha)} C_k^{(\alpha)^*}$$

and a similar matrix

$$\tilde{\rho}_{lk} = \frac{1}{Z} \sum_{\alpha=1}^{Z} C_l^{(\alpha)} C_k^{(\alpha)^*}$$

Thus, ρ_{lk} is a sum over the ensemble members of a product of expansion coefficients, while $\tilde{\rho}_{lk}$ is an average over the ensemble of this product. Also, let $A_{kl} = \langle \phi_k | A | \phi_l \rangle$. Then, the expectation value can be written as follows:

$$\langle A \rangle = \frac{1}{Z} \sum_{k,l} \rho_{lk} A_{kl} = \frac{1}{Z} \sum_k (\rho A)_{kk} = \frac{1}{Z} \text{Tr}(\rho A) = \text{Tr}(\tilde{\rho} A)$$

where ρ and A represent the matrices with elements ρ_{lk} and A_{kl} in the basis of vectors $\{|\phi_k\rangle\}$. The matrix ρ_{lk} is known as the *density matrix*. There is an abstract operator corresponding to this matrix that is basis-independent. It can be seen that the operator

$$\rho = \sum_{\alpha=1}^{Z} |\Psi^{(\alpha)}\rangle \langle \Psi^{(\alpha)}|$$

and similarly

$$\tilde{\rho} = \frac{1}{Z} \sum_{\alpha=1}^{Z} |\Psi^{(\alpha)}\rangle \langle \Psi^{(\alpha)}|$$

have matrix elements ρ_{lk} when evaluated in the basis set of vectors $\{|\phi_k\rangle\}$.

$$\langle\phi_l|\rho|\phi_k\rangle = \sum_{\alpha=1}^{Z}\langle\phi_l|\Psi^{(\alpha)}\rangle\langle\Psi^{(\alpha)}|\phi_k\rangle = \sum_{\alpha=1}^{Z} C_l^{(\alpha)} C_k^{(\alpha)^*} = \rho_{lk}$$

Note that ρ is a hermitian operator

$$\rho^\dagger = \rho$$

so that its eigenvectors form a complete orthonormal set of vectors that span the Hilbert space. If w_k and $|w_k\rangle$ represent the eigenvalues and eigenvectors of the operator $\tilde{\rho}$, respectively, then several important properties they must satisfy can be deduced.

Firstly, let A be the identity operator I. Then, since $\langle I \rangle = 1$, it follows that

$$1 = \frac{1}{Z}\mathrm{Tr}(\rho) = \mathrm{Tr}(\tilde{\rho}) = \sum_k w_k$$

Thus, the eigenvalues of $\tilde{\rho}$ must sum to 1. Next, let A be a projector onto an eigenstate of $\tilde{\rho}$, $A = |w_k\rangle\langle w_k| \equiv P_k$. Then

$$\langle P_k \rangle = \mathrm{Tr}(\tilde{\rho}|w_k\rangle\langle w_k|)$$

But, since $\tilde{\rho}$ can be expressed as

$$\tilde{\rho} = \sum_k w_k |w_k\rangle\langle w_k|$$

and the trace, being basis set independent, can be therefore be evaluated in the basis of eigenvectors of $\tilde{\rho}$, the expectation value becomes

$$\langle P_k \rangle = \sum_j \langle w_j| \sum_i w_i |w_i\rangle\langle w_i|w_k\rangle\langle w_k|w_j\rangle$$
$$= \sum_{i,j} w_i \delta_{ij}\delta_{ik}\delta_{kj}$$
$$= w_k$$

However,

$$\langle P_k \rangle = \frac{1}{Z}\sum_{\alpha=1}^{Z}\langle\Psi^{(\alpha)}|w_k\rangle\langle w_k|\Psi^{(\alpha)}\rangle$$
$$= \frac{1}{Z}\sum_{\alpha=1}^{Z}|\langle\Psi^{(\alpha)}|w_k\rangle|^2 \geq 0$$

Thus, $w_k \geq 0$. Combining these two results, we see that, since $\sum_k w_k = 1$ and $w_k \geq 0$, $0 \leq w_k \leq 1$, so that w_k satisfy the properties of probabilities.

With this in mind, we can develop a physical meaning for the density matrix. Let us now consider the expectation value of a projector $|a_i\rangle\langle a_i| \equiv \mathcal{P}_{a_i}$ onto one of the eigenstates of the operator A. The expectation value of this operator is given by

$$\langle\mathcal{P}_{a_i}\rangle = \frac{1}{Z}\sum_{\alpha=1}^{Z}\langle\Psi^{(\alpha)}|\mathcal{P}_{a_i}|\Psi^{(\alpha)}\rangle = \frac{1}{Z}\sum_{\alpha=1}^{Z}\langle\Psi^{(\alpha)}|a_i\rangle\langle a_i|\Psi^{(\alpha)}\rangle = \frac{1}{Z}\sum_{\alpha=1}^{Z}|\langle a_i|\Psi^{(\alpha)}\rangle|^2$$

But $|\langle a_i|\Psi^{(\alpha)}\rangle|^2 \equiv P_{a_i}^{(\alpha)}$ is just probability that a measurement of the operator A in the αth member of the ensemble will yield the result a_i. Thus,

$$\langle\mathcal{P}_{a_i}\rangle = \frac{1}{Z}\sum_{\alpha=1}^{P} P_{a_i}^{(\alpha)}$$

or the expectation value of $\mathcal{P}_{a_i}$ is just the ensemble averaged probability of obtaining the value a_i in each member of the ensemble. However, note that the expectation value of $\mathcal{P}_{a_i}$ can also be written as

$$\langle \mathcal{P}_{a_i} \rangle = \text{Tr}(\tilde{\rho}\mathcal{P}_{a_i}) = \text{Tr}(\sum_k w_k |w_k\rangle\langle w_k|a_i\rangle\langle a_i|) = \sum_{k,l} \langle w_l|w_k|w_k\rangle\langle w_k|a_i\rangle\langle a_i|w_l\rangle$$

$$= \sum_{k,l} w_k \delta_{kl} \langle w_k a_i\rangle\langle a_i w_l\rangle$$

$$= \sum_k w_k |\langle a_i|w_k\rangle|^2$$

Equating the two expressions gives

$$\frac{1}{Z}\sum_{\alpha=1}^{Z} \langle P_{a_i}^{(\alpha)}\rangle = \sum_k w_k |\langle a_i|w_k\rangle|^2$$

The interpretation of this equation is that the ensemble averaged probability of obtaining the value a_i if A is measured is equal to the probability of obtaining the value a_i in a measurement of A if the state of the system under consideration were the state $|w_k\rangle$, weighted by the average probability w_k that the system in the ensemble is in that state. Therefore, the density operator ρ (or $\tilde{\rho}$) plays the same role in quantum systems that the phase space distribution function $f(\mathbf{\Gamma})$ plays in classical systems.

B. Time evolution of the density operator

The time evolution of the operator ρ can be predicted directly from the Schrödinger equation. Since $\rho(t)$ is given by

$$\rho(t) = \sum_{\alpha=1}^{Z} |\Psi^{(\alpha)}(t)\rangle\langle\Psi^{(\alpha)}(t)|$$

the time derivative is given by

$$\frac{\partial\rho}{\partial t} = \sum_{\alpha=1}^{Z} \left[\left(\frac{\partial}{\partial t}|\Psi^{(\alpha)}(t)\rangle\right)\langle\Psi^{(\alpha)}(t)| + |\Psi^{(\alpha)}(t)\rangle\left(\frac{\partial}{\partial t}\langle\Psi^{(\alpha)}(t)|\right) \right]$$

$$= \frac{1}{i\hbar}\sum_{\alpha=1}^{Z} \left[\left(H|\Psi^{(\alpha)}(t)\rangle\right)\langle\Psi^{(\alpha)}(t)| - |\Psi^{(\alpha)}(t)\langle\left(\langle\Psi^{(\alpha)}(t)|H\right) \right]$$

$$= \frac{1}{i\hbar}(H\rho - \rho H)$$

$$= \frac{1}{i\hbar}[H,\rho]$$

$$\frac{\partial\rho}{\partial t} = \frac{1}{i\hbar}[H,\rho]$$

where the second line follows from the fact that the Schrödinger equation for the bra state vector $\langle\Psi^{(\alpha)}(t)|$ is

$$-i\hbar\frac{\partial}{\partial t}\langle\Psi^{(\alpha)}(t)| = \langle\Psi^{(\alpha)}(t)|H$$

Note that the equation of motion for $\rho(t)$ differs from the usual Heisenberg equation by a minus sign! Since $\rho(t)$ is constructed from state vectors, it is not an observable like other hermitian operators, so there is no reason to expect that its time evolution will be the same. The general solution to *its* equation of motion is

$$\rho(t) = e^{-iHt/\hbar}\rho(0)e^{iHt/\hbar}$$

$$= U(t)\rho(0)U^\dagger(t)$$

The equation of motion for $\rho(t)$ can be cast into a quantum Liouville equation by introducing an operator

$$iL = \frac{1}{i\hbar}[..., H]$$

In term of iL, it can be seen that $\rho(t)$ satisfies

$$\frac{\partial \rho}{\partial t} = -iL\rho$$
$$\rho(t) = e^{-iLt}\rho(0)$$

What kind of operator is iL? It acts on an operator and returns another operator. Thus, it is not an operator in the ordinary sense, but is known as a *superoperator* or *tetradic operator* (see S. Mukamel, *Principles of Nonlinear Optical Spectroscopy*, Oxford University Press, New York (1995)).

Defining the evolution equation for ρ this way, we have a perfect analogy between the density matrix and the state vector. The two equations of motion are

$$\frac{\partial}{\partial t}|\Psi(t)\rangle = -\frac{i}{\hbar}H|\Psi(t)\rangle$$
$$\frac{\partial}{\partial t}\rho(t) = -iL\rho(t)$$

We also have an analogy with the evolution of the classical phase space distribution $f(\mathbf{\Gamma}, t)$, which satisfies

$$\frac{\partial f}{\partial t} = -iLf$$

with $iL = \{..., H\}$ being the classical Liouville operator. Again, we see that the limit of a commutator is the classical Poisson bracket.

C. The quantum equilibrium ensembles

At equilibrium, the density operator does not evolve in time; thus, $\partial\rho/\partial t = 0$. Thus, from the equation of motion, if this holds, then $[H, \rho] = 0$, and $\rho(t)$ is a constant of the motion. This means that it can be simultaneously diagonalized with the Hamiltonian and can be expressed as a pure function of the Hamiltonian

$$\rho = f(H)$$

Therefore, the eigenstates of ρ, the vectors, we called $|w_k\rangle$ are the eigenvectors $|E_i\rangle$ of the Hamiltonian, and we can write H and ρ as

$$H = \sum_i E_i |E_i\rangle\langle E_i|$$
$$\rho = \sum_i f(E_i)|E_i\rangle\langle E_i|$$

The choice of the function f determines the ensemble.

1. The microcanonical ensemble

Although we will have practically no occasion to use the quantum microcanonical ensemble (we relied on it more heavily in classical statistical mechanics), for completeness, we define it here. The function f, for this ensemble, is

$$f(E_i)\delta E = \theta(E_i - (E + \delta E)) - \theta(E_i - E)$$

where $\theta(x)$ is the Heaviside step function. This says that $f(E_i)\delta E$ is 1 if $E < E_i < (E + \delta E)$ and 0 otherwise. The partition function for the ensemble is $\text{Tr}(\rho)$, since the trace of ρ is the number of members in the ensemble:

$$\Omega(N, V, E) = \text{Tr}(\rho) = \sum_i [\theta(E_i - (E + \delta E)) - \theta(E_i - E)]$$

The thermodynamics that are derived from this partition function are exactly the same as they are in the classical case:

$$S(N, V, E) = -k \ln \Omega(N, V, E)$$
$$\frac{1}{T} = -k \left(\frac{\partial \ln \Omega}{\partial E} \right)_{N,V}$$

etc.

2. The canonical ensemble

In analogy to the classical canonical ensemble, the quantum canonical ensemble is defined by

$$\rho = e^{-\beta H}$$
$$f(E_i) = e^{-\beta E_i}$$

Thus, the quantum canonical partition function is given by

$$Q(N, V, T) = \text{Tr}(e^{-\beta H}) = \sum_i e^{-\beta E_i}$$

and the thermodynamics derived from it are the same as in the classical case:

$$A(N, V, T) = -\frac{1}{\beta} \ln Q(N, V, T)$$

$$E(N, V, T) = -\frac{\partial}{\partial \beta} \ln Q(N, V, T)$$

$$P(N, V, T) = \frac{1}{\beta} \frac{\partial}{\partial V} \ln Q(N, V, T)$$

etc. Note that the expectation value of an observable A is

$$\langle A \rangle = \frac{1}{Q} \text{Tr}(A e^{-\beta H})$$

Evaluating the trace in the basis of eigenvectors of H (and of ρ), we obtain

$$\langle A \rangle = \frac{1}{Q} \sum_i \langle E_i | A e^{-\beta H} | E_i \rangle = \frac{1}{Q} \sum_i e^{-\beta E_i} \langle E_i | A | E_i \rangle$$

The quantum canonical ensemble will be particularly useful to us in many things to come.

3. Isothermal-isobaric and grand canonical ensembles

Also useful are the isothermal-isobaric and grand canonical ensembles, which are defined just as they are for the classical cases:

$$\Delta(N, P, T) = \int_0^\infty dV e^{-\beta PV} Q(N, V, T) = \int_0^\infty dV \text{Tr}(e^{-\beta(H+PV)})$$

$$\mathcal{Z}(\mu, V, T) = \sum_{N=0}^\infty e^{\beta \mu N} Q(N, V, T) = \sum_{N=0}^\infty \text{Tr}(e^{-\beta(H-\mu N)})$$

D. A simple example – the quantum harmonic oscillator

As a simple example of the trace procedure, let us consider the quantum harmonic oscillator. The Hamiltonian is given by

$$H = \frac{P^2}{2m} + \frac{1}{2}m\omega^2 X^2$$

and the eigenvalues of H are

$$E_n = \left(n + \frac{1}{2}\right)\hbar\omega, \qquad n = 0, 1, 2, ...$$

Thus, the canonical partition function is

$$Q(\beta) = \sum_{n=0}^{\infty} e^{-\beta(n+1/2)\hbar\omega} = e^{-\beta\hbar\omega/2} \sum_{n=0}^{\infty} \left(e^{-\beta\hbar\omega}\right)^n$$

This is a geometric series, which can be summed analytically, giving

$$Q(\beta) = \frac{e^{-\beta\hbar\omega/2}}{1 - e^{-\beta\hbar\omega}} = \frac{1}{e^{\beta\hbar\omega/2} - e^{-\beta\hbar\omega/2}} = \frac{1}{2}csch(\beta\hbar\omega/2)$$

The thermodynamics derived from it as as follows:

1. **Free energy**:
   The free energy is

   $$A = -\frac{1}{\beta}\ln Q(\beta) = \frac{\hbar\omega}{2} + \frac{1}{\beta}\ln\left(1 - e^{-\beta\hbar\omega}\right)$$

2. **Average energy**:
   The average energy $E = \langle H \rangle$ is

   $$E = -\frac{\partial}{\partial\beta}\ln Q(\beta) = \frac{\hbar\omega}{2} + \frac{\hbar\omega e^{-\beta\hbar\omega}}{1 - e^{-\beta\hbar\omega}} = \left(\frac{1}{2} + \langle n \rangle\right)\hbar\omega$$

3. **Entropy**
   The entropy is given by

   $$S = k\ln Q(\beta) + \frac{E}{T} = -k\ln\left(1 - e^{-\beta\hbar\omega}\right) + \frac{\hbar\omega}{T}\frac{e^{-\beta\hbar\omega}}{1 - e^{-\beta\hbar\omega}}$$

Now consider the classical expressions. Recall that the partition function is given by

$$Q(\beta) = \frac{1}{h}\int dpdx e^{-\beta\left(\frac{p^2}{2m} + \frac{1}{2}m\omega^2 x^2\right)} = \frac{1}{h}\left(\frac{2\pi m}{\beta}\right)^{1/2}\left(\frac{2\pi}{\beta m\omega^2}\right)^{1/2} = \frac{2\pi}{\beta\omega h} = \frac{1}{\beta\hbar\omega}$$

Thus, the classical free energy is

$$A_{\mathrm{cl}} = \frac{1}{\beta}\ln(\beta\hbar\omega)$$

In the classical limit, we may take $\hbar$ to be small. Thus, the quantum expression for A becomes, approximately, in this limit:

$$A_{\mathrm{Q}} \longrightarrow \frac{\hbar\omega}{2} + \frac{1}{\beta}\ln(\beta\hbar\omega)$$

and we see that

$$A_{\mathrm{Q}} - A_{\mathrm{cl}} \longrightarrow \frac{\hbar\omega}{2}$$

The residual $\hbar\omega/2$ (which truly vanishes when $\hbar \to 0$) is known as the quantum *zero point* energy. It is a pure quantum effect and is present because the lowest energy quantum mechanically is not $E = 0$ but the ground state energy $E = \hbar\omega/2$.