

Physics 127a: Class Notes

Lecture 12: Quantum Statistical Mechanics

Basic Expressions

The results we have found for classical statistical mechanics have corresponding ones when quantum mechanics is important.

The average value of an observable is given by a statistical and quantum expectation value

$$A_{obs} = \sum_n P_n \langle n | A | n \rangle \quad (1)$$

where $|n\rangle$ are the energy eigenstates (with energies E_n), A is the Hermitian operator corresponding to the observable, and P_n are probabilities depending on the ensemble:

microcanonical ensemble: $P_n = 1/\Omega$ over the Ω accessible states (e.g. E_n between energy E and $E + \Delta$);

canonical ensemble: $P_n = Q_N^{-1} e^{-\beta E_n}$ with $Q_N = \sum_n e^{-\beta E_n}$;

grand canonical ensemble: $P_n = Q^{-1} e^{-\beta(E_n - \mu N_n)}$ with $Q = \sum_{n,N} e^{-\beta(E_n^{(N)} - \mu N)}$;

Note there are no “cross terms” such as $\langle m | A | n \rangle$. This corresponds to a *random phase* assumption as well as the *equal probability* assumption for the microcanonical ensemble.

I will first review the notation appearing in these results, and then discuss where they come from starting from quantum mechanics.

Notation

The Dirac notation, using a *bra* and a *ket*, proves convenient, and if you are not familiar with it you should consult your favorite quantum book (or another one if it is not included!).

Basically $|\psi\rangle$ represents a quantum state. Technically $|\psi\rangle$ is a vector in *Hilbert space*, and the usual rules of linear algebra apply. We call this a *ket* vector. For example for a single electron in a $1/r$ potential this might be a $1s$ state or a $2p_x$ state with respect to specified coordinate directions. We often think about the state in terms of its *position* representation or *wavefunction*

$$|\psi\rangle \Leftrightarrow \psi(\mathbf{x}) \quad (2)$$

but we could also use a momentum representation $\tilde{\psi}(\mathbf{p})$ which would be the Fourier transform of $\psi(\mathbf{x})$. The notation $|\psi\rangle$ does not presuppose any particular way of representing the state.

The *scalar product* between two states $|\psi_1\rangle$ and $|\psi_2\rangle$ is denoted by

$$\langle \psi_2 | \psi_1 \rangle = \text{some complex number} = \langle \psi_1 | \psi_2 \rangle^* \quad (3)$$

In terms of the position representation

$$\langle \psi_2 | \psi_1 \rangle = \int d\mathbf{x} \psi_2^*(x) \psi_1(x). \quad (4)$$

This introduces the *bra vector* or *dual space* vector $\langle \psi |$. It is introduced though the formation of the scalar product, but it is often useful to “liberate it” from this construction.

Any complete set of states $|\phi_i\rangle$ provides a basis

$$|\psi\rangle = \sum_i c_i |\phi_i\rangle. \quad (5)$$

It is often easiest to use an *orthonormal basis* with

$$\langle\phi_j|\phi_i\rangle = \delta_{ij} \quad (\text{zero if } i \neq j, \text{ unity if } i = j). \quad (6)$$

Then

$$c_i = \langle\phi_i|\psi\rangle \quad \text{and} \quad |\psi\rangle = \sum_i |\phi_i\rangle \langle\phi_i|\psi\rangle. \quad (7)$$

Since the latter expression is true for any $|\psi\rangle$, we have the completeness relation

$$\sum_i |\phi_i\rangle \langle\phi_i| = 1, \quad (8)$$

(put bras and kets on either side of this to make sense of it).

The position states $|\mathbf{x}\rangle$ (particle “at” \mathbf{x}) are a basis for single particle states

$$|\psi\rangle = \int d\mathbf{x} \psi(\mathbf{x}) |\mathbf{x}\rangle, \quad (9)$$

$$\psi(x) = \langle x|\psi\rangle. \quad (10)$$

The average value of measurements of an observable in the state $|\psi\rangle$ is

$$A_{obs} = \langle\psi|A|\psi\rangle \equiv \langle\psi|(A|\psi\rangle) \quad (11)$$

where A is the Hermitian operator that corresponds to the observable. For example the energy is

$$E_{obs} = \langle\psi|H|\psi\rangle. \quad (12)$$

We can translate this into a more familiar form for a single particle in one dimension in a potential $V(x)$ for example by inserting complete sets of position states $\int dx |x\rangle \langle x| = 1$

$$E_{obs} = \int dx' \int dx \langle\psi|x'\rangle \langle x'|H|x\rangle \langle x|\psi\rangle \quad (13)$$

$$= \int dx \psi^*(x) \left(-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + V(x) \right) \psi(x). \quad (14)$$

Density Matrix

Definition Write the quantum state of system plus rest of universe in terms of a complete orthonormal set of system states $|\phi_i\rangle$ and rest of universe states $|\theta_a\rangle$

$$|\psi\rangle = \sum_{i,\alpha} C_{i\alpha} |\phi_i\rangle |\theta_\alpha\rangle. \quad (15)$$

Consider a Hermitian operator A corresponding to an observable of the system, and therefore depending only on system variables. The expectation value is

$$\langle A \rangle = \langle\psi|A|\psi\rangle = \sum_{i,j} \langle\phi_i|A|\phi_j\rangle \rho_{ji} \quad (16)$$

where the *density matrix* ρ_{ij} is

$$\rho_{ij} = \sum_{\alpha} C_{i\alpha}^* C_{j\alpha} \quad (17)$$

and we have used $\langle \theta_{\alpha} | \theta_{\beta} \rangle = \delta_{\alpha\beta}$. We can define the *density operator* ρ as the operator that has these matrix elements $\langle \phi_j | \rho | \phi_i \rangle = \rho_{ij}$, and then can write the expectation value as

$$\langle A \rangle = Tr(A\rho) = Tr(\rho A) \quad (18)$$

(where Tr denotes the trace—the sum of diagonal elements) using the completeness result $\sum_j |j\rangle \langle j| = 1$ in Eq. (16) and the cyclic invariance of the trace.

Properties

- From the definition Eq. (17) ρ is Hermitian: $\rho_{ij}^* = \rho_{ji}$;
- Therefore ρ has real eigenvalues w_i and the corresponding eigenvectors $|i\rangle$ form a complete orthonormal set. We can use this set as a basis to diagonalize ρ , so that in this basis

$$\rho_{ij} = w_i \delta_{ij} \quad \text{or} \quad \rho = \sum_i w_i |i\rangle \langle i|. \quad (19)$$

- Using $\langle \psi | A | \psi \rangle = Tr(\rho A)$ for special choices of A we show

– choosing $A = 1$ leads to

$$\sum_i w_i = 1. \quad (20)$$

– choosing $A = |j\rangle \langle j|$ (the projection operator onto the j th state) leads to

$$w_j = |\langle j | \psi \rangle|^2 \geq 0 \quad \text{for all } j. \quad (21)$$

Together Eqs. (20) and (21) show that the w_i can be considered to be a *probability*, so that w_i is the interpreted as the probability that the system is in the quantum state $|i\rangle$. In the diagonal basis Eq. (18) reduces to

$$\langle A \rangle = \sum_i w_i \langle i | A | i \rangle \quad (22)$$

telling us that the expectation value of an observable is given by the weighted average of the expectation value of the observable in the states $|i\rangle$. Note that *this only applies in the special choice of basis that diagonalizes ρ* . In a general basis we just have the expression Eq. (18)

$$\langle A \rangle = \sum_{p,q} \langle p | A | q \rangle \rho_{qp}. \quad (23)$$

If all but one w_i are zero, and then the non-zero value must be unity, the system is in a *pure state* and the results correspond to the elementary formulation of quantum mechanics in terms of wave functions. If more than one w_i is nonzero, the system is in a *mixed state*, and ρ combines the necessary quantum and statistical information. A necessary and sufficient condition for ρ to describe a pure state is $\rho^2 = \rho$, so that ρ is then a projection operator.

Time dependence The system is supposed to be sufficiently weakly coupled to the rest of the universe that the time dependence is given by the internal Hamiltonian H . Then each state evolves according to the Schrodinger equation, which for the bra and ket vectors reads

$$-i\hbar \frac{\partial}{\partial t} |i\rangle = H |i\rangle, \quad (24a)$$

$$i\hbar \frac{\partial}{\partial t} \langle i| = \langle i| H. \quad (24b)$$

Then in the diagonal basis we have

$$-i\hbar \frac{\partial}{\partial t} \rho = \sum_i w_i \left(-i\hbar \frac{\partial}{\partial t} |i\rangle \langle i| \right) \quad (25)$$

$$= \sum_i w_i (H |i\rangle \langle i| - |i\rangle \langle i| H) = [H, \rho] \quad (26)$$

with $[H, \rho]$ the commutator. (Note the time dependence of ρ comes from the time dependence of the states, and the equation for $\partial\rho/\partial t$ has the opposite sign from the equation for the time dependence of an operator corresponding to a physical observable in the Heisenberg picture.)

Statistical Mechanics

So far ρ just encapsulates how much of the behavior of the rest of the universe we need to know to calculate expectation values of system observables. We now need to evaluate ρ for various physical situations, and in particular for systems in the microcanonical, canonical, and grand canonical ensemble.

If we allow the system to come into equilibrium we must have $\partial\rho/\partial t = 0$ and so

$$\text{Equilibrium } \rho: \quad [H, \rho] = 0. \quad (27)$$

A sufficient condition for this is $\rho = \rho(H, C_1, C_2 \dots)$ with C_i conserved quantities i.e. observables satisfying $[H, C_i] = 0$, and as in the classical case we assume this is the only way of satisfying Eq. (27) in physical systems. This gives us the very important result that ρ and H are diagonalized in the same basis, i.e. *the states for which the density matrix is diagonal are in fact the energy eigenstates* $|n\rangle$ (with energies E_n), so that

$$\rho = \sum_n w_n |n\rangle \langle n| \quad \text{and} \quad \langle A \rangle = \sum_n w_n \langle n | A | n \rangle. \quad (28)$$

Again these simple algebraic statements hide, or sensibly sidestep, many complicated processes. We could, for example, imagine setting up our system in some pure state as an initial condition. This would be some simple state that we could imagine establishing (e.g. all spins pointing in the x direction). We could imagine following the time dependence of this pure state under the internal Hamiltonian: in general it would evolve into a complicated linear combination of such simple basis states (which are not eigenstates of the Hamiltonian in general). Unfortunately we would be fooling ourselves to imagine this is a good description of the system, since the delicate phase relationships between these different components are too easily disrupted by the residual interactions with the rest of the universe. The “quantum coherence” of the system is lost, because macroscopic systems are inevitably coupled to the rest-of-the-universe degrees of freedom, and even very weak coupling readily destroys the delicate phase relationships between different components of the wavefunction. Thus it is profitable to sidestep these difficulties, and use the observed fact that macroscopic systems come to equilibrium and so are defined by a time independent density matrix (approximately, to sufficient precision...), and use this to calculate what we need to know.

The major stumbling block to building quantum computers is precisely to reduce this loss of quantum coherence (the tendency of the system to evolve from a pure state to a mixed state because of inevitable residual couplings). It is *very hard* to build systems of even a moderate number of particles, spins etc., for which the quantum coherence is maintained over a useful time period. There is a great deal of work examining the loss of coherence for particular models of simple systems, the rest of the universe, and the residual couplings. For such simple systems the process can be followed in detail. Of course, for macroscopic systems the processes are much too complicated and varied to follow in detail, and we immediately jump to the mixed state described by the equilibrium ρ .

Microcanonical ensemble The energy and conserved quantities are all fixed for every allowed state

$$H |n\rangle = E |n\rangle \quad (29)$$

$$C_i |n\rangle = c_i |n\rangle \quad (30)$$

and so H and all the C_i and therefore ρ itself are proportional to unit matrices between these states, i.e.

$$\rho = \left[\begin{array}{c} \frac{1}{\Omega} \left[\begin{array}{cccc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & 1 \end{array} \right] \\ [0] \end{array} \right] \quad (31)$$

where the first block runs over the Ω accessible states. Or in terms of the weights

$$w_n = \begin{cases} \frac{1}{\Omega} & \text{for the } \Omega \text{ accessible states} \\ 0 & \text{otherwise} \end{cases} \quad (32)$$

Canonical ensemble The same arguments as in the classical case lead to

$$w_n = \frac{e^{-\beta E_n}}{\sum_n e^{-\beta E_n}} \quad (33)$$

We can also write the result as

$$\rho = \frac{e^{-\beta H}}{\text{Tr} e^{-\beta H}} \quad (34)$$

and then we have

$$Q_N = \text{Tr} e^{-\beta H} \quad \text{and} \quad S = -k \text{Tr}(\rho \ln \rho). \quad (35)$$

Coordinate representation Although we will use the diagonal (energy eigenstate) representation for most purposes, we can use other basis states instead. For example consider the very simple example of a single particle moving along a line of length L (with periodic boundary conditions). Define the unnormalized density matrix in the canonical ensemble

$$\bar{\rho} = e^{-\beta H}. \quad (36)$$

Let us evaluate $\bar{\rho}$ in the coordinate representation $\bar{\rho}(x, x'; \beta)$. A convenient way to do this is to differentiate Eq. (36) with respect to β

$$-\frac{\partial \bar{\rho}}{\partial \beta} = H \bar{\rho}. \quad (37)$$

which for the free particle becomes in the position basis

$$-\frac{\partial \bar{\rho}(x, x'; \beta)}{\partial \beta} = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} \bar{\rho}(x, x'; \beta) \quad (38)$$

which must be solve with the “boundary condition” $\bar{\rho}(\beta = 0) = 1$, i.e. $\bar{\rho}(x, x'; \beta) = \delta(x - x')$ (the delta function). Notice what we are doing is calculating $\bar{\rho}$ at some finite temperature by integrating from infinite temperature where the density matrix is simple. We can solve Eq. (38) by recognizing it is the *diffusion equation* where β plays the role of time. The solution is a Gaussian that spreads as β increases

$$\bar{\rho}(x, x'; \beta) = \sqrt{\frac{m}{2\pi\hbar^2\beta}} \exp\left[-\left(\frac{m}{2\hbar^2\beta}\right)(x - x')^2\right], \quad (39)$$

where the prefactor is given by matching to the delta function as β goes to zero (e.g. integrate over all x). Notice we can write the exponential as $e^{-\pi(x-x')^2/\lambda^2}$ with $\lambda = \sqrt{\hbar^2/2\pi mkT}$ the thermal length introduced in our discussion of the classical gas. (Be careful of the \hbar and h symbols!)

The partition function is

$$Q_1 = Tr \bar{\rho} = \int_0^L dx \bar{\rho}(x, x; \beta) = L \sqrt{\frac{mkT}{2\pi\hbar^2}}. \quad (40)$$

You should be able to see how this relates to the partition function Q_N of the ideal gas calculated previously.

Equation (39) is the starting point for the *path integral* formulation of quantum statistical problems—see Feynman’s book *Statistical Mechanics* for more on this.

Grand canonical ensemble Similarly, in this case

$$w_n = \frac{e^{-\beta(E_n^{(N)} - \mu N)}}{\sum_n e^{-\beta(E_n^{(N)} - \mu N)}} \quad (41)$$

and

$$\rho = \frac{e^{-\beta(H - \mu N)}}{Tr e^{-\beta(H - \mu N)}} \quad (42)$$

with N here the number operator. The grand canonical partition function is

$$Q = Tr e^{-\beta(H - \mu N)}. \quad (43)$$