Sept. 26

Physics 212 – Statistical Mechanics

Basic Formulae of Stantistical Mechanics

Our principal object of study in this course is the statistical mechanical partition function

$$Z = \int e^{-\beta\mathcal{H}} \tag{1}$$

where $\mathcal{H}$ is the Hamiltonian, $\beta = 1/T$, and T is the temperature. In classical statistical mechanics, the integral is taken over all of phase space for the system in question. In quantum statistical mechanics, there is a sum over all energy eigenstates. In both cases, the statistical weight of a given configuration is

$$\exp[-\mathcal{H}(q,p)/T] \quad \text{or} \quad \exp[-E_i/T] \tag{2}$$

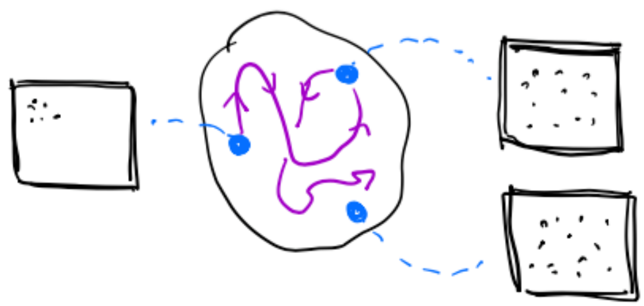
In the above, $\beta = 1/T$ and T is the temperature. In this course, I will measure temperature in energy unit, setting Boltzmann's constant k_B equal to 1. This sum or integral and weighting defines the *canonical ensemble.*

In this lecture, I will discuss the derivation of the canonical ensemble and describe some of its properties. This will mainly be a reminder; I hope that you have seen a full treatment of this material in an earlier course. In this lecture, I will generally use the language of classical systems, but at some points I will quote their generalization to quantum systems.

Let's first very briefly review the foundations of Statistical Mechanics and the origin of the canonical ensemble. I hope that this is a reminder of arguments you have heard before.

The basic postulates of Statistical Mechanics are two. The first is well-appreciated: conservation of energy. The energy of an isolated system is constant. This is the *first law of thermodynamics.*

The second, due to Boltzmann, is more subtle. According to the first law, a dynamical system evolves along its energy surface, that is, the set of configurations in phase space $\{q_i, p_i\}$ with fixed energy $\mathcal{H}(q_i, p_i) = E$. On this surface, there are special points with simple interpretation, for example, for an ideal gas, the points where all of the atoms are in the same small volume with almost identical energies. But, very quickly, this point will evolve to a more generic point at which the atoms are scattered through the volume and have a distribution of momenta. Boltzmann's insight is that, first, these generic points are macroscopically indistinguishable, and, second, that any one (that is, almost any point on the energy surface) has the macroscopic properties of the state of thermal equilibrium. In the small, these states are constantly changing into one another, but, in the large, there is no visible change.



(3)

This being so, *the properties of the state of thermal equilibrium are given by an average over all of the microscopic states on the energy surface.*

In this course, I will take this statement as an axiom. Over more than a century, there have been many attempts to prove it, and proofs have been given for very idealized systems. There are also idealized systems ("integrable systems") for which the statement is not true. In general, this postulate has been very successful in allowing us to compute the properties of real systems with a large number of interacting particles or degrees of freedom.

It follows from this axiom that the properties of a macroscopic system in thermal equilibrium are given by averages in the measure

$$\Omega(E) = \prod_i \int \frac{dq_i dp_i}{2\pi}\ \delta(\mathcal{H}(q,p) - E) \tag{4}$$

This is called the *microcanonical ensemble.*

Our axiom applies only to very large systems. Usually, I will speak ideally of systems with infinite volume V or containing an infinite number of particles N. In practice, though, since Avogadro's number is so large, systems of the size of a few microns are already large enough that this picture applies.

To go further, I would like to make an important distinction of the levels of infinity that we meet in Statistical Mechanics.

- *intenstive quantities* are quantities that remain constant as $N \to \infty$.
- *extensive quantities* are quantities that grow proportional to N as $N \to \infty$.
- *super-extensive quantities* are quantities that grow as to e^{cN} as $N \to \infty$.

Super-extensive quantities play an important role in the analysis that I am going to describe now.

The quantity $\Omega(E)$ is super-extensive. A good example to think about is system of N coupled harmonic oscillators. The partition function is

$$\Omega(E) = \prod_i \int \frac{dx_i dp_i}{(2\pi)} \, \delta\Big(E - \sum_i [\frac{p_i^2}{2m} + \frac{kx_i^2}{2}]\Big) \tag{5}$$

Making the change of variables

$$y_{2i-1} = \frac{1}{\sqrt{2mE}} \, p_i \qquad y_{2i} = \sqrt{\frac{k}{2E}} \, x_i \ , \tag{6}$$

this becomes

$$\Omega(E) = \Big(2\sqrt{\frac{m}{k}}E\Big)^N \cdot \frac{1}{E} \cdot \mathcal{A}(2N) \tag{7}$$

where $\mathcal{A}(2N)$ is the area of the unit sphere in $2N$ dimensions. The important part here is

$$\Omega(E) = (\text{const}) \cdot \exp[N \log E] \ . \tag{8}$$

The microcanoncal integral $\Omega(E)$ is often written

$$\Omega(E) = \exp[S(E)] \tag{9}$$

The quantity $S(E)$ is the *entropy*. It is a properly extensive quantity that characterizes the number of states available to the system on a surface of constatnt energy.

In quantum mechanics, we would replace the delta function in energy by a shell of small thickness in E. Let $\mathcal{N}$ be the number of states in this shell. Then

$$S(E) = \log \Omega(E) = \log \mathcal{N} \tag{10}$$

The probability of any given quantum state in this ensemble is

$$P_i = \frac{1}{\mathcal{N}} \tag{11}$$

Then

$$S = -\log P_i \quad \text{for any } i \tag{12}$$

or

$$S = -\sum_i P_i \log P_i \tag{13}$$

This expression for S can be viewed more generally as a measure of the *Shannon entropy*, which measures the (lack of) information in a probability distribution

When $E \to E_0$, the ground state energy of the system, there will be only one state, or a finite number of states, in the energy shell. Then the macroscopic $S(E) \to 0$. This is the *third law of thermodynamics.*

Now we are ready to define the canonical ensemble. It is very useful to transform the basic variable of our construction from the energy to a surrogate intensive quantity, the temperature T or the inverse temperature β. To do this, we integrate the microcanonical ensemble with a comparable super-extensive quantity,

$$\int dE\ \Omega(E) e^{-\beta E} = \int dE\ \ \exp[S[E] - \beta E] \tag{14}$$

Both exponents are truly enormous, so, if the integrand has a pea,k it will be very sharply peaked and we can evaluate the integral by steepest descents. The maximum of the integral occurs at

$$\left.\frac{dS}{dE}\right|_* = \beta\ . \tag{15}$$

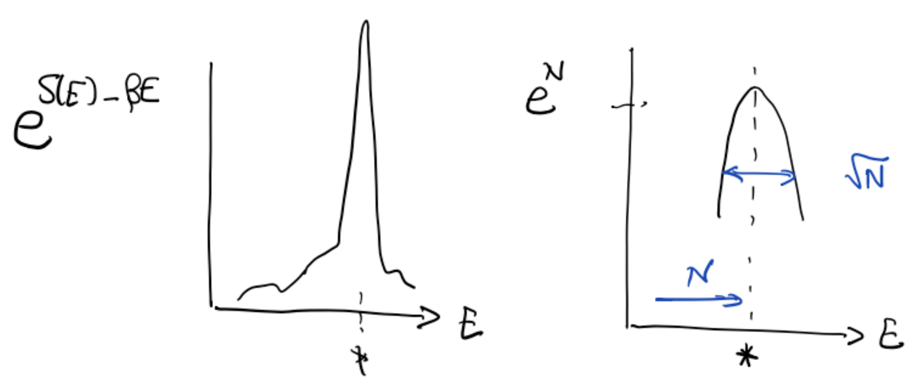
The fluctuations around this peak are described by

$$\exp[-\frac{1}{2}\mathcal{C}(\delta E)^2]\ , \tag{16}$$

where

$$\mathcal{C} = -\frac{d^2 S}{dE^2} \tag{17}$$

Note that d^2S/dE^2 must be < 0 for this to be sensible. The peak of the integral corresponds very closely to a fixed energy, and we can trade the variable E for the corresponding value of β.



(18)

Thus, define

$$Z = e^{-\beta F} = \int e^{-\beta \mathcal{H}} \tag{19}$$

as I wrote at the beginning of the lecture. The (extensive) quantity $F(\beta)$ is the *free energy.* Macroscopically— that is, ignoring fluctuations about the maximum—the relation between F and S is a *Legendre transformation*,

$$-\beta F = S - \beta E \tag{20}$$

with

$$\beta = \frac{\partial S}{\partial E} \tag{21}$$

Ordinarily, we write $T = 1/\beta$, so that this relation becomes

$$F = E - TS \tag{22}$$

The spirit of the Legendre transforamtion is that we can now consider F to be a function of β or T rather than E: $F = F(T)$ and we can consider the free energy to be a function $F(T)$. Note that

$$\begin{aligned}\frac{\partial F}{\partial T} &= \frac{\partial E}{\partial T} - S - T\frac{\partial S}{\partial E}\frac{\partial E}{\partial T} \\ &= \frac{\partial E}{\partial T} - S - T\frac{1}{T}\frac{\partial E}{\partial T}\end{aligned} \tag{23}$$

Then

$$\frac{\partial F}{\partial T} = -S \tag{24}$$

and we can recover the entropy from knowledge of the free energy $F(T)$.

Viewed as a function of β or temperature, the canonical ensemble portrays Statistical Mechanics as a battle of energy against entropy. Changing the temperature shifts the balance: At low temperature, $\beta \to \infty$ and energy completely dominates. At high temperature, $\beta \to 0$ and entropy is the whole story. We will return to this intuition again and again during the course.

As an application of the canonical ensemble, consider a thermodynamic system consisting of two subsystems that are weakly coupled in such a way that they are able to maintain thermal equilibrium. The canonical ensemble for this system is

$$\begin{aligned}Z &= \int dE_1 dE_2 \; \exp[S(E_1) + S(E_2) - \beta(E_1 + E_2)] \\ &= \int dE d\Delta E \; \exp[S(E + \Delta E/2) + S(E - \Delta E/2) - \beta E]\end{aligned} \tag{25}$$

When both systems are large, both integrals are well approximated by steepest descents. Then, at the peak,

$$\frac{dS_1}{dE_1} = \frac{dS_2}{dE_2} \quad \text{or} \quad \beta_1 = \beta_2 \tag{26}$$

Thus, *two coupled systems are in thermal equilibrium when they are at the same temperature. In the approach to equilibrium, energy flows from one system to another in* *such a way that the total entropy is maximized.* This the second law of thermodynamics.

The formulae apply equally well when only one system, say, system 2, is large and the other is small. In that case, the temperature is determined by

$$\frac{dS_2}{dE_2} = \beta \tag{27}$$

and the properties of the small system are described by a canonical ensemble corresponding to this temperature. We call the large system a "heat bath" that regulates the small system. If the small system has a number of degrees of freedom N that is so small that its energy fluctuations are relevant, the size of these fluctuations will be of order

$$\sqrt{N} \sim E/\sqrt{N} \tag{28}$$

The third ensemble traditionally treated in Statistical Mechanics courses is the *grand canonical ensemble.* Here we consider the number of particles N to be indefinite and regulated by an external reservoir of particles. The integral over configurations is written

$$\Xi = \sum_N \int_{\text{fixed } N} e^{-\beta(\mathcal{H}-\mu N)} \tag{29}$$

where μ is called the *chemical potential.* The grand free energy is Φ, given by

$$\Xi = e^{-\beta\Phi} \tag{30}$$

Φ is connected to F by another Legendre transformation

$$\Phi = F - \mu N \ . \tag{31}$$

This motivates another change of variables, from N to μ. Using the same logic as above, the particle number is recovered from $\Phi(T, \mu)$ by

$$\left.\frac{\partial\Phi}{\partial\mu}\right|_T = N \tag{32}$$

In this course, we will meet systems with an indefinite number of particles, even systems where the number of particles has quantum-mechanical uncertainty.

This completes my review of the basic principles of Statistical Mechanics. If you are unfamiliar with these principles, please take time to review them from the books of Landau and Lifshitz, Sethna, or others that you may have studied. An important application of these principles is the calculation of the properties of the classical and quantum ideal gases. As a physicist, you will need to be totally familiar with this topic, but I will not cover it in this course.

I will now switch to a discussion of some formal aspects of the Statistical Mechanical ensembles that I will need to make use of during the course. I will discuss two topics, first, the variational principle of Statistical Mechanics, second, correlation functions and susceptibilities.

I hope you are familiar with the following principle from Quantum Mechanics: Let $\mathcal{H}$ be a Hamiltonian that we would like to diagonalize but probably cannot. As a model for $\mathcal{H}$, we consider a family of Hamiltonians $\mathcal{H}(a)$ that we can diagonalize. Let $|0\rangle$ be the ground state of $\mathcal{H}$ with ground state energy E_0, and let $|a\rangle$ be the ground state of $\mathcal{H}(a)$ with ground state energy E_a. Then

$$E_0 \leq E_a + \langle a| \left(\mathcal{H} - \mathcal{H}_a\right) |a\rangle \tag{33}$$

We can use this as a variational principle to find the best choice of a to model $\mathcal{H}$ by minimizing the right-hand side with respect to a.

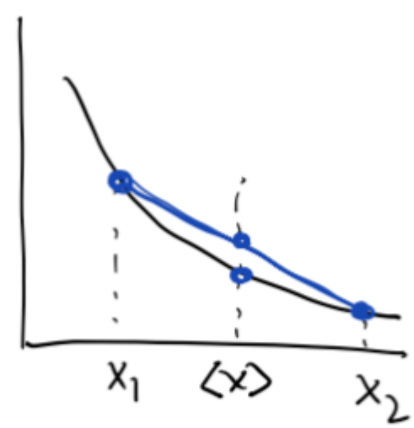
There is a similar principle in Statistical Mechanics. To derive it, note that, for any positive probability distribution $P(x)$ with

$$\int dx P(x) = 1 \ , \quad \langle f(x)\rangle = \int dx P(x) f(x) \tag{34}$$

it is true that

$$\left\langle e^{-x} \right\rangle \geq e^{-\langle x\rangle} \tag{35}$$

For example, if we average the values of e^{-x} at two points x_1 and x_2, we always get a larger value that that from evaluating e^{-x} at the average of x_1 and x_2.



(36)

Now let $\mathcal{H}$ be a Hamiltonian for which we would like to compute the partition function Z but probably cannot. As a model for $\mathcal{H}$, we consider a family of Hamiltonians $\mathcal{H}(a)$ for which we can compute the partition function. Let $\langle f\rangle_a$ be the expectation value of f computed with the partition function of $\mathcal{H}_a$. Then

$$Z = \int e^{-\beta\mathcal{H}} = \int e^{-\beta\mathcal{H}_a} e^{-\beta(\mathcal{H}-\mathcal{H}_a)} \tag{37}$$

From the inequality above, we have

$$Z/Z_a \geq \exp[-\beta \langle \mathcal{H} - \mathcal{H}_a \rangle_a] \tag{38}$$

or, for the corresponding free energies

$$F \leq F_a + \langle \mathcal{H} - \mathcal{H}_a \rangle_a \tag{39}$$

Again, we can vary a to find the $\mathcal{H}_a$ that is the best model for $\mathcal{H}$. We will use this method at various points in the course.

Now consider another generalization of the canonical partition function. Often, in the systems we will consider, there are macroscopic fields whose values are determined by the thermodynamics. I will most often use the example of the local magnetization $m(x)$ in a magnet, though the same language can be applied to the density of a fluid, the fraction of a given constituent, the orientatons or stretching of component molecules, or any other type of local variable. We can also define an extensive global quantity

$$M = \int d^3x \; m(x) \tag{40}$$

or the analogue for any of these other choices. For a magnet, the global magnetization can be shifted by applying a constant magnetic field H. In other cases, it might not be physically possible to apply this field, but we can still imagine that we can affect to global density by such a field in principle. We can also imagine imposing a field that varies from one position to another in any way that is convenient for our discussion. Note that M is extensive and H is intensive. In a real magnet M might be a vector quantity, but here I will treat it as a scalar.

Chemists usually begin the study of thermodynamics by consider the particle density or number and using the parameter pressure to regulate it. Density gets mixed up with volume and there are some complications that are important to chemists. I use magnetization here to present a somewhat cleaner situation.

Define, then, the partition function of an idealized magnet in a constant external magnetic field,

$$Z(H) = \int e^{-\beta(\mathcal{H} - HM)} \tag{41}$$

Then

$$\frac{1}{Z}\frac{\partial}{\partial H} Z = \frac{\partial}{\partial H} \log Z = \beta M \tag{42}$$

and we can recover the value of the magnetization in thermal equilibrium by

$$\left.\frac{\partial F}{\partial H}\right|_T = -M \tag{43}$$

It is very convenient to define the *Gibbs free energy* by a Legendre transformation

$$G = F + HM \tag{44}$$

The quantity G is formally a function of T and M. The value of H is recovered by

$$\left.\frac{\partial G}{\partial M}\right|_T = H \tag{45}$$

The quantity $G(T, M)$ will be a very important object in our study.

Now let's modify the partition function to allow H to vary from point to point,

$$Z[H(x)] = \int e^{-\beta(\mathcal{H} - \int d^3x H(x) m(x))} \tag{46}$$

If you would like, you can think of an array of atoms, with H taking an adjustible value on each atom,

$$\int d^3x H(x) m(x) \to \sum_i H_i m_i \tag{47}$$

Taking the derivative with respect to H_i, we find $\langle m_i \rangle$. In the continuum, we describe this operatiopn as taking the *variational derivative* with respect to the local magnetic field $\delta/\delta H(x)$ and write the operation as

$$\frac{1}{Z}\frac{\delta}{\delta H(x)} Z = \frac{\int e^{-\beta(\mathcal{H} - \int Hm)} \beta m(x)}{\int e^{-\beta(\mathcal{H} - \int Hm)}} = \beta \langle m(x) \rangle \ . \tag{48}$$

Similarly,

$$\frac{1}{Z}\frac{\delta}{\delta H(x)\delta H(y)} Z = \beta^2 \langle m(x) m(y) \rangle \ . \tag{49}$$

Then $Z[T, H(x)]$ is the *generating function of correlation functions* of $m(x)$. The successive derivatives of Z with respect to local fields $H(x)$ compute the successive correlation functions.

A more interesting set of relations is found by thinking about the successive derivatives of $\log Z$. First,

$$\frac{\delta}{\delta H(x)} \log Z = \frac{\int e^{-\beta(\mathcal{H} - \int d^3z H(z) m(z))} \beta m(x)}{\int e^{-\beta(\mathcal{H} - \int d^3z H(z) m(z))}} = \beta \langle m(x) \rangle \ . \tag{50}$$

Now take $\partial/\partial H(y)$ of the expression. Note that we must differentiate both the numerator and the denominator. The result is

$$\begin{aligned} \frac{\delta^2}{\delta H(x)\delta H(y)} \log Z = & \frac{\int e^{-\beta(\mathcal{H} - \int d^3z H(z) m(z))} \beta m(x) \cdot \beta m(y)}{\int e^{-\beta(\mathcal{H} - \int d^3z H(z) m(z))}} \\ & - \frac{\int e^{-\beta(\mathcal{H} - \int d^3z H(z) m(z))} \beta m(x)}{\int e^{-\beta(\mathcal{H} - \int d^3z H(z) m(z))}} \cdot \frac{\int e^{-\beta(\mathcal{H} - \int d^3z H(z) m(z))} \beta m(y)}{\int e^{-\beta(\mathcal{H} - \int d^3z H(z) m(z))}} \end{aligned} \tag{51}$$

or

$$\frac{\delta^2}{\delta H(x)\delta H(y)} \log Z = \beta^2 (\langle m(x) m(y) \rangle - \langle m(x) \rangle \langle m(y) \rangle) \ . \tag{52}$$

What is this object? If there is a background H field, then in general there will be a nonzero value of $m(x)$ that responds to this. The quantity

$$\langle m(x)m(y)\rangle_c = \langle m(x)m(y)\rangle - \langle m(x)\rangle \langle m(y)\rangle \tag{53}$$

subtracts out the contribution to the joint expectation value of $m(x)$ and $m(y)$ that comes from this background value, leaving only the effect of the actual statistical fluctuation. This object is called the *connected correlation function.*

Take another derivative. After some simplification, you should find

$$\begin{aligned}\frac{\delta^2}{\delta H(x)\delta H(y)\delta H(z)} \log Z = \beta^3 \Big[&\langle m(x)m(y)m(z)\rangle - \langle m(x)m(y)\rangle \langle m(z)\rangle \\ &- \langle m(x)m(z)\rangle \langle m(y)\rangle - \langle m(y)m(z)\rangle \langle m(x)\rangle \\ &+2 \langle m(x)\rangle \langle m(y)\rangle \langle m(z)\rangle \Big]\end{aligned} \tag{54}$$

The right-hand side contains only contributions that are joinly sensitive to fluctuations at x, y, and z. In general, $\log Z$ is the generating function of connected correlations functions. All of the derivatives of $\log Z$ are increasingly complex connected correlators.

The expectation value of the energy in the canonical ensemble is given by

$$\langle E\rangle = -\frac{\partial}{\partial\beta}\log Z \tag{55}$$

We can also see this from the formula for the free energy as a Legendre transformation,

$$\begin{aligned}\frac{\partial}{\partial\beta}(\beta F) &= F - \beta\frac{\partial F}{\partial\beta} \\ &= F = T\frac{\partial F}{\partial T} = F + TS\end{aligned} \tag{56}$$

Now take another derivative with respect to temperature. This gives the response of the energy to a change in temperature, the *specific heat* C.

$$C = \frac{\partial E}{\partial T} = -\frac{\beta}{T}\frac{\partial}{\partial\beta}(-\frac{\partial}{\partial\beta}\log Z) \tag{57}$$

Computing this,

$$C = \frac{1}{T^2}\frac{\partial^2}{\partial\beta^2}\log Z \tag{58}$$

or

$$C = \frac{1}{T^2}[\left\langle E^2\right\rangle - \langle E\rangle^2) \tag{59}$$

So, the response of the energy of the system to a change in temperature is also given by a connected correlator that measures the fluctuations of the energy in equilibrium.

Each term in the above equation is of order N^2, but after subtraction the difference C is proportional to N, because the size of the fluctuations of E is order $\sqrt{N}$. Thus, C is properly extensive, as it should be.

Similarly, we can define the *magnetic susceptibility* as

$$\chi = \frac{\partial M}{\partial H}\bigg|_T \,. \tag{60}$$

Computing this similarly, we find

$$\chi = \frac{1}{\beta}\frac{\partial^2}{\partial H^2}\log Z \tag{61}$$

or

$$\chi = \frac{1}{\beta}(\left\langle M^2 \right\rangle - \langle M \rangle^2) \tag{62}$$

This connection between fluctuations of thermodynamic quantitues and the response of the system to external fields is completely general. We will use it at many points in the course.

There is one more issue that I would like to discuss as part of our introductory material. Although this is a course in *classical Statistical Mechanics*, we will encounter a surprising amount of quantum mechanics. We will discuss superfluidity and superconductivity which are essentially quantum phenomena. But also, there is a close analogy between thermal fluctuations and quantum fluctuations that we will exploit at certain points. Also, since we will deal with fields like the magnetization $m(x)$ that have thermal fluctuations, we will be studying field theories with fluctuations that are closely related to quantum field theories.

I hope you are not discouraged by analogies to quantum field theory. Think, instead, that this is the easiest quantum field theory course you will ever take.

In the course of this discussion, we will encounter a formalism that I hope you have seen before but which may be new to you, called "second quantization". I would like to give you a simple introduction to this, which will suffice for this course.

Consider, then, a cavity supporting waves with resonant frequencies ω_j. You can think about electromagnetic waves in a cavity, but here I will use the notation of scalar (e.g., sound) waves. However, to help your intuition, I will call the quantum excitations photons. The classical description of this system is as a system of harmonic oscillators, one for each resonant mode. The Hamiltonian can be written as a sum of harmonic oscillators.

$$\mathcal{H} = \sum_j (\frac{1}{2}p_j^2 + \frac{1}{2}\omega_j^2 q_j^2) \tag{63}$$

It is not hard to relate the p_js and $q_j's$ to integrals over the waveform, but that will not be necessary for this discussion.

In your quantum mechanics class, you encountered the operator quantization of the harmonic oscillator Hamiltonian. Let the ground state of the oscillator be $|0\rangle$. Then there are operators a, $a^\dagger$ such that

$$[a, a^\dagger] = 1 \ , \text{and} \quad a\,|0\rangle = 0 \ . \tag{64}$$

These properties allow us to construct the full spectrum of the oscillator as the set of states

$$|n\rangle = \frac{1}{\sqrt{n!}}(a^\dagger)^n\,|0\rangle \tag{65}$$

The Hamiltonian can be written

$$\mathcal{H} = \omega(a^\dagger a + \frac{1}{2}) = \omega(\mathcal{N} + \frac{1}{2}) \ , \tag{66}$$

where $\mathcal{N}$ is the number operator

$$\mathcal{N} = a^\dagger a \quad \text{which satisfies} \quad \mathcal{N}\,|n\rangle = n\,|n\rangle \ . \tag{67}$$

I will consistently set $\hbar = 1$.

We can now write the Hamiltonian of the system of waves in a cavity in a useful form. Quantizing each oscillator in this way, we have

$$\mathcal{H} = \sum_j \omega_j(a_j^\dagger a_j + \frac{1}{2}) \tag{68}$$

where

$$[a_j, a_k^\dagger] = \delta_{jk} \quad [a_j, a_k] = [a_j^\dagger, a_k^\dagger] = 0 \tag{69}$$

A typical state in the Hilbert space is

$$|1_i, 3_2, 0_3, 1_4, \cdots\rangle = a_1^\dagger (a_2^\dagger)^3 a_4^\dagger \cdots |0\rangle \tag{70}$$

with energy

$$E - E_0 = \omega_1 + 3\omega_2 + \omega_4 + \cdots \tag{71}$$

where $E_0 = \sum_j \omega_j/2$ is the ground state energy. This state has a simple description. In a given state, each mode has a definite occupation number, and the energy of the state is given by the sum of the values ω_j for each occupied state. It is very natural to refer to the excitations as quantum particles ("photons"). Then the state above has 1 photon in the model 1, 3 photons in the mode 2, 1 photon in the mode 4, etc., and the energy of the state (above the ground state) is the sum of the photon energies. Notice that the photons here are treated as indistinguishable particles. They obey Bose-Einstein statistics; that is, we count each set of occupation numbers precisely once.

Now consider the limit of a very big cavity. In the interior, away from the boundaries, we can approximate the eigenfunctions as plane waves of definte wavenumber k or momentum p ($p = \hbar k = k$). In a box with periodic boundary conditions, the eigenfunctions are exactly plane waves with quantized momenta, but the quantized values approximate a continuum as the box is made large. If each mode is a harmonic oscillator, we will have operators a and $a^\dagger$ for each mode. It is useful to give the eigenmodes a continuum normalization, or to modify the commutation relation (69) to

$$[a_p, a_k^\dagger] = (2\pi)^d \delta^{(d)}(p - k) \tag{72}$$

for a box in d space dimensions. The operator a_p lowers the number of photons in the mode p by 1, and the operator $a_p^\dagger$ raises the number of photons in the model p by 1, so we call these *creation and annihilation operators*.

Let $|0\rangle$ be the state with no photons. Then a typical state of the Hilbert space is

$$|p, q, k\rangle = a_p^\dagger a_q^\dagger a_k^\dagger |0\rangle \tag{73}$$

and, if $\epsilon(p)$ is the energy of a photon in the model p, the energy of this state is

$$\mathcal{H}|p, q, k\rangle - \mathcal{H}_0 = \epsilon(p) + \epsilon(q) + \epsilon(k) \ . \tag{74}$$

This space of states is called *Fock space*, after Vladimir Fock (also one of the inventors of the Hartree-Fock approximation).

The Fourier transform of a_p gives an operator $a(x)$ which represents the wavefunction of a photon localized at the point x,

$$a(\vec{x}) = \int \frac{d^d p}{(2\pi)^d} e^{i\vec{p}\cdot\vec{x}} a_p \ . \tag{75}$$

It is not difficult to show that

$$[a(x), a^\dagger(y)] = \delta^{(d)}(x - y) \tag{76}$$

Now consider the operators

$$\Phi(t, \vec{x}) = \int \frac{d^d p}{(2\pi)^d} \left(a_p e^{-i\epsilon(p)t + i\vec{p}\cdot\vec{x}}\right) \qquad \Phi^\dagger(t, \vec{x}) = \int \frac{d^d p}{(2\pi)^d} \left(a_p^\dagger e^{+i\epsilon(p)t - i\vec{p}\cdot\vec{x}}\right) \tag{77}$$

where $\epsilon = p^2/2m$. With a little effort, you can show that $\Phi(x, t)$ satisfies the Schrödinger equation

$$\left[i\frac{\partial}{\partial t} + \frac{1}{2m}\nabla^2\right]\Phi(t, \vec{x}) = 0 \tag{78}$$

The field $\Phi(t, \vec{x})$ has the interesting property that it destroys one photon, and that the value of the destruction matrix element

$$\langle 0| \Phi(t, \vec{x}) |p\rangle = e^{-i\epsilon(p)t + i\vec{p}\cdot\vec{x}} \tag{79}$$

is the wavefuntion from which the photon was taken. With a little more effort, you can show that

$$\mathcal{H} = \int d^d x \Phi^\dagger \left[-\frac{1}{2m} \nabla^2 \right] \Phi \tag{80}$$

is exactly equal to the sum of oscillator Hamiltonians

$$\mathcal{H} = \int \frac{d^3 p}{(2\pi)^d} \, \frac{p^2}{2m} a_p^\dagger a_p \ . \tag{81}$$

Creation and annihilation operators and Fock space appear ubiquitously in the analysis of systems with many degrees of freedom. So if these concepts are unfamiliar to you, please consult a graduate quantum mechanics book to learn more.