DENSITY OPERATORS AND PURE VERSUS MIXED ENSEMBLES

Polarized Versus Unpolarized Beams

The formalism of quantum mechanics developed so far makes statistical predictions on an ensemble, that is, a collection, of identically prepared physical systems.

More precisely, in such an ensemble all members are supposed to be characterized by the same state ket $|\alpha\rangle$.

A good example of this is a beam of silver atoms coming out of an SG filtering apparatus.

Every atom in the beam has its spin pointing in the same direction, namely, the direction determined by the inhomogeneity of the magnetic field of the filtering apparatus.

We have not yet discussed how to describe quantum mechanically an ensemble of physical systems for which some, say 60%, are characterized by $|\alpha>$, and the remaining 40% are characterized by some other ket $|\beta>$.

To illustrate vividly the incompleteness of the formalism developed so far, let us consider silver atoms coming directly out of a hot oven, yet to be subjected to a filtering apparatus of the Stern-Gerlach type.

On symmetry grounds we expect that such atoms have random spin orientations; in other words, there should be no preferred direction associated with such an ensemble of atoms.

According to the formalism developed so far, the most general state ket of a spin ½ system is given by

$$|\alpha\rangle = c_+|+\rangle + c_-|-\rangle$$

Is this equation capable of describing a collection of atoms with random spin orientations?

The answer is clearly no;

The above equation characterizes a state ket whose spin is pointing in some definite direction, namely, in the direction of $\hat{n}$, whose polar and azimuthal angles, β and α, respectively, are obtained by solving

We had

$$\frac{c_+}{c_-} = \frac{\cos(\beta/2)}{e^{i\alpha}\sin(\beta/2)}$$

$$\chi = \begin{pmatrix} \langle +|\alpha\rangle \\ \langle -|\alpha\rangle \end{pmatrix} \equiv \begin{pmatrix} c_+ \\ c_- \end{pmatrix} = \begin{pmatrix} \cos\left(\frac{\beta}{2}\right)e^{-i\alpha/2} \\ \sin\left(\frac{\beta}{2}\right)e^{i\alpha/2} \end{pmatrix}$$

To cope with a situation of this kind we introduce the concept of fractional population, or probability weight.

An ensemble of silver atoms with completely random spin orientation can be viewed as a collection of silver atoms in which 50% of the members of the ensemble are characterized by |+> and the remaining 50% by |->.

We specify such an ensemble by assigning

$$w_+ = 0.5, \qquad w_- = 0.5$$

where w_+ and w_- are the fractional population for spin-up and -down, respectively.

Because there is no preferred direction for such a beam, it is reasonable to expect that this same ensemble can be regarded equally well as a 50-50 mixture of $|S_X;+>$ and $|S_X;->$.

The mathematical formalism needed to accomplish this will appear shortly.

It is very important to note that we are simply introducing here two real numbers w_+ and w_-.

There is no information on the relative phase between the spin-up and the spin-down ket.

Quite often we refer to such a situation as an incoherent mixture of spin-up and spin-down states.

What we are doing here is to be clearly distinguished from what we did with a coherent linear superposition, for example,

$$\left(\frac{1}{\sqrt{2}}\right)|+\rangle+\left(\frac{1}{\sqrt{2}}\right)|-\rangle$$

where the phase relation between |+> and |-> contains vital information on the spin orientation in the xy-plane, in this case in the positive x-direction.

In general, we should not confuse w_+ and w_- with $|c_+|^2$ and $|c_-|^2$.

The probability concept associated with w_+ and w_- is much closer to that encountered in classical probability theory.

The situation encountered in dealing with silver atoms directly from the hot oven may be compared with that of a graduating class in which 50% of the graduating seniors are male, the remaining 50% female.

When we pick a student at random, the probability that the particular student is male (or female) is 0.5.

Whoever heard of a student referred to as a coherent linear superposition of male and female with a particular phase relation?

The beam of silver atoms coming directly out of the oven is an example of a completely random ensemble; **unpolarized** because there is no preferred direction for spin orientation.

In contrast, the beam that has gone through a selective Stern-Gerlach-type measurement is an example of a **pure ensemble**; the beam is said to be **polarized** because all members of the ensemble are characterized by a single common ket that describes a state with spin pointing in some definite direction.

To appreciate the difference between a completely random ensemble and a pure ensemble, let us consider a rotatable SG apparatus where we can vary the direction of the inhomogeneous **B** just by rotating the apparatus.

When a completely unpolarized beam directly out of the oven is subjected to such an apparatus, we always obtain two emerging beams of *equal intensity no matter what the orientation of the apparatus may be.*

In contrast, if a polarized beam is subjected to such an apparatus, the relative intensities of the two emerging beams vary as the apparatus is rotated.

For some particular orientation the ratio of the intensities actually becomes one to zero.

In fact, the formalism we developed in Chapter 1 tells us that the relative intensities are simply $\cos^2(\beta/2)$ and $\sin^2(\beta/2)$, where β is the angle between the spin direction of the atoms and the direction of the inhomogeneous magnetic field in the SG apparatus.

A complete random ensemble and a pure ensemble can be regarded as the extremes of what is known as a **mixed ensemble**.

In a mixed ensemble a certain fraction—for example, 70%—of the members are characterized by a state ket $|\alpha\rangle$, the remaining 30% by $|\beta\rangle$.

In such a case the beam is said to be **partially polarized**.

Here $|\alpha\rangle$ and $|\beta\rangle$ need not even be orthogonal; we can, for example, have 70% with spin in the positive x-direction and 30% with spin in the negative z-direction.

Ensemble Averages and Density Operator

We now present the density operator formalism, pioneered by J. von Neumann in 1927, that quantitatively describes physical situations with mixed as well as pure ensembles.

Our general discussion here is not restricted to spin ½ systems, but for illustrative purposes we return repeatedly to spin ½ systems.

A pure ensemble by definition is a collection of physical systems such that every member is characterized by the same ket $|\alpha\rangle$.

In contrast, in a mixed ensemble, a fraction of the members with relative population w_1 are characterized by $|\alpha^{(1)}\rangle$, some other fraction with relative population w_2, by $|\alpha^{(2)}\rangle$, and so on.

Roughly speaking, a mixed ensemble can be viewed as a mixture of pure ensembles, just as the name suggests.

The fractional populations are constrained to satisfy the normalization condition

$$\sum_i w_i = 1$$

As we mentioned previously, $|\alpha\rangle$ and $|\beta\rangle$ need not be orthogonal.

Furthermore, the number of terms in the i sum of $\sum_i w_i = 1$ need not coincide with the dimensionality N of the ket space; it can easily exceed N.

For example, for spin ½ systems with N = 2, we may consider 40% with spin in the positive z-direction, 30% with spin in the positive x-direction, and the remaining 30% with spin in the negative y-direction.

Suppose we make a measurement on a mixed ensemble of some observable A.

We may ask what is the average measured value of A when a large number of measurements are carried out.

The answer is given by the ensemble average of A, which is defined by

$$[A] \equiv \sum_i w_i \langle \alpha^{(i)} | A | \alpha^{(i)} \rangle$$
$$= \sum_{\iota} \sum_{a'} w_\iota |\langle a' | \alpha^{(\iota)} \rangle|^2 a'$$

where |a'> is an eigenket of A.

Recall that $\langle \alpha^{(i)} | A | \alpha^{(i)} \rangle$ is the usual quantum mechanical expectation value of A taken with respect to state $|\alpha^{(i)}>$.

The last equation for [A] tells us that these expectation values must further be weighted by the corresponding fractional populations w_i.

Notice how probabilistic concepts enter twice; first in $|\langle a' | \alpha^{(i)} \rangle|^2$ for the quantum-mechanical probability for state $|\alpha^{(i)}>$ to be found in an A eigenstate |a'> second, in the probability factor w_i for finding in the ensemble a quantum-mechanical state characterized by$|\alpha^{(i)}>$.

We can now rewrite ensemble average [A] using a more general basis, {|b'>}:

$$[A] = \sum_i w_i \sum_{b'} \sum_{b''} \langle \alpha^{(i)}|b'\rangle\langle b'|A|b''\rangle\langle b''|\alpha^{(i)}\rangle$$

$$= \sum_{b'} \sum_{b''} \left(\sum_i w_i \langle b''|\alpha^{(i)}\rangle\langle \alpha^{(i)}|b'\rangle \right) \langle b'|A|b''\rangle$$

The number of terms in the sum of the b' (b") is just the dimensionality of the ket space, while the number of terms in the sum of the *i* depends on how the mixed ensemble is viewed as a mixture of pure ensembles.

Notice that in this form the basic property of the ensemble which does not depend on the particular observable A is factored out.

This motivates us to define the density operator ρ as follows:

$$\rho \equiv \sum_i w_i |\alpha^{(i)}\rangle\langle\alpha^{(i)}|$$

The elements of the corresponding density matrix have the following form:

$$\langle b''|\rho|b'\rangle = \sum_i w_i \langle b''|\alpha^{(i)}\rangle\langle\alpha^{(i)}|b'\rangle$$

The density operator contains all the physically significant information we can possibly obtain about the ensemble in question.

Returning to

$$[A] = \sum_i w_i \sum_{b'} \sum_{b''} \langle \alpha^{(i)} | b' \rangle \langle b' | A | b'' \rangle \langle b'' | \alpha^{(i)} \rangle$$

$$= \sum_{b'} \sum_{b''} \left(\sum_i w_i \langle b'' | \alpha^{(i)} \rangle \langle \alpha^{(i)} | b' \rangle \right) \langle b' | A | b'' \rangle$$

we see that the ensemble average can be written as

$$[A] = \sum_{b'} \sum_{b''} \langle b'' | \rho | b' \rangle \langle b' | A | b'' \rangle$$

$$= \mathrm{tr}(\rho A).$$

Because the trace is independent of representations, tr(ρA) can be evaluated using any convenient basis.

As a result, [A]=tr(ρA) is an extremely powerful relation.

There are two properties of the density operator worth recording.

First, the density operator is Hermitian, as is evident from $\rho = \sum_i w_i |\alpha^{(i)}\rangle\langle\alpha^{(i)}|$.

Second, the density operator satisfies the normalization condition

$$\begin{aligned}\mathrm{tr}(\rho) &= \sum_i \sum_{b'} w_i \langle b'|\alpha^{(i)}\rangle\langle\alpha^{(i)}|b'\rangle \\ &= \sum_i w_i \langle \alpha^{(i)}|\alpha^{(i)}\rangle \\ &= 1.\end{aligned}$$

Because of the Hermiticity and the normalization condition, for spin ½ systems with dimensionality 2 the density operator, or the corresponding density matrix, is characterized by three independent real parameters.

Four real numbers characterize a 2×2 Hermitian matrix.

However, only three are independent because of the normalization condition.

The three numbers needed are $[S_x]$, $[S_y]$, and $[S_z]$; the reader may verify that knowledge of these three ensemble averages is sufficient to reconstruct the density operator.

The manner in which a mixed ensemble is formed can be rather involved.

We may mix pure ensembles characterized by all kinds of $|\alpha^{(i)}>$'s with appropriate w_i's; yet for spin ½ systems three real numbers completely characterize the ensemble in question.

This strongly suggests that a mixed ensemble can be decomposed into pure ensembles in many different ways.

A problem to illustrate this point appears at the end of this chapter.

A pure ensemble is specified by $w_i = 1$ for some $|\alpha^{(i)}\rangle$—with i = n, for instance—and $w_i = 0$ for all other conceivable state kets, so the corresponding density operator is written as

$$\rho = |\alpha^{(n)}\rangle\langle\alpha^{(n)}|$$

with no summation.

Clearly, the density operator for a pure ensemble is idempotent, that is,

$$\rho^2 = \rho$$

Or equivalently

$$\rho(\rho - 1) = 0$$

Thus, for a pure ensemble only, in addition to $tr(\rho) = 1$ we have

$$\mathrm{tr}(\rho^2) = 1$$

The eigenvalues of the density operator for a pure ensemble are zero or one, as can be seen by inserting a complete set of base kets that diagonalize the Hermitian operator ρ between ρ and $(\rho - 1)$ of $\rho(\rho - 1) = 0$.

When diagonalized, the density matrix for a pure ensemble must therefore look like

$$\rho \doteq \begin{pmatrix} 0 & & & & & & & & 0 \\ & 0 & & & & & & & \\ & & \ddots & & & & & & \\ & & & 0 & & & & & \\ & & & & 1 & & & & \\ & & & & & 0 & & & \\ & & & & & & 0 & & \\ & & & & & & & 0 & \\ & & & & & & & & \ddots \\ 0 & & & & & & & & 0 \end{pmatrix}$$

It can be shown that $tr(\rho^2)$ is maximal when the ensemble is pure;

for a mixed ensemble $tr(\rho^2)$ is a positive number less than one.

Given a density operator, let us see how we can construct the corresponding density matrix in some specified basis.

To this end we first recall that

$$|\alpha\rangle\langle\alpha| = \sum_{b'}\sum_{b''}|b'\rangle\langle b'|\alpha\rangle\langle\alpha|b''\rangle\langle b''|$$

this shows that we can form the square matrix corresponding to $|\alpha^{(i)}\rangle\langle\alpha^{(i)}|$ by combining, in the sense of outer product, the column matrix formed by $\langle b'|\alpha^{(i)}\rangle$ with the row matrix formed by $\langle\alpha^{(i)}|b''\rangle$, which, of course, is equal to $\langle b''|\alpha^{(i)}\rangle^{*}$.

The final step is to sum such square matrices with weighting factors w_i as indicated in $\rho = \sum_i w_i |\alpha^{(i)}\rangle\langle\alpha^{(i)}|$.

The final form agrees with:

$$\langle b''|\rho|b'\rangle = \sum_i w_i \langle b''|\alpha^{(i)}\rangle\langle\alpha^{(i)}|b'\rangle$$

as expected

It is instructive to study several examples, all referring to spin ½ systems.

Example 1

A completely polarized beam with Sz +.

$$\rho = |+\rangle\langle +| \doteq \begin{pmatrix} 1 \\ 0 \end{pmatrix}(1,0)$$

$$= \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

Example 2

A completely polarized beam with Sx +.

$$\rho = |S_x; \pm\rangle\langle S_x; \pm| = \left(\frac{1}{\sqrt{2}}\right)(|+\rangle \pm |-\rangle)\left(\frac{1}{\sqrt{2}}\right)(\langle + | \pm \langle - |)$$

$$\doteq \begin{pmatrix} \frac{1}{2} & \pm\frac{1}{2} \\ \pm\frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

The ensembles of Examples 1 and 2 are both pure.

Example 3 An unpolarized beam.

This can be regarded as an incoherent mixture of a spin-up ensemble and a spin-down ensemble with equal weights (50% each):

$$\rho = \left(\tfrac{1}{2}\right)|+\rangle\langle+| + \left(\tfrac{1}{2}\right)|-\rangle\langle-|$$

$$\doteq \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{pmatrix},$$

which is just the identity matrix divided by 2.

As we remarked earlier, the same ensemble can also be regarded as an incoherent mixture of an Sx+ ensemble and an Sx- ensemble with equal weights.

It is gratifying that our formalism automatically satisfies the expectation

$$\begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{pmatrix} = \frac{1}{2}\begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} + \frac{1}{2}\begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

where we see from Example 2 that the two terms on the right-hand side are the density matrices for pure ensemble with Sx+ and Sx-.

Because ρ in this case is just the identity operator divided by 2 (the dimensionality), we have

$$\mathrm{tr}(\rho S_x) = \mathrm{tr}(\rho S_y) = \mathrm{tr}(\rho S_z) = 0$$

where we used the fact that S_k is traceless.

Thus for the ensemble average of **S** we have

$$[\mathbf{S}] = 0$$

This is reasonable because there should be no preferred spin direction in a completely random ensemble of spin ½ systems.

Example 4

As an example of a partially polarized beam, let us consider a 75-25 mixture of two pure ensembles, one with Sz+ and the other with Sx+:

$$w(S_z+) = 0.75, \qquad w(S_x+) = 0.25$$

The corresponding ρ can be represented by

$$\rho \doteq \tfrac{3}{4}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \tfrac{1}{4}\begin{pmatrix} \tfrac{1}{2} & \tfrac{1}{2} \\ \tfrac{1}{2} & \tfrac{1}{2} \end{pmatrix}$$

$$= \begin{pmatrix} \tfrac{7}{8} & \tfrac{1}{8} \\ \tfrac{1}{8} & \tfrac{1}{8} \end{pmatrix},$$

$$[S_x] = \frac{\hbar}{8}, \qquad [S_y] = 0, \qquad [S_z] = \frac{3\hbar}{8}$$

Time Evolution of Ensembles

How does the density operator ρ change as a function of time?

Let us suppose that at some time t_0 the density operator is given by

$$\rho(t_0) = \sum_i w_i |\alpha^{(i)}\rangle\langle\alpha^{(i)}|$$

If the ensemble is to be left undisturbed, we cannot change the fractional population w_i.

So the change in ρ is governed solely by the time evolution of state ket $|\alpha^{(i)}\rangle$:

$$|\alpha^{(i)}\rangle \quad \text{at} \quad t_0 \rightarrow |\alpha^{(i)}, t_0; t\rangle$$

From the fact that $|\alpha^{(i)}, t_0; t\rangle$ satisfies the Schrodinger equation we obtain

$$i\hbar\frac{\partial \rho}{\partial t} = \sum_i w_i \left(H|\alpha^{(i)}, t_0; t\rangle\langle\alpha^{(i)}, t_0; t| - |\alpha^{(i)}, t_0; t\rangle\langle\alpha^{(i)}, t_0; t|H \right) = -[\rho, H]$$

This looks like the Heisenberg equation of motion except that the sign is wrong!

This is not disturbing because ρ is not a dynamic observable in the Heisenberg picture.

It is amusing that $i\hbar \frac{\partial \rho}{\partial t} = -[\rho, H]$ can be regarded as the quantum-mechanical analogue of Liouville's theorem in classical statistical mechanics,

$$\frac{\partial \rho_{\text{classical}}}{\partial t} = -[\rho_{\text{classical}}, H]_{\text{classical}}$$

where $\rho_{\text{classical}}$ stands for the density of representative points in phase space.

Thus the name density operator for the ρ appearing in $i\hbar \frac{\partial \rho}{\partial t} = -[\rho, H]$ is indeed appropriate.

The classical analogue of $[A] = tr(\rho A)$ for the ensemble average of some observable A is given by

$$A_{\text{average}} = \frac{\int \rho_{\text{classical}} A(q, p)\, d\Gamma_{q,p}}{\int \rho_{\text{classical}}\, d\Gamma_{q,p}}$$

where $d\Gamma_{q,p}$ stands for a volume element in phase space.

Continuum Generalizations

So far we have considered density operators in ket space where the base kets are labeled by the discrete-eigenvalues of some observable.

The concept of density matrix can be generalized to cases where the base kets used are labeled by continuous eigenvalues.

In particular, let us consider the ket space spanned by the position eigenkets $|x'\rangle$.

The analogue of

$$[A] = \sum_{b'} \sum_{b''} \langle b'' | \rho | b' \rangle \langle b' | A | b'' \rangle$$

$$= \mathrm{tr}(\rho A).$$

is given by

$$[A] = \int d^3x' \int d^3x'' \langle \mathbf{x}'' | \rho | \mathbf{x}' \rangle \langle \mathbf{x}' | A | \mathbf{x}'' \rangle$$

The density matrix here is actually a function of x' and x", namely,

$$\langle \mathbf{x}'' | \rho | \mathbf{x}' \rangle = \langle \mathbf{x}'' | \left(\sum_i w_i | \alpha^{(i)} \rangle \langle \alpha^{(i)} | \right) | \mathbf{x}' \rangle$$

$$= \sum_i w_i \psi_i(\mathbf{x}'') \psi_i^*(\mathbf{x}'),$$

where ψ_i is the wave function corresponding to the state ket $|\alpha^{(i)}\rangle$.

Notice that the diagonal element (that is, x' = x") of this is just the weighted sum of the probability densities.

Once again, the term density matrix is indeed appropriate.

In continuum cases, too, it is important to keep in mind that the same mixed ensemble can be decomposed in different ways into pure ensembles.

For instance, it is possible to regard a "realistic" beam of particles either as a mixture of plane-wave states (monoenergetic free-particle states) or as a mixture of wave-packet states.

Quantum Statistical Mechanics

We conclude this section with a brief discussion on the connection between the density operator formalism and statistical mechanics.

Let us first record some properties of completely random and of pure ensembles.

The density matrix of a completely random ensemble looks like

$$\rho \doteq \frac{1}{N}\begin{pmatrix} 1 & & & & & & 0 \\ & 1 & & & & & \\ & & 1 & & & & \\ & & & \ddots & & & \\ & & & & 1 & & \\ & & & & & 1 & \\ 0 & & & & & & 1 \end{pmatrix}$$

in any representation [see Example 3]: ρ

$$= \left(\frac{1}{2}\right)|+\rangle\langle+| + \left(\frac{1}{2}\right)|-\rangle\langle-| = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

This follows from the fact that all states corresponding to the base kets with respect to which the density matrix is written are equally populated.

In contrast, in the basis where ρ is diagonalized, we have $\rho \doteq \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & 1 & \vdots \\ 0 & \cdots & 0 \end{pmatrix}$ for the matrix representation of the density operator for a pure ensemble.

The two diagonal matrices $\rho \doteq \frac{1}{N}\begin{pmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{pmatrix}$ and $\rho \doteq \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & 1 & \vdots \\ 0 & \cdots & 0 \end{pmatrix}$, both satisfying the normalization requirement $tr(\rho)=1$, cannot look more different.

It would be desirable if we could somehow construct a quantity that characterizes this dramatic difference.

Thus we define a quantity called σ by

$$\sigma = -\operatorname{tr}(\rho \ln \rho)$$

The logarithm of the operator ρ may appear rather formidable, but the meaning of $\sigma = -tr(\rho ln \rho)$ is quite unambiguous if we use the basis in which ρ is diagonal:

$$\sigma = -\sum_k \rho_{kk}^{(\mathrm{diag})} \ln \rho_{kk}^{(\mathrm{diag})}$$

Because each element $\rho_{kk}^{(diag)}$ is a real number between 0 and 1, σ is necessarily positive semidefinite.

For a completely random ensemble $\rho \doteq \frac{1}{N}\begin{pmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{pmatrix}$, we have

$$\sigma = -\sum_{k=1}^{N} \frac{1}{N} \ln\left(\frac{1}{N}\right) = \ln N$$

In contrast, for a pure ensemble $\rho \doteq \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & 1 & \vdots \\ 0 & \cdots & 0 \end{pmatrix}$ we have

$$\sigma = 0$$

where we have used $\rho_{kk}^{(diag)} = 0$ or $ln\rho_{kk}^{(diag)} = 0$ for each term in $\sigma = -\sum_k \rho_{kk}^{(diag)} ln\rho_{kk}^{(diag)}$.

We now argue that physically σ can be regarded as a quantitative measure of disorder.

A pure ensemble is an ensemble with a maximum amount of order because all members are characterized by the same quantum-mechanical state ket;

it may be likened to marching soldiers in a well-regimented army.

According to $\sigma = 0$, a vanishes for such an ensemble.

On the other extreme, a completely random ensemble, in which all quantum-mechanical states are equally likely, may be likened to drunken soldiers wandering around in random directions.

According to $\sigma = -\sum_{k=1}^{N} \frac{1}{N} lnN = lnN$, σ is large;

indeed, we will show later that lnN is the maximum possible value for a subject to the normalization condition

$$\sum_k \rho_{kk} = 1$$

In thermodynamics we learn that a quantity called entropy measures disorder.

It turns out that our σ is related to the entropy per constituent member, denoted by S, of the ensemble via

$$S = k\sigma$$

where k is a universal constant identifiable with the Boltzmann constant.

In fact, S=kσ may be taken as the definition of entropy in quantum statistical mechanics.

We now show how the density operator ρ can be obtained for an ensemble in thermal equilibrium.

The basic assumption we make is that nature tends to maximize a subject to the constraint that the ensemble average of the Hamiltonian has a certain prescribed value.

To justify this assumption would involve us in a delicate discussion of how equilibrium is established as a result of interactions with the environment, which is beyond the scope of this book.

In any case, once thermal equilibrium is established, we expect $\frac{\partial \rho}{\partial t} = 0$.

Because of $i\hbar \frac{\partial \rho}{\partial t} = -[\rho, H]$, this means that ρ and H can be simultaneously diagonalized.

So the kets used in writing $\sigma = -\sum_k \rho_{kk}^{(diag)} ln \rho_{kk}^{(diag)}$ may be taken to be energy eigenkets.

With this choice ρ_{kk} stands for the fractional population for an energy eigenstate with energy eigenvalue E_k.

Let us maximize σ by requiring that

$$\delta\sigma = 0$$

However, we must take into account the constraint that the ensemble average of H has a certain prescribed value.

In the language of statistical mechanics, [H] is identified with the internal energy per constituent denoted by U:

$$[H] = \mathrm{tr}(\rho H) = U$$

In addition, we should not forget the normalization constraint $\sum_k \rho_{kk} = 1$.

So our basic task is to require $\delta\sigma=0$ subject to the constraints

$$\delta[H]=\sum_k \delta\rho_{kk}E_k=0$$

$$\delta(\mathrm{tr}\,\rho)=\sum_k \delta\rho_{kk}=0$$

We can most readily accomplish this by using Lagrange multipliers.

We obtain

$$\sum_k \delta\rho_{kk}\left[(\ln\rho_{kk}+1)+\beta E_k+\gamma\right]=0$$

which for an arbitrary variation is possible only if

$$\rho_{kk} = \exp(-\beta E_k - \gamma - 1)$$

The constant γ can be eliminated using the normalization condition $\sum_k \rho_{kk} = 1$, and our final result is

$$\rho_{kk} = \frac{\exp(-\beta E_k)}{\sum_l^N \exp(-\beta E_l)}$$

We recognize the denominator of $\rho_{kk} = \frac{exp(-\beta E_k)}{\sum_l^N exp(-\beta E_l)}$ as the partition function

It is to be understood throughout that the sum is over distinct energy eigenstates;

if there is degeneracy we must sum over states with the same energy eigenvalue.

The density matrix element $\rho_{kk} = \frac{exp(-\beta E_k)}{\sum_l^N exp(-\beta E_l)}$ is appropriate for what is known in statistical mechanics as a canonical ensemble.

Had we attempted to maximize σ without the internal-energy constraint $\delta[H] = \sum_k \delta\rho_{kk}E_k = 0$, we would have obtained instead

$$\rho_{kk} = \frac{1}{N}, \quad \text{(independent of } k\text{)}$$

which is the density matrix element appropriate for a completely random ensemble.

Comparing $\rho_{kk} = \frac{exp(-\beta E_k)}{\sum_l^N exp(-\beta E_l)}$ with $\rho_{kk} = \frac{1}{N}$, we infer that a completely random ensemble can be regarded as the $\beta \to 0$ limit (physically the high-temperature limit) of a canonical ensemble.

We recognize the denominator of $\rho_{kk} = \frac{exp(-\beta E_k)}{\sum_l^N exp(-\beta E_l)}$ as the partition function in statistical mechanics:

$$Z = \sum_k^N \exp(-\beta E_k)$$

It can also be written as

$$Z = \mathrm{tr}(e^{-\beta H})$$

Knowing ρ_{kk} given in the energy basis, we can write the density operator as

$$\rho = \frac{e^{-\beta H}}{Z}$$

This is the most basic equation from which everything follows.

We can immediately evaluate the ensemble average of any observable A:

$$[A] = \frac{\mathrm{tr}(e^{-\beta H} A)}{Z}$$

$$= \frac{\left[\sum_k^N \langle A \rangle_k \exp(-\beta E_k)\right]}{\sum_k^N \exp(-\beta E_k)}$$

In particular, for the internal energy per constituent we obtain

$$U = \frac{\left[\sum_k^N E_k \exp(-\beta E_k)\right]}{\sum_k^N \exp(-\beta E_k)}$$

$$= -\frac{\partial}{\partial \beta}(\ln Z),$$

a formula well known to every student of statistical mechanics.

The parameter β is related to the temperature T as follows:

$$\beta = \frac{1}{kT} \tag{3.4.55}$$

where k is the Boltzmann constant. It is instructive to convince ourselves of this identification by comparing the ensemble average $[H]$ of simple harmonic oscillators with the kT expected for the internal energy in the classical limit, which is left as an exercise. We have already commented that in the high-temperature limit, a canonical ensemble becomes a completely random ensemble in which all energy eigenstates are equally populated. In the opposite low-temperature limit ($\beta \to \infty$), (3.4.48) tells us that a canonical ensemble becomes a pure ensemble where only the ground state is populated.

As a simple illustrative example, consider a canonical ensemble made up of spin ½ systems, each with a magnetic moment $e\hbar/2m_e c$ subjected to a uniform magnetic field in the z-direction.

$$\rho \doteq \frac{\begin{pmatrix} e^{-\beta\hbar\omega/2} & 0 \\ 0 & e^{\beta\hbar\omega/2} \end{pmatrix}}{Z}$$

$$Z = e^{-\beta\hbar\omega/2} + e^{\beta\hbar\omega/2}$$

Partition function

$$[S_x] = [S_y] = 0, \quad [S_z] = -\left(\frac{\hbar}{2}\right)\tanh\left(\frac{\beta\hbar\omega}{2}\right)$$

The ensemble average of the magnetic moment component is just $e/m_e c$ times $[S_z]$. The paramagnetic susceptibility x may be computed from

$$\left(\frac{e}{m_e c}\right)[S_z] = \chi B$$

In this way we arrive at Brillouin's formula for x:

$$\chi = \left(\frac{|e|\hbar}{2m_e cB}\right)\tanh\left(\frac{\beta\hbar\omega}{2}\right)$$

EXERCISE: An ensemble of silver atoms (each with spin ½) is prepared so that 60% of the atoms are in the $S_z = \frac{\hbar}{2}$ eigenstate of S_z and 40% of the atoms are in the $S_x = -\frac{\hbar}{2}$ eigenstate of Sx (Sx and Sz are the x and z components of the spin angular momentum operator), (a) Compute the density matrix at time t=0 in the basis of eigenstates of , S_z. (b) Assume that the silver atoms sit in a magnetic field, $\boldsymbol{B} = B_0\hat{\boldsymbol{y}}$, and have a magnetic Hamiltonian, H = μ**S•B**, where μ is the magnetic moment of a silver atom. Compute the density matrix at time t (in the basis of eigenstates of S_z). (c) Compute $\langle S_z(t) \rangle$ at time t = 0 and at time t.

■ **EXERCISE 6.6.** An ensemble of silver atoms (each with spin $\frac{1}{2}$) is prepared so that 60% of the atoms are in the $S_z = +\frac{\hbar}{2}$ eigenstate of $\hat{S}_z$ and 40% of the atoms are in the $S_x = -\frac{\hbar}{2}$ eigenstate of $\hat{S}_x$ ($\hat{S}_x$ and $\hat{S}_z$ are the x and z components of the spin angular momentum operator). (a) Compute the density matrix at time $t = 0$ in the basis of eigenstates of $\hat{S}_z$. (b) Assume that the silver atoms sit in a magnetic field, $\mathbf{B} = B_0\hat{\mathbf{y}}$, and have a magnetic Hamiltonian, $\hat{H} = \mu\mathbf{S} \cdot \mathbf{B}$, where μ is the magnetic moment of a silver atom. Compute the density matrix at time t (in the basis of eigenstates of $\hat{S}_z$). (c) Compute $\langle S_z(t) \rangle$ at time $t = 0$ and at time t.

(a) Let $|i_\pm\rangle$ denote the eigenstates of $\hat{S}_k (k = x, y, z)$ with eigenvalues $\pm(\hbar/2)$ so $(\hat{S}_k|k_\pm\rangle = \pm(\hbar/2)|k_\pm\rangle)$. The density operator at time $t = 0$ is the mixed state

$$\hat{\rho}(0) = \frac{6}{10}|z_+\rangle\langle z_+| + \frac{4}{10}|x_-\rangle\langle x_-|. \tag{1}$$

Now note that the matrix representation of the components $\hat{S}_x$, $\hat{S}_y$, and $\hat{S}_z$ in the basis of eigenstates of $\hat{S}_z$ are

$$S_x = \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad S_y = \frac{\hbar}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \text{and} \quad S_z = \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{2}$$

The eigenstates of $\hat{S}_x$ and $\hat{S}_y$, in the basis of eigenstates of $\hat{S}_z$ are

$$\begin{pmatrix} \langle z_+|x_\pm\rangle \\ \langle z_-|x_\pm\rangle \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ \pm 1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} \langle z_+|y_\pm\rangle \\ \langle z_-|y_\pm\rangle \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ \pm i \end{pmatrix}, \tag{3}$$

respectively. The eigenstates of $\hat{S}_z$ in the basis of eigenstates of $\hat{S}_z$ are

$$\begin{pmatrix} \langle z_+|z_+\rangle \\ \langle z_-|z_+\rangle \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} \langle z_+|z_-\rangle \\ \langle z_-|z_-\rangle \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{4}$$

Using these results we find the initial density matrix in the $|z_\pm\rangle$ basis,

$$\rho(0) = \begin{pmatrix} \frac{4}{5} & -\frac{1}{5} \\ -\frac{1}{5} & \frac{1}{5} \end{pmatrix}. \tag{5}$$

(**b**) The Hamiltonian is $\hat{H} = \frac{1}{2}\mu\hbar B_0(|y_+\rangle\langle y_+| - |y_-\rangle\langle y_-|)$ (this is its spectral decomposition). If we let $\rho_{y++} \equiv \langle y_+|\hat{\rho}|y_+\rangle$, then the Liouville equation for various matrix elements of the density matrix in basis $|y_\pm\rangle$ is given by

$$i\frac{\partial\rho_{y++}(t)}{\partial t} = 0, \quad i\frac{\partial\rho_{y+-}(t)}{\partial t} = \mu B_0\rho_{y+-}(t), \quad i\frac{\partial\rho_{y-+}(t)}{\partial t} = -\mu B_0\rho_{y-+}(t) \quad \text{and} \quad i\frac{\partial\rho_{y--}(t)}{\partial t} = 0, \tag{6}$$

Now note that the initial density matrix in the basis $|y_{\pm}\rangle$ is given by

$$\rho(0) = \begin{pmatrix} \frac{1}{2} & \frac{3+2i}{10} \\ \frac{3-2i}{10} & \frac{1}{2} \end{pmatrix}. \tag{7}$$

Using Eq. (7), we can solve Eqs. (6) and write the density matrix at time t in the basis $|y_{\pm}\rangle$. We find

$$\rho(t) = \begin{pmatrix} \frac{1}{2} & \frac{3+2i}{10} e^{-i\mu B_0 t} \\ \frac{3-2i}{10} e^{+i\mu B_0 t} & \frac{1}{2} \end{pmatrix}. \tag{8}$$

We now can transform Eq. (8) to the basis $|z_{\pm}\rangle$ to obtain

$$\rho(t) = \begin{pmatrix} \frac{1}{2}+\frac{3}{10}\cos(\mu B_0 t)+\frac{1}{5}\sin(\mu B_0 t) & \frac{3}{10}\sin(\mu B_0 t)-\frac{1}{5}\cos(\mu B_0 t) \\ \frac{3}{10}\sin(\mu B_0 t)-\frac{1}{5}\cos(\mu B_0 t) & \frac{1}{2}-\frac{3}{10}\cos(\mu B_0 t)-\frac{1}{5}\sin(\mu B_0 t) \end{pmatrix}. \tag{9}$$

(**c**) The average z-component of spin angular momentum at time, $t = 0$ is

$$\langle S_z(0)\rangle = \mathrm{Tr}\,\hat{S}_z\hat{\rho}(0) = \frac{\hbar}{2}\mathrm{Tr}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\begin{pmatrix}\frac{4}{5} & -\frac{1}{5}\\ -\frac{1}{5} & \frac{1}{5}\end{pmatrix} = \frac{3}{10}\hbar. \quad (10)$$

The average z component of spin angular momentum at time t is

$$\langle S_z(t)\rangle = \frac{3}{10}\hbar\cos(\mu B_0 t) + \frac{1}{5}\hbar\sin(\mu B_0 t). \quad (11)$$