15 Identical-Particle Systems: Quantum Statistics. The Density Matrix

In previous chapters we studied some effects of the electron spin that could be treated as perturbations of an otherwise spinless system. This could lead one to believe that spin effects are mere corrections that can be taken into account or not, depending on the required precision of the calculation or the observation one wishes to explain. This conclusion is certainly wrong, since the spin (whether integer or half-integer) can have crucial effects that are not only impossible to obtain by approximate methods, but in many cases determine fundamental features of the system's behavior. In this chapter we will focus on such effects, collectively called exchange effects. Their existence is characteristic of states of two or more identical particles that are invariant with respect to their exchange. This provides an opportunity to briefly discuss some quantum collective phenomena that are a defining feature of condensed matter.

Most real quantum systems are composed of subsystems in different quantum states. For a systematic study of such composite systems – which are mixtures of pure states – it is necessary to adopt a statistical approach and to introduce the relevant concepts for their description. For this reason, in Section 15.5 we present an introduction to the density matrix as a convenient tool for dealing with statistical quantum mixtures. The chapter then concludes with a brief discussion of decoherence.

15.1 Multipartite Systems

We will begin our analysis by studying the structure of the Hamiltonian for a general multipartite system. This can contain from a few particles, such as the electrons of an atom, to an arbitrarily large number, such as the molecules in a macroscopic volume of a homogeneous gas.

The Hamiltonian of a system of N particles (identical or different, for the moment) subject to both internal interactions and external fields, is

$$\widehat{H} = -\hbar^2 \sum_{i=1}^{N} \frac{1}{2m_i} \nabla_i^2 + U(\{\boldsymbol{r}_i\}). \tag{15.1}$$

In the absence of magnetic fields, which we will assume to be the case, U does not depend on the spin operators. We can write U as the sum of two terms

$$U(\{\boldsymbol{r}_i\}) = \sum_i V_i^{\text{ext}}(\boldsymbol{r}_i) + V(\{\boldsymbol{r}_i\}), \tag{15.2}$$

https://doi.org/10.1017/9781009679633.016 Published online by Cambridge University Press

where the potential due to the external fields, V_i^{ext}, is a function of the position of particle i, whereas the potential due to the internal interactions, V, depends normally on the relative coordinates,

$$V(\{\boldsymbol{r}_i\}) = V(\{\boldsymbol{r}_i - \boldsymbol{r}_j\}). \tag{15.3}$$

This can be seen by turning off the external fields. The total momentum must then be conserved, which means that the momentum operator $\widehat{\boldsymbol{P}}$ must commute with the Hamiltonian. Indeed, with

$$\widehat{\boldsymbol{P}} = \sum_i \widehat{\boldsymbol{p}}_i = -i\hbar \sum_i \nabla_i, \tag{15.4}$$

and V written in the form (15.3), we have

$$\begin{aligned}[\widehat{\boldsymbol{P}}, H] = [\widehat{\boldsymbol{P}}, V] &= -i\hbar \sum_i \nabla_i V(\{\boldsymbol{r}_i - \boldsymbol{r}_j\}) \\ &= -i\hbar(V' - V' + V' - V' + \cdots + V' - V') = 0.\end{aligned} \tag{15.5}$$

The result is, in fact, obvious, since it reduces to the demand that the Hamiltonian of a system not acted upon by external fields must be invariant with respect to arbitrary translations of the origin of the coordinates. Hence, V can only depend on internal (relative) coordinates.

15.1.1 Equal-particle systems: Exchange degeneracy

Let us now consider a system where all particles are equal;[1] in this case, (15.1) becomes

$$\widehat{H} = \sum_i \widehat{H}_i(\boldsymbol{r}_i) + V(\{\boldsymbol{r}_i - \boldsymbol{r}_j\}), \tag{15.6}$$

where $\widehat{H}_i(\boldsymbol{r}_i)$ is the Hamiltonian of particle i moving in the common potential. Since V depends only on the relative coordinates, (15.6) is not affected by the exchange of particles (only the order in which the terms appear in the sum changes). We conclude that the Hamiltonian of a system of identical particles is invariant with respect to the permutation of any pair of them, simply because they are all identical.[2]

The invariance of the Hamiltonian with respect to permutations leads to a new type of degeneracy: the exchange degeneracy, which consists in the fact that all the eigenfunctions of $\widehat{H}$ obtained by pairwise exchange (interchange) of equal particles are degenerate (they all have the same energy). However, not all of them are invariant upon permutation of particles. Our main objective here is to precisely construct the simultaneous eigenfunctions of the Hamiltonian and permutation operators. Before we proceed, it is useful to obtain some preliminary results.

[1] Often the system contains one or more particles that are different from the rest, such as the nucleus of an atom and its electrons; those different particles are not considered in the discussion that follows.

[2] We are assuming that we can label the particles in order to identify and number them. This hypothesis is not as unobjectionable in the quantum case as it might seem at first sight, since the quantum (statistical) description is oblivious to any notion of individuality; a consistent description of identical particles is only one that does not identify them, but rather treats them with full equivalence, as will be seen later. This is the reason why quantum theory has only quantum statistics, that is, those based on the identity (actually statistical indifference) of particles. It is common in the quantum literature to find the rather subjective term indistinguishability to refer to this property.

https://doi.org/10.1017/9781009679633.016 Published online by Cambridge University Press

Let $\widehat{P}_{ij}$ be the permutation operator that interchanges the particles i and j; the Hamiltonian $\widehat{H}_P$ with the permuted particles is $\widehat{H}_P = \widehat{P}_{ij}\widehat{H}\widehat{P}_{ij}^{-1}$. From the invariance condition $\widehat{H} = \widehat{H}_P$ we see that $\widehat{P}_{ij}$ and $\widehat{H}$ commute, $\widehat{P}_{ij}\widehat{H} = \widehat{H}\widehat{P}_{ij}$, which means that simultaneous eigenstates of $\widehat{H}$ and $\widehat{P}_{ij}$ can be constructed.

When $\widehat{P}_{ij}$ acts on a function, it interchanges all the variables of the particles i and j, including the spin variables. If $\xi_i = (\boldsymbol{r}_i, \sigma_i)$ is the set of space and spin variables of particle i, we write

$$\widehat{P}_{12} f(\xi_1, \xi_2, \xi_3 \dots) = f(\xi_2, \xi_1, \xi_3 \dots), \tag{15.7}$$

and so on. The eigenvalues of $\widehat{P}_{ij}$ are easily determined as follows. Let $\Psi = \Psi(\xi_1, \xi_2, \xi_3 \dots)$ be an eigenfunction of $\widehat{P}_{ij}$; from the eigenvalue equation

$$\widehat{P}_{ij}\Psi = \lambda_{ij}\Psi; \tag{15.8}$$

we can see that

$$\widehat{P}_{ij}^2\Psi = \lambda_{ij} P_{ij}\Psi = \lambda_{ij}^2\Psi = \Psi, \tag{15.9}$$

that is, after two successive exchanges of i and j, the original function is recovered. Therefore,

$$\lambda_{ij} = \pm 1. \tag{15.10}$$

Consequently, the eigenfunctions of $\widehat{P}_{ij}$ are either invariant or change sign upon permutation of particles i and j. In the first case the wave function is symmetric (S) and in the second, it is antisymmetric (A) with respect to particles i and j. If the wave function is S or A at $t = 0$, it will be so at all t, since λ is an integral of motion.[3]

To illustrate, consider the case of three particles numbered 1, 2 and 3, two of which are in state m with energy E_m and the other in state n with energy E_n. The energy of the system is $E = 2E_m + E_n$. However, the state vector is not well defined because we do not know which particle has which energy. By renumbering the particles in all possible ways (without distinguishing between those with the same energy E_m) we get three different vectors $|E_m E_m E_n\rangle, |E_m E_n E_m\rangle, |E_n E_m E_m\rangle$. The most general state vector is an arbitrary (normalized) linear combination of these,

$$|E\rangle = a|E_m E_m E_n\rangle + b|E_m E_n E_m\rangle + c|E_n E_m E_m\rangle. \tag{15.11}$$

This state vector does not have a well-defined symmetry with respect to $\widehat{P}_{12}$, for example, and does not in general satisfy condition (15.8).[4] This means that $|E\rangle$ changes upon the permutation even though physically nothing has been done; therefore it does

[3] In the quantum literature it is common to blame the indistinguishability of identical particles for the need to introduce the total (anti)symmetrization characteristic of quantum statistics. However, apart from the fact that this notion introduces an unwanted subjective element into the physical description, as already mentioned, there are classical situations where the distinguishability of equal components is irrelevant for statistical purposes. Consider, for example, statistics on the occupancy of hospital beds in a city, a case in which it matters only whether a patient occupies a bed at a given time, and not who occupies it. If, in one of the hospitals (or in all of them, for that matter) the use of floors by specialty is changed from one day to the next, the statistics do not change at all. This indistinguishability does not imply quantum statistics. What is important in the quantum case is that identical particles have the same relevant physical properties.

[4] A side note. This example illustrates a very common phenomenon, namely that the solution to a differential equation may have fewer symmetries than the equation it solves. In such cases it is usually said that there has been a spontaneous breaking of the symmetry (note the metaphorical use of the term

https://doi.org/10.1017/9781009679633.016 Published online by Cambridge University Press

not represent a physical solution in general. We can construct physical solutions by demanding them to be either symmetric ($\lambda_{12} = 1$) or antisymmetric ($\lambda_{12} = -1$) eigenfunctions of $\widehat{P}_{12}$; it suffices to take $b = c$, with a arbitrary in the first instance,

$$|E\rangle_S = a|E_1E_1E_2\rangle + b(|E_1E_2E_1\rangle + b|E_2E_1E_1\rangle) = \widehat{P}_{12}|E\rangle_S, \tag{15.12}$$

and $c = -b$, with $a = 0$ in the second case,

$$|E\rangle_A = b(|E_1E_2E_1\rangle - |E_2E_1E_1\rangle) = -\widehat{P}_{12}|E\rangle_A. \tag{15.13}$$

However, the physical solutions must also be symmetric or antisymmetric with respect to $\widehat{P}_{13}$ and $\widehat{P}_{23}$. To complicate matters, the exchange operators do not necessarily commute with each other. For example,

$$\widehat{P}_{13}\widehat{P}_{12}\Psi(123) = \widehat{P}_{13}\Psi(213) = \Psi(231), \tag{15.14}$$

whereas

$$\widehat{P}_{12}\widehat{P}_{13}\Psi(123) = \widehat{P}_{12}\Psi(321) = \Psi(312). \tag{15.15}$$

In a sense, identical entities are being treated differently. This raises a key question: how do we construct the eigenfunction that describes the physical state of the system?

Let us consider first the simple case of only two identical particles, so that

$$\widehat{H}(\xi_1, \xi_2) = \widehat{H}(\xi_2, \xi_1). \tag{15.16}$$

Due to the symmetry of the problem, we can write the Schrödinger equation

$$\widehat{H}(\xi_1, \xi_2)\Psi(\xi_1, \xi_2) = E\Psi(\xi_1, \xi_2) \tag{15.17}$$

in the equivalent form, obtained by interchanging the two particles,

$$\widehat{H}(\xi_2, \xi_1)\Psi(\xi_2, \xi_1) = E\Psi(\xi_2, \xi_1) \tag{15.18}$$

and use Eq. (15.16) to get

$$\widehat{H}(\xi_1, \xi_2)\Psi(\xi_2, \xi_1) = E\Psi(\xi_2, \xi_1). \tag{15.19}$$

For Eqs. (15.17) and (15.19) to be consistent, it is required that

$$\Psi(\xi_2, \xi_1) = \pm\Psi(\xi_1, \xi_2), \tag{15.20}$$

which is Eq. (15.8) for this particular case. Therefore, the eigenfunctions of $\widehat{H}$ must be symmetrized. This is easily done by considering the symmetrized combinations $\Psi(\xi_1, \xi_2) + \Psi(\xi_2, \xi_1)$ and $\Psi(\xi_1, \xi_2) - \Psi(\xi_2, \xi_1)$ after normalization.

"spontaneous"). For example, the Schrödinger equation of the H atom has spherical symmetry, but only its s-state eigenfunctions have this symmetry.

A more interesting example, which helps to clarify the concept of spontaneous symmetry breaking, occurs in a high-temperature ferromagnetic system with spherical symmetry. If the temperature is slowly lowered beyond the Curie point, the initially disordered spins spontaneously order themselves (i.e., through their interactions) and the material acquires a certain magnetization. The system has undergone a phase transition, moving from its initial spherical state to a state with axial symmetry, that is, the phase transition is accompanied by a spontaneous reduction (breaking) of the symmetry.The density matrix (see Section 4) describing the final state has less symmetry than the Hamiltonian.

https://doi.org/10.1017/9781009679633.016 Published online by Cambridge University Press

15.1.2 Symmetric and antisymmetric states

Similarly, in the general case of more than two particles, the total wave function must be either S or A. These fully symmetrized wave functions are simultaneous eigenfunctions of all exchange operators $\widehat{P}_{ij}$ with the same eigenvalue, so they all commute. It is clear that the states with all λ_{ij} equal (either $+1$ or -1) are those with the maximum symmetry that can be constructed. They are the only ones that treat equal particles in an equivalent way. Not surprisingly, these functions are the most physically interesting, as we will see in the next section.

The symmetric N-particle wave function can easily be constructed from the $N!$ degenerate wave functions $\Psi(\xi_i)$ as follows,

$$\Psi^S = C \sum_{P_k} \widehat{P}_k \Psi(\xi_1, \xi_2, \ldots, \xi_N), \tag{15.21}$$

where the sum runs over all possible pairwise permutations P_k; C is the normalization constant. The antisymmetric wave function is constructed in a similar way, but alternating the signs of each term according to the parity of the permutation. More specifically, if r_k is the number of pairs of particles that must be interchanged to achieve the permutation P_k, we have

$$\Psi^A = C \sum_{k} (-1)^{r_k} \widehat{P}_k \Psi(\xi_1, \xi_2, \ldots, \xi_N). \tag{15.22}$$

The coefficient $(-1)^{r_k}$ is 1 if P_k is even and -1 if P_k is odd; therefore, if we perform a single permutation of any pair of particles in (15.22), Ψ^A changes sign.

The wave functions (15.21) and (15.22) are generally very complicated, and it is difficult to say more about them without going into detail. However, there is one case of great practical importance where the wave functions can be expressed more explicitly, namely when the interaction between the particles is weak enough to be considered a perturbation. In this case, the unperturbed wave functions correspond to $V = 0$, and as Eq. (15.6) shows, the Hamiltonian reduces to the sum of one-particle Hamiltonians, so that Ψ can be written as the product of N one-particle eigenfunctions,

$$\Psi(\xi_1, \xi_2 \ldots, \xi_N) = \psi_{n_1}(\xi_1)\psi_{n_2}(\xi_2)\ldots\psi_{n_N}(\xi_N). \tag{15.23}$$

The N-particle wave function must then be symmetrized or antisymmetrized according to (15.21) or (15.22).

Exercise E15.1. Projectors for bipartite states

Let $|n_1 n_2\rangle = \frac{1}{\sqrt{2}}\left(|n_1\rangle|n_2\rangle + |n_2\rangle|n_1\rangle\right)$ if $|n_1\rangle \neq |n_2\rangle$, $|n_1 n_2\rangle = |n_1\rangle|n_2\rangle$ if $|n_1\rangle = |n_2\rangle$, describe a symmetric state of two identical particles. Prove that $\widehat{\Omega} = \sum_{n_1 \geq n_2} |n_1 n_2\rangle\ \langle n_1 n_2|$ is the projector on symmetric states.

Solution. For convenience we write the state $|n_1 n_2\rangle$ in the form

$$|n_1 n_2\rangle = \frac{C_n}{\sqrt{2}}\left(|n_1\rangle|n_2\rangle + |n_2\rangle|n_1\rangle\right), \tag{15.24}$$

$$C_n = \begin{cases} 1 & \text{when } n_1 \neq n_2, \\ 1/\sqrt{2} & \text{when } n_1 = n_2. \end{cases}$$

https://doi.org/10.1017/9781009679633.016 Published online by Cambridge University Press

To study the effect of $\widehat{\Omega}$ on an arbitrary state $|\psi\rangle = |m_1 m_2\rangle$, we go to the x-representation,

$$\langle x_1 x_2|\widehat{\Omega}|m_1 m_2\rangle = \int dx_1' dx_2' \langle x_1 x_2|\widehat{\Omega}|x_1' x_2'\rangle \langle x_1' x_2'|m_1 m_2\rangle. \tag{15.25}$$

We first determine the matrix element

$$\begin{aligned}
&\langle x_1 x_2|\widehat{\Omega}|x_1' x_2'\rangle \qquad (15.26)\\
&\quad = \sum_{n_1 \geq n_2} \tfrac{1}{2} C_n^2 \langle x_1 x_2| \left(|n_1\rangle|n_2\rangle + |n_2\rangle|n_1\rangle\right) \left(\langle n_1|\langle n_2| + \langle n_2|\langle n_1|\right) |x_1' x_2'\rangle\\
&\quad = \sum_{n_1 \geq n_2} \tfrac{1}{2} C_n^2 \left(\varphi_{n_1}(x_1)\varphi_{n_2}(x_2) + \varphi_{n_2}(x_1)\varphi_{n_1}(x_2)\right)\\
&\qquad \times \left(\varphi_{n_1}^*(x_1')\varphi_{n_2}^*(x_2') + \varphi_{n_2}^*(x_1')\varphi_{n_1}^*(x_2')\right)\\
&\quad = \tfrac{1}{2}\cdot\tfrac{1}{2}\cdot 4\varphi_n(x_1)\varphi_n(x_2)\varphi_n^*(x_1')\varphi_n^*(x_2')\\
&\qquad + \sum_{n_1>n_2} \tfrac{1}{2}\varphi_{n_1}(x_1)\varphi_{n_2}(x_2)\left(\varphi_{n_1}^*(x_1')\varphi_{n_2}^*(x_2') + \varphi_{n_1}^*(x_2')\varphi_{n_2}^*(x_1')\right)\\
&\qquad + \sum_{n_1>n_2} \tfrac{1}{2}\varphi_{n_2}(x_1)\varphi_{n_1}(x_2)\left(\varphi_{n_2}^*(x_1')\varphi_{n_1}^*(x_2') + \varphi_{n_1}^*(x_1')\varphi_{n_2}^*(x_2')\right).
\end{aligned}$$

In the last term we swap n_1 and n_2, $\sum_{n_1>n_2} \to \sum_{n_1<n_2}$, and combine with the previous term to write

$$\begin{aligned}
&\langle x_1 x_2|\widehat{\Omega}|x_1' x_2'\rangle \qquad (15.27)\\
&\quad = \tfrac{1}{2}\sum_{n_1}\sum_{n_2} \varphi_{n_1}(x_1)\varphi_{n_2}(x_2)\left(\varphi_{n_1}^*(x_1')\varphi_{n_2}^*(x_2') + \varphi_{n_1}^*(x_2')\varphi_{n_2}^*(x_1')\right)\\
&\quad = \tfrac{1}{2}\sum_{n_1} \varphi_{n_1}(x_1)\varphi_{n_1}^*(x_1')\sum_{n_2}\varphi_{n_2}(x_2)\varphi_{n_2}^*(x_2')\\
&\qquad + \tfrac{1}{2}\sum_{n_1} \varphi_{n_1}(x_1)\varphi_{n_1}^*(x_2')\sum_{n_2}\varphi_{n_2}(x_2)\varphi_{n_2}^*(x_1')\\
&\quad = \tfrac{1}{2}\delta(x_1 - x_1')\delta(x_2 - x_2') + \tfrac{1}{2}\delta(x_1 - x_2')\delta(x_2 - x_1').
\end{aligned}$$

Introducing this result in (15.25), we get

$$\begin{aligned}
&\langle x_1 x_2|\widehat{\Omega}|m_1 m_2\rangle \qquad (15.28)\\
&\quad = \tfrac{1}{2}\int dx_1' dx_2' \left(\delta(x_1 - x_1')\delta(x_2 - x_2')\right.\\
&\qquad \left. + \delta(x_1 - x_2')\delta(x_2 - x_1')\right) \langle x_1' x_2'|m_1 m_2\rangle\\
&\quad = \tfrac{1}{2}\langle x_1 x_2|m_1 m_2\rangle + \tfrac{1}{2}\langle x_2 x_1|m_1 m_2\rangle = \tfrac{1}{2}\left(\langle x_1 x_2| + \langle x_2 x_1|\right)|m_1 m_2\rangle,
\end{aligned}$$

which can be written as

$$\langle x_1 x_2|\widehat{\Omega}|\psi\rangle = \langle x\ |\psi^S\rangle, \tag{15.29}$$

where $|\psi^S\rangle$ is the symmetric component of $|\psi\rangle$.

https://doi.org/10.1017/9781009679633.016 Published online by Cambridge University Press

15.2 Bosons and Fermions. The Pauli Principle

We have just seen that physical stationary states of identical particles are those in which all particles play an equal role. This observation can be elevated to the category of principle by demanding that identical particles receive identical treatment in the theory. The only situation consistent with the principle of equal treatment is that the eigenvalues of all exchange operators $\hat{P}_{ij}$ are well defined, that is, that the wave functions are S or A.[5] It is important to note that these N-particle wave functions are nonfactorizable, therefore they describe entangled states, as discussed in Section 14.5 in connection with the two-particle states.

Which of the two possibilities is realized depends on the type of particles in question. It turns out that:

(i) Particles with integer spin are described by S wave functions in the stationary state ($\lambda_{ij} = 1$). These particles are generically known as bosons, and include pions (spin 0), photons (spin 1), etc.

(ii) Particles with half-integer spin are described by A wave functions in the stationary state ($\lambda_{ij} = -1$). These particles are called fermions. They include electrons, protons, and neutrons (all of which have spin 1/2), etc.[6]

This fundamental and very important correspondence, known as the *spin-statistics theorem*, was postulated by Pauli in 1939 and there are no known exceptions to it. For example, atoms and molecules will have S or A wave functions depending on whether their total electronic spin is integer or half-integer, respectively. This property is considered a postulate because it cannot be derived from other principles within the framework of nonrelativistic QM.[7]

According to the spin-statistics theorem, the stationary states of a system of identical particles of spin s are those whose wave function Ψ is transformed according to

$$\widehat{P}_{ij}\Psi(1,2,\ldots,i,\ldots,j\ldots,N) = (-1)^{2s}\Psi(1,2,\ldots,j,\ldots,i\ldots,N). \tag{15.30}$$

Let us take first the case of bosons. The wave function must be of the form of (15.21), with the normalization coefficient determined as follows: The states n_j can be the same or different. Assume there are N_1 particles in state n_1, N_2 in state n_2 and so on, with $N_1 + N_2 + \cdots = N$; the occupancy numbers N_i fully specify the state. The S wave function contains $N!/N_1!N_2!\ldots$ terms, so, considering the orthonormality of the functions $\psi_{n_j}(\xi_i)$, we write

[5] Heisenberg and Dirac independently considered the possibility of using fully symmetrized or antisymmetrized wave functions as early as 1926.

[6] The neutron has a magnetic moment, despite being electrically neutral. This was an early indication that the neutron is not an elementary particle; in fact, it is made up of quarks, which are charged particles. Neutral elementary particles with spin, such as neutrinos, photons and the Z boson, do not have a magnetic moment; we know they have a spin because of the conservation of angular momentum.

[7] We recall from Section 14.5.1 that SED provides an explanation for the antisymmetry of the electron wave function as a consequence of the antiphase pair coupling to the background field modes. More generally, the spin-statistics theorem can be derived within relativistic quantum field theory, which involves postulates that are not part of QM; see Tomonaga (1997).

https://doi.org/10.1017/9781009679633.016 Published online by Cambridge University Press

$$\Psi^S = \sqrt{\frac{N_1!N_2!\dots}{N!}} \sum_{P_k} \widehat{P}_k \psi_{n_1}(\xi_1)\psi_{n_2}(\xi_2)\dots, \tag{15.31}$$

where the N_i can take any value between 0 and N, with the restriction that the total sum be N.

In the case of fermions, the wave function must be of the form of (15.22), which can be written as

$$\Psi^A = C \begin{vmatrix} \psi_{n_1}(\xi_1) & \psi_{n_1}(\xi_2) & \cdots & \psi_{n_1}(\xi_N) \\ \psi_{n_2}(\xi_1) & \psi_{n_2}(\xi_2) & \cdots & \psi_{n_2}(\xi_N) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_{n_N}(\xi_1) & \psi_{n_N}(\xi_2) & \cdots & \psi_{n_N}(\xi_N) \end{vmatrix}. \tag{15.32}$$

This is the *Slater determinant*.[8] Permuting any two particles here means swapping the corresponding columns, which changes the sign of the determinant, that is, of Ψ, as it should. However, we observe an extremely important new phenomenon: if in (15.32) we make any two states equal (e.g., if we put $n_1 = n_2$), the determinant will have two equal rows and give zero. This is always the case: there is no (stationary) state of a physical system in which two (or more) equal fermions share the *same* quantum numbers. This most important result is the Pauli principle or exclusion principle referred to in Section 14.1, which allows us to understand the periodic table of the elements as a consequence of the atomic structure of matter, as will be shown in Section 16.2. Pauli was able to infer this principle from an analysis of atomic structure and other data; we will apply it later to precisely solve this problem, although we have used it previously to explain metallic conductivity (Section 6.3). Although the Pauli principle is a particular result of a more general principle, the one that assigns fully antisymmetrized functions to any stationary system of identical fermions, it continues to be called a principle.

From the total antisymmetry of the wave function it follows that the fermion occupancy number of each state can only be 0 (empty state) or 1 (fully occupied state). Therefore $N_i! = 1$ (since $0! = 1! = 1$) and the normalized wave function is obtained by putting into (15.32)

$$C = \frac{1}{\sqrt{N!}}, \tag{15.33}$$

as corresponds to the case in which we have $N!$ different states of equal weights. Note that the S wave function can be obtained with the help of the Slater determinant for the A wave function, but using only + signs in the expansion.

15.2.1 The two quantum statistics

Bose–Einstein (B–E) statistics. Consider an ideal gas of bosons in equilibrium; this could even be a "gas" of photons. In a given volume at arbitrary temperature $T > 0$, one finds a certain number of particles in each of the possible quantum states, which

[8] John Clarke Slater (1900–1976) was a US-born theoretical physicist and chemist, who introduced the important Slater determinants for physical and chemical calculations. He wrote several highly recommended textbooks (if you are lucky enough to find them).

https://doi.org/10.1017/9781009679633.016 Published online by Cambridge University Press

gives the quantum gas a characteristic statistical behavior. In the electromagnetic case, for example, the equilibrium distribution is given by Planck's law. An extreme case of this distribution occurs at very low temperatures (of the order of millionths of a Kelvin), at which the particles tend to fill the ground state and move very slowly, so that the de Broglie wavelength associated with their motion, $\lambda = h/p$, can reach dimensions similar to those of the container. As a result, virtually all the waves associated with the particles in the volume superimpose, leading to a high correlation between most of the bosons in the system. When this almost complete correlation is reached, a *Bose–Einstein condensate* spontaneously emerges and the system can exhibit vivid macroscopic effects of coherence and interference, as discussed in Section 15.4.

Fermi–Dirac (F–D) statistics. An ideal gas of fermions in equilibrium (e.g., electrons) has a very different statistical behavior. Each fermion in the system occupies a different quantum state, as can be seen from the vanishing of the wave function (15.32) when $\xi_1 = \xi_2$. In other words, the probability of two identical fermions being in the same state in the same system is zero. This makes the statistical behavior of the fermion gas sui generis. We had the opportunity to study some basic features of the F–D statistics in Section 6.3 (see Figure 6.6) and to apply it to the calculation of the specific heat of metals.

A simple example will allow us to clarify the idea behind the different statistics. Consider a system of two identical particles that can be in either state n or m, and assume that their interaction can be neglected. There are thus four possible product states of independent particles: $|nn\rangle = |n\rangle\,|n\rangle$, $|nm\rangle = |n\rangle\,|m\rangle$, $|mn\rangle = |m\rangle\,|n\rangle$, and $|mm\rangle = |m\rangle\,|m\rangle$. In a macroscopic problem, these four states would be counted as four possible and different cases, according to the methods of classical statistical mechanics, which lead to the Maxwell–Boltzmann distribution (M–B statistics). Suppose instead that the particles are bosons; we must construct fully symmetric, mutually orthogonal state functions, which gives three different possible states,

$$|nn\rangle^S = |n\rangle\,|n\rangle\,, \tag{15.34}$$

$$|nm\rangle^S = \tfrac{1}{\sqrt{2}}(\,|n\rangle\,|m\rangle + |m\rangle\,|n\rangle\,), \tag{15.35}$$

$$|mm\rangle^S = |m\rangle\,|m\rangle\,. \tag{15.36}$$

For fermions, instead, we must construct totally antisymmetric combinations, of which there is only one,

$$|nm\rangle^A = \tfrac{1}{\sqrt{2}}(\,|n\rangle\,|m\rangle - |m\rangle\,|n\rangle\,). \tag{15.37}$$

Thus for the given conditions of the problem, the M–B (classical) statistics admits four different states, the B–E statistics admits three, and the F–D statistics admits only one.

If the two particles are in the same (m or n) state, this happens in 2 out of 4 cases (relative probability 1/2 for each one) in the M–B case; in 2 out of 3 cases (relative probability 2/3) in the B–E case, and not at all in the F–D case (probability 0). It is therefore more likely to find clusters of bosons in the same quantum state than classical statistics predicts, and impossible to find in the F–D case. This is expressed by saying that bosons in the same state "attract each other," while fermions "repel each other." These effective interactions are not explicitly contained in the Hamiltonian,

https://doi.org/10.1017/9781009679633.016 Published online by Cambridge University Press

but have been introduced into the preceding description in a phenomenological way, by symmetrization or antisymmetrization of the state functions.

The fact that the state function is a symmetric or antisymmetric sum of terms introduces quantum correlations between the particles, as we learned from the discussion in Section 14.5 and in particular from the case of bipartite singlet spin studied in 14.5.5. The following examples illustrate the influence of the different statistics on the correlation between the position variables of the particles.

Exercise E15.2. Correlation between position variables

Analyze the correlation between the position variables of two equal particles whose entangled wave function is given by

$$|nm\rangle = \frac{1}{\sqrt{2}}(\,|n\rangle\,|m\rangle + \lambda\,|m\rangle\,|n\rangle\,), \quad \lambda = \pm 1, \tag{15.38}$$

and discuss the effect of the parameter λ.

Solution. As a measure of the correlation between $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$, we shall calculate the covariance, defined as

$$\mathrm{cov}(\boldsymbol{r}_1, \boldsymbol{r}_2) = \langle(\boldsymbol{r}_1 - \langle\boldsymbol{r}_1\rangle)\cdot(\boldsymbol{r}_2 - \langle\boldsymbol{r}_2\rangle)\rangle. \tag{15.39}$$

After some simple algebra, we obtain

$$\mathrm{cov}(\boldsymbol{r}_1, \boldsymbol{r}_2) = \lambda|\langle n|\boldsymbol{r}|m\rangle|^2, \tag{15.40}$$

where we took into account that the expectation values are the same for particle 1 and particle 2.

Note that the particle positions are only correlated if the state functions $|n\rangle$ and $|m\rangle$ overlap, i.e., if there are values of $\boldsymbol{r}$ for which both functions are simultaneously different from zero. In other words, the symmetrization effects are only important when the overlap of the wave functions is important. The covariance is negative for the antisymmetric wave function and positive for the symmetric one, as would be expected: in the first case the particles tend to avoid each other, while in the second case they tend to approach each other.

Exercise E15.3. Mean square distance between particles

Consider two entangled particles whose orbital wave function is given by (15.38). Calculate the expectation value of the square of the distance between the particles and discuss the result as a function of λ.

Solution. Using Eq. (15.38), the calculation is straightforward and gives

$$\begin{aligned} &\langle nm|(\boldsymbol{r}_1 - \boldsymbol{r}_2)^2|nm\rangle_\lambda \\ &= \langle n|\boldsymbol{r}^2|n\rangle + \langle m|\boldsymbol{r}^2|m\rangle - 2\langle n|\boldsymbol{r}|n\rangle\cdot\langle m|\boldsymbol{r}|m\rangle - 2\lambda|\langle n|\boldsymbol{r}|m\rangle|^2. \end{aligned} \tag{15.41}$$

The first three terms on the right do not depend on the symmetrization and represent the expectation value of $\boldsymbol{r}_1^2 + \boldsymbol{r}_2^2 - 2\boldsymbol{r}_1\cdot\boldsymbol{r}_2$ as if the particles were independent; therefore we can write

https://doi.org/10.1017/9781009679633.016 Published online by Cambridge University Press

$$\langle nm|(\boldsymbol{r}_1 - \boldsymbol{r}_2)^2|nm\rangle_\lambda = \langle nm|(\boldsymbol{r}_1 - \boldsymbol{r}_2)^2|nm\rangle_{\lambda=0} - 2\lambda|\langle n|\boldsymbol{r}|m\rangle|^2. \tag{15.42}$$

The term with $\lambda = 0$ would be the mean square distance if the system followed the M–B statistics. To this classical result we must add the positive quantity $2|\langle n|\boldsymbol{r}|m\rangle|^2$ if $\lambda = -1$ (antisymmetric wave function) or subtract the same quantity if $\lambda = +1$ (symmetric wave function). This shows that bosons tend to stay closer together than classical particles, while equal-spin fermions tend to stay further apart.

15.3 Effects of Statistics on the Energy Spectrum

The correlations discussed in the previous section have important consequences on the atomic energy levels. Let us consider the case of two electrons whose spin-independent interaction can be taken as a perturbation. (This calculation is used in Section 16.3 in connection with the study of the He atom.) The individual orbital wave functions in the absence of interaction are $\psi_n(\boldsymbol{r})$ and $\psi_m(\boldsymbol{r})$; the fully symmetrized orbital wave functions are given by Eq. (15.38), which we rewrite as

$$\psi_\pm = \frac{1}{\sqrt{2}}\left[\psi_n(\boldsymbol{r}_1)\psi_m(\boldsymbol{r}_2) \pm \psi_m(\boldsymbol{r}_1)\psi_n(\boldsymbol{r}_2)\right]. \tag{15.43}$$

The complete wave function is the product of an orbital function and a spin factor, since by hypothesis, the Hamiltonian is spin-independent. This wave function must be antisymmetric because we are dealing with electrons, so one of its factors must be symmetric and the other antisymmetric. From Eqs. (11.175), (11.177), (11.181), and (11.182) we see that the spin factors are symmetric for $S = 1$ and antisymmetric for $S = 0$. Therefore, the sign $\pm$ in Eq. (15.43) can be written as $(-1)^S$, and $\psi_\pm = \psi_S$.

Now let $V(|\boldsymbol{r}_1 - \boldsymbol{r}_2|)$ be the perturbation potential. Since this potential commutes with $\widehat{P}_{12}$, it does not mix states of different permutational symmetry and we can treat the problem as if it were nondegenerate. Therefore, to first order in perturbation theory the energy of the symmetric orbital state ($S = 0$) is

$$E_+ = E^{(0)} + \langle\psi_+|V|\psi_+\rangle, \tag{15.44}$$

and for the state with $S = 1$ it is

$$E_- = E^{(0)} + \langle\psi_-|V|\psi_-\rangle. \tag{15.45}$$

As we learned from the previous exercises, in the region $|\boldsymbol{r}_1 - \boldsymbol{r}_2| \approx 0$, where $V(|\boldsymbol{r}_1 - \boldsymbol{r}_2|)$ is normally important, these matrix elements have very different values. It is therefore convenient to write the correction as the sum of two terms,

$$\langle\psi_\pm|V|\psi_\pm\rangle = J \pm K. \tag{15.46}$$

The J term, given by

$$J = \int |\psi_n(\boldsymbol{r}_1)|^2|\psi_m(\boldsymbol{r}_2)|^2 V d^3\boldsymbol{r}_1 d^3\boldsymbol{r}_2, \tag{15.47}$$

https://doi.org/10.1017/9781009679633.016 Published online by Cambridge University Press

is usually interpreted as the expectation value of the interaction potential when particle 1 (not one of the particles) is in state n and particle 2 in state m. Since particle 1 is not in state n all the time, etc., this interaction energy has to be corrected. The correction term, the exchange energy, is given by

$$K = \int \psi_n^*(\boldsymbol{r}_1)\psi_m^*(\boldsymbol{r}_2)V\psi_n(\boldsymbol{r}_2)\psi_m(\boldsymbol{r}_1)d^3\boldsymbol{r}_1 d^3\boldsymbol{r}_2. \tag{15.48}$$

The total energy,[9]

$$E_\pm = E^{(0)} + J \pm K, \tag{15.49}$$

shows that there is an energy difference of $2K$ between the states with $S = 0$ and $S = 1$, an effect known as exchange splitting (see Problem 15.5.2).

The previous result can be rewritten as an effective-coupling effect between the spins of the electrons, introducing Dirac's exchange operator

$$\widehat{V}_{\text{exchange}} = -\tfrac{1}{2}K\left(1 + \frac{\widehat{\boldsymbol{s}}_1 \cdot \widehat{\boldsymbol{s}}_2}{\hbar^2}\right). \tag{15.50}$$

Indeed, from $\hat{S}^2 = \hat{s}_1^2 + \hat{s}_2^2 + 2\widehat{\boldsymbol{s}}_1 \cdot \widehat{\boldsymbol{s}}_2$ we get

$$\frac{\widehat{\boldsymbol{s}}_1 \cdot \widehat{\boldsymbol{s}}_2}{\hbar^2} = 2S(S+1) - 3 = \begin{cases} -3 & \text{for } S = 0 \\ 1 & \text{for } S = 1 \end{cases} \tag{15.51}$$

and therefore the eigenvalues of the exchange operator are K for $S = 0$ and $-K$ for $S = 1$, in accordance with Eq. (15.49).

The following are three additional examples of gross effects on the energy spectrum of a composite system, formally produced by the exchange interaction.

1. Two-particle systems. Take first a two-boson system; its wave function is totally symmetric. On the other hand, we know that the wave function is symmetric only for even orbital momentum. Therefore, the solutions for odd l, although mathematically legitimate, are not realized.

Now take a two-spin 1/2 system; the total spin can be 0 or 1. In the first case (antisymmetric spinor wave function) the orbital wave function is symmetric and corresponds to even orbital momenta. In the spin 1 case, in turn, only wave functions corresponding to odd orbital momentum can be accepted. In both cases only some of the energy levels predicted by the Schrödinger equation are realized, depending on the total spin, even when the interaction does not depend on spin. This example shows that, in general, the spectrum of an atomic system depends on its total spin.

2. Three-particle systems. Although still simple to analyze, the tripartite case already shows the complications to be expected when dealing with a larger number of particles. We assume that the interaction between the particles is negligible and consider the 1D case for simplicity. The unsymmetrized orbital wave functions can then be expressed as the product of three single-particle wave functions

$$\psi_{lmn}(x_1, x_2, x_3) = \psi_l(x_1)\psi_m(x_2)\psi_n(x_3), \tag{15.52}$$

[9] Note that this division of the interaction energy into two parts J and K, although mathematically convenient, is physically meaningless. In fact, the calculation must be done with the joint probability that either particle is in state n and the other in state m. Therefore, it makes no sense to physically interpret any of the terms of the expression (15.46) separately, although this is very often done.

https://doi.org/10.1017/9781009679633.016 Published online by Cambridge University Press

with total energy given by

$$E \equiv E_{lmn} = E_l + E_m + E_n. \tag{15.53}$$

If the particles are spin-zero bosons, the following cases may occur:
(i) $l = n = m$, $E = 3E_n$. The wave function (15.52) is already symmetric.This is the case in particular for the ground state of the system.
(ii) $l = n \neq m$. The only physically realizable state is

$$\psi^S_{nnm} = \frac{1}{\sqrt{3}}(\psi_{nnm} + \psi_{nmn} + \psi_{mnn}), \quad E = 2E_n + E_m. \tag{15.54}$$

This is the case for the system's first excited state.
(iii) l, n, m different. The only fully symmetric state that can be constructed from the six states with exchange degeneracy is

$$\psi^S_{nlm} = \frac{1}{\sqrt{6}}(\psi_{lnm} + \psi_{nml} + \psi_{mln} + \psi_{nlm} + \psi_{lmn} + \psi_{mnl}), \quad E = E_l + E_n + E_m. \tag{15.55}$$

Now consider three fermions with spin 1/2. In addition to the space coordinates $x_i, i = 1, 2, 3$, we use the spin coordinates S_{3i} ; we denote the set of variables $(x_i, S_{3i}) = (\xi_i)$. With the individual wave function (for spin-independent Hamiltonians) written as the product of the orbital function $\psi_n(x_i)$ times the spin wave function $\chi_\pm(S_{3i})$,

$$\psi^\pm_n(\xi_i) = \psi_n(x_i)\chi_\pm(S_{3i}), \tag{15.56}$$

and the spin orientation denoted by a, b, c, the unsymmetrized wave function is

$$\psi^{abc}_{lnm} = \psi^a_l(\xi_1)\psi^b_n(\xi_2)\psi^c_m(\xi_3), \tag{15.57}$$

and the fully antisymmetric wave function is given by the Slater determinant,

$$(\psi^{abc}_{nlm})^A = \frac{1}{\sqrt{6}}\begin{vmatrix} \psi^a_l(\xi_1) & \psi^a_l(\xi_2) & \psi^a_l(\xi_3) \\ \psi^b_n(\xi_1) & \psi^b_n(\xi_2) & \psi^b_n(\xi_3) \\ \psi^c_m(\xi_1) & \psi^c_m(\xi_2) & \psi^c_m(\xi_3) \end{vmatrix}. \tag{15.58}$$

While the l, n, m orbital quantum numbers can be the same or different, the spin indices a, b, c cannot all be different (since they can only take the values ± 1); this fact has a profound effect on the spectrum, as shown in the following.
(i) $l = n = m$. Since necessarily at least two spin indices are equal, the determinant has at least two equal rows, so this state is not realized.
(ii) $l = n \neq m$. The equality of l and n requires that a and b are different. There are four states that satisfy this condition: $(+-+, +--, -++, -+-)$, corresponding to the energy $E = 2E_n + E_m$; one of these states would be, for example

$$(\psi^{-++}_{nnm})^A = \frac{1}{\sqrt{6}}(\psi^{-++}_{nnm} + \psi^{+-+}_{mnn} + \psi^{++-}_{nmn} - \psi^{++-}_{mnn} - \psi^{+-+}_{nnm} - \psi^{-++}_{nmn}). \tag{15.59}$$

All these states have spin projection 1/2 and are doubly degenerate. In particular, the ground state is one of them. We conclude that if the spin is nonzero, the orbital wave function of the ground state can have nodes.
(iii) l, n, m different. All possible combinations of spin indices are acceptable, so there are 2^3 degenerate states, corresponding to one spin-3/2 quadruplet and two spin-1/2 doublets. Among these states there are those with well-defined symmetry with respect

https://doi.org/10.1017/9781009679633.016 Published online by Cambridge University Press

to the exchange of space or spin coordinates separately, such as the following state, which is symmetric in the spin indices and antisymmetric in the orbital indices,

$$\left(\psi_{lnm}^{+++}\right)^A = \tfrac{1}{\sqrt{6}}\left(\psi_{lnm}^{+++} + \psi_{nml}^{+++} + \psi_{mln}^{+++} - \psi_{nlm}^{+++} - \psi_{lmn}^{+++} - \psi_{mnl}^{+++}\right). \quad (15.60)$$

There are states that are antisymmetric only for the simultaneous exchange of both sets of indices, as in case (ii) above. The energy of these states is $E = E_l + E_n + E_m$.

3. The ground state of N bosons or fermions in a 1D infinite well. The individual wave functions and energies are given by

$$\psi_n(x) = \sqrt{\frac{2}{a}} \sin \frac{\pi n}{a} x, \quad E_n = \frac{\pi^2 \hbar^2}{2ma^2} n^2 \equiv E_1 n^2. \quad (15.61)$$

For N bosons that do not interact with each other, the ground state corresponds to all particles with $n = 1$ and total energy $E = NE_1$, so that the energy per particle is

$$\frac{E}{N} = E_1. \quad (15.62)$$

Now suppose we have N fermions with spin 1/2. In this case there are only two electrons for each value of n (with two possible spin-1/2 states), so that in the ground state of the system all energy levels with n from 1 to $N/2$ are filled. The ground-state energy is therefore given by

$$E = 2E_1 \sum_{n=1}^{N/2} n^2 = \tfrac{1}{12} E_1 N(N+1)(N+2), \quad (15.63)$$

as follows from the formula $\sum_{k=1}^{n} k^2 = \frac{1}{6} n(n+1)(2n+1)$. If N is very large, we can write with good approximation

$$E = \tfrac{1}{12} E_1 N^3. \quad (15.64)$$

The average energy per particle can now be considerably higher than E_1, since it is proportional to the square of the total number of particles,

$$\frac{E}{N} = \tfrac{1}{12} E_1 N^2. \quad (15.65)$$

The energy of the top occupied level is the Fermi energy, as explained in Section 6.3,

$$E_F \equiv E_{\frac{N}{2}} = \tfrac{1}{4} E_1 N^2 = \frac{\pi^2 \hbar^2}{8m} \rho^2. \quad (15.66)$$

The last equality was written for the 1D case, with the fermion density ρ defined as $\rho = N/a$ (see Eq. (6.91), which applies to the 3D case). As a physical example of this situation, we recall the problem discussed in Section 6.3, where we saw that the conduction electrons of a metal can be considered to a first approximation as forming a gas of particles moving in a common potential produced by the entire structure. An energy diagram of this metal at low temperatures is essentially that shown in Figure 6.6 for the 1D metal model at $T = 0$. More generally, the particle distribution as a function of temperature is discussed in the following section.

https://doi.org/10.1017/9781009679633.016 Published online by Cambridge University Press

15.3.1 The boson and fermion distributions as a function of temperature

The formulas for both the fermion and the boson distributions are derived using the techniques of statistical mechanics; see Section 15.5.6 for an introduction. The distribution of fermions as a function of energy and temperature, called the Fermi–Dirac distribution,

$$\bar{n}_{FD}(E) = \frac{2}{e^{(E-E_F)/kT} + 1}, \tag{15.67}$$

is obtained taking into account the exclusion principle. At $T = 0$, all energy levels up to E_F are occupied and those above E_F are empty. With increasing temperature the curve rounds off as the excited electrons near E_F occupy higher energy states and leave lower energy states. This shows that at very low temperatures (close to absolute zero) the energy of a metal is essentially independent of its temperature and is close to the minimum energy.

The Bose–Einstein distribution, in turn, is of the general form

$$\bar{n}_{BE}(E) = \frac{1}{e^{(E-\mu)/kT} - 1} \tag{15.68}$$

with $E > \mu$, where μ is the chemical potential.[10] In the case of photons, $\mu = 0$ and Eq. (15.68) coincides with Planck's formula. A comparison with the F–D formula shows that the key difference is the sign of the 1 in the denominator.

As can be seen from Fig. 15.1, at very low energies the contrast between the distributions is high: while the bosons tend to concentrate, fermions tend to stay apart, as discussed earlier. For very large energies compared to kT, both distributions approach the classical M–B distribution, $\bar{n}(E) \sim e^{-E/kT}$.

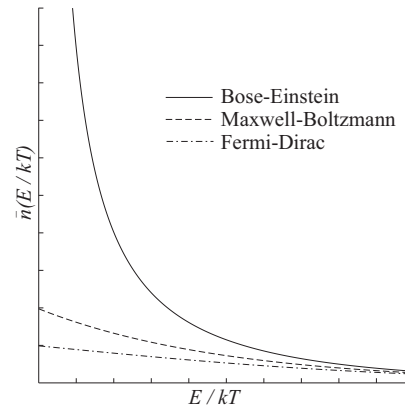


Figure 15.1 A comparison of the B–E, M–B, and F–D distribution functions at different temperatures, for $\mu/kT = -0.1$.

[10] The chemical potential of a given substance is the chemical energy per mole of the substance. The chemical potential marks the energy at which Bose–Einstein condensation occurs (see Section 15.4).

https://doi.org/10.1017/9781009679633.016 Published online by Cambridge University Press

Exercise E15.4. The Bose distribution

Consider a mode of frequency ω of the radiation field in equilibrium with matter at temperature T. Using the quantum description in terms of photons, determine: (a) the probability $P_n(T)$ with which the (field) eigenstate n contributes to the state of equilibrium; (b) the dispersion σ_n^2 of the number of photons of frequency ω.

Solution. The probability of state n is given by

$$P_n = \frac{1}{Z} e^{-\beta E_n} = e^{-\beta \hbar \omega n}(1 - e^{-\beta \hbar \omega}),$$

where $Z = e^{-\beta E_0}/(1 - e^{-\beta \hbar \omega})$. Since the average number of photons is

$$\bar{n} = 1/\left(e^{\beta \hbar \omega} - 1\right),$$

it follows that $e^{\beta \hbar \omega} = 1 + (1/\bar{n})$, so $1 - e^{-\beta \hbar \omega} = 1/(\bar{n} + 1)$, which gives

$$P_n = \frac{1}{\bar{n} + 1} \left(\frac{\bar{n}}{\bar{n} + 1}\right)^n. \tag{15.69}$$

This result is known as the Bose distribution and applies to all systems consisting of bosons. A distribution of the form $P_n = Ng^n$ (here $g = \bar{n}/(\bar{n} + 1) = e^{-\beta \hbar \omega}$) is a geometric distribution.[11]

The mean square of n is obtained as follows.

$$\begin{aligned} \langle n^2 \rangle &= \sum n^2 P_n = P_0 \sum [n\,(n-1) + n]\, g^n \\ &= P_0 g^2 \sum n\,(n-1)\, g^{n-2} + \bar{n} \\ &= P_0 g^2 \frac{d^2}{dg^2} \frac{1}{1-g} + \bar{n} = \frac{2g^2}{(1-g)^2} + \bar{n}, \end{aligned}$$

i.e.,

$$\langle n^2 \rangle = \sum n^2 P_n = 2\bar{n}^2 + \bar{n} = \bar{n}\,(2\bar{n} + 1)\,. \tag{15.70}$$

The dispersion of the number n of photons of frequency ω is thus

$$\sigma_n^2 = \langle n^2 \rangle - \bar{n}^2 = \bar{n}^2 + \bar{n}. \tag{15.71}$$

Note that with $\overline{E} = \hbar \omega \bar{n}$, the result just obtained gives for the fluctuations of the energy

$$\sigma_E^2 = \hbar^2 \omega^2 \sigma_n^2 = \overline{E}^2 + \hbar \omega \overline{E}, \tag{15.72}$$

[11] Since $P_1 = P_0 g$ we can also write $P_n = P_0 (P_1/P_0)^n$, $P_{n+s} = P_n (P_1/P_0)^s$. This property can be taken as the definition of a geometric distribution. In particular, the probability that the ω modes contain n or more photons is

$$f(n) = \sum_{r=n}^{\infty} P_r = \sum_{s=0}^{\infty} P_{n+s} = P_n \sum_{s=0}^{\infty} (1/\xi)^s = P_n \frac{\xi}{\xi - 1},$$

where $\xi = P_0/P_1 > 1$.

https://doi.org/10.1017/9781009679633.016 Published online by Cambridge University Press

therefore a linear and a quadratic term contribute to the fluctuations. At very high temperatures the quadratic term dominates over the linear one, so that in the classical limit we get the Rayleigh–Jeans distribution

$$\sigma_E^2 = \pi^4 c^6 \frac{\rho^2(\omega)}{\omega^4} = \overline{E}^2 = \hbar^2\omega^2\bar{n}^2 = k_B^2 T^2, \quad kT \gg \hbar\omega. \tag{15.73}$$

This allows us to identify the quadratic term as the classical prediction for the energy fluctuations of the Maxwell field. In the low-temperature limit the linear term is dominant,

$$\sigma_E^2 = \hbar\omega\overline{E} = \hbar^2\omega^2\bar{n}, \quad kT \ll \hbar\omega, \tag{15.74}$$

or

$$\sigma_n^2 = \bar{n}. \tag{15.75}$$

Einstein interpreted this result in 1905 as follows. Consider a gas composed of molecules moving independently with fixed energy $\mathcal{E} = \hbar\omega$; suppose that the number of these molecules in a unit volume follows a Poisson distribution (which is consistent with the independence hypothesis). The fluctuations of this number are then given by $\sigma_n^2 = \bar{n}$, and their contribution to the energy fluctuations is $\sigma_E^2 = \mathcal{E}\bar{n} = \hbar\omega\bar{n}$.[12] This led Einstein to propose that the linear term reflects an independent "granular" structure of the radiation field, that is, that at very low intensities the field must be considered as consisting of independent radiation quanta – the current photons.[13]

15.4 Collective Phenomena. Phonons and Bose-Einstein Condensates

When large numbers of atoms or electrons in a solid interact, their local quantum mechanics can be very complex. But sometimes strikingly simple collective properties appear, which cannot exist for single atoms, or even small collections of them: they emerge at a higher level. *Quantum collective phenomena* include the quantum Hall effect, superconductivity and magnetism in strongly correlated metals, B–E condensation of dilute gases, phonons and excitons, topologically ordered quantum states and quantum critical phenomena.

Here we briefly present some of the most important examples of quantum collective phenomena; for more in-depth discussions, the student is referred to modern textbooks on condensed matter.

15.4.1 Lattice dynamics: Phonons

Atoms vibrate in every material: in crystals, atomic clusters, or more complicated structures. In crystals, the vibrational modes are related to a number of properties, such as

[12] A Poisson distribution has the form $P(n) = e^{-a}a^n/n!$, and $\bar{n} = a, \sigma_n^2 = a = \bar{n}$. The relative fluctuations of the number n are given by $(\sigma_n/\bar{n})^2 = 1/\bar{n}$ and quickly become negligible as $\bar{n}$ grows.

[13] As noted in Section 2.7, these "simple" results provided the young Einstein with the basis for proposing the corpuscular (photonic) structure of a low-density radiation field.

https://doi.org/10.1017/9781009679633.016 Published online by Cambridge University Press

the temperature dependence of the lattice parameters and elastic constants; they determine heat and electric conductivity, critical temperature, and other thermodynamic functions. Vibrational modes enhance the diffusion processes and play a crucial role in superconductivity, ferroelectricity, shape memory alloys, and so on.

The basic theory of the lattice dynamics was given by M. Born and K. Huang (1919–2005), who developed the theory for crystal lattices with translational symmetry. The interaction of electrons with vibrating atoms in the lattice can be conveniently treated as an interaction with *phonons*,[14] which are collective excitations, that is, excited states of the correlated vibrational modes of the structure.

For simplicity, consider a 1D crystalline lattice or linear chain consisting of a large number of particles N, which can be atoms or molecules. If the lattice is rigid enough, the atoms must exert forces on each other to maintain their equilibrium position. (These forces can be van der Waals forces, covalent bonds, electrostatic attractions, and others, all of which are ultimately due to the electric force.) The potential energy of the whole lattice is the sum of all pairwise potential energies $V\left(x_i - x_j\right)$.

In order to solve this many-body problem, two important approximations are usually made: the sum is performed only over neighboring atoms (which is legitimate because the fields produced by distant atoms are effectively screened out), and the potentials V are considered as harmonic-oscillator potentials (which is legitimate as long as the atoms remain close to their equilibrium positions). Assuming that all atoms have the same mass m and the same natural frequency of oscillation ω, the Hamiltonian for this system is

$$H = \frac{1}{2m}\sum_{1=1}^{N} p_i^2 + \frac{1}{2}m\omega^2 \sum_{i=j+1} (x_i - x_j)^2. \tag{15.76}$$

The normal coordinates and momenta in Fourier space, defined as

$$Q_k = \frac{1}{\sqrt{N}}\sum_l e^{ikal} x_l, \;\; P_k = \frac{1}{\sqrt{N}}\sum_l e^{ikal} p_l, \tag{15.77}$$

where a is the lattice spacing, satisfy the commutation relations

$$[Q_k, P_{k'}] = \frac{1}{N}\sum_{l,m} e^{ikal} e^{-ik'am} \left[x_l, p_m\right], \tag{15.78}$$

$$[Q_k, Q_{k'}] = 0, \qquad [P_k, P_{k'}] = 0. \tag{15.79}$$

The Hamiltonian can thus be written in wave vector space as (see Problem P15.10)

$$H = \frac{1}{2m}\sum_k P_k P_{-k} + \frac{m}{2}\sum_k \omega_k^2 Q_k Q_{-k}, \tag{15.80}$$

where

$$\omega_k = 2\omega \left|\sin\frac{ka}{2}\right|. \tag{15.81}$$

[14] The concept of the phonon was introduced in 1932 by I. Tamm. The name phonon is derived from the Greek word φωνή (sound), because long-wavelength phonons produce sound.

https://doi.org/10.1017/9781009679633.016 Published online by Cambridge University Press

Equation (15.80) represents a set of uncoupled harmonic oscillators in the reciprocal space; these oscillators are the phonons, whose oscillation frequency depends on k. According to Eq. (15.81), the allowed values for k depend on the boundary conditions. If, for simplicity, periodic boundary conditions are imposed, the resulting quantization is

$$k_s = \frac{2\pi s}{Na}, \quad s = 0, \pm 1, \pm 2, \ldots \pm \frac{N}{2},$$

with the upper limit coming from the minimum wavelength, which is twice the lattice spacing a. The corresponding energy levels are given by

$$E_n = \left(\frac{1}{2} + n\right)\hbar\omega_k, \quad n = 0, 1, 2, \ldots.$$

The relation between frequency and wave vector, $\omega = \omega(k)$, is known from classical optics as a *dispersion relation*.

15.4.1.1 Dispersion relations

In general, dispersion can be caused either by geometric boundary conditions or by the interaction of the waves with the medium. Given the dispersion relation, the phase velocity and group velocity of each wave component can be calculated as a function of frequency.

The speed of propagation (group velocity) of an acoustic phonon, which is also the speed of sound in the lattice, is given by the slope of the dispersion relation, $d\omega_k/dk$. At low values of k (i.e., long wavelengths), the dispersion relation is nearly linear according to Eq. (15.81), and the speed of sound is approximately ωa. As a result, packets of phonons with different (but long) wavelengths can travel long distances through the lattice without breaking apart. This is the reason why sound propagates through solids without significant distortion. However, for large values of k, i.e., short wavelengths, this behavior fails due to the microscopic details of the lattice.

For a one-dimensional alternating array of two types of ions or atoms of mass m_1, m_2 connected by springs of spring constant K, two modes of vibration (or branches) result, namely

$$\omega_{\pm}^2 = K\left(\frac{1}{m_1} + \frac{1}{m_2}\right) \pm K\sqrt{\left(\frac{1}{m_1} + \frac{1}{m_2}\right)^2 - \frac{4\sin^2\frac{ka}{2}}{m_1 m_2}}, \tag{15.82}$$

where k is the wave vector of the vibration related to its wavelength by $k = 2\pi/\lambda$. In this case the dispersion relations exhibit two types of phonons, namely optical and acoustic, corresponding to the upper and lower curves in Fig. 15.2. More optical and acoustic dispersion branches appear in a unit cell with a larger number of atoms. Typical maximal phonon frequencies range from 10 to 30 THz (1 THz $= 10^{12}\,\mathrm{s}^{-1}$).[15]

Since the phonons are described as harmonic oscillators, we can use the second quantization technique that is introduced in Section 17.6 and write the Hamiltonian of the system as

[15] Efficient heat transport by phonons in this frequency range is important for temperature control in nanoscale devices, where it replaces classical heat diffusion.

https://doi.org/10.1017/9781009679633.016 Published online by Cambridge University Press

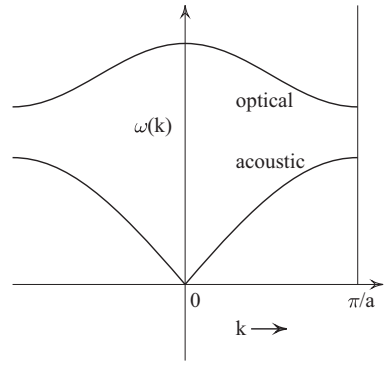


Figure 15.2 Phonon dispersion curves of a linear diatomic chain with two different types of atoms, showing an optical branch (upper curve) in addition to the acoustic branch (lower curve). The vertical axis is the energy or frequency of the phonon and the horizontal axis is the wave vector k.

$$\mathcal{H} = \sum_k \hbar\omega_k \left(b_k^\dagger b_k + \tfrac{1}{2} \right),$$

where $b_k^\dagger b_k = n_k$ is the occupation number operator and the $b_k^\dagger$ and b_k, respectively, create and destroy a single field excitation – a phonon – with an energy of $\hbar\omega_k$. Note that the ground state of each phonon has an associated energy of $\frac{1}{2}\hbar\omega_k$, indicating that the motion of the atoms is not frozen; the atoms move randomly about their equilibrium position. At $T > 0$, the thermally excited states of the lattice vibrations can be considered as a *phonon gas*. This means that the chemical potential for adding a phonon is zero, just as in the case of a photon gas inside a cavity. Therefore, the phonon gas obeys the B–E statistics, that is, the probability of having phonons of a given frequency in a given state is

$$n\left(\omega_{k,s}\right) = \frac{1}{e^{\hbar\omega_{k,s}/kT} - 1}, \tag{15.83}$$

where ω_{ks} is the frequency of the phonons in the state.

15.4.2 Superconductivity

Superconductivity occurs when electrical resistance disappears and magnetic fields are expelled from the material. Soon after K. Onnes[16] succeeded in liquefying helium, he discovered that the electrical resistance of mercury samples suddenly dropped to zero below the critical temperature of 4.15 K; this is the defining property of superconductivity. It was soon realized that superconductivity was quite common in metals at low temperatures – with a few exceptions. Twenty-two years later W. Meissner (1882–1974) found that when a magnetic field is applied to a superconducting sample, the magnetic flux is expelled from its bulk; this is the *Meissner effect.*

[16] H. Kamerlingh Onnes (1853–1926) was a Dutch physicist famous for his discovery of superconductivity. In 1908 he was the first to liquefy helium and cool it to nearly 1.5 Kelvin, for which he was awarded the Nobel Prize in Physics in 1913.

https://doi.org/10.1017/9781009679633.016 Published online by Cambridge University Press

A breakthrough in understanding the structure of the superconducting ground state and the nature of its excitations was the BCS theory proposed in 1957,[17] which explained the transport properties in terms of a condensate of Cooper pairs.[18] In a superconductor, electrons are bound together into Cooper pairs by a weak attractive force. In a conventional superconductor, this attraction is caused by an exchange of phonons between the electrons. The evidence that phonons – the vibrations of the ionic lattice described above – are relevant for superconductivity is provided by the *isotope effect*, that is, the dependence of the superconducting critical temperature on the mass of the ions.

A few years later, B. D. Josephson predicted that a supercurrent could flow between two pieces of superconductor separated by a thin layer of insulator. The Josephson effect is exploited by superconducting quantum interference devices (SQUIDs) and is used to accurately measure the magnetic flux quantum $\Phi_0 = h/2e$ (see Section 15.4.4).

In superconductors there is a characteristic length scale ξ called the coherence length, which is a measure of the size of a Cooper pair (i.e., the distance between the two electrons that make up the pair). The electron near or at the Fermi surface moving through the lattice of a metal produces an attractive potential behind it over a distance of the order of 3×10^{-6} cm, which is large compared with the lattice spacing ($\sim 10^{-8}$ cm). In a low-temperature superconductor the coherence length is of the order of 10^{-4} cm, covering a few thousand atoms, while in the high-temperature superconductors it is only of the order of $10^{-8} - 10^{-6}$ cm.

Another important parameter is the London penetration depth (usually denoted as λ or λ_L), which characterizes the distance to which a magnetic field penetrates into the superconductor and becomes equal to e^{-1} times the magnetic field at the surface of the superconductor. Typical values of λ range from 50 to 500 nm. When $\lambda < \xi/\sqrt{2}$ the superconductor is of type I, when $\lambda > \xi/\sqrt{2}$ the superconductor is of type II.[19] In the first case the superconductivity is abruptly destroyed when a strong magnetic field is applied (as is usually the case with pure metals); type II superconductors have two critical magnetic fields and a more complex magnetic behavior.

Superconductivity is a vibrant field of research, especially since the advent of so-called high-temperature superconductors (with critical temperatures of the order of -140 C or lower at ambient pressure). Promising large-scale applications include electric power transmission, transformers, and other devices that would have significant economic and environmental benefits.

[17] Named after John Bardeen (1908–1991), Leon Cooper (1930), and Robert Schrieffer (1931–2019), who were awarded the Nobel Prize in 1972 "for their jointly developed theory of superconductivity." Bardeen had already been awarded the Nobel Prize in 1956 for his work with W. Shockley and W. H. Brattain on semiconductors, which led to the history-making invention of the transistor.

[18] *Superfluidity* in liquid ^{4}He, discovered by P. Kapitza and (independently) by J. Allen and D. Misener in 1937, had already been explained in 1938 by F. and H. London and L. Titza (with further elaborations by L. Landau) as a consequence of Bose–Einstein condensation.

[19] The coherence length ξ was introduced by Vitaly Ginzburg (1916–2009; Nobel Prize 2003) and L. Landau as part of their theory of superconductivity, according to which the electrons contributing to superconductivity form a superfluid. The ratio $\kappa = \lambda/\xi$ is known as the *Ginzburg–Landau parameter*.

https://doi.org/10.1017/9781009679633.016 Published online by Cambridge University Press

15.4.3 Bose–Einstein condensate

A Bose–Einstein condensate (BEC) is a state of matter in which extremely cold atoms exhibit collective behavior and act like a single "super atom." To reach such a state, the system, composed of a few thousand bosons (atoms), must be very close to zero Kelvin ($\sim 10^{-9}$ K). Under such conditions, a large fraction of the bosons occupy the lowest quantum states in which microscopic quantum mechanical phenomena, in particular wave-function interference, become macroscopically visible. The low speeds of motion of the bosons imply a large de Broglie wavelength, as large as the dimensions of the container, which leads to extended correlations between the variables of the particles. Indeed, the BEC is in a highly entangled state, due to the fact that the particles are coherently distributed in space.

As mentioned in Chapter 2, this new phase of matter was discovered by Einstein in 1925, following a related suggestion by Bose, hence the name Bose–Einstein condensate. For almost 70 years, the technical difficulties involved in producing BEC kept them in the laboratory, but they are now being produced conventionally.

The basic BEC concepts, such as the critical temperature or the critical particle density, can be obtained by considering an ideal gas of noninteracting bosons. The transition to BEC occurs below the critical temperature T_c, which for a uniform 3D gas composed of particles with no apparent internal degrees of freedom is

$$T_{\mathrm{c}} \simeq 3.3125\, \frac{\hbar^2 n_c^{2/3}}{m k_{\mathrm{B}}}, \tag{15.84}$$

where m is the boson mass and n_c is the particle density, given in terms of the thermal de Broglie length λ_c by

$$n_c \simeq 2.612 \frac{1}{\lambda_c^3}. \tag{15.85}$$

Since $n_c^{-1/3}$ is the average distance between particles, Eq. (15.85) means that condensation occurs when this distance becomes comparable to λ_c at the given temperature. While in air at atmospheric pressure and room temperature the number of molecules (N_2 and O_2) per cm^3 is 10^{19}, in a condensate the density is much lower, only 10^{12}–10^{13} atoms /cm^3.

A particularly spectacular quantum phenomenon occurs when the condensate is split into two separate parts, which are then brought closer until they overlap, producing a macroscopic quantum wave interference pattern. It is also possible to prepare a BEC with fermions by first pairing up fermions of opposite spin, similarly as the electrons do to form Cooper pairs.[20]

15.4.4 The quantum Hall effect

The Hall effect, discovered in 1879 by Edwin Hall (1855–1938), is a simple consequence of the motion of charged particles in a magnetic field. The underlying physics of the

[20] The superfluidity of two different phases of ^{3}He is due to the similar Bose–Einstein condensation of pairs of its fermion atoms.

https://doi.org/10.1017/9781009679633.016 Published online by Cambridge University Press

Hall effect is the following: a current-carrying plate is subject to a constant, perpendicular magnetic field. If the field points in the z-direction and the electrons are restricted to move in the xy-plane, we get the coupled equations of motion

$$m\ddot{x} = -eB\dot{y},\ m\ddot{y} = eB\dot{x}. \tag{15.86}$$

The general solution is

$$x(t) = X - R\sin(\omega_B t + \phi),\ \ y(t) = Y + R\cos(\omega_B t + \phi), \tag{15.87}$$

where $\omega_B = eB/m$ is the cyclotron frequency. When a constant current I is made to flow by applying an electric field in the x-direction, this current is deflected by the magnetic field and bent in the y-direction. In a finite material, this results in an accumulation of charge along the edge and an associated electric field in the y-direction. This continues until the electric field cancels out the bending due to the magnetic field and the electrons then only move in the x-direction. The induced electric field E_y is responsible for the Hall voltage V_H along the y-direction. The Hall coefficient is defined by

$$R_H = \frac{\rho_{xy}}{B}, \tag{15.88}$$

where ρ_{xy} is the off-diagonal component of the resistivity tensor. In the classical case, R_H depends only on the charge and density of the conducting particles, $R_H = 1/Ne$.

However, at low temperatures and in strong magnetic fields, the resistivity exhibits a peculiar behavior, as was first demonstrated by K. von Klitzing in 1980.[21] The electrons moving in the metal plate can be regarded as a confined two-dimensional electron gas experiencing quantum correlations. At low temperatures the Hall resistance becomes quantized,

$$R_H = \frac{\hbar}{ne^2}, \tag{15.89}$$

where n is an integer.[22] The Coulomb interactions lead to additional values of R_H with odd fractional numbers in the denominator, $n = 1/3, 2/3, 3/5, \ldots$; this is the *fractional quantum Hall effect*. The value of R_H remains strictly constant in a finite range of B.

These discoveries revealed the very rich physics of 2D systems at the nanoscale, which has since become an active field of theoretical and experimental research with important technological applications.

15.5 The Density Matrix. Pure States and Mixtures

As mentioned in Chapter 8, in nature we most often find quantum systems that require combinations of noninterfering states to describe the complete ensemble. Consider, for example, a variable A whose value A_i is well defined in each sub-ensemble i with

[21] Klaus von Klitzing (1943) is a German physicist best known for the discovery of the integer quantum Hall effect, for which he was awarded the Nobel Prize in Physics in 1985.

[22] It is a remarkable fact that the quantized resistance is determined only by fundamental constants. In fact, the original proposal by Klitzing et al. (1980) was to use the quantum Hall effect as a method for a highly accurate determination of the fine-structure constant $\alpha = e^2/\hbar c$.

https://doi.org/10.1017/9781009679633.016 Published online by Cambridge University Press

weight w_i; its average value $\langle A\rangle = \sum_{i=1}^{N} w_i A_i$ depends on the relative weight of the independent subensembles. For instance, in a macroscopic volume of H_2 in equilibrium the molecules can have total electron spin 0 or 1, for the rest they are entirely equivalent. So there are two different but essentially equivalent subsystems coexisting, forming what we take to be a homogeneous gas. Their corresponding wave functions do not interfere (because they are independent subsystems) and the system as a whole must be described by something that treats them equally. In the more general case, we have a set of states with different weights,

$$\begin{array}{llllll} \text{State} & |1\rangle & |2\rangle & |3\rangle & \cdots & |N\rangle \\ \text{Weight} & w_1 & w_2 & w_3 & \cdots & w_N \end{array} . \tag{15.90}$$

In such cases the system as a whole is not characterized by one wave function, but by several wave functions $|i\rangle$ and their corresponding statistical weights w_i. The system is then said to be in a mixture of pure states $|i\rangle$. In the H_2 example, we have a mixture of w_0 molecules with spin 0 and w_1 molecules with spin 1.

A simple and elegant tool to describe such situations is the density operator or density matrix, which was proposed simultaneously but independently in 1927 by Landau and von Neumann,

$$\widehat{\rho} = \sum_i w_i |i\rangle\langle i|. \tag{15.91}$$

In writing this expression it is not assumed that the states $|i\rangle$ are orthogonal, although normally they are.[23] It is often convenient to use a different basis (e.g., the eigenvectors of the operator $\hat{A}$, or the energy eigenfunctions, and so on.). We therefore expand the state $|i\rangle$ in terms of a convenient basis $\{|n\rangle\}$ as

$$|i\rangle = \sum_n a_{in}|n\rangle, \qquad a_{in} = \langle n|i\rangle. \tag{15.92}$$

Now the average value over the complete ensemble is

$$\begin{aligned} \langle\widehat{A}\rangle = \sum_i w_i A_i &= \sum_i w_i \langle i|\widehat{A}|i\rangle = \sum_i \sum_{m,n} w_i \langle n|a_{in}^* \widehat{A} a_{im}|m\rangle \\ &= \sum_i \sum_{m,n} w_i a_{in}^* a_{im} \langle n|\widehat{A}|m\rangle. \end{aligned} \tag{15.93}$$

Rewriting the coefficient of $\langle n|\widehat{A}|m\rangle$ in this expression as follows:

$$\begin{aligned} \sum_i w_i a_{in}^* a_{im} &= \sum_i w_i \langle m|i\rangle\langle i|n\rangle = \langle m|\left(\sum_i w_i |i\rangle\langle i|\right)|n\rangle \\ &= \langle m|\widehat{\rho}|n\rangle = \rho_{mn}, \end{aligned} \tag{15.94}$$

and substituting in the preceding expression for $\langle\widehat{A}\rangle$, we get

$$\langle\widehat{A}\rangle = \sum_{n,m} \rho_{mn} A_{nm} = \sum_m (\widehat{\rho}\widehat{A})_{mm} = \operatorname{tr}\widehat{\rho}\widehat{A}. \tag{15.95}$$

[23] A notable exception is well known in quantum optics, where it is useful to express the density matrix in the Glauber representation in terms of coherent states, which are not orthogonal; see Section 9.4.3.

https://doi.org/10.1017/9781009679633.016 Published online by Cambridge University Press

This is a very important result: it tells us that the ensemble average of $\widehat{A}$ is given by the trace of the product $\widehat{\rho}\widehat{A}$. Note that swapping A and ρ gives the same result for $\langle\widehat{A}\rangle$, even if they do not commute,

$$\langle\widehat{A}\rangle = \mathrm{tr}\,\widehat{\rho}\widehat{A} = \sum_{n,m} A_{nm}\rho_{mn} = \sum_n (\widehat{A}\widehat{\rho})_{nn} = \mathrm{tr}\,\widehat{A}\widehat{\rho}. \tag{15.96}$$

This is an instance of the general theorem that establishes that in the calculation of a trace it is possible to perform any cyclic exchange of the factors, even if they do not commute. For example, for noncommuting $\widehat{A}, \widehat{B}, \widehat{C}$, $\mathrm{tr}\widehat{A}\widehat{B}\widehat{C} = \mathrm{tr}\widehat{B}\widehat{C}\widehat{A} = \mathrm{tr}\widehat{C}\widehat{A}\widehat{B}$.

An alternative way of introducing the density matrix goes as follows. When the system is in the pure state $|i\rangle$, all w_k are zero except w_i. The average of $\widehat{A}$ is then simply given by

$$A_i = \langle i|\widehat{A}|i\rangle = \sum_n \langle i|\widehat{A}|n\rangle\langle n|i\rangle = \sum_n \langle n|i\rangle\langle i|\widehat{A}|n\rangle = \mathrm{tr}\,\widehat{P}_i\widehat{A}, \tag{15.97}$$

where $\widehat{P}_i = |i\rangle\langle i|$ is the projector onto the state $|i\rangle$, and $\{|n\rangle\}$ is any basis. Now we generalize this result to a mixture of states with different weights, introducing $\sum_n |n\rangle\langle n| = 1$,

$$\begin{aligned}\langle\widehat{A}\rangle &= \sum_i w_i \langle i|\widehat{A}|i\rangle = \sum_i \sum_n w_i \langle i|\widehat{A}|n\rangle\langle n|i\rangle \\ &= \sum_n \langle n| \sum_i w_i |i\rangle\langle i|\widehat{A}|n\rangle = \sum_n \langle n| \left(\sum_i w_i \widehat{P}_i\right) \widehat{A}|n\rangle \\ &= \mathrm{tr} \sum_i w_i \widehat{P}_i \widehat{A}.\end{aligned} \tag{15.98}$$

The density matrix $\hat{\rho}$ can therefore be obtained from first principles by comparing Eqs. (15.95) and (15.98),

$$\langle A\rangle = \sum_i w_i A_i = \sum_i w_i \mathrm{tr}\,\widehat{P}_i\widehat{A} = \mathrm{tr}\left(\sum_i w_i \hat{P}_i\right)\widehat{A} = \mathrm{tr}\,\hat{\rho}\widehat{A}, \tag{15.99}$$

which gives

$$\hat{\rho} = \sum_i w_i \hat{P}_i = \sum_i w_i |i\rangle\langle i|, \tag{15.100}$$

This result is the direct generalization of (15.97) to mixed states.

15.5.1 The density matrix in quantum statistical mechanics

For a more comprehensive conceptual understanding, it is useful to see the density matrix from the point of view of statistical mechanics, a subject in which it plays an important role.

Any real physical system of interest is inevitably open, that is, it is necessarily in contact with its environment, with which it can exchange energy, momentum, and even particles.[24] Therefore, none of its dynamical variables can have a perfectly defined and

[24] In particular, there is no known way of getting rid of ubiquitous fields such as the ZPF and the gravitational field.

https://doi.org/10.1017/9781009679633.016 Published online by Cambridge University Press

constant value. Even if the system is in equilibrium, only the mean values of the dynamical variables will be constant, but their instantaneous values will fluctuate constantly around their respective averages. The random, unpredictable nature of these fluctuations is due to the enormous complexity of the environment, which requires a large number of variables, often themselves subject to fluctuations, to fully characterize it.

Under these conditions, if we want to study a complex quantum system, which can be in different states depending on external conditions, we have to pay attention not only to the system, but also to those elements of the physical world to which it is related and which most strongly influence its behavior, at least from the point of view of our actual interest (for other studies, other external elements may be relevant). By including these elements as part of our "complete system," we can consider it to be disconnected from the rest of the physical world (i.e., we consider the influence of the rest of the universe to be negligible for our purposes). In this approximation, and by characterizing the state of the relevant external part with a set of vectors $|\alpha\rangle$, the state vector of the "complete system" can be written in the form

$$|\Psi\rangle = \sum_{i,\alpha} c_{i\alpha}|\alpha\rangle|i\rangle = \sum_i c_i(t)|i\rangle, \tag{15.101}$$

where the random coefficients $c_i(t) = \sum_\alpha c_{i\alpha}(t)|\alpha\rangle$ are complicated combinations of the exterior states. This description corresponds to the case where the state $|\alpha\rangle$ contributes with fluctuating probability $|c_i|^2$ to the state $|\Psi\rangle$; while $|i\rangle$ characterizes the state of the system of interest, the c_i are determined by the outside with which it is in contact.

It is not difficult to establish the relationship between the coefficients c_i and the weights w_i. This relationship must be statistical, since the c_i fluctuate, whereas the weights w_i are fixed for a given situation. In fact, we should expect the number of particles that are instantly in a given state (e.g., with energy in the interval $(E_i, E_i + \Delta E_i)$), to vary erratically. The quantum expectation value of a variable $\widehat{A}$, here denoted $\langle\widehat{A}\rangle_{\text{QM}}$, is given by

$$\langle\widehat{A}\rangle_{\text{QM}} = \langle\Psi|\widehat{A}|\Psi\rangle = \sum_{i,j} c_i^* c_j \langle i|\widehat{A}|j\rangle. \tag{15.102}$$

To determine $\langle\widehat{A}\rangle$ we still have to average this expression over the set of possible values of the c_i's, an operation that we will denote with an upper bar; note carefully that this is generally not a quantum operation. We thus obtain

$$\langle\widehat{A}\rangle = \overline{\langle\widehat{A}\rangle_{\text{QM}}} = \sum_{i,j} \overline{c_i^* c_j} \langle i|\widehat{A}|j\rangle. \tag{15.103}$$

To go further, we need to make some assumptions about the statistical behavior of the coefficients c_i. Given the enormous number of degrees of freedom involved in determining the instantaneous value of each coefficient, the most natural hypotheses are the following:

a) Postulate of a priori weights; this means

$$\overline{c_i} = 0, \quad \overline{|c_i|^2} = w_i. \tag{15.104}$$

https://doi.org/10.1017/9781009679633.016 Published online by Cambridge University Press

b) Postulate of statistical independence; if $i \neq j$, then c_i and c_j are statistically independent, that is,

$$\overline{c_i^* c_j} = 0, \quad \text{if } i \neq j. \tag{15.105}$$

Combining the two expressions, we get

$$\overline{c_i^* c_j} = w_i \delta_{ij}. \tag{15.106}$$

The resulting weights

$$w_i = \overline{c_i^* c_i} \tag{15.107}$$

are determined by physical factors belonging to the external system and the interaction mechanism; they are, therefore, outside the theoretical apparatus related to the quantum system under study here, and so we take them as given. Their determination usually corresponds to other branches of physics; for example, statistical mechanics provides rules for determining these functions under specified conditions. Substituting (15.106) into (15.103) we get

$$\langle \widehat{A} \rangle = \sum_{i,j} w_i \delta_{ij} \langle i | \widehat{A} | j \rangle = \sum_i w_i A_i. \tag{15.108}$$

We have thus recovered Eq. (15.91) as a *statistical description* of the system of interest.

Formula (15.106) gives the desired relationship between the a priori weights and the coefficients c_i of the expansion of the state vector of the complete system. On the other hand, it follows from (15.91) (considering the basis as orthonormal) that

$$\begin{aligned} \langle i | \widehat{\rho} | j \rangle &= \sum_k w_k \langle i | k \rangle \langle k | j \rangle \\ &= \sum_k w_k \langle k | j \rangle \delta_{ik} = w_i \langle i | j \rangle = w_i \delta_{ij}, \end{aligned} \tag{15.109}$$

which introduced in (15.106) gives

$$\overline{c_i^* c_j} = \langle j | \widehat{\rho} | i \rangle. \tag{15.110}$$

This expression establishes the connection between the coefficients of the state function Ψ and the matrix elements of $\widehat{\rho}$, giving the latter a direct statistical meaning. Substituting this into Eq. (15.103), we obtain the fundamental formula (15.95),

$$\langle \widehat{A} \rangle = \sum_{i,j} \langle j | \widehat{\rho} | i \rangle \langle i | \widehat{A} | j \rangle = \operatorname{tr} \widehat{\rho} \widehat{A}. \tag{15.111}$$

Equation (15.108) shows that the different states $|i\rangle$ contribute independently to the value of $\langle \widehat{A} \rangle$, as a result of the hypothesis of statistical independence (or random phases) of their coefficients, expressed in (15.105). In fact, with $\psi(x) = \sum c_i \varphi_i(x)$, the spatial density averaged over the external parameters is

$$\rho(x) = \overline{|\psi(x)|^2} = \sum_{i,j} \overline{c_i^* c_j} \varphi_i^* \varphi_j = \sum_i w_i |\varphi_i|^2 = \sum_i w_i \rho_i(x). \tag{15.112}$$

https://doi.org/10.1017/9781009679633.016 Published online by Cambridge University Press

The ensemble is manifestly an incoherent superposition, that is, probabilities (not amplitudes) are added. If the c_i were fixed and no averaging over their realizations were required, the result would be

$$\rho(x) = \sum_{ij} c_i^* c_j \varphi_i^* \varphi_j = \sum_i |c_i|^2 \rho_i(x) + \sum_{i \neq j} c_i^* c_j \varphi_i^* \varphi_j, \tag{15.113}$$

which has all the interference terms with $i \neq j$ characteristic of the coherent superpositions (of amplitudes) representing the states described by a wave function. This confirms the need to introduce the density matrix if one wishes to be able to study incoherent systems, which are more complex than those described by a single state vector.

Before concluding this section, an additional observation should be made. The weights (a priori probabilities) w_i are determined, as we have seen, by considerations external to the system under study. Their origin can be either quantum or classical, and their determination is a completely different physical problem, although not unrelated to our immediate interest. The w_i are usually taken to be constant, so that the time evolution of the density matrix is determined solely by the states it contains, that is, by the quantum rules. Conceptually, this need not be the case; we could imagine, for example, that the outside of the system is changing because the temperature of the thermal bath is changing, or for some other reason. In such cases, the evolution of the system depends on the outside, and the problem that makes sense to study is that of the whole system. For this reason, we will normally restrict ourselves to the case where the external conditions remain unchanged and the weights maintain their constant value. From now on, we will assume that this is the case.

15.5.2 Basic properties of the density matrix

All the basic properties of the density matrix follow from its definition (15.91) and from the basic rule for calculating expectation values, Eq. (15.95). Taking here $\widehat{A} = 1$, we get

$$\operatorname{tr} \widehat{\rho} = 1. \tag{15.114}$$

This result is a direct consequence of the condition

$$\sum_i w_i = 1, \tag{15.115}$$

as follows from (15.91) in combination with (15.114),[25]

$$\sum_i w_i = 1 = \sum_i w_i \langle i | \, i \rangle = \operatorname{tr} \sum_i w_i |i\rangle\langle i| = \operatorname{tr} \hat{\rho}.$$

Since, by definition, the weights w_i are positive and real,

$$w_i \geq 0, \quad w_i^* = w_i, \tag{15.116}$$

[25] If the states $|i\rangle$ are not orthogonal, Eq. (15.115) is not necessarily valid.

https://doi.org/10.1017/9781009679633.016 Published online by Cambridge University Press

$\widehat{\rho}$ is automatically self-adjoint,

$$\widehat{\rho}^{\dagger} = \left(\sum_i w_i |i\rangle\langle i| \right)^{\dagger} = \sum_i w_i |i\rangle\langle i| = \widehat{\rho}. \tag{15.117}$$

This result also follows from Eq. (15.94), which shows that $\rho_{mn} = \rho_{nm}^*$. From this expression we get

$$\rho_{nn} = \sum_i w_i |a_{in}|^2 \geq 0.$$

Since (15.114) establishes that $\sum_n \rho_{nn} = 1$, it follows that the diagonal elements of the density matrix are not only real but also satisfy the following constraint,

$$0 \leq \rho_{nn} \leq 1. \tag{15.118}$$

It follows that if the mixture contains N independent pure states, it is characterized by a Hermitian matrix of dimension $N \times N$, positive definite[26] and with trace one. Since this matrix contains $N^2 - 1$ real parameters, these are the ones needed to specify the state of the mixture characterized by $\widehat{\rho}$. The square matrix $N \times N$ contains N^2 complex elements; the hermiticity condition reduces the number of independent real parameters to N^2, and the condition on the trace reduces this number to $N^2 - 1$. For example, an arbitrary mixture of spin-1/2 states contains two independent pure states ($N = 2$), corresponding to spin down and spin up. Therefore, the polarization state of these particles (electrons, nucleons, etc.) is specified by three parameters that can be chosen as the components of the polarization vector.

Any positive definite Hermitian matrix of trace 1 can be taken as a density matrix. The proof is as follows. Let $\widehat{\rho}'$ be an operator that satisfies these requirements; in its own representation we can do the spectral decomposition, so if ρ'_n are the eigenvalues of $\widehat{\rho}'$, we write $\widehat{\rho}' = \sum_n \rho'_n |n\rangle\langle n|$; since the trace is 1, we have that $\sum_n \rho'_n = 1$, and since the eigenvalues of a Hermitian operator are real (and in this case positive definite by hypothesis), it is also true that $0 \leq \rho'_n \leq 1$. Consequently, the eigenvalues ρ'_n have the properties required to be used as statistical weights and $\widehat{\rho}'$ can be interpreted as a statistical operator.

The matrix $\widehat{\rho}$ also has the following property: in the representation in which it is diagonal, $\rho_{nm} = \rho_n \delta_{mn}$, we may write

$$\operatorname{tr} \widehat{\rho}^2 = \sum_{n,m} \rho_{nm} \rho_{mn} = \sum_n \rho_n^2 \leq \left(\sum_n \rho_n \right)^2 = (\operatorname{tr} \rho)^2 = 1. \tag{15.119}$$

Since the trace is representation-independent, we get that, in general,

$$\operatorname{tr} \widehat{\rho}^2 \leq 1. \tag{15.120}$$

The equal sign applies only to pure states.

[26] A matrix ρ is positive definite if for any vector with components A_n, $\sum \rho_{nm} A_n^* A_m > 0$ (i.e., $A^{\dagger} \rho A > 0$). That this is true here is immediately clear from Eq. (15.94), from which it follows that $\sum_{nm} \rho_{nm} A_n^* A_m = \sum_i w_i \left| \sum_n a_{in} A_n \right|^2 \geq 0$. The eigenvalues of a positive definite matrix are nonnegative.

https://doi.org/10.1017/9781009679633.016 Published online by Cambridge University Press

On the other hand, it can be shown that when changing the representation, the matrix $\widehat{\rho}$ is transformed according to the general rules of transformation of an operator. In fact, when going from the $|i\rangle$ representation to a $|p\rangle$ representation by means of the unitary operator $\widehat{U}$, $|p\rangle = \widehat{U}|i\rangle$, we get

$$\widehat{\rho} = \sum_i w_i |i\rangle\langle i| = \sum_i w_i \widehat{U}^{-1}|p\rangle\langle p|\widehat{U} = \widehat{U}^{-1}\left(\sum_i w_i |p\rangle\langle p|\right)\widehat{U}. \tag{15.121}$$

In the last equality we took into account that the eigenvalues of an operator do not change with the representation. The quantity in parentheses is the density matrix in the new representation, which we call $\widehat{\rho}'$, $\widehat{\rho} = \widehat{U}^{-1}\widehat{\rho}'\widehat{U}$, and inversion confirms the usual rule,

$$\widehat{\rho}' = \widehat{U}\widehat{\rho}\widehat{U}^{-1}. \tag{15.122}$$

From here it is easy to show that not only the eigenvalues, but, in general, the expectation values of an operator are independent of the representation. Indeed, by denoting with A' the variables in the new representation, we get

$$\langle\widehat{A}'\rangle = \operatorname{tr}\widehat{\rho}'\widehat{A}' = \operatorname{tr}\widehat{U}\widehat{\rho}\widehat{U}^{-1}\widehat{U}\widehat{A}\widehat{U}^{-1} = \operatorname{tr}\widehat{U}\widehat{\rho}\widehat{A}\widehat{U}^{-1} = \operatorname{tr}\widehat{\rho}\widehat{A} = \langle\widehat{A}\rangle. \tag{15.123}$$

In the derivation we used $\operatorname{tr}\widehat{A}\,\widehat{B} = \operatorname{tr}\widehat{B}\widehat{A}$.

Before going further, and by way of illustration, it is useful to apply the general theory to the particular case of the coordinate representation (assuming this makes sense in the case we are interested in). We write $\widehat{\rho}$ in the form

$$\widehat{\rho} = \sum_{i,m,n} w_i a_{im} a_{in}^* |m\rangle\langle n| = \sum_{m,n} \rho_{mn}|m\rangle\langle n|, \tag{15.124}$$

where

$$\rho_{mn} = \sum_i w_i a_{im} a_{in}^*. \tag{15.125}$$

From here we obtain

$$\rho(x', x) \equiv \langle x'|\widehat{\rho}|x\rangle = \sum_{m,n} \rho_{mn}\langle x'|m\rangle\langle n|x\rangle, \tag{15.126}$$

and denoting with $\varphi_n(x) = \langle x|n\rangle$ the wave function of state $|n\rangle$, we obtain the important result

$$\rho(x', x) = \sum_{m,n} \rho_{mn}\varphi_m(x')\varphi_n^*(x), \tag{15.127}$$

which gives for the expectation value of $\widehat{A}$

$$\langle\widehat{A}\rangle = \operatorname{tr}\widehat{\rho}\widehat{A} = \int dx\,\langle x|\widehat{\rho}\widehat{A}|x\rangle. \tag{15.128}$$

Writing the integrand in the form

$$\langle x|\widehat{\rho}\widehat{A}|x\rangle = \int dx'\langle x|\rho|x'\rangle\langle x'|\widehat{A}|x\rangle = \int dx'\rho(x, x')\langle x'|\widehat{A}|x\rangle, \tag{15.129}$$

we get

$$\langle\widehat{A}\rangle = \int dx dx'\rho(x, x')\langle x'|\widehat{A}|x\rangle. \tag{15.130}$$

https://doi.org/10.1017/9781009679633.016 Published online by Cambridge University Press

This is just the specific form of the general rule $\langle \widehat{A} \rangle = \mathrm{tr}\left(\widehat{\rho}\widehat{A}\right)$ for $\widehat{\rho}$ and $\widehat{A}$ continuous matrices. In general, we can write the matrix element of $\widehat{A}$ – which acts on x – as

$$\langle x'|\widehat{A}|x\rangle = \widehat{A}_x \langle x'|x\rangle = \widehat{A}_x \delta(x'-x), \tag{15.131}$$

where the subscript x means that $\widehat{A}$ acts only on the variable x. For example, setting $\hat{A} = 1$ yields $\langle x'|1|x\rangle = \delta(x'-x)$, which when introduced in (15.130) confirms that the trace of ρ is 1,

$$1 = \int dx dx' \rho(x,x')\delta(x'-x) = \int \rho(x,x)dx. \tag{15.132}$$

Similarly, with $\widehat{A} = f(x)$, an arbitrary function of x, we have $\langle x'|f(x)|x\rangle = f(x)\delta(x'-x)$ and

$$\langle f(x)\rangle = \int f(x)\rho(x,x)dx. \tag{15.133}$$

From these results it follows that the spatial density of particles is given by the diagonal element of $\rho(x,x')$,

$$\rho(x) = \rho(x,x). \tag{15.134}$$

On the other hand, for $\widehat{A} = \widehat{p}$, from (15.131) we obtain

$$\langle x'|\widehat{p}|x\rangle = -i\hbar \frac{\partial}{\partial x}\delta(x'-x), \tag{15.135}$$

from which, after integration by parts, we get

$$\langle \widehat{p}\rangle = -i\hbar \int dx dx' \rho(x,x')\frac{\partial}{\partial x}\delta(x'-x) = i\hbar \int dx \left[\frac{\partial \rho(x,x')}{\partial x}\right]_{x'=x}. \tag{15.136}$$

15.5.3 Time evolution of the density matrix

In order to construct the equation for the time evolution of $\widehat{\rho}$ we write the state $|i\ (t)\rangle$ at time t in terms of the state $|i\ (t_0)\rangle$ at an initial time t_0 and apply the evolution operator $\widehat{S}(t,t_0)$,

$$|i\ (t)\rangle = \widehat{S}(t,t_0)|i(t_0)\rangle, \qquad \widehat{S}(t) = e^{-i\widehat{H}t/\hbar}. \tag{15.137}$$

In turn, the evolution of $\widehat{S}$ satisfies the rule

$$i\hbar \frac{\partial \widehat{S}}{\partial t} = \widehat{H}\widehat{S}. \tag{15.138}$$

Equation (15.137) allows us to relate $\widehat{\rho}(t)$ to $\widehat{\rho}(t_0)$, assuming constant weights (the result is immediate using Eq. (15.122) with $\hat{U} = \hat{S}$),

$$\begin{aligned}\widehat{\rho}(t) &= \sum_i w_i |i(t)\rangle\langle i(t)| = \sum_i w_i \widehat{S}|i(t_0)\rangle\langle i(t_0)|\widehat{S}^\dagger \\ &= \widehat{S}(t,t_0)\left[\sum_i w_i |i(t_0)\rangle\langle i(t_0)|\right]\widehat{S}^\dagger(t,t_0) = \widehat{S}(t,t_0)\widehat{\rho}(t_0)\widehat{S}^\dagger(t,t_0).\end{aligned} \tag{15.139}$$

https://doi.org/10.1017/9781009679633.016 Published online by Cambridge University Press

By taking the partial derivative with respect to time, we get, using Eq. (15.138) and its adjoint,

$$\frac{\partial \widehat{\rho}}{\partial t} = \frac{\partial \widehat{S}}{\partial t}\widehat{\rho}(t_0)\widehat{S}^\dagger + \widehat{S}\widehat{\rho}(t_0)\frac{\partial \widehat{S}^\dagger}{\partial t}$$
$$= \frac{1}{i\hbar}\left[\widehat{H}\,\widehat{S}\widehat{\rho}(t_0)\widehat{S}^\dagger - \widehat{S}\widehat{\rho}(t_0)\widehat{S}^\dagger\widehat{H}\right] = \frac{1}{i\hbar}\left[\widehat{H}\widehat{\rho}(t) - \widehat{\rho}(t)\widehat{H}\right], \tag{15.140}$$

that is,

$$i\hbar\frac{\partial \widehat{\rho}}{\partial t} = [\widehat{H}, \widehat{\rho}\,]. \tag{15.141}$$

This is known as the von Neumann equation. Note that the sign in this expression is the opposite of that corresponding to the equation of motion for a dynamical variable in Heisenberg's description, Eq. (8.102). The reason is that (15.140) actually means that $i\hbar\,(d\widehat{\rho}/dt) = i\hbar\,(\partial\widehat{\rho}/\partial t) + [\widehat{\rho}, \widehat{H}] = 0$, that is,

$$\frac{d\widehat{\rho}}{dt} = 0, \tag{15.142}$$

which is the quantum version of Liouville's theorem.[27] If we define the Liouville operator as $\hat{L}\cdot = [\widehat{H}, \cdot]$, we get $\hat{L}\hat{\rho} = [\widehat{H}, \widehat{\rho}\,]$, and the von Neumann equation takes the form

$$i\hbar\frac{\partial \hat{\rho}}{\partial t} = \hat{L}\hat{\rho}, \tag{15.143}$$

which is known as the quantum Liouville equation. In the particular, but very important case that the Hamiltonian is time-independent, we have $\widehat{S}(t, t_0) = e^{-\frac{i}{\hbar}\widehat{H}(t-t_0)}$, whence

$$\widehat{\rho}(t) = e^{-\frac{i}{\hbar}\widehat{H}(t-t_0)}\widehat{\rho}(t_0)e^{\frac{i}{\hbar}\widehat{H}(t-t_0)}. \tag{15.144}$$

In the energy representation $\widehat{H}|n\rangle = E_n|n\rangle$, one gets

$$\langle n|\rho(t)|n'\rangle = \langle n|e^{-\frac{i}{\hbar}\widehat{H}(t-t_0)}\rho(t_0)e^{\frac{i}{\hbar}\widehat{H}(t-t_0)}|n'\rangle$$
$$= e^{-i\omega_{nn'}(t-t_0)}\langle n|\widehat{\rho}(t_0)|n'\rangle, \tag{15.145}$$

indicating that the time evolution of $\widehat{\rho}(t)$ is reduced in this representation to an oscillation of the nondiagonal matrix elements, with constant diagonal matrix elements.

From Eq. (15.140) it follows that if $\widehat{\rho}$ depends only on integrals of motion, the system is in equilibrium, $\partial\widehat{\rho}/\partial t = 0$. This result has its classical counterpart, which states that any distribution that depends only on integrals of motion corresponds to a state of equilibrium. This parallelism results from the fact that the evolution equation for the density matrix (15.140) can be regarded as the quantum analog of the Liouville equation, as we have seen. However, it must still be emphasized that the latter equation corresponds to a description in phase space, whereas in the quantum problem the evolution takes place in Hilbert space.

[27] In its classical Hamiltonian version, Liouville's theorem says that the number of states contained in a given region of phase space is conserved.

https://doi.org/10.1017/9781009679633.016 Published online by Cambridge University Press

Exercise E15.5. Definition of a stationarity time

Assume a "stationarity time" T_s defined by the relation

$$\frac{1}{T_s^2} = \langle(\Delta\widehat{\dot{\rho}})^2\rangle, \tag{15.146}$$

where $\widehat{\rho}$ is the density matrix. Discuss the validity of this definition and show that for any quantum system,

$$T_s^2\langle(\Delta\widehat{H})^2\rangle \geq \hbar^2. \tag{15.147}$$

When does this equality hold?

Solution. Although for a nonstationary system, $\widehat{\dot{\rho}} \neq 0$, the mean value $\langle\widehat{\dot{\rho}}\rangle$ is always zero (see Problem P15.19). This suggests using the dispersion of $\widehat{\dot{\rho}}$ as a measure of the degree of stationarity of the system, as in Eq. (15.146). Since T_s is given directly by the density matrix and does not refer to a particular dynamical variable, this stationarity time characterizes the (state of the) system globally. Note carefully the difference with times that refer to a specific variable, which may also satisfy Heisenberg's inequalities, but are related to the evolution of that variable.

Taking into account that $\langle\widehat{\dot{\rho}}\rangle = 0$, and using $\widehat{\rho} = \sum_n w_n|n\rangle\langle n|$, we get

$$\begin{aligned}\langle(\Delta\widehat{\dot{\rho}})^2\rangle &= \langle\widehat{\dot{\rho}}^2\rangle = \operatorname{tr}\widehat{\rho}\widehat{\dot{\rho}}^2 = \sum_{m,n}\langle m|w_n|n\rangle\langle n|\widehat{\dot{\rho}}^2|m\rangle \\ &= \sum_n w_n\langle n|\widehat{\dot{\rho}}^2|n\rangle = \sum_{n,k} w_n\langle n|\widehat{\dot{\rho}}|k\rangle\langle k|\widehat{\dot{\rho}}|n\rangle.\end{aligned} \tag{15.148}$$

Using now the evolution equation for the density matrix (15.140), we note that

$$\begin{aligned}i\hbar\langle n|\widehat{\dot{\rho}}|k\rangle &= \langle n|\widehat{H}\widehat{\rho} - \widehat{\rho}\widehat{H}|k\rangle \\ &= \sum_m\langle n|\widehat{H}w_m|m\rangle\langle m|k\rangle - \sum_m\langle n|w_m|m\rangle\langle m|\widehat{H}|k\rangle = (w_k - w_n)H_{nk}.\end{aligned} \tag{15.149}$$

Substituting, we can write

$$\langle(\Delta\widehat{\dot{\rho}})^2\rangle = \frac{1}{\hbar^2}\sum_{n,k} w_n(w_n - w_k)^2 H_{nk}H_{kn} \leq \frac{1}{\hbar^2}\sum_{n,k\neq n} w_n H_{nk}H_{kn}. \tag{15.150}$$

The inequality follows from the fact that $(w_n - w_k)^2 \leq 1$, since $0 \leq w_k \leq 1$. On the other hand, the following relations hold:

$$\begin{aligned}\langle(\Delta\widehat{H})^2\rangle &= \sum_n w_n H_{nn}^2 - \left(\sum\nolimits_n w_n H_{nn}\right)^2 \\ &= \sum_n w_n\langle(\Delta\widehat{H})^2\rangle_n + \left[\sum\nolimits_n w_n(H_{nn})^2 - \left(\sum\nolimits_n w_n H_{nn}\right)^2\right] \\ &\geq \sum_n w_n\langle(\Delta\widehat{H})^2\rangle_n,\end{aligned} \tag{15.151}$$

where $\langle(\Delta\widehat{H})^2\rangle_n$ is the dispersion of the energy in the pure state n. Since the dispersion can be written as the sum over the nondiagonal elements,

https://doi.org/10.1017/9781009679633.016 Published online by Cambridge University Press

$$\langle(\Delta\widehat{H})^2\rangle_n = \langle n|\widehat{H}^2|n\rangle - \left(\langle n|\widehat{H}|n\rangle\right)^2$$
$$= \sum_k \langle n|\widehat{H}|k\rangle\langle k|\widehat{H}|n\rangle - \langle n|\widehat{H}|n\rangle\langle n|\widehat{H}|n\rangle = \sum_{k\neq n} H_{nk}H_{kn}, \tag{15.152}$$

we get

$$\langle(\Delta\widehat{H})^2\rangle \geq \sum_{n,k\neq n} w_n \widehat{H}_{nk}\widehat{H}_{kn}. \tag{15.153}$$

Substituting into the expression for the dispersion of $\widehat{\rho}$ and using the definition of T_s, Eq. (15.146), we obtain Eq. (15.147). From this it follows that the equals sign occurs if and only if the system is in a pure state.

Exercise E15.6. Evolution equation for $\widehat{\rho}$ in the interaction representation
Write the equation of evolution of a quantum ensemble described by a density matrix in the interaction description.

Solution. If V is the interaction potential, the Hamiltonian can be written in the form $\widehat{H} = \widehat{H}_0 + \widehat{V}$, so the density matrix satisfies the Heisenberg equation

$$i\hbar\frac{\partial\widehat{\rho}}{\partial t} = [\widehat{H},\widehat{\rho}\,] = \widehat{H}_0\widehat{\rho} - \widehat{\rho}\widehat{H}_0 + \widehat{V}\widehat{\rho} - \widehat{\rho}\widehat{V}. \tag{15.154}$$

In the interaction description we have

$$\widehat{\rho}_{\rm int} = e^{i\widehat{H}_0 t/\hbar}\widehat{\rho}e^{-i\widehat{H}_0 t/\hbar}, \tag{15.155}$$

$$\widehat{H}_{0\rm int} = \widehat{H}_0, \quad \widehat{V}_{\rm int} = e^{i\widehat{H}_0 t/\hbar}\widehat{V}e^{-i\widehat{H}_0 t/\hbar}, \tag{15.156}$$

and the time derivative gives

$$i\hbar\frac{\partial\widehat{\rho}_{\rm int}}{\partial t} = i\hbar\left(\frac{i}{\hbar}\widehat{H}_0\widehat{\rho}_{\rm int} - \widehat{\rho}_{\rm int}\frac{i}{\hbar}\widehat{H}_0\right) + i\hbar e^{i\widehat{H}_0 t/\hbar}\frac{\partial\widehat{\rho}}{\partial t}e^{-i\widehat{H}_0 t/\hbar}. \tag{15.157}$$

Using Eq. (15.154) and simplifying, the desired result is obtained,

$$i\hbar\frac{\partial\widehat{\rho}_{\rm int}}{\partial t} = [\widehat{V}_{\rm int},\widehat{\rho}_{\rm int}]. \tag{15.158}$$

This is the quantum Liouville equation in the interaction representation; if we replace the commutator multiplied by $1/i\hbar$ by a Poisson bracket, we obtain the classical Liouville equation in the interaction description.

15.5.4 Pure states

From the preceding discussion it is clear that the description of a physical system in terms of a density matrix is the most general quantum description available and includes as a special case the description of pure states.

https://doi.org/10.1017/9781009679633.016 Published online by Cambridge University Press

When the system is in the pure state $|p\rangle$, the mixture is reduced to only this state. The pure state is therefore defined by the condition

$$w_i = \delta_{pi} \tag{15.159}$$

and the corresponding density matrix $\widehat{\rho} = \sum_i w_i |i\rangle\langle i| = \sum_i \delta_{pi} |i\rangle\langle i|$ coincides with the projector for the state $|p\rangle$,

$$\widehat{\rho} = |p\rangle\langle p| = \widehat{P}_p. \tag{15.160}$$

For example, in the coordinate representation, the density matrix of the pure state $\psi_p(x)$ is

$$\rho_p(x', x) = \langle x'|p\rangle\langle p|x\rangle = \psi_p(x')\psi_p^*(x). \tag{15.161}$$

Note that in particular

$$\rho_p(x, x) = \psi_p(x)\psi_p^*(x) = \rho(x). \tag{15.162}$$

From Eq. (15.160) it follows that

$$\widehat{\rho}^2 = |p\rangle\langle p|p\rangle\langle p| = |p\rangle\langle p| = \widehat{\rho}, \tag{15.163}$$

so the density matrix of a pure state is idempotent. This property is used to define (construct) a pure state.

Since an algebraic relation between matrices also holds for their eigenvalues, the eigenvalues ρ' of the density matrix of a pure state satisfy $\rho'^2 = \rho'$, so they can only be 0 and 1.

For a mixture, $\mathrm{tr}\,\widehat{\rho}^2 \leq 1$; the equal sign only applies if it is a pure state ($\mathrm{tr}\,\widehat{\rho} = 1$ always). This allows us to define a degree of impurity of a mixture as

$$\mathcal{I} \equiv 1 - \mathrm{tr}\widehat{\rho}^2 = \mathrm{tr}\left(\widehat{\rho} - \widehat{\rho}^2\right) = \mathrm{tr}\widehat{\rho}\,(1 - \widehat{\rho})\,, \tag{15.164}$$

representing the expectation value of $\langle 1 - \widehat{\rho}\rangle$, which is equal to 0 for a pure state and tends to 1 as the impurity increases (when $\mathrm{tr}\,\widehat{\rho}^2 \ll 1$). Equation (15.163) can be written as $\sum w_n^2 |n\rangle\langle n| = \sum w_n |n\rangle\langle n|$, whence $w_n^2 = w_n$, that is, an idempotent density matrix can contain weights of value 0 or 1 only.

The extreme opposite of the pure case is when the complete lack of information about the possible state of any of the N subsystems leads to an equal (a priori) weighting of all of them,

$$\mathrm{tr}\hat{\rho}^2 = \sum^N w_n^2 = \sum^N \frac{1}{N^2} = \frac{1}{N}. \tag{15.165}$$

For $N \to \infty$ this quantity can be arbitrarily small. This is the minimum value that $\mathrm{tr}\,\hat{\rho}^2$ can take. In fact, if two of the weights take the values $\frac{1}{N} + \delta$ and $\frac{1}{N} - \delta$ (the trace is one), there is a change in $\mathrm{tr}\,\hat{\rho}^2$ given by $\left(\frac{1}{N} + \delta\right)^2 + \left(\frac{1}{N} - \delta\right)^2 - \frac{2}{N^2} = 2\delta^2 \geq 0$, and the value of $\mathrm{tr}\,\hat{\rho}^2$ increases for any $\delta \neq 0$.

https://doi.org/10.1017/9781009679633.016 Published online by Cambridge University Press

15.5.5 Spin polarization

When the density matrix involves a small number of parameters, it is possible, and even convenient, to express it in terms of results directly accessible from the experiment. This possibility often feeds an operational view for the construction of the density matrix that pervades (and even distorts) the literature on the subject. The following topic is an instructive example of considerable physical interest, since it deals with the construction of a density matrix to describe the polarization states of particles with spin 1/2. The procedure is applicable to any problem for which a two-dimensional Hilbert space is appropriate.

We write the spin operator in terms of the Pauli matrices, namely

$$\hat{\boldsymbol{S}} = \tfrac{1}{2}\hbar\hat{\boldsymbol{\sigma}}. \tag{15.166}$$

A general ensemble of spin-1/2 particles is in a superposition of the different possible polarization states, so the expectation value of the spin is

$$\langle\widehat{\boldsymbol{S}}\rangle = \tfrac{1}{2}\hbar\langle\hat{\boldsymbol{\sigma}}\rangle = \tfrac{1}{2}\hbar\boldsymbol{P}. \tag{15.167}$$

This defines the polarization vector,

$$\boldsymbol{P} \equiv \langle\hat{\boldsymbol{\sigma}}\rangle. \tag{15.168}$$

We now express the polarization vector as the average over the density matrix of the corresponding polarization state,

$$\boldsymbol{P} = \langle\hat{\boldsymbol{\sigma}}\rangle = \operatorname{tr}\widehat{\rho}\hat{\sigma}. \tag{15.169}$$

To construct the density matrix that goes into this formula, we take into account that the three Pauli matrices plus the identity matrix 2×2 form a basis for the representation of the operators acting in a 2D space; then $\widehat{\rho}$, which has the dimension 2×2, can be written in the form

$$\widehat{\rho} = a_0\mathbb{I} + a_1\widehat{\sigma}_1 + a_2\widehat{\sigma}_2 + a_3\widehat{\sigma}_3. \tag{15.170}$$

Taking the trace of this expression and using $\operatorname{tr}\widehat{\sigma}_i = 0$, we get $\operatorname{tr}\widehat{\rho} = 1 = a_0\operatorname{tr}\mathbb{I} = 2a_0$, so that $a_0 = 1/2$. Now taking the i component of (15.169) and introducing (15.170), we get

$$P_i = \operatorname{tr}\widehat{\rho}\,\widehat{\sigma}_i = \operatorname{tr}\left[a_0\widehat{\sigma}_i + \sum\nolimits_k a_k\widehat{\sigma}_k\widehat{\sigma}_i\right] = \sum_k a_k\operatorname{tr}\widehat{\sigma}_k\widehat{\sigma}_i = 2a_i. \tag{15.171}$$

The latter result follows from the fact that if i and k are different, the product $\widehat{\sigma}_i\widehat{\sigma}_k$ is another Pauli matrix, therefore the only contributing term is $i = k$, with $\widehat{\sigma}_i^2 = \mathbb{I}$. This gives $a_i = P_i/2$. Collecting results, we get the revealing expression

$$\widehat{\rho} = \tfrac{1}{2}(\mathbb{I} + \boldsymbol{P}\cdot\hat{\sigma}), \tag{15.172}$$

or, in explicit matrix form,

$$\widehat{\rho} = \tfrac{1}{2}\begin{pmatrix} 1+P_3 & P_1 - iP_2 \\ P_1 + iP_2 & 1-P_3 \end{pmatrix}. \tag{15.173}$$

https://doi.org/10.1017/9781009679633.016 Published online by Cambridge University Press

We see that knowing the three components P_i of the polarization vector is sufficient to completely determine the density matrix. This means that the results of three independent measurements (three experiments) are sufficient to determine the spin density matrix. The pragmatic application of this point of view provides the basis for the operational interpretation of the density matrix mentioned earlier.

Exercise E15.7. Density operator and spin polarization

A spin polarized beam is in a pure state described by the spinor $|p\rangle = \begin{pmatrix} \sqrt{1/3} \\ \sqrt{2/3} \end{pmatrix}$. Find the corresponding density matrix and spin polarization vector.

Solution. In terms of the basis

$$|+\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |-\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \tag{15.174}$$

the pure state represented by $|p\rangle$ is

$$|p\rangle \equiv \begin{pmatrix} \sqrt{1/3} \\ \sqrt{2/3} \end{pmatrix} = \sqrt{\tfrac{1}{3}}\,|+\rangle + \sqrt{\tfrac{2}{3}}\,|-\rangle\,. \tag{15.175}$$

Its density matrix is therefore given by

$$\hat{\rho} = |p\rangle\langle p| = \begin{pmatrix} \sqrt{1/3} \\ \sqrt{2/3} \end{pmatrix} \begin{pmatrix} \sqrt{1/3} & \sqrt{2/3} \end{pmatrix} = \frac{1}{3}\begin{pmatrix} 1 & \sqrt{2} \\ \sqrt{2} & 2 \end{pmatrix}. \tag{15.176}$$

Since it is a pure state, this matrix must be idempotent, and indeed it is,

$$\hat{\rho}^2 = \tfrac{1}{9}\begin{pmatrix} 3 & 3\sqrt{2} \\ 3\sqrt{2} & 6 \end{pmatrix} = \tfrac{1}{3}\begin{pmatrix} 1 & \sqrt{2} \\ \sqrt{2} & 2 \end{pmatrix} = \hat{\rho}. \tag{15.177}$$

On the other hand, this density matrix can be written in terms of the polarization vector $\boldsymbol{P} = \langle\hat{\boldsymbol{\sigma}}\rangle$ in the form (15.173). Comparing (15.176), we get

$$\tfrac{1}{2}(1 + P_3) = \tfrac{1}{3}, \quad \tfrac{1}{2}(1 - P_3) = \tfrac{2}{3}; \tag{15.178}$$

$$\tfrac{1}{2}(P_1 - iP_2) = \tfrac{\sqrt{2}}{3}, \quad \tfrac{1}{2}(P_1 + iP_2) = \tfrac{\sqrt{2}}{3}. \tag{15.179}$$

The components of the polarization vector are therefore

$$P_1 = \tfrac{2\sqrt{2}}{3}, \quad P_2 = 0, \quad P_3 = -\tfrac{1}{3}, \tag{15.180}$$

that is,

$$\boldsymbol{P} = \left(\tfrac{2\sqrt{2}}{3},\, 0,\, -\tfrac{1}{3}\right). \tag{15.181}$$

This vector lies in the xz-plane and its magnitude is $P^2 = 1$. The latter is to be expected, since every pure state with spin 1/2 is fully polarized in one direction.

---

https://doi.org/10.1017/9781009679633.016 Published online by Cambridge University Press

15.5.6 The density matrix and quantum statistics

We have seen that the density matrix allows us to make the most general description possible of an open quantum system and that it can be considered as the quantum counterpart of the classical distribution function, since its evolution equation is the quantum version of the Liouville equation. We will now use this perspective to build a theory of systems that require both statistical and quantum treatment, such as those that obey quantum statistics. We will therefore deal simultaneously with classical statistics and the statistical properties of quantum systems.

15.5.6.1 The canonical and grand canonical ensembles

Let us consider a system of macroscopic dimensions – which we will call a macrosystem – composed of a large number of (micro) systems, such as a volume of gas, and so on. If the macrosystem is in equilibrium at a fixed temperature T and E_n represents the eigenvalue of the Hamiltonian $\widehat{H}$ of one of its component subsystems, then the probability that it is in the energy state E_n is given by the Maxwell–Boltzmann distribution

$$w_n = \frac{1}{Z} e^{-\beta E_n}, \tag{15.182}$$

with $\beta = 1/kT$, where $Z = \sum_n e^{-\beta E_n}$ is the partition function, such that $\sum w_n = 1$. Note that each state contributes to the value of Z with a number of terms equal to the order of the degeneracy of the state. The density matrix now takes the form (for simplicity, we consider nondegenerate states)

$$\widehat{\rho} = \frac{1}{Z} \sum_n e^{-\beta E_n} |n\rangle\langle n|, \tag{15.183}$$

where $|n\rangle$ represents one of the eigenvectors of $\widehat{H}$, $\widehat{H}|n\rangle = E_n|n\rangle$. Note that we have used classical statistics to describe the states of the macrosystem, and now we apply quantum statistics to the result. Note also that for the partial eigenstates $|k\rangle$ of the Hamiltonian we get

$$\widehat{\rho}\,|k\rangle = \frac{1}{Z} \sum_n e^{-\beta \widehat{H}} |n\rangle \, \langle n \, |k\rangle = \frac{1}{Z} e^{-\beta E_k} |k\rangle = w_k |k\rangle, \tag{15.184}$$

that is, they are simultaneously eigenstates of the density matrix, with eigenvalue equal to the statistical weight of the state. It is possible to rewrite Eq. (15.183) independently of any representation, since

$$\widehat{\rho} = \frac{1}{Z} \sum_n e^{-\beta \widehat{H}} |n\rangle\langle n| = \frac{1}{Z} e^{-\beta \widehat{H}} \sum_n |n\rangle\langle n|, \tag{15.185}$$

and taking into account that $|n\rangle$ is a member of a complete set,

$$\widehat{\rho} = \frac{1}{Z} e^{-\beta \widehat{H}}. \tag{15.186}$$

The ensemble of systems described by this density matrix is by definition, a canonical ensemble. A canonical ensemble is the statistical ensemble that describes the possible states of a mechanical system in thermal equilibrium at a fixed temperature. Since

https://doi.org/10.1017/9781009679633.016 Published online by Cambridge University Press

$\widehat{\rho}$ depends only on the Hamiltonian, which is an integral of motion, it is an ensemble in equilibrium, whose thermodynamic properties are determined by the partition function Z. Since

$$\operatorname{tr}\widehat{\rho} = 1 = \frac{1}{Z}\operatorname{tr} e^{-\beta\widehat{H}}, \tag{15.187}$$

the partition function is

$$Z = \operatorname{tr} e^{-\beta\widehat{H}}. \tag{15.188}$$

Substituting this result into Eq. (15.186) we get

$$\widehat{\rho} = \frac{e^{-\beta\widehat{H}}}{\operatorname{tr} e^{-\beta\widehat{H}}}. \tag{15.189}$$

The weights w_n are given by Eq. (15.183),

$$w_n = \frac{e^{-\beta E_n}}{\sum_n e^{-\beta E_n}}. \tag{15.190}$$

It is important to note that the sum in the denominator of this expression is over the states, not over the eigenvalues, which is not the same when there is degeneracy.

If the ensemble has the additional property of allowing the number of particles in the system to vary, it is called a grand canonical ensemble. The density matrix then has the form

$$\widehat{\rho} = \frac{1}{Z}e^{\beta(\mu\hat{n}-\widehat{H})}, \quad Z = \sum_n e^{\beta(\mu\hat{n}-\widehat{H})}, \tag{15.191}$$

where μ is the chemical potential and $\hat{n}$ is the number operator.

Without going further into statistical quantum mechanics, which is the subject of a more advanced course, it is important to emphasize here that, with the proposed methods, it is possible to construct a statistical description of quantum systems without introducing the notion of phase space.

Exercise E15.8. The Wigner distribution function

As shown in Section 3.3.1, the Weyl transform of a function of two variables $f(x, y)$ is $\frac{1}{\pi\hbar}\int_{-\infty}^{\infty} f(x - z, y + z)e^{-2ipz/\hbar}dz$. Show that the Wigner distribution function, defined as the Weyl transform of the density matrix, has properties that allow it to be interpreted as the quantum equivalent of a probability density in phase space.

Solution. The Wigner distribution $P_W(x, p)$ associated with the density matrix $\widehat{\rho}$, which in the coordinate representation is $\rho(x', x'')$, is obtained by writing $x'' = x - z, x' = x + z$ and performing a Fourier transformation on z (the Fourier variable is $k = p/\hbar$). Therefore,

$$P_W(x, p) = \frac{1}{\pi\hbar}\int_{-\infty}^{\infty}\langle x + z|\widehat{\rho}\,|x - z\rangle e^{-2ipz/\hbar}dz. \tag{15.192}$$

For a pure state, $\langle x'|\widehat{\rho}\,|x''\rangle = \psi^*(x'')\psi(x')$ and (15.192) reduce to

$$P_W(x, p) = \frac{1}{\pi\hbar}\int_{-\infty}^{\infty}\psi^*(x - z)\psi(x + z)e^{-2ipz/\hbar}dz. \tag{15.193}$$

https://doi.org/10.1017/9781009679633.016 Published online by Cambridge University Press

To prove that $P_W(x,p)$ plays the role of a probability density in phase space, we first calculate the marginal probabilities. Integrating over p, we get

$$\int P_W(x,p)dp = \frac{1}{\pi\hbar}\int \rho(x+z,x-z)e^{-2ipz/\hbar}dp$$
$$= \frac{1}{2\pi}2\pi\int \rho(x+z,x-z)\delta(z) = \rho(x,x) = \rho(x). \tag{15.194}$$

To calculate $\int_{-\infty}^{\infty} P_W(x,p)dx$ it is preferable to first obtain the density matrix in momentum space. Taking $\phi_i(p) = \widetilde{\psi}_i(p)$, we obtain

$$\rho(p,p') = \sum_i w_i\varphi_i(p)\varphi_i^*(p') = \frac{1}{2\pi\hbar}\int\left[\sum_i w_i\psi_i(x)\psi_i^*(x')\right]e^{-\frac{i}{\hbar}(px-p'x')}dxdx', \tag{15.195}$$

that is,

$$\rho(p,p') = \frac{1}{2\pi\hbar}\int \rho(x,x')e^{-\frac{i}{\hbar}(px-p'x')}dxdx'. \tag{15.196}$$

The probability in momentum space is therefore given by

$$P(p) = \rho(p,p) = \frac{1}{2\pi\hbar}\int \rho(x',x'')e^{-\frac{i}{\hbar}p(x'-x'')}dx'dx''$$
$$= \frac{1}{2\pi\hbar}\int \rho(x+z,x-z)e^{-2ipz/\hbar}2dxdz, \tag{15.197}$$

so, comparing with (15.192), we get

$$\int_{-\infty}^{\infty} P_W(x,p)dx = P(p). \tag{15.198}$$

Since $\rho(x)$ and $P(p)$ are both normalized to unity, it follows from (15.194) or (15.198) that

$$\int_{-\infty}^{\infty} dx\int_{-\infty}^{\infty} dp P_W(x,p) = 1. \tag{15.199}$$

The properties (15.194), (15.198), and (15.199) suggest that $P_W(x,p)$ can be considered the quantum equivalent of the probability density in phase space. This is a frequent and useful application of the Wigner distribution; however, care must be taken with this interpretation, since the function $P_W(x,p)$ can take negative values.

A variety of distributions are found in the literature, although the Wigner distribution is the most widely used. In particular, it is used extensively in the development of stochastic optics pioneered by T. W. Marshall and E. Santos in the 1980s in order to understand the role of the ZPF in quantum optics experiments.

15.5.7 Decoherence

For many years, quantum measurement theory has attempted to provide a convincing answer to the problem posed by the intertwining of the states of the instrument used for measurement and those of the microsystem being measured.

https://doi.org/10.1017/9781009679633.016 Published online by Cambridge University Press

Suppose a microsystem is described by an initial wave function which is a coherent superposition of all possible eigenstates of a certain variable we want to measure. Since a single result is obtained, after the measurement the microsystem is assigned a single state, namely an eigenstate of the measured variable. After many similar measurements, the ensemble will no longer be in a pure state, but will be described by a statistical mixture (a density matrix) of the different outcomes, with their respective probabilities.[28] This transition is not unitary and therefore does not conform to all the laws of QM. Because of its importance for the present discussion, we will briefly present the central idea here.

Consider a quantum system described in terms of the set of eigenstates $\{|\varphi_k\rangle\}$ of a particular operator whose value is to be measured with an instrument A, characterized by eigenstates $|A_k\rangle$ and eigenvalues a_k.[29] If the meter reads a_1 after the measurement, the microsystem is said to be in state $|\varphi_1\rangle$, and so on. Assume that the initial state of the whole system (microsystem + instrument) is described by the vector

$$|\Psi\rangle = \sum_k c_k \, |A_k\rangle \, |\varphi_k\rangle \, . \tag{15.200}$$

This pure state corresponds to the density matrix

$$\hat{\rho}_{\text{in}} = |\Psi\rangle \, \langle\Psi| = \sum_{k,k'} c_k c_{k'}^* \, (|A_k\rangle \, \langle A_{k'}|) \, (|\varphi_k\rangle \, \langle\varphi_{k'}|) \, . \tag{15.201}$$

The different variants of measurement theory propose that carrying out the measurement is equivalent in one way or another to averaging over the states of the instrument, which is taken to correspond to the partial trace over the orthonormal vectors $|A_l\rangle$, thus obtaining a reduced density matrix,

$$\begin{aligned} \hat{\rho}_{\text{red}} = \sum_l \langle A_l \, |\Psi\rangle \, \langle\Psi| \, A_l\rangle &= \sum_l \sum_{k,k'} c_k c_{k'}^* \, (\langle A_l \, |A_k\rangle \, \langle A_{k'}| \, A_l\rangle) \, (|\varphi_k\rangle \, \langle\varphi_{k'}|) \\ &= \sum_k |c_k|^2 \, |\varphi_k\rangle \, \langle\varphi_k| \, . \end{aligned} \tag{15.202}$$

This result tells us that the measurement will give the value k with probability $|c_k|^2$, and so on. The point is that the reduced $\hat{\rho}_{\text{red}}$ no longer describes a pure state, but a mixture. This is clear from the absence of interference terms in $\hat{\rho}_{\text{red}}$. The transition $\hat{\rho}_{\text{in}} \to \hat{\rho}_{\text{red}}$ corresponds to what is called reduction or collapse of the state vector (or the wave function, if you prefer). But this transition is not unitary, since a unitary evolution transforms a pure state into another pure state and never into a mixture. We see that the operation of taking the partial trace over a portion of the whole system, thereby reducing a density matrix, is not part of Schrödinger's theory.

Entangled states such as the one described by Eq. (15.200) are, as we have seen, very common in the description of composite quantum systems. The entangled components can often be more than two, and all of them can correspond to microsystems governed by quantum laws. Because they are pure, entangled states imply coherence

[28] The theory of the density matrix and the notion of pure states are discussed in Section 15.5. A previous reading of this section is useful for a thorough understanding of the following discussion.

[29] This may viewed as an extrapolation of the cat example, with the consequences it implies.

https://doi.org/10.1017/9781009679633.016 Published online by Cambridge University Press

between the different terms that make up the wave function, as is clear from the example of Eq. (15.201) with the presence of interference terms. These states are ubiquitous in quantum theory, but they are not observed in macroscopic systems because of the huge number of components, which means that coherence must be lost in the transition from micro to macro (in particular to a cat; see Section 14.5.3). We have already seen that a system governed by quantum laws evolves in a unitary manner, so such a transition process cannot occur within the framework of this theory. In short, QM is not the place for reduced matrices; they belong to other fields, such as information theory.

In an attempt to resolve this contradiction – and to avoid having to introduce the unconvincing notion of (nonlocal) collapse – it has been argued that an isolated microsystem is an idealization and that, in reality, every system coexists with an extremely complex and chaotic environment. Therefore, a realistic description should be done using a state vector or a density matrix given in its simplest form by Eqs. (15.200) and (15.201), where now the states $|A_k\rangle$ refer to the environment in which the microsystem is inserted. Averaging over the states of the environment is justified not only by their expected relevance, but more fundamentally by the fact that their chaotic nature makes them unknown, inviting us to average over their possible realizations. By averaging, we get a reduced result as given by Eq. (15.202), where all initial coherence has been erased. The density matrix thus obtained refers to description-independent probabilities, typical of macroscopic systems.

The enormous complexity of the environment means that the eigenvalues of interest are effectively a continuum, so that any coupling of the microsystem to the environment produces important changes. For example, estimates of the effects of the cosmic microwave background radiation at a temperature as low as 3 Kelvin show that it is capable of producing decoherence in times as short as 10^{-23} s, thus apparently, but only apparently, settling the problem of achieving consistency between quantum predictions and empirical data.

We conclude that appealing to our ignorance of the specific and detailed conditions of a part of the total system does not solve the problem in principle. Nor is it solved by considering that the real state is the entangled one, but that for all practical purposes this can be interpreted as if the reduction had taken place. A bona fide description of the system must include all accessible information about the whole system (such as that given by Eqs. (15.200) and (15.201)), and cannot be arbitrarily replaced by the reduced description unless such a reduction has a direct physical meaning and corresponds to an irreversible effect.

Although there are several variants of the decoherence program, none of them has gained sufficient acceptance to be considered a solution to the problem. An attempt of a slightly different nature is to modify the Schrödinger equation itself by adding nonlinear terms in such a way that the behavior of microsystems is practically unaffected, but a loss of coherence is obtained when the body attains macroscopic dimensions. An ingenious and remarkable version of such tentative theories, known as the spontaneous localization model, is due to Ghirardi, Rimini, and Weber (1986). However, it should be clear that these theories are ad hoc, since the form of the added stochastic nonlinear terms, as well as the value of their parameters, are chosen to obtain the desired results, so that they can only be considered as provisional.

https://doi.org/10.1017/9781009679633.016 Published online by Cambridge University Press

Problems

P15.1 Show that the particle exchange operator $\widehat{P}_{ij}$ is Hermitian and that it commutes with $\widehat{P}_{nm}$ only when (i, j) and (n, m) refer to different pairs of particles (there is no common particle).

P15.2 Show that the operators

$$\widehat{P}_{ij}^{\pm} \equiv \tfrac{1}{2}\left(1 \pm \widehat{P}_{ij}\right) \tag{15.203}$$

are projectors. What is the effect of these operators on a fully symmetric state function $\Psi(\xi_1, \ldots, \xi_i, \xi_j, \ldots, \xi_N)$?

P15.3 Obtain the solution (15.49) using perturbation theory for degenerate systems. The degenerate wave functions are considered to be $\psi_{\pm}$.

P15.4 Show that if $\psi_{Nn}(x_1, x_2)$ is the wave function of the two linearly coupled oscillators (see Eq. (9.125)), then

$$\widehat{P}_{12}\psi_{Nn} = (-1)^n \psi_{Nn}, \tag{15.204}$$

where N, n represent the excitation quantum number of the center of mass and relative motion, respectively.

P15.5 The Hamiltonian of a system of three linearly coupled identical spin-zero bosons is

$$\widehat{H} = \sum_{i=1}^{3} \frac{\widehat{p}_i^2}{2m} + \sum_{i=1}^{3} \tfrac{1}{2} m^2\omega^2 x_i^2 + \sum_{j>i} \sum_{i=1}^{3} \tfrac{1}{2} m\omega^2\beta \left(x_i - x_j\right)^2. \tag{15.205}$$

In terms of the normal coordinates

$$Z = \tfrac{1}{3}(x_1 + x_2 + x_3); \quad z_1 = x_1 - x_2; \quad z_2 = x_3 - \tfrac{1}{2}\left(x_1 + x_2\right), \tag{15.206}$$

the Hamiltonian takes the form

$$\widehat{H} = \frac{\widehat{P}^2}{2M} + \tfrac{1}{2} M\omega^2 Z^2 + \frac{\widehat{p}_{z_1}^2}{2m_1} + \tfrac{1}{2} m_1\omega_1^2 z_1^2 + \frac{\widehat{p}_{z_2}^2}{2m_2} + \tfrac{1}{2} m_2\omega_2^2 z_2^2, \tag{15.207}$$

where

$$\widehat{P} = \widehat{p}_1 + \widehat{p}_2 + \widehat{p}_3; \quad \widehat{p}_{z_1} = \tfrac{1}{2}(\widehat{p}_1 - \widehat{p}_2); \quad \widehat{p}_{z_2} = \tfrac{2}{3}(\widehat{p}_3 - \frac{\widehat{p}_1 + \widehat{p}_2}{2});$$
$$\omega_1^2 = \omega_2^2 = \omega^2(1 + 3\beta); \quad M = 3m; \quad m_1 = \tfrac{1}{2}m; \quad m_2 = \tfrac{2}{3}m. \tag{15.208}$$

(a) Find the solutions of the Schrödinger equation.
(b) Determine the eigenvalues of the Hamiltonian.
(c) Indicate which states are physically realizable.
(d) Show that solutions with exchange degeneracy are not necessarily orthogonal.

P15.6 Consider a system of four uncoupled, equal and collinear harmonic oscillators. Construct the wave functions and specify the energy eigenvalues of the physically realizable steady states when:

(a) All four particles are bosons with spin zero.
(b) All four particles are fermions with spin 1/2.

https://doi.org/10.1017/9781009679633.016 Published online by Cambridge University Press

P15.7 The deuteron (consisting of a proton and a neutron) has spin 1. List the possible states of spin and total angular momentum of a system of two deuterons in a state of angular momentum L.

P15.8 Three particles of spin zero are rigidly joined to form an equilateral triangle that rotates on a circle of radius r. Determine the rotational levels of this system.

P15.9 Let $\widehat{F}$ be an operator of the form

$$\widehat{F}(1,2,\ldots,N) = \widehat{f}(1) + \widehat{f}(2) + \cdots + \widehat{f}(N) = \sum_{i=1}^{N} \widehat{f}(i), \tag{15.209}$$

where $\widehat{f}(i)$ operates only on the coordinates of particle i in a system of N identical particles. Show that

$$\left\langle \Psi^A \mid \widehat{F} \mid \Psi^A \right\rangle = \sum_{i=1}^{N} \left\langle \psi_i \mid \widehat{f}(i) \mid \psi_i \right\rangle, \tag{15.210}$$

where Ψ^A is the antisymmetric wave function (15.32) and the wave functions ψ_i of a particle are orthonormal.

P15.10 Prove that the lattice Hamiltonian (15.76) transforms into (15.80) in the reciprocal space.

P15.11 Prove that in the limit $m_2 \to m_1$, Eq. (15.82) is reduced to (15.81).

P15.12 Prove that $\mathrm{tr}\widehat{A}\widehat{B}\widehat{C} = \mathrm{tr}\,\widehat{C}\widehat{A}\widehat{B}$ and that $\mathrm{tr}\left[\widehat{A},\widehat{B}\right]\widehat{C} = tr\widehat{A}\left[\widehat{B},\widehat{C}\right]$.

P15.13 State whether the following operators can be taken as density matrices and whether they describe or can describe pure states and under what conditions:

$$\begin{pmatrix} \frac{1}{2} & ia \\ -ia & \frac{1}{2} \end{pmatrix}, \quad \begin{pmatrix} \frac{1}{1+b^2} & e^{i\alpha} \\ e^{-i\alpha} & \frac{b^2}{1+b^2} \end{pmatrix}, \quad \begin{pmatrix} 0.3 & iZ\beta \\ iZ\beta & 0.7 \end{pmatrix}. \tag{15.211}$$

The parameters a, b, Z, α are real, but β can be complex.

P15.14 A physical system can be in two independent states. Show that the most general density matrix describing this situation has the form

$$\widehat{\rho} = \begin{pmatrix} a & be^{i\varphi} \\ be^{-i\varphi} & 1-a \end{pmatrix}. \tag{15.212}$$

Determine the requirements that a and b must meet in general and the conditions under which the described state is pure.

P15.15 Prove that the operators $\Lambda_n^{\pm}=(1\pm\hat{\sigma}\cdot\widehat{\mathbf{n}})/2$ are idempotent and mutually orthogonal, that is, that they are projection operators. Study their action on a density matrix, both for pure states and for mixtures.

P15.16 Construct the density operator and the polarization vector that correspond to the pure state described by the spinor $\begin{pmatrix} \sqrt{1/3} \\ \sqrt{2/3} \end{pmatrix}$.

P15.17 A physical system can be in three states $|1\rangle, |2\rangle, |3\rangle$ with probabilities $\frac{1}{2}, \frac{3}{8}, \frac{1}{8}$, respectively. Build the corresponding density matrix. How many additional conditions can be imposed?

https://doi.org/10.1017/9781009679633.016 Published online by Cambridge University Press

P15.18 Show that the free-particle density matrix in the momentum representation is

$$\rho(p, p') = \delta(p - p')e^{-\beta p^2/2m}. \tag{15.213}$$

P15.19 Show that $\langle\widehat{\hat{\rho}}\rangle = 0$ is always true.

P15.20 Consider a mixture of the form $\hat{\rho} = \lambda\hat{\rho}_1 + (1 - \lambda)\,\hat{\rho}_2,\ 0 < \lambda < 1$. Prove that the dispersion of a generic dynamic variable $\hat{A}$ satisfies the condition

$$\left\langle(\Delta_\rho \hat{A})^2\right\rangle \geq \lambda\left\langle(\Delta_{\rho_1} \hat{A})^2\right\rangle + (1 - \lambda)\left\langle(\Delta_{\rho 2}\hat{A})^2\right\rangle.$$

When does the equality hold?

Bibliographical Notes

Tomonaga's Story of Spin (1997) is especially enjoyable to read. In particular, it contains a detailed discussion of the relationship between spin and statistics.

For a comprehensive theoretical discussion of phonons, see Srivastava (2022).

For an early authoritative explanation of superconductivity and Cooper pairs based on physical intuition, see Weisskopf (1979).

A classic reference on superconductivity is Tinkham (1996).

There are a number of textbooks on condensed matter which contain extensive descriptions of the quantum Hall effect. Several very good lecture notes on the quantum Hall effect can be found on www.damtp.cam.ac.uk/user/tong/qhe.html.

For a review of the Wigner distribution theory, see Hillery et al. (1984).

For the application of the Wigner distribution in stochastic optics, see Marshall and Santos (1988) and Casado et al. (2019).

For an extensive review of the many dynamical models proposed over the years to elucidate quantum measurements, see Allaverdyan et al. (2013). This review paper also contains a flexible and realistic model to account for the emergence of classicality in the measurement process.

https://doi.org/10.1017/9781009679633.016 Published online by Cambridge University Press