14 The Electron Spin. Entangled States

The introduction of spin in QM has far-reaching consequences. The most obvious of them have to do with the Pauli exclusion principle, which is responsible for the orbital structure of atoms and molecules, the subject of Chapter 16. More generally, as will be seen in Chapter 15, the statistics of quantum particles is determined by whether their spin is integer or half-integer.

This chapter is essentially devoted to the spin of the electron. The first part deals with its groundbreaking discovery, its introduction into quantum formalism, and some of its most important effects on atomic spectra, notably the anomalous Zeeman effect.

Historically, in QM spin has been considered as an angular momentum that particles can have by the mere fact of their existence, which is called "intrinsic" and does not allow any explanation. To address this shortcoming, in Section 4 we present a possible explanation for the origin of electron spin as a result of its interaction with the ZPF.

Section 5 introduces the entangled system of two particles with spin, which provides an opportunity for a discussion – necessarily schematic – of the Schrödinger cat and the Einstein–Podolsky–Rosen (EPR) thought experiment, as well as the Bell inequalities.

14.1 Discovery of the Electron Spin

In the years prior to the establishment of QM, there had been intensive efforts to explain the structure of atomic spectra. In particular, some progress had been made in describing the Zeeman effect, but a full explanation of the phenomenon had to wait until electron spin could be incorporated into Schrödinger's theory. As a reminder of the confusion that prevailed in those years, a certain form of the phenomenon, precisely the part that could not be explained in terms of classical models, is still called the *anomalous Zeeman effect* to distinguish it from what is known as the *normal Zeeman effect*, which corresponds to the part due to orbital angular momentum and was studied in Chapter 13.

A detailed analysis of the anomalous Zeeman effect led A. Landé[1] in 1921 to propose a modification of the classical formula for the gyromagnetic ratio, that is, the

[1] Alfred Landé (1888–1976) was a German-US theoretical physicist who carried out a deep and successful study of the anomalous Zeeman effect, in parallel with a search during forty years for an objectively realistic description of the quantum systems. He became a constant and fierce opponent of the Copenhagen interpretation and of the dualistic (particle and wave) view, and wrote several books on the foundations of the theory.

https://doi.org/10.1017/9781009679633.015 Published online by Cambridge University Press

relationship between the magnetic moment of a circulating charge (e.g., an orbiting electron) and its corresponding angular momentum. Specifically, instead of the classical formula (12.139)

$$\frac{\mu_z}{L_z} = \frac{e}{2mc} = -\frac{e_0}{2mc}, \tag{14.1}$$

Landé proposed to write

$$\frac{\mu_z}{L_z} = -g\frac{e_0}{2mc}, \tag{14.2}$$

where g is a numerical factor. Landé himself was able to determine the value of g as a function of the quantum numbers of the atomic system on empirical grounds, following an exhaustive study of the known experimental data on the Zeeman effect, but without theoretical support. Furthermore, the formula for g had the surprising peculiarity of requiring the use of half-integer quantum numbers (see Eq. (14.41)).

Prior to Landé's work, in 1915, the formula (14.2) had been used experimentally in the case of electrons by Einstein in collaboration with the Dutch physicist Wander Johannes de Haas (1878–1960). The idea behind the experiment was to test Ampère's hypothesis that the magnetization of a ferromagnetic core is due to the elementary magnets generated by the currents in each of its molecules. To test the theory, Einstein and de Haas suspended a magnetizable cylinder from a quartz filament and applied an alternating current to a coil around the cylinder, ensuring that the frequency of the induced current matched the frequency of the torsional oscillation to take advantage of resonance. With this device they verified that the change in magnetization produces a similar change in angular momentum, so that μ_z/L_z remains constant and negative. However, as they focused on the order of magnitude of the phenomenon rather than its details, their reported numerical value was not very precise. Subsequent careful measurements of the Einstein–de Haas effect conducted by different authors [Barnett (1917); Stewart (1918)] gave $g = 2$. Therefore, the anomalous value 2 of the gyromagnetic ratio for the electrons of a ferromagnetic body was indeed known before Landé's proposal and could be taken as independent evidence for the necessity to adopt Eq. (14.2) instead of (14.1) in the general atomic case.

At about the same time (1921), O. Stern[2] proposed an experiment to verify the recently discovered angular-momentum quantization rules in the presence of a magnetic field. As we saw in Chapter 11, the component of the orbital momentum in the direction of the magnetic field can only have the value $\hbar m$, where m is an integer. Therefore, the projection μ_z of the magnetic moment can only have one of the $(2l+1)$ eigenvalues given by

$$\mu_z = -\mu_0 m, \tag{14.3}$$

where μ_0 is the Bohr magneton. This surprising result, obtained by Sommerfeld in 1916 within the framework of Bohr's theory, became known as space quantization,[3] and it was this result that Stern proposed to verify. Following Stern's proposal, the

[2] Otto Stern (1888–1969) was a German–US American physicist who, with Walther Gerlach (1889–1979), contributed to the discovery of spin. He was awarded the Nobel Prize in Physics in 1943 for his contribution to the development of the molecular beam method and his discovery of the magnetic moment of the proton.

[3] This quantization displeased Stern so much that he proposed to verify it experimentally.

https://doi.org/10.1017/9781009679633.015 Published online by Cambridge University Press

German physicist Walter Gerlach (1889–1979) carried out an experiment in which a beam of monovalent atoms was passed through a highly inhomogeneous magnetic field. According to electrodynamics, the force acting on a magnetic moment (a dipole or a circulating current) passing through a magnetic field $\boldsymbol{B}$ is given by

$$\boldsymbol{F} = \nabla(\boldsymbol{\mu} \cdot \boldsymbol{B}). \tag{14.4}$$

If the z axis is oriented along the field and y along the direction of the beam, the rapid precession of the magnetic moment around the axis causes the component μ_x to average to zero, so that the force acting on the magnetic moment in the z-direction can be approximated as

$$F_z = \mu_x \frac{\partial B_x}{\partial z} + \mu_z \frac{\partial B_z}{\partial z} \simeq \mu_z \frac{\partial B_z}{\partial z} = -\mu_0 \frac{\partial B_z}{\partial z} m = -F_0 m, \tag{14.5}$$

with $F_0 = \mu_0 \partial B_z / \partial z$. The angular momentum is randomly oriented, so classical theory would predict a uniform distribution for μ_z from $-\mu$ to $+\mu$, producing a single extended patch on the screen. However, since according to Eq. (14.5) the force on the dipole depends on the quantum number m, the beam must split into as many branches as there are different values of this number, that is, $2l + 1$; furthermore, since l is an integer according to Bohr's theory, $2l + 1$ should be odd.

The experiment was first carried out with atoms from group I of the periodic table (alkali metals H, Li, Na, K, Rb, Cs, Fr) in the s-state, for which $l = 0$ and $m = 0$. It turned out that *two* separate beams formed, indicating that the atoms had a nonzero magnetic moment in the s-state! With an angular momentum $\hbar s$ associated with the magnetic moment, $2s + 1 = 2$ gives $s = 1/2$. Although the result was consistent with Landé's empirical theory, the origin of the half-integer angular momentum remained unexplained.

In parallel with this series of works, various researchers were working on the origin and structure of the periodic table of elements. In particular, generalizing Bohr's views and developing a proposal by Edmund Stoner (1899–1968) concerning the possible number of stationary states of an atom in an external magnetic field, Pauli concluded in 1925 that it was possible to understand the structure of the periodic table by considering that each electronic orbital state was determined by four quantum numbers instead of the three known n, l, m, and by simultaneously assuming that each state could be occupied by only one electron. According to Pauli, the proposed fourth number should be no more than a dichotomous variable in order to avoid an excessive increase in the size of the atomic Hilbert space.

This momentous proposal, known as the *Pauli exclusion principle*, has far-reaching consequences, some of which we will have the opportunity to analyze in due course. In Chapter 6 we saw that the free electrons in a metal cannot all occupy the lowest energy level as a result of the exclusion principle. At this point, it is interesting to highlight the discovery of the need for a new quantum number to characterize the atomic state of an electron, which, according to Pauli, should have only two possible values. In fact, the exclusion principle not only solved the problem of the number of electrons that could be accommodated in each atomic shell[4] with a well-defined angular momentum,

[4] The organization of the atomic electrons into shells was first proposed by Bohr and later extended to all atoms.

https://doi.org/10.1017/9781009679633.015 Published online by Cambridge University Press

but also gave a qualitative answer to a much more fundamental old question, namely, what prevents the electrons of a heavy atom (which can contain dozens of them) from all falling into the lowest-energy state, which would seem to be a more stable solution?[5]

In 1925, the young Dutch physicists G. E. Uhlenbeck[6] and S. Goudsmit[7] jointly proposed a reinterpretation of the previous results that resolved several of the difficulties in a very simple way. Instead of considering the new dichotomous quantum number as a characteristic of the state of the atom, they proposed to interpret it as an intrinsic angular momentum of the electron, with a value of 1/2 in units of $\hbar$. This is precisely the spin of the electron, whose formal quantum treatment was formulated shortly afterwards (1927) by Pauli himself, and whose description we will develop in what follows. The value 1/2 is uniquely fixed by the condition that the spin projection in a certain direction, identified with the new quantum number, can take only the two values $+\frac{1}{2}\hbar$ and $-\frac{1}{2}\hbar$, as required by the Pauli exclusion principle and by the Stern–Gerlach experiments.

14.2 The Quantum Theory of Spin

Since the electron spin is $\hbar/2$, we can apply to it the angular-momentum theory developed in Section 11.3 and express it in terms of the 2×2 Pauli matrices

$$\widehat{\boldsymbol{S}} = \tfrac{1}{2}\hbar\widehat{\boldsymbol{\sigma}}, \tag{14.6}$$

where $\widehat{\boldsymbol{\sigma}}$ is the 3D vector with components $\widehat{\sigma}_i$, the Pauli matrices,

$$\hat{\sigma}_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \hat{\sigma}_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \hat{\sigma}_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{14.7}$$

and

$$[\sigma_i, \sigma_j] = 2i\epsilon_{ijk}\sigma_k, \tag{14.8}$$

Furthermore, the Pauli matrices satisfy the relation

$$\sigma_i\sigma_j = \delta_{ij} + i\epsilon_{ijk}\sigma_k, \tag{14.9}$$

which shows in particular that the square of each one of them is the unit 2×2 matrix. Multiplying (14.9) by $\widehat{A}_i\widehat{B}_j$, where $\widehat{A}_i$ and $\widehat{B}_j$ are the Cartesian components of two arbitrary vector operators that commute with $\widehat{\boldsymbol{\sigma}}$, and adding over repeated indices, we obtain the useful formula

$$(\widehat{\boldsymbol{\sigma}} \cdot \widehat{\boldsymbol{A}})(\widehat{\boldsymbol{\sigma}} \cdot \widehat{\boldsymbol{B}}) = \widehat{\boldsymbol{A}} \cdot \widehat{\boldsymbol{B}} + i\widehat{\boldsymbol{\sigma}} \cdot (\widehat{\boldsymbol{A}} \times \widehat{\boldsymbol{B}}), \qquad \text{for } \left[\widehat{\boldsymbol{A}}, \widehat{\boldsymbol{\sigma}}\right] = \left[\widehat{\boldsymbol{B}}, \widehat{\boldsymbol{\sigma}}\right] = 0. \tag{14.10}$$

[5] The present argument is a posteriori. In fact, Pauli's arguments (and Stoner's at the same time) concerned the anomalous Zeeman effect (discussed in Section 3) and, incidentally, the atomic configuration of rare gases. We recall that Pauli, guided by his positivist views, referred to the dichotomous variable he had discovered as an ambivalence, avoiding any use of models that he considered superfluous – a point of view that has permeated the literature until today. This issue is revisited in the next section.

[6] George Eugene Uhlenbeck (1900–1988) was a Dutch–US theoretical physicist, best known for having introduced together with Goudsmit, the electron spin as an intrinsic angular momentum.

[7] Samuel Abraham Goudsmit (1902–1978) was a Dutch–US theoretical physicist, famous for jointly proposing the concept of electron spin as an intrinsic angular momentum.

https://doi.org/10.1017/9781009679633.015 Published online by Cambridge University Press

Since the electron spin is represented by a 2×2 matrix operator, the wave function of the electron with spin requires two components; this structure is called a spinor, and has the form $\begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}$, see Eq. (11.144). Spinors are mathematical objects that must be clearly distinguished from other objects with several components, such as a vector on the plane. In particular, the transformation properties under rotation of the coordinate system are so peculiar to spinors that the former are used to define the latter. As there is not enough space to develop this point fully, we will limit ourselves to explaining its basic aspects.

We recall that Eq. (8.167) is the general formula for the operator of rotation by the angle θ around the $\widehat{\mathbf{n}}$ axis through the origin. With the angular-momentum operator $\widehat{\boldsymbol{S}}$ given by Eq. (14.6), we obtain for the rotation of a spin 1/2

$$\widehat{T}_{\widehat{\mathbf{n}}}(\theta) = e^{-\frac{i}{\hbar}\theta\widehat{\mathbf{n}}\cdot\widehat{\boldsymbol{S}}} = e^{-\frac{1}{2}i\theta\widehat{\mathbf{n}}\cdot\widehat{\boldsymbol{\sigma}}}. \tag{14.11}$$

Expanding the exponential into a power series and taking into account that according to (14.10) (see Eq. (14.107))

$$(\widehat{\mathbf{n}}\cdot\widehat{\boldsymbol{\sigma}})^{2k} = \mathbb{I}, \quad (\widehat{\mathbf{n}}\cdot\widehat{\boldsymbol{\sigma}})^{2k+1} = \widehat{\mathbf{n}}\cdot\widehat{\boldsymbol{\sigma}}, \tag{14.12}$$

we get

$$\widehat{T}_{\widehat{\mathbf{n}}}(\theta) = \mathbb{I}\cos\frac{\theta}{2} - i(\widehat{\mathbf{n}}\cdot\widehat{\boldsymbol{\sigma}})\sin\frac{\theta}{2}. \tag{14.13}$$

Note the half angle, characteristic of spinors. In particular, for the operator representing an infinitesimal rotation around the direction $\widehat{\mathbf{n}}$ by the angle $\delta\theta$, this gives

$$\widehat{T}_{\widehat{\mathbf{n}}}(\delta\theta) = \mathbb{I} - \tfrac{1}{2}i\delta\theta\widehat{\mathbf{n}}\cdot\widehat{\boldsymbol{\sigma}}. \tag{14.14}$$

Now let Ψ be the spinor that describes the electron state of interest

$$\Psi = \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}, \tag{14.15}$$

where ψ_1 and ψ_2 are one-component wave functions (solutions of the Schrödinger equation). According to the general theory of transformations discussed in Section 8.6, an infinitesimal rotation of the coordinate system described by $\widehat{T}_{\widehat{\mathbf{n}}}(\delta\theta)$ transforms the Ψ spinor into Ψ', where

$$\Psi' = \widehat{T}_{\widehat{\mathbf{n}}}(\delta\theta)\Psi = \left(1 - i\tfrac{1}{2}\delta\theta\widehat{\mathbf{n}}\cdot\widehat{\boldsymbol{\sigma}}\right)\Psi. \tag{14.16}$$

We see that the transformation is carried out with half the angle of rotation. This property defines spinors and essentially distinguishes them from tensors (scalars A, vectors A_i, second-rank tensors A_{ij}, etc.), since tensors are transformed by integral multiples of the angle. For example, a scalar remains unchanged after a rotation around the z axis, $A = A'$, while the x-component of a 3D vector $\boldsymbol{A}$ becomes $A'_x = A_x\cos\theta + A_y\sin\theta$, and so on. In contrast, the spinor Ψ becomes

$$\Psi' = \left(1 - i\tfrac{1}{2}\delta\theta\hat{\sigma}_3\right)\Psi = \begin{pmatrix} \left(1 - i\frac{1}{2}\delta\theta\right)\psi_1 \\ \left(1 + i\frac{1}{2}\delta\theta\right)\psi_2 \end{pmatrix} = \Psi - i\tfrac{1}{2}\delta\theta\begin{pmatrix} \psi_1 \\ -\psi_2 \end{pmatrix}. \tag{14.17}$$

https://doi.org/10.1017/9781009679633.015 Published online by Cambridge University Press

An immediate consequence of this property is that a spinor is not invariant under a rotation by the angle 2π. Indeed, from (14.13) it follows that

$$\Psi' = \widehat{T}_z(2\pi)\Psi = (\cos\pi)\Psi = -\Psi. \tag{14.18}$$

Therefore a spinor cannot directly represent a physical quantity (an observable); the physical quantities must be bilinear combinations of spinors, which do remain invariant against rotations by an angle of 2π due to the double change of sign. This is not a limitation, since we know from the general rules of QM that the quantities of interest, such as the expectation values

$$\langle\widehat{A}\rangle = \int \Psi^\dagger \widehat{A}\Psi dx, \tag{14.19}$$

involve bilinear (quadratic) combinations of spinors.

It is clear that an expression like (14.19) is to be understood as a matrix formula, with $\widehat{A}$ a 2×2 matrix whose components may themselves be operators acting on the Hilbert space spanned by the ψ_i. For example, the particle density is

$$\Psi^\dagger\Psi = \begin{pmatrix}\psi_1^* & \psi_2^*\end{pmatrix}\begin{pmatrix}\psi_1 \\ \psi_2\end{pmatrix} = \psi_1^*\psi_1 + \psi_2^*\psi_2; \tag{14.20}$$

the expectation value of $\widehat{S}_z$ is

$$\begin{aligned}\langle\widehat{S}_z\rangle &= \tfrac{1}{2}\hbar\int dx\,\Psi^\dagger\widehat{\sigma}_z\Psi = \tfrac{1}{2}\hbar\int dx\begin{pmatrix}\psi_1^* & \psi_2^*\end{pmatrix}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\begin{pmatrix}\psi_1\\ \psi_2\end{pmatrix}\\ &= \tfrac{1}{2}\hbar\int dx(\psi_1^*\psi_1 - \psi_2^*\psi_2);\end{aligned} \tag{14.21}$$

the expectation value of $\widehat{S}_x$ is

$$\langle\widehat{S}_x\rangle = \tfrac{1}{2}\hbar\int dx\begin{pmatrix}\psi_1^* & \psi_2^*\end{pmatrix}\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}\begin{pmatrix}\psi_1\\ \psi_2\end{pmatrix} = \tfrac{1}{2}\hbar\int dx(\psi_1^*\psi_2 + \psi_2^*\psi_1), \tag{14.22}$$

and so on.

14.2.1 The Pauli equation

We now proceed to construct an evolution equation for Ψ which replaces (or rather, generalizes) the Schrödinger equation to the case of the electron with spin. The usual approach, introduced by Pauli, is to take into account that when an electron is placed in a magnetic field, its spin magnetic moment is coupled to the field. The magnetic moment in the z direction is $\widehat{\mu}_z = -g\frac{e_0}{2mc}\widehat{S}_z$, as follows from Eq. (14.2) by replacing $\widehat{L}_z$ with $\widehat{S}_z$. In vector notation we can then write, with $g = 2$ and $\mu_0 = e_0\hbar/2mc$ (the Bohr magneton),

$$\widehat{\boldsymbol{\mu}} = -2\frac{e_0}{2mc}\widehat{\mathbf{S}} = -\frac{e_0}{2mc}\hbar\widehat{\boldsymbol{\sigma}} = -\mu_0\widehat{\boldsymbol{\sigma}}. \tag{14.23}$$

This result shows that the magnitude of μ is (by definition) one Bohr magneton. The coupling to a field $\boldsymbol{B}$ then gives rise to the mutual energy

$$\widehat{V} = -\widehat{\boldsymbol{\mu}}\cdot\boldsymbol{B} = \mu_0\boldsymbol{B}\cdot\widehat{\boldsymbol{\sigma}}. \tag{14.24}$$

https://doi.org/10.1017/9781009679633.015 Published online by Cambridge University Press

This interaction term is to be added to the Hamiltonian of the Schrödinger equation, now written as a matrix equation for the spinor Ψ (see Eq. (14.28)). The resulting equation is called the Pauli equation: it correctly describes the behavior of the nonrelativistic electron with spin, as we will have the opportunity to see when we study, for example, the anomalous Zeeman effect in the following section.

The Pauli equation is a postulate that introduces the spin and the correct value of the intrinsic magnetic moment of the electron ad hoc, with $g = 2$. However, this equation can be derived from Dirac's fundamental theory of the relativistic electron, which naturally includes spin (see Chapter 19). In order not to remain on a purely phenomenological level, we will derive the Pauli equation from a more fundamental postulate than the mere introduction of the correct intrinsic magnetic moment, which then follows from the formalism (see the bibliographical note at the end of the chapter).

When an electron (more generally, a charged particle) is immersed in an electromagnetic field, the mechanical momentum $\boldsymbol{p}$ appearing in the Hamiltonian must be replaced by the canonical momentum $\boldsymbol{P} = \boldsymbol{p} - (e/c)\boldsymbol{A}$, as we have seen in Section 12.6. On the other hand, it follows from Eq. (14.10) that $(\hat{\boldsymbol{\sigma}} \cdot \hat{\boldsymbol{p}})^2 = \hat{\boldsymbol{p}}^2$. This means that in the Schrödinger equation we can replace $\hat{\boldsymbol{p}}^2$ by $(\hat{\boldsymbol{\sigma}} \cdot \hat{\boldsymbol{p}})^2$; each of the two components of the new wave function Ψ then satisfies the Schrödinger equation. We therefore postulate that to describe an electron with spin 1/2 in the presence of an external magnetic field, the substitution $\hat{\boldsymbol{p}}^2 \to \hat{\boldsymbol{P}}^2 \to (\hat{\boldsymbol{\sigma}} \cdot \hat{\boldsymbol{P}})^2$ must be made in the Hamiltonian.

The operator $\hat{\boldsymbol{p}}$ acts on the Hilbert space of the Schrödinger eigenfunctions, $\mathcal{H}_H$, while $\hat{\boldsymbol{\sigma}}$ acts on a newly introduced Hilbert space for the spin eigenfunctions, $\mathcal{H}_s$. The substitution thus expands the initial Hilbert space $\mathcal{H}_H$ to the product space $\mathcal{H}_H \otimes \mathcal{H}_s$ and transforms the Schrödinger scalar wave functions into two-component spinors. This is a simple way of extending the Hilbert space to include spinor functions, which is further justified by confirming that the resulting equations maintain their invariance against the Galileo group of transformations (3D translations and rotations), as is consistent with a nonrelativistic theory. Moreover, it is the nonrelativistic version of a procedure used to construct the Dirac equation, as discussed in Section 19.2.

According to the proposed procedure, the Pauli (or Pauli–Schrödinger) equation for an electron with spin 1/2 immersed in an electromagnetic field is

$$i\hbar \frac{\partial \Psi}{\partial t} = \frac{1}{2m}\left[\widehat{\boldsymbol{\sigma}} \cdot \left(\hat{\boldsymbol{p}} - \frac{e}{c}\boldsymbol{A}\right)\right]^2 \Psi + V\Psi. \tag{14.25}$$

We will now rewrite this equation in a more suggestive way. To do this, we use Eqs. (14.9) and (14.10) to write

$$\begin{aligned}\left[\widehat{\boldsymbol{\sigma}} \cdot \left(\hat{\boldsymbol{p}} - \frac{e}{c}\boldsymbol{A}\right)\right]^2 &= \sum_{ij}\left(\widehat{p}_i - \frac{e}{c}A_i\right)\left(\widehat{p}_j - \frac{e}{c}A_j\right)\widehat{\sigma}_i\widehat{\sigma}_j \\ &= \left(\hat{\boldsymbol{p}} - \frac{e}{c}\boldsymbol{A}\right)^2 + i\widehat{\boldsymbol{\sigma}} \cdot \left(\hat{\boldsymbol{p}} - \frac{e}{c}\boldsymbol{A}\right) \times \left(\hat{\boldsymbol{p}} - \frac{e}{c}\boldsymbol{A}\right).\end{aligned} \tag{14.26}$$

Since

$$\begin{aligned}\left(\hat{\boldsymbol{p}} - \frac{e}{c}\boldsymbol{A}\right) \times \left(\hat{\boldsymbol{p}} - \frac{e}{c}\boldsymbol{A}\right) &= -\frac{e}{c}(\hat{\boldsymbol{p}} \times \boldsymbol{A} + \boldsymbol{A} \times \hat{\boldsymbol{p}}) \\ &= -\frac{e}{c}(-i\hbar\nabla \times \boldsymbol{A}) = \frac{i\hbar e}{c}\boldsymbol{B}\end{aligned}$$

https://doi.org/10.1017/9781009679633.015 Published online by Cambridge University Press

(in the Coulomb gauge $\nabla \cdot \mathbf{A} = 0$), substituting in the previous expression, we obtain

$$\left[\widehat{\boldsymbol{\sigma}} \cdot \left(\hat{\boldsymbol{p}} - \frac{e}{c}\boldsymbol{A}\right)\right]^2 = \left(\hat{\boldsymbol{p}} - \frac{e}{c}\boldsymbol{A}\right)^2 - \frac{\hbar e}{c}\boldsymbol{B} \cdot \widehat{\boldsymbol{\sigma}}. \tag{14.27}$$

This result allows us to write Eq. (14.25) as

$$i\hbar\frac{\partial \Psi}{\partial t} = \frac{1}{2m}\left(\hat{\boldsymbol{p}} - \frac{e}{c}\boldsymbol{A}\right)^2\Psi + \left(V - \frac{\hbar e}{2mc}\boldsymbol{B} \cdot \widehat{\boldsymbol{\sigma}}\right)\Psi, \tag{14.28}$$

which is a more common version of the Pauli equation. We see that the postulate $\hat{\boldsymbol{P}}^2 \rightarrow (\hat{\boldsymbol{\sigma}} \cdot \hat{\boldsymbol{P}})^2$ leads naturally to the appearance of the interaction term $-\widehat{\boldsymbol{\mu}} \cdot \boldsymbol{B}$ and assigns to the electron the intrinsic magnetic moment

$$\widehat{\boldsymbol{\mu}} = -\mu_0\widehat{\boldsymbol{\sigma}}, \tag{14.29}$$

where $\mu_0 = \hbar e_0/2mc$ is the Bohr magneton. This result agrees with Eqs. (14.23) and (14.24), that is, it actually implies that the gyromagnetic ratio predicted for the electron spin contains the required factor of 2.

Expanding the Hamiltonian of (14.28) as is done, for example, in Section 12.7 and using $\widehat{\boldsymbol{J}} = \widehat{\boldsymbol{L}} + \widehat{\boldsymbol{S}}$, we can give to the Pauli equation the alternative form,

$$i\hbar\frac{\partial \Psi}{\partial t} = \widehat{H}_0\Psi + \frac{\mu_0}{\hbar}\boldsymbol{B} \cdot (\widehat{\boldsymbol{L}} + 2\widehat{\boldsymbol{S}})\Psi. \tag{14.30}$$

It is important to note that the sum $\widehat{\boldsymbol{L}} + 2\widehat{\boldsymbol{S}} = \widehat{\boldsymbol{J}} + \widehat{\boldsymbol{S}}$ appears in the Hamiltonian, that is, that $\hat{\boldsymbol{S}}$ contributes twice as much as $\widehat{\boldsymbol{L}}$ to the magnetic energy (and to the magnetic moment) because for $\widehat{\boldsymbol{L}}$, $g = 1$, but for the spin, $g = 2$. This means, in particular, that the total magnetic moment $-\mu_0(\widehat{\boldsymbol{L}}+2\widehat{\boldsymbol{S}})$ and the total angular momentum vector $\widehat{\boldsymbol{L}}+\widehat{\boldsymbol{S}}$ are not parallel.

14.3 Effects of Spin on Atomic Spectra

As mentioned in Section 14.1, it was not until the electron spin was introduced into Schrödinger's theory that the structure of atomic spectra in the presence of a magnetic field could be adequately explained. As we shall see, several other important properties of atomic spectra can only be accounted for when the spin of the electron and the spin of the proton are taken into account. Our first example is the anomalous Zeeman effect, that is, the splitting of spectral lines in paramagnetic materials with unpaired electron spins.

14.3.1 The anomalous Zeeman effect

Atoms whose total electronic spin is zero have integer angular momentum $\widehat{\boldsymbol{J}} = \widehat{\boldsymbol{L}}$ and the theory of the Zeeman effect discussed in Section 12.7 can be applied to them. The simplest example is He because it has two orbital electrons, whose spins can couple to 0 (or to 1; this topic is discussed in Section 16.3). However, if an atom has an electronic spin other than zero, the associated magnetic moment will also couple to the external magnetic field and contribute to the Zeeman effect, causing the spectral lines to split

https://doi.org/10.1017/9781009679633.015 Published online by Cambridge University Press

not into triplets, but into multiplets that can be of a more complicated structure, giving rise to the anomalous Zeeman effect.

Take an atom with Hamiltonian $\widehat{H}_0$, immersed in a weak, uniform magnetic field along the z-axis; the stationary Pauli–Schrödinger equation is

$$E\Psi = \widehat{H}_0\Psi + \frac{\mu_0 B}{\hbar}(\widehat{J}_z + \widehat{S}_z)\Psi. \tag{14.31}$$

This equation is very complicated because $\widehat{S}_z$ represents the total electron spin operator, which can be $0, \frac{1}{2}, 1, \frac{3}{2}, \ldots$ and the Ψ wave function should have the corresponding number $(2S+1)$ of components. However, we can solve this problem with the help of a theorem from the theory of tensor operators, called the Wigner–Eckart theorem, which will not be proved here (see the references), but a simple version of it will be constructed now.

Equation (14.33) is a particular case of the Wigner–Eckart theorem. We can understand the meaning of this important theorem as follows. In a system with spherical symmetry, $\widehat{\boldsymbol{J}}$ determines the only direction of reference; it follows that the expectation value $\left\langle \hat{\boldsymbol{A}} \right\rangle$ of any vector $\boldsymbol{A}$, if different from zero, is proportional to $\left\langle \hat{\boldsymbol{J}} \right\rangle$, hence $\left\langle \hat{\boldsymbol{A}} \right\rangle = C \left\langle \hat{\boldsymbol{J}} \right\rangle$, and $\left\langle \hat{\boldsymbol{A}} \cdot \hat{\boldsymbol{J}} \right\rangle = C \left\langle \hat{\boldsymbol{J}}^2 \right\rangle$. Therefore, $C = \left\langle \hat{\boldsymbol{A}} \cdot \hat{\boldsymbol{J}} \right\rangle / \left\langle \hat{\boldsymbol{J}}^2 \right\rangle$, which, when substituted into the first expression, gives

$$\langle \widehat{A}_\alpha \rangle = \langle \widehat{J}_\alpha \rangle \left[\left\langle \hat{\boldsymbol{A}} \cdot \hat{\boldsymbol{J}} \right\rangle / \left\langle \hat{\boldsymbol{J}}^2 \right\rangle \right]. \tag{14.32}$$

The angular dependence of $\langle \widehat{A}_\alpha \rangle$ is therefore determined by $\langle \widehat{J}_\alpha \rangle$ (the other factors in (14.32) are scalars). Applying this result to our case, the matrix elements of $\hat{\boldsymbol{A}}$ are given by

$$\langle j m_j' | \widehat{A}_q | j m_j \rangle = \frac{\langle j m_j | \hat{\boldsymbol{A}} \cdot \hat{\boldsymbol{J}} | j m_j \rangle}{\hbar^2 j(j+1)} \langle j m_j' | \widehat{J}_q | j m_j \rangle, \tag{14.33}$$

where $\widehat{A}_q$ and $\widehat{J}_q$ are the spherical components of the respective vectors, defined for $q = -1, 0, 1$ as

$$\widehat{A}_{\pm 1} = \tfrac{1}{\sqrt{2}}(\widehat{A}_x \pm i\widehat{A}_y), \ \widehat{A}_0 = \widehat{A}_z. \tag{14.34}$$

A comparison with (11.102) shows in particular that $\widehat{J}_1 = \widehat{J}_+$, $\widehat{J}_0 = \widehat{J}_z$, $\widehat{J}_{-1} = \widehat{J}_-$. Furthermore, the states $|jm_j\rangle$ and $|jm_j'\rangle$ may still correspond to different values n and n', respectively, of additional quantum numbers that may be required by the specific problem, but are independent of the angular description.

We now return to the problem of solving Eq. (14.31), and consider the case where B is not too large, so that we can treat the magnetic term as a perturbation. Since $\widehat{H}_0$ has spherical symmetry for an atom, we can take as unperturbed eigenfunctions the eigenstates of $\widehat{\boldsymbol{J}}^2$ and $\widehat{J}_z$, so that the first-order correction to the energy is given by

$$\delta E = \frac{\mu_0 B}{\hbar} \langle j m_j | \widehat{J}_z + \widehat{S}_z | j m_j \rangle. \tag{14.35}$$

The matrix element of $\widehat{J}_z$ is immediate,

$$\langle j m_j | \widehat{J}_z | j m_j \rangle = \hbar m_j. \tag{14.36}$$

https://doi.org/10.1017/9781009679633.015 Published online by Cambridge University Press

To compute the matrix element of $\widehat{S}_z = \widehat{S}_0$ we use (14.33), which gives

$$\begin{aligned}\langle jm_j|\widehat{S}_z|jm_j\rangle &= \frac{\langle jm_j|\hat{\boldsymbol{S}}\cdot\hat{\boldsymbol{J}}|jm_j\rangle}{\hbar^2 j(j+1)}\langle jm_j|\widehat{J}_z|jm_j\rangle \\ &= \frac{m_j}{\hbar j(j+1)}\langle jm_j|\hat{\boldsymbol{S}}\cdot\hat{\boldsymbol{J}}|jm_j\rangle. \end{aligned} \tag{14.37}$$

For the calculation of the remaining matrix element we proceed as follows. So far, the states are only required to be eigenstates of $\hat{\boldsymbol{J}}^2$ and $\widehat{J}_z$. But the required j can be obtained by combining different values of l and s, and since, by hypothesis, the atom possesses well-defined total spin, we must characterize the eigenstates with the quantum numbers j, l, s, m_j, that is, $|jm_j\rangle \to |jlsm_j\rangle$. Using this basis, the calculation is considerably simplified, since from $\widehat{\boldsymbol{J}} = \widehat{\boldsymbol{L}} + \hat{\boldsymbol{S}}$ it follows that $\hat{\boldsymbol{S}}\cdot\hat{\boldsymbol{J}} = \frac{1}{2}\left(\widehat{\boldsymbol{J}}^2 + \widehat{\boldsymbol{S}}^2 - \widehat{\boldsymbol{L}}^2\right)$, so that

$$\langle jlsm_j|\hat{\boldsymbol{S}}\cdot\hat{\boldsymbol{J}}|jlsm_j\rangle = \frac{\hbar^2}{2}[j(j+1) + s(s+1) - l(l+1)]. \tag{14.38}$$

Substituting this value into (14.37), we obtain

$$\langle jm_j|\widehat{S}_z|jm_j\rangle = \hbar\frac{j(j+1) + s(s+1) - l(l+1)}{2j(j+1)}m_j. \tag{14.39}$$

Finally, adding (14.36) and (14.39) and substituting into (14.35), we obtain the correction to the energy due to the Zeeman effect,

$$\delta E = g\mu_0 B m_j, \tag{14.40}$$

where the Landé g factor is

$$g = 1 + \frac{j(j+1) + s(s+1) - l(l+1)}{2j(j+1)}, \tag{14.41}$$

in exact agreement with the empirical expression obtained by Landé before the introduction of QM.

Since g takes different values for different atomic states (j, l, s, m_j), the spectral lines can be decomposed into very complicated multiplets. Consider the simplest case, $s = 1/2$, when there are two different eigenstates of $\widehat{\boldsymbol{L}}^2$ for a given j, namely $l = j + \frac{1}{2}$ and $l = j - \frac{1}{2}$. In both cases, Eq. (14.41) gives

$$g = \frac{j + \frac{1}{2}}{l + \frac{1}{2}}. \tag{14.42}$$

This is the case for the hydrogen atom, which we will use as an example. For $s_{1/2}$-states (the notation is L_J), it follows from (14.42) that $g = (1/2 + 1/2)/(0 + 1/2) = 2$; for $p_{1/2}$-states, $g = 2/3$; for $p_{3/2}$-states, $g = 4/3$; and so on. If g_0 is the Landé factor of the initial state and g that of the final state, the frequency of the radiation emitted in the atomic transition is

$$\omega = \omega_0 + (g_0 m_j^0 - g m_j)\,\omega_L, \tag{14.43}$$

where ω_0 is the emission frequency in the absence of the magnetic field, $\omega_L = \mu_0 B/\hbar$ is the Larmor frequency, and the quantum numbers m_j and m_j^0 are related by the selection rules $m_j - m_j^0 = 0, \pm 1$. Some transitions corresponding to the Lyman series are

https://doi.org/10.1017/9781009679633.015 Published online by Cambridge University Press

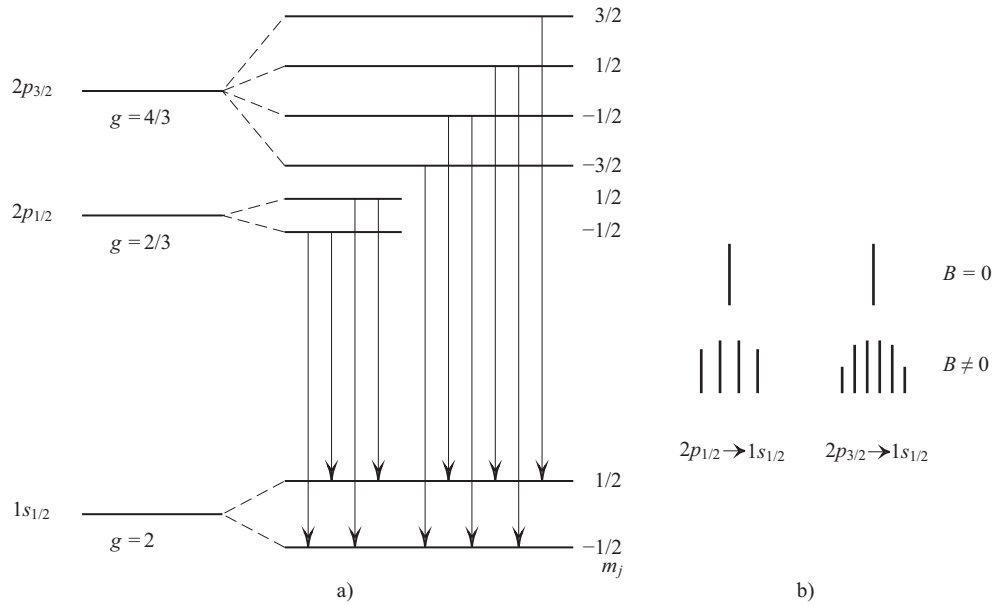


Figure 14.1 Diagram of the Zeeman spectrum showing the transitions between the lowest energy levels of the H atom. (a) The states $2p_{1/2}$ and $2p_{3/2}$ are separated by the spin–orbit coupling. (b) Line splitting for the allowed transitions $2p_{1/2} \rightarrow 1s_{1/2}$ and $2p_{3/2} \rightarrow 1s_{1/2}$.

shown in Fig. 14.1. It is clear that the complex dependence of the Landé factor on the atomic quantum numbers produces a variety of multiplets, the study of which provides information about j and l for the states involved in each transition.

When the total electron spin is zero (as in the case of He), the Landé factor reduces to $g = 1$; Eq. (14.40) reduces to (12.135) for the normal Zeeman effect, and the spectral lines split into the triplets predicted by the (classical) Lorentz theory. The electron spin is thus identified as the cause of the failure of the classical description of the Zeeman effect for atoms with an odd number of valence electrons.

Exercise E14.1. Atomic diamagnetism and paramagnetism

Investigate the diamagnetism and the paramagnetism of atoms immersed in a constant, homogeneous magnetic field.

Solution. According to Eq. (14.40) and up to linear terms in $\boldsymbol{A}$, the magnetic interaction energy is

$$\delta E_1 = \mu_0 B g m_j. \tag{14.44}$$

On the other hand, if in the Hamiltonian of Eq. (14.28) the quadratic terms in $\boldsymbol{A}$ are preserved, a second correction is obtained, which for the case of a uniform external magnetic field in the z-direction is

$$\delta E_2 = \frac{e^2}{2mc^2}\langle \boldsymbol{A}^2 \rangle = \frac{e^2}{8mc^2}\langle x^2 + y^2 \rangle B^2 = \frac{e^2 B^2}{12mc^2}\langle r^2 \rangle \tag{14.45}$$

https://doi.org/10.1017/9781009679633.015 Published online by Cambridge University Press

(it was assumed that in the absence of $\boldsymbol{A}$, the system possesses spherical symmetry, so $\langle x^2+y^2\rangle = \frac{2}{3}\langle r^2\rangle$). Therefore, the magnetic energy of the interaction between the atom and the field is

$$\Delta E_{\text{mag}} = \mu_0 g B m_j + \frac{e^2B^2}{12mc^2}\langle r^2\rangle. \tag{14.46}$$

From this result it follows that the magnetic moment of the atom in the magnetic field is given by

$$\mu \equiv -\frac{\partial \Delta E_{\text{mag}}}{\partial B} = -\mu_0 g m_j - \frac{e^2 B}{6mc^2}\langle r^2\rangle. \tag{14.47}$$

The first term represents the intrinsic magnetic moment of the atom and characterizes its paramagnetic properties, while the second term, proportional to B, represents the induced magnetic moment and characterizes the diamagnetic properties of the atom ($\mu_{\text{mag}} < 0$). If N is Avogadro's number, the diamagnetic susceptibility per mole is

$$\chi_{\text{diam}} \equiv N\frac{\partial \mu_{\text{diam}}}{\partial B} = -\frac{e^2 N}{6mc^2}\langle r^2\rangle. \tag{14.48}$$

This term is always nonzero and negative, which means that all atoms have diamagnetic properties (i.e., they react against the external field). Equation (14.48) is called the Langevin formula. This formula allows us to determine $\langle r^2\rangle$ from the experimental value of χ_{diam}; the values obtained coincide with those provided by the kinetic theory of gases for atomic radii.

With regard to the paramagnetic moment, it should be noted that, since it is proportional to m_j, it can have either sign; however, in a state of thermodynamic equilibrium (characterized by the Boltzmann distribution) negative values of m_j prevail, since they correspond to lower energy states and the average over m_j is not zero. This leads to an average paramagnetic susceptibility (which tends to strengthen the external field) given by the Curie–Langevin formula:

$$\chi_{\text{param}} = \frac{\mu_0 g^2}{3kT} j(j+1). \tag{14.49}$$

If $j \neq 0$, χ_{param} is much larger than $|\chi_{\text{diam}}|$ and the atom has paramagnetic properties (e.g., hydrogen is always paramagnetic, since $j_{\min} = 1/2$). However, atoms with paired electron spins show diamagnetic properties (e.g., He in its ground state, or, more generally, the noble gases).

In a paramagnetic atom, the dipole moments tend to align in the direction of the field, inducing a field that strengthens the applied one. However, the alignment tends to be destroyed by thermal atomic motion, so the susceptibility should decrease with temperature; the Curie–Langevin formula (which is only valid at relatively low temperatures) shows that this is the case. A more accurate calculation predicts that the magnetization reaches a temperature-independent saturation value, corresponding to the case where all unpaired atomic magnetic dipoles are aligned with the external field; it is the intensity of the saturation field that increases with temperature. Exceptionally paramagnetic materials that retain their acquired magnetization when the external field is removed (Fe, Co, etc.) are called ferromagnetic materials.

https://doi.org/10.1017/9781009679633.015 Published online by Cambridge University Press

14.3.2 *Fine and hyperfine structure of the H spectrum

The spin of the orbital electrons has a small but nonnegligible effect on the atomic spectrum, even in the absence of an external magnetic field. This is due to the fact that the electron in its orbital motion generates a magnetic field, to which the intrinsic magnetic moment of the electron is coupled. The resulting interaction term is commonly called the *spin–orbit interaction*. To this must be added a relativistic correction to the kinetic energy, which is of the same order of magnitude. The corresponding formulas are more conveniently derived in detail from the Dirac equation (see Chapter 19).

In general, the predictions of these formulas are in excellent agreement with the experimental data on the so-called fine spectrum. However, with the refinement of experimental techniques, discrepancies appeared on a smaller scale. In particular, in 1947 W. Lamb[8] together with Robert Retherford (1912–1981) observed a small separation between the $2s_{1/2}$ and $2p_{1/2}$ levels of the H atom, which is about one-tenth of the separation between the $2p_{1/2}$ and $2p_{3/2}$ levels. This Lamb shift can be explained as essentially due to the radiative effects of the electron, which were not taken into account in the previous theory or in Dirac's, but are fully accounted for in QED and in SED.[9]

There are even finer corrections to the atomic spectrum, called hyperfine corrections, which have various causes. The most important are related to the nucleus. We can think first of the effect on the atomic energy levels due to the finite size of the nucleus, which modifies the Coulomb potential in the nuclear region; but this correction, even for s-states, is at most of the order of 0.1% of the Lamb shift; see Eq. (14.111). A more important nuclear effect (in addition to the mass correction already made) comes from the magnetic moment associated with the spin of the nuclear proton, which generates a magnetic field to which the magnetic moment of the electron couples. This correction is of the order of $m/M \sim 10^{-3}$ times that due to the fine effects (relativistic and spinorial) and is even smaller than the Lamb effect (ca. 20%); it is therefore called a hyperfine correction.

Transitions in H atoms between the slightly excited state $1s_{1/2}$ with total spin $S = 1$ and the ground state $1s_{1/2}$ with $S = 0$ produce radiation of frequency $\omega \sim 1420$ MHz, of the order of the Lamb shift. This radiation, with a wavelength of 21.4 cm, has played an important role in the study of the distribution of H in intergalactic space. The reason for this is that the state with $S = 1$ is not produced by electromagnetic transitions (the radiative transition $S = 0$ to $S = 1$ is forbidden by the selection rule $\Delta l = 1$), but by collisions; therefore, the astronomical study of the 21 cm radiation produced when this state of H is de-excited allows us to know the density and the average temperature of the H clouds in interstellar space.

[8] Willis Eugene Lamb (1913–2008) was a US physicist who won the Nobel Prize in Physics in 1955 for his discoveries concerning the fine structure of the hydrogen spectrum. In the latter part of his career he devoted increasing attention to the field of quantum measurement. Lamb was openly critical of many of the trends in QM interpretation, and in particular of the use of the term *photon*.

[9] In Chapter 17 we will look at radiative corrections, including the Lamb shift.

https://doi.org/10.1017/9781009679633.015 Published online by Cambridge University Press

Exercise E14.2. Effect of spin–spin interaction between nucleons
Consider the interaction between two equal nucleons in the CM system and determine the simplest form in which the central spin–spin interaction potential can be expressed.

Solution. The system is conveniently described in terms of the relative variables $\boldsymbol{r} = \boldsymbol{r}_1 - \boldsymbol{r}_2$, $\boldsymbol{p} = \frac{1}{2}(\boldsymbol{p}_1 - \boldsymbol{p}_2)$. The interaction potential between the nucleons can be modeled as the sum of a scalar contribution $V_0(r)$, a spin–spin coupling term $V_{ss}\left(\hat{\sigma}_1 \cdot \hat{\sigma}_2\right)$, plus additional spin-interaction terms. We shall calculate the contribution of the first two terms to the energy.

From $4\hat{\boldsymbol{S}}^2 = \hat{\sigma}_1^2 + \hat{\sigma}_2^2 + 2\hat{\sigma}_1 \cdot \hat{\sigma}_2$ it follows that $\hat{\sigma}_1 \cdot \hat{\sigma}_2 = 2\widehat{\mathbf{S}}^2 - 3$, so that for a potential of the form

$$V = V_0(r) + V_{ss}\left(\hat{\sigma}_1 \cdot \hat{\sigma}_2\right) \tag{14.50}$$

the Hamiltonian commutes with $\hat{\boldsymbol{L}}$ and $\hat{\boldsymbol{S}}$ and we will have eigenfunctions of $\hat{\boldsymbol{L}}^2, \hat{L}_3, \hat{\boldsymbol{S}}^2, \widehat{S}_3$, The total spin can take any of the values $S = 0, 1$, giving respectively

for $S = 0$, $\quad \hat{\sigma}_1 \cdot \hat{\sigma}_2 = -3$, $\quad V = V_0 - 3V_{ss}$;
for $S = 1$, $\quad \hat{\sigma}_1 \cdot \hat{\sigma}_2 = +1$, $\quad V = V_0 + V_{ss}$.

The problem has been reduced to a Schrödinger problem for a spinless particle in a central potential whose value depends on the total spin.

14.3.3 Magnetic resonance

Atomic electrons or nuclei with nonzero spin can absorb or emit electromagnetic radiation in response to magnetic fields. This is because the spin magnetic moment precesses at a frequency proportional to the field that produces the Zeeman splitting (the Larmor frequency $\omega_L = \mu_0 B/\hbar$, in the case of the electron), and it is therefore possible to induce transitions between spin states via resonance coupling to an oscillating electromagnetic field related to this frequency.

The first induced transitions between nuclear Zeeman levels were obtained by I. Rabi[10] in 1933 using radiofrequency waves in an atomic beam. In 1945 the Swiss physicist Felix Bloch (1905–1983) and the US physicist Edward Purcell (1912–1997) independently performed the first nuclear magnetic resonance (NMR) experiments in condensed matter, directly detecting the response of the H proton to the electromagnetic field. Today, NMR has a wide range of applications in physics, chemistry, medicine, geology, and other fields.

In electron paramagnetic resonance (EPR) – first observed by the Soviet physicist Yuri K. Zavoysky in 1944 – a material composed of atoms or molecules carrying an unpaired electron is placed in a magnetic field that polarizes the material's bulk magnetic moment. When the material is exposed to the radiofrequency of the field, the

[10] Isidor Isaac Rabi (1898–1988) was a US physicist who won the Nobel Prize in Physics in 1944 for his discovery of nuclear magnetic resonance. He was also one of the first scientists in the United States to work on the cavity magnetron, which is used in microwave radar and microwave ovens (more generally, to generate microwaves).

https://doi.org/10.1017/9781009679633.015 Published online by Cambridge University Press

spin is "flipped" by the transfer of energy from the field to the sample. EPR has made it possible to study phenomena such as the defects that give certain crystals their color, the formation of free radicals in bulk materials, and the behavior of free or conduction electrons in metals.

To simplify the analysis we consider the Hamiltonian for an electron spin in the presence of a magnetic field $\boldsymbol{B}$ given by Eq. (14.24), where $\boldsymbol{B}$ consists of two terms: the static field B_0 along the z-axis, and the radiofrequency field $B_1(t)$ along the x-axis, that is,

$$\widehat{H} = -\widehat{\boldsymbol{\mu}} \cdot (B_0 \hat{\mathbf{k}} + B_1 \hat{\mathbf{i}} \cos \omega t). \tag{14.51}$$

From Eq. (14.23) we have $\widehat{\boldsymbol{\mu}} = -\mu_0 \widehat{\boldsymbol{\sigma}}$, where the components of $\widehat{\boldsymbol{\sigma}}$ are the 2×2 Pauli matrices σ_i. Therefore, in matrix form $\hat{H}$ becomes

$$\widehat{H} = \hbar \begin{pmatrix} \omega_L & \Omega \cos \omega t \\ \Omega \cos \omega t & -\omega_L \end{pmatrix}, \tag{14.52}$$

where

$$\omega_L = \mu_0 B_0 / \hbar, \quad \Omega = \mu_0 B_1 / \hbar. \tag{14.53}$$

The response to the oscillatory term is most conveniently described by splitting the Hamiltonian into zero- and first-order terms in Ω,

$$\widehat{H} = \widehat{H}_0 + \widehat{H}_1 = \hbar \omega_L \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + \hbar \Omega \cos \omega t \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \tag{14.54}$$

and writing the first-order term in the interaction representation studied in Section 8.6.6,

$$\widehat{H}_{1I} = e^{i \hat{H}_0 t/\hbar} \widehat{H}_1 e^{-i \hat{H}_0 t/\hbar} = \hbar \Omega \cos \omega t \left(e^{i \omega_L t \hat{\sigma}_z} \widehat{\sigma}_x e^{-i \omega_L t \hat{\sigma}_z} \right). \tag{14.55}$$

After some matrix algebra, we get

$$e^{i \omega_L t \hat{\sigma}_z} \widehat{\sigma}_x e^{-i \omega_L t \hat{\sigma}_z} = \begin{pmatrix} 0 & e^{2i\omega_L t} \\ e^{-2i\omega_L t} & 0 \end{pmatrix}, \tag{14.56}$$

which introduced into (14.55) gives, using $\cos \omega t = \frac{1}{2} \left(e^{i\omega t} + e^{-i\omega t} \right)$ and neglecting the rapidly oscillating terms $e^{\pm i(\omega + 2\omega_L)t}$,

$$\widehat{H}_{1I} \simeq \frac{\hbar \Omega}{2} \begin{pmatrix} 0 & e^{-i(\omega - 2\omega_L)t} \\ e^{i(\omega - 2\omega_L)t} & 0 \end{pmatrix}. \tag{14.57}$$

This approximation is called the rotating-wave approximation, RWA. It is a very good approximation when the frequency ω of the applied field is close to $2\omega_L$, because then the system responds resonantly to the field, while the rapidly oscillating terms average out. As an effect of the resonant response, the field periodically transfers the population between the two Zeeman levels.

The name of the approximation derives precisely from the form of the Hamiltonian in the interaction picture, where only the part of the electromagnetic wave that approximately co-rotates is retained and the counter-rotating component is discarded. The RWA is widely used in magnetic resonance studies and in quantum dynamics of two-level systems more generally.[11]

[11] See Section 17.3.2 for a more detailed treatment of the RWA.

https://doi.org/10.1017/9781009679633.015 Published online by Cambridge University Press

14.4 Modeling the Electron Spin

In the quantum literature it is common to refrain from assigning a specific model to the electron spin. This is because the electron is considered to be a point particle and it is not clear how a magnetic moment or an intrinsic rotation can be associated with a (mathematical) point. Nevertheless, there are various models that try to give it a specific physical meaning. The following is an example that is in line with the principles of SED, based on the experimental observation that electrons interact with circularly polarized modes of the radiation field.

The circularly polarized modes of the ZPF have, in addition to an energy $\frac{1}{2}\hbar\omega$ and a linear momentum $\frac{1}{2}\hbar\boldsymbol{k}$, an "intrinsic" angular momentum given by

$$J = \pm\frac{\hbar}{2}\hat{\mathbf{k}} \tag{14.58}$$

along the direction of propagation $\hat{\mathbf{k}}$. The permanent interaction of the electron with the ZPF impresses on it an "internal" rotation, which can be regarded as a nonrelativistic version of the *zitterbewegung* discussed in Section 19.3 in connection with Dirac's equation.

To calculate the angular momentum of this rotation, we use the evolution equation for the average value of a dynamical function $\mathcal{G}(\boldsymbol{x}, \boldsymbol{p})$ derived in Section 10.3, Eq. (10.86), and apply it to the Cartesian components $l_k = \epsilon_{ijk} x_i p_j$. This leads to the evolution equation for the average angular momentum (with the external torque given by $\boldsymbol{M} = \boldsymbol{x} \times \boldsymbol{f}$),

$$\frac{d}{dt}\langle \boldsymbol{L}\rangle = \langle \boldsymbol{M}\rangle + m\tau\,\langle \boldsymbol{x}\times\dddot{\boldsymbol{x}}\rangle - e^2\,\langle \boldsymbol{x}\times\mathcal{D}\rangle\,, \tag{14.59}$$

which in the quantum regime (i.e., in the Markovian limit) must be interpreted in terms of expectation values. For every Cartesian component we get

$$\frac{d}{dt}\left\langle \hat{L}_{ij}\right\rangle = \left\langle \hat{M}_{ij}\right\rangle + m\tau\left\langle \hat{x}_i\dddot{\hat{x}}_j - \hat{x}_j\dddot{\hat{x}}_i\right\rangle - \left\langle \hat{D}^{px}_{ij} - \hat{D}^{px}_{ji}\right\rangle. \tag{14.60}$$

This equation reveals the mechanism that leads to a mean effective angular momentum through the antisymmetric part of the diffusion coefficient $\hat{D}^{px}$. In the absence of an external torque, the angular momentum is conserved when the following balance equation between the two radiative terms is satisfied,

$$m\tau\left\langle \hat{x}_i\dddot{\hat{x}}_j - \hat{x}_j\dddot{\hat{x}}_i\right\rangle = \left\langle D^{px}_{ij} - D^{px}_{ji}\right\rangle. \tag{14.61}$$

For simplicity, we consider an isotropic HO with frequency ω in its ground state. Then $m\dddot{x}_j = -\omega^2 p_j$, and from (14.61) we have

$$\langle L_{ij}\rangle = \langle x_i p_j - x_j p_i\rangle = \frac{-1}{\tau\omega^2}\left\langle D^{px}_{ij} - D^{px}_{ji}\right\rangle. \tag{14.62}$$

The usual numerical value $\langle L_{ij}\rangle = 0$ for the ground state represents the average net angular momentum induced by the full ZPF. However, since the electron responds separately to modes of left or right polarization, we should consider the separate effect of these modes on the angular momentum.

https://doi.org/10.1017/9781009679633.015 Published online by Cambridge University Press

The circular polarizations are characterized by the vectors

$$\boldsymbol{\epsilon}_{\boldsymbol{k}\pm} = \frac{1}{\sqrt{2}}\left(\boldsymbol{\epsilon}_{\boldsymbol{k}i} \pm i\boldsymbol{\epsilon}_{\boldsymbol{k}j}\right), \tag{14.63}$$

with $\boldsymbol{\epsilon}_{\boldsymbol{k}i}, \boldsymbol{\epsilon}_{\boldsymbol{k}j}$ unit Cartesian vectors orthogonal to $\mathbf{k}$. The appropriate variables to describe the response of the particle to the (circularly) polarized field modes are the spherical variables (x^+, x^-, x_k), which are given by

$$x^{\pm} = \tfrac{1}{\sqrt{2}}\left(x_i \mp i x_j\right), \tag{14.64}$$

$$x_i = \tfrac{1}{\sqrt{2}}\left(x^+ + x^-\right), \quad x_j = i\tfrac{1}{\sqrt{2}}\left(x^+ - x^-\right). \tag{14.65}$$

The nonzero matrix elements of the oscillator are $x_{i01} = \left(x_{01}^+ + x_{01}^-\right)/\sqrt{2}$; and so on. Further, since $x_{10}^{\pm} = \left(x_{01}^{\mp}\right)^*$, Eq. ((14.62)) gives

$$\left\langle L_{ij}\right\rangle_0 = m\omega_0\left(x_{01}^+ x_{10}^- - x_{01}^- x_{10}^+\right) = m\omega_0\left(\left|x_{01}^+\right|^2 - \left|x_{01}^-\right|^2\right). \tag{14.66}$$

In the ground state, $\left\langle L_{ij}\right\rangle_0 = 0$; hence the two terms, $\left|x_{01}^+\right|^2$ and $\left|x_{01}^-\right|^2$, contribute with equal magnitude and opposite sign to the k-component of the oscillator's angular momentum, as should be the case for a nonpolarized vacuum. These separate contributions are

$$\left\langle L_{ij}\right\rangle_0^+ = m\omega_0\left|x_{01}^+\right|^2, \quad \left\langle L_{ij}\right\rangle_0^- = -m\omega_0\left|x_{01}^-\right|^2. \tag{14.67}$$

Using $x_{i01} = x_{j01} = \sqrt{\hbar/(2m\omega_0)}$ for the harmonic oscillator, we get

$$m\omega_0\left|x_{01}^{\pm}\right|^2 = \frac{\hbar}{2}, \tag{14.68}$$

whence the size of each separate contribution to the angular momentum in (14.66) is just $\hbar/2$. In order to distinguish this contribution from the orbital component of the angular momentum we write $\left\langle S_{ij}\right\rangle^{\pm}$ instead of $\left\langle L_{ij}\right\rangle_0^{\pm}$, so that

$$\left\langle S_{ij}\right\rangle^{\pm} = \pm\frac{\hbar}{2}. \tag{14.69}$$

Furthermore, by applying Eq. (10.86) to the square of the angular momentum, one obtains the correct quantum value for $\left\langle \boldsymbol{L}^2\right\rangle_0$, independent of the position coordinates and consisting of two terms due to each of the polarized ZPF modes (see the bibliographic notes at the end of the chapter),

$$\left\langle \boldsymbol{S}^2\right\rangle^+ = \left\langle \boldsymbol{S}^2\right\rangle^- = \frac{3}{4}\hbar^2. \tag{14.70}$$

The fact that the results for $\left\langle S_{ij}\right\rangle$ and $\left\langle \boldsymbol{S}^2\right\rangle$ do not depend on the oscillator's frequency ω_0, suggests that they hold in a more general case, and in particular for the free particle. We therefore conclude that this angular momentum induced by the ZPF can be identified with the spin of the electron. The term "intrinsic," which is usually attached to it, indicates its permanence, since the ZPF is always and everywhere present.

https://doi.org/10.1017/9781009679633.015 Published online by Cambridge University Press

14.5 Quantum Correlations and Apparent Nonlocality

Nonlocality is in conflict with general physical principles. However, it is a fact that the quantum-mechanical *description* has nonlocal features. We have discussed one aspect of this issue in Chapter 5 in connection with quantum tunneling, and in Chapter 10 in connection with the quantum potential. Here we deal with another manifestation of nonlocality, specifically related to the (apparent) nonlocal (and noncausal) influence between objects – what Einstein called "spooky action at a distance." Such nonlocality manifests itself characteristically in the form of (quantum) correlations between dynamical variables of composite systems. Taken at face value, the Pauli principle, which prohibits two noninteracting electrons from occupying the same quantum state, is a reflection of such quantum nonlocality. Where do these nonlocalities come from, what is the underlying physics?

Long before the advent of QM, the concept of field was introduced by Faraday[12] to explain the apparent "action at a distance" of noncontact forces affecting the dynamics of particles. In classical physics, the neglect of the field (whether gravitational, electric, magnetic or other) as a mediator between the particles leads to an apparent nonlocality (taken by an "effective" one) resulting from the incomplete consideration of the system. It is therefore reasonable to assume that also in quantum physics, the nonlocality of the description may be the result of an incompleteness. In other words, that an explanation for the (apparently) nonlocal behavior of quantum systems can be found by taking into account the presence of a field that acts as a mediator between the particles.

This is precisely the idea behind the explanation of quantum nonlocality provided by SED, with the introduction of the radiation field, including the ZPF, as the mediator. The phenomenon of quantum correlation, which is conventionally regarded as a signature of nonlocality, is a particularly useful example to illustrate this role of the field. By being connected to the separate (not directly interacting) particles that make up a system, the field can act as a mediator and lead to causal correlations between them – which can persist as long as the system is not perturbed by external factors that can cause decoherence (see Section 15.5.7).

For the sake of clarity, the discussion will focus on the specific case of a system composed of two identical particles, that is, particles that, by virtue of having the same physical properties, exhibit the same behavior under the same circumstances. The results obtained for the bipartite case can then be extended to the analysis of a multipartite system of identical particles.

[12] Michael Faraday (1791–1867) was a most eminent English scientist dedicated to the study and development of electromagnetism and electrochemistry. He is one of the most brilliant and prolific scientists in history, yet he never went to school. In his laboratory he constructed (invented) the first electric motor, the first electric generator, and the first induction transformer. He introduced the notion of (electromagnetic) field in order to eliminate the appeal to actions at a distance from science. Maxwell turned this proposal into the most transcendental principle of science: nature is made up of matter and fields.

https://doi.org/10.1017/9781009679633.015 Published online by Cambridge University Press

14.5.1 Bipartite system correlated by the background field

As a background to the study of the correlations between two particles, we recall how a single quantum particle responds to the background field, according to SED. In Section 10.4 we have seen that in the quantum regime, the kinematics of the particle is determined by the field modes with which the particle is in interaction, which by default are those of the ZPF; specifically, according to (10.119),

$$x_n(t) = x_{nn} + \sum_{k\neq n} x_{nk}a_{nk}e^{-i\omega_{kn}t} + \text{c.c.}, \qquad p_n(t) = p_{nn} + \sum_{k\neq n} p_{nk}a_{nk}e^{-i\omega_{kn}t} + \text{c.c.}, \tag{14.71}$$

for any state n, with x_{nk} and p_{nk} the elements of matrices $\hat{x}$ and $\hat{p}$, respectively, satisfying the canonical commutator

$$\left[\widehat{x}, \widehat{p}\right] = i\hbar. \tag{14.72}$$

Let us now consider a system consisting of two equal, noninteracting particles, subject to a common binding potential. The fact that they are part of one and the same system means that they are also subject to the same background field. Furthermore, since the particles are identical, they satisfy the same equation of motion and therefore *they respond to common modes of the field*.

Both particles together have an energy $\mathcal{E}$ which is the sum of the individual energies $\mathcal{E}_n$ and $\mathcal{E}_m$ corresponding to states n and m. The total energy level is degenerate, since $\mathcal{E}_n$ may be the energy of particle 1 and $\mathcal{E}_m$ that of particle 2, or vice versa. To distinguish between the two possibilities we name them A and B, and write, for $m \neq n$ and $\mathcal{E}_n \neq \mathcal{E}_m$,

$$\mathcal{E}_A = \mathcal{E}_n + \mathcal{E}_m,\ \mathcal{E}_B = \mathcal{E}_m + \mathcal{E}_n,\ \mathcal{E}_A = \mathcal{E}_B = \mathcal{E}, \tag{14.73}$$

with

$$\mathcal{E}_n - \mathcal{E}_m = \hbar\omega_{nm}. \tag{14.74}$$

In Eq. (14.73) and the following expressions, the first term refers to particle 1 and the second term to particle 2; for the sake of simplicity in writing, the labels are omitted whenever this does not lead to confusion.

Our aim is to calculate the correlation between the variables $x(t)$ of the two particles. The product corresponding to state A is, according to Eq. (14.71),

$$(x_1x_2)_A(t,\phi) = \sum_{k,l} x_{nk}a_{nk}e^{-i\omega_{kn}t}x_{ml}e^{i\gamma_{ml}\phi}a_{ml}e^{-i\omega_{lm}t}, \tag{14.75}$$

where $\gamma_{ml}\phi$ is a phase difference between the coupling of particle 2 and that of particle 1 to the respective field mode, the value of which will be given in due course. Similarly, the product corresponding to state B is

$$(x_1x_2)_B(t,\phi) = \sum_{k,l} x_{mk}a_{mk}e^{-i\omega_{km}t}x_{nl}e^{i\gamma_{nl}\phi}a_{nl}e^{-i\omega_{ln}t}. \tag{14.76}$$

https://doi.org/10.1017/9781009679633.015 Published online by Cambridge University Press

In Eq. (14.75), all products $a_{nk}a_{ml}$ average to zero except two: (i) when $k = n$ and $l = m$, and (ii) when $l = n$ and $k = m$, because $\langle a_{nk}a_{kn}\rangle = \langle a_{nk}a_{nk}^*\rangle = 1$ and hence also $a_{nn} = 1$. Therefore, the average contains only two terms,

$$\langle (x_1x_2)_A\rangle = x_{nn}x_{mm} + x_{nm}x_{mn}e^{i\gamma_{mn}\phi}. \tag{14.77}$$

Similarly, the only products $a_{mk}a_{nl}$ in Eq. (14.76) that do not average to zero are: (i) when $k = m$ and $l = n$, and (ii) when $l = m$ and $k = n$, whence

$$\langle (x_1x_2)_B\rangle = x_{mm}x_{nn} + x_{mn}x_{nm}e^{i\gamma_{nm}\phi}. \tag{14.78}$$

In the terms of the state vectors in the product Hilbert space,

$$|A\rangle = |n\rangle\,|m\rangle\,,\ |B\rangle = |m\rangle\,|n\rangle\,, \tag{14.79}$$

where the left ket refers to particle 1 and the right ket to particle 2, Eqs. (14.77) and (14.78) have the form

$$\langle (x_1x_2)_A\rangle = \langle A|\,\hat{x}_1\hat{x}_2\,|A\rangle + e^{i\gamma_{mn}\phi}\,\langle A|\,\hat{x}_1\hat{x}_2\,|B\rangle\,, \tag{14.80}$$

$$\langle (x_1x_2)_B\rangle = \langle B|\,\hat{x}_1\hat{x}_2\,|B\rangle + e^{-i\gamma_{mn}\phi}\,\langle B|\,\hat{x}_1\hat{x}_2\,|A\rangle\,, \tag{14.81}$$

or regrouping,

$$\langle (x_1x_2)_A\rangle = \langle A|\,\hat{x}_1\hat{x}_2\left|\left(A + e^{i\gamma_{mn}\phi}B\right)\right\rangle, \tag{14.82}$$

$$\langle (x_1x_2)_B\rangle = e^{-i\gamma_{mn}\phi}\,\langle B|\,\hat{x}_1\hat{x}_2\left|\left(A + e^{i\gamma_{mn}\phi}B\right)\right\rangle. \tag{14.83}$$

Combining the two expressions and taking into account the invariance of the correlation with respect to the exchange of A and B, we obtain

$$\begin{aligned}\langle (x_1x_2)_{AB}\rangle &= \frac{1}{2}\left[\langle (x_1x_2)_A\rangle + \langle (x_1x_2)_B\rangle\right]\\ &= \langle \Psi_{AB}|\,\hat{x}_1\hat{x}_2\,|\Psi_{AB}\rangle\,,\end{aligned} \tag{14.84}$$

where

$$|\Psi_{AB}\rangle = \frac{1}{\sqrt{2}}\left|A + e^{i\gamma_{mn}\phi}B\right\rangle \tag{14.85}$$

is the state vector of the bipartite system with energy $\mathcal{E}_A = \mathcal{E}_B = \mathcal{E}$. Furthermore, since the exchange of A and B can at most change the sign of $|\Psi_{AB}\rangle$, the phase factor must be such that $e^{i\gamma_{mn}\phi} = \pm 1$, that is,

$$\gamma_{mn}\phi = n\pi, \quad n = 0, 1, 2, \ldots \tag{14.86}$$

which means that $|\Psi_{AB}\rangle$ is *either symmetric or antisymmetric* with respect to the exchange of A and B, depending on whether n is even or odd,

$$\left|\Psi_{AB}^{\pm}\right\rangle = \frac{1}{\sqrt{2}}\,|A \pm B\rangle\,. \tag{14.87}$$

In Chapter 15 it will be shown that the even or odd parity of $|\Psi_{AB}\rangle$ is crucial in determining the statistical properties of systems of identical particles.

According to Eq. (14.84), by taking the average of the product of functions $x_1(t)$ and $x_2(t)$, we have obtained the quantum expectation value of the product $\hat{x}_1\hat{x}_2$ calculated

https://doi.org/10.1017/9781009679633.015 Published online by Cambridge University Press

in the state described by (14.87). With respect to the individual state vectors, this is a *nonfactorizable* vector. The cause of the nonfactorizability can be traced back to the second term on the right-hand side of Eqs. (14.80) and (14.81), which is due to the common field mode (nm), the only mode to which both particles respond coherently. In other words, it is a consequence of their being coupled to the same field mode (nm), *either in phase or in antiphase*.

Note that the Hilbert space to which $|\Psi_{AB}\rangle$ belongs is a new Hilbert space formed by the tensor product of the individual Hilbert spaces spanned by the sets of state vectors belonging to particles 1 and 2. This new Hilbert space is spanned by the set $\{|\Psi_{AB}\rangle\}$ of the two (orthonormal) eigenvectors corresponding to the total energy $\mathcal{E}_{AB}$, given by Eq. (14.87).

Thus, in addition to elucidating the physical mechanism that correlates the particles, SED explains the emergence of entangled state vectors. The procedure used here to calculate the correlation of x_1 and x_2 can be logically extended to other products of dynamical variables, as required.

14.5.2 Entanglement

A quantum system consisting of two or more particles can be in an entangled state, which has the property that it cannot be written down as a single product of the states of the individual particles. Entanglement – which Schrödinger originally called *Verschränkung* (interlacing) – involves a relationship between the components of the system, expressed by the appearance of characteristic correlations between their variables, as in the example of the previous section. This property has significant potential applications in quantum cryptography, quantum computing, particle interferometry, and teleportation, and has given rise to a whole new multidisciplinary branch called *quantum information*, which deals with information subject to the laws of QM. The current literature on this subject is growing rapidly, not least because entanglement has opened up the possibility of dense coding for calculation and remote transport of quantum information.

As these topics constitute a vast and promising field that is very active today, it is convenient to extend our introductory discussion of the bipartite entangled state a little.

The elementary quantum information unit is the qubit, which is an arbitrary (single-particle) unit vector in a 2D Hilbert space, $|\chi\rangle = a\,|+\rangle + b\,|-\rangle = a\binom{1}{0} + b\binom{0}{1} = \binom{a}{b}$, with $|a|^2 + |b|^2 = 1$. If we identify the one-particle eigenstates with $|n\rangle = \binom{1}{0}$ and $|m\rangle = \binom{0}{1}$, as in the previous case, the unit vector is $|\chi\rangle = a\,|n\rangle + b\,|m\rangle$, and the possible bipartite entangled eigenstates are, according to Eq. (14.87),

$$|\Psi_{\pm}\rangle = \frac{1}{\sqrt{2}}\left(|n\rangle\,|m\rangle \pm |m\rangle\,|n\rangle\right). \tag{14.88}$$

It is important to note that when the particles move independently, their individual states $|n\rangle, |m\rangle$ are well defined and can be combined to form the product states $|n\rangle\,|m\rangle$ and $|m\rangle\,|n\rangle$. It is only when both particles form a *single* system in a state characterized by a well-defined eigenvalue – $\mathcal{E} = \mathcal{E}_n + \mathcal{E}_m$ with $\mathcal{E}_n \neq \mathcal{E}_m$, in the preceding example – that they are entangled.

https://doi.org/10.1017/9781009679633.015 Published online by Cambridge University Press

The two-state system of a spin-1/2 particle is a particularly simple case of a qubit. For this reason, the terminology of spin-1/2 particles is often used in quantum information; but note that there are other physical realizations of two-state systems, such as *quantum dot systems*, *superconducting quantum interference devices* (SQUID), and so on.

In Section 11.4 we introduced the (entangled, with total spin $S = 0$) singlet state formed by two spin-1/2 particles, Eq. (11.182),

$$|00\rangle = \frac{1}{\sqrt{2}}\Big(|+\rangle\,|-\rangle - |-\rangle\,|+\rangle\Big). \tag{14.89}$$

Such states have been widely used to exhibit peculiar quantum properties ever since they were discussed by Einstein, Podolsky, and Rosen in their famous 1935 EPR paper and by Schrödinger in the series of papers in which he introduced the equally famous cat that bears his name. The EPR paper challenged the notion – popular at the time and still held to by many today – of QM as a complete theory, in the (particular) sense that every element of physical reality has a counterpart in the theory; see Section 14.5.4. In the same year, at Einstein's suggestion, Schrödinger elaborated on entanglement and proposed the cat paradox, which is discussed in the following section.

As mentioned earlier, entangled states can involve more than two particles forming a composite system, resulting in different degrees of entanglement. A general n-particle entangled state occurs when each product state $|a\rangle\,|b\rangle\cdots|z\rangle$ satisfies the conditions specified by the problem as a whole, for example, by representing an eigenstate of a given total spin or total energy,

$$|\Psi\rangle = \sum_{(a,b,\ldots z)} C_{ab\ldots z}\,|a\rangle_1\,|b\rangle_2\cdots|z\rangle_n\,. \tag{14.90}$$

We see that unless all $C_{ab\ldots z}$ except one are zero, it is not possible to assign a (unique) state to the individual components; one then speaks of nonfactorizable states. Nonfactorizable states invariably imply nonclassical correlations.

14.5.3 Schrödinger's cat paradox

To expose the peculiar nature of nonfactorizable wave functions, Schrödinger took the quantum description to its extreme and applied it to a macroscopic system, for which it becomes manifestly paradoxical. Schrödinger concluded that it is in the entangled states that the definitive separation between the behavior of classical and quantum systems becomes apparent and the deepest conceptual problems arise, which is why these states must be studied exhaustively to unravel their mysteries.

The system analyzed by Schrödinger consists of a radioactive nucleus and a cat, all enclosed in a box. The box contains a detector that triggers a mechanism capable of poisoning the cat if the nucleus decays. Both the nucleus and the cat are then considered to be in one of two possible states: (a) for the nucleus, $|0\rangle_N$ if it has not decayed and $|1\rangle_N$ if it has already decayed; (b) for the cat, $|1\rangle_C$ if it is alive and $|0\rangle_C$ if it has died. Assuming the joint system behaves quantum mechanically, its wave function as a function of time is

$$|\Psi\rangle = C_{01}(t)\,|0\rangle_N\,|1\rangle_C + C_{10}(t)\,|1\rangle_N\,|0\rangle_C\,. \tag{14.91}$$

https://doi.org/10.1017/9781009679633.015 Published online by Cambridge University Press

This is an "entangled state of cat+nucleus."

If, instead of a cat, another nucleus or atomic system were considered, the description would be perfectly acceptable. But describing the state of the cat in a superposition of dead and alive cats is an unwarranted extrapolation: multimolecular macroscopic entities like cats do not get entangled (with nuclei or whatever). Moreover, one must consider that not only the cat but also the detector and the poison dispenser are macroscopic entities.

The cat paradox is further accentuated by "measuring," that is, "looking inside the box," since we know from experience that we will always find a dead or alive cat, but never the superposition (14.91). In an attempt to escape the paradox, and on the assumption that $|\Psi\rangle$ describes the state of each individual system, the postulate of the collapse or reduction of the wave function is introduced, according to which it is instantly reduced (collapsed) by the act of observation to $|0\rangle_N\,|1\rangle_C$ if the nucleus has not yet decayed (and the cat is still alive), or to $|1\rangle_N\,|0\rangle_C$ if the nucleus has decayed and the cat is found dead.

From the ensemble perspective, however, there is no physical collapse, nor is the situation paradoxical. Even before measuring, the cat is either alive or dead. The observation in each individual instance (i.e., each element of the ensemble) gives information that can be used to construct a new description; for example, we can construct states representing separate ensembles of dead and alive cats. This happens with every statistical description (classical or quantum) as a result of using the new information. In other words, a statistical description is never complete nor does it exhaust the possibilities, so that the increase in information allows the description to be improved.

We conclude that what is paradoxical is not QM itself, but its application to a system that is obviously not quantum. It is inappropriate to conceive of a system consisting of a macroscopic object composed by a large number of particles, such as a real cat (plus the decaying atomic nucleus) as being in a superposition state as described by Eq. (14.91), called a Schrödinger cat state, because of the unavoidable lack of quantum coherence; see Section 15.5.7.

Schrödinger cats have become a paradigm of nonclassical, purely quantum mechanical states and a subject of intense investigation. Cat states with crystals weighing a few micrograms have already been produced; however, realizing a coherent superposition of larger objects in macroscopically distinct states, remains a challenge due to the inevitable decoherence induced by the environment.

14.5.4 The EPR *Gedankenexperiment*

For the sake of clarity in the following discussion, we will use a variation of the original EPR example given by Bohm and Aharonov, called EPRB. Consider a composite particle with spin zero at rest, which spontaneously decays into two particles, each with spin 1/2, moving in opposite directions along the x-axis. The wave function describing the spin state of the bipartite system is given by Eq. (14.89), which represents a scalar with spin 0, that is, with no privileged direction; the state is completely isotropic.

Let us consider this system from the EPR point of view. When the particles are so far apart that there is no interaction between them, we measure (with a Stern–Gerlach apparatus) the spin of particle 2 along an arbitrary direction z, perpendicular to x;

https://doi.org/10.1017/9781009679633.015 Published online by Cambridge University Press

suppose the result is +1/2. This tells us that if the z-projection of the spin of particle 1 (which we have not measured) were measured, it would be $-1/2$. For EPR this result, predicted with probability 1, means that particle 1 has, in some appropriate sense, a spin $-1/2$ on the z-axis.

We now decide to measure spin 2 along the y-axis instead of the z-axis. Suppose we get +1/2 again. This means that the y-projection of spin 1 would be $-1/2$ with probability 1, a prediction that is just as legitimate as the previous one. This imaginary procedure leads us to conclude that the spin of particle 1 has (in some sense) a well-defined value simultaneously in both the y and the z directions. In QM, however, these values cannot coexist, since there are no simultaneous eigenstates of the noncommuting operators σ_y and σ_z. The conclusion of EPR is that not all objective properties of a quantum system are contained in the description provided by the wave function, so the theory is incomplete – like any statistical theory. The assertion of the incompleteness of QM is known as the EPR theorem.

At present, the same thought experiment and a similar analysis are commonly used to conclude, not that QM is incomplete, but that the dynamical variables do not possess well-defined values (or have no values at all) prior to their measurement, or, in plain and more frequent words, that it is the observation that precisely defines or assigns the value. The observer and his measuring instruments enter actively into the description, but none of this has any place in the Hamiltonian of the system.

14.5.5 The Bell inequalities

The EPR debate leads naturally to the question: Can the quantum description be completed? One of the root causes of the belief about the impossibility of completing the description is the no-hidden-variable theorem first formulated by J. von Neumann[13] in 1932. Shortly after its publication, in 1935, the proof was shown to be flawed by the young German mathematician Grete Hermann (1901–1984), but her work went largely unrecognized. After Bohm (1952) gave a constructive proof of the possibility of a hidden-variable (HV) theory, Bell rediscovered the error in von Neumann's proof in 1964 and began to formulate his famous inequalities.[14]

Thirty years after EPR, Bell (1964) used the thought-experiment method to establish his theorem, which has become a major attraction for theoretical and experimental work on the foundations of quantum theory. His aim was to obtain a result that would allow testing the hypothesis that the quantum description can be completed by introducing HV. The introduction of such variables would serve to restore determinism in the completed description. For example, if measuring the momentum of a particle in two equivalent experiments yields p_1 in one case and p_2 in the other, it is assumed that

[13] János (John) Lajos Neuman (1903–1957) was an outstanding Hungarian–US polymath who made fundamental contributions to mathematics, theoretical physics, computation, economics, cybernetics and other fields; he is considered one of the most important mathematicians of the twentieth century.

[14] Over the years, a number of theorems have appeared claiming the impossibility of constructing HV theories consistent with QM. They all ended up specifying the kind of HV that were forbidden or allowed. However, this is not a distinguishing feature of quantum theory; no-go theorems appear in many branches of knowledge, from social issues such as national censuses to mathematical questions (as is well known thanks to Gödel's theorems), with a relevant position in current information theory (see, e.g., no-go theorems in Wikipedia).

https://doi.org/10.1017/9781009679633.015 Published online by Cambridge University Press

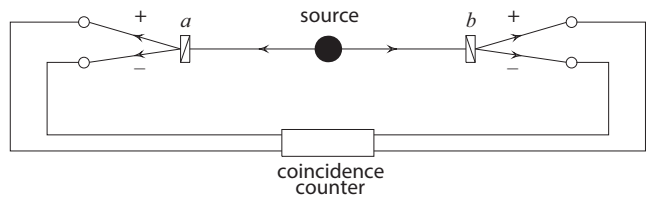


Figure 14.2 Diagram of Bell's test with coupled spin-1/2 particles in a singlet (total spin 0) state. The orientations a and b of the Stern–Gerlach apparatuses can be adjusted by the experimenter. The individual coincidences (++, −, +- and -+) registered are added to obtain the experimental correlation $C_E(a, b)$.

this is because the variables that determine the precise value of p in the experiment, which are hidden to the current quantum description, had a value λ_1 in the first case and a different value λ_2 in the second. If we had access to the λ variables (which would then no longer be hidden and belong to a more detailed theory than current QM) we would be able to predict or know exactly the value of p in each case.

Following Bell's line of argument, let us resort to the EPRB version and analyze the correlation of the projections of spins 1 and 2 along directions $\boldsymbol{a}$, $\boldsymbol{b}$, respectively, for a two-spin entangled system described by Eq. (14.89), see Fig. 14.2. The condition $S = 0$ clearly indicates that the two 1/2-spins are correlated; the purpose is to determine whether the quantum correlation can be explained in terms of HV and, if so, to draw some conclusion about the nature of these variables.

Let us first outline the experimental procedure. Since the singlet state is isotropic, we can make the z-axis coincide with the direction of $\boldsymbol{a}$. Suppose we measure the projection of spin 1 along $\boldsymbol{a}$; the result, A, will be either +1 or −1. We do the same with spin 2, this time measuring it along $\boldsymbol{b}$; the result, B, will be either +1 or −1. The large separation between the two particles means that the measurements are independent and there is no a priori reason to consider the possibility of a mutual influence. From the experimental data collected after a large number N of measurements, we can calculate the quantity

$$C_E(\boldsymbol{a}, \boldsymbol{b}) = \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} A_n B_n, \tag{14.92}$$

where A_n, B_n are the results ± 1 of measurement n. This expression gives the experimental value of the correlation of the projections of spins 1 and 2 measured along $\boldsymbol{a}$ and $\boldsymbol{b}$, respectively.

To establish an inequality, we consider four series of experiments that consist in measuring first in the directions $\boldsymbol{a}$, $\boldsymbol{b}$, then $\boldsymbol{a}$, $\boldsymbol{b}'$, then $\boldsymbol{a}'$, $\boldsymbol{b}$, and finally $\boldsymbol{a}'$, $\boldsymbol{b}'$. Assuming that the values of the four outcomes A, A', B, B' are *predetermined* – that is, that the HV λ assigned to each and every particle is independent of the setup and of the measurement[15] – we write, for every λ,

$$S_E(\boldsymbol{a}, \boldsymbol{a}', \boldsymbol{b}, \boldsymbol{b}') = A(\boldsymbol{a})B(\boldsymbol{b}) - A(\boldsymbol{a})B'(\boldsymbol{b}') + A'(\boldsymbol{a}')B(\boldsymbol{b}) + A'(\boldsymbol{a}')B'(\boldsymbol{b}'), \tag{14.93}$$

or in shorthand notation,

[15] More explicitly, this assumption means that the values of each outcome do not depend on the specific series (or measurement context), be it $(\boldsymbol{a}, \boldsymbol{b})$, $(\boldsymbol{a}, \boldsymbol{b}')$, $(\boldsymbol{a}', \boldsymbol{b})$ or $(\boldsymbol{a}', \boldsymbol{b}')$. This assumption – which has important implications, as will be seen below – is often referred to in the literature as measurement independence or context independence.

https://doi.org/10.1017/9781009679633.015 Published online by Cambridge University Press

$$S_E = A(B - B') + A'(B + B'). \tag{14.94}$$

Since $B' = \pm B$, one of the expressions in the parentheses always vanishes, which gives

$$S_E = \pm 2. \tag{14.95}$$

Therefore, the average of $|S_E|$ taken over many measurements must satisfy the inequality

$$\overline{|S_E|} = \overline{|A(B - B') + A'(B + B')|} \leqslant 2. \tag{14.96}$$

The theoretical equivalent of (14.92), assuming the existence of a set of variables λ that determine the result of each measurement, with $\rho(\lambda)$ being the corresponding probability, is

$$C_B(\boldsymbol{a}, \boldsymbol{b}) = \int_\Lambda A(\boldsymbol{a}, \lambda) B(\boldsymbol{b}, \lambda) \rho(\lambda) d\lambda, \tag{14.97}$$

where the subindex B refers to the correlation under the Bell conditions, that is, the assumption of deterministic HV and the requirement of factorization,

$$(AB)(\boldsymbol{a}, \boldsymbol{b}; \lambda) = A(\boldsymbol{a}, \lambda) B(\boldsymbol{b}, \lambda), \tag{14.98}$$

which implies that the probability distributions for the two particles are independent of each other: $A(\boldsymbol{a}, \lambda)$ does not depend on the direction $\boldsymbol{b}$ in which spin 2 is measured, and likewise $B(\boldsymbol{b}, \lambda)$ does not depend on $\boldsymbol{a}$, in accordance with the general principle of locality.[16] The procedure used to arrive at the inequality for $\left|\overline{S_E}\right|$, Eq. (14.96), can now be applied to (14.97) to obtain

$$\overline{|S_B|} = \left| C_B(\boldsymbol{a}, \boldsymbol{b}) - C_B(\boldsymbol{a}, \boldsymbol{b}') + C_B(\boldsymbol{a}', \boldsymbol{b}) + C_B(\boldsymbol{a}', \boldsymbol{b}') \right| \leq 2. \tag{14.99}$$

This result, called the CHSH inequality after its authors, J. Clauser,[17] M. Horne, A. Shimony and R. Holt, is a variant of the original Bell inequality (BI).[18]

On the other hand, the quantum formula for the correlation of the spin projections is obtained as the average value of the product $\left(\hat{\boldsymbol{\sigma}}_1 \cdot \boldsymbol{a}\right)\left(\hat{\boldsymbol{\sigma}}_2 \cdot \boldsymbol{b}\right)$, taken over the singlet state of spin zero given by Eq. (14.89),

$$C_Q(\boldsymbol{a}, \boldsymbol{b}) = \langle 00 | \, \hat{\boldsymbol{\sigma}}_1 \cdot \boldsymbol{a} \hat{\boldsymbol{\sigma}}_2 \cdot \boldsymbol{b} \, | 00 \rangle \, . \tag{14.100}$$

For the calculation of this quantity we may take, without loss of generality (because of the isotropy of the singlet state), the direction $\boldsymbol{a}$ along the z-axis. This diagonalizes the operator $\hat{\boldsymbol{\sigma}}_1 \cdot \boldsymbol{a}$, with $\langle \pm | \, \hat{\boldsymbol{\sigma}}_1 \cdot \boldsymbol{a} \, | \pm \rangle = \pm 1$, and therefore

$$C_Q(\boldsymbol{a}, \boldsymbol{b}) = \tfrac{1}{2}\Big(\langle - | \, \hat{\boldsymbol{\sigma}}_2 \cdot \boldsymbol{b} \, | - \rangle - \langle + | \, \hat{\boldsymbol{\sigma}}_2 \cdot \boldsymbol{b} \, | + \rangle \Big) = -b_z = -\cos\theta,$$

[16] As mentioned earlier, according to the general principle of locality there is no reason to consider a mutual influence between the results of two spins measured in spatially separated regions. The idea that two particles can interact instantaneously over a distance – accepted until now by not a few quantum physicists – was criticized by Einstein in a letter to M. Born, with the qualifier *spukhafte Fernwirkung* (spooky action at a distance).

[17] John Francis Clauser (1942) is a US-born theoretical and experimental physicist known for his contributions to the foundations of QM, for which he received the 2022 Nobel Prize.

[18] Several variants of the original Bell inequality have been proposed; here they are generally referred to as BI. The topic has undergone much development and has been the subject of considerable controversy and objection, including the derivation of concrete local and realistic counterexamples that violate Bell's inequalities.

https://doi.org/10.1017/9781009679633.015 Published online by Cambridge University Press

or, in more general terms,

$$C_Q(\boldsymbol{a},\boldsymbol{b}) = -\boldsymbol{a}\cdot\boldsymbol{b} = -\cos\theta_{ba}, \tag{14.101}$$

where θ_{ba} is the angle formed by the unit vectors $\boldsymbol{a}$ and $\boldsymbol{b}$.

Bell's theorem is established by showing that the result (14.101) predicted by QM is incompatible with the CHSH inequality (14.99). This incompatibility can be checked by selecting certain directions $\boldsymbol{a},\boldsymbol{a}',\boldsymbol{b},\boldsymbol{b}'$. Specifically, for $\boldsymbol{a} = 90°$, $\boldsymbol{a}' = 0$, $\boldsymbol{b} = 225°$, $\boldsymbol{b}' = 135°$ (all lying on the yz plane), you get

$$\overline{|S_Q|} = \frac{4}{\sqrt{2}} = 2\sqrt{2}, \tag{14.102}$$

which evidently is in contradiction with (14.99).

Numerous experiments over the past 50 years have confirmed beyond reasonable doubt the quantum prediction (14.101) and thus the violation of the BI. A prevailing conclusion is that QM cannot be completed with local HV, which in turn is commonly interpreted to mean that QM is incompatible with a local realistic interpretation. But this is not a foregone conclusion.

As we have seen earlier, a *central assumption* in the derivation of the BI is that the outcome of every measurement on particle 1 and on particle 2 is predetermined by a specific value of the HV λ; thus, for example, the value $A(a,\lambda)$ is independent of whether it is measured along with B or with B'. This assumption allowed grouping the terms in the inequality (14.96) or its equivalent in terms of correlations, the BI (14.99), which conflicts with QM (and with experiment). Therefore, (14.99) tells us that QM cannot be completed with (local) *deterministic* HV.

If the introduction of a distribution of deterministic HV cannot complete the quantum mechanical description, the natural conclusion is that a random element is still there. Throughout this text we have referred several times to the fluctuations inherent in the quantum mechanical description and have identified the ZPF as a source. Only a finer and deeper description, aimed at explaining the ultimate cause of this (then only apparent) randomness, could remove the indeterminism of the current quantum formalism.[19] Locality and realism are still alive; there is no "spooky action at a distance," but we have to be content with a statistical description for now.

Problems

P14.1 Show that

$$[\widehat{\boldsymbol{L}}\cdot\widehat{\boldsymbol{S}},\widehat{J}_i] = 0, \tag{14.103}$$

$$[\widehat{\boldsymbol{L}}\cdot\widehat{\boldsymbol{S}},\widehat{\boldsymbol{J}}^2] = [\widehat{\boldsymbol{L}}\cdot\widehat{\boldsymbol{S}},\widehat{\boldsymbol{L}}^2] = [\widehat{\boldsymbol{L}}\cdot\widehat{\boldsymbol{S}},\widehat{\boldsymbol{S}}^2] = 0. \tag{14.104}$$

[19] This is the idea behind the superdeterministic theories advocated by some authors, notably by G. 't Hooft (2016) in the framework of his cellular automaton theory.

https://doi.org/10.1017/9781009679633.015 Published online by Cambridge University Press

P14.2 Any analytic function of the Pauli matrices can be written as a linear expression of these matrices plus the unit 2×2 matrix. Give the following expressions this form,

$$(1+\sigma_x)^n, \quad (1+\sigma_x)^{1/2}, \quad (a\sigma_x + b\sigma_y)^2; \tag{14.105}$$

$$\sigma_x^{-k}, \quad (a+\sigma_x)^{-1} \quad (a \neq 1). \tag{14.106}$$

P14.3 Show in detail that

$$e^{i\widehat{\boldsymbol{\sigma}}\cdot\widehat{\mathbf{n}}\theta} = \cos\theta + i\widehat{\boldsymbol{\sigma}}\cdot\widehat{\mathbf{n}}\sin\theta. \tag{14.107}$$

P14.4 Construct the spinors that are eigenfunctions of the operator

$$\widehat{S} = \widehat{S}_x\cos\theta + \widehat{S}_y\sin\theta \tag{14.108}$$

and show that their eigenvalues are $\lambda = \pm 1$. Apply your result to the particular cases $\theta = 0, \frac{\pi}{2}, \pi, 2\pi$.

P14.5 Solve the Pauli equation for a free particle. Which are the conserved quantities?

P14.6 Show that if the magnetic field varies only with time, the Pauli wave function can be factored into a spatial part that satisfies the Schrödinger equation with no magnetic field and a spinorial part $|\chi\rangle$, the solution of the equation

$$i\hbar\frac{\partial\,|\chi(t)\rangle}{\partial t} = \mu_0(\widehat{\boldsymbol{\sigma}}\cdot\boldsymbol{B})\,|\chi(t)\rangle\,. \tag{14.109}$$

P14.7 Show that

$$\langle JJ|\widehat{S}_z|JJ\rangle = \frac{J(J+1)+S(S+1)-L(L+1)}{2(J+1)}. \tag{14.110}$$

Hint: First prove that $\widehat{\boldsymbol{J}}\cdot\widehat{\boldsymbol{S}} = (\widehat{J}_z+1)\widehat{S}_z + \widehat{S}_-\widehat{J}_+ + \widehat{J}_-\widehat{S}_+$, in units $\hbar = 1$.

P14.8 Show that the mass appearing in the relativistic correction to the kinetic energy of the electron of a hydrogen-like atom, $-p^4/8m^3c^2$, is the mass of the electron and not the reduced mass.

P14.9 Show that the correction to the energy of a hydrogen-like atom whose nucleus is modeled as a sphere of radius R_N with uniform charge distribution is given by

$$\delta E = \frac{2}{5}\frac{Z^4e^2R_N^2}{a_0^3n^3}\delta_{l0}. \tag{14.111}$$

Note: In the nuclear region we can make the approximation $|\psi(r)|^2 \simeq |\psi(0)|^2$.

P14.10 Calculate the spin–orbit coupling effect on the energy of an isotropic 3D harmonic oscillator of mass m and frequency ω.

P14.11 Determine the energy eigenvalues and eigenfunctions for a charged particle (without spin) moving in a space occupied by an electric field and a magnetic field, both uniform and constant, whose directions are perpendicular to each other.

P14.12 Determine exactly the energy spectrum of a charged (spinless) isotropic oscillator immersed in a constant, uniform magnetic field.

https://doi.org/10.1017/9781009679633.015 Published online by Cambridge University Press

P14.13 An electron is in a magnetic field that varies with time according to the law

$$B_x = B\sin\theta\cos\omega t, \quad B_y = B\cos\theta\sin\omega t, \quad B_z = B\cos\theta. \tag{14.112}$$

At time $t = 0$ the projection of the spin in the direction of the field is $+1/2$. Find the probability as a function of time that the particle is in the $-1/2$ spin state in the direction of the field, for $t > 0$.

P14.14 Three particles of spin $1/2$, placed at the corners of an equilateral triangle, are described by the interaction Hamiltonian

$$\widehat{H} = \tfrac{1}{3}\lambda\left(\hat{\sigma}_1\cdot\hat{\sigma}_2 + \hat{\sigma}_1\cdot\hat{\sigma}_3 + \hat{\sigma}_2\cdot\hat{\sigma}_3\right). \tag{14.113}$$

List the energy levels and their degeneracy.

P14.15 Consider a neutron interference experiment in which the two beams have opposite polarizations on the z-axis, so that the superposition of the beams at the output of the instrument has the spin in the xy-plane. Under these conditions, if the spin of one of the two branches is inverted by means of a magnetic field, the final spin will be completely polarized on the z-axis. Show that this is so.

Bibliographical Notes

For a detailed discussion of entanglement and some of its applications, see Auletta, Fortunato, and Parisi (2009).

For Bohm's variant of the EPR thought experiment with an entangled spin pair (EPRB), see Bohm (1951), Bell (1964).

Von Neumann's theorem of 1932 was followed by a series of increasingly refined variants. For a survey, see Avis, Moriyama, and Owari (2009).

For a concrete derivation of the entangled spin pair with local HV that reproduces the quantum correlation and thus violates the BI, see Cetto (2022).

https://doi.org/10.1017/9781009679633.015 Published online by Cambridge University Press