8 Local reality and the Einstein–Bohr debate

- According to the orthodox interpretation of quantum mechanics (Niels Bohr being its leading voice), the attributes of a physical object (position, momentum, spin, etc.) can be assigned only when they have been measured. Einstein advocated, as more reasonable, the local realist viewpoint that a physical object has definite attributes whether they have been measured or not.
- The orthodox view that measurement actually produces an object's property implies that the measurement of one part of an entangled quantum state would instantaneously produce the value of another part, no matter how far the two parts have been separated. Einstein, Podolsky, and Rosen devised a thought experiment in order to shine a light on this "spooky action-at-a-distance" feature of the orthodoxy; its discussion and debate have illuminated some of the fundamental issues related to the meaning of quantum mechanics.
- Such discussions led later to Bell's theorem showing that these seemingly philosophical questions could lead to observable results. The experimental vindication of the orthodox interpretation has sharpened our appreciation of the nonlocal features of quantum mechanics. Nevertheless, the counter-intuitive picture of objective reality as offered by quantum mechanics still troubles many, leaving one to wonder whether quantum mechanics is ultimately a complete theory.

8.1 Quantum mechanical basics—superposition and probability

Recall our discussion of wave–particle duality in Section 6.1. Physical objects are found to be neither simply waves nor simply particles, but to have wave and particle attributes simultaneously—two seemingly contradictory properties at the same time. In the new quantum mechanics (QM) they are represented by quantum states, which are taken to be vectors in a linear algebra space, called

the Hilbert space. These vectors can be added and obey equations that are linear—hence display the property of waves. These waves are interpreted as probability waves.

In less abstract language, a central quantity in quantum mechanics is the wavefunction, for example, the position wavefunction $\psi(x)$. It satisfies the Schrödinger wave equation, which is a linear differential equation; any linear combination of its solutions is still a solution. The wavefunction $\psi(x)$ has the interpretation that when we make a measurement of a particle's position, $|\psi(x)|^2\,dx$ is the probability of finding the particle in the interval $(x, x + dx)$. Namely, there is the possibility of a quantum mechanical position state in which the particle is in more than two, in fact an infinite number of, positions simultaneously.

Before a measurement is made, the particle can be in all these positions simultaneously. A measurement of the position would yield one particular value, say x_A. To make a measurement of a particle's position finding any particular value, according to the orthodox interpretation, is a random process, only subject to the likelihood as predicted by the probability distribution given by the wavefunction. Many physicists were ill at ease with the probability feature being built right into the foundation of the theory. Einstein famously objected: "God does not play dice!"

8.2 The Copenhagen interpretation

If a measurement of the position finds the particle to be at x_A, then immediately afterwards one should find the particle at x_A as well. Namely, due to the measurement, the particle "jumps" from the state being simultaneously in all positions to a state with definite position at x_A. According to the interpretation of the wavefunction as given above, the measurement causes the "collapse" of the wavefunction. Thus there are two fundamentally distinctive categories of physical processes in quantum mechanics:

1. **Smooth evolution of the wavefunction:** The Schrödinger equation completely determines the behavior of the wavefunction. There is nothing random about this description.
2. **Quantum mechanical measurement:** A measurement to obtain a particular result, according to quantum mechanics, is a random process. The theory only predicts the probability of getting any particular outcome. A measurement, which involves the interaction between the micro and macro realms of physics, collapses the wavefunction. This collapse of the wavefunction is not described by the Schrödinger equation; it is necessarily a non-local process as the wavefunction changes its value everywhere instantaneously.

8.2.1 The Copenhagen vs. the local realist interpretations

The first category of processes is noncontroversial, while questions related to measurement (the second category) bring out the strangeness of the QM theory.

One asks the question: "If the measurement finds the particle at x_A, where **was** the particle **just before** the measurement?"

- **Local realists:** "A particle has an objective reality; it has a set of attributes whether they have been measured or not. Therefore, it was at x_A just before the measurement finding it at x_A." According to Einstein, Schrödinger, *et al.,* this is the answer a reasonable theory would have given.
- **Orthodox quantum mechanical interpretation:** Bohr, Heisenberg, Born, and Jordan, *et al*. (the Copenhagen School) would have replied: "The particle was not really anywhere." Not only it is impossible to know, but it's **not even meaningful** to ask such a question (like asking the marital status of a table). The framework of every theory determines the relevant issue; in quantum mechanics, such a question should not have even been asked! Thus, a Copenhagen theorist would say: "The measurement compels the particle to assume a position. Observations not only disturb what's measured, they produce it."

To local realists, like Einstein, a particle must have objective reality (mass, spin, position, etc.) *independent* of whether these properties are being measured or not. (The moon is there whether you look at it or not.) Thus quantum mechanics must be an incomplete theory—the particle is at x_A, yet the theory cannot tell us it is so.

8.3 EPR paradox: Entanglement and nonlocality

From 1928 onward, Einstein had engaged Niels Bohr (the leading proponent of the Copenhagen school) in a series of debates (some public, but mostly private) as to the meaning of measurements in quantum mechanics. To sharpen his argument, to bring out the strangeness of the theory more clearly, Einstein, with his collaborators Boris Podolsky and Nathan Rosen, published in 1935 a paper in which a thought experiment was discussed in order to bring out clearly the underlying nonlocal nature of quantum mechanics.[1] The influence of this paper has grown over the years as subsequent developments showed that the question it raised was of fundamental importance to the meaning of quantum mechanics.

[1]Bohr's rejoinder can be found in Bohr (1935).

Local reality in physics Let us recall briefly the history of the locality concept in physics. One aspect of Newton's theory of gravitation that he himself found unsatisfactory is the invocation of the "action-at-a-distance" force. Somehow the source particle can act instantaneously on the test particle some distance away. The same situation holds for Coulomb's law. This was later remedied with the introduction of the Faraday–Maxwell field. In a field theory such an interaction is pictured as a two-step process: the source particle brings about a field everywhere (with the field emanating from the source and propagating outward at a finite speed). The field then acts on the test

particle **locally**. Thus, through field theory, locality was restored back to physics. Einstein now points out that quantum mechanics brought about a new form of nonlocality.

The EPR thought experiment We shall present the Einstein–Podolsky–Rosen (EPR) "paradox" as simplified and sharpened by David Bohm (1917–92). Consider the decay of a spin-zero particle, for definiteness take it be a neutral pion (a spin-zero elementary particle), into an electron and positron pair (both have spin one-half):

$$e^+ \longleftarrow \pi^0 \longrightarrow e^-.$$

The decay products will speed away from each other in opposite directions. To have angular momentum conservation, the spins of the daughter pair must be opposite each other in order for their sum to be a spin-zero state. In quantum mechanics such a final state is a superposition of two states: in one the electron spin is up, in another it's down:[2]

$$\psi_0 = \frac{1}{\sqrt{2}} \left(\psi_\uparrow^{(-)} \psi_\downarrow^{(+)} - \psi_\downarrow^{(-)} \psi_\uparrow^{(+)} \right) \tag{8.1}$$

where $\psi^{(-)}$ is the wavefunction of the electron e^-, and $\psi^{(+)}$ is the wavefunction of the positron e^+. The subscripts indicate their respective spin orientation (in some definite direction, say the z direction). Since these two terms have coefficients with equal magnitude, there is equal probability for either outcome to take place. These two terms, the two (product) wavefunctions $\psi_\uparrow^{(-)} \psi_\downarrow^{(+)}$ and $\psi_\downarrow^{(-)} \psi_\uparrow^{(+)}$, superpose to make up one quantum state as the final state of this decay process.

[2]The minus sign, which is irrelevant for our discussion, reflects the relative phases of Clebsch–Gordan coefficients in the addition of two spin-1/2 states to form a spin-zero angular momentum state.

The entangled states

We now measure the spin orientations of the e^+e^- pair. Let us concentrate on the electron spin. There is a 50% chance of finding the electron's spin being up and 50% chance down. But once the electron spin is measured, say finding it to be spin up, we are 100% sure that the positron spin must be down—this is so no matter how far away the positron has traveled: to the other side of the lab bench, or to the other side of the galaxy. One can perform such a measurement repeatedly and the spin orientations of these widely separated particle pairs are always 100% correlated. Such a correlation[3] of the spin states is described as being "entangled". In quantum mechanics, these two particles are in one entangled quantum state; any change will affect both particles together instantaneously.

[3]Mathematically, it has the feature that the probability of finding any particular combinations of electron and positron spin states is not a simple product of probabilities of finding electron and positron spin states separately.

Entanglement as viewed by local realists To a local realist this entanglement by itself does necessarily represent a deep puzzle. We encounter this sort of total correlation often in our daily experience. Let's say Chris and Alex have two coins, one gold and the other silver. They also have two boxes, each holding one coin so that Chris carries one and Alex the other. After they departed from each other, Chris opens his box finding a silver coin. It does not matter how far he had travelled, he knows immediately that Alex has the gold one.

Thus to a local realist, the perfect correlation of the e^+e^- spins simply means they were correlated **before** they were measured. Namely, since the beginning, the silver coin had been in the box Chris took and the gold coin in the box Alex had.

Copenhagen interpretation of entanglement However the interpretation given by the Copenhagen school is very different. This orthodox view would say that, before the measurement, the electron and positron were not in definite spin states. Their quantum state is not just one or the other, but a superposition of these two possibilities: the electron is in both spin up and down states, while the positron also does not have definite spin, but always has its spin pointing in the opposite direction to the electron spin. Because the electron and positron are entangled, according to the orthodox interpretation, a measurement of the electron spin **compels** the positron, no matter how far away it had traveled, to jump into some definite spin state (opposite to that of the measured electron). Einstein found such an instantaneous effect so strange that he called it "spooky action-at-a-distance". This comes about because of the claim that the states do not have any definite attributes until they have been measured. It is the measurement here that compels the positron to jump into its spin state over there!

Local realist hidden-variable theories To the local realists, the two particles always had some definite spin orientation, yet quantum mechanics can only predict it with some probability. This just means that quantum mechanics is an incomplete theory. The suggestion is made that there is a set of yet unknown variables; their specification in a more complete theory would then lead to definitive predictions. Such local realist theories are often referred to as "hidden-variable theories".

Hidden-variable theories can account for the quantum mechanical result in simple situations In simple situations, for example, measuring the spin components in the same orientation or in two perpendicular orientations, hidden-variable theories can account for the quantum mechanical result Putting this in more quantitative terms, both hidden-variable theories and QM will find the average (as denoted by $\langle \ldots \rangle$) product value of electron and positron spins (in units of $\hbar/2$) in any particular direction, whether in the z direction or x direction, to be

$$\left\langle S_z^{(-)} S_z^{(+)} \right\rangle = \left\langle S_x^{(-)} S_x^{(+)} \right\rangle = -1. \tag{8.2}$$

The precise quantum mechanical calculation is presented in SuppMat Section 8.4, see Eq. (8.23). The same result can be understood in the framework of the local realist interpretation: as any pair of electrons and positrons is produced, its members have opposite spin immediately after their birth. This clearly holds as well when we average over all the measured values.

In the above we discussed only one spin orientation at a time. Now consider measuring spins in two independent directions, say the z and x directions. A quantum mechanical calculation [Eq. (8.24) below] shows that

$$\left\langle S_z^{(-)} S_x^{(+)} \right\rangle = 0. \tag{8.3}$$

It is not difficult for hidden-variable theories to reproduce this result. For example, one can have a theory that allows independent spin orientations in the two perpendicular directions to have electron spin up and positron spin down, randomly (namely, equally likely). They all average out to zero.

8.3.1 The post-EPR era and Bell's inequality

Many (most?) working physicists took an agonistic viewpoint. Since these issues concern the interpretation of the situation before measurements, one can adopt a "shut-up-and-calculate" attitude. One just uses the Schrödinger equation to compute quantities of practical interest and ignores the "philosophical puzzles". Then came the surprise when John S. Bell (1928–90) published a paper (Bell 1964) showing that what were assumed for the situations "before the measurement" (e.g. whether e^+e^- spins were already correlated) actually have experimentally observable consequences (Bell's inequality). Since the 1980s a whole series of experimental results have demonstrated that the local realist viewpoint[4] is not supported by observation. The orthodox way of interpreting entanglement has gained ground. It is interesting to note that currently the researchers whose work has more bearing on these "philosophical issues" are actually the ones pursuing the very practical ends of constructing "quantum computers", for which QM entanglement is of paramount importance. One way or another, Einstein's thoughts on the deep meanings of quantum mechanics still exerts an influence on current investigations.

[4] Here we ignore the "many-worlds interpretation" of quantum mechanics, advocated by Hugh Everrett. Some would argue this interpretation as the ultimate "realist" theory.

Bell's inequality derived

Basically, what Bell did was to extend the above discussion of spin values S_z and S_x to more than two directions. In such richer systems, one can deduce relations that can distinguish the local realist interpretation from that of QM, independent of the assumed forms of any hidden-variable theory.

Let us again consider the spin measurement of an electron and positron produced by a parent system having zero angular momentum. We can measure the spin in any direction perpendicular to the e^+e^- pair's motion (call it the $\hat{\mathbf{y}}$ direction); we broaden our consideration from just the $\hat{\mathbf{z}}$ and $\hat{\mathbf{x}}$ directions to three directions $\left(\hat{\mathbf{a}}, \hat{\mathbf{b}}, \hat{\mathbf{c}}\right)$ in the x-z plane. According to the local realists, the particles must have definite spin values at all times: the electron's spin in the $\hat{\mathbf{a}}$ direction can take on value of $S_a^{(-)} = \pm 1$ (in units of $\hbar/2$), similarly $S_b^{(-)} = \pm 1$ and $S_c^{(-)} = \pm 1$. For notational simplicity we will write the electron spin values as $S_a^{(-)} \equiv E(a, \lambda) = \pm 1$, $S_b^{(-)} \equiv E(b, \lambda) = \pm 1$, and $S_c^{(-)} \equiv E(c, \lambda) = \pm 1$, respectively. We have explicitly displayed their dependence on the hidden variable λ. For positron spins, $S_a^{(+)}$, $S_b^{(+)}$, and $S_c^{(+)}$, we write $P(a, \lambda) = \pm 1$, $P(b, \lambda) = \pm 1$, and $P(c, \lambda) = \pm 1$, respectively. Since they must form a spin-zero system, we must have $E(a, \lambda) = -P(a, \lambda)$, $E(b, \lambda) = -P(b, \lambda)$, and $E(c, \lambda) = -P(c, \lambda)$.

Instead of presenting Bell's original derivation, we shall present one only involving simple arithmetic (d' Espagnat 1979). The electron spin in three directions $\left(\hat{\mathbf{a}}, \hat{\mathbf{b}}, \hat{\mathbf{c}}\right)$ has $2^3 = 8$ possible configurations. Thus in the **local**

realist's approach we can think[5] of the following possible electron and positron spin configurations as soon as the particles are produced:

[5]From the Copenhagen viewpoint, it is not allowed even in principle to think of any definite spin states before their measurement.

$$\begin{array}{cccccc|ccc} & E_a & E_b & E_c & \longleftrightarrow & & P_a & P_b & P_c \\ N_1 & + & + & + & & & - & - & - \\ N_2 & + & + & - & & & - & - & + \\ N_3 & + & - & + & & & - & + & - \\ N_4 & + & - & - & & & - & + & + \\ N_5 & - & + & + & & & + & - & - \\ N_6 & - & + & - & & & + & - & + \\ N_7 & - & - & + & & & + & + & - \\ N_8 & - & - & - & & & + & + & + \end{array} \tag{8.4}$$

N_i is the number of events having the spin configuration in the ith row. Thus there are N_1 events with $S_a^{(-)} = E(a,\lambda) = +1$, $E(b,\lambda) = +1$, and $E(c,\lambda) = +1$, etc. (The exact values of N_i are to be determined, hopefully, in some hidden-variable theory.) Because the electron and positron spins must be anti-aligned (in order to have zero total angular momentum), the positron spin configuration in the second group of columns must be exactly opposite to those in the electron column (first group): thus $P_a = -E_a$ and $P_b = -E_b$, etc.

The probability of having the ith row configuration is $p_i = N_i/\Sigma N$ with ΣN being the total number of events. Thus, according to the local realists (lr) approach, the average value of a spin product $\left\langle S_a^{(-)} S_b^{(+)} \right\rangle_{\text{lr}} \equiv \langle a,b\rangle$ is

$$\langle a,b\rangle = \sum_i (p_i)\,[E(a)\,P(b)]_i = \sum_i N_i\,[E(a)\,P(b)]_i\,/\,\Sigma N. \tag{8.5}$$

From the table in (8.4) we see that the values of the spin products are $[E(a)\,P(b)]_{1,2,7,8} = -1$ and $[E(a)\,P(b)]_{3,4,5,6} = +1$ (with the row number being indicated by the subscript). This allows us to write out weighted sums such as (8.5) explicitly,

$$\langle a,b\rangle = (-N_1 - N_2 + N_3 + N_4 + N_5 + N_6 - N_7 - N_8)\,/\,\Sigma N.$$

Similarly, we can calculate the average product value for spins in the a and c direction:

$$\langle a,c\rangle = (-N_1 + N_2 - N_3 + N_4 + N_5 - N_6 + N_7 - N_8)\,/\,\Sigma N;$$

thus

$$\langle a,b\rangle - \langle a,c\rangle = 2\,(-N_2 + N_3 + N_6 - N_7)\,/\,\Sigma N. \tag{8.6}$$

We also have

$$\langle b,c\rangle = (-N_1 + N_2 + N_3 - N_4 - N_5 + N_6 + N_7 - N_8)\,/\,\Sigma N,$$

$$1 = (+N_1 + N_2 + N_3 + N_4 + N_5 + N_6 + N_7 + N_8)\,/\,\Sigma N,$$

or

$$1 + \langle b,c\rangle = 2\,(N_2 + N_3 + N_6 + N_7)\,/\,\Sigma N. \tag{8.7}$$

A comparison of (8.6) and (8.7) leads to **Bell's inequality**,

$$|\langle a,b\rangle - \langle a,c\rangle| \leq [1 + \langle b,c\rangle]\,, \tag{8.8}$$

or, in the explicit notation of spin products, the average values according to the local realists (lr) must obey

$$\left|\left\langle S_a^{(-)} S_b^{(+)}\right\rangle_{\rm lr} - \left\langle S_a^{(-)} S_c^{(+)}\right\rangle_{\rm lr}\right| \leq \left[1 + \left\langle S_b^{(-)} S_c^{(+)}\right\rangle_{\rm lr}\right]. \tag{8.9}$$

This result is independent of any assumption of the N_i values. The significance of Bell's inequality is that any realist hidden-variable theory must satisfy such a relation.

The quantum mechanical result

How does Bell's inequality compare to quantum mechanics result? The quantum mechanical (QM) result for the correlation of spin components in two general directions (with angle θ_{ab} between them) is calculated in (8.27) below:

$$\left\langle S_a^{(-)} S_b^{(+)}\right\rangle_{\rm QM} = -\cos\theta_{ab}. \tag{8.10}$$

The two sides of Bell's inequality (8.9) have the quantum mechanical values

$$\text{LHS} = |-\cos\theta_{ab} + \cos\theta_{ac}| \tag{8.11}$$

and

$$\text{RHS} = 1 - \cos\theta_{bc}. \tag{8.12}$$

To see that Bell's inequality is incompatible with this QM result, consider, for example, the case of $(\hat{\mathbf{a}}, \hat{\mathbf{b}}, \hat{\mathbf{c}})$ with $\hat{\mathbf{a}}$ and $\hat{\mathbf{b}}$ being perpendicular, $\theta_{ab} = \pi/2$, and $\hat{\mathbf{c}}$ being 45° from $\hat{\mathbf{a}}$ and $\hat{\mathbf{b}}$: $\theta_{ac} = \pi/4$ and $\theta_{bc} = \pi/4$. Thus, the LHS (8.11) would be $\left|-\cos\frac{\pi}{2} + \cos\frac{\pi}{4}\right| = 1/\sqrt{2} \simeq 0.7$; and the RHS (8.12) would be $1 - \cos\frac{\pi}{4} \simeq 0.3$, which is clearly **not** greater than the LHS, as required by Bell's inequality. That is, in this situation, no matter what choice one makes of the N_i values, the hidden-variable theory will not be able to mimic the QM prediction.

8.3.2 Local reality vs. quantum mechanics—the experimental outcome

John Clauser (1942–) and his collaborators were the first ones to carry out, in 1972, an experimental test of Bell's inequality and found that the spooky prediction of QM do occur (Clauser and Shimony 1978). Alain Aspect (1947–) and his collaborator performed experiments (Aspect *et al.* 1981) and were able to show more convincingly that such entanglement connections also take effect instantaneously (at a speed faster than light speed), with results in agreement with QM. Thus the nonlocal feature, what Einstein termed the "spooky action-at-a-distance" effect, does seem to be a fundamental part of nature. Quantum entanglement seems to say that if you have a system composed of more than one particle, the individual particles are actually not individual. It leaves us

with a rather strange picture of reality as it seems one is not allowed in principle to assign objective attributes, independent of actual measurements.

John Bell, when referring to the implications of Aspect's experiment, speaks for many when he said,

For me it's a dilemma. I think it's a deep dilemma, and the resolution of it will not be trivial; it will require a substantial change in the way we look at things.

An aside: Quantum computer

Nowadays, physicists do not regard the peculiarities of quantum systems as a problem, but rather an opportunity. A proper appreciation of the profound counter-intuitive properties of quantum multiparticle systems and the nature of entanglement allows the possibility of using this peculiar behavior for potential applications such as quantum computing. A quantum computer is a device that makes direct use of QM phenomena, such as superposition and entanglement, to perform operations on data. Quantum computing will be a revolutionary new form of computation.

Conventional computers manipulate **bits**, each of which can take on values of 0 or 1. Thus two bits can be in four states: 00, 01, 10, 11, and n bits encompass 2^n states. But a classical computer can only be in any one of these states sequentially.

Quantum computers manipulate quantum bits, called **qubits**, which are quantum states, i.e. a **superposition** of the classical states. Namely, a quantum computer can be in many of the classical states simultaneously. For example, using an electron spin (with spin up or down), a qubit can be in $|\downarrow\rangle \equiv |0\rangle$ and $|\uparrow\rangle \equiv |1\rangle$ states, as well as in the state of $a\,|0\rangle + b\,|1\rangle$; and a two-qubit system in $a\,|00\rangle + b\,|10\rangle + c\,|01\rangle + d\,|11\rangle$, etc. The complex numbers a, b, c, d are the relative phases and amplitudes within the superposition. Thus, while a classical computer acts on binary numbers stored in the input register to output another number, a quantum computer acts on the **whole superposition** in qubits of its input register, thus achieving enormous parallelism.

* * *

The 2005 reprint of Pais' Einstein biography (Pais 1982) includes a new Foreword by Roger Penrose (1931–). This short essay is, in this author's opinion, a particularly insightful appraisal of Einstein's scientific achievement. Relevant to our discussion of Einstein's view of quantum mechanics, Penrose has this to say:

It must be said that some of Einstein's objections to quantum theory have not really stood the test of time—most notably it was "unreasonable" that the theory should possess strange **non-local** *aspects (puzzling features Einstein correctly pointed out). Yet, his most fundamental criticism does, I believe, remain valid. This objection is that the theory seems not to present us with any fully* **objective** *picture of physical reality. Here, I would myself side with Einstein (and with certain other key figures in the development of the theory, notably Schrödinger and Dirac) in the belief that quantum theory is not yet complete.*

8.4 SuppMat: Quantum mechanical calculation of spin correlations

In quantum mechanics, a state $|\psi\rangle$ is a vector[6] in the Hilbert space. It can be expanded in term of a complete set of basis vectors $\{|i\rangle\}$:

$$|\psi\rangle = \sum_i \psi_i |i\rangle . \tag{8.13}$$

The basis vectors are usually taken to be the eigenvectors of some operator A (representing some observable): $A|i\rangle = a_i|i\rangle$, where a_i is a number (the eigenvalue). This means that if the system is in the state $|i\rangle$, a measurement of the observable A is certain to obtain the result of a_i. The coefficient of expansion $\psi_i = \langle\psi\ |i\rangle$ is interpreted as the probability amplitude. A measurement of A of the system in the general state of $|\psi\rangle$ will result in obtaining one, say a_j, of the possible eigenvalues $\{a_i\}$, with probability $p_j = \left|\psi_j\right|^2$. The familiar wavefunction $\psi(x)$ is simply the coefficient of expansion in the representation space having position eigenstates $\{|x\rangle\}$ as basis vectors. In this case Eq. (8.13) becomes

$$|\psi\rangle = \int dx\psi(x)|x\rangle . \tag{8.14}$$

The orthonormality condition of the basis vectors $\langle i|\ j\rangle = \delta_{ij}$ means we have $\langle j|A|\ i\rangle = \delta_{ij}a_i$. Thus the average value of an observable A:

$$\langle A\rangle = \sum_i p_i a_i = \sum_i |\psi_i|^2 a_i, \tag{8.15}$$

can be obtained efficiently by taking the expectation value of an operator (i.e. sandwich the operator between the bra and ket vectors of the state):

$$\langle A\rangle = \sum_{i,j} \psi_j^* \psi_i \delta_{ij} a_i = \sum_{i,j} \psi_j^* \psi_i \langle j|A|\ i\rangle = \langle\psi\ |A|\ \psi\rangle . \tag{8.16}$$

To reach the last expression we have used the expansion of Eq. (8.13).

8.4.1 Quantum mechanical calculation of spin average values

Spin states

Here we shall mostly deal with spin eigenstates $|s,m\rangle$, which are labeled by eigenvalues of the total spin $S^2 = S_x^2 + S_y^2 + S_z^2$ and one spin component, say S_z, respectively:

$$S^2|s,m\rangle = s(s+1)\hbar^2|s,m\rangle , \qquad S_z|s,m\rangle = m\hbar|s,m\rangle . \tag{8.17}$$

In our notation of setting $\frac{1}{2}\hbar \equiv 1$, we shall simply label the spin state by suppressing the s value and concentrate on S_z with $m = \pm\frac{1}{2}$:

$$S_z|S_z\pm\rangle = \pm|S_z\pm\rangle . \tag{8.18}$$

In the representation space spanned by the basis vectors $|S_z\pm\rangle$

$$|S_z+\rangle \equiv |\uparrow\rangle \doteq \binom{1}{0}, \qquad |S_z-\rangle \equiv |\downarrow\rangle \doteq \binom{0}{1}, \tag{8.19}$$

[6] A reader who is not familiar with Dirac notation may simply think of the "ket" vector $|\psi\rangle$ as a column vector, and the "bra" vector $\langle\phi|$ as a row vector. The inner product is represented by a bracket $\langle\phi|\ \psi\rangle$ which is the scalar resulting from the multiplication of a row and a column vector. Similarly, an operator is represented by a matrix, and the expectation value $\langle\psi|A|\ \psi\rangle$ as the multiplication of a row vector and a matrix, then with another column vector.

the spin operators $S_{x,y,z}$ are represented by the Pauli matrices $\sigma_{x,y,z}$

$$S_z \doteq \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad S_x \doteq \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \tag{8.20}$$

as can be checked by

$$S_z \left|\uparrow\right\rangle = \left|\uparrow\right\rangle \qquad \text{and} \qquad S_z \left|\downarrow\right\rangle = -\left|\downarrow\right\rangle. \tag{8.21}$$

Spinless state resulting from adding two spin $\frac{1}{2}$ states

Adding the electron and positron spin operators $\mathbf{S}^{(-)} + \mathbf{S}^{(+)} = \mathbf{S}$ we can have total spin $S = 1$ or $S = 0$. Concentrating on the $S = 0$ state with its z component $S_z = S_z^{(-)} + S_z^{(+)}$ and spin value $m^{(-)} + m^{(+)} = M_s = 0$, the total spin state is labeled as $|S = 0, M_s = 0\rangle \equiv |0, 0\rangle$. This total spin-zero state is related to the individual electron/positron S_z eigenstates $\left|S_z^{(-)}+\right\rangle \left|S_z^{(+)}-\right\rangle \equiv \left|\uparrow^{(-)}\downarrow^{(+)}\right\rangle$, etc. as

$$|0, 0\rangle = \frac{1}{\sqrt{2}} \left(\left|\uparrow^{(-)}\downarrow^{(+)}\right\rangle - \left|\uparrow^{(-)}\downarrow^{(+)}\right\rangle\right). \tag{8.22}$$

This is an example of the expansion discussed in (8.13). Namely, the final state is a superposition of the electron/positron states with the z component of electron spin up and positron spin down and vice versa, with respective expansion coefficients $\pm 1/\sqrt{2}$. This is the same relation as (8.1) but expressed in terms of Dirac notation. Thus the probability of finding the state with the electron spin up and positron spin down is $1/2$ and the probability of finding the state with the electron spin down and positron spin up is also $1/2$.

8.4.2 Spin correlation in one direction

From this we can check that the quantum mechanical formalism yields the average value

$$\left\langle S_z^{(-)} S_z^{(+)} \right\rangle_{\text{QM}} = \left\langle 0, 0 \left| S_z^{(-)} S_z^{(+)} \right| 0, 0 \right\rangle = -1, \tag{8.23}$$

showing that the electron–positron spins must point in the opposite directions. Because of (8.22), this involves calculating the type of terms such as

$$S_z^{(-)} S_z^{(+)} \left|\uparrow^{(-)}\downarrow^{(+)}\right\rangle = \left(S_z^{(-)} \left|\uparrow^{(-)}\right\rangle\right) \left(S_z^{(+)} \left|\downarrow^{(+)}\right\rangle\right) = -\left|\uparrow^{(-)}\downarrow^{(+)}\right\rangle,$$

leading to $S_z^{(-)} S_z^{(+)} |0, 0\rangle = -|0, 0\rangle$ and the claimed result of (8.23) because of the normalization condition $\langle 0, 0| \, 0, 0\rangle = 1$. This calculation expresses the fact that $S_z^{(-)}$ and $S_z^{(+)}$ are in different spin spaces so we can have $S_z^{(-)}$ act directly on $\left|\uparrow^{(-)}\right\rangle$ and $S_x^{(+)}$ act directly on $\left|\downarrow^{(+)}\right\rangle$ as in (8.21).

8.4.3 Spin correlation in two directions

Two directions that are perpendicular

We also expect

$$\left\langle S_z^{(-)} S_x^{(+)} \right\rangle_{\rm QM} = \langle 0,0|\, S_z^{(-)} S_x^{(+)}\, |0,0\rangle = 0. \tag{8.24}$$

Namely, spin values in perpendicular directions are uncorrelated. In contrast to (8.21), the spin operator S_x flips the spin in the z direction:

$$S_x\,|\uparrow\rangle = |\downarrow\rangle \qquad \text{and} \qquad S_x\,|\downarrow\rangle = |\uparrow\rangle, \tag{8.25}$$

because, as the representations in (8.19) and (8.20) show,

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

A simple exercise shows that $\left\langle 0,0 \left| S_x^{(-)} S_x^{(+)} \right| 0,0 \right\rangle = -1$. We next calculate the action of $S_z^{(-)} S_x^{(+)}$ on the state $|0,0\rangle$ of (8.22):

$$\begin{aligned} S_z^{(-)} S_x^{(+)}\, |0,0\rangle &= S_z^{(-)} S_x^{(+)} \left(\left|\uparrow^{(-)}\downarrow^{(+)}\right\rangle - \left|\downarrow^{(-)}\uparrow^{(+)}\right\rangle \right) / \sqrt{2} \\ &= \left(\left|\uparrow^{(-)}\uparrow^{(+)}\right\rangle + \left|\downarrow^{(-)}\downarrow^{(+)}\right\rangle \right) / \sqrt{2}. \end{aligned} \tag{8.26}$$

When multiplied with the bra vector $\langle 0,0|$ of (8.22), the orthogonality conditions[7] such as $\left\langle \uparrow^{(-)} \middle|\, \downarrow^{(-)} \right\rangle = \left\langle \downarrow^{(+)} \middle|\, \uparrow^{(+)} \right\rangle = 0$ lead to $\langle 0,0|\, S_z^{(-)} S_x^{(+)}\, |0,0\rangle = 0$.

[7] An even simpler way is to note that $\left|\uparrow^{(-)}\uparrow^{(+)}\right\rangle = |S=1, M_s=+1\rangle \equiv |1,+1\rangle$ and $\left|\downarrow^{(-)}\downarrow^{(+)}\right\rangle = |1,-1\rangle$; then the orthogonality condition is $\langle 0,0|\ 1,+1\rangle = \langle 0,0|\ 1,-1\rangle = 0$.

Two general directions

Here we calculate the quantum mechanical expectation value of a product of spins in two general directions, calling them $\hat{\mathbf{a}}$ and $\hat{\mathbf{b}}$**.** We are free to choose $\hat{\mathbf{a}} = \hat{\mathbf{z}}$ and $\hat{\mathbf{b}}$ in the x-z plane: $\hat{\mathbf{b}} = \cos\theta \hat{\mathbf{z}} + \sin\theta \hat{\mathbf{x}}$. Thus $S_a^{(-)} = S_z^{(-)}$ and $S_b^{(+)} = \cos\theta S_z^{(+)} + \sin\theta S_x^{(+)}$:

$$\left\langle 0,0 \left| S_a^{(-)} S_b^{(+)} \right| 0,0 \right\rangle = \cos\theta \left\langle 0,0 \left| S_z^{(-)} S_z^{(+)} \right| 0,0 \right\rangle + \sin\theta\, \langle 0,0|\, S_z^{(-)} S_x^{(+)}\, |0,0\rangle.$$

Knowing the results of (8.23) and (8.24), we immediately have

$$\left\langle 0,0 \left| S_a^{(-)} S_b^{(+)} \right| 0,0 \right\rangle = \cos\theta(-1) + \sin\theta(0) = -\cos\theta. \tag{8.27}$$