14 Quantum mechanics

Our purpose in this chapter is to present the key concepts of quantum mechanics in the language of Hilbert spaces. The reader who has not previously met the physical ideas motivating quantum mechanics, and some of the more elementary applications of Schrödinger's equation, is encouraged to read any of a number of excellent texts on the subject such as [1–4]. Otherwise, the statements given here must to a large extent be taken on trust – not an altogether easy thing to do, since the basic assertions of quantum theory are frequently counterintuitive to anyone steeped in the classical view of physics. Quantum mechanics is frequently presented in the form of several postulates, as though it were an axiomatic system such as Euclidean geometry. As often presented, these postulates may not meet the standards of mathematical rigour required for a strictly logical set of axioms, so that little is gained by such an approach. We will do things a little more informally here. For those only interested in the mathematical aspects of quantum mechanics and the role of Hilbert space see [5–8].

Many of the standard applications, such as the hydrogen atom, will be omitted here as they can be found in all standard textbooks, and we leave aside the enormous topic of measurement theory and interpretations of quantum mechanics. This is not to say that we need be totally comfortable with quantum theory as it stands. Undoubtedly, there are some philosophically disquieting features in the theory, often expressed in the form of so-called paradoxes. However, to attempt an 'interpretation' of the theory in order to resolve these apparent paradoxes assumes that there are natural metaphysical concepts. Suitable introductions to this topic can be found in [3, chap. 11] or [4, chap. 5].

14.1 Basic concepts

Photon polarization experiments

To understand how quantum mechanics works, we look at the outcome of a number of *Gedanken experiments* involving polarized light beams. Typically, a monochromatic plane wave solution or Maxwell equations (see Problem 9.23) has electric field

$$\mathbf{E} \propto \mathrm{Re}\,(\alpha\mathbf{e}_x + \beta\mathbf{e}_y)\mathrm{e}^{i(kz-\omega t)}$$

where $\mathbf{e}_x$ and $\mathbf{e}_y$ are the unit vectors in the x- and y-directions and α and β are complex numbers such that

$$|\alpha|^2 + |\beta|^2 = 1.$$

When α/β is real we have a linearly polarized wave, as for example

$$\mathbf{E} \propto \frac{1}{\sqrt{2}}(\mathbf{e}_x + \mathbf{e}_y)e^{i(kz-\omega t)}.$$

If $b = \pm ia = \pm i/\sqrt{2}$ the wave is circularly polarized; the $+$ sign is said to be right circularly polarized, and the $-$ sign is left circularly polarized. In all other cases it is said to be elliptically polarized. If we pass a polarized beam through a polarizer with axis of polarization $\mathbf{e}_x$, then the beam is reduced in intensity by the factor $|a|^2$ and the emergent beam is $\mathbf{e}_x$-polarized. Thus, if the resultant beam is passed through another $\mathbf{e}_x$-polarizer it will be 100% transmitted, while if it is passed through an $\mathbf{e}_y$-polarizer it will be totally absorbed and nothing will come through. This is the classical situation.

As was discovered by Planck and Einstein at the turn of the twentieth century, light beams come in discrete packets called *photons*, having energy $E = h\nu = \hbar\omega$ where $h \approx 6.625 \times 10^{-27}\text{g cm}^2\,\text{s}^{-1}$ is Planck's constant and $\hbar = h/2\pi$. What happens if we send the beams through the polarizers one photon at a time? Since the frequency of each photon is unchanged it emerges with the same energy, and since the intensity of the beam is related to the energy, it must mean that the number of photons is reduced. However, the most obvious conclusion that the beam consists of a mixture of photons consisting of a fraction $|\alpha|^2$ polarized in the $\mathbf{e}_x$-direction and $|\beta|^2$ in the $\mathbf{e}_y$-direction will not stand up to scrutiny. For, if a beam with $\alpha = \beta = 1/\sqrt{2}$ were passed through a polarizer designed to only transmit waves linearly polarized in the $(\mathbf{e}_x + \mathbf{e}_y)/\sqrt{2}$ direction, then it should be 100% transmitted. However, on the mixture hypothesis only half the $\mathbf{e}_x$-polarized photons should get through, and half the $\mathbf{e}_y$-polarized photons, leaving a total fraction $\frac{1}{2}\cdot\frac{1}{2} + \frac{1}{2}\cdot\frac{1}{2} = \frac{1}{2}$ being transmitted.

In quantum mechanics it is proposed that each photon is a 'complex superposition' $\alpha\mathbf{e}_x + \beta\mathbf{e}_y$ of the two polarization states $\mathbf{e}_x$ and $\mathbf{e}_y$. The probability of transmission by an $\mathbf{e}_x$-polarizer is given by $|\alpha|^2$, while the probability of transmission by an $\mathbf{e}_y$-polarizer is $|\beta|^2$. The effect of the $\mathbf{e}_x$-polarizer is essentially to 'collapse' the photon into an $\mathbf{e}_x$-polarized state. The polarizer can be regarded both as a measuring device or equally as a device for preparing photons in an $\mathbf{e}_x$-polarized state. If used as a measuring device it returns the value 1 if the photon is transmitted, or 0 if not – in either case the act of measurement has changed the state of the photon being measured.

An interesting arrangement to illustrate the second point of view is shown in Fig. 14.1. Consider a beam of photons incident on an $\mathbf{e}_x$-polarizer, followed by an $\mathbf{e}_y$-polarizer; the net result is that no photons come out of the second polarizer. Now introduce a polarizer for the direction $(1/\sqrt{2})(\mathbf{e}_x + \mathbf{e}_y)$ – in other words a device that should block some of the photons between the two initial polarizers. If the mixture theory were correct, it is inconceivable that this could increase transmission. Yet the reality is that half the photons emerge from this intermediary polarizer with polarization $(1/\sqrt{2})(\mathbf{e}_x + \mathbf{e}_y)$, and a further half of these, namely a quarter in all, are now transmitted by the $\mathbf{e}_y$-polarizer.

How can we find the transmission probability of a polarized state $\mathbf{A}$ with respect to an arbitrary polarization direction $\mathbf{B}$? The following argument is designed to be motivational rather than rigorous. Let $\mathbf{A} = \alpha\mathbf{e}_x + \beta\mathbf{e}_y$, where α and β are complex numbers subject to $|\alpha|^2 + |\beta|^2 = 1$. We write $\alpha = \langle\mathbf{e}_x \,|\, \mathbf{A}\rangle$, called the *amplitude* for $\mathbf{e}_x$-transmission of an

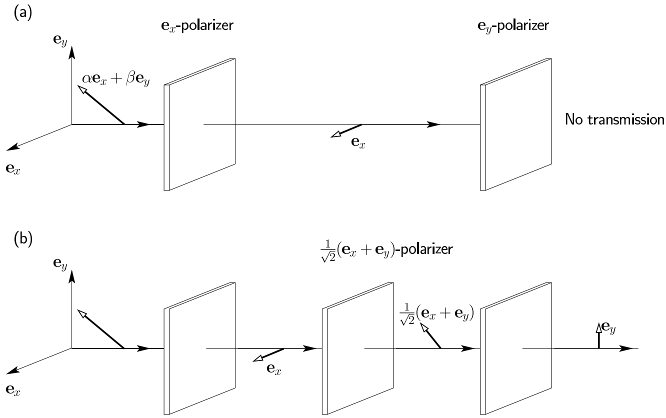

Figure 14.1 Photon polarization experiment

A-polarized photon. It is a complex number having no obvious physical interpretation of itself, but its magnitude square $|\alpha|^2$ is the probability of transmission by an $\mathbf{e}_x$-polarizer. Similarly $\beta = \langle \mathbf{e}_y \,|\, \mathbf{A} \rangle$ is the amplitude for $\mathbf{e}_y$-transmission, and $|\beta|^2$ the probability of transmission by an $\mathbf{e}_y$-polarizer. What is the polarization $A^{\perp}$ such that a polarizer of this type allows for no transmission, $\langle \mathbf{A}^{\perp} \,|\, \mathbf{A} \rangle = 0$? For linearly polarized waves with α and β both real we expect it to be geometrically orthogonal, $\mathbf{A}^{\perp} = \beta\mathbf{e}_x - \alpha\mathbf{e}_y$. For circularly polarized waves, the orthogonal 'direction' is the opposite circular sense. Hence

$$\frac{1}{\sqrt{2}}(\mathbf{e}_x \pm i\mathbf{e}_y)^{\perp} = \frac{1}{\sqrt{2}}(\mathbf{e}_x \mp i\mathbf{e}_y) \equiv \mp\frac{1}{\sqrt{2}}(i\mathbf{e}_x - \mathbf{e}_y)$$

since phase factors such as $\pm i$ are irrelevant. In the general elliptical case we might guess that $\mathbf{A}^{\perp} = \overline{\beta}\mathbf{e}_x - \overline{\alpha}\mathbf{e}_y$, since it reduces to the correct answer for linear and circular polarization. Solving for $\mathbf{e}_x$ and $\mathbf{e}_y$ we have

$$\mathbf{e}_x = \overline{\alpha}\mathbf{A} + \beta\mathbf{A}^{\perp}, \qquad \mathbf{e}_y = \overline{\beta}\mathbf{A} - \alpha\mathbf{A}^{\perp}.$$

Let $\mathbf{B} = \gamma\mathbf{e}_x + \delta\mathbf{e}_y$ be any other polarization, then substituting for $\mathbf{e}_x$ and $\mathbf{e}_y$ gives

$$\mathbf{B} = (\gamma\overline{\alpha} + \delta\overline{\beta})\mathbf{A} + (\gamma\beta - \alpha\delta)\mathbf{A}^{\perp}.$$

Setting $\mathbf{B} = \mathbf{A}$ gives the normalization condition $|\alpha|^2 + |\beta|^2 = 1$. Hence, since $\langle \mathbf{A} \,|\, \mathbf{A} \rangle = 1$ (transmission probability of 1),

$$\langle \mathbf{B} \,|\, \mathbf{A} \rangle = (\gamma\overline{\alpha} + \delta\overline{\beta}) = \overline{\langle \mathbf{A} \,|\, \mathbf{B} \rangle}.$$

Other systems such as the Stern–Gerlach experiment, in which an electron of magnetic moment $\boldsymbol{\mu}$ is always deflected in a magnetic field $\mathbf{H}$ in just two directions, exhibit a completely analogous formalism. The conclusion is that the quantum mechanical states of a system form a complex vector space with inner product $\langle \phi \,|\, \psi \rangle$ satisfying the usual conditions

$$\langle \phi \mid \alpha\psi_1 + \beta\psi_2 \rangle = \alpha\langle \phi \mid \psi_1 \rangle + \beta\langle \phi \mid \psi_2 \rangle \quad \text{and} \quad \langle \phi \mid \psi \rangle = \overline{\langle \psi \mid \phi \rangle}.$$

The probability of obtaining a value corresponding to ϕ in a measurement is

$$P(\phi, \psi) = |\langle \phi \mid \psi \rangle|^2.$$

As will be seen, states are in fact normalized to $\langle \psi \mid \psi \rangle = 1$, so that only linear combinations $\alpha\psi_1 + \beta\psi_2$ with $|\alpha|^2 + |\beta|^2 = 1$ are permitted.

The Hilbert space of states

We will now assume that every physical system corresponds to a separable Hilbert space $\mathcal{H}$, representing all possible states of the system. The Hilbert space may be finite dimensional, as for example the states of polarization of a photon or electron, but often it is infinite dimensional. A **state** of the system is represented by a non-zero vector $\psi \in \mathcal{H}$, but this correspondence is not one-to-one, as any two vectors ψ and ψ' that are proportional through a non-zero complex factor, $\psi' = \lambda\psi$ where $\lambda \in \mathbb{C}$, will be assumed to represent identical states. In other words, a state is an equivalence class or *ray* of vectors $[\psi]$ all related by proportionality. A state may be represented by any vector from the class, and it is standard to select a representative having unit norm $\|\psi\| = 1$. Even this restriction does not uniquely define a vector to represent the state, as any other vector $\psi' = \lambda\psi$ with $|\lambda| = 1$ will also satisfy the unit norm condition. The angular freedom, $\lambda = e^{ic}$, is sometimes referred to as the *phase* of the state vector. Phase is only significant in a relative sense; for example, $\psi + e^{ic}\phi$ is in general a different state to $\psi + \phi$, but $e^{ic}(\psi + \phi)$ is not.

In this chapter we will adopt Dirac's bra-ket notation which, though slightly quirky, has largely become the convention of choice among physicists. Vectors $\psi \in \mathcal{H}$ are written as **kets** $|\psi\rangle$ and one makes the identification $|\lambda\psi\rangle = \lambda|\psi\rangle$. By the Riesz representation theorem 13.10, to each linear functional $f : \mathcal{H} \to \mathbb{C}$ there corresponds a unique vector $\phi \equiv |\phi\rangle \in \mathcal{H}$ such that $f(\psi) = \langle \phi \mid \psi \rangle$. In Dirac's terminology the linear functional is referred to as a **bra**, written $\langle \phi|$. The relation between bras and kets is antilinear,

$$\langle \lambda\psi + \phi| = \overline{\lambda}\langle \psi| + \langle \phi|.$$

In Dirac's notation it is common to think of a linear operator (A, D_A) as acting to the left on kets (vectors), while acting to the right on bras (linear functionals):

$$A|\psi\rangle \equiv |A\psi\rangle,$$

and if $\phi \in D_{A^*}$

$$\langle \phi|A \equiv \langle A^*\phi|.$$

The following notational usages for the matrix elements of an operator between two vectors are all equivalent:

$$\langle \phi|A|\psi\rangle \equiv \langle \phi \mid A\psi\rangle = \langle A^*\phi \mid \psi\rangle = \overline{\langle \psi \mid A^*\phi\rangle} = \overline{\langle \psi|A^*|\phi\rangle}.$$

If $|e_i\rangle$ is an o.n. basis of kets in a separable Hilbert space then we may write

$$A|e_j\rangle = \sum_i a_{ij}|e_i\rangle \quad \text{where} \quad a_{ij} = \langle e_i|A|e_j\rangle.$$

Observables

In classical mechanics, *physical observables* refer to quantities such as position, momentum, energy or angular momentum, which are real numbers or real multicomponented objects. In quantum mechanics **observables** are represented by self-adjoint operators on the Hilbert space of states. We first consider the case where A is a hermitian operator (bounded and continuous). Such an observable is said to be **complete** if the corresponding hermitian operator A is complete, so that there is an orthonormal basis made up of eigenvectors $|\psi_1\rangle, |\psi_2\rangle, \dots$ such that

$$A|\psi_n\rangle = \alpha_n|\psi_n\rangle \quad \text{where} \quad \langle\psi_m\,|\,\psi_n\rangle = \delta_{mn}. \tag{14.1}$$

The result of measuring a complete observable is always one of the eigenvalues α_n, and the fact that these are real numbers provides a connection with classical physics. By Theorem 13.2 every state $|\psi\rangle$ can be written uniquely in the form

$$|\psi\rangle = \sum_{n=1}^{\infty} c_n|\psi_n\rangle \quad \text{where} \quad c_n = \langle\psi_n\,|\,\psi\rangle, \tag{14.2}$$

or, since the vector $|\psi\rangle$ is arbitrary, we can write

$$I \equiv \mathrm{id}_{\mathcal{H}} = \sum_{n=1}^{\infty}|\psi_n\rangle\langle\psi_n|. \tag{14.3}$$

Exercise: Show that the operator A can be written in the form

$$A = \sum_{n=1}^{\infty}\alpha_n|\psi_n\rangle\langle\psi_n|.$$

The matrix element of the identity operator between two states $|\phi\rangle$ and $|\psi\rangle$ is

$$\langle\phi\,|\,\psi\rangle = \langle\phi|I|\psi\rangle = \sum_{n=1}^{\infty}\langle\phi\,|\,\psi_n\rangle\langle\psi_n\,|\,\psi\rangle.$$

Its physical interpretation is that $|\langle\phi\,|\,\psi\rangle|^2$ is the probability of realizing a state $|\psi\rangle$ when the system is in the state $|\phi\rangle$. Since both state vectors are unit vectors, the Cauchy–Schwarz inequality ensures that the probability is less than one,

$$|\langle\phi\,|\,\psi\rangle|^2 \le \|\phi\|^2\|\psi\|^2 = 1.$$

If A is a complete hermitian operator with eigenstates $|\psi_n\rangle$ satisfying Eq. (14.1) then, according to this assumption, the probability that the eigenstate $|\psi_n\rangle$ is realized when the system is in the state $|\psi\rangle$ is given by $|c_n|^2 = |\langle\psi_n\,|\,\psi\rangle|^2$ where the c_n are the coefficients in the expansion (14.2). Thus $|c_n|^2$ is the probability that the value α_n be obtained on measuring the observable A when the system is in the state $|\psi\rangle$. By Parseval's identity (13.7) we have

$$\sum_{n=1}^{\infty} |c_n|^2 = \|\psi\|^2 = 1,$$

and the **expectation value** of the observable A in a given state $|\psi\rangle$ is given by

$$\langle A\rangle \equiv \langle A\rangle_\psi = \sum_{n=1}^{\infty} |c_n|^2 \alpha_n = \langle\psi|A|\psi\rangle. \tag{14.4}$$

The act of measuring the observable A 'collapses' the system into one of the eigenstates $|\psi_n\rangle$, with probability $|c_n|^2 = |\langle\psi_n\,|\,\psi\rangle|^2$. This feature of quantum mechanics, that the result of a measurement can only be known to within a probability, and that the system is no longer in the same state after a measurement as before, is one of the key differences between quantum and classical physics, where a measurement is always made delicately enough so as to minimally disturb the system. Quantum mechanics asserts that this is impossible, even in principle.

The **root mean square deviation** ΔA of an observable A in a state $|\psi\rangle$ is defined by

$$\Delta A = \sqrt{\langle(A - \langle A\rangle I)^2\rangle}.$$

The quantity under the square root is positive, for

$$\langle(A - \langle A\rangle I)^2\rangle = \langle\psi\,|(A - \langle A\rangle I)^2\psi\rangle = \|(A - \langle A\rangle I)\psi\|^2 \geq 0$$

since A is hermitian and $\langle A\rangle$ is real. A useful formula for the RMS deviation is

$$\begin{aligned}(\Delta A)^2 &= \langle\psi|A^2 - 2A\langle A\rangle + \langle A\rangle^2 I|\psi\rangle \\ &= \langle A^2\rangle - \langle A\rangle^2.\end{aligned} \tag{14.5}$$

Hence $|\psi\rangle$ is an eigenstate of A if and only if it is **dispersion-free**, $\Delta A = 0$. For, by (14.5), if $A|\psi\rangle = \alpha|\psi\rangle$ then $\langle A\rangle = \alpha$ and $\langle A^2\rangle = \alpha^2$ immediately results in $\Delta A = 0$, and conversely if $\Delta A = 0$ then $\|(A - \langle A\rangle I)\psi\|^2$, which is only possible if $A|\psi\rangle = \langle A\rangle|\psi\rangle$. Dispersion-free states are sometimes referred to as **pure states** with respect to the observable A.

Theorem 14.1 (Heisenberg) *Let A and B be two hermitian operators, then for any state $|\psi\rangle$*

$$\Delta A \Delta B \geq \tfrac{1}{2}\big|\langle[A, B]\rangle\big| \tag{14.6}$$

where $[A, B] = AB - BA$ is the **commutator** *of the two operators.*

Proof: Let

$$|\psi_1\rangle = (A - \langle A\rangle I)|\psi\rangle, \qquad |\psi_2\rangle = (B - \langle B\rangle I)|\psi\rangle,$$

so that $\Delta A = \|\psi_1\|$ and $\Delta B = \|\psi_2\|$. Using the Cauchy–Schwarz inequality,

$$\begin{aligned}\Delta A\,\Delta B = \|\psi_1\|\,\|\psi_2\| &\geq \big|\langle\psi_1\,|\,\psi_2\rangle\big| \\ &\geq \big|\mathrm{Im}\,\langle\psi_1\,|\,\psi_2\rangle\big| \\ &= \Big|\frac{1}{2i}\big(\langle\psi_1\,|\,\psi_2\rangle - \langle\psi_2\,|\,\psi_1\rangle\big)\Big|.\end{aligned}$$

Now

$$\begin{aligned}\langle\psi_1\,|\,\psi_2\rangle &= \langle(A-\langle A\rangle I)\psi\,|\,(B-\langle B\rangle I)\psi\rangle \\ &= \langle\psi|AB|\psi\rangle - \langle A\rangle\langle B\rangle.\end{aligned}$$

Hence

$$\Delta A\,\Delta B \geq \tfrac{1}{2}\big|\langle\psi|AB-BA|\psi\rangle\big| = \tfrac{1}{2}\big|\langle[A,B]\rangle\big|.$$

■

Exercise: Show that for any two hermitian operators A and B, the operator $i[A, B]$ is hermitian.

Exercise: Show that $\langle[A, B]\rangle = 0$ for any state $|\psi\rangle$ that is an eigenvector of either A or B.

A particularly interesting case of Theorem 14.1 occurs when A and B satisfy the *canonical commutation relations*,

$$[A, B] = i\hbar I, \tag{14.7}$$

where $\hbar = h/2\pi$ is Planck's constant divided by 2π. Such a pair of observables are said to be **complementary**. With some restriction on admissible domains they hold for the position operator $Q = A_x$ and the momentum operator $P = -i\hbar\mathrm{d}/\mathrm{d}x$ discussed in Examples 13.19 and 13.20. For, let f be a function in the intersection of their domains, then

$$[Q, P] = x\Big(-i\hbar\frac{\mathrm{d}f}{\mathrm{d}x}\Big) + i\hbar\frac{\mathrm{d}(xf)}{\mathrm{d}x} = i\hbar f,$$

whence

$$[Q, P] = i\hbar I. \tag{14.8}$$

Theorem 14.1 results in the classic **Heisenberg uncertainty relation**

$$\Delta Q\,\Delta P \geq \frac{\hbar}{2}.$$

Sometimes it is claimed that this relation has no effect at a macroscopic level because Planck's constant h is so 'small' ($h \approx 6.625\times 10^{-27}\,\mathrm{g\,cm^2\,s^{-1}}$). Little could be further from the truth. The fact that we are supported by a solid Earth, and not collapse in towards its centre, can be traced to this and similar relations.

Exercise: Show that Eq. (14.7) cannot possibly hold in a finite dimensional space. [*Hint*: Take the trace of both sides.]

Theorem 14.2 *A pair of complete hermitian observables A and B commute, $[A, B] = 0$, if and only if there exists a complete set of common eigenvectors. Such observables are said to be* **compatible**.

Proof: If there exists a basis of common eigenvectors $|\psi_1\rangle, |\psi_2\rangle, \dots$ such that

$$A|\psi_n\rangle = \alpha_n|\psi_n\rangle, \quad B|\psi_n\rangle = \beta_n|\psi_n\rangle,$$

then $AB|\psi_n\rangle = \alpha_n\beta_n|\psi_n\rangle = BA|\psi_n\rangle$ for each n. Hence for arbitrary vectors ψ we have from (14.2)

$$[A, B]|\psi\rangle = (AB - BA)\sum_n |\psi_n\rangle\langle\psi_n\,|\,\psi\rangle = 0.$$

Conversely, suppose that A and B commute. Let α be an eigenvalue of A with eigenspace $M_\alpha = \{|\psi\rangle \mid A|\psi\rangle = \alpha|\psi\rangle\}$, and set P_α to be the projection operator into this subspace. If $|\psi\rangle \in M_\alpha$ then $B|\psi\rangle \in M_\alpha$, since

$$AB|\psi\rangle = BA|\psi\rangle = B\alpha|\psi\rangle = \alpha B|\psi\rangle.$$

For any $|\phi\rangle \in \mathcal{H}$ we therefore have $BP_\alpha|\phi\rangle \in M_\alpha$. Hence

$$P_\alpha B P_\alpha|\phi\rangle = BP_\alpha|\phi\rangle$$

and since $|\phi\rangle$ is an arbitrary vector,

$$P_\alpha B P_\alpha = BP_\alpha.$$

Taking the adjoint of this equation, and using $B^* = B$, $P_\alpha = P_\alpha^*$, gives

$$P_\alpha B P_\alpha = P_\alpha^* B^* P_\alpha^* = P_\alpha^* B^* = P_\alpha B$$

and it follows that $P_\alpha B = BP_\alpha$, the operator B commutes with all projection operators P_α.

If β is any eigenvalue of B with projection map P_β, then since P_α is a hermitian operator that commutes with B the above argument shows that it commutes with P_β,

$$P_\alpha P_\beta = P_\beta P_\alpha.$$

Hence, the operator $P_{\alpha\beta} = P_\alpha P_\beta$ is hermitian and idempotent, and using Theorem 13.14 it is a projection operator. The space it projects into is $M_{\alpha\beta} = M_\alpha \cap M_\beta$. Two such spaces $M_{\alpha\beta}$ and $M_{\alpha'\beta'}$ are clearly orthogonal unless $\alpha = \alpha'$ and $\beta = \beta'$. Choose an orthonormal basis for each $M_{\alpha\beta}$. The collection of these vectors is a complete o.n. set consisting entirely of common eigenvectors to A and B. For, if $|\phi\rangle \neq 0$ is any non-zero vector orthogonal to all $M_{\alpha\beta}$, then $P_\alpha P_\beta|\phi\rangle \neq 0$ for all α, β. Since A is complete this implies $P_\beta|\phi\rangle = 0$ for all eigenvalues β of B, and since B is complete we must have $|\phi\rangle = 0$. ∎

Example 14.1 Consider spin $\frac{1}{2}$ electrons in a Stern–Gerlach device for measuring spin in the z-direction. Let σ_z be the operator for the observable 'spin in the z-direction'. It can only take on two values – up or down. This results in two eigenvalues ± 1, and the eigenvectors are written

$$\sigma_z|+z\rangle = |+z\rangle, \qquad \sigma_z|-z\rangle = -|-z\rangle.$$

Thus

$$\sigma_z = |+z\rangle\langle +z| - |-z\rangle\langle -z|, \qquad I = |+z\rangle\langle +z| + |-z\rangle\langle -z|,$$

and setting $|e_1\rangle = |+z\rangle$, $|e_2\rangle = |-z\rangle$ results in the matrix components

$$(\sigma_z)_{ij} = \langle e_1|\sigma_z|e_j\rangle = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{14.9}$$

Every state of the system can be written

$$|\psi\rangle = \psi_1|{+z}\rangle + \psi_2|{-z}\rangle \quad \text{where} \quad \psi_i = \langle e_i \,|\, \psi\rangle.$$

The operator $\sigma_{\mathbf{n}}$ representing spin in an arbitrary direction

$$\mathbf{n} = \cos\theta\,\mathbf{e}_z + \sin\theta\cos\phi\,\mathbf{e}_x + \sin\theta\sin\phi\,\mathbf{e}_y$$

has expectation values in different directions given by the classical values

$$\begin{aligned}\langle +z|\sigma_{\mathbf{n}}|+z\rangle &= \cos\theta,\\ \langle +x|\sigma_{\mathbf{n}}|+x\rangle &= \sin\theta\cos\phi,\\ \langle +y|\sigma_{\mathbf{n}}|+y\rangle &= \sin\theta\sin\phi,\end{aligned}$$

where $|\pm x\rangle$ refers to the pure states in the x direction, $\sigma_x|\pm x\rangle = \pm|\pm x\rangle$, etc.

Since $\sigma_{\mathbf{n}}$ is hermitian with eigenvalues $\lambda_i = \pm 1$ its matrix with respect to any o.n. basis has the form

$$(\sigma_{\mathbf{n}})_{ij} = \begin{pmatrix}\alpha & \beta\\ \overline{\beta} & \delta\end{pmatrix}$$

where

$$\alpha + \delta = \lambda_1 + \lambda_2 = 0, \qquad \alpha\delta - \beta\overline{\beta} = \lambda_1\lambda_2 = -1.$$

Hence $\delta = -\alpha$ and $\alpha^2 = 1 - |\beta|^2$. The expectation value of $\sigma_{\mathbf{n}}$ in the $|+z\rangle$ state is given by

$$\langle +z|\sigma_{\mathbf{n}}|+z\rangle = (\sigma_{\mathbf{n}})_{11} = \alpha = \cos\theta$$

so that $\beta = \sin\theta\,\mathrm{e}^{-ic}$ where c is a real number. For $\mathbf{n} = \mathbf{e}_x$ and $\mathbf{n} = \mathbf{e}_y$ we have $\cos\theta = 0$,

$$\sigma_x = \begin{pmatrix}0 & \mathrm{e}^{-ia}\\ \mathrm{e}^{ia} & 0\end{pmatrix}, \qquad \sigma_y = \begin{pmatrix}0 & \mathrm{e}^{-ib}\\ \mathrm{e}^{ib} & 0\end{pmatrix}. \tag{14.10}$$

The states $|\pm x\rangle$ and $|\pm y\rangle$ are the eigenstates of σ_x and σ_y with normalized components

$$|\pm x\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}\mathrm{e}^{-ia}\\ \pm 1\end{pmatrix}, \qquad |\pm y\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}\mathrm{e}^{-ib}\\ \pm 1\end{pmatrix}$$

and as the expection values of σ_x in the orthogonal states $|\pm y\rangle$ vanish,

$$\langle +y|\sigma_x|+y\rangle = \tfrac{1}{2}\big(\mathrm{e}^{i(b-a)} + \mathrm{e}^{i(a-b)}\big) = \cos(b-a) = 0.$$

Hence $b = a + \pi/2$. Applying the unitary operator U

$$U = \begin{pmatrix}\mathrm{e}^{ia} & 0\\ 0 & 1\end{pmatrix}$$

results in $a = 0$, and the spin operators are given by the *Pauli representation*

$$\sigma_x = \sigma_1 = \begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}, \qquad \sigma_y = \sigma_2 = \begin{pmatrix}0 & -i\\ i & 0\end{pmatrix}, \qquad \sigma_z = \sigma_3 = \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}. \tag{14.11}$$

For a spin operator in an arbitrary direction the expectation values are $\langle +x|\sigma_{\mathbf{n}}|+x\rangle = \sin\theta\cos\phi$, etc., from which it is straightforward to verify that

$$\sigma_{\mathbf{n}} = \begin{pmatrix} \cos\theta & \sin\theta e^{-i\phi} \\ \sin\theta e^{i\phi} & -\cos\theta \end{pmatrix} = \sin\theta\cos\phi\sigma_x + \sin\theta\sin\phi\sigma_y + \cos\theta\sigma_z.$$

Exercise: Find the eigenstates $|+\mathbf{n}\rangle$ and $|-\mathbf{n}\rangle$ of $\sigma_{\mathbf{n}}$.

Unbounded operators in quantum mechanics

An important part of the framework of quantum mechanics is the *correspondence principle*, which asserts that to every classical dynamical variable there corresponds a quantum mechanical observable. This is at best a sort of guide – for example, as there is no natural way of defining general functions $f(Q, P)$ for a pair of non-commuting operators such as Q and P, it is not clear what operators correspond to generalized position and momentum in classical canonical coordinates. For rectangular cartesian coordinates x, y, z and momenta $p_x = m\dot{x}$, etc. experience has taught that the Hilbert space of states corresponding to a one-dimensional dynamical system is $\mathcal{H} = L^2(\mathbb{R})$, and the position and momentum operators are given by

$$Q\psi(x) = x\psi(x) \quad \text{and} \quad P\psi(x) = -i\hbar\frac{d\psi}{dx}.$$

These operators are unbounded operators and have been discussed in Examples 13.17, 13.19 and 13.20 of Chapter 13.

As these operators are not defined on all of $\mathcal{H}$ it is most common to take domains

$$D_Q = \Big\{\psi(x) \in L^2(\mathbb{R}) \;\Big|\; \int_{-\infty}^{\infty} x^2|\psi(x)|^2\,dx < \infty\Big\},$$
$$D_P = \Big\{\psi(x) \in L^2(\mathbb{R}) \;\Big|\; \psi(x) \text{ is differentiable and } \int_{-\infty}^{\infty}\Big\|\frac{d\psi(x)}{dx}\Big\|^2 dx < \infty\Big\}.$$

These domains are dense in $L^2(\mathbb{R})$ since the basis of functions $\phi_n(x)$ constructed from hermite polynomials in Example 13.7 (see Eq. (13.6)) belong to both of them. As shown in Example 13.19 the operator (Q, D_Q) is self-adjoint, but (P, D_P) is a symmetric operator that is not self-adjoint (see Example 13.20). To make it self-adjoint it must be extended to the domain of absolutely continuous functions.

Example 14.2 The position operator Q has no eigenvalues and eigenfunctions in $L^2(\mathbb{R})$ (see Example 13.14). For the momentum operator P the eigenvalue equation reads

$$\frac{d\psi}{dx} = i\lambda\psi(x) \implies \psi(x) = e^{i\lambda x},$$

and even when λ is a real number the function $\psi(x)$ does not belong to D_P,

$$\int_{-\infty}^{\infty} |e^{i\lambda x}|^2\,dx = \int_{-\infty}^{\infty} 1\,dx = \infty.$$

For each real number k set $\varepsilon_k(x) = e^{ikx}$, and

$$\langle \varepsilon_k \,|\, \psi \rangle = \int_{-\infty}^{\infty} e^{-ikx}\, \psi(x)\, dx$$

is a linear functional on $L^2(\mathbb{R})$ – in fact, it is the Fourier transform of Section 12.3. This linear functional can be thought of as a tempered distribution $\langle \varepsilon_k|$ on the space of test functions of rapid decrease D_P. It is a bra that corresponds to no ket vector $|\varepsilon_k\rangle$ (this does not violate the Riesz representation theorem 13.10 since the domain D_P is not a closed subspace of $L^2(\mathbb{R})$). In quantum theory it is common to write equations that may be interpreted as

$$\langle \varepsilon_k|P = k\langle \varepsilon_k|,$$

which hold in the distributional sense,

$$\langle \varepsilon_k|P|\psi\rangle = k\langle \varepsilon_k \,|\, \psi\rangle \qquad \text{for all } \psi \in D_P.$$

Its integral version holds if we permit integration by parts, as for distributions,

$$\int_{-\infty}^{\infty} e^{-ikx}\Big(-i\frac{d\psi(x)}{dx}\Big)dx = \int_{-\infty}^{\infty} i\frac{de^{-ikx}}{dx}\psi(x)\, dx = k\int_{-\infty}^{\infty} e^{-ikx}\psi(x)\, dx.$$

Similarly, for each real a define the linear functional $\langle \delta_a|$ by

$$\langle \delta_a \,|\, \psi\rangle = \psi(a)$$

for all kets $|\psi\rangle \equiv \psi(x) \in D_Q$. These too can be thought of as distributions on a set of test functions of rapid decrease. They behave as 'eigenbras' of the position operator Q,

$$\langle \delta_a|Q = a\langle \delta_a|$$

since

$$\begin{aligned}\langle \delta_a|Q|\psi\rangle &= \langle \delta_a \,|\, x\psi(x)\rangle \\ &= a\psi(a) = a\langle \delta_a \,|\, \psi\rangle\end{aligned}$$

for all $|\psi\rangle \in D_Q$. While there is no function $\delta_a(x)$ in $L^2(\mathbb{R})$ having $\langle \delta_a|$, the Dirac delta function $\delta_a(x) = \delta(x - a)$ can be thought of as fulfilling this role in a distributional sense (see Chapter 12).

For a self-adjoint operator A we may apply Theorem 13.25. Let E_λ be the spectral family of increasing projection operators defined by A, such that

$$A = \int_{-\infty}^{\infty} \lambda\, dE_\lambda \quad \text{and} \quad I = \int_{-\infty}^{\infty} dE_\lambda.$$

The latter relation follows from

$$1 = \langle \psi \,|\, \psi\rangle = \int_{-\infty}^{\infty} d\langle \psi \,|\, E_\lambda \psi\rangle \tag{14.12}$$

for all $\psi \in D_A$.

Exercise: Prove Eq. (14.12).

If S is any measurable subset of $\mathbb{R}$ then the probability of the measured value of A lying in S, when the system is in a state $|\psi\rangle$, is given by

$$P_S(A) = \int_S \mathrm{d}\langle\psi \,|\, E_\lambda \psi\rangle.$$

The expectation value and RMS deviation are given by

$$\langle A\rangle = \int_{-\infty}^{\infty} \lambda \, \mathrm{d}\langle\psi \,|\, E_\lambda \psi\rangle$$

and

$$(\Delta A)^2 = \int_{-\infty}^{\infty} (\lambda - \langle A\rangle)^2 \mathrm{d}\langle\psi \,|\, E_\lambda \psi\rangle.$$

Example 14.3 The spectral family for the position operator is defined as the 'cut-off' operators

$$(E_\lambda \psi)(x) = \begin{cases} \psi(x) & \text{if } x \le \lambda, \\ 0 & \text{if } x > \lambda. \end{cases}$$

Firstly, these operators are projection operators since they are idempotent ($E_\lambda^2 = E_\lambda$) and hermitian:

$$\langle\phi \,|\, E_\lambda \psi\rangle = \int_{-\infty}^{\lambda} \overline{\phi(x)}\psi(x)\,\mathrm{d}x = \langle E_\lambda \phi \,|\, \psi\rangle$$

for all $\phi, \psi \in L^2(\mathbb{R})$. They are an increasing family since the image spaces are clearly increasing, and $\mathbb{E}_{-\infty} = O$, $E_\infty = I$. The function $\lambda \mapsto \langle\phi \,|\, E_\lambda \psi\rangle$ is absolutely continuous, since

$$\langle\phi \,|\, E_\lambda \psi\rangle = \int_{-\infty}^{\lambda} \overline{\phi(x)}\psi(x)\,\mathrm{d}x$$

and has generalized derivative $'$ with respect to λ given by

$$\langle\phi \,|\, E_\lambda \psi\rangle' = \overline{\phi(\lambda)}\psi(\lambda).$$

Hence

$$\langle\phi \,|\, Q\psi\rangle = \int_{-\infty}^{\infty} \overline{\phi(\lambda)}\lambda\psi(\lambda)\,\mathrm{d}\lambda = \int_{-\infty}^{\infty} \lambda\,\mathrm{d}\langle\phi \,|\, E_\lambda \psi\rangle,$$

which is equivalent to the required spectral decomposition

$$Q = \int_{-\infty}^{\infty} \lambda \, \mathrm{d}E_\lambda.$$

Exercise: Show that for any $-\infty \le a < b \le \infty$ for the spectral family of the previous example

$$\int_a^b \mathrm{d}\langle\psi \,|\, E_\lambda \psi\rangle = \int_a^b |\psi(x)|^2 \mathrm{d}x.$$

Problems

Problem 14.1 Verify for each direction

$$\mathbf{n} = \sin\theta\cos\phi\mathbf{e}_x + \sin\theta\sin\phi\mathbf{e}_y + \cos\theta\mathbf{e}_z$$

the spin operator

$$\sigma_{\mathbf{n}} = \begin{pmatrix} \cos\theta & \sin\theta e^{-i\phi} \\ \sin\theta e^{i\phi} & -\cos\theta \end{pmatrix}$$

has eigenvalues ± 1. Show that up to phase, the eigenvectors can be expressed as

$$|+\mathbf{n}\rangle = \begin{pmatrix} \cos\frac{1}{2}\theta e^{-i\phi} \\ \sin\frac{1}{2}\theta \end{pmatrix}, \qquad |-\mathbf{n}\rangle = \begin{pmatrix} -\sin\frac{1}{2}\theta e^{-i\phi} \\ \cos\frac{1}{2}\theta \end{pmatrix}$$

and compute the expectation values for spin in the direction of the various axes

$$\langle\sigma_i\rangle_{\pm\mathbf{n}} = \langle\pm\mathbf{n}|\sigma_i|\pm\mathbf{n}\rangle.$$

For a beam of particles in a pure state $|+\mathbf{n}\rangle$ show that after a measurement of spin in the $+x$ direction the probability that the spin is in this direction is $\frac{1}{2}(1 + \sin\theta\cos\phi)$.

Problem 14.2 If $\mathbf{A}$ and $\mathbf{B}$ are vector observables that commute with the Pauli spin matrices, $[\sigma_i, A_j] = [\sigma_i, B_j] = 0$ (but $[A_i, B_j] \neq 0$ in general) show that

$$(\sigma\cdot\mathbf{A})(\sigma\cdot\mathbf{B}) = \mathbf{A}\cdot\mathbf{B} + i(\mathbf{A}\times\mathbf{B})\cdot\sigma$$

where $\sigma = (\sigma_1, \sigma_2, \sigma_3)$.

Problem 14.3 Prove the following commutator identities:

$$[A,[B,C]] + [B,[C,A]] + [C,[A,B]] = 0 \quad \text{(Jacobi identity)}$$
$$[AB, C] = A[B,C] + [A,C]B$$
$$[A, BC] = [A,B]C + B[A,C]$$

Problem 14.4 Using the identities of Problem 14.3 show the following identities:

$$\begin{aligned} &[Q^n, P] = ni\hbar Q^{n-1}, \\ &[Q, P^m] = mi\hbar P^{m-1}, \\ &[Q^n, P^2] = 2ni\hbar Q^{n-1}P + n(n-1)\hbar^2 Q^{n-2}, \\ &[L_m, Q_k] = i\hbar\epsilon_{mkj}Q_j, \quad [L_m, P_k] = i\hbar\epsilon_{mkj}P_j, \end{aligned}$$

where $L_m = \epsilon_{mij}Q_iP_j$ are the **angular momentum operators**.

Problem 14.5 Consider a one-dimensional wave packet

$$\psi(x,t) = \frac{1}{\sqrt{2\pi\hbar}}\int_{-\infty}^{\infty} e^{i(xp - p^2t/2m)/\hbar}\Psi(p)\,dp$$

where

$$\Psi(p) \propto e^{-(p-p_0)^2/2(\Delta p)^2}.$$

Show that $|\psi(x,t)|^2$ is a Gaussian normal distribution whose peak moves with velocity p/m and whose spread Δx increases with time, always satisfying $\Delta x\Delta p \geq \hbar/\sqrt{2}$.

If an electron ($m = 9 \times 10^{-28}$ g) is initially within an atomic radius $\Delta x_0 = 10^{-8}$ cm, after how long will Δx be (a) 2×10^{-8} cm, (b) the size of the solar system (about 10^{14} cm)?

14.2 Quantum dynamics

The discussion of Section 14.1 refers only to *quantum statics* – the essential framework in which quantum descriptions are to be set. The dynamical evolution of quantum systems is determined by a hermitian operator H, possibly but not usually a function of time $H = H(t)$, such that the time development of any state $|\psi(t)\rangle$ of the system is given by **Schrödinger's equation**

$$i\hbar \frac{\mathrm{d}}{\mathrm{d}t}|\psi\rangle = H|\psi\rangle. \tag{14.13}$$

The operator H is known as the **Hamiltonian** or **energy operator**. Equation (14.13) guarantees that all inner products are preserved for, taking the adjoint gives

$$-i\hbar \frac{\mathrm{d}}{\mathrm{d}t}\langle\psi| = \langle\psi|H^* = \langle\psi|H$$

and for any pair of states $|\psi\rangle$ and $|\phi\rangle$,

$$\begin{aligned} i\hbar \frac{\mathrm{d}}{\mathrm{d}t}\langle\psi\,|\,\phi\rangle &= i\hbar\Big(\Big(\frac{\mathrm{d}}{\mathrm{d}t}\langle\psi|\Big)|\phi\rangle + \langle\psi|\Big(\frac{\mathrm{d}}{\mathrm{d}t}|\phi\rangle\Big)\Big) \\ &= -\langle\psi|H|\phi\rangle + \langle\psi|H|\phi\rangle = 0. \end{aligned}$$

In particular the normalization $\|\psi(t)\| = \|\phi(t)\| = 1$ is preserved by Schrödinger's equation. Since

$$\langle\psi(t)\,|\,\phi(t)\rangle = \langle\psi(0)\,|\,\phi(0)\rangle$$

for all pairs of states, there exists a unitary operator $U(t)$ such that

$$|\psi(t)\rangle = U(t)|\psi(0)\rangle. \tag{14.14}$$

If H is independent of t then

$$U(t) = \mathrm{e}^{(-i/\hbar)Ht} \tag{14.15}$$

where the exponential function can be defined as in the comments prior to Theorem 13.26 at the end of Chapter 13. If H is a complete hermitian operator $H = \sum_n \lambda_n |\psi_n\rangle\langle\psi_n|$ then

$$\mathrm{e}^{(-i/\hbar)Ht} = \sum_n \mathrm{e}^{(-i/\hbar)\lambda t}|\psi_n\rangle\langle\psi_n|$$

and for a self-adjoint operator

$$H = \int_{-\infty}^{\infty} \lambda\,\mathrm{d}E_\lambda \Longrightarrow \mathrm{e}^{(-i/\hbar)Ht} = \int_{-\infty}^{\infty} \mathrm{e}^{(-i/\hbar)\lambda t}\,\mathrm{d}E_\lambda.$$

To prove Eq. (14.15) substitute (14.14) in Schrödinger's equation

$$i\hbar \frac{\mathrm{d}U}{\mathrm{d}t}|\psi(0)\rangle = HU|\psi(0)\rangle,$$

and since $|\psi(0)\rangle$ is an arbitrary initial state vector,

$$i\hbar \frac{\mathrm{d}U}{\mathrm{d}t} = HU.$$

Setting $U(t) = \mathrm{e}^{(-i/\hbar)Ht} V(t)$ (always possible since the operator $\mathrm{e}^{(-i/\hbar)Ht}$ is invertible with inverse $\mathrm{e}^{(i/\hbar)Ht}$) we obtain

$$\mathrm{e}^{(-i/\hbar)Ht} HV(t) + i\hbar \mathrm{e}^{(-i/\hbar)Ht} \frac{\mathrm{d}V(t)}{\mathrm{d}t} = H\mathrm{e}^{(-i/\hbar)Ht} V(t).$$

As H and $\mathrm{e}^{(-i/\hbar)Ht}$ commute it follows that $V(t) = \text{const.} = V(0) = I$ since $U(0) = I$ on setting $t = 0$ in Eq. (14.14).

The Heisenberg picture

The above description of the evolution of a quantum mechanical system is called the **Schrödinger picture**. There is an equivalent version called the **Heisenberg picture** in which the states are treated as constant, but observables undergo a dynamic evolution. The idea is to perform a unitary transformation on $\mathcal{H}$, simultaneously on states and operators:

$$\begin{aligned} |\psi\rangle \mapsto |\psi'\rangle &= U^*|\psi\rangle \qquad (|\psi\rangle = U|\psi'\rangle) \\ A \mapsto A' &= U^* A U, \end{aligned}$$

where U is given by (14.15). This transformation has the effect of bringing every solution of Schrödinger's equation to rest, for if $|\psi(t)\rangle$ is a solution of Eq. (14.13) then

$$|\psi'\rangle = U^*|\psi(t)\rangle = U^*U|\psi(0)\rangle = |\psi(0)\rangle \implies \frac{\mathrm{d}}{\mathrm{d}t}|\psi'(t)\rangle = 0.$$

It preserves all matrix elements, and in particular all expectation values:

$$\langle A'\rangle_{\psi'} = \langle\psi'|A'|\psi'\rangle = \langle\psi|UU^*AUU^*|\psi\rangle = \langle\psi|A|\psi\rangle = \langle A\rangle_{\psi}.$$

Thus the states and observables are physically equivalent in the two pictures.

We derive a dynamical equation for the Heisenberg operator A':

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}A' &= \frac{\mathrm{d}}{\mathrm{d}t}(U^*AU) \\ &= \frac{\mathrm{d}U^*}{\mathrm{d}t}AU + U^*\frac{\mathrm{d}A}{\mathrm{d}t}U + U^*A\frac{\mathrm{d}U}{\mathrm{d}t} \\ &= \frac{1}{-i\hbar}U^*HAU + U^*\frac{\mathrm{d}A}{\mathrm{d}t}U + \frac{1}{i\hbar}U^*AHU \end{aligned}$$

since, by Eq. (14.15),

$$\frac{\mathrm{d}U}{\mathrm{d}t} = \frac{1}{i\hbar}HU, \qquad \frac{\mathrm{d}U^*}{\mathrm{d}t} = \frac{1}{-i\hbar}U^*H^* = \frac{1}{-i\hbar}U^*H.$$

Hence

$$\frac{\mathrm{d}}{\mathrm{d}t}A' = \frac{1}{i\hbar}[A', H'] + \frac{\partial A'}{\partial t} \tag{14.16}$$

where

$$\frac{\partial A'}{\partial t} = U^* \frac{\mathrm{d}A}{\mathrm{d}t} U = \Big(\frac{\mathrm{d}A}{\mathrm{d}t}\Big)'.$$

The motivation for this identification is the following: if $|\psi_i\rangle$ is a rest basis of $\mathcal{H}$ in the Schrödinger picture, so that $\mathrm{d}|\psi_i\rangle/\mathrm{d}t = 0$, and if $|\psi_i'\rangle = U^*|\psi_i\rangle$ is the 'moving basis' obtained from it, then

$$\langle\psi_i'|A'|\psi_j'\rangle = \langle\psi_i|A|\psi_j\rangle$$

and

$$\langle\psi_i'|\frac{\partial A'}{\partial t}|\psi_j'\rangle = \langle\psi_i|U\frac{\partial A'}{\partial t}U^*|\psi_j\rangle = \langle\psi_i|\frac{\mathrm{d}A}{\mathrm{d}t}|\psi_j\rangle = \frac{\mathrm{d}}{\mathrm{d}t}\langle\psi_i|A|\psi_j\rangle.$$

Thus the matrix elements of $\partial A'/\partial t$ measure the *explicit* time rate of change of the matrix elements of the operator A in the Schrödinger representation.

If A is an operator having no explicit time dependence, so that $\partial A'/\partial t = 0$, then A' is a constant of the motion if and only if it commutes with the Hamiltonian, $[A', H'] = 0$,

$$\frac{\mathrm{d}A'}{\mathrm{d}t} = 0 \iff [A'H'] = [A, H] = 0.$$

In particular, since every operator commutes with itself, the Hamiltonian H' is a constant of the motion if and only if it is time independent, $\partial H'/\partial t = 0$.

Example 14.4 For an electron of charge e, mass m and spin $\frac{1}{2}$, notation as in Example 14.1, the Hamiltonian in a magnetic field $\mathbf{B}$ is given by

$$H = \frac{-e\hbar}{2mc}\sigma\cdot\mathbf{B}.$$

If $\mathbf{B}$ is parallel to the z-axis then $H = -(e\hbar/2mc)\sigma_z B$ and setting $|\psi\rangle = \psi_1(t)|{+z}\rangle + \psi_2(t)|{-z}\rangle$, Schrödinger's equation (14.13) can be written as the two differential equations

$$i\hbar\,\mathrm{d}\dot{\psi}_1 = -\frac{e\hbar}{2mc}B\psi_1, \qquad i\hbar\,\mathrm{d}\dot{\psi}_2 = \frac{e\hbar}{2mc}B\psi_2$$

with solutions

$$\psi_1(t) = \psi_{10}\mathrm{e}^{i(\omega/2)t}, \qquad \psi_2(t) = \psi_{20}\mathrm{e}^{-i(\omega/2)t},$$

where $\omega = eB/mc$. Substituting in the expectation values

$$\langle\psi(t)|\sigma|\psi(t)\rangle = \begin{pmatrix}\sin\theta\cos\phi(t)\\ \sin\theta\sin\phi(t)\\ \cos\theta(t)\end{pmatrix}$$

results in $\theta(t) = \theta_0 = \text{const.}$ and

$$\cos\phi(t) = \frac{\mathrm{e}^{i(\phi_0-\omega t)} + \mathrm{e}^{-i(\phi_0-\omega t)}}{2} = \cos(\phi_0 - \omega t).$$

Hence $\phi(t) = \phi_0 - \omega t$, and the motion is a precession with angular velocity ω about the direction of the magnetic field.

In the Heisenberg picture, set $\sigma_x = \sigma_x(t)$, etc., where $\sigma_x(0) = \sigma_1, \sigma_y(0) = \sigma_2, \sigma_z(0) = \sigma_3$ are the Pauli values, (14.11). From the commutation relations

$$[\sigma_1, \sigma_2] = 2i\sigma_3, \qquad [\sigma_2, \sigma_3] = 2i\sigma_1, \qquad [\sigma_3, \sigma_1] = 2i\sigma_2 \tag{14.17}$$

and $\sigma_x = U^*\sigma_1 U$, etc., it follows that

$$[\sigma_x, \sigma_y] = 2i\sigma_z, \text{ etc.}$$

Heisenberg equations of motion are

$$\begin{aligned}\dot{\sigma}_x &= \frac{1}{i\hbar}[\sigma_x, H] = \omega\sigma_y,\\ \dot{\sigma}_y &= -\omega\sigma_x,\\ \dot{\sigma}_z &= 0.\end{aligned}$$

Hence $\ddot{\sigma}_x = \omega\dot{\sigma}_y = -\omega^2\sigma_x$ and the solution of Heisenberg's equation is

$$\sigma_x = \mathsf{A}e^{i\omega t} + \mathsf{B}e^{-i\omega t}, \qquad \sigma_y = \frac{1}{\omega}\dot{\sigma}_x.$$

The 2×2 matrices A, B are evaluated by initial values at $t = 0$, resulting in

$$\begin{aligned}\sigma_x(t) &= \cos\omega t\, \sigma_1 + \sin\omega t\, \sigma_2,\\ \sigma_y(t) &= -\sin\omega t\, \sigma_1 + \cos\omega t\, \sigma_2\\ \sigma_z &= \sigma_3 = \text{const.}\end{aligned}$$

Correspondence with classical mechanics and wave mechanics

For readers familiar with Hamiltonian mechanics (see Section 16.5), the following correspondence can be set up between classical and quantum mechanics:

	Quantum mechanics	*Classical mechanics*
State space	Hilbert space $\mathcal{H}$	Phase space Γ
States	Normalized kets $\lvert\psi\rangle \in \mathcal{H}$	Points $(q_i, p_j) \in \Gamma$
Observables	Self-adjoint operators in $\mathcal{H}$; multiple values $\lambda_b i$ in each state $\lvert\psi\rangle$ with probability $P = \langle\psi_i \mid \psi\rangle^2$	Real functions $f(q_i, p_j)$ on phase space; one value for each state
Commutators	Bracket commutators $[A, B]$	Poisson brackets (f, g)
Dynamics	1. Schrödinger picture $i\hbar\frac{\mathrm{d}}{\mathrm{d}t}\lvert\psi\rangle = H\lvert\psi\rangle$ 2. Heisenberg picture $\dot{A} = \frac{1}{i\hbar}[A, H] + \frac{\partial A}{\partial t}$	1. Hamilton's equations $\dot{q}_i = \frac{\partial H}{\partial p_i},\ \dot{p}_i = -\frac{\partial H}{\partial q_i}$ 2. Poisson bracket form $\dot{f} = (f, H) + \frac{\partial f}{\partial t}$

If f and g are classical observables with quantum mechanical equivalents F and G then, from Heisenberg's equation of motion, the proposal is that the commutator $[F, G]$

corresponds to the $i\hbar$ times the Poisson bracket,

$$[F, G] \longleftrightarrow i\hbar(f, g) = i\hbar\Big(\frac{\partial f}{\partial q_i}\frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial q_i}\frac{\partial g}{\partial q_i}\Big).$$

For example if Q_i are position operators representing classical variable q_i and $P_i = -i\hbar\partial/\partial q_i$ the momentum operators, then the classical canonical commutation relations imply

$$\begin{aligned}(q_i, q_j) = 0 &\Longrightarrow [Q_i, Q_j] = 0,\\ (p_i, p_j) = 0 &\Longrightarrow [P_i, P_j] = 0,\\ (q_i, p_j) = \delta_{ij} &\Longrightarrow [Q_i, P_j] = i\hbar\delta_{ij} I.\end{aligned}$$

Generalizing from the one-dimensional case, we assume $\mathcal{H}$ is the set of differentiable functions in $L^2(\mathbb{R}^n)$ such that $x_i\psi(x_1, \dots, x_n)$ belongs to $L^2(\mathbb{R}^n)$ for each x_i. The above commutation relations are satisfied by the standard operators:

$$Q_i\psi(x_1, \dots, x_n) = x_i\psi(x_1, \dots, x_n), \qquad P_i\psi(x_1, \dots, x_n) = -i\hbar\frac{\partial\psi}{\partial x_i}.$$

For a particle in a potential $V(x, y, z)$ the Hamiltonian is $H = \mathbf{p}^2/2m + V(x, y, z)$, which corresponds to the quantum mechanical Schrödinger equation

$$i\hbar\frac{\partial\psi}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\psi + V(\mathbf{r})\psi. \tag{14.18}$$

Exercise: Show that the probability density $P(\mathbf{r}, t) = \psi\overline{\psi}$ satisfies the conservation equation

$$\frac{\partial P}{\partial t} = -\nabla\mathbf{J} \quad \text{where} \quad \mathbf{J} = \frac{i\hbar}{2m}(\psi\nabla\overline{\psi} - \overline{\psi}\nabla\psi).$$

A trial solution of Eq. (14.18) by separation of variables, $\psi = T(t)\phi(\mathbf{x})$, results in

$$\psi = \mathrm{e}^{-i\omega t}\phi(\mathbf{x})$$

where $\phi(\mathbf{x})$ satisfies the **time-independent Schrödinger equation**

$$H\phi(\mathbf{x}) = -\frac{\hbar^2}{2m}\nabla^2\phi(\mathbf{x}) + V(\mathbf{r})\phi(\mathbf{x}) = E\phi(\mathbf{x})$$

where E is given by Planck's relation, $E = \hbar\omega = h\nu$. From its classical analogue, the eigenvalue E of the Hamiltonian is interpreted as the energy of the system, and if the Hamiltonian is a complete operator with discrete spectrum E_n then the general solution of the Schrödinger equation is given by

$$\psi(\mathbf{x}, t) = \sum_n c_n\phi_n(\mathbf{x})\mathrm{e}^{-iE_n t/\hbar}$$

where

$$H\phi_n(\mathbf{x}) = E_n\phi_n(\mathbf{x}).$$

Harmonic oscillator

The classical one-dimensional harmonic oscillator has Hamiltonian

$$H_{\mathrm{cl}} = \frac{1}{2m}p^2 + \frac{k}{2}q^2.$$

Its quantum mechanical equivalent should have energy operator

$$H = \frac{1}{2m}P^2 + \frac{k}{2}Q^2 \tag{14.19}$$

where

$$P = -i\hbar\frac{\mathrm{d}}{\mathrm{d}x}, \qquad [Q, P] = i\hbar I.$$

Set

$$A = \frac{1}{\sqrt{\omega\hbar}}\Big(\frac{1}{\sqrt{2m}}P + i\sqrt{\frac{k}{2}}Q\Big)$$

where $\omega = \sqrt{k/m}$ and we find

$$H = \omega\hbar(N + \tfrac{1}{2}I) \tag{14.20}$$

where N is the self-adjoint operator $N = AA^* = N^*$.

It is not hard to show that

$$[A, A^*] = AA^* - A^*A = -I, \tag{14.21}$$

and from the identities in Problem 14.3 it follows that

$$[N, A] = [AA^*, A] = A[A^*, A] + [A, A]A^* = A, \tag{14.22}$$

$$[N, A^*] = [A, N^*]^* = -[N, A]^* = -A^*. \tag{14.23}$$

All eigenvalues of N are non-negative, $n \geq 0$, for if $N|\psi_n\rangle = n|\psi_n\rangle$ then

$$0 \leq \|A^*\psi_n\|^2 = \langle\psi_n|AA^*|\psi_n\rangle = \langle\psi_n|N|\psi_n\rangle = n\langle\psi_n\,|\,\psi_n\rangle. \tag{14.24}$$

Let $n_0 \geq 0$ be the lowest eigenvalue. Using (14.23), the state $A^*|\psi_n\rangle$ is an eigenstate of N with eigenvalue $(n-1)$

$$NA^*|\psi_n\rangle = (A^*N - A^*)|\psi_n\rangle = (n-1)A^*|\psi_n\rangle.$$

Hence $A^*|\psi_{n_0}\rangle = 0$, else $n_0 - 1$ would be an eigenvalue, contradicting n_0 being lowest, and setting $n = n_0$ in Eq. (14.24) gives $n_0 = 0$. Furthermore, if n is an eigenvalue then $A|\psi_n\rangle \neq 0$ is an eigenstate with eigenvalue $(n+1)$ for, by Eqs. (14.22) and (14.21)

$$NA|\psi_n\rangle = (AN + A)|\psi_n\rangle = (n+1)|\psi_n\rangle,$$

$$\|A|\psi_n\rangle\|^2 = \langle\psi_n|A^*A|\psi_n\rangle = \langle\psi_n|AA^* + I|\psi_n\rangle = (n+1)\langle\psi_n\,|\,\psi_n\rangle > 0.$$

The eigenvalues of N are therefore $n = 0, 1, 2, 3, \ldots$ and the eigenvalues of H are $\frac{1}{2}\hbar\omega$, $\frac{3}{2}\hbar\omega, \ldots, (n + \frac{1}{2})\hbar\omega, \ldots$

Angular momentum

A similar analysis can be used to find the eigenvalues of the angular momentum operators $L_i = \epsilon_{ijk} Q_j P_k$. Using the identities in Problems 14.3 and 14.4 it is straightforward to derive the commutation relations of angular momentum

$$[L_i, L_j] = i\hbar\epsilon_{ijk} L_k \tag{14.25}$$

and

$$[L^2, L_i] = 0 \tag{14.26}$$

where $L^2 = (L_1)^2 + (L_2)^2 + (L_3)^2$ is the **total angular momentum**.

Exercise: Prove the identities (14.25) and (14.26).

Any set of three operators L_1, L_2, L_3 satisfying the commutation relations (14.25) are said to be **angular momentum operators**. If they are of the form $L_i = \epsilon_{ijk} Q_j P_k$ then we term them **orbital angular momentum**, else they are called **spin angular momentum**, or a combination thereof. If we set $J_i = L_1/\hbar$ and $J^2 = (J_1)^2 + (J_2)^2 + (J_3)^2$ then

$$[J_i, J_j] = i\epsilon_{ijk} J_k, \qquad [J^2, J_i] = 0$$

and since J_3 and J^2 commute there exist, by Theorem 14.2, a common set of eigenvectors $|j^2 m\rangle$ such that

$$J^2 |j^2 m\rangle = j^2 |j^2 m\rangle, \qquad J_3 |j^2 m\rangle = m |j^2 m\rangle.$$

Thus

$$\langle j^2 m | J^2 | j^2 m\rangle = \sum_{k=1}^{3} \langle j^2 m | (J_k)^2 | j^2 m\rangle = \sum_{k=1}^{3} \| J_k | j^2 m\rangle \|^2 \geq m^2 \| | j^2 m\rangle \|^2.$$

Since the left-hand side is equal to $j^2 \| |j^2 m\rangle \|^2$ we have $j^2 \geq m^2 \geq 0$, and there is an upper and lower bound to the eigenvalue m for any fixed j^2. Let this upper bound be l.

If we set $J_\pm = J_1 \pm iJ_2$, then it is simple to show the identities

$$[J_3, J_\pm] = \pm J_\pm, \qquad J_\pm J_\mp = J^2 - (J_3)^2 \pm J_3. \tag{14.27}$$

Exercise: Prove the identities (14.27).

Hence $J_\pm$ are raising and lowering operators for the eigenvalues of J_3,

$$J_3 J_\pm |j^2 m\rangle = \big(J_\pm J_3 \pm J_\pm\big) |j^2 m\rangle = (m \pm 1) J_\pm |j^2 m\rangle$$

while they leave the eigenvalue of J^2 alone,

$$J^2 \big(J_\pm |j^2 m\rangle\big) = J_\pm J^2 |j^2 m\rangle = j^2 J_\pm J^2 |j^2 m\rangle.$$

Since l is the maximum possible value of m, we must have $J_\pm |j^2 l\rangle = 0$ and using the second identity of (14.27) we have

$$J_- J_+ |j^2 l\rangle = \big(J^2 - (J_3)^2 - J_3\big) |j^2 l\rangle = (j^2 - l^2 - l) |j^2 l\rangle = 0,$$

whence $j^2 = l(l+1)$. Since for each integer n, $(J_-)^n|j^2l\rangle$ is an eigenket of J_3 with eigenvalue $(l-n)$ and the eigenvalues of J_3 are bounded below, there exists an integer n such that $(J_-)^n|j^2l\rangle \neq 0$ and $(J_-)^{n+1}|j^2l\rangle = 0$. Using (14.27) we deduce

$$\begin{aligned} 0 = J_+J_-(J_-)^n|j^2l\rangle &= \big(J^2 - (J_3)^2 \pm J_3\big)(J_-)^n|j^2l\rangle \\ &= \big(j^2 - (l-n)^2 + (l-n)\big)\big(J_-\big)^n|j^2l\rangle \end{aligned}$$

so that

$$j^2 = (l-n)(l-n-1) = l^2 + l$$

from which it follows that $l = \frac{1}{2}n$. Thus the eigenvalues of total angular momentum L^2 are of the form $l(l+1)\hbar^2$ where l has integral or half integral values. The eigenspaces are $(2l+1)$-degenerate, and the simultaneous eigenstates of L_3 have eigenvalues $m\hbar$ where $-l \leq m \leq m$. For orbital angular momentum it turns out that the value of l is always integral, but spin eigenstates may have all possible eigenvalues, $l = 0, \frac{1}{2}, 1, \frac{3}{2}, \dots$ depending on the particle in question.

Problems

Problem 14.6 In the Heisenberg picture show that the time evolution of the expection value of an operator A is given by

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle A'\rangle_{\psi'} = \frac{1}{i\hbar}\langle [A', H']\rangle_{\psi'} + \langle \frac{\partial A'}{\partial t}\rangle_{\psi'}.$$

Convert this to an equation in the Schrödinger picture for the time evolution of $\langle A\rangle_\psi$.

Problem 14.7 For a particle of spin half in a magnetic field with Hamiltonian given in Example 14.4, show that in the Heisenberg picture

$$\begin{aligned} \langle \sigma_x(t)\rangle_{\mathbf{n}} &= \sin\theta\cos(\phi - \omega t), \\ \langle \sigma_y(t)\rangle_{\mathbf{n}} &= \sin\theta\sin(\phi - \omega t), \\ \langle \sigma_z(t)\rangle_{\mathbf{n}} &= \cos\theta. \end{aligned}$$

Problem 14.8 A particle of mass m is confined by an infinite potential barrier to remain within a box $0 \leq x, y, z \leq a$, so that the wave function vanishes on the boundary of the box. Show that the energy levels are

$$E = \frac{1}{2m}\frac{\pi^2\hbar^2}{a^2}\big(n_1^2 + n_2^2 + n_3^2\big),$$

where $n_1,\ n_2,\ n_3$ are positive integers, and calculate the stationary wave functions $\psi_E(\mathbf{x})$. Verify that the lowest energy state is non-degenerate, but the next highest is triply degenerate.

Problem 14.9 For a particle with Hamiltonian

$$H = \frac{\mathbf{p}^2}{2m} + V(\mathbf{x})$$

show from the equation of motion in the Heisenberg picture that

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle \mathbf{r}\cdot\mathbf{p}\rangle = \left\langle \frac{\mathbf{p}^2}{m}\right\rangle - \langle \mathbf{r}\cdot\nabla V\rangle.$$

This is called the *Virial theorem*. For stationary states, show that

$$2\langle T \rangle = \langle \mathbf{r} \cdot \nabla V \rangle$$

where T is the kinetic energy. If $V \propto r^n$ this reduces to the classical result $\langle 2T + nV \rangle = 0$.

Problem 14.10 Show that the nth normalized eigenstate of the harmonic oscillator is given by

$$|\psi_n\rangle = \frac{1}{(n!)^{1/2}} A^n |\psi_0\rangle.$$

Show from $A^* \psi_0 = 0$ that

$$\psi_0 = c\mathrm{e}^{-\sqrt{km}x^2/2\hbar} \quad \text{where} \quad c = \left(\frac{m\omega}{\pi\hbar}\right)^{1/4}$$

and the nth eigenfunction is

$$\psi_n(x) = \frac{i^n}{(2^n n!)^{1/2}} \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} \mathrm{e}^{-m\omega x^2/2\hbar} H_n\left(\sqrt{\frac{m\omega}{\hbar}}x\right),$$

where $H_n(y)$ is the nth hermite polynomial (see Example 13.7).

Problem 14.11 For the two-dimensional harmonic oscillator define operators A_1, A_2 such that

$$[A_i, A_j] = [A_i^*, A_j^*] = 0, \qquad [A_i, A_j^*] = -\delta_{ij}, \qquad H = \hbar\omega(2N + I)$$

where $i, j = 1, 2$ and N is the number operator

$$N = \tfrac{1}{2}(A_1 A_1^* + A_2 A_2^*).$$

Let J_1, J_2 and J_3 be the operators

$$J_1 = \tfrac{1}{2}(A_2 A_1^* + A_1 A_2^*), \qquad J_2 = \tfrac{1}{2}i(A_2 A_1^* - A_1 A_2^*), \qquad J_3 = \tfrac{1}{2}(A_1 A_1^* - A_2 A_2^*),$$

and show that:

(a) The J_i satisfy the angular momentum commutation relations $[J_1, J_2] = iJ_3$, etc.
(b) $J^2 = J_1^2 + J_2^2 + J_3^2 = N(N + 1)$.
(c) $[J^2, N] = 0, \quad [J_3, N] = 0$.

From the properties of angular momentum deduce the energy levels and their degeneracies for the two-dimensional harmonic oscillator.

Problem 14.12 Show that the eigenvalues of the three-dimensional harmonic oscillator have the form $(n + \frac{3}{2})\hbar\omega$ where n is a non-negative integer. Show that the degeneracy of the nth eigenvalue is $\frac{1}{2}(n^2 + 3n + 2)$. Find the corresponding eigenfunctions.

14.3 Symmetry transformations

Consider two observers O and O' related by a symmetry transformation such as a translation $\mathbf{x}' = \mathbf{x} - \mathbf{a}$, or rotation $\mathbf{x}' = \mathsf{A}\mathbf{x}$ where $\mathsf{A}\mathsf{A}^T = \mathsf{I}$, etc. For any state corresponding to a ray $[|\psi\rangle]$ according to O let O' assign the ray $[|\psi'\rangle]$, and for any observable assigned the self-adjoint operator A by O let O' assign the operator A'. Since the physical elements are determined by the modulus squared of the matrix elements between states (being the probability of transition between states) and the expectation values of observables, this correspondence is said to be a **symmetry transformation** if

$$|\langle \phi' \,|\, \psi' \rangle|^2 = |\langle \phi \,|\, \psi \rangle|^2 \tag{14.28}$$

$$\langle A' \rangle_{\psi'} = \langle \psi' | A' | \psi' \rangle = \langle A \rangle_{\psi} = \langle \psi | A | \psi \rangle \tag{14.29}$$

for all states $|\psi\rangle$ and observables A.

Theorem 14.3 (Wigner) *A ray correspondence that satisfies (14.28) for all rays is generated up to a phase by a transformation $U : \mathcal{H} \to \mathcal{H}$ that is either unitary or anti-unitary.*

A unitary transformation was defined in Chapter 13 as a linear transformation, $U(|\psi\rangle + \alpha|\phi\rangle) = U|\psi\rangle + \alpha U|\phi\rangle$, which preserves inner products

$$\langle U\psi \,|\, U\phi \rangle = \langle \psi \,|\, \phi \rangle \iff UU^* = U^*U = I.$$

A transformation V is said to be **antilinear** if

$$V(|\psi\rangle + \alpha|\phi\rangle) = V|\psi\rangle + \overline{\alpha} V|\phi\rangle.$$

The adjoint V^* is defined by

$$\langle V^*\psi \,|\, \phi \rangle = \overline{\langle \psi \,|\, V\phi \rangle} = \langle V\phi \,|\, \psi \rangle,$$

in order that it too will be antilinear (we pay no attention to domains here). An operator V is called **anti-unitary** if it is antilinear and $V^*V = VV^* = I$. In this case

$$\langle V\psi \,|\, V\phi \rangle = \overline{\langle \psi \,|\, \phi \rangle} = \langle \phi \,|\, \psi \rangle$$

for all vectors $|\psi\rangle$ and $|\phi\rangle$.

Outline proof of Theorem 14.3: We prove Wigner's theorem in the case of a two-dimensional Hilbert space. The full proof is along similar lines. If $|e_1\rangle,\ |e_2\rangle$ is an orthonormal basis then, up to phases, so is $|e'_1\rangle,\ |e'_2\rangle$,

$$\delta_{ij} = |\langle e_i \,|\, e_j \rangle|^2 = |\langle e'_i \,|\, e'_j \rangle|^2.$$

Let $|\psi\rangle = a_1|e_1\rangle + a_2|e_2\rangle$ be any unit vector, $\langle \psi \,|\, \psi \rangle = |a_1|^2 + |a_2|^2 = 1$. Set

$$|\psi'\rangle = a'_1|e'_1\rangle + a'_2|e'_2\rangle$$

and we have, from Eq. (14.28),

$$|a'_1|^2 = |\langle e'_1 \,|\, \psi' \rangle|^2 = |\langle e_1 \,|\, \psi \rangle|^2 = |a_1|^2,$$

and similarly $|a'_2|^2 = |a_2|^2$. Hence we can set real angles α, θ, φ, etc. such that

$$a_1 = \cos\alpha \mathrm{e}^{i\theta}, \quad a'_1 = \cos\alpha \mathrm{e}^{i\theta'},$$
$$a_2 = \sin\alpha \mathrm{e}^{i\varphi}, \quad a'_2 = \sin\alpha \mathrm{e}^{i\varphi'}.$$

Let $|\psi_1\rangle$ and $|\psi_2\rangle$ be an arbitrary pair of unit vectors,

$$|\psi_i\rangle = \cos\alpha_i \mathrm{e}^{i\theta_1}|e_1\rangle + \sin\alpha_i \mathrm{e}^{i\varphi_1}|e_2\rangle,$$

then $|\langle\psi_1'\,|\,\psi_2'\rangle|^2 = |\langle\psi_1\,|\,\psi_2\rangle|^2$ implies

$$\cos(\theta_2' - \varphi_2' - \theta_1' + \varphi_1') = \cos(\theta_2 - \varphi_2 - \theta_1 + \varphi_1).$$

Hence

$$\theta_2' - \varphi_2' - (\theta_1' - \varphi_1') = \pm\big(\theta_2 - \varphi_2 - (\theta_1 - \varphi_1)\big). \tag{14.30}$$

Define an angle δ by

$$\theta_1' - \varphi_1' = \delta \pm (\theta_1 - \varphi_1),$$

and it follows from (14.30) that

$$\theta_2' - \varphi_2' = \delta \pm (\theta_2 - \varphi_2).$$

Hence for an arbitrary vector $|\psi\rangle$,

$$\theta' - \varphi' = \delta \pm (\theta - \varphi).$$

For the $+$ sign this results in the transformation

$$|\psi\rangle \mapsto |\psi'\rangle = \mathrm{e}^{i(\varphi'-\varphi)}\big(a_1\mathrm{e}^{i\delta}|e_1'\rangle + a_2|e_2'\rangle\big)$$

while for the $-$ sign it is

$$|\psi\rangle \mapsto |\psi'\rangle = \mathrm{e}^{i(\varphi'+\varphi)}\big(\overline{a_1}\mathrm{e}^{i\delta}|e_1'\rangle + \overline{a_2}|e_2'\rangle\big).$$

These transformations are, up to a phase e^{if}, unitary and anti-unitary respectively. This is Wigner's theorem. ■

Exercise: Show that the phase $f = \varphi' \pm \varphi$ is independent of the state.

If U is a unitary transformation and A is an observable, then

$$\langle A'\rangle_{\psi'} = \langle\psi'|A'|\psi'\rangle = \langle\psi|U^*\mathrm{e}^{-if}A'\mathrm{e}^{if}U|\psi\rangle = \langle U^*A'U\rangle_\psi$$

and the requirement (14.29) implies that this holds for arbitrary vectors $|\psi\rangle$ if and only if

$$A' = UAU^*.$$

Performing two symmetries g and h in succession results in a symmetry transformation of $\mathcal{H}$ satisfying

$$U(g)U(h) = \mathrm{e}^{i\varphi}U(gh)$$

where the phase φ may depend on g and h. This is called a **projective** or **ray representation** of the group G on the Hilbert space $\mathcal{H}$. It is not in general possible to choose the phases such that all $\mathrm{e}^{i\varphi} = 1$, giving a genuine representation. For a continuous group (see Section 10.8), elements in the component of the identity G_0 must be unitary since they are connected continuously with the identity element, which is definitely unitary. Anti-unitary transformations can only correspond to group elements in components that are not continuously connected with the identity.

Infinitesimal generators

If G is a Lie group, the elements of which are unitary transformations characterized as in Section 6.5 by a set of real parameters $U = U(a_1, \dots, a_n)$ such that $U(0, \dots, 0) = I$, we define the **infinitesimal generators** by

$$S_j = -i\frac{\partial U}{\partial a_j}\bigg|_{\mathbf{a}=\mathbf{0}}. \tag{14.31}$$

These are self-adjoint operators since $UU^* = I$ implies that

$$0 = \frac{\partial U}{\partial a_j}\bigg|_{\mathbf{a}=\mathbf{0}} + \frac{\partial U^*}{\partial a_j}\bigg|_{\mathbf{a}=\mathbf{0}} = i\big(S_j - S_j^*\big).$$

Note that self-adjoint operators do not form a Lie algebra, since their commutator is not in general self-adjoint. However, as seen in Section 6.5, Problem 6.12, the operators iS_j do form a Lie algebra,

$$[iS_i, iS_j] = \sum_{k=1}^{n} C^k_{ij}\, iS_k.$$

Exercise: Show that an operator S satisfies $S^* = -S$ iff it is of the form $S = iA$ where A is self-adjoint. Show that the commutator product preserves this property.

Example 14.5 If S is a hermitian operator the set of unitary transformations $U(a) = \mathrm{e}^{iaS}$ where $-\infty < a < \infty$ is a one-parameter group of unitary transformations,

$$U(a)U^*(a) = \mathrm{e}^{iaS}\mathrm{e}^{-iaS} = I, \qquad U(a)U(b) = \mathrm{e}^{iaS}\mathrm{e}^{ibS} = \mathrm{e}^{i(a+b)S} = U(a+b).$$

Its infinitesimal generator is

$$-i\frac{\partial}{\partial a}\mathrm{e}^{iaS}\bigg|_{a=0} = S.$$

Example 14.6 Let O and O' be two observers related by a displacement of the origin through the vector $\mathbf{a}$. Let the state vectors be related by

$$|\psi'\rangle = T(\mathbf{a})|\psi\rangle.$$

If $\mathbf{Q}$ is the position operator then

$$\mathbf{q}' = \langle\mathbf{Q}\rangle_{\psi'} = \langle\psi'|\mathbf{Q}|\psi'\rangle = \mathbf{q} - \mathbf{a} = \langle\psi|\mathbf{Q}|\psi\rangle - \mathbf{a}\langle\psi|I|\psi\rangle,$$

whence

$$T^*(\mathbf{a})\mathbf{Q}T(\mathbf{a}) = \mathbf{Q} - \mathbf{a}I. \tag{14.32}$$

Taking the partial derivative with respect to a_i at $\mathbf{a} = \mathbf{0}$ we find

$$-iS_iQ_j + iQ_jS_i = -\delta_{ij} \quad \text{where} \quad S_j = -i\frac{\partial T}{\partial a_j}\bigg|_{\mathbf{a}=\mathbf{0}}.$$

Hence $[S_i, Q_j] = -i\delta_{ij}$, and we may expect

$$S_i = \frac{1}{\hbar}P_i$$

where P_i are the momentum operators

$$P_i = -i\hbar\frac{\partial}{\partial q_i}, \qquad [Q_i, P_j] = i\hbar\delta_{ij}.$$

This is consistent with $\mathbf{P}' = \mathbf{P}$, since

$$T^*(\mathbf{a})P_i T(\mathbf{a}) = P_i \implies [S_j, P_i] \propto [P_j, P_i] = 0.$$

To find the translation operators $T(\mathbf{a})$, we use the group property $T(\mathbf{a})T(\mathbf{b}) = T(\mathbf{a}+\mathbf{b})$, and take the derivative with respect to b_i at $\mathbf{b} = \mathbf{0}$:

$$iT(\mathbf{a})S_i = T(\mathbf{a})\frac{\partial T}{\partial b_i}\bigg|_{\mathbf{b}=\mathbf{0}} = \frac{\partial T}{\partial a_i}(\mathbf{a}).$$

The solution of this operator equation may be written

$$\begin{aligned} T(\mathbf{a}) &= \mathrm{e}^{i\sum_i a_i S_i} = \mathrm{e}^{i\mathbf{a}\cdot\mathbf{S}} \\ &= \mathrm{e}^{i\mathbf{a}\cdot\mathbf{P}/\hbar} = \mathrm{e}^{ia_1P_1/\hbar}\mathrm{e}^{ia_2P_2/\hbar}\mathrm{e}^{ia_3P_3/\hbar} \end{aligned}$$

since the P_i commute with each other. It is left as an exercise (Problem 14.14) to verify Eq. (14.32).

Example 14.7 Two observers related by a rotation through an angle θ about the z-axis

$$q_1' = q_1\cos\theta + q_2\sin\theta, \qquad q_2' = -q_1\sin\theta + q_2\cos\theta, \qquad q_3' = q_3$$

are related by a unitary operator $R(\theta)$ such that

$$|\psi'\rangle = R(\theta)|\psi\rangle, \qquad R^*(\theta)R(\theta) = I.$$

In order to arrive at the correct transformation of expectation values we require that

$$R^*(\theta)Q_1R(\theta) = Q_1\cos\theta + Q_2\sin\theta, \tag{14.33}$$
$$R^*(\theta)Q_2R(\theta) = -Q_1\sin\theta + Q_2\cos\theta, \tag{14.34}$$
$$R^*(\theta)Q_3R(\theta) = Q_3. \tag{14.35}$$

Setting

$$J = -i\frac{\mathrm{d}R}{\mathrm{d}\theta}\bigg|_{\theta=0}, \qquad J^* = J$$

we find on taking derivatives at $\theta = 0$ of Eqs. (14.33)–(14.35)

$$[J, Q_1] = iQ_2, \qquad [J, Q_2] = -iQ_1, \qquad [J, Q_3] = 0.$$

A solution is the z-component of angular momentum,

$$J = \frac{1}{\hbar}L_3 = \frac{1}{\hbar}\big(Q_1P_2 - Q_2P_1\big),$$

since $[J, Q_1] = iQ_2$, etc. (see Problem 14.4).

It is again easy to verify the group property $R(\theta_1)R(\theta_2) = R(\theta_1+\theta_2)$ and as in the translational example above,

$$iR(\theta)J = \frac{\mathrm{d}R}{\mathrm{d}\theta} \implies R\theta = \mathrm{e}^{i\theta J} = \mathrm{e}^{i\theta L_3/\hbar}.$$

It is again left as an exercise to show that this operator satisfies Eqs. (14.33)–(14.35). For a rotation of magnitude θ about an axis $\mathbf{n}$ the rotation operator is

$$R(\mathbf{n}, \theta) = \mathrm{e}^{i\theta \mathbf{n}\cdot\mathbf{L}/\hbar}$$

where $\mathbf{L}$ is the angular momentum operator having components $L_i = \epsilon_{ijk} Q_j P_k$. Since these operators do not commute, satisfying the commutation relations (14.25), we have in general

$$R(\mathbf{n}, \theta) \neq \mathrm{e}^{i\theta_1 L_1/\hbar} \mathrm{e}^{i\theta_2 L_2/\hbar} \mathrm{e}^{i\theta_3 L_3/\hbar}.$$

Exercise: Show that the transformation of momentum components under a rotation with infinitesimal generator J is

$$[J, P_1] = iP_2, \qquad [J, P_2] = -iP_1, \qquad [J, P_3] = 0.$$

Example 14.8 Under a time translation $t' = t - \tau$, we have $|\psi'(t')\rangle = |\psi(t)\rangle$, so that

$$|\psi'(t)\rangle = |\psi(t + \tau)\rangle = T(\tau)|\psi(t)\rangle.$$

Hence, by Schrödinger's equation

$$iS|\psi(t)\rangle = \frac{\partial T}{\partial \tau}\Big|_{\tau=0} |\psi(t)\rangle = \frac{\mathrm{d}}{\mathrm{d}t}|\psi(t)\rangle = \frac{1}{i\hbar} H|\psi(t)\rangle.$$

The infinitesimal generator of the time translation is essentially the Hamiltonian, $S = -H/\hbar$. If the Hamiltonian H is time-independent,

$$T(\tau) = \mathrm{e}^{-iH\tau/\hbar}. \tag{14.36}$$

Conserved quantities

Under a time-dependent unitary transformation

$$|\psi'\rangle = U(t)|\psi\rangle.$$

Schrödinger's equation (14.13) results in

$$\begin{aligned} i\hbar \frac{\mathrm{d}}{\mathrm{d}t}|\psi'\rangle &= i\hbar \frac{\partial U}{\partial t}|\psi\rangle + i\hbar U(t) \frac{\mathrm{d}}{\mathrm{d}t}|\psi\rangle \\ &= H'|\psi'\rangle \end{aligned}$$

where

$$H' = UHU^* + i\hbar \frac{\partial U}{\partial t} U^*. \tag{14.37}$$

Exercise: Show that under an anti-unitary transformation

$$H' = -UHU^* + i\hbar \frac{\partial U}{\partial t} U^*.$$

U is called a **Hamiltonian symmetry** if $H' = H$. Then, multiplying Eq. (14.37) by U on the right gives

$$[U, H] + i\hbar \frac{\partial U}{\partial t} = 0. \tag{14.38}$$

If U is independent of time then U commutes with the Hamiltonian, $[U, H] = 0$.

If G is an n-parameter Lie group of unitary Hamiltonian symmetries $U(t, a_1, a_2, \dots, a_n)$, having infinitesimal generators

$$S_i(t) = -i \frac{\partial U}{\partial a_i}\Big|_{\mathbf{a}=\mathbf{0}},$$

then differentiating Eq. (14.38) with respect to a_i gives

$$[S_i, H] + i\hbar \frac{\partial S_i}{\partial t} = 0. \tag{14.39}$$

Any hermitian operator $S(t)$ satisfying this equation is said to be a **constant of the motion** or **conserved quantity**, for Schrödinger's equation implies that its expection values are constant:

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\langle S\rangle_\psi = \frac{\mathrm{d}}{\mathrm{d}t}\langle\psi|S|\psi\rangle &= \langle\frac{\mathrm{d}\psi}{\mathrm{d}t}|S|\psi\rangle + \langle\psi|S\frac{\mathrm{d}}{\mathrm{d}t}|\psi\rangle + \langle\psi|\frac{\partial S}{\partial t}|\psi\rangle \\
&= \frac{-1}{i\hbar}\langle\psi|HS|\psi\rangle + \frac{1}{i\hbar}\langle\psi|SH|\psi\rangle + \langle\psi|\frac{\partial S}{\partial t}|\psi\rangle \\
&= \frac{1}{i\hbar}\langle\psi|[S, H] + i\hbar\frac{\partial S}{\partial t}|\psi\rangle = 0.
\end{aligned}$$

Exercise: Show that in the Heisenberg picture, this is equivalent to

$$\frac{\mathrm{d}S_{\mathrm{H}}}{\mathrm{d}t} = 0 \quad \text{where} \quad S_{\mathrm{H}} = \mathrm{e}^{-iHt/\hbar} S \mathrm{e}^{iHt/\hbar}.$$

From Examples 14.6 and 14.7 it follows that invariance of the Hamiltonian under translations and rotations is equivalent to conservation of momentum and angular momentum respectively. In both cases the infinitesimal generators are time-independent. If the Hamiltonian is invariant under time translations, having generator $S = -H/\hbar$ (see Example 14.8), then Eq. (14.39) reduces to

$$-\frac{1}{\hbar}[H, H] - i\frac{\partial H}{\partial t} = 0,$$

which is true if and only if H has no explicit time dependence, $\partial H/\partial t = 0$.

Discrete symmetries

There are a number of important symmetries of a more discrete nature, illustrated in the following examples.

Example 14.9 Consider a spatial inversion $\mathbf{r} \mapsto \mathbf{r}' = -\mathbf{r}$, which can be thought of as a rotation by 180° about the z-axis followed by a reflection $x' = x,\ y' = y,\ z' = -z$. Let Π

be the operator on $\mathcal{H}$ induced by such an inversion, satisfying

$$\Pi^* Q_i \Pi = -Q_i, \qquad \Pi^* P_i \Pi = -P_i.$$

By Wigner's theorem,

$$\Pi^*[Q_i, P_j]\Pi = \Pi^* i\hbar\delta_{ij}\Pi = \begin{cases} i\hbar\delta_{ij} & \text{if } \Pi \text{ is unitary,} \\ -i\hbar\delta_{ij} & \text{if } \Pi \text{ is anti-unitary.} \end{cases}$$

It turns out that Π must be a unitary operator, for

$$\Pi^*[Q_i, P_j]\Pi = [\Pi^* Q_i \Pi, \Pi^* P_j \Pi] = [-Q_i, -P_j] = [Q_i, P_j] = i\hbar\delta_{ij}.$$

Note also that angular momentum operators are invariant under spatial reflections,

$$\Pi^* L_i \Pi = L_i \quad \text{where} \quad \mathbf{L} = \mathbf{Q} \times \mathbf{P}.$$

Since successive reflections result in the identity $\Pi^2 = I$, we have $\Pi^* = \Pi$. Hence Π is a hermitian operator, corresponding to an observable called **parity**, having eigenvalues ± 1. States of eigenvalue 1, $\Pi|\psi\rangle = |\psi\rangle$, are said to be of **even parity**, while those of eigenvalue -1 are called **odd parity**, $\Pi|\psi\rangle = -|\psi\rangle$. Every state can be decomposed as a sum of an even and an odd parity state,

$$|\psi\rangle = \tfrac{1}{2}(I + \Pi)|\psi\rangle + \tfrac{1}{2}(I - \Pi)|\psi\rangle = |\psi_+\rangle + |\psi_-\rangle.$$

Exercise: Show that if $[\Pi, H] = 0$, the parity of any state is preserved throughout its motion, and eigenstates of H with non-degenerate eigenvalue have definite parity.

Example 14.10 In classical physics, if $\mathbf{q}(t)$ is a solution of Newton's equations then so is the reverse motion $\mathbf{q}_{\text{rev}}(t) = \mathbf{q}(t)$ having opposite momentum $\mathbf{p}_{\text{rev}}(t) = -\mathbf{p}(-t)$. If O' is an observer having time in the reversed direction $t' = -t$ to that of an observer O, let the time-reversed states be

$$|\psi'\rangle = \Theta|\psi\rangle,$$

where Θ is the time-reversal operator. Since we require

$$\Theta^* Q_i \Theta = Q_i, \qquad \Theta^* P_i \Theta = -P_i$$

a similar discussion to that in Example 14.9 gives

$$\begin{aligned} \Theta^*[Q, P]\Theta &= \Theta^* i\hbar I \Theta = \pm i\hbar I \\ &= [Q, -P] = -i\hbar I. \end{aligned}$$

Hence time-reversal Θ is an anti-unitary operator.

If the Hamiltonian H is invariant under time reversal, $[H, \Theta] = 0$, then applying Θ to Schrödinger's equation (14.13) gives

$$-i\hbar\frac{\mathrm{d}}{\mathrm{d}t}\Theta|\psi(t)\rangle = \Theta i\hbar\frac{\mathrm{d}}{\mathrm{d}t}|\psi\rangle = \Theta H|\psi\rangle = H\Theta|\psi(t)\rangle.$$

Changing the time variable t to $-t$,

$$i\hbar\frac{\mathrm{d}}{\mathrm{d}t}\Theta|\psi(-t)\rangle = H\Theta|\psi(-t)\rangle.$$

It follows that $|\psi_{\mathrm{rev}}(t)\rangle = \Theta|\psi(-t)\rangle$ is a solution of Schrödinger's equation, which may be thought of as the time-reversed solution. In this sense, the dynamics of quantum mechanics is time-reversable, but note that because of the anitilinear nature of the operator Θ, a complex conjugation is required in addition to time inversion. For example in the position representation, if $\psi(\mathbf{r}, t)$ is a solution of Schrödinger's wave equation (14.18), then $\psi(\mathbf{r}, -t)$ is not in general a solution. However, taking the complex conjugate shows that $\psi_{\mathrm{rev}}(t) = \overline{\psi(\mathbf{r}, -t)}$ is a solution of (14.18),

$$i\hbar\frac{\partial}{\partial t}\overline{\psi(\mathbf{r}, -t)} = -\left(\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{r})\right)\overline{\psi(\mathbf{r}, -t)}.$$

Identical particles

Consider a system consisting of N indistinguishable particles. If the Hilbert space of each individual particle is $\mathcal{H}$ we take the Hilbert space of the entire system to be the tensor product

$$\mathcal{H}^N = \mathcal{H} \otimes \mathcal{H} \otimes \cdots \otimes \mathcal{H}.$$

As in Chapter 7 this may be regarded as the tensor space spanned by free formal products

$$|\psi_1\psi_2 \dots \psi_N\rangle \equiv |\psi_1\rangle|\psi_2\rangle \dots |\psi_N\rangle$$

where each $|\psi_i\rangle \in \mathcal{H}$, subject to identifications

$$\big(\alpha|\psi_1\rangle + \beta|\phi\rangle\big)|\psi_2\rangle \cdots = \alpha|\psi_1\rangle|\psi_2\rangle \cdots + \beta|\phi\rangle|\psi_2\rangle \dots, \text{ etc.}$$

The inner product on $\mathcal{H}^N$ is defined by

$$\langle\psi_1\psi_2 \dots \psi_N \,|\, \phi_1\phi_2 \dots \phi_N\rangle = \langle\psi_1 \,|\, \phi_1\rangle\langle\psi_2 \,|\, \phi_2\rangle \dots \langle\psi_N \,|\, \phi_N\rangle.$$

For each pair $1 \leq i < j \leq N$ define the permutation operator $P_{ij} : \mathcal{H}^N \to \mathcal{H}^N$ by

$$\langle\phi_1 \dots \phi_i \dots \phi_j \dots \phi_N \,|\, P_{ij}\psi\rangle = \langle\phi_1 \dots \phi_j \dots \phi_i \dots \phi_N \,|\, \psi\rangle$$

for all $|\phi_i\rangle \in \mathcal{H}$. This is a linear operator that 'interchanges particles' i and j,

$$P_{ij}|\psi_1 \dots \psi_i \dots \psi_j \dots \psi_N\rangle = |\psi_1 \dots \psi_j \dots \psi_i \dots \psi_N\rangle.$$

If $|\psi_a\rangle$ $(a = 1, 2, \dots)$ is an o.n. basis of the Hilbert space $\mathcal{H}$ then the set of all vectors

$$|\psi_{a_1}\psi_{a_2} \dots \psi_{a_N}\rangle \equiv |\psi_{a_1}\rangle|\psi_{a_2}\rangle \dots |\psi_{a_N}\rangle$$

forms an o.n. basis of $\mathcal{H}^N$. Since $P_i j$ transforms any such o.n. basis to an o.n. basis it must be a unitary operator. These statements extend to a general permutation P, since it can be written as a product of interchanges

$$P = P_{ij}P_{kl} \dots$$

As there is no dynamical way of detecting an interchange of identical particles, the expection values of the Hamiltonian must be invariant under permutations, $\langle H\rangle_{P\psi} = \langle H\rangle_{\psi}$, so that

$$\langle\psi|P^*HP|\psi\rangle = \langle\psi|H|\psi\rangle$$

for all $|\psi\rangle \in \mathcal{H}$. Hence $P^*HP = H$ and as P is unitary, $PP^* = I$,

$$[H, P] = 0.$$

This is yet another example of a discrete symmetry. In classical mechanics it is taken as given that all particles have an individuality and are in principle distinguishable. It is basic to the philosophy of quantum mechanics, however, that since no physical procedure exists for 'marking' identical particles such as a pair of electrons in order to keep track of them, there can be no method even in principle of distinguishing between them.

All interchanges have the property $(P_{ij})^2 = I$, from which they are necessarily hermitian, $P_{ij} = P^*_{ij}$. Every interchange therefore corresponds to an observable. It has eigenvalues $\epsilon \pm 1$ and, since P_{ij} is a constant of the motion, any eigenstate

$$P_{ij}|\psi\rangle = \epsilon|\psi\rangle$$

remains an eigenstate corresponding to the same eigenvalue, for

$$\epsilon = \langle\psi|P_{ij}|\psi\rangle = \langle P_{ij}\rangle_\psi = \text{const.}$$

Since no physical observable can distinguish between states related by a permutation, a similar argument to that used for the Hamiltonian shows that every observable A commutes with all permutation operators,

$$[A, P] = 0.$$

Hence, if $|\psi\rangle$ is a non-degenerate eigenstate of A, it is an eigenstate of every permutation operator P,

$$\begin{aligned} A|\psi\rangle = a|\psi\rangle &\implies AP|\psi\rangle = PA|\psi\rangle = aP|\psi\rangle \\ &\implies P|\psi\rangle = p|\psi\rangle \end{aligned}$$

for some factor $p = \pm 1$. If, as is commonly assumed, every state is representable as a sum of non-degenerate common eigenvectors of a commuting set of complete observables $A, B, C, \dots$ we must then assume that every physical state $|\psi\rangle$ of the system is a common eigenvector of all permutation operators. In particular, for all interchanges P_{ij}

$$P_{ij}|\psi\rangle = p_{ij}|\psi\rangle \quad \text{where} \quad p_{ij} = \pm 1.$$

All p_{ij} are equal for the state $|\psi\rangle$, since for any pair k, l

$$P_{ij}|\psi\rangle = P_{ki}P_{lj}P_{kl}P_{lj}P_{ki}|\psi\rangle,$$

from which it follows that $p_{ij} = p_{kl}$ since

$$p_{ij}|\psi\rangle = (p_{ki})^2(p_{lj})^2 p_{kl}|\psi\rangle = p_{kl}|\psi\rangle.$$

Thus for all permutations either $P|\psi\rangle = |\psi\rangle$ or $P|\psi\rangle = (-1)^P|\psi\rangle$. In the first case, $p_{ij} = 1$, the state is said to be **symmetrical** and the particles are called **bosons** or to obey **Bose–Einstein statistics**. If $p_{ij} = -1$ the state is **antisymmetrical**, the particles are said to be **fermions** and obey **Fermi–Dirac statistics**. It turns out that bosons are always particles of integral spin such as photons or mesons, while fermions such as electrons or protons always have half-integral spin. This is known as the **spin-statistics theorem**, but lies beyond the scope of this book (see, for example, [9]).

The celebrated **Pauli exclusion principle** asserts that two identical fermions cannot occupy the same state, for if

$$|\psi\rangle = |\psi_1\rangle \dots |\psi_i\rangle \dots |\psi_j\rangle \dots |\psi_N\rangle$$

and $|\psi_i\rangle = |\psi_j\rangle$ then

$$P_{ij}|\psi\rangle = |\psi\rangle = -|\psi\rangle$$

since every state has eigenvalue -1. Hence $|\psi\rangle = 0$.

Problems

Problem 14.13 If the operator K is complex conjugation with respect to a complete o.n. set,

$$K\Big(\sum_i \alpha|e_i\rangle\Big) = \sum_i \overline{\alpha_i}|e_i\rangle,$$

show that every anti-unitary operator V can be written in the form $V = UK$, where U is a unitary operator.

Problem 14.14 For any pair of operators A and B show by induction on the coefficients that

$$\mathrm{e}^{aB}A\mathrm{e}^{-aB} = A + a[B, A] + \frac{a^2}{2!}[B, [B, A]] + \frac{a^3}{3!}[B, [B, [B, A]]] + \dots$$

Hence show the relation (14.32) holds for $T(\mathbf{a}) = \mathrm{e}^{i\mathbf{a}\cdot\mathbf{P}/\hbar}$.

Problem 14.15 Using the expansion in Problem 14.14 show that $R(\theta) = \mathrm{e}^{i\theta L_3/\hbar}$ satisfies Eqs. (14.33)–(14.35).

Problem 14.16 Show that the time reversal of angular momentum $\mathbf{L} = \mathbf{Q} \times \mathbf{P}$ is $\Theta^* L_i \Theta = -L_i$, and that the commutation relations $[L_i, L_j] = i\hbar\epsilon_{ijk}L_k$ are only preserved if Θ is anti-unitary.

14.4 Quantum statistical mechanics

Statistical mechanics is the physics of large systems of particles, which are usually identical. The systems are generally so large that only averages of physical quantities can be accurately dealt with. This section will give only the briefest introduction to this enormous and far-ranging subject.

Density operator

Let a quantum system have a complete o.n. basis $|\psi_i\rangle$. If we imagine the rest of the universe (taken in a somewhat restricted sense) to be spanned by an o.n. set $|\theta_a\rangle$, then the general state of the combined system can be written

$$|\Psi\rangle = \sum_i \sum_a c_{ia} |\psi_i\rangle |\theta_a\rangle.$$

An operator A acting on the system only acts on the vectors $|\psi_i\rangle$, hence

$$\begin{aligned}
\langle A\rangle_\Psi &= \langle \Psi | A | \Psi \rangle \\
&= \sum_i \sum_a \sum_j \sum_b \overline{c_{ia}} \langle \theta_a | \langle \psi_i | A | \psi_j \rangle | \theta_b \rangle c_{jb} \\
&= \sum_i \sum_a \sum_j \sum_b A_{ij} \overline{c_{ia}} c_{jb} \delta_{ab} \\
&= \sum_i \sum_j A_{ij} \rho_{ji}
\end{aligned}$$

where

$$A_{ij} = \langle \psi_i | A | \psi_j \rangle, \qquad \rho_{ji} = \sum_a c_{ja} \overline{c_{ia}}.$$

The operator A can be written

$$A = \sum_i \sum_j A_{ij} |\psi_i\rangle \langle \psi_j|.$$

Exercise: Verify that for any $|\phi\rangle \in \mathcal{H}$,

$$A|\phi\rangle = \sum_i \sum_j A_{ij} |\psi_i\rangle \langle \psi_j | \phi \rangle.$$

Define the **density operator** ρ as that having components ρ_{ij},

$$\rho = \sum_i \sum_j \rho_{ij} |\psi_i\rangle \langle \psi_j|,$$

which is hermitian since

$$\overline{\rho_{ji}} = \sum_a \overline{c_{ia}} c_{ja} = \rho_{ij}.$$

A useful expression for the expectation value of A is

$$\langle A \rangle = \mathrm{tr}(A\rho) = \mathrm{tr}(\rho A) \tag{14.40}$$

where the trace of an operator is given by

$$\mathrm{tr}\, B = \sum_i \langle \psi_i | B | \psi_i \rangle = \sum_i B_{ii}.$$

Exercise: Show that the trace of an operator is independent of the o.n. basis $|\psi_i\rangle$.

Setting $A = I$ we have

$$\langle I \rangle = \langle \Psi \mid \Psi \rangle = \|\Psi\|^2 = 1 \implies \mathrm{tr}(\rho) = 1,$$

and setting $A = |\psi_k\rangle\langle\psi_k|$ gives

$$\langle A \rangle = \langle \Psi \mid \psi_k \rangle \langle \psi_k \mid \Psi \rangle = \|\langle \Psi \mid \psi_k \rangle\|^2 \geq 0.$$

On the other hand

$$\begin{aligned} \mathrm{tr}(A\rho) &= \sum_i \langle \psi_i | A\rho | \psi_i \rangle = \sum_i \langle \psi_i \mid \psi_k \rangle \langle \psi_k | \rho | \psi_i \rangle \\ &= \sum_i \delta_{ik} \langle \psi_k | \rho | \psi_i \rangle = \langle \psi_k | \rho | \psi_k \rangle = \rho_{kk}. \end{aligned}$$

Hence all diagonal elements of the density matrix are positive, $\rho_{kk} = \langle \psi_k | \rho | \psi_k \rangle \geq 0$. Assuming ρ is a complete operator, select $|\psi_i\rangle$ to be eigenvectors, $\rho|\psi_i\rangle = w_i|\psi_i\rangle$, so that ρ is diagonalized

$$\rho = \sum_i w_i |\psi_i\rangle\langle\psi_i|. \tag{14.41}$$

We then have

$$\sum_i w_i = 1, \quad w_i \geq 0.$$

The interpretation of the density operator ρ, or its related state $|\Psi\rangle$, is as a **mixed state** of the system, with the ith eigenstate $|\psi_i\rangle$ having probability w_i. A **pure state** occurs when there exists k such that $w_k = 1$ and $w_i = 0$ for all $i \neq k$. In this case $\rho^2 = \rho$ and the density operator is idempotent – it acts as a projection operator into the one-dimensional subspace spanned by the associated eigenstate ψ_k.

Exercise: Show the converse: if $\rho^2 = \rho$ then all $w_i = 1$ or 0, and there exists k such that $\rho = |\psi_k\rangle\langle\psi_k|$.

Exercise: Show that the probability of finding the system in a state $|\chi\rangle$ is $\mathrm{tr}\, \rho|\chi\rangle\langle\chi|$.

Example 14.11 Consider a beam of photons in the z-direction. Let $|\psi_1\rangle$ be the state of a photon polarized in the x-direction, and $|\psi_2\rangle$ be the state of a photon polarized in the y-direction. The general state is a linear sum of these two,

$$|\psi\rangle = a|\psi_1\rangle + b|\psi_2\rangle \quad \text{where} \quad |a|^2 + |b|^2 = 1.$$

The pure state represented by this vector has density operator $\rho = |\psi\rangle\langle\psi|$, having components

$$\rho_{ij} = \langle \psi_i | \rho | \psi_j \rangle = \langle \psi_i \mid \psi \rangle \overline{\langle \psi_j \mid \psi \rangle} = \begin{pmatrix} a\bar{a} & a\bar{b} \\ b\bar{a} & b\bar{b} \end{pmatrix}.$$

For example, the pure states corresponding to $45°$-polarization $\left(a = b = \frac{1}{\sqrt{2}}\right)$ and

135°-polarization $\left(a = -b = -\frac{1}{\sqrt{2}}\right)$ have respective density operators

$$\rho = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} \quad \text{and} \quad \rho = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

A half–half mixture of 45°- and 135°-polarized photons is indistinguishable from an equal mixture of x-polarized photons and y-polarized photons, since

$$\Big(\frac{1}{\sqrt{2}}|\psi_1\rangle + \frac{1}{\sqrt{2}}|\psi_2\rangle\Big)\Big(\frac{1}{\sqrt{2}}\langle\psi_1| + \frac{1}{\sqrt{2}}\langle\psi_2|\Big) + \Big(\frac{1}{\sqrt{2}}|\psi_1\rangle - \frac{1}{\sqrt{2}}|\psi_2\rangle\Big)\Big(\frac{1}{\sqrt{2}}\langle\psi_1| - \frac{1}{\sqrt{2}}\langle\psi_2|\Big)$$
$$= \tfrac{1}{2}|\psi_1\rangle\langle\psi_1| + \tfrac{1}{2}|\psi_2\rangle\langle\psi_2|.$$

From Schrödinger's equation (14.13),

$$i\hbar\frac{\mathrm{d}}{\mathrm{d}t}|\psi_i\rangle = H|\psi_i\rangle, \qquad -i\hbar\frac{\mathrm{d}}{\mathrm{d}t}\langle\psi_i| = \langle\psi_i|H.$$

It follows that the density operator satisfies the evolution equation

$$\frac{\mathrm{d}\rho}{\mathrm{d}t} = \frac{-i}{\hbar}[H, \rho]. \tag{14.42}$$

From the solution (14.14), (14.15) of the Schrödinger equation, the solution of (14.42) is

$$\rho(t) = \mathrm{e}^{(-i/\hbar)Ht}\rho(0)\mathrm{e}^{(i/\hbar)Ht}. \tag{14.43}$$

Hence for any function $f(\rho) = \sum_i f(w_i)|\psi_i\rangle\langle\psi_i|$ the trace is constant,

$$\mathrm{tr}\big(f(\rho)\big) = \mathrm{tr}\big(\mathrm{e}^{(-i/\hbar)Ht} f(\rho(0))\mathrm{e}^{(i/\hbar)Ht}\big) = \mathrm{tr}\big(f(\rho(0))\big) = \text{const}.$$

A mixed state is said to be **stationary** if $\mathrm{d}\rho/\mathrm{d}t = 0$. From Eq. (14.42) this implies $[H, \rho] = 0$, and for any pair of energy eigenvectors

$$H|E_j\rangle = E_j|E_j\rangle, \qquad H|E_k\rangle = E_k|E_k\rangle$$

we have

$$0 = \langle E_j|\rho H - H\rho|E_k\rangle = (E_k - E_j)\langle E_j|\rho|E_k\rangle.$$

Hence, if $E_j \neq E_k$ then $\langle E_j|\rho|E_k\rangle = \rho_{jk} = 0$, and if H has no degenerate energy levels then

$$\rho = \sum_i w_i|E_i\rangle\langle E_i|,$$

which is equivalent to the assertion that the density operator is a function of the Hamiltonian, $\rho = \rho(H)$. If H has degenerate energy levels then ρ and H can be simultaneously diagonalized, and it is possible to treat this as a limiting case of non-degenerate levels. It is reasonable therefore to assume that $\rho = \rho(H)$ in all cases.

Ensembles

An **ensemble** of physical systems is another way of talking about the density operator. Essentially, we consider a large number of copies of the same system, within certain constraints, to represent a statistical system of particles. Each member of the ensemble is a possible state of the system; it is an eigenstate of the Hamiltonian and the density operator tells us its probability within the ensemble.

One of the simplest examples is the **microcanonical ensemble**, where ρ is constant for energy values in a narrow range, $E < E_k < E + \Delta E$, and $w_j = 0$ for all energy values E_j outside this range. For those energy values within the allowed range, we set $w_k = w = 1/s$ where s is the number of energy values in the range $(E, E + \Delta E)$. Let $j(E)$ be the number of states with energy $E_k < E$, then

$$s = j(E + \Delta E) - j(E) = \Sigma(E)\Delta E,$$

where

$$\Sigma(E) = \frac{\mathrm{d}j(E)}{\mathrm{d}E} = \text{the density of states.}$$

For the microcanonical ensemble all $w_k = 0$ for $E_k \le E$ or $E_k \ge E + \Delta E$, while

$$w_k = \frac{1}{\Sigma(E)\Delta E} \quad \text{if} \quad E < E_k < E + \Delta E.$$

The **canonical ensemble** can be thought of as a system embedded in a heat reservoir consisting of the external world. Let H be the Hamiltonian of the system and H_R that of the reservoir. The total Hamiltonian of the universe is $H_U = H_R + H$. Suppose the system is in the eigenstate $|\psi_m\rangle$ of energy E_m

$$H|\psi_m\rangle = E_m|\psi_m\rangle$$

and let $|\Psi\rangle = |\theta\rangle|\psi_m\rangle$ be the total state of the universe. If we assume the universe to be in a microcanonical ensemble, then

$$H_U|\Psi\rangle = E_U|\Psi\rangle \quad \text{where} \quad E < E_U < E + \Delta E.$$

Using the decomposition $H_U = H_R + H$ we have

$$H_R|\theta\rangle|\psi_m\rangle + |\theta\rangle E_m|\psi_m\rangle = E_U|\theta\rangle|\psi_m\rangle,$$

whence

$$H_R|\theta\rangle = (E_U - E_m)|\theta\rangle.$$

Thus $|\theta\rangle$ is an eigenstate of H_R with energy $E_U - E_m$. If $\Sigma_R(E_R)$ is the density of states in the reservoir, then

$$w_m \Sigma_U(E_U)\Delta E = \Sigma_R(E_U - E_m)\Delta E,$$

whence

$$w_m = \frac{\Sigma_R(E_U - E_m)}{\Sigma_U(E_U)}.$$

For $E_m \ll E_U$, as expected of a system in a much larger reservoir,

$$\ln w_m = \text{const.} - \beta E_m,$$

most commonly written in the form

$$w_m = \frac{1}{Z} e^{-\beta E_m} \quad \text{where} \quad Z = \sum_{m=0}^{\infty} e^{-\beta E_m}, \tag{14.44}$$

where the last identity follows from $\sum_m w_m = 1$. The density operator for the canonical ensemble is thus

$$\rho = \frac{1}{Z} e^{-\beta H} \tag{14.45}$$

where, by the identity $\operatorname{tr} \rho = 1$,

$$Z = \operatorname{tr} e^{-\beta H} = \sum_{m=0}^{\infty} e^{-\beta E_m}, \tag{14.46}$$

known as the **canonical partition function**. The average energy is

$$U = \langle E \rangle = \operatorname{tr}(\rho H) = \frac{1}{Z} \sum_k E_k e^{-\beta E_k} = -\frac{\partial \ln Z}{\partial \beta}. \tag{14.47}$$

Example 14.12 Consider a linear harmonic oscillator having Hamiltonian given by Eq. (14.19). The energy eigenvalues are

$$E_m = \tfrac{1}{2}\hbar\omega + m\hbar\omega$$

and the partition function is

$$\begin{aligned} Z &= \sum_{m=0}^{\infty} e^{-\beta E_m} \\ &= e^{-\frac{1}{2}\beta\hbar\omega} \sum_{m=0}^{\infty} \left(e^{-\beta\hbar\omega}\right)^m. \\ &= \frac{e^{\beta\hbar\omega/2}}{e^{\beta\hbar\omega} - 1}. \end{aligned}$$

From Eq. (14.47) the average energy is

$$\begin{aligned} U &= -\frac{\partial \ln Z}{\partial \beta} \\ &= \frac{1}{2}\hbar\omega + \frac{\hbar\omega}{e^{\beta\hbar\omega} - 1}. \end{aligned}$$

As $\beta \to 0$ we have $U \approx \beta^{-1}$. This is the classical limit $U = kT$, where T is the temperature, and is an indication of the identity $\beta = 1/kT$. As $\beta \to \infty$ we arrive at the low temperature limit,

$$U \to \tfrac{1}{2}\hbar\omega + \hbar\omega e^{-\beta\hbar\omega} \approx \tfrac{1}{2}\hbar\omega.$$

The **entropy** is defined as

$$S = -k\,\mathrm{tr}(\rho \ln \rho) = -k \sum_i w_i \ln w_i. \tag{14.48}$$

For a pure state, $w_i = 1$ or 0, we have $S = 0$. This is interpreted as a state of maximum order. For a completely random state, $w_i = \text{const.} = 1/N$ where N is the total number of states in the ensemble (assumed finite here), the entropy is

$$S = k \ln N.$$

This state of maximal disorder corresponds to a maximum value of S, as may be seen by using the method of Lagrange multipliers: the maximum of S occurs where $dS = 0$ subject to the constraint $\sum_i w_i = 1$,

$$\begin{aligned} dS = 0 &\Longrightarrow d\sum_i w_i \ln w_i - \lambda \sum_i dw_i = 0 \\ &\Longrightarrow \sum_i (1 + \ln w_i - \lambda)\, dw_i = 0. \end{aligned}$$

Since dw_i is arbitrary the Lagrange multiplier is $\lambda = 1 + \ln w_i$, so that

$$w_1 = w_2 = \cdots = e^{\lambda - 1}.$$

Exercise: Two systems may be said to be independent if their combined density operator is $\rho = \rho_1\rho_2 = \rho_2\rho_1$. Show that the entropy has the additive property for independent systems, $S = S_1 + S_2$.

If the Hamiltonian depends on a parameter a, we define 'generalized force' A conjugate to a by

$$A = \left\langle -\frac{\partial H}{\partial a}\right\rangle = -\,\mathrm{tr}\left(\rho \frac{\partial H}{\partial a}\right). \tag{14.49}$$

For example, for a gas in a volume V, the **pressure** p is defined as

$$p = \left\langle -\frac{\partial H}{\partial V}\right\rangle = -\,\mathrm{tr}\left(\rho \frac{\partial H}{\partial V}\right).$$

If $H|\psi_k\rangle = E_k|\psi_k\rangle$ where $E_k = E_k(a)$ and $|\psi_k\rangle = |\psi_k(a)\rangle$, then

$$\frac{\partial H}{\partial a}|\psi_k\rangle + H\frac{\partial}{\partial a}|\psi_k\rangle = \frac{\partial E_k}{\partial a}|\psi_k\rangle + E_k\frac{\partial}{\partial a}|\psi_k\rangle$$

so that

$$
\begin{aligned}
A = -\left\langle \frac{\partial H}{\partial a} \right\rangle &= -\operatorname{tr}\Big(\rho \frac{\partial H}{\partial a}\Big) \\
&= -\sum_k w_k \langle \psi_k | \frac{\partial H}{\partial a} | \psi_k \rangle \\
&= -\sum_k w_k \Big(\langle \psi_k | \frac{\partial E_k}{\partial a} | \psi_k \rangle + \langle \psi_k | E_k - H | \frac{\partial}{\partial a} \psi_k \rangle \Big) \\
&= -\sum_k w_k \Big(\frac{\partial E_k}{\partial a} \| \psi_k \|^2 + \langle \psi_k | E_k - E_k | \frac{\partial}{\partial a} \psi_k \rangle \Big) \\
&= -\sum_k w_k \frac{\partial E_k}{\partial a}.
\end{aligned}
$$

The total work done under a change of parameter da is defined to be

$$
\mathrm{d}W = -\mathrm{d}U = -\sum_k w_k \frac{\partial E_k}{\partial a} = A \,\mathrm{d}a.
$$

For a change in volume this gives the classical formula $\mathrm{d}W = p\,\mathrm{d}V$.

For the canonical ensemble we have, by Eq. (14.44),

$$
A = -\sum_k w_k \frac{\partial E_k}{\partial a} = \frac{1}{\beta} \frac{\partial \ln Z}{\partial a}, \tag{14.50}
$$

and as the entropy is given by

$$
S = -k \sum_k w_k \ln w_k = k \sum_k w_k (\ln Z + \beta E_k) = k(\ln Z + \beta U)
$$

we have

$$
\mathrm{d}S = k(\mathrm{d} \ln Z + \beta \,\mathrm{d}U) = \frac{1}{T}(\mathrm{d}U + A \,\mathrm{d}a) \tag{14.51}
$$

where

$$
\beta = \frac{1}{kT}.
$$

This relation forms the basic connection between statistical mechanics and thermodynamics (see Section 16.4); the quantity T is known as the **temperature** of the system.

Systems of identical particles

For a system of N identical particles, bosons or fermions, let h be the Hamiltonian of each individual particle, having eigenstates

$$
h|\varphi_a\rangle = \varepsilon_a |\varphi_a\rangle \quad (a = 0, 1, \dots).
$$

The Hamiltonian of the entire system is $H : \mathcal{H}^N \to \mathcal{H}^N$ given by

$$H = h_1 + h_2 + \cdots + h_N$$

where

$$h_i|\psi_1\rangle \ldots |\psi_i\rangle \ldots |\psi_N\rangle = |\psi_1\rangle \ldots h|\psi_i\rangle \ldots |\psi_N\rangle.$$

The eigenstates of the total Hamiltonian,

$$H|\Phi_k\rangle = E_k|\Phi_k\rangle,$$

are linear combinations of state vectors

$$|\psi_1\rangle|\psi_2\rangle \ldots |\psi_N\rangle$$

such that

$$E_k = \varepsilon_0 + \varepsilon_1 + \cdots + \varepsilon_N.$$

If n_0 particles are in state $|\varphi_0\rangle$, n_1 particles in state $|\varphi_1\rangle$, etc. then the energy eigenstates are determined by the set of occupation numbers $(n_0, n_1, n_2, \ldots)$ such that

$$E_k = \sum_{a=0}^{\infty} n_a \varepsilon_a.$$

If we are looking for eigenstates that are simultaneously eigenstates of the permutation operators P, then they must be symmetric states $P|\Psi_k\rangle = |\Psi_k\rangle$ for bosons, and antisymmetric states $P|\Psi_k\rangle = (-1)^P|\Psi_k\rangle$ in the case of fermions. Let S be the symmetrization operator and A the antisymmetrization operator

$$S = \frac{1}{N!}\sum_P P, \qquad A = \frac{1}{N!}\sum_P (-1)^P P.$$

Both are hermitian and idempotent

$$S^* = \frac{1}{N!}\sum_P P^* = \frac{1}{N!}\sum_P P^{-1} = S, \qquad A^* = A \text{ since } (-1)^{P^*} = (-1)^P$$

$$S^2 = S, \qquad A^2 = A, \qquad AS = SA = 0.$$

Thus S and A are orthogonal projection operators, and for any state $|\Psi\rangle$

$$PS|\Psi\rangle = S|\Psi\rangle, \qquad PA|\Psi\rangle = (-1)^P A|\Psi\rangle$$

for all permutations P. For bosons the eigenstates are of the form $S|\varphi_{a_1}\rangle|\varphi_{a_2}\rangle \ldots |\varphi_{a_N}\rangle$, while for fermions they are $A|\varphi_{a_1}\rangle|\varphi_{a_2}\rangle \ldots |\varphi_{a_N}\rangle$. In either case the state of the system is completely determined by the occupation numbers $n_0, n_1, \ldots$ For bosons the occupation numbers run from 0 to N, while the Pauli exclusion principle implies that fermionic occupation numbers only take on values 0 or 1. Thus, for the canonical distribution

$$w(n_0, n_1, \ldots) = \frac{1}{Z} e^{-\beta \sum_a n_a \varepsilon_a}, \qquad \sum_a n_a = N,$$

where

$$Z_{\text{Bose}} = \sum_{n_0=0}^{N} \sum_{n_1=0}^{N} \dots \mathrm{e}^{-\beta \sum_a n_a \varepsilon_a},$$

$$Z_{\text{Fermi}} = \sum_{n_0=0}^{1} \sum_{n_1=0}^{1} \dots \mathrm{e}^{-\beta \sum_a n_a \varepsilon_a}.$$

The constraint $\sum_a n_a = N$ makes these sums quite difficult to calculate directly.

In the classical version where particles are distinguishable, all ways of realizing a configuration are counted separately,

$$\begin{aligned} Z_{\text{Boltzmann}} &= \sum_{n_0=0}^{N} \sum_{n_1=0}^{N} \dots \mathrm{e}^{-\beta \sum_a n_a \varepsilon_a} \frac{N!}{n_0! n_1! \dots} \\ &= \left(\mathrm{e}^{-\beta \varepsilon_0} + \mathrm{e}^{-\beta \varepsilon_1} + \dots \right)^N \\ &= (Z_1)^N, \end{aligned}$$

where Z_1 is the one-particle partition function

$$Z_1 = \sum_a \mathrm{e}^{-\beta \varepsilon_a}.$$

The average energy is

$$U = \langle E \rangle = -\frac{\partial \ln Z}{\partial \beta} = \frac{N \sum_a \varepsilon_a \mathrm{e}^{-\beta \varepsilon_a}}{\sum_a \mathrm{e}^{-\beta \varepsilon_a}} = N \langle \varepsilon \rangle.$$

It is generally accepted that $Z_{\text{Boltzmann}}$ should be divided by $N!$, discounting all possible permutations of particles, in order to avoid the Gibbs paradox.

In the quantum case, it is easier to consider an even larger distribution, wherein the number of particles is no longer fixed. Assuming an *open system*, allowing exchange of particles between the system and reservoir, an argument similar to that used to arrive at the canonical ensemble gives

$$w_{mN} = \frac{1}{Z_g} \mathrm{e}^{\alpha N - \beta E_m},$$

where

$$Z_g = \sum_{N=0}^{\infty} \sum_m \mathrm{e}^{\alpha N - \beta E_m}.$$

This is known as the partition function for the **grand canonical ensemble**. In terms of the density operator,

$$\rho = \frac{1}{Z_g} \mathrm{e}^{-\beta H + \alpha N}, \qquad Z_g = \mathrm{tr}\, \mathrm{e}^{-\beta H + \alpha N}.$$

For a system of identical particles,

$$w(n_0, n_1, \dots) = \frac{1}{Z} \mathrm{e}^{\alpha \sum_a n_a - \beta \sum_a n_a \varepsilon_a},$$

where $\sum_a n_a = N$ is no longer fixed. We can therefore write

$$\begin{aligned} Z_{\text{Bose}} &= \sum_{n_0=0}^{\infty}\sum_{n_1=0}^{\infty}\dots e^{\alpha\sum_a n_a - \beta\sum_a n_a\varepsilon_a} \\ &= \sum_{n_0=0}^{\infty} e^{(\alpha n_0 - \beta\varepsilon_0)n_0} \sum_{n_1=0}^{\infty} e^{(\alpha n_1 - \beta\varepsilon_1)n_1}\dots \\ &= \frac{1}{1-e^{\alpha-\beta\varepsilon_0}}\frac{1}{1-e^{\alpha-\beta\varepsilon_1}}\dots \end{aligned}$$

and

$$\ln Z_{\text{Bose}} = -\sum_{a=0}^{\infty}\ln\bigl(1-\lambda e^{-\beta\varepsilon_a}\bigr), \quad \lambda = e^{\alpha}.$$

Similarly

$$Z_{\text{Fermi}} = \sum_{n_0=0}^{1}\sum_{n_1=0}^{1}\dots e^{\alpha\sum_a n_a - \beta\sum_a n_a\varepsilon_a}\dots$$

results in

$$\ln Z_{\text{Fermi}} = \sum_{a=0}^{\infty}\ln\bigl(1+\lambda e^{-\beta\varepsilon_a}\bigr), \quad \lambda = e^{\alpha}.$$

Summarizing, we have

$$\ln Z = \pm\sum_{a=0}^{\infty}\ln\bigl(1\pm\lambda e^{-\beta\varepsilon_a}\bigr) \tag{14.52}$$

where the + sign occurs for fermions, the − sign for bosons.

The average occupation numbers are

$$\begin{aligned} \langle n_a\rangle &= \sum_{n_0}\sum_{n_1}\dots n_a w(n_0, n_1, \dots) \\ &= \frac{1}{Z}\sum_{n_0}\sum_{n_1}\dots n_a e^{\sum_a(\alpha-\beta\varepsilon_a)n_a} \\ &= -\frac{1}{\beta}\frac{\partial}{\partial\varepsilon_a}\ln Z. \end{aligned}$$

Using Eq. (14.52),

$$\langle n_a\rangle = -\frac{1}{\beta}\frac{\partial}{\partial\varepsilon_a}\Bigl(\pm\ln\bigl(1\pm\lambda e^{-\beta\varepsilon_a}\bigr)\Bigr) = \frac{1}{\lambda^{-1}e^{\beta\varepsilon_a}\pm 1}, \tag{14.53}$$

where the + sign applies to Fermi particles, and the − to Bose. The parameter λ is often written $\lambda = e^{\beta\mu}$, where μ is known as the **chemical potential**. It can be shown, using the method of steepest descent (see [10]) that the formulae (14.53) are also valid for the canonical ensemble. In this case the total particle number is fixed so that the chemical potential may be found from

$$\sum_a\langle n_a\rangle = \frac{1}{Z}\sum_{n_0}\sum_{n_1}\dots\Bigl(\sum_a n_a\Bigr)e^{\sum_b(\alpha-\beta\varepsilon_b)n_b} = N\sum_{n_0}\sum_{n_1}\dots w(n_0, n_1, \dots) = N.$$

That is,

$$N = \sum_a \frac{1}{e^{\beta(\varepsilon_a - \mu)} \pm 1}$$

and

$$U = \langle E \rangle = -\frac{\partial Z}{\partial \beta} = \sum_a \frac{\varepsilon_a}{e^{\beta(\varepsilon_a - \mu)} \pm 1} = \sum_a \varepsilon_a \langle n_a \rangle.$$

Application to perfect gases, black body radiation and other systems may be found in any standard book on statistical mechanics [1, 10–12].

Problems

Problem 14.17 Show that the correctly normalized fermion states are

$$\frac{1}{\sqrt{N!}} \sum_P (-1)^P P |\varphi_{a_1}\rangle |\varphi_{a_2}\rangle \ldots |\varphi_{a_N}\rangle$$

and normalized boson states are

$$\frac{1}{\sqrt{N!}\sqrt{n_0!}\sqrt{n_1!}\ldots} \sum_P P |\varphi_{a_1}\rangle |\varphi_{a_2}\rangle \ldots |\varphi_{a_N}\rangle.$$

Problem 14.18 Calculate the canonical partition function, mean energy U and entropy S, for a system having just two energy levels 0 and E. If $E = E(a)$ for a parameter a, calculate the force A and verify the thermodynamic relation $\mathrm{d}S = \frac{1}{T}(\mathrm{d}U + A\,\mathrm{d}a)$.

Problem 14.19 let $\rho = e^{-\beta H}$ be the unnormalized canonical distribution. For a free particle of mass m in one dimension show that its position representation form $\rho(x, x'; \beta) = \langle x | \rho | x' \rangle$ satisfies the diffusion equation

$$\frac{\partial \rho(x, x'; \beta)}{\partial \beta} = \frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} \rho(x, x'; \beta)$$

with 'initial' condition $\rho(x, x'; 0) = \delta(x - x')$. Verify that the solution is

$$\rho(x, x'; \beta) = \left(\frac{m}{2\pi \hbar^2 \beta} \right)^{1/2} e^{-m(x-x')^2/2\hbar^2\beta}.$$

Problem 14.20 A solid can be regarded as being made up of $3N$ independent quantum oscillators of angular frequency ω. Show that the canonical partition function is given by

$$Z = \left(\frac{e^{-\beta\hbar\omega/2}}{1 - e^{-\beta\hbar\omega}} \right)^{3N},$$

and the specific heat is given by

$$C_V = \frac{\mathrm{d}U}{\mathrm{d}T} = 3Nk \left(\frac{T_0}{T} \right)^2 \frac{e^{T_0/T}}{(e^{T_0/T} - 1)^2} \quad \text{where} \quad kT_0 = \hbar\omega.$$

Show that the high temperature limit $T \gg T_0$ is the classical value $C_V = 3Nk$.

Problem 14.21 Show that the average occupation numbers for the classical distribution, $Z_{\mathrm{Boltzmann}}$ are given by

$$\langle n_a \rangle = -\frac{1}{\beta} \frac{\partial}{\partial \varepsilon_a} \ln Z_{\mathrm{Boltzmann}} = \lambda e^{-\beta \varepsilon_a}.$$

Hence show that

$$\langle n_a \rangle_{\text{Fermi}} < \langle n_a \rangle_{\text{Boltzmann}} < \langle n_a \rangle_{\text{Bose}}$$

and that all three types agree approximately for low occupation numbers $\langle n_a \rangle \ll 1$.

Problem 14.22 A *spin system* consists of N particles of magnetic moment μ in a magnetic field B. When n particles have spin up, $N - n$ spin down, the energy is $E_n = n\mu B - (N - n)\mu B = (2n - N)\mu B$. Show that the canonical partition function is

$$Z = \frac{\sinh\big((N + 1)\beta\mu B\big)}{\sinh \beta\mu B}.$$

Evaluate the mean energy U and entropy S, sketching their dependence on the variable $x = \beta\mu B$.

References

[1] R. H. Dicke and J. P. Wittke. *Introduction to Quantum Mechanics*. Reading, Mass., Addison-Wesley, 1960.

[2] P. Dirac. *The Principles of Quantum Mechanics*. Oxford, Oxford University Press, 1958.

[3] J. M. Jauch. *Foundations of Quantum Mechanics*. Reading, Mass., Addison-Wesley, 1968.

[4] A. Sudbery. *Quantum Mechanics and the Particles of Nature*. Cambridge, Cambridge University Press, 1986.

[5] R. D. Richtmyer. *Principles of Advanced Mathematical Physics, Vol. 1*. New York, Springer-Verlag, 1978.

[6] M. Schlechter. *Operator Methods in Quantum Mechanics*. New York, Elsevier-North Holland, 1981.

[7] J. von Neumann. *Mathematical Foundations of Quantum Mechanics*. Princeton, N.J., Princeton University Press, 1955.

[8] E. Zeidler. *Applied Functional Analysis*. New York, Springer-Verlag, 1995.

[9] R. F. Streater and A. S. Wightman. *PCT, Spin and Statistics, and All That*. New York, W. A. Benjamin, 1964.

[10] E. Schrödinger. *Statistical Thermodynamics*. Cambridge, Cambridge University Press, 1962.

[11] L. D. Landau and E. M. Lifshitz. *Statistical Physics*. Oxford, Pergamon, 1980.

[12] F. Reif. *Statistical and Thermal Physics*. New York, McGraw-Hill, 1965.