

**APAS 5110. Internal Processes in Gases. Fall 1999.**  
**Mathematical Structure of Nonrelativistic Quantum Mechanics**

1. POSTULATES

The following is intended as an outline of the basic mathematical structure of nonrelativistic quantum mechanics. A rigorous exposition is given by J. Von Neumann (1955, 'Mathematical Foundations of Quantum Mechanics', Princeton University Press).

**Postulate 1.** Particles are described by **wavefunctions**  $\psi(t, \mathbf{x}, \sigma)$  which are complex-valued, square-integrable functions of space  $\mathbf{x}$  and time  $t$  and of additional (discrete) variables  $\sigma$  which characterize the particles' spin and other internal properties. Square-integrable means that the integral over all space of the (sum over all  $\sigma$  of the) absolute value squared of the wavefunction

$$\int_V \sum_{\sigma} |\psi(t, \mathbf{x}, \sigma)|^2 d^3x \quad (1.1)$$

exists and is finite. The domain, the volume  $V$ , over which the wavefunctions is defined may be chosen arbitrarily, although usually  $V$  is taken to be all of space. In the bra-ket notation introduced in Postulate 3 below, square-integrability means that  $\langle \psi | \psi \rangle$  exists and is finite.

Two wavefunctions  $\phi$  and  $\psi$  are considered equal,  $\phi = \psi$ , if and only if

$$\int_V \sum_{\sigma} |\phi - \psi|^2 d^3x = 0 \quad (1.2)$$

so that  $\phi$  may differ from  $\psi$  on a set of measure zero.

Wavefunctions can be defined formally as **vectors**  $\psi_{x\sigma}(t)$  in a **Hilbert space**, which is a countably infinite-dimensional vector space. von Neumann takes Hilbert space as the starting point for defining wavefunctions.

**Postulate 2.** The wavefunction of a pair of particles is a function

$$\psi(1, 2) \equiv \psi(t, \mathbf{x}_1, \sigma_1, \mathbf{x}_2, \sigma_2) \quad (1.3)$$

of the coordinates of each of the particles at time  $t$ . If the particles are identical fermions (particles of half-integral spin), then the wavefunction is antisymmetric in the coordinates: it changes sign under the exchange  $1 \leftrightarrow 2$  of particles

$$\psi(2, 1) = -\psi(1, 2) \quad (\text{fermions}) \quad (1.4)$$

while if the particles are identical bosons (particles of integral spin), then the wavefunction is symmetric, i.e. unchanged, under particle exchange

$$\psi(2, 1) = \psi(1, 2) \quad (\text{bosons}) . \quad (1.5)$$

More generally, the wavefunction of a system of many particles is a function of the coordinates of each of the particles. The wavefunction is antisymmetric under exchange of the coordinates of any pair of identical fermions, and symmetric under exchange of the coordinates of any pair of identical bosons.

The space of (anti)symmetric space of many-particle wavefunctions is called the (anti)symmetric **Fock space** associated with a Hilbert space.

**Postulate 3.** The Hilbert space of wavefunctions, and more generally the Fock space of many-particle wavefunctions, is equipped with an **inner product**

$$\langle \phi | \psi \rangle = \sum_{\sigma_1, \dots, \sigma_N} \int \phi(t, \mathbf{x}_1, \sigma_1, \dots, \mathbf{x}_N, \sigma_N)^\dagger \psi(t, \mathbf{x}_1, \sigma_1, \dots, \mathbf{x}_N, \sigma_N) d^{3N}x, \quad (1.6)$$

where the superscript  $\dagger$  on the wavefunction  $\phi$  denotes its **Hermitian conjugate**, which is the transpose of the complex conjugate of  $\phi$

$$\phi^\dagger \equiv (\phi^*)^T. \quad (1.7)$$

Why the transpose? The inner product (1.6) is essentially a generalization of the scalar product in finite-dimensional vector spaces; in matrix notation, the scalar product of  $\phi^*$  and  $\psi$  would be written as the transpose of the column vector  $\phi^*$  multiplied by the column vector  $\psi$ :

$$\langle \phi | \psi \rangle = \left( \dots \phi^* \dots \right) \cdot \begin{pmatrix} \cdot \\ \cdot \\ \psi \\ \cdot \\ \cdot \end{pmatrix}. \quad (1.8)$$

The convenient notation  $\langle \phi | \psi \rangle$  is Dirac's bra-ket notation: the ket  $|\psi\rangle$  denotes the vector  $\psi$  in Hilbert space, while the bra  $\langle \phi |$  is the Hermitian conjugate of the vector  $\phi$ .

Notice that in nonrelativistic quantum mechanics space  $\mathbf{x}$  and time  $t$  appear on an unequal footing. Space and time appear more symmetrically when one considers scattering amplitudes for particles to scatter between specified initial and final states, which involve integrals over both space and time. Such scattering amplitudes are a forerunner of the Feynman diagrams of relativistic quantum field theory, where space and time appear in a manifestly Lorentz covariant fashion.

Two wavefunctions are defined to be equal,  $\phi = \psi$ , if the 'distance' between the two wavefunctions is zero,

$$\langle \phi - \psi | \phi - \psi \rangle = 0 \quad (1.9)$$

which is the same as equation (1.2).

**Postulate 4.** Each physically observable property of a particle or system of particles is associated with some **linear, Hermitian operator**  $A$ , acting on Hilbert space, which possesses a **complete orthonormal** set of **eigenfunctions**  $\psi_n$  with **real eigenvalues**  $a_n$

$$A\psi_n = a_n\psi_n. \quad (1.10)$$

The eigenvalues  $a_n$  represent the possible values of the physical property associated with  $A$ . The property has a definite value  $a$  if the wavefunction  $\psi$  is an eigenfunction of  $A$  with eigenvalue  $a$ , but in general  $\psi$  is not an eigenfunction, and then the property does not have a definite value.

The Hermitian property of operators is defined and discussed in §3. The word orthonormal means that the eigenfunctions  $\psi_n$  are mutually orthogonal and normalized to unity

$$\langle \psi_m | \psi_n \rangle = \delta_{mn}. \quad (1.11)$$

Equation (1.11) is true if the spectrum of eigenvalues is discrete; in the case of a continuous spectrum of eigenvalues, the normalization condition must be modified to (2.6), as discussed

in §2. The word complete means that an arbitrary wavefunction  $\psi$  in the Hilbert space can be expanded as a linear combination of eigenfunctions

$$\psi = \sum_n c_n \psi_n \quad (1.12)$$

where  $c_n$  are complex numbers. According to the definition (1.9) of equality of wavefunctions, equation (1.12) is to be understood in the sense that  $\langle (\psi - \sum_n c_n \psi_n) | (\psi - \sum_n c_n \psi_n) \rangle = 0$ . Taking the inner product of the expansion (1.12) with  $\langle \psi_m |$  and using the orthonormality (1.11) of the eigenfunctions implies that the complex coefficients  $c_n$  are given by

$$c_n = \langle \psi_n | \psi \rangle . \quad (1.13)$$

When a wavefunction  $\psi$  is expanded in terms of some specific set of orthonormal eigenfunctions  $\psi_n$ , as in equation (1.12), one sometimes refers to  $\psi$  as being ‘in the  $\psi_n$  representation’. The origin and significance of this nomenclature, which comes from group theory, is clarified in §6.

The quantities  $\langle \phi | A | \psi \rangle$  are called **matrix elements** of the operator  $A$ . The properties of an operator are completely specified by its matrix elements  $\langle \psi_m | A | \psi_n \rangle$  in some representation  $\psi_n$  (not necessarily eigenfunctions of  $A$ ).

**Postulate 5.** If the wavefunction is  $\psi$ , normalized to unity  $\langle \psi | \psi \rangle = 1$ , then the probability that the physical property associated with  $A$  takes the eigenvalue  $a_n$  is

$$|c_n|^2 = |\langle \psi_n | \psi \rangle|^2 . \quad (1.14)$$

The unit normalization of  $\psi$  ensures that the sum of the probabilities is one

$$1 = \langle \psi | \psi \rangle = \sum_n |c_n|^2 . \quad (1.15)$$

The mean value of the physical property associated with  $A$  is

$$\bar{A} = \sum_n a_n |c_n|^2 = \langle \psi | A | \psi \rangle . \quad (1.16)$$

**Postulate 6.** The operators of energy  $E$  and momentum  $p$  are, in the spatial representation,

$$E = i\hbar \frac{\partial}{\partial t} , \quad (1.17)$$

and

$$\mathbf{p} = -i\hbar \frac{\partial}{\partial \mathbf{x}} . \quad (1.18)$$

**Postulate 7.** There exists a Hermitian operator  $H$ , called the Hamiltonian, which determines the evolution of the wavefunction through the linear equation

$$i\hbar \frac{\partial \psi}{\partial t} = H \psi . \quad (1.19)$$

Equation (1.19) is called Schrödinger’s equation.

## 2. CONTINUOUS SPECTRA OF EIGENVALUES

In quantum mechanics some observables, such as angular momentum, take on a discrete spectrum of eigenvalues, but others, such as position and momentum, have a continuous spectrum. In the discrete case, the eigenfunctions  $\psi_n$  can be normalized to unity,  $\langle \psi_n | \psi_n \rangle = 1$ , equation (1.11), but in the continuous case, this normalization must be modified, equation (2.6) below. The fact that this normalization requires individual eigenfunctions to be normalized to infinity, equation (2.10), leads to the conclusion that values of physical quantities corresponding to operators with continuous spectra cannot be measured with infinite precision. This is related to the circumstance (von Neumann 1955) that the dimension of Hilbert space is countably infinite (i.e. any linearly independent set of square-integrable wavefunctions can be labeled by the integers), whereas the continuum is uncountable. That the spectrum of physically realizable eigenvalues must be countable in reality is of some philosophical interest. Notwithstanding the difficulties, operators with continuous spectra of eigenvalues are of considerable practical use and importance, witness the spatial and momentum operators.

Suppose that  $A$  is a Hermitian operator having a complete set of orthogonal eigenfunctions  $\psi_k$  with  $k$  a continuous parameter. In going from the discrete to the continuous case, one wishes to replace both the discrete amplitudes  $c_n$  and the discrete probabilities  $|c_n|^2$  by continuous analogs

$$c_n \rightarrow c_k dk , \quad (2.1)$$

$$|c_n|^2 \rightarrow |c_k|^2 dk . \quad (2.2)$$

In accordance with (2.1), an arbitrary wavefunction  $\psi$  in Hilbert space is written as an integral instead of a sum (1.12) over eigenfunctions

$$\psi = \int c_k \psi_k dk . \quad (2.3)$$

In accordance with (2.2), for wavefunctions  $\psi$  with unit normalization  $\langle \psi | \psi \rangle = 1$ , one wishes to interpret  $|c_k|^2 dk$  as the probability that the parameter  $k$  lies in an interval  $dk$ , the unit normalization of  $\psi$  ensuring, as in the discrete case (1.15), that the sum of the probabilities is one

$$\int |c_k|^2 dk = \langle \psi | \psi \rangle = 1 . \quad (2.4)$$

But (2.3) implies that  $\langle \psi | \psi \rangle$  is

$$\langle \psi | \psi \rangle = \int c_{k'} c_k \langle \psi_{k'} | \psi_k \rangle dk' dk . \quad (2.5)$$

Comparing the two expressions (2.4) and (2.5), which must be equal for arbitrary  $c_k$ , one concludes that the eigenfunctions  $\psi_k$  must satisfy the orthonormality condition

$$\langle \psi_{k'} | \psi_k \rangle = \delta(k' - k) , \quad (2.6)$$

where  $\delta(k)$  is the **Dirac delta-function**. The Dirac delta-function is defined by the properties that it is zero everywhere except at the origin  $k = 0$ , where it is infinite,

$$\delta(k) = 0 \quad (k \neq 0) , \quad \delta(0) = \infty , \quad (2.7)$$

in such a way that its integral over any interval containing the origin is unity

$$\int \delta(k) dk = 1 . \quad (2.8)$$

Equations (2.3) and (2.6) imply that the coefficients  $c_k$  are given by

$$c_k = \langle \psi_k | \psi \rangle , \quad (2.9)$$

which is the same as the discrete case (1.13).

It is apparent from the orthonormality condition (2.6) that the eigenfunctions  $\psi_k$  are themselves not square integrable quantities, since

$$\langle \psi_k | \psi_k \rangle = \infty . \quad (2.10)$$

This means that the eigenfunctions  $\psi_k$  are not admissible wavefunctions for single particles (nor for any finite number of particles), whose wavefunctions must be normalizable to unity to admit a finite total probability. The infinite normalization (2.10) of the eigenfunctions  $\psi_k$  can be interpreted as requiring an infinite number of particles, which cannot occur in reality. For example, in the spatial representation, the eigenfunctions of the spatial operator  $\mathbf{x}$  satisfying the orthonormality condition (2.6) are delta-functions  $\delta(\mathbf{x} - \mathbf{x}_0)$  with eigenvalues  $\mathbf{x}_0$ , but these eigenfunctions cannot represent actual particles, since that would require an infinite number of particles at the position  $\mathbf{x}_0$ . Similarly, in the momentum representation, the orthonormal eigenfunctions of the momentum operator  $\mathbf{p}$  are Fourier modes  $(2\pi)^{-3/2} e^{i\mathbf{k}\cdot\mathbf{x}}$ , which can be interpreted as representing an infinite train of particles of momentum  $\mathbf{p} = \hbar\mathbf{k}$  distributed uniformly over all space.

Two conclusions can be drawn from the argument of the previous paragraph. The first is that if an operator has a continuous spectrum of eigenvalues, then a wavefunction cannot physically be in a single eigenstate, but rather it must be in a superposition of eigenstates. The distribution of the wavefunction may be arbitrarily narrow over some interval of eigenstates, but it cannot be infinitely narrow. For example, a wavefunction  $\psi(\mathbf{x})$  can be regarded as an expansion in eigenfunctions of the spatial operator,

$$\psi(\mathbf{x}) = \int \psi(\mathbf{x}') \delta(\mathbf{x}' - \mathbf{x}) d^3x' , \quad (2.11)$$

with amplitude  $\psi(\mathbf{x})$  and probability  $|\psi(\mathbf{x})|^2$  for the particle to be at position  $\mathbf{x}$ . The wavefunction  $\psi(\mathbf{x})$  can be made narrow about some position  $\mathbf{x}_0$  say, but it cannot physically be made infinitely narrow. The condition on a physically realizable wavefunction  $\psi(\mathbf{x})$  is that it must be square-integrable, as asserted in Postulate 1 of §1.

The fact that particles cannot be confined to single eigenstates of a continuous spectrum can be viewed as a reflection of the uncertainty principle. For example, isolating a particle to a definite position  $\mathbf{x}$  in space requires infinite momentum, which is physically impossible, while isolating a particle to a definite momentum  $\mathbf{p}$  requires infinite space, which is likewise physically impossible.

The second conclusion to be drawn is that quantities which are truly physically observable, in the sense that they can take definite values, must correspond to operators having discrete spectra of eigenvalues. Nonetheless, operators having continuous spectra of great practical use and importance, witness the spatial and momentum operators. The way out

of this difficulty is to regard operators with continuous spectra as limiting cases of operators with discrete spectra. While reality presumably finds its own way to discretize itself, mathematically an operator with a continuous spectrum can be discretized for example by the simple device of restricting the space of eigenfunctions to a discrete grid, and replacing integrations by sums over the grid. The continuous spectrum is attained in the limit of an infinitely fine grid. That the same continuous limit is attained independent of how the discretization is done follows from the fact that physically measurable quantities are eigenvalues, and eigenfunctions whose eigenvalues differ only infinitesimally cannot be distinguished observationally. In the case of the momentum operator, discretizing the momentum eigenvalues on a grid is equivalent to putting particles in a periodic cubic box. A continuous spectrum of momenta is attained in the limit of an infinitely large box.

A similar conclusion applies wherever there is an accumulation point in a discrete spectrum of eigenvalues, a point where an infinite sequence of eigenvalues  $a_n$  converges to a finite value  $a$  as  $n \rightarrow \infty$ .

Thus where an operator with a continuous spectrum of eigenvalues exists, it should be regarded as being attainable as the limit of a sequence of operators (and boundary conditions) with discrete spectra and no accumulation points.

### 3. HERMITIAN OPERATORS

The **Hermitian conjugate** of an operator  $A$  is defined to be the transpose of the complex conjugate of  $A$

$$A^\dagger \equiv (A^*)^T . \quad (3.1)$$

Evidently the Hermitian conjugate of the Hermitian conjugate of an operator is itself,  $(A^\dagger)^\dagger = A$ . The definition (3.1) is to be understood as equivalent to the statement that

$$\phi^\dagger A^\dagger = (A\phi)^\dagger \quad (3.2)$$

for all wavefunctions  $\phi$  in the Hilbert space. An operator  $A$  is said to be **Hermitian** if it equals its Hermitian conjugate  $A^\dagger$

$$A = A^\dagger . \quad (3.3)$$

The statement (3.3) should be understood as equivalent to the statement that

$$\langle \phi | A | \psi \rangle = \langle \phi | A^\dagger | \psi \rangle \quad (3.4)$$

for all wavefunctions  $\phi$  and  $\psi$  in Hilbert space. Equation (3.4) is the same as  $\langle \phi | A \psi \rangle = \langle A \phi | \psi \rangle$ , in view of the definition (3.2) of the Hermitian conjugate.

For an operator to be Hermitian generally requires that the wavefunctions satisfy appropriate boundary conditions. In the case of the momentum operator, for example,

$$\begin{aligned} \langle \phi | \mathbf{p} \psi \rangle &= \int \phi^* \left( \frac{-i\hbar \partial \psi}{\partial \mathbf{x}} \right) d^3x = [-i\hbar \phi^* \psi] + \int \left( \frac{-i\hbar \partial \phi}{\partial \mathbf{x}} \right)^* \psi d^3x \\ &= [-i\hbar \phi^* \psi] + \langle \mathbf{p} \phi | \psi \rangle \end{aligned} \quad (3.5)$$

shows that  $\langle \phi | \mathbf{p} \psi \rangle$  is equal to  $\langle \mathbf{p} \phi | \psi \rangle$  only if the surface term  $[-i\hbar \phi^* \psi]$  vanishes. Thus the momentum operator  $\mathbf{p}$  is Hermitian only with respect to (a Hilbert space of) wavefunctions for which the aforesaid surface term vanishes.

Associated with a Hermitian operator are three important properties:

- (1) its eigenvalues are real;
- (2) its eigenfunctions are orthonormal;
- (3) its eigenfunctions form a complete set.

In the theory of finite-dimensional vector spaces, the property of being Hermitian is equivalent to properties (1)-(3): a finite-dimensional matrix of complex numbers has a complete orthonormal set of eigenfunctions with real eigenvalues if and only if it is Hermitian. In Hilbert space, the correspondence is almost but not quite true: properties (1)-(3) imply an operator is Hermitian, and being Hermitian implies properties (1) and (2), and ‘usually’ but not always the completeness property (3).

It is straightforward to prove that the combination of properties (1)-(3) implies that an operator is necessarily Hermitian. For suppose that an operator  $A$  has a complete set of orthonormal eigenfunctions  $\psi_n$  with real eigenvalues  $a_n$ . Let  $\phi$  and  $\psi$  be any two wavefunctions, with expansions say  $\phi = \sum b_n \psi_n$  and  $\psi = \sum c_n \psi_n$  in terms of the complete set  $\psi_n$ . Then

$$\begin{aligned} \langle \phi | A | \psi \rangle &= \sum_{m,n} b_m^* c_n \langle \psi_m | A | \psi_n \rangle = \sum_{m,n} b_m^* c_n a_n \langle \psi_m | \psi_n \rangle \\ &= \sum_n b_n^* c_n a_n = \sum_n b_n^* c_n a_n^* = \langle \phi | A^\dagger | \psi \rangle \end{aligned} \quad (3.6)$$

proves that  $A$  is Hermitian, as claimed. Step 1 of (3.6) follows from completeness, step 2 from  $\psi_n$  being eigenfunctions, step 3 from the orthonormality of the eigenfunctions, step 4 from the reality of the eigenvalues,  $a = a^*$ , and the final step 5 is steps 1 to 3 in inverse order.

It is also straightforward to prove conversely that a Hermitian operator necessarily satisfies properties (1) and (2). To prove property (1), that the eigenvalues of a Hermitian operator  $A$  are necessarily real, suppose that  $\phi$  is an eigenfunction of  $A$ , with eigenvalue  $a$ . Then

$$a \langle \phi | \phi \rangle = \langle \phi | A | \phi \rangle = \langle \phi | A^\dagger | \phi \rangle = a^* \langle \phi | \phi \rangle . \quad (3.7)$$

Since  $\langle \phi | \phi \rangle$  is necessarily real, positive, and finite, equation (3.7) implies  $a = a^*$ , i.e.  $a$  is real, as claimed.

To prove property (2), that the eigenfunctions of a Hermitian operator  $A$  form an orthonormal set, suppose that  $\psi_m$  and  $\psi_n$  are two eigenfunctions of  $A$  with (necessarily real, as above) eigenvalues  $a_m$  and  $a_n$ . Then

$$0 = \langle \psi_m | A^\dagger | \psi_n \rangle - \langle \psi_m | A | \psi_n \rangle = (a_m - a_n) \langle \psi_m | \psi_n \rangle . \quad (3.8)$$

If the eigenvalues are unequal,  $a_m \neq a_n$ , then equation (3.8) shows that  $\psi_m$  and  $\psi_n$  are necessarily mutually orthogonal

$$\langle \psi_m | \psi_n \rangle = 0 . \quad (3.9)$$

If on the other hand the eigenvalues are the same,  $a_m = a_n$ , then  $\phi_m$  and  $\phi_n$  need not be orthogonal. However,  $\phi_m$  can be made orthogonal to  $\phi_n$  by Gram-Schmidt orthogonalization, that is, by subtracting from  $\phi_m$  an appropriate amount of  $\phi_n$ ,

$$\phi_m \rightarrow \phi_m - \frac{\langle \phi_m | \phi_n \rangle}{\langle \phi_n | \phi_n \rangle} \phi_n . \quad (3.10)$$

More generally, any set of two or more degenerate (i.e. having the same eigenvalue) eigenfunctions can be orthogonalized by the Gram-Schmidt procedure. An orthogonal set of eigenfunctions can always be made orthonormal by normalizing the eigenfunctions suitably. Gram-Schmidt orthogonalization is not unique, but any two orthonormal sets of eigenfunctions, say  $\phi_m$  and  $\psi_n$ , which span the space of degenerate eigenfunctions must be related to each other

$$\phi_m = \sum_n U_{nm}^\dagger \psi_n, \quad \psi_n = \sum_m U_{nm} \phi_m \quad (3.11)$$

by a transformation  $U$

$$U_{nm} = \langle \phi_m | \psi_n \rangle \quad (3.12)$$

which evidently must be unitary (see §4).

It follows from the argument of the previous paragraph that, if a Hermitian operator has a complete orthonormal set of eigenfunctions, then that set is unique up to arbitrary unitary transformations among degenerate eigenfunctions.

This leaves property (3): do the eigenfunctions of a Hermitian operator necessarily form a complete set? The answer is no.

As an example of a Hermitian operator which has no eigenfunctions, consider a particle confined by an infinite potential to a box with perfectly reflecting walls. The momentum operator  $\mathbf{p} \equiv -i\hbar\partial/\partial\mathbf{x}$  remains a well-defined Hermitian operator acting on wavefunctions which satisfy the appropriate boundary conditions, namely those which vanish at (and beyond) the boundary of the box. However, the momentum operator has no eigenfunctions satisfying the boundary conditions, since if it did then that would imply that there exist states of definite momentum for the particle in the box, whereas in fact the particle is constantly changing its momentum by bouncing off the walls. Indeed, the average value of the momentum is zero for all wavefunctions confined to the box,  $\langle \psi | \mathbf{p} | \psi \rangle = 0$ .

In the above example, the momentum operator failed to have eigenfunctions because the space of wavefunctions was too restricted: that is, the condition that the wavefunctions vanished at the boundary was too restrictive. The condition for the momentum operator to remain Hermitian with respect to wavefunctions defined over the box is that the surface term in equation (3.5) vanishes for all wavefunctions. In particular, the momentum operator remains Hermitian if the space of wavefunctions is enlarged by admitting wavefunctions which are periodic over the box (this is not the only way of enlarging the space of wavefunctions, but it is the simplest way). Here, instead of bouncing off the walls, a particle which exits at one face is considered to reemerge through the opposite face. As is well known, the momentum operator does have a complete orthonormal set of eigenfunctions defined over the space of wavefunctions periodic in a box, namely the Fourier modes  $\psi_{\mathbf{k}} = e^{2\pi i \mathbf{k} \cdot \mathbf{x}/L}$  where  $L$  is the length of each side of the cubic box, and each component of  $\mathbf{k} = (k_x, k_y, k_z)$  runs over the integers,  $k_i = \dots, -1, 0, 1, \dots$ .

The above example illustrates the fact that in Hilbert space a Hermitian operator must be accompanied by appropriate boundary conditions for there to be a possibility that it possesses a complete set of eigenfunctions. In fact the momentum operator  $\mathbf{p}$  does have a complete set of orthonormal eigenfunctions  $\psi = (2\pi)^{-3/2} e^{i\mathbf{k} \cdot \mathbf{x}}$ , provided that the domain of definition of the wavefunctions is extended over all space (or that periodic boundary conditions are imposed on the finite rectangular box, so that a particle exiting at one face

reemerges through the opposite face, rather than being reflected). The boundary conditions here reflect the physical requirement — Heisenberg’s uncertainty principle — that a precise measurement of momentum requires infinite spatial extent.

If an operator is Hermitian, are there always appropriate nontrivial boundary conditions such that the operator possesses a complete set of eigenfunctions? Apparently the answer is still ‘not necessarily’. However, for an important class of operators, namely those which are bounded below or above, a complete set of orthonormal eigenfunctions always exists. A Hermitian operator  $A$  is said to be bounded below if the mean value of  $A$  exceeds some (real) constant  $a_0$

$$\langle \psi | A | \psi \rangle \geq a_0 \quad (3.13)$$

for all wavefunctions  $\psi$  with unit normalization  $\langle \psi | \psi \rangle = 1$ . Bounded above is the same thing but with a  $\leq$  sign. For example, any operator which is the absolute value squared  $A^\dagger A$  of some operator  $A$  is bounded below, since  $\langle \psi | A^\dagger A | \psi \rangle = \langle A\psi | A\psi \rangle$  is positive for any  $\psi$ . For example, in the previously considered case of a particle confined to a box with perfectly reflecting walls, the square  $p^2$  of the momentum operator has a complete set of eigenfunctions, even though the momentum operator  $\mathbf{p}$  itself has no eigenfunctions. For bounded Hermitian operators, the boundary conditions affect the existence of eigenfunctions only insofar as they suffice to ensure that the operator is indeed Hermitian and bounded.

The Hamiltonian, which represents the energy of a system, is expected physically to be bounded below in all cases, so the Hamiltonian should always have a complete set of eigenfunctions.

The proof that a bounded Hermitian operator  $A$  has a complete set of eigenfunctions involves an explicit procedure for constructing eigenfunctions, a problem of interest in its own right, and goes along the following lines. An operator which is bounded above can be converted into one which is bounded below by taking its negative, so it suffices to consider operators which are bounded below. It suffices also to consider operators with discrete spectra of eigenvalues, since according to the argument at the end of §2 operators with continuous spectra must be attainable as the limit of a sequence of operators with discrete spectra. Let  $\psi_0$  be that normalized wavefunction,  $\langle \psi_0 | \psi_0 \rangle = 1$ , for which the mean value of  $A$  acquires its minimum value,  $a_0$  say, so that

$$\langle \psi_0 | A | \psi_0 \rangle = a_0 . \quad (3.14)$$

If there are several wavefunctions which attain the minimum (3.14), just choose one of them. The idea is to show that  $\psi_0$  is an eigenfunction of  $A$ , with eigenvalue  $a_0$ . Consider perturbing the wavefunction,  $\psi_0 \rightarrow \psi_0 + \epsilon\phi$ , by a small perturbation  $\epsilon\phi$  in which  $\epsilon$  is a small (infinitesimal) complex number, and  $\phi$  is orthogonal to  $\psi_0$  but otherwise arbitrary. The orthogonality condition  $\langle \phi | \psi_0 \rangle = 0$  ensures that unit normalization  $\langle \psi_0 + \epsilon\phi | \psi_0 + \epsilon\phi \rangle = 1$  is preserved to first order in  $\epsilon$ . To first order, the perturbed value of the average of  $A$  is

$$\langle \psi_0 + \epsilon\phi | A | \psi_0 + \epsilon\phi \rangle = a_0 + 2 \operatorname{Re} (\epsilon^* \langle \phi | A | \psi_0 \rangle) , \quad (3.15)$$

the left hand side of which must be greater than or equal to  $a_0$  since  $a_0$  is by definition the minimum value of the average of  $A$ . Therefore the last term on the right hand side of equation (3.15) must be greater than or equal to zero, and since the phase of  $\epsilon$  can be

chosen arbitrarily, it follows that the term must be identically zero, so that

$$\langle \phi | A | \psi_0 \rangle = 0 . \quad (3.16)$$

Equation (3.16) is true for arbitrary wavefunctions  $\phi$  orthogonal to  $\psi_0$ . Now  $A\psi_0$  can always be written as the sum of a piece proportional to  $\psi_0$  and a piece  $\phi$  orthogonal to  $\psi_0$ ,

$$A\psi_0 = a\psi_0 + b\phi , \quad (3.17)$$

for some complex coefficients  $a$  and  $b$ . Taking  $\langle \psi_0 |$  of equation (3.17) implies that  $a = a_0$  from equation (3.14), while taking  $\langle \phi |$  of equation (3.17) implies that  $b = 0$  from equation (3.16), which proves that

$$A\psi_0 = a_0\psi_0 , \quad (3.18)$$

so  $\psi_0$  is an eigenfunction of  $A$  with eigenvalue  $a_0$ , as was to be shown. Successive eigenfunctions of  $A$  can be constructed iteratively. Suppose that orthonormal eigenfunctions  $\psi_0, \psi_1, \dots, \psi_{N-1}$  of  $A$  have been constructed, with eigenvalues  $a_n = \langle \psi_n | A | \psi_n \rangle$ . Then the  $N$ 'th eigenfunction  $\psi_N$  is that normalized wavefunction which minimizes  $\langle \psi_N | A | \psi_N \rangle$  among all normalized wavefunctions orthogonal to all the preceding  $\psi_n, n = 0$  to  $N - 1$ . The associated eigenvalue is  $\langle \psi_N | A | \psi_N \rangle = a_N$ . It is evident by construction that  $a_N$  must be less than or equal to any subsequent eigenvalue  $a_n$  with  $n > N$ . Thus the above prescription yields the eigenfunctions  $\psi_n$  ordered by their eigenvalues,  $a_0 \leq a_1 \leq \dots$ .

It remains to demonstrate the completeness of the set of eigenfunctions constructed according to the prescription of the previous paragraph. Suppose that  $\psi$  is some arbitrary wavefunction, normalized without loss of generality to unity,  $\langle \psi | \psi \rangle = 1$ , and suppose that  $\langle \psi | A | \psi \rangle = a$ . Completeness will have been demonstrated if it can be shown that

$$\sum_{n=0}^N c_n \psi_n \rightarrow \psi \text{ as } N \rightarrow \infty \quad (3.19)$$

in the sense of equation (1.11), with coefficients  $c_n = \langle \psi_n | \psi \rangle$ . Define  $\phi$  to be the difference between  $\psi$  and the partial sum on the left hand side of equation (3.19)

$$\phi \equiv \psi - \sum_{n=0}^N c_n \psi_n . \quad (3.20)$$

Taking  $\langle \psi_n |$  of equation (3.20) shows that  $\phi$  must be orthogonal to all the  $\psi_n$ . If the construction of eigenfunctions yields only a finite number of eigenfunctions, then by construction there can be no wavefunctions orthogonal to all the  $\psi_n$ , so in that case  $\phi$  must be zero when the sum (3.20) is over all the eigenfunctions. So assume that the construction yields an infinite number of eigenfunctions. From the definition (3.20) of  $\phi$ ,

$$\langle \phi | \phi \rangle = \left\langle \psi - \sum_{n=0}^N c_n \psi_n \left| \psi - \sum_{n=0}^N c_n \psi_n \right. \right\rangle = 1 - \sum_{n=0}^N |c_n|^2 , \quad (3.21)$$

and also

$$\langle \phi | A | \phi \rangle = \left\langle \psi - \sum_{n=0}^N c_n \psi_n \left| A \right| \psi - \sum_{n=0}^N c_n \psi_n \right\rangle = a - \sum_{n=0}^N |c_n|^2 a_n . \quad (3.22)$$

But since  $\phi$  is orthogonal to all the  $\psi_n$ , it follows that

$$\langle \phi | A | \phi \rangle \geq a_N , \quad (3.23)$$

since otherwise there would exist an eigenfunction orthogonal to all the  $\psi_n$  for  $n = 0$  to  $N$  with eigenvalue less than  $a_N$ , which by construction of the eigenfunctions  $\psi_n$  cannot occur. Combining (3.21)-(3.23) implies

$$\langle \phi | \phi \rangle \leq \frac{a - a_0 - \sum_{n=0}^N |c_n|^2 (a_n - a_0)}{a_N - a_0} \leq \frac{a - a_0}{a_N - a_0} . \quad (3.24)$$

But  $a_N \rightarrow \infty$  as  $N \rightarrow \infty$ , since the spectrum of eigenvalues is discrete and without accumulation points. Therefore  $\langle \phi | \phi \rangle \rightarrow 0$  as  $N \rightarrow \infty$ , which proves the completeness relation (3.19) as desired. This proves the proposition claimed, that a bounded Hermitian operator has a complete set of orthonormal eigenfunctions.

As a final comment, it is worth noting that the case of operators with complex-valued eigenvalues is no more general than the case of real-valued eigenvalues. It is not hard to show that  $A$  is an operator having a complete orthonormal set of eigenfunctions with complex eigenvalues if and only if  $A$  is a sum  $A = \text{Re}A + i\text{Im}A$  of two commuting Hermitian operators,  $[\text{Re}A, \text{Im}A] = 0$ , having a complete orthonormal set of simultaneous eigenfunctions. The if is obvious; in proving the only if, one defines the real and imaginary parts of  $A$  by  $\text{Re}A \equiv (A + A^\dagger)/2$  and  $\text{Im}A \equiv (A - A^\dagger)/(2i)$ .

#### 4. UNITARY OPERATORS

An operator  $U$  is **unitary** if its Hermitian conjugate is its inverse,

$$U^\dagger U = U U^\dagger = 1 . \quad (4.1)$$

It is to be noted that if  $A$  is a Hermitian operator with a complete set of orthonormal eigenfunctions, then  $e^{iA}$  is a unitary operator, and also  $(A-i)/(A+i)$ , the Cayley transform of  $A$ , is a unitary operator.

#### 5. COMMUTATION OF OPERATORS

The **commutator**  $[A, B]$  of two operators  $A$  and  $B$  is defined by

$$[A, B] \equiv AB - BA . \quad (5.1)$$

An important theorem states that, if  $A$  and  $B$  are two linear Hermitian operators with complete orthonormal sets of eigenfunctions, then  $A$  and  $B$  have a complete orthonormal set of **simultaneous** eigenfunctions if and only if  $A$  and  $B$  commute

$$[A, B] = 0 . \quad (5.2)$$

Physically, the theorem means that two properties  $A$  and  $B$  are simultaneously measurable in quantum mechanics if and only if the corresponding operators commute.

The proof of the theorem is as follows. First, suppose that  $A$  and  $B$  have a complete set of simultaneous eigenfunctions  $\psi_n$  with eigenvalues  $a_n$  and  $b_n$ . Then for each eigenfunction  $\psi_n$

$$AB\psi_n = Ab\psi_n = bA\psi_n = ba\psi_n = ab\psi_n = aB\psi_n = Ba\psi_n = BA\psi_n \quad (5.3)$$

so that  $[A, B]\psi_n = 0$  for all  $n$ . Since the eigenfunctions  $\psi_n$  form a complete set, it follows that  $[A, B]\psi = 0$  for all wavefunctions  $\psi$ , from which it may be concluded that  $A$  and  $B$  commute,  $[A, B] = 0$ . Conversely, suppose that  $[A, B] = 0$ . Let  $\psi_n$  be a complete set of orthonormal eigenfunctions of  $A$  with eigenvalues  $a_n$ . Since the set is complete, the action of the operator  $B$  on each eigenfunction  $\psi_n$  of  $A$  must be expressible as some linear combination of the eigenfunctions, that is,  $B\psi_n = \sum_m b_{nm}\psi_m$ , with coefficients  $b_{nm} = \langle \psi_m | B | \psi_n \rangle$ . Then

$$0 = [A, B]\psi_n = (AB - BA)\psi_n = A \sum_m b_{nm}\psi_m - B a_n \psi_n = \sum_m b_{nm}(a_m - a_n)\psi_m. \quad (5.4)$$

Since the eigenfunctions  $\psi_n$  are orthogonal, the vanishing of the sum on the right hand side of (5.4) implies that each coefficient must vanish,  $b_{nm}(a_m - a_n) = 0$  for all  $m$  and  $n$ . It follows that  $b_{nm} = 0$  for all offdiagonal elements corresponding to unequal eigenvalues,  $a_m \neq a_n$ . In the case of equal eigenvalues,  $a_m = a_n$ , on the other hand,  $b_{nm}$  can be diagonalized over the space of degenerate eigenfunctions, yielding a revised set of simultaneous eigenfunctions of  $A$  and  $B$ . With all sets of degenerate eigenfunctions so diagonalized, then  $b_{nm} = 0$  for all  $n \neq m$ , yielding a complete set of orthonormal simultaneous eigenfunctions of  $A$  and  $B$ , as was to be shown. of operators with a discrete spectrum.

## 6. REPRESENTATIONS

In mathematics, a **group** is a set of elements together with a binary operation, called group multiplication, such that the group product  $gh$  of any two elements  $g$  and  $h$  in the group is also a member of the group. The other defining properties of a group are: (i) **associativity**,  $f(gh) = (fg)h$ ; (ii) the existence of a **unit** element  $1$  satisfying  $1g = g1 = g$  for all elements  $g$ ; and (iii) the existence of an **inverse** element  $g^{-1}$  satisfying  $g^{-1}g = gg^{-1} = 1$  for each element  $g$ . An **abelian** group is a group all of whose elements commute,  $gh = hg$  for any  $g$  and  $h$ , but in general a group need not be abelian.

The term **representation** comes from group theory, where it is defined as a mapping from a group onto a set of (complex-valued, generally) matrices, such that group multiplication corresponds to ordinary matrix multiplication. Thus each group element  $g$  corresponds to some invertible matrix  $g_{mn}$ , and the group product  $gh$  of two elements corresponds to the matrix product  $(gh)_{mn} = \sum_l g_{ml}h_{ln}$ . The **dimension** of a representation is the rank of the matrices of the representation (i.e.  $N$ , for  $N \times N$  matrices), which may be infinite.

The matrices of a representation act in the usual way, by matrix multiplication, on vectors of the same dimension. A group element  $g$  with representation  $g_{mn}$ , acting on a vector  $\psi_n$ , transforms that vector to another vector  $\sum_n g_{mn}\psi_n$ . Thus each element of the group in a given representation can be regarded as defining a transformation — a ‘rotation’ — of the vectors in an  $N$ -dimensional vector space.

The matrices  $g_{mn}$  of a group representation can be regarded as being defined (as regards their components) with respect to an orthonormal basis of vectors  $(1, 0, \dots, 0)$ ,  $(0, 1, \dots, 0)$ ,  $\dots$ ,  $(0, 0, \dots, 1)$  which is called the **basis** of the representation. If the basis is rotated by some arbitrary unitary transformation  $U$  into another orthonormal basis, then it defines a new representation of the group, with matrix elements  $g' = U^\dagger g U$ . Representations related in this fashion are called **equivalent**.

Sets of  $N$ -dimensional vectors are transformed into each other by the action of the group. If the vectors fall into disjoint sets such that the members of a set transform among each

other, but not between different sets, then the representation is said to be **reducible**. If not, then the representation is said to be **irreducible**. A representation can always be decomposed into irreducible parts.

Suppose that  $A$  is a Hermitian operator equipped with a complete orthonormal set  $\psi_n$  of eigenfunctions with eigenvalues  $a_n$ . The Hermitian operator  $A$  can be regarded as defining a group whose elements are the unitary operators  $e^{i\alpha A}$ , with  $\alpha$  some angle. The irreducible representations of the group are the eigenfunctions  $\psi_n$ , since these eigenfunctions transform into themselves, modulo a phase factor, under the action of the group,  $e^{i\alpha A}\psi_n = e^{i\alpha a_n}\psi_n$ . The irreducible representations here are all one-dimensional, and therefore in a sense trivial.

Two important examples of representations of the kind described in the previous paragraph are the **spatial representation** and the **momentum representation**. The spatial representation is defined by the set of irreducible representations of the group of unitary transformations  $e^{i\alpha \cdot \mathbf{x}}$  obtained by exponentiating the spatial operator  $\mathbf{x}$ . These irreducible representations are the eigenfunctions  $\delta(\mathbf{x} - \mathbf{x}_0)$  of the spatial operator  $\mathbf{x}$ , with eigenvalues  $\mathbf{x}_0$ . Similarly, the momentum representation is defined by the set of irreducible representations of the group of transformations  $e^{i\alpha \cdot \mathbf{p}}$  based on the momentum operator  $\mathbf{p} = -i\hbar\partial/\partial\mathbf{x}$ . The irreducible representations here are the eigenfunctions of the momentum operator  $\mathbf{p}$ , the Fourier modes  $(2\pi)^{-3/2}e^{i\mathbf{k} \cdot \mathbf{x}}$ , with eigenvalues  $\hbar\mathbf{k}$ .

The unitary groups obtained by exponentiating the spatial and momentum operators have a physical interpretation. In the case of the momentum operator, the group is the group of spatial translations, an element  $e^{i\alpha \cdot \mathbf{p}}$  of the group corresponding to translation by spatial distance  $\Delta\mathbf{x} = \hbar\alpha$ . This is true because the momentum operator  $\mathbf{p}$  (divided by  $\hbar$ ) can be regarded as generating an infinitesimal spatial translation by  $\delta\mathbf{x}$ ,

$$\psi(\mathbf{x} + \delta\mathbf{x}) = \left(1 + \delta\mathbf{x} \cdot \frac{\partial}{\partial\mathbf{x}}\right) \psi(\mathbf{x}) = (1 + i\delta\mathbf{x} \cdot \mathbf{p}/\hbar) \psi(\mathbf{x}), \quad (6.1)$$

repeated application of which yields a finite translation by  $\Delta\mathbf{x}$

$$\psi(\mathbf{x} + \Delta\mathbf{x}) = e^{i\Delta\mathbf{x} \cdot \mathbf{p}/\hbar} \psi(\mathbf{x}). \quad (6.2)$$

Similarly, the group obtained by exponentiating the spatial operator  $\mathbf{x} = i\hbar\partial/\partial\mathbf{p}$  is the group of translations in momentum space, an element  $e^{i\alpha \cdot \mathbf{x}}$  of the group corresponding to translation in momentum space by  $\Delta\mathbf{p} = -\hbar\alpha$ .

In the theory of continuous unitary groups, a set of linearly independent operators  $A_j$  such that every element of the group can be expressed in the form  $\exp\left(i\sum_j \alpha_j A_j\right)$  is called a set of **generators** of the group. In the example of the group based on exponentiation of the spatial operator  $\mathbf{x}$ , the generators are the three components  $x_j$  of the three-dimensional spatial operator. Similarly, the generators of the group based on exponentiation of the momentum operator  $\mathbf{p}$  are the three component operators  $p_j$ . In both these examples, the generators commute with each other,  $[x_i, x_j] = 0$  and  $[p_i, p_j] = 0$ . As a result of this commutation, each element of the spatial group can be factored into the product of three commuting elements,  $\exp\left(i\sum_j \alpha_j x_j\right) = \prod_j \exp(i\alpha_j x_j)$ , and similarly for the momentum group. When every element of a group can be factored uniquely into commuting products of elements from disjoint subsets of the group, then it is said to be a **direct product** of the subsets. Thus the spatial and momentum groups are direct products of three groups,

corresponding to transformations in each of the three dimensions. In fact it is plain that the spatial and momentum groups are abelian, meaning that all elements of the group commute with each other. An abelian group with a single generator is called a **cyclic** group. Every abelian group can be expressed as a direct product of cyclic groups. The fact that the irreducible representations of the spatial and momentum groups are all one-dimensional is intimately related to the fact that the group is abelian.

Groups with higher than one-dimensional irreducible representations are necessarily non-abelian, and the generators of such groups do not all commute. The commutators  $[A_i, A_j]$  of the generators of a continuous unitary group can always be expressed as some linear combination of the generators

$$[A_i, A_j] = i \sum_k f_{ijk} A_k . \quad (6.3)$$

This is true because by definition every element of a continuous unitary group can be expressed as an exponential of some linear combination of the generators  $A_k$ , and in particular there exist constants  $f_{ijk}$  such that

$$\exp(i\delta\alpha_i A_i) \exp(i\delta\alpha_j A_j) \exp(-i\delta\alpha_i A_i) \exp(-i\delta\alpha_j A_j) = \exp\left(i\delta\alpha_i \delta\alpha_j \sum_k f_{ijk} A_k\right) , \quad (6.4)$$

which leads to equation (6.3) when expanded to second order in small angles  $\delta\alpha_i$  and  $\delta\alpha_j$ . The quantities  $f_{ijk}$  are called the **structure constants** of the group, and they essentially define the group.

A prominent example of a non-abelian continuous group is the three-dimensional orthogonal rotation group  $O(3)$ , whose generators are the angular momentum operators  $L_j$  (divided by  $\hbar$ ).