APPLIED(9) COURSE

Quantum Mechanics

For the Student of

Faculty of Education Level 4

Major Math

First Term 2022-2023

Contents

Education Is Not The Learning of Facts; I's Rather The Traininig Of The Mind To Think (Albert Einstein)

Chapter 1

Dirac Formulation of Quantum Mechanics

The failure of classical mechanics to account for many experimental resultssuch as the stability of atoms and matter, blackbody radiation, specific heatof solids, wave-particle duality of light and material particles, and such, led physicists to the realization that classical concepts were inherently inadequate to describe the physical behavior of events on an atomic scale. To ex-

plainthese phenomena, a fundamental departure from classical mechanics wasnecessary. This departure took the form of postulating, as a fundamental law of nature, that there is a limit to the accuracy with which a measurement(or observation) on a physical system can be made. That is, the aciual measurement itself disturbs the system being measured in an uncontrollable way, regardless of the care, skill, or ingenuity of the experimenter. The disturbance produced by the measurement in turn requires modification ofthe classical concept of causality, since, in the classical sense, there is a causalconnection between the system and the measurement. This leads to a theoryin which one

can predict only the probability of obtaining a certain result when a measurement is made on a system rather than an exact value, as inthe classical case.

Classical mechanics must be contained as a limiting case in quantum mechanics because, if the disturbance caused by an observation may beneglected, classical mechanics is valid. The quantum description of a system must shift to a classical description in this limit, provided the quantum system has a classical analog. This is called the correspondence principle andrestricts the possible forms that a quantum theory may have.

In the following we give a simplified treatment of the

Dirac formulation ofnonrelativistic quantum mechanics. We restrict ourselves to one-dimensionalproblems, for the most part, since the extension to three dimensions isfairly straightforward.

The Dirac formulation involves the concept of vectors (and operators) ina space that may have a finite or an infinite number of dimensions. Let usgive a simple illustration of the way in which such vectors arise in the theory. We shall consider a particle of mass m constrained to move in one dimensionin a potential $V(q)$, where q is the coordinate of the particle which may have any value from $-\infty$ to $+\infty$; that is, the particle may be anywhere in theone-dimensional

space. According to the Schrödinger formulation of wave mechanics, the state of the particle at time t is described by a wavefunction in the position representation, $\psi(q,t)$. If no intervening measurements are made, this state develops in a completely causal way from thestate at time t_0, $\psi(q,t_0)$, according to the postulated Schrödinger wave equation

$$\left[-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial q^2}+V(q)\right]\psi(q,t)=i\hbar\frac{\partial\psi}{\partial t} \tag{1.0.1}$$

where $\hbar$ is Planck's constant divided by 2π, The probability interpretation (necessary when a measurement is made to determine the position of the particle) of $\psi(q,t)$ is as follows: $|\psi(q,t)|^2dq$ gives the probability

of finding the particle between q and $q + dq$ at time t when a measurement of positionis made.

We may take the Fourier transform of $\psi(q, t)$ to obtain another wave function

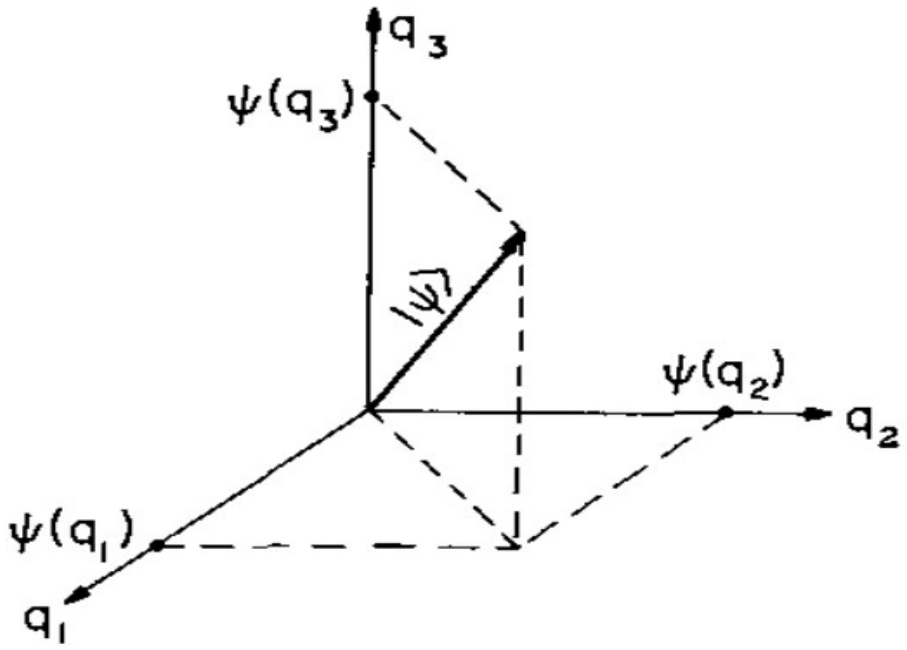
$$\phi(p, t) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} \psi(q, t) e^{-ipq/\hbar} dq \qquad (1.0.2)$$

This is called the wave function in the momentum representation, where prepresents the momentum of the particle. It is completely determined by $\psi(q, t)$, which represents the state of the system at time t. It is therefore reasonable to say that $\phi(p, t)$ represents the same dynamical state as $\psi(q, t)$.It is just another way of describing the same state. For the momentum wavefunc-

tion the probability interpretation is that $|\phi(p,t)|^2 dp$ gives the probabilitythat a measurement of the momentum will yield a value between p and $p + dp$. The theory can be developed in an entirely equivalent way in either theposition or the momentum representation. In fact, the representation playsa role analogous to a coordinate system in geometry. Since, in ordinary geometry, problems may be solved by means of vectors, without the use of acoordinate system (and in more generality), it is interesting to ask if quantum mechanics may be formulated without the use of a particular representation.The results would be independent of any particular representation then. The

obvious advantages of using a representation in such a formulation wouldnot be lost, however. A convenient representation could always be used tocarry out a calculation just as a coordinate system may be chosen when vectors are used. This is the goal of the Dirac formulation of quantum mechanics: to develop the theory independent of any specific representation. To see how to go about this program, let us attempt to give a geometrical interpretation to the wave function $\psi(q)$ at time t to take advantage of the concept of vectors. The coordinate q can have any value from $-\infty$ to $+\infty$, as noted earlier. For each specific value, say $q_1, q_2, q_3, \ldots$, the wave function has a

value $\psi(q_1), \psi(q_2), \psi(q_3), \ldots$. We may imagine that an infinite-dimensional space has a set of mutually perpendicular axes each labeled by one of the values of $q(q_1, q_2, q_3, \ldots)$, and that $\psi(q_1)$ is the projection of some vector on the q_1 axis, $\psi(q_2)$ is the projection of the same vector on the q_2 axis, and so on. The



fg.1

Figure 1.1: Ket vector and three of its coordinate representatives.

vector then represents the state of the system just as its components do. This vector is not an ordinary vector since it has a complex character, and we must have a special notation to designate it, just as we do for an ordinary vector. Dirac uses the symbol $|\ \rangle$ to designate a vector of this type and calls it a ket vector, or simply a ket, to distinguish it from ordinary vectors. The particular vector whose components are $\psi(q_1), \psi(q_2), \ldots$ is called ket ψ and written $|\psi\rangle$. Figure 1.1 shows a diagrammatic sketch of the vector $|\psi\rangle$ and its "components" along the mutuallyperpendicular axes described above.

By way of analogy, if A is an ordinary vector and

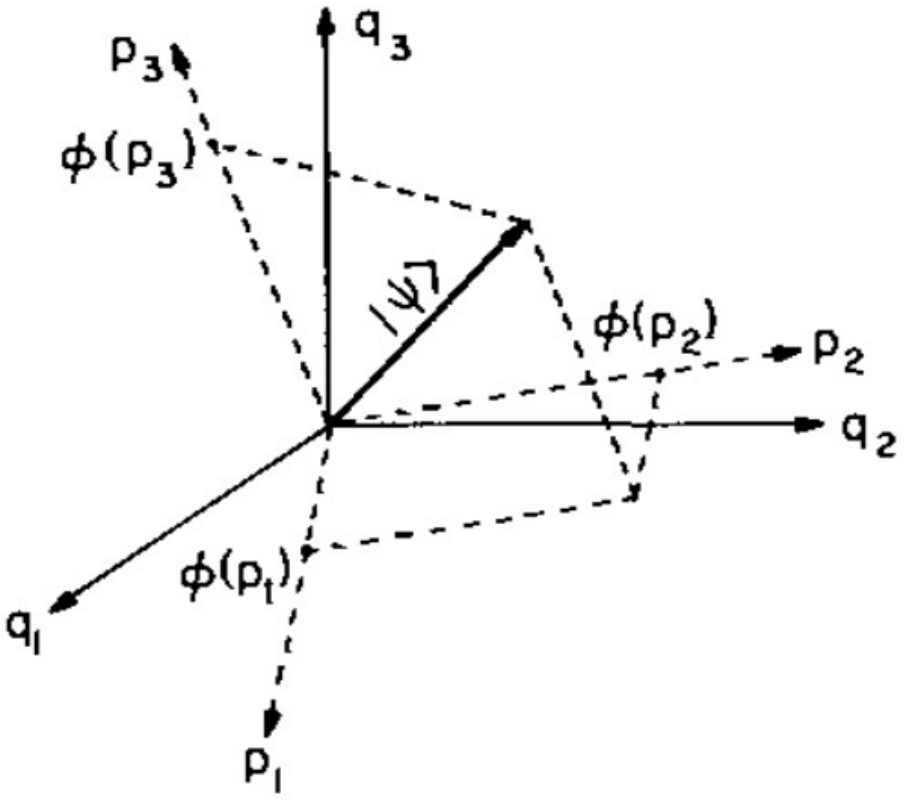
(x, y, z) represent a cartesian coordinate system, A may be specified by giving its components along these axes: A $= (A_x, A_y, A_z)$; that is, A can be represented by its components. Similarly, $|\psi\rangle$ may be specified by its components along the orthogonal q axes: $|\psi\rangle = [\psi(q_1), \psi(q_2), \psi(q_3), \ldots]$. Thus A represents the vector equally as well as its components along certain axes, and $|\psi\rangle$ represents thestate of the system just as well as its components. The vector in this case is said to be given in the position representation. The vector A may also be specified by giving its components along anothercartesian coordinate system (x', y', z') rotated with respect to (x, y, z) :A $= (A_{x'}, A_{y'}, A_{z'})$. So too

$|\psi\rangle$ may be expressed in another representation: $|\psi\rangle = [\phi(p_1), \phi(p_2), \phi(p_3), \dots]$. This is called the momentum representationand is visualized roughly as the components of the same vector on a rotatedorthogonal set of axes; this is shown in Fig. (1.2). The relation between the q and p axes is given by the Fourier transform.

1.1 KET VECTORS

As noted above, Dirac calls vectors designated by the symbol $|a\rangle$, $|x\rangle$, andsuch ket vectors. A general ket is denoted by $|\ \rangle$, and the labels insidedesignate particular kets.

From the discussion above, we associate a ket vector



fg.2

Figure 1.2: Ket vector and three of its coordinate representatives.

with each state ofthe dynamical system under study. Since we shall postulate that a linearsuperposition of states of the system is also a state of the system, the ketvector space must be a linear vector space. A vector space is said to belinear in the following sense. If ct and c2 are complex numbers and $|a\rangle$ and $|b\rangle$ are two kets, the linear combination

$$|u\rangle = c_1|a\rangle + c_2|b\rangle \tag{1.1.1}$$ EQ3

is also a ket vector, since a linear combination of two states associated with $|a\rangle$ and $|b\rangle$ is also a state of the system. If a ket depends on a parameter q', which may take on any value in a certain range, $q'_1 \leq q' \leq q'_2$, we

may generalize (1.1.1) to read

$$|v\rangle = \int_{q_1'}^{q_2'} c(q')|q'\rangle dq' \tag{1.1.2}$$

where $c(q')$ is an ordinary (complex) function of q' and the vector $|v\rangle$ is in ket space. Kets such as $|u\rangle$ (and $|v\rangle$) above are said to be linearly dependent on $|a\rangle$ and $|b\rangle$ (or $|q'\rangle$)). If, in a certain set of ket vectors (two or more), none of them can be expressed as a linear combination of the others, the vectorsare said to be linearly independent.

Although the classical and quantum superposition principles are different,as we shall see below, it may be stated by way of analogy that, if i, j, and k are three

mutually perpendicular unit vectors in ordinary space, any other vector may be written as a linear combination of these three; that is, anyother constant vector A may be written as $\mathbf{A} = c_1\mathbf{j} + c_2\mathbf{j} + c_3\mathbf{k}$. On the otherhand, i cannot be expressed as a linear combination of J and k and is said to be linearly independent of j and k.

Another assumption in the theory is that if a state is superimposed withitself, there results not a new state vector but only the original state again;that is, when $c_1|a\rangle$ and $c_2|a\rangle$ are added, where c_1 and c_2 are arbitrary complex numbers, the result is

$$c_1|a\rangle + c_2|a\rangle = (c_1 + c_2)|a\rangle \tag{1.1.3}$$

and the kets $c_1|a\rangle$, $c_1|a\rangle$, $(c_1 + c_2)|a\rangle$ all represent the same state of thesystem, with the exception of the case $c_1 + c2 = 0$, which corresponds tono state at all. Thus a state is specified entirely by the direction of the ketvector. It may be concluded that $+|a\rangle$ and $-|a\rangle$ represent the same state. Therefore, there is a one-to-one correspondence between a state of a syste-mand a direction in ket vector space. This assumption is a departure fromclassical mechanics and shows that the classical and quantum superpositionprinciples are different.

The ket vector have a finite or an infinite number of-space maydimensions. The dimensionality is deter-

mined by the number of linearlyindependent kets in the space. Since independent states of a quantum systemare represented by independent kets, the dimensionality is determined bythe number of independent states of the quantum system.

1.2 SCALAR PRODUCT; BRA VECTORS

We have introduced ket vectors in an abstract linear vector space by sayingthat their projection on a given set of orthogonal axes in an infinite-dimensional space gives the values of the wave function y)(q, t) in the position representation at time *t*.

The essential definition of kets is that a direction in

ket space andevery state of the system are in one-to-one correspondence.

In the study of ordinary vector analysis, we may define the scalar productof A and B as follows: with every two vectors A and Bin the space, there isassociated a real number *f*, which is written

$$\mathbf{f = A \cdot B} \tag{1.2.1}$$

The scalar product of any two vectors is then defined, since the number toassociate with any pair of them is known. This definition may seem strange at first but a little reflection shows that it is a more general definition thanany formulas we might give for finding

the number f, having been given A and B. One such formula is $f = |A||B|\cos\theta$, where the first two factors arethe magnitudes of A and B, and 6 is the angle between them. But the length itself is defined only in terms of the scalar product of the vector with itself,and so the formula does not serve as an effective definition of a scalarproduct, although it is very useful in practice.

More generally, the scalar product of a particular vector B with all other vectors A in the space may be regarded as a way of defining B. If the set ofnumbers f(B) for all A's is given, B is determined. For three-dimensional space, it is sufficient to take for A

the three unit vectors i, j, and k, which arelinearly independent, and define B by giving its scalar product with each. Thus

$$B_x = \mathrm{B} \cdot \mathrm{i}, \qquad \mathrm{B_y} = \mathrm{B} \cdot \mathrm{j}, \qquad \mathrm{B_z} = \mathrm{B} \cdot \mathrm{k}, \tag{1.2.2}$$

and the three numbers B_x, B_y, and B_z define B.

It is a postulate of the theory of ordinary vectors that the function $f(\mathrm{B})$ a linear function of B. This means that, if $\mathrm{B_1}$ and $\mathrm{B_2}$ are two vectors,

$$\mathrm{A} \cdot (\mathrm{B_1} + \mathrm{B_2}) = \mathrm{A} \cdot \mathrm{B_1} + \mathrm{A} \cdot \mathrm{B_2} \tag{1.2.3}$$

$$\mathrm{A} \cdot (\mathrm{cB}) = \mathrm{c}(\mathrm{A} \cdot \mathrm{B}) \tag{1.2.4}$$

where c is a number. It is clear that the numbers $f(\mathrm{B})$ may be considered a function of B since for every A

there is a number, f(B). This is what is meant by the expression a function $\phi(x)$ of a continuous variable x: with each x is associated a number $\phi(x)$.

After this introduction, we now define scalar products of ketvectors in the following way. With each ket $|a\rangle$ is associated a complexnumber f. (In the examples above the numbers were real but ket vectors aremore general vectors than those in ordinary space.) The set of numbersassociated with different $|a\rangle$'s is a function of $|a\rangle$. This function must be a linear function, which means that if $|a_1\rangle$ and $|a_2\rangle$ are two kets, the number associated with $|a_1\rangle$ and $|a_2\rangle$ is the sum of the numbers associated with $|a_1\rangle$ and $|a_2\rangle$ separately, and the num-

ber associated with $c|a\rangle$, where c is a complex number, is c times the number associated with $|a\rangle$, that is,

$$f(|a_1\rangle + |a_2\rangle) = f(|a_1\rangle) + f(|a_2\rangle) \tag{1.2.5}$$ EQ5

$$f(c|a_1\rangle) = cf(|a_1\rangle) \tag{1.2.6}$$ EQ6

Dirac calls the vectors denotedby the symbol $\langle$ | bra vectors. We may write the scalar product of ($\langle f|$ and $|a\rangle$ as

$$f(|a\rangle) = \langle f|a\rangle \tag{1.2.7}$$ EQ7

If we give all the numbers f for each ket $|a\rangle$, we have defined $\langle f|$. The space of bra vectors is different from the space of ket vectors, just as the reciprocal lattice space was different from the original crystal space.

The definition here is more general, however, because f may be a complex number in (1.2.7) whereas it was real in the crystal example.

When we use the scalar product notation of (1.2.7), we may rewrite (1.2.5, 1.2.6) as

$$\langle f|(|a_1\rangle + |a_2\rangle) = \langle f|a_1\rangle + \langle f|a_2\rangle \tag{1.2.8}$$

$$\langle f|(c|a_1\rangle) = c\langle f|a_1\rangle \tag{1.2.9}$$

Since a bra is defined by its scalar product with a ket, $\langle b| = 0$ if $\langle b|a\rangle = 0$ for every ket $|a\rangle$. Similarly, $\langle b_1| = \langle b_2|$ if $\langle b_1|a\rangle = \langle b_2|a\rangle$ for every $|a\rangle$.The sum of two bras is defined by its scalar product with $|a\rangle$.

Thus

$$(\langle b_1| + \langle b_2|)|a\rangle = \langle b_1|a\rangle + \langle b_2|a\rangle \tag{1.2.10}$$

$$(c\langle b_1|)|a\rangle = c\langle b_1|a\rangle \tag{1.2.11}$$

Thus far we have defined bras only in terms of their scalar products withkets, and there is no definite relation between them. To give a connection,we make the following assumption: each ket may be associated with a singlebra in a unique way; that is, a one-to-one correspondence between kets andbras is assumed. It is therefore reasonable to give the bra the same label asthe ket with which it is associated. Thus $\langle a|$ is the

bra associated with $|a\rangle$ .Similarly, with the ket

$$|u\rangle = |a\rangle + |b\rangle \tag{1.2.12}$$

there is associated the bra

$$\langle u| = \langle a| + \langle b| \tag{1.2.13}$$

and with the ket

$$|v\rangle = c|a\rangle \tag{1.2.14}$$

where c is a complex number, there is associated the bra

$$\langle v| = c^*\langle a| \tag{1.2.15}$$

where c^* is the complex conjugate of c. We shall not go into the reason for taking c^* instead of c but just accept it as a new assumption for simplicity. It is

therefore reasonable to call the bra associated with a ket its hermitianadjoint, and vice versa, and write

$$\langle u| = (|u\rangle)^\dagger, \qquad |u\rangle = (\langle u|)^\dagger, \tag{1.2.16}$$

where the dagger means that the bra is changed to its associated ket (andvice versa) and the complex conjugate of any numbers involved.

Since by assumption there is a unique correspondence between bras andkets, the direction of a bra vector may represent the state of a quantumsystem equally as well as does the direction of a ket. They are said to beduals of one another.

As yet we have not defined the length of a bra or ket.

We shall considertwo kets $|a\rangle$ and $|b\rangle$ and the associated bras $\langle a|$ and $\langle b|$. From these vectorswe may form four numbers $\langle a|b\rangle$, $\langle b|a\rangle$, $\langle a|a\rangle$, and $\langle b|b\rangle$. In general, $\langle a|b\rangle$ and $\langle b|a\rangle$) will be complex, and we make the additional assumption that theyare related by

$$\langle a|b\rangle = \langle b|a\rangle^* \qquad (1.2.17)$$ EQ17

where the asterisk means complex conjugate in the future. With thisassumption, if we let $|b\rangle = |a\rangle$, we conclude that $\langle a|a\rangle$ is real. We define the length,or norm, of $|a\rangle$ as $\langle a|a\rangle$, and so assumption (1.2.17) is necessary if we want thevectors to have a real norm. We make the further assumption that the lengthof a

vector must be positive or zero, that is,

$$\langle a|a\rangle \geq 0 \tag{1.2.18}$$ EQ18

The equality holds only if $|a\rangle = 0$.

The assumptions (1.2.17) and (1.2.18) may be given motivation from aconsideration of the wave function $\psi(q,t)$ and its complex conjugate $\psi^*(q,t)$. We visualized $\psi(q,t)$ as components of $|\psi\rangle$ in ket space. Likewise we mayvisualize $\psi^*(q,t)$ as the components of $\langle\psi|$ in bra space. We then know from wave mechanics that the complex numbers $\psi^*(q,t)\chi(q,t)$ and $\chi^*(q,t)\psi(q,t)$ are related by

$$\psi^*(q)\chi(q) = [\chi^*(q)\psi(q)]^* \tag{1.2.19}$$

and

$$\int_{-\infty}^{+\infty} |\psi(q)|^2 dq \geq 0 \tag{1.2.20}$$

Similar relations should hold for bras and kets since they can be intimatelyrelated to wave functions. This motivated the assumptions (1.2.17) and (1.2.18).

The concept of orthogonality is also important where vectors are concerned. In the case of bras and kets, if the scalar product $\langle a|b\rangle = 0$, the vectors are orthogonal. In wave mechanics, $\psi^*(q)$ and $\chi(q)$ are orthogonal if $\int \psi^*(q)\chi(q)dq = 0$. The orthogonality involved here is different from the orthogonality of two ordinary vectors A and B. If $\mathrm{A} \cdot \mathrm{B} = 0$, A and B are at right angles to one another. But A and B are in the

same vector space. In the present case, $\langle a|$ and $|b\rangle$ are in different vector spaces. (See the crystal-lattice example treated earlier.) Nevertheless, if $\langle a|b\rangle = 0$, it may be said that $|a\rangle$ and $|b\rangle$ are orthogonal as well as $\langle a|$ and $\langle b|$. When $\langle a|b\rangle = 0$, it may also be said that the associated quantum states of the system that they represent are orthogonal.

If the norm of all vectors in the space is finite, the space is called Hilbertspace. The theory must include vectors of infinite norm, as we shall seelater. The space of these vectors forms an even more general vector spacewhich is called ket or bra space. Including vectors of infinite norm requiresthe introduction of the

Dirac δ function at a later stage.

1.3 LINEAR OPERATORS

sec1.3

The concept of linear operators is already familiar to the reader. For example, if $f(t)$ is a square integrable function of a continuous variable t,the function belongs to Hilbert space. We may then define the linearoperator djdt in this space by associating another function $g(t)$ with $f(t)$ and write

$$g(t) = \frac{d}{dt} f(t) \tag{1.3.1}$$

If, with every $f(t)$ in the space, we associate another $g(t)$, we have defined the operator d/dt. If, further-

more, we require that

$$\frac{d}{dt}[f_1(t) + f_2(t)] = g_1(t) + g_2(t) \tag{1.3.2}$$

$$\frac{d}{dt}cf(t) = cg(t) \tag{1.3.3}$$

where $g_1 g_2$, and g are the three functions associated f_1, f_2 and f respectively, and c is a number, then d/dt is a linear operator.

We may similarly define other linear operators such as integration, multiplication by a constant, and many others and build up a whole schemeof linear operators. Clearly, such operators are needed also in vector spaceto extend its range of applicability.

We must now introduce linear operators in the space

of ket and bra vectors. If with each ket $|a\rangle$ in the space we associate another ket $|b\rangle$, the association may be used to define an operator D which we may write in the form

$$|b\rangle = D|a\rangle \tag{1.3.4}$$

where D might mean differentiation, integration, or something else. Note the convention that an operator appears to the left of the ket on which it operates.

We are interested only in linear operators; this means that if $|a_1\rangle, |a_2\rangle$ and $|a\rangle$ are any three kets and c is a number, D must satisfy the relations

$$D(|a_1\rangle + |a_2\rangle) = D|a_1\rangle + D|a_2\rangle \tag{1.3.5}$$ EQ3.5

$$D(c|a_\rangle) = cD|a\rangle \qquad (1.3.6)$$ EQ3.6

Since an operator is completely defined when its effect on every ket in thespace is known, two operators D_1 and $D_2 2$ are equal if $D_1|a\rangle = D_2|a\rangle$ forevery $|a\rangle$. The null operator, $D = 0$, is defined by $D|a\rangle = 0$ for every $|a\rangle$.

The identity operator, $D = I$, is defined by $D|a\rangle = |a\rangle$ for every $|a\rangle$.

At this stage we may build up an algebra of linear operators. We maydefine the sum of two operators $D_1 + D_2$ by their action on $|a\rangle$:

$$(D_1 + D_2)|a\rangle = D_1|a\rangle + D_2|a\rangle \qquad (1.3.7)$$ EQ3.7

a product

$$(\boldsymbol{D_1 D_2})|\boldsymbol{a}\rangle = \boldsymbol{D_1}(\boldsymbol{D_2})|\boldsymbol{a}\rangle \qquad (1.3.8)$$ EQ3.8

From this, if $D_1 = D_2$, we can define powers of operators, and so on.

We also have, for example,

$$(\boldsymbol{D_1} + \boldsymbol{D_2})|\boldsymbol{a}\rangle = (\boldsymbol{D_2} + \boldsymbol{D_1})|\boldsymbol{a}\rangle \qquad (1.3.9)$$ EQ3.9

$$[(\boldsymbol{D_1} + \boldsymbol{D_2}) + \boldsymbol{D_3}]|\boldsymbol{a}\rangle = [(\boldsymbol{D_1} + (\boldsymbol{D_2} + \boldsymbol{D_3})|\boldsymbol{a}\rangle \qquad (1.3.10)$$ EQ3.10

$$[(\boldsymbol{D_1}(\boldsymbol{D_2} + \boldsymbol{D_3})|\boldsymbol{a}\rangle = \boldsymbol{D_1 D_2}|\boldsymbol{a}\rangle + \boldsymbol{D_1 D_3}|\boldsymbol{a}\rangle \qquad (1.3.11)$$ EQ3.11

The commutator of two operators D1 and D2 is written $[D_1, D_2]$ and isdefined by

$$[\boldsymbol{D_1}, \boldsymbol{D_2}] = \boldsymbol{D_1}, \boldsymbol{D_2} - \boldsymbol{D_2}, \boldsymbol{D_1} \qquad (1.3.12)$$

In general, $D_1 D_2 \neq D_2 D_1$, which is a property held in common with matrices. The algebra of quantum mechanics is a noncommutative algebra. Two familiar linear operators that do not commute are $D_1 = x$ (multiplication by x) and $D_2 = d/dx$ (differentiation). It is easily verified that, if $f(x)$ is a continuous function of x,

$$\left[x, \frac{d}{dx}\right] f(x) = \left(x\frac{d}{dx} - \frac{d}{dx}x\right) f(x) = -f(x) \quad (1.3.13)$$

so that noncommutating operators are already familiar.

Multiplication by a constant is a linear operation. A constant operatorcommutes with all linear operators.

If two operators D_1 and D_2 satisfy the equations

$$D_1 D_2 = D_2 D_1 = I \tag{1.3.14}$$

where I is the identity operator, then D_2 is the inverse of D_1 and D_1 is the inverse of D_2, if the inverse exists. This is written as

$$D_2 = D_1^{-1}, \qquad D_1 = D_2^{-1}, \tag{1.3.15}$$

The inverse of a product of operators is

$$(D_1 D_2 D_3)^{-1} = D_3^{-1} D_2^{-1} D_1^{-1} \tag{1.3.16}$$

As noted earlier, these properties of operators are common to finite squarematrices. In fact, later we shall represent operators by matrices.

We have defined the action of linear operators on kets;

we must now givemeaning to their operation on a bra. We shall consider the ket

$$|b\rangle = D|a\rangle \tag{1.3.17}$$

We may take the scalar product of this ket with any bra, say $\langle c|$; this scalar product $\langle c|b\rangle = \langle c|(D|a\rangle)$ depends linearly on $|a\rangle$ since D is linear. From the definition of a bra, the scalar product $\langle c|b\rangle$ may be considered as the scalar product of $|a\rangle$ with some bra, say $\langle d|$. Then for each $\langle c|$ there corresponds abra $\langle d|$. The bra $\langle d|$ depends linearly on $\langle c|$ so that $\langle d|$ is obtained from $\langle c|$ by the application of a linear operator to $\langle c|$. Since this operator is uniquelydetermined by D, we

may reasonably write

$$\langle d| = \langle c|D \tag{1.3.18}$$

We adopt the convention that operators always appear to the right of brasand summarize the definition above by the relation

$$\langle c|(D|a\rangle) = (\langle c|D)|a\rangle \tag{1.3.19}$$

It therefore is unnecessary to use the parentheses, and either side may be written $\langle c|D|a\rangle$. Therefore, D may first operate on $\langle c|$ and the result applied to $|a\rangle$, or vice versa. The operator properties given in (1.3.5) to (1.3.11) are equally valid whether they are applied to bras or kets. Note also that $\langle c|D|a\rangle$ is a closed-

bracket expression and is therefore a complex number ingeneral.

A simple example of a linear operator that occurs frequently in the quantumtheory is $|a\rangle\langle b| = P$. We see that P may operate on a ket to give

$$P|c\rangle = |a\rangle\langle b|c\rangle \tag{1.3.20}$$

which is a ket $|a\rangle$ multiplied by the number $\langle b|c\rangle$, and

$$\langle c|P = \langle c|a\rangle\langle b| \tag{1.3.21}$$

is a bra $\langle b|$ multiplied by the number $\langle c|a\rangle$. It is left as an exercise to show that P satisfies the requirements of a linear operator. An example in ordinaryvector analysis that corresponds approximately to an opera-

tor such as P isthe dyadic ij. In this case, $\mathbf{ij} \cdot \mathbf{k} = 0$, $\mathbf{i} \cdot \mathbf{ij} = \mathbf{j}$, and so on.

Linear operators play a central role in the physical interpretation of thetheory. Following Dirac, we make the assumption that each quantity thatcan be measured for a physical system (which is called a dynamical variable)can be represented by a particular kind of linear operator, to be described inthe following section. Examples of dynamical variables associated with linear operators are position (q), momentum (p), angular momentum L, energy (H), and such which occur in classical mechanics, as well as spin angular momentum (σ) which has no classical analog. Classi-

cally these variables commute with each other, but quantum-mechanically it may be postulated that some of these operators do not commute. The commutation relations determine the type of algebra the operators obey and mark thedeparture of quantum mechanics from classical mechanics.

1.4 HERMITIAN OPERATORS

Linear operators are, in general, complex quantities; if we let themcorrespond to dynamical variables, they would be complex. However, physically,quantities such as momentum, position, and the like give real numbers whenthey are measured. Therefore, the linear

operators that represent dynamicalvariables must be restricted to real linear operators. Such operators are saidto be hermitian and are defined as follows:

The bra associated with the ket $|q\rangle = L|p\rangle$, where L is a linear operator,is written

$$\langle q| = \langle p|L^\dagger = (L|p\rangle)^\dagger = (|q\rangle)^\dagger \tag{1.4.1}$$

The symbol $L^\dagger$ is called the hermitian adjoint of L; that is, the bra $\langle q|$, which is the hermitian adjoint of $|p\rangle$, may be considered the result of some linearoperator acting on $\langle p|$, which is designated by $L^\dagger$.

1.5 Examples

Example: Prove that $\boldsymbol{L^{\dagger\dagger} = L}$

Proof

We let

$$|\boldsymbol{b}\rangle = \boldsymbol{L}|\boldsymbol{p}\rangle \tag{1.5.1}$$ EQ5.1

where $|\boldsymbol{p}\rangle$ is an arbitrary ket. Its adjoint (associated bra) is

$$\langle \boldsymbol{b}| = \langle \boldsymbol{p}|\boldsymbol{L}^{\dagger} \tag{1.5.2}$$

If we take the adjoint again, we obtain

$$|\boldsymbol{b}\rangle = \boldsymbol{L}^{\dagger\dagger}|\boldsymbol{p}\rangle \tag{1.5.3}$$ EQ5.3

If we take the scalar product of an arbitrary bra $\langle a|$ with both (1.5.1) and(1.5.1) we have

$$\langle a|b\rangle = \langle a|L|p\rangle = \langle a|L^{\dagger\dagger}|p\rangle \tag{1.5.4}$$

Since $\langle a|$ and $|p\rangle$ are arbitrary, we have

$$L = L^{\dagger\dagger} \tag{1.5.5}$$

If in (1.2.17), we let $\langle a| = \langle p|L^{\dagger}$ and $|a\rangle = L|p\rangle$ we have

$$\langle p|L^{\dagger}|b\rangle = \langle b|L|p\rangle^{*} \tag{1.5.6}$$ EQ5.6

If a linear operator is self-adjoint,

$$L = L^{\dagger} \tag{1.5.7}$$

the operator is said to be hermitian. From (1.5.13), if L is hermitian, it mustsatisfy the equation

$$\langle p|L|b\rangle = \langle b|L|p\rangle^* \tag{1.5.8}$$

for any $|b\rangle$ and $|p\rangle$. Therefore, any operator that satisfies (1.5.8) is hermitian.The following properties may be proved for any linear operator:

$$(cL|a\rangle)^\dagger = c^*\langle a|L^\dagger \tag{1.5.9}$$

where c is a constant,

$$[(L_1 + L_2)|a\rangle]^\dagger = \langle a|(L_1^\dagger + L_2^\dagger) \tag{1.5.10}$$

$$(L_1L_2)|a\rangle]^\dagger = \langle a|L_2^\dagger L_1^\dagger \tag{1.5.11}$$

$$(\langle a|L_1L_2L_3)^\dagger = L_3^\dagger L_2^\dagger l_1^\dagger|a\rangle \tag{1.5.12}$$

$$\langle a|L_1 L_2|b\rangle^* = \langle b|L_2^\dagger L_1^\dagger|a\rangle \qquad (1.5.13)$$ EQ5.6

$$(|a\rangle\langle b|)^\dagger = |b\rangle\langle a| \qquad (1.5.14)$$

The algebra of adjoints of operators is the same as for finite square matrices.

1.6 THE EIGENVALUE PROBLEM

Bras and ket vectors, or rather directions of bras and kets, are associated with states of a system, and linear hermitian operators are associated with dynamical variables that describe the system. In a next section we show howto relate these mathematical concepts to physical measurements made onthe system. Before this, we must introduce the concept of eigenvalues of

hermitian operators.

An eigenvalue problem is a familiar one in classical mathematics as well asin classical physics. One of the simplest examples involves the solution of the equation

$$\boldsymbol{Lu(x) = \lambda u(x)} \tag{1.6.1}$$

where L is known to be $|d^2/dx^2$ and $u(x)$ and λ are unknown. If we add the boundary conditions that $u(0) = u(l) = 0$, we find that λ can take on only a certain discrete set of eigenvalues given by $\lambda_n = \pi^2 n^2/l2$, where $n = 0, \pm 1, \pm 2, \ldots$. The associated eigenfunctions $u_n(x)$ are $u_n(x) = \sin(n\pi x/l)$. Note that the effect of an operator L on an eigenfunction $u_n(x)$ is to

reproduce $u_n(x)$. If L operates on an arbitrary function $u(x)$, it will not,in general, reproduce $u(x)$ times a number.

We may similarly formulate an eigenvalue problem for operators in ket (and bra) space. We let L be a linear operator and $|a\rangle$ a ket. If L operates on $|a\rangle$ and gives $|a\rangle$ multiplied by a number l, then $|a\rangle$ is an eigenket of L and l is the associated eigenvalue. This may be written

$$L|a\rangle = l|a\rangle \tag{1.6.2}$$

This is an eigenvalue problem: L is a known operator, and l and $|a\rangle$ are unknown, and we are asked to find kets which, when acted on by L,reproduce the same

ket times a number subject to a set of boundary conditions.It is customary to label an eigenket with its eigenvalue; with this convention,we may rewrite the eigenvalue equation as

$$\boldsymbol{L}|l\rangle = l|l\rangle \tag{1.6.3}$$ EQ6.3

The eigenvalue problem may equally well be formulated in terms of bras:

$$\langle d|\boldsymbol{D} = d\langle d| \tag{1.6.4}$$

For simplicity, in this book we shall usually consider cases in which thereis only one eigenvalue for each eigenvector. If more than one independent eigenvector can be associated with a given eigenvalue, the sys-

tem is saidto be degenerate.

If $|l\rangle$ is an eigenket of $\boldsymbol{L}$, then, by (1.6.3), any constant c times $|l\rangle$ is also an eigenket with the same eigenvalue. In line with earlier assumptions, the states represented by $|l\rangle$ and $c|l\rangle$ are the same state.

We shall be interested usually in the solution of the eigenvalue problem forlinear hermitian operators for reasons that should become clear in the nextsection. Before attempting the solutions of any specific eigenvalue problem, we shall prove two very important theorems valid for all linear hermitian operators.

1.7 THEOREMS

sec.1.7

Theorem 1

The eigenvalues of a linear hermitian operator are real.

Proof

We let L be a linear hermitian operator. The eigenvalues of L satisfy the equation

$$L|l\rangle = l|l\rangle \tag{1.7.1}$$ EQ7.1

If we form the scalar product of both sides with $\langle l|$, we have

$$\langle l|L|l\rangle = l\langle l|l\rangle \tag{1.7.2}$$ EQ7.2

If we take the complex conjugate of both sides, we obtain,

$$\langle l|\boldsymbol{L}|l\rangle^* = \langle l|\boldsymbol{L}^\dagger|l\rangle = l^*\langle l|l\rangle \qquad (1.7.3)$$ EQ7.3

But since $\boldsymbol{L}^\dagger = \boldsymbol{L}$, and $\langle l|l\rangle \neq 0$, comparison of (1.7.3) and (1.7.3 shows that $l = l^*$, and the theorem is proved. We see that $\langle l|l\rangle = 0$ only in the trivial case, in which$|l\rangle = 0$. Note that the norm $\langle l|l\rangle$ is real.

Theorem 2

Two eigenvectors of a linear hermitian operator $\boldsymbol{L}$ belonging to different eigenvalues are orthogonal.

Proof

We let l' and l'' be two eigenvalues of L and $|l'\rangle$ and $|l''\rangle$ be the associated eigenkets.Then ($L = L^\dagger; l'$ and l'' are real).

$$L|l'\rangle = l'|l'\rangle \tag{1.7.4}$$ EQ7.4

$$\langle l''|L = l''\langle l''| \tag{1.7.5}$$ EQ7.5

If we form the scalar product of (1.7.4) with $\langle l''|$, the scalar product of (1.7.5) with $|l'\rangle$, and subtract, we find that

$$(l' - l'')\langle l'|l''\rangle = 0 \tag{1.7.6}$$

Since $l' \neq l''$ by assumption, then $\langle l'|l''\rangle = 0$, and the theorem is proved. From (1.7.4) and (1.7.5) we see that the eigenvalues associated with eigenkets are the

sameas those associated with eigenbras.

The solution of an eigenvalue problem in many cases is complicated. Weshall now solve a particularly simple one to illustrate the method. Later, we may discuss a physical system that may be described by this exam-plebut for the moment we shall consider it merely as a mathematical example.

We suppose a linear hermitian operator $\boldsymbol{\sigma_z}$ that satisfies an auxiliary condition

$$\boldsymbol{\sigma_z^2 = I} \tag{1.7.7}$$ EQ7.7

where $\boldsymbol{I}$ is the identity operator, and we wish to solve the eigenvalue problem

$$\boldsymbol{\sigma_z|s\rangle = s|s\rangle} \tag{1.7.8}$$ EQ7.8

From Theorem 1, we know that s is real, and from Theorem 2 we know that $\langle s'|s''\rangle = 0$ $s' \neq s''$.

To solve for the eigenvalues and eigenvectors, we multiply both sides of (1.7.8) from the left by σ_z, use (1.7.7) and (1.7.8), and obtain

$$\boldsymbol{\sigma_z^2|s\rangle = |s\rangle = s\sigma_z|s\rangle = s^2|s\rangle} \tag{1.7.9}$$

or

$$\boldsymbol{(s^2 - 1)|s\rangle = 0} \tag{1.7.10}$$

If we form the scalar product of this with $\langle s|$, we see that, since $\langle s|s\rangle$ is positive and not zero, the eigenvalues of az are given by

$$s = \pm 1 \tag{1.7.11}$$

Since by assumption there can be no degeneracy (two eigenvalues the same),there are only two eigenvalues, and so we may rewrite (1.7.8) as

$$\sigma_z|+1\rangle = +1|+1\rangle, \qquad \sigma_z|-1\rangle = -1|-1\rangle \tag{1.7.12}$$ EQ7.12

By Theorem 2 we know that

$$\langle +1|-1\rangle = \langle -1|+1\rangle = 0 \tag{1.7.13}$$ EQ7.13

These are the orthogonality relations obeyed by eigenvectors belonging todifferent eigenvalues.

As we know, any eigenket multiplied by a constant is also an eigenketbelonging to the same eigenvalue. We may therefore choose a constant sothat the norm of the eigenvectors is unity as long as the norm is finite and write

$$\langle +1|+1\rangle = \langle -1|-1\rangle = 1 \qquad (1.7.14)$$ EQ7.14

These are the normalization conditions. Normalization does not specify the vector uniquely; we may still multiply $|+1)$ by $e^{i\alpha}$ since $\langle +1|$ will bemultiplied by $e^{-i\alpha}$, where α is real, and (1.7.14) will be left unchanged. Such a phase shift is of no physical significance in the theory, and we shallusually choose $\alpha = 0$.

For any eigenvalue problem in which the norm of the vectors is finite, theeigenvectors may always be normalized and (1.7.13) and (1.7.14) combined into the orthogonality relations

$$\langle l'|l''\rangle = \delta_{l'l''} \tag{1.7.15}$$

where $\delta_{l'l''}$ is the Kronecker δ defined by

$$\delta_{ij} = \begin{cases} 1 & if \quad i = j \\ 0 & if \quad i \neq j \end{cases} \tag{1.7.16}$$

When the vectors have an infinite norm, these results have to be generalized,as we discuss later.

Anticipating future work, we shall now show that az may be representedby a 2×2 matrix given by

$$\sigma_z = \begin{bmatrix} 1 & 0 \\ \\ 0 & 1 \end{bmatrix} \tag{1.7.17}$$

To show this, we form the scalar products of both equations (1.7.12) with $\langle +1|$ and $\langle -1|$, respectively. If we use (1.7.13) and (1.7.14), we obtain theso-called "matrix elements" of $\boldsymbol{\sigma}_z$ given by

$$\langle +1|\boldsymbol{\sigma}_z|+1\rangle = 1 \quad \langle +1|\boldsymbol{\sigma}_z|-1\rangle = 0 \tag{1.7.18}$$

$$\langle -1|\boldsymbol{\sigma}_z|+1\rangle = 0 \quad \langle -1|\boldsymbol{\sigma}_z|-1\rangle = -1 \tag{1.7.19}$$

We then may group these results into a matrix such as (1.7.17), with the convention that the rows are labeled by the eigenbras and the columns bythe eigenkets.

Any ket in the space may be expressed in terms of the eigenkets $|+1\rangle$ and $|-1\rangle$. When this can be done, it is said that the eigenkets form a completeset by def-

inition. Again we are anticipating the results of the next section.

To show that any ket $|P\rangle$ in the space may be expanded in terms of $|+1\rangle$ and $|+1\rangle$, we write the identity

$$|P\rangle = I|P\rangle = \frac{1}{2}(I + \sigma_z + I - \sigma_z)|P\rangle \qquad (1.7.20)$$

$$= \frac{1}{2}(I + \sigma_z)|P\rangle + \frac{1}{2}(I - \sigma_z)|P\rangle \qquad (1.7.21)$$ EQ7.21

We consider each factor separately. On using (1.7.7), we have

$$\sigma_z\Big[\frac{1}{2}(I + \sigma_z)|P\rangle\Big] = 1\Big[\frac{1}{2}(\sigma_z + I)|P\rangle\Big] \qquad (1.7.22)$$

so that $\frac{1}{2}(I + \sigma_z)|P\rangle$ is an eigenket of σ_z with eigenvalue +1. It may therefore differ from $|+1\rangle$ only by

a constant, and we may write

$$\frac{1}{2}(I+\sigma_z)|P\rangle = c_1|+1\rangle \qquad (1.7.23)$$ EQ7.23

where c_1 is a constant. Similarly, we see that the last term in (1.7.21) is given by

$$\sigma_z\Big[\frac{1}{2}(I-\sigma_z)|P\rangle\Big] = -1\Big[\frac{1}{2}(I-\sigma_z)|P\rangle\Big] \qquad (1.7.24)$$

so that we may write

$$\frac{1}{2}(I-\sigma_z)|P\rangle = c_2|-1\rangle \qquad (1.7.25)$$ EQ7.25

where c_2 is another constant. Thus (1.7.21) may be written by using (1.7.23) and (1.7.25) as

$$|P\rangle = c_1|+1\rangle + c_2|-1\rangle \qquad (1.7.26)$$ EQ7.26

as originally stated. Any ket is therefore linearly dependent on the kets $|+1\rangle$ and $|-1\rangle$, and we have

proved that the set $\{|+1\rangle, |-1\rangle\}$ is complete.

We may also derive the so-called completeness relation in this simple example. We multiply (1.7.26) from the left alternatively by $\langle +1|$ and $\langle -1|$,use the orthonormality relations (1.7.13) and (1.7.14), and see that

$$c_1 = \langle +1|P\rangle \qquad c_2 = \langle -1|P\rangle \tag{1.7.27}$$

If we substitute these into (1.7.26), we obtain

$$|P\rangle = \Big(|+1\rangle\langle +1| + |-1\rangle\langle -1|\Big)|P\rangle \tag{1.7.28}$$

Since $|P\rangle$ is arbitrary, this equation will be satisfied if

$$|+1\rangle\langle +1| + |-1\rangle\langle -1| = I \tag{1.7.29}$$

which is the completeness or closure relation. We discuss the completenessrelation for general hermitian

operators in the next section.

The Hilbert space in this example is two-dimensional because weconsidered only nondegenerate eigenvalues. If we substitute (1.7.27) in (1.7.23) and (1.7.25), we have the results

$$\frac{1}{2}(\boldsymbol{I}+\boldsymbol{\sigma}_z)|\boldsymbol{P}\rangle = |+1\rangle\langle+1| \qquad (1.7.30)$$

$$\frac{1}{2}(\boldsymbol{I}-\boldsymbol{\sigma}_z)|\boldsymbol{P}\rangle = |-1\rangle\langle-1| \qquad (1.7.31)$$

We may subtract these equations to obtain $\boldsymbol{\sigma}_z$:

$$\boldsymbol{\sigma}_z = |+1\rangle\langle+1| - |-1\rangle\langle-1| \qquad (1.7.32)$$

so that we have expressed az in terms of operators of the type $|a\rangle\langle a|$ mentioned near the end of Section (1.3). Eigenvalues of an operator are sometimesre-

ferred to as its spectrum.

Chapter 2

ORBITAL ANGULAR MOMENTUM; ELECTRON SPIN

2.1 EIGENVALUES AND EIGENVECTORS OF ANGULAR MOMENTUM

Befpre we proceed, it essential to note that, if a single measurement of an observable is made, one of its eigenvalues is obtained. The ability to solve eigenvalue problems is therefore essential torelate the theory to experiment. Thus far we have solved only one such problem, in Section sec.1.7 **1.7, where the observable satis-**

fied the algebraic equationa $\sigma_z^2 = 1$. In that case the Hilbert space consisted of only two eigenvectors,and the eigenvalue spectrum had only the two discrete values $+1$ and -1.

Now for a single particle of mass m constrained to move in one dimension in a field of force. Classically, this system may be described completely by a position coordinate q and a momentum p. If we specify both these quantities at a certain time, we have specified the classical state of thesystem at this time.

Alternatively, to treat this system quantum-mechanically, according to the theory wehave developed thus far, we associate with each of these dynamical variables(since

they are observable) a linear hermitian operator which we shall call q and p. As operators, they satisfy

$$p = p^\dagger \qquad q = q^\dagger \tag{2.1.1}$$

After the operators needed to describe the physical system are enumerated,the next step in setting up the quantum problem is to specify the algebra that the operators must obey. This requires an additional postulate for the theory;it is given in terms of the commutation relations for p and q, namely,

$$[q, q] = [p, p] = 0 \tag{2.1.2}$$ EQ2.1.2

$$[q, p] = qp - pq = i\hbar \tag{2.1.3}$$ EQ2.1.3

where $\hbar$ is Planck's constant divided by 2π; that is, we postulate that p and q satisfy the commutation relations above. Classically, q and p commute sothat, to the extent they do not commute, the quantum and classical systems differ. Accordingly, the classical system is quantized when the observablesq and p satisfy (2.1.2, 2.1.3). The justification for the quantum postulate is theremarkable agreement between theory and experiment. It is possibly the mostprofound and fundamental postulate in the theory.

If $\hbar \to 0$, q and p will commute so that classical mechanics should becontained in the quantum formulation in the limit as $\hbar \to 0$. This is just the correspon-

dence principle.

It is said that q and p obey a noncommutative algebra. Before proceeding,let us develop a few useful algebraic relations. If l is an integer, we mayprove by mathematical induction from (2.1.2, 2.1.3) that

$$[q, p^l] = i\hbar l p^{l-1} = i\hbar \frac{\partial}{\partial p} p^l \tag{2.1.4}$$ EQ2.1.4

$$[p, q^l] = -i\hbar l q^{l-1} = -i\hbar \frac{\partial}{\partial q} q^l \tag{2.1.5}$$ EQ2.1.5

Bearing in mind these above attributes of the operators q and p, we will be ready to proceed towards the next subjet, namely, Orbital angular momentum. Orbital angular momentum plays an essential role in quantum mechanics just as it does in classical mechan-

ics. Classically, the angular momentum about a point 0 is defined by

$$\mathbf{I} = \mathbf{r} \times \mathbf{p} \tag{2.1.6}$$ EQ2.1.6

where r is the radius vector from 0 to the particle and p is its linear momentum. Since I is an observable, we postulate that I is a hermitian operator defined by (2.1.6) where r and p are the coordinate and momentum operators. We'et $[q_1, q_2, q_3]$ be the three coordinate operators corresponding to r and $[p_1, p_2, p_3] =$ p be the corresponding momentum operators. We postulate as in (2.1.2, 2.1.3) that these operators obey the

commutation relations

$$[q_i, p_j] = i\hbar\delta_{ij}; \quad [q_i, q_j] = [p_i, p_j] = 0 \tag{2.1.7}$$ EQ2.1.7

where i and $j = 1, 2,$ or 3. This says that q_1 and p_2 commute, for example. In other words, measurements of a coordinate in one direction does notinterfere with the measurement of the momentum in an orthogonal directionas it does in the same direction.

From (2.1.6) and (2.1.7), we see that

$$l_1 = q_2 p_3 - q_3 p_2 \tag{2.1.8}$$ EQ2.1.8

$$l_2 = q_3 p_1 - q_1 p_3 \tag{2.1.9}$$ EQ2.1.9

$$l_3 = q_1 p_2 - q_2 p_1 \tag{2.1.10}$$ EQ2.1.10

If we use (2.1.7), we may easily show that

$$[l_x, l_y] = i\hbar l_z \tag{2.1.11}$$ EQ2.1.11

$$[l_y, l_z] = i\hbar l_x \tag{2.1.12}$$ EQ2.1.12

$$[l_z, l_x] = i\hbar l_y \tag{2.1.13}$$ EQ2.1.13

so that no additional postulates are needed to quantize

I. Note also since, for example, q_2 and p_3 commute as do q_3 and p_2. we do not have to worry aboutordering of the separate factors in l_1, l_2, and l_3.

The total angular momentum is

$$\mathbf{I}^2 = l_1^2 + l_3^2 + l_3^2 \tag{2.1.14}$$

Also,

$$[\mathbf{I}^2, l_i] = 0 \qquad \textbf{Try to prove} \tag{2.1.15}$$ EQ2.1.15

That is, each component of the angular momentum separately commutes with $\mathbf{I}^2$.

It is convenient to define two nonhermitian operators $l_\pm$ by

$$l_\pm = l_1 \pm i l_2 \tag{2.1.16}$$

or

$$l_1 = \frac{1}{2}(l_- + l_+) \tag{2.1.17}$$ EQ2.1.17

$$l_2 = \frac{1}{2} i (l_- - l_+) \tag{2.1.18}$$ EQ2.1.18

Since l_1 and l_2 are hermitian it follows that

$$l_+ = l_-^\dagger \tag{2.1.19}$$

In terms of $l_\pm$, we see that

$$\mathbf{I}^2 = l_3^2 + \frac{1}{2}(l_+ l_- + l_- l_+) \tag{2.1.20}$$ EQ2.1.20

We will leave as an exercise to show that

$$[\mathbf{I}^2, \boldsymbol{l}_\pm] = \mathbf{0} \tag{2.1.21}$$ EQ2.1.21

$$[\boldsymbol{l}_z, \boldsymbol{l}_\pm] = \pm\hbar \boldsymbol{l}_\pm \tag{2.1.22}$$ EQ2.1.22

$$[\boldsymbol{l}_+, \boldsymbol{l}_-] = \mathbf{2}\hbar \boldsymbol{l}_\mathbf{3} \tag{2.1.23}$$ EQ2.1.23

If we alternatively add and subtract (2.1.23) from (2.1.20) we obtain

$$\boldsymbol{l}_+\boldsymbol{l}_- = \mathbf{I}^2 - \boldsymbol{l}_\mathbf{3}^2 + \hbar \boldsymbol{l}_\mathbf{3} \tag{2.1.24}$$ EQ2.1.24

$$\boldsymbol{l}_-\boldsymbol{l}_+ = \mathbf{I}^2 - \boldsymbol{l}_\mathbf{3}^2 - \hbar \boldsymbol{l}_\mathbf{3} \tag{2.1.25}$$ EQ2.1.25

We have shown that $\mathbf{I}^2$ commutes with $\boldsymbol{l}_1, \boldsymbol{l}_2$, and $\boldsymbol{l}_3$ but the components donot commute with each other. We have shown that any two operators thatcommute may be simultaneously diagonalized. We may there-

fore obtain a representation in which both are diagonal; that is, we may find eigenvectors which are simultaneous eigenvectors of two commuting operators. Considerthe eigenvalue problems

$$l_z|\mu;\nu\rangle = \mu\hbar|\mu;\nu\rangle \tag{2.1.26}$$ EQ2.1.26

$$\mathbf{I}^2|\mu;\nu\rangle = \nu\hbar^2|\mu;\nu\rangle \tag{2.1.27}$$ EQ2.1.27

Since $[l_z, \mathbf{I}^2] = 0$ we see that

$$\mathbf{I}^2 l_z|\mu;\nu\rangle = \mu\hbar\mathbf{I}^2|\mu;\nu\rangle = \mu\nu\hbar^3|\mu;\nu\rangle \tag{2.1.28}$$ EQ2.1.28

$$= l_z\mathbf{I}^2|\mu;\nu\rangle \tag{2.1.29}$$ EQ2.1.29

We wish to obtain the eigenvalues μ and ν by techniques similar to those used for the harmonic oscillator.

By (2.1.21) we have

$$\mathbf{I}^2 \boldsymbol{l}_\pm = \boldsymbol{l}_\pm \mathbf{I}^2 \tag{2.1.30}$$

By (2.1.26) we see that

$$\mathbf{I}^2\{\boldsymbol{l}_\pm|\boldsymbol{\mu};\boldsymbol{\nu}\rangle\} = \boldsymbol{\nu}\hbar^2\{\boldsymbol{l}_\pm|\boldsymbol{\mu};\boldsymbol{\nu}\rangle\} \tag{2.1.31}$$ EQ2.1.31

This says that if $|\boldsymbol{\mu};\boldsymbol{\nu}\rangle$ is an eigenket of $\mathbf{I}^2$ with eigenvalue $\boldsymbol{\nu}\hbar^2$, then $\boldsymbol{l}_+|\boldsymbol{\mu};\boldsymbol{\nu}\rangle$ and $\boldsymbol{l}_-|\boldsymbol{\mu};\boldsymbol{\nu}\rangle$ are also eigenkets with the same eigenvalue.

Consider next (2.1.22):

$$\boldsymbol{l}_z \boldsymbol{l}_\pm = \boldsymbol{l}_\pm \boldsymbol{l}_z + \hbar \boldsymbol{l}_\pm \tag{2.1.32}$$ EQ2.1.32

It therefore follows from (2.1.26,2.1.27) that

$$\boldsymbol{l}_z\{\boldsymbol{l}_\pm|\boldsymbol{\mu};\boldsymbol{\nu}\rangle\} = (\boldsymbol{\mu} \pm 1)\hbar\{\boldsymbol{l}_\pm|\boldsymbol{\mu};\boldsymbol{\nu}\rangle\} \tag{2.1.33}$$ EQ2.1.33

Thus if $|\mu;\nu\rangle$ is an eigenvector of l_z with eigenvalue $\mu\hbar$ then $l_+|\mu;\nu\rangle$ is an eigenvector of l_z with eigenvalue $(\mu+1)\hbar$ and $l_-|\mu;\nu\rangle$ is an eigenvector of l_z with eigenvalue $(\mu-1)\hbar$, both with the same ν by (2.1.31). We have thus generated two additional eigenvectors of l_z from the original whose eigenvalues differ by $\pm\hbar$. We may obviously continue this process and obtain the infinite sequence

$$|\mu;\nu\rangle \quad l_+|\mu;\nu\rangle \quad l_+^2|\mu;\nu\rangle \quad l_+^3|\mu;\nu\rangle \ldots \tag{2.1.34}$$

$$\mu\hbar \quad (\mu+1)\hbar \quad (\mu+2)\hbar \quad (\mu+3)\hbar \ldots \tag{2.1.35}$$

$$|\mu;\nu\rangle \quad l_-|\mu;\nu\rangle \quad l_-^2|\mu;\nu\rangle \quad l_-^3|\mu;\nu\rangle \ldots \tag{2.1.36}$$

$$\mu\hbar \quad (\mu-1)\hbar \quad (\mu-2)\hbar \quad (\mu-3)\hbar \ldots \tag{2.1.37}$$

where ν is unchanged.

Since the norm of a vector must be greater than or equal to zero, we assume that

$$\langle \mu; \nu | \mu; \nu \rangle > 0 \tag{2.1.38}$$ EQ2.1.38

That is, the original vector exists. Then since $l_- = l_+^\dagger$, we have

$$\langle \mu; \nu | l_- l_+ | \mu; \nu \rangle = \langle \mu; \nu | (\mathbf{I}^2 - l_z^2 - \hbar l_z) | \mu; \nu \rangle \tag{2.1.39}$$

$$= \hbar^2 (\nu - \mu^2 - \mu) \langle \mu; \nu | \mu; \nu \rangle \geq 0 \tag{2.1.40}$$

where we used (EQ2.1.25 2.1.25) and (EQ2.1.26 2.1.26, EQ2.1.27 2.1.27). By (EQ2.1.38 2.1.38), it follows that

$$\nu - \mu^2 - \mu \geq 0 \tag{2.1.41}$$ EQ2.1.41

where ν is fixed. This tells us that for any given ν if μ gets arbitrarily big either positively or negatively the vector $l_+|\mu;\nu\rangle$ would develop a negative norm which is forbidden. We must therefore anticipate an upper and lower bound on μ for each ν. The equality is satisfied above for

$$\mu = -\frac{1}{2} \pm \sqrt{\frac{1}{4} + \nu} \tag{2.1.42}$$

which are the two bounds on μ for fixed ν. Let l be the largest value μ may have so that

$$l \equiv -\frac{1}{2} + \sqrt{\frac{1}{4} + \nu} \tag{2.1.43}$$

or

$$\nu = l(l+1) \tag{2.1.44}$$

When $\mu = l$, we have

$$l_+|l;\nu\rangle = 0 \tag{2.1.45}$$ EQ2.1.45

since otherwise we would generate an eigenvector with eigenvalue $\mu = l + 1$ which would violate (2.1.41).

If we start with state $|l;\nu\rangle$ and apply l_- k times then we generate state $|l-k;\nu\rangle$ by (2.1.34-2.1.37). The length of $l_-|l-k;\nu\rangle$ is

$$\langle l-k;\nu|l_+l_-|l-k;\nu\rangle = \langle l-k;\nu|(\mathbf{l}^2 - l_3^2 - \hbar l_3)|l-k;\nu\rangle$$

$$(2.1.46)$$ EQ2.1.46

$$= \{\nu-(l-k)^2+(l-k)\}\hbar^2\langle l-k;\nu|l-k;\nu\rangle \geq 0 \tag{2.1.47}$$

If $|l-k;\nu\rangle \neq 0$, then since $\nu = l(l+1)$, we conclude

$$l(l+1) - (l-k)^2 + (l+k) \geq 0 \tag{2.1.48}$$

This puts a limit on the size of k

$$(l-k) = \frac{1}{2} \pm \sqrt{\frac{1}{4} + l(l+1)} \tag{2.1.49}$$

or k_{max} is determined by

$$l - k_{max} = \frac{1}{2} - \sqrt{\frac{1}{4} + l(l+1)} = \frac{1}{2} - \sqrt{\left(l + \frac{1}{2}\right)^2} \tag{2.1.50}$$

so that

$$k_{max} = 2l \tag{2.1.51}$$

But k is a positive integer so $2l$ must be a positive integer. Therefore, l may have only the values

$$l = 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \frac{5}{2}, \ldots \tag{2.1.52}$$

When $k_{max} = 2l$, we conclude that

$$l_-|l - k_{max}; \nu\rangle = l_-| - l; \nu\rangle = 0 \tag{2.1.53}$$ EQ2.1.53

Therefore μ ranges between $+1$ and $|1$ in unit steps. It is conventional to let $\mu = m$ and to designate ν by l since $\nu = l(l+1)$ and write

$$l_3|m;l\rangle = m\hbar|m;l\rangle \tag{2.1.54}$$ EQ2.1.54

$$I^2|m;l\rangle = l(l+1)\hbar^2|m;l\rangle \tag{2.1.55}$$ EQ2.1.55

where the eigenvalues are

$$l = 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \frac{5}{2}, \ldots \tag{2.1.56}$$

$$m = -l, -l+1, -l+2, \ldots, l-2, l-1, l \tag{2.1.57}$$

The eigenvectors are orthogonal since I^2 and l_3 are hermitian so that

$$\langle m';l'|m;l\rangle = \delta_{ll'}\delta_{mm'} \tag{2.1.58}$$ EQ2.1.58

so that the matrix elements of l_3 and I^2 are

$$\langle m'; l'|l_3|m; l\rangle = m\hbar\delta_{ll'}\delta_{mm'} \qquad (2.1.59)$$ EQ2.1.59

$$\langle m'; l'|I^2|m; l\rangle = l(l+1)\hbar^2\delta_{ll'}\delta_{mm'} \qquad (2.1.60)$$ EQ2.1.60

Let us next obtain the matrix elements of $l_\pm$ in the representation in which l_3 and I^2 are diagonal. We have shown that $l_+|m; l\rangle$ is an eigenvector of I^2 with eigenvalue $l(l+1)\hbar^2$ and also an eigenvector of l_3 with eigenvalue $(m+1)\hbar$. Therefore, $l_+|m; l\rangle$ can differ from $|m+1; l\rangle$ by a complex constant. We may therefore write

$$l_+|m; l\rangle = \lambda_{l,m}\hbar|m+1; l\rangle \qquad (2.1.61)$$ EQ2.1.61

So that

$$\langle \boldsymbol{m+1; l}|\boldsymbol{l_+}|\boldsymbol{m; l}\rangle = \boldsymbol{\lambda_{l,m}}\hbar \tag{2.1.62}$$

If we take the complex conjugate of both sides, we have

$$\langle \boldsymbol{m; l}|\boldsymbol{l_-}|\boldsymbol{m+1; l}\rangle = \boldsymbol{\lambda^*_{l,m}}\hbar \tag{2.1.63}$$

This relation is satisfied if

$$\boldsymbol{l_-}|\boldsymbol{m+1; l}\rangle = \boldsymbol{\lambda^*_{l,m}}\hbar|\boldsymbol{m; l}\rangle \tag{2.1.64}$$ EQ2.1.64

Consider next

$$\boldsymbol{l_-l_+}|\boldsymbol{m; l}\rangle = \boldsymbol{l_-\lambda_{l,m}}\hbar|\boldsymbol{m+1; l}\rangle \tag{2.1.65}$$

$$= \hbar^2|\boldsymbol{\lambda_{l,m}}|^2|\boldsymbol{m; l}\rangle \tag{2.1.66}$$

$$= [\boldsymbol{I^2 - l_3^2 - \hbar l_3}]|\boldsymbol{m; l}\rangle \tag{2.1.67}$$

$$= [\boldsymbol{l(l+1) - m^2 - m}]\hbar^2|\boldsymbol{m; l}\rangle \tag{2.1.68}$$

where we used (2.1.61), (2.1.64), (2.1.24), and (2.1.55).

Therefore, we conclude that

$$\boldsymbol{\lambda_{l,m} = \sqrt{l(l+1) - m(m+1)}} \tag{2.1.69}$$

$$\boldsymbol{= \sqrt{(l-m)(l+m+1)}} \tag{2.1.70}$$

and

$$\boldsymbol{l_+|m;l\rangle = \hbar\sqrt{(l-m)(l+m+1)}|m+1;l\rangle} \tag{2.1.71}$$

$$\boldsymbol{l_-|m;l\rangle = \hbar\sqrt{(l-m+1)(l+m)}|m-1;l\rangle} \tag{2.1.72}$$

so the matrix elements are

$$\boldsymbol{\langle m';l'|l_+|m;l\rangle = \hbar\sqrt{(l-m)(l+m+1)}\delta_{ll'}\delta_{m'm+1}}$$

(2.1.73) EQ2.1.73

$$\boldsymbol{\langle m';l'|l_-|m;l\rangle = \hbar\sqrt{(l-m+1)(l+m)}\delta_{ll'}\delta_{m'm-1}}$$

(2.1.74) EQ2.1.74

If we use (EQ2.1.17 2.1.17, EQ2.1.18 2.1.18), we obtain the nonvanishing matrix elements

$$\langle m+1; l|l_1|m; l\rangle = \frac{\hbar}{2}\sqrt{(l-m)(l+m+1)} \quad (2.1.75)$$ EQ2.1.75

$$\langle m-1; l|l_1|m; l\rangle = \frac{\hbar}{2}\sqrt{(l-m+1)(l+m)} \quad (2.1.76)$$ EQ2.1.76

$$\langle m+1; l|l_2|m; l\rangle = -i\frac{\hbar}{2}\sqrt{(l-m)(l+m+1)} \quad (2.1.77)$$ EQ2.1.77

$$\langle m-1; l|l_2|m; l\rangle = i\frac{\hbar}{2}\sqrt{(l-m+1)(l+m)} \quad (2.1.78)$$ EQ2.1.78

We shall write out a few of these explicitly. For $l = 0$, we have explicitly the null matrices

$$l_3 = 0, \qquad \boldsymbol{I}^2 = 0, \qquad l_1 = 0 = l_2 \quad (2.1.79)$$

Next for $l = \frac{1}{2}$, $m = \pm\frac{1}{2}$, the matrix elements are

$$l_3 = \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \qquad \boldsymbol{I}^2 = \frac{3\hbar^2}{4}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad (2.1.80)$$

$$\boldsymbol{l_1 = \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}} \qquad \boldsymbol{l_2 = \frac{\hbar}{2}\begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}} \tag{2.1.81}$$

while the state vectors become

$$\boldsymbol{\left| +\frac{1}{2}; \frac{1}{2} \right\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix}} \qquad \boldsymbol{\left| -\frac{1}{2}; \frac{1}{2} \right\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix}} \tag{2.1.82}$$

For $l = 1$, $m = -1, 0, +1$, and we have

$$\boldsymbol{l_3 = \hbar \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix}} \qquad \boldsymbol{I^2 = 2\hbar^2 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}} \tag{2.1.83}$$

$$\boldsymbol{l_1 = \frac{\hbar}{\sqrt{2}} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}} \quad \boldsymbol{l_2 = \frac{\hbar}{\sqrt{2}} \begin{bmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{bmatrix}} \tag{2.1.84}$$

while

$$\boldsymbol{|+1;1\rangle = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}} \qquad \boldsymbol{|0;1\rangle = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}} \qquad \boldsymbol{|-1;1\rangle = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}} \tag{2.1.85}$$

It was left to the reader to proceed for computing the previous matrices for $l = \frac{3}{2}$, $m = \pm\frac{3}{2}, \pm\frac{1}{2}$ so the matrices are 4×4.

We have chosen I^2 and l_3 as the two commuting operators to diagonalize.We could have also chosen I^2 and l_1 or I^2 and l_2. We say that the 3 or z-axis is the axis of quantization when we diagonalize I^2 and l_3. There is obviously nothing unique about the z-axis here. If we, for example, applied a uniform magnetic field in a certain direction, then it would usually be advantageous tochoose this direction as the axis of quantization.

We have shown that l_z may have integral or half-

integral multiples of has its eigenvalues. Simply this result arises because of the commutation relations (2.1.11-2.1.13) and (2.1.15) and has nothing to do with the definitions (2.1.8-2.1.10) of I in terms of the coordinates and momentum. However, if I is to representorbital angular momentum, then the eigenvectors of I^2 and l_3 must have coordinate or momentum representatives. That is, we must be able to expressthe I matrices in terms of coordinate and momentum matrices. We show inthe next section that this is only possible if we restrict the eigenvalues of l_z to be an integer times $\hbar$. The half-integer values do not have a classical analog in that $|m;l\rangle$ does not have a coordi-

nate representative when m is a half-integer.

We have no reason for throwing out half-integral values if we say thesecorrespond to intrinsically quantum mechanical effects which have no classicalanalog. We call such intrinsic angular momentum spin angular momentum.It turns out indeed that some particles have not only orbital angularmomentum but in addition are born with spin angular momentum. Electrons areborn with a spin $l = \frac{1}{2}$ so $m = \pm\frac{1}{2}$. Effects due to this can be measuredexperimentally and all attempts to explain these effects classically have failed.