APPLIED(9) COURSE

Quantum Mechanics

For the Student of

Faculty of Education Level 4

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

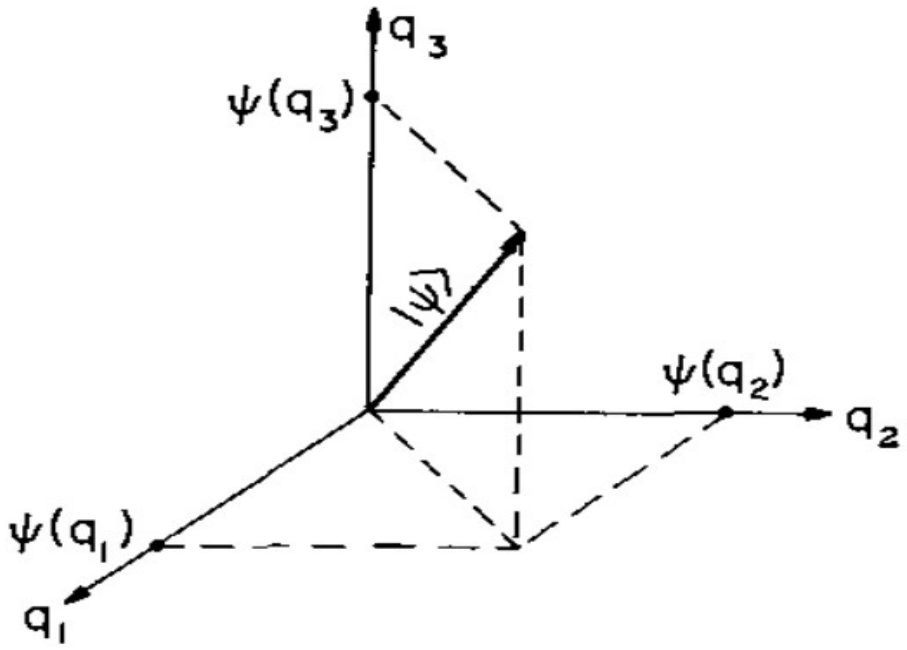

Figure 1.1: Ket vector and three of its coordinate representatives.

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

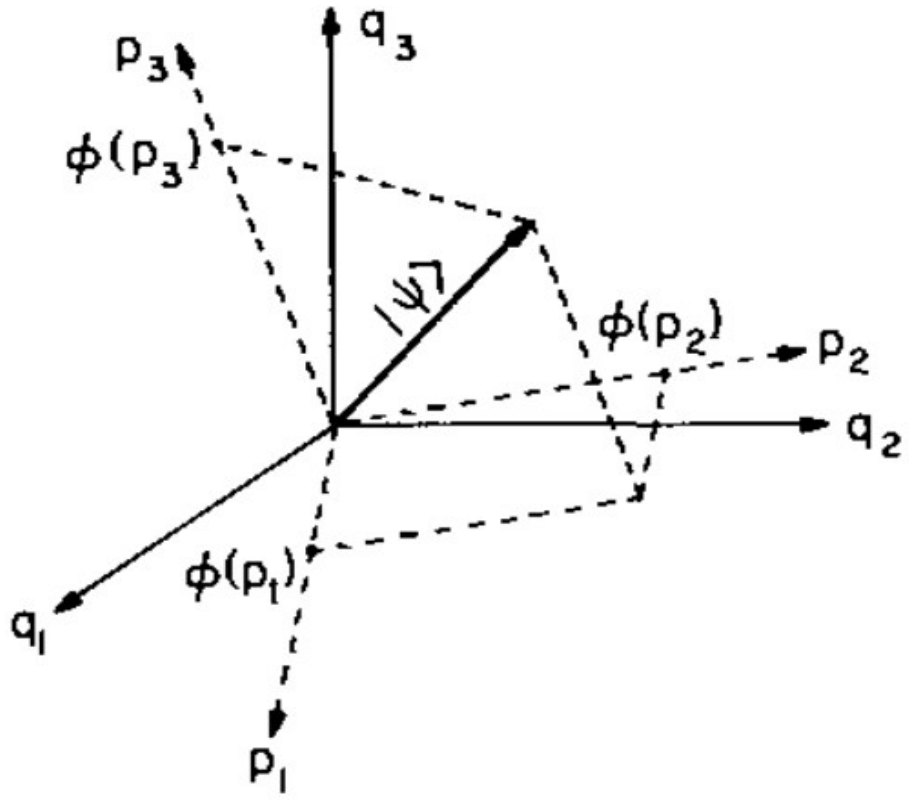

Figure 1.2: Ket vector and three of its coordinate representatives.

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

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

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,

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

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

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

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

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

$$\langle a|b\rangle = \langle b|a\rangle^* \tag{1.2.17}$$

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,

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

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

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

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

$$D(c|a_\rangle) = cD|a\rangle \tag{1.3.6}$$

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 $|a\rangle$:

$$(D_1 + D_2)|a\rangle = D_1|a\rangle + D_2|a\rangle \tag{1.3.7}$$

a product

$$(\boldsymbol{D_1 D_2})|\boldsymbol{a}\rangle = \boldsymbol{D_1}(\boldsymbol{D_2})|\boldsymbol{a}\rangle \tag{1.3.8}$$

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

We also have, for example,

$$(\boldsymbol{D_1 + D_2})|\boldsymbol{a}\rangle = (\boldsymbol{D_2 + D_1})|\boldsymbol{a}\rangle \tag{1.3.9}$$

$$[(\boldsymbol{D_1 + D_2}) + \boldsymbol{D_3}]|\boldsymbol{a}\rangle = [(\boldsymbol{D_1} + (\boldsymbol{D_2 + D_3})|\boldsymbol{a}\rangle \tag{1.3.10}$$

$$[(\boldsymbol{D_1}(\boldsymbol{D_2 + D_3})|\boldsymbol{a}\rangle = \boldsymbol{D_1 D_2}|\boldsymbol{a}\rangle + \boldsymbol{D_1 D_3}|\boldsymbol{a}\rangle \tag{1.3.11}$$