The Mathematical Language of Quantum Mechanics

Dirac Language

We now introduce the standard way of writing(from Dirac) vectors in Quantum Physics.

The **ket** $|A\rangle$ will be the symbol representing the vector $\vec{A}$ (the old notation) where the **ket-label** A contains **all the information** we know about the ket vector.

For example, we might have labels like (if we were talking about ordinary 2-dimensional vectors) $|Q\rangle = |r_Q, \theta_Q\rangle = |Q_x, Q_y\rangle$.

Vector addition is written as $|A\rangle + |B\rangle = |C\rangle$. This addition property is very important.

These ket vectors will represent **real physical systems** in the universe, that is, their mathematical properties will be so extensive that real physical systems can be completely represented by them!

The addition of ket vectors will be connected with the idea of **superposition** of physical systems that will be central to our understanding of quantum physics.

This property of two vectors that they can be added together to form a third vector in the same space will allow us, later on, to construct physical systems that are **sums** of other physical systems with strange properties.

In particular, it will allow us to construct **magenta electrons as superpositions of hard and soft electrons**!

So this very simple and basic property of vectors will be connected to powerful properties of real physical systems.

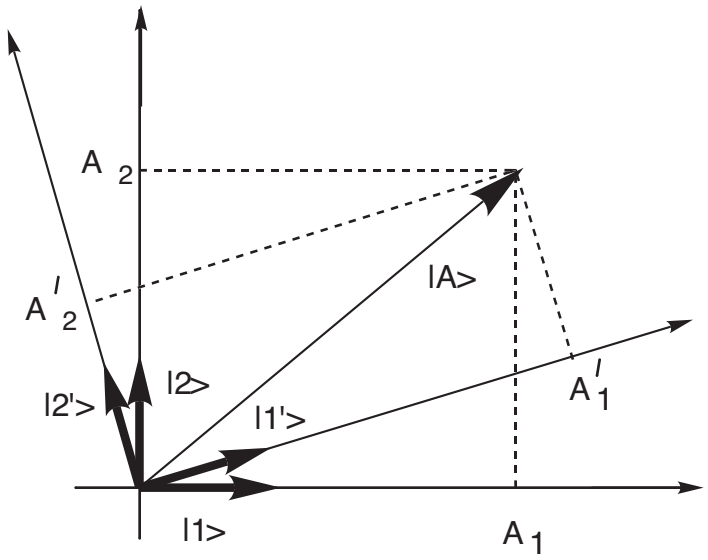
For the moment, however, the Dirac language **simply** looks like a change of notation.

Let us continue.

The equations that we have developed so far are rewritten in the following way in Dirac language:

$$\begin{aligned}
&\hat{e}_i \rightarrow |i\rangle \\
&\hat{e}_i \cdot \hat{e}_j = (|i\rangle, |j\rangle) = \langle i|j\rangle = \text{"}bracket\text{"} \\
&\vec{A} = \sum_{j=1}^{3} A_j \hat{e}_j \rightarrow |A\rangle = \sum_{j=1}^{3} A_j |j\rangle \\
&\hat{e}_k \cdot \vec{A} = A_k \rightarrow \langle k|A\rangle = A_k \\
&\vec{A} = \sum_{j=1}^{3} (\hat{e}_j \cdot \vec{A}) \hat{e}_j \rightarrow |A\rangle = \sum_{j=1}^{3} \langle j|A\rangle |j\rangle = \sum_{j=1}^{3} |j\rangle\langle j|A\rangle
\end{aligned} \qquad (29)$$

The actual choice of any basis set is arbitrary as shown below



where we have

$$|A\rangle = A_1|1\rangle + A_2|2\rangle = A'_1|1'\rangle + A'_2|2'\rangle \qquad (30)$$

Clearly, the vector is the same for different bases - only the components change; any **orthonormal** basis set is equivalent in quantum mechanics, i.e., any basis set can be used to represent any vector (or physical system).

As we shall see, in quantum physics, the A_i numbers (the vector components) will represent **real physically measurable** quantities.

To generalize to an N-dimensional space(this is no longer any real space that you can move around in), we must use N mutually orthonormal vectors to describe it, say the set $|1\rangle,|2\rangle,|3\rangle,.....,|N-1\rangle,|N\rangle$ where the orthonormality is expressed by the relations $\langle i|j\rangle = \delta_{ij}$.

Property: for scalar products we have $\langle A|(|B\rangle + |C\rangle) = \langle A|B\rangle + \langle A|C\rangle$

We now use this property in an ordinary 2-dimensional space example, i.e., suppose we have the pair of vectors $|x\rangle, |y\rangle$ where $\langle x|x\rangle = 1 = \langle y|y\rangle$, $\langle x|y\rangle = 1 = \langle y|x\rangle$ and where the (x,y) labels indicate the direction of the vector in real 2-dimensional space - a real physically measurable thing. Since the vectors $|x\rangle, |y\rangle$ form a orthonormal pair of vectors, we can use the pair of vectors $|x\rangle, |y\rangle$ as basis vectors for the 2-dimensional plane (all vectors in the plane can be written in terms of them).

Now suppose we have two vectors

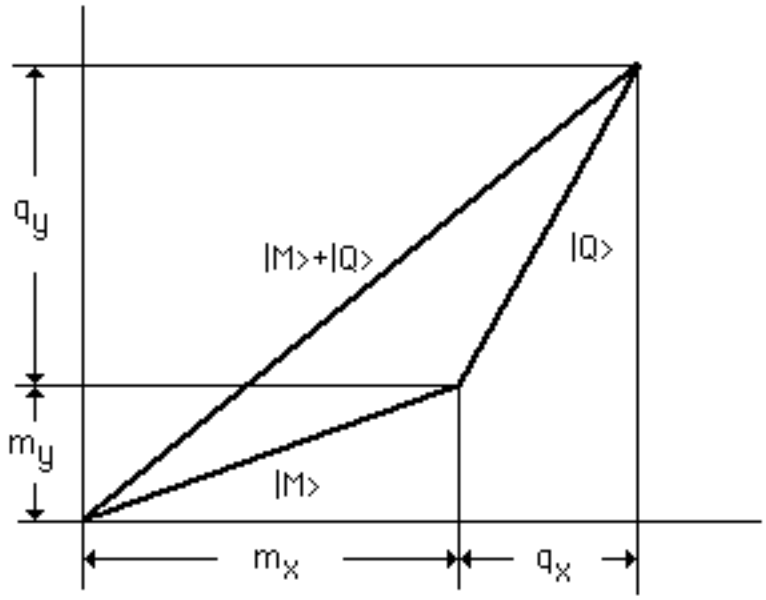
$$|M\rangle = m_x|x\rangle + m_y|y\rangle$$

$$m_x = \sqrt{\langle M|M\rangle}\cos\theta_M = \langle x|M\rangle = L_M\cos\theta_M\,,\; m_y = \sqrt{\langle M|M\rangle}\sin\theta_M = \langle y|M\rangle = L_M\sin\theta_M$$

$$|Q\rangle = q_x|x\rangle + q_y|y\rangle$$

$$q_x = \sqrt{\langle Q|Q\rangle}\cos\theta_Q = \langle x|Q\rangle = L_Q\cos\theta_Q\,,\; q_y = \sqrt{\langle Q|Q\rangle}\sin\theta_Q = \langle y|Q\rangle = L_Q\sin\theta_Q$$

as shown below:



Now

$$|M\rangle + |Q\rangle = (m_x + q_x)|x\rangle + (m_y + q_y)|y\rangle$$

where we have just added components.

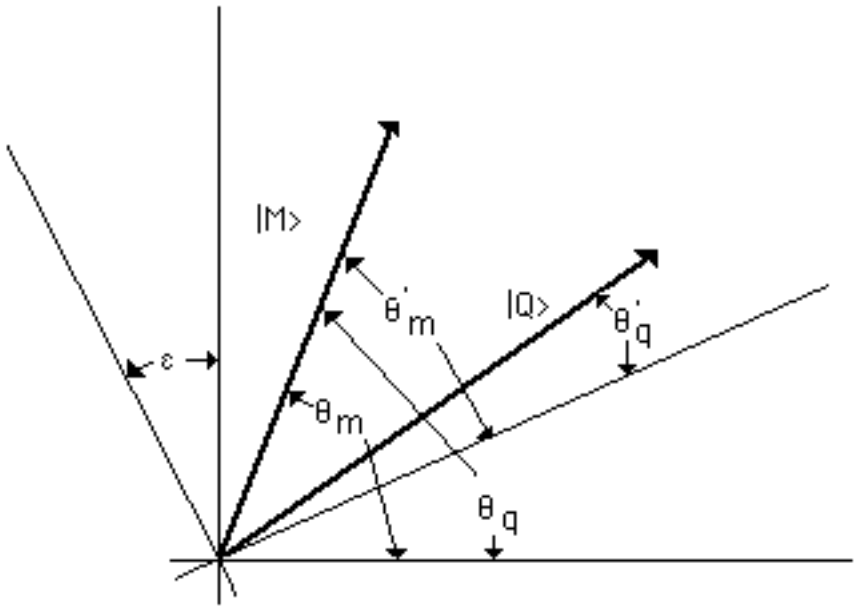
The scalar product of these two vectors is given by

$$\langle M|Q\rangle = m_x q_x + m_y q_y \qquad (31)$$

= **invariant** (**independent** of basis choice)

Proof:

If we choose a rotated axis pair (rotate by angle ε as shown below),

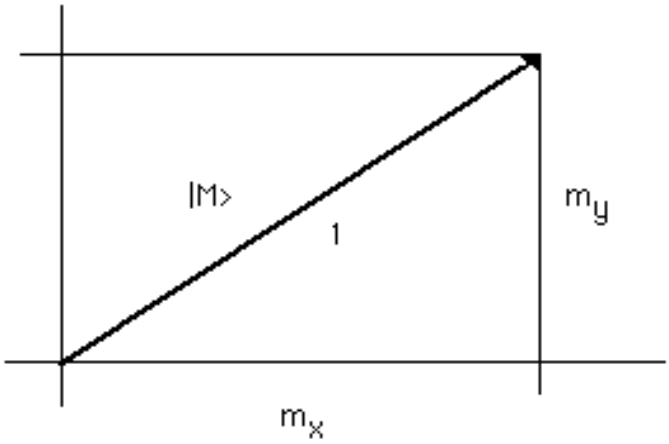


then we get

$$\langle M'|Q'\rangle = L_M L_Q \cos\left(\theta'_M - \theta'_Q\right) = L_M L_Q \cos\left((\theta_M - \varepsilon) - (\theta_Q - \varepsilon)\right) = L_M L_Q \cos\left(\theta_M - \theta_Q\right) = \langle M|Q\rangle$$

so that the scalar product is **independent of choice of basis vectors.**

Now suppose we want $\langle M|M\rangle = 1$ (vector normalized to 1). What does that mean? If we let $Q \to M$ in Eq.(31), then we must have $\langle M|M\rangle = m_x^2 + m_y^2 = 1$ or the sum of the squares of the components = 1. This is just the Pythagorean theorem(see diagram)



If $\langle M|M\rangle = m_x^2 + m_y^2 \neq 1$, then we can **normalize** the vector by dividing as shown below:

$$|M\rangle = \frac{1}{\sqrt{m_x^2 + m_y^2}}(m_x|x\rangle + m_y|y\rangle) \qquad (32)$$

So we only need to know the two numbers (the two orthogonal components) for each vector and everything else can be determined.

The components seem to contain all the possible information about a vector.

Now, generalizing to N dimensions (where we would need N mutually orthonormal basis vectors). Let us designate the basis vectors by

$$|1\rangle, |2\rangle, |3\rangle, ..., |N-1\rangle, |N\rangle$$

Suppose we have the two vectors

$$|M\rangle = m_1|1\rangle + m_2|2\rangle + ...m_N|N\rangle$$
$$|Q\rangle = q_1|1\rangle + q_2|2\rangle + ...q_N|N\rangle$$

then(as before),

$$|M\rangle + |Q\rangle = (m_1 + q_1)|1\rangle + (m_2 + q_2)|2\rangle + ... + (m_N + q_N)|N\rangle$$

and

$$\langle M|Q\rangle = m_1 q_1 + m_2 q_2 + ... + m_N q_N$$

= **invariant** (**independent** of basis choice)

Our discussion up to this point has involved what is called a **real vector space**, i.e., the vector components are real numbers. Quantum mechanics involves a complex vector space, so we will need a generalization to create the correct vectors and vector properties appropriate for quantum theory.

In particular, the coefficients (components) of the basis vectors will need to be complex numbers, i.e., in the expression

$$|M\rangle = m_x|x\rangle + m_y|y\rangle$$

m_x and m_y will need to be complex numbers.

To allow for the possibility that quantum vectors might have complex components, we **generalize** the definition of the scalar product to

$$\langle M|Q\rangle = m^*_x\, q_x + m^*_y\, q_y \tag{33}$$

and

$$\begin{aligned}\langle M|M\rangle &= m^*_x\, m_x + m^*_y\, m_y \\ &= |m_x|^2 + |m_y|^2\end{aligned} \tag{34}$$

which means that the length is still a real number!

An alternative way of looking at the inner products goes as follows: given two **ket** vectors $|A\rangle, |B\rangle$ we define new mathematical objects given by the symbols $\langle A|, \langle B|$ called **bra** vectors. If the ket vectors $|A\rangle, |B\rangle$ belong to a vector space, then the bra vectors $\langle A|, \langle B|$ also belong to a vector space called the **dual** space and

$$\langle A| \quad \text{is the dual vector for the vector} \quad |A\rangle$$

If we have the ket vector $|A\rangle = A_1|1\rangle + A_2|2\rangle$ then its dual or bra vector is

$$\langle A| = A^*_1\langle 1| + A^*_2\langle 2| \tag{35}$$

and the bra(-c-)ket or inner product is defined as

$$\langle A|B\rangle = \big(\langle A|\big)|B\rangle = \big(|A\rangle, |B\rangle\big) \tag{36}$$

The bra vector is called a **linear functional.** When acting on a ket vector it produces a number (the inner product as in Eq.(36)). Using this idea, Eq.(34) above can be written, explicitly, in this way

$$\begin{aligned}\langle M|M\rangle &= \Big(m^*_x\,\langle x| + m^*_y\,\langle y|\Big)\Big(m_x|x\rangle + m_y|y\rangle\Big) \\ &= m^*_x\, m_x\langle x|x\rangle + m^*_x\, m_y\langle x|y\rangle + m^*_y\, m_x\langle y|x\rangle + m^*_y\, m_y\langle y|y\rangle \quad \text{(property \#1)} \\ &= m^*_x\, m_x + m^*_y\, m_y \qquad \text{(using orthonormality)} \\ &= |m_x|^2 + |m_y|^2\end{aligned}$$

We also have that

$$\langle V|W\rangle = \langle W|V\rangle^* \tag{37}$$

Illustration using Color and Hardness Properties

The above formalism is easily able to describe the world of color and hardness.

Suppose that we choose the ket vector $|g\rangle$ to represent the state of an electron with color = green, the ket vector $|m\rangle$ to represent the state of an electron with color = magenta, the ket vector $|h\rangle$ to represent the state of an electron with hardness = hard, and the ket vector $|s\rangle$ to represent the state of an electron with hardness = soft.

We will assume that

$$\begin{aligned} \langle g|g\rangle &= 1 = \langle m|m\rangle \\ \langle g|m\rangle &= 0 = \langle m|g\rangle \end{aligned} \tag{38}$$

i.e., they are an orthonormal basis set and

$$\begin{aligned} \langle h|h\rangle &= 1 = \langle s|s\rangle \\ \langle h|s\rangle &= 0 = \langle s|h\rangle \end{aligned} \tag{39}$$

so that they are an orthonormal basis set also.

Since we observed that hard electrons can appear to be both green and magenta electrons we might make **a first guess** (this is what theoretical physicists do) that

$$|hard\rangle = |h\rangle = a|g\rangle + b|m\rangle \tag{40}$$

= SUPERPOSITION of both color states

= LINEAR COMBINATION of both color states

where

$$a = \langle g|h\rangle \quad \text{and} \quad b = \langle m|h\rangle \tag{41}$$

Can this simple correspondence really work? Is it possible that SUPERPOSITION is just a linear combination or just vector addition!

We shall see later that it does work!!

We will find many different mathematical ways to represent these fundamental vector objects.

If we know the basis we are using (we shall see later that it will be chosen according to what we are **trying to measure**) then we only need the two numbers a and b, i.e., the components .

Two different ways to represent the superposition are

$$|hard\rangle = a|g\rangle + b|m\rangle = \begin{pmatrix} a \\ b \end{pmatrix} \tag{42}$$

where $\begin{pmatrix} a \\ b \end{pmatrix}$ is a column vector or 1 x 2 matrix (more about this later).

That is almost all we will need to know about vectors in order to do quantum theory.

So lots of new words, but really just high school ideas of vectors, components, length(Pythagorus), perpendicular, etc along with an open mind willing to accept new uses for these old ideas and new ways of expressing them.

As we shall see, **ket vectors will represent physical systems in quantum mechanics.**

IV. Operators in a Vector Space

There is one other kind of mathematical object we will need to know about to do quantum theory. They are called **operators.** This one may not be familiar to you from H.S. but the concept should be, as we will now see.

In a vector space, an **operator** is a **definite rule for taking every vector in the space into some other vector in the space.**

An operator will be represented by the ^ symbol, i.e., $\hat{Q}$, so that the action of the operator $\hat{Q}$ is given by the relation

$$|B'\rangle = \hat{Q}|B\rangle \qquad (43)$$

Think about this moving from **right to left ---**

vector $|B\rangle$ acted upon by operator $\hat{Q}$ is changed into vector $|B'\rangle$

Think of an operator as some kind of box, where you put a vector in and get another(**either different or the same**) vector out. This is similar to the definition of a function for the space of numbers.

Some properties of operators in QM:

$$\hat{Q}(|A\rangle + |B\rangle) = \hat{Q}|A\rangle + \hat{Q}|B\rangle \quad \text{(linearity)}$$
$$\hat{Q}(c|A\rangle) = c\hat{Q}|A\rangle \quad , \quad c = \textit{complex number}$$
$$\langle C|\left(\hat{Q}|B\rangle\right) = \langle C|B'\rangle = \textit{number} \equiv \langle C|\hat{Q}|B\rangle \quad \text{(matrix element)} \qquad (44)$$
$$(\hat{Q}_1 + \hat{Q}_2)|A\rangle = \hat{Q}_1|A\rangle + \hat{Q}_2|A\rangle \quad \text{(linearity)}$$
$$(\hat{Q}_1\hat{Q}_2)|A\rangle = \hat{Q}_1(\hat{Q}_2|A\rangle) \quad \text{(order matters)}$$

These properties imply that all of our operators are what the mathematicians call **LINEAR** operators. The fact that quantum mechanics can be understood using only linear operators is truly amazing because they are the simplest kind of operator that mathematicians can think of.

All observables or "quantities that we can measure" will be represented by operators in quantum mechanics.

Now we introduce a **completely new property** using vectors and operators.

This new property of operators involves mathematical objects called **eigenvalues and eigenvectors.**

In some special cases it turns out that we find the following result:

$$\hat{B}|B\rangle = b|B\rangle \qquad (45)$$

where b = number, that is, we get the **exact same vector** back when using the operator (possibly multiplied by a number $\neq 1$).

In this case, $|B\rangle$ is an **eigenvector** of $\hat{B}$ with **eigenvalue** b.

These ideas will be very important in quantum mechanics because the **only possible results that can be obtained from the measurement of a physical quantity or observable** that can be represented by the operator $\hat{W}$ will be the **eigenvalues** of $\hat{W}$ (**a postulate we will make later**)!

The set of eigenvalues is called the **spectrum** of the operator.

It will also turn out that the **eigenvectors** of the operator representing an **observable** can **always** be used as the **basis** for the vector space.

Average Values in Quantum Mechanics
or Another Way to Represent Operators

Suppose that we have N **identically prepared** physical systems, each represented by the **same** state vector $|\psi\rangle$.

Suppose that we make measurements of an observable represented by the operator $\hat{B}$ with eigenvalues/eigenvectors given by

$$\hat{B}|b_j\rangle = b_j|b_j\rangle \quad \text{j} = 1,2,3,4,.....$$

Suppose the measurement results(remember they must be the eigenvalues as we shall see later) are

$$b_k \quad , \quad n_k \text{ times} \qquad \text{k} = 1,2,3,4,.....$$

where

$$\sum_k n_k = N = \text{ total number of measurements}$$

Now from the definition we derived earlier, the average value of $\hat{B}$ is

$$\left\langle \hat{B} \right\rangle = \text{average or expectation value of } \hat{B}$$

$$= \frac{1}{N}\sum_k n_k b_k = \sum_k \frac{n_k}{N} b_k = \sum_k b_k \,\mathrm{Prob}(b_k) \tag{46}$$

In quantum theory, this will given by (**another postulate**) the expression

$$\left\langle \hat{B} \right\rangle = \sum_k b_k \,\mathrm{Prob}(b_k) = \sum_k b_k \left|\langle b_k|\psi\rangle\right|^2 \tag{47}$$

Since the set $\{|b_k\rangle\}$ is a basis we can write

$$|\psi\rangle = \sum_k \langle b_k|\psi\rangle |b_k\rangle \tag{48}$$

so that we will have the important result

$$\mathrm{Prob}(b_k) = \left|\langle b_k|\psi\rangle\right|^2 = \text{absolute value squared of the component of } |\psi\rangle \text{ along } |b_k\rangle \tag{49}$$

Now by definition

$$\left|\langle b_k|\psi\rangle\right|^2 = \langle b_k|\psi\rangle^* \langle b_k|\psi\rangle \tag{50}$$

and from Eq.(37) and the definition of the bra vector we have

$$\langle b_k|\psi\rangle^* = \langle \psi|b_k\rangle \tag{51}$$

Therefore

$$\left|\langle b_k|\psi\rangle\right|^2 = \langle \psi|b_k\rangle\langle b_k|\psi\rangle \tag{52}$$

and

$$\left\langle \hat{B} \right\rangle = \sum_k b_k \langle \psi|b_k\rangle\langle b_k|\psi\rangle = \langle \psi| \left(\sum_k b_k |b_k\rangle\langle b_k| \right) |\psi\rangle \tag{53}$$

Now the quantity $\sum_k b_k |b_k\rangle\langle b_k| = \hat{Q}$ must be some operator in the vector space since when it acts on any vector the result is another vector.

Proof:

Now we can always write

$$|a\rangle = \sum_s d_s |b_s\rangle = \text{arbitrary vector}$$

since the set $\{|b_s\rangle\}$ is a basis. This gives

$$\hat{Q}|a\rangle = \left(\sum_k b_k |b_k\rangle\langle b_k| \right) |a\rangle = \left(\sum_k b_k |b_k\rangle\langle b_k| \right) \left(\sum_s d_s |b_s\rangle \right)$$

$$= \sum_{k,s} d_s b_k |b_k\rangle\langle b_k|b_s\rangle = \sum_{k,s} d_s b_k |b_k\rangle \delta_{ks} = \sum_k b_k d_k |b_k\rangle$$

so that the operation by $\hat{Q}$ on an arbitrary vector returns a vector indicating that $\hat{Q}$ is an operator.

But we also have

$$\hat{B}|a\rangle = \sum_k d_k \hat{B}|b_k\rangle = \sum_k b_k d_k |b_k\rangle$$

This is the same result as for the operator $\hat{Q}$, which says (since the vector $|a\rangle$ is arbitrary) that we can always write

$$\hat{B} = \sum_k b_k |b_k\rangle\langle b_k| \qquad (54)$$

This is another(very important) way of representing the operator $\hat{B}$, i.e., **any operator can be written in terms of its eigenvalues and eigenvectors.** It is called the **spectral decomposition.**

Therefore, the average value in Eq.(53) becomes

$$\langle \hat{B} \rangle = \langle \psi | \left(\sum_k b_k |b_k\rangle\langle b_k| \right) |\psi\rangle = \langle \psi | \hat{B} | \psi \rangle \qquad (55)$$

Therefore, when we do an experiment of this type, i.e., identical measurements on many identical systems, the (expected) average of those measurements is just this **average** or **expectation** value given by $\langle \psi | \hat{B} | \psi \rangle$.

Projection Operators

Operators of the form

$$\hat{P} = |b_k\rangle\langle b_k| \qquad (56)$$

where

$$\hat{B}|b_j\rangle = b_j |b_j\rangle \quad \text{j} = 1,2,3,4,.....$$

are called **projection operators.** They have some very interesting and useful properties.

First, what are the eigenvalues of a projection operator? We have

$$\hat{P}^2 = (|b_k\rangle\langle b_k|)(|b_k\rangle\langle b_k|) = |b_k\rangle\langle b_k|b_k\rangle\langle b_k| = |b_k\rangle\langle b_k| = \hat{P}$$

or if

$$\hat{P}|\alpha\rangle = \lambda|\alpha\rangle$$

so that $|\alpha\rangle = \text{eigenvector}$ and $\lambda = \text{corresponding eigenvalue}$, then we have

$$\hat{P}^2|\alpha\rangle = \lambda\hat{P}|\alpha\rangle = \lambda^2|\alpha\rangle = \hat{P}|\alpha\rangle = \lambda|\alpha\rangle$$

$$(\lambda^2 - \lambda)|\alpha\rangle = 0$$

$$\lambda^2 - \lambda = 0 \rightarrow \lambda = 0 \text{ or } 1 \rightarrow \text{eigenvalues}$$

Second, consider the following operator $\hat{I} = \sum_k |b_k\rangle\langle b_k|$. We have

$$\hat{I}|\beta\rangle = \sum_k |b_k\rangle\langle b_k|\beta\rangle = \sum_k \langle b_k|\beta\rangle |b_k\rangle = |\beta\rangle$$

or

$$\hat{I} = \sum_k |b_k\rangle\langle b_k| \qquad (57)$$

is the i**dentity** operator. The identity operator is defined as that operator which does not change anything!

Our earlier results now makes sense since

$$\hat{B} = \hat{B}I = \hat{B}\sum_k |b_k\rangle\langle b_k| = \sum_k \hat{B}|b_k\rangle\langle b_k| = \sum_k b_k |b_k\rangle\langle b_k|$$

$$|\psi\rangle = \sum_k \langle b_k|\psi\rangle |b_k\rangle = \sum_k |b_k\rangle\langle b_k|\psi\rangle = \left(\sum_k |b_k\rangle\langle b_k|\right)|\psi\rangle = \hat{I}|\psi\rangle$$

Operator Examples:

Suppose we define an operator $\hat{G} = |g\rangle\langle g| = |green\rangle\langle green|$. We then have these properties

$$\hat{G}|g\rangle = |g\rangle\langle g|g\rangle = |g\rangle$$

$$\hat{G}|m\rangle = |g\rangle\langle g|m\rangle = 0$$

or the states representing the green and magenta electrons are the eigenvectors with eigenvalues 1 and 0, respectively.

We find that

$$\langle g|\hat{G}|g\rangle = |\langle g|g\rangle|^2 = 1 = \text{expectation value of } \hat{G} \text{ in the green state}$$

and we also have

$$\langle m|\hat{G}|m\rangle = |\langle g|m\rangle|^2 = 0 = \text{expectation value of } \hat{G} \text{ in the magenta state}$$

These results make sense if we **interpret** $\hat{G} = |g\rangle\langle g|$ as the operator corresponding to a measurement of the green property of electrons, i.e., an observer looking at the output of a green aperture of a color box.

The first result then says if we measure the probability that the color of a green electron is green we get the value 1 as expected and the second result then says if we measure the probability that the color of a green electron is magenta we get the value 0 as expected.

Pushing this strange idea even further, if we assume that

$$|hard\rangle = |h\rangle = \frac{1}{\sqrt{2}}|g\rangle + \frac{1}{\sqrt{2}}|m\rangle \text{ = SUPERPOSITION of green and magenta}$$

then we have

$$\langle h|\hat{G}|h\rangle = \text{expectation value of } \hat{G} \text{ in the hard state}$$

$$= \left(\frac{1}{\sqrt{2}}\langle g| + \frac{1}{\sqrt{2}}\langle h|\right)\hat{G}\left(\frac{1}{\sqrt{2}}|g\rangle + \frac{1}{\sqrt{2}}|m\rangle\right)$$

$$= \frac{1}{2}\langle g|\hat{G}|g\rangle + \frac{1}{2}\langle g|\hat{G}|m\rangle + \frac{1}{2}\langle m|\hat{G}|g\rangle + \frac{1}{2}\langle m|\hat{G}|m\rangle$$

$$= \frac{1}{2}\langle g|\hat{G}|g\rangle = \frac{1}{2}$$

Another way of saying this is (using Eq.(47))

$$\langle h|\hat{G}|h\rangle = \text{expectation value of } \hat{G} \text{ in the hard state}$$

$$= \sum(\text{eigenvalue g})(\text{probability of g in } |h\rangle)$$

$$= (1)\left|\langle eigenvalue = 1|h\rangle\right|^2 + (0)\left|\langle eigenvalue = 0|h\rangle\right|^2$$

$$= (1)\left|\langle g|h\rangle\right|^2 + (0)\left|\langle m|h\rangle\right|^2$$

$$= (1)\frac{1}{2} + (0)\frac{1}{2} = \frac{1}{2}$$

which again makes sense i.e., if we have a beam of hard electrons, then we will measure an electron to be green 1/2 of the time as we observed earlier!

Clearly, this formalism is both neat and very powerful and certainly seems to have the potential to describe our earlier observations. We will see shortly that the formalism can completely represent quantum systems and quantum measurements.

More Useful Mathematical Ideas

Matrices

We have been representing an operator by the way it acts on vectors.

An operator can also be represented by an array of numbers. In mathematics such an array is called a **matrix.**,i.e., a matrix is a mathematical object which takes the form

$$[O] = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} = \text{2x2 matrix representing operator } \hat{O} \qquad (58)$$

where the numbers

$$\alpha, \beta, \gamma, \delta \text{ are called matrix elements of the operator } \hat{O}$$

and their numerical values depend on which set of basis vectors we are using and the corresponding operator $\hat{O}$.

Let us assume the two basis vectors are labelled by $|1\rangle, |2\rangle$.

We then define

$$\alpha = O_{11} = \langle 1|\hat{O}|1\rangle = O_{\textit{row label,column label}}$$

$$\beta = O_{12} = \langle 1|\hat{O}|2\rangle$$

$$\gamma = O_{21} = \langle 2|\hat{O}|1\rangle$$

$$\delta = O_{22} = \langle 2|\hat{O}|2\rangle$$

or

$$[O] = \begin{pmatrix} O_{11} & O_{12} \\ O_{21} & O_{22} \end{pmatrix} = \begin{pmatrix} \langle 1|\hat{O}|1\rangle & \langle 1|\hat{O}|2\rangle \\ \langle 2|\hat{O}|1\rangle & \langle 2|\hat{O}|2\rangle \end{pmatrix} \qquad (59)$$

This is the matrix representation of the operator $\hat{O}$ in the $|1\rangle, |2\rangle$ basis.

A vector in the space is then represented by a 2-element column matrix (vector) of the form

$$|a\rangle = a_1|1\rangle + a_2|2\rangle = \langle 1|a\rangle|1\rangle + \langle 2|a\rangle|2\rangle \equiv \begin{pmatrix} \langle 1|a\rangle \\ \langle 2|a\rangle \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \qquad (60)$$

This implies that the basis vectors themselves are represented (in the $|1\rangle, |2\rangle$ basis) by the particular column matrices

$$|1\rangle = \begin{pmatrix} \langle 1|1\rangle \\ \langle 2|1\rangle \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \text{ and } |2\rangle = \begin{pmatrix} \langle 1|2\rangle \\ \langle 2|2\rangle \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \qquad (61)$$

Now consider the color basis. We have

$$|1\rangle = |g\rangle = \begin{pmatrix} \langle g|g\rangle \\ \langle m|g\rangle \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \text{ and } |2\rangle = |m\rangle = \begin{pmatrix} \langle g|m\rangle \\ \langle m|m\rangle \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

and

$$|hard\rangle = |h\rangle = \frac{1}{\sqrt{2}}|g\rangle + \frac{1}{\sqrt{2}}|m\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix} \qquad (62)$$

We also have

$$|soft\rangle = |s\rangle = \frac{1}{\sqrt{2}}|g\rangle - \frac{1}{\sqrt{2}}|m\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

Note that the hard and soft states are also orthonormal. Both sets are bases as can be seen explicitly by the example below:

$$|\psi\rangle = \begin{pmatrix} a \\ b \end{pmatrix} = a\begin{pmatrix} 1 \\ 0 \end{pmatrix} + b\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \frac{1}{\sqrt{2}}(a+b)\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix} + \frac{1}{\sqrt{2}}(a-b)\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

The operator $\hat{\mathrm{G}}$, was defined to be

$$\mathrm{G} = |g\rangle\langle g| \qquad (63)$$

and it represented measuring the "**green**" property of electrons.

The matrix representing $\hat{\mathrm{G}}$ (in the color basis) can now be written down:

$$(\hat{G})_{11} = \langle g|(|g\rangle\langle g|)|g\rangle = 1$$

$$(\hat{G})_{12} = \langle g|(|g\rangle\langle g|)|m\rangle = 0$$

$$(\hat{G})_{21} = \langle m|(|g\rangle\langle g|)|g\rangle = 0$$

$$(\hat{G})_{22} = \langle m|(|g\rangle\langle g|)|m\rangle = 0$$

or

$$\left[\hat{G}\right] = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \tag{64}$$

We then have

$$\hat{G}|g\rangle = (|g\rangle\langle g|)|g\rangle = |g\rangle$$

or

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \rightarrow \text{matrix multiplication}$$

where matrix multiplication is defined by these relations: if

$$[A] = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \quad , \quad [B] = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$$

then

$$[A][B] = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}\begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{pmatrix}$$

or

$$[AB]_{ij} = \sum_k a_{ik}b_{kj} \tag{65}$$

So we have **two equivalent ways** of doing calculations, i.e., using **operator algebra** or **matrix algebra.**

Another Example in a Finite Dimensional Vector Space

Let us consider the 2-dimensional vector space spanned by the orthonormal basis set

$$|1\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad , |2\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

We can define two projection operators as

$$\hat{P}_1 = |1\rangle\langle 1| \quad , \quad \hat{P}_2 = |2\rangle\langle 2|$$

The matrix representation of these two projection operator is easily found using $\langle 1|1\rangle = \langle 2|2\rangle = 1$ and $\langle 1|2\rangle = \langle 2|1\rangle = 0$ and $\hat{Q}_{ki} = \langle k|\hat{Q}|i\rangle$. We have

$$\left(\hat{P}_1\right)=\begin{pmatrix}\langle 1|\hat{P}_1|1\rangle & \langle 1|\hat{P}_1|2\rangle\\ \langle 2|\hat{P}_1|1\rangle & \langle 2|\hat{P}_1|2\rangle\end{pmatrix}=\begin{pmatrix}\langle 1|1\rangle\langle 1|1\rangle & \langle 1|1\rangle\langle 1|2\rangle\\ \langle 2|1\rangle\langle 1|1\rangle & \langle 2|1\rangle\langle 1|2\rangle\end{pmatrix}=\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}$$

$$\left(\hat{P}_2\right)=\begin{pmatrix}\langle 1|\hat{P}_2|1\rangle & \langle 1|\hat{P}_2|2\rangle\\ \langle 2|\hat{P}_2|1\rangle & \langle 2|\hat{P}_2|2\rangle\end{pmatrix}=\begin{pmatrix}\langle 1|2\rangle\langle 2|1\rangle & \langle 1|2\rangle\langle 2|2\rangle\\ \langle 2|2\rangle\langle 2|1\rangle & \langle 2|2\rangle\langle 2|2\rangle\end{pmatrix}=\begin{pmatrix}0 & 0\\ 0 & 1\end{pmatrix}$$

Now consider an arbitrary vector in this space

$$|a\rangle=\begin{pmatrix}a_1\\ a_2\end{pmatrix}=a_1|1\rangle\ +a_2|2\rangle$$

We then have (using both Dirac and matrix language)

$$\hat{P}_1|a\rangle=|1\rangle\langle 1|a\rangle=a_1|1\rangle\langle 1|1\rangle\ +a_2|1\rangle\langle 1|2\rangle=a_1|1\rangle$$

or

$$\left(\hat{P}_1\right)\begin{pmatrix}a_1\\ a_2\end{pmatrix}\ =\ \begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}\begin{pmatrix}a_1\\ a_2\end{pmatrix}=\begin{pmatrix}a_1\\ 0\end{pmatrix}=a_1\begin{pmatrix}1\\ 0\end{pmatrix}$$

and the projection operator performs as advertised.

We note that (at least in this **special case**)

$$\left(\hat{P}_1\right)+\left(\hat{P}_2\right)=\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}+\begin{pmatrix}0 & 0\\ 0 & 1\end{pmatrix}=\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}=\hat{I}=\textit{identity operator}$$

or

$$(\hat{P}_1+\hat{P}_2)|a\rangle=(|1\rangle\langle 1|+|2\rangle\langle 2|)|a\rangle=\sum_{j=1}^{2}|j\rangle\langle j|a\rangle=|a\rangle=\hat{I}|a\rangle$$

where we have made use of the expansion formula for an arbitrary state in an orthonormal basis.

Commutator

We define the **commutator** between two operators as

$$\left[\hat{A},\hat{B}\right]=\hat{A}\hat{B}-\hat{B}\hat{A} \qquad (66)$$

For ordinary numbers the commutator is zero, but for operators it is not always zero.

If two operators have a **zero** commutator, then we say that they **commute.**

If we have

$$\hat{A}=\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\ ,\quad \hat{B}=\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}$$

then

$$\left[\hat{A},\hat{B}\right]=\hat{A}\hat{B}-\hat{B}\hat{A}=\begin{pmatrix}1&0\\0&-1\end{pmatrix}\begin{pmatrix}0&1\\1&0\end{pmatrix}-\begin{pmatrix}0&1\\1&0\end{pmatrix}\begin{pmatrix}1&0\\0&-1\end{pmatrix}$$

$$=\begin{pmatrix}0&-1\\-1&0\end{pmatrix}-\begin{pmatrix}0&1\\1&0\end{pmatrix}=-2\begin{pmatrix}1&0\\0&1\end{pmatrix}\neq 0\rightarrow \text{don't commute}$$

Let us now devise a **general procedure** for finding the eigenvalues and eigenvectors of an operator. We will use the matrices that represent operators. In particular, we consider the matrices (operators)

$$\hat{A}=\begin{pmatrix}1&0\\0&-1\end{pmatrix}\ ,\quad \hat{B}=\begin{pmatrix}0&1\\1&0\end{pmatrix}\ ,\quad \hat{C}=\begin{pmatrix}1&2\\2&1\end{pmatrix}$$

We consider the operator $\hat{A}$ first. The eigenvalue/eigenvector problem can be written in the following way:

$$\hat{A}|\alpha\rangle=\alpha|\alpha\rangle$$

or

$$\begin{pmatrix}1&0\\0&-1\end{pmatrix}\begin{pmatrix}a\\b\end{pmatrix}=\alpha\begin{pmatrix}a\\b\end{pmatrix}$$

The last matrix equation corresponds to two simple algebraic equations

$$a=\alpha a$$

$$b=-\alpha b$$

The solutions to these equations are:

$$a\neq 0\ \rightarrow\alpha=1$$

$$\rightarrow b=-b\rightarrow b=0$$

and

$$b\neq 0\rightarrow\alpha=-1$$

$$\rightarrow a=-a\ \rightarrow a=0$$

This says that the eigenvalues of $\hat{A}$ are $\alpha=\pm 1$ and the corresponding eigenvectors (normalized) to 1 are

$$|+1\rangle=\begin{pmatrix}1\\0\end{pmatrix}\ ,\quad |-1\rangle=\begin{pmatrix}0\\1\end{pmatrix}$$

So, for diagonal matrices (where the only non-zero matrix elements lie along the diagonal) the eigenvalue/eigenvector problem is trivial. The eigenvalues are just the diagonal matrix elements and the eigenvectors are just column vectors with a single entry = 1 and all other entries = 0. This generalizes to any number of dimensions.

Let us use this simple case to illustrate the general procedure for non-diagonal matrices. We had

$$\begin{pmatrix}1&0\\0&-1\end{pmatrix}\begin{pmatrix}a\\b\end{pmatrix}=\alpha\begin{pmatrix}a\\b\end{pmatrix}$$

This can be written as

$$\begin{pmatrix} 1-\alpha & 0 \\ 0 & -1-\alpha \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = 0$$

This set of equations is called homogeneous (zeroes on the right-hand side). A trivial solution always exists - it is $a = b = 0$. This solution is of no interest to physicists. We will have a non-trivial solution if and only if

$$\text{determinant}\begin{pmatrix} 1-\alpha & 0 \\ 0 & -1-\alpha \end{pmatrix} = \det\begin{pmatrix} 1-\alpha & 0 \\ 0 & -1-\alpha \end{pmatrix} = \begin{vmatrix} 1-\alpha & 0 \\ 0 & -1-\alpha \end{vmatrix} = 0$$

where

$$\det\begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc$$

In the above case we get

$$\begin{vmatrix} 1-\alpha & 0 \\ 0 & -1-\alpha \end{vmatrix} = 0 = (1-\alpha)(-1-\alpha) \rightarrow \alpha = \pm 1 = \text{eigenvalues}$$

Proof: Suppose that we have a pair of linear equations in two unknowns

$$a\alpha_1 + b\alpha_2 = q_1$$
$$c\alpha_1 + d\alpha_2 = q_2$$

We can write these equations in matrix notation as

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} = \begin{pmatrix} q_1 \\ q_2 \end{pmatrix}$$

or

$$A\alpha = Q$$

where

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad , \quad \alpha = \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} \quad , \quad Q = \begin{pmatrix} q_1 \\ q_2 \end{pmatrix}$$

The inverse of a matrix A^{-1} (or an operator) is defined by

$$A^{-1}A = I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = identity \qquad (67)$$

If the inverse exists, then we can write

$$A^{-1}A\alpha = I\alpha = \alpha = A^{-1}Q$$

which represents a solution for the unknowns α_1 and α_2.

For a 2x2 matrix

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

the inverse is given by

$$A^{-1} = \frac{1}{\det A}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \qquad \text{where} \qquad \det A = ad - bc$$

since

$$A^{-1}A = \frac{1}{\det A}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{68}$$

The solution of the original equation is then given by

$$\begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} = \frac{1}{\det A}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}\begin{pmatrix} q_1 \\ q_2 \end{pmatrix} \tag{69}$$

Example:

Suppose we have

$$x + 2y = 5$$
$$2x - 3y = 4$$

or

$$\begin{pmatrix} 1 & 2 \\ 2 & -3 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 5 \\ 4 \end{pmatrix}$$

We have

$$A = \begin{pmatrix} 1 & 2 \\ 2 & -3 \end{pmatrix} \quad , \quad \det A = -7 \quad , \quad A^{-1} = -\frac{1}{7}\begin{pmatrix} -3 & -2 \\ -2 & 1 \end{pmatrix}$$

Check:

$$A^{-1}A = -\frac{1}{7}\begin{pmatrix} -3 & -2 \\ -2 & 1 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 2 & -3 \end{pmatrix} = -\frac{1}{7}\begin{pmatrix} -7 & 0 \\ 0 & -7 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

as it should.

We therefore have the solution

$$\begin{pmatrix} x \\ y \end{pmatrix} = -\frac{1}{7}\begin{pmatrix} -3 & -2 \\ -2 & 1 \end{pmatrix}\begin{pmatrix} 5 \\ 4 \end{pmatrix} = -\frac{1}{7}\begin{pmatrix} -7 \\ -14 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

or

$$x = 1, y = 2$$

Now let us return to the case of the eigenvalue/eigenvector equation where we have

$$\hat{A}|\alpha\rangle = \alpha|\alpha\rangle = \alpha\hat{I}|\alpha\rangle$$

or

$$(\hat{A} - \alpha\hat{I})|\alpha\rangle = 0$$

If we define

$$\hat{R} = \hat{A} - \alpha\hat{I}$$

then we have the equation

$$\hat{R}|\alpha\rangle = 0$$

or

$$\left[\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} - \begin{pmatrix} \alpha & 0 \\ 0 & \alpha \end{pmatrix}\right]\begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} A_{11}-\alpha & A_{12} \\ A_{21} & A_{22}-\alpha \end{pmatrix}\begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Thus, we have a solution

$$\begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} = \frac{1}{\det\begin{pmatrix} A_{11}-\alpha & A_{12} \\ A_{21} & A_{22}-\alpha \end{pmatrix}}\begin{pmatrix} A_{11}-\alpha & A_{12} \\ A_{21} & A_{22}-\alpha \end{pmatrix}^{-1}\begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

The only possibility that this gives a nonzero(or nontrivial) result is when

$$\det\begin{pmatrix} A_{11}-\alpha & A_{12} \\ A_{21} & A_{22}-\alpha \end{pmatrix} = 0$$

which completes the proof.

A more general procedure(matrix not diagonal) for finding eigenvalues goes like the following:

$$\hat{G} = \begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{pmatrix}$$

$$\hat{G}|\lambda\rangle = \begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{pmatrix}\begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = \lambda\begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = \lambda|\lambda\rangle$$

$$\begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{pmatrix}\begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} - \lambda\begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = \begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{pmatrix}\begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} - \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}\begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = \begin{pmatrix} g_{11}-\lambda & g_{12} \\ g_{21} & g_{22}-\lambda \end{pmatrix}\begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = 0$$

Once again, this last set of equation is pair of homogeneous linear equations. Again, there is a always the trivial solution $(\lambda_1 = \lambda_2 = 0)$. A non-trivial solution exists only if

$$\text{determinant}\begin{pmatrix} g_{11}-\lambda & g_{12} \\ g_{21} & g_{22}-\lambda \end{pmatrix} = (g_{11}-\lambda)(g_{22}-\lambda) - g_{12}g_{21} = 0 \qquad (70)$$

which is a **quadratic** equation for the eigenvalues.

Let us apply this procedure to the other two operators. First we work with $\hat{B}$.

$$\hat{B} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \rightarrow \begin{vmatrix} -\alpha & 1 \\ 1 & -\alpha \end{vmatrix} = 0 = \alpha^2 - 1 \rightarrow \alpha = \pm 1 = \text{eigenvalues}$$

To find the eigenvectors we then proceed as follows:

$$\alpha = +1$$

$$\rightarrow \hat{B}|+1\rangle_B = |+1\rangle_B \rightarrow \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix}$$

$$\rightarrow b = a$$

To normalize the eigenvector we must have

$$ {}_B\langle +1|+1\rangle_B = 1 = a^2 + b^2 = 2a^2 \rightarrow a = \frac{1}{\sqrt{2}} = b $$

so that

$$ |+1\rangle_B = \begin{pmatrix} a \\ b \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} $$

and similarly

$$ |-1\rangle_B = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} $$

Second, we work with $\hat{C}$.

$$ \hat{C} = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \rightarrow \begin{vmatrix} 1-\alpha & 2 \\ 2 & 1-\alpha \end{vmatrix} = 0 = (1-\alpha)(1-\alpha) - 4 \rightarrow \alpha^2 - 2\alpha - 3 = 0 \rightarrow \alpha = 3, -1 $$ = eigenvalues

To find the eigenvectors we then proceed as follows:

$$ \alpha = +3 $$

$$ \rightarrow \hat{C}|+3\rangle_C = 3|+3\rangle_C \rightarrow \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = 3 \begin{pmatrix} a \\ b \end{pmatrix} $$

$$ \rightarrow a + 2b = 3a $$

$$ \rightarrow 2a + b = 3b $$

$$ \rightarrow b = a \quad , \quad a^2 + b^2 = 1 $$

$$ \rightarrow a = \frac{1}{\sqrt{2}} = b $$

so that

$$ |+3\rangle_C = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} $$

and

$$ |-1\rangle_C = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} $$

A strange thing has happened. Both operators $\hat{B}$ and $\hat{C}$ have the same eigenvectors (different eigenvalues).

It turns out that the following important property is true:

if two operators commute, then the share a common set of eigenvectors

This the case for $\hat{B}$ and $\hat{C}$.

Special Matrices (operators)

Two types of operators will dominate our discussion of quantum mechanics.

Hermitian : $\hat{A} = \hat{A}^{+} = \hat{A}^{*Tr}$ or $A_{ij} = A_{ji}^{*}$

Unitary : $\hat{A}^{-1} = \hat{A}^{+} = \hat{A}^{*Tr}$ or $(A^{-1})_{ij} = A_{ji}^{*}$

More Mathematical Details

Definition : A set of vectors is said to be **linearly independent** if a linear relation of the form

$$\sum_{k=1}^{n} c_k |k\rangle = |\mathrm{O}\rangle$$

implies that $c_k = 0$ for **all** k**;** otherwise the set of vectors is **linearly dependent.**

If a set of vectors is linearly dependent, then we can express a member of the set as a linear combination of the other members of the set.

Examples:

(1) Consider the set of vectors(3-tuples in this case)

$$|1\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad , \quad |2\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad , \quad |3\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

This set is linearly independent since

$$a_1|1\rangle + a_2|2\rangle + a_3|3\rangle = \begin{pmatrix} a_1 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ a_2 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ a_3 \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = |\mathrm{O}\rangle = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

implies that the only solution is

$$a_1 = a_2 = a_3 = 0$$

(2) Consider the set of vectors(3-tuples in this case)

$$|1\rangle = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \quad , \quad |2\rangle = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \quad , \quad |3\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

This set is linearly independent since

$$a_1|1\rangle + a_2|2\rangle + a_3|3\rangle = \begin{pmatrix} a_1 \\ a_1 \\ 0 \end{pmatrix} + \begin{pmatrix} a_2 \\ -a_2 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ a_3 \end{pmatrix} = \begin{pmatrix} a_1 + a_2 \\ a_1 - a_2 \\ a_3 \end{pmatrix} = |\mathrm{O}\rangle = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

implies that the solution is

$$a_1 + a_2 = a_1 - a_2 = a_3 = 0 \quad \text{or} \quad a_1 = a_2 = a_3 = 0$$

(3) Consider the set of vectors(3-tuples in this case)

$$|1\rangle = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \ , \quad |2\rangle = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \ , \quad |3\rangle = \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}$$

This set is linearly dependent since

$$a_1|1\rangle + a_2|2\rangle + a_3|3\rangle = \begin{pmatrix} a_1 \\ a_1 \\ a_1 \end{pmatrix} + \begin{pmatrix} a_2 \\ -a_2 \\ 0 \end{pmatrix} + \begin{pmatrix} 2a_3 \\ 0 \\ a_3 \end{pmatrix} = \begin{pmatrix} a_1 + a_2 + 2a_3 \\ a_1 - a_2 \\ a_1 + a_3 \end{pmatrix} = |\mathrm{O}\rangle = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

implies that the solution is

$$a_1 + a_2 + 2a_3 = a_1 - a_2 = a_1 + a_3 = 0 \quad \text{or} \quad a_3 = -a_1 \ , \ a_2 = a_1$$

and we have (for example)

$$|1\rangle = -|2\rangle + |3\rangle$$

We say that an infinite set of vectors is linearly independent if **every finite subset** is linearly independent.

Definition : The maximum number of linearly independent vectors in a space V is called the **dimension** of the space $\dim(V)$.

Definition : A set of vectors $\{|k\rangle \ , \ k = 1,2,3,.....,n\}$ **spans** the space if **every** vector $|Q\rangle$ in the space can be written as a linear combination of vectors in the set

$$\sum_{k=1}^{n} q_k |k\rangle = |Q\rangle$$

This linear combination, which is given by the coefficients $q_k \ , k = 1,2,.....,n$, is **unique.**

Definition : A set of vectors is a **basis** for the space if it is a linearly independent set **and** spans the space, that is, if $\dim(V) = m$, a set of m linearly independent vectors is called a **basis** on V.

The set of vectors

$$|1\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad , \quad |2\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad , \quad |3\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

is the maximal set of linearly independent vectors since any other vector $|g\rangle$ in the space can always be written as a linear combination of them as

$$|g\rangle = a_1|1\rangle + a_2|2\rangle + a_3|3\rangle = \begin{pmatrix} a_1 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ a_2 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ a_3 \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$$

Therefore the dimension of this vector space is 3. This set of vectors is also a basis. The basis is **not unique** since the set of linearly independent vectors

$$|1\rangle = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \quad , \quad |2\rangle = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \quad , \quad |3\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

also spans the space, i.e.,

$$|g\rangle = c_1|1\rangle + c_2|2\rangle + c_3|3\rangle = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} c_1 + c_2 + 2c_3 \\ c_1 - c_2 \\ c_3 \end{pmatrix}$$

implies that

$$c_1 = a_1 + a_2 - 2a_3 \quad , \quad c_2 = a_1 - 2a_3 \quad , \quad c_3 = a_3$$

and, thus, this set is also a basis. Clearly, a basis spans the whole of V.

Example: In the space of 3-tuples, a basis is represented by the three vectors

$$|1\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad , \quad |2\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad , \quad |3\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

so that an arbitrary vector in the space

$$|g\rangle = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$$

can be written

$$|g\rangle = a_1|1\rangle + a_2|2\rangle + a_3|3\rangle = \begin{pmatrix} a_1 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ a_2 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ a_3 \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$$

so that $a_1, a_2, and\ a_3$ are the components.

Gram-Schmidt Orthogonalization Process

An orthonormal basis set for an n-dimensional vector space can always be constructed from any set of n linearly independent vectors using the Gram-Schmidt orthogonalization method.

Suppose that we have a set of n linearly independent vectors $|\alpha_i\rangle$, $i = 1,2,.....,n$ that are not a mutually orthonormal set. We can construct a mutually orthonormal set $|\beta_i\rangle$, $i = 1,2,.....,n$ using the following steps:

(1) let $|\beta_1\rangle = |\alpha_1\rangle$

(2) let $|\beta_2\rangle = |\alpha_2\rangle + a_1|\beta_1\rangle$ where we choose a_1 such that $\langle\beta_1|\beta_2\rangle = 0$

(3) this gives

$$\langle\beta_1|\beta_2\rangle = 0 = \langle\beta_1|\alpha_2\rangle + a_1\langle\beta_1|\beta_1\rangle$$

$$a_1 = -\frac{\langle\beta_1|\alpha_2\rangle}{\langle\beta_1|\beta_1\rangle}$$

Now proceed by **induction.**

Suppose we have constructed k mutually orthogonal vectors $|\beta_i\rangle$, $i = 1,2,.....,k$. If we let

$$|\beta_{k+1}\rangle = |\alpha_k\rangle + \sum_{j=1}^{k} a_j|\beta_j\rangle$$

with

$$a_j = -\frac{\langle\beta_j|\alpha_{k+1}\rangle}{\langle\beta_j|\beta_j\rangle}$$

then we have $\langle\beta_j|\beta_{k+1}\rangle = 0$ for $j = 1,2,....,k$. These steps are repeated until we have n mutually orthogonal vectors. We then normalize them to 1 and create an orthonormal set.

For example, suppose we have the set

$$|\alpha_1\rangle = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \quad , \quad |\alpha_2\rangle = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \quad , \quad |\alpha_3\rangle = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$$

These vectors are not orthonormal.

(1) let

$$|\beta_1\rangle = |\alpha_1\rangle = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \quad , \quad \langle\beta_1|\beta_1\rangle = 2$$

(2) let

$$|\beta_2\rangle = |\alpha_2\rangle + a_1|\beta_1\rangle$$

with

$$a_1 = -\frac{\langle\beta_1|\alpha_2\rangle}{\langle\beta_1|\beta_1\rangle} = -\frac{1}{2}$$

and thus

$$|\beta_2\rangle = |\alpha_2\rangle - \frac{1}{2}|\alpha_1\rangle = \frac{1}{2}\begin{pmatrix}-1\\1\\2\end{pmatrix} \quad , \quad \langle\beta_2|\beta_2\rangle = \frac{3}{2} \quad , \quad \langle\beta_1|\beta_2\rangle = 0$$

(3) let

$$|\beta_3\rangle = |\alpha_3\rangle + a_1|\beta_1\rangle + a_2|\beta_2\rangle$$

with

$$a_1 = -\frac{\langle\beta_1|\alpha_3\rangle}{\langle\beta_1|\beta_1\rangle} = -\frac{1}{2} \quad \text{and} \quad a_2 = -\frac{\langle\beta_2|\alpha_3\rangle}{\langle\beta_2|\beta_2\rangle} = -\frac{1}{3}$$

and thus

$$|\beta_3\rangle = \frac{2}{3}\begin{pmatrix}1\\-1\\1\end{pmatrix} \quad , \quad \langle\beta_3|\beta_3\rangle = \frac{4}{3} \quad , \quad \langle\beta_1|\beta_3\rangle = 0 \quad , \quad \langle\beta_2|\beta_3\rangle = 0$$

We normalize the vectors by dividing by their respective **norms,**

$$|\gamma_i\rangle = \frac{|\beta_i\rangle}{\|\beta_i\rangle\|} = \frac{|\beta_i\rangle}{|\langle\beta_i|\beta_i\rangle|^{1/2}}$$

The orthonormal set is then

$$|\gamma_1\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\\0\end{pmatrix} \; , \; |\gamma_2\rangle = \frac{1}{\sqrt{6}}\begin{pmatrix}-1\\1\\2\end{pmatrix} \; , \; |\gamma_3\rangle = \frac{1}{\sqrt{3}}\begin{pmatrix}1\\-1\\1\end{pmatrix}$$

Examples - Functions of Operators

Suppose that we have the eigenvector/eigenvalue equations for a self-adjoint operator

$$\hat{A}|k\rangle = a_k|k\rangle \quad , \quad k = 1,2,....,N$$

We then assume that

$$f(\hat{A})|k\rangle = f(a_k)|k\rangle \quad , \quad k = 1,2,....,N$$

for the eigenvectors.

We can show that this works for polynomials and power series as follows:

$$|\psi\rangle = \sum_{k=1}^{N} |k\rangle\langle k|\psi\rangle$$

$$\hat{A}|\psi\rangle = \hat{A}\sum_{k=1}^{N} |k\rangle\langle k|\psi\rangle = \sum_{k=1}^{N} \hat{A}|k\rangle\langle k|\psi\rangle = \sum_{k=1}^{N} a_k|k\rangle\langle k|\psi\rangle = \left(\sum_{k=1}^{N} a_k|k\rangle\langle k|\right)|\psi\rangle$$

$$\rightarrow \hat{A} = \sum_{k=1}^{N} a_k|k\rangle\langle k| \rightarrow \text{spectral resolution of the operator}$$

Now define the **projection** operator

$$\hat{P}_k = |k\rangle\langle k| \rightarrow \hat{P}_k\hat{P}_j = \hat{P}_k\delta_{kj}$$

We then have

$$\hat{A} = \sum_{k=1}^{N} a_k|k\rangle\langle k| = \sum_{k=1}^{N} a_k\hat{P}_k$$

or any operator is represented by a sum over its eigenvalues and corresponding projection operators.

We then have

$$\hat{A}^2 = \left(\sum_{k=1}^{N} a_k\hat{P}_k\right)\left(\sum_{j=1}^{N} a_j\hat{P}_j\right) = \sum_{k,j=1}^{N} a_k a_j\hat{P}_k\hat{P}_j = \sum_{k,j=1}^{N} a_k a_j\hat{P}_k\delta_{kj} = \sum_{k=1}^{N} a_k^2\hat{P}_k$$

$$\rightarrow \hat{A}^n = \sum_{k=1}^{N} a_k^n\hat{P}_k$$

Therefore, for

$$f(x) = \sum_{n=1}^{N} q_n x^n$$

we have

$$f(\hat{A}) = \sum_{n=1}^{N} q_n\hat{A}^n = \sum_{n=1}^{N} q_n\sum_{k=1}^{N} a_k^n\hat{P}_k = \sum_{k=1}^{N}\left(\sum_{n=1}^{N} q_n a_k^n\right)\hat{P}_k = \sum_{k=1}^{N} f(a_k)\hat{P}_k$$

This says that, in general, we have

$$f(\hat{A})|\psi\rangle = f(\hat{A})\sum_{k=1}^{N} |k\rangle\langle k|\psi\rangle = \sum_{k=1}^{N} f(\hat{A})|k\rangle\langle k|\psi\rangle = \sum_{k=1}^{N} f(a_k)|k\rangle\langle k|\psi\rangle = \left(\sum_{k=1}^{N} f(a_k)|k\rangle\langle k|\right)|\psi\rangle$$

$$\rightarrow f(\hat{A}) = \sum_{k=1}^{N} f(a_k)|k\rangle\langle k| \rightarrow \text{spectral resolution of a function of an operator}$$

Numerical example:

$$\hat{A} = \begin{pmatrix} 4 & 3 \\ 3 & 4 \end{pmatrix}$$

has eigenvalues 7,1 with eigenvectors

$$|7\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix}, |1\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\end{pmatrix}$$

This gives

$$\hat{P}_7 = |7\rangle\langle 7| = \frac{1}{2}\begin{pmatrix}1 & 1\\1 & 1\end{pmatrix}, \hat{P}_1 = |1\rangle\langle 1| = \frac{1}{2}\begin{pmatrix}1 & -1\\-1 & 1\end{pmatrix}$$

and therefore

$$\hat{A} = 7\hat{P}_7 + \hat{P}_1 = \frac{7}{2}\begin{pmatrix}1 & 1\\1 & 1\end{pmatrix} + \frac{1}{2}\begin{pmatrix}1 & -1\\-1 & 1\end{pmatrix} = \begin{pmatrix}4 & 3\\3 & 4\end{pmatrix}$$

$$\hat{A}^2 = 7^2\hat{P}_7 + \hat{P}_1 = \frac{7^2}{2}\begin{pmatrix}1 & 1\\1 & 1\end{pmatrix} + \frac{1}{2}\begin{pmatrix}1 & -1\\-1 & 1\end{pmatrix} = \begin{pmatrix}25 & 24\\24 & 25\end{pmatrix} = \begin{pmatrix}4 & 3\\3 & 4\end{pmatrix}\begin{pmatrix}4 & 3\\3 & 4\end{pmatrix}$$

$$\log\hat{A} = \log(7)\hat{P}_7 + \log(1)\hat{P}_1 = \frac{\log(7)}{2}\begin{pmatrix}1 & 1\\1 & 1\end{pmatrix}$$

$$\sqrt{\hat{A}} = \sqrt{7}\hat{P}_7 + \hat{P}_1 = \frac{\sqrt{7}}{2}\begin{pmatrix}1 & 1\\1 & 1\end{pmatrix} + \frac{1}{2}\begin{pmatrix}1 & -1\\-1 & 1\end{pmatrix} = \frac{1}{2}\begin{pmatrix}\sqrt{7}+1 & \sqrt{7}-1\\\sqrt{7}-1 & \sqrt{7}+1\end{pmatrix}$$

Clearly we then have

$$\log\hat{A}|7\rangle = \frac{1}{\sqrt{2}}\frac{\log(7)}{2}\begin{pmatrix}1 & 1\\1 & 1\end{pmatrix}\begin{pmatrix}1\\1\end{pmatrix} = \log(7)\frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix} = \log(7)|7\rangle$$

$$\log\hat{A}|1\rangle = \frac{1}{\sqrt{2}}\frac{\log(7)}{2}\begin{pmatrix}1 & 1\\1 & 1\end{pmatrix}\begin{pmatrix}1\\-1\end{pmatrix} = \frac{1}{\sqrt{2}}\frac{\log(7)}{2}\begin{pmatrix}0\\0\end{pmatrix} = 0 = \log(1)|1\rangle$$

as expected.

The Dirac δ-Function

History

In the development of quantum mechanics by P. Dirac , the following sequence of ideas occurred (as we shall describe later).

(1) Observable = measurable quantity <--> Hermitian operator
(2) Physical states are linear combinations of eigenvectors
[requires complete orthonormal basis]
(3) Possible measurements are represent by the eigenvalues
[must be real numbers]

$\hat{O}|\lambda\rangle = \lambda|\lambda\rangle$ where $|\lambda\rangle$ = eigenvector and λ = eigenvalue

(4) Some observables have a discrete spectrum (finite number or denumerably infinite number)

$\langle\lambda'|\lambda\rangle = \delta_{\lambda'\lambda}$ for different eigenvalues this is a well-defined statement

(5) Other observables have a continuous spectrum (non-denumerably infinite number of eigenvalues).

For example , the position operator $\hat{X}$, which we will talk about later, is such that

$$\hat{X}|x\rangle = x|x\rangle \quad \text{where} \quad |x\rangle = \text{eigenvector} \quad \text{and} \quad x = \text{eigenvalue}$$

We now ask the question, what is $\langle x'|x\rangle$? Dirac assumed that

$$\langle x'|x\rangle \equiv \delta_{x'x} \equiv \delta(x - x')$$

where

$$\int_{-\infty}^{\infty} \delta(x - x')dx' = 1$$

$$\delta(x - x') = 0 \quad \text{if} \quad x' \neq x$$

$$\int_{-\infty}^{\infty} f(x')\delta(x - x')dx' = f(x)$$

This gives the correct physical theory in the sense that all predictions agree with experiment (1929).

Eventually (1960) the mathematicians, who initially vehemently disputed Dirac's assumption of this new "function", caught up to the physicists and proved all of its properties in the Theory of Distributions.

Here is how Dirac(in his own words) introduced this function.

Our work in Section 10 led us to consider quantities involving a certain kind of infinity. To get a precise notation for dealing with these infinities, we introduce a quantity $\delta(x)$ depending on a parameter x satisfying the conditions

$$\int_{-\infty}^{\infty} \delta(x)dx = 1$$

$$\delta(x) = 0 \quad \text{for} \quad x \neq 0$$

To get a picture of $\delta(x)$, take a function of the real variable x which vanishes everywhere except inside a small domain, of length ε say, surrounding the origin x = 0, and which is so large inside this domain that its integral over the domain is unity. The exact shape of the function inside this domain does not matter, provided there are no unnecessary wild variations (for example provided the function is always of order ε^{-1}). Then in the limit $\varepsilon \to 0$ this function will go over into $\delta(x)$.

$\delta(x)$ is not a function of x according to the usual mathematical definition of a function, which requires a function to have a definite value for each point in its domain, but is something more general, which we may call an "improper function" to show up the difference from a function defined by the usual definition. Thus $\delta(x)$ is not a quantity which can be generally used in mathematical analysis like an ordinary function, but its use must be confined to certain simple types of expressions for which it is obvious that no inconsistency can arise.

Complicated Arguments

(1) Consider

$$\int_{0-}^{0+} dt'\,\delta(-t')f(t') = -\int_{-0-}^{-0+} dt'\,\delta(t')f(-t') = -\int_{0+}^{0-} dt'\,\delta(t')f(-t') = \int_{0-}^{0+} dt'\,\delta(t')f(-t') = f(0)$$

which implies that $\delta(-t)=\delta(t)$ in the sense of using integrals, i.e.,

$$f(0) = \int \delta(-t)\,f(t)dt = \int \delta(t)\,f(t)dt$$

(2) Consider

$$\int_{0-}^{0+} dt'\,\delta(at')f(t') = \frac{1}{|a|}\int_{0-/a}^{0+/a} dt''\,\delta(t'')f(\frac{t''}{a}) = \frac{1}{|a|}f(0)$$

which implies that $\delta(at)=\frac{1}{|a|}\delta(t)$ in the sense of using integrals, i.e.,

$$\frac{1}{|a|}f(0) = \int \delta(at)\,f(t)dt = \frac{1}{|a|}\int \delta(t)\,f(t)dt$$

(3) Consider

$$\int_{0-}^{0+} dt'\,t'\,\delta(t')f(t') = f(0)(0) = 0$$

If $f(0)\neq 0$, then this implies that

$$t\delta(t) = 0$$

(4) Consider

$$\int_{-L}^{L} f(x)\delta(x^2-b^2)dx = \int_{-L}^{0} f(x)\delta(x^2-b^2)dx + \int_{0}^{L} f(x)\delta(x^2-b^2)dx$$

Now near $x=-b$, we can write

$$x^2-b^2 = -2b(x+b)$$

and near $x=b$, we can write

$$x^2-b^2 = 2b(x-b)$$

therefore we have

$$\int_{-L}^{L} f(x)\delta(x^2-b^2)dx = \int_{-L}^{0} f(x)\delta(-2b(x+b))dx + \int_{0}^{L} f(x)\delta(2b(x-b))dx$$

$$= -\frac{1}{2|b|}f(-b) + \frac{1}{2|b|}f(b)$$

or

$$\delta(x^2 - b^2) = \frac{[\delta(x-b) + \delta(x+b)]}{2|b|}$$

(5) Now consider a function $g(x)$ which has a single zero at $x = x_0$. This implies that near x_0 we have

$$g(x) = (x - x_0)g'(x_0)$$

In this case we can write

$$\int_{-\infty}^{\infty} f(x)\delta(g(x))dx = \int_{-\infty}^{\infty} f(x)\frac{\delta(x-x_0)}{|g'(x_0)|}dx = \frac{f(x_0)}{|g'(x_0)|}$$

and generalizing, if $g(x)$ has N zeroes at $x = x_i$, $i = 1,2,3,.....,N$, then near each zero we can write

$$g(x) = (x - x_i)g'(x_i)$$

and we have

$$\delta(g(x)) = \sum_{i=1}^{N} \frac{\delta(x-x_i)}{|g'(x_i)|}$$

Examples :

$$\int_{-\infty}^{\infty} e^{-x}\delta(x^2 - a^2)dx = \int_{-\infty}^{\infty} e^{-x}\frac{\delta(x-a) + \delta(x+a)}{2a}dx = \frac{e^{-a} + e^{a}}{2a} = \frac{\cosh(a)}{a}$$

$$\int_{-\infty}^{\infty} e^{-x^2}\delta(\sin x)dx = \int_{-\infty}^{\infty} e^{-x^2}\sum_{n=-\infty}^{\infty}\frac{\delta(x-n\pi)}{|\cos n\pi|}dx = \sum_{n=-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-x^2}\delta(x-n\pi)dx = \sum_{n=-\infty}^{\infty} e^{-(n\pi)^2}$$

We will discuss delta functions later when we have to deal with the Schrodinger equation.

We will use this formalism to describe the theory/experiments in quantum mechanics. These ideas about vectors/operators will be sufficient for us to develop a quantum theory.