

§12. MATHEMATICS OF QUANTUM MECHANICS

In order for the science of physics to progress, physical theory needs to contain ideas based on experimental evidence and be described by mathematics. Theories are usually built in three stages, (i) the basic idea is formulated and interpreted and the physical foundation of the theory is built through experiments, (ii) a mathematical foundation is built which describes the experimental evidence in some form acceptable to peer review process. During this stage the physical meaning of the mathematical symbols come to represent the physical laws, (iii) a final stage in which the mathematical apparatus is *set to work* [Bohr63] to produce new results which can then be confirmed through experiments. This third stage leads to further understanding of the physical content of the theory and to further development the mathematical apparatus.

When Newton formulated the laws of mechanics, the corresponding mathematical tools also had to be formulated. These tools included differential and integral calculus. When quantum mechanics was formulated, the mathematical mechanisms were already in place in the form of the theory of linear operators.

This apparatus was summed up by Neils Bohr (1855–1962) as,

... in quantal formalism, the quantities by which the state of a physical system is ordinarily defined are replaced by symbolic operators subjected to a non-commutative algorithm involving Planck's constant. This procedure presents a fixation of such quantities to the extent which would be required for the deterministic description of classical physics, but allows us to determine their spectral distribution as revealed by evidence about atomic processes. In conformity with the non-pictorial character of the formalism, its physical interpretation finds expression in laws of an essentially statistical type... [Bohr63].

The mathematics of Quantum Mechanics and its relationship to the electromagnetic force — quantum electrodynamics is based on a few simple principals. Before these principals can be presented some background on the notation needs to be developed. It may seem unnecessary to define the simple mathematical term of a vector, but there is benefit, since the terminology has unique meaning in the field of quantum mechanics. ^[1]

... all science as it grow toward perfection becomes mathematical in its ideas.

— Alfred North Whitehead(1861–1947) (1911)

§12.1. VECTORS AND VECTOR SPACES

In quantum mechanics the theory abstract vector spaces provides the same tools as differential calculus does in classical mechanics. Although the concept of three dimensional vector spaces is familiar, the concept of an abstract vector space may not. There are two aspects to a vector space — geometric and algebraic. In the geometric aspect a vector can be described as a directed line segment which can represent a quantity with both magnitude and direction. A vector is distinguished from a scalar which is a quantity which has only magnitude. ^[2]

In algebraic aspect a vector can be described by a one-to-one correspondence between the unique set of vectors radiating from the origin of a coordinate system and the coordinates of the terminal points of each vector. This correspondence is represented by an ordered pair of real numbers (x_1, x_2, x_3) in 3 dimensions or simply x_i where the subscript $i=1, 2, 3$. In higher dimensions the algebraic notation for a vector is given by an n -tuple of real numbers (x_1, x_2, \dots, x_n) . Even though there is no longer a physical analog of the n -dimensional vector, the language of these vectors can be maintained through the algebra of abstract vector spaces.

The collection of vector associated with all points in the coordinate space is called a *vector space*. The *dimension* of a given vector space is the dimension of the associated coordinate space of points. This dimension is the number of coordinates needed to define a geometric point in the vector space. A one dimensional vector space describes a line of points, while a two dimensional vector space described a *plane* of points and a three dimensional vector space describes the world in which we live (in non-relativistic terms). Vector space is higher than three dimensions become difficult to visualize, but the formal description of such vector spaces is not a problem.

In the algebraic representation a vector can be expressed as a linear combination of non-coplanar vectors b_i or *basis* vectors. Using the basis of the vector space any vector can be *built up* from the basis elements of the vector space. It is the convention of abstract vector analysis though to

choose the basis vectors perpendicular to each other. When this occurs the basis is called *orthonormal* otherwise it is an oblique basis.

The basis are then equivalent to a *coordinate system* for an n -dimensional point in the vector space. The vectors which have a length of 1 (defined in Eq. (12.4)) and which point along the perpendicular coordinate axes of an n -dimensional space constitute an orthonormal basis of the associated n -dimensional vector space. For any vector space larger than a single dimension, there will be an infinite number of orthonormal basis. Any vector in this space can be written in terms of any of those basis. In a 3 dimensional vector space the vector \mathbf{x} can be described as,

$$\mathbf{x} = \alpha b_1 + \beta b_2 + \gamma b_3, \quad (12.1)$$

where $\alpha, \beta,$ and γ are scalars and called the expansion coefficients of the vector \mathbf{x} in the vector containing the basis vectors $b_1, b_2,$ and b_3 [Brau70].

§12.1.1. Abstract Vector Algebra

If vector analysis is to become the calculus of quantum mechanics, the *algebra* of vectors needs to be defined. Addition of vectors is defined as:

$$\mathbf{z} = \mathbf{x} + \mathbf{y}. \quad (12.2)$$

The sum of any two vectors is another vector. ^[3]

The multiplication of vectors is called the scalar or *inner* product. Given two vectors of 3 dimensions, $\mathbf{x} = (x_1, x_2, x_3)$ and $\mathbf{y} = (y_1, y_2, y_3)$ as:

$$\mathbf{x} \circ \mathbf{y} = \sum_{i=1}^3 x_i \cdot y_i. \quad (12.3)$$

The product of two vectors is a *number*. If the vector occupy a real vector space than the inner product results in a real number. If the vector occupy a complex vector space then the result is a complex number.

Using this notation, the length of a vector is denoted by $\|\mathbf{x}\|$ ^[5] and is given by,

$$\|\mathbf{x}\| = \sqrt{\mathbf{x} \circ \mathbf{x}} = \sqrt{\sum_{i=1}^3 x_i \cdot x_i}. \quad (12.4)$$

The angle between two vectors \mathbf{x} and \mathbf{y} is defined as,

$$\theta = \cos^{-1} \left[\frac{\mathbf{x} \circ \mathbf{y}}{\|\mathbf{x}\| \cdot \|\mathbf{y}\|} \right]. \quad (12.5)$$

A vector can also be multiplied by a number. This operation simply creates a new vector whose length is now n times the length of the original vector, such that,

$$3 \cdot \mathbf{x} = \mathbf{x} \circ \mathbf{x} \circ \mathbf{x}. \quad (12.6)$$

So long as the two vectors \mathbf{x} and \mathbf{y} occupy a *real* vector space of finite dimension then the inner product can be generalized to,

$$\mathbf{x} \circ \mathbf{y} = \sum_{i=1}^n x_i \cdot y_i. \quad (12.7)$$

If the vectors \mathbf{x} and \mathbf{y} are to be used to define the inner product in a *complex* vector space, there are difficulties with the definition of the length of a vector. If the distance function defined above in Eq. (12.4) is to be used, then the length of a complex vector will be,

$$\|\mathbf{x}\| = \sqrt{\mathbf{x} \circ \mathbf{x}} = \sqrt{-\mathbf{x}^2} \quad (12.8)$$

which is an imaginary number if \mathbf{x} is real. This situation causes problems since the distance function should be a real value. Since quantum mechanics is based on the use of complex numbers, restricting the algebra of vectors to real numbers would eliminate its use in quantum theory. The definition of the inner product can be *fixed* with the following definition. Given two complex vectors \mathbf{x} and \mathbf{y} whose complex components are defined as $\mathbf{x} = (\xi_1, \xi_2, \dots, \xi_n)$ and $\mathbf{y} = (\eta_1, \eta_2, \dots, \eta_n)$ which refer to some basis, then the inner product of \mathbf{x} and \mathbf{y} is given as,

$$\mathbf{x} \circ \mathbf{y} = \sum_{i=1}^n \xi_i^* \cdot \eta_i \quad (12.9)$$

where ξ^* is the complex conjugate of ξ .^[6]

Using this definition the *length* of a vector in a complex vector space is not given as,

$$\|\mathbf{x}\| \equiv \sqrt{\mathbf{x} \circ \mathbf{x}} = \sqrt{\sum_{i=1}^n \xi_i^* \cdot \xi_i} = \sqrt{\sum_{i=1}^n |\xi_i|^2} \quad (12.10)$$

where $|\xi_i|$ is the absolute value of ξ_i . Since $|\xi_i|^2 \geq 0$, $\|\mathbf{x}\|$ is always real. If the vectors \mathbf{x} and \mathbf{y} are in a real vector space then $\xi_i = \xi_i^*$ which reduces the definition of the inner product to that in Eq. (12.7).

Unlike real vector spaces in complex vector spaces the inner product is not symmetrical, so that,

$$\mathbf{x} \circ \mathbf{y} \neq \mathbf{y} \circ \mathbf{x}, \quad (12.11)$$

rather,

$$(\mathbf{x} \circ \mathbf{y}) = (\mathbf{y} \circ \mathbf{x})^*. \quad (12.12)$$

This property will be the foundation of the commutator algebra used throughout quantum mechanics.

With the concept of length of a vector and the angle between two vectors defined, the theory of vector spaces can be applied to physical problems. One important concept in quantum mechanics is that of an *orthogonal* set of vectors (and later a set of functions represented by the vectors).

Two vectors \mathbf{x} and \mathbf{y} are orthogonal if and only $\mathbf{x} \circ \mathbf{y} = 0$. In two a three dimensional vector space two vector are orthogonal if the angle between them is 90° . This definition can be simply generalized to an n -dimensional vector space. If $\mathbf{x} \circ \mathbf{y} = 0$ then $(\mathbf{x} \circ \mathbf{y}) = (\mathbf{y} \circ \mathbf{x})^* = 0$. Although the inner product is not symmetric, the orthogonal condition is symmetric.

For a set of vectors $\{\mathbf{x}_1, \mathbf{x}_2, \dots\}$ the orthogonal condition is met when $\mathbf{x}_i \circ \mathbf{x}_j = \delta_{i,j}$ for all i and j .

§12.2. LINEAR FUNCTIONALS

Corresponding to any linear vector space \mathbf{V} there exists a *dual* space of *linear functionals* on \mathbf{V} . The set of linear functionals on the vector space is itself an n -dimensional vector space. This principal can be applied to diverse applications of mathematics — from economics to combinatorics. A linear functional assigns a scalar function $F(\mathbf{x})$ for each vector \mathbf{x} such that,

$$F(a\mathbf{x} + b\mathbf{y}) = aF(\mathbf{x}) + bF(\mathbf{y}), \quad (12.13)$$

for any vector \mathbf{x} and \mathbf{y} and any scalar a and b .^[7] The linear functionals form a vector space \mathbf{V} . An important theorem in linear functionals is Riesz's Theorem which states there is a one-to-one correspondence between linear functionals F in \mathbf{V}' and the vector \mathbf{x} in \mathbf{V} such that all linear functionals have the form,

$$F(\mathbf{x}) = f \circ \mathbf{x}, \quad (12.14)$$

where f is a fixed vector and \mathbf{x} is an arbitrary vector in the vector space \mathbf{V} [Ries55], [Byro69].^[8] If this condition holds, the two vector spaces \mathbf{V} and \mathbf{V}' are said to be isomorphic.

§12.2.1. Linear Operators

An operator on a vector space maps one vector onto another vector or vectors. If \mathbf{L} is an operator and \mathbf{x} is a vector, then $\mathbf{y} = \mathbf{L}\mathbf{x}$ is another vector. An operator is defined by specifying its actions on every vector in the vector space. A linear operator satisfies the relation,

$$\mathbf{L}(c_1\mathbf{x}_1 + c_2\mathbf{x}_2) = c_1(\mathbf{L}\mathbf{x}_1) + c_2(\mathbf{L}\mathbf{x}_2) \quad (12.15)$$

Two operators are said to be equal, $\mathbf{L} = \mathbf{M}$ if $\mathbf{L}\mathbf{x} = \mathbf{M}\mathbf{x}$ for all vectors in the common domain of \mathbf{L} and \mathbf{M} . The sum of two operators can be defined as,

$$(\mathbf{L} + \mathbf{M})\mathbf{x} = \mathbf{L}\mathbf{x} + \mathbf{M}\mathbf{x} \quad (12.16)$$

and the product of two operators can be defined as,

$$\mathbf{L}\mathbf{M}\mathbf{x} = \mathbf{M}(\mathbf{M}\mathbf{x}) \quad (12.17)$$

From this equation multiplication of operators is associative since $\mathbf{L}(\mathbf{M}\mathbf{N}) = (\mathbf{L}\mathbf{M})\mathbf{N}$, but as will become the foundation of quantum mechanics, the multiplication of operators is not necessarily commutative, since $\mathbf{L}\mathbf{M} \neq \mathbf{M}\mathbf{L}$.

§12.3. DIRAC NOTATION AND LINEAR OPERATORS

There is a *shorthand* notation that is used in quantum mechanics — the Dirac notation.^[9] In this notation the vectors in the vector space \mathbf{V} are called *ket* vectors and are denoted as $|\mathbf{x}\rangle$. The linear functionals in the dual space \mathbf{V}' are called *bra* vectors and are denoted as $\langle F|$. The numerical values of the linear functionals are denoted as,

$$F(\mathbf{x}) \equiv \langle F | \mathbf{x} \rangle. \quad (12.18)$$

According to Riesz's Theorem, there is a one-to-one correspondence between *bras* and *kets*. This notation allows a simplification of the previous linear functional equations where f denoted a fixed vector and F the function applied to the arbitrary vector \mathbf{x} . In the *bra-ket* notation the *bra* $\langle F |$ or the *ket* $|F\rangle$ can be used to determine which vector space \mathbf{V}' or \mathbf{V} is referred to, so that,

$$\langle F | \mathbf{x} \rangle = F \circ \mathbf{x} \quad (12.19)$$

where $|F\rangle$ is the vector previously denoted as f .

The Reisz Theorem establishes an *antilinear* correspondence between *bras* and *kets* such that if $\langle F | \leftrightarrow |F\rangle$ then,

$$c_1^* \langle F | + c_2^* \langle F | \leftrightarrow c_1 |F\rangle + c_2 |F\rangle, \quad (12.20)$$

where the constant c^* is the complex conjugate of constant c .

In a vector space of discrete vectors it is possible to represent the vectors as column or row vectors. In the previous notation the distinction between the row and column vector was not important. In the Dirac notation, this distinction becomes the method of defining the inner product of two vectors. If a column vector is given as,

$$|\mathbf{x}\rangle \equiv \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \end{pmatrix}, \quad (12.21)$$

and a row vector is given by,

$$\langle \mathbf{x} | \equiv (x_1 \quad x_2 \quad x_3 \quad \dots), \quad (12.22)$$

then the *inner product* can be defined as,

$$\langle \mathbf{x} | \mathbf{x} \rangle \equiv \sum_{i=1}^n x_i^* \cdot x_i. \quad (12.23)$$

The algebra of the Dirac notation can nor be summarized. The inner product of a *ket* is always real and given by,

$$\langle \mathbf{x} | \mathbf{x} \rangle \geq 0. \quad (12.24)$$

The distribution law holds for *kets* such that,

$$\langle \mathbf{x} | \mathbf{y} + \mathbf{z} \rangle = \langle \mathbf{x} | \mathbf{y} \rangle + \langle \mathbf{x} | \mathbf{z} \rangle. \quad (12.25)$$

The inner product with one *ket* multiplied by a constant can be restated as,

$$\langle \mathbf{x} | \alpha \mathbf{y} \rangle = \alpha \langle \mathbf{x} | \mathbf{y} \rangle \quad (12.26)$$

The inner product of $\langle \mathbf{x} |$ with $|\mathbf{y}\rangle$ is the complex conjugate of the inner product of $\langle \mathbf{y} |$ with $|\mathbf{x}\rangle$, such that,

$$\langle \mathbf{x} | \mathbf{y} \rangle = \langle \mathbf{y} | \mathbf{x} \rangle^*. \quad (12.27)$$

The result if this inner product $\langle \mathbf{x} | \mathbf{y} \rangle$ is a complex number that is associated with each pair of vectors \mathbf{x} and \mathbf{y} .

The *norm* of a *ket* vector is the positive square root of their inner product, such that,

$$\|\mathbf{x}\| = \sqrt{\langle \mathbf{x} | \mathbf{x} \rangle} \quad (12.28)$$

Using the Dirac notation and changing the notation for a vector to the symbols used in quantum mechanics, the description of a linear operator can be refined. ^[10]

A linear operator can then be represented as a matrix \mathbf{O} which operates on the column vector ψ to produce the another column vector ϕ , such that,

$$\mathbf{O}|\psi\rangle = |\phi\rangle, \quad (12.29)$$

where the operator \mathbf{O} in an n -dimensional vector space is given in the matrix representation as,

$$\mathbf{O} = O_{i,j} = \begin{bmatrix} O_{1,1} & O_{1,2} & \cdots & O_{1,n} \\ O_{2,1} & O_{2,2} & \cdots & O_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ O_{n,1} & O_{n,2} & \cdots & O_{n,n} \end{bmatrix}. \quad (12.30)$$

Choosing an orthonormal basis $\{|u_i\rangle, i=1\dots n\}$ in which to expand the vectors $|\psi\rangle$ and $|\phi\rangle$ gives,

$$|\psi\rangle = \sum_{j=1}^n a_j |u_j\rangle \text{ and } |\phi\rangle = \sum_{k=1}^n b_k |u_k\rangle. \quad (12.31)$$

Operating on Eq. (12.30) with the basis vector $\langle u_i|$ results in,

$$\sum_{j=1}^n \langle u_i | \circ | u_j \rangle a_j = \sum_{k=1}^n \langle u_i | u_k \rangle b_k = b_i.$$

which has the form of a matrix equation given as,

$$\sum_{j=1}^n o_{i,j} a_j = b_i, \quad (12.32)$$

with the expression,

$$o_{i,j} = \langle u_i | \circ | u_j \rangle, \quad (12.33)$$

denoting the *matrix element* of the operator \mathbf{O} . The expression in Eq. (12.33) says that the number $o_{i,j}$ is the vector $\circ | u_j \rangle$ multiplied by the vector $\langle u_i |$.

Using this matrix representation for the operator and a column vector any linear operator can be uniquely specified, given a basis, by specifying the equation of Eq. (12.30) and Eq. (12.33). The effect of the operator \mathbf{O} on a vector ψ can be determined by multiplying the vector by the operator, such that,

$$\begin{aligned} \mathbf{O}|\psi\rangle &= |\phi\rangle \\ &= \begin{bmatrix} O_{1,1} & O_{1,2} & \cdots & O_{1,n} \\ O_{2,1} & O_{2,2} & \cdots & O_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ O_{n,1} & O_{n,2} & \cdots & O_{n,n} \end{bmatrix} \times \begin{bmatrix} \psi_1 \\ \psi_2 \\ \vdots \\ \psi_n \end{bmatrix}, \quad (12.34) \\ &= (O_{1,1} \cdot \psi_1 + O_{1,2} \cdot \psi_2 + \cdots + O_{1,n} \cdot \psi_n) |u_1\rangle + \cdots + (O_{n,1} \cdot \psi_1 + O_{n,2} \cdot \psi_2 + \cdots + O_{n,n} \cdot \psi_n) |u_n\rangle, \\ &= |\phi\rangle. \end{aligned}$$

This leads to a useful property for the analysis of quantum mechanical system. If for some particular operator \mathbf{O} and a particular vector $|\psi\rangle$, that the relation,

$$\mathbf{O}|\psi\rangle = \lambda|\psi\rangle \quad (12.35)$$

exists where λ is a complex number then $|\psi\rangle$ is said to be an *eigenvector* of \mathbf{O} , with an *eigenvalue* of λ , which is the length of the new vector relative to the length of $|\psi\rangle$. Since vector will be eigenvectors of certain operators and certain operators will have some vectors as eigenvectors and not others. The operator–eigenvector relationship will depend only on the vector and the operator, not on the basis in which the vectors are formed. [11]

In the mathematics of quantum mechanics the act of *measuring* is performed by *operating* on the eigenvector to produce a *measurement* represented by an eigenvalue. The possible set of results of measuring an observable is the set of eigenvalues of the corresponding operator [Polk85].

§12.3.1. Measurable Properties

The measurable properties of a quantum system are represented by a linear operators. A particular physical state of the system is represented by a vector in the vector space. If this vector is an eigenvalue of an operator associated with a measurable property of the system, then the state represented by the vector is an eigenstate of the measured property.

The act of *measuring* is performed by *operating* on the eigenvector to produce a *measurement* represented by the eigenvalue. The possible set of results of measuring an observable is the set of eigenvalues of the corresponding operator [Polk89].

If an operator \mathbf{O} , represents a physical dynamical variable, then this operator must be Hermitian. Suppose an operator \mathbf{O} acts on the vector $|\psi\rangle$ to produce another vector $|\phi\rangle$. When the operator acts on the *ket* vector to its right, it is denoted as,

$$\mathbf{O}|\psi\rangle = |\phi\rangle, \quad (12.36)$$

while the operator acts on the *bra* vector $\langle\psi|$ to its left it is denoted as,

$$\langle\psi|\mathbf{O} = \langle\phi|. \quad (12.37)$$

Using the expression in Eq.(12.36) and Eq.(12.37) the products between $\langle\psi|$ and $\mathbf{O}|\psi\rangle$ and $\langle\psi|\mathbf{O}$ and $|\psi\rangle$ are not in general equal such that, $\langle\psi_1|\mathbf{O}|\psi_2\rangle$ usually $\neq \langle\psi_2|\mathbf{O}|\psi_1\rangle$, since one is the product of $\mathbf{O}|\psi_1\rangle$ with $|\psi_2\rangle$ and the other is the product of $\langle\psi_2|\mathbf{O}$ with $|\psi_1\rangle$. If however the products are equal then the operator \mathbf{O} is said to be Hermitian. Hermitian

operators are important in quantum mechanics because their eigenvalue are always *real*.

Using the matrix representation a Hermitian operator can be defined as a complex valued matrix \mathbf{A} . Then the matrix $\bar{\mathbf{A}}$ denotes a matrix obtained from \mathbf{A} by replacing each element of the complex valued matrix, $z = a + ib$ with its conjugate $z = a - ib$. The matrix $\bar{\mathbf{A}}$ is the conjugate of the matrix \mathbf{A} . The matrix \mathbf{A} is *real* $\mathbf{A} = \bar{\mathbf{A}}$. The transpose of the conjugate of the matrix \mathbf{A} is denoted by \mathbf{A}^* , that is

$$\mathbf{A}^* = (\bar{\mathbf{A}})^T \quad (12.38)$$

If

$$\mathbf{A} = (a_{i,j}), \quad (12.39)$$

then the transpose of \mathbf{A} is given by,

$$\mathbf{A}^T = (a_{j,i}), \quad (12.40)$$

and the conjugate of \mathbf{A} is given by,

$$\bar{\mathbf{A}} = (\bar{a}_{i,j}), \quad (12.41)$$

and the transpose of the conjugate is given by,

$$(\bar{\mathbf{A}})^T = (\bar{a}_{j,i}) = \bar{\mathbf{A}}^T \quad (12.42)$$

A matrix \mathbf{A} such that $\mathbf{A} = \mathbf{A}^*$ is a Hermitian matrix. There are several attributes of Hermitian matrices that are important in the mathematics of quantum mechanics: (i) if the eigenvalues of a matrix are distinct, then the associated eigenvectors are linearly independent; (ii) if \mathbf{A} is a Hermitian matrix, then the eigenvalues of \mathbf{A} are real, (iii) if \mathbf{A} is a real symmetric matrix, then the eigenvalues of \mathbf{A} are real; (iv) if \mathbf{A} is a Hermitian matrix, then the eigenvectors associated with the distinct eigenvalues are mutually orthogonal vectors [Pett78], [Aki77].

§12.3.2. Quantum Operators

In the linear vector space algebra, an operator F is defined which acts on a vector $|u\rangle$ to produce another vector $|v\rangle$ such that $|v\rangle = F|u\rangle$. The eigenvalue problem can then be stated as the search to find the vectors $|u\rangle$

and the numbers λ , such that, $F|u\rangle = \lambda|u\rangle$, that is, to find vectors which are transformed by the operator F into multiples of themselves. The vectors satisfying the expression are eigenvectors and the corresponding numbers λ , the eigenvalues. ^[12]

If an operator F is to represent a physical dynamical variable, then it will be Hermitian. Suppose that an operator F , acting on a vector $|u\rangle$, produces a vector $|v\rangle$. When the operators acts on the vector to its right, it is denoted as, $F|u\rangle = |v\rangle$. The corresponding operator which transforms $\langle u|$ into $\langle v|$ is denoted by, $F\langle u| = \langle v|$.

Given the quantities $\langle u_2|F|u_1\rangle$ and $|u_1\rangle F\langle u_2|$ they are not in general equal; the first is the product of $F|u_1\rangle$ with $|u_2\rangle$ and the second is the product of $|u_1\rangle$ with $F|u_2\rangle$. An operator for which these quantities are equal is called *Hermitian*.^[13] The vector $\langle u|$ is then the Hermitian conjugate of the vector $|v\rangle$. Hermitian operators are important because their eigenvalues are *real* and therefore represent observable physical variables. If *any* dynamical variable (operator) F has eigenvectors $|f_n\rangle$ such that $F|f_n\rangle = f_n|f_n\rangle$, then operators representing *other* dynamical variables operate on the eigenstates of F . This relation can be stated in terms of the *commutator*,

$$[F_1, F_2] = F_1F_2 - F_2F_1, \quad (12.43)$$

of the pair of operators F_1 and F_2 .^[14] If F_1 and F_2 posses common eigenstates their commutator is zero (0), such that $F_1F_2 = F_2F_1$. If the states $|n\rangle$ are common eigenstates of F_1 and F_2 , having the respective sets of eigenvalues f_{1n} and f_{2n} , then $F_1F_2 = F_2F_1$ is satisfied for *all* operators.

It should be noted that the commutator defined in Eq. (12.43) is not a Hermitian operator. In order to illustrate the, the general matrix elements of the product k_1k_2 is given as,

$$\langle n|F_1F_2|n\rangle = \sum_{n'} \langle n|F_1|n'\rangle \langle n'|F_2|n\rangle. \quad (12.44)$$

A Hermitian operator is defined as any matrix whose elements are equal to their transpose conjugate. The transpose conjugate of Eq. (12.44) is,

$$\langle n' | F_1 F_2 | n \rangle^* = \sum_{n''} \langle n' | F_1 | n'' \rangle^* \langle n'' | F_2 | n \rangle^* \quad (12.45)$$

which is, since F_1 and F_2 are Hermitian,

$$\sum_{n''} \langle n | F_1 | n'' \rangle \langle n'' | F_2 | n' \rangle = \langle n | F_2 F_1 | n' \rangle. \quad (12.46)$$

Similarly, the transpose conjugate of $\langle n | F_2 F_1 | n' \rangle$ is $\langle n | F_1 F_2 | n' \rangle$, therefore,

$$\langle n' | F_1 F_2 - F_2 F_1 | n \rangle = -\langle n | F_1 F_2 - F_2 F_1 | n' \rangle^*. \quad (12.47)$$

This inequality, due to the *minus sign* may be corrected by multiplying the commutator by i , that is,

$$i[F_1 F_2] = i(F_1 F_2 - F_2 F_1), \quad (12.48)$$

which is Hermitian.

Restating the commutator algebra in terms of the Hamiltonian gives,

$$[u, v] = \sum_r \left\{ \frac{\partial u}{\partial q_r} \frac{\partial v}{\partial p_r} - \frac{\partial u}{\partial p_r} \frac{\partial v}{\partial q_r} \right\}, \quad (12.49)$$

where p_r and q_r are any set of canonical variables [Dira25], with the following algebra,^[15]

$$[u, v] = -[v, u], \quad (12.50)$$

$$[u, c] = 0, \text{ where } c \text{ is a constant,} \quad (12.51)$$

$$\left. \begin{aligned} [u_1 + u_2, v] &= [u_1, v] + [u_2, v], \\ [u, v_1 + v_2] &= [u, v_1] + [u, v_2]. \end{aligned} \right\} \quad (12.52)$$

$$\left. \begin{aligned} [u_1 u_2, v] &= \sum_r \left\{ \left(\frac{\partial u_1}{\partial q_r} u_2 + u_1 \frac{\partial u_2}{\partial q_r} \right) \frac{\partial v}{\partial p_r} - \left(\frac{\partial u_1}{\partial p_r} u_2 + u_1 \frac{\partial u_2}{\partial p_r} \right) \frac{\partial v}{\partial q_r} \right\}, \\ &= [u_1, v] u_2 + u_1 [u_2, v], \\ [u, v_1 v_2] &= [u, v_1] v_2 + v_1 [u, v_2]. \end{aligned} \right\} \quad (12.53)$$

$$[u, [v, w]] + [v, [w, u]] + [w, [u, v]] = 0. \quad (12.54)$$

§12.3.3. *Commutators and Poisson Brackets*

During the development of the Lagrangian and Hamiltonian formalisms, the Poisson bracket notation was presented. The dynamical variables u and v in the Poisson bracket are functions on noncommuting p 's and q 's. This can be shown by defining a set of commuting variables $\alpha_1, \dots, \alpha_k$ and their associated partial derivatives as,

$$\frac{\partial u}{\partial \alpha_i} = \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} [u(\alpha_1, \dots, \alpha_i + \varepsilon, \dots, \alpha_k) - u(\alpha)] \quad (12.55)$$

With noncommuting variables a general derivative can be defined which involves an arbitrary function $c(\alpha)$, such that,

$$\frac{\partial u}{\partial \alpha_i} c(\alpha) = \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} [u(\alpha_1, \dots, \alpha_i + \varepsilon, \dots, \alpha_k) - u(\alpha)] c(\alpha) \quad (12.56)$$

This derivative is called the Fréchet derivative [Park79]. If the function $c(\alpha)$ commutes with all the α 's then,

$$\frac{\partial u}{\partial \alpha_i} c(\alpha) = c(\alpha) \frac{\partial u}{\partial \alpha_i} \quad (12.57)$$

The Fréchet derivative follows the usual derivative rules, such that,

$$\frac{\partial (u+v)}{\partial \alpha_i} c(\alpha) = \frac{\partial u}{\partial \alpha_i} c(\alpha) + \frac{\partial v}{\partial \alpha_i} c(\alpha), \quad (12.58)$$

and,

$$\frac{\partial (au)}{\partial \alpha_i} c(\alpha) = a \frac{\partial u}{\partial \alpha_i} c(\alpha). \quad (12.59)$$

A theorem provided by [Park79] will be used to *equate* the Poisson bracket and the commutator. The theorem is given as,

$$\frac{\partial u}{\partial \alpha_i} [\alpha_i, v] = [u, v]. \quad (12.60)$$

The proof of Park's theorem is given by induction by writing the general function $u(\alpha)$ as a sum of the *pieces*, each of which is of the form $a \cdot \alpha_3 \cdot \alpha_2 \cdot \alpha_1 \cdot \alpha_3 \cdot \alpha_1 \cdot \alpha_4 \cdots$ and it is assumed that the commuting variables u meets this criteria. By adding another factor of α_j to the left–and side of Eq. (12.60) gives,

$$\begin{aligned}\frac{\partial(\alpha_j u')}{\partial \alpha_i}[\alpha_i, v] &= \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} (\alpha_j + \varepsilon [\alpha_j, v] - \alpha_j) - u(\alpha) u' + \alpha_j \frac{\partial u'}{\partial \alpha_i}[\alpha_i, v], \\ &= [\alpha_j, v] u' + \alpha_j [u', v], \\ &= [\alpha_j u', v]\end{aligned}\tag{12.61}$$

If Eq. (12.61) holds for u' , then it holds for $\alpha_j u'$. By setting $u' = \alpha_k$, the theorem holds for the trivial case and has been proved by induction.

The relation between the Poisson bracket and the commutator can now be made by replacing i by j , v by α_i and u by v to give,

$$\frac{\partial u}{\partial \alpha_i}[\alpha_j, \alpha_i] = [v, \alpha_i],\tag{12.62}$$

or

$$\frac{\partial u}{\partial \alpha_i}[\alpha_i, \alpha_j] = [\alpha_i, v].\tag{12.63}$$

The expression can now be substituted into Eq. (12.61) to give,

$$\frac{\partial u}{\partial \alpha_i} \frac{\partial v}{\partial \alpha_j}[\alpha_i, \alpha_j] = [u, v]\tag{12.64}$$

This identity yields the relationship between the Poisson bracket and the commutator by dividing the dynamical variables α_i into two sets, p_1, \dots, p_f and q_1, \dots, q_f with,

$$[q_m, q_n] = i\hbar \delta_{m,n}\tag{12.65}$$

as α_i and α_j progress through the set of p 's and q 's,

$$\frac{\partial u}{\partial q_n} \frac{\partial v}{\partial p_n} i\hbar + \frac{\partial u}{\partial p_n} \frac{\partial v}{\partial q_n} (-i\hbar) = [u, v]\tag{12.66}$$

or using Eq. (12.64),

$$[u, v] = i\hbar \left(\frac{\partial u}{\partial q_n} \frac{\partial v}{\partial p_n} + \frac{\partial u}{\partial p_n} \frac{\partial v}{\partial q_n} \right)\tag{12.67}$$

§12.3.4. *Commutators and the Electromagnetic Field*

A practical application of the commutator is useful at this point. The classical electromagnetic field contained in a *bound* cavity can be *decomposed* into eigenmodes using the commutator process. The i^{th} complex amplitude of the electromagnetic field can be given as,

$$a_i(t) = \frac{(m_i \omega_i q_i(t) + i p_i(t))}{(2\hbar m_i \omega_i)^{1/2}}, \quad (12.68)$$

where m_i is the characteristic mass of the charged particle in the electromagnetic field and q_i and p_i are real canonical variables. In the description ω_i is the eigenfrequency for each eigenmode i . The real and imaginary components of the complex amplitude $a_i(t)$ are related to the canonical variables $q_i(t)$ and $p_i(t)$ respectively.

The electromagnetic field can now be described by defining the complex amplitude as,

$$a_i(t) = a_i e^{-i\omega_i t}. \quad (12.69)$$

This expression describes a complex variable $a_i(t)$ moving with velocity $-\omega$ in a circle around the origin in the complex plane. The real and imaginary components correspond to the electric and magnetic components of the electromagnetic field and the circular motion corresponds to the oscillations of the energy contained in the field.

When the complex amplitude is quantized $a_i(t)$ becomes the field operator $\hat{a}_i(t)$ which obeys the commutator rules,

$$[\hat{a}_i(t), \hat{a}_j^*(t)] = \delta_{i,j}. \quad (12.70)$$

Each eigenmode i has a discrete set of eigenstates $|n_i\rangle$ where n_i is the number of *photons* in mode i .

I like relativity and quantum theories because I don't understand them and they make me feel as if space has shifted about like a swan that can't settle, returning to sit still and be measured; and as if the atom were an impulsive thing always changing its mind.

— D. H. Lawrence in [Lede93]

¹ The description of the mathematical techniques in this section is meant to provide a very broad overview of the subject in preparation for the description of the interaction between the electromagnetic field and charged particles in conductors. The study of the quantum mechanics of fields — quantum field theory — is very complex and even the basics of the concepts are outside the scope of this monograph.

² The notation used in the development of vector algebra follows that found in texts on vector spaces. This notation makes use of the following: (i) italic fonts (\mathbf{x}) represent components of a vector; (ii) bold fonts (\mathbf{x}) represent a vector, real or complex; (iii) bold Arial fonts (\mathbf{O}) represent operators in the vector space; (iv) outline fonts (\mathcal{V}) represent vector spaces. When the vector algebra is used in the description of quantum mechanics ϕ and ψ represent vectors in the complex vector space.

Since the approach taken in this monograph is not based on the specifics of the Heisenberg or Schrödinger formalisms, but rather the operator algebra of P.A.M. Dirac, the generic Dirac notation is presented without the background development found in [Dira58].

³ In quantum mechanics the sum of two vectors will be used to represent the *superposition* of two states of a system.

⁴ The notation used for the scalar and vector product varies dramatically. In this monograph the symbol $\mathbf{a} \circ \mathbf{b}$ will be used to denote the product of the vectors \mathbf{a} and \mathbf{b} , while the symbol $a \cdot b$ will be used to denote the product of the scalars a and b .

⁵ In the analysis of abstract vector spaces the length or magnitude of a vector \mathbf{x} is given by $\|\mathbf{x}\|$ so that the absolute value of a complex number can be denoted by $|\mathbf{x}|$.

⁶ If the number ξ is a complex variable such that $\xi = a + ib$ then the complex conjugate of ξ is given by $\xi^* = a - ib$, where a and b are real scalars.

⁷ The set of linear functionals on an n -dimensional vector space is itself an n -dimensional vector space with respect to the common algebraic operations. The purpose for developing the *dual* space is to describe a Hilbert space in which the coordinates of the vector space take the form of functions whose variables are themselves coordinates in the vector space. An infinite dimensional vector space, with inner product properties can then be used to describe the behavior of quantum mechanical objects [Dunf57], [Youn86].

⁸ The Reisz–Fischer Theorem states that the vector space of square integrable functions, that are functions with a finite norm, are complete. This can be proved by letting

the functions $f_1(x), f_2(x), \dots$ be elements in a function space. If $\lim_{n,m \rightarrow \infty} \|f_n - f_m\|^2 = \lim_{n,m \rightarrow \infty} \int_a^b |f_n(x) - f_m(x)|^2 dx = 0$ then there exists a square (Lebesgue) integrable function $f(x)$ to which the sequence $f_n(x)$ converges to a mean — that is there exists an f such that $\lim_{n,m \rightarrow \infty} \int_a^b |f_n(x) - f_m(x)|^2 dx = 0$.

⁹ In his original publication [Dira58], Dirac assumed a one-to-one correspondence between *bras* and *kets* and it was stated this was a mathematical or physical assumption. The Reisz Theorem shows that this assumption is not needed [Byro70], [Youn86]. In quantum mechanics more general vector space are used, including the Rigged Hilbert space [Böhm78] which the one-to-one correspondence between *bras* and *kets* does not hold [Ball90].

¹⁰ At his point in the monograph the notation will change from that found in mathematics texts to a notation used in quantum mechanics texts.

¹¹ Associated with each square matrix \mathbf{A} , $\mathbf{A} = ((a_{i,j}))$ of order n is a function $f(\lambda) = |\mathbf{A} - \lambda \mathbf{I}|$ called the characteristic function. The equation $f(\lambda) = |\mathbf{A} - \lambda \mathbf{I}| = 0$ can be expressed in the polynomial $c_0 \lambda^n + c_1 \lambda^{n-1} + c_2 \lambda^{n-2} + \dots + c_{n-1} \lambda = 0$ and is the characteristic equation of the matrix \mathbf{A} , where the λ 's are the eigenvalues of the matrix \mathbf{A} . Any non-zero column vector \mathbf{X}_i , such that $(\mathbf{A} - \lambda_i \mathbf{I}) \mathbf{X}_i = 0$ is an eigenvector of the matrix \mathbf{A} .

¹² Associated with each square matrix A , $A = ((a_{ij}))$ of order n is a function $f(\lambda) = |A - \lambda I|$ called the characteristic function. The equation $f(\lambda) = |A - \lambda I| = 0$ can be expressed in the polynomial $c_0 \lambda^n + c_1 \lambda^{n-1} + c_2 \lambda^{n-2} + \dots + c_{n-1} \lambda + c_n = 0$ and is called the characteristic equation of matrix A are called the eigenvalues of A . Any non-zero column vector, X_i , such that $(A - \lambda_i I) X_i = 0$ is called an eigenvector of matrix A .

¹³ If A is a complex matrix, then \bar{A} denotes the matrix obtained from A by replacing each element $z = a + bi$ with its conjugate $\bar{z} = a - bi$. The matrix \bar{A} is the conjugate of matrix A . The matrix A is real iff $A = \bar{A}$. The transpose of the conjugate of matrix A is denoted by A^* that is $A^* = (\bar{A})^T$. If $A = ((a_{ij}))$ then $A^T = ((a_{ji}))$, $\bar{A} = ((\bar{a}_{ji}))$ and $(\bar{A})^T = ((\bar{a}_{ji})) = (\bar{A}^T)$. The transpose of the conjugate of a matrix is equal to the conjugate of the transpose of the matrix.

A matrix A such that $A = A^*$ is a *Hermitian* matrix iff $a_{ij} = \bar{a}_{ji}$ for all ij . Since $a_{ii} = \bar{a}_{ii}$ only if a_{ii} is real. Some theorem important to this discussion are:

- (1) If the eigenvalues of a matrix are distinct, then the associated eigenvectors are linearly independent.
- (2) If A is a Hermitian matrix, then the eigenvalues of A are real.
- (3) If A is a real symmetric matrix, then the eigenvalues of A are real.
- (4) If A is a Hermitian matrix, then the eigenvectors associated with distinct eigenvalues are mutually orthogonal vectors. [Pett78], [Aki77].

¹⁴The use of Heisenberg's *canonical commutation relation* to describe the behavior of quantum mechanical *observable* did not come about through direct development. In three papers on the matrix formulation of quantum mechanics [Heis25], [Born25] and [Born25a] together with papers of Heisenberg [Heis25a] and Born [Born26], the use of Hilbert space concepts formed the beginnings of the theory of quantum mechanics *independent* of the classical description of nature. This process was similar (in analogy only) to Fourier's use in an analytical model to describe the behavior of thermodynamics independent of the underlying physical process. It is the *abstract modeling* process that provides little or no *natural understanding* of quantum mechanics in terms of everyday experience.

¹⁵The method just used to derive the commutator algebra of the Hamiltonian, is based on transforming the classical *Poisson Bracket* formulation of classical mechanics. A more modern derivation based on obtaining operators for particular dynamical variables, does not depend on *quantizing* a classical theory. There are objections to the *classical* quantization method just presented. The first is an epistemological one. If the quantum mechanical equations can be obtained from classical mechanics, then the content of quantum theory must be logically contained in classical mechanics. The second objection is technical. The Poisson Bracket equations of classical mechanics are independent of the particular choice of generalized coordinates. There are several theorems proving the impossibility of such general coordinates quantization [Abra78], [Aren65], [Marg67].