Contents

Chapter 4

Mathematical Background for Quantum Mechanics

4.1 Linear Algebra

Linear algebra is a generalization and abstraction from ordinary vectors. The scalar quantities are allowed to be complex and the vector spaces can have more than three dimensions, and in fact they can be of infinite dimension.

4.1.1 Vectors

An abstract concept that we call a vector will be defined by its algebraic properties. The abstract vector is represented by the symbol $|\alpha\rangle$. A vector space is a set of vectors $(|\alpha\rangle,\ |\beta\rangle,\ |\gamma\rangle\cdots)$. Scalars (which are complex numbers) will be represented by roman letters, $(a, b, c, \cdots)$. The scalars and vectors are required to satisfy a set of algebraic rules.

Addition: Vectors can be added. The sum of two vectors is a vector. Addition is commutative and associative:

$$|\alpha\rangle + |\beta\rangle = |\beta\rangle + |\alpha\rangle = |\gamma\rangle \tag{4.1}$$

$$|\alpha\rangle + (|\beta\rangle + |\gamma\rangle) = (|\alpha\rangle + |\beta\rangle) + |\gamma\rangle \tag{4.2}$$

There exists a **Null vector** defined as that vector which when added to another vector does not change it:

$$|\alpha\rangle + |0\rangle = |\alpha\rangle \tag{4.3}$$

Inverse vector. Every vector has an inverse. The sum of a vector and its inverse is the null vector:

$$|\alpha\rangle + |-\alpha\rangle = |0\rangle \tag{4.4}$$

Scalar multiplication Multiplication by a scalar yields a vector:

$$a|\alpha\rangle = |\gamma\rangle \tag{4.5}$$

$$a(|\alpha\rangle + |\beta\rangle\rangle) = a|\alpha\rangle + b|\beta\rangle \tag{4.6}$$

Linear Independence and dimension of the vector space: An arbitrary linear combination of vectors is a vector:

$$|\lambda\rangle = a|\alpha\rangle + b|\beta\rangle + \cdots \tag{4.7}$$

If there is a vector $|\lambda\rangle$ that can not be expressed as a linear combination of other vectors in the set, it is said to be linearly independent of the set. A set of vectors is **linearly independent** if each one is linearly independent of all the rest. A collection is said to **span the space** if every vector in the space can be written as a linear combination of the members of the collection. The number of vectors needed to span the space is called the **dimension** of the space.

Basis vectors: Once we have a linearly independent set that spans the space, we can express all vectors as linear combinations of them. We will represent a basis set by the symbols $|e_1\rangle$, $|e_2\rangle$, etc. Just as for three dimensional vectors, it is frequently more convenient to work with the components of a vector than with the vector itself. The components of a vector $|a\rangle$ relative to the basis set are defined as the coefficients in the expression:

$$|\alpha\rangle = a_1|e_1\rangle + a_2|e_2\rangle + \cdots \tag{4.8}$$

In this way the vector is specified by its components, $(a_1, a_2, \cdots)$. Multiplication of a vector by a scalar is equivalent to multiplying all its components by the scalar: $c|\alpha\rangle = (ca_1, ca_2, \cdots)$. The sum of two vectors is conveniently expressed in terms of the sums of the components: $|\alpha\rangle + |\beta\rangle$ has components $(a_1 + b_1, a_2 + b_2, \cdots)$.

4.1.2 The inner product

For three dimensional vectors two vectors can be multiplied to yield the scalar product, the length of the first times the length of the projection of the second onto the first. We require that the abstract vectors satisfy a generalized version of the scalar product. First, we define **adjoint** vectors, $\langle\alpha|, \langle\beta|, \cdots$, whose representation will be clear in a while. The adjoint vector contains the same information as the vector, and the two can be converted back and forth into each other. The adjoint set is used to define the **inner product** of two vectors as:

$$\langle\beta||\alpha\rangle = \langle\alpha||\beta\rangle^* \tag{4.9}$$

The inner product, which is the generalization of the three-dimensional scalar product, is a complex number (or scalar). From the definition the inner product of a vector with itself is real. We add the requirement that it be positive:

$$\langle\alpha||\alpha\rangle \quad \geq \; 0 \tag{4.10}$$

and that the answer is 0 only if $|\alpha\rangle = |0\rangle$. A vector space for which the inner product can be defined is called an **inner product space**.

If:

$$|\gamma\rangle = c|\alpha\rangle$$

then the definition of inner product requires that:

$$\langle\gamma| = c^*\langle\alpha|$$

In the future we shall replace the two adjacent vertical lines in the scalar product by one. No confusion will result from this short-hand.

We define the **norm** of a vector to be the scalar:

$$||\alpha|| = \sqrt{\langle\alpha||\alpha\rangle} = \sqrt{\langle a|a\rangle} \tag{4.11}$$

If the norm is 1, the vector is said to be normalized. We can always multiply a vector by the inverse of its norm to normalize it. From now on we shall assume that the basis vectors $|e_i\rangle$ are normalized.

Orthogonal basis sets

If the inner product of two vectors is 0, the vectors are said to be orthogonal. Suppose in our basis set, we have two normalized vectors, $|e_1\rangle$ and $|e_2\rangle$, which are linearly independent but not orthogonal. We can define a new pair of vectors:

$$|e_1'\rangle = |e_1\rangle, \qquad |e_2'\rangle = \alpha\left(c|e_1\rangle + |e_2\rangle\right) \tag{4.12}$$

with their adjoints:

$$\langle e_1'| = \langle e_1|, \qquad \langle e_2'| = \alpha^*\left(c^*\langle e_1| + \langle e_2|\right) \tag{4.13}$$

Since $|e_1\rangle$ and $|e_2\rangle$ are linearly independent, the primed vectors will be linearly independent provided $c \neq 0$. Their scalar product is:

$$\langle e_1'|e_2'\rangle = \alpha\left(c\langle e_1|e_1\rangle + \langle e_1|e_2\rangle\right) = \alpha\left(c + \langle e_1|e_2\rangle\right) \tag{4.14}$$

We can make this 0 by choosing:

$$c = -\langle e_1|e_2\rangle \tag{4.15}$$

Thus from two independent non-orthogonal vectors we can construct two independent orthogonal vectors. The second vector can be normalized by:

$$\langle e_2'|e_2'\rangle = |\alpha|^2\left(c^*\langle e_1| + \langle e_2|\right)\left(c|e_1\rangle + |e_2\rangle\right)$$
$$= |\alpha|^2\left(|c|^2\langle e_1|e_1\rangle + c^*\langle e_1|e_2\rangle + c\langle e_2|e_1\rangle + \langle e_2|e_2\rangle\right)$$
$$= |\alpha|^2\left(1 - |c|^2\right)$$

so that:

$$|\alpha| = \frac{1}{\sqrt{1-|c|^2}}$$

Since $|\alpha|$ is real, it follows that $|c| < 1$. This process can be generalized to any basis set by the Gram-Schmidt process to obtain a normalized and orthogonal basis set. We shall generally assume that a basis set is orthonormal.

We can now write any vector in terms of the orthonormal basis set as:

$$|\gamma\rangle = \sum_i c_i|e_i\rangle$$

To satisfy the rules of the inner product:

$$\langle\gamma| = \sum_i c_i^*\langle e_i|$$

and the inner product of $|\gamma\rangle$ with itself is:

$$\langle\gamma|\gamma\rangle = \sum_i\sum_j c_i c_j^*\langle e_j|e_i\rangle = \sum_i\sum_j c_i c_j^*\delta_{ij} = \sum_i |c_i|^2$$

so that:

$$||\gamma|| = \sqrt{\sum_i |c_i|^2}$$

which is the same way we obtain the length of ordinary vectors in three-dimensions in terms of their components.

Generalization of angle in three-dimensions

In ordinary vector analysis the angle θ between two vectors can be defined by $\cos\ \theta = (\vec{a} \cdot \vec{b})/|\vec{a}||\vec{b}|$. In the abstract vector space, this generalizes to:

$$\frac{\langle\alpha|\beta\rangle}{\sqrt{\langle\alpha|\alpha\rangle\langle\beta|\beta\rangle}}$$

Since the inner product $\langle\alpha|\beta\rangle$ can be complex this expression can't define a real angle. Nevertheless, certain geometrical properties remain valid. For instance the absolute value of the inner product is less than or equal to 1. To prove this, let:

$$|\gamma\rangle = |\beta\rangle - \frac{\langle\alpha|\beta\rangle}{\langle\alpha|\alpha\rangle}|\alpha\rangle$$

with:

$$\langle\gamma| = \langle\beta| - \frac{\langle\alpha|\beta\rangle^*}{\langle\alpha|\alpha\rangle}\langle\alpha|$$

Then, we form the non-negative quantity:

$$\langle\gamma|\gamma\rangle = \langle\beta|\beta\rangle + \frac{|\langle\alpha|\beta\rangle|^2}{\langle\alpha|\alpha\rangle^2}\langle\alpha|\alpha\rangle - \frac{\langle\alpha|\beta\rangle^*\langle\alpha|\beta\rangle}{\langle\alpha|\alpha\rangle} - \frac{\langle\beta|\alpha\rangle\langle\alpha|\beta\rangle}{\langle\alpha|\alpha\rangle}$$

$$= \langle\beta|\beta\rangle + \frac{|\langle\alpha|\beta\rangle|^2}{\langle\alpha|\alpha\rangle} - 2\frac{|\langle\alpha|\beta\rangle|^2}{\langle\alpha|\alpha\rangle}$$

$$= \langle\beta|\beta\rangle - \frac{|\langle\alpha|\beta\rangle|^2}{\langle\alpha|\alpha\rangle}$$

Since $\langle\gamma|\gamma\rangle \geq 0$ it follows that:

$$\frac{|\langle\alpha|\beta\rangle|^2}{\langle\alpha|\alpha\rangle\langle\beta|\beta\rangle} \leq 1$$

This is usually written in the form of the **Schwarz Inequality**:

$$|\langle\alpha|\beta\rangle|^2 \leq \langle\alpha|\alpha\rangle\langle\beta|\beta\rangle$$

4.1.3 Linear Transformations

If a vector is multiplied by a scalar, or if it is rotated, or in some other way converted to a new vector in the vector space, we call the process a transformation. A transformation $\hat{T}$ is linear if:

$$\hat{T}(a|\alpha\rangle + b|\beta\rangle) = a\hat{T}|\alpha\rangle + b\hat{T}|\beta\rangle \tag{4.16}$$

for any vectors $|\alpha\rangle$ and $|\beta\rangle$ and any scalars, a and b. Since the result of the transformation is a vector, it must be possible to represent it in terms of the basis set. We can characterize the effect of the transformation on the basis vectors by:

$$\hat{T}|e_1\rangle = T_{11}|e_1\rangle + T_{21}|e_2\rangle + T_{31}|e_3\rangle + \cdots \tag{4.17}$$

$$\hat{T}|e_2\rangle = T_{12}|e_1\rangle + T_{22}|e_2\rangle + T_{32}|e_3\rangle + \cdots \tag{4.18}$$

etc. We can represent all the equations as:

$$\hat{T}|e_j\rangle = \sum_{i=1}^{n} T_{ij}|e_i\rangle, \qquad (j = 1, 2, \cdots, n) \tag{4.19}$$

For an orthonormal basis, we see that:

$$\langle e_k|\hat{T}|e_j\rangle = \sum_{i=1}^{n} T_{ij}\langle e_k|e_i\rangle = T_{kj} \tag{4.20}$$

With this set of $n \times n$ numbers, we can express the effect of the transformation on any vector:

$$|\alpha\rangle = \sum_{j=1}^{n} a_j|e_j\rangle \tag{4.21}$$

as:

$$\hat{T}|\alpha\rangle = \sum_j a_j\hat{T}|e_j\rangle = \sum_j a_j \sum_i T_{ij}|e_i\rangle = \sum_i \left(\sum_j T_{ij}a_j\right)|e_i\rangle = \sum_i a_i'|e_i\rangle \tag{4.22}$$

The transformation converts the components of the vector a_i into new components relative to the same basis system:

$$a_i' = \sum_j T_{ij}a_j \tag{4.23}$$

The set T_{ij} allows us to calculate the effect of the transformation on any vector. It is convenient to represent this set in the form of a matrix:

$$\mathbf{T} = \begin{pmatrix} T_{11} & T_{12} & \cdots & T_{1n} \\ T_{21} & T_{22} & \cdots & T_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ T_{n1} & T_{n2} & \cdots & T_{nn} \end{pmatrix} \tag{4.24}$$

The sum of two transformations is obtained by adding the matrix representations of each one to obtain the matrix for the sum.

The product of two transformations $\hat{S}\hat{T}$ is defined as follows: First perform the transformation $\hat{T}$ on the starting vector to obtain a new vector. Then perform the transformation $\hat{S}$ on the new vector to obtain the final vector. If:

$$\mathbf{U} = \mathbf{ST} \tag{4.25}$$

then:

$$U_{ij} = \sum_k S_{ik} T_{kj} \tag{4.26}$$

When representing transformations as matrices, it is convenient to represent vectors as columns of numbers. For example, we can represent $|\alpha\rangle$ as:

$$\mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \\ \cdot \\ \cdot \\ \cdot \\ a_n \end{pmatrix} \tag{4.27}$$

With this notation, the transformation rule for vector components can be written as a matrix equation:

$$\mathbf{a}' = \mathbf{Ta} \tag{4.28}$$

The transpose $\tilde{\mathbf{T}}$ of a matrix $\mathbf{T}$ is obtained by interchanging all the rows and columns. This is equivalent to:

$$\tilde{T}_{ij} = T_{ji} \tag{4.29}$$

The transpose of a column vector is a row vector.

A matrix is **symmetric** if the transpose is the same as the original matrix. A matrix is **anti-symmetric** if the transpose is the negative of the original. An anti-symmetric matrix must have all zeros on the diagonal.

The **complex** conjugate of a matrix or a vector is obtained by replacing all the elements with their complex conjugates. The **adjoint** of a matrix is obtained by taking the complex conjugate of its transpose:

$$\mathbf{T}^\dagger = \tilde{\mathbf{T}}^* \tag{4.30}$$

The adjoint of a column vector is a row vector consisting of the complex conjugates of the elements in the column vector. Thus, if we represent $|\alpha\rangle$ as the column vector of eqn. 4.27, we can represent its adjoint $\langle\alpha|$ as the row vector:

$$\mathbf{a} = \begin{pmatrix} a_1^* & a_2^* & \cdot & \cdot & \cdot & a_n^* \end{pmatrix} \tag{4.31}$$

The specific representation for the inner product of vectors is:

$$\langle\alpha|\beta\rangle = \mathbf{a}^\dagger \mathbf{b} \tag{4.32}$$

This definition satisfies all the formal properties that we established for the inner product, and for the definition of adjoint vectors. In particular:

$$\langle\alpha|\beta\rangle = (a_1^*, a_2^*, \cdots, a_n^*) \begin{pmatrix} b_1 \\ b_2 \\ \cdot \\ \cdot \\ \cdot \\ b_n \end{pmatrix} = \sum_i a_i^* b_i \tag{4.33}$$

For the inner product of a vector with itself:

$$\langle\alpha|\alpha\rangle = \sum_i |a_i|^2 \geq 0. \tag{4.34}$$

Only if all the components $a_i = 0$ does the equal sign hold. This is a straightforward generalization of the three-dimensional scalar product of real vectors.

If the adjoint of a matrix is the same as the matrix, the matrix and its associated operation are said to be Hermitian. Hermitian transformations play a central role in quantum mechanics.

Matrix Multiplication

Matrix multiplication is not in general commutative, i.e. $\mathbf{ST} \neq \mathbf{TS}$. We define the difference obtained by reversing the order as the **commutator**:

$$[\mathbf{S}, \mathbf{T}] = \mathbf{ST} - \mathbf{TS} \tag{4.35}$$

The **transpose of a product** is the product of the transposes in the reverse order::

$$\tilde{\mathbf{ST}} = \tilde{\mathbf{T}}\,\tilde{\mathbf{S}} \tag{4.36}$$

The same rule applies to the **adjoint of a product**:

$$(\mathbf{ST})^\dagger = \mathbf{T}^\dagger \mathbf{S}^\dagger \tag{4.37}$$

The **unit transformation**, $\hat{I}$, leaves all vectors unchanged, so that :

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

For an arbitrary transformation $\hat{T}$, it follows that:

$$\hat{T}\hat{I}|\alpha\rangle \equiv \hat{T}\left(\hat{I}|\alpha\rangle\right) = \hat{T}|\alpha\rangle$$

and:

$$\hat{I}\hat{T}|\alpha\rangle = \hat{T}|\alpha\rangle,$$

so that:

$$\hat{I}\hat{T} = \hat{T}\hat{I} = \hat{T}.$$

In matrix form:

$$\mathbf{I} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ 0 & 0 & \cdots & 1 \end{pmatrix} \tag{4.38}$$

Determinant of a matrix:

The determinant of a matrix is a number obtained by adding together all possible different products of n elements, with one element taken from each row and column. The terms in the sum are multiplied by ± 1. To determine this sign, first arrange the elements in each product according to the first index. The product is multiplied by ± 1 depending on whether the second index can be brought into natural order by an even or odd number of permutations of the terms. For example the term $T_{11}T_{22}T_{33}\cdots T_{nn}$ is multiplied by $+1$, and the term $T_{12}T_{21}T_{33}\cdots T_{nn}$ is multiplied by -1. For an N dimensional matrix there are $N!$ terms in the determinant.

Example: 2 x 2 Determinant:

$$D = T_{11}T_{22} - T_{12}T_{21}$$

3 x 3 Determinant:

$$D = T_{11}T_{22}T_{33} + T_{12}T_{23}T_{31} + T_{13}T_{21}T_{32} - T_{13}T_{22}T_{31} - T_{11}T_{23}T_{32} - T_{12}T_{21}T_{33}$$

$N \times N$ determinant:

$$D = \sum_{i=1}^{N} (-1)^{i+1} T_{1i} C_{1i}$$

where C_{1i} is the determinant obtained by eliminating the first row and the i^{th} column from the original array.

Inverse of a matrix: $\mathbf{T^{-1}T} = \mathbf{TT^{-1}} = \mathbf{I}$. A matrix has an inverse only if its determinant is non-zero:

$$(\mathbf{ST})^{-1} = \mathbf{T^{-1}S^{-1}} \tag{4.39}$$

Similarity transformations

Consider a transformation that connects two sets of basis vectors (not necessarily orthonormal):

$$|e_i\rangle = \sum_j S_{ji} |f_j\rangle \tag{4.40}$$

The components of an arbitrary vector are different in the two bases. If the components of $|\alpha\rangle$ are represented in the two bases as $\{a_i^f\}$ and $\{a_i^e\}$, then:

$$a_i^f = \sum_j S_{ij} a_j^e \tag{4.41}$$

Using matrix notation:

$$\mathbf{a^f} = \mathbf{S a^e} \tag{4.42}$$

The components of a transformation matrix also change when the basis is changed. Consider a transformation described in the basis $\{e\}$ by:

$$\mathbf{b^e} = \mathbf{T^e a^e} \tag{4.43}$$

Then in the new basis obtained with the transformation $\mathbf{S}$:

$$\mathbf{b^f} = \mathbf{S b^e} = \mathbf{S T^e a^e} \tag{4.44}$$

But:

$$\mathbf{a^f} = \mathbf{S a^e}, \qquad \mathbf{a^e} = \mathbf{S^{-1} a^f}$$

so:

$$\mathbf{b^f} = \mathbf{S T^e S^{-1} a^f} \equiv \mathbf{T^f a^f}$$

So in the f basis, we have:

$$T^f = S T^e S^{-1} \tag{4.45}$$

The matrix $\mathbf{T^f}$ has the same effect in the basis $\{f\}$ as the matrix T^e in the basis $\{e\}$. Two matrices related in this way are said to be similar, and the transformation is a similarity transformation.

If the two basis sets are both orthonormal, then:

$$\langle e_k| = \sum_l S^*_{lk} \langle e_l| = \sum_l (S^\dagger)_{kl} \langle f_l| \tag{4.46}$$

and:

$$\langle e_k|e_i\rangle = \delta_{ik} = \sum_l S^\dagger_{kl} \sum_j S_{ji} \langle f_l|f_j\rangle = \sum_{l,j} S^\dagger_{kl} S_{ji} \delta_{lj} = \sum_j S^\dagger_{kj} S_{ji} \tag{4.47}$$

so that:

$$\mathbf{S^\dagger S} = \mathbf{I} \tag{4.48}$$

A transformation $\hat{U}$ with the property $\mathbf{U^\dagger} = \mathbf{U^{-1}}$ is called a **unitary transformation**. Thus, a similarity transformation brought about by a unitary transformation preserves the orthogonality of the basis vectors.

Invariants of a similarity transformation:
determinant:
We use a theorem from linear algebra:

$$det(\mathbf{AB}) = det(\mathbf{A}) \cdot det(\mathbf{B}) = det(\mathbf{BA}) \tag{4.49}$$

$$det(\mathbf{ABC}) = det(\mathbf{AB})det(\mathbf{C}) = det(\mathbf{A})det(\mathbf{B})det(\mathbf{C}) \tag{4.50}$$

$$det(\mathbf{T^f}) = det(\mathbf{ST^eS^{-1}}) = det(\mathbf{S^{-1}ST^e}) = det(\mathbf{T^e}) \tag{4.51}$$

The determinant is invariant under a similarity transformation.

Trace:

The trace of a matrix is the sum of the diagonal elements of the matrix.

The trace of a product of matrices can be obtained as follows. Using:

$$(\mathbf{AB})_{ij} = \sum_k A_{ik}B_{kj} \tag{4.52}$$

we can obtain the trace of **AB** as:

$$Tr(\mathbf{AB}) = \sum_i (AB)_{ii} = \sum_i \sum_k A_{ik}B_{ki} \tag{4.53}$$

$$Tr(\mathbf{BA}) = \sum_i (BA)_{ii} = \sum_i \sum_k B_{ik}A_{ki} = \sum_i \sum_k A_{ki}B_{ik} = \sum_k \sum_i A_{ik}B_{ki} \tag{4.54}$$

Thus:

$$Tr(\mathbf{AB}) = Tr(\mathbf{BA}) \tag{4.55}$$

We can apply this identity to the trace of a matrices related by a similarity transformation:

$$Tr(T^f) = Tr(ST^eS^{-1}) = Tr(T^eS^{-1}S) = Tr(T^e) \tag{4.56}$$

The trace is invariant under a similarity transformation.

4.1.4 Eigenvectors and Eigenvalues

For any transformation there are usually certain vectors which are only stretched or contracted by the transformation. These vectors have the property that:

$$\mathbf{Ta} = \lambda\mathbf{a} \tag{4.57}$$

or:

$$(\mathbf{T} - \lambda\mathbf{I})\mathbf{a} = \mathbf{0} \tag{4.58}$$

This equation can be solved for $\mathbf{a} \neq 0$ only if:

$$\det(\mathbf{T} - \lambda\mathbf{I}) = 0 \tag{4.59}$$

This is an n^{th} order polynomial in λ, so that there are n roots. However some of them, or even all of them, could be the same. There is at least one root and no more than n distinct roots. For each root you can solve the simultaneous equations for the components of the corresponding vector.

If the eigenvectors span the space, then they can be used as a set of basis vectors. However, if they are not orthogonal then their value is rather limited. In general, a linear combination of eigenvectors associated with different eigenvalues will not itself be an eigenvector. Thus, the use of the Gram-Schmidt process to orthogonalize the vectors would leave you with a set that are not eigenvectors.

If the set were orthogonal:

$$\hat{T}|\alpha_i\rangle = \lambda_i|\alpha_i\rangle \tag{4.60}$$

$$T_{ij} = \langle\alpha_i|\hat{T}|\alpha_j\rangle = \langle\alpha_i|\lambda_j|\alpha_j\rangle = \lambda_j\langle\alpha_i|\alpha_j\rangle = \lambda_j\delta_{ij} \tag{4.61}$$

Using the eigenvectors as the basis system, the matrix is diagonal:

$$\mathbf{T} = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ . & . & & . \\ . & . & & . \\ . & . & & . \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} \tag{4.62}$$

The eigenvectors in their own basis system are:

$$a^{(1)} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ . \\ . \\ . \\ 0 \end{pmatrix}, \quad a^{(2)} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ . \\ . \\ . \\ 0 \end{pmatrix}, \quad \cdots \quad, a^{(n)} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ . \\ . \\ . \\ 1 \end{pmatrix} \tag{4.63}$$

We will see that a certain class of transformations does permit the above representation.

4.1.5 Hermitian Transformations

Let $|b\rangle = \hat{T}|\alpha\rangle = |\hat{T}\alpha\rangle$ be the vector obtained from $|\alpha\rangle$ by applying the transformation $\hat{T}$. The adjoint of the transformed vector is $\langle b| = \langle\hat{T}\alpha| \equiv$

$\langle\alpha|\hat{T}^\dagger$. To see that this definition of the adjoint is consistent with the matrix representation of vectors and transformations, the i^{th} element of $|b\rangle$ is given by:

$$b_i = \sum_j T_{ij} a_j$$

The i^{th} element of the adjoint of this vector is the complex conjugate:

$$b_i^* = \sum_j T_{ij}^* a_j^* = \sum_j a_j^* T_{ji}^\dagger$$

provided the adjoint matrix is defined by $T_{ji}^\dagger = T_{ij}^*$. In matrix notation:

$$\mathbf{b} = \mathbf{Ta}, \qquad \mathbf{b}^\dagger = \mathbf{a}^\dagger \mathbf{T}^\dagger$$

We can form the inner products:

$$\langle b|a\rangle = \langle a|T^\dagger|a\rangle$$

$$\langle a|b\rangle = \langle a|T|a\rangle$$

Since the definition of the inner product requires that $\langle b|a\rangle = \langle a|b\rangle^*$ it follows that:

$$\langle a|T^\dagger|a\rangle = \langle a|T|a\rangle^*$$

If $|a\rangle$ is an eigenvector of $\hat{T}$ then $T|a\rangle = \lambda|a\rangle$, Then:

$$\langle a|T|a\rangle = \lambda\langle a|a\rangle$$

and:

$$\langle a|T^\dagger|a\rangle = \lambda^*\langle a|a\rangle$$

The transformation $\hat{T}$ is said to be Hermitian when $\hat{T}^\dagger = \hat{T}$, i.e. the transformation is self-adjoint. Hermitian transformations, $\hat{H}$, have three important properties

1. All the eigenvalues are real. From the above equation, if $\hat{T}^\dagger = \hat{T}$ it follows that $\lambda^* = \lambda$ and the eigenvalues are real.

2. The eigenvectors corresponding to distinct eigenvalues are orthogonal:

$$H|h_i\rangle = \lambda_i|h_i\rangle$$

$$\langle h_j|H = \lambda_j\langle h_i|$$

$$\langle h_j|H|h_i\rangle = \lambda_i\langle h_j|h_i\rangle$$

$$\langle h_j|H|h_i\rangle = \lambda_j\langle h_j|h_i\rangle$$

$$(\lambda_i - \lambda_j)\langle h_j|h_i\rangle = 0$$

so if $\lambda_i \neq \lambda j$, $\langle h_j|h_i\rangle = 0$. Q.E.D.

3. For *finite vector spaces*, the eigenvectors span the space. See a math book for the proof.

4. Linearly independent eigenvectors associated with the same eigenvalue can be orthogonalized by the Gram-Schmidt process. Linearly independent eigenvectors associated with the same eigenvalue are said to be **degenerate**.

Since the eigenvectors of a Hermitian transformation span the space and can be made to be orthogonal, they can be used as an orthonormal basis set. In this basis the matrix representation of the transformation associated with the eigenvectors will be diagonal. Different basis sets are connected by a symmetry transformation so the matrix representation of any Hermitian transformation in some arbitrary basis set can be diagonalized by a symmetry transformation. The elements of the diagonal matrix will be the eigenvalues of H:

$$H|h_j\rangle = \lambda|h_j\rangle, \qquad H_{ij} = \langle h_i|H|h_j\rangle = \lambda_j\langle h_i|h_j\rangle = \lambda_j\delta_{ij} \tag{4.64}$$

4.2 Function Spaces

4.2.1 Functions as Vectors

Functions can also be represented in terms of linear algebra. For example, suppose we let $|f\rangle$ and $|g\rangle$ represent the functions $f(x)$ and $g(x)$. These abstract vectors satisfy all the rules of linear algebra, i.e. they can be added, multiplied by a constant, and their inverses are obvious. We can define an inner product as follows:

$$\langle f|g\rangle \equiv \int_a^b f^*(x)g(x)\ dx \tag{4.65}$$

where the limits of the integral are determined by the domain of the functions:

$$\langle f|f\rangle = \int |f(x)|^2\ dx \geq 0, \qquad \text{equality only if f(x)} \equiv 0 \tag{4.66}$$

For this definition to be valid, the integrals must be finite. In particular:

$$\langle f|f\rangle = \int |f(x)|^2\ dx < \infty \tag{4.67}$$

Clearly, not all functions will satisfy this requirement. It turns out that if we restrict the space of functions to those that do (they are said to be square integrable), then $\langle f|g\rangle$ will also be finite.

As an example, we could consider the space P(N) of all polynomials of degree less than N in the domain from $-1 \leq x \leq 1$. Clearly a set of functions that span this space would be:

$$|f_1\rangle = 1, \quad |f_2\rangle = x, \quad |f_3\rangle = x^2, \quad \cdots \quad |f_N\rangle = x^{N-1} \tag{4.68}$$

Not only are these not normalized, they are not orthogonal. We can fix both problems. Let's construct a few of the polynomials. It turns out that if we construct the set for $P(N)$, the same set is valid for $P(N+1)$, provided we add one of order N. We begin with:

$$f_1 = c, \quad |c|^2 \int_{-1}^{1} dx = 2|c|^2 = 1, \quad f_1 = \frac{1}{\sqrt{2}} \tag{4.69}$$

Now let $f_2 = a + bx$. Since f_1 is an even function, we must set $a = 0$ so that f_2 is odd, and therefore orthogonal to f_1. We choose b to normalize f_2:

$$1 = |b|^2 \int_{-1}^{1} x^2 \, dx = \frac{2}{3}|b|^2 \tag{4.70}$$

so:

$$f_2 = \sqrt{\frac{3}{2}}x$$

Next let $f_3 = a(1 + bx + cx^2)$. To make it orthogonal to f_2 we must set $b = 0$. Then choose c to make it orthogonal to f_1:

$$0 = \int_{-1}^{1} (1 + cx^2) \, dx = 2 + 2c/3 \tag{4.71}$$

so that $c = -1/3$, and $f_3 = a(1 - 3x^2)$. To normalize it:

$$1 = |a|^2 \int_{-1}^{1} (1 - 6x^2 + 9x^4) \, dx = |a|^2(2 - 4 + 18/5) = 8/5|a|^2 \tag{4.72}$$

Following convention, we choose $a = -\sqrt{5/8}$ so:

$$f_3 = \sqrt{\frac{5}{8}}(3x^2 - 1)$$

We can continue to build each increasing order in this way. The set of polynomials are called Legendre polynomials, which however were defined long before quantum mechanics, and not normalized as here. Instead the overall constant in each one is chosen so that $P_l(1) = 1$. They alternate in parity, so that the polynomials in even powers of x are even functions, and the polynomials in odd powers are odd functions.

4.2.2 Operators as Linear Transformations

Linear transformations simply convert one function in the function space to some other function in the space and satisfy the linearity condition:

$$\hat{T}(a|f\rangle + b|g\rangle) = a\hat{T}|f\rangle + b\hat{T}|g\rangle \tag{4.73}$$

Simple operations on functions such as multiplication by a power of x, differentiation, or the indefinite integral, transform one function into another. These operations can be treated as linear transformations provided they satisfy the above linearity criterion. As a simple example we can consider the operation: $\hat{x}$ defined as multiplication by x. This clearly does not work in $P(N)$, since it would carry polynomials of order $(N-1)$ out of the space. The operator $\hat{D} \equiv \frac{d}{dx}$ transforms a function to its derivative. If we apply this to the polynomials in $P(N)$ we will get a new polynomial in the same space. Thus for this space $\hat{D}$ is a linear transformation but $\hat{x}$ is not.

In function space, the eigenvectors of a transformation are called *eigenfunctions*. For example the eigenfunctions of $\hat{T}$ must satisfy:

$$\hat{T}|f(x)\rangle = \lambda|f(x)\rangle \tag{4.74}$$

For the derivative:

$$\hat{D}|f(x)\rangle = \lambda|f(x)\rangle \quad \rightarrow \quad \frac{df_\lambda}{dx} = \lambda f_\lambda(x) \tag{4.75}$$

The solution for $f(x)$ yields the eigenfunctions of $\hat{D}$, namely $f_\lambda = Ae^{\lambda x}$. Since the exponential function can not be expressed in terms of a polynomial of finite order, the only eigenfunction of $\hat{D}$ in $P(N)$ is $f_0 = A$, with $\lambda = 0$.

Recall that a Hermitian transformation satisfies:

$$\langle f|\hat{T}g\rangle = \langle \hat{T}f|g\rangle \tag{4.76}$$

Clearly $\hat{x}$ is Hermitian. However, $\hat{D}$ is not:

$$\begin{aligned}\langle f|\hat{D}g\rangle &= \int_a^b f^* \frac{dg}{dx}\, dx = \int_a^b \frac{d(f^*g)}{dx}\, dx - \int_a^b \left(\frac{df}{dx}\right)^* g\, dx \\ &= f^*(b)g(b) - f^*(a)g(a) - \langle \hat{D}f|g\rangle \end{aligned} \tag{4.77}$$

We need to do a couple of things to obtain an Hermitian operator from $\hat{D}$. First, use $i\hat{D}$ instead. Then:

$$\langle f|i\hat{D}g\rangle = i\int_a^b \frac{d(f^*g)}{dx}\, dx - i\int_a^b \left(\frac{df}{dx}\right)^* g\, dx$$

$$= i[f^*(b)g(b) - f^*(a)g(a)] + \langle i\hat{D}f|g\rangle \tag{4.78}$$

Next, restrict the function space to functions for which $f(a) = f(b)$, then the operator $i\hat{D}$ is Hermitian.

Note that the functions of $P(N)$ do not satisfy this condition. In most QM problems we are interested in the domain from $-\infty$ to ∞, and the wave functions must vanish at these limits if they are to be square integrable. For these spaces, $i\hat{D}$ is Hermitian.

The operator $\hat{x}$ is a valid transformation in the infinite polynomial space, $P(\infty)$. It is also Hermitian, as we already pointed out. However, the eigenfunction equation:

$$x(a_0 + a_1x + a_2x^2 + \cdots) = \lambda(a_0 + a_1x + a_2x^2 + \cdots) \tag{4.79}$$

cannot be solved except by setting all the coefficients to 0. The vector $|0\rangle$ is not a legitimate eigenfunction. Hence, even though it is Hermitian, $\hat{x}$ has no eigenfunctions in $P(\infty)$. Does it have eigenfunctions in any space?

$$xf(x) = \lambda f(x) \tag{4.80}$$

The only solution is obtained if $\lambda = x$, from which we conclude that:

$$f_\lambda(x) = A\delta(x - \lambda) \tag{4.81}$$

Thus, in an infinite space $\hat{x}$ has an infinite number of eigenfunctions, one for each point in space. However, the eigenfunctions seem to be a bit pathological.

4.2.3 Hilbert Space

To include the wave functions of quantum mechanics in a system of functions that can be treated with the methods of linear algebra, we need to include a much larger set of functions than are included in the set $P(N)$. To do so we will work with a set of all functions that are square integrable. The set will obviously be an infinite set. An infinite number of basis vectors are required to span the space. A complete inner product space is called a Hilbert space. The Hilbert space of functions in the range from $-1 < x < 1$ is labeled $L_2(-1, 1)$. The set of all square-integrable functions in the interval $a < x < b$ is called $L_2(a, b)$. Mostly we will work in the infinite space, $L_2(-\infty, \infty)$.

The eigenfunctions of $i\hat{D}$ and $\hat{x}$ are:

$$f_\lambda(x) = A_\lambda e^{-i\lambda x}, \qquad g_\lambda(x) = B_\lambda\delta(x - \lambda) \tag{4.82}$$

The eigenvalues are continuously distributed over all real numbers. The set of all eigenvalues are called the spectrum of the operator. These two operators have continuous spectra.

These functions are not however square-integrable and are not therefore normalizable. However they each satisfy an orthogonality relation:

$$\int_{-\infty}^{\infty} f_\lambda(x)^* f_\mu(x)\ dx = A_\lambda^* A_\mu \int_{-\infty}^{\infty} e^{i(\lambda-\mu)x}\ dx = |A_\lambda|^2 2\pi\delta(\lambda-\mu) \tag{4.83}$$

$$\int_{-\infty}^{\infty} g_\lambda(x)^* g_\mu(x)\ dx = B_\lambda^* B_\mu \int_{-\infty}^{\infty} \delta(x-\lambda)\delta(x-\mu)\ dx = |B_\lambda|^2 \delta(\lambda-\mu) \tag{4.84}$$

Although these functions are not normalizable and are not part of the Hilbert space, following the example of Dirac their use is extremely convenient. They are normalized by setting $|A_\lambda|^2 2\pi = 1$ and $|B_\lambda|^2 = 1$.

$$f_\lambda = \frac{1}{\sqrt{2\pi}} e^{-i\lambda x} \tag{4.85}$$

$$g_\lambda = \delta(x-\lambda) \tag{4.86}$$

Can we use these eigenfunctions as basis vectors? Let an arbitrary function $h(x)$ in the L_2 space be represented by the two expansions:

$$|h\rangle = \int_{-\infty}^{\infty} a_\lambda |f_\lambda\rangle\ d\lambda : \qquad |h\rangle = \int_{-\infty}^{\infty} b_\lambda |g_\lambda\rangle\ d\lambda \tag{4.87}$$

where we have replaced the sum over eigenvectors with an integral since the eigenvalues are continuously distributed. We can obtain the expansion coefficients a_λ and b_λ in exactly the same way as we did for a Fourier series:

$$\langle f_\mu | h\rangle = \int_{-\infty}^{\infty} a_\lambda \langle f_\mu | f_\lambda\rangle\ d\lambda\ dx = \int_{-\infty}^{\infty} a_\lambda \delta(\lambda-\mu)\ d\lambda = a_\mu \tag{4.88}$$

so that:

$$a_\mu = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{i\mu x} h(x)\ dx = H(\mu) \tag{4.89}$$

where $H(\mu)$ is just the Fourier transform of $h(x)$.

The components in the $\hat{x}$ representation are:

$$\langle g_\mu | h\rangle = \int_{-\infty}^{\infty} b_\lambda \langle g_\mu | g_\lambda\rangle\ d\lambda\ dx = b_\mu \tag{4.90}$$

so that:

$$b_\mu = \int \delta(x-\mu)\ h(x)\ dx = h(\mu) \tag{4.91}$$

The component b_μ is simply the value of the function at $x = \mu$.

Thus, in the $\hat{D}$ representation the components of $|h\rangle$ are the values of the Fourier transform, and in the $\hat{x}$ representation the components are the values of the function itself. Both representations are equally valid representations of the vector $|h\rangle$. In this sense $h(x)$ is the collection of the components of $|h\rangle$ in the position representation.

We already know that the momentum operator is $\frac{\hbar}{i}\frac{d}{dx}$. Thus, we can write $\hat{p} = -i\hbar\hat{D}$. The eigenfunctions of $\hat{p}$ can be obtained directly from those for $\hat{D}$, or we can re-derive them as follows:

$$\frac{\hbar}{i}\frac{df_p}{dx} = p f_p(x)$$

$$\frac{df_p}{dx} = \frac{ip}{\hbar} f_p(x)$$

$$f_p(x) = \alpha e^{\frac{ipx}{\hbar}}$$

Again we choose α so that $\langle p'|p\rangle = \delta(p-p')$. Thus:

$$\langle p'|p\rangle = |\alpha|^2 \int_{-\infty}^{\infty} e^{i(p-p')x/\hbar}\, dx = |\alpha|^2 \hbar \int_{-\infty}^{\infty} e^{i(p-p')u}\, du = |\alpha|^2 2\pi\hbar\delta(p-p')$$

so that:

$$f_p(x) = \frac{1}{\sqrt{2\pi\hbar}} e^{\frac{ipx}{\hbar}}$$

With the eigenvalues continuously distributed we can not represent operators in terms of matrices, but it turns out that the quantities $\langle f_\lambda|\hat{T}|f_\mu\rangle$ are still perfectly well defined and this is generally called the $\lambda\mu$ component of $\hat{T}$ just as if we could represent it in matrix form.

4.3 Quantum Mechanics

1. The state of a particle is represented by a normalized vector in the Hilbert space, L_2.

2. Observable quantities, which are necessarily real, are represented by Hermitian operators, $\hat{Q}(\hat{x}, \hat{p}, t)$. The expectation value of $\hat{Q}$ in the quantum state $|\Psi\rangle$ is $\langle\Psi|\hat{Q}\Psi\rangle = \langle\Psi|\hat{Q}|\Psi\rangle$. Eigenstates of the operator are given by:
$$\hat{Q}|\Psi_q\rangle = q|\Psi_q\rangle \tag{4.92}$$

3. A measurement of the observable $\hat{Q}$ is certain to return a particular value only if the state is an eigenstate of $\hat{Q}$:

$$\langle\Psi_q|(Q|\Psi_q\rangle) = q \tag{4.93}$$

4. For the measurement of a physical observable to be meaningful, a second measurement of the same observable must get the same result. This means that after a measurement that yields the value q the system must be in the eigenstate of $\hat{Q}$ that has the eigenvalue, q. Moreover, **the result of a physical measurement must be one of the eigenvalues of the operator representing the measurement.** The set of eigenvalues are called the spectrum of the operator.

5. The eigenvectors of an operator $\hat{Q}$ that represents a physical observable evidently must be a complete set (but they need not be in L_2). The eigenvectors of a Hermitian operator in infinite space do not necessarily span the space. The theorem quoted earlier applies only to finite spaces. Thus, physical operators must be Hermitian **and** their eigenvectors must be a complete set.

The case of operators like $\hat{p}$ and $\hat{x}$ needs more discussion. The eigenvalues of these operators are contiuously distributed the eigenvectors will not be in L_2 and can not represent a physical state. This is related to the fact that it is not possible to measure the position or momentum of a particle exactly. An exact measurement of the momentum would require either a finite universe or a very large detector. An exact measurement of position would require the detection of the particle in a box of vanishing size. The measurement is therefore not exact and the wave-function doesn't need to collapse to an eigenfucntion (which is a delta function) but instead to a wave-function that extends over a very small range in values of the measured quantity. All operators with continuously distributed eigenvalues will be of this sort.

If the eigenvalues are discrete:

$$\hat{Q}|e_n\rangle = q_n|e_n\rangle, \quad n = 1, 2, 3, \cdots \tag{4.94}$$

The eigenvectors can always be chosen so that:

$$\langle e_m|e_n\rangle = \delta_{mn} \tag{4.95}$$

The completeness of the eigenvectors means that any quantum state can be represented as:

$$|\Psi\rangle = \sum_n c_n |e_n\rangle \tag{4.96}$$

where:

$$c_n = \langle e_n|\Psi\rangle \tag{4.97}$$

The normalization condition requires that:

$$\langle\Psi|\Psi\rangle = \sum_n |c_n|^2 = 1 \tag{4.98}$$

The expectation value for the measurement of $\hat{Q}$ is:

$$< \hat{Q} >= \langle\Psi|\hat{Q}|\Psi\rangle = \sum(c_m^* \langle e_m|)\hat{Q}\sum(c_n|e_n\rangle) =$$

$$\sum c_m^* c_n \langle e_m|\hat{Q}|e_n\rangle = \sum c_m^* c_n q_n \delta_{mn} = \sum |c_n|^2 q_n = \sum p_n q_n \tag{4.99}$$

where p_n is the probability that a measurement of $\hat{Q}$ on this state yields the particular value q_n:

$$p_n = \sum |c_n|^2 = |\langle e_n|\Psi\rangle|^2 \tag{4.100}$$

If the spectrum is continuous:

$$\hat{Q}|e_k\rangle = \lambda_k |e_k\rangle, \qquad -\infty < k < \infty \tag{4.101}$$

The most commonly chosen bases are the position and momentum eigenvectors, which, as we have observed, are not normalizable and are therefore not even in the L_2 space. However, they satisfy:

$$\langle e_k|e_l\rangle = \delta(k-l) \tag{4.102}$$

Moreover, an arbitrary physical state in L_2 can be represented as:

$$|\Psi\rangle = \int_{-\infty}^{\infty} c(k)|e_k\rangle \; dk \tag{4.103}$$

where:

$$c(k) = \langle e_k|\Psi\rangle \tag{4.104}$$

The probability that a measurement of $\hat{Q}$ will yield a value between k and $k + dk$ is:

$$|c(k)|^2 \; dk = |\langle e_k|\Psi\rangle|^2 \; dk \tag{4.105}$$

which will be finite.

If we choose $\hat{x}$ for the operator $\hat{Q}$, the eigenvectors are:

$$|e_{x'}\rangle = \delta(x - x') \tag{4.106}$$

Then:

$$c(x') = \langle e_{x'}|\Psi\rangle = \int \delta(x - x')\Psi(x,t)\ dx' = \Psi(x',t) \tag{4.107}$$

$$|c(x')|^2\ dx' = |\Psi(x',t)|^2\ dx' \tag{4.108}$$

If we use the momentum eigenvectors:

$$|e_p\rangle = \frac{1}{\sqrt{2\pi\hbar}} e^{ipx/\hbar} \tag{4.109}$$

$$c(p) = \langle e_p|\Psi\rangle = \frac{1}{\sqrt{2\pi\hbar}} \int e^{ipx/\hbar}\Psi(x,t)\ dx \equiv \Phi(p,t) \tag{4.110}$$

The probability that a measurement of the momentum will yield a number between p and $p + dp$ is:

$$|\Phi(p,t)|^2\ dp = |c_p|^2\ dp \tag{4.111}$$

We can represent a quantum state with either the function $\Psi(x,t)$ (**position representation**) or $\Phi(p,t)$ (**momentum representation**). Since the eigenvectors of $\hat{x}$ and $\hat{p}$ do not lie in L_2, no physical state can be an eigenstate of $\hat{x}$ or $\hat{p}$. This will be true for any observable whose eigenvalues are continuously distributed.

Momentum representation of $\hat{x}$

We know that in the x basis $\hat{x}$ is represented by x and $\hat{p}$ is represented by $\frac{\hbar}{i}\frac{d}{dx}$. How is $\hat{x}$ represented in the momentum representation?

$$\langle x\rangle = \int \Psi^*(x,t) x \Psi(x,t) dx$$

$$= \int \left[\frac{1}{\sqrt{2\pi\hbar}} \int \Phi^*(p,t)e^{-ipx/\hbar}dp\right] x \left[\frac{1}{\sqrt{2\pi\hbar}} \int \Phi(p',t)e^{ip'x/\hbar}dp'\right] dx$$

$$= \frac{1}{2\pi\hbar} \int\int \Phi^*(p) \left[\int xe^{-i(p-p')x/\hbar}\ dx\right] \Phi(p')dp'\ dp$$

$$= \frac{1}{2\pi\hbar} \int\int \Phi^*(p) \left[\hbar^2 \int ue^{-i(p-p')u}du\right] \Phi(p')dp'\ dp$$

$$= \int \int \Phi^*(p) \left[i\hbar \frac{d}{dp} \frac{1}{2\pi} \int e^{i(p'-p)u} du \right] \Phi(p') dp' \; dp$$

$$= \int \Phi^*(p) \left[\int \; i\hbar \frac{d\delta(p'-p)}{dp} \Phi(p') dp' \right] \; dp$$

$$= \int \Phi^*(p) \left[i\hbar \frac{d}{dp} \int \delta(p'-p) \Phi(p') dp' \right] \; dp$$

$$= \int \Phi^*(p) \left(i\hbar \frac{d}{dp} \right) \Phi(p) \; dp$$

In the momentum representation, $\hat{x} = i\hbar \frac{d}{dp}$ and $\hat{p} = p$.

4.4 The Uncertainty Principle

Expectation value of a measurement associated with the operator $\hat{A}$ is $\langle \hat{A} \rangle = \langle \Psi | \hat{A} | \Psi \rangle$. For many measurements on identically prepared states, the expectation value of the squared deviation is:

$$\sigma_A^2 = \langle (\hat{A} - \langle \hat{A} \rangle)^2 \rangle = \langle \Psi | [(\hat{A} - \langle \hat{A} \rangle)(\hat{A} - \langle \hat{A} \rangle)] | \Psi \rangle \qquad (4.112)$$

Using the fact that $\hat{A}$ is Hermitian:

$$= \Big(\langle \Psi | (\hat{A} - \langle \hat{A} \rangle) \Big) \Big((\hat{A} - \langle \hat{A} \rangle) | \Psi \rangle \Big)$$

$$= \langle (\hat{A} - \langle \hat{A} \rangle) \Psi | (\hat{A} - \langle \hat{A} \rangle) \Psi \rangle = \langle f | f \rangle \quad \geq 0 \qquad (4.113)$$

Consider two physical observables, $\hat{A}$ and $\hat{B}$, applied to the ensemble of identically prepared states. Let:

$$U = \hat{A} - \langle A \rangle, \qquad \hat{V} = \hat{B} - \langle B \rangle, \qquad \hat{W} = \hat{U} + i\lambda \hat{V} \qquad (4.114)$$

where λ is an arbitrary real parameter. The operators $\hat{U}$ and $\hat{V}$ are Hermitian. We can calculate the inner product of the vector $|W\Psi\rangle$, which is a non-negative quantity:

$$\langle W^2 \rangle = \langle (U + i\lambda V)\Psi | (U + i\lambda V)\Psi \rangle \;\; \geq \;\; 0 \qquad (4.115)$$

$$\langle \Psi | (U - i\lambda V)(U + i\lambda V) | \Psi \rangle \;\; \geq \;\; 0 \qquad (4.116)$$

$$\langle U^2 \rangle - i\lambda(\langle VU \rangle - \langle UV \rangle) + \lambda^2 \langle V^2 \rangle \geq \;\; 0$$

$$\sigma_A^2 + \lambda^2 \sigma_B^2 + i\lambda \langle [U, V] \rangle \geq \;\; 0$$

Since this quantity is real, we see that $\langle [U,V]\rangle$ is 0 or imaginary, The positive quantity on the left is minimum when $\lambda = -i\langle [U,V]\rangle/(2\sigma_B^2)$. Putting in its minimum value:

$$\sigma_A^2 + \frac{\langle [U,V]\rangle^2}{4\sigma_B^2} \geq 0 \tag{4.117}$$

Using the fact that $\langle \hat{A}\rangle$ and $\langle \hat{B}\rangle$ are just numbers, it is easy to show that:

$$[\hat{U},\hat{V}] = [\hat{A},\hat{B}]$$

Like $[\hat{U},\hat{V}]$, , $[\hat{A},\hat{B}]$, the commutator of two arbitrary Hermitian operators, must be zero or pure imaginary.

We now have:

$$\sigma_A^2\sigma_B^2 \geq -\frac{\langle [A,B]\rangle^2}{4} = \left(\frac{\langle [A,B]\rangle}{2i}\right)^2 \tag{4.118}$$

Taking the square root, we obtain Heisenberg's uncertainty principle:

$$\sigma_A\sigma_B \geq |\frac{\langle [A,B]\rangle}{2i}| \tag{4.119}$$

Physical observables associated with commuting operators can both be measured simultaneously to arbitrary precision, but two observables associated with non-commuting operators can not. Another way of saying this is that there is no basis set in which non-commuting operators are both diagonal. Therefore, they cannot share eigenvectors.

As a now familiar example, consider $\hat{x}$ and $\hat{p}$:

$$[\hat{x},\hat{p}] = i\hbar, \qquad \sigma_x\sigma_p \geq \frac{1}{2i}i\hbar = \frac{\hbar}{2} \tag{4.120}$$

Wave function with minimum uncertainty

The equality in the uncertainty principle holds when:

$$\langle W^2\rangle = \langle (U - i\lambda V)\Psi|(U + i\lambda V)\Psi\rangle = 0 \tag{4.121}$$

where:

$$\lambda = -i\frac{\langle [\hat{U},\hat{V}]\rangle}{2\sigma_B^2}$$

This means that:

$$(\hat{U} + i\lambda\hat{V})|\Psi\rangle = 0 \tag{4.122}$$

which by substitution is:

$$(\hat{A} - \langle A \rangle)|\Psi\rangle + i\lambda(\hat{B} - \langle B \rangle)|\Psi\rangle = 0 \tag{4.123}$$

For momentum and position, $\hat{A} = \hat{p}$ and $\hat{B} = \hat{x}$, and $[\hat{A}, \hat{B}] = [\hat{p}, \hat{x}] = -i\hbar$, so that:

$$\lambda = -\frac{\hbar}{2\sigma_x^2}$$

Operator equation 4.123 translates to the differential equation:

$$\frac{\hbar}{i}\frac{d\Psi}{dx} - \langle p \rangle \Psi + i\lambda(x - \langle x \rangle)\Psi = 0 \tag{4.124}$$

$$\frac{d\Psi}{dx} - \frac{\lambda}{\hbar}\left(x - \langle x \rangle + \frac{i}{\lambda}\langle p \rangle\right)\Psi = 0 \tag{4.125}$$

Substituting:

$$u \equiv x - \langle x \rangle \frac{i\langle p \rangle}{\lambda}$$

we obtain:

$$\frac{d\Psi}{du} = \frac{\lambda}{\hbar} u \Psi$$

whose solution is:

$$\Psi = Ae^{\frac{\lambda u^2}{2\hbar}}$$

$$\Psi = Ae^{\frac{\lambda}{2\hbar}[x - \langle x \rangle + i\frac{\langle p \rangle}{\lambda}]^2}$$

$$\Psi = Ae^{\frac{\lambda}{2\hbar}[x - \langle x \rangle]^2} e^{i\frac{\langle p \rangle}{\hbar}[x - \langle x \rangle]} e^{-\frac{\langle p \rangle^2}{2\lambda\hbar}}$$

$$\Psi = A'e^{\frac{\lambda}{2\hbar}[x - \langle x \rangle]^2} e^{i\frac{\langle p \rangle}{\hbar}[x - \langle x \rangle]} = A'e^{-\frac{(x - \langle x \rangle)^2}{4\sigma_x^2}} e^{i\frac{\langle p \rangle (x - \langle x \rangle)}{\hbar}}$$

$$|\Psi|^2 = |A'|^2 e^{-\frac{(x - \langle x \rangle)^2}{2\sigma^2}} \tag{4.126}$$

The minimum uncertainty wave function is the familiar Gaussian. For bound state problems, $\langle p \rangle = 0$, and the wave function is real.

4.4.1 Time Dependence of Expectation Values

$$\frac{d\langle Q \rangle}{dt} = \frac{d}{dt}\left(\langle \Psi | Q\Psi \rangle\right)$$

$$= \langle \frac{\partial \Psi}{\partial t} | Q\Psi \rangle + \langle \Psi | \frac{\partial Q}{\partial t} \Psi \rangle + \langle \Psi | Q \frac{\partial \Psi}{\partial t} \rangle$$

Using the S.E.:

$$H\Psi = i\hbar \frac{\partial \Psi}{\partial t} \tag{4.127}$$

$$\frac{d\langle Q\rangle}{dt} = +\frac{i}{\hbar}\langle H\Psi|Q\Psi\rangle + \langle\Psi|\frac{\partial Q}{\partial t}\Psi\rangle - \frac{i}{\hbar}\langle\Psi|QH\Psi\rangle \tag{4.128}$$

$$\frac{d\langle Q\rangle}{dt} = +\frac{i}{\hbar}\langle\Psi|(HQ - QH)\Psi\rangle + \langle\Psi|\frac{\partial Q}{\partial t}\Psi\rangle$$

$$\frac{d\langle Q\rangle}{dt} = \frac{i}{\hbar}\langle[\hat{H},\hat{Q}]\rangle + \langle\frac{\partial \hat{Q}}{\partial t}\rangle \tag{4.129}$$

This important equation tells us that if $\hat{Q}$ has no explicit time dependence, and if it commutes with the Hamiltonian, its expectation value is a constant of the motion. In particular, if the Hamiltonian itself has no explicit time-dependence, then its expectation value, usually the energy of the system, is a constant of the motion.

We can use the above equation to obtain:

$$\langle[\hat{H},\hat{Q}]\rangle = i\hbar\left(\frac{d\langle Q\rangle}{dt} - \langle\frac{\partial \hat{Q}}{\partial t}\rangle\right) \tag{4.130}$$

If we assume that $\hat{Q}$ has no explicit time-dependence:

$$\sigma_H^2\sigma_Q^2 \geq \frac{-1}{4}\langle[\hat{H},\hat{Q}]\rangle^2 = \frac{\hbar^2}{4}\left[\frac{d\langle Q\rangle}{dt}\right]^2 \tag{4.131}$$

$$\sigma_H\sigma_Q \geq \frac{\hbar}{2}\left|\frac{d\langle Q\rangle}{dt}\right| \tag{4.132}$$

Let the time interval Δt be the time over which $\langle Q\rangle$ changes by one standard deviation, σ_Q:

$$\sigma_Q = |\frac{d\langle Q\rangle}{dt}|\Delta t \tag{4.133}$$

Denote σ_H by ΔE. The above equation then reads:

$$\Delta E\ \Delta t \geq \frac{\hbar}{2} \tag{4.134}$$

which is the famous time-energy uncertainty principle. If the quantum state has a well-defined energy with small ΔE it follows that Δt will be large for all observables $\hat{Q}$, i.e. all properties of the system will change slowly. It also follows that if any property changes rapidly so that its Δt is small, then ΔE cannot be small.

4.4.2 Lifetime and energy of an unstable state

Low-lying excited atomic states have typical lifetimes as short as 10^{-8} s. The life-time is the average time that the state survives. We can label that time as the uncertainty Δt. There is necessarily an uncertainty in the energy associated with this life time:

$$\Delta E \Delta t > \frac{\hbar}{2}; \qquad \Delta E > \frac{\hbar}{2\Delta t} = \frac{3 \times 10^{-16}}{10^{-8}} = 3 \times 10^{-8} \text{ eV} \tag{4.135}$$

Since the energy of the excitation energy of the states if of the order of several eV, the uncertainty in the energy is very small. However, life-times of this order can be measured, by measuring the variation in the energy of the light emitted when the state decays to a lower energy state. In other words, the energy spread ΔE is measured directly.

On a smaller distance scale a proton can be excited to a state with energy about 300 MeV above the ground state. However when the energy of this state is measured, the experiments yield a relatively large spread of energies of the order of $\sigma_E = 50$ MeV. If we interpret this spread as the result of the life-time of the state, we obtain:

$$\Delta t \approx \frac{\hbar}{2\Delta E} = \frac{3 \times 10^{-22}}{50} = 6 \times 10^{-24} \ s \tag{4.136}$$

What does it mean to talk about a state with such a short life-time? Does it really exist at all? The answer to this question is obtained by finding whether this is enough time for the quantum state to be formed. Can the quarks communicate with each other during such a short time? The radius of the proton is about 10^{-15} m. A particle traveling with the velocity of light can travel 2×10^{-15} m in 6×10^{-24} s. Thus, the uncertainty principle tells us that the minimum life-times of the excited states of a proton may be about as short as we can imagine them to be and still qualify as quantum states. Of course, the uncertainty relation does not prevent them from lasting longer, but in fact the large uncertainty in the energy of the states can generally be associated with their short life-times.

4.5 Appendix: Delta Function

The delta function is made respectable by treating it as the limit of the sequence:

$$D_n(x) = \sqrt{\frac{n}{\pi}} e^{-nx^2}, \qquad n = 1, 2, 3 \cdots$$

As n increases the Gaussian function becomes more and more peaked around $x = 0$. The function is normalized:

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

Consequently, for a smooth function $f(x)$:

$$\int_{-\infty}^{\infty} D_n(x) f(x)\ dx \approx f(0) \int_{-\infty}^{\infty} D_n(x)\ dx = f(0)$$

The approximation gets better and better as n gets larger and larger. We can express D_n in terms of the Gaussian integral as:

$$D_n(x) = \sqrt{\frac{n}{\pi}} e^{-nx^2} = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{(itx - t^2/4n)} dt$$

In the limit of $n \to \infty$, we obtain the representation:

$$\delta(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{itx} dt$$

and:

$$\delta(x - x') = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{it(x - x')} dt$$

This representation is consistent with the method of Fourier transform:

$$F(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-ikx} f(x)\ dx$$

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{ikx} F(k)\ dk$$

Taking $f(x) = \delta(x - x')$:

$$F(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-ikx} \delta(x - x')\ dx = \frac{1}{\sqrt{2\pi}} e^{-ikx'}$$

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{ikx} \frac{e^{-ikx'}}{\sqrt{2\pi}}\ dk$$

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ik(x - x')}\ dk$$