The formalism of Quantum Mechanics

I. INTRODUCTION

The formalism of quantum mechanics is discussed in chapter 3 in Griffiths. There is also an Appendix on Linear Algebra. You might read these sections in addition to these notes, and in a few places I will include comments on Griffiths’ presentation. Although these notes are self contained, they cover a more general case than Griffiths, so I recommend that you study section 3.4 on the statistical interpretation of the wave function. Section 3.5 on the generalized uncertainty relation is not covered in these notes.

Here you will directly be introduced to the Dirac formalism, which Griffiths treats in section 3.6 but never uses fully. It is a very powerful formalism and not difficult to learn. It brings out the structure of quantum mechanics in a simple way and has also become increasingly important in many modern application of quantum mechanics such as quantum information.

The mathematics of quantum mechanics is linear algebra. However, there are a couple of differences compared to the linear algebra that you have learned in your mathematics courses. These are:

1. A vector is denoted as $|\alpha\rangle$ (or $\langle\beta|$). This is a trivial change of notation, but it is surprisingly useful.

2. The vectors are complex—i.e. they have complex components. This is a simple generalization. Just some * here and there (well...).

3. The vectors may have an infinite number of components, $a_i, i = 1, 2, 3, \ldots \infty$ and, more importantly, they may have a continuum of components, $a(x), -\infty < x < \infty$ note that here $x$ is a real number. The set of components is in this case a function of the real variable $x$ (and we use the standard notation for functions $a(x)$ instead of $a_x$, which would be more analogous to $a_i$).

Item 3 above is in general a highly non-trivial generalization of linear algebra, and some of the properties that hold in finite dimensional vector spaces may no longer be true. Fortunately, however, for the vectors that occur in quantum mechanics this does not happen and the
generalization is simple! Thus we need not worry about the things a mathematician would start to think about. We will just generalize the standard linear algebra to vectors whose components are functions. Why is it that the vectors of quantum mechanics have this nice property? Well, it is really something required by physics. The properties in question are namely related to the Born interpretation of the wave function as a probability amplitude, and what can go wrong is that probabilities would not add up to one. This is of course not acceptable for a physical theory, so in quantum mechanics we only use those vector spaces for which this holds true. It might then be a tricky mathematical problem to prove these properties, i.e. to prove that we are using a consistent mathematical model of a physical system. For the cases you will encounter in this course you can rest assured that these problems have been sorted out.[1]

How linear algebra enters quantum mechanics will be explored and explained later, but it might be good to know from the start what to look for. Here is a brief preview and an illustrative example:

1. The state of a system is described by a vector.

2. Observables, i.e. things that can be measured, correspond to operators that implement linear transformations of the vectors.

3. A crucial feature of quantum mechanics is the principle of superposition, the linear combination \( a \psi(x) + b \phi(x) \) of two wave functions is a possible wave function for the system. In linear algebra this is just the basic property of a vector space, which says that a linear combination of vectors is a new vector in the space (see (6) below).

Example: You have already encountered the eigenfunctions \( \psi_n(x) \) and energy eigenvalues \( E_n = \hbar \omega (n + \frac{1}{2}) \) for the harmonic oscillator hamiltonian. It turned out to be very useful to introduce the ladder operators, \( a \) and \( a^\dagger \) satisfying the commutation relation:

\[
[a, a^\dagger] = 1 \quad (1)
\]

Their action on the normalized eigenfunctions were,

\[
a |n\rangle = \sqrt{n} |n-1\rangle \quad (2)
\]

\[
a^\dagger |n\rangle = \sqrt{n + 1} |n+1\rangle
\]
From these we formed the number operator $\hat{N} = a^\dagger a$. Note that all these operators are differential operators acting on wave functions $\psi(x)$, and that the origin of the commutation relation (1) is the operator identity $\frac{d}{dx}x = x \frac{d}{dx} + 1$. Now, there is another type of objects that in general do not commute, namely matrices. So it is quite natural to ask whether the commutation relation (1) could be represented by matrices. It is not hard to prove that this is in fact not possible using finite matrices. However, if we allow for infinite matrices and define,

$$a = \begin{pmatrix} 0 & 1 & 0 & \sqrt{2} & 0 & \sqrt{3} & 0 & \ldots \end{pmatrix}, \quad a^\dagger = \begin{pmatrix} 0 & 1 & 0 & \sqrt{2} & 0 & \sqrt{3} & 0 & \ldots \end{pmatrix},$$

(3)

it is clear (just do the multiplication!) that (1) is satisfied, if the ”1” on the LHS is interpreted as the infinite diagonal unit matrix, $1 = \text{diag}(1,1,\ldots)$. The number operator takes the form,

$$\hat{N} = \begin{pmatrix} 0 & 1 & 2 & \ldots \end{pmatrix}.$$  

(4)

If we now let these matrices act on the column vectors

$$\psi_0 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \psi_1 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad \psi_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad \text{etc.}$$

(5)

the relations (2) are fulfilled. Since the hamiltonian is given by the operator relation $H = \hbar\omega(\hat{N} + \frac{1}{2})$ we see that the energy eigenvalue for the state $\psi_n$ is $E_n = \hbar\omega(n + \frac{1}{2})$ as expected.
We thus have two alternative descriptions of the harmonic oscillator states, either in terms of wave functions $\psi_n(x)$ and differential operators acting on these functions, or in terms of (infinite) vectors, $\psi_n$ and matrices acting upon them. We say that these two descriptions are different representations of the quantum mechanical system. Mathematically, the function $\psi_n(x)$ and the infinite column vector $\psi_n$ are just the components of the same abstract state vector $|n\rangle$ in different basis. The first correspond to the position basis $|x\rangle$ and the second to the energy basis $|E\rangle$. All this will be explained in detail in the following.

**II. DIRAC’S NOTATION**

A vector is denoted as $|\alpha\rangle$ (called a ket-vector or simply a ket), or, as $\langle\beta|$ (called a bra-vector or simply a bra). This is standard quantum mechanics notation introduced by Dirac.[2] Note that a bra and ket together gives a bracket $\langle\alpha|\beta\rangle$ (this is a physicists sense of a joke). $\langle\alpha|\beta\rangle$ denotes the scalar product of the vectors $|\alpha\rangle$ and $|\beta\rangle$.

**III. COMPLEX VECTORS**

A vector space consists of vectors and scalars. Two vectors $|\alpha\rangle$, $|\beta\rangle$ can be added and multiplied by scalars $a, b$ and this gives a new vector $|\gamma\rangle$

$$|\gamma\rangle = a|\alpha\rangle + b|\beta\rangle$$  \hfill (6)

in the vector space. As noted above this is the mathematical formulation of the quantum mechanical superposition principle. A vector can be expanded in a set of linearly independent basis vectors $|e_i\rangle$, $i = 1, 2, ... n$,

$$|\alpha\rangle = \sum_{i=1}^{n} a_i |e_i\rangle .$$  \hfill (7)

The coefficients $a_i$ are complex scalars and are called the components of the vector, they describe the vector completely [3] [4]. $n$ is the dimension of the vector space. For now we shall take $n$ to be finite, but later we shall also discuss vector space with infinitely many components, as the one considered in the example in section I. You are probably used to the scalars being real numbers. In the vector spaces that occur in quantum mechanics the scalars are, in general, complex (they must be since the wave function is complex, and this is an example of components of a vector). This leads to some simple changes.
The scalar product of two vectors \( |\alpha\rangle, |\beta\rangle \) is denoted \( \langle \alpha|\beta \rangle \) and has the properties

\[
\langle \beta|\alpha \rangle = \langle \alpha|\beta \rangle^* \quad \text{(8)}
\]

\[
\langle \alpha|\alpha \rangle \geq 0, \quad \text{and} \quad \langle \alpha|\alpha \rangle = 0 \text{ if and only if } |\alpha\rangle = 0 \quad \text{(9)}
\]

\[
\langle \gamma|(a|\alpha\rangle + b|\beta\rangle) = a\langle \gamma|\alpha \rangle + b\langle \gamma|\beta \rangle \quad \text{(10)}
\]

Note the complex conjugation that enters in the first relation.

Let \( |e_i\rangle \) be an orthonormal basis,

\[
\langle e_i|e_j \rangle = \delta_{ij} \quad \text{(11)}
\]

The components of \( |\alpha\rangle = \sum_{i=1}^{n} a_i|e_i\rangle \) are then

\[
a_i = \langle e_i|\alpha \rangle \quad \text{(12)}
\]

and the scalar product of \( |\alpha\rangle \) and \( |\beta\rangle = \sum_{i=1}^{n} b_i|e_i\rangle \) is

\[
\langle \alpha|\beta \rangle = \sum_{i=1}^{n} a_i^*b_i \quad \text{(13)}
\]

and the square of the norm of the vector \( |\alpha\rangle \) is

\[
\langle \alpha|\alpha \rangle = \sum_{i=1}^{n} a_i^*a_i \quad \text{(14)}
\]

To each ket \( |\alpha\rangle = \sum_{i=1}^{n} a_i|e_i\rangle \), there is a corresponding bra

\[
\langle \alpha| = \sum_{i=1}^{n} \langle e_i|a_i^* \quad \text{(15)}
\]

Note the complex conjugation that enters on the components of \( \langle \alpha| \), this follows from (8) and makes the norm, i.e. the length of the vector, real and non-negative.

I will frequently use the simplified notation \( |i\rangle = |e_i\rangle \).

IV. LINEAR TRANSFORMATIONS

A linear transformation \( \hat{T} \) takes a vector \( |\alpha\rangle \) and transforms it into another vector \( |\alpha'\rangle = \hat{T}|\alpha\rangle \) in such a way that the linearity condition

\[
\hat{T}(a|\alpha\rangle + b|\beta\rangle) = a\hat{T}|\alpha\rangle + b\hat{T}|\beta\rangle \quad \text{(16)}
\]
is fulfilled. $\hat{T}$ is called an operator. Because of this linearity relation, the action of the operator $\hat{T}$ on an arbitrary vector can be obtained from its action on the basis vectors $|i\rangle$:

$$\hat{T}|j\rangle = \sum_{i=1}^{n} T_{ij}|i\rangle \quad \text{Note order of indices!}$$  \hspace{1cm} (17)

$$T_{ij} = \langle i|\hat{T}|j\rangle .$$  \hspace{1cm} (18)

$T_{ij}$ are the components of the operator $\hat{T}$, they form an $n \times n$-matrix $T$. If $|\alpha\rangle = \sum_{i} a_{i}|i\rangle$ and $|\alpha'\rangle = \sum_{i} a'_{i}|i\rangle$ we introduce the notation $a$ and $a'$ for the column vectors (i.e. the $n \times 1$-matrices) with components $a_{i}$ and $a'_{i}$ respectively. Then we can write $|\alpha'\rangle = \hat{T}|\alpha\rangle$ in matrix language as

$$\sum_{i} a'_{i}|i\rangle = \hat{T} \sum_{j} a_{j}|j\rangle = \sum_{i,j} a_{j}T_{ij}|i\rangle$$  \hspace{1cm} (19)

$$a'_{i} = \sum_{j} T_{ij}a_{j}$$  \hspace{1cm} (20)

$$a' = Ta$$  \hspace{1cm} (21)

Note that matrix equations, such as (19), depend on the particular basis $|i\rangle$ used.

V. HERMITIAN TRANSFORMATIONS

If $\hat{T}$ is an operator then one defines the hermitian conjugate operator $\hat{T}^{\dagger}$ as follows

$$\langle \alpha|\hat{T}^{\dagger}|\beta\rangle = \langle \beta|\hat{T}^{\dagger}|\alpha\rangle^{*}$$  \hspace{1cm} (22)

(for all vectors $|\alpha\rangle, |\beta\rangle$).

[The definition (22) is equivalent to Griffiths [A.87] $\langle \hat{T}^{\dagger}|\alpha|\beta\rangle = \langle \beta|\hat{T}|\alpha\rangle$ ([3.83] in the first edition). His notation is confusing and I will try to avoid it, but you need to know it in order to read and understand the textbook. What stands inside $|...\rangle$ is just a name for the vector, and a natural notation is then $|\hat{T}\beta\rangle = \hat{T}|\beta\rangle$. Using this notation one finds

$$\langle \hat{T}^{\dagger}|\alpha|\beta\rangle = \langle \beta|\hat{T}^{\dagger}|\alpha\rangle^{*} = \langle \beta|\hat{T}\dagger|\alpha\rangle^{*}$$  \hspace{1cm} (23)

which, when inserted in Griffiths’ eq. [A.87] gives (the complex conjugate of) our definition (22). Note that $\langle \hat{T}^{\dagger}|\alpha\rangle$ is not equal to $\langle \alpha|\hat{T}^{\dagger}$, whereas $|\hat{T}\alpha\rangle = \hat{T}|\alpha\rangle$.]

6
A very important type of operators in quantum mechanics are those which hermitian conjugate equals the operator itself:

\[ \hat{T}^\dagger = \hat{T} \quad . \]  

(24)

Such an operator is said to be **hermitian** or **self-adjoint**. It follows form (22) that the components of a hermitian operator obey \((T_{ji})^* = T_{ij}\). The matrix \(T\) is then said to be hermitian and this condition is written as \(T^\dagger = T\).

For a product of operators one has the useful rule

\[ (\hat{S}\hat{T})^\dagger = \hat{T}^\dagger\hat{S}^\dagger \quad . \]  

(25)

This can be proven using (22), or by using components, *i.e.* matrix notation. Note the change of order of the operators.

Hermitian operators are of crucial importance in quantum mechanics since any observable, such as *e.g.* position, momentum, angular momentum, energy, spin, *etc.* ... for any physical system corresponds to a hermitian operator. [5]

### VI. EIGENVALUES AND EIGENVECTORS

Given an operator \(\hat{T}\), consider the equation

\[ \hat{T}|\alpha\rangle = \lambda|\alpha\rangle \quad , \]  

(26)

where \(\lambda\) is a scalar. This is an eigenvalue equation. Vectors \(|\alpha\rangle\) that solve this equation are called **eigenvectors** of the operator \(\hat{T}\) since they are mapped into themselves by \(\hat{T}\), and \(\lambda\) are called **eigenvalues**. In matrix notation this equation becomes

\[ Ta = \lambda a \quad \text{or, in components,} \quad \sum_{j=1}^{n} T_{ij}a_j = \lambda a_i \quad . \]  

(27)

You have encountered this equation in your linear algebra courses. Rewriting it as \((T - \lambda 1)a = 0\), one sees that it has solutions \(a \neq 0\) only if \(\det(T - \lambda 1) = 0\). Solving this characteristic equation ('sekularekvationen'), which is an \(n\)th order polynomial in \(\lambda\), determines the \(n\) eigenvalues \(\lambda_i, \ i = 1 \ldots n\). Having obtained the eigenvalues one can determine the corresponding eigenvectors by solving (26). If \(T\) happened to be a diagonal matrix then, obviously, the eigenvalues are simply the diagonal elements. In general, to find the
eigenvalues is equivalent to diagonalizing the matrix $\mathbf{T}$ by a change of basis vectors. $\mathbf{T}$ is diagonal when the basis vectors are the eigenvectors.

The set of all eigenvalues $\{\lambda_i\}$ and eigenvectors $\{|i_s\rangle\}$ for an operator is called the spectrum of the operator. If there is only one eigenvector, $|i\rangle$ with a particular eigenvalue $\lambda_i$ then this eigenvalue is said to be non-degenerate, if there are $d$ linearly independent eigenvectors, $|i_s\rangle$, $s = 1 \ldots d$ with eigenvalue $\lambda_i$ this eigenvalue is said to be $d$-fold degenerate. From now on we shall consider the non-degenerate case.

If $\hat{T}$ is a hermitian operator, then the following holds

1. All eigenvalues are real.

2. Eigenvectors corresponding to different eigenvalues $\lambda \neq \lambda'$ are orthogonal.

3. The eigenvectors span the vector space.

Item 3 means that the eigenvectors of a hermitian operator can be chosen as a basis $|i\rangle$, $i = 1, 2 \ldots n$ in the vector space. If all eigenvalues are non-degenerate, the corresponding eigenvectors are automatically orthogonal because of 2. For a degenerate eigenvalue, one may have to take linear combinations of the obtained eigenvectors to obtain orthonormal basis vectors: $\langle i|j\rangle = \delta_{ij}$.

Properties 1 and 2 are simple to prove, see Griffiths. The property 3 is obvious if all eigenvalues are different and can be shown to hold in general.

In quantum mechanics, an observable, e.g. the energy, is represented by a hermitian operator and the possible results of the measurement of this observable are the eigenvalues of this operator. These must be real and they are since the operator is hermitian. After one has measured the observable and obtained one of the eigenvalues the state of the system is described by the corresponding eigenvector—this is the so called collapse of the wave function.
VII. SUMMARY OF RESULTS AND THE RESOLUTION OF THE IDENTITY

For a hermitian operator we have from the above

\[ \hat{T} = \hat{T}^\dagger \]  
(28)

\[ \hat{T}|i\rangle = \lambda_i|i\rangle \]  
(29)

|\rangle, \quad i = 1, 2, \ldots, n \text{ basis of eigenvectors, } \lambda_i \text{ real eigenvalues} \]  
(30)

\[ \langle i|j\rangle = \delta_{ij} \quad \text{orthonormal basis} \]  
(31)

|\psi\rangle = \sum_{i=1}^{n} \psi_i|i\rangle , \quad \psi_i = \langle i|\psi\rangle \]  
(32)

\[ |\psi\rangle = \sum_{i=1}^{n} |i\rangle\langle i|\psi\rangle = (\sum_{i=1}^{n} |i\rangle\langle i|)|\psi\rangle . \]  
(33)

From the last equation we read off the operator identity (called the ‘resolution of the identity’)

\[ 1 = \sum_{i=1}^{n} |i\rangle\langle i| . \]  
(34)

Note that here ”1” is an operator (and should perhaps be denoted by \( \hat{1} \)). The identity (34) is a consequence of the vectors |\rangle forming a basis. This identity is very useful, it can be inserted anywhere in an equation since it is just the identity operator and many useful formulae can be derived almost without effort, for example the components of |\psi\rangle above.

As a further example we derive the components of an arbitrary operator \( \hat{A} \) in the basis \{\rangle\} defined by the hermitian operator \( \hat{T} \):

\[ \hat{A} = 1 \cdot \hat{A} \cdot 1 = \sum_{i,j} |i\rangle\langle i|\hat{A}|j\rangle\langle j| = \sum_{i,j} |i\rangle A_{ij} \langle j| \]  
(35)

\[ A_{ij} = \langle i|\hat{A}|j\rangle . \]  

Note that \( \hat{A} \) does not have to be hermitian. In particular, if we calculate the components of the operator \( \hat{T} \) itself in its own basis we get,

\[ T_{ij} = \langle i|\hat{T}|j\rangle = \lambda_i \langle i|j\rangle = \lambda_i \delta_{ij} \]  
(36)

which just expresses the trivial fact that the operator \( \hat{T} \) is diagonal in the basis defined by its own eigenvectors.
So far everything was simple, every hermitian operator in a finite dimensional vector space has a complete set of eigenvectors that can be chosen as basis vectors (and the eigenvalues are real). As mentioned above the only new things compared to the linear algebra you have met in your mathematics course is that the vector space is complex but this was easily taken care of.

We now want to extend the theory of linear algebra to infinite dimensional vector spaces. The vector index $i$ may run from 1 to $\infty$ (i.e. $n \rightarrow \infty$ in the formulae above) but it may also be that the vector index becomes a continuous variable $x$ which takes for example all real values. The components of a vector $\psi_i$ then becomes a function of $x$, $\psi(x)$—thus functions can be thought of as components of vectors.

Formally, the generalization to infinite dimensional vector spaces is simple. However, mathematical difficulties may arise. For example, it may happen that the scalar product of two vectors is infinite. Moreover, the theorems that hold for a hermitian operator in a finite dimensional vector space—that the eigenvectors span the space—do not in general hold for a hermitian operator in an infinite dimensional vector space. Thus infinite dimensional vector spaces are mathematically much trickier than their finite dimensional counterparts. See Griffiths for a discussion of these problems. However, it so happens that the infinite dimensional vector spaces that occur in quantum mechanics are simple and behave exactly as their finite dimensional brethren! In particular, the hermitian operators that occur in quantum mechanics do have eigenvectors that span the vector space and hence can be used as a basis. These eigenvectors may have an infinite norm, however this infinity is of the same fairly innocent type as we encountered for the free particle.

Thus, for the purposes of quantum mechanics one can generalize linear algebra to infinite dimensional vector spaces almost trivially without worrying about whether things are well-defined or not. Quantum mechanics guarantees that this is the case. By using the bra, ket notation of Dirac the formulae and the calculations will look almost the same also when there is a continuum of components.

In my opinion, Griffiths puts too much emphasis on the potential mathematical difficulties in infinite dimensional vector spaces and he also stops using the Dirac notation that he has introduced. This obscures the close analogy with the finite dimensional case. Here, we will use Dirac’s vector notation and trust that the mathematics works just as in the finite dimensional case as is required for the consistency of quantum mechanics.
VIII. THE INFINITE DIMENSIONAL DISCRETE SPECTRUM: \( i = 1, 2, \ldots, \infty \)

As noted above this generalization is trivial, just make the replacement \( n \to \infty \) in the formulae above.

IX. THE CONTINUOUS SPECTRUM: \( i \to x \)

There are two new things: The sum is replaced by an integral and the Kronecker delta is replaced by Dirac’s delta function[6]

\[
\sum_{i=1}^{n} \to \int_{-\infty}^{\infty} dx \\
\delta_{ij} \to \delta(x - x')
\]

The equations (28-33) become

\[
\hat{T} = \hat{T}^\dagger \\
\hat{T}|x\rangle = \lambda_x |x\rangle \\
|x\rangle, \quad -\infty < x < \infty \text{ basis of eigenvectors, } \lambda_x \text{ real eigenvalues} \\
\langle x|x'\rangle = \delta(x - x') \text{ orthonormal basis} \\
|\psi\rangle = \int_{-\infty}^{\infty} dx \psi(x)|x\rangle, \quad \psi(x) = \langle x|\psi\rangle \\
|\psi\rangle = \int_{-\infty}^{\infty} dx |x\rangle \langle x|\psi\rangle = \left( \int_{-\infty}^{\infty} dx |x\rangle \langle x| \right) |\psi\rangle
\]

\( x \) here can be the position of a particle, but can also be any other continuum variable such as the momentum or the energy for the scattering states.\[7\] From the last equation we read off the operator identity (called the ‘resolution of the identity’)

\[
1 = \int_{-\infty}^{\infty} dx |x\rangle \langle x|
\]

Compare this to the finite dimensional case above! The resolution of the identity can be used to derive results just as above, using the defining property of Dirac’s delta function

\[
\int dx' \delta(x - x')f(x') = f(x)
\]

Note that Dirac’s delta function vanishes when \( x \neq x' \) and is \( \infty \) when \( x = x' \). Thus eigenvectors corresponding to different eigenvalues are orthogonal. However, the norm of the eigenvectors is infinite. (Compare the notes on the free particle.)
Comment: It is also possible to have an operator whose spectrum \((i.e.\) set of eigenvalues) consists of both a discrete part and a continuous part—the Hamiltonian for the Hydrogen atom is an example of this and so is the delta-potential well. In these cases, one will have both an integral \(\int\) and a \(\sum\) in the formulae.

X. THE SCALAR PRODUCT

It is straightforward to derive the expression for the scalar product in terms of the components of the vectors, Eq [3.6] in Griffiths:

\[
|\psi\rangle = \int dx |x\rangle \psi(x), \quad |\phi\rangle = \int dx |x\rangle \phi(x)
\]

\[
\langle \phi|\psi \rangle = (\int dx \phi^*(x)|x\rangle)(\int dx' |x'\rangle \psi(x')
\]

\[
= \int dx \int dx' \phi^*(x) \psi(x') \langle x|x' \rangle
\]

using \(\langle x|x' \rangle = \delta(x - x')\) and the defining property of the delta function (46), we arrive at the important result

\[
\langle \phi|\psi \rangle = \int dx \phi^*(x) \psi(x).
\]

Compare this expression to the finite dimensional one eq.(13).

Note that Griffiths takes the expression (50), in terms of the components \(\psi(x)\) of the vector \(\psi\) as the definition of the scalar product. This is possible but obscures the general structure and the similarities to the finite dimensional case. The integral in (50) must exist, thus this leads to restrictions on the allowed functions \(\psi(x)\), see Griffiths.

XI. CHANGE OF BASIS

Using the resolution of the identity it is very easy to perform a change of basis. We illustrate this here with the discrete case, but the formulae translate immediately to the general case.

Let \(|e_i\rangle, i = 1, ...n\) and \(|f_i\rangle, i = 1, ...n\) be two sets of normalized basis vectors (they may for example be the eigenvectors of two operators corresponding to two physical observables).
There are now two resolutions of the identity

\[ 1 = \sum_i |e_i\rangle\langle e_i| = \sum_i |f_i\rangle\langle f_i| \quad . \tag{51} \]

Using these one finds immediately the linear transformation that relates the two sets of basis vectors

\[ |e_i\rangle = \sum_j |f_j\rangle\langle f_j|e_i\rangle \quad . \tag{52} \]

A vector \( |\psi\rangle \) can be expanded as

\[ |\psi\rangle = \sum_i |e_i\rangle\langle e_i|\psi\rangle = \sum_i |f_i\rangle\langle f_i|\psi\rangle \tag{53} \]

and its components are related as

\[ \langle e_i|\psi\rangle = \sum_j \langle e_i|f_j\rangle\langle f_j|\psi\rangle \quad , \tag{54} \]

again using the resolution of the identity. Defining the matrix \( S \) with components \( S_{ij} = \langle e_i|f_j\rangle \), this can be written in matrix notation as \( \psi^{(e)} = S\psi^{(f)} \), where \( \psi^{(e/f)} \) is the vector with elements \( \langle e_i|f_j\rangle\psi \). For the components of an operator \( \hat{T} \) one finds

\[ \langle e_i|\hat{T}|e_j\rangle = \sum_{k,m} \langle e_i|f_k\rangle\langle f_k|\hat{T}|f_m\rangle\langle f_m|e_j\rangle \quad . \tag{55} \]

In matrix notation this becomes

\[ T^{(e)} = ST^{(f)}S^{-1} \quad , \tag{56} \]

where \( S^{-1} \) is the matrix with components \( S_{ij}^{-1} = \langle f_i|e_j\rangle \). Using the resolution of the identity one verifies that \( S^{-1}S = 1 \), hence \( S^{-1} \) is the inverse of the matrix \( S \).

The transformation (56) is called a similarity transformation. Diagonalizing the matrix \( T^{(f)} \) by finding a similarity transformation \( S \) such that \( T^{(e)} \) is diagonal is equivalent to transforming to a basis consisting of the eigenvectors of \( \hat{T} \). (We assume that \( \hat{T} \) corresponds to a physical observable so that a basis of eigenvectors exists.)

XII. THE SCHRÖDINGER EQUATION

The time dependent Schrödinger equation in Dirac notation reads

\[ \hat{H}|\Psi\rangle = i\hbar \frac{\partial}{\partial t} |\Psi\rangle \quad , \tag{57} \]
where $\hat{H}$ is the Hamiltonian describing the system, and $|\Psi(t)\rangle$ is the state vector that contains all information about the system. Written in this abstract vector form it applies to any physical system: a particle moving along the $x$-axis, a particle moving in three dimensions, a spin, several particles, a molecule, a solid, a superconductor...... you name it!

Just as we did before we can derive a time-independent equation by separating variables (we assume that $\hat{H}$ has no explicit $t$ dependence). Making the ansatz

$$|\Psi(t)\rangle = e^{-iEt/\hbar}|\psi\rangle,$$

where $|\psi\rangle$ is $t$-independent, we find the time-independent Schrödinger equation

$$\hat{H}|\psi\rangle = E|\psi\rangle.$$

The above is the Schrödinger equation in vector form, we can obtain it in component form by inserting the resolution of the identity in suitable places. In the discrete case, with $\sum_i |i\rangle\langle i| = 1$, we obtain

$$\sum_{i,j} |i\rangle\langle i|\hat{H}|j\rangle\langle j|\psi\rangle = E \sum_i |i\rangle\langle i|\psi\rangle,$$

which gives the component equation

$$\sum_j H_{ij}\psi_j = E\psi_i,$$

where we have defined, in the standard way,

$$H_{ij} = \langle i|\hat{H}|j\rangle, \quad \psi_i = \langle i|\psi\rangle.$$

The Schrödinger equation for, for example, a spin is of this form.

In the continuum case, with $\int dx |x\rangle\langle x| = 1$, we obtain instead

$$\int dx \int dy |x\rangle\langle x|\hat{H}|y\rangle\langle y|\psi\rangle = E \int dx |x\rangle\langle x|\psi\rangle,$$

which gives the component equation

$$\int dy H(x,y)\psi(y) = E\psi(x),$$

where

$$H(x,y) = \langle x|\hat{H}|y\rangle, \quad \psi(x) = \langle x|\psi\rangle.$$
The Schrödinger equation \[ -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + V(x) \psi(x) = E \psi(x) \] that describes the motion of a particle along the \( x \)-axis is of the form (64), \textit{i.e.}, it is given in component form. We see immediately that the right hand sides of the two equations agree. The left hand sides will also agree if we make the identification \( H(x, y) = \delta(x - y) \left[ -\frac{\hbar^2}{2m} \frac{d^2}{dy^2} + V(y) \right] \).

The variable \( x \) in the continuum case need not be the position of a particle moving along the \( x \)-axis, but can be any continuous observable, such as the momentum. The formulae can also easily be generalized to the case of several continuous variables, which is needed for, for example, one particles motion in three dimensional space or several particles or...... All one has to do is to replace \( x \) above by a set of variables and the integral \( \int dx \) by a multiple integral. Formally, everything will be the same.

**XIII. THE OPERATORS \( \hat{x} \) AND \( \hat{p} \)**

Here we will discuss the position, \( \hat{x} \), and momentum, \( \hat{p} \), operators for a particles motion. Consider the eigenvalue equations

\[ \hat{x} |x\rangle = x |x\rangle \quad (66) \]
\[ \hat{p} |p\rangle = p |p\rangle \quad (67) \]

Here, \( \hat{x}, \hat{p} \) are operators, \( x, p \) are eigenvalues and \( |x\rangle, |p\rangle \) are eigenvectors. There is a unique eigenvector \( |x\rangle \) \((|p\rangle) \) for each real \( x \) \((p) \) and they are complete:

\[ \int_{-\infty}^{\infty} dx |x\rangle \langle x| = 1 \quad (68) \]
\[ \int_{-\infty}^{\infty} dp |p\rangle \langle p| = 1 \quad (69) \]

This is always true for operators in quantum mechanics that correspond to physical observables. Sometimes it can be proven, but in general, we simply assume this to be true, without proof, for any quantum mechanical operator. The completeness is intimately tied to the foundations of quantum mechanics as we will see when discussing the statistical interpretation below.

A general state (describing a particles motion in one dimension) can be expanded in either the \( |x\rangle \) or the \( |p\rangle \) basis:

\[ |\psi\rangle = \int_{-\infty}^{\infty} dx |x\rangle \langle x| \psi\rangle = \int_{-\infty}^{\infty} dp |p\rangle \langle p| \psi\rangle \quad . \quad (70) \]
This gives two wave functions that describe the system: \( \psi(x) = \langle x|\psi \rangle \) and \( \psi(p) = \langle p|\psi \rangle \). These are just components of the same vector \( |\psi \rangle \) in different bases, \textit{i.e.} they are related by a change of basis, see the section above on this topic. \( \psi(x) \) is the wave function in position space, this is the usual wave function which enters the usual Schrödinger equation and which gives the probability of finding the particle at position \( x \). \( \psi(p) \) on the other hand, is the wave function in momentum space and gives the probability to find the particle with momentum \( p \). We will see below that \( \psi(x) \) and \( \psi(p) \) are related by a Fourier transformation.

Let us determine the components of \( |x\rangle \) and \( |p\rangle \) in the \( x \)-basis. For \( |x\rangle \), we have

\[
|y\rangle = \int dx \, |x\rangle \langle x|y \rangle = \int dx \, |x\rangle \delta(x - y) \tag{71}
\]

hence,

\[
\psi_y(x) \equiv \langle x|y \rangle = \delta(x - y) \tag{72}
\]

Thus, the wave function in \( x \)-space for the state \( |x\rangle \) is Dirac’s delta-function. This makes sense, since, according to the probability interpretation of the wave function, it means that the particle can only be found at \( x = y \). The operator \( \hat{x} \) becomes in \( x \)-basis

\[
\hat{x} = \int dx \int dy \, |x\rangle \langle x|\hat{x}|y \rangle \langle y| = \int dx \int dy \, |x\rangle x \delta(x - y) \langle y| \tag{73}
\]

where we have used (66), (72) and the properties of the \( \delta \)-function. Notice the similarity with the finite dimensional expression eq.(35), and that in its own eigenbasis, \( \hat{x} \) is represented by a diagonal ’matrix’ \( x\delta(x - y) \) just as in eq.(36).

We now turn to the components of \( |p\rangle = \int dx \, |x\rangle \langle x|p \rangle \). \( \psi_p(x) \equiv \langle x|p \rangle \) is the wave function for the \( |p\rangle \) state. To proceed we need to know what \( \hat{p} \) is. We defined this operator in the \( x \)-basis as

\[
\hat{p}_x \psi(x) = -i\hbar \frac{\partial}{\partial x} \psi(x) \tag{74}
\]

(Here, we use \( \hat{p} \) for the abstract operator, \textit{e.g.} the one in (67), whereas \( \hat{p}_x \) denotes the representation of this operator in the \( x \)-basis (\( \hat{p}_x = -i\hbar \partial / \partial x \)). Usually, we use \( \hat{p} \) also for the latter operator.) Let us recall the solution of the eigenvalue equation for \( \hat{p} \) in the
$x$-representation:

\[ \hat{p}_x \psi_p(x) = p \psi_p(x) \]  
\[ -i\hbar \frac{\partial}{\partial x} \psi_p(x) = p \psi_p(x) \]  
\[ \psi_p(x) = A e^{i px/\hbar}, \quad p \text{ real} . \]  

Thus we have

\[ |p\rangle = \frac{1}{\sqrt{2\pi\hbar}} \int dx \, |x\rangle e^{i px/\hbar}, \quad \psi_p(x) \equiv \langle x|p\rangle = \frac{1}{\sqrt{2\pi\hbar}} e^{i px/\hbar} , \]  

where we have chosen $A = 1/\sqrt{2\pi\hbar}$. We can now evaluate the scalar product of $|p\rangle$ states:

\[ \langle p'|p\rangle = \frac{1}{2\pi\hbar} \int dx \int dx' \langle x'|x\rangle e^{i(px' - px)/\hbar} \]  
\[ = \frac{1}{2\pi\hbar} \int dx \, e^{ix(p-p')/\hbar} \]  
\[ = \delta(p'-p) . \]  

In the last step we used the useful relation (see Problem 2.26 in Griffiths)

\[ \frac{1}{2\pi} \int_{-\infty}^{\infty} dk \, e^{ikx} = \delta(x) . \]  

This is proven as follows. Let $F(k)$ be the Fourier transform of $f(x)$:

\[ F(k) = \frac{1}{\sqrt{2\pi}} \int dx \, f(x) e^{ikx} \]  
\[ f(x) = \frac{1}{\sqrt{2\pi}} \int dk \, F(k) e^{-ikx} . \]  

One then finds

\[ \int dx \, f(x) \left[ \frac{1}{2\pi} \int dk \, e^{ikx} \right] = \frac{1}{\sqrt{2\pi}} \int dk \, F(k) = f(0) , \]  
from which (80) follows.

We summarize some of the important results of this section:

\[ |y\rangle = \int dx \, \delta (y-x)x\rangle \]  
\[ \langle x|y\rangle = \delta (x-y) \]  
\[ |p\rangle = \frac{1}{\sqrt{2\pi\hbar}} \int dx \, e^{i px/\hbar} |x\rangle \]  
\[ \langle x|p\rangle = \frac{1}{\sqrt{2\pi\hbar}} e^{i px/\hbar} \]  
\[ \langle p'|p\rangle = \delta(p'-p) . \]  

Comments
• $|p\rangle$ (as well as $|x\rangle$) are non-normalizable but note that they are orthogonal and almost normalized in the sense that $\delta_{ij} \to \delta(x-y)$: $\langle x|y \rangle = \delta(x-y)$ and $\langle p'|p \rangle = \delta(p'-p)$ (Griffiths calls this ‘delta-function normalized’), and this is good enough, see the notes on the free particle.

• The wave functions in $x$ and $p$ space $\psi(x) = \langle x|\psi \rangle$ and $\psi(p) = \langle p|\psi \rangle$ are related by a Fourier transform:

$$\langle x|\psi \rangle = \int dp \langle x|p \rangle \langle p|\psi \rangle = \frac{1}{\sqrt{2\pi\hbar}} \int dp \, e^{ipx/\hbar} \langle p|\psi \rangle . \quad (89)$$

We end this section by proving that the operators $\hat{x}$ and $\hat{p}$ are hermitian. For $\hat{x}$ we have:

$$\langle \phi|\hat{x}|\psi \rangle = \int dx \int dy \langle \phi|x \rangle \langle x|\hat{x}|y \rangle \langle y|\psi \rangle = \int dx \, \phi^*(x)x\psi(x) \quad (90)$$

$$\langle \phi|\hat{x}^\dagger|\psi \rangle \equiv \langle \psi|\hat{x}|\phi \rangle^* = (\int dx \, \psi^*x\phi)^* = \int dx \, \psi^*x\phi^* = \langle \phi|\hat{x}|\psi \rangle , \quad (91)$$

hence $\hat{x}$ is hermitian, see (22).

For $\hat{p}$, we find

$$\langle \phi|\hat{p}|\psi \rangle = \int dx \, \phi^*(x)(-i\hbar \frac{d}{dx})\psi(x) \quad (92)$$

$$\langle \phi|\hat{p}^\dagger|\psi \rangle \equiv \langle \psi|\hat{p}|\phi \rangle^* = (\int dx \, \psi^*(x)(-i\hbar \frac{d}{dx})\phi(x))^*$$

$$= (-i\hbar \psi^*|\phi|_\infty^* + (\int dx \, i\hbar \frac{d\psi^*}{dx}\phi)^*$$

$$= \int dx \, \phi^*(x)(-i\hbar \frac{d}{dx})\psi(x) + (\ldots)|_\infty$$

$$= \langle \phi|\hat{p}|\psi \rangle \quad , \quad (93)$$

provided the wave functions go to zero fast enough at infinity so that the surface term vanishes (this is the case for normalizable wave functions). This shows that $\hat{p}$ is hermitian (22).

**XIV. THE HILBERT SPACE**

The set of possible vectors $|\psi\rangle$ describing the quantum mechanical state of a particular system form a Hilbert space. Different systems have different Hilbert spaces. We here discuss briefly the properties of such a space.
A space is said to be *complete* if all convergent series of elements in the space also belong to the space. (Note that this has nothing to do with the completeness of the eigenvectors to an operator discussed above.)

A *Hilbert space* $\mathcal{H}$ is a vector space, with a scalar product, that is complete:

$$|\alpha\rangle = \sum_{j=1}^{\infty} a_j |\alpha_j\rangle$$

is an element in $\mathcal{H}$ provided $|\alpha_i\rangle$ are elements in $\mathcal{H}$ and the series is convergent

$$\langle \alpha | \alpha \rangle = \sum_{j=1}^{\infty} |a_j|^2 < \infty \quad .$$

For a particle moving along the $x$-axis, the Hilbert space is $L_2(-\infty, \infty)$, this is the space of all functions $\psi(x)$, such that

$$\int_{-\infty}^{\infty} dx |\psi(x)|^2 < \infty \quad .$$

These functions are called square-integrable. (Note that $\psi(x)$ are the components of the vector and that the Hilbert space here is defined in terms of the components.)

For one spin 1/2, the Hilbert space is a two dimensional vector space, for spin one 1 it is a three dimensional vector space, for $N$ spin 1/2 it is a $2^N$ dimensional vector space etc.

**XV. THE STATISTICAL INTERPRETATION**

Griffiths discussion of the statistical interpretation in Section 3.4 is restricted to the motion of one particle in one dimension. Here we consider the general case.

1. The *system* is whatever we describe by quantum mechanics. It can be a particle moving in one dimension, in three dimensions, several particles, a spin, particles with spin, a molecule, a solid etc.

2. The system is completely described by a vector $|\psi\rangle$ in a Hilbert space $\mathcal{H}$. Each vector in $\mathcal{H}$ describes a possible state for the system.

   - Note that the principle of superposition is implied by this.
   - A vector in $\mathcal{H}$ is called a state vector since it describes a quantum state.
3. Each observable (i.e. each measurable quantity) $Q$ corresponds to a hermitian operator $\hat{Q}$ acting on the vectors in $\mathcal{H}$.

- For the particle: $Q(x, p, t) \rightarrow \hat{Q}(\hat{x}, \hat{p}, t)$.

4. The expectation value of $Q$ is:

$$\langle Q \rangle = \frac{\langle \psi | \hat{Q} | \psi \rangle}{\langle \psi | \psi \rangle} \quad (97)$$

$$= \langle \psi | \hat{Q} | \psi \rangle, \quad \text{if} \quad \langle \psi | \psi \rangle = 1 \quad . \quad (98)$$

- Repeated measurements on identically prepared systems (i.e. each one of them in state $|\psi\rangle$) gives this mean value.

Let $\hat{Q}$ have eigenvectors $|e_i\rangle$ with eigenvalues $\lambda_i$ (We assume here that the spectrum $\{\lambda_i\}$ is discrete, for the generalization, see below.):

$$\hat{Q}|e_i\rangle = \lambda_i|e_i\rangle, \quad i = 1, 2, ... \quad (99)$$

- $\lambda_i$ are real (since $\hat{Q}$ is hermitian).

- $\{|e_i\rangle\}$ is a basis: $\langle e_i | e_j \rangle = \delta_{ij}, \quad \sum_i |e_i\rangle \langle i| = 1$

5. A measurement of the observable $Q$ gives one of the eigenvalues $\lambda_i$.

If $|\psi\rangle = \sum_i |e_i\rangle \langle e_i | \psi \rangle$ then the probability, $P_i$ to get $\lambda_i$ as a result of the measurement is

$$P_i = \frac{|\langle e_i | \psi \rangle|^2}{\langle \psi | \psi \rangle} \quad (100)$$

$$= |\langle e_i | \psi \rangle|^2 \quad \text{if} \quad \langle \psi | \psi \rangle = 1. \quad (101)$$

- Compare this to the statistical interpretation of $\psi(x) = \langle x | \psi \rangle$.

- The probability for the system to be found in the general (normalized) state $|\chi\rangle$ is obtained by replacing $|e_i\rangle$ by $|\chi\rangle$ above.
6. Immediately after the measurement of the observable $Q$ that resulted in $\lambda_i$, the state of the system is $|\psi_{\text{after}}\rangle = |e_i\rangle$.

- This is the collapse of the wave function (or rather the state vector) $|\psi\rangle$.
- Repeated measurement of $Q$ immediately after the first measurement gives the same result $\lambda_i$ with probability one.

**Continuous spectrum**

For a general hermitian operator $\hat{Q}$ with a continuous spectrum one is, in general, not guaranteed that the eigenvectors span the space. From the above we infer that this property is of fundamental importance in quantum mechanics—it is closely tied to the statistical interpretation of the state vector. One assumes that all hermitian operators that correspond to physical observables have the property that their eigenvectors span the Hilbert space. In some cases, this can be proven, in other cases one simply takes this as an assumption.

In the discussion above we assumed a discrete spectrum, but 1-6 hold also when the spectrum is continuous, provided one makes the standard replacements: $\sum_i \rightarrow \int dx$ and $\delta_{ij} \rightarrow \delta(x - y)$.[8]

---

[1] For the mathematically interested, more advanced texts on quantum mechanics will contain some discussion of these points.

[2] Dirac was one of the most famous physicists of the twentieth century. His perhaps most important contribution was the Dirac equation which combines quantum mechanics and special relativity to give a very accurate description of electrons. His 1930 textbook, *The Principles of Quantum Mechanics*, is still recommended reading for the serious student of quantum mechanics.

[3] The vector space contains a *null vector*, $0$, *i.e.* a vector with the property that $|\alpha\rangle + 0 = |\alpha\rangle$ for any $|\alpha\rangle$. Obviously, this is the vector with all components 0, I will denote this vector simply by 0. Griffiths uses the notation $|0\rangle$ for this 0-vector, see [A.4]—this is very bad notation. Normally, in physics texts $|0\rangle$ is used to mean a vector that is not the 0-vector. Avoid using $|0\rangle$ for the 0-vector!
Griffiths sometimes puts the vector equal to its components. This is OK if you know what it means but I will try to avoid this.

Another important type of operator in quantum mechanics are operators that obey $\hat{U}\hat{U}^\dagger = 1$. This means that the hermitian conjugate operator is equal to the inverse operator $\hat{U}^\dagger = \hat{U}^{-1}$, where the inverse is defined as the operator that obeys $\hat{U}\hat{U}^{-1} = 1$, where 1 is the identity operator. In matrix language this becomes $UU^\dagger = 1$.

Sometimes, like in the infinite square well, we are interested in wave functions defined on a finite segment of the line, then the correspondence becomes $\sum_{i=1}^{n} \rightarrow \int_{a}^{b} dx$ and to define the theory one must supply proper boundary conditions.

It can also easily be generalized to the case of several continuous variables, which is necessary, eg for describing a particle in three dimensional space or several particles. All one has to do is to replace $x$ by a set of variables $\{x_1, x_2, \ldots\}$ and $\int dx$ by a multiple integral over these variables $\int dx_1 \int dx_2 \ldots$.

There are also operators where the spectrum has both a discrete part and a continuous part, the energy for a the hydrogen atom is an example, this can also be accommodated in the formalism here with minor changes.