$$i\hbar\frac{\partial}{\partial t}|\psi(t)\rangle = \hat{H}|\psi(t)\rangle$$

University of Oslo

Quantum Mechanics

A Compendium

Gabriel S. Cabrera

Intended for
FYS3110

October 20, 2018

Contents

CHAPTER 1
Formalism

Introduction

The purpose of this chapter is to give the reader a thorough understanding of the notation used in modern quantum mechanics. Since this compendium is intended for those taking a second-level course, it is important that one be familiar with this written representation of our microscopic world otherwise the rest of the text will be incomprehensible[1].

$$\int_{-\infty}^{\infty} \Psi_i^* \Psi_j \, dx \qquad \langle \Psi_i | \Psi_j \rangle$$

(a) Old Notation (b) Dirac Notation

Figure 1.1: A side-by-side comparison showcasing the compactness of Dirac Notation.

In many first-level quantum mechanics courses, students are exposed to semi-familiar notation containing unintuitive and mystical elements – eventually the concepts are understood, and students may notice that the old notation starts appearing clunky and messy. To remedy this, this chapter will be all about learning how to use the more streamlined **Dirac Notation**.

$$\langle \ | \qquad | \ \rangle$$

(a) A *bra* (b) A *ket*

Figure 1.2: The two components used in *braket* notation.

Back in the early days, two different types of notation were used; a form of vector notation known as the *matrix formulation of quantum mechanics* was developed by Werner Heisenberg. His counterpart Erwin Schrödinger approached the same problems using a different kind of notation, known as *wavefunction formalism*, which uses complex functions in the form of *kets*. In the end however it was Paul Dirac who unified both systems with the abstract notation and formalism of brakets.

1.1 Quantum States

Introduction

Before discussing *quantum states*, we should discuss the general definition of a *classical state* for any arbitrary system:

[1] The phrase "it's all Greek to me" comes to mind, though the Greek alphabet shouldn't be so foreign at this point. As the Greeks would say: "*Αυτά μου φαίνονται αλαμπουρνὲζικα.*"

A *state* is defined as the complete description/specification of a system/object at a given time.

Being a subcategory of the above, quantum states can be described more precisely:

A quantum state *exclusively* contains *all* the possible probabilities of *all* different outcomes in all different measurements.

In terms of notation, a quantum state is a complex vector represented by a ket.

1.1.1 Measurements

It is very important at this point not to confuse quantum states with measurements; the main distinction is that states are complex vectors, or *kets*, while measurements are eigenvalues/numbers.

So, what is a measurement? According to the Copenhagen interpretation of quantum mechanics:

> "During an observation, the system must interact with a laboratory device" [1], so a measurement is defined as an interaction with a macroscopic object.

Once we make a measurement, we go from a system of probabilities to a system with some certainty[2]. This process all takes place at a fixed instance in time – over time however the state changes smoothly in such a way governed by the system's *Hamiltonian*.

1.1.2 Complex Vectors aka Kets

An interesting property of quantum states is our ability to add them to others; such an addition is the equivalent of a *change of probabilities*. Let us take a further look at the algebra underlying these systems of *complex vectors*, or *kets*.

Complex numbers are composed of a *real* and *imaginary* part of the following form:

$$z = x + iy, \quad z \in \mathbb{C},\ x, y \in \mathbb{R} \tag{1.1}$$

In polar coordinates, we can rewrite (1.1) as follows:

$$z = re^{i\theta}, \quad z \in \mathbb{C},\ r = \sqrt{x^2 + y^2},\ \theta = \arctan\left(\frac{y}{x}\right) \tag{1.2}$$

In (1.2), we have that r is the number's *amplitude* and θ is its *phase*. We need to understand these concepts in order to work with complex vectors, as they are very much like the real vectors in a lot of ways, but differ in terms of how they operate with each other. First of all, we have something called a *vector identifier* ($|\quad\rangle$) which tells us that a vector is complex; anything placed within the parentheses, such as Q in $|Q\rangle$, is known as its *name*. This entire structure, as mentioned previously, is called a *ket*.

[2] This type of occurrence is an abrupt, non-deterministic change.

In a complex vector space, we need to be able to make sense of the multiplication of complex numbers, since the normal physical interpretation of real vectors does not apply here in the normal sense. The properties of complex vector spaces are listed below[3]:

1. $\alpha\,|u\rangle = |u'\rangle = |u\rangle\,\alpha$ where $\alpha, u' \in \mathbb{C}$
2. $|u_1\rangle + |u_2\rangle = |u_3\rangle$
3. There exists a null vector
4. The associative property applies to kets such that: $(|u_1\rangle + |u_2\rangle) + |u_3\rangle = |u_1\rangle + (|u_2\rangle + |u_3\rangle)$

Other than this, the same mathematical principles used in real vectors can often be applied to complex vectors – with some scrutiny and care, of course. This is all very abstract for the moment, so let's take a look at an example using the *matrix formulation of quantum mechanics*:

Example 1.1.1: Complex Vectors using the Matrix Formulation

We can use the principles listed above to manipulate concrete objects; in this case, we will be performing operations on sets of column matrices, which are essentially column-vectors with complex elements:

$$|u\rangle = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix},\ u_1, u_2, u_3 \in \mathbb{C} \tag{1.3}$$

Let us take our ket $|a\rangle$ and apply some of the aforementioned properties of a complex vector space first via multiplication, then addition:

1. Firstly, we see that multiplying a ket by a complex number simply leaves us with another ket:
$$\alpha\,|u\rangle = \alpha \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} \alpha u_1 \\ \alpha u_2 \\ \alpha u_3 \end{bmatrix} = |u'\rangle \tag{1.4}$$
2. Secondly, we see that the same applies to addition; a ket plus a ket equals a ket:
$$|u\rangle + |v\rangle = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} + \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} u_1 + v_1 \\ u_2 + v_2 \\ u_3 + v_3 \end{bmatrix} = |v'\rangle \tag{1.5}$$

The above examples don't only work for 3-vectors, these processes can be applied to vectors of *any* length and are identical regardless.

Now that we've seen an example using Heisenberg's notation, let us take a look at an example using Schrödinger's notation based on *functions* rather than *matrices*:

[3] In this case, we are now generalizing the numbers we multiply with our complex vectors as also being complex.

Example 1.1.2: Complex Functions of 1-Variable using Wavefunction Formalism

We are given a function $f(x)$ such that $x \in \mathbb{R}$ and $f(x) \in \mathbb{C}$. But how can a *real* input give us a *complex* output? To understand this, we can decompose our function into its *real* and *imaginary* parts:

$$f(x) = f_{\Re}(x) + i f_{\Im}(x) \tag{1.6}$$

In (1.6), we have three components: the two real-valued functions $f_{\Re}(x)$ and $f_{\Im}(x)$, and the imaginary number i. As we clearly see, the range of our function will therefore be complex!
Let's now perform the same operations as in Example 1.1.2: scalar multiplication and addition.

1. First of all, we can clearly see that our function is a ket such that $|f\rangle = f(x)$ since multiplying it with a complex coefficient returns another ket:

$$\begin{aligned} \alpha f(x) &= (\alpha_{\Re} + i\alpha_{\Im})\,(f_{\Re}(x) + f_{\Im}(x)) \\ &= \alpha_{\Re} f_{\Re}(x) - \alpha_{\Im} f_{\Im}(x) + i\,(\alpha_{\Im} f_{\Re}(x) + \alpha_{\Re} f_{\Im}(x)) \end{aligned} \tag{1.7}$$

2. Secondly, the same applies to addition once more as we see in the following example:

$$f_1(x) + f_2(x) = \underbrace{f_{1,\Re}(x) + f_{2,\Re}(x)}_{f_{3,\Re}(x)} + i\underbrace{(f_{1,\Im}(x) + f_{2,\Im}(x)}_{f_{3,\Im}(x)} = f_3(x) \tag{1.8}$$

1.2 The Inner Product

Introduction

As with the linear algebra of real-valued vectors, we can evaluate the *inner product* of two kets in a way distinct from but reminiscent of the dot product – however, we must take care not to ignore the effects of working with complex values. The inner product space in this case has a specific name: *Hilbert space*. Notation wise, we will begin by representing the inner product of two kets $|u\rangle$ and $|v\rangle$ as follows:

$$(|u\rangle, |v\rangle) \tag{1.9}$$

As in the previous operations, the inner product of two kets will return a complex value; this requires a more complex process than the dot product. We can make sure we don't overlook such complexities via the following axioms:

1. The inner product is *not* commutative – switching the kets is insufficient, you must also conjugate the result:

$$\underbrace{(|u\rangle, |v\rangle)}_{\in\mathbb{C}} = (|v\rangle, |u\rangle)^* \tag{1.10}$$

2. The inner product is linear exclusively in its *second* argument:

$$(|u\rangle, \alpha|v_1\rangle + \beta|v_2\rangle) = \alpha(|u\rangle, |v_1\rangle) + \beta(|u\rangle, |v_2\rangle) \tag{1.11}$$

3. The inner product of a ket with itself results in a product that is both *real* and ≥ 0:

$$(|u\rangle, |u\rangle) \in \mathbb{R} \geq 0 \tag{1.12}$$

Note that the *only* case where our product equals zero is when our ket itself equals zero.

1.2.1 Dirac Notation

We've already introduced the basics of Dirac notation earlier in Figures 1.1 and 1.2, though its usefulness thus far has not yet been made clear, so we will investigate how it is used in practice. One of the most common uses of Dirac notation is in the representation of inner products, where it is defined as follows:

$$\langle u|v\rangle = (|u\rangle, |v\rangle) \tag{1.13}$$

The left hand side of (1.13) is called a braket: a mathematical object composed of a *bra* and a *ket*. It is acceptable (but generally unnecessary) to separate these components as follows:

$$\langle u|v\rangle = \langle u|(|v\rangle) \tag{1.14}$$

As seen above, a major advantage of this notation is the fact that we can split the inner product into two workable components. In addition to this, a braket always represents a complex number – this means that we can think of a *bra* (such as $\langle u|$) as something that acts upon a *ket* (such as $|v\rangle$) to give an output in the form of a complex number: the complete braket.

Although bras do technically operate on kets, they are *not* considered operators. This will be explained later in more detail, but for the moment we will simply clarify that an operator should always take a ket as input, and return a *ket* as an output[4]; since a bra returns a *complex number* as output, it is not an operator!

Example 1.2.1: Axioms with Brakets

We've previously discussed the axioms that make up our Hilbert space in 1.2, but not in Dirac notation! Here is the equivalent using brakets:

1.
$$\langle u|v\rangle = \langle v|u\rangle^* \tag{1.15}$$

2. Let $|v'\rangle = \alpha|v_1\rangle + \beta|v_2\rangle$ – we then have that:
$$\langle u|v'\rangle = \alpha\langle u|v_1\rangle + \beta\langle u|v_2\rangle \tag{1.16}$$

3.
$$\langle u|u\rangle \in \mathbb{R} \geq 0,\ \langle u|u\rangle = 0 \text{ iff } |u\rangle = 0 \tag{1.17}$$

[4]Recall that operators are used to find eigenvalues, so they must therefore return an eigenvector-eigenvalue product.

We can use brakets to further simplify our notation by removing the need to use parentheses for coefficients. All three of the following are therefore equivalent:

This: $\langle u| (\alpha |v\rangle)$

Can be written as this: $\langle u|\alpha|v\rangle$

Or this: $\alpha \langle u|v\rangle$

1.2.2 Bras vs. Kets

So what is the difference between a bra and a ket? The answer is subtle: the identity of the bra is based in the complex conjugation of the ket. They are related by the following:

$$|v'\rangle = \alpha |v_1\rangle + \beta |v_2\rangle$$
$$\langle v'| = \alpha^* \langle v_1| + \beta^* \langle v_2|$$

Let's investigate further by once again by investigating different aspects of these objects through the lenses of Schrödinger and Heisenberg's notations. We will begin by using Schrödinger's interpretation to visualize these mathematical objects:

Example 1.2.2: Bras and Kets in the Matrix Formulation

At this point, we've already discussed kets in great detail, so the following representation of an arbitrary ket should be of no surprise[a]:

$$|u\rangle = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}, \; u_1, u_2, u_3 \in \mathbb{C} \tag{1.18}$$

To understand how bras and kets relate to each other, it is best to represent the corresponding[b] bra $\langle u|$ for our aforementioned ket $|u\rangle$ in the form of a matrix:

$$\langle u| = \begin{bmatrix} u_1^* & u_2^* & u_3^* \end{bmatrix} \tag{1.19}$$

We've already mentioned that brakets are complex numbers, so let us see why this is the case. Assume we now have an arbitrary ket $|v\rangle$, what would happen if we used matrix multiplication to compute the complete braket $\langle u|v\rangle$? Let's take a look:

$$\langle u|v\rangle = \begin{bmatrix} u_1^* & u_2^* & u_3^* \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = u_1^* v_1 + u_2^* v_2 + u_3^* v_3 \in \mathbb{C} \tag{1.20}$$

So as we see, brakets are complex numbers, just as expected. In conclusion: we see that kets can be represented as column vectors, while corresponding bras are transposed conjugated versions of their respective kets in the form of row vectors.

[a] Check out Exercise 1.1.2 if you are unfamiliar with the material presented.
[b] In mathematical terms, a $\langle u|$ is called the *adjoint* of a complex vector $|u\rangle$.

We've now described the finite vector representation of ket, but have neglected to investigate how it is representable in the form of a function; to begin, let us create two kets $|u\rangle$ and $|v\rangle$. These

kets are no longer simple vectors, but the continuous functions $U(x)$ and $V(x)$, respectively, meaning we need to use integration to evaluate the braket $\langle u|v\rangle$. First of all, we can represent a single bra as follows[5]:

$$\langle u| = \int dx\, U^*(x) \tag{1.21}$$

We can then complete the braket as such:

$$\langle u|v\rangle = \int dx\, U^*(x)V(x) \tag{1.22}$$

In addition to this, we can also infer the following:

$$\langle v|u\rangle = \int dx\, V^*(x)U(x) = \left(\int dx\, U^*(x)V(x)\right)^* = \langle u|v\rangle^* \tag{1.23}$$

With all these tools at our disposal, brakets can be used in a multitude of interesting and unique ways.

1.3 The Outer Product

This section will be brief, as its purpose is to explain what occurs when a ket is multiplied by a bra, such as:

$$|a\rangle\langle b| \tag{1.24}$$

What the above represents is the **outer product** of two vectors, which can be visualized using a vector representation:

$$\begin{pmatrix} a_1 & a_2 & a_3 \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} a_1b_1 & a_2b_1 & a_3b_1 \\ a_1b_2 & a_2b_2 & a_3b_2 \\ a_1b_3 & a_2b_3 & a_3b_3 \end{pmatrix} \tag{1.25}$$

1.4 The Dirac Delta Function

The Dirac delta function can be intuitively understood as follows:

$$\delta(x) = \begin{cases} \infty, & x = 0 \\ 0, & x \neq 0 \end{cases} \tag{1.26}$$

It is also imperative it follows the following as well:

[5] We've placed the dx on the left-hand side of our function, but we must still integrate over it! This is simply to keep proper track of our bras and kets.

$$\int_{-\infty}^{\infty} \delta(x)\, dx = 1. \tag{1.27}$$

Though it is called a function, it is technically more correctly defined as being the "limit of a distribution." In addition to this, it is mostly meaningless to use it outside of an integral due to its infinite nature.

There are plenty of useful ways to use the Dirac delta, here is an example of a useful formula:

Example 1.4.1: Dirac Delta Trickery

Here is a fun example of how the Dirac delta function can be used:

$$f(x) = \int dx'\, f(x')\delta\left(x - x'\right) \tag{1.28}$$

To show that this is true, we can use the following:

$$\delta\left(x - x'\right) = 0, \text{ for } x \neq x' \tag{1.29}$$

Using the above principle, we can show that (1.28) is valid:

$$\int_{x-\epsilon}^{x+\epsilon} dx'\, f(x')\delta\left(x - x'\right) = f(x) \underbrace{\int_{x-\epsilon}^{x+\epsilon} dx'\, \delta\left(x - x'\right)}_{1} = f(x) \tag{1.30}$$

By solving the above integral for $x = x'$, we can show that $\delta\left(x - x'\right)$ becomes infinite in such a case!

As mentioned previously, we cannot technically classify the Dirac delta as a function, but it is possible to use a limit combined with an integral as a proper definition:

$$\int dx'\, f(x')\delta\left(x - x'\right) = \lim_{\epsilon \to 0} \int dx\, f(x')\delta_\epsilon\left(x - x'\right) \tag{1.31}$$

Note that we are not using the Dirac delta, but rather an ordinary function $\delta_\epsilon(x - x')$ whose parameters:

1. Satisfy a scaling requirement:

$$\delta_{\frac{\epsilon}{k}}(x) = k\delta_\epsilon\left(kx\right) \tag{1.32}$$

2. Are normalized:

$$\int dx'\, \delta_\epsilon(x') = 1\ \forall\, \epsilon \tag{1.33}$$

There are multiple ways to define $\delta_\epsilon(x - x')$ (these are known as *delta sequences*) such as:

$$\delta_\epsilon\left(x - x'\right) = \frac{1}{\pi} \frac{\epsilon}{\epsilon^2 + \left(x - x'\right)^2} \tag{1.34}$$

$$\delta_\epsilon (x - x') = \frac{1}{\pi} \frac{\sin\left(\frac{x-x'}{\epsilon}\right)}{x - x'} \tag{1.35}$$

The essence of these delta sequences is that they always tend towards $\delta(x - x')$ as $\epsilon \to 0$; as a result, they are extremely useful when needing to work directly with the Dirac delta!

1.5 Sets

So far, the theory we've discussed has consisted of individual brakets, bras, and kets; now that we've gone through the basics, its time to discuss the theory of entire sets of complex vectors, i.e. sets of *kets*. Initially, one might think to represent a set of kets in the following manner:

$$|u\rangle = \{|u_1\rangle, |u_2\rangle, ..., |u_n\rangle\} \tag{1.36}$$

The above can be represented in a more elegant manner:

$$|u\rangle = |u_i\rangle \text{ for } i = 1, 2, ..., n \tag{1.37}$$

1.5.1 Discrete and Continuous Infinite Sets

For a discrete set, we can reuse the same notation used for finite sets:

$$|f\rangle = \sum_{i=1}^{\infty} f_i |i\rangle, \text{ for } f_i \in \mathbb{C} \text{ and } i = 1, 2, ... \tag{1.38}$$

If this basis is orthonormal, then we have that:

$$\langle i|j\rangle = \delta_{ij} \implies f_i = \langle i|f\rangle \tag{1.39}$$

We also want this relation to be true in the continuous case: $\langle x|f\rangle = f(x)$. We begin by setting up an integral representative of our aforementioned ket $|f\rangle$:

$$|f\rangle = \int_0^L dx' \, f(x') |x'\rangle \tag{1.40}$$

We can then complete the braket, and evaluate the product of $\langle x|$ and the above function:

$$\begin{aligned} \langle x|f\rangle = \int_0^L dx' \, f(x') \langle x|x'\rangle = f(x) \\ \implies \langle x|x'\rangle = \delta(x - x') \end{aligned} \tag{1.41}$$

It is interesting to note that the Kronecker delta and Dirac delta each play the same role – if our system is discrete, we use the Kronecker delta. If it is continuous, we use the Dirac delta.

Given an orthonormal basis,

$$\textbf{Discrete:} \quad \langle i|j\rangle = \underbrace{\delta_{ij}}_{\text{Kronecker delta}} \tag{1.42}$$

$$\textbf{Continuous:} \quad \langle x|x'\rangle = \underbrace{\delta\left(x-x'\right)}_{\text{Dirac delta}} \tag{1.43}$$

1.6 The Basis

Introduction

Before getting into the bases of Hilbert spaces, it is best to be given a refresher as to what a basis represents in a general sense. With this in mind, let $\mathbf{V}$ be a vector space, and let $\mathbf{H}$ be a subspace of $\mathbf{V}$. Let $\mathcal{B} = \{\vec{v}_1, \vec{v}_2, ..., \vec{v}_p\} \subset \mathbf{V}$.

Given the above, $\mathcal{B}$ is a **basis** for $\mathbf{H}$ if:

- $\mathcal{B}$ is a *linearly independent* set.
- $\text{Span}\{\vec{v}_1, \vec{v}_2, ..., \vec{v}_p\} = \mathbf{H}$

1.6.1 In Quantum Mechanics

The ideas presented in the introduction transfer over to Hilbert spaces with relative ease – there are a few additional interesting properties to also keep in mind. First of all, let us assume we have a set of kets $\{|i\rangle\}$ for $i = 1, 2, ..., N$. Given this information, we have that:

1. A set is linearly independent if none of its elements can be expressed as linear combinations of the others:

$$\sum_{i\neq j} \alpha_i |i\rangle \neq |j\rangle\,, \text{ for any choice of } \alpha \tag{1.44}$$

2. A set spans its vector space if any vector in the space can be expressed as a linear combination of some "$|i\rangle$"s:

$$|v\rangle = \sum_{i=1}^{N} v_i |i\rangle\,, \text{ where } \underbrace{v_i}_{\in\mathbb{C}} \in \mathbb{C} \text{ for any } |v\rangle \tag{1.45}$$

3. A linearly independent spanning set is called a *basis*
4. The dimension of a Hilbert space has to do with the difference in outcomes of measurements; the number of possible outcomes dictates the dimension of the system, so the dimension of a Hilbert space is equal to the number of basis kets.

5. If $\langle i|j\rangle = 0$ for all $i \neq j$, then the basis is *orthogonal*.
6. Furthermore, if $\langle i|i\rangle = 1$ for all $i = 1, 2, ..., N$, the basis is *orthonormal*. Mathematically, we can state that a basis is orthonormal if:

$$\langle i|j\rangle = \delta_{ij}, \text{ where } \delta_{ij} = \begin{cases} 1, & \text{if } i = j, \\ 0, & \text{if } i \neq j \end{cases} \tag{1.46}$$

In addition to the above, note that i will never equal zero since this simply represents the zero vector and is thereby *not* part of the basis.

1.6.2 Change of Basis – General

Let $\mathbf{V}$ be a finite dimensional vector space. We define two of its bases $\mathcal{B}$ and $\mathcal{C}$ as:

$$\mathcal{B} = \left\{\vec{b}_1, \vec{b}_2, ..., \vec{b}_n\right\} \tag{1.47}$$

$$\mathcal{C} = \{\vec{c}_1, \vec{c}_2, ..., \vec{c}_n\} \tag{1.48}$$

In addition, we have the vector $\vec{v} \in \mathbf{V}$ with the basis $\mathcal{B}$ coordinates $[\vec{v}]_{\mathcal{B}} = \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_n \end{bmatrix}$

How do we find the coordinates for $\vec{v}$ in terms of the basis $\mathcal{C}$? To do this, we need a transformation matrix $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$, which gives us our formula:

$$[\vec{v}]_{\mathcal{C}} = \underset{\mathcal{C}\leftarrow\mathcal{B}}{P}[\vec{v}]_{\mathcal{B}} \tag{1.49}$$

This can also be reversed:

$$[\vec{v}]_{\mathcal{B}} = \underset{\mathcal{B}\leftarrow\mathcal{C}}{P}[\vec{v}]_{\mathcal{C}} \tag{1.50}$$

In addition, the transformation matrices are each others' inverses:

$$\underset{\mathcal{B}\leftarrow\mathcal{C}}{P} = \left(\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}\right)^{-1} \tag{1.51}$$

Finding a Transformation Matrix Let $\mathbf{V} \in \mathbb{R}^n$ and $\mathbf{W} \in \mathbb{R}^m$ be vector spaces. In addition to this, we have a linear transformation $T : \mathbf{V} \to \mathbf{W}$ such that $T(\vec{x}) = A\vec{x}$, where A is an $m \times n$ matrix and $\vec{x} \in V$.

Now, let $\mathcal{B}$ and $\mathcal{C}$ be bases for $\mathbf{V}$ and $\mathbf{W}$ respectively, such that $\vec{b}_i \in \mathcal{B}$; if we want to find a transformation matrix $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ such that:

$$[\vec{v}]_{\mathcal{C}} = \underset{\mathcal{C}\leftarrow\mathcal{B}}{P}[\vec{v}]_{\mathcal{B}} \tag{1.52}$$

Then we can write it as:

$$\underset{\mathcal{C}\leftarrow\mathcal{B}}{P} = \left[\left[T\left(\vec{b_1}\right)\right]_{\mathcal{C}}\left[T\left(\vec{b_2}\right)\right]_{\mathcal{C}} \cdots \left[T\left(\vec{b_n}\right)\right]_{\mathcal{C}}\right] \tag{1.53}$$

This can be rewritten as:

$$\underset{\mathcal{C}\leftarrow\mathcal{B}}{P} = \left[\left[A\vec{b_1}\right]_{\mathcal{C}}\left[A\vec{b_2}\right]_{\mathcal{C}}\cdots\left[A\vec{b_n}\right]_{\mathcal{C}}\right] \tag{1.54}$$

To find each individual column $\left[A\vec{b_i}\right]_{\mathcal{C}}$, we must find a linear combination for the vector $A\vec{b_i} \in \mathbb{R}^m$ in terms of the basis $\mathcal{C}$:

$$A\vec{b_i} = \beta_{i,1}\vec{c}_1 + \beta_{i,2}\vec{c}_2 + \cdots + \beta_{i,m}\vec{c}_m \tag{1.55}$$

The column we are looking for will be composed of our $\beta_{i,j}$ coefficients – to find them, solve for the generated augmented matrix. We are looking for an answer of the form:

$$\vec{\beta_i} = \begin{bmatrix}\beta_{i,1}\\ \beta_{i,2}\\ \vdots \\ \beta_{i,m}\end{bmatrix} \tag{1.56}$$

Corollary Let $\mathcal{B} = \left\{\vec{b}_1, \vec{b}_2, ..., \vec{b}_n\right\}$ be a basis for the vector space $\mathbf{V}$.

We also have the transformation $\mathbf{V} \to \mathbb{R}^n$ and the coordinate transformation $\vec{v} \to [\vec{v}]_{\mathcal{B}}$. The transformation is linear, injective and surjective.

In addition, we have the sets $S = \{\vec{u}_1, \vec{u}_2, ..., \vec{u}_p\} \subset \mathbf{V}$ and $S_{\mathcal{B}} = \left\{[\vec{u}_1]_{\mathcal{B}}, [\vec{u}_2]_{\mathcal{B}}, ..., [\vec{u}_p]_{\mathcal{B}}\right\} \subset \mathbb{R}^n$.

Given all these conditions, the following equivalencies can be deduced:

- The set S spans $\mathbf{V} \iff$ the set $S_{\mathcal{B}}$ spans $\mathbb{R}^n$
- The set S is *linearly independent* $\iff$ the set $S_{\mathcal{B}}$ is *linearly independent*
- The set S is a basis for $\mathbf{V} \iff$ the set $S_{\mathcal{B}}$ is a basis for $\mathbb{R}^n \iff p = n$ and the matrix $[S_{\mathcal{B}}] = \left[[\vec{u}_1]_{\mathcal{B}}, [\vec{u}_2]_{\mathcal{B}}, ..., [\vec{u}_p]_{\mathcal{B}}\right]$ is an invertible $n \times n$ matrix $\iff p = n$ and $\det(S_{\mathcal{B}}) \neq 0$

One more equivalency If $\mathbf{S}$ is a vector space of dimension N, $S = \{\vec{u}_1, \vec{u}_2, ..., \vec{u}_n\}$ and $\mathcal{B} = \left\{\vec{b}_1, \vec{b}_2, ..., \vec{b}_n\right\}$ is a basis, then the following statements are all equivalent:

- S is a basis
- The matrix $[S_{\mathcal{B}}] = [[\vec{u}_1]_{\mathcal{B}}, [\vec{u}_2]_{\mathcal{B}}, ..., [\vec{u}_n]_{\mathcal{B}}]$ is invertible
- rref$([S_{\mathcal{B}}])$ has a pivot element in each column
- rref$([S_{\mathcal{B}}])$ has a pivot element in each row

1.6.3 Change of Basis – Dirac Notation

Say we are given two different bases $\{|n\rangle\}$ and $\{|n'\rangle\}$ – we can represent an arbitrary state ψ in terms of either basis:

$$|\psi\rangle = \sum_n \psi_n |n\rangle \text{ where } \psi_n = \langle n|\psi\rangle \tag{1.57}$$

The above is represented in the $\{|n\rangle\}$ basis, while the following equation is represented using the $\{|n'\rangle\}$ basis:

$$|\psi\rangle = \sum_{n'} \psi_{n'} |n'\rangle \text{ where } \psi_{n'} = \langle n'|\psi\rangle \tag{1.58}$$

Now, let's assume we are only give $|\psi\rangle$ in terms of our first basis (implying we know the values of all ψ_n), how can we use this information to calculate the values of each $\psi_{n'}$? First of all, we have the following:

$$\underbrace{\langle n'|\psi\rangle}_{\psi_{n'}} = \sum_n \psi_n \underbrace{\langle n'|n\rangle}_{S_{n'n}} \tag{1.59}$$

Let's rewrite the above:

$$\psi_{n'} = \sum_n S_{n'n}\psi_n \tag{1.60}$$

So in conclusion, we can use a transformation matrix S to solve for each $\psi_{n'}$, which can be found using the methods outlined in 1.6.2.

In addition to the above, one should also note that S is known as a *unitary matrix*; in short, unitary matrices abide by the following:

$$S^\dagger \equiv (S^{\mathrm{T}})^* = S^{-1} \tag{1.61}$$

Notice the small dagger † placed above S: this will be explained in further detail later, but for now we can simply say that it represents the *adjoint* of a matrix.

1.7 Operators

Introduction

As we've discussed previously, a ket can be used to represent a particular *quantum state*. We also mentioned that a quantum state is not the same as an *observable* or *measurement* – so how can we represent these observables and their *values*?

To accomplish this, we can use **operators** – these can be thought of as the representatives of observables, but not their values. The whole concept of operators is rooted in the ideas of eigenvectors and eigenvalues, and is best shown mathematically. Let's say we have an operator $\hat{K}$ and a ket $|\lambda\rangle$; the resulting product of these two will give us an eigenvalue-eigenvector pair as shown below:

$$\hat{K}|\lambda\rangle = \lambda|\lambda\rangle \tag{1.62}$$

If this were physically meaningful, $\hat{K}$ would be representing some kind of observable, while $|\lambda\rangle$ would be a quantum state, and λ would be the actual value of the observable.

1.7.1 Linear Operators

Some operators behave *linearly* when exposed to multiple kets:

$$\hat{L}\left(\alpha\left|u\right\rangle + \beta\left|v\right\rangle\right) = \alpha\hat{L}\left|u\right\rangle + \beta\hat{L}\left|v\right\rangle \tag{1.63}$$

An example of this is the *identity operator* $\hat{I}$:

Example 1.7.1: The Identity Operator

Generally, the identity of a mathematical object is something that, when operated[a] together with the object, returns the original object itself.
In quantum mechanics, the identity operator works as follows:

$$\hat{I}\left|u\right\rangle = \left|u\right\rangle \ \forall\ \left|u\right\rangle \tag{1.64}$$

And is defined by:

$$\hat{I} = \sum_{i=1}^{N} \left|i\right\rangle\left\langle i\right| \ \forall\ \left\langle i|j\right\rangle = \delta_{ij} \tag{1.65}$$

Given an arbitrary ket $\left|u\right\rangle$, we can show that this definition is correct:

$$\left|u\right\rangle = \hat{I}\left|u\right\rangle = \sum_{i=1}^{N} \left|i\right\rangle\left\langle i|u\right\rangle = \sum_{i=1}^{N} \left\langle i|u\right\rangle\left|i\right\rangle \tag{1.66}$$

We know that the above sum will evaluate to zero for each $i \neq u$, so the only remaining element will be $\left|u\right\rangle$.

[a] The type of operation is not specified, since the identity might only function for a specific type of operator. In terms of simple numbers, zero is the identity for addition, but one is the identity for multiplication.

In addition to the above, the matrix representation of a linear operator will always be a *square matrix*!

1.7.2 Representations of Operators

So far, our operators have been represented quite abstractly, so let's examine their matrix and function representations:

The Matrix Representation of Operators

As stated previously, we can represent a ket as follows:

$$\left|\lambda\right\rangle = \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{bmatrix} \tag{1.67}$$

Since the above ket contains n elements, the corresponding operator must be of the following shape:

$$\hat{L} = \begin{bmatrix} L_{11} & L_{12} & \dots & L_{1n} \\ L_{21} & L_{22} & \dots & L_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ L_{n1} & L_{n2} & \dots & L_{nn} \end{bmatrix} \tag{1.68}$$

Taking the product of $\hat{L}$ and $|\lambda\rangle$ would then yield another column vector with n elements:

$$\hat{L}\,|\lambda\rangle = \begin{bmatrix} L_{11} & L_{12} & \dots & L_{1n} \\ L_{21} & L_{22} & \dots & L_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ L_{n1} & L_{n2} & \dots & L_{nn} \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{bmatrix} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} \tag{1.69}$$

This new ket $|u\rangle$ is in fact an eigenvalue-eigenvector pair, such that:

$$\hat{L}\,|\lambda\rangle = |u\rangle = \lambda\,|\lambda\rangle \tag{1.70}$$

Where λ is the *eigenvalue* and $|\lambda\rangle$ is the *eigenvector*.

The Function Representation of Operators

Another way to represent operators is by using functions. We begin by defining the following:

$$|u\rangle = U\,(x) \tag{1.71}$$

Then, we can represent an operator $\hat{L}$ as:

$$\hat{L} = c_1 \frac{\partial}{\partial x} + c_2 \frac{\partial^2}{\partial x^2} \tag{1.72}$$

1.7.3 Getting Complex Numbers from Operators

Let's say we are given a discrete basis[6] $\{|n\rangle\}$ for $n = 1, 2, ..., N$. In addition to this, we are given an operator $\hat{K}$ in the form of a matrix. We then have that:

$$\langle m|\hat{K}|n\rangle = K_{mn} \tag{1.73}$$

In other words, a braket can be used to extract individual elements from an operator. Next, let's say we are given a ket $|v\rangle$ such that:

$$|v\rangle = \sum_{n=1}^{N} v_n\,|n\rangle \tag{1.74}$$

[6] It is important to keep in mind that a basis will be orthogonal!

Then, we can complete the braket:

$$\langle m|\hat{K}|v\rangle = \langle n|\,\hat{K}\sum_{n=1}^{N} v_n\,|n\rangle = \sum_{n=1}^{N} v_m\,\langle m|\hat{K}|n\rangle = \sum_{n=1}^{N} K_{mn}v_m \tag{1.75}$$

In the end, this results in a single complex number.

Example 1.7.2: Extracting Matrix Elements

In 1.2.2, recall that we'd concluded that kets could be represented by column vectors, while bras could be represented as row vectors. We've also just mentioned that $\langle m|\hat{K}|n\rangle = K_{mn}$ to extract an element from a matrix $\hat{K}$.

Let's now use two vectors from the standard basis for $\mathbb{R}^3$: $\vec{a} = \begin{bmatrix}1\\0\\0\end{bmatrix}$ and $\vec{b} = \begin{bmatrix}0\\1\\0\end{bmatrix}$. We can define $\vec{a}$ as being the first standard basis vector, while $\vec{b}$ is the second standard basis vector – this means that they represent indices 1 and 2 respectively. In Dirac notation, these are $|a\rangle$ and $|b\rangle$.

Using these vectors, we can extract an arbitrary element from a matrix K; before we begin however, we need to convert our $|a\rangle$ into its adjoint $\langle a|$. As mentioned previously, the corresponding *bra* for an arbitrary *ket* is its transposed complex-conjugate, meaning that $\langle a| = \begin{bmatrix}1 & 0 & 0\end{bmatrix}$. Using this information, let's set up our equation:

$$\langle a|K|b\rangle = \begin{bmatrix}1 & 0 & 0\end{bmatrix}\begin{bmatrix}K_{11} & K_{12} & K_{13}\\ K_{21} & K_{22} & K_{23}\\ K_{31} & K_{32} & K_{33}\end{bmatrix}\begin{bmatrix}0\\1\\0\end{bmatrix} \tag{1.76}$$

The entire purpose of this example is to create an intuition for $\langle m|\hat{K}|n\rangle = K_{mn}$. In our particular case, we've already mentioned that $\langle a|$ is the adjoint of the *first* standard basis vector, while $|b\rangle$ is the *second* standard basis vector. This means that we should expect to extract element K_{12} from our matrix K – let's continue our multiplication:

$$= \begin{bmatrix}1 & 0 & 0\end{bmatrix}\begin{bmatrix}K_{12}\\ K_{22}\\ K_{32}\end{bmatrix} = K_{12} \tag{1.77}$$

So our prediction was correct, and we were able to extract the expected element from our matrix.

So far, we've seen what happens when a braket composed of basis vectors contains an operator/matrix, but what happens when we are dealing with a product of operators? We will be working with the operators $\hat{K}$ and $\hat{L}$:

$$\langle m|\hat{K}\hat{L}|n\rangle = \sum_{r=1}^{N}\langle m|\hat{K}|r\rangle\,\langle r|\hat{L}|n\rangle = \sum_{r=1}^{N} K_{mr}L_{rn} \tag{1.78}$$

So we can conclude that there is an element of distributivity involved in such operations, though it involves a third index r.

1.7.4 Operators in Different Bases

Previously in 1.6.3, we discussed the methods used to implement a change of basis, but what about changing the basis of an operator? We will be using the $\hat{K}$ operator to demonstrate the process.

Firstly, we've already established that we can extract individual elements from our operator matrix via the following:

$$K_{m'n'} = \langle m'|\hat{K}|n'\rangle \tag{1.79}$$

Taking the product of the operator above and the identity operator will leave the value of the above unchanged:

$$= \langle m'|\hat{I}\hat{K}\hat{I}|n'\rangle \tag{1.80}$$

We can then use the definition of the identity operator as shown in (1.65) to expand the above:

$$= \langle m'|\sum_m |m\rangle\langle m|\sum_n |n\rangle\langle n|n'\rangle = \sum_{mn} \underbrace{\langle m'|m\rangle}_{S_{m'm}} \underbrace{\langle m|\hat{K}|n\rangle}_{K_{mn}} \underbrace{\langle n|n'\rangle}_{S^\dagger_{nn'}} \tag{1.81}$$

As a short aside, you may be wondering how we determined that $\langle n|n'\rangle = S^\dagger_{nn'}$. This is due to the following:

$$\langle n|n'\rangle = \langle n'|n\rangle^* = (S_{n'n})^* = (S^{\mathrm{T}}_{nn'})^* = S^\dagger_{nn'} \tag{1.82}$$

Anyways, we can then conclude that changing the basis of an operator can be accomplished via the following usage of our transformation matrix S:

$$\hat{K}' = S\hat{K}S^\dagger \tag{1.83}$$

1.7.5 Commutation Relations

In quantum mechanics we must use commutation relations quite often, for commutativity is no longer something we can take for granted. We are used to assuming that the commutative property is always valid:

$$ab - ba = 0,\ \forall\, a, b \in \mathbb{R} \tag{1.84}$$

However, we are no longer dealing with real numbers – instead, we are dealing with *operators* where this is not always the case! Operators that *commute* are very important, and the reasons for this are explained in Section 1.7.6. For now, let's introduce some new notation:

$$\left[\hat{A}, \hat{B}\right] = \hat{A}\hat{B} - \hat{B}\hat{A} \tag{1.85}$$

The above is basically shorthand for the commutation operation for two arbitrary operators $\hat{A}$ and $\hat{B}$. Also, given the following special case:

$$\left[\hat{A}, \hat{B}\right] = \hat{A}\hat{B} - \hat{B}\hat{A} = 0 \tag{1.86}$$

We say that our above operators *commute*[7].

As for the other properties of commutative relations, one to definitely keep in mind is the *associativity* of commutation relations:

$$\left[\hat{A}\hat{B}, \hat{C}\right] = \hat{A}\left[\hat{B}, \hat{C}\right] + \left[\hat{A}, \hat{C}\right]\hat{B} \tag{1.87}$$

You may also want to take note of the *distributivity* of commutation relations:

$$\left[\hat{A}, \hat{B} + \hat{C}\right] = \left[\hat{A}, \hat{B}\right] + \left[\hat{A}, \hat{C}\right] \tag{1.88}$$

1.7.6 Compatibility

Two or more operators are considered *compatible* if they share a *common complete set of eigenstates* – ie the same set of eigenkets, but with different eigenvalues.

Let's say we have an operator $\hat{A}$ with a set of eigenstates $|\psi_n\rangle$; if we now bring in another compatible operator $\hat{B}$, then it will share the same set of eigenkets such that:

$$\hat{A}\,|\psi_n\rangle = a_n\,|\psi_n\rangle \implies \hat{B}\,|\psi_n\rangle = b_n\,|\psi_n\rangle \tag{1.89}$$

Since these operators are compatible, this implies that they commute:

$$\left[\hat{A}, \hat{B}\right]|\psi_n\rangle = \left(\hat{A}\hat{B} - \hat{B}\hat{A}\right)|\psi_n\rangle = \cancelto{0}{(a_n b_n - b_n a_n)}\,|\psi_n\rangle = 0 \tag{1.90}$$

The above holds for all $|\psi_n\rangle$ and their superpositions, and we can therefore conclude that they commute, since their commutator is zero. A state labeled by eigenvalues of mutually compatible operators is written as follows:

$$|\psi_n\rangle \equiv |a_n, b_n\rangle \tag{1.91}$$

Conversely, if $\left[\hat{A}, \hat{B}\right] \neq 0$, this implies that $\hat{A}$ and $\hat{B}$ do *not* share a common complete set of eigenstates; in other words they would be *incompatible*.

Compatibility is a very useful property, since these eigenstates are values for states that are certain. If they are compatible, it would behoove you to choose these eigenstates – if you do, your observables will be both well-defined and certain!

One final note of great importance: if we have three arbitrary operators $\hat{A}$, $\hat{B}$, and $\hat{C}$ such that $\hat{A}$ and $\hat{B}$ commute and $\hat{B}$ and $\hat{C}$ commute, then $\hat{A}$ and $\hat{C}$ will also commute *unless* $\hat{A}$ is *degenerate*:

[7] Alternatively, we may say that they are *commutators*.

If $\left[\hat{A}, \hat{B}\right] = 0$ and $\left[\hat{B}, \hat{C}\right] = 0$ then $\left[\hat{A}, \hat{C}\right] \neq 0 \implies \hat{A}$ is **degenerate**.

1.7.7 The Hermitian Conjugate of an Operator

Up to this point, we've spent a great deal of time working with complex conjugates and transposed matrices, and we've even mentioned the *adjoint operator* (also known as *dagger*, represented by a superscript †). Let's finally explore some of the properties of the Hermitian conjugate!

Firstly, you may already be familiar with the following definition of the adjoint:

$$\hat{K}^\dagger \equiv \left(\hat{K}^{\mathrm{T}}\right)^* \tag{1.92}$$

Let's begin by expanding upon this without Dirac notation; recall the notation used in the introduction to 1.2, which was used to represent the inner product of two vectors. Using this notation, the Hermitian conjugate of an operator exhibits the following property:

$$\left(|u\rangle, \hat{K}\,|v\rangle\right) = \left(\hat{K}^\dagger\,|u\rangle, |v\rangle\right) \tag{1.93}$$

Using Dirac notation, we can rewrite the above with brakets:

$$\langle u|\hat{K}|v\rangle = \langle \hat{K}^\dagger u|v\rangle \tag{1.94}$$

We also have that:

$$\langle u|\hat{K}|v\rangle = \langle v|\hat{K}^\dagger|u\rangle^* \iff \langle v|\hat{K}^\dagger|u\rangle = \langle u|\hat{K}|v\rangle^* \tag{1.95}$$

Using a *discrete basis* $\{|m\rangle\}$, we have that:

$$\underbrace{\langle n|\hat{K}^\dagger|m\rangle}_{K^\dagger_{nm}} = \underbrace{\langle m|\hat{K}|n\rangle^*}_{K^*_{mn}} \tag{1.96}$$

In other words, we have that $K^\dagger_{nm} = K^*_{mn}$.

Example 1.7.3: The Adjoint of a 2×2 Matrix

Let's say we are given a matrix $\hat{K} = \begin{bmatrix} 1 & 2i \\ i & i \end{bmatrix}$. If we wish to find its *adjoint*, we can simply use (1.92). With this knowledge, plugging in our own $\hat{K}$ then gives us that:

$$\hat{K}^\dagger = \left(\hat{K}^{\mathrm{T}}\right)^* = \left(\begin{bmatrix} 1 & 2i \\ i & i \end{bmatrix}^{\mathrm{T}}\right)^* = \left(\begin{bmatrix} 1 & i \\ 2i & i \end{bmatrix}\right)^* = \begin{bmatrix} 1 & -i \\ -2i & -i \end{bmatrix} \tag{1.97}$$

One more useful relation to keep in mind is the following:

$$\hat{K}\left|v\right\rangle = \left\langle v\right|\hat{K}^{\dagger} \tag{1.98}$$

1.8 Hermitian Operators

Introduction

Although we've already gone through the basics regarding operators, it is best to allocate an entire section for Hermitian operators, since they make up a large portion of the operators we will be working with. As a result, this section will be dedicated to the theory behind Hermitian operators.

As we discussed in 1.7.7, the Hermitian conjugate of an operator is defined as follows:

$$\hat{K}^{\dagger} = \left(\hat{K}^{\mathrm{T}}\right)^{*} \tag{1.99}$$

Hermitian operators are a special class of operators, in the sense that they remain unchanged over adjunction:

$$\hat{K} = \hat{K}^{\dagger} \tag{1.100}$$

This allows for many interesting possiblities – let's take a look at an example:

Example 1.8.1: Is it Hermitian?

We are given the following operator:

$$\hat{K} = i\frac{d}{dx} \tag{1.101}$$

If it is Hermitian, it will abide by the following:

$$\hat{K} = \hat{K}^\dagger \tag{1.102}$$

To determine whether or not we are dealing with a Hermitian operator, we must create a pair of test states $|\psi\rangle$ and $|\phi\rangle$, which we will represent with the functions $\Psi(x)$ and $\Phi(x)$ respectively. Using these kets, we are interested in proving that the following is true:

$$\langle\psi|\hat{K}^\dagger|\phi\rangle = \langle\phi|\hat{K}|\psi\rangle^* \tag{1.103}$$

To show that the above holds, we must evaluate the right hand side via integration:

$$\langle\phi|\hat{K}|\psi\rangle^* = \left[\int_{-\infty}^{\infty} dx\, \Phi^*(x) i\frac{d}{dx}\Psi(x)\right]^* \tag{1.104}$$

Through partial integration, we can rewrite the above:

$$= \left[\cancelto{0}{\Phi^*(x)i\Psi(x)\Big|_{-\infty}^{\infty}} - \int_{-\infty}^{\infty} dx\, i\left(\frac{d}{dx}\Phi^*(x)\right)\Psi(x)\right]^* \tag{1.105}$$

Given than $|\psi\rangle$ and $|\phi\rangle$ are normalizable, and are therefore zero at $\pm\infty$, we can further simplify the above:

$$= \left[-\int_{-\infty}^{\infty} dx\, i\left(\frac{d}{dx}\Phi^*(x)\right)\Psi(x)\right]^* \tag{1.106}$$

Now, let's apply our conjugation to the integrand, and rewrite our integral using brakets:

$$= \int_{-\infty}^{\infty} dx\, \Psi^*(x) i\frac{d}{dx}\Phi(x) = \langle\psi|\hat{K}^\dagger|\phi\rangle \tag{1.107}$$

We clearly see that our original assumption holds, and that $\hat{K}$ is therefore Hermitian. □

1.8.1 Eigenvalues and Eigenvectors for Hermitian Operators

We will now discuss a special class of kets known as *eigenvectors* (or *eigenkets*), that exhibit a very interesting and useful property: when their operators act upon them, the operation is equivalent to multiplication by a number.

For example, let's assume that the operator $\hat{K}$ has an eigenvector $|\lambda\rangle$ – this would then imply the following:

$$\hat{K}\,|\lambda\rangle = \lambda\,|\lambda\rangle\,,\ \text{where}\,\lambda \in \mathbb{C} \tag{1.108}$$

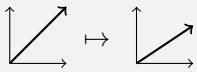
As we see above[8], we end up with a factor λ – this is our eigenvalue!

Example 1.8.2: Pointer Analogy

Let $\hat{M}$ be an operator such that:

$$\hat{M} : (x, y) \mapsto (2x, y) \tag{1.109}$$

We can represent this visually using pointers:

$\hat{M}$:  (1.110)

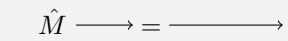
The eigenpointers[a] in this case would be pointers that remain unchanged in *direction* under the transformation given by $\hat{M}$; the factor by which the pointers are stretched will be the analogous eigenvalue.

In this case, we have two eigenpointers:  and

Why are these considered eigenpointers? It's because our operator only stretches our pointers in the x-direction; for instance, we clearly see that the vertical arrow has no x-component, which means it cannot be stretched out in that direction and therefore remains unchanged:

$\hat{M}$  (1.111)

On the other hand, the horizontal arrow has no y-component, which means that our operator would simply act as a stretching factor, doubling the length of our arrow:

$\hat{M}$  (1.112)

In other words, the vertical eigenpointer has an eigenvalue of 1, while the horizontal eigenpointer has an eigenvalue of 2. These principles can be applied to linear algebra (and therefore quantum mechanics) in myriad ways, and although the concepts presented will be far more abstract and unintuitive, this is a useful example which can make the world of linear algebra that much more relatable.

[a] An analogy for eigenvectors

1.8.2 Spectra and Degeneracy

We denote a collection of eigenvalues for an operator as the operator's **spectrum**; these can be *discrete*, *continuous*, or some combination of both, with variance in different regions.

One property of a spectrum is degeneracy, which is defined as follows:

A spectrum is termed **degenerate** if 2 or more *linearly independent eigenkets* have the *same eigenvalues*.

There are several ways to describe degeneracy in a system; let's say that we have a system with

[8] We often label the eigenkets by their corresponding eigenvalue.

g linearly independent eigenkets who share the same eigenvalue λ. We would then be able to state that:

> λ is g-fold degenerate.

Alternatively, we could also say that:

> The degree of degeneracy of λ is g.

Example 1.8.3: Pointer Degeneracy

Let's take another look at our pointer analogy; this time, we have an operator $\hat{M}$ such that:

$$\hat{M} : (x, y) \mapsto (x, y) \tag{1.113}$$

In Example 1.8.1, we found that our operator had two different eigenvalues 1 and 2 for its two eigenpointers.
This time, *every single conceivable pointer* is an eigenpointer, each with the eigenvalue 1. This is due to the fact that our operator has no effect on the input pointer, and returns the input as output.

$\hat{M}$ ↗ = ↗ $\qquad$ $\hat{M}$ → = → $\qquad$ $\hat{M}$ ⟶ = ⟶

One may be tempted to state that the eigenvalue 1 is infinitely degenerate, but this is *completely false*! The reason this is incorrect is due to the definition of degeneracy, which states that we only count *linearly independent eigenkets*.
In conclusion, since we have two coordinates x and y, our basis consists of *two* linearly independent eigenpointers. Thus, we can safely state that:

> The eigenvalue 1 is *twofold degenerate* because all *linearly independent* pointers are eigenvectors with eigenvalue 1.

1.8.3 Properties of Hermitian Operators

We've already discussed the basics of Hermitian operators, but there are other important characteristics to consider outside of their self-adjoint property. In particular:

1. The *eigenvalues* of a Hermitian operator are *real*.
2. The different *eigenvalues* of a Hermitian operator have *orthogonal eigenkets*[9].
3. The *eigenkets* of a Hermitian operator form a *basis*.[10]

[9] Some states have a total probability of 1, which implies an eigenvalue that is known with 100% certainty.
[10] Any state can be represented by a combination of eigenkets.

1.8.4 The Eigendecomposition of Hermitian Operators

To begin, let's say we have a Hermitian operator $\hat{K}$ such that:

$$\hat{K}\,|\lambda_i\rangle = \lambda_i\,|\lambda_i\rangle \tag{1.114}$$

Using the third property defined in 1.8.3 and (1.73), we can represent an arbitrary element K_{ij} of the matrix representation of $\hat{K}$ via:

$$K_{ij} = \langle\lambda_i|\hat{K}|\lambda_j\rangle = \langle\lambda_i|\lambda_j|\lambda_j\rangle \tag{1.115}$$

We can then move our central λ_j to the left of our braket:

$$= \lambda_j\,\langle\lambda_i|\lambda_j\rangle = \lambda_j\delta_{ij} \tag{1.116}$$

The above can be represented as a *diagonal matrix*:

$$\hat{K} = \begin{bmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ 0 & 0 & \lambda_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \end{bmatrix} \tag{1.117}$$

Or as a sum:

$$\hat{K} = \sum_r \lambda_r\,|\lambda_r\rangle\,\langle\lambda_r| \tag{1.118}$$

Where $|\lambda_r\rangle$ represents an arbitrary *eigenket* of $\hat{K}$.

Example 1.8.4: Explaining the Summation

We wish to show that our aforementioned sum does truly represent our Hermitian operator $\hat{K}$. Let's begin with the equation for extracting an individual matrix element K_{ij}:

$$K_{ij} = \langle \lambda_i | \hat{K} | \lambda_j \rangle \tag{1.119}$$

Let's now plug (1.118) for $\hat{K}$ into the above:

$$= \langle \lambda_i | \sum_r \lambda_r | \lambda_r \rangle \langle \lambda_r | \lambda_j \rangle \tag{1.120}$$

We can then move our central elements out of their brakets, and show that this results in the possibility of replacing two brakets with Kronecker deltas:

$$= \sum_r \lambda_r \underbrace{\langle \lambda_i | \lambda_r \rangle}_{\delta_{ir}} \underbrace{\langle \lambda_r | \lambda_j \rangle}_{\delta_{rj}} \tag{1.121}$$

Taking the sum over these Kronecker deltas shows that our original assumption was true:

$$= \sum_r \lambda_r \delta_{ir} \delta_{rj} = \lambda_j \delta_{ij} \tag{1.122}$$

CHAPTER 2

Interpretations of Quantum Mechanics

Introduction

So far, we've mostly been focused on pure linear algebra; this is because we need the tools given to us by Dirac notation in order to make quantum mechanical calculations. We can now connect the abstract theory from Chapter 1 to real physical phenomena and begin to model systems representing everything from subatomic particles to molecular dynamics.

2.1 The Postulates of Quantum Mechanics

Introduction

In this section, we will define and elaborate upon the six postulates of quantum mechanics – these are a manner by which we can use our knowledge of the formalism in Chapter 1 to describe aspects of a quantum mechanical system. Let us begin by summarizing them:

1. **Quantum states** are *normalized complex vectors.*
2. **Physical observables** are *Hermitian operators.*
3. The **measurement values** of a physical observable K are the *eigenvalues of* $\hat{K}$.
4. The **probability of getting the value** λ_i in state $|\psi\rangle$ is $|\langle\lambda_i|\psi\rangle|^2$. This is known as the *Born rule* for non-degenerate states.
5. The **time evolution of a state** is described using the *Schrödinger equation*:

$$i\hbar\frac{d}{dt}|\Psi(\vec{r},t)\rangle = \hat{H}|\Psi(\vec{r},t)\rangle \tag{2.1}$$

6. The **state of a wavefunction** after an ideal measurement yielding λ_i is reduced to $|\lambda_i\rangle$.

Now that we've seen all the postulates, we will go through each one in detail.

2.1.1 Postulate I - Quantum States

> Quantum states are normalized complex vectors.

We've already been exposed to aspects of the first postulate of quantum mechanics earlier in Section 1.1, where we used their properties as a way of giving context to the subsequent formalism. At this point, we already know that quantum states are represented by *complex vectors*, but an aspect we have not yet discussed is the fact that these complex vectors must be *normalized*. In other words, given a quantum state $|\psi\rangle$:

$$\langle\psi|\psi\rangle = 1 \tag{2.2}$$

Another important definition is that of the **Hilbert space**:

> A *Hilbert space* is defined as a *complex vector space* with *normalizable vectors*.

An addition interesting property of quantum states is that two states that differ by a *phase factor*[1] correspond to the *same physical state*:

$$|\psi\rangle = e^{i\theta}\,|\psi\rangle \tag{2.3}$$

This allows us to visualize quantum states as *rays* in Hilbert space!

2.1.2 Postulate II - Physical Observables

> Physical observables are Hermitian operators.

The details behind this statement require a knowledge of classical mechanics, so let us instead look at two examples of operators:

[1] The phase factor θ is a real number.

Example 2.1.1: Useful Operators

In quantum mechanics, we cannot simultaneously know a system's position and momentum with absolute certainty due to *Heisenberg's uncertainty principle*. To access the probabilistic information in our system regarding these properties, we can make use of:

The **position operator** $\hat{x} = x$.

The **momentum operator** $\hat{p} = -i\hbar\frac{d}{dx}$

According to Heisenberg's uncertainty principle, we can only obtain a sharp measurement for two observables simultaneously *if they commute*. That is to say, we can have an exact measurement for two arbitrary operators $\hat{A}$ and $\hat{B}$ *only* if:

$$\left[\hat{A}, \hat{B}\right] = 0 \tag{2.4}$$

In this case, we have that *position and momentum do not commute*, and so they cannot be exactly measured simultaneously:

$$[\hat{x}, \hat{p}] = \hat{x}\hat{p} - \hat{p}\hat{x} = x \cdot -i\hbar\frac{d}{dx} + i\hbar\frac{d}{dx}x = i\hbar \neq 0 \quad \square \tag{2.5}$$

As an added point, we can always test whether our system is correctly evaluated by using *Ehrenfest's theorem*:

The expectation value of quantum operators obey the classical equations of motion.

So if you are unsure about a result, test it against Ehrenfest's theorem to see whether or not it obeys classical laws of motion – if it does not, it may be mathematically correct, but not physically!

2.1.3 Postulate III - Measurement Values

The measurement values of a physical observable K are the eigenvalues of $\hat{K}$.

Alternatively, one could say that the measurement values of an observable are an eigenvalue of the corresponding operator. There are a few more points to be made here, namely:

- In each measurement, the outcome is *one* of the eigenvalues.
- For *Hermitian operators*, eigenvalues are *real*.
- The number of unique eigenvalues is *less than or equal to* the number of eigenvectors in (ie. dimension D of) the Hilbert space.
- The dimension of the Hilbert space D is *greater than or equal to* the number of mutually exclusive measurement values of any observables.

The aforementioned principles can be used to one's advantage in many situations, so keep them in mind!

2.1.4 Postulate IV - Probability Amplitudes and Born's Rule

The probability of getting the value λ_i in state $|\psi\rangle$ is $|\langle\lambda_i|\psi\rangle|^2$.
This is known as the *Born rule* for non-degenerate states.

The above statement is very useful as a starting point, but we must elaborate upon it since it only applies to non-degenerate states[2].

The Non-Degenerate Discrete Case

Let's begin by defining an operator $\hat{K}$ such that:

$$\hat{K}\,|\lambda_i\rangle = \lambda_i\,|\lambda_i\rangle \tag{2.6}$$

If we are given a *discrete orthogonal set of eigenvalues* such that $\langle\lambda_i|\lambda_j\rangle = \delta_{ij}$, then the probability amplitude of measuring the value λ_i for a wavefunction ψ is λ_i is $p_{\lambda_i} = \langle\lambda_i|\psi\rangle$. This means that the probability is:

$$P_{\lambda_i} = |\langle\lambda_i|\psi\rangle|^2 \tag{2.7}$$

Example 2.1.2: The Sum of Multiple Probabilities

We wish to show that taking the sum over all probabilites P_{λ_i} gives us one. Let's begin by setting up a summation over all i:

$$\sum_i P_{\lambda_i} = \sum_i |\langle\lambda_i|\psi\rangle|^2 = \sum_i \langle\psi|\lambda_i\rangle\,\langle\lambda_i|\psi\rangle = \langle\psi|\underbrace{\left(\sum_i |\lambda_i\rangle\,\langle\lambda_i|\right)}_{\text{Identity } \hat{I}}|\psi\rangle \tag{2.8}$$

Since the expression in (2.8) is orthonormal, it contains the *identity operator*; we are then left with the following result:

$$\implies \langle\psi|\psi\rangle = 1 \quad \square \tag{2.9}$$

In addition to the above, we also have that if $|\psi\rangle = |\lambda_j\rangle$ and $|\langle\lambda_i|\lambda_j\rangle|^2 = \delta_{ij}$, this implies that $|\lambda_j\rangle$ is a state that gives the measurement value λ_j with 100% certainty.

The Non-Degenerate Continuous Case

What if we have a set of eigenvalues that is continuous? Then, the probability will always be zero unless taken over an interval; so given an interval $[\lambda, \lambda + \Delta\lambda]$, the probability that our eigenvalue lies within this interval is defined as follows:

$$P_{\lambda_i} = \int_{\lambda}^{\lambda+\Delta\lambda} d\lambda\,|\langle\lambda|\psi\rangle|^2 \tag{2.10}$$

[2]Degeneracy is introduced in Section 2.1.4.

Where $|\langle\lambda|\psi\rangle|^2$ is the probability density.

The Expectation Value of an Observable K

If we wish to find the expectation value of an observable with the operator $\hat{K}$, we begin by taking the weighted mean of our probabilites as follows:

$$\underbrace{\langle K\rangle_\psi}_{\text{Expectation value in state } \psi} = \sum_i \lambda_i P_{\lambda_i} = \sum_i \langle\psi|\lambda_i\rangle\, \lambda_i\, \langle\lambda_i|\psi\rangle \tag{2.11}$$

We can use the *eigendecomposition*[3] of the operator for K to our advantage:

$$= \langle\psi| \underbrace{\left(\sum_i |\lambda_i\rangle\, \lambda_i\, \langle\lambda_i|\right)}_{\text{Eigendecomposition of } \hat{K}} |\psi\rangle = \langle\psi|\hat{K}|\psi\rangle \tag{2.12}$$

So the expectation value of an observable K in state ψ is defined as:

$$\langle K\rangle_\psi = \langle\psi|\hat{K}|\psi\rangle \tag{2.13}$$

The Degenerate Case

So far, we've only accounted for Born's rule in cases of *non-degeneracy*; so what happens when our system[4] has several eigenstates corresponding to the same eigenvalues? In such a case, we begin by defining the **projection operator**:

$$\hat{P}_\lambda = \sum_g |\lambda^{(g)}\rangle\, \langle\lambda^{(g)}| \tag{2.14}$$

In the above, we have several eigenkets $|\lambda^{(g)}\rangle$ with the same eigenvalue λ. Using this information, we can derive Born's rule for a degenerate set of eigenstates:

$$P_\lambda = \langle\psi|\hat{P}_\lambda|\psi\rangle = \sum_g \left|\langle\lambda^{(g)}|\psi\rangle\right|^2 \tag{2.15}$$

[3] This is introduced in Section 1.8.4

[4] Our system may be discrete or continuous.

2.1.5 Postulate V - The Time Evolution of a State

The time evolution of a state is described using the Schrödinger equation:

$$i\hbar\frac{d}{dt}\left|\Psi\left(\vec{r},t\right)\right\rangle = \hat{H}\left|\Psi\left(\vec{r},t\right)\right\rangle \tag{2.16}$$

The Schrödinger equation is used to relate a state's energy to its change over time. It is non-chaotic, meaning that it is well-behaved, linear, and deterministic.

2.1.6 Postulate VI - Ideal Measurements

The state of a wavefunction after an ideal measurement yielding λ_i is reduced to $|\lambda_i\rangle$.

Where an ideal measurement only *minimally* disturbs the system.

This postulate is considered dependent on **Postulate V**, since it is derived from the Schrödinger equation. Regardless, it is important as it introduces us to the idea of a *collapsed wavefunction*, which "is one of two processes by which quantum systems evolve in time; the other is continuous evolution via the Schrödinger equation" [2].

A wavefunction is said to collapse when it is observed; initially, such a wavefunction would be composed of a superposition of eigenstates. Once the observation is complete and an individual eigenstate is observed, it becomes the sole remaining eigenstate.

Example 2.1.3: Modeling the Collapse

We can use the projection operator $\hat{P}_\lambda$ for an eigenstate λ to model the collapse of a wavefunction ψ:

$$|\psi\rangle \xrightarrow[\text{collapse}]{} \frac{\hat{P}_\lambda |\psi\rangle}{\langle\psi|\hat{P}_\lambda^\dagger \hat{P}_\lambda|\psi\rangle^{\frac{1}{2}}} \tag{2.17}$$

The projection operator is Hermitian, so $\hat{P}_\lambda^\dagger = \hat{P}_\lambda$. The projection operator is defined as follows:

$$\hat{P}_\lambda = |\lambda\rangle\langle\lambda| \tag{2.18}$$

Let's plug this into (2.17) and see what happens:

$$= \frac{|\lambda\rangle\langle\lambda|\psi\rangle}{\left(\langle\psi|\lambda\rangle \cancelto{1}{\langle\lambda|\lambda\rangle} \langle\lambda|\psi\rangle\right)^{\frac{1}{2}}} = \frac{\langle\psi|\lambda\rangle}{|\langle\psi|\lambda\rangle|}|\lambda\rangle \tag{2.19}$$

Take a close look at our final expression – it is composed of a phase factor $\frac{\langle\psi|\lambda\rangle}{|\langle\psi|\lambda\rangle|}$ multiplied by our observed eigenvalue $|\lambda\rangle$. If we recall the conclusion from (2.3), we know that two states that differ by a phase factor correspond to the same physical state. In other words, we no longer have a superposition of eigenkets, and are left only with the measured physical state.

2.2 Multiple Interpretations

Introduction

So far in this chapter, our discussion has revolved around six universally accepted postulates; these have developed through observation and the study of Hilbert spaces. Accepting these postulates is considered a *minimal approach*, as it does not make any assumptions about the nature of the quantum world – simply put, these postulates simply make statements about the outcomes of measurements and their probabilities during repeated experiments.

Since these theories are relatively young, making further assumptions about what occurs on the quantum level requires a leap of faith. As a result, there are multiple competing interpretations as to what is *truly* occuring on the subatomic level – for example, is the world inherently probabilistic, or are we simply unable to measure their exact values due to technological limitations? Let's begin with the instrumentalist perspective!

2.2.1 The Copenhagen Interpretation

(a) Niels Bohr [3] (b) Werner Heisenberg [4]

Figure 2.1

Originating with *Niels Bohr* and *Werner Heisenberg*, the Copenhagen (or instrumentalist) interpretation of quantum mechanics assumes that an individual system does not have definite and hidden values for its physical properties; from their viewpoint, measurements are fundamental and QM is only concerned with predicting the outcomes

This diverges from any previous physical science; even systems in statistical mechanics are composed of particles with unknown positions and velocities, but they nevertheless *exists* in a microstate[5]; those probabilites are simply convenient as mathematical tools. According to the Copenhagen interpretation, a system cannot be described with certainty no matter what - *probabilities are fundamental*!

One issue with this interpretation is one of boundaries: where does a measurement go from being quantum to classical? For the moment, there is no well defined scale to determines this.

2.2.2 The Realist Interpretation

On the polar opposite side of things, we have the realists; this interpretation assumes that an individual system has *definite values* for **all** its physical properties. This perspective is one in

[5]These are simply not considered due to limits in measuring equipment.

which we claim ignorance – we assume these probabilites in quantum mechanics reflect our lack of knowledge, and that measurements play *no fundamental role*.

This naturally leads to the conclusion that quantum mechanics is incomplete. Through such a conclusion, we must also assume that there exist *hidden variables*, and that obtaining them would give a complete picture of the system.

An issue with this interpretation is the fact that given the existence of hidden variables, we can formulate the Bell inequalities in such a way that they are satisfied by all such hidden variable theories, but not by quantum mechanics[6]! One must accept non-locality to believe in realism.

2.2.3 The Many Worlds Interpretation

There is still a question to raise, and that is: what is the difference between the *classical world* and the *quantum world*? As we discussed in Section 2.2.1, there is no such sharp distinction between these – so what are the limits of the quantum world?

The first person to seriously consider the notion that measurements are also quantum was *Hugh Everett* in 1957; this lead to a wild conclusion. Take the following system for instance:

$$|S\rangle = \alpha\,|S_1\rangle + \beta\,|S_2\rangle \tag{2.20}$$

If measurements are also quantum, then that means we must include our measurement apparatus in the equation once a measurement occurs, since the environment is also taken into account as a quantum state. We can call this measurement apparatus $|M_0\rangle$. When these interact, the new state of the system is then as follows:

$$|S\rangle = |S_1\rangle\,|M_1\rangle + |S_2\rangle\,|M_2\rangle \tag{2.21}$$

Following this, if our measurement apparatus shows the result M_1, our system becomes the first state S_1. If our apparatus displays M_2, we have S_2. This interpretation implies that we ourselves are also a part of M_1 (or M_2), so we cannot know the result of the other measurement since it would logically reside outside of our universe. In plainer terms: measurements don't collapse wavefunction, it only leads to the *proliferation of worlds*!

One of the main issues regarding this interpretation is as follows: how can we explain the origin of the Born rule, if we cannot know the required coefficients when we only exist within a single M_i?

[6] According to the Copenhagen interpretation, quantum theory is complete, while the realists believe that we should develop a new theory that includes hidden variables and locality.

CHAPTER 3

One Dimensional Quantum Mechanics

3.1 Time Evolution of States

Introduction

According to the fifth[1] postulate of quantum mechanics, a state will evolve as time passes. Imagine for instance that we are given the following state:

$$|\alpha(t_0)\rangle = |\alpha\rangle \tag{3.1}$$

As time passes, we can represent the evolution of this state at a time t via a simple product:

$$|\alpha(t)\rangle = \hat{U}(t, t_0)\,|\alpha\rangle \tag{3.2}$$

Where $\hat{U}$ is defined as the *time evolution operator*. Alternatively, we can also call it a *propagator*.

3.1.1 Conditions for Û

In Section 2.1.5, we also discussed the fact that the time evolution of a state had to be linear and non-chaotic. This puts constraints on the behavior of our propagator $\hat{U}$:

1. The time evolution operator does not affect our state at t_0:

$$\lim_{t \to t_0} \hat{U}(t, t_0) = \hat{I} \tag{3.3}$$

2. $\hat{U}$ should preserve the norm of states:

$$\langle \alpha(t) | \alpha(t) \rangle = \langle \alpha | \cancelto{\hat{I}}{\hat{U}^\dagger(t, t_0)\hat{U}(t, t_0)} | \alpha \rangle = \langle \alpha | \alpha \rangle \tag{3.4}$$

   This can be summarized by stating that $\hat{U}$ must be a *unitary operator*[2], and that:

$$\hat{U}^\dagger(t, t_0)\hat{U}(t, t_0) = \hat{I} \tag{3.5}$$

3. The *composition property* for $\hat{U}$ implies that:

$$\hat{U}(t_2, t_0) = \hat{U}(t_2, t_1)\hat{U}(t_1, t_0) \tag{3.6}$$

[1] Introduced in Section 2.1.5.

[2] If $\hat{A}$ is a unitary operator, then $\hat{A} = \hat{A}^\dagger = \hat{A}^{-1}$

3.1.2 Derivation of the Schrödinger Equation

Let's use these conditions to derive the Schrödinger equation. Since this equation describes change over time, we can start by considering an *infinitessimal time evolution*:

$$\hat{U}(t_0 + dt, t_0) = \hat{I} - i\hat{\Omega}(t_0)dt \tag{3.7}$$

Where $\hat{\Omega}(t_0) = \hat{\Omega}(t_0)^\dagger$ is *Hermitian*. To continue, we need to confirm that our three aforementioned conditions are satisfied for the given $\hat{U}$.

The Identity

Let's begin by showing that:

$$\lim_{t \to t_0} \hat{U}(t, t_0) = \hat{I} \tag{3.8}$$

Since we are using an infinitessimal with a limit, this will be straightforward:

$$\lim_{dt \to 0} \hat{U}(t_0 + dt, t_0) = \lim_{t \to t_0} \hat{U}(t, t_0) = \hat{I} \quad \square \tag{3.9}$$

Preserving the Norm

Now, we must show that:

$$\langle \alpha | \hat{U}^\dagger(t, t_0) \hat{U}(t, t_0) | \alpha \rangle = \langle \alpha | \alpha \rangle \tag{3.10}$$

Rather than dealing with the whole equation, it will suffice to check that:

$$\hat{U}^\dagger(t, t_0) \hat{U}(t, t_0) = \hat{I} \tag{3.11}$$

Since we are checking this for $\hat{U}(t_0 + dt, t_0) = \hat{I} - i\hat{\Omega}(t_0)dt$, let's simply plug it into (3.11) and see what happens:

$$\hat{U}^\dagger(t, t_0) \hat{U}(t, t_0) = \Big(\hat{I} + i\hat{\Omega}(t_0)dt\Big)\Big(\hat{I} - i\hat{\Omega}(t_0)dt\Big) \tag{3.12}$$

To simplify our next step, we can set $\mathcal{O}\left(dt^2\right) \equiv \hat{\Omega}(t_0)^2 dt^2$. We can then take the following product:

$$= \hat{I} + \cancelto{0}{i\hat{\Omega}(t_0)dt - i\hat{\Omega}(t_0)dt} + \mathcal{O}\left(dt^2\right) \tag{3.13}$$

We can then conclude that $\hat{U}(t, t_0)$ is unitary with a deviation of $\mathcal{O}\left(dt^2\right)$. Since we are assuming that $dt \to 0$, this leads to $\mathcal{O}\left(dt^2\right) \to 0$, and we've therefore fulfilled our second condition. $\square$

The Composition Property

Let show that our last condition is fulfilled:

$$\hat{U}(t_2, t_0) = \hat{U}(t_2, t_1)\hat{U}(t_1, t_0) \tag{3.14}$$

Let's start with the left side of the equation:

LHS Let $t_2 = t_0 + 2dt$ and let $t_1 = t_0 + dt$, then:

$$\hat{U}\left(t_2, t_0\right) = \hat{I} - i\hat{\Omega}(t_0)2dt \tag{3.15}$$

Now let's take a look at the right side:

RHS

$$\hat{U}\left(t_2, t_1\right)\hat{U}\left(t_1, t_0\right) = \left(\hat{I} - i\hat{\Omega}\left(t_0 + dt\right)dt\right)\left(\hat{I} - i\hat{\Omega}\left(t_0\right)dt\right) \tag{3.16}$$

As seen previously, $\mathcal{O}\left(dt^2\right) \equiv \hat{\Omega}(t_0)^2 dt^2$; expanding the above and plugging this in gives us:

$$= \hat{I} - i\hat{\Omega}(t_0)2dt + \mathcal{O}\left(dt^2\right) \tag{3.17}$$

And so we see that our composition property is fulfilled, once again with a deviation of $\mathcal{O}\left(dt^2\right)$ that approaches zero. □

Final Steps

It is at this point we should discuss our magical little operator $\hat{\Omega}$ – what does it actually represent? In short, it is a step away from the Hamiltonian operator:

$$\hat{\Omega}(t) = \frac{\hat{H}(t)}{\hbar} \tag{3.18}$$

So with our now complete picture, we can plug (3.18) into (3.7), yielding the following:

$$\hat{U}\left(t_0 + dt, t_0\right) = \hat{I} - i\frac{\hat{H}(t_0)}{\hbar}dt \tag{3.19}$$

According to the *composition property*, we know that the following is true:

$$\hat{U}\left(t + dt, t_0\right) = \underbrace{\hat{U}\left(t + dt, t\right)}_{= \hat{I} - i\frac{\hat{H}(t)}{\hbar}dt}\hat{U}\left(t, t_0\right) \tag{3.20}$$

We can subtract $\hat{U}(t, t_0)$ from both sides:

$$\hat{U}\left(t + dt, t_0\right) - \hat{U}\left(t, t_0\right) = -i\frac{\hat{H}(t)}{\hbar}dt\,\hat{U}(t, t_0) \tag{3.21}$$

And then multiply both sides by $\frac{i\hbar}{dt}$:

$$i\hbar\frac{d}{dt}\hat{U}(t,t_0) = \hat{H}(t)\hat{U}(t,t_0) \tag{3.22}$$

Now, let's have our equation operate on an arbitrary state $|\alpha\rangle$:

$$i\hbar\frac{d}{dt}\hat{U}(t,t_0)\,|\alpha\rangle = \hat{H}(t)\hat{U}(t,t_0)\,|\alpha\rangle \tag{3.23}$$

Finally, we apply $\hat{U}(t,t_0)$ to our state as shown in (3.2), leaving us with the Schrödinger equation:

$$i\hbar\frac{d}{dt}\,|\alpha(t)\rangle = \hat{H}(t)\,|\alpha(t)\rangle \quad \square \tag{3.24}$$

3.2 The Schrödinger Equation

Introduction

We first introduced the Schrödinger equation in Section 2.1.5, where we discussed its properties – specifically, that it would evolve in a deterministic and non-chaotic way.

$$i\hbar\frac{d}{dt}\,|\alpha(t)\rangle = \hat{H}(t)\,|\alpha(t)\rangle \tag{3.25}$$

(3.25) is known as the **time-dependent Schrödinger equation** or **TDSE**, since we see that it is a function of time – in some cases, however, it is possible to work with a time-independent version of this equation. Also, take note of the operator $\hat{H}$; this is known as the **Hamiltonian** of a system – simply put, it can be considered representative of the *total energy in a given system.*

3.2.1 The Time-Independent Schrödinger Equation

Derivation

Let's now make an assumption about our system – let's define our Hamiltonian $\hat{H}$ such that it is not explicitly dependent on time:

$$\hat{U}(t,t_0) = e^{-\frac{i}{\hbar}\hat{H}(t-t_0)} \tag{3.26}$$

Let's implement a taylor expansion for our exponent such that $e^x = \sum_{n=0}^{\infty}\frac{x^n}{n!}$:

$$= e^{-\frac{i}{\hbar}\hat{H}(t-t_0)} = \hat{I} - \frac{i}{\hbar}\hat{H}(t-t_0) + \frac{1}{2}\left(-\frac{i}{\hbar}\right)^2\hat{H}\hat{H}(t-t_0) - \cdots \tag{3.27}$$

Now, let's take a look at our wavefunction $|\alpha\rangle$, which can be expanded as follows:

$$|\alpha\rangle = \sum_n \alpha_n |n\rangle \tag{3.28}$$

Where the energy eigenstate $|n\rangle$ is characterized by $\hat{H}|n\rangle = E_n|n\rangle$. Using this, we then have that the state at $|\alpha(t)\rangle$ is:

$$|\alpha(t)\rangle = \hat{U}(t,0)|\alpha\rangle = e^{-\frac{i\hat{H}t}{\hbar}} \sum_n \alpha_n |n\rangle = \sum_n \alpha_n e^{-\frac{i\hat{H}t}{\hbar}} |n\rangle \tag{3.29}$$

Let's ignore the sum and our elements α_n for now, and Taylor-expand an individual *exponent-ket* pair:

$$e^{-\frac{i\hat{H}t}{\hbar}} |n\rangle = \left(\hat{I} - \frac{i}{\hbar}\hat{H}t + \frac{1}{2}\left(-\frac{i}{\hbar}\right)^2 t^2 \hat{H}\hat{H} + \cdots\right)|n\rangle \tag{3.30}$$

Next, we use our aforementioned eigenvalue E_n:

$$= \left(\hat{I} - \frac{iE_n t}{\hbar} + \frac{1}{2}\left(-\frac{i}{\hbar}\right)^2 t^2 E_n^2 + \cdots\right)|n\rangle \tag{3.31}$$

And finally, we return our expansion to exponent form:

$$= e^{-\frac{iE_n t}{\hbar}} |n\rangle \tag{3.32}$$

So in conclusion, if we assume that our Hamiltonian is time-independent, we can conclude that our time-dependent state $|\alpha(t)\rangle$ can be written as a sum of products of the *time-independent states* α_n and a propagator:

$$|\alpha(t)\rangle = \sum_n \alpha_n e^{-\frac{iE_n t}{\hbar}} |n\rangle \tag{3.33}$$

Let's plug this into the *TDSE*, and see what happens:

$$i\hbar\frac{d}{dt}\sum_n \alpha_n e^{-\frac{iE_n t}{\hbar}} |n\rangle = \hat{H}\sum_n \alpha_n e^{-\frac{iE_n t}{\hbar}} |n\rangle \tag{3.34}$$

The sum of all α_n can be represented by $|\alpha\rangle$, so let's rewrite the above:

$$i\hbar|\alpha\rangle\sum_n \frac{d}{dt} e^{-\frac{iE_n t}{\hbar}} |n\rangle = \hat{H}|\alpha\rangle\sum_n e^{-\frac{iE_n t}{\hbar}} |n\rangle \tag{3.35}$$

We can then take the derivative on the left-hand side and cancel out some terms:

$$-\frac{iE_n}{\hbar}\cdot i\hbar|\alpha\rangle \cancel{\sum_n e^{-\frac{iE_n t}{\hbar}} |n\rangle} = \hat{H}|\alpha\rangle \cancel{\sum_n e^{-\frac{iE_n t}{\hbar}} |n\rangle} \tag{3.36}$$

And we are left with the **time-independent Schrödinger equation**, or **TISE**:

$$E_n |\alpha\rangle = \hat{H} |\alpha\rangle \tag{3.37}$$

Stationary States

The time-dependence of the expectation value of an arbitrary operator $\hat{O}$ is given by:

$$\langle\alpha(t)|\hat{O}|\alpha(t)\rangle = \langle\alpha|\hat{U}^\dagger(t,t_0)\hat{O}\hat{U}(t,t_0)|\alpha\rangle \tag{3.38}$$

If our operator $\hat{O}$ is time-independent, then we have that:

$$\langle\alpha(t)|\hat{O}|\alpha(t)\rangle = \langle\alpha|\hat{O}|\alpha\rangle \tag{3.39}$$

Given that $\hat{O}$ is time-independent, then the above implies that $|\alpha\rangle$ is also *time-independent* – for instance, given the following:

$$\hat{U}(t,t_0)\,|\alpha\rangle = e^{i\varphi(t,t_0)}\,|\alpha\rangle \tag{3.40}$$

Then $\hat{U}(t,t_0)$ is solely changing the *phase* of $|\alpha\rangle$, implying that $|\alpha\rangle$ is a **stationary state**. One example of a stationary state is the *energy eigenstate*.

Example 3.2.1: Time-Independent Expectation Values

Let's show that the expectation value of any operator is *time-independent* given the case in (3.40). To accomplish this, we must begin by setting up the complete expectation value equation shown in (3.38) and plug (3.40) in:

$$\langle\alpha|\hat{U}^\dagger(t,t_0)\hat{O}\hat{U}(t,t_0)|\alpha\rangle = \cancelto{1}{e^{-i\varphi(t,t_0)}e^{i\varphi(t,t_0)}}\,\langle\alpha|\hat{O}|\alpha\rangle \tag{3.41}$$

As we see above, the propagators cancel out, and we are left with a time-independent result, proving our original assumption. □

3.2.2 The Schrödinger and Heisenberg Picture

Introduction

So far, we've approached all concepts surrounding operators and states from one direction – we've assumed that *states* evolve over time, and that operators remain unchanged over time; this is known as the **Schrödinger picture**, and is the basis upon all the work we've done in this section.

Alternatively, it is possible to approach quantum mechanics from another direction: we may choose to assume our states remain unchanged over time, and that it is our *operators* that are time-dependent. This perspective is known as the **Heisenberg picture**.

The process by which we move from one picture to another is relatively straightforward; initially, we begin with (3.38) in the *Schrödinger picture*:

$$\langle\alpha(t)|\hat{O}|\alpha(t)\rangle = \left(\langle\alpha|\,\hat{U}^{\dagger}(t,t_0)\right)\hat{O}\left(\hat{U}(t,t_0)\,|\alpha\rangle\right) \tag{3.42}$$

Next, we change the order by which we operate our propagators, giving us the *Heisenberg picture*:

$$= \langle\alpha|\left(\hat{U}^{\dagger}(t,t_0)\hat{O}\hat{U}(t,t_0)\right)|\alpha\rangle = \langle\alpha|\hat{O}(t,t_0)|\alpha\rangle \tag{3.43}$$

The Heisenberg Law of Motion

As mentioned in the introduction, the Heisenberg picture assumes that kets are time-independent, and that operators depend on time. As shown previously, this implies that:

$$\hat{O}(t,t_0) = \hat{U}^{\dagger}(t,t_0)\hat{O}\hat{U}(t,t_0) \tag{3.44}$$

We can apply this conclusion to derive the Heisenberg law of motion, an equivalent alternative to the *Schrödinger equation*:

$$\hat{O}(t+dt,t) = \left(\hat{I} + i\frac{\hat{H}(t)}{\hbar}\right)\hat{O}(t)\left(\hat{I} - i\frac{\hat{H}(t)}{\hbar}\right) \tag{3.45}$$

Expanding this gives us the following:

$$= \hat{O} + \frac{i}{\hbar}\left(\hat{H}(t)\hat{O}(t) - \hat{O}(t)\hat{H}(t)\right)dt + \mathcal{O}\left(dt^2\right) \tag{3.46}$$

As explained in Section 3.1.2, we can disregard $\mathcal{O}\left(dt^2\right)$ since $dt \to 0$. In addition to this, recall from Section 1.7.5 that we can use a more compact notation for commutation such that:

$$\left[\hat{H}(t),\hat{O}(t)\right] \equiv \hat{H}(t)\hat{O}(t) - \hat{O}(t)\hat{H}(t) \tag{3.47}$$

Given these simplifications, the final step is to take the limit as $dt \to 0$, divide by dt, and move $\hat{O}(t)$ to the left-hand side. This leaves us with the **Heisenberg law of motion**:

$$\frac{d}{dt}\hat{O}(t) = \frac{i}{\hbar}\left[\hat{H}(t),\hat{O}(t)\right] \tag{3.48}$$

3.3 Position and Momentum

As a short aside, it is important to be familiar with two operators that were briefly mentioned in Example 2.1.2 – position and momentum.

First of all, these are defined as:

The **position operator**:

$$\hat{x} = x \tag{3.49}$$

The **momentum operator**:

$$\hat{p} = -i\hbar\frac{d}{dx} \tag{3.50}$$

Although these do not commute with each other, the result of performing a commutation between them yields an interesting result:

$$[\hat{x}, \hat{p}] = i\hbar \tag{3.51}$$

This operation is known as the **canonical commutation relation**, and is of vital importance in many calculations.

3.4 Harmonic Oscillators

Introduction

In classical mechanics, we can use relatively simple mathematics to describe the behavior of waves and oscillations. One common model is that of the **harmonic oscillator**: a system that always trends towards a point of equilibrium by an amount proportional to its distance from the point of equilibrium.

We can construct a simple harmonic oscillator potential via a generalized second-order Taylor expansion, excluding the first element:

$$V(x) = V_0 + \frac{1}{2}V^{''}(0)x^2 + \mathcal{O}\left(x^3\right) \tag{3.52}$$

Then, we simply redefine the above in a more desirable format:

$$V_{\text{H}}(x) = V_0 + \frac{1}{2}V^{''}(0)x^2 \tag{3.53}$$

Let's use this to derive an equation for the total energy of our system. First, we use the force-potential relation:

$$F(x) = -\nabla V_{\text{H}}(x) = -\frac{d}{dx}\left(V_0 + \frac{1}{2}V^{''}(0)x^2\right) = -V^{''}(0)x \tag{3.54}$$

We see that this equation is of the same for as that of a spring force, so we can set $k \equiv V^{''}(0)$. Classically speaking, there is a way to parameterize k, and that is in terms of the oscillation frequency ω and the mass m:

$$\omega = \sqrt{\frac{k}{m}} \implies k = m\omega^2 \tag{3.55}$$

Using all the above allows us to create our total energy equation:

$$E = U + K = \frac{p^2}{2m} + \frac{1}{2}kx^2 = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 x^2 \tag{3.56}$$

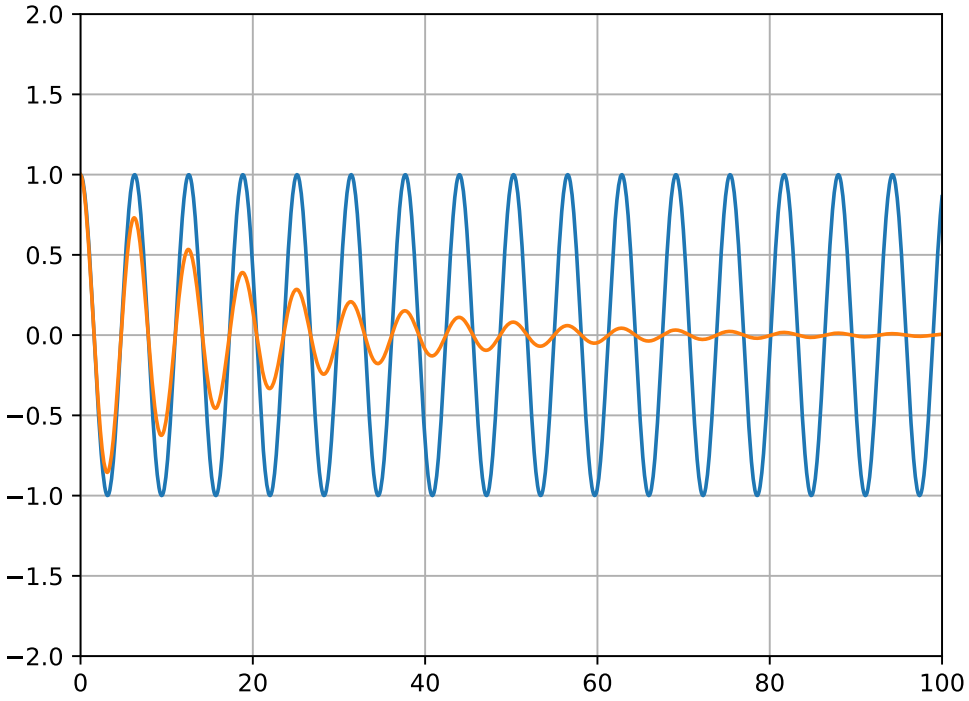


Figure 3.1: A **simple harmonic oscillator** (in blue) and a **damped harmonic oscillator** (in orange).

However, we are not here to discuss classical mechanics, we are here to discuss *quantum mechanics*! To convert our energy equation (3.56) from classical to quantum, we simply plug in our operators for position and momentum, and turn the full expression into a *Hamiltonian* operator[3]:

$$\hat{H} = \frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega^2\hat{x}^2 \tag{3.57}$$

If need be, this can be expanded upon further by using the definitions for the momentum and position operators:

$$\hat{H} = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + \frac{1}{2}m\omega^2 x^2 \tag{3.58}$$

[3]Recall that the Hamiltonian corresponds to the *sum of the potential and kinetic energy* in a system

Constructing the Schrödinger equation with this operator thus generates a system with an oscillating potential!

3.4.1 Ladder Operators

A useful property of the quantum harmonic oscillator is its **creation** and **annihilation** operators: $\hat{a}^\dagger$ and $\hat{a}$ respectively. These are members of a class of operators called **raising** and **lowering** operators:

> "A **raising** or **lowering** operator (collectively known as **ladder operators**) is an operator that increases or decreases the *eigenvalue* of another operator." [5]

For our ladder operators $\hat{a}^\dagger$ or $\hat{a}$, it is *only* meaningful to apply them to an eigenstate of the Hamiltonian $\hat{H}$ for the harmonic oscillator. Let's take a look at their properties:

The quantum harmonic oscillator **creation operator**:

$$\hat{a}^\dagger \equiv \frac{1}{\sqrt{2\hbar\omega m}}\left(m\omega\hat{x} - i\hat{p}\right) \tag{3.59}$$

The quantum harmonic oscillator **annihilation operator**:

$$\hat{a} \equiv \frac{1}{\sqrt{2\hbar\omega m}}\left(m\omega\hat{x} + i\hat{p}\right) \tag{3.60}$$

Our harmonic oscillator Hamiltonian $\hat{H}$ can also be defined in terms of $\hat{a}^\dagger$ and $\hat{a}$:

$$\hat{H} = \hbar\omega\left(\hat{a}^\dagger\hat{a} + \frac{1}{2}\right) \tag{3.61}$$

We will also need the **commutation relation** for $\hat{a}^\dagger$ and $\hat{a}$:

$$\left[\hat{a}, \hat{a}^\dagger\right] = 1 \tag{3.62}$$

This can be used to find a potentially useful relation:

$$\hat{a}\hat{a}^\dagger - \hat{a}^\dagger\hat{a} = 1 \implies \hat{a}\hat{a}^\dagger = 1 + \hat{a}^\dagger\hat{a} \tag{3.63}$$

Example 3.4.1: Redefining the Hamiltonian

In (3.61) we claimed that we could replace portions of the harmonic oscillator Hamiltonian by using a product of its ladder operators. Let's prove it:

$$\hat{a}^\dagger \hat{a} = \frac{1}{2\hbar\omega m} \left(m\omega\hat{x} - i\hat{p}\right)\left(m\omega\hat{x} + i\hat{p}\right) \tag{3.64}$$

Things get a bit messy as we multiply and expand:

$$= \frac{1}{2\hbar\omega m}\left(m^2\omega^2\hat{x}^2 + \hat{p}^2 + im\omega\,\cancelto{[\hat{x},\hat{p}]}{\left(\hat{x}\hat{p} - \hat{p}\hat{x}\right)}\right) \tag{3.65}$$

We are then able to apply the *canonical commutation relation*[a], and find a very simple solution:

$$= \frac{1}{\hbar\omega}\left(\cancel{\frac{m^2\omega^2}{2m}\hat{x}^2} + \cancelto{\hat{H}}{\frac{\hat{p}^2}{2m}} - \frac{m\omega\hbar}{2m}\right) = \frac{1}{\hbar\omega}\hat{H} - \frac{1}{2} \tag{3.66}$$

We can finally prove that the harmonic oscillator Hamiltonian can be written in terms of the product of its ladder operators:

$$\hat{a}^\dagger\hat{a} = \frac{1}{\hbar\omega}\hat{H} - \frac{1}{2} \iff \hat{H} = \hbar\omega\left(\hat{a}^\dagger\hat{a} + \frac{1}{2}\right) \quad \square \tag{3.67}$$

[a] Introduced in (3.51) on p.50

Since we are dealing with an analytically solvable harmonic oscillator potential, this implies that our energy eigenvalue can also be defined analytically:

The **energy eigenvalue** for an eigenstate $|n\rangle$:

$$E_n = \hbar\omega\left(n + \frac{1}{2}\right) \tag{3.68}$$

Example 3.4.2: Deriving the Energy Eigenvalue for the Harmonic Oscillator Potential

Let's begin with a general statement that describes the relation between a Hamiltonian operator and its energy eigenvalue:

$$\hat{H}\left|E\right\rangle = E\left|E\right\rangle \tag{3.69}$$

How can we use this to derive the result found in (3.78), giving us the answer to $\hat{a}\left|E\right\rangle$? We can use (3.61) to accomplish this:

$$\hat{H}\left(\hat{a}\left|E\right\rangle\right) = \hbar\omega\left(\underbrace{\hat{a}^\dagger\hat{a}}_{\hat{a}\hat{a}^\dagger - 1} + \frac{1}{2}\right)\left|E\right\rangle = \hat{a}\,\underbrace{\hbar\omega\left(\hat{a}^\dagger\hat{a} + \frac{1}{2}\right)}_{\hat{H}}\left|E\right\rangle - \hbar\omega\hat{a}\left|E\right\rangle = (E - \hbar\omega)\left(\hat{a}\left|E\right\rangle\right) \tag{3.70}$$

This implies that $\hat{a}\left|E\right\rangle$ is an eigenstate of $\hat{H}$ with the eigenvalue $E - \hbar\omega$. As a result, we could reapply this process indefinitely on a step-wise basis, starting with $E \to E = E - \hbar\omega$. Regardless, let's press forward and find our boundary conditions by calculating the norm of $\hat{a}\left|E\right\rangle$:

$$\langle E|\hat{a}^\dagger\hat{a}|E\rangle = \langle E|\left(\frac{\hat{H}}{\hbar\omega} - \frac{1}{2}\right)|E\rangle = \left(\frac{E}{\hbar\omega} - \frac{1}{2}\right)\cancel{\langle E|E\rangle}^{1} \tag{3.71}$$

The above implies that we would get *negative norm states* if $E < \frac{\hbar\omega}{2}$. Since there must be a $|0\rangle$, we can then infer that $E = \frac{\hbar\omega}{2}$ is our **ground state**[a] such that:

$$\hat{a}\left|E\right\rangle = \hat{a}\left|\frac{\hbar\omega}{2}\right\rangle = 0 \tag{3.72}$$

Since we have now discovered our ground state energy E_0, we can insert $\hat{a}^\dagger$ into (3.69) to find an expression for E_1:

$$\hat{H}\left(\hat{a}^\dagger\left|E\right\rangle\right) = \hbar\omega\left(\hat{a}^\dagger\hat{a} + \frac{1}{2}\right)\hat{a}^\dagger\left|E\right\rangle \tag{3.73}$$

Next, we can reapply (3.61) and factor our expression into two terms:

$$= \hat{a}^\dagger\hbar\omega\left(\hat{a}\hat{a}^\dagger + \frac{1}{2}\right)\left|E\right\rangle = \hat{a}^\dagger\hbar\omega\left(\hat{a}^\dagger\hat{a} + \frac{1}{2}\right)\left|E\right\rangle + \hat{a}^\dagger\hbar\omega\left|E\right\rangle = (E + \hbar\omega)\left(\hat{a}^\dagger\left|E\right\rangle\right) \tag{3.74}$$

This proves that $\hat{a}^\dagger\left|E\right\rangle$ is also an eigenstate of $\hat{H}$ with an eigenvalue $E + \hbar\omega$. Let's use this fact to derive the solution to an arbitrary E_n; we begin by defining a new operator:

$$\hat{N} \equiv \hat{a}^\dagger\hat{a} \tag{3.75}$$

Given that $\hat{H}\left|n\right\rangle = E_n\left|n\right\rangle$ for $n = 0, 1, 2, ...$, we have that:

$$\hat{H} = \hbar\omega\left(\hat{N} + \frac{1}{2}\right) \implies E_n = \hbar\omega\left(n + \frac{1}{2}\right) \quad \square \tag{3.76}$$

[a] Implying that anything below this is a *zero-point energy*.

Let's now see what happens when we actually apply these operators to the energy eigenstate for a harmonic oscillator potential. In short, $\hat{a}^\dagger$ and $\hat{a}$ will *raise* and *lower* an energy eigenstate $|n\rangle$ by a single quantum such that:

Applying the **creation operator**:

$$\hat{a}^\dagger |n\rangle = \sqrt{n+1}\,|n+1\rangle \tag{3.77}$$

Applying the **annihilation** operator:

$$\hat{a}\,|n\rangle = \sqrt{n}\,|n-1\rangle \tag{3.78}$$

Any energy eigenstate can be constructed by applying $\hat{a}^\dagger$ to the ground state:

$$|n\rangle = \frac{1}{\sqrt{n!}}\left(\underbrace{a^\dagger \cdots a^\dagger}_{n}\right)|0\rangle \tag{3.79}$$

Where each state $|n\rangle$ is *orthogonal*.

3.4.2 Position and Momentum

The operators $\hat{x}$ and $\hat{p}$ for position and momentum were introduced earlier in this chapter, in Section 3.3. Just as we were able to redefine the Hamiltonian $\hat{H}$ for the harmonic oscillator potential in terms of its ladder operators $\hat{a}^\dagger$ and $\hat{a}$, so can we redefine $\hat{x}$ and $\hat{p}$ such that:

Redefining the **position** operator:

$$\hat{x} = \sqrt{\frac{\hbar}{2m\omega}}\left(\hat{a}^\dagger + \hat{a}\right) \tag{3.80}$$

Redefining the **momentum** operator:

$$\hat{p} = i\sqrt{\frac{\hbar\omega m}{2}}\left(\hat{a}^\dagger - \hat{a}\right) \tag{3.81}$$

CHAPTER 4

Three Dimensional Quantum Mechanics

Introduction

We will soon be introducing a new operator: **angular momentum** $\hat{\vec{L}}$. This operator is unique with respect to all aforementioned operators, due to the fact that it cannot exist in a 1-D system. Just as classical angular momentum must occur along an orthogonal axis of rotation, so does the angular momentum operator; as a result, we must now begin working in three dimensions.

With added dimensions come new considerations – operators are now vectors, and are therefore composed of three components; how do these interact with each other now that direction is a concern? We will answer this question in this chapter by summarizing the commutation relations and eigenvalues for several operators, and discuss a few of them in following sections.

4.1 Commutation Relations and Eigenvalues

This section is mostly intended to be used as a quick reference section – for the sake of saving space, all pairs of commutation relations will only be shown in a single order. That is to say, if we have have two operator components $\hat{a}_m$ and $\hat{b}_n$, we will only be showing the commutator relation for $\left[\hat{a}_m, \hat{b}_n\right]$, and *not* for $\left[\hat{b}_n, \hat{a}_m\right]$.

The reason for this simplification is that flipping the order of commutation simply reverses the sign of the result. So given that $\left[\hat{a}_m, \hat{b}_n\right] = c$, we would automatically know that $\left[\hat{b}_n, \hat{a}_m\right] = -c$.

4.1.1 Position and Momentum

Let's begin with the 3-D position and momentum operators: $\hat{\vec{x}}$ and $\hat{\vec{p}}$ respectively (defined for one dimension in Section 2.1.2.) For the sake of accessibility, we will create a reference table:

Table 4.1: Commutation relations for the 3-D *position* and *momentum* operators.

	$\hat{p}_x$	$\hat{p}_y$	$\hat{p}_z$
$\hat{x}$	$[\hat{x}, \hat{p}_x] = i\hbar$	$[\hat{x}, \hat{p}_y] = 0$	$[\hat{x}, \hat{p}_z] = 0$
$\hat{y}$	$[\hat{y}, \hat{p}_x] = 0$	$[\hat{y}, \hat{p}_y] = i\hbar$	$[\hat{y}, \hat{p}_z] = 0$
$\hat{z}$	$[\hat{z}, \hat{p}_x] = 0$	$[\hat{z}, \hat{p}_y] = 0$	$[\hat{z}, \hat{p}_z] = i\hbar$

4.1.2 Angular Momentum

Between Components

Next, we have the three commutators for the components of **angular momentum** $\hat{\vec{L}}$:

Table 4.2: Commutation relations for the 3 components of the 3-D *angular momentum* operator.

$$[\hat{L}_x, \hat{L}_y] = i\hbar\hat{L}_z$$
$$[\hat{L}_y, \hat{L}_z] = i\hbar\hat{L}_x$$
$$[\hat{L}_z, \hat{L}_x] = i\hbar\hat{L}_y$$

Squares

It is also *very* important to note that the **square of the angular momentum** operator commutes with all components of $\hat{\vec{L}}$. That is to say:

Table 4.3: Commutation relations for the 3 components of the 3-D *angular momentum* operator and $\hat{\vec{L}}^2$.

$$[\hat{\vec{L}}^2, \hat{L}_x] = 0$$
$$[\hat{\vec{L}}^2, \hat{L}_y] = 0$$
$$[\hat{\vec{L}}^2, \hat{L}_z] = 0$$

Ladder Operators

Finally, we must take a look at the **ladder operator** for angular momentum, which is by convention centered around the z-component of $\hat{\vec{L}}$; for this reason, we define the ladder operator $\hat{L}_\pm \equiv \hat{L}_x \pm i\hat{L}_y$. In this form, it can be applied to $\hat{L}_z$ and $\hat{\vec{L}}^2$:

Table 4.4: Commutation relations for the *angular momentum ladder operator*.

$$[\hat{L}_z, \hat{L}_\pm] = \pm i\hbar\hat{L}_\pm$$
$$[\hat{\vec{L}}^2, \hat{L}_\pm] = 0$$

Eigenvalues

Given a simultaneous eigenstate $|\lambda, \mu\rangle$ for $\hat{\vec{L}}^2$ and $\hat{L}_z$, the quantum numbers $l = \{0, 1, 2, ...\}$ and $m = \{-l, -l+1, ..., l-1, l\}$, and that:

$$\hat{L}_\pm |l, m\rangle = \hbar\sqrt{l(l+1) - m(m \pm 1)}\, |l, m \pm 1\rangle \tag{4.1}$$

We have that:

Table 4.5: Eigenvalues for components of the *angular momentum operator*.

$$\begin{aligned}
\hat{\vec{L}}^2 |\lambda, \mu\rangle &= \lambda |\lambda, \mu\rangle &\implies \quad \hat{\vec{L}}^2 \hat{L}_\pm^n |\lambda, \mu\rangle &= \lambda \hat{L}_\pm^n |\lambda, \mu\rangle \\
\hat{\vec{L}}^2 |l, m\rangle &= \hbar^2 l(l+1) |l, m\rangle &\implies \quad \hat{\vec{L}}^2 \hat{L}_\pm^n |l, m\rangle &= \hbar^2 l(l+1) \hat{L}_\pm^n |l, m\rangle \\
\hat{L}_z |\lambda, \mu\rangle &= \mu |\lambda, \mu\rangle &\implies \quad \hat{L}_z \hat{L}_\pm^n |\lambda, \mu\rangle &= (\mu \pm n\hbar)\, \hat{L}_\pm^n |\lambda, \mu\rangle \\
\hat{L}_z |l, m\rangle &= \hbar m |l, m\rangle &\implies \quad \hat{L}_z \hat{L}_\pm^n |l, m\rangle &= \hbar(m+n) \hat{L}_\pm^n |l, m\rangle
\end{aligned}$$

4.1.3 Spin

Between Components

For the **spin** operator $\hat{\vec{S}}$, the commutation relation is analogous to that of angular momentum:

Table 4.6: Commutation relations for the 3 components of the 3-D *spin* operator.

$$[\hat{S}_x, \hat{S}_y] = i\hbar\hat{S}_z$$
$$[\hat{S}_y, \hat{S}_z] = i\hbar\hat{S}_x$$
$$[\hat{S}_z, \hat{S}_x] = i\hbar\hat{S}_y$$

Squares

As is the relation between its square and its components:

Table 4.7: Commutation relations for the 3 components of the 3-D *spin* operator and $\hat{\vec{S}}^2$.

$$[\hat{\vec{S}}^2, \hat{S}_x] = 0$$
$$[\hat{\vec{S}}^2, \hat{S}_y] = 0$$
$$[\hat{\vec{S}}^2, \hat{S}_z] = 0$$

Ladder Operators

The **ladder operator** for spin follows the same conventions as that of angular momentum; we define the ladder operator $\hat{S}_\pm \equiv \hat{S}_x \pm i\hat{S}_y$. It can be applied to $\hat{S}_z$ and $\hat{\vec{S}}^2$:

Table 4.8: Commutation relations for the *spin ladder operator*.

$$[\hat{S}_z, \hat{S}_\pm] = \pm i\hbar\hat{S}_\pm$$
$$[\hat{\vec{S}}^2, \hat{S}_\pm] = 0$$

Eigenvalues

Given a simultaneous eigenstate $|\lambda, \mu\rangle$ for $\hat{\vec{S}}^2$ and $\hat{S}_z$, the quantum numbers $s = \left\{0, \frac{1}{2}, 1, \frac{3}{2}, ...\right\}$ and $m = \{-s, -s+1, ..., s-1, s\}$, and that:

$$\hat{S}_\pm |s, m\rangle = \hbar\sqrt{s(s+1) - m(m\pm 1)}\, |s, m\pm 1\rangle \tag{4.2}$$

We have that:

Table 4.9: Eigenvalues for components of the *spin operator*.

$$\hat{\vec{S}}^2 |\lambda, \mu\rangle = \lambda |\lambda, \mu\rangle \implies \hat{\vec{S}}^2 \hat{S}_\pm^n |\lambda, \mu\rangle = \lambda \hat{S}_\pm^n |\lambda, \mu\rangle$$
$$\hat{\vec{S}}^2 |s, m\rangle = \hbar^2 s(s+1) |s, m\rangle \implies \hat{\vec{S}}^2 \hat{S}_\pm^n |s, m\rangle = \hbar^2 s(s+1) \hat{S}_\pm^n |s, m\rangle$$
$$\hat{S}_z |\lambda, \mu\rangle = \mu |\lambda, \mu\rangle \implies \hat{S}_z \hat{S}_\pm^n |\lambda, \mu\rangle = (\mu \pm n\hbar)\, \hat{S}_\pm^n |\lambda, \mu\rangle$$
$$\hat{S}_z |s, m\rangle = \hbar m |s, m\rangle \implies \hat{S}_z \hat{S}_\pm^n |s, m\rangle = \hbar(m+n) \hat{S}_\pm^n |s, m\rangle$$

4.1.4 Total Angular Momentum

Between Components

The **total angular momentum** of a state is defined as $\hat{\vec{J}} = \hat{\vec{L}} + \hat{\vec{S}}$, and follows the same pattern as angular momentum and spin:

Table 4.10: Commutation relations for the 3 components of the 3-D *total angular momentum* operator.

$$[\hat{J}_x, \hat{J}_y] = i\hbar\hat{J}_z$$
$$[\hat{J}_y, \hat{J}_z] = i\hbar\hat{J}_x$$
$$[\hat{J}_z, \hat{J}_x] = i\hbar\hat{J}_y$$

Squares

Once more, this applies to the square of $\hat{\vec{J}}$ and its components:

Table 4.11: Commutation relations for the 3 components of the 3-D *total angular momentum* operator and $\hat{\vec{J}}^2$.

$$[\hat{\vec{J}}^2, \hat{J}_x] = 0$$
$$[\hat{\vec{J}}^2, \hat{J}_y] = 0$$
$$[\hat{\vec{J}}^2, \hat{J}_z] = 0$$

Ladder Operators

And finally, the **ladder operator** for total angular momentum continues to behave as the two aforementioned ladder operators; we define the ladder operator $\hat{J}_\pm \equiv \hat{J}_x \pm i\hat{J}_y$. In this form, it can be applied to $\hat{J}_z$ and $\hat{\vec{J}}^2$:

Table 4.12: Commutation relations for the *total angular momentum ladder operator*.

$$[\hat{J}_z, \hat{J}_\pm] = \pm i\hbar\hat{J}_\pm$$
$$[\hat{\vec{J}}^2, \hat{J}_\pm] = 0$$

Eigenvalues

Given a simultaneous eigenstate $|\lambda, \mu\rangle$ for $\hat{\vec{J}}^2$ and $\hat{J}_z$, the quantum numbers $j = \left\{0, \frac{1}{2}, 1, \frac{3}{2}, ...\right\}$ and $m = \{-j, -j+1, ..., j-1, j\}$, and that:

$$\hat{J}_\pm |j, m\rangle = \hbar\sqrt{j(j+1) - m(m \pm 1)}\, |j, m \pm 1\rangle \tag{4.3}$$

We have that:

Table 4.13: Eigenvalues for components of the *total angular momentum operator*.

$$\hat{\vec{J}}^2 |\lambda, \mu\rangle = \lambda |\lambda, \mu\rangle \implies \hat{\vec{J}}^2 \hat{J}_\pm^n |\lambda, \mu\rangle = \lambda \hat{J}_\pm^n |\lambda, \mu\rangle$$
$$\hat{\vec{J}}^2 |j, m\rangle = \hbar^2 j(j+1) |j, m\rangle \implies \hat{\vec{J}}^2 \hat{J}_\pm^n |j, m\rangle = \hbar^2 j(j+1) \hat{J}_\pm^n |j, m\rangle$$
$$\hat{J}_z |\lambda, \mu\rangle = \mu |\lambda, \mu\rangle \implies \hat{J}_z \hat{J}_\pm^n |\lambda, \mu\rangle = (\mu \pm n\hbar)\, \hat{J}_\pm^n |\lambda, \mu\rangle$$
$$\hat{J}_z |j, m\rangle = \hbar m |j, m\rangle \implies \hat{J}_z \hat{J}_\pm^n |j, m\rangle = \hbar(m+n) \hat{J}_\pm^n |j, m\rangle$$

4.2 Angular Momentum

Introduction

Although we will begin to explain the processes underlying quantum angular momentum through analogies from classical mechanics, the truth of the situation is that the reverse would more accurately reflect reality. Regardless, we are taught to understand macroscopic phenomena before quantum phenomena, so we will go ahead and start with the following comparison – we are all familiar with angular momentum in the classical sense:

$$\vec{L} = \vec{r} \times \vec{p} \tag{4.4}$$

Where $\vec{L}$ is the classical angular momentum of an object at a position $\vec{r}$ relative to the origin, with a momentum $\vec{p}$.

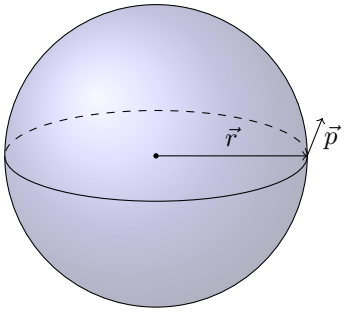


Figure 4.1: Visually representing $\vec{r}$ and a possible $\vec{p}$ for a sphere.

In quantum mechanics, we define the 3-D **angular momentum** $\hat{\vec{L}}$ separately for its three components, in the same way:

$$\hat{\vec{L}} = \hat{\vec{x}} \times \hat{\vec{p}} \tag{4.5}$$

Which gives us the three components:

$$\hat{L}_x = \hat{y}\hat{p}_z - \hat{z}\hat{p}_y$$

$$\hat{L}_y = \hat{z}\hat{p}_x - \hat{x}\hat{p}_z$$

$$\hat{L}_z = \hat{x}\hat{p}_y - \hat{y}\hat{p}_x$$

Example 4.2.1: Deriving a Commutation Relation

As defined in Section 4.1.2, we know that:

$$\left[\hat{L}_x, \hat{L}_y\right] = i\hbar\hat{L}_z \tag{4.6}$$

Let's begin by using (4.5) and the definition of a commutation relation (1.85) on p.23:

$$\left[\hat{L}_x, \hat{L}_y\right] = \hat{L}_x\hat{L}_y - \hat{L}_y\hat{L}_z \tag{4.7}$$

Things get hairy quickly:

$$= (\hat{y}\hat{p}_z - \hat{z}\hat{p}_y)(\hat{z}\hat{p}_x - \hat{x}\hat{p}_z) - (\hat{z}\hat{p}_x - \hat{x}\hat{p}_z)(\hat{y}\hat{p}_z - \hat{z}\hat{p}_y) \tag{4.8}$$

Especially when we expand our products:

$$= \hat{y}\hat{p}_z\hat{z}\hat{p}_x - \hat{y}\hat{p}_z\hat{x}\hat{p}_z + \hat{z}\hat{p}_y\hat{x}\hat{p}_z - \hat{z}\hat{p}_y\hat{z}\hat{p}_x - \hat{z}\hat{p}_x\hat{y}\hat{p}_z + \hat{z}\hat{p}_x\hat{z}\hat{p}_y - \hat{x}\hat{p}_z\hat{z}\hat{p}_y + \hat{x}\hat{p}_z\hat{y}\hat{p}_z \tag{4.9}$$

Let's reorder the above:

$$= (\hat{x}\hat{p}_z\hat{y}\hat{p}_z - \hat{y}\hat{p}_z\hat{x}\hat{p}_z) + (\hat{y}\hat{p}_z\hat{z}\hat{p}_x - \hat{z}\hat{p}_x\hat{y}\hat{p}_z) + (\hat{z}\hat{p}_x\hat{z}\hat{p}_y - \hat{z}\hat{p}_y\hat{z}\hat{p}_x) + (\hat{z}\hat{p}_y\hat{x}\hat{p}_z - \hat{x}\hat{p}_z\hat{z}\hat{p}_y) \tag{4.10}$$

You may have noticed that we have created a series of commutation relations! Compressing our notation makes it obvious that some of these pairs are commutators, allowing use to eliminate half of our expression:

$$= \cancelto{0}{[\hat{x}\hat{p}_z, \hat{y}\hat{p}_z]} + [\hat{y}\hat{p}_z, \hat{z}\hat{p}_x] + \cancelto{0}{[\hat{z}\hat{p}_x, \hat{z}\hat{p}_y]} + [\hat{z}\hat{p}_y, \hat{x}\hat{p}_z] \tag{4.11}$$

Next, let's rewrite the above in expanded notation:

$$= (\hat{y}\hat{p}_z\hat{z}\hat{p}_x - \hat{z}\hat{p}_x\hat{y}\hat{p}_z) + (\hat{z}\hat{p}_y\hat{x}\hat{p}_z - \hat{x}\hat{p}_z\hat{z}\hat{p}_y) \tag{4.12}$$

We can now factor out any position-momentum products that commute with all other elements:

$$= \hat{y}\hat{p}_x(\hat{p}_z\hat{z} - \hat{z}\hat{p}_z) + \hat{p}_y\hat{x}(\hat{z}\hat{p}_z - \hat{p}_z\hat{z}) \tag{4.13}$$

We can rewrite these in commutator form, using the known commutation relations for position and momentum:

$$= \hat{y}\hat{p}_x[\hat{p}_z, \hat{z}] + \hat{p}_y\hat{x}[\hat{z}, \hat{p}_z] = -\hat{y}\hat{p}_x i\hbar + \hat{p}_y\hat{x}i\hbar \tag{4.14}$$

Finally, factoring out our expression gives us our original equation:

$$i\hbar(\hat{p}_y\hat{x} - \hat{y}\hat{p}_x) = i\hbar\hat{L}_z \quad \square \tag{4.15}$$

4.2.1 Squared Operator

It is important to note that the result of the commutation relations for angular momentum imply that $\hat{L}_x$, $\hat{L}_y$, and $\hat{L}_z$ do *not* share a common set of eigenstates, and therefore *cannot* have simultaneous sharp values.

On the flipside, there does exist another operator that commutes with any component of $\hat{\vec{L}}$ – its square $\hat{\vec{L}}^2$. We can define the square in two ways:

$$\hat{\vec{L}}^2 = \hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2 = \hat{L}_x\hat{L}_x + \hat{L}_y\hat{L}_y + \hat{L}_z\hat{L}_z \tag{4.16}$$

Example 4.2.2: Proving the Commutability of $\hat{\vec{L}}^2$ with the components of $\hat{\vec{L}}$

We are interested in proving the following:

$$\left[\hat{\vec{L}}^2, \hat{L}_z\right] = 0 \tag{4.17}$$

To do so, we can use (4.16) as follows:

$$\left[\hat{\vec{L}}^2, \hat{L}_z\right] = \left[\hat{L}_x^2, \hat{L}_z\right] + \left[\hat{L}_y^2, \hat{L}_z\right] + \cancelto{0}{\left[\hat{L}_z^2, \hat{L}_z\right]} \tag{4.18}$$

Via the associativity property of commutation relations[a] we have that:

$$= \hat{L}_x\left[\hat{L}_x, \hat{L}_z\right] + \left[\hat{L}_x, \hat{L}_z\right]\hat{L}_x + \hat{L}_y\left[\hat{L}_y, \hat{L}_z\right] + \left[\hat{L}_y, \hat{L}_z\right]\hat{L}_y \tag{4.19}$$

Plugging in our known commutation relations between components of $\hat{\vec{L}}$ shows that our initial assumption was correct:

$$= \hat{L}_x\left(-i\hbar\hat{L}_y\right) + \left(-i\hbar\hat{L}_y\right)\hat{L}_x + \hat{L}_y\left(i\hbar\hat{L}_x\right) + \left(i\hbar\hat{L}_x\right)\hat{L}_y = 0 \quad \square \tag{4.20}$$

[a] This is originally mentioned in (1.87), on p.24.

4.2.2 Ladder Operators

The **angular momentum ladder operator** $\hat{L}_\pm$ is by its very nature a relative quantity; as such, it is necessary to rely upon convention. In this case, convention dictates that it can only operate on $\hat{L}_z$, which means we must use the following definition:

$$\hat{L}_\pm = \hat{L}_x \pm i\hat{L}_y \tag{4.21}$$

The ladder operator can be represented in several ways, which can be summarize as follows:

$$\hat{L}_- = \left(\hat{L}_+\right)^\dagger = \left(\hat{L}_x + i\hat{L}_y\right)^\dagger = \left(\hat{L}_x\right)^\dagger - \left(i\hat{L}_y\right)^\dagger = \hat{L}_x - i\hat{L}_y \tag{4.22}$$

Example 4.2.3: Proving Another Commutation Relation

In Section 4.1.2, we saw that:

$$\left[\hat{L}_z, \hat{L}_+\right] = \hbar \hat{L}_+ \tag{4.23}$$

This is relatively straightforward to prove, as we can use the distributivity of commutation relations[a] and the definition of $\hat{L}_+$ to expand our left-hand side:

$$\left[\hat{L}_z, \hat{L}_+\right] = \left[\hat{L}_z, \hat{L}_x + i\hat{L}_y\right] = \left[\hat{L}_z, \hat{L}_x\right] + i\left[\hat{L}_z, \hat{L}_y\right] \tag{4.24}$$

Next, we can use the commutation relations for our two terms, which proves our initial assumption:

$$= i\hbar\hat{L}_y + i\left(-i\hbar\hat{L}_x\right) = \hbar\left(\hat{L}_x + i\hat{L}_y\right) = \hbar\hat{L}_+ \tag{4.25}$$

[a] This is originally introduced in (1.88) on p.24.

We've already discussed the basics regarding **eigenvalues** throughout this compendium, but something we haven't discussed is the effect of our ladder operator $\hat{L}_\pm$ on the eigenvalues of $\hat{\vec{L}}^2$ and $\hat{L}_z$. Before we begin, we should recall an important property: if two operators are commutators, they will have a *common* and *complete* set of eigenstates. For $\hat{\vec{L}}^2$ and $\hat{L}_z$, we know that:

$$\left[\hat{\vec{L}}^2, \hat{L}_z\right] = 0 \tag{4.26}$$

This implies that we can construct a common set of eigenstates $|\lambda, \mu\rangle$ such that:

$$\hat{\vec{L}}^2 |\lambda, \mu\rangle = \lambda |\lambda, \mu\rangle \tag{4.27}$$

And:

$$\hat{L}_z |\lambda, \mu\rangle = \mu |\lambda, \mu\rangle \tag{4.28}$$

Where λ and μ are the *eigenvalues* for $\hat{\vec{L}}^2$ and $\hat{L}_z$ respectively.

Let's now introduce a new state involving the raising operator, and take a look at a few examples:

$$\hat{L}_+ |\lambda, \mu\rangle \tag{4.29}$$

Example 4.2.4: Ladder Operators 1

Let's find out what happens when we evaluate $\hat{\vec{L}}^2\hat{L}_+ |\lambda,\mu\rangle$. We know from previous experience that:

$$\left[\hat{\vec{L}}^2, \hat{L}_\pm\right] = 0 \tag{4.30}$$

This implies that $\hat{\vec{L}}^2\hat{L}_\pm = \hat{L}_\pm\hat{\vec{L}}^2$. We can use this information to our benefit:

$$\hat{\vec{L}}^2\hat{L}_+ |\lambda,\mu\rangle = \hat{L}_+\hat{\vec{L}}^2 |\lambda,\mu\rangle = \lambda\hat{L}_+ |\lambda,\mu\rangle \tag{4.31}$$

We can therefore conclude that the ladder operator *does not affect* $\hat{\vec{L}}^2$.

Example 4.2.5: Ladder Operators 2

Let's now evaluate the following expression, and find its eigenvalues:

$$\hat{L}_z\hat{L}_+ |\lambda,\mu\rangle \tag{4.32}$$

These operators do not commute, and instead behave according to $\left[\hat{L}_z, \hat{L}_\pm\right] = \pm\hbar\hat{L}_\pm$. This implies the following:

$$\hat{L}_z\hat{L}_+ - \hat{L}_+\hat{L}_z = \hbar\hat{L}_+ \iff \hat{L}_z\hat{L}_+ = \hat{L}_+\hat{L}_z + \hbar\hat{L}_+ \tag{4.33}$$

Let's see what happens when we use the above conclusion in our expression:

$$\hat{L}_z\hat{L}_+ |\lambda,\mu\rangle = \hat{L}_+\hat{L}_z |\lambda,\mu\rangle + \hbar\hat{L}_+ |\lambda,\mu\rangle = \hat{L}_+\mu |\lambda,\mu\rangle + \hbar\hat{L}_+ |\lambda,\mu\rangle \tag{4.34}$$

After some algebraic manipulation, we find our result:

$$= \mu\hat{L}_+ |\lambda,\mu\rangle + \hbar\hat{L}_+ |\lambda,\mu\rangle = (\mu+\hbar)\hat{L}_+ |\lambda,\mu\rangle \tag{4.35}$$

4.2.3 Quantum Numbers

Now that we've understood the *mathematical* process by which one can derive the eigenvalues of our various operators, we need to start setting some *physical* restrictions on our systems. One such restriction that must be taken into account is that of *normalization* – in particular, it is vital that our normalized operator-eigenstate pairs be positive. Let's begin by normalizing $\hat{L}_\pm |\lambda,\mu\rangle$:

$$\text{norm}\left(\hat{L}_\pm |\lambda,\mu\rangle\right) = \langle\lambda,\mu|\hat{L}_+^\dagger\hat{L}_+|\lambda,\mu\rangle = \langle\lambda,\mu|\hat{L}_-\hat{L}_+|\lambda,\mu\rangle \tag{4.36}$$

We will need to evaluate $\hat{L}_-\hat{L}_+$ before continuing:

$$\hat{L}_-\hat{L}_+ = \left(\hat{L}_x - i\hat{L}_y\right)\left(\hat{L}_x + i\hat{L}_y\right) = \hat{L}_x^2 + \hat{L}_y^2 + i\cancel{\left[\hat{L}_x, \hat{L}_y\right]}^{\hbar\hat{L}_z} = \hat{\vec{L}}^2 - \hat{L}_z^2 + i\hbar\hat{L}_z \tag{4.37}$$

Let's plug this back into our original equation:

$$\langle \lambda, \mu | \hat{L}_- \hat{L}_+ | \lambda, \mu \rangle = \langle \lambda, \mu | \left(\hat{\vec{L}}^2 - \hat{L}_z^2 + i\hbar \hat{L}_z \right) | \lambda, \mu \rangle \tag{4.38}$$

We can then use our eigenvalue relations and replace our three operator-eigenstate pairs:

$$= \left(\lambda - \mu^2 - \hbar\mu \right) \cancelto{1}{\langle \lambda, \mu | \lambda, \mu \rangle} = \lambda - \mu^2 - \hbar\mu \geq 0 \tag{4.39}$$

Notice a problem? In short, if our norm must always evaluate to zero or greater, there will always exist a μ for any λ such that $\left(\lambda - \mu^2 - \hbar\mu \right) < 0$. This is why we need to implement a restriction for μ in terms of λ – in particular, it is necessary that:

$$\lambda - \mu^2 - \hbar\mu \geq 0 \iff \lambda \geq \mu(\mu + \hbar) \tag{4.40}$$

Now we must consider a very important detail; the above expression does not imply the existence of a single boundary, but rather the need for *two boundary conditions*. We will call these $\mu_{\min}$ and $\mu_{\max}$, and they will have to fulfill the following:

$$\lambda \geq \mu_{\min}(\mu_{\min} + \hbar) = \mu_{\max}(\mu_{\max} + \hbar) \tag{4.41}$$

There are two *mathematically possible* solutions to this boundary pair, but only one *physically sound* solution. The solution we are interested in is:

$$\mu_{\max} = -\mu_{\min} \tag{4.42}$$

We might choose to stop here, but there is another important consideration we must make: we are dealing with quantized systems, whereby the application of a ladder operator cannot occur continuously. As a result, we will need to begin parameterizing our eigenvalues to more accurately mirror reality. We will set up the following:

Table 4.14: Defining our *azimuthal* quantum number l and *magnetic* quantum number m for angular momentum.

$$\begin{aligned} \mu &\equiv m\hbar \\ m_{\max} &\equiv l \\ m_{\min} &= -l \\ \mu_{\max} &= l\hbar \\ \mu_{\min} &= -l\hbar \end{aligned}$$

With these new parameterizations, we can recalculate our eigenvalues λ and μ:

$$\lambda = \mu_{\max} \left(\mu_{\max} + \hbar \right) = l\hbar \left(l\hbar + \hbar \right) = \hbar^2 l \left(l + 1 \right) \tag{4.43}$$

$$\mu = m\hbar \tag{4.44}$$

Now that we've redefined our eigenvalues in terms of our two quantum numbers, we should get back to the root of the issue – namely, the quantization brought upon by the ladder operator.

The situation is as follows: since $\mu_{\max}$ should be reached by $\mu_{\min}$ through the application of $\hat{L}_+$ an integer N number of times. This process is shown below:

$$\mu_{\max} = \mu_{\min} + \hbar N \iff \hbar l = -\hbar l + \hbar N \implies l = \frac{N}{2} \tag{4.45}$$

And so l is able to take *half-integer values*, as summarized below:

$$l = \left\{0, \frac{1}{2}, 1, \frac{3}{2}, ...\right\} \tag{4.46}$$

Finally, we must find the possible values for m; we already know that $m_{\min} = -l$ and $m_{\max} = l$, so we must simply consider the values in between – how do these behave? To figure this out, all we need to consider is the behavior of the eigenvalue μ in terms of m, as we apply the $\hat{L}_+$ operator:

$$\hat{L}_z \left|\lambda, \mu\right\rangle = \mu \left|\lambda, \mu\right\rangle \implies \hat{L}_z \left|l, m\right\rangle = \hbar m \left|l, m\right\rangle \tag{4.47}$$

$$\hat{L}_z \hat{L}_+ \left|\lambda, \mu\right\rangle = (\mu + \hbar) \hat{L}_+ \left|\lambda, \mu\right\rangle \implies \hat{L}_z \hat{L}_+ \left|l, m\right\rangle = \hbar(m+1) \hat{L}_+ \left|l, m\right\rangle \tag{4.48}$$

As we clearly see, our eigenvalue m always increases by a step size of 1, this means that m can have the following values:

$$m = \{-l, -l+1, ..., l-1, l\} \tag{4.49}$$

4.3 Spherical Angular Momentum

Introduction

So far, we've spent a portion of our time examining 1-D and 3-D *Cartesian* systems, but this does not lend itself well to certain properties; in particular, angular momentum and spin are far more naturally described as functions of[1] angle φ and azimuth θ, rather than x, y, and z. How these are related is shown in a visually intuitive manner in Figure 4.2:

[1] These are independent of the distance from the origin: r.

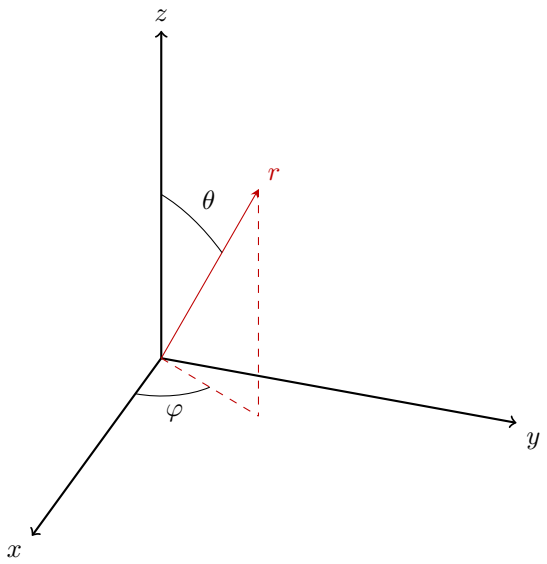


Figure 4.2: Converting from a *Cartesian* to a *spherical* coordinate system.

To perform such a conversion mathematically, we can use the following relations:

$$\begin{aligned} r &= \sqrt{x^2+y^2+z^2} & x &= r\sin\theta\cos\varphi \\ \varphi &= \arctan\left(\frac{y}{x}\right) & y &= r\sin\theta\sin\varphi \\ \theta &= \arccos\left(\frac{z}{r}\right) & z &= r\cos\theta \end{aligned}$$

We will also need a way to describe our previously Cartesian gradient $\nabla\vec{x}$ in terms of r, φ, and θ:

$$\begin{aligned} \frac{\partial}{\partial x} &= \sin\theta\cos\varphi\frac{\partial}{\partial r} + \frac{\cos\theta\cos\varphi}{r}\frac{\partial}{\partial\theta} - \frac{\sin\varphi}{r\sin\theta}\frac{\partial}{\partial\varphi} \\ \frac{\partial}{\partial y} &= \sin\theta\sin\varphi\frac{\partial}{\partial r} + \frac{\cos\theta\sin\varphi}{r}\frac{\partial}{\partial\theta} + \frac{\cos\varphi}{r\sin\theta}\frac{\partial}{\partial\varphi} \\ \frac{\partial}{\partial z} &= \cos\theta\frac{\partial}{\partial r} - \frac{\sin\theta}{r}\frac{\partial}{\partial\theta} \end{aligned}$$

4.3.1 Derivation of $\hat{L}_z$

Now that we've laid the groundwork for our new coordinate system, we should start investigating how the angular momentum eigenstate $|l, m\rangle$ for $\hat{L}_z$ and $\hat{\vec{L}}^2$ is represented in terms of φ, and θ. As a result, we will assume that it is possible to represent such a state as follows:

$$|l, m\rangle = f(\theta, \varphi) \tag{4.50}$$

Let's focus on $\hat{L}_z$ for now, and use a combination of (4.5), (3.49), and (3.50) to create the following:

$$\hat{L}_z = \hat{x}\hat{p}_y\hat{y}\hat{p}_x = \frac{\hbar}{i}\left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right) \tag{4.51}$$

Using spherical coordinates implies the following:

$$\begin{aligned} = \frac{\hbar}{i}\Bigg(& r\sin\theta\cos\varphi\left(\cancel{\sin\theta\sin\varphi\frac{\partial}{\partial r}} + \cancel{\frac{\cos\theta\sin\varphi}{r}\frac{\partial}{\partial\theta}} + \frac{\cos\varphi}{r\sin\theta}\frac{\partial}{\partial\varphi}\right) \\ -\ & r\sin\theta\sin\varphi\left(\cancel{\sin\theta\cos\varphi\frac{\partial}{\partial r}} + \cancel{\frac{\cos\theta\cos\varphi}{r}\frac{\partial}{\partial\theta}} - \frac{\sin\varphi}{r\sin\theta}\frac{\partial}{\partial\varphi}\right)\Bigg) \end{aligned} \tag{4.52}$$

Fortunately, the majority of the expression cancels out, and leaves us with a simpler result:

$$= \frac{\hbar}{i}\left(\cancel{r\sin\theta}\cos\varphi\frac{\cos\varphi}{\cancel{r\sin\theta}}\frac{\partial}{\partial\varphi} - \frac{\cancel{r\sin\theta}\sin\varphi\sin\varphi}{\cancel{r\sin\theta}}\frac{\partial}{\partial\varphi}\right) = \frac{\hbar}{i}\frac{\partial}{\partial\varphi}\left(\cancel{\cos^2\varphi + \sin^2\varphi}^{\,1}\right) \tag{4.53}$$

Once again, we can remove a significant portion of our expression. With a little bit of trigonometry, we find our result:

The *spherical* **angular momentum** operator $\hat{L}_z$:

$$\hat{L}_z = \frac{\hbar}{i}\frac{\partial}{\partial\varphi} \tag{4.54}$$

4.3.2 Spherical Harmonics

We will now introduce a function that can be plugged into (4.50) called **Laplace's spherical harmonics**:

The *eigenstate* of $\hat{L}_z$ and $\hat{\vec{L}}^2$:

$$|l,m\rangle = Y_l^m(\theta,\varphi) \tag{4.55}$$

With **Laplace's spherical harmonics** as defined by:

$$Y_l^m(\theta,\varphi) \equiv \Theta(\theta)e^{im\varphi} \tag{4.56}$$

Where the **azimuthal** component is:

$$\Theta(\theta) \equiv A_l^m P_l^m(\cos\theta) \tag{4.57}$$

And the **associated Legendre polynomial**:

$$P_l^m(x) \equiv \left(1-x^2\right)^{\frac{|m|}{2}} \left(\frac{d}{dx}\right)^{|m|} P_l(x) \tag{4.58}$$

Defined by the **Legendre polynomial**:

$$P_l(x) \equiv \frac{1}{2^l l!}\left(\frac{d}{dx}\right)^l \left(x^2-l\right)^l \tag{4.59}$$

Spherical harmonics can be used to represent *any* function on the sphere such that:

$$f(\theta,\varphi) = \sum_{l=0}^{\infty}\sum_{m=-l}^{l} a_l^m Y_l^m(\theta,\varphi) \tag{4.60}$$

Derivation Process

Separation of Variables In Section 4.1.2, we introduced the eigenvalue for $\hat{L}_z$; we can replace our eigenstate with a spherical harmonic:

$$\hat{L}_z|l,m\rangle = \hbar m|l,m\rangle \implies \hat{L}_z Y_l^m(\theta,\varphi) = \hbar m Y_l^m(\theta,\varphi) \tag{4.61}$$

We can then plug in the spherical coordinate value for $\hat{L}_z$:

$$\iff \frac{\hbar}{i}Y_l^m(\theta,\varphi) = \hbar m Y_l^m(\theta,\varphi) \implies Y_l^m(\theta,\varphi) = imY_l^m(\theta,\varphi) \tag{4.62}$$

Y_l^m remains unchanged when multiplied by a factor im meaning it is **single-valued**. As a result, it can be separated into a product of two independent factors:

$$Y_l^m(\theta,\varphi) = \Theta_l^m(\theta)e^{im\varphi} \tag{4.63}$$

The significance of the above is made clear when we see the manifestation of single-valuedness stemming from e:

$$Y_l^m(\theta,\varphi) = Y_l^m(\theta,\varphi+2\pi) \tag{4.64}$$

This makes sense, because we constructed the above with the following condition in mind:

$$e^{im2\pi} = 1 \text{ because } e^{im(\varphi+2\pi)} = e^{im\varphi} \tag{4.65}$$

Finally, this implies that m must be an *integer* – a conclusion we already made through a separate process in (4.49).

Legendre's Equation For the next step of the process, we will need to define $\hat{\vec{L}}^2$ in spherical coordinates and use its eigenvalue-eigenstate relation as we did in the previous step:

$$\hat{\vec{L}}^2 = -\hbar^2\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\varphi^2}\right] \tag{4.66}$$

We introduced the following eigenvalue in Section 4.1.2:

$$\hat{\vec{L}}^2|l,m\rangle = \hbar^2 l(l+1)|l,m\rangle \tag{4.67}$$

Combining the above with (4.66) and (4.63) gives us:

$$-\hbar^2\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\varphi^2}\right]\Theta_l^m(\theta)e^{im\varphi} = \hbar^2 l(l+1)\Theta_l^m(\theta)e^{im\varphi} \tag{4.68}$$

Since the only part of the equation that depends on φ is $e^{im\varphi}$, we can take the derivative on the left such that:

$$-\hbar^2\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) - \frac{m^2}{\sin^2\theta}\right]\Theta_l^m(\theta)\cancel{e^{im\varphi}} = \hbar^2 l(l+1)\Theta_l^m(\theta)\cancel{e^{im\varphi}} \tag{4.69}$$

This allows us to cancel out some parts of the remaining expression. Before continuing, let's redefine some variables such that:

$$x = \cos\theta \implies dx = -\sin\theta d\theta \tag{4.70}$$

It is *very important* to note that x does not represent **position** in Cartesian coordinates! It is simply representative of the quantity defined above, and will be very useful when plugged in. Let's redefine a few components of (4.69) in terms of x before continuing:

$$\frac{\partial}{\partial\theta} = \frac{\partial x}{\partial\theta}\frac{\partial}{\partial x} = -\sin\theta\frac{\partial}{\partial x} \implies \frac{1}{\sin\theta}\frac{\partial}{\partial\theta} = -\frac{\partial}{\partial x} \tag{4.71}$$

Let input a few more terms from (4.69) into the above expression:

$$\implies -\frac{\partial}{\partial x}\left(\sin\theta\frac{\partial}{\partial\theta}\right) = \frac{\partial}{\partial x}\left(\frac{\partial}{\partial x}\right) = \frac{\partial}{\partial x}\left(\left(1-x^2\right)\frac{\partial}{\partial x}\right) \tag{4.72}$$

Plugging everything back into (4.69) then leaves us with **Legendre's equation**:

$$\left[\frac{\partial}{\partial x}\left(\left(1-x^2\right)\frac{\partial}{\partial x}\right) - \frac{m^2}{1-x^2} + l(l+1)\right]\Theta_l^m(x) = 0 \tag{4.73}$$

The solutions to Legendre's equation are the **associated Legendre polynomials**[2] multiplied by a factor dependent on l and m:

$$\Theta_l^m(x) = A_l^m P_l^m(x) \tag{4.74}$$

In conclusion, there exists a *constant* A_l^m for each individual Y_l^m such that:

$$Y_l^m(\theta,\varphi) = A_l^m P_l^m(\cos\theta)e^{im\varphi} \tag{4.75}$$

Finding $\mathbf{A_l^m}$ We are nearly finished with this process, as all that is left is to explain how one may find A_l^m – this process is *conceptually* straightforward, but can be *algebraically* difficult depending on the system. In short, A_l^m is a *normalization constant*, which means we must simply take the *norm* of our eigenstate to find it:

$$\langle l,m|l',m'\rangle = \int_0^\pi d\theta\sin\theta\int_0^{2\pi} d\varphi\left(Y_l^m(\theta,\varphi)\right)^* Y_l^m(\theta,\varphi) = \delta_{ll'}\delta_{mm'} \tag{4.76}$$

Solving the above for a particular system can then yield A_l^m, which are determined such that $\langle l,m|l',m'\rangle = \delta_{ll'}\delta_{mm'}$ holds.

Example 4.3.1: A Few Spherical Harmonics

Here are a few examples that may come in handy:

$$Y_0^0 = \frac{1}{\sqrt{4\pi}} \tag{4.77}$$

$$Y_1^0 = \sqrt{\frac{3}{4\pi}}\cos\theta \tag{4.78}$$

For any Y_l^0 (ie when $m=0$), our spherical harmonic will only depend on θ.

$$Y_1^{\pm1} = \mp\sqrt{\frac{3}{8\pi}}\sin\theta e^{\pm i\varphi} \tag{4.79}$$

[2] These are defined in (4.58) on p.70.

4.4 Central-Symmetric Potentials

Introduction

Of the endless number of potentials that exist in our universe, one of the most fundamental are the **central-symmetric potential**; these manifest themselves in all facets of physics – be it an electric field generated by a point charge, or a gravitational field generated by a celestial object, the central-symmetric potential is a versatile and useful tool that also has applications within quantum mechanics.

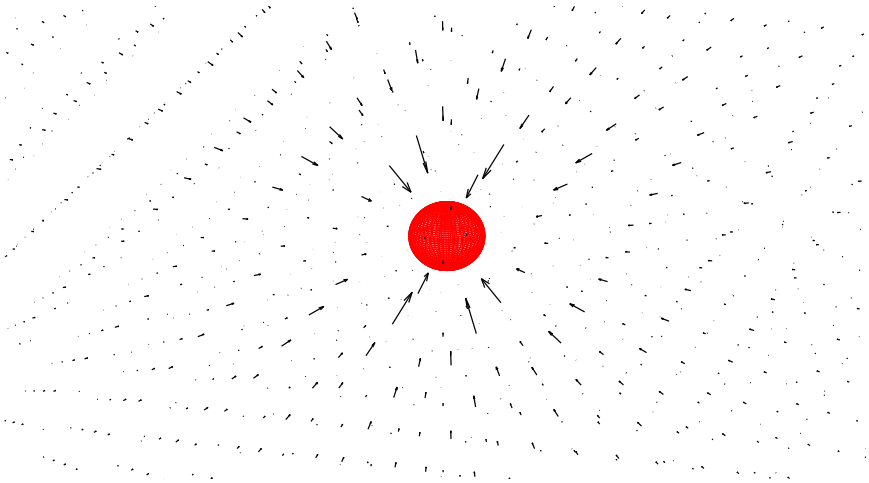


Figure 4.3: The gradient of a **central symmetric potential**.

To properly delve into the quantum central-symmetric potential, we must once again turn our eyes to the Hamiltonian of our new system, and investigate the *TISE*.

4.4.1 The Central-Symmetric Radial Equation

In Section 4.3.2, we separated our *angular momentum* into factors dependent on φ and θ. In that case, we were able to disregard the distance from the origin r since angular momentum doesn't depend on this – now, we must do the opposite. Since the Hamiltonian is based upon a *central-symmetric* potential, that means it can *only* depend on r:

For a central symmetric potential, the eigenvalue of the system's Hamiltonian (ie its **total energy** and **TISE**) is given by the **radial equation for central symmetric potentials**:

$$\left[-\frac{\hbar^2}{2M}\frac{\partial^2}{\partial r^2} + V(r) + \underbrace{\frac{\hbar^2}{2Mr^2}l(l+1)}_{\text{centrifugal barrier}}\right] U(r) = EU(r) \tag{4.80}$$

4.4.2 Derivation of the Radial Equation

Preparation

To understand the origin of (4.80), we must begin with an expression for the Hamiltonian:

$$\hat{H} = \frac{\hat{\vec{p}}^2}{2M} + V(r) = -\frac{\hbar^2}{2M}\nabla^2 + V(r) \tag{4.81}$$

Next, we can plug in the expression for the *gradient* in *spherical coordinates*:

$$= -\frac{\hbar^2}{2M}\left(\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{1}{r^2}\underbrace{\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta} + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\varphi^2}\right]}_{-\frac{\hat{\vec{L}}^2}{\hbar^2}}\right) + V(r) \tag{4.82}$$

Let's rewrite the above in a more simple manner; this will allow us to separate our equation into a *radial* and *angular* part:

$$= \underbrace{-\frac{\hbar^2}{2M}\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + V(r)}_{\text{only acts on } r} + \frac{\hat{\vec{L}}^2}{2Mr^2} \tag{4.83}$$

Since our first two terms only act on r, that means that we can derive a useful set of commutations:

$$\begin{gathered}\left[\hat{H}, \hat{\vec{L}}^2\right] = 0 \text{ and } \left[\hat{\vec{L}}^2, \hat{L}_z\right] = 0 \\ \implies \left[\hat{H}, \hat{L}_z\right] = 0\end{gathered} \tag{4.84}$$

This implies that $\hat{H}$, $\hat{\vec{L}}^2$, and $\hat{L}_z$ have a *common complete* set of eigenstates, which means that we can create a new eigenstate ψ with separated components as shown below:

$$\psi(r,\theta,\varphi) = R(r)Y_l^m(\theta,\varphi) \tag{4.85}$$

For future simplicity, we will set $U(r) \equiv rR(r)$.

Finding Eigenvalues/Eigenfunctions

We already know from past experience that the following is true:

$$\hat{H}\psi = E\psi \tag{4.86}$$

Let's plug the values we discovered in (4.83) into the above:

$$\left[-\frac{\hbar^2}{2M}\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + V(r) + \frac{\overbrace{\hat{\vec{L}}^2}^{\hbar^2 l(l+1)}}{2Mr^2}\right]\frac{U(r)}{r}Y_l^m(\theta,\varphi) = E\frac{U(r)}{r}Y_l^m(\theta,\varphi) \tag{4.87}$$

Since $\dfrac{\partial}{\partial r}Y_l^m = 0$, taking the product rule allows us to eliminate Y_l^m:

$$\left[-\frac{\hbar^2}{2M}\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + V(r) + \frac{\hbar^2 l(l+1)}{2Mr^2}\right]\frac{U(r)}{r} = E\frac{U(r)}{r} \tag{4.88}$$

Before moving on, we must make sure to properly take our derivatives, so we will evaluate a portion of our equation separately:

$$\begin{aligned}\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right)\frac{U(r)}{r} &= \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\left(\frac{\partial U}{\partial r}\frac{1}{r} - \frac{U(r)}{r^2}\right)\right) = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r\frac{\partial U(r)}{\partial r} - U(r)\right)\\ &= \frac{1}{r^2}\left(\frac{\partial U}{\partial r} + r\frac{\partial^2 U}{\partial r^2} - \frac{\partial U}{\partial r}\right) = \frac{1}{r}\frac{\partial^2 U}{\partial r^2}\end{aligned} \tag{4.89}$$

We can then plug the result back into (4.88):

$$-\frac{\hbar^2}{2M}\frac{1}{r}\frac{\partial^2}{\partial r^2}U(r) + V(r)\frac{U(r)}{r} + \frac{\hbar^2}{2Mr^2}l(l+1)\frac{U(r)}{r} = E\frac{U(r)}{r} \tag{4.90}$$

Finally, we can multiply both sides by r and factor out $U(r)$, leaving us with our energy eigenvalue and **central-symmetric radial equation**:

$$\left[-\frac{\hbar^2}{2M}\frac{\partial^2}{\partial r^2} + V(r) + \frac{\hbar}{2Mr^2}l(l+1)\right]U(r) = EU(r) \tag{4.91}$$

In other words, the above is the central-symmetric version of the time-independent Schrödinger equation, but has different boundary conditions than the Cartesian TISE; for instance, r cannot be negative.

CHAPTER 5

The Hydrogen Atom

Introduction

Up to this point, everything we've worked with has been purely theoretical – now, we are finally going to apply our knowledge to something real:

The **Hydrogen atom** is a specialization of the *Coulomb potential* and other properties, such as *angular momentum* and *spin-orbit interactions*.

First of all, let's define our Coulomb potential:

$$V(r) = -\frac{e^2}{4\pi\epsilon_0 r} \tag{5.1}$$

For the remainder of this chapter, we will be using the *reduced electron mass* for M in the *radial equation*[1]:

$$M = \frac{m_e m_p}{m_e + m_p} \approx \frac{m_e m_p}{m_p} = m_e \tag{5.2}$$

Since we have that $m_e << m_p$, we can use the above approximation rather than the full expression. This gives us a radial equation with a factor of $\frac{2m_e}{\hbar^2}$ multiplied on both sides:

$$\left[-\frac{\partial^2}{\partial r^2} - \frac{e^2 m_e}{2\pi\epsilon_0\hbar^2 r} + \frac{l(l+1)}{r^2}\right] U(r) = \frac{2m_e}{\hbar} E U(r) \tag{5.3}$$

Defining $\varrho \equiv \frac{r}{a_0}$ gives us a more elegant expression:

$$\left[-\frac{\partial^2}{\partial \varrho^2} - \frac{2}{\varrho} + \frac{l(l+1)}{\varrho^2}\right] U(\varrho) = \frac{2m_e a_0^2}{\hbar^2} E U(\varrho) \tag{5.4}$$

Where a_0 is the **Bohr radius**:

$$a_0 \equiv \frac{4\pi\epsilon_0\hbar^2}{e^2 m_e} \tag{5.5}$$

[1] Defined in (4.80) on p.74

5.1 Ladder Operators

To raise and lower the energy eigenvalue of our system, we can use the following *ladder operators*:

The **creation operator** for the Hydrogen atom

$$\hat{A}_l^\dagger = -\frac{l+1}{\varrho} + \frac{1}{l+1} + \frac{\partial}{\partial \varrho} \tag{5.6}$$

The **annihilation operator** for the Hydrogen atom

$$\hat{A}_l = -\frac{l+1}{\varrho} + \frac{1}{l+1} - \frac{\partial}{\partial \varrho} \tag{5.7}$$

These will *raise* and *lower* an electron's energy eigenvalue E_n, where $E_1 = -13.6\text{eV}$

How did this come to be? Let's use (5.3) to derive the above ladder operators!

5.1.1 Derivation of the Ladder Operators

In (5.3), we see that all our terms (except for the middle one) represent units of the dimension [1/length]; this implies that our center term should also follow suit. As a result, we can infer that $\frac{4\pi\epsilon_0\hbar^2}{e^2 m_e}$ must also be a length, known as the *Bohr radius*:

$$a_0 = \frac{4\pi\epsilon_0\hbar^2}{e^2 m_e} \approx 0.5\text{Å} \tag{5.8}$$

We will now make (5.3) dimensionless by redefining length in terms of the Bohr radius:

$$\varrho \equiv \frac{r}{a_0} \implies r = a_0\varrho \tag{5.9}$$

We can rewrite (5.3) in a more elegant manner, by using a_0 and ϱ:

$$\left[-\frac{\partial^2}{a_0^2 \partial \varrho^2} - \frac{2}{a_0^2 \varrho} + \frac{l(l+1)}{a_0^2 \varrho^2}\right] U(\varrho) = \frac{2m_e}{\hbar} E U(\varrho) \tag{5.10}$$

Multiplying both sides by a_0 then gives us a new operator $\hat{h}_l$: the radial Hamiltonian for a central symmetric potential, as well as (5.4):

$$\left[\underbrace{-\frac{\partial^2}{\partial \varrho^2} - \frac{2}{\varrho} + \frac{l(l+1)}{\varrho^2}}_{\hat{h}_l}\right] U(\varrho) = \frac{2m_e a_0^2}{\hbar^2} E U(\varrho) \tag{5.11}$$

To continue any further, we will need a physical connection to the above expression – specifically, the fact that $\frac{2m_e a_0^2}{\hbar^2} = -\frac{1}{E_1}$, where $E_1 = -13.6\text{eV}$[2] is the lowest energy level possible for an electron in the Hydrogen atom. This therefore implies the following:

$$\frac{2m_e a_0^2}{\hbar^2} E = \frac{E}{E_1} \tag{5.12}$$

This means we can redefine our expression in (5.11):

$$\hat{h}_l U(\varrho) = \frac{E}{E_1} U(\varrho) \tag{5.13}$$

We can also rewrite the above operator-eigenstate equation in terms of a quantized value λ, giving us:

$$\frac{E}{E_1} = \frac{1}{\lambda^2} > 0 \implies \hat{h}_l U_{\lambda,l} = -\frac{1}{\lambda^2} U_{\lambda,l} \tag{5.14}$$

We will return to the above later, when we derive the form of our possible eigenstates.

[2]The energy must be negative, since we have a negative sign in the previous expression!

Example 5.1.1: Product of Ladder Operators

Let's see what happens when we take the product of our ladder operators $\hat{A}_l$ and $\hat{A}_l^\dagger$. We begin with the definitions shown in (5.7) and (5.6):

$$\hat{A}_l\hat{A}_l^\dagger = \left(-\frac{l+1}{\varrho} + \frac{1}{l+1} - \frac{\partial}{\partial\varrho}\right)\left(-\frac{l+1}{\varrho} + \frac{1}{l+1} + \frac{\partial}{\partial\varrho}\right) \tag{5.15}$$

Expanding our product yields two squares and a commutation relation:

$$= -\frac{\partial^2}{\partial\varrho^2} + \left(-\frac{l+1}{\varrho} + \frac{1}{l+1}\right)^2 - \left[\frac{\partial}{\partial\varrho}, -\frac{l+1}{\varrho} + \frac{1}{l+1}\right] \tag{5.16}$$

Looking at our commutation relation, $\frac{1}{l+1}$ commutes with $\frac{\partial}{\partial\varrho}$ and can be extracted. Before continuing, let's solve for the remaining commutation:

$$\left[\frac{\partial}{\partial\varrho}, \frac{l+1}{\varrho}\right] \tag{5.17}$$

So as to correctly implement the product rule of derivation, we will insert a temporary "dummy function" $f(\varrho)$ and expand:

$$\left[\frac{\partial}{\partial\varrho}, \frac{l+1}{\varrho}\right] f(\varrho) = \frac{\partial}{\partial\varrho}\frac{l+1}{\varrho}f(\varrho) - \frac{l+1}{\varrho}\frac{\partial}{\partial\varrho}f(\varrho) \tag{5.18}$$

We can then take our derivatives:

$$= -\frac{l+1}{\varrho^2}f(\varrho) + \cancel{\frac{l+1}{\varrho}\frac{\partial}{\partial\varrho}f(\varrho)} - \cancel{\frac{l+1}{\varrho}\frac{\partial}{\partial\varrho}f(\varrho)} = -\frac{l+1}{\varrho^2}f(\varrho) \tag{5.19}$$

We can then plug the above result into (5.16):

$$\hat{A}_l\hat{A}_l^\dagger = -\frac{\partial^2}{\partial\varrho^2} + \left(-\frac{l+1}{\varrho} + \frac{1}{l+1}\right)^2 - \frac{l+1}{\varrho^2} = -\frac{\partial^2}{\partial\varrho^2} - \frac{2}{\varrho} + \frac{l(l+1)}{\varrho^2} + \frac{1}{(l+1)^2} \tag{5.20}$$

Recognize something in the expression above? As seen in (5.11), we have found an expression in terms of $\hat{h}_l$:

$$\hat{A}_l\hat{A}_l^\dagger = \hat{h}_l + \frac{1}{(l+1)^2} \tag{5.21}$$

As seen in the above example, we can use the ladder operators to redefine $\hat{h}_l$:

$$\hat{h}_l = \hat{A}_l\hat{A}_l^\dagger - \frac{1}{(l+1)^2} \tag{5.22}$$

Alternatively, we can define the products of our ladder operators as follows:

$$\hat{A}_l \hat{A}_l^\dagger = \hat{h}_l + \frac{1}{(l+1)^2} \tag{5.23}$$

Conversely, this also implies that:

$$\hat{A}_l^\dagger \hat{A}_l = \hat{h}_{l+1} + \frac{1}{(l+1)^2} \tag{5.24}$$

5.1.2 Quantum Numbers

Recall that we introduced (5.14), with the intent to derive the possible energy eigenstates for the Hydrogen atom. To accomplish this we need to find our quantum numbers. We begin by applying the creation operator $\hat{A}_l^\dagger$ to $U_{\lambda,l}$, giving us a new state:

$$\hat{h}_l \hat{A}_l^\dagger U_{\lambda,l} = \left(\hat{A}_l \hat{A}_l^\dagger - \frac{1}{(l+1)^2} \right) \hat{A}_l^\dagger U_{\lambda,l} \tag{5.25}$$

We can rearrange the right-hand side, and simplify our equation:

$$\hat{A}_l^\dagger \underbrace{\left(\hat{A}_l \hat{A}_l^\dagger - \frac{1}{(l+1)^2} \right)}_{\hat{h}_l} U_{\lambda,l} = -\lambda^{-2} \left(\hat{A}_l^\dagger U_{\lambda,l} \right) \tag{5.26}$$

We see that the eigenvalue remains the same, implying that $\hat{A}_l^\dagger U_{\lambda,l}$ is an eigenstate of $\hat{h}_{l+1}$ with the eigenvalue $-\lambda^{-2}$. Let's now normalize our new state[3]:

$$1 = \left| \hat{A}_l^\dagger U_{\lambda,l} \right|^2 = \int_0^\infty d\varrho \, U_{\lambda,l}^* \hat{A}_l \hat{A}_l^\dagger U_{\lambda,l} \tag{5.27}$$

We can plug in the value for $\hat{A}_l \hat{A}_l^\dagger$ computed in (5.23), and then replace $\hat{h}_l$ by its eigenvalue $-\lambda^{-2}$

$$= \int_0^\infty d\varrho \, U_{\lambda,l}^* \left(\hat{h}_l + \frac{1}{(l+1)^2} \right) U_{\lambda,l} = \int_0^\infty d\varrho \, U_{\lambda,l}^* \left(-\lambda^{-2} + \frac{1}{(l+1)^2} \right) U_{\lambda,l} \tag{5.28}$$

Most of the above expression is constant with respect to ρ, so we can simplify:

$$= \left(-\lambda^{-2} + \frac{1}{(l+1)^2} \right) \int_0^\infty d\varrho \, \underbrace{U_{\lambda,l}^* U_{\lambda,l}}_{\geq 0} \tag{5.29}$$

There is a problem to consider now; our norm is not permitted to be negative, but this will inevitably occur for a large enough value of l! So for a fixed λ, we must find a maximum value of l which we will call $l_{\max}$. Since we are keeping λ fixed, $l_{\max}$ will depend on λ such that:

[3]Note that we integrate from *zero* to infinity, since we are dealing with spherical coordinates!

$$\hat{A}^{\dagger}_{l_{\max}} U_{\lambda, l_{\max}} = 0 \tag{5.30}$$

This means that $\lambda = l_{\max} + 1$, which remains unchanged as we apply $\hat{A}^{\dagger}_{l}$. Let's now redefine λ such that is becomes our *principal quantum number*:

$$n \equiv \lambda = l_{\max} + 1 \tag{5.31}$$

Based on (5.14), we can infer the following:

$$\frac{E}{E_1} = \frac{1}{n^2}, \text{ for } n = 1, 2, 3, ... \tag{5.32}$$

Which in turn, implies that $l = 0, 1, 2, ..., n-1$. Finally, we must implement one last quantum number m, which takes the values $m = -l, -l+1, ..., l-1, l$.

These three quantum numbers combined are known as the **nlm-state**, of the form $|nlm\rangle$. These states are always degenerate for $n > 1$, where the degree of degeneracy is given by $d = 2n^2 - n$. As a wavefunction, this state can be represented in the form ψ_{nlm}.

CHAPTER 6

Spin-1/2

6.1 Introduction to Spin

Up to this point, we've dealt with an algebra for *angular momentum*, which dictates the behavior of the azimuthal quantum number l and defines the eigenvalues for the angular momentum operator. Now, we will introduce the *spin* operator $\hat{\vec{S}}$ as well as the quantum numbers s and m_s, where s is the *spin quantum number* and m_s is the corresponding magnetic quantum number, analogous to m and l.

Since $\hat{\vec{L}}$ and $\hat{\vec{S}}$ share the same algebra, we have the following commutation relations available to us:

$$\left[\hat{S}_x, \hat{S}_y\right] = i\hbar\hat{S}_z, \ \left[\hat{S}_y, \hat{S}_z\right] = i\hbar\hat{S}_x, \ \left[\hat{S}_z, \hat{S}_x\right] = i\hbar\hat{S}_y \tag{6.1}$$

There does exist an exception to the analogy of angular momentum and spin, and that is the fact that spin does not represent an actual physical rotation, but an *internal degree of freedom* that cannot be related to a particle's position or momentum!

Back to our analogy, we can also define an operator $\hat{\vec{S}}^2$ such that:

$$\hat{\vec{S}}^2 = \hat{S}_x^2 + \hat{S}_y^2 + \hat{S}_z^2 \tag{6.2}$$

The eigenvalue for this operator is similar to that of $\hat{\vec{L}}^2$:

$$\hat{\vec{S}}^2 \left|sm_s\right\rangle = \hbar s(s+1) \left|sm_s\right\rangle \tag{6.3}$$

The eigenvalue for the z-component of $\hat{\vec{S}}$ is also similar to that of $\hat{L}_z$:

$$\hat{S}_z \left|sm_s\right\rangle = \hbar m \left|sm_s\right\rangle \tag{6.4}$$

As for our quantum numbers, the allowed values of s are not dependent on n in the way l depends on n, and increases in half-integer steps:

$$s = \left\{0, \frac{1}{2}, 1, \frac{3}{2}, ...\right\} \tag{6.5}$$

On the other hand, m_s and m_l behave similarly to each other, in the sense that:

$$m_s = \{-s, -s+1, ..., s-1, s\} \tag{6.6}$$

Particles that have half-integer spin are called *fermions*, while particles with integer spin are called *bosons*. For example, *electrons* are fermions with *spin-1/2*.

6.2 Introduction to Spin-1/2

As mentioned in the previous subsection, particles with half-integer spins are fermions – one property of fermions is that they abide by the Pauli exclusion principle:

> "The **Pauli exclusion principle** is the quantum mechanical principle which states that two or more identical fermions (particles with half-integer spin) cannot occupy the same quantum state within a quantum system simultaneously." [6]

As a result, there are considerations to be made when working with spin-1/2 particles, so this section will be dedicated exclusively to these. In addition, electrons being spin-1/2 particles make this a particularly important subject to discuss.

Before we begin, we will define a new notation for the eigenstates of $\hat{S}_z$, which is of the form $|sm_s\rangle$:

Eigenstates for $\hat{S}_z$:	$\left\|\frac{1}{2}\frac{1}{2}\right\rangle$	$\left\|\frac{1}{2}-\frac{1}{2}\right\rangle$
New notation:	$\|\uparrow\rangle$	$\|\downarrow\rangle$
Vector representation:	$\begin{pmatrix}1\\0\end{pmatrix}$	$\begin{pmatrix}0\\1\end{pmatrix}$

Before moving on, there are four useful relations between these states and their adjoints we must mention, as they will be the basis of many simplifications in the coming sections:

$$\langle\uparrow|\uparrow\rangle = \langle\downarrow|\downarrow\rangle = 1, \qquad \langle\uparrow|\downarrow\rangle = \langle\downarrow|\uparrow\rangle = 0 \tag{6.7}$$

6.3 Operators

6.3.1 Spin Operator

We can use (6.4) to find the eigenvalues of $\hat{S}_z$ for each possible spin-1/2 state:

$$\hat{S}_z |\uparrow\rangle = \frac{\hbar}{2} |\uparrow\rangle , \qquad \hat{S}_z |\downarrow\rangle = -\frac{\hbar}{2} |\downarrow\rangle \tag{6.8}$$

This information allows us to derive the *eigendecomposition* of the operator:

$$\hat{S}_z = \frac{\hbar}{2} \left(|\uparrow\rangle \langle\uparrow| - |\downarrow\rangle \langle\downarrow| \right) \tag{6.9}$$

Example 6.3.1: The Eigendecomposition

Let's apply the eigendecomposition of $\hat{S}_z$ as presented in (6.9) to our two spin-1/2 states:

$$\hat{S}_z |\uparrow\rangle = \frac{\hbar}{2} \left(|\uparrow\rangle \cancelto{1}{\langle\uparrow|\uparrow\rangle} - |\downarrow\rangle \cancelto{0}{\langle\downarrow|\uparrow\rangle} \right) = \frac{\hbar}{2} |\uparrow\rangle \tag{6.10}$$

$$\hat{S}_z |\downarrow\rangle = \frac{\hbar}{2} \left(|\uparrow\rangle \cancelto{0}{\langle\uparrow|\downarrow\rangle} - |\downarrow\rangle \cancelto{1}{\langle\downarrow|\downarrow\rangle} \right) = -\frac{\hbar}{2} |\downarrow\rangle \tag{6.11}$$

We clearly see that they return the correct eigenvalues, as shown in (6.8).

6.3.2 Ladder Operators

The ladder operators for spin follow the same patters as those for angular momentum, such that $S_+^\dagger = S_-$ and:

$$\hat{S}_\pm = \hat{S}_x \pm \hat{S}_y \tag{6.12}$$

There are four possible combinations of these operators $\hat{S}_+$, $\hat{S}_-$ and the eigenstates $|\uparrow\rangle$, $|\downarrow\rangle$:

$$\hat{S}_+ |\downarrow\rangle = \hbar |\uparrow\rangle , \qquad \hat{S}_- |\uparrow\rangle = \hbar |\downarrow\rangle \tag{6.13}$$

The two other possible combinations give us zero, since you cannot lower $m_s = -1/2$ nor raise $m_s = 1/2$:

$$\hat{S}_+ |\uparrow\rangle = \hat{S}_- |\downarrow\rangle = 0 \tag{6.14}$$

Example 6.3.2: Deriving an Eigenvalue

Let us show that (6.13) is true for the first case; this can be accomplished by normalizing both sides of the following:

$$\hat{S}_+ \left|\downarrow\right\rangle = \hbar \left|\uparrow\right\rangle \tag{6.15}$$

We start with the right-hand side:

$$\left|\hbar \left|\uparrow\right\rangle\right|^2 = \left\langle\uparrow\right| \left|\hbar\hbar\right| \left|\uparrow\right\rangle = \hbar^2 \cancelto{1}{\left\langle\uparrow\middle|\uparrow\right\rangle} = \hbar^2 \tag{6.16}$$

And now the left-hand side:

$$\left|\hat{S}_+ \left|\downarrow\right\rangle\right|^2 = \left\langle\downarrow\right| \left|\hat{S}_+^\dagger \hat{S}_+\right| \left|\downarrow\right\rangle = \left\langle\downarrow\right| \left|\hat{S}_- \hat{S}_+\right| \left|\downarrow\right\rangle \tag{6.17}$$

We apply the definition of the ladder operator and substitute the resulting commutation relation:

$$= \left\langle\downarrow\right| \left| \left(\hat{S}_x - i\hat{S}_y\right)\left(\hat{S}_x + i\hat{S}_y\right) \right| \left|\downarrow\right\rangle = \left\langle\downarrow\right| \left| \underbrace{\hat{S}_x^2 + \hat{S}_y^2}_{\hat{\vec{S}}^2 - \hat{S}_z^2} + i \underbrace{\left[\hat{S}_x, \hat{S}_y\right]}_{i\hbar\hat{S}_z} \right| \left|\downarrow\right\rangle \tag{6.18}$$

The next step is to apply each operator to the given state:

$$= \left\langle\downarrow\right| \left(\hat{\vec{S}}^2 \left|\downarrow\right\rangle - \hat{S}_z^2 \left|\downarrow\right\rangle - \hbar\hat{S}_z \left|\downarrow\right\rangle\right) = \hbar^2 \left(\frac{3}{4} - \frac{1}{4} + \frac{1}{2}\right) \cancelto{1}{\left\langle\downarrow\middle|\downarrow\right\rangle} \tag{6.19}$$

Simplifying the last expression gives us our result:

$$\left|\hat{S}_+ \left|\downarrow\right\rangle\right|^2 = \hbar^2 \quad \square \tag{6.20}$$

We can also represent our ladder operators as eigendecompositions:

$$\hat{S}_+ = \hbar \left|\uparrow\right\rangle \left\langle\downarrow\right|, \qquad \hat{S}_- = \hbar \left|\downarrow\right\rangle \left\langle\uparrow\right| \tag{6.21}$$

Example 6.3.3: Eigendecompositions of the Ladder Operators

Let's compute the eigenvalues of our ladder operators using the eigendecompositions presented in (6.21), starting with $\hat{S}_+$:

$$\hat{S}_+ |\uparrow\rangle = \hbar |\uparrow\rangle \langle\downarrow | \uparrow\rangle = 0 \tag{6.22}$$

$$\hat{S}_+ |\downarrow\rangle = \hbar |\uparrow\rangle \langle\downarrow | \downarrow\rangle = \hbar |\uparrow\rangle \tag{6.23}$$

Next, we will check the eigenvalues for $\hat{S}_-$:

$$\hat{S}_- |\uparrow\rangle = \hbar |\downarrow\rangle \langle\uparrow | \uparrow\rangle = \hbar |\downarrow\rangle \tag{6.24}$$

$$\hat{S}_- |\downarrow\rangle = \hbar |\downarrow\rangle \langle\uparrow | \downarrow\rangle = 0 \tag{6.25}$$

We see that these are correct as per (6.13) and (6.14). □

6.4 Pauli Spin Matrices

As mentioned briefly in the introduction, we can represent our spin-up and spin-down states with the vectors $\begin{pmatrix}1\\0\end{pmatrix}$ and $\begin{pmatrix}0\\1\end{pmatrix}$ respectively. Before we continue, recall the way we represented bras and kets as matrices in Example 1.2.2.

The purpose of this brief review is intended to prepare us for the introduction of the **Pauli spin matrices**, a way to represent the components $\hat{S}_x$, $\hat{S}_y$, and $\hat{S}_z$ of the spin operator $\hat{\vec{S}}$. To derive these matrices, we must begin by representing each component such that they are composed of our ladder operators – this is because these can each be represented with an eigendecomposition, which can in turn be represented by vectors!

Using (6.12), we can set up the two first components:

$$\hat{S}_x = \frac{1}{2}\left(\hat{S}_+ + \hat{S}_-\right) = \frac{\hbar}{2}\left(|\uparrow\rangle\langle\downarrow| + |\downarrow\rangle\langle\uparrow|\right) \tag{6.26}$$

$$\hat{S}_y = \frac{1}{2i}\left(\hat{S}_+ - \hat{S}_-\right) = \frac{\hbar}{2i}\left(|\uparrow\rangle\langle\downarrow| - |\downarrow\rangle\langle\uparrow|\right) \tag{6.27}$$

We already know the eigendecomposition for $\hat{S}_z$ from (6.9):

$$\hat{S}_z = \frac{\hbar}{2}\left(|\uparrow\rangle\langle\uparrow| - |\downarrow\rangle\langle\downarrow|\right) \tag{6.28}$$

We can then insert our vectors into each expression and compute the outer product:

$$\hat{S}_x = \frac{\hbar}{2}\left(\begin{pmatrix}1 & 0\end{pmatrix}\begin{pmatrix}0\\1\end{pmatrix} + \begin{pmatrix}0 & 1\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix}\right) = \frac{\hbar}{2}\begin{pmatrix}0 & 1\\1 & 0\end{pmatrix} \tag{6.29}$$

$$\hat{S}_y = \frac{\hbar}{2i}\left(\begin{pmatrix}1 & 0\end{pmatrix}\begin{pmatrix}0\\1\end{pmatrix} - \begin{pmatrix}0 & 1\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix}\right) = \frac{\hbar}{2}\begin{pmatrix}0 & i\\-i & 0\end{pmatrix} \tag{6.30}$$

$$\hat{S}_z = \frac{\hbar}{2}\left(\begin{pmatrix}1 & 0\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix} - \begin{pmatrix}0 & 1\end{pmatrix}\begin{pmatrix}0\\1\end{pmatrix}\right) = \frac{\hbar}{2}\begin{pmatrix}1 & 0\\0 & -1\end{pmatrix} \tag{6.31}$$

These are the Pauli spin matrices, with an extra factor $\frac{\hbar}{2}$.

The **Pauli spin matrices**:

$$\vec{S} = \frac{\hbar}{2}\vec{\sigma} \tag{6.32}$$

Where $\vec{\sigma}$ is a set comprising of the following components:

$$\vec{\sigma} = \{\sigma_x, \sigma_y, \sigma_z\} = \left\{\begin{pmatrix}0 & 1\\1 & 0\end{pmatrix}, \begin{pmatrix}0 & i\\-i & 0\end{pmatrix}, \begin{pmatrix}1 & 0\\0 & -1\end{pmatrix}\right\} \tag{6.33}$$

These matrices are *traceless*[1] and *idempotent*, implying that $\sigma_x\sigma_x = \sigma_y\sigma_y = \sigma_z\sigma_z = I_2$.

6.5 Magnetic Fields

6.5.1 Classical Introduction

We are now sufficiently equipped to study a new real-world scenario, an electron suspended in a magnetic field. We can begin by approaching this from a classical standpoint: imagine we have a charge q orbiting an electric current at a constant radius r due to magnetic forces. This period of rotation T of this current would depend on the current, and so we can create an expression for this current:

$$I = \frac{q}{T} = \frac{q}{\frac{2\pi r}{v}} = \frac{qv}{2\pi r} \tag{6.34}$$

A circulating current gives us the **magnetic moment** $\vec{\mu}$, and the above expression for I can be inserted into the result:

$$\vec{\mu} = \frac{1}{2}\oint \vec{r} \times I d\vec{l} = \pi r^2 I\hat{z} = \frac{qvr}{2}\hat{z} \tag{6.35}$$

We also know that the norm of the system's orbital angular momentum is as follows:

$$L = mvr \tag{6.36}$$

This gives us our final expression for the magnetic moment:

$$\vec{\mu} = \frac{q}{2m}\vec{L} \tag{6.37}$$

[1] The **trace** of a matrix is the sum of its diagonal elements.

6.5.2 The Hamiltonian

Now, imagine we wish to approach this from a quantum perspective – we now consider our electron a spin-1/2 particle whose Hamiltonian is defined as follows:

$$\hat{H} = -\vec{B} \cdot \vec{\mu} \tag{6.38}$$

Where $\vec{B}$ is an external magnetic field, and $\vec{\mu}$ is the *quantum* magnetic moment, which is defined in terms of a particle's spin operator:

$$\vec{\mu} = \gamma \vec{S} \tag{6.39}$$

Here, γ is known as the **gyromagnetic ratio**. As we plainly see, spin has the properties of the $\hat{L}$ operator; still, we must not forget that this occurs *without* the existence of a physical rotation.

For an electron, the gyromagnetic ratio is $\gamma = -\frac{e}{2m_e} \approx 2.0023193043171(52)$ [7]. In conclusion, we can state that the Hamiltonian of a magnetic moment in a magnetic field is:

$$\hat{H} = -\gamma \vec{B} \cdot \vec{S} \tag{6.40}$$

6.5.3 Energy Eigenvalues

To calculate our energy eigenvalues, we will take our Hamiltonian and use (6.9) to express it as an eigendecomposition:

$$\hat{H} = -\gamma \vec{B} \cdot \vec{S} = -\gamma \vec{B} \frac{\hbar}{2} \left(|\uparrow\rangle \langle\uparrow| - |\downarrow\rangle \langle\downarrow| \right) \tag{6.41}$$

The possible eigenstates of this $\hat{H}$ are $|\uparrow\rangle$ and $|\downarrow\rangle$, such that:

$$\hat{H} |\uparrow\rangle = -\gamma B \frac{\hbar}{2} |\uparrow\rangle \tag{6.42}$$

$$\hat{H} |\downarrow\rangle = \gamma B \frac{\hbar}{2} |\downarrow\rangle \tag{6.43}$$

6.5.4 Change Over Time

Now, let's look at a system with a spin-1/2 placed in a magnetic field, and observe the changes in the system over time. To accomplish this, we must create a superposition of states, which will represent a stationary state $|\psi\rangle$:

$$|\psi\rangle = a |\uparrow\rangle + b |\downarrow\rangle \tag{6.44}$$

As always, we set a normalization condition such that $|a|^2 + |b|^2 = 1$; this time, however, we will set up a parameterization that suits our purposes, while still respecting the aforementioned boundaries:

$$a = \cos\left(\frac{\alpha}{2}\right) e^{i\varphi}, \qquad b = \sin\left(\frac{\alpha}{2}\right) e^{i\theta} \tag{6.45}$$

Next, we assume that the stationary state defined in (6.44) represents our time-dependent state $|\Psi(t)\rangle$ at $t = 0$, giving us the following:

$$|\Psi(t)\rangle = e^{-i\frac{\hat{H}}{\hbar}t} |\psi\rangle = ae^{i\frac{\gamma B}{2}t} |\uparrow\rangle + be^{-i\frac{\gamma B}{2}t} |\downarrow\rangle \tag{6.46}$$

Now, imagine we were to take a series of energy measurements and take the average; as the number of measurements approaches infinity, the average will approach the expectation value. Let's compute this expectation value, and see how the result changes as a function of time:

$$\left\langle \hat{S}_z \right\rangle = \langle \Psi(t) | \hat{S}_z | \Psi(t) \rangle \tag{6.47}$$

Plugging in our expression for $\Psi(t)$ and its adjoint, as well as the eigendecomposition of $\hat{S}_z$ gives us:

$$= \left(a^* e^{-i\frac{\gamma B}{2}t} |\uparrow\rangle + b^* e^{i\frac{\gamma B}{2}t} |\downarrow\rangle\right) \frac{\hbar}{2} \left(|\uparrow\rangle \langle\uparrow| - |\downarrow\rangle \langle\downarrow|\right) \left(ae^{i\frac{\gamma B}{2}t} |\uparrow\rangle + be^{-i\frac{\gamma B}{2}t} |\downarrow\rangle\right) = \frac{\hbar}{2}\left(|a|^2 - |b|^2\right) \tag{6.48}$$

In conclusion, our final expression for our expectation value is as follows:

$$\left\langle \hat{S}_z \right\rangle = \frac{\hbar}{2} \cos(\alpha) \tag{6.49}$$

We clearly see that the expression is *independent* of time, implying that the energy eigenvalue of our magnetic moment is constant over time.

Bibliography

[1] H. J. Folse, "The copenhagen interpretation of quantum theory and whitehead's philosophy of organism," *Tulane Studies in Philosophy*, vol. 23, p. 32–47, 1974.

[2] J. V. Neumann and J. V. Neumann, *Mathematical foundations of quantum mechanics*. Princeton Univ. Pr., 1955.

[3] By The American Institute of Physics credits the photo to AB Lagrelius & Westphal, which is the Swedish company used by the Nobel Foundation for most photos of its book series Les Prix Nobel. (Niels Bohr's Nobel Prize biography, from 1922) [Public domain], via Wikimedia Commons.

[4] Bundesarchiv, Bild 183-R57262 / Unknown author / CC-BY-SA 3.0 [CC BY-SA 3.0 de (https://creativecommons.org/licenses/by-sa/3.0/de/deed.en)], via Wikimedia Commons.

[5] Wikipedia contributors, "Ladder operator — Wikipedia, the free encyclopedia," 2017, [Online; accessed 15-September-2018]. [Online]. Available: https://en.wikipedia.org/w/index.php?title=Ladder_operator&oldid=813942979

[6] ——, "Pauli exclusion principle — Wikipedia, the free encyclopedia," 2018, [Online; accessed 19-October-2018]. [Online]. Available: https://en.wikipedia.org/w/index.php?title=Pauli_exclusion_principle&oldid=864812567

[7] B. Odom, D. Hanneke, B. D'Urso, and G. Gabrielse, "New measurement of the electron magnetic moment using a one-electron quantum cyclotron," *Physical Review Letters*, vol. 97, no. 3, 2006.