STEPHEN WIGGINS

ELEMENTARY QUANTUM MECHANICS

SCHOOL OF MATHEMATICS
UNIVERSITY OF BRISTOL

Contents

List of Figures

Preface

These notes have grown out of a one term course on quantum mechanics in the School of Mathematics at the University of Bristol. I have taught this course several times, but these notes benefit immensely from the previous notes of Martin Sieber, Noah Linden, and Toby Cubitt. I am grateful that they have shared their materials with me.

Books and Other Reference Materials

There are literally hundreds of books on quantum mechanics, and I have clearly not reviewed them all. There are some good ones, and some I don't like at all. Below are some personal comments on a few books that you might find helpful.

To begin with the following is an excellent book on quantum mechanics for mathematics students by Hannabuss[1]. This book has some nice choices of topics and it can be used to fill in background material for which there is no time in the course.

[1] K. Hannabuss. *An introduction to quantum theory*, volume 1. Clarendon Press, 1997

By the time of the late 1970's it was probably regarded by many physicists and chemists that quantum mechanics was a "mature" subject whose foundations were unlikely to change substantially in the future. This turned out to be far from true. Nevertheless, a book (actually, two volumes) first appeared in 1977[2] that was encyclopedic in its scope (for the time) and still remains today as an extremely useful text with excellent discussions about many topics: The sections on the mathematical foundations of quantum mechanics (including Dirac's notation), the postulates of quantum mechanics, angular momentum, and tensor products are all excellent (and all in the first volume). If you are going on to do further work in quantum mechanics then you will need to have a good reference on your shelf. This book is an excellent choice for that purpose. Many of the explanations are at a deeper level than you will find in most books.

[2] C. Cohen-Tannoudji, B. Diu, and F. Laloe. *Quantum Mechanics*. Wiley-VCH, 1992

David Bohm was one of the great thinkers about what quantum mechanics means (but from the point of view of the physics). He wrote a book in the 1950's that was reprinted by Dover in 1989[3]. The sections on measurement and the Stern=Gerlach experiment are excellent.

[3] D. Bohm. *Quantum theory*. Courier Corporation, 1951

I mentioned that by the late 1970s many viewed quantum mechanics as a "mature subject", but big changes were coming. Working on the foundations of quantum mechanics was not really a "mainstream" topic in the 1950's, 60's, and 70's, but this did not stop a number of pioneers. I mentioned Bohm, but Bell, Leggett, Clauser, Zurek are some other names that contributed fundamental ideas that are now definitely part of the "mainstream". The ability to test many of these seemingly abstract ideas exploded with experimental advances throughout the 1980's and 90's (continuing to this day) and this has completely reinvigorated our thinking about the entire "structure" of quantum mechanics, and has likely also completely changed the way that it will be taught in the future. Textbooks are beginning to appear with this new point of view and, in my view, and an excellent example is the book by Schumacher and Westmoreland[4].

Also worth mentioning is the book of David Tannor[5]. There are extremely useful topics that you will find in this book that you will not find in any other book, and they are developed in a very clear manner. One of the criticisms of beginning quantum mechanics courses is that students tend not to know how to "do anything" with it once they have finished the course. There is more than a little truth in this assessment. Tannor is a theoretical chemist, and theoretical chemists' job is to understand how quantum mechanical systems "do things". Tannor's book will show you, better than most, how to use quantum mechanics to "get asnwers" about atomic and molecular systems.

[4] B. Schumacher and M. Westmoreland. *Quantum processes systems, and information*. Cambridge University Press, 2010

[5] D. J Tannor. *Introduction to quantum mechanics: a time-dependent perspective*. University Science Books, 2007

Internet Resources

The internet offers a wealth of (free) resources related to quantum mechanics. There is some very good material on the internet. There is some material that has good points, but these might be offset by some errors that will lead you astray in your understanding. And then there is some material that is plain wrong. Anyone can publish on the internet, and no one is "certifying" something as "correct", like is done in the peer review process. Even the famous Wikipedia is a "mixed bag" when it comes to quality. There are some topics that I think are very well discussed and make excellent references (e.g. the overview of linear differential equations and the discussion of Dirac's bra ket notation). However, there are some discussions of mathematical topics that contain errors. So be careful!

However, there is one internet source in quantum mechanics that I will use that is absolutely outstanding. These are the lecture notes of Professor Robert Littlejohn of the Physics department at the University of California at Berkeley. Professor Littlejohn himself has made fundamental contributions to quantum physics and he has developed these

notes over the years in his course at Berkeley. They can be obtained here:

http://bohr.physics.berkeley.edu/classes/221/1011/221.html

Please reference them if you use them in any source. I will refer to them throughout these notes when I use them as "Littlejohn's notes". In general, these notes are at a more advance level than this course. However, I would particularly recommend Notes 1: The Mathematical Formalism of Quantum Mechanics, Notes 2: The Postulates of Quantum Mechanics, and Notes 13: Representations of the Angular Momentum Operators and Rotations. You will not find better discussions of angular momentum and spin, and their mathematical and physical descriptions, than exist in these notes.

1
The Mathematical Structure of Quantum Mechanics

In this chapter we develop the develop the mathematics that is necessary to describe and compute the quantum mechanical concepts and quantities used in this course. In many cases we will not give complete proofs of mathematical results (especially in the infinite dimensional setting). Hopefully, you have had other courses that cover these concepts rigorously. The focus here will be describing the mathematical setting and results, and using them to describe and compute quantities in quantum mechanics.

1.1 Vector Spaces and Inner Products

Vector spaces play an important role in defining the notion of ''quantum state", so we begin by giving the definition of a vector space.

Definition 1 (Vector Space over a Field)**.** *A vector space, denote V, over a field $\mathcal{F}$, is a set of elements (the elements are called "vectors") on which two (binary) operations are defined:*

Vector addition. For any $\phi, \psi \in V$, then $\phi + \psi \in V$.

Scalar multiplication. For any $\alpha \in \mathcal{F}$, $\phi \in V$, then $\alpha\phi \in V$.

Moreover, the operations of vector addition and scalar multiplication satisfy the following eight axioms.

Associativity of vector addition. $\psi + (\phi + \chi) = (\psi + \phi) + \chi, \quad \forall\, \psi, \phi, \chi \in V$.

Commutativity of vector addition. $\psi + \phi = \phi + \psi, \quad \forall\, \psi, \phi \in V$.

Identity element for vector addition ("existence of a zero vector"). There exists a vector $0 \in V$ such that $\psi + 0 = \psi, \quad \forall \psi \in V$.

Inverse element for vector addition. $\forall \psi \in V$ there exists a vector $-\psi \in V$ such that $\psi + (-\psi) = 0$.

Distributivity of scalar multiplication with respect to vector addition. $\alpha(\psi + \phi) = \alpha\psi + \alpha\phi, \quad \forall\alpha \in \mathcal{F}, \psi, \phi \in V.$

Distributivity of scalar multiplication with respect to field addition. $(\alpha + \beta)\psi = \alpha\psi + \beta\psi, \quad \forall\alpha, \beta \in \mathcal{F}, \psi \in V.$

Compatibility of scalar multiplication with field multiplication. $\alpha(\beta\psi) = (\alpha\beta)\psi, \forall\alpha, \beta \in \mathcal{F}, \psi, \in V.$

Identity element of scalar multiplication. There exists an element $1 \in \mathcal{F}$ *such that* $1\,\psi = \psi, \quad \forall\psi \in V.$

For the purposes of this course, the field $\mathcal{F}$ *will be the complex numbers.*. A vector space where the field is the complex numbers is often referred to as a *complex vector space*.

We will also require an additional structure defined on a complex vector space–a complex inner product, which we now define.

Definition 2 (Inner Product)**.** *An inner product on a vector space* V *is a map of an ordered pair of vectors,* (ψ, ϕ) *to the complex numbers that satisfies the following properties:*

i) $(\psi, \alpha\phi + \beta\chi) = \alpha(\psi, \phi) + \beta(\psi, \chi), \quad \alpha, \beta \in \mathbb{C}.$

ii) $(\psi, \phi) = \overline{(\phi, \psi)}.$

iii) $(\psi, \psi) \geq 0$*, and equality holds if and only if* $\psi = 0$*,*

for any $\alpha, \beta \in \mathbb{C}$, $\psi, \phi, \chi \in V$.

Using the inner product, we can construct a norm defined for vectors in V. Geometrically, the norm provides us with a measure of the "length" of vectors in V. The definition of the "norm induced by the inner product" is as follows.

Definition 3 (Norm Induced by the Inner Product)**.** *For any vector* $\psi \in V$*, we define the norm of* ψ*, denoted* $\| \psi \|$*, as follows:*

$$\| \psi \| = \sqrt{(\psi, \psi)}. \tag{1.1}$$

Mathematically, norms satisfy certain properties (similar to inner products). In this case, those properties are inherited from the properties of the inner product. Rather than discussing in detail the properties satisfied by the norm, we will use Definition 3 without further comment.

We make the following remarks.

- Almost always, we will consider vectors $\psi \in V$ having norm one, i.e. $\| \psi \| = 1$. For any (nonzero) vector $\psi \in V$, we can divide it by its norm, and the result will be a vector having norm one. Such vectors are said to be *normalised*[1].

[1] "Normalizaton" will play a central role in our interpretation of the meaning of quantum mechanics in terms of observable quantities. We will see this in the next chapter.

- The vectors are said to be orthogonal if their inner product is zero, i.e. ψ, $\phi \in V$ are said to be orthogonal if $(\psi, \phi) = 0$. If ψ and ϕ have norm one (or "unit norm") then they are said to be *orthonormal*.
- The complex inner product is *antilinear* (or "conjugate linear") in the first argument. The meaning of this phrase is made clear in the following calculation.

$$(\alpha\phi + \beta\chi, \psi) = \overline{(\psi, \alpha\phi + \beta\chi)} = \overline{\alpha\,(\psi, \phi) + \beta\,(\psi, \chi)} = \bar{\alpha}\,(\phi, \psi) + \bar{\beta}\,(\chi, \psi)\,, \tag{1.2}$$

which holds for any α, $\beta \in \mathbb{C}$, ψ, ϕ, $\chi \in V$.

- The Schwarz inequality (or Cauchy-Schwarz inequality, or Cauchy-Bunyakovsky-Schwarz inequality) is a fundamental inequality that has many uses in quantum mechanics. It bounds the magnitude of the inner product of two vectors by their norm, i.e.

$$|(\phi, \psi)| \leq \| \phi \| \| \psi \|, \quad \forall \psi, \phi \in V. \tag{1.3}$$

We prove this inequality. Consider the quantity:

$$(\psi + \alpha\phi, \psi + \alpha\phi) \geq 0. \tag{1.4}$$

Using the properties of the inner product to expand this expression gives:

$$\begin{aligned} (\psi + \alpha\phi, \psi + \alpha\phi) &= (\psi + \alpha\phi, \psi) + \alpha(\psi + \alpha\phi, \phi), \\ &= (\psi, \psi) + \bar{\alpha}(\phi, \psi) + \alpha(\psi, \phi) + \alpha\bar{\alpha}(\phi, \phi) \geq 0. \end{aligned} \tag{1.5}$$

Now we make the choice:

$$\alpha = -\frac{(\phi, \psi)}{(\phi, \phi)}. \tag{1.6}$$

Substituting (1.6) into (1.5) gives:

$$(\psi, \psi) - \frac{(\psi, \phi)(\phi, \psi)}{(\phi, \phi)} - \frac{(\phi, \psi)(\psi, \phi)}{(\phi, \phi)} + \frac{(\phi, \psi)(\psi, \phi)}{(\phi, \phi)} \geq 0. \tag{1.7}$$

or

$$(\psi, \psi) - \frac{(\psi, \phi)(\phi, \psi)}{(\phi, \phi)} \geq 0. \tag{1.8}$$

Multiplying (1.8) by (ϕ, ϕ) gives:

$$(\psi, \psi)(\phi, \phi) \geq (\phi, \psi)(\psi, \phi), \tag{1.9}$$

which is the same as:

$$(\| \psi \| \| \phi \|)^2 \geq |(\phi, \psi)|^2, \tag{1.10}$$

which verifies the inequality.

- The norm satisfies the triangle inequality:

$$\| \psi + \phi \| \leq \| \psi \| + \| \phi \|, \quad \forall \psi, \phi \in V. \tag{1.11}$$

(We leave it as an exercise to show that this is true.)

Some examples of complex inner product spaces are the following.

Example. $\mathbb{C}^n$ denotes the space of n-tuples of complex numbers, $x \equiv (x_1, \ldots, x_n)$, $x_i \in \mathbb{C}$. It is straightforward to verify that $\mathbb{C}^n$ is a complex vector space. We define the following inner product on $\mathbb{C}^n$:

$$(x, y) = \sum_{j=1}^{n} \bar{x}_j y_j, \quad x, y \in \mathbb{C}^n. \tag{1.12}$$

We leave it as an exercise to verify that (1.12) satisfies the properties of an inner product.

Example. Let $L^2(D, dx)$ denote the set of all complex valued functions of a real variable, x, defined on some domain $D \in \mathbb{R}$, denoted by $\psi(x)$, that satisfy the following:

$$\int_D |\psi(x)|^2 dx < \infty. \tag{1.13}$$

$L^2(D, dx)$ becomes a complex vector space by defining vector addition and scalar multiplication "pointwise" as follows:

$$(\psi + \phi)(x) = \psi(x) + \phi(x), \tag{1.14}$$

$$(\alpha\psi)(x) = \alpha\psi(x). \tag{1.15}$$

We leave it as an exercise to verify that vector addition and scalar multiplication defined in this way satisfy the axioms of a complex vector space.

The complex vector space $L^2(D, dx)$ can be made into a complex inner product space by defining an inner product on $L^2(D, dx)$ as follows:

$$(\psi, \phi) = \int_D \overline{\psi(x)}\phi(x)dx. \tag{1.16}$$

We leave it as an exercise to verify that (1.16) satisfies the properties of an inner product.

Thus far, our discussion of complex inner product spaces has been extremely general. In order to carry out the type of computations required in quantum mechanics, we need to introduce some additional structure on V. We begin with the notion of *linear independence*.

Definition 4 (Linear Independence)**.** *A set of vectors $\psi_1, \ldots, \psi_n \in V$ is said to be linearly independent if the only scalars satisfying the equation*

$$\alpha_1\psi_1 + \cdots + \alpha_n\psi_n = 0,$$

are $\alpha_1 = \alpha_2 = \ldots = \alpha_n = 0$. Otherwise, the set of vectors is said to be linearly dependent.

The notion of linear independence leads naturally to the notion of *dimensionality* of a vector space.

Definition 5 (Dimensionality)**.** *A vector space is n-dimensional if it contains n linearly independent vectors, but not $n + 1$. If it contains n linearly independent vectors for every positive integer n then it is infinite dimensional.*

Next, we have the notion of a *spanning set* and a *basis* for a subset of V.

Definition 6 (Spanning Set, Basis)**.** *A set of vectors $\{\psi_1, \ldots, \psi_n\} \in V$ spans a finite dimensional vector space $W \subset V$ if every $\psi \in W$ can be written as $\psi = \alpha_1\psi_1 + \cdots + \alpha_n\psi_n$. A set of vectors $\{\psi_1, \ldots, \psi_n\} \in W$ is a basis for W if they are linearly independent and they span W.*

For infinite dimensional complex vector spaces we will need to give meaning to expressions like:

$$\sum_{i=1}^{\infty} \alpha_i\psi_i.$$

This notion is dealt with in the following definition.

Definition 7 (Convergence)**.** *A sequence of vectors $\psi_n \in V$ converges (strongly) to $\psi \in V$ if*

$$\| \psi - \psi_n \| \to 0 \quad as \quad n \to \infty.$$

A Typical Example in Quantum Mechanics of an Infinite Dimensional Complex Vector Space. We consider the 2π-periodic complex valued functions of a real variable, θ, satisfying:

$$\int_0^{2\pi} |\psi(\theta)|^2 d\theta < \infty.$$

We denote the space of such functions by $L^2(S^1, d\theta)$. We state (without proof) that a basis for this space is given by:

$$\{1, \cos(n\theta), \sin(n\theta), \quad n = 1, \ldots\}$$

Hence, an arbitrary function $\psi(\theta) \in L^2(S^1, d\theta)$ can be expressed as:

$$\psi(\theta) = \frac{a_0}{2} + \sum_{n=1}^{\infty} \left(a_n \cos(n\theta) + b_n \sin(n\theta)\right).$$

In discussing the convergence of sequences, the notion of a Cauchy sequence is fundamental.

Definition 8 (Cauchy Sequence)**.** *A sequence of vectors $\{\psi_n\}$ is said to be a Cauchy sequence if $\forall \epsilon > 0$ $\exists N \in \mathbb{N}$ such that for all m, $n > N$*

$$\| \psi_m - \psi_n \| < \epsilon.$$

We now are at the point where we can define the notion of a *Hilbert space*.

Definition 9 (Hilbert Space)**.** *A vector space, V with an inner product that is "complete" (i.e. every Cauchy sequence converges strongly to a vector in V) is called a Hilbert space.*

We now return to the notion of a spanning set and basis that incorporates the inner product.

Definition 10 (Spanning Set, Basis)**.** *A set of vectors S spans an inner product space V if there is no other non-zero vector $\psi \in V$, with $(\psi, \phi) = 0, \forall \phi \in S$. It forms a basis if no vector can be removed from S without changing the span.*

With the mathematical apparatus surrounding complex inner product spaces defined we can now define *the state space and state vectors of a quantum mechanical system*.

Definition 11 (State Space and State Vectors of a Quantum Mechanical System)**.** *The state space of a quantum mechanical system is a (complex) Hilbert space. A state vector of a quantum mechanical system is a vector in the state space having unit magnitude.*

From the point of view of physics, two state vectors are "the same" if one is equal to the other multiplied by a complex number. In this sense a state vector is often referred to as *a ray in Hilbert space*[2].

[2] This is a very important point, and is related to "normalization".

1.2 Linear Operators

Linear operators on the state space describe both measurable physical quantities (i.e. "observables") and how state vectors change in time (i.e. "dynamics"). In particular, we have the following definition.

Definition 12 (Observable)**.** *Observables in quantum mechanics are described by self-adjoint linear operators.*

However, we are getting a bit ahead of ourselves with this definition. First, we need to define the notion of "linear operator" and then the notion of a "self-adjoint linear operator"[3].

Definition 13 (Linear Operator)**.** *A linear operator A on a vector space V assigns to vectors $\psi \in V$ a vector $A\psi \in V$ such that*

$$A(\alpha\psi + \beta\phi) = \alpha A\psi + \beta A\phi, \quad \forall \alpha, \beta \in \mathbb{C},\ \psi, \phi \in V.$$

We now consider some examples of linear operators.

Example. Consider an appropriately defined (i.e. it can be done, but we are not going to work out the details at this point) vector space of complex valued differentiable functions of a real, scalar variable x. Then a linear operator on this vector space can be defined by differentiation as follows:

$$(A\psi)(x) = -i\frac{d\psi}{dx}(x).$$

We leave it as an exercise to show that this operator satisfies the condition of linearity given in Definition 13.

Example. Consider the complex vector space $\mathbb{C}^3$, with basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$. Recall from linear algebra (and this is a very important point to recall) that a linear operator on a vector space is defined through the action of the linear operator on each basis element.[4] In particular, for this example suppose that we have:

$$\begin{aligned} A\mathbf{e}_1 &= 3\mathbf{e}_1 + 2i\mathbf{e}_2, \\ A\mathbf{e}_2 &= -2i\mathbf{e}_1 - \mathbf{e}_2, \\ A\mathbf{e}_3 &= \mathbf{e}_3. \end{aligned} \tag{1.17}$$

More generally, suppose that $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ is a basis of $\mathbb{C}^n$, and let $A : \mathbb{C}^n \to \mathbb{C}^n$ be a linear operator. Now $A\mathbf{e}_i$ is a vector in $\mathbb{C}^n$,

[3] We want to address the terminology of "self-adjoint" versus "Hermitian". Mathematicians tend to prefer "self-adjoint" and physicists prefer "Hermitian". Do they mean the same thing? For finite dimensional, complex vector spaces equipped with an inner product (which encompasses a great deal of the course) there is absolutely no difference in what these two terms mean. In infinite dimensions, however, there are "subtle" differences depending on issues related to boundedness versus unboundedness of the operators, and the nature of the domains on which the operators are defined (which may be dictated by boundary conditions). This is something that you should be aware of, but it is not necessary for us to go into such details here. More details about this issue, including a number of enlightening examples, can be found here https://arxiv.org/pdf/quant-ph/9907069.pdf.

[4] This is an important remark that lies at the foundations of much of what we would like to achieve in quantum mechanics. The judicious (and sometimes fortuitous) choice of a basis is central to our ability to understand and ??solve?? quantum mechanical problems. We will see this in numerous situations throughout this course.

for each $i = 1, \ldots, n$. Therefore, it can be represented as a linear combination of the basis elements as follows:

$$A\mathbf{e}_i = \sum_{j=1}^{n} A_{ji}\mathbf{e}_j, \tag{1.18}$$

where $\{A_{ij}\}$ (note the reversal of i and j) is the matrix representation of A with respect to the basis $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$.

It should be familiar that the matrix representation of an operator (with respect to a chosen basis) can be used to transform the components of a vector that is represented in the same basis. We demonstrate this fact.

Let

$$\mathbf{x} = \sum_{i=1}^{n} x_i\mathbf{e}_i \quad \text{and} \quad \mathbf{y} = A\mathbf{x} = \sum_{j=1}^{n} y_j\mathbf{e}_j.$$

Then, using (1.18), we can express $\mathbf{y} = A\mathbf{x}$ as:

$$\mathbf{y} = \sum_{i=1}^{n} x_i A\mathbf{e}_i = \sum_{i,j=1}^{n} x_i A_{ji}\mathbf{e}_j,$$

from which it follows that:[5]

$$y_j = \sum_{i=1}^{n} A_{ji}x_i. \tag{1.19}$$

[5] Compare (1.18) and (1.19). This shows that basis vectors transform differently than the components of vectors. This is important to keep in mind when computing matrix representations of linear operators with respect to a basis.

Note that if $\{\mathbf{e}_i\}$ is an orthonormal basis then we have:

$$(\mathbf{e}_j, A\mathbf{e}_i) = \left(\mathbf{e}_j, \sum_{k=1}^{n} A_{ki}\mathbf{e}_k\right) = \sum_{k=1}^{n} A_{ki}(\mathbf{e}_j, \mathbf{e}_k) = \sum_{k=1}^{n} A_{ki}\delta_{jk} = A_{ji}.$$

Example. Suppose $\{\mathbf{e}_1, \mathbf{e}_2\}$ is an orthonormal basis of $\mathbb{C}^2$ and let $A : \mathbb{C}^2 \to \mathbb{C}^2$ be a linear operator. Suppose we have:

$$\begin{aligned} A\mathbf{e}_1 &= 3\mathbf{e}_1 + 2i\mathbf{e}_2, \\ A\mathbf{e}_2 &= \mathbf{e}_2. \end{aligned} \tag{1.20}$$

Then the matrix representation of A with respect to the basis $\{\mathbf{e}_1, \mathbf{e}_2\}$ is given by:

$$A = \begin{pmatrix} (\mathbf{e}_1, A\mathbf{e}_1) & (\mathbf{e}_1, A\mathbf{e}_2) \\ (\mathbf{e}_2, A\mathbf{e}_1) & (\mathbf{e}_2, A\mathbf{e}_2) \end{pmatrix} = \begin{pmatrix} 3 & 0 \\ 2i & 1 \end{pmatrix}$$

Now let

$$\mathbf{x} = 2\mathbf{e}_1 + i\mathbf{e}_2,$$

and let us choose the following explicit representations for $\mathbf{e}_1$ and $\mathbf{e}_2$:

$$\mathbf{e}_1 \to \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad \mathbf{e}_2 \to \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Then

$$\mathbf{x} = \begin{pmatrix} 2 \\ i \end{pmatrix},$$

and

$$\mathbf{y} = A\mathbf{x} = \begin{pmatrix} 3 & 0 \\ 2i & 1 \end{pmatrix} \begin{pmatrix} 2 \\ i \end{pmatrix} = \begin{pmatrix} 6 \\ 5i \end{pmatrix} = 6\mathbf{e}_1 + 5i\mathbf{e}_2.$$

It is important to keep in mind that the matrix representation of a linear operator depends on the chosen basis, i.e. for two bases $\{\mathbf{e}_i\}$ and $\{\mathbf{f}_i\}$, in general $(\mathbf{f}_i, A\mathbf{f}_j) \neq (\mathbf{e}_i, A\mathbf{e}_j)$.

The following idea will be useful.

Definition 14 (Bounded Linear Operator)**.** *A linear operator A defined on V is bounded if there is a positive number d such that $\| A\psi \| \leq d \| \psi \|$ for every $\psi \in V$. The smallest number with this property is called the operator norm of A and is denoted by $\| A \|$.*

Every linear operator on a *finite dimensional* Hilbert space is bounded, as we now show. Let A be a linear operator defined on an n dimensional vector space V that is equipped with an orthonormal basis $\{e_1, \ldots, e_n\}$. Let $A_{ij} \equiv (e_i, Ae_j)$ and let b be the largest of the numbers $|\sum_{i=1}^{n} \overline{A}_{ij} A_{ik}|$ for $j, k = 1, \ldots, n$. Then for any $\mathbf{x} \in V$ we set $\mathbf{y} = A\mathbf{x}$, and consider the following calculation:

$$
\begin{aligned}
\| A\mathbf{x} \| &= \sum_{i=1}^{n} \bar{y}_i y_i, \\
&= \sum_{i=1}^{n} \sum_{k=1}^{n} \bar{A}_{ik} \bar{x}_k \sum_{j=1}^{n} A_{ij} x_j, \\
&= \sum_{j,k=1}^{n} \bar{x}_k \left(\sum_{i=1}^{n} \bar{A}_{ik} A_{ij} \right) x_j, \\
&\leq \sum_{j,k=1}^{n} | \bar{x}_k | \left| \sum_{i=1}^{n} \bar{A}_{ik} A_{ij} \right| | x_j |, \\
&\leq b \left(\sum_{k=1}^{n} |x_k|^2 \right), \\
&\leq b \max_k |x_k|^2 \, n^2, \\
&\leq b n^2 \sum_{k=1}^{n} |x_k|^2, \\
&= b n^2 \| \mathbf{x} \| . \qquad (1.21)
\end{aligned}
$$

1.3 *Self-Adjoint Operators, Eigenvalues and Eigenvectors*

As we noted earlier, self-adjoint operators[6] play a central role since they are the mathematical manifestation of "observables" in quantum mechanics. In this section we will discuss self-adjoint operators in more detail. We begin with the definition.

[6] When we use the word "operator" in this course we will always mean "linear operator".

Definition 15 (Adjoint of an Operator)**.** *Let V be a vector space equipped with an inner product and suppose $A : V \to V$ is a bounded linear operator. The adjoint of A, denoted by $A^\dagger$ is defined by:*

$$(\phi, A^\dagger \psi) = (A\phi, \psi), \qquad (1.22)$$

for all $\phi, \psi \in V$.

With A known, we can view (1.22) as an equation for the "unknown" $A^\dagger$. Hence, the question arises, does (1.22) have a solution, i.e. does the adjoint of a bounded linear operator exist? The answer is "yes", but answering this question is beyond the scope of this course. You will learn about the issues in more detail in a course on functional analysis.

In many situations, when we work with the adjoint of a linear operator on a finite dimensional (complex) vector space, V, we will work with its matrix representation. Here we derive an important property of the matrix elements.

Using (1.22) and the properties of the (complex) inner product, we have:

$$(\phi, A^\dagger \psi) = (A\phi, \psi) = \overline{(\psi, A\phi)}. \tag{1.23}$$

Now let $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ be an orthonormal basis of V. Using (1.23), we have:

$$A^\dagger_{jk} = (\mathbf{e}_j, A^\dagger \mathbf{e_k}) = \overline{(\mathbf{e}_k, A\mathbf{e}_j)} = \overline{A}_{kj}. \tag{1.24}$$

Hence, given a matrix A, to compute the adjoint of A we take the transpose of A and the complex conjugate of each matrix element.

If A is a bounded linear operator with a bounded inverse, A^{-1}, then $A^\dagger$ has an inverse given by $(A^\dagger)^{-1} = (A^{-1})^\dagger$. This follows from the following calculation and the definition of the adjoint given in Definition 15:

$$\begin{aligned} (\phi, \left(A^{-1}\right)^\dagger \psi) &= (A^{-1}\phi, \psi) = (A^{-1}\phi, A^\dagger (A^\dagger)^{-1}\psi), \\ &= (AA^{-1}\phi, (A^\dagger)^{-1}\psi) = (\phi, (A^\dagger)^{-1}\psi). \end{aligned}$$

Since this calculation holds for all ϕ, ψ, it follows that:

$$\left(A^{-1}\right)^\dagger = (A^\dagger)^{-1}.$$

Definition 16 (Self-adjoint, or Hermitian Operator)**.** *A bounded linear operator, A, defined on vector space V equipped with an inner product is said to be self-adjoint, or Hermitian, if*

$$(\psi, A\phi) = (A\psi, \phi), \qquad \forall \psi, \phi \in V. \tag{1.25}$$

Example. Let $\{\mathbf{e}_1, \mathbf{e}_2\}$ be an orthonormal basis defined on $\mathbb{C}^2$. We define a linear operator on $\mathbb{C}^2$ as follows:

$$\begin{aligned} A\mathbf{e}_1 &= 3\mathbf{e}_1 + 2i\mathbf{e}_2, \\ A\mathbf{e}_2 &= -2i\mathbf{e}_1 + \mathbf{e}_2. \end{aligned} \tag{1.26}$$

The matrix representation of A with respect to the basis $\{\mathbf{e}_1, \mathbf{e}_2\}$ is given by:

$$A = \begin{pmatrix} 3 & -2i \\ 2i & 1 \end{pmatrix} = \overline{A^T} = A^\dagger,$$

and we see that A is self-adjoint.

A useful identity is the following. Let A and B be bounded linear operators on a complex vector space equipped with an inner product. Then we have:

$$(AB)^\dagger = B^\dagger A^\dagger.$$

The proof of this identity is left as an exercise.

The notion of eigenvalues and eigenvectors plays a central role in both the mathematical structure and the physical interpretation of quantum mechanics.

Definition 17 (Eigenvector, Eigenvalue)**.** *ψ_n is an eigenvector of an operator A with eigenvalue λ_n if*

$$A\psi_n = \lambda_n \psi_n, \quad \lambda_n \in \mathbb{C}.$$

The position and momentum operators in three dimensions are fundamental operators in quantum mechanics.

Definition 18 (The Position and Momentum Operators in Three Dimensions)**.** *The position operator, X, in three dimensions has components X_j, $j = 1, 2, 3$, that are defined on functions ψ in a Hilbert space $\mathcal{H}$ by:*

$$(X_j\psi)(\mathbf{x}) = x_j\psi(\mathbf{x}). \tag{1.27}$$

The momentum operator P in three dimensions has components P_j, $j = 1, 2, 3$ that are defined on differentiable ψ in a Hilbert space $\mathcal{H}$ by

$$(P_j\psi)(\mathbf{x}) = \frac{\hbar}{i}\frac{\partial}{\partial x_j}\psi(\mathbf{x}). \tag{1.28}$$

One Space Dimension. In one dimension the position and momentum operators have particularly simple forms:

$$\begin{aligned} (X\psi)(x) &= x\psi(x), \\ P\psi(x) &= \frac{\hbar}{i}\frac{d\psi}{dx}. \end{aligned} \tag{1.29}$$

A Hilbert space that will be useful for describing one dimensional quantum mechanical systems is the space of functions defined on $\mathbb{R}$ (with particular functions denoted by $\psi(x)$) satisfying:

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

In other words, X and P are defined only for those functions $\psi(x)$ such that:

$$\begin{aligned} \| P\psi \|^2 &= \int_{\mathbb{R}} \left| \frac{\hbar}{i}\frac{d\psi}{dx} \right|^2 dx < \infty, \\ \| X\psi \|^2 &= \int_{\mathbb{R}} x^2 |\psi(x)|^2 dx < \infty. \end{aligned} \tag{1.30}$$

It is a simple calculation to show that the position and momentum operators are linear operators.[7]

The one dimensional position and momentum operators illustrate the issues that occur with unbounded operators. In particular, on the space $L^2(\mathbb{R}, dx)$ of square integrable functions X and P are not bounded. They are not defined for all functions in $L^2(\mathbb{R}, dx)$, but only on a subset of this Hilbert space. The problem of determining whether or not an unbounded operator is self-adjoint is complicated by the subset of the Hilbert space on which they are defined. On the subsets on which they are defined, these operators are self-adjoint. This can be understood from the following calculations:

$$(X\phi, \psi) = \int_{\mathbb{R}} \bar{x}\overline{\phi(x)}\psi(x)dx = \int_{\mathbb{R}} \overline{\phi(x)}x\psi(x)dx = (\phi, X\psi), \tag{1.31}$$

and

$$\begin{aligned}
(P\phi, \psi) &= \int_{\mathbb{R}} \overline{\left(\frac{\hbar}{i}\frac{d\phi}{dx}\right)}\psi dx, \\
&= -\frac{\hbar}{i}\int_{\mathbb{R}} \frac{d\bar{\phi}}{dx}\psi dx, \\
&= -\frac{\hbar}{i}\left([\bar{\phi}\psi]_{-\infty}^{\infty} - \int_{\mathbb{R}} \bar{\phi}\frac{d\psi}{dx}dx\right), \\
&= -\frac{\hbar}{i}[\bar{\phi}\psi]_{-\infty}^{\infty} + \int_{\mathbb{R}} \bar{\phi}\frac{\hbar}{i}\frac{d\psi}{dx}dx, \\
&= (\phi, P\psi).
\end{aligned} \tag{1.32}$$

These formal calculations "make sense" for functions $\chi(x) \in L^2(\mathbb{R}, dx)$ having the property that:

$$\lim_{x\to\pm\infty} \chi(x) = 0.$$

Next we define another common operator in quantum mechanics using the position and momentum operators that we have just defined.

Definition 19 (Hamiltonian Operator)**.** *For twice differentiable functions $\psi(\mathbf{x})$ the Hamiltonian operator is defined as:*

$$(H\psi)(\mathbf{x}) = -\frac{\hbar^2}{2m}\nabla^2\psi(\mathbf{x}) + V(\mathbf{x})\psi(\mathbf{x}), \tag{1.33}$$

where $V(\mathbf{x})$ is the potential energy function. In terms of the position and momentum operators defined in (1.27) and(1.28), respectively, (1.33) can be re-written as:

$$H = \frac{1}{2m}P^2 + V(X). \tag{1.34}$$

[7] The proof that the one dimensional position operator is linear is straightforward. For appropriate functions, $\psi(x)$, $\phi(x)$ and complex numbers α, $\beta \in \mathbb{C}$, we have:

$$\begin{aligned}
X(\alpha\psi(x) + \beta\phi(x)) &= x(\alpha\psi(x) + \beta\phi(x)), \\
&= x\alpha\psi(x) + x\beta\phi(x), \\
&= \alpha x\psi(x) + \beta x\phi(x), \\
&= \alpha X\psi(x) + \beta X\phi(x).
\end{aligned}$$

The proof that the one dimensional momentum operator is linear follows from the fact that differentiation is a linear operation.

It is easy to see that (1.34) is self-adjoint from the following calculation:

$$\begin{aligned} H^\dagger &= \left(\frac{1}{2m}P^2 + V(X)\right)^\dagger, \\ &= \frac{1}{2m}\left(P^2\right)^\dagger + (V(X))^\dagger, \\ &= \frac{1}{2m}(P^\dagger)^2 + V(X^\dagger), \\ &= \frac{1}{2m}(P)^2 + V(X) = H. \end{aligned}$$

The equality, $(V(X))^\dagger = V(X^\dagger)$, which is essential for the validity of the result, is taken as an *assumption*. It will be straightforward to verify that the potential energy functions considered in this course satisfy this requirement. It is an interesting mathematics problem to determine conditions on potential energy functions for which this equality holds, but that is beyond the scope of this course. [8]

1.4 Dirac Notation

In this section we will introduce a notation, due to Dirac[9], that renders many of the conceptual issues associated with computing quantum mechanical quantities in linear vector spaces simple and transparent. Throughout the rest of the course we will, almost exclusively, use *Dirac notation*, which we now describe.[10]

Let V be a complex vector space equipped with an inner product. Rather than denote vectors in V by ψ, we will denote them by

$$|\,\psi\rangle \in V, \tag{1.35}$$

which we refer to as a "ket", or "ket vector". Hence, the state vector of a quantum mechanical system is given by a normalised ket.

As an example, we express (1.26) in Dirac notation. In particular, we consider the complex vector space $\mathbb{C}^2$. We denote an orthonormal basis on $\mathbb{C}^2$ by $\{\mathbf{e}_1, \mathbf{e}_2\}$, or $\{|\,\mathbf{e}_1\rangle, |\,\mathbf{e}_2\rangle\}$. We can define a linear operator, A, on $\mathbb{C}^2$ by defining it's action on the basis vectors as follows[11]:

$$\begin{aligned} A\,|\,\mathbf{e}_1\rangle &= 3\,|\,\mathbf{e}_1\rangle + 2i\,|\,\mathbf{e}_2\rangle, \\ A\,|\,\mathbf{e}_2\rangle &= -2i\,|\,\mathbf{e}_1\rangle + |\,\mathbf{e}_2\rangle. \end{aligned} \tag{1.36}$$

For $|\,\phi\rangle, |\,\psi\rangle \in V$ we denote their inner product by:

$$\langle\phi\,|\,\psi\rangle \equiv (|\,\phi\rangle, |\,\psi\rangle)\,, \quad \text{(where the previous notation was } (\phi,\psi)\text{)}\,, \tag{1.37}$$

[8] In this definition we see a procedure for associating a quantum mechanical Hamiltonian operator with a classical Hamiltonian function. In particular, in the classical Hamiltonian function we merely replace the momentum and position variables by their quantum mechanical operator counterparts. This method of canonical quantization starting from the classical Hamiltonian framework was pioneered by Dirac .

P. A. M. Dirac. *The principles of quantum mechanics*. Number 27. Oxford university press, 1981

[9] P. A. M. Dirac. A new notation for quantum mechanics. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 35, pages 416–418. Cambridge University Press, 1939

[10] Computations involving the inner product and the adjoint play an important role in quantum mechanics. Dirac notation serves to simplify the set-up for these computations in a way that makes them almost obvious. You should get a sense of that in this section, but it should become particularly apparent in Chapter 3.

[11] Note that this example is just a reformulation of (1.26) in Dirac notation.

where $\langle\phi \mid$ is referred to as a "bra". Hence the inner product pairing defined by (1.37) is referred to as a "bra-ket" (a misspelling of "bracket").

Example. Consider the complex vector space $\mathbb{C}^2$ and let $\{\mid \mathbf{e}_1\rangle, \mid \mathbf{e}_2\rangle\}$ denote an orthonormal basis on $\mathbb{C}^2$, i.e. $\langle \mathbf{e}_i \mid \mathbf{e}_j\rangle = \delta_{ij}$. Then arbitrary vectors $\mid \psi\rangle, \mid \phi\rangle \in \mathbb{C}^2$ can be written as:

$$\begin{aligned} \mid \psi\rangle &= a \mid \mathbf{e}_1\rangle + b \mid \mathbf{e}_2\rangle, \\ \mid \phi\rangle &= c \mid \mathbf{e}_1\rangle + d \mid \mathbf{e}_2\rangle, \end{aligned} \tag{1.38}$$

and we have:

$$\begin{aligned} \langle\phi \mid \psi\rangle = (\mid \phi\rangle, \mid \psi\rangle) &= (c \mid \mathbf{e}_1\rangle + d \mid \mathbf{e}_2\rangle, a \mid \mathbf{e}_1\rangle + b \mid \mathbf{e}_2\rangle)\,, \\ &= a\,(c \mid \mathbf{e}_1\rangle + d \mid \mathbf{e}_2\rangle, \mid \mathbf{e}_1\rangle) \\ &\quad + b\,(c \mid \mathbf{e}_1\rangle + d \mid \mathbf{e}_2\rangle, \mid \mathbf{e}_2\rangle)\,, \\ &= a\bar{c}\,(\mid \mathbf{e}_1\rangle, \mid \mathbf{e}_1\rangle) + a\bar{d}\,(\mid \mathbf{e}_2\rangle, \mid \mathbf{e}_1\rangle) \\ &\quad + b\bar{c}\,(\mid \mathbf{e}_1\rangle, \mid \mathbf{e}_2\rangle) + b\bar{d}\,(\mid \mathbf{e}_2\rangle, \mid \mathbf{e}_2\rangle)\,, \\ &= a\bar{c}\,\langle \mathbf{e}_1 \mid \mathbf{e}_1\rangle + a\bar{d}\,\langle \mathbf{e}_2 \mid \mathbf{e}_1\rangle \\ &\quad + b\bar{c}\,\langle \mathbf{e}_1 \mid \mathbf{e}_2\rangle + b\bar{d}\,\langle \mathbf{e}_2 \mid \mathbf{e}_2\rangle, \\ &= a\bar{c} + b\bar{d} = (\bar{c}\ \bar{d}) \begin{pmatrix} a \\ b \end{pmatrix} \end{aligned} \tag{1.39}$$

With respect to the basis $\{\mid \mathbf{e}_1\rangle, \mid \mathbf{e}_2\rangle\}$ we can make the following identifications that should be familiar from your experience with matrix algebra:

$$\mid \phi\rangle \quad \leftarrow \quad \begin{pmatrix} c \\ d \end{pmatrix}, \tag{1.40}$$

$$\langle\phi \mid \quad \leftarrow \quad (\bar{c}\ \bar{d}). \tag{1.41}$$

In the context of this example, we can examine linear operators and Dirac notation. Let A be a linear operator on $\mathbb{C}^2$. We consider $A \mid \psi\rangle$ and the inner product of this vector with another vector $\mid \phi\rangle$:

$$(\mid \phi\rangle,\, A \mid \psi\rangle) = \langle\phi \mid A \mid \psi\rangle. \tag{1.42}$$

Note that we have:

$$A_{ij} = \langle \mathbf{e}_i \mid A \mid \mathbf{e}_j\rangle, \quad \text{which is the same as} \quad (\mathbf{e}_i . A\mathbf{e}_j),\, (\mid \mathbf{e}_i\rangle, A\mathbf{e}_j\rangle).$$

We derived this expression for the matrix elements of a linear operator with respect to a given basis earlier. Here we show how the same relation holds in Dirac notation. We can express the action of A on each basis element as follows:

$$A \mid \mathbf{e}_j\rangle = \sum_k A_{kj} \mid \mathbf{e}_k\rangle.$$

Using this relation, it follows that:

$$\langle \mathbf{e}_i \mid A \mid \mathbf{e}_j\rangle = \sum_k A_{kj}\langle \mathbf{e}_i \mid \mathbf{e}_k\rangle = \sum_k A_{kj}\delta_{ik} = A_{ij}. \tag{1.43}$$

As noted earlier, any $\mid \phi\rangle, \mid \psi\rangle \in \mathbb{C}^2$ can be expressed as:

$$\begin{aligned} \mid \phi\rangle &= c \mid \mathbf{e}_1\rangle + d \mid \mathbf{e}_2\rangle, \\ \mid \psi\rangle &= a \mid \mathbf{e}_1\rangle + b \mid \mathbf{e}_2\rangle. \end{aligned} \tag{1.44}$$

Then if:

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \tag{1.45}$$

we have:

$$\langle \phi \mid A \mid \psi\rangle = (\bar{c}\,\bar{d}) \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} \tag{1.46}$$

Operators in Dirac Notation. Let V be a complex vector space equipped with an inner product, and consider the "kets" $\mid u\rangle, \mid v\rangle, \mid w\rangle \in V$. Using a bra and a ket, we define a linear operator as follows:

$$B = \mid u\rangle\langle v \mid .$$

B acts on kets, $\mid w\rangle \in V$, as follows:

$$B \mid w\rangle = \mid u\rangle \underbrace{\langle v \mid w\rangle}_{\text{complex number}} = \underbrace{\langle v \mid w\rangle}_{\text{complex number}} \mid u\rangle.$$

In particular, let $V = \mathbb{C}^n$ and let $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ be an orthonormal basis of $\mathbb{C}^n$, i.e. $\langle \mathbf{e}_i \mid \mathbf{e}_j\rangle = \delta_{ij}$. Then define B as follows:

$$B = \mid \mathbf{e}_2\rangle\langle \mathbf{e}_1 \mid .$$

Then we have:

$$\begin{aligned} B \mid \mathbf{e}_1\rangle &= \mid \mathbf{e}_2\rangle\langle \mathbf{e}_1 \mid \mathbf{e}_1\rangle = \mid \mathbf{e}_2\rangle, \\ B \mid \mathbf{e}_2\rangle &= \mid \mathbf{e}_2\rangle\langle \mathbf{e}_1 \mid \mathbf{e}_2\rangle = \mid \mathbf{e}_2\rangle \cdot 0 = 0. \end{aligned} \tag{1.47}$$

Since the basis is orthonormal we can represent an arbitrary linear operator on $\mathbb{C}^n$ as follows:

$$A = \sum_{ij} A_{ij} \mid \mathbf{e}_i\rangle\langle \mathbf{e}_j \mid, \tag{1.48}$$

and using the fact that a linear operator can be defined by specifying how it acts on each basis element. In this case, it is as follows:

$$\begin{aligned} A \mid \mathbf{e}_k\rangle &= \sum_{ij} A_{ij} \mid \mathbf{e}_i\rangle\langle \mathbf{e}_j \mid \mathbf{e}_k\rangle, \\ &= \sum_{ij} A_{ij} \mid \mathbf{e}_i\rangle \delta_{jk}, \\ &= \sum_{i} A_{ik} \mid \mathbf{e}_i\rangle. \end{aligned} \tag{1.49}$$

In particular, we consider the special case of $\mathbb{C}^2$. In this case we specify a linear operator acting on each basis element in the following manner:

$$\begin{aligned} A \mid \mathbf{e}_1\rangle &= a \mid \mathbf{e}_1\rangle + c \mid \mathbf{e}_2\rangle, \\ A \mid \mathbf{e}_2\rangle &= b \mid \mathbf{e}_1\rangle + d \mid \mathbf{e}_2\rangle, \end{aligned} \tag{1.50}$$

from which it follows that the matrix representation of the operator with respect to this particular basis is given by:

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

From (1.49), we have:

$$A = a \mid \mathbf{e}_1\rangle\langle \mathbf{e}_1 \mid + b \mid \mathbf{e}_1\rangle\langle \mathbf{e}_2 \mid + c \mid \mathbf{e}_2\rangle\langle \mathbf{e}_1 \mid + d \mid \mathbf{e}_2\rangle\langle \mathbf{e}_2 \mid . \tag{1.51}$$

This expression can be verified by considering the action of A in this representation on each basis element:

$$\begin{aligned} A \mid \mathbf{e}_1\rangle &= a \mid \mathbf{e}_1\rangle\langle \mathbf{e}_1 \mid \mathbf{e}_1\rangle + b \mid \mathbf{e}_1\rangle\langle \mathbf{e}_2 \mid \mathbf{e}_1\rangle \\ &\quad + c \mid \mathbf{e}_2\rangle\langle \mathbf{e}_1 \mid \mathbf{e}_1\rangle + d \mid \mathbf{e}_2\rangle\langle \mathbf{e}_2 \mid \mathbf{e}_1\rangle, \\ &= a \mid \mathbf{e}_1\rangle + c \mid \mathbf{e}_2\rangle. \end{aligned} \tag{1.52}$$

Similarly,

$$A \mid \mathbf{e}_2\rangle = b \mid \mathbf{e}_1\rangle + d \mid \mathbf{e}_2\rangle. \tag{1.53}$$

Adjoints in Dirac Notation. Recall that we showed earlier that given the matrix representation of a linear operator, the matrix representation of the adjoint of the linear operator (with respect to the same basis) is obtained by taking the transpose of the matrix, and then the complex conjugate of each matrix element. For example, if we consider a linear operator A on a two dimensional complex vector space and its matrix representation with respect to some basis:

$$A \leftrightarrow \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \tag{1.54}$$

then the matrix representation of the adjoint of A with respect to the same basis is given by:

$$A^\dagger \leftrightarrow \begin{pmatrix} \bar{a} & \bar{c} \\ \bar{b} & \bar{d} \end{pmatrix}. \tag{1.55}$$

We can now see how this fact would be manifested in Dirac notation. Consider

$$A = a \mid \mathbf{e}_1\rangle\langle\mathbf{e}_1 \mid + b \mid \mathbf{e}_1\rangle\langle\mathbf{e}_2 \mid + c \mid \mathbf{e}_2\rangle\langle\mathbf{e}_1 \mid + d \mid \mathbf{e}_2\rangle\langle\mathbf{e}_2 \mid, \tag{1.56}$$

then in order for the matrix representation to agree with (1.55) we must have:

$$A^\dagger = \bar{a} \mid \mathbf{e}_1\rangle\langle\mathbf{e}_1 \mid + \bar{b} \mid \mathbf{e}_2\rangle\langle\mathbf{e}_1 \mid + \bar{c} \mid \mathbf{e}_1\rangle\langle\mathbf{e}_2 \mid + \bar{d} \mid \mathbf{e}_2\rangle\langle\mathbf{e}_2 \mid . \tag{1.57}$$

Note how each term in (1.56) transforms under the adjoint operation. For example, if we have $A = a \mid \psi\rangle\langle\phi \mid$, then:

$$A^\dagger = (a \mid \psi\rangle\langle\phi \mid)^\dagger = \bar{a} \mid \phi\rangle\langle\psi \mid .$$

We can also see this from the following calculation:

$$\begin{aligned}
\langle u \mid A^\dagger \mid v\rangle &= \langle Au \mid v\rangle, \quad \text{(definition of the adjoint in Dirac notation)}, \\
&= \overline{\langle v \mid A \mid u\rangle}, \quad \text{(property of the inner product in Dirac notation)}, \\
&= \overline{\langle v \mid (a \mid \psi\rangle\langle\phi \mid) \mid u\rangle}, \\
&= \bar{a}\overline{\langle v \mid \psi\rangle\langle\phi \mid u\rangle}, \\
&= \bar{a}\overline{\langle v \mid \psi\rangle}\ \overline{\langle\phi \mid u\rangle}, \\
&= \bar{a}\langle\psi \mid v\rangle\langle u \mid \phi\rangle, \\
&= \bar{a}\langle u \mid \phi\rangle\langle\psi \mid v\rangle, \\
&= \langle u \mid (\bar{a} \mid \phi\rangle\langle\psi \mid) \mid v\rangle,
\end{aligned} \tag{1.58}$$

and since this is true for any $\mid u\rangle$, $\mid v\rangle$ we can conclude that:

$$A^\dagger = (a \mid \psi\rangle\langle\phi \mid)^\dagger = \bar{a} \mid \phi\rangle\langle\psi \mid .$$

In particular, if we apply this result to the operator:

$$\mid \mathbf{e}_1\rangle\langle\mathbf{e}_2 \mid + \mid \mathbf{e}_2\rangle\langle\mathbf{e}_1 \mid,$$

we see that this operator is self-adjoint (you will need to use the fact that the "adjoint of the sum of two operators is the sum of the adjoint of each operator", but you showed that in the exercises).

Moreover, this calculation is also consistent with the following identity:

$$(\mid \psi\rangle)^\dagger = \langle\psi \mid .$$

Now we want to determine the bra associated with the ket $A \mid \psi\rangle$. Towards this end, we let $\mid u\rangle = A \mid \psi\rangle$, and then we perform the following calculation:

$$\begin{aligned} \langle u \mid v\rangle &= \overline{\langle v \mid u\rangle}, \\ &= \overline{\langle v \mid A \mid \psi\rangle}, \\ &= \langle\psi \mid A^\dagger \mid v\rangle, \end{aligned} \tag{1.59}$$

which is true for any $\mid v\rangle$. Hence, we have:

$$\langle u \mid = (A \mid \psi\rangle)^\dagger = \langle\psi \mid A^\dagger. \tag{1.60}$$

1.5 *Projection Operators and the Spectral Theorem*

Next we define the notion of a projection operator. These play an important role in quantum mechanics, especially with respect to the notion of *measurement*.

Definition 20 (Projection Operator)**.** *Let $P : V \to V$ be a linear operator on a vector space V. Then P is said to be a projection operator if $P^2 = P$.*

In Dirac notation an example of a projection operator would be a "bra-ket combination" of the form $\mid \phi\rangle\langle\phi \mid$, where the ket $\mid \phi\rangle$ is normalised to have norm one. Note that projection operators of this form are *self-adjoint*.

Example. Let $\{\mid \mathbf{e}_1\rangle, \mid \mathbf{e}_2\rangle\}$ be an orthonormal basis of $\mathbb{C}^2$. and consider the projection operator defined as $P = \mid \mathbf{e}_1\rangle\langle\mathbf{e}_1 \mid$ and let it act on $\mid \psi\rangle = a \mid \mathbf{e}_1\rangle + b \mid \mathbf{e}_2\rangle$. Then we have:

$$P(a \mid \mathbf{e}_1\rangle + b \mid \mathbf{e}_2\rangle) = (\mid \mathbf{e}_1\rangle\langle\mathbf{e}_1 \mid)(a \mid \mathbf{e}_1\rangle + b \mid \mathbf{e}_2\rangle) = a \mid \mathbf{e}_1\rangle.$$

Note that we have:

$$P^2 =| \mathbf{e}_1\rangle\langle\mathbf{e}_1 \mid \mathbf{e}_1\rangle\langle\mathbf{e}_1 |=| \mathbf{e}_1\rangle\langle\mathbf{e}_1 |= P \quad \text{and} \quad P^\dagger =| \mathbf{e}_1\rangle\langle\mathbf{e}_1 |= P.$$

More generally, it is easy to see that for an orthonormal basis of $\mathbb{C}^n$:

$$| \mathbf{e}_1\rangle, \ldots, | \mathbf{e}_n\rangle \in \mathbb{C}^n,$$

the operator defined by:

$$\sum_{k=1}^{d} | \mathbf{e}_k\rangle\langle\mathbf{e}_k |,$$

is a self adjoint projection operator onto the subspace of dimension d spanned by $\mathbf{e}_1, \ldots, \mathbf{e}_d$, $d \leq n$.

As another example, consider $\mathbb{C}^3$ with the orthonormal basis $\{| \mathbf{e}_1, | \mathbf{e}_2\rangle, | \mathbf{e}_3\rangle\}$. Then define the self adjoint operator

$$P =| \mathbf{e}_1\rangle\langle\mathbf{e}_1 | + | \mathbf{e}_2\rangle\langle\mathbf{e}_2 |,$$

and consider it action on an arbitrary vector

$$| \psi\rangle = a | \mathbf{e}_1\rangle + b | \mathbf{e}_2\rangle + c | \mathbf{e}_3\rangle.$$

Doing this calculation, we obtain:

$$\begin{aligned} P | \psi\rangle &= (| \mathbf{e}_1\rangle\langle\mathbf{e}_1 | + | \mathbf{e}_2\rangle\langle\mathbf{e}_2 |)(a | \mathbf{e}_1\rangle + b | \mathbf{e}_2\rangle + c | \mathbf{e}_3\rangle), \\ &= = a | \mathbf{e}_1\rangle + b | \mathbf{e}_2\rangle, \end{aligned} \tag{1.61}$$

which is the projection of $| \psi\rangle$ onto the subspace of $\mathbb{C}^3$ spanned by $| \mathbf{e}_1\rangle$ and $| \mathbf{e}_2\rangle$. Moreover, it should be easy to see that for this example we have: $P^2 = P$ and $P^\dagger = P$.

Remark. If $\{| \mathbf{e}_1\rangle, \ldots, | \mathbf{e}_n\rangle\}$ is an orthonormal basis of $\mathbb{C}^n$, then the operator:

$$\mathbb{I} = \sum_{i=1}^{n} | \mathbf{e}_i\rangle\langle\mathbf{e}_i |,$$

is a self adjoint projection operator onto the entire space. This can be seen as follows. Consider an arbitrary vector, $| \psi\rangle \in \mathbb{C}^n$:

$$| \psi\rangle = \sum_{i=1}^{n} a_i | \mathbf{e}_i\rangle.$$

Then we have:

$$\mathbb{I} | \psi\rangle = \sum_{ij} a_j | \mathbf{e}_i\rangle\langle\mathbf{e}_i | \mathbf{e}_j\rangle = \sum_{ij} a_j | \mathbf{e}_i\rangle\delta_{ij} = \sum_{i} a_i | \mathbf{e}_i\rangle =| \psi\rangle. \tag{1.62}$$

The following result describes some extremely important properties of self-adjoint operators that play a fundamental role in quantum mechanics.

Theorem 1 (Spectral Theorem for Finite Dimensional Self-Adjoint Operators)**.** *Let $A : V \to V$ be a self-adjoint linear operator on a finite dimensional complex inner product space. Then:*

i) A has real eigenvalues, and the eigenvectors of A corresponding to distinct eigenvalues are orthogonal.

i) The eigenvectors of A span V.

Proof. We begin by showing that A has real eigenvalues. Let $|\,\mathbf{e}\rangle$ be a normalised eigenvector of A with eigenvalues λ, i.e.,

$$A\,|\,\mathbf{e}\rangle = \lambda\,|\,\mathbf{e}\rangle, \qquad \langle\mathbf{e}\,|\,\mathbf{e}\rangle = 1.$$

Then we have:

$$\begin{aligned}
\lambda = \langle\mathbf{e}\,|\,A\,|\,\mathbf{e}\rangle &= \langle A^\dagger\mathbf{e}\,|\,\mathbf{e}\rangle, \\
&= \overline{\langle\mathbf{e}\,|\,A^\dagger\,|\,\mathbf{e}\rangle}, \\
&= \overline{\langle\mathbf{e}\,|\,A\,|\,\mathbf{e}\rangle}, \quad \text{since} \quad A = A^\dagger, \\
&= \overline{\langle\mathbf{e}\,|\,\lambda\,|\,\mathbf{e}\rangle}, \\
&= \bar{\lambda}\,\overline{\langle\mathbf{e}\,|\,\mathbf{e}\rangle}, \\
&= \bar{\lambda}, \qquad (1.63)
\end{aligned}$$

from which it follows that λ is real.

Next we show that eigenvectors corresponding to distinct eigenvalues are orthogonal, i.e. we consider eigenvectors $|\,\mathbf{e}_1\rangle$, $|\,\mathbf{e}_2\rangle$ such that:

$$\begin{aligned}
A\,|\,\mathbf{e}_1\rangle &= \lambda_1\,|\,\mathbf{e}_1\rangle, \\
A\,|\,\mathbf{e}_2\rangle &= \lambda_2\,|\,\mathbf{e}_2\rangle. \qquad (1.64)
\end{aligned}$$

Then we have:

$$\begin{aligned}
\lambda_1\langle\mathbf{e}_2\,|\,\mathbf{e}_1\rangle &= \langle\mathbf{e}_2\,|\,A\,|\,\mathbf{e}_1\rangle \\
&= \langle A^\dagger\mathbf{e}_2\,|\,\mathbf{e}_1\rangle \\
&= \overline{\langle\mathbf{e}_1\,|\,A^\dagger\,|\,\mathbf{e}_2\rangle}, \\
&= \overline{\langle\mathbf{e}_1\,|\,A\,|\,\mathbf{e}_2\rangle}, \quad \text{since} \quad A = A^\dagger, \\
&= \bar{\lambda}_2\overline{\langle\mathbf{e}_1\,|\,\mathbf{e}_2\rangle}, \\
&= \lambda_2\langle\mathbf{e}_2\,|\,\mathbf{e}_1\rangle. \qquad (1.65)
\end{aligned}$$

Hence, we have:

$$(\lambda_1 - \lambda_2)\langle \mathbf{e}_2 \mid \mathbf{e}_1\rangle = 0 \quad \Rightarrow \langle \mathbf{e}_2 \mid \mathbf{e}_1\rangle = 0.$$

Now we return to the final part of this result–" eigenvectors of A span V". The issue we need to deal with here is the possibility of two (or more) eigenvectors having the same eigenvalue, i.e. *degeneracy*, since if we have n distinct eigenvalues, we have n distinct eigenvectors, and we have proven above that the eigenvectors are orthogonal.

First, we note that in linear algebra it is proven that a self-adjoint operator on a n-dimensional, complex inner produce space has n orthogonal eigenvectors. What is *not* necessarily true is that the eigenvalues corresponding to each of these eigenvectors are all different. What we have shown above is that if we have n distinct eigenvalues, then we have n orthogonal eigenvectors.

We examine this situation more closely assuming that we have n orthogonal eigenvectors. Suppose the eigenvectors $\mid \mathbf{e}_1\rangle$, $\mid \mathbf{e}_2\rangle$ have the same eigenvalue, λ, i.e.

$$\begin{aligned} A \mid \mathbf{e}_1\rangle &= \lambda \mid \mathbf{e}_1\rangle, \\ A \mid \mathbf{e}_2\rangle &= \lambda \mid \mathbf{e}_2\rangle. \end{aligned} \tag{1.66}$$

Then it follows that any linear combination of $\mid \mathbf{e}_1\rangle$ and $\mid \mathbf{e}_2\rangle$ is also an eigenvector having eigenvalue λ:

$$A(\alpha \mid \mathbf{e}_1\rangle + \beta \mid \mathbf{e}_2\rangle) = \lambda(\alpha \mid \mathbf{e}_1\rangle + \beta \mid \mathbf{e}_2\rangle).$$

This calculation can easily be generalised to the case where more than two eigenvectors have the same eigenvalue. Hence we can conclude that the the set of eigenvectors corresponding to the same eigenvalue spans a subspace, V_λ, of V. We let P_λ denote the projection onto the subspace V_λ. It can be proven that P_λ is self adjoint. This same procedure can be carried out for each eigenvalue having more than one eigenvector. In this way V is represented as the direct sum of subspaces corresponding to the span of the eigenvectors corresponding to a given eigenvalue. We can also construct self adjoint projection operators corresponding to each subspace. The case where the eigenvalues are distinct is particularly simple. Each of the subspaces in that case are just one dimensional, and the projection operators are easily constructed from the bra and ket associated with each eigenvalue in the manner that we have already described.

□

Before proceeding with the remainder of the proof of this theorem, we make an important remark that uses what we have shown up to this point.

In completing this section we describe a very important idea for representing a self adjoint operator based on the above theorem.

Spectral representation of A in Dirac notation. We have shown that if A is a self-adjoint linear operator on $\mathbb{C}^n$ having distinct eigenvalues, then its normalised eigenvectors, denoted $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ form a basis of $\mathbf{C}^n$. We can therefore write A in the form:[12]

$$A = \sum_{i=1}^{n} \lambda_i \mid \mathbf{e}_i\rangle\langle \mathbf{e}_i \mid, \tag{1.67}$$

[12] Note that the spectral representation of an operator in Dirac notation is the Dirac notation manifestation of the diagonalization of an operator in a basis of eigenvectors.

where λ_i is the eigenvalue corresponding to the eigenvector $\mid \mathbf{e}_i\rangle$ and $\mid \mathbf{e}_i\rangle\langle \mathbf{e}_i \mid$ is the projection operator onto the subspace spanned by $\mid \mathbf{e}_i\rangle$. This can be verified with the following "check" of how A acts on each basis vector (as discussed above):

$$\begin{aligned} A \mid \mathbf{e}_k\rangle &= \sum_{i=1}^{n} \lambda_i \mid \mathbf{e}_i\rangle\langle \mathbf{e}_i \mid \mathbf{e}_k\rangle, \\ &= \sum_{i=1}^{n} \lambda_i \mid \mathbf{e}_i\rangle\delta_{ik}, . \\ &= \lambda_k \mid \mathbf{e}_k\rangle \end{aligned} \tag{1.68}$$

In the case where the eigenvalues are not necessarily distinct we can write A in the form:

$$A = \sum_{\alpha} \lambda_\alpha P_\alpha, \tag{1.69}$$

where P_α is the projection operator onto the subspace of eigenvectors corresponding to the eigenvalue λ_α. You should convince yourself that (1.69) takes the form (1.67) when the eigenvalues are distinct.

Problems

1. *Triangle inequality, bounded operators.*

   (a) Use Schwarz's inequality for the inner product on a Hilbert space, to prove the triangle inequality, $||\psi + \phi|| \leq ||\psi|| + ||\phi||$.

   (b) If A and B are bounded linear operators with norms a and b respectively, show that $A + B$ and AB are also bounded.

2. *Matrix and operator notations.*

   Let $\{e_1, e_2\}$ be a basis for $\mathbb{C}^2$ (e_1 and e_2 need not be normalised or orthogonal). Let A and B be operators on $\mathbb{C}^2$ defined by

$$\begin{aligned} Ae_1 &= e_2; & Ae_2 &= 2e_1 + e_2 \\ Be_1 &= ie_2; & Be_2 &= -ie_1. \end{aligned} \tag{1.70}$$

   (a) Calculate the matrices $(A)_{ij}$ and $(B)_{ij}$, of A and B with respect to the basis $\{e_1, e_2\}$.

   (b) Calculate ABe_1 and ABe_2 using (1.70) directly and hence calculate the matrix of AB with respect to the basis $\{e_1, e_2\}$.

   (c) Calculate the matrix of AB by doing matrix multiplication of $(A)_{ij}$ and $(B)_{ij}$ and confirm that your answer agrees with the result of part (b).

   (d) Assume now that e_1 and e_2 are orthonormal. Let $\phi = 2e_1 + 3ie_2$.

   (i) Use operator notation to calculate $A\phi$ and then the inner product $\Big(\phi, A\phi\Big)$.

   (ii) Calculate the inner product $\Big(\phi, A\phi\Big)$ using matrix notation; confirm that your answer agrees with that in (i).

3. *Properties of the adjoint.*

   Let A and B be bounded linear operators and α be a complex scalar. Check, using the definition of the adjoint of a bounded linear operator, that

   (a) $(A^\dagger)^\dagger = A$.

   (b) $(A + B)^\dagger = A^\dagger + B^\dagger$.

   (c) $(\alpha A)^\dagger = \alpha^* A^\dagger$.

   (d) $(AB)^\dagger = B^\dagger A^\dagger$.

   (e) Assuming that A has a bounded inverse A^{-1}, show that $(A^\dagger)^{-1} = (A^{-1})^\dagger$.

4.

(a) Use the definition of the adjoint of an operator to show that the adjoint of the operator $|\psi\rangle\langle\phi|$ is $|\phi\rangle\langle\psi|$, where $|\psi\rangle$ and $|\phi\rangle$ are vectors (not necessarily normalised).

(b) Let P_k be the projection operator $\sum_{i=1}^{k}|e_i\rangle\langle e_i|$, where $\{|e_i\rangle\}$ form an orthonormal basis for $\mathbb{C}^n$, and $k \leq n$.

Show that P_k is Hermitian and that $P_k^2 = P_k$.

(c) Using the result above to show that the eigenvalues of P_k are 1 and 0.

5. Let H be the operator on $\mathbb{C}^2$ defined by

$$H|1\rangle = -\frac{1}{2}|1\rangle + \frac{\sqrt{3}}{2}|2\rangle\,; \quad H|2\rangle = \frac{\sqrt{3}}{2}|1\rangle + \frac{1}{2}|2\rangle \tag{1.71}$$

where $|1\rangle$ and $|2\rangle$ give an orthonormal basis for $\mathbb{C}^2$.

(a) Find the matrix of H with respect to the basis $\{|1\rangle, |2\rangle\}$ and show that H is self-adjoint.

(b) Find the expression for H in Dirac notation in this basis (i.e. write H in the form $\sum_{ij=1}^{2}|i\rangle\langle j|$) and confirm that your expression gives the correct vectors when acting on $|1\rangle$ and $|2\rangle$.

(c) Show that $H^2 = I$, where I is the identity operator on $\mathbb{C}^2$.

(d) Calculate the eigenvalues of H and find orthonormal eigenvectors $|e_1\rangle$ and $|e_2\rangle$ for H.

(e) Write H in Dirac notation with respect to the basis of eigenvectors (i.e. write H in the form $\sum_{ij=1}^{2}|e_i\rangle\langle e_j|$) and, by expanding this expression out, confirm that it is equal to the expression you found in (b).

6. Let H be the operator defined in (1.71).

(a) We define e^{-itH} by

$$e^{-itH} = \sum_{n=0}^{\infty}\frac{(-it)^n H^n}{n!}.$$

Using $H^2 = I$, or otherwise, write e^{-itH} in the form

$$e^{-itH} = a(t)I + ib(t)H$$

where a and b are real functions of t which you should find.

(b) Show that e^{-itH} is unitary.

(c) The state of a quantum system $|\psi(t)\rangle$ at time t is defined by

$$|\psi(t)\rangle = e^{-itH}\,|\psi_0\rangle\,.$$

Consider the case $|\psi_0\rangle = |1\rangle$. Write $|\psi_0\rangle$ in terms of eigenstates of H and hence, or otherwise, calculate $|\psi(t)\rangle$. Show that $|\psi(t)\rangle$ satisfies Schrödinger's equation (in units in which $\hbar = 1$):

$$i\frac{\partial\,|\psi(t)\rangle}{\partial t} = H\,|\psi(t)\rangle\,.$$

7. Consider a Hamiltonian defined on $\mathbb{C}^2$ where in some orthonormal basis $\{|\,v_1\rangle, |\,v_2\rangle\}$ it has the following matrix representation:

$$H = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}. \tag{1.72}$$

(a) Show that H is self-adjoint.

(b) Express H in Dirac notation in the basis $\{|\,v_1\rangle\ |\,v_2\rangle\}$.

(c) Compute the expectation value of H in the state $|\,v_1\rangle$.

(d) Compute the eigenvalues and the eigenstates of H.

(e) Denoting the eigenstates of H by $\{|\,E_1\rangle, |\,E_2\rangle\}$, express H in this basis using Dirac notation.

(f) Consider the operator:

$$e^{-\frac{iHt}{\hbar}}. \tag{1.73}$$

Express this operator in the basis $\{|\,E_1\rangle, |\,E_2\rangle\}$ using Dirac notation.

8. Let $|\,\phi\rangle, |\,\psi\rangle$ denote kets and let A denote a linear operator acting on the kets.

(a) Show that

$$(A\,|\,\psi\rangle)^\dagger = \langle\psi\,|\,A^\dagger.$$

(b) For the following four expressions state if it is a scalar, ket, bra, or operator and compute its adjoint:

(i) $\langle\psi\,|\,A\,|\,\phi\rangle\langle\psi\,|\,\phi\rangle$.

(ii) $\langle\psi\,|\,\phi\rangle\langle\psi\,|\,A$.

(iii) $\langle\psi\,|\,\phi\rangle A\,|\,\phi\rangle\langle\psi\,|$.

(iv) $A \mid \psi\rangle\langle\phi \mid A \mid \psi\rangle$.

Now suppose that A is self-adjoint.

(c) Show that the eigenvalues of A are real.

(d) Show that the eigenvectors of A corresponding to distinct eigenvalues are orthogonal.

9. Consider the state space $\mathbb{C}^2$ with orthonormal basis,

$$\mid 1\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad \mid 2\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

We define the operators A and B on $\mathbb{C}^2$ with respect to this basis as follows:

$$\begin{aligned} A \mid 1\rangle &= 2 \mid 1\rangle - i \mid 2\rangle, \\ A \mid 2\rangle &= i \mid 1\rangle + 2 \mid 2\rangle \end{aligned}$$

$$\begin{aligned} B \mid 1\rangle &= \mid 1\rangle - i \mid 2\rangle, \\ B \mid 2\rangle &= i \mid 1\rangle + \mid 2\rangle \end{aligned}$$

(a) Write down the matrix representations of A and B with respect to the basis $\mid 1\rangle, \mid 2\rangle$.

(b) Show that A and B are self-adjoint.

(c) Show that A and B commute.

(d) Show that the eigenvalues of A are 1 and 3. Let $\mid e_1\rangle$ denote the normalized eigenstate corresponding to the eigenvalue 1 and let $\mid e_2\rangle$ denote the normalized eigenstate corresponding to the eigenvalue 3. Compute expressions for $\mid e_1\rangle$ and $\mid e_2\rangle$.

(e) Express the basis vectors $\mid 1\rangle$ and $\mid 2\rangle$ in terms of $\mid e_1\rangle$ and $\mid e_2\rangle$.

(f) Express the normalized eigenstates of B in terms of the normalized eigenstates of A.

(g) Express B in Dirac notation in terms of $\mid e_1\rangle$ and $\mid e_2\rangle$.

(h) Compute the expectation value of B in the state $\mid e_2\rangle$.

2
Dynamics of a Quantum Particle

2.1 *The Schrödinger Equation*

The non-relativistic, time dependent Schrödinger equation for a single particle of (constant) mass m in d dimensions (where, practically speaking d will be 1, 2, or 3) has the form:

$$-\frac{\hbar^2}{2m}\nabla^2\psi(\mathbf{r},t)+V(\mathbf{r},t)\psi(\mathbf{r},t)=i\hbar\frac{\partial}{\partial t}\psi(\mathbf{r},t), \tag{2.1}$$

where $\mathbf{r}\in\mathbb{R}^d$ (or some subset of $\mathbb{R}^d$), $\hbar=\frac{h}{2\pi}$, where h is Planck's constant, and $V(\mathbf{r},t)$ is the potential energy. Schrödinger's original paper on the topic[1] should be mostly understandable to you and provides interesting background leading to the development of the equation (but, in this regard the book of Stone[2] is absolutely superb). Many books offer a description of a type of derivation of the Schrödinger equation. For our purposes, we will just start with the equation, study its structure, solve it in certain situations, and learn for to interpret the results. But concerning the derivation of the Schrödinger equation, the following quote of Richard Feynman[3] is particularly insightful and captures some of the mystery surrounding the development of quantum mechanics in Feynman's inimitable manner.

> Where did we get that (equation) from? Nowhere. It is not possible to derive it from anything you know. It came out of the mind of Schrödinger.

The Schrödinger equation is a linear partial differential equation. Boundary conditions and initial conditions are essential for specifying the particular setting for which a solution is sought. These will be dealt with in the particular problems that we study. The solution of (2.1) is a complex valued function, $\psi(\mathbf{r},t)$, which is referred to as the *wavefunction*. We will discuss the meaning of the wavefunction for a particle of mass m moving under the influence of a potential energy

[1] E. Schrödinger. An undulatory theory of the mechanics of atoms and molecules. *Physical review*, 28(6):1049, 1926

[2] A D. Stone. *Einstein and the quantum: The quest of the valiant Swabian*. Princeton University Press, 2015

[3] A. J. G. Hey, T. Hey, and P. Walters. *The new quantum universe*. Cambridge University Press, 2003

function $V(\mathbf{r},t)$ shortly. But first we discuss some general features of solving this linear partial differential equation.

The difficulty in solving (2.1) is directly related to the form of the potential energy, $V(\mathbf{r},t)$. For example, if $V(\mathbf{r},t)=0$ (a ''free particle") the solution of (2.1) is straightforward. We will treat this situation shortly. However, first we consider the general case where the potential energy is independent of time, i.e.,

$$V(\mathbf{r},t)=V(\mathbf{r}). \tag{2.2}$$

In this situation the Schrödinger equation can be solved using the method of separation of variables. To apply this method we assume a solution of the form:

$$\psi(\mathbf{r},t)=\phi(\mathbf{r})f(t) \tag{2.3}$$

Substituting this into (2.1) gives:

$$f(t)\left(-\frac{\hbar^2}{2m}\nabla^2+V(\mathbf{r})\right)\phi(\mathbf{r})=\phi(\mathbf{r})\,i\hbar\frac{d}{dt}f(t) \tag{2.4}$$

We divide both sides by $\phi(\mathbf{r})f(t)$ and obtain:

$$\frac{1}{\phi(\mathbf{r})}\left(-\frac{\hbar^2}{2m}\nabla^2\phi(\mathbf{r})+V(\mathbf{r})\phi(\mathbf{r})\right)=\frac{i\hbar}{f(t)}\frac{df(t)}{dt} \tag{2.5}$$

The left hand side of (2.5) is a function of $\mathbf{r}$ and the right hand side is a function of t. Since $\mathbf{r}$ and t are independent variables the two sides of (2.5) must be equal to the same constant, which we call E, and we write the resulting two ordinary differential equations separately below:

$$-\frac{\hbar^2}{2m}\nabla^2\phi(\mathbf{r})+V(\mathbf{r})\phi(\mathbf{r})=E\phi(\mathbf{r}). \tag{2.6}$$

$$\frac{d}{dt}f(t)=\frac{E}{i\hbar}f(t)=-\frac{i}{\hbar}Ef(t)\Rightarrow f(t)=f(0)e^{-\frac{i}{\hbar}Et}. \tag{2.7}$$

The equation (2.6) is referred to as the *time independent Schrödinger equation*.

Hence a solution of (2.1) is given by:

$$\psi(\mathbf{r},t)=f(t)\phi(\mathbf{r})=\phi(\mathbf{r})e^{-\frac{i}{\hbar}Et}. \tag{2.8}$$

States having the form of (2.7) are referred to as *stationary states*. This terminology may appear a bit curious, at first, because the word stationary generally means not changing (in time) and (2.7) clearly has a time dependence as a result of the term $e^{-\frac{i}{\hbar}Et}$. The origin of the term *stationary state* comes from the interpretation of the magnitude squared of the wavefunction (2.20) as a probability density described in Section

2.2. Clearly, the magnitude squared of wavefunctions of the form (2.7) is independent of time. It is important to recall why this is the case. It is a direct consequence of the form of the solution of Schrödinger's equation given in (2.7). The form arose from the method that we used for solving Schrödinger's equation, *separation of variables*. This method was possible because the potential energy function was independent of time. Situations where the potential energy are time dependent are important and fascinating, but are beyond the scope of this course. Some background and examples on this topic can be found in the book by David Tannor[4].

[4] D. J Tannor. *Introduction to quantum mechanics: a time-dependent perspective*. University Science Books, 2007

While (2.8) is a solution of (2.1), it is not the most general solution. It is merely a *particular solution*. We explain this statement in a bit more detail. There are two points to consider:

- We have not imposed any boundary and/or initial conditions on the solution.

- The solution (2.8) is given for a particular separation constant E.

The first point is dealt with on a problem-by-problem basis. The second point is more fundamental and underlies all of our approaches to ''quantum problems".

The mathematical framework for the solution of the time independent Schrödinger equation is an example of a (linear) eigenvalue problem. In particular, we have shown earlier that the mapping:

$$\phi(\mathbf{r}) \mapsto -\frac{\hbar^2}{2m}\nabla^2\phi(\mathbf{r}) + V(\mathbf{r})\phi(\mathbf{r}), \tag{2.9}$$

is a linear operator on the complex inner product space (i.e. complex Hilbert space) $L^2(D)$, where $D \subset \mathbb{R}^d$ is the spatial domain. Hence, (2.9) has the form of an eigenvalue problem where $\phi(\mathbf{r})$ is the eigenvector with corresponding eigenvalue E. We denote the eigenvector-eigenvalue pairs by $(\phi_k(\mathbf{r}), E_k)$. Since (2.9) is a self-adjoint operator it has a complete set of eigenvectors in $L^2(D)$. Therefore every function $\psi(\mathbf{r}, t)$ in $L^2(D)$ can be represented as follows:

$$\psi(\mathbf{r}, t) = \sum_k c_k(t)\phi_k(\mathbf{r}). \tag{2.10}$$

Substituting this into (2.1) gives:

$$-\frac{\hbar^2}{2m}\sum_k c_k(t)\nabla^2\phi_k(\mathbf{r}) + \sum_k c_k(t)V(\mathbf{r})\phi_k(\mathbf{r}) = i\hbar\sum_k \phi_k(\mathbf{r})\frac{dc_k(t)}{dt}, \tag{2.11}$$

or

$$\sum_k c_k(t) \underbrace{\left(-\frac{\hbar^2}{2m}\nabla^2\phi_k(\mathbf{r}) + V(\mathbf{r})\phi_k(\mathbf{r}) \right)}_{E_k\phi_k(\mathbf{r})} = i\hbar \sum_k \phi_k(\mathbf{r}) \frac{dc_k(t)}{dt}. \tag{2.12}$$

Equating the coefficients on the basis vectors on the left and right hand sides gives:

$$i\hbar \frac{dc_k}{dt} = E_k c_k \Rightarrow c_k(t) = c_k(0) e^{-\frac{iE_k}{\hbar}t}. \tag{2.13}$$

Substituting this into (2.10) gives:

$$\psi(\mathbf{r}, t) = \sum_k c_k(0) e^{-\frac{iE_k}{\hbar}t} \phi_k(\mathbf{r}). \tag{2.14}$$

Physical Requirements for the Mathematical Structure of the Wavefunction

For the wavefunction to be physically meaningful we will require it to be a continuous and single valued function. We will make the need for this clear when we describe the physical interpretation of the wavefunction in Section 2.2. Furthermore, the wavefunction has additional properties that come from the structure of the partial differential equation that it satisfies. In particular, if there is a boundary across which the potential energy changes values then the wave function must be continuous across that boundary and remain single valued. Moreover, if the potential energy function is piecewise continuous it can be shown that all partial derivatives of the wavefunction with respect to the spatial variables must be continuous across the boundary. It is possible to find solutions of the Schrödinger equation that do not satisfy these properties, but they are not considered to be physically meaningful. We will see an example of such a wavefunction shortly.

For the remainder of this section we restrict ourselves to the one dimensional Schrödinger equation, i.e. $\mathbf{r} \in \mathbb{R}$*, and we will refer to the spatial variable as* x*. However, the general concepts and arguments apply in three dimensions.* The mathematical structure of the time-independent Schrödinger and the structure of the potential energy function implies that the solutions satisfy a *parity property* that we now describe in more detail.

We begin by defining the parity operator acting on functions.

Definition 21 (Parity Operator)**.** *Given a function* $f(x)$*, the parity operator is defined by:*

$$Pf(x) = f(-x) \tag{2.15}$$

The following are some examples of the action of the parity operator on a function.

$$\begin{aligned} P\sin(x) &= \sin(-x) = -\sin(x), \\ P\cos(x) &= \cos(-x) = \cos(x), \\ P(\cos(x)+\sin(x)) &= \cos(x) - \sin(x) \end{aligned} \tag{2.16}$$

If $Pf(x) = f(x)$, i.e. $f(x)$ is an eigenfunction of P with eigenvalue 1, then $f(x)$ is said to be an *even function*. If $Pf(x) = -f(x)$, i.e. $f(x)$ is an eigenfunction of P with eigenvalue -1, then $f(x)$ is said to be an *odd function*.

Then we have the following theorem.

Theorem 2. *Let $\psi(x)$ be a solution of the time independent Schrödinger equation*

$$-\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2} + V(x)\psi(x) = E\psi(x), \tag{2.17}$$

with $V(x) = V(-x)$ and boundary conditions that are symmetric in x. Then either $\psi(x) = \psi(-x)$ or $\psi(x) = -\psi(-x)$ are solutions.

Proof. Suppose $\psi(x)$ is a solution of the time independent Schrödinger equation:

$$-\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2}(x) + V(x)\psi(x) = E\psi(x) \tag{2.18}$$

Suppose we let

$$x \to -x$$

and use $V(x) = V(-x)$. Under this transformation the time independent Schrödinger equation becomes:

$$-\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2}(-x) + V(x)\psi(-x) = E\psi(-x). \tag{2.19}$$

Hence, $\psi(x)$ and $\psi(-x)$ satisfy the same equation. Therefore they must be the same solution, except for an overall multiplicative constant, i.e.,

$$\psi(x) = a\psi(-x)$$

for some constant a. Requiring the wavefunctions to be normalized implies $a = \pm 1$, from which the result follows.

□

2.2 *The Interpretation of the Wave Function*

In quantum mechanics the wavefunction provides the description for the motion of a particle of (constant) mass m in d dimensions, $\psi(\mathbf{r}, t)$, $\mathbf{r} \in \mathbb{R}^d$. In contrast to classical mechanics, rather than providing position and velocity as a function of time, quantum mechanics is an inherently statistical theory, and

$$\rho(\mathbf{r}, t) = |\psi(\mathbf{r}, t)|^2. \tag{2.20}$$

represents the probability density for the position of a particle with wavefunction $\psi(\mathbf{r}, t)$ at time t. More explicitly, the probability to find the particle in a volume $V \subset \mathbb{R}^d$ at time t is:

$$\int_V |\psi(\mathbf{r}, t)|^2 d^3\mathbf{r}. \tag{2.21}$$

Note that the probabilistic interpretation requires the wavefunction to be normalized, i.e.,

$$\int_{\mathbb{R}^d} |\psi(\mathbf{r}, t)|^2 d^d\mathbf{r} = 1. \tag{2.22}$$

for bound systems.

Also, note that because of the absolute value in the integrand of (2.21) multiplying the wavefunction by a factor $e^{i\alpha}$, $\alpha \in \mathbb{R}$, does not change the value of the probability density. Often the phrase, ''an overall constant phase is unobservable" is used to describe this feature.

The probabilistic interpretation of the wavefunction gives rise to other statistical quantitites that can be used to describe the motion. These are collected below.

The average position of a particle described by a wavefunction $\psi(\mathbf{r}, t)$ at time t is

$$\langle \mathbf{r} \rangle \equiv \mathbf{r}_{\text{av}} = \int_{\mathbb{R}^d} \mathbf{r} |\psi(\mathbf{r}, t)|^2 d^d\mathbf{r}, \tag{2.23}$$

and the variance of the position around this average is:

$$\begin{aligned} \langle (\mathbf{r} - \langle \mathbf{r} \rangle)^2 \rangle &= \langle |\mathbf{r}|^2 \rangle - |\langle \mathbf{r} \rangle|^2 \\ &= \int_{\mathbb{R}^d} |\mathbf{r} - \langle \mathbf{r} \rangle|^2 |\psi(\mathbf{r}, t)|^2 d^d\mathbf{r}. \end{aligned} \tag{2.24}$$

2.3 *The Free Particle*

A ''free particle" is a particle that moves in a region where it is not subjected to an external force. This means that the potential energy is constant, which we will take to be zero in this discussion. We will treat both the classical and quantum mechanical free particle.

Classical Mechanical Free Particle

Newton's equations of motion for a free particle of (constant) mass m are:

$$m\ddot{\mathbf{r}} = 0, \tag{2.25}$$

or, since m is constant:

$$\frac{d}{dt}(m\dot{\mathbf{r}}) = 0. \tag{2.26}$$

Integrating once gives:

$$m\dot{\mathbf{r}} \equiv \mathbf{p} = \text{constant}, \tag{2.27}$$

where $\mathbf{p}$ is referred to as the (linear) momentum. Hence (2.27) expresses the well-known result that (linear) momentum is constant (in time) for a classical particle of constant mass that is not acted on by a force. Moreover, (2.27) can be integrated to obtain the position of the free particle as a function ot time:

$$\mathbf{r}(t) = \mathbf{r}(0) + \frac{\mathbf{p}}{m}t. \tag{2.28}$$

The total energy of the particle is purely kinetic energy and is given by:

$$E = \frac{1}{2}m\dot{\mathbf{r}}\cdot\dot{\mathbf{r}} \equiv \frac{1}{2}mv^2 = \frac{\mathbf{p}\cdot\mathbf{p}}{2m} \equiv \frac{p^2}{2m} \tag{2.29}$$

where v denotes the magnitude of $\dot{\mathbf{r}}$ and p denotes the magnitude of $\mathbf{p}$.

Hence, we see from (2.27), (2.28). and (2.29) that the momentum, position, and total energy of a classical free particle can be determined simultaneously.

Quantum Mechanical Free Particle

Now we treat the quantum mechanical free particle of (constant) mass m. In this case we will restrict ourselves to the one dimensional case. The three dimensional case is straightforward, but the details tend to distract from the main ideas in the first time that one sees it.

The time independent Schrödinger equation (2.6) in this case is given by:

$$-\frac{\hbar^2}{2m}\frac{d^2}{dx^2}\phi(x) = E\phi(x). \tag{2.30}$$

Initially, there are three cases to consider: $E > 0$, $E < 0$ and the trivial case $E = 0$. It can be shown that the solution for $E < 0$ gives rise to a wavefunction that is non-differentiable at the origin, i.e. it has a

cusp. Therefore it is not a physically interesting situation, and we will therefore not consider this case. We will focus on the case $E > 0$, for which we have:

$$\frac{d^2\phi}{dx^2} + k^2\phi = 0, \tag{2.31}$$

where

$$\frac{2mE}{\hbar^2} \equiv k^2 \Rightarrow E = \frac{\hbar^2 k^2}{2m} > 0. \tag{2.32}$$

The solution can be written as:

$$\phi(x) = Ae^{ikx} + Be^{-ikx}, \tag{2.33}$$

where A and B are (complex) constants. The functions.

Insight into the nature of (2.33) can be obtained by noting that the functions e^{ikx} and e^{-ikx} are eigenfunction of the momentum operator. This can be seen by a direct calculation. Recall (see Definition 18) that the action of the momentum operator, $P \equiv \frac{\hbar}{i}\frac{d}{dx}$ on a function $\phi(x)$ is of the form:

$$P\phi(x) \equiv \frac{\hbar}{i}\frac{d\phi(x)}{dx}. \tag{2.34}$$

Then we have:

$$\frac{\hbar}{i}\frac{d}{dx}e^{\pm ikx} = \frac{\hbar}{i}\left(\pm ik\right)e^{\pm ikx} = \pm\hbar k e^{\pm ikx}. \tag{2.35}$$

Hence e^{ikx} is an eigenfunction with eigenvalue $\hbar k$. This corresponds to a particle moving to the right with positive momentum $\hbar k$. Similarly, e^{-ikx} is an eigenfunction with eigenvalue $-\hbar k$, corresponding to a particle moving to the left with negative momentum $-\hbar k$.

The quantum mechanical free particle moving to the right with momentum $\hbar k$ is described by the wavefunction $\psi(x) = Ae^{ikx}$. The momentum and the energy, (2.32), are known precisely, and the average position can be determined from the probability density:

$$|\psi(x)|^2 = |A|^2. \tag{2.36}$$

This expression raises a number of issues. The probability density is constant. Hence the probability of finding a particle in any interval on the x axis is the same as that for any other interval of equal length, and does not change with time. In other words, there are no special places for a free particle to be found. However, the concept of ''probabilty" is questionable in this case since the wavefunction is not normalizable, i.e. the integral of the probability over the entire line is infinite. So while the momentum and energy can be known precisely, the position is unknown.

Despite the fact that functions of the form $e^{\pm ikx}$ are not normalizable on their entire domain they still play an important role in quantum mechanics. We will see this shortly when we introduce the concept of a wave packet. However, there has been significant effort in providing mathematical meaning for the normalization of such ''non-normalizable" functions. This is described in detail in .

M. Amaku, F. A. B. Coutinho, and F. M. Toyama. The normalization of wave functions of the continuous spectrum. *Revista Brasileira de Ensino de Física*, 42, 2020

So how can we describe a quantum mechanical free particle in a way that bears some resemblance to our description of a classical free particle? This can be done using the notion of a *wave packet*. .

Wave Packets

A wave packet is constructed by considering a sum of functions of the form $a(k)e^{ikx}$, where the sum is over k (hence, particle momenta). Because k is unrestricted, the sum actually is an integral and we write

$$\psi(x,0) = \int_{-\infty}^{\infty} a(k)e^{ikx}dk. \tag{2.37}$$

The amplitudes $a(k)$ of each e^{ikx} determine the so-called spectral content of the wave packet. We will assume $a(k)$ is a Gaussian function of the following form:

$$a(k) = \frac{C\alpha}{\sqrt{\pi}}e^{-\alpha^2k^2}, \tag{2.38}$$

where C and α are real constants that can be used to control the shape of the wave packet. The constant C is referred to as the amplitude. Note that at $x = \pm 2\alpha$ the amplitude is reduced by a factor $\frac{1}{e}$. This leads us to identify α with the width of the Gaussian.

With the choice of $a(k)$ to be a Gaussian function, the phrase *Gaussian wavepacket* is often used. Substituting into gives:

$$\psi(x,0) = \int_{-\infty}^{\infty} a(k)e^{ikx}dk = \frac{C\alpha}{\sqrt{\pi}}\int_{-\infty}^{\infty} e^{\left(ikx-\alpha^2k^2\right)}dk. \tag{2.39}$$

To evaluate the integral, we first complete the square in the exponent as follows:

$$ikx - \alpha^2k^2 = -\left(\alpha k - \frac{ix}{2\alpha}\right)^2 - \frac{x^2}{4\alpha^2}. \tag{2.40}$$

The second term on the right is constant for the integration over k. To integrate the first term we change variables with the substitution $z = \alpha k - \frac{ix}{2\alpha}$, we obtain

$$\psi(x,0) = \frac{C}{\sqrt{\pi}}e^{-\frac{x^2}{4\alpha^2}}\int_{-\infty}^{\infty} e^{-z^2}dz. \tag{2.41}$$

The integral is well-known[5] and is equal to $\sqrt{\pi}$, which yields

$$\psi(x,0) = Ce^{-\frac{x^2}{4\alpha^2}} = Ce^{-\left(\frac{x}{2\alpha}\right)^2}, \tag{2.42}$$

and the constant C can be chosen so that the wavefunction is normalized. In other words, we compute:

$$\int_{-\infty}^{\infty}(\psi(x,0))^2dx = C^2\int_{-\infty}^{\infty} e^{-2\frac{x^2}{4\alpha^2}}dx = 1 \tag{2.43}$$

[5] $\int_{-\infty}^{\infty} e^{-x^2}dx = \sqrt{\pi}$ is the famous Gaussian integral. It is famous because it turns up over and over again in (seemingly) diverse areas of mathematics and physics.

Using the expression for the Gaussian integral, we find that

$$C = \frac{1}{\sqrt{\alpha}\,(2\pi)^{\frac{1}{4}}},$$

and, hence,

$$\psi(x,0) = \frac{1}{\sqrt{\alpha}\,(2\pi)^{\frac{1}{4}}} e^{-\left(\frac{x}{2\alpha}\right)^2}. \tag{2.44}$$

Wave packets are the basic building blocks of a number of useful tools for understanding the time evolution of quantum systems. More background, and many examples, can be found in the book of Tannor[6]. We will learn more about the role played by the wave packet description of particles in determining position and momentum when we consider the issue of uncertainty in Section 3.2.

[6] D. J Tannor. *Introduction to quantum mechanics: a time-dependent perspective*. University Science Books, 2007

2.4 *The Square Well*

We consider a particle moving in the interval $[0, a]$, where $V(x) = 0$ in this interval, and $V(x) = \infty$ for $x > a$ and $x < 0$. The time independent Schrödinger equation inside the well is:

$$-\frac{\hbar^2}{2m}\frac{d^2}{dx^2}\psi = E\psi, \tag{2.45}$$

where we will take as boundary conditions $\psi(a) = \psi(0) = 0$. The general solution of (2.45) is:

$$\psi(x) = \begin{cases} A\cosh\left(\sqrt{2m|E|}\,\frac{x}{\hbar}\right) + B\sinh\left(\sqrt{2m|E|}\,\frac{x}{\hbar}\right) & \text{if } E < 0 \\ A + Bx & \text{if } E = 0 \\ A\cos\left(\sqrt{2mE}\,\frac{x}{\hbar}\right) + B\sin\left(\sqrt{2mE}\,\frac{x}{\hbar}\right) & \text{if } E > 0 \end{cases} \tag{2.46}$$

when

$$E \leq 0 \quad \psi(0) = \psi(a) = 0 \quad \Rightarrow A = B = 0, \tag{2.47}$$

and when

$$E > 0 \quad \begin{array}{lll} \psi(0) = 0 & \Rightarrow & A = 0, \\ \psi(a) = 0 & \Rightarrow & \sqrt{2mE}\,\frac{a}{\hbar} = n\pi. \end{array} \tag{2.48}$$

This implies that we have nontrivial solutions if and only if[7]:

$$E = E_n = \frac{n^2\pi^2\hbar^2}{2ma^2}, \quad n = 1, 2, \ldots, \tag{2.49}$$

and therefore

$$\psi_n(x) = B_n \sin\left(\frac{n\pi x}{a}\right). \tag{2.50}$$

[7] The ''energy eigenvalue" expression for the square well, (2.49), highlights the quantum aspect of the square well. The energy only exists in discrete amounts. Note that the energy cannot be zero. The integer n is often referred to as the *principal quantum number*.

The constant B_n is determined by requiring

$$\int_0^a |\psi_n(x)|^2 dx = 1. \tag{2.51}$$

which gives:

$$\frac{1}{2}|B_n|^2 a = 1, \tag{2.52}$$

or

$$B_n = \sqrt{\frac{2}{a}}, \tag{2.53}$$

and therefore

$$\psi_n(x) = \sqrt{\frac{2}{a}} \sin\left(\frac{n\pi x}{a}\right). \tag{2.54}$$

Using (2.8), the time evolution of $\psi_n(x)$ is given by:

$$\psi_n(x,t) = \sqrt{\frac{2}{a}} \sin\left(\frac{n\pi x}{a}\right) e^{-\frac{in^2\pi^2\hbar t}{2ma^2}}. \tag{2.55}$$

Linearity of the Schrödinger equation implies that the general wavefunction can be obtained by a superposition of the eigenfunctions:

$$\psi(x,t) = \sqrt{\frac{2}{a}} \sum_{n=1}^{\infty} c_n e^{-\frac{in^2\pi^2\hbar t}{2ma^2}} \sin\left(\frac{n\pi x}{a}\right). \tag{2.56}$$

The expansion coefficients c_n can be obtained as follows. Noting that

$$\psi(x,0) = \sqrt{\frac{2}{a}} \sum_{n=1}^{\infty} c_n \sin\left(\frac{n\pi x}{a}\right), \tag{2.57}$$

therefore

$$c_n = \sqrt{\frac{2}{a}} \int_0^a \psi(x,0) \sin\left(\frac{n\pi x}{a}\right) dx. \tag{2.58}$$

Summarizing: for the square well the time-independent Schrödinger equation has eigenvalues $E_n = \frac{n^2\pi^2\hbar^2}{2ma^2}$, $n = 1, 2, \ldots$ with corresponding eigenfunctions $\psi_n(x) = \sqrt{\frac{2}{a}} \sin\left(\frac{n\pi x}{a}\right)$. The general solution of the time-dependent Schrödinger equation can be written as $\psi(x,t) = \sum_n c_n e^{\frac{-iE_n t}{\hbar}} \psi_n(x)$ with $c_n = \int_0^a \psi(x,0)\psi_n(x)dx$.

Interpretation of the Wavefunction for the Square Well

Now we apply the general probabilistic interpretation of the wavefunction described in Section 2.2 to the quantum particle in the square well.

Using (2.20) and (2.55), the probability density for a particle in the n^{th} eigenstate of a square well is:

$$|\psi_n(x)|^2 = \frac{2}{a}\sin^2\frac{n\pi x}{a} = \frac{1}{a}\left(1-\cos\frac{2n\pi x}{a}\right). \tag{2.59}$$

1. We can integrate (2.59) to obtain the probability distribution function, which gives the probability of finding a particle between $x = 0$ and $x = X \leq a$ is:

$$\int_0^X |\psi_n(x)|^2 dx = \frac{X}{a} - \frac{1}{2n\pi}\sin\frac{2n\pi X}{a} \equiv F_n(X). \tag{2.60}$$

2. The probability of finding the particle in the interval $[\frac{a}{4}, \frac{3a}{4}]$ is

$$\begin{aligned}\int_{\frac{a}{4}}^{\frac{3a}{4}} |\psi_n(x)|^2 dx &= F_n\left(\frac{3a}{4}\right) - F_n\left(\frac{a}{4}\right)\\ &= \frac{1}{2} - \frac{1}{2n\pi}\left(\sin\frac{3n\pi}{2} - \sin\frac{n\pi}{2}\right)\\ &= \begin{cases}\frac{1}{2} & \text{n even}\\ \frac{1}{2} + \frac{1}{n\pi}(-1)^{\frac{n-1}{2}} & \text{n odd}\end{cases}\end{aligned} \tag{2.61}$$

3. The mean position of the particle is:

$$\int_0^a x|\psi_n(x)|^2 dx = \frac{1}{a}\int_0^a \left(x - x\cos\frac{2n\pi x}{a}\right)dx = \frac{a}{2}. \tag{2.62}$$

4. The variance of the position is:

$$\int_0^a \left(x - \frac{1}{2}a\right)^2 |\psi_n(x)|^2 dx = \left(\frac{1}{12} - \frac{1}{12n^2\pi^2}\right)a^2. \tag{2.63}$$

Now note that the general time-dependent solution of the Schrödinger equation for the square well is given by:

$$\psi(x,t) = \sum_n c_n\psi_n(x)\exp\left(-i\frac{E_n t}{\hbar}\right) = \sqrt{\frac{2}{a}}\sum_n c_n \sin\frac{n\pi x}{a}\exp\left(-i\frac{n^2\pi^2\hbar t}{2ma^2}\right). \tag{2.64}$$

Then, using orthogonality for the $\psi_n(x)$, the probability density for the square well is given by:

$$\int_0^a |\psi(x,t)|^2 dx = \int_0^a \sum_{n,m} c_n\bar{c}_m\psi_n(x)\overline{\psi_m(x)}\exp\left(-i\frac{(E_n - E_m)t}{\hbar}\right)dx = \sum_m |c_m|^2. \tag{2.65}$$

If we require

$$\int_0^a |\psi(x,t)|^2 dx = 1, \tag{2.66}$$

then this implies that

$$\sum_m |c_m|^2 = 1. \tag{2.67}$$

This leads to the following interpretation. The probability of measuring the energy to be E_n or, equivalently, of finding the particle to be in the n^{th} eigenstate, $\psi_n(x)$, is $|c_n|^2$[8].

[8] This is a very important point. It says that the probability of measuring the value E_n is $|c_n|^2$. This interpretation is due to Heisenberg and Born.

2.5 *Probability Current and the Conservation of Probability*

We have seen that the probability density, (2.20), plays a central role in how we relate the wavefunction to observable properties of a quantum particle. In this section we explore properties of the time evolution of the probability density.

The first question we consider is the following. If

$$\int_{\mathbb{R}^d} |\psi(\mathbf{r},0)|^2 d^d\mathbf{r} = 1. \tag{2.68}$$

then under what conditions do we have

$$\int_{\mathbb{R}^d} |\psi(\mathbf{r},t)|^2 d^d\mathbf{r} = 1, \qquad \forall\, t > 0\,? \tag{2.69}$$

If the wavefunction from which the probability density is constructed is an eigenstate then the answer is clear

$$|\psi_n(\mathbf{r},t)|^2 = |\psi_n(\mathbf{r})e^{-\frac{iE_nt}{\hbar}}|^2 = |\psi_n(\mathbf{r})|^2, \tag{2.70}$$

i.e., the probabilty density is independent of time (but it depends on the spatial variable).

Now we will derive a partial differential equation that describes the evolution of the probability density. We begin by computing the time derivative of ρ:

$$\frac{\partial}{\partial t}\rho = \frac{\partial}{\partial t}(\psi\bar{\psi}) = \frac{\partial\psi}{\partial t}\bar{\psi} + \psi\frac{\partial\bar{\psi}}{\partial t}. \tag{2.71}$$

Recall from (2.1) that the Schrödinger equation and its complex conjugate are given by:

$$\frac{\partial\psi}{\partial t} = \frac{1}{i\hbar}\left(-\frac{\hbar^2}{2m}\nabla^2\psi + V\psi\right), \tag{2.72}$$

and

$$\frac{\partial \bar{\psi}}{\partial t} = -\frac{1}{i\hbar}\left(-\frac{\hbar^2}{2m}\nabla^2\bar{\psi} + V\bar{\psi}\right).. \tag{2.73}$$

Substituting these two equations into (2.71) gives:

$$\begin{aligned}\frac{\partial \rho}{\partial t} &= -\frac{\hbar}{2mi}\left(\bar{\psi}\nabla^2\psi - \psi\nabla^2\bar{\psi}\right), \\ &= -\frac{\hbar}{2mi}\nabla\cdot\left(\bar{\psi}\nabla\psi - \psi\nabla\bar{\psi}\right).\end{aligned} \tag{2.74}$$

This leads to the following definition:

Definition 22 (Probability Current)**.** *The probability current,* $\mathbf{j}(\mathbf{r},t)$*, is defined as*

$$\mathbf{j}(\mathbf{r},t) = \frac{\hbar}{2mi}\left(\bar{\psi}\nabla\psi - \psi\nabla\bar{\psi}\right) = \frac{1}{m}\mathrm{Re}\left(\bar{\psi}\frac{\hbar}{i}\nabla\psi\right). \tag{2.75}$$

The calculations above can be collected together into the following theorem.

Theorem 3. $\rho(\mathbf{r},t)$ *and* $\mathbf{j}(\mathbf{r},t)$ *satisfy*

$$\frac{\partial \rho}{\partial t} + \nabla\cdot\mathbf{j}(\mathbf{r},t) = 0. \tag{2.76}$$

(2.76) has the same form as the continuity equation from fluid mechanics, which expresses the conservation of mass density under a flow. In this setting the role of mass density is played by the probability density and the role of the flow is played by the probability current. Next we describe how the partial differential equation implies the conservation of the probability.

First we establish some notation. Let $B(R)$ denote the ball of radius R in $\mathbb{R}^d$ and let $S \equiv \partial B(R)$ denote the surface of this ball. Our result will require an assumption on the class of functions under consideration. We will assume the following behavior at infinity for the probability current:

$$\mathbf{j}(\mathbf{r},t)|\mathbf{r}|^{d-1} \to 0 \quad \text{as} \quad |\mathbf{r}| \to \infty. \tag{2.77}$$

Integrating (2.76) over $B(R)$ gives:

$$\frac{\partial}{\partial t}\int_{B(R)}\rho(\mathbf{r},t)d^d\mathbf{r} = -\int_{B(R)}\nabla\cdot\mathbf{j}(\mathbf{r},t)d^d\mathbf{r} = -\int_{\partial B(R)}\mathbf{j}(\mathbf{r},t)\cdot d\mathbf{S}, \tag{2.78}$$

where the last equality follows from the application of the divergence theorem. Now taking $R \to \infty$ and using (2.77) gives:

$$\frac{\partial}{\partial t}\int_{\mathbb{R}^d}\rho(\mathbf{r},t)d^d\mathbf{r} = 0, \tag{2.79}$$

Therefore

$$\int_{\mathbb{R}^d} \rho(\mathbf{r},t)d^d\mathbf{r}, \tag{2.80}$$

is independent of t. In particular, if it is 1 for $t = 0$ then it is 1 for all t.

The probability current has a suggestive interpretation in terms of classical mechanics. Rewriting the probability current here for easy reference:

$$\mathbf{j}(\mathbf{r},t) = \frac{1}{m}\mathrm{Re}\left(\bar{\psi}\left(\frac{\hbar}{i}\nabla\psi\right)\right). \tag{2.81}$$

Recall that $\frac{\hbar}{i}\nabla\psi$ is interpreted as the product of the momentum with the wavefunction. Therefore $\mathbf{j}(\mathbf{r},t)$ has the form of $\frac{\text{momentum}}{m}|\psi|^2$, i.e. the product of the velocity and the probability density.

2.6 *Some Additional Properties of the Square Well Directly from the Structure of the Schrödinger Equation*

We derived the following properties of the square well directly from the solutions of the Schrödinger equation for the square well:

1. The energy levels are non-degenerate, i.e., $E_n \neq E_m$ if $n \neq m$.
2. The energy levels are real, i.e. $E_n \in \mathbb{R}$, $\forall n$.
3. The eigenfunctions are orthonormal, i.e.,

$$\int_0^a \psi_n(x)\overline{\psi_m(x)}dx = \delta_{n,m}. \tag{2.82}$$

We now want to show that these these properties follow directly from the structure of the Schrödinger equation.

Theorem 4. *The energy levels of the one dimensional square well are non-degenerate.*

Proof. We argue by contradiction.

Assume that eigenfunctions $\psi_m(x)$ and $\psi_n(x)$ corresponding to different states ($n \neq m$) have $E_n = E_m$:

$$\begin{aligned} -\frac{d^2\psi_n}{dx^2} &= \frac{2mE_n}{\hbar^2}\psi_n, \\ -\frac{d^2\psi_m}{dx^2} &= \frac{2mE_m}{\hbar^2}\psi_m, \qquad E_n = E_m. \end{aligned} \tag{2.83}$$

Multiplying the first equation in (2.83) by $\psi_m(x)$ and the second equation in (2.83) by $\psi_n(x)$ and subtracting gives:

$$\psi_m\frac{d^2\psi_n}{dx^2} - \psi_n\frac{d^2\psi_m}{dx^2} = 0, \tag{2.84}$$

or

$$\frac{d}{dx}\left(\psi_m \frac{d\psi_n}{dx} - \psi_n \frac{d\psi_m}{dx}\right) = 0, \tag{2.85}$$

from which it follows that:

$$\psi_m \frac{d\psi_n}{dx} - \psi_n \frac{d\psi_m}{dx} = \text{constant}. \tag{2.86}$$

If we evaluate (2.86) at $x = 0$ where $\psi_m(0) = \psi_n(0) = 0$ we conclude that constant $= 0$. Therefore (2.86) becomes:

$$\psi_m \frac{d\psi_n}{dx} - \psi_n \frac{d\psi_m}{dx} = 0. \tag{2.87}$$

Dividing (2.87) by $\psi_n(x)\psi_m(x)$ gives:

$$\frac{1}{\psi_n}\frac{d\psi_n}{dx} = \frac{1}{\psi_m}\frac{d\psi_m}{dx}. \tag{2.88}$$

After integrating this equation we obtain:

$$\log|\psi_n| = \log|\psi_m| + \text{constant}, \tag{2.89}$$

or

$$\psi_n(x) = \text{constant}\,\psi_m(x). \tag{2.90}$$

from which it follows that $\psi_n(x)$ and $\psi_m(x)$ represent the same state, since they are equal *after normalization*.. This contradicts our original asssumption. □

The following theorem is concerned with non-degeneracy of the energy levels of a more general one dimensional potential energy function, $V(x)$, that has the property that the Schrödinger equation with this potential has a set of eigenfunctions and eigenvalues, $(\phi_k(x), E_k)$.

Theorem 5. *The energy levels of the one dimensional time independent Schrödinger equation:*

$$-\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2} + V(x)\psi(x) = E\psi(x), \tag{2.91}$$

with boundary condition $\psi(x) \to 0$ as $|x| \to \infty$ are all non-degenerate.

Proof. We will leave the proof of this result as an exercise.

□

Non-degeneracy of the energy levels is a general result in one dimension, but it is not necessarily true in higher dimensions.

Next we will prove a basic result that will enable us to proved several properties of the eigenstates of the square well.

First, let $\alpha(x)$ and $\beta(x)$ be any two functions defined on $[0, a]$ which vanish at $x = 0$ and $x = a$ and which are twice differentiable. Then we have the following theorem.

Theorem 6 (Hermiticity)**.** *The differential operator:*

$$H = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2}, \tag{2.92}$$

satisfies

$$\int_0^a \bar{\alpha} H\beta dx = \int_0^a \overline{(H\alpha)}\beta dx, \tag{2.93}$$

where the overline indicates the complex conjugate[9]*.*

The property embodied by (2.93) is often referred to as *Hermiticity*.[10]

Proof. We use integration by parts:

$$-\frac{\hbar^2}{2m}\int_0^a \bar{\alpha}\frac{d^2\beta}{dx^2}dx = -\frac{\hbar^2}{2m}\left\{\left[\bar{\alpha}\frac{d\beta}{dx}\right]_0^a - \int_0^a \frac{d\bar{\alpha}}{dx}\frac{d\beta}{dx}dx,\right\}, \tag{2.94}$$

where the first term in the curly brackets is zero since $\overline{\alpha(a)} = \overline{\alpha(0)} = 0$. We then integrate the remaining terms by parts:

$$-\frac{\hbar^2}{2m}\left\{-\left[\frac{d\alpha}{dx}\beta\right]_0^a + \int_0^a \frac{d^2\bar{\alpha}}{dx^2}\beta dx\right\}, \tag{2.95}$$

where, again, the first term in the curly brackets is zero since $\overline{\beta(a)} = \overline{\beta(0)} = 0$. Hence, we finally obtain:

$$= -\frac{\hbar^2}{2m}\int_0^a \overline{\left(\frac{d^2\alpha}{dx^2}\right)}\beta dx, \tag{2.96}$$

which was the result to be proved.

□

Now we describe some further consequences of this theorem for the square well.

1. The energy levels of the square well are all real. This can be seen by a direct calculation and the use of Theorem 6. We have:

$$H\psi_n = E_n\psi_n \quad \text{and} \quad \int_0^a \bar{\psi}_n\psi_n dx = 1, \tag{2.97}$$

from which it follows that:

$$E_n = \int_0^a \bar{\psi}_n H\psi_n dx. \tag{2.98}$$

[9] Note that (2.92) is the Hamiltonian for the square well.

[10] You should compare (2.93) with the general definition of self-adjoint of Hermitian operator given in Definition **??**.

Now we apply Theorem 6 to this result with $\alpha = \beta = \psi_n$ (which satisfies the hypotheses of the theorem) to obtain:

$$E_n = \int_0^a \left(\overline{H\psi_n}\right) \psi_n dx = \int_0^a \bar{E}_n \bar{\psi}_n \psi_n dx = \bar{E}_n, \tag{2.99}$$

from which it follows that the energy levels are real.[11]

[11] You should compare this calculation with (1.63).

2. Now we show that the eigenstates satisfy an orthonormality condition, i.e.,

$$\int_0^a \bar{\psi}_m \psi_n dx = \delta_{mn}. \tag{2.100}$$

We can assume that we have normalized the eigenstates and that we have

$$\int_0^a \bar{\psi}_n \psi_n dx = 1, \tag{2.101}$$

therefore we only need to consider the case $n \neq m$. Integrating the expression $H\psi_n = E_n\psi_n$ gives:

$$E_n \int_0^a \bar{\psi}_m \psi_n dx = \int_0^a \bar{\psi}_m H \psi_n dx. \tag{2.102}$$

Applying Theorem 6 to this last expression, and using the fact that $\overline{E_n} = E_n$ gives

$$\int_0^a \left(\overline{H\psi_m}\right) \psi_n dx = E_m \int_0^a \bar{\psi}_m \psi_n dx \tag{2.103}$$

Hence

$$\left(E_n - E_m\right) \int_0^a \bar{\psi}_m \psi_n dx = 0. \tag{2.104}$$

Since the energy levels are non-degenerate we have:[12]

[12] You should compare this calculation with (1.65).

$$\int_0^a \bar{\psi}_m \psi_n dx = 0 \quad \text{if} \quad m \neq m. \tag{2.105}$$

This result for the square well can be generalized as follows. Let $H = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + V(x)$, and $\alpha(x)$ and $\beta(x)$ be any two functions, defined on $\mathbb{R}$, that are twice differentiable and satisfy $\alpha(x) \to 0$, $\beta(x) \to 0$ as $|x| \to \infty$. Then we can prove the following result

Theorem 7.

$$\int_{-\infty}^{+\infty} \bar{\alpha} H \beta dx = \int_{-\infty}^{+\infty} \left(\overline{H\alpha}\right) \beta dx \tag{2.106}$$

Proof. We will leave the proof of this result as an exercise. □

Theorem 7 has consequences similar to the consequences of Theorem 6 for the square well. In particular, The energy levels E_n and the corresponding eigenfunctions ψ_n satisfying the time independent Schrödinger equation

$$H\psi_n = E_n\psi_n, \tag{2.107}$$

have the property that:

1. $E_n \in \mathbb{R}$
2. $\int_{-\infty}^{+\infty} \bar{\psi}_m \psi_n dx = \delta_{mn}$

2.7 *Expectation Values*

Recall that the average position of a particle in the square well at time t is

$$\langle x \rangle = \int_0^a x|\psi(x,t)|^2 dx. \tag{2.108}$$

$\langle x \rangle$ is called the expectation value of x.

Definition 23 (Expectation Value of Momentum)**.** *We define the expectation value of the momentum p (i.e. the average value of the momentum) at time t as*

$$\langle p \rangle = m\frac{d\langle x \rangle}{dt}. \tag{2.109}$$

We have the following characterization of the expectation value of the momentum of a particle in the square well.

Theorem 8. *For a particle in a square well represented by a normalized wave function $\psi(x,t)$*

$$\langle p \rangle = \int_0^a \bar{\psi}\left(-i\hbar\frac{\partial\psi}{\partial x}\right) dx. \tag{2.110}$$

Proof. Differentiating (2.108) with respect to time gives:

$$\frac{d\langle x \rangle}{dt} = \int_0^a x\frac{\partial\bar{\psi}}{\partial t}\psi dx + \int_0^a x\bar{\psi}\frac{\partial\psi}{\partial t}dx \tag{2.111}$$

Recall that the Shr$'$odinger equation and its complex conjugate are given by the following two equations:

$$\frac{\partial\psi}{\partial t} = \frac{1}{i\hbar}\left(-\frac{\hbar^2}{2m}\frac{\partial^2\psi}{\partial x^2} + V\psi\right) \tag{2.112}$$

$$\frac{\partial\bar{\psi}}{\partial t} = -\frac{1}{i\hbar}\left(-\frac{\hbar^2}{2m}\frac{\partial^2\bar{\psi}}{\partial x^2} + V\bar{\psi}\right) \tag{2.113}$$

Substituting these two equations into (2.111) and multiplying by m gives (note that $V = 0$ for the square well):

$$\langle p \rangle = \frac{\hbar}{2i} \int_0^a x\psi \frac{\partial^2 \bar{\psi}}{\partial x^2} dx - \frac{\hbar}{2i} \int_0^a x\bar{\psi} \frac{\partial^2 \bar{\psi}}{\partial x^2} dx \tag{2.114}$$

Using Theorem 6 for the first integral gives:

$$\begin{aligned} \langle p \rangle &= \frac{\hbar}{2i} \int_0^a \bar{\psi} \left(\frac{\partial^2}{\partial x^2}(x\psi) - x\frac{\partial^2 \psi}{\partial x^2} \right) dx, \\ &= \frac{\hbar}{2i} \int_0^a \bar{\psi} \left(x\frac{\partial^2 \psi}{\partial x^2} + 2\frac{\partial \psi}{\partial x} - x\frac{\partial^2 \psi}{\partial x^2} \right) dx, \\ &= \frac{\hbar}{i} \int_0^a \bar{\psi} \frac{\partial \psi}{\partial x} dx. \end{aligned} \tag{2.115}$$

□

We make the following remarks.

1. The proof uses only the Schrödinger equation and Theorem 6, and it generalizes to higher dimensions in a straightforward manner.
2. The generalization to higher dimensions is

$$\langle \mathbf{p} \rangle = \int_{\mathbb{R}^d} \bar{\psi} \left(\frac{\hbar}{i} \nabla \psi \right) d^d \mathbf{r}. \tag{2.116}$$

3. Note that this result is consistent with the observation that $\frac{\hbar}{i}\nabla\psi =$ momentum $\times\, \psi$.

This suggests that the quantum expectation value of the energy of a particle in a square well in a state represented by $\psi(x,t)$ is:

$$\langle E \rangle = \int_0^a \bar{\psi} \left(-\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} \psi \right) dx = \int_0^a \bar{\psi} i\hbar \frac{\partial \psi}{\partial t} dx. \tag{2.117}$$

Note that is we expand $\psi(x,t)$ in terms of eigenfunctions $\psi_n(x)$

$$\psi(x,t) = \sum_n c_n \psi_n(x) e^{-i\frac{E_n t}{\hbar}}. \tag{2.118}$$

we have, using $\int_0^a \psi_n(x)\psi_m(x)dx = \delta_{nm}$

$$\langle E \rangle = \sum_n |c_n|^2 E_n. \tag{2.119}$$

which is consistent with the assumption that $|c_n|^2$ is the probability of measuring the particle to be in the n^{th} eigenfunction.

The generalization to a particle described by the classical Hamiltonian

$$H(q,p) = \frac{p^2}{2m} + V(q), \tag{2.120}$$

is given by:

$$\langle E\rangle = \int_{\mathbb{R}^d} \bar{\psi}\left(-\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{q})\right)\psi d^d\mathbf{q} = \int_{\mathbb{R}^d} \bar{\psi}\left(i\hbar\frac{\partial\psi}{\partial t}\right) d^d\mathbf{q}. \quad (2.121)$$

2.8 Scattering and Tunneling

In this section several piecewise constant potential energy functions are considered that allow us to describe and analyze the phenomena of scattering and tunneling. Towards this end, the notion of the probability current will play an important role.

However, first we begin with an important observation related to classical mechanical motion. Consider one dimensional motion defined by the Hamiltonian function:

$$H = \frac{p^2}{2m} + V(x), \quad (x,p) \in \mathbb{R}^2. \quad (2.122)$$

For a fixed energy, E, i.e., $\frac{p^2}{2m} + V(x) = E$, classical motion is allowable for $E \geq V$ since

$$p = \pm\sqrt{2m(E-V)}. \quad (2.123)$$

Quantum mechanically, this constraint can be violated, as we will see.

Constant Potential

We begin by considering a simple case that will allow us to establish some of the basic ideas that we will use in more complicated situations afterwards. We consider $V(x) = V =$ constant, for which the time independent Schrödinger equation is given by

$$-\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2} + V\psi = E\psi. \quad (2.124)$$

We consider the two cases.

$\boxed{E > V.}$

We define the wavenumber as $k = \frac{1}{\hbar}\sqrt{2m(E-V)}$ In this case the time independent Schrödinger equation takes the form:

$$-\frac{d^2\psi}{dx^2} = k^2\psi, \quad (2.125)$$

which has the solution

$$\psi(x) = Ae^{ikx} + Be^{-ikx}. \quad (2.126)$$

As discussed in Section 2.3, it follows from

$$-i\hbar\frac{d}{dx}e^{\pm ikx} = \pm\hbar k e^{\pm ikx}. \tag{2.127}$$

that $-i\hbar\frac{d}{dx} = \pm\hbar$ is the momentum. The two parts of the wavefunction have the following interpretation:

$$\begin{array}{lll} Ae^{ikx}: & \text{particle moving to the right with momentum} & \sqrt{2m(E-V)} \\ Be^{-ikx}: & \text{particle moving to the left with momentum} & -\sqrt{2m(E-V)} \end{array} \tag{2.128}$$

The probability current has the form:

$$\begin{aligned} j(x,t) = \frac{\hbar}{m}\mathrm{Re}\bar{\psi}\frac{1}{i}\frac{\partial\psi}{\partial x} &= \frac{\hbar}{m}\mathrm{Re}\left(k|A|^2 - k|B|^2 + kA\bar{B}e^{2ikx} - k\bar{A}Be^{-2ikx}\right) \\ &= \frac{\hbar k}{m}\left(|A|^2 - |B|^2\right) \end{aligned} \tag{2.129}$$

Note that the plane wave Ae^{ikx} extends over all of $\mathbb{R}$. They cannot be normalized because $|\psi(x)|^2 = A^2 = \text{constant}$. Moreover, recall that the notion of probability current was developed in the context of normalized wavefunctions. Nevertheless, we can think of this notion on intervals of finite length. In particular, for our applications we will consider the probability current in a finite interval containing the step. This enables the probability current to play a role in determining matching conditions for the wavefunction on either side of an obstacle. We will see this in Section 2.8. We also remark that in these applications the wavefunction can be thought of as representing a beam of particles of finite flux j.

$\boxed{E < V.}$

In this case we set

$$\tilde{k} = \frac{1}{\hbar}\sqrt{2m(V-E)}, \tag{2.130}$$

and the time independent Schrödinger equation takes the form:

$$\frac{d^2\psi}{dx^2} = \tilde{k}^2\psi, \tag{2.131}$$

which has the solution

$$\psi(x) = Ae^{-\tilde{k}x} + Be^{\tilde{k}x}. \tag{2.132}$$

Hence, the wavefunction has exponentially growing solutions as $|x| \to \infty$ for nonzero A and B. This will require careful consideration when choosing an allowable form of the wavefunction in specific settings, as we will see.

Step Potentials

We next consider a potential of the form

$$V(x) = \begin{cases} 0 & x < 0 \\ V & x > 0 \end{cases} \tag{2.133}$$

with $V > 0$.

Solutions of the time independent Schrödinger equation were given in Section 2.8. Matching conditions for the wavefunction at the step are obtained from the continuity equation for the probability current (recall Theorem 3 and Section 2.1):

$$\frac{\partial}{\partial t}|\psi|^2 + \frac{\partial j}{\partial x} = 0. \tag{2.134}$$

Requiring the probability current to be independent of time gives:

$$\frac{\partial j}{\partial x} = 0 \Rightarrow j = \text{constant}. \tag{2.135}$$

Conservation of probability across the step implies that we have the following ''matching condition" for the probability current at the step:

$$\frac{\hbar}{m}\text{Re}\left(\bar{\psi}\frac{1}{i}\frac{\partial \psi}{\partial x}\right)|_{x=0-} = \frac{\hbar}{m}\text{Re}\left(\bar{\psi}\frac{1}{i}\frac{\partial \psi}{\partial x}\right)|_{x=0+} \tag{2.136}$$

The matching condition will be satisfied if we require $\psi(x)$ and $\frac{\partial \psi}{\partial x}(x)$ to be continuous across a finite step.

$\boxed{E > V.}$

For definiteness, consider particles incident from left to right. Classically all particles continue from $x < 0$ to $x > 0$ with momenta:

$$p = \begin{cases} \sqrt{2mE} & x < 0 \\ \sqrt{2m(E-V)} & x > 0 \end{cases} \tag{2.137}$$

Quantum mechanically, the solution of the time independent Schrödinger equation is given by:

$$\psi(x) = \begin{cases} A_0 e^{ik_0 x} + B_0 e^{-ik_0 x}, & x < 0, \\ A_1 e^{ik_1 x} + B_1 e^{-ik_1 x}, & x > 0, \end{cases} \tag{2.138}$$

where

$$k_0 = \frac{1}{\hbar}\sqrt{2mE} \qquad k_1 = \frac{1}{\hbar}\sqrt{2m(E-V)} \tag{2.139}$$

Since the particles are incoming from left to right there is no component of the wavefunction moving from right to left in $x > 0$. This implies that $B_1 = 0$

Continuity of $\psi(x)$ at $x = 0$ implies

$$A_0 + B_0 = A_1. \tag{2.140}$$

Continuity of $\frac{\partial \psi}{\partial x}(x)$ at $x = 0$ implies

$$k_0(A_0 - B_0) = k_1 A_1. \tag{2.141}$$

Using (2.140) and (2.141), after some algebra we obtain:

$$\begin{aligned} A_1 &= \tfrac{2k_0}{k_0+k_1} A_0, \\ B_0 &= A_1 - A_0 = \tfrac{k_0-k_1}{k_0+k_1} A_0. \end{aligned} \tag{2.142}$$

We saw in Section 2.8 that for a wavefunction of the form $\psi(x) = Ae^{ikx} + Be^{-ikx}$ the probability current has the form:

$$j = \frac{\hbar k}{m} \left(|A|^2 - |B|^2 \right). \tag{2.143}$$

Using the fact that the probability current is conserved, and therefore it has the same value on each side of the step, we obtain:

$$k_0 |A_0|^2 - k_0 |B_0|^2 = k_1 |A_1|^2. \tag{2.144}$$

The terms in (2.144) have the following interpretation:

$$\begin{aligned} k_0|A_0|^2 &: \text{ probability current from left to right for } x < 0, \\ k_0|B_0|^2 &: \text{ probability current from right to left for } x < 0, \\ k_1|A_1|^2 &: \text{ probability current from left to right for } x > 0. \end{aligned} \tag{2.145}$$

We rewrite (2.144) as:

$$1 - \frac{|B_0|^2}{|A_0|^2} = \frac{k_1}{k_0} \frac{|A_1|^2}{|A_0|^2}. \tag{2.146}$$

We have the following definitions.

Definition 24 (Reflection Coefficient)**.** *The reflection coefficient is defined as:*

$$R = \frac{|B_0|^2}{|A_0|^2} = \frac{\textit{reflected probability current}}{\textit{incoming probability current}}. \tag{2.147}$$

Definition 25 (Transmission Coefficient)**.** *The transmission coefficient is defined as*

$$T = \frac{k_1 |A_1|^2}{k_0 |A_0|^2} = \frac{\textit{transmitted probability current}}{\textit{incoming probability current}}. \tag{2.148}$$

Using these definitions, along with the conservation of probability current given in (2.146), we can express the conservation of probability current as:

$$R + T = 1. \tag{2.149}$$

$0 < E < V$.

For this case, for a wave incident from left to right, classically all particles are reflected from the step. You can only find particles for $x < 0$. The situation for quantum mechanics is different, as we will see.

The solution of the time independent Schrödinger equation is given by:

$$\psi(x) = \begin{cases} A_0 e^{ik_0 x} + B_0 e^{-ik_0 x}, & x < 0 \\ A_1 e^{-\tilde{k}_1 x} + B_1 e^{\tilde{k}_1 x}, & x > 0 \end{cases} \tag{2.150}$$

where

$$k_0 = \frac{1}{\hbar}\sqrt{2mE}, \qquad \tilde{k}_1 = \frac{1}{\hbar}\sqrt{2m(V-E)}. \tag{2.151}$$

For $\psi(x)$ to remain bounded as $x \to \infty$ we require $B_1 = 0$. The matching conditions at the step for ψ are:

$$\begin{aligned} A_0 + B_0 &= A_1, \\ ik_0(A_0 - B_0) &= -\tilde{k}_1 A_1. \end{aligned} \tag{2.152}$$

Using simple algebraic manipulations, (2.152) can be rewritten as:

$$A_1 = \frac{2k_0}{k_0 + i\tilde{k}_1} A_0 \qquad B_0 = \frac{k_0 - i\tilde{k}_1}{k_0 + i\tilde{k}_1} A_0, \tag{2.153}$$

where the latter expression can be simplified to

$$|B_0| = |A_0|, \tag{2.154}$$

which can be written as:

$$\frac{|B_0|^2}{|A_0|^2} = 1. \tag{2.155}$$

From this relation, and the definition of the reflection coefficient, we have:

$$R = \frac{|B_0|^2}{|A_0|^2} = 1, \tag{2.156}$$

and therefore

$$T = 1 - R = 0 \tag{2.157}$$

Nevertheless, for $x > 0$

$$|\psi|^2 = |A_1|^2 e^{-2\tilde{k}_1 x}. \tag{2.158}$$

Hence there is a finite probability that the particle can be found in $x > 0$. This is related to the quantum mechanical phenomenon of tunneling.

While our analysis shows that probability density can be found in the classically forbidden region, i.e. ??inside the barrier??, since the barrier is infinite in length, and the probability density decays at an exponential rate with respect to x, it can never ??get out the other side??. Next we will consider the case of a finite barrier where the possibility of probability current being transmitted through the barrier exists. This is referred to as *quantum tunnelling*.

A Barrier of Finite Width: Quantum Tunnelling

We next consider the case of a barrier of finite width on the line. This is described by the potential

$$V(x) = \begin{cases} 0 & x < 0, \\ V & 0 < x < a, \\ 0 & x > a. \end{cases} \tag{2.159}$$

The particle is incident from left to right in $x < 0$, and we will only consider the case $0 < E < V$. In this case the solution to the time independent Schrödinger equation is given by:

$$\psi(x) \begin{cases} A_0 e^{ik_0 x} + B_0 e^{-ik_0 x}, & x < 0, \\ A_1 e^{-\tilde{k}_1 x} + B_1 e^{\tilde{k}_1 x}, & 0 < x < a, \\ A_2 e^{ik_0 x} + B_2 e^{-ik_0 x}, & x > a, \end{cases} \tag{2.160}$$

where

$$k_0 = \frac{1}{\hbar}\sqrt{2mE}, \qquad \tilde{k}_1 = \frac{1}{\hbar}\sqrt{2m(V-E)}. \tag{2.161}$$

We note that $B_2 = 0$ since there is no wave moving to the right for $x > a$.

We will compute the reflection coefficient, R from the boundary $x = 0$, and the transmission coefficient, T, across the boundary $x = a$. These are given by:

$$R = \frac{|B_0|^2}{|A_0|^2}, \quad T = \frac{|A_2|^2}{|A_0|^2}. \tag{2.162}$$

We apply the matching conditions at the boundaries to (2.160) in order to obtain relations amongst the coefficients. The matching conditions give:

$$\text{at } x = 0 \quad \begin{matrix} \psi: & A_0 + B_0 = A_1 + B_1, & \text{(i)}, \\ \frac{d\psi}{dx}: & ik_0(A_0 - B_0) = -\tilde{k}(A_1 - B_1), & \text{(ii)}, \end{matrix} \tag{2.163}$$

$$\text{at } x = a \quad \begin{matrix} \psi: & A_1 e^{-\tilde{k}_1 a} + B_1 e^{\tilde{k}_1 a} = A_2 e^{ik_0 a}, & \text{(iii)}, \\ \frac{d\psi}{dx}: & -\tilde{k}_1 \left(A_1 e^{-\tilde{k}_1 a} - B_1 e^{\tilde{k}_1 a} \right) = ik_0 A_2 e^{ik_0 a}, & \text{(iv)}. \end{matrix} \tag{2.164}$$

It is clear from the form of (2.162) that if we can express $B_0 = \text{constant} A_0$ and $A_2 = \text{constant}' A_0$, where constant and constant$'$ are functions of the system parameters a, k_0, $\tilde{k}_1$ then expressions for the reflection and transmission coefficients in terms of these same system parameters easily follow. After some algebra, we obtain:

$$A_1 = \frac{1}{2}(1 - \frac{ik_0}{\tilde{k}_1}) A_2 e^{ik_0 a} e^{\tilde{k}_1 a} \quad B_1 = \frac{1}{2}(1 + \frac{ik_0}{\tilde{k}_1}) A_2 e^{ik_0 a} e^{-\tilde{k}_1 a} \tag{2.165}$$

$$\begin{aligned} A_0 &= \frac{1}{2} A_1 \left(1 - \frac{\tilde{k}_1}{ik_0} \right) + \frac{1}{2} B_1 \left(1 + \frac{\tilde{k}_1}{ik_0} \right), \\ &= \frac{1}{4} \left(1 - \frac{ik_0}{\tilde{k}_1} + 1 + \frac{i\tilde{k}_1}{k_0} \right) A_2 e^{ik_0 a} e^{\tilde{k}_1 a} + \frac{1}{4} \left(1 + \frac{ik_0}{\tilde{k}_1} + 1 - \frac{i\tilde{k}_1}{k_0} \right) A_2 e^{ik_0 a} e^{-\tilde{k}_1 a}, \\ &= \left(\cosh(\tilde{k}_1 a) + \frac{i}{2} \left(\frac{\tilde{k}_1}{k_0} - \frac{k_0}{\tilde{k}_1} \right) \sinh(\tilde{k}_1 a) \right) A_2 e^{ik_0 a}, \end{aligned} \tag{2.166}$$

$$B_0 = \frac{1}{2} A_1 \left(1 + \frac{\tilde{k}_1}{ik_0} \right) + \frac{1}{2} B_1 \left(1 - \frac{\tilde{k}_1}{ik_0} \right) = \frac{i}{2} \left(-\frac{\tilde{k}_1}{k_0} - \frac{k_0}{\tilde{k}_1} \right) \sinh(\tilde{k}_1 a) A_2 e^{ik_0 a}. \tag{2.167}$$

Therefore the reflection coefficient is given by

$$R = \frac{|B_0|^2}{|A_0|^2} = \frac{\frac{(\tilde{k}_1^2 + k_0^2)^2}{4k_0^2 \tilde{k}_1^2} \sinh^2(\tilde{k}_1 a)}{\cosh^2(\tilde{k}_1 a) + \frac{(\tilde{k}_1^2 - k_0^2)^2)}{4k_0^2 \tilde{k}_1^2} \sinh^2(\tilde{k}_1 a)} = \frac{\frac{(\tilde{k}_1^2 + k_0^2)^2}{4k_0^2 \tilde{k}_1^2} \sinh^2(\tilde{k}_1 a)}{1 + \frac{(\tilde{k}_1^2 + k_0^2)^2}{4k_0^2 \tilde{k}_1^2} \sinh^2(\tilde{k}_1 a)}, \tag{2.168}$$

and the transmission coefficient is given by

$$T = \frac{|A_2|^2}{|A_0|^2} = \frac{1}{1 + \frac{(\tilde{k}_1^2 + k_0^2)^2}{4k_0^2 \tilde{k}_1^2} \sinh^2(\tilde{k}_1 a)}. \tag{2.169}$$

Hence we have quantum transmission, even though classically it is impossible. This is called *quantum tunneling*. . As a check on these calculations we easily see that $R + T = 1$.

It is insightful to consider some limiting cases. For $\tilde{k}_1 a$ "very large" we have

$$T \approx \frac{4k_0^2\tilde{k}_1^2}{(\tilde{k}_1^2 + k_0^2)^2} e^{-2\tilde{k}_1 a}. \tag{2.170}$$

Hence the transmission coefficient becomes "small", and therefore reflection dominates.. Alternately, for
and for $\tilde{k}_1 a$ "very small" the transmission coefficient takes the form

$$T \approx 1 - \frac{(\tilde{k}_1^2 + k_0^2)^2}{4k_0^2} a^2, \tag{2.171}$$

and transmission dominates over reflection.

Problems

1. A particle of mass m moves freely in the interval $[0, a]$ on the x-axis. The wavefunction describing it may be assumed to vanish at $x = 0$ and $x = a$. Initially the wavefunction is given by:

$$\psi(x,0) = \frac{1}{\sqrt{a}} \sin\left(\frac{\pi x}{a}\right)\left(1 + 2\cos\left(\frac{\pi x}{a}\right)\right). \tag{2.172}$$

Show that at a later time t the wavefunction is given by:

$$\psi(x,t) = \frac{1}{\sqrt{a}} \exp\left(-\frac{i\pi^2\hbar t}{2ma^2}\right) \sin\left(\frac{\pi x}{a}\right)\left(1 + 2\exp\left(-\frac{i3\pi^2\hbar t}{2ma^2}\right)\cos\left(\frac{\pi x}{a}\right)\right). \tag{2.173}$$

Find the probability that at time t the particle lies in the interval $[0, a/2]$. What is the probability that the particle's energy will be measured to be that of the n^{th} energy level of the system?

2. Repeat the calculations in Problem 1 for the case when the initial condition is given by:

$$\psi(x,0) = \frac{12}{\sqrt{a^3}} \begin{cases} x & \text{if} \quad 0 < x \leq \frac{a}{2} \\ a - x & \text{if} \quad \frac{a}{2} \leq x < a, \end{cases} \tag{2.174}$$

and verify that the wavefunction is normalized.

2. A particle of mass m, moving freely between impenetrable barriers at $x = 0$ and $x = a$ (you may assume that quantum wavefunctions vanish at such barriers), is in the eigenstate corresponding to the lowest energy level when the barrier at $x = a$ is suddenly displaced to $x = 2a$, at time $t = 0$. By expanding the original wavefunction in terms of the eigenfunctions for motion within $[0, 2a]$, find the wavefunction at a subsequent time t, and show that it is a superposition of states of energies

$$E = \frac{n^2\pi^2\hbar^2}{8ma^2}. \tag{2.175}$$

for $n = 2$ and $n = 1, 3, 5, \ldots$ Show that the probability of finding the particleâĂŹs energy to be unchanged is $\frac{1}{2}$.

4. Consider the time-independent Schrödinger equation

$$-\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2}(x) + V(x)\psi(x) = E\psi(x), \tag{2.176}$$

with the boundary condition that $\psi(x) \to 0$ as $|x| \to \infty$, sufficiently fast for the solutions to be normalizable. Prove that the energy levels are all non-degenerate.

5. Let

$$H = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + V(x), \tag{2.177}$$

and $\alpha(x)$ and $\beta(x)$ be any two functions that are twice differentiable and satisfy $\alpha(x) \to 0$ and $\beta(x) \to 0$ as $|x| \to \infty$. Prove that

$$\int_{-\infty}^{\infty} \alpha^* H\beta \, dx = \left(\int_{-\infty}^{\infty} \beta^* H\alpha \, dx \right)^*. \tag{2.178}$$

6. Let $\psi(x,t)$ be a solution of the Sch$'$odinger equation

$$H\psi = i\hbar\frac{\partial\psi}{\partial t}. \tag{2.179}$$

where H is as in (2.177) and $V(x) \to 0$ as $|x| \to \infty$. . Show that the expectation value of the momentum for a particle in the state represented by ψ is given by

$$\langle p \rangle = -i\hbar \int_{-\infty}^{\infty} \psi^* \frac{\partial\psi}{\partial x} \, dx. \tag{2.180}$$

7. Let $\psi(x,t)$ be a solution of the Sch$'$odinger equation in the previous problem. Show that for operators A that do not depend explicitly on time t,

$$i\hbar\frac{d}{dt}\int_{-\infty}^{\infty} \psi^*(x,t) A\psi(x,t) \, dx = \int_{-\infty}^{\infty} \psi^*(x,t)[A,H]\psi(x,t) \, dx, \tag{2.181}$$

where $[A,H] = AH - HA$. Hence prove that

1. the expectation value of the energy is a constant of the (quantum) dynamics when the potential V is time-independent;
2. $$\frac{d}{dt}\langle p \rangle = -\int_{-\infty}^{\infty} \psi^*(x,t)\frac{dV}{dx}\psi(x,t) \, dx. \tag{2.182}$$

Comment on the relationship with the corresponding results in classical mechanics.

8. A particle of mass m moves freely in the interval $[0,a]$ on the x-axis. The wavefunction describing it may be assumed to vanish at $x = 0$ and $x = a$. Initially the wavefunction is given by:

$$\psi(x,0) = \frac{1}{\sqrt{a}} \sin\left(\frac{\pi x}{a}\right)\left(1 + 2\cos\left(\frac{\pi x}{a}\right)\right). \tag{2.183}$$

In problem 1 we showed that at a later time t the wavefunction is given by:

$$\psi(x,t) = \frac{1}{\sqrt{a}} \exp\left(-\frac{i\pi^2\hbar t}{2ma^2}\right) \sin\left(\frac{\pi x}{a}\right) \left(1 + 2\exp\left(-\frac{i3\pi^2\hbar t}{2ma^2}\right) \cos\left(\frac{\pi x}{a}\right)\right). \tag{2.184}$$

Compute the expectation value of the energy in this state, first by calculating

$$i\hbar \int_0^a \psi^* \frac{\partial \psi}{\partial t} dx, \tag{2.185}$$

and second by using the probability for the particle to be found in the n^{th} energy eigenstate of the system.

9. Let H be an operator which commutes with the parity operator P (i.e. $[P, H] = 0$). Show that any eigenfunction of H is also an eigenfunction of P, with eigenvalue either 1 or 1.

10. A beam of particles with energy $E > 0$, quantum mechanically described by a plane wave incident from $-\infty$, moves along the x-axis in the potential

$$V(x) = \begin{cases} 0 & \text{if} \quad x < 0, \\ -V & \text{if} \quad x > 0 \end{cases} \tag{2.186}$$

where $V > 0$. Find the reflection and transmission coefficients, R and T. Show that, contrary to the classical result, $R \to 1$ in the limit as $\frac{V}{E} \to \infty$.

11. A beam of particles with energy E, quantum mechanically described by a plane wave incident from $-\infty$, move along the x-axis in the potential

$$V(x) = \begin{cases} 0 & \text{if} \quad x < 0, \\ V & \text{if} \quad 0 < x < a, \\ 0 & x > a \end{cases} \tag{2.187}$$

where $E > V > 0$. Find the reflection and transmission coefficients, R and T. Show that when

$$\sqrt{2m(E-V)} = n\pi\hbar, \tag{2.188}$$

where $n = 0, 1, 2, \ldots$ then $T = 1$ (and hence $R = 0$).

12. Consider a particle in the infinite square well of width 1 (the potential is $V(x) = 0$ on $[0, 1]$ and ∞ everywhere else). Recall that the time-independent Schrödinger equation has eigenvalues at

$$E_n = \frac{n^2\pi^2\hbar^2}{2m},$$

with associated wavefunctions

$$\psi_n(x) = \sqrt{2}\sin(n\pi x).$$

Suppose that the particle in this well has as its initial wave function an even mixture of the first two stationary states:

$$\Psi(x, 0) = A(\psi_1(x) + \psi_2(x)),$$

for some normalization constant A.

(a) Normalize $\Psi(x, 0)$.

(b) Find $\Psi(x, t)$ and $|\Psi(x, t)|^2$. Express the latter as a sinusoidal function of time. To simplify the result, let $\omega = \pi^2\hbar/2m$.

(c) Compute $\langle x \rangle$.

Hint. Recall that

$$\cos^2(\theta/2) = \frac{1 + \cos\theta}{2},$$

and integrate by parts. Furthermore,

$$\sin(\pi x)\sin(2\pi x) = \frac{1}{2}(\cos(\pi x) - \cos(3\pi x)).$$

(d) Compute $\langle p \rangle$.

(e) If you measured the energy of this particle, what values might you get, and what is the probability of getting each of them? Find the expectation value of H.

3
Measurement, Uncertainty, Time Evolution, and the Harmonic Oscillator

We will use the mathematical structure developed in the first chapter to study the quantum theory of measurement, uncertainty, and time evolution. We will conclude this chapter with a study of one of the fundamental quantum mechanical systems–the harmonic oscillator.

3.1 Measurement

Now we will describe some aspects of the "quantum theory of measurement". This will play an important role in the remainder of the course. We begin by recalling two facts that were mentioned earlier.

- The state of a physical system is described by a normalised ket, $|\psi\rangle$, in a Hilbert space $\mathcal{H}$.
- Every measurable physical quantity of a physical system ("observable") is described by a self-adjoint operator, A, acting on the Hilbert space $\mathcal{H}$.

We now give a brief description of the three "rules" or postulates of measurement. Often these, collectively go by the name of the "Born rule", since they were formulated in a paper of Max Born published in 1926.

Zur Quantenmechanik der Stoßvorgänge, Max Born, Zeitschrift für Physik, 37, #12 (Dec. 1926), pp. 863-867 (German); English translation, On the quantum mechanics of collisions, in Quantum theory and measurement, section I.2, J. A. Wheeler and W. H. Zurek, eds., Princeton, NJ: Princeton University Press, 1983, ISBN 0-691-08316-9.

Outcome of a Measurement. The only possible outcome of the measurement of a physical observable A is an eigenvalue of A, λ.

Probability for a Particular Outcome of a Measurement. Suppose the system is in the state, $\mid \psi\rangle$, and the observable A is measured. Then the eigenvalue of A, λ, occurs with the probability:

$$\text{prob}_{\psi}(\lambda) \equiv \langle \psi \mid P_\lambda \mid \psi \rangle = \| P_\lambda \mid \psi \rangle \|^2, \tag{3.1}$$

where P_λ is the orthogonal projection onto the subspace spanned by the eigenvectors corresponding to λ.

State of the System After Measurement. If the measurement of A on the system that is in the state $\mid \psi\rangle$ gives the outcome λ, then the state of the system after the measurement is given by:

$$\frac{P_\lambda \mid \psi\rangle}{\sqrt{\langle \psi \mid P_\lambda \mid \psi \rangle}}, \tag{3.2}$$

where, recall

$$\| P_\lambda \mid \psi \rangle \| = \sqrt{\langle \psi \mid P_\lambda^\dagger P_\lambda \mid \psi \rangle} = \sqrt{\langle \psi \mid P_\lambda^2 \mid \psi \rangle} = \sqrt{\langle \psi \mid P_\lambda \mid \psi \rangle}. \tag{3.3}$$

Summary: *Measurement of an observable yields an eigenvalue. If we know the state before the measurement, we can calculate the probability of obtaining a particular eigenvalue as a result of measurement in that state. However, the measurement changes the state, but we can calculate that state after the measurement, provided we know the state in which the measurement is made, and the outcome of the measurement.*

Examples. We now consider some examples that illustrate the calculations and formalism in the quantum theory of measurement.

Let λ_i and $\mid e_i\rangle$ denoted the eigenvalues and (normalised) eigenvectors of a self-adjoint linear operator A defined on a finite dimensional Hilbert space. Let us suppose that there are no degeneracies, then the eigenspace corresponding to an eigenvalue λ_i is one dimensional and the associated projection operators are given by:

$$P_{\lambda_i} = \mid e_i\rangle\langle e_i \mid, \tag{3.4}$$

and A can be represented in the following spectral form:

$$A = \sum_i \lambda_i \mid e_i\rangle\langle e_i \mid, \quad \langle e_i \mid e_i \rangle = 1. \tag{3.5}$$

- Suppose the system is in the *normalised* state:

$$\mid \psi\rangle = \sum_i a_i \mid e_i\rangle, \tag{3.6}$$

and suppose A is measured on the state $\mid \psi\rangle$. Then, using (3.1), the probability that the outcome of the measurement of A in $\mid \psi\rangle$ is λ_i is given by:

$$\text{prob}_\psi(\lambda_i) = \langle \psi \mid P_{\lambda_i} \mid \psi \rangle = \langle \psi \mid e_i \rangle \langle e_i \mid \psi \rangle = \mid \langle e_i \mid \psi \rangle \mid^2 = \mid a_i \mid^2 . \tag{3.7}$$

Using (3.2), if λ_i is measured, then after the measurement the state is given by:

$$\frac{P_{\lambda_i} \mid \psi\rangle}{\sqrt{\langle \psi \mid P_{\lambda_i} \mid \psi \rangle}} = \frac{\mid e_i \rangle \langle e_i \mid \psi \rangle}{\mid \langle e_i \mid \psi \rangle \mid} = \frac{\langle e_i \mid \psi \rangle}{\mid \langle e_i \mid \psi \rangle \mid} \mid e_i \rangle = \frac{a_i}{\mid a_i \mid} \mid e_i \rangle. \tag{3.8}$$

- Note that we have:

$$\begin{aligned} \sum_i \text{prob}_\psi(\lambda_i) &= \sum_{\lambda_i} \langle \psi \mid P_{\lambda_i} \mid \psi \rangle = \sum_i \langle \psi \mid e_i \rangle \langle e_i \mid \psi \rangle, \\ &= \langle \psi \mid \left(\sum_i \mid e_i \rangle \langle e_i \mid \right) \mid \psi \rangle = \langle \psi \mid \psi \rangle = 1. \end{aligned} \tag{3.9}$$

- Suppose $\mid \psi\rangle$ is an eigenstate of A, i.e.,

$$\mid \psi \rangle = \mid e_n \rangle.$$

Then we have

$$\text{prob}_\psi(\lambda_n) = \mid \langle e_n \mid \psi \rangle \mid^2 = 1,$$

and

$$\text{prob}_\psi(\lambda_m) = \mid \langle e_m \mid \psi \rangle \mid^2 = 0, \quad m \neq n.$$

In other words, if we measure A in an eigenstate of A the outcome of the measurement is always the eigenvalue corresponding to the eigenstate.

- Recall that the *expectation value of A in the state* $\mid \psi\rangle$, denoted $E_\psi(A) \equiv \langle \psi \mid A \mid \psi \rangle$, is the predicted *mean value* of the measurement of A in the state ψ. Using (3.5), we have:

$$\begin{aligned} E_\psi(A) &= \langle\psi \mid A \mid \psi\rangle, \\ &= \sum_i \lambda_i \langle\psi \mid e_i\rangle\langle e_i \mid \psi\rangle, \\ &= \sum_i \lambda_i \langle\psi \mid P_{\lambda_i} \mid \psi\rangle, \\ &= \sum_i \lambda_i \text{prob}_\psi(\lambda_i). \end{aligned} \tag{3.10}$$

- The *dispersion of an operator A in the state $\mid\psi\rangle$* is defined as follows:

$$\begin{aligned} \Delta_\psi(A) &= \left[E_\psi\left(\left(A - E_\psi(A)\right)^2\right)\right]^{\frac{1}{2}}, \\ &= \left[\langle\psi \mid (A^2 - 2A\langle\psi \mid A \mid \psi\rangle + \langle\psi \mid A \mid \psi\rangle^2) \mid \psi\rangle\right]^{\frac{1}{2}}, \\ &= \left[\langle\psi \mid A^2 \mid \psi\rangle - \langle\psi \mid A \mid \psi\rangle^2\right]^{\frac{1}{2}}, \\ &= \left[E_\psi(A^2) - (E_\psi(A))^2\right]^{\frac{1}{2}}. \end{aligned} \tag{3.11}$$

  This quantity will play an important role when we discuss the *generalised uncertainty relation* a bit later.

It is important to keep in mind that the expectation value of an operator and the dispersion of an operator, in general, depend on the state $\mid\psi\rangle$.

Example. Suppose $\mathcal{H} = \mathbb{C}^2$, and let

$$\{\mid 1\rangle, \mid 2\rangle\} = \left\{\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right\}$$

denote an orthonormal basis of $\mathbb{C}^2$. We define a self adjoint linear operator on $\mathbb{C}^2$ as follows:

$$\begin{aligned} A \mid 1\rangle &= 4 \mid 1\rangle - 2 \mid 2\rangle \\ A \mid 2\rangle &= -2 \mid 1\rangle + 4 \mid 2\rangle. \end{aligned} \tag{3.12}$$

In the basis $\{\mid 1\rangle, \mid 2\rangle\}$ the matrix representation of this linear operator is given by:

$$A = \begin{pmatrix} 4 & -2 \\ -2 & 4 \end{pmatrix}, \tag{3.13}$$

and we easily verify that $A = A^\dagger$.

We can also express this linear map in Dirac notation in the basis $\{|\,1\rangle, |\,2\rangle\}$ as follows:

$$A = 4\,|\,1\rangle\langle 1\,| - 2\,|\,1\rangle\langle 2\,| - 2\,|\,2\rangle\langle 1\,| + 4\,|\,2\rangle\langle 2\,|\,. \tag{3.14}$$

The eigenvalues of A are determined from:

$$\det \begin{pmatrix} 4-\lambda & -2 \\ -2 & 4-\lambda \end{pmatrix} \tag{3.15}$$

or

$$(4-\lambda)^2 - 4 = 0. \tag{3.16}$$

From which is follows that:

$$\lambda_1 = 6, \quad \lambda_2 = 2,$$

with the (normalised) eigenvectors corresponding to these eigenvalues given by:

$$|\,e_1\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad |\,e_2\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

By inspection, one can see that $|\,e_1\rangle$ and $|\,e_2\rangle$ can be expressed in terms of the basis vectors $\{|\,1\rangle, |\,2\rangle\}$ as follows:

$$|\,e_1\rangle = \frac{|\,1\rangle - |\,2\rangle}{\sqrt{2}}, \quad |\,e_2\rangle = \frac{|\,1\rangle + |\,2\rangle}{\sqrt{2}}.$$

Then it is straightforward to represent A in Dirac notation in either basis as follows:

$$\begin{aligned} A &= 6\,|\,e_1\rangle\langle e_1\,| + 2\,|\,e_2\rangle\langle e_2\,|, \\ &= 6\left(\frac{|\,1\rangle - |\,2\rangle}{\sqrt{2}}\right)\left(\frac{\langle 1\,| - \langle 2\,|}{\sqrt{2}}\right) + 2\left(\frac{|\,1\rangle + |\,2\rangle}{\sqrt{2}}\right)\left(\frac{\langle 1\,| + \langle 2\,|}{\sqrt{2}}\right), \\ &= 4\,|\,1\rangle\langle 1\,| - 2\,|\,1\rangle\langle 2\,| - 2\,|\,2\rangle\langle 1\,| + 4\,|\,2\rangle\langle 2\,|\,. \end{aligned} \tag{3.17}$$

Suppose the system is in the state $|\,1\rangle$, and we wish to measure A in this state. We know that the only possibility for the outcome of a measurement will be an eigenvalue of A. Using (3.1), the probabilities for the two possible outcomes are given by:

$$\begin{aligned} \text{prob}_{|1\rangle}(\lambda = 6) &= \langle 1\,|\,P_{\lambda_1}\,|\,1\rangle, \\ &= \langle 1\,|\,e_1\rangle\langle e_1\,|\,1\rangle, \\ &= \langle 1\,|\left(\frac{|\,1\rangle - |\,2\rangle}{\sqrt{2}}\right)\left(\frac{\langle 1\,| - \langle 2\,|}{\sqrt{2}}\right)|\,1\rangle, \\ &= \frac{1}{2}, \end{aligned} \tag{3.18}$$

and

$$\begin{aligned} \text{prob}_{|1\rangle}(\lambda = 2) &= \langle 1 \mid P_{\lambda_2} \mid 1\rangle, \\ &= \langle 1 \mid e_2\rangle\langle e_2 \mid 1\rangle, \\ &= \langle 1 \mid \left(\frac{\mid 1\rangle + \mid 2\rangle}{\sqrt{2}}\right)\left(\frac{\langle 1 \mid + \langle 2 \mid}{\sqrt{2}}\right) \mid 1\rangle, \\ &= \frac{1}{2}. \end{aligned} \tag{3.19}$$

Note that the probabilities add to one (as they should).

Using (3.10), the expectation value of A in the state $\mid 1\rangle$ is given by:

$$\begin{aligned} E_{|1\rangle}(A) &= \langle 1 \mid A \mid 1\rangle \\ &= \langle 1 \mid (4 \mid 1\rangle\langle 1 \mid -2 \mid 1\rangle\langle 2 \mid -2 \mid 2\rangle\langle 1 \mid +4 \mid 2\rangle\langle 2 \mid) \mid 1\rangle = 4. \end{aligned} \tag{3.20}$$

Using (3.11), the dispersion of A in the state $\mid 1\rangle$ is given by:

$$\Delta_{|1\rangle}(A) = \left[\langle 1 \mid A^2 \mid 1\rangle - \langle 1 \mid A \mid 1\rangle^2\right]^{\frac{1}{2}}. \tag{3.21}$$

Now we compute the various terms that go into this quantity. From (3.13) we have:

$$A^2 = \begin{pmatrix} 20 & -16 \\ -16 & 20 \end{pmatrix}, \tag{3.22}$$

which in Dirac notation is:

$$A^2 = 20 \mid 1\rangle\langle 1 \mid -16 \mid 1\rangle\langle 2 \mid -16 \mid 2\rangle\langle 1 \mid +20 \mid 2\rangle\langle 2 \mid, \tag{3.23}$$

and therefore;

$$\langle 1 \mid A^2 \mid 1\rangle = \langle 1 \mid (20 \mid 1\rangle\langle 1 \mid -16 \mid 1\rangle\langle 2 \mid -16 \mid 2\rangle\langle 1 \mid +20 \mid 2\rangle\langle 2 \mid) \mid 1\rangle = 20.$$

Assembling the different quantities, we obtain:

$$\Delta_{|1\rangle}(A) = \left[\langle 1 \mid A^2 \mid 1\rangle - \langle 1 \mid A \mid 1\rangle^2\right]^{\frac{1}{2}} = (20 - 16)^{\frac{1}{2}} = 2. \tag{3.24}$$

If the result of the measurement of A in the state $\mid 1\rangle$ is $\lambda = 6$ then, using (3.2), the state of the system after the measurement is given by:

$$\frac{\mid e_1\rangle\langle e_1 \mid 1\rangle}{\mid \langle e_1 \mid 1\rangle \mid} = \mid e_1\rangle \frac{\frac{1}{\sqrt{2}}}{\frac{1}{\sqrt{2}}} = \mid e_1\rangle$$

Interference

Interference is a unique characteristic of quantum systems that illustrates the wave-like nature of matter. We will illustrate this characteristic in the context of the example above.

Suppose we measure A in the state $\mid \psi\rangle$. If $\mid \psi\rangle = \mid 1\rangle$, then we showed that the probability that we measure $\lambda = 6$ in the state $\mid 1\rangle$ is $\frac{1}{2}$. In summary:

$$\mid \psi\rangle = \mid 1\rangle \Rightarrow \text{prob}_{|1\rangle}(\lambda = 6) = \frac{1}{2},$$

Now if $\mid \psi\rangle = \mid 2\rangle$ then the probability that we measure $\lambda = 6$ is given by:

$$\begin{aligned} \text{prob}_{|2\rangle}(\lambda = 6) &= \langle 2 \mid e_1\rangle\langle e_1 \mid 2\rangle = \langle 2 \mid \left(\frac{\mid 1\rangle - \mid 2\rangle}{\sqrt{2}}\right)\left(\frac{\langle 1 \mid -\langle 2 \mid}{\sqrt{2}}\right) \mid 2\rangle. \\ &= \frac{1}{2}. \end{aligned}$$

Now suppose that the system is in the state:

$$\mid \psi\rangle = \frac{\mid 1\rangle + \mid 2\rangle}{\sqrt{2}},$$

i.e., a linear superposition of $\mid 1\rangle$ and $\mid 2\rangle$. Then let us compute the probability to measure $\lambda = 6$ in the state $\mid \psi\rangle$:

$$\begin{aligned} \text{prob}_{|\psi\rangle}(\lambda = 6) &= \langle \psi \mid P_{\lambda=6} \mid \psi\rangle, \\ &= \langle \psi \mid e_1\rangle\langle e_1 \mid \psi\rangle, \\ &= \left(\frac{\langle 1 \mid + \langle 2 \mid}{\sqrt{2}}\right)\left(\frac{\mid 1\rangle - \mid 2\rangle}{\sqrt{2}}\right)\left(\frac{\langle 1 \mid - \langle 2 \mid}{\sqrt{2}}\right)\left(\frac{\mid 1\rangle + \mid 2\rangle}{\sqrt{2}}\right) = 0. \end{aligned} \tag{3.25}$$

Since the new state $\mid \psi\rangle$ is a (normalised) linear superposition of $\mid 1\rangle$ and $\mid 2\rangle$, with equal amplitudes, we might "guess" that the probability for measuring $\lambda = 6$ in this state is the sum of the probabilities for measuring $\lambda = 6$ in each state of the linear superposition. However, the calculation (3.25) shows that probabilities "do not add" in this way, i.e. there is destructive interference that is indicative of the "wave-like" characteristic of quantum mechanics.

More generally, suppose we consider the state

$$\mid \psi\rangle = \frac{\mid 1\rangle + e^{i\theta} \mid 2\rangle}{\sqrt{2}},$$

and we compute the probability to measure $\lambda = 6$ in this state. Following the same procedure as above, this is found to be:

$$\begin{aligned}
\text{prob}_{|\psi\rangle}(\lambda = 6) &= \langle\psi \mid P_{\lambda=6} \mid \psi\rangle, \\
&= \langle\psi \mid e_1\rangle\langle e_1 \mid \psi\rangle, \\
&= \left(\frac{\langle 1 \mid + e^{-i\theta}\langle 2 \mid}{\sqrt{2}}\right)\left(\frac{\mid 1\rangle - \mid 2\rangle}{\sqrt{2}}\right)\left(\frac{\langle 1 \mid - \langle 2 \mid}{\sqrt{2}}\right)\left(\frac{\mid 1\rangle + e^{i\theta} \mid 2\rangle}{\sqrt{2}}\right), \\
&= \frac{\left(1 - e^{-i\theta}\right)\left(1 - e^{i\theta}\right)}{4}, \\
&= \frac{1 - \cos\theta}{2}, \qquad (3.26)
\end{aligned}$$

which indicates that the *relative* phase between $\mid 1\rangle$ and $\mid 2\rangle$ is observable. However, a "global" phase is not observable, i.e. if we considered the state:

$$\mid \psi\rangle = e^{i\theta}\left(\frac{\mid 1\rangle + \mid 2\rangle}{\sqrt{2}}\right),$$

and computed the probability to measure $\lambda = 6$ in this state, we would get the same answer as (3.25), i.e. zero.

3.2 *Unitary Operators, Time Evolution*

Unitary operators play a fundamental role in quantum mechanics, as we now discuss in this section. We begin with the definition.

Definition 26 (Unitary Operator)**.** *Let U be a linear operator on a Hilbert space $\mathcal{H}$. The U is said to be a unitary operator if:*

$$U^\dagger U = UU^\dagger = \mathbb{I},$$

where $\mathbb{I}$ is the identity operator on $\mathcal{H}$.

An important property of unitary operators is that they "preserve the inner product". This means the following:

$$(U \mid \psi\rangle, U \mid \phi\rangle) = \langle\psi \mid U^\dagger U \mid \phi\rangle = \langle\psi \mid \phi\rangle.$$

The significance of unitary operators in quantum mechanics is that they "govern time evolution of the wave function" in the sense of the following calculations. Let

$$U(t) = \exp\left(-\frac{itH}{\hbar}\right), \qquad (3.27)$$

where, generally, the exponential of an operator is defined through the usual series expansion:

$$\exp A = \sum_{n=0}^{\infty} \frac{A^n}{n!}. \qquad (3.28)$$

(We leave it as an exercise to show that (3.27) is unitary.) Of course, one must prove that the exponential series defining a unitary operator converges. This is "relatively easy" in finite dimensions since all finite dimensional operators are bounded. Of course, things are more technical in the infinite dimensional case, but we will not consider these issues in this course.

Now we want to show that the time evolution of state vectors is governed by (3.27). Consider

$$| \psi(t)\rangle = U(t) \mid \psi_0\rangle, \quad \text{for some constant} \quad | \psi_0\rangle \in \mathcal{H}. \tag{3.29}$$

Now we show that $| \psi(t)\rangle$ satisfies the Schrödinger equation:

$$i\hbar \frac{d}{dt} \mid \psi(t)\rangle = H \mid \psi(t)\rangle. \tag{3.30}$$

First, note that:

$$\begin{aligned}
\frac{d}{dt}U(t) &= \frac{d}{dt} \exp\left(-\frac{iHt}{\hbar}\right), \\
&= \frac{d}{dt} \sum_{n=0}^{\infty} \left(\frac{-it}{\hbar}\right)^n \frac{H^n}{n!}, \\
&= \sum_{n=1}^{\infty} \left(\frac{-i}{\hbar}H\right) \left(\frac{-it}{\hbar}\right)^{n-1} n \frac{H^{n-1}}{n!}, \\
&= \left(\frac{-iH}{\hbar}\right) \sum_{n=1}^{\infty} \left(\frac{-itH}{\hbar}\right)^{n-1} \frac{1}{(n-1)!}, \\
&= \left(-\frac{iH}{\hbar}\right) \exp\left(-\frac{iHt}{\hbar}\right), \\
&= \left(-\frac{iH}{\hbar}\right) U(t).
\end{aligned} \tag{3.31}$$

Then we have:

$$\begin{aligned}
i\hbar \frac{d}{dt} \mid \psi(t)\rangle &= i\hbar \left(\frac{d}{dt}U(t)\right) \mid \psi_0\rangle + i\hbar U(t) \frac{d}{dt} \mid \psi_0\rangle, \\
&= i\hbar \frac{-iH}{\hbar} U(t) \mid \psi_0\rangle \\
&= HU(t) \mid \psi_0\rangle, \\
&= H \mid \psi(t)\rangle,
\end{aligned} \tag{3.32}$$

which shows that an initial condition $| \psi_0\rangle \in \mathcal{H}$ for the Schrödinger equation evolves in time by application of the unitary operator, $| \psi(t)\rangle = U(t) \mid \psi_0\rangle = \exp\left(-\frac{itH}{\hbar}\right) \mid \psi_0\rangle$.

Finally, note that if $\mid e_i\rangle$ is an eigenstate of H with eigenvalue E_i, then we have:

$$U(t) \mid e_i\rangle = \exp\left(-\frac{itH}{\hbar}\right) \mid e_i\rangle = \exp\left(-\frac{itE_i}{\hbar}\right) \mid e_i\rangle. \tag{3.33}$$

(This really requires a proof, using the exponential series expression for a unitary operator, that you should be able to provide.) So if we express $\mid \psi_0\rangle$ in a basis of eigenstates of H:

$$\mid \psi_0\rangle = \sum_i a_i \mid e_i\rangle, \tag{3.34}$$

Then we have:

$$\mid \psi(t)\rangle = \exp\left(-\frac{itH}{\hbar}\right) \sum_i a_i \mid e_i\rangle = \sum_i a_i \exp\left(-\frac{itE_i}{\hbar}\right) \mid e_i\rangle. \tag{3.35}$$

Commutation Relations

In general, two linear operators, A and B, do not *commute*, i.e.

$$AB \neq BA.$$

Recall that in quantum mechanics measurable quantities correspond to self adjoint operators. Related to this notion, whether or not two self adjoint linear operators commute has important physical consequences, as we will see. First, we consider an example that is particularly relevant to quantum mechanics.

Example. Recall the one dimensional momentum and position operators:

$$P = \frac{\hbar}{i}\frac{d}{dx}, \quad X.$$

Then we have:

$$\begin{aligned}(PX\psi)(x) &= \frac{\hbar}{i}\frac{d}{dx}(x\psi(x)),\\ &= \frac{\hbar}{i}\psi(x) + \frac{\hbar}{i}x\frac{d}{dx}\psi(x),\\ &= \frac{\hbar}{i}\psi(x) + (XP\psi)(x),\end{aligned} \tag{3.36}$$

or

$$PX = \frac{\hbar}{i} + XP. \tag{3.37}$$

To be more precise, we have the following definition.

Definition 27 (Commutator)**.** *The commutator of two linear operators A and B is defined as:*

$$[A, B] = AB - BA. \tag{3.38}$$

With this notation (3.37) becomes:

$$[X, P] = i\hbar. \tag{3.39}$$

Using the same idea, it is not hard to show that the following commutation relations hold in three dimensions.

$$[P_i, P_j] = 0, \quad [X_i, X_j] = 0, \quad [X_i, P_j] = i\hbar\delta_{ij}. \tag{3.40}$$

Commutators of linear operators satisfy a number of useful identities that we list below.

(i) $[A, B] = -[B, A]$

(ii) $[A, B]$ is linear in both A and B.

(iii) $[A, BC] = B[A, C] + [A, B]C$.

(iv) $[A, [B, C]] + [B, [C, A]] + [C, [A, B]] = 0$

The proofs of these identities will be left as exercises.

Simultaneous Measurability

Suppose A and B are self adjoint operators representing "observables", i.e. quantities that can be measured. Recall that the outcome of a measurement of A (or B) can only be known with certainty if it is in an eigenstate. We are concerned here with the possibility of measuring A and B in the same basis. The following result is fundamental, and it highlights the significance of the commutator.

Theorem 9. *Suppose A and B are self adjoint operators on a finite dimensional, complex inner product space. Then A and B have a joint orthonormal basis of eigenvectors if and only if $[A, B] = 0$.*

Proof. First, let is suppose that

$$| n\rangle,\ n = 1, 2, 3, \ldots,$$

is an orthonormal basis of eigenvectors *for both* A and B, i.e.,

$$A \mid n\rangle = a_n \mid n\rangle, \tag{3.41}$$
$$B \mid n\rangle = b_n \mid n\rangle. \tag{3.42}$$

A general state can be represented as follows:

$$| \psi \rangle = \sum_n c_n \mid n \rangle.$$

Then we have:

$$\begin{aligned}
AB \mid \psi \rangle &= \sum_n ABc_n \mid n \rangle \\
&= \sum_n c_n AB \mid n \rangle, \\
&= \sum_n c_n Ab_n \mid n \rangle, \\
&= \sum_n c_n b_n A \mid n \rangle, \\
&= \sum_n c_n b_n a_n \mid n \rangle, \\
&= \sum_n c_n a_n b_n \mid n \rangle, \\
&= \sum_n c_n a_n B \mid n \rangle, \\
&= \sum_n c_n B a_n \mid n \rangle, \\
&= \sum_n c_n BA \mid n \rangle, \\
&= \sum_n BAc_n \mid n \rangle, \\
&= BA \mid \psi \rangle. \qquad (3.43)
\end{aligned}$$

Since the state $| \psi \rangle$ was completely arbitrary, this implies that

$$AB = BA.$$

Now let us suppose that $AB = BA$ and that $| n \rangle$, $n = 1, 2, 3, \ldots$ is an orthonormal basis for A (i.e. $A \mid n \rangle = a_n \mid n \rangle$). We will show that this implies that $| n \rangle$, $n = 1, 2, 3, \ldots$ is an orthonormal basis for B.

Now since $AB = BA$ we have:

$$AB \mid n \rangle = BA \mid n \rangle = a_n B \mid n \rangle. \qquad (3.44)$$

This calculation implies that $B \mid n \rangle$ is also an eigenvector of A with eigenvalue a_n. There are two cases to consider.

$\boxed{a_n \text{ is non-degenerate.}}$ In this case there is only one eigenvector corresponding to the eigenvalue a_n. Therefore since $B \mid n \rangle$ is also an eigenvector of A, it must be proportional to $| n \rangle$, i.e.

$$B \mid n \rangle = b_n \mid n \rangle.$$

In other words, $| n \rangle$ is also an eigenvector of B. Hence, $| n \rangle$, $n = 1, 2, 3, \ldots$ are also eigenvectors for B.

Now we consider the other case.

a_n is degenerate.

Suppose $\mid n_j\rangle$, $j = 1, \ldots, d$ are all the orthonormal eigenvectors of A corresponding to the eigenvalue a_n. We know from (3.44) that $B \mid n\rangle$ is an eigenvector of A, with eigenvalue a_n, but it is not necessarily an eigenvector of B since it need not be proportional to $\mid n\rangle$

Now let

$$W = \text{span } \{\mid n_j\rangle, j = 1, \ldots, d\}.$$

You should be able to verify to yourself that W is an invariant subspace for A, i.e. $A : W \to W$. Now you need to go a step further and, using (3.44) and argue that W is an invariant subspace for B, i.e. $B : W \to W$. In this case, it makes sense to restrict B to W. Now we want to argue that B restricted to W is self-adjoint. This can be seen as follows. Express B as follows:

$$B \mid n_j\rangle = \sum_{i=1}^{d} \mid n_i\rangle\langle n_i \mid B \mid n_j\rangle.$$

Then the matrix elements of B satisfy the following:

$$B_{ij} = \langle n_i \mid B \mid n_j\rangle = \overline{\langle n_j \mid B^\dagger \mid n_i\rangle} = \overline{\langle n_j \mid B \mid n_i\rangle} = \overline{B_{ji}}.$$

Hence, B, restricted to W, is self adjoint. Therefore, $B\Big|_W$ can be diagonalised, i.e. we can find d orthonormal eigenvectors, $\mid \tilde{n}_j\rangle$, $j = 1, \ldots, d$, and $B \mid \tilde{n}_j\rangle = b_{\tilde{n}_j} \mid \tilde{n}_j\rangle$. Now it follows from (3.44) that

$$A \mid \tilde{n}_j\rangle = a_{\tilde{n}_j} \mid \tilde{n}_j\rangle.$$

In this way, we are able to find a set of orthonormal eigenvectors for both A and B , restricted to W. This procedure can be repeated for every degenerate eigenvalue, if necessary, to construct a set of orthonormal eigenvectors for A and B on the entire space.

□

(Mathematically, this proof is a bit "sloppy". See if you can figure out where it should be "cleaned up".)

This is an important result since it often turns out that in order to specify the state of a system, we need to measure more than one quantity, i.e. we require more than one observable. We will see a concrete example of this when we study angular momentum. For the observables to be "compatible" we need to be able to express them in a common basis, i.e. the observables must commute.

Uncertainty Relations

In this section we will derive the general form of the uncertainty principle due to H. P. Robertson for self-adjoint operators [1]. As a preamble to a discussion of the uncertainty principle in quantum mechanics it is difficult to do much better than the description that Robertson gave in his original paper, from which we quote:

> The uncertainty principle is one of the most characteristic and important consequences of the new quantum mechanics. This principle, as formulated by Heisenberg for two conjugate quantum-mechanical variables, states that the accuracy with which two such variables can be measured simultaneously is subject to the restriction that the product of the uncertainties in the two measurements is at least of order h (Planck's constant).

[1] H. P. Robertson. The uncertainty principle. *Physical Review*, 34(1):163, 1929

Prior to Robertson's work, the uncertainty principle began to be established in quantum theory through the work of Heisenberg [2], Kennard [3], and Weyl [4]. Heisenberg described [5] a very insightful and intriguing thought experiment (Heisenberg's microscope) that is well-worth considering in order to grasp the deeper physical aspects of the uncertainly principle. A description of the uncertainty principle from a contemporary point of view is given in the paper of Furuta[6], which also contains a description of its historical development.

[2] W. Heisenberg. Über den anschaulichen inhalt der quantentheoretischen kinematik und mechanik. In *Original Scientific Papers Wissenschaftliche Originalarbeiten*, pages 478–504. Springer, 1985

[3] E. H. Kennard. Zur quantenmechanik einfacher bewegungstypen. *Zeitschrift für Physik*, 44(4-5):326–352, 1927

[4] H. Weyl. Gruppentheorie und quantenmechanik, hirzel, leipzig. *Theory of Groups and Quantum Mechanics, 2nd ed.(1931), transl. H. P. Robertson, Dover, NY (1950)*, pages 100–101, 1928

[5] W. Heisenberg. *The physical principles of the quantum theory*. Courier Corporation, 1949

[6] A. Furuta. One thing is certain: Heisenberg's uncertainty principle is not dead. *Scientific American*, 2012

We now give the proof of Robertson's mathematical formulation of the uncertainty principle. Recall that the dispersion of A in the state $\mid \psi\rangle$ is defined as:

$$\Delta_\psi(A) = \left[E_\psi(A^2) - E_\psi(A)^2\right]^{\frac{1}{2}}, \tag{3.45}$$

where

$$E_\psi(A) = \langle \psi \mid A \mid \psi\rangle, \qquad \langle \psi \mid \psi\rangle = 1.$$

We now derive the following "generalised" uncertainty relation.

Theorem 10. *Let A and B be self-adjoint operators, then*

$$\Delta_\psi(A)\Delta_\psi(B) \geq \frac{1}{2} \mid E_\psi([A,B]) \mid . \tag{3.46}$$

We have equality if

$$\left\{(A - E_\psi(A)) - it(B - E_\psi(B))\right\} \mid \psi\rangle = 0. \tag{3.47}$$

for some $t \in \mathbb{R}$.

Proof. Define

$$[A,B] = iC.$$

Then it is a simple exercise to show that

$$C^\dagger = C.$$

We will use this result in the course of the proof.

Now consider

$$(A - itB), \quad t \in \mathbb{R}.$$

Then

$(A - itB)^\dagger = A + itB$, (note that we used the fact that t is real here.)

and therefore:

$$\begin{aligned} (A - itB)^\dagger (A - itB) &= (A + itB)(A - itB), \\ &= A^2 - it(AB - BA) + t^2 B^2, \\ &= A^2 + tC + t^2 B^2. \end{aligned} \tag{3.48}$$

Using this result, it follows that:

$$\begin{aligned} E_\psi(A^2) + tE_\psi(C) + t^2 E_\psi(B^2) &= \langle \psi \mid (A - itB)^\dagger (A - itB) \mid \psi \rangle, \\ &= \| (A - itB) \mid \psi \rangle \|^2, \\ &\geq 0, \end{aligned} \tag{3.49}$$

and this expression is zero if and only if:

$$(A - itB) \mid \psi \rangle = 0. \tag{3.50}$$

Now we choose t such that the left hand side of (3.49) is minimal. The choice of t that accomplishes this is obtained by differentiating the left hand side of (3.49) with respect to t and setting the result to zero:

$$\begin{aligned} &\frac{d}{dt}\left(E_\psi(A^2) + tE_\psi(C) + t^2 E_\psi(B^2)\right) = E_\psi(C) + 2tE_\psi(B^2) = 0, \\ &\Rightarrow \quad t = -\frac{E_\psi(C)}{2E_\psi(B^2)}. \end{aligned}$$

Substituting this value of t into the left hand side of (3.49) gives:

$$E_\psi(A^2) - \frac{E_\psi(C)^2}{2E_\psi(B^2)} + \frac{E_\psi(C)^2}{4E_\psi(B^2)} \geq 0,$$

which simplifies to

$$E_\psi(A^2) - \frac{1}{4}\frac{E_\psi(C)^2}{E_\psi(B^2)} \geq 0,$$

or

$$E_\psi(A^2)E_\psi(B^2) \geq \frac{1}{4}E_\psi(C)^2. \tag{3.51}$$

Now (3.51) is valid for any self-adjoint operators, A and B. We therefore make the following substitution into (3.51):

$$A \to A - E_\psi(A), \quad B \to B - E_\psi(B). \tag{3.52}$$

It then follows from straightforward calculations that:

$$E_\psi\left(A - E_\psi(A)\right) = E_\psi(A^2) - E_\psi(A)^2 = \Delta_\psi(A)^2, \tag{3.53}$$

$$E_\psi\left(B - E_\psi(B)\right) = E_\psi(B^2) - E_\psi(B)^2 = \Delta_\psi(B)^2, \tag{3.54}$$

and

$$[A - E_\psi(A),\, B - E_\psi(B)] = [A, B] = iC. \tag{3.55}$$

Substituting (3.53) and (3.54) into the left side of (3.51), and applying (3.55) to the right side of (3.51) gives:

$$\Delta_\psi(A)^2\,\Delta_\psi(B)^2 \geq \frac{1}{4}E_\psi\left(\frac{1}{i}[A, B]\right)^2. \tag{3.56}$$

Then taking the square root of the result gives:

$$\Delta_\psi(A)\,\Delta_\psi(B) \geq \left|\frac{1}{2i}E_\psi([A, B])\right| = \frac{1}{2}\left|E_\psi([A, B])\right|. \tag{3.57}$$

Finally, note that (3.47) is obtained by substituting (3.52) into (3.50).

□

It is important to realize that, strictly speaking, this is purely a mathematical result describing statistical properties of self-adjoint operators. The physics is introduced when the self-adjoint operators correspond to physical observables, and we next consider examples of this

Generalized Uncertainty Relation for the Position and Momentum Operators. We now compute the generalised uncertainty relation for the position and momentum operators;

$$\Delta_\psi(P)\Delta_\psi(X) \geq \frac{1}{2}\left|E_\psi[P, X]\right|.$$

Recall that:

$$[P, X] = -i\hbar,$$

and therefore

$$|E_\psi([P, X])| = \hbar.$$

So we have:

$$\Delta_\psi(P)\Delta_\psi(X) \geq \frac{\hbar}{2}.$$

Recall that, in general, the dispersion of an observable depends upon the state in which the measurement is made. However, the right hand side of this relation is independent of the state. So this uncertainty relation is true *for any* (normalised) state $|\psi\rangle$ in which the measurements of X and P are made.

A State of Minimal Uncertainty. Recall that we have equality of the generalised uncertainty relation (3.46) provided (3.47) is satisfied. In other words, substituting (3.52) into (3.49) gives the following criterion for minimizing the uncertainty:

$$\{A - E(A) - it(B - E(B))\} \mid \psi\rangle = 0. \tag{3.58}$$

Now we consider a specific example. We consider the (one dimensional) position and momentum operators, and we want to find the state, $|\psi\rangle$ for which the uncertainty is minimal. In particular, we have:

$$A = P = \frac{\hbar}{i}\frac{d}{dx},\ B = X,\ \tilde{p} \equiv E_\psi(P),\ \tilde{x} \equiv E_\psi(X).$$

With these substitutions (3.58) becomes:

$$\left(\frac{\hbar}{i}\frac{d}{dx} - \tilde{p}\right)\psi = it(x - \tilde{x})\psi \tag{3.59}$$

or

$$\frac{d}{dx}\psi = \frac{i}{\hbar}\left(\tilde{p} + it(x - \tilde{x}\right)\psi. \tag{3.60}$$

This equation can be solved to give:

$$\psi(x) = c\ \exp\left(\frac{i}{\hbar}\tilde{p}x - \frac{t}{2\hbar}(x - \tilde{x})^2\right), \tag{3.61}$$

and the constant c can be chosen to satisfy normalisation. You have seen this particular function before (and you will continue to see it in your mathematics studies). It is referred to as a "Gaussian", and wave functions of this form play an important role in quantum mechanics. Recall that we first introduced the Gaussian function in Section 2.3 in our study of the free particle.

The Harmonic Oscillator

The (unforced, undamped) harmonic oscillator is a basic, "paradigm" dynamical system that is essential to understand.

The Hamitonian for the classical harmonic oscillator is given by:

$$H = \frac{1}{2m}p^2 + \frac{1}{2}m\omega^2 x^2, \tag{3.62}$$

where m denotes the mass of the (point) particle, ω is the frequency of oscillation, and p and x are the coordinates corresponding the the momentum and position, respectively. Hamilton's differential equations are given by:

$$\begin{aligned} \dot{x} &= \frac{\partial H}{\partial p} = \frac{p}{m}, \\ \dot{p} &= -\frac{\partial H}{\partial x} = -m\omega^2 x, \end{aligned} \tag{3.63}$$

and "solving" the classical harmonic oscillator might mean finding the solutions to Hamilton's differential equations.

We now turn our attention to the quantum mechanical harmonic oscillator. The quantum mechanical Hamiltonian is obtain from the classical Hamiltonian by replacing p and x with the corresponding operators:

$$H = \frac{1}{2m}P^2 + \frac{1}{2}m\omega^2 X^2. \tag{3.64}$$

For us, "solving" the quantum mechanical oscillator equation will mean finding the eigenvalues and eigenstates of the time-independent Schrdingier equation:

$$H \mid \psi\rangle = E \mid \psi\rangle, \tag{3.65}$$

where E is the eigenvalue and H is given by (3.64). It will be useful the recall the commutation relation between X and P, which we re-write here:

$$[X, P] = i\hbar. \tag{3.66}$$

We will introduce a method that is algebraic in nature, and heavily uses the commutation relation (3.66) (we will revisit the same method when we study quantum mechanical angular momentum). It goes by the names of "raising and lowering operators". "creation and annihilation operators", "ladder operators" or the method of factorization. The different names are used in the different areas in which it is used. It is surprising that it is not taught as a general mathematical approach in ordinary differential equations and partial differential

equations courses. More details on the methoc can be found in this reference[7].

[7] B. Mielnik and O. Rosas-Ortiz. Factorization: little or great algorithm? *Journal of Physics A: Mathematical and General*, 37 (43):10007, 2004

We begin by introducing the operators:

$$\begin{aligned} a &= P - im\omega X, \\ a^\dagger &= P + im\omega X. \end{aligned} \tag{3.67}$$

It is straightforward to verify that they satisfy the following commutation relation:

$$\begin{aligned} [a, a^\dagger] &= [P - im\omega X, P + im\omega X], \\ &= im\omega[P, X] - im\omega[X, P], \\ &= 2m\omega\hbar. \end{aligned} \tag{3.68}$$

Moreover, it is easy to verify that:

$$\begin{aligned} a^\dagger a &= (P + im\omega X)(P - im\omega X), \\ &= P^2 + im\omega(XP - PX) + m^2\omega^2 X^2, \\ &= 2m\left(H - \frac{\omega\hbar}{2}\right), \end{aligned} \tag{3.69}$$

and

$$\begin{aligned} aa^\dagger &= [a, a^\dagger] + a^\dagger a, \\ &= 2m\left(H + \frac{\omega\hbar}{2}\right). \end{aligned} \tag{3.70}$$

We define the *number operator* as follows:

$$N = \frac{1}{2m\hbar\omega} a^\dagger a. \tag{3.71}$$

Note that using (3.69) and (3.71), the Hamiltonian, (3.64), can be expressed in terms of the number operator as follows:

$$H = \hbar\omega\left(N + \frac{1}{2}\right). \tag{3.72}$$

Below we collect together some useful properties of the number operator.

Properties of the number operator.

- N is self adjoint. This follows from the calculation:

$$N^\dagger = \frac{a^\dagger a}{2m\hbar\omega} = N. \tag{3.73}$$

- N is positive. This follows from the calculation:

$$\langle \psi \mid N \mid \psi \rangle = \langle \psi \mid \frac{a^\dagger a}{2m\hbar\omega} \mid \psi \rangle = \frac{1}{2m\hbar\omega} \left(a \mid \psi \rangle, a \mid \psi \rangle \right) \geq 0. \quad (3.74)$$

- Commutators:

$$[N, a] = \frac{1}{2m\hbar\omega}[a^\dagger a, a] = \frac{1}{2m\hbar\omega} \left\{ a^\dagger [a, a] + [a^\dagger, a] a \right\} = -a, \quad (3.75)$$

$$[N, a^\dagger] = \frac{1}{2m\hbar\omega}[a^\dagger a, a^\dagger] = \frac{1}{2m\hbar\omega} \left\{ a^\dagger [a, a^\dagger] + [a^\dagger, a^\dagger] a \right\} = a^\dagger. \quad (3.76)$$

Now consider the general equation for the eigenvalues and eigenvectors of N: [8]

$$N \mid \eta \rangle = \eta \mid \eta \rangle. \quad (3.77)$$

[8] Note that when we write down such general equations, we are *assuming* that $\mid \eta \rangle \neq 0$, since zero eigenvectors are not interesting. If we give an argument for the existence of a general eigenvector, we then need to prove afterward (if not done during the course of the proof) that the eigenvector is not zero.

Our goal is to determine the eigenvalues η and the eigenstates $\mid \eta \rangle$ of N by using the following properties:

$$\langle \eta \mid N \mid \eta \rangle \geq 0, \qquad \eta = \frac{\langle \eta \mid N \mid \eta \rangle}{\langle \eta \mid \eta \rangle} \geq 0. \quad (3.78)$$

It follows from (3.72) that the eigenvalues and eigenstates of the number operator are also eigenvalues and eigenstates of the Hamiltonian of the harmonic oscillator.

$a^\dagger \mid \eta \rangle$ is an eigenstate of N with eigenvalue $\eta + 1$.

Using (3.76) we have:

$$\begin{aligned} N a^\dagger \mid \eta \rangle &= \left(a^\dagger N + a^\dagger \right) \mid \eta \rangle, \\ &= a^\dagger N \mid \eta \rangle + a^\dagger \mid \eta \rangle, \\ &= (\eta + 1) a^\dagger \mid \eta \rangle, \end{aligned} \quad (3.79)$$

which implies that $a^\dagger \mid \eta \rangle$ is an eigenstate of N with eigenvalue $\eta + 1$, *provided* $a^\dagger \mid \eta \rangle \neq 0$, which we show is the case.

Technical Point: $a^\dagger \mid \eta \rangle \neq 0$**.**

$$\begin{aligned} \left(a^\dagger \mid \eta \rangle, a^\dagger \mid \eta \rangle \right) &= \langle \eta \mid a a^\dagger \mid \eta \rangle = \langle \eta \mid a^\dagger a + 2m\hbar\omega \mid \eta \rangle, \\ &= \langle \eta \mid (2m\hbar\omega) N + 2m\hbar\omega \mid \eta \rangle, \\ &= 2m\hbar\omega \langle \eta \mid N \mid \eta \rangle + 2m\hbar\omega \langle \eta \mid \eta \rangle, \\ &= 2m\hbar\omega \eta \langle \eta \mid \eta \rangle + 2m\hbar\omega \langle \eta \mid \eta \rangle, \\ &= 2m\hbar\omega (\eta + 1) \langle \eta \mid \eta \rangle, \\ &> 0, \end{aligned} \quad (3.80)$$

from which it follows that:

$$a^\dagger \mid \eta\rangle \neq 0, \tag{3.81}$$

$a \mid \eta\rangle$ is an eigenstate of N with eigenvalue $\eta - 1$.

Using (3.75), we have:

$$\begin{aligned} Na \mid \eta\rangle &= (aN - a) \mid \eta\rangle, \\ &= (\eta - 1)a \mid \eta\rangle, \end{aligned} \tag{3.82}$$

which implies that $a \mid \eta\rangle$ is an eigenstate of N with eigenvalue $\eta - 1$, *provided* $a \mid \eta\rangle \neq 0$, which we show is the case with the following calculation:

Technical Point: $a \mid \eta\rangle \neq 0$**.**

$$(a \mid \eta\rangle, a \mid \eta\rangle) = \langle \eta \mid a^\dagger a \mid \eta\rangle = 2m\hbar\omega\langle \eta \mid N \mid \eta\rangle = 2m\hbar\omega\eta\langle \eta \mid \eta\rangle \geq 0, \tag{3.83}$$

from which it follows that $a \mid \eta\rangle = 0$ if and only if $\eta = 0$ (since we are assuming that $\mid \eta\rangle \neq 0$).

From these calculations we see that it is natural to refer to $a^\dagger$ as the *raising operator* and a as the *lowering operator*.

N has a smallest eigenvalue, which is zero.

Since $N \geq 0$, it has a smallest eigenvalue, η_{min}. We now argue that $\eta_{min} = 0$. Suppose

$$a \mid \eta_{min}\rangle \neq 0. \tag{3.84}$$

Then we have:

$$Na \mid \eta_{min}\rangle = (\eta_{min} - 1)\, a \mid \eta_{min}\rangle. \tag{3.85}$$

This implies that $a \mid \eta_{min}\rangle$ is an eigenstate of N with eigenvalue $\eta_{min} - 1$, which contradicts the fact that η_{min} is the minimum eigenvalue of N. This contradiction means that the assumption (3.84) is false. Hence we have:

$$a \mid \eta_{min}\rangle = 0. \tag{3.86}$$

From this calculation, and the conclusion of (3.83), it follows that:

$$\eta_{min} = 0. \tag{3.87}$$

We summarise our results.

- N has eigenvalues $0, 1, 2, 3, \ldots$.
- It follows from (3.79) that $a^\dagger \mid n\rangle$ is proportional to $\mid n+1\rangle$, and the proportionality constant is chosen so that the eigenstates are normalised
- It follows from (3.83) that $a \mid n\rangle$ is proportional to $\mid n-1\rangle$, and the proportionality constant is chosen so that the eigenstates are normalised.

The eigenstate with smallest eigenvalue, $\mid 0\rangle$ is referred to as the *ground state* or *vacuum state*, and we normalise this eigenstate such that:

$$\langle 0 \mid 0\rangle = 1. \tag{3.88}$$

All other eigenstates can be obtained from $\mid 0\rangle$ by acting on it with the raising operator, $a^\dagger$:

$$\mid n\rangle = c_n (a^\dagger)^n \mid 0\rangle. \tag{3.89}$$

It follows from a calculation that requiring

$$\langle n \mid n\rangle = 1,$$

implies that

$$c_n^2 = \frac{1}{\langle 0 \mid a^n (a^\dagger)^n \mid 0\rangle}, \tag{3.90}$$

or

$$c_n = \frac{1}{\sqrt{n!}\,(2m\hbar\omega)^{\frac{n}{2}}}. \tag{3.91}$$

The calculation that leads to this normalisation factor are described in Problem 5 at the end of this chapter..

Now recall the form of the Hamiltonian for the harmonic oscillator expressed in terms of the number operator given in (3.72). Hence, we have shown that:

$$H|n\rangle = \hbar\omega \left(n + \frac{1}{2}\right) \mid n\rangle \qquad n = 0, 1, 2, \ldots \tag{3.92}$$

In other words, we are able to determine the eigenvalues of H from the properties of the raising and lowering operators (in particular, from the number operator). Now we turn our attention to finding the eigenstates explicitly.

First, we determine $\mid 0\rangle$. From (3.86) we have:

$$a \mid 0\rangle = 0. \tag{3.93}$$

From the form of a given in (3.67), (3.93) has the form:

$$(P - im\omega X) \mid 0\rangle = 0, \tag{3.94}$$

or, substituting into this equation the explicit expressions for the position and momentum operators:

$$\frac{\hbar}{i}\psi_0'(x) - im\omega x\psi_0(x) = 0, \tag{3.95}$$

or

$$\psi_0'(x) = -\frac{m\omega}{\hbar}x\psi_0(x). \tag{3.96}$$

This ordinary differential equation can be solved explicitly to give:

$$\psi_0(x) = c\exp\left(-\frac{m\omega}{2\hbar}x^2\right). \tag{3.97}$$

The constant c is evaluated through the normalisation condition:

$$1 = \int_{-\infty}^{+\infty} \mid \psi_0(x) \mid^2 dx = \int_{-\infty}^{+\infty} c^2\exp\left(-\frac{m\omega}{\hbar}x^2\right)dx = c^2\sqrt{\frac{\pi\hbar}{m\omega}}, \tag{3.98}$$

which gives:

$$c = \left(\frac{m\omega}{\pi\hbar}\right)^{\frac{1}{4}}. \tag{3.99}$$

With $\psi_0(x)$ determined, the remaining eigenstates can be computed by acting on $\psi_0(x)$ with the raising operator:

$$\mid n\rangle = c_n\left(a^\dagger\right)^n \mid 0\rangle, \quad c_n = \frac{1}{\sqrt{n!}(2m\hbar\omega)^{\frac{m}{2}}}. \tag{3.100}$$

or, more explicitly,

$$\psi_n(x) = c_n\left(\frac{\hbar}{i}\frac{d}{dx} + im\omega x\right)^n \psi_0(x). \tag{3.101}$$

Problems

1. *General properties of commutators.* Check, using the definition of the commutator, that

   (a) $[A, B] = -[B, A]$,

   (b) $[A, B]$ is linear in both A and B,

   (c) $[A, BC] = B[A, C] + [A, B]C$,

   (d) $[A, [B, C]] + [B, [C, A]] + [C, [A, B]] = 0$,

   for operators A, B and C.

2. *Canonical commutation relations.* *(a)* Prove by induction that

$$[X^n, P] = i\hbar n X^{n-1}$$

   and that

$$[X, P^n] = i\hbar n P^{n-1}$$

   where n is a positive integer.

   (b) Using these results show that if $f(x)$ can be expanded in a polynomial in x and $g(p)$ can be expanded in a polynomial in p, then

$$[f(X), P] = i\hbar f'(X)$$

   and

$$[X, g(P)] = i\hbar g'(P)$$

   This result is derived for any differentiable f and g in Hannabuss, problem 7.4.

   (c) More generally, prove by induction, that if two operators, A and B, commute with their commutator (i.e. $[A, [A, B]] = [B, [A, B]] = 0$) then

$$[A, B^n] = nB^{n-1}[A, B] \quad \text{and} \quad [B, A^n] = -nA^{n-1}[A, B]$$

3. *Commutators again.* Consider two operators N and A which satisfy $[N, A] = -A$. And define the operator $C(\theta)$ depending on a real parameter θ by

$$C(\theta) = e^{\theta N} A e^{-\theta N}$$

(a) Using the definition of the exponential of an operator, show that $\frac{d}{d\theta}(e^{\theta N}) = Ne^{\theta N}$.

(b) Show that $\frac{dC(\theta)}{d\theta} = -C(\theta)$.

(c) Assuming that the differential equation in part (b) has a unique solution given $C(0)$, show that $C(\theta) = e^{-\theta}A$.

4. *Creation and annihilation operators, 1.* Let X and P be the standard position and momentum observables, satisfying $[X, P] = i\hbar$ ($\hbar$ is Planck's constant). The creation and annihilation operators $a^\dagger$ and a are defined by

$$a^\dagger = P + im\omega X; \quad a = P - im\omega X$$

where m and ω are constants.

(a) Show that $[a, a^\dagger] = 2m\omega\hbar$.

(b) Show that

$$a(a^\dagger)^n = (a^\dagger)^n a + n2m\omega\hbar(a^\dagger)^{(n-1)},$$

where n is a non-negative integer.

(c) Show also that

$$ae^{\lambda a^\dagger} = e^{\lambda a^\dagger}(a + \lambda 2m\omega\hbar)$$

where λ is a complex constant and $e^{\lambda a^\dagger}$ is defined by

$$e^{\lambda a^\dagger} = \sum_{n=0}^{\infty} \lambda^n (a^\dagger)^n / n!.$$

5. *Creation and annihilation operators, 2.* Let a and $a^\dagger$ be the operators defined in the previous question, and let $|0\rangle$ be the normalised vacuum state defined by

$$a\,|0\rangle = 0.$$

The normalised state $|n\rangle$ is defined by

$$|n\rangle = c_n (a^\dagger)^n\,|0\rangle\,,$$

where c_n is a positive constant, and n is a non-negative integer.

(a) Show that $2m\omega\hbar n c_n^2 = c_{n-1}^2$, and hence, or otherwise, calculate c_n.

(b) Using part (b) of the previous question, show that $a\,|n\rangle = d_n\,|n-1\rangle$, where d_n is a constant, and calculate d_n.

(c) Show that $\langle n_1\,|n_2\rangle = 0$ if $n_1 \neq n_2$.

(d) By writing X and P in terms of a and $a^\dagger$, calculate $\langle n_1|\,X\,|n_2\rangle$ and $\langle n_1|\,P\,|n_2\rangle$.

6. Prove that the dispersion of an operator A in the state $|\psi\rangle$ vanishes *if and only if* $|\psi\rangle$ is an eigenvector of A.

7. On a space spanned by the orthonormal basis $|1\rangle, |2\rangle, |3\rangle$, let G be the operator defined by

$$\begin{aligned} G\,|1\rangle &= \frac{5}{4}\,|1\rangle + \frac{\sqrt{3}}{4}\,|2\rangle \\ G\,|2\rangle &= \frac{\sqrt{3}}{4}\,|1\rangle + \frac{7}{4}\,|2\rangle \\ G\,|3\rangle &= 2\,|3\rangle \end{aligned}$$

(a) Find the eigenvalues of G and eigenvectors of G.

(b) Write G in a spectral decomposition $G = \lambda_1 P_1 + \lambda_2 P_2$, where one of the projection operators is two-dimensional; you should give P_1 and P_2 explicitly in Dirac notation.

(c) Let a particle initially be in the state $|\phi\rangle = \cos\theta\,|1\rangle + \sin\theta\,|3\rangle$, where θ is a constant. Find the probability that the outcome, +2, is found when G is measured, and find the state immediately after measurement.

(d) Find the expected value of G in the state $|\phi\rangle$ and the dispersion of G in this state.

(e) Is there a value of θ for which the dispersion is zero? Comment on this.

8. Let H denote a self-adjoint Hamiltonian operator and $|\,\psi\rangle$ an eigenstate of H, i.e.

$$H\,|\,\psi\rangle = E\,|\,\psi\rangle.$$

For any other operator A, $\langle A\rangle \equiv \langle\psi\,|\,A\,|\,\psi\rangle$ denotes the expectation value of A in the state $|\,\psi\rangle$.

(a) For any operator A show that

$$\langle \psi \mid [H, A] \mid \psi \rangle = 0. \tag{3.102}$$

(b) If X and P denote the position and momentum operators, respectively, show that:

(i) $[P^2, X] = -2i\hbar P$,

(ii) $[P^2, XP] = -2i\hbar P^2$,

(iii) $[X^N, XP] = Ni\hbar X^N$, where N is an integer, $N \geq 1$.

Now suppose we consider a self-adjoint Hamiltonian operator of the form:

$$H = T + V = \frac{P^2}{2m} + kX^N, \tag{3.103}$$

where X and P denote the position and momentum operators, respectively, and N is an integer, $N \geq 1$, $m > 0$, and k is a real number.

(c) Using (3.102) and $A = X$, show that $\langle P \rangle = 0$.

(d) Using (3.102) and $A = XP$, show that $2\langle T \rangle = N\langle V \rangle$.

(e) Compute the variance, $\langle (\Delta P)^2 \rangle = \langle P^2 \rangle - \langle P \rangle^2$, of the momentum operator P in the state $\mid \psi \rangle$ and express it in terms of the expectation value of the kinetic energy, T, in the state $\mid \psi \rangle$.

9. Consider the state space $\mathbb{C}^2$ with orthonormal basis,

$$\mid 1\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad \mid 2\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

We define the operator A on $\mathbb{C}^2$ with respect to this basis as follows:

$$\begin{aligned} A \mid 1\rangle &= 2 \mid 1\rangle - i \mid 2\rangle, \\ A \mid 2\rangle &= i \mid 1\rangle + 2 \mid 2\rangle. \end{aligned}$$

(a) Write down the matrix representations of A with respect to the basis $\mid 1\rangle, \mid 2\rangle$.

(b) Show that A is self-adjoint.

(c) Show that the eigenvalues of A are 1 and 3. Let $\mid e_1\rangle$ denote the normalized eigenstate corresponding to the eigenvalue 1 and let $\mid e_2\rangle$ denote the normalized eigenstate corresponding to the eigenvalue 3. Compute expressions for $\mid e_1\rangle$ and $\mid e_2\rangle$.

(d) Express the basis vectors $\mid 1\rangle$ and $\mid 2\rangle$ in terms of $\mid e_1\rangle$ and $\mid e_2\rangle$.

(e) Show that $|\, e_1\rangle$ and $|\, e_2\rangle$ are orthonormal.

(f) Express A in the basis $\{|\, 1\rangle, |\, 2\rangle\}$ using Dirac notation.

(g) Express A in the basis $\{|\, e_1\rangle, |\, e_2\rangle\}$ using Dirac notation.

(h) Suppose the system is in the state $|1\rangle$ and we wish to measure A in this state. Compute the probability of measuring 1 in the state $|1\rangle$, $\Pr_{|1\rangle}(1)$.

(i) Suppose the system is in the state $|1\rangle$ and we wish to measure A in this state. Compute the probability of measuring 3 in the state $|1\rangle$, $\Pr_{|1\rangle}(3)$.

(j) Suppose the system is in the state $|2\rangle$ and we wish to measure A in this state. Compute the probability of measuring 1 in the state $|2\rangle$, $\Pr_{|2\rangle}(1)$.

(k) Suppose the system is in the state $|2\rangle$ and we wish to measure A in this state. Compute the probability of measuring 3 in the state $|2\rangle$, $\Pr_{|2\rangle}(3)$.

(l) If the result of the measurement of A in the state $|1\rangle$ is 1, what is the state of the system after measurement?

(m) If the result of the measurement of A in the state $|1\rangle$ is 3, what is the state of the system after measurement?

(n) If the result of the measurement of A in the state $|2\rangle$ is 1, what is the state of the system after measurement?

(o) If the result of the measurement of A in the state $|2\rangle$ is 3, what is the state of the system after measurement?

4
Quantum Mechanics of Angular Momentum

Angular Momentum: Classical to Quantum

Why is angular momentum interesting? We are not going to spend a great deal of time answering this question, but you should think about it and come up with your own answers. However, you could ask the same question of "energy" and "linear momentum". One answer might be that they have proven to be very useful quantities for describing the behaviour of mechanical systems (at this point we have not distinguished classical and quantum mechanical systems–you should think about that). This is true. However, you might think of some simple mechanical systems (e.g. a point mass vibrating at the end of a spring, a point mass on the end of a string that you are swinging in a circle) and think about what quantities would best describe the motion of that system. Once you have answered this question to your satisfaction, then you could go on to a "deeper" question (that's what scientists do), and you could ask "where do energy, linear momentum, and angular momentum come from?" It's not entirely clear what would comprise a satisfactory answer to this question. However, one important, and deep, answer to this question lies in understanding the notion of *symmetry* and it's role in mechanics, both classical and quantum. There is a remarkable theorem due to the mathematician Emmy Noether called (not surprisingly) "Noether's theorem" that relates specific symmetries to "conserved quantities" (such as energy. linear momentum, and angular momentum)[1]. This will not be discussed in this course, but if you go further in your studies of classical and quantum mechanics, you will undoubtedly encounter this topic. For now, we will begin with the "traditional" way of introducing angular momentum in quantum mechanics.

We begin with the expression for the *classical* angular momentum of a single particle about a point (which we will refer to as "the origin",

[1] Emmy Noether was one of the greatest mathematicians of the twentieth century. The breadth of her contributions to mathematics is truly remarkable. Her Wikipedia entry gives a nice overview of her life, https://en.wikipedia.org/wiki/Emmy_Noether. Recently, a book has been published solely devoted to Noether's theorem (*Neuenschwander, D. E. (2017). Emmy Noether's wonderful theorem. JHU Press*).

if required):

$$\mathbf{L} = \mathbf{x} \times \mathbf{p} \tag{4.1}$$

where $\mathbf{x} \equiv (x_1, x_2, x_3)$ is the position of the particle with respect to the origin and $\mathbf{p} \equiv (p_1, p_2, p_3)$ is the (linear) momentum of the particle. The components of the angular momentum vector are given by:

$$\begin{aligned} L_1 &= x_2 p_3 - x_3 p_2, \\ L_2 &= x_3 p_1 - x_1 p_3, \\ L_3 &= x_1 p_2 - x_2 p_1, \end{aligned} \tag{4.2}$$

Now we will introduce some notation that will enable us to more concisely write the expression for angular momentum. It will also enable us to much more simply manipulate various quantities associated with angular momentum (both classical and quantum).

The Levi-Civita symbol (sometimes also referred to as the "permutation symbol", "antisymmetric symbol", or "alternating symbol") is defined as:

$$\epsilon_{ijk} = \begin{cases} 1 \quad \text{if} \quad (i,j,k) \quad \text{is} \quad (1,2,3),\ (3,1,2), \text{ or } (2,3,1), \\ -1 \quad \text{if} \quad (i,j,k) \quad \text{is} \quad (1,3,2),\ (3,2,1), \text{ or } (2,1,3), \\ 0 \quad \text{if} \quad i = j \quad \text{or} \quad j = k \quad \text{or} \quad k = i. \end{cases} \tag{4.3}$$

Stated more concisely, ϵ_{ijk} is 1 if (i,j,k) is an even permutation of $(1,2,3)$, -1 if it is an odd permutation, and 0 if any index is repeated.

Using the Levi-Civita symbol we have:

$$L_i = \sum_{j=1}^{3} \sum_{k=1}^{3} \epsilon_{ijk} x_j p_k, \qquad i = 1,2,3. \tag{4.4}$$

Of course, the first time you see this, you really should write it out completely so that you are sure that the compact notation really does give (4.2). Doing this gives:

$$\begin{aligned} L_i &= \sum_{j=1}^{3}\sum_{k=1}^{3}\epsilon_{ijk}x_j p_k \\ &= \sum_{j=1}^{3}\left(\epsilon_{ij1}x_j p_1 + \epsilon_{ij2}x_j p_2 + \epsilon_{ij3}x_j p_3\right) \\ &= \epsilon_{i11}x_1 p_1 + \epsilon_{i21}x_2 p_1 + \epsilon_{i31}x_3 p_1 \\ &\quad +\epsilon_{i12}x_1 p_2 + \epsilon_{i22}x_2 p_2 + \epsilon_{i32}x_3 p_2 \\ &\quad +\epsilon_{i13}x_1 p_3 + \epsilon_{i23}x_2 p_3 + \epsilon_{i33}x_3 p_3 \\ &= 0 + \epsilon_{i21}x_2 p_1 + \epsilon_{i31}x_3 p_1 \\ &\quad +\epsilon_{i12}x_1 p_2 + 0 + \epsilon_{i32}x_3 p_2 \\ &\quad +\epsilon_{i13}x_1 p_3 + \epsilon_{i23}x_2 p_3 + 0. \end{aligned} \tag{4.5}$$

From which it follows immediately that

$$L_1 = \epsilon_{132}x_3 p_2 + \epsilon_{123}x_2 p_3 = x_2 p_3 - x_3 p_2, \tag{4.6}$$

$$L_2 = \epsilon_{231}x_3 p_1 + \epsilon_{213}x_1 p_3 = x_3 p_1 - x_1 p_3, \tag{4.7}$$

$$L_3 = \epsilon_{321}x_2 p_1 + \epsilon_{312}x_1 p_2 = x_1 p_2 - x_2 p_1. \tag{4.8}$$

Now that we are sure that (4.4) "encodes" (4.2) we can simplify the notion even further. Instead of writing the "double sum" in (4.2) we write:

$$L_i = \sum_{jk}\epsilon_{ijk}x_j p_k, \quad i = 1, 2, 3. \tag{4.9}$$

We can simplify the expression further by writing (4.9) as:

$$L_i = \epsilon_{ijk}x_j p_k, \quad i = 1, 2, 3, \tag{4.10}$$

where it is understood that repeated indices are summed from 1 to 3. This "summation over repeated indices" is referred to as the "Einstein summation convention". To summarize, the compact notation of (4.10) means the same as (4.4). We will use the Einstein summation convention throughout these notes.

Now passing to the quantum mechanical version of (orbital) angular momentum, we apply the quantization rules of associating to the position variables x_1, x_2, x_3 the Hermitian position operators X_1, X_2, X_3, respectively, and to the momentum variables p_1, p_2, p_3 the Hermitian momentum operators P_1, P_2, P_3, respectively and form the quantities leading to the following definition.

Definition 28 (Quantum Mechanical (Orbital) Angular Momentum)**.**

$$L_1 = X_2P_3 - X_3P_2, \tag{4.11}$$
$$L_2 = X_3P_1 - X_1P_3, \tag{4.12}$$
$$L_3 = X_1P_2 - X_2P_1, \tag{4.13}$$

which we denote more concisely using the Levi-Civita symbol as follows:

$$L_i = \epsilon_{ijk}X_jP_k, \quad i = 1,2,3. \tag{4.14}$$

The operators L_i, $i = 1,2,3$ defined in this way are Hermitian, as we now show.

Proposition 1. *The operators L_1, L_2, L_3 are Hermitian operators.*

Proof. Initially, one might think that this statement is "obvious" since the L_i are the sum of the product of two Hermitian operators. However, it is not necessarily true that the product of two Hermitian operators is Hermitian, and in the course of showing that the L_i are Hermitian, we will see the crucial issue.

We carry out the proof for $L_1 = X_2P_3 - X_3P_2$. The proof for L_2 and L_3 is analogous. We have

$$L_1^\dagger = (X_2P_3 - X_3P_2)^\dagger = (X_2P_3)^\dagger - (X_3P_2)^\dagger = P_3^\dagger X_2^\dagger - P_2^\dagger X_3^\dagger = P_3X_2 - P_2X_3. \tag{4.15}$$

Now, if we could "reverse the order" of P_3X_2 and P_2X_3 we would be finished. In other words, is it true that $P_3X_2 = X_2P_3$ and $P_2X_3 = X_3P_2$? In other words, do X_2 and P_3 commute, and do X_3 and P_2 commute? The answer is "yes". Recall the commutation relations:

$$[X_j, P_k] = i\hbar\delta_{j,k}, \quad [X_j, X_k] = [P_j, P_k] = 0, \quad j,k = 1,2,3. \tag{4.16}$$

Hence, we have

$$L_1^\dagger = (X_2P_3 - X_3P_2)^\dagger = P_3X_2 - P_2X_3 = X_2P_3 - X_3P_2 = L_1. \tag{4.17}$$

□

The next result concerns the commutation relations of (orbital) angular momentum and will be fundamental to the rest of the material on angular momentum.

Proposition 2.

1. $[L_j, P_k] = i\hbar\epsilon_{jkl}P_l$
2. $[L_j, X_k] = i\hbar\epsilon_{jkl}X_l$

3. $[L_j, L_k] = i\hbar\epsilon_{jkl}L_l$

Before proving this result we need a preliminary lemma on commutators in general.

Lemma 1. *For any operators A, B, C,*

$$i)\ [AB, C] = A[B, C] + [A, C]B \qquad ii)\ [C, AB] = A[C, B] + [C, A]B$$

Proof. The proof uses the basic "trick" of adding and subtracting the same quantity to an expression, but combining the two quantities with different terms in the expression. For i) we have:

$$\begin{aligned}
[AB, C] &= ABC - CAB \\
&= A\left(BC - CB + CB\right) - CAB \\
&= A\left([B, C] + CB\right) - CAB \\
&= A[B, C] + \left(AC - CA + CA\right)B - CAB \\
&= A[B, C] + \left([A, C] + CA\right)B - CAB \\
&= A[B, C] + [A, C]B
\end{aligned}$$

The proof of ii) follows immediately if we note that $[AB, C] = -[C, AB]$.

□

We now return to the proof of Proposition 2.

Proof. We begin with 1.

$$\begin{aligned}
[L_j, P_k] &= [\epsilon_{jmn}\, X_m P_n, P_k] \\
&= \epsilon_{jmn}\, [X_m P_n, P_k] \\
&= \epsilon_{jmn}\, \left(X_m[P_n, P_k] + [X_m, P_k]P_n\right) \quad \text{using Lemma 1 i)} \\
&= \epsilon_{jmn}\, \left(0 + i\hbar\delta_{mk}P_n\right) \quad \text{using (4.16)} \\
&= i\hbar\epsilon_{jkn}\, P_n
\end{aligned}$$

The proof of 2 proceeds along the same lines.

$$\begin{aligned}
[L_j, X_k] &= [\epsilon_{jmn}\, X_m P_n, X_k] \\
&= \epsilon_{jmn}\, [X_m P_n, X_k] \\
&= \epsilon_{jmn}\, \left(X_m[P_n, X_k] + [X_m, X_k]P_n\right) \quad \text{using Lemma 1 i)} \\
&= \epsilon_{jmn}\, \left(-i\hbar\delta_{kn}X_m + 0\right) \quad \text{using (4.16)} \\
&= -i\hbar\epsilon_{jmk}\, X_m, \\
&= i\hbar\epsilon_{jkm}\, X_m
\end{aligned}$$

We prove 3 for L_1 and L_2. The proof for the other components is easily obtained by permuting the indices appropriately.

$$\begin{aligned}
[L_1, L_2] &= [L_1, X_3P_1 - X_1P_3] \\
&= [L_1, X_3P_1] - [L_1, X_1P_3] \\
&= X_3[L_1, P_1] + [L_1, X_3]P_1 - X_1[L_1, P_3] - [L_1, X_1]P_3 \quad \text{using Lemma 1 ii)} \\
&= 0 + i\hbar\epsilon_{13l}X_lP_1 - X_1i\hbar\epsilon_{13l}P_l + 0 \quad \text{using parts 1 and 2 of this Proposition} \\
&= i\hbar\epsilon_{132}\left(X_2P_1 - X_1P_2\right) \\
&= i\hbar(X_1P_2 - X_2P_1) = i\hbar L_3
\end{aligned}$$

□

The commutation relations in statement 3 of Proposition 2 are really the "heart" of the quantum theory of angular momentum. You will see that they form the basis of essentially all of the results that we develop related to angular momentum. In more advanced treatments of this topic it can be shown that operators satisfying these commutation relations have their origins in the properties of rotations in three dimensional space. Therefore it is natural to define any three Hermitian operators that satisfy such commutation related as quantum angular momentum operators.

Definition 29 (General Definition of Quantum Angular Momentum Operators)**.** *Any three Hermitian operators J_1, J_2, J_3 that satisfy:*

$$[J_j, J_k] = i\hbar\epsilon_{jkl}J_l, \tag{4.18}$$

are called (quantum) angular momentum operators.

An obvious question is "why do we do this"? The quantization of classical orbital angular momentum was fairly straightforward. Why does this need to be "generalized" in any way? It turns out that there are other "types of angular momentum", in addition to the quantization of classical orbital angular momentum. "Spin" is one type of "other" angular momentum. We will study spin explicitly at the end of our study of angular momentum. However, for now we merely state that spin is a property of particles (such as electrons, protons, neutrons, photons, etc) that has no classical analog. However, it is described by three Hermitian operators satisfying the same commutation relations that we derived for orbital angular momentum. This is a good place to say something about terminology.

Terminology. Eigenvalues and eigenvectors are common terms associated with linear operators. However, in quantum mechanics some synonyms are associated with the term "eigenvector" that are specific to

the particular physics and/or notation context. For example, the word "eigenstate" or "eigenket" (reflecting the influence of the Dirac notation) is often used synonomously for the term "eigenvector".

Our next goal is to construct a basis of angular momentum eigenvectors (what does this mean?). Initially, this would seem to pose a problem since we cannot construct a basis consisting of simultaneous eigenvectors of J_1, J_2, and J_3 since these three operators do not commute with each other (we will comment on this shortly). However, we can motivate this by analogy with classical orbital angular momentum. Classical orbital angular momentum is a vector in three dimensions, and therefore it can be described by a magnitude and a direction. Quantum mechanically we can think of the operator defined by the square of the vector of three operators defining the quantum mechanical angular momentum, i.e. the square of

$$\mathbf{J} \equiv (J_1, J_2, J_3),$$

as defining the square of the magnitude of total angular momentum:

$$\mathbf{J} \cdot \mathbf{J} \equiv \mathbf{J}^2 = J_1^2 + J_2^2 + J_3^2. \tag{4.19}$$

It is easy to see that $\mathbf{J}^2$ is Hermitian since it is the sum of the squares of three Hermitian operators (we don't have the problem that we had earlier since J_i commutes with J_i, $i = 1, 2, 3$). Moreover, we have the following result.

Proposition 3.

$$[\mathbf{J}^2, J_i] = 0, \quad i = 1, 2, 3.$$

Proof.

$$\begin{aligned}
[\mathbf{J}^2, J_i] &= [J_j J_j, J_i], \\
&= J_j [J_j, J_i] + [J_j, J_i] J_j, \quad \text{using Lemma 1 i)} \\
&= i\hbar J_j \epsilon_{jik} J_k + i\hbar \epsilon_{jik} J_k J_j, \quad \text{using (4.18)} \\
&= i\hbar \epsilon_{jik} \left(J_j J_k + J_k J_j \right) \\
&= 0,
\end{aligned}$$

since ϵ_{jik} is antisymmetric in j, k and the term in parentheses is symmetric in j, k. We show this last step explicitly. We have

$$\begin{aligned}
\epsilon_{jik} J_k J_j &= \epsilon_{kij} J_j J_k, \quad \text{where we have just interchanged the summation indices } k \text{ and } j. \\
&= -\epsilon_{jik} J_j J_k,
\end{aligned}$$

where in the last step we have used the fact that antisymmetry implies we can interchange the order of two indices, but this requires us to

multiply the resulting term by a minus sign. Adding the left and right hand sides gives:

$$\epsilon_{jik} J_j J_k + \epsilon_{jik} J_k J_j = \epsilon_{jik} (J_j J_k + J_k J_j) = 0.$$

□

This result implies that we can find simultaneous eigenvectors of $\mathbf{J}$ and one component of angular momentum, which we shall take as J_3 (mainly for traditional reasons). Now there is a *lot* in this result, and in the statements leading up to it, that we now want to discuss.

Recall that two of the nicest and most useful properties of Hermitian operators are that their eigenvalues are real, and eigenvectors corresponding to distinct eigenvalues are orthogonal (and we will always be normalizing our eigenvectors so that they have unit length–if this is not familiar, you should go back and review this from earlier in the course). It could happen that two different eigenvectors have the same eigenvalue. This situation is referred to as a *degeneracy*. We will not deal with degeneracies explicitly, at the moment, but it is important to realize that they can arise and whether or not our algebraic manipulations allow for their possibility. Even when degeneracies exist, it is still possible to find a "complete set" of orthonormal eigenvectors. You just have to work a bit harder.

Now let's consider two Hermitian operators, and let's rule out the possibility of degeneracies for the moment (it makes getting the main idea across easier–we can come back and consider the case of degeneracies afterwards). We know that each Hermitian operator, individually, possesses a *complete set of orthonormal eigenvectors* (make sure you know what the word "complete" means in this statement). However, if the two operators commute (and there are no degeneracies, for the moment) then it is possible to find a complete set of orthonormal vectors that are eigenvectors for each Hermitian operator simultaneously. Commutation of the operators is essential here. So, getting back to angular momentum, we know that there exists a complete set of orthonormal vectors that are eigenvectors for both $\mathbf{J}^2$ and J_3. We just need to be able to compute them, which is the next topic. The same result can be shown to hold when there are degeneracies, but we will postpone dealing with that situation, for the moment. The book by Cohen-Tannoudji et al. has a very nice discussion of these ideas in the chapter on "Mathematical Tools of Quantum Mechanics". The Littlejohn notes also have a good treatment of this topic. (Recall that you proved this result earlier in the course, and even dealt with the case of degeneracies.)

Before going on to compute eigenvalues and eigenvectors associated with angular momentum, it is worth making some final remarks.

Recall that in quantum mechanics Hermitian operators are referred to as "observables". They are the mathematical manifestations of quantities that we can "observe", i.e. measure. The latter part of this course will be very much concerned with this topic. However, recall that two observables can be measured simultaneously if and only if the corresponding operators commute. So we have shown that $\mathbf{J}^2$ commutes with J_1, J_2 and J_3. Does this mean that $\mathbf{J}^2$, J_1, J_2, and J_3 can all be measured simultaneously? But J_1, J_2, and J_3 do not commute. What is the situation here?

Eigenvalues and Eigenvectors of Angular Momentum

Now we are going to do three things.

- Compute eigenvalues of $\mathbf{J}^2$ and J_3.
- Compute eigenvectors that are simultaneously eigenvectors of $\mathbf{J}^2$ and J_3.
- Use the eigenvectors to compute matrix representations of the components of $\mathbf{J}$ for "certain physical situations" (which will be more fully described when we get to it).

The material below follows closely the "Littlejohn Notes" ("Notes 13"). We denote the eigenvalues of $\mathbf{J}^2$ by $a\hbar^2$ and the eigenvalues of J_3 by $m\hbar$. The factors of $\hbar$ may seem a bit mysterious. However, recall that $\hbar$ has the units of angular momentum, i.e. Joule $\cdot$ second or kilogram $\cdot$ meter^2 /second. Therefore the factors of $\hbar$ serve to make the eigenvalues a and m dimensionless. We will label the eigenvectors by the eigenvalues a and m as $|am\rangle$, and therefore we have:

$$\begin{aligned} \mathbf{J}^2|am\rangle &= a\hbar^2|am\rangle, \\ J_3|am\rangle &= m\hbar|am\rangle. \end{aligned} \tag{4.20}$$

We are assuming nondegeneracy, so that a and m are unique labels for the simultaneous eigenvectors of $\mathbf{J}$ and J_3, up to normalization and the choice of a phase. However, we will always assume that the eigenvectors are normalized, i.e.,

$$\langle am|am\rangle = 1. \tag{4.21}$$

Later we will consider the consequences of allowing for degeneracies.

The following operators will play an important role in our construction of eigenvalues and eigenvectors.

Definition 30 (Ladder Operators for Quantum Angular Momentum)**.** *The "ladder operators" J_+ and J_- are defined by:*

$$J_+ = J_1 + iJ_2, \tag{4.22}$$
$$J_- = J_1 - iJ_2 \tag{4.23}$$

It should be clear that these operators are Hermitian conjugates in the sense that:

$$(J_\pm)^\dagger = J_\mp. \tag{4.24}$$

The following result gives some fundamental commutation relations among $\mathbf{J}^2$, J_3 and the ladder operators that will play important roles in our understanding of the "eigenstructure" of $\mathbf{J}^2$ and J_3.

Proposition 4.

1. $[J_3, J_\pm] = \pm\hbar J_\pm$,
2. $[J^2, J_\pm] = 0$,
3. $[J_+, J_-] = 2\hbar J_3$,
4. $J_\pm J_\mp = J^2 - J_3^2 \pm \hbar J_3$.

Proof. The proofs of these results uses the definition of the ladder operators (Definition 30), the angular momentum commutation relations given in (4.18), and Proposition 3. The necessary calculations proceed as follows.

1.

$$\begin{aligned}
[J_3, J_\pm] &= [J_3, J_1 \pm iJ_2] \\
&= [J_3, J_1] \pm i[J_3, J_2] \\
&= i\hbar J_2 \pm i(-i\hbar)J_1 \\
&= i\hbar J_2 \pm \hbar J_1 \\
&= \pm\hbar(J_1 \pm iJ_2) \\
&= \pm\hbar J_\pm
\end{aligned}$$

2.

$$\begin{aligned}
[J^2, J_\pm] &= [J^2, J_1 \pm iJ_2] \\
&= [J^2, J_1] \pm i[J^2, J_2] \\
&= 0 \pm i\,0 = 0
\end{aligned}$$

3.

$$\begin{aligned}
[J_+, J_-] &= [J_1 + iJ_2\,, J_1 - iJ_2] \\
&= [J_1 + iJ_2\,, J_1] + [J_1 + iJ_2\,, -iJ_2] \\
&= [J_1, J_1] + [iJ_2, J_1] + [J_1, -iJ_2] + [iJ_2, -iJ_2] \\
&= 0 + i(-i\hbar)J_3 - i(i\hbar)J_3 + 0 \\
&= 2\hbar J_3
\end{aligned}$$

4.

$$\begin{aligned}
J_+J_- &= (J_1 + iJ_2)(J_1 - iJ_2) \\
&= J_1^2 + J_2^2 - i(J_1J_2 - J_2J_1) \\
&= J^2 - J_3^2 - i[J_1, J_2] \\
&= J^2 - J_3^2 + \hbar J_3 \qquad (4.25)
\end{aligned}$$

$$\begin{aligned}
J_-J_+ &= (J_1 - iJ_2)(J_1 + iJ_2) \\
&= J_1^2 + J_2^2 + i(J_1J_2 - J_2J_1) \\
&= J^2 - J_3^2 + i[J_1, J_2] \\
&= J^2 - J_3^2 - \hbar J_3 \qquad (4.26)
\end{aligned}$$

□

We will also need the results given in the following Corollary.

Corollary 1.

$$\langle am|J_-J_+|am\rangle = \hbar^2\left(a - m(m+1)\right) \geq 0 \qquad (4.27)$$
$$\langle am|J_+J_-|am\rangle = \hbar^2\left(a - m(m-1)\right) \geq 0 \qquad (4.28)$$

Proof. We first prove (4.27).

Using (4.26), we have

$$\begin{aligned}
\langle am|J_-J_+|am\rangle &= \langle am|\left(J^2 - J_3^2 - \hbar J_3\right)|am\rangle \\
&= \langle am|J^2|am\rangle - \langle am|J_3^2|am\rangle - \hbar\langle am|J_3|am\rangle \\
&= \hbar^2 a\langle am|am\rangle - \hbar^2 m^2\langle am|am\rangle - \hbar^2 m\langle am|am\rangle \\
&= \hbar^2\left(a - m(m+1)\right)\langle am|am\rangle \\
&= \hbar^2\left(a - m(m+1)\right)
\end{aligned}$$

It remains to argue that this expression is nonnegative. Using (4.24), we have

$$(J_+|am\rangle)^\dagger = \langle am|J_+^\dagger = \langle am|J_-,$$

and therefore $\langle am|J_-J_+|am\rangle$ is the square of the norm of $J_+|am\rangle$, which is clearly nonnegative.

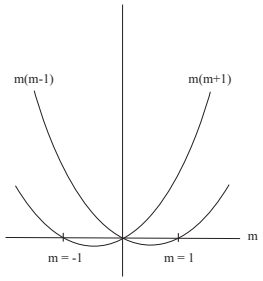
The proof of (4.28) is completely analogous to the proof of (4.27). □

Now we have the tools to prove our main result.

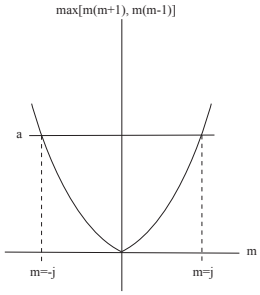
First, note that from (4.27) and (4.28) that we have:

$$a \geq \max\left[m(m+1), m(m-1)\right]. \tag{4.29}$$

The functions $m(m-1)$ and $m(m+1)$ are plotted in Fig. 4.1(a), and the maximum of these functions is plotted in Fig. 4.1(b).



(a) Graphs of the functions $m(m-1)$ and $m(m+1)$.



(b) Graph of the function max (m(m+1), m(m-1)) with maximum and minimum values of m shown for a given value of a.

It is not hard to verify that the maximum function is symmetric about $m = 0$ and ≥ 0 everywhere. We choose a value of $a \geq 0$ and draw a horizontal line in Fig. 4.1(b) corresponding to $a = \max[m(m+1), m(m-1)]$. This shows that for any $a \geq 0$ there is a maximum and a minimum value of m (which, by symmetry, are of equal magnitude) for which (4.29) is satisfied for all values in between the maximum and minimum. For $a \geq 0$ we denote the maximum value of m by j and the minimum value of m by $-j$. Therefore we have:

$$-j \leq m \leq j, \tag{4.30}$$

Clearly, j is a function of a, and $j \geq 0$ since $a \geq 0$. From Fig. 4.1(b) we see that:

$$a = j(j+1). \tag{4.31}$$

We will find it more convenient to parametrize the eigenvectors of $\mathbf{J}^2$ by j, rather than a. Therefore we will write $j(j+1)$ for the eigenvalue of $\mathbf{J}^2$, rather than a, and we will denote the simultaneous eigenvectors of $\mathbf{J}^2$ and J_3 by $|\, jm\rangle$, rather than $|\, am\rangle$. Then (4.20) is rewritten as follows:

$$\begin{aligned} \mathbf{J}^2|jm\rangle &= j(j+1)\hbar^2|jm\rangle, \\ J_3|jm\rangle &= m\hbar|jm\rangle. \end{aligned} \tag{4.32}$$

Using this new notation, we also rewrite (4.27) and (4.28) as (where we have also factored the right hand side of each expression):

$$\langle jm|J_-J_+|jm\rangle = \hbar^2\left(j(j+1) - m(m+1)\right) = \hbar^2(j-m)(j+m+1) \geq 0, \tag{4.33}$$

$$\langle jm|J_+J_-|jm\rangle = \hbar^2\left(j(j+1) - m(m-1)\right) = \hbar^2(j+m)(j-m+1) \geq 0, \tag{4.34}$$

Now we want to consider the conditions under which the inequalities (4.33) and (4.34) become equalities, i.e. under what conditions do the matrix elements on the left hand side of the equals sign become zero? First we consider (4.33) . In this case we have $J_+|jm\rangle = 0$ if and only if

$$j - m = 0 \quad \text{or} \quad j + m + 1 = 0, \tag{4.35}$$

or, equivalently

$$m = j \quad \text{or} \quad m = -j - 1, \tag{4.36}$$

However, from (4.30), it follows that $m = -j-1$ cannot occur, and therefore,

$$J_+ \mid jm\rangle = 0 \quad \text{if and only if} \quad m = j. \tag{4.37}$$

A similar analysis of (4.34) shows that:

$$J_- \mid jm\rangle = 0 \quad \text{if and only if} \quad m = -j. \tag{4.38}$$

We assume that $\mid jm\rangle$ is a normalized eigenvector of $\mathbf{J}^2$ and J_3 with eigenvalues $j(j+1)\hbar^2$ and $m\hbar$. Using Proposition 4, part 1, we have:

$$J_3\left(J_\pm \mid jm\rangle\right) = \left(J_\pm J_3 \pm \hbar J_\pm\right) \mid jm\rangle = m\hbar J_\pm \mid jm\rangle \pm \hbar J_\pm \mid jm\rangle = (m\pm 1)\hbar\left(J_\pm \mid jm\rangle\right), \tag{4.39}$$

and, using Proposition 4, part 2 we have

$$\mathbf{J}^2\left(J_\pm \mid jm\rangle\right) = J_\pm \mathbf{J}^2 \mid jm\rangle = j(j+1)\hbar^2\left(J_\pm \mid jm\rangle\right). \tag{4.40}$$

From these two calculations we make the following conclusions.

- If $J_+ \mid jm\rangle$ does not vanish, then it is an eigenvector of $\mathbf{J}^2$ and J_3 with eigenvalues $j(j+1)\hbar^2$ and $(m+1)\hbar$, respectively. In other words, J_+ does not change j, but it increases m by 1.
- If $J_- \mid jm\rangle$ does not vanish, then it is an eigenvector of $\mathbf{J}^2$ and J_3 with eigenvalues $j(j+1)\hbar^2$ and $(m-1)\hbar$, respectively. In other words, J_- does not change j, but it decreases m by 1.

Therefore, repeated application of J_+ on $\mid jm\rangle$ generates the sequence of eigenvectors:

$$\mid jm\rangle, \mid j, m+1\rangle. \ldots, \mid j, m+n_1\rangle,$$

which must terminate for some integer $n_1 \geq 0$, or else the bounds (4.30) would be violated. Hence we have:

$$j = m + n_1, \quad \text{or} \quad m = j - n_1. \tag{4.41}$$

Similarly, repeated application of J_- on $\mid jm\rangle$ generates the sequence of eigenvectors:

$$\mid jm\rangle, \mid j, m-1\rangle. \ldots, \mid j, m-n_2\rangle,$$

which must terminate for some integer $n_2 \geq 0$, or else the bounds (4.30) would be violated. Hence we have:

$$-j = m - n_2, \quad \text{or} \quad m = -j + n_2. \tag{4.42}$$

Taken together, (4.41) and (4.42) imply that:

$$2j = n_1 + n_2 \geq 0, \tag{4.43}$$

This implies that the only allowed values of j are:

$$j \in \{0, \frac{1}{2}, 1, \frac{3}{2}, \ldots\} \tag{4.44}$$

Key point: It is important to realize that (4.44) only tells us that these are the values of j allowed by the angular momentum commutation relations. It does not tell us which particular values occur in a specific application. This highlights the strength of considering "angular momentum in general, as defined by the commutation relations, and then worrying about the particular type of angular momentum when considering specific applications.
For example, when the angular momentum under consideration corresponds to spin (something we have not considered yet) a "spin $\frac{1}{2}$" particle only has $j = \frac{1}{2}$. For a particle moving in a central force potential orbital angular momentum is the relevant type of angular momentum, and it can only possess eigenvalues having integer values of j (which we will discuss shortly).

We summarize the main result of this section as follows.

Theorem 11. *The operator $\mathbf{J}^2$ has eigenvalues $j(j+1)\hbar^2$, with $j \in \{0, \frac{1}{2}, 1, \frac{3}{2}, \ldots\}$. For each j we have $2j+1$ eigenvectors, $\mid jm\rangle$, $m = -j, -j+1, \ldots, j-1, j$ which are eigenvectors of J_3 with respective eigenvalue $m\hbar$, i.e.*

$$\mathbf{J}^2 \mid jm\rangle = j(j+1)\hbar^2 \mid jm\rangle, \qquad J_3 \mid jm\rangle = m\hbar \mid jm\rangle. \tag{4.45}$$

The Eigenvalues of $J_\pm$ and a Phase Convention

We have shown that $J_+ \mid jm\rangle$ and $J_- \mid jm\rangle$ are eigenvectors of J_3 with eigenvalues $(m+1)\hbar$ and $(m-1)\hbar$, respectively. Assuming nondegeneracy (why?), we can write

$$J_+ \mid jm\rangle = c \mid j, m+1\rangle \tag{4.46}$$
$$J_- \mid jm\rangle = c' \mid j, m-1\rangle, \tag{4.47}$$

where c and c' are complex numbers that are arbitrary and therefore are chosen to be real and positive. The specific choices are a matter of convention, and are chosen as follows. Recalling (4.33) and (4.34), we have:

$$|c|^2\langle j,m+1 \mid j,m+1\rangle = \langle jm \mid J_-J_+ \mid jm\rangle \quad = \quad \mid c \mid^2 = \hbar^2(j-m)(j+m+1)$$
$$|c'|^2\langle j,m-1 \mid j,m-1\rangle = \langle jm \mid J_+J_- \mid jm\rangle \quad = \quad \mid c' \mid^2 = \hbar^2(j+m)(j-m+1)$$

Hence, with the chosen phase convention, we can take:

$$c = \hbar\sqrt{(j-m)(j+m+1)}, \tag{4.48}$$
$$c' = \hbar\sqrt{(j+m)(j-m+1)}, \tag{4.49}$$

and therefore we have:

$$J_+ \mid jm\rangle = \hbar\sqrt{(j-m)(j+m+1)}\mid j,m+1\rangle \tag{4.50}$$
$$J_- \mid jm\rangle = \hbar\sqrt{(j+m)(j-m+1)}\mid j,m-1\rangle \tag{4.51}$$

Degeneracies and Multiplicities

In several places throughout our discussions we have explicitly assumed that there were "no degeneracies". Now we will consider this question more carefully (look in the section just above and convince yourself why the assumption of nondegeneracy was necessary).

First, we recall what it means for the simultaneous eigenvectors of $\mathbf{J}^2$ and J_3 to be degenerate, but now in a bit more detail. We denote the eigenspace of $\mathbf{J}^2$ and J_3 corresponding to eigenvalues $j(j+1)\hbar^2$ and $m\hbar$, respectively, by $\mathcal{E}_{jm}$ and we denote its dimension by:

$$\dim \mathcal{E}_{jm} = N_{jm}. \tag{4.52}$$

Then the simultaneous eigenvectors of $\mathbf{J}^2$ and J_3 are said to be degenerate if $N_{jm} > 1$. Stated another way, for a fixed j and m, there is more than one simultaneous eigenvector of $\mathbf{J}^2$ and J_3 corresponding to that fixed j and m. Here is the main general result on degeneracies for $\mathbf{J}^2$ and J_3.

> *For fixed j all the eigenspaces $\mathcal{E}_{jm}$ for $m = -j, \ldots, +j$ have the same dimension. We denote this dimension by N_j, which is referred to as the multiplicity of the j value. The multiplicity can take on any value from 0 (which corresponds to that case of that particular j value not occurring) to ∞.*

We give an outline of the proof that is almost exactly that given in the "Littlejohn notes", but just we a few comments concerning where more details could be added.

The ladder operators will play an important role in this argument.

We start with the eigenspace $\mathcal{E}_{jj}$ which has dimension N_{jj} , and then we choose a set of N_{jj} linearly independent vectors in this space (why can we make such a choice?). If we apply J_- to these vectors, we obtain a set of N_{jj} vectors that are eigenvectors of $\mathbf{J}^2$ and J_3 with eigenvalues $j(j+1)\hbar^2$ and $(m-1)\hbar$ (that is, with a lowered value of m). These vectors must lie in the eigenspace $\mathcal{E}_{j,j-1}$, and, as one can show, they are also linearly independent (you should show that this is true). Thus, $\dim \mathcal{E}_{j,j-1} = N_{j,j-1} \geq N_{jj}$ (why do we have $\geq$ here and not $=$?).

Now we go "back up" with the other ladder operator. Choose a set of $N_{j,j-1}$ linearly independent vectors in $\mathcal{E}_{j,j-1}$ and apply the raising operator J_+ to them. This creates a set of $N_{j,j-1}$ vectors that lie in the eigenspace $\mathcal{E}_{jj}$, which, as one can show, are also linearly independent (you should understand how this can be shown) . Thus, $N_{jj} \geq N_{j,j-1}$. But this is consistent with $N_{j,j-1} \geq N_{jj}$ only if $N_{j,j-1} = N_{jj}$.

This argument can be repeated for all of the m values, and in this way we see that all the eigenspaces $\mathcal{E}_{jm}$ for $m = -j, \ldots, j$ have the same dimension. We denote this dimension by N_j , which we call the multiplicity of the given j value. The multiplicity can take on any value from 0 (in which case the j value does not occur) to ∞.

Examples of Matrix Representations for Angular Momentum

Using eigenvectors we can construct matrix representations for systems with a fixed j.

First, we collect together some matrix elements that we have already computed (you need to go back in the notes and make sure of this) and that we will need.

$$\langle jm' \mid J_3 \mid jm \rangle = m\hbar\delta_{m'm}, \tag{4.53}$$

$$\langle jm' \mid \mathbf{J}^2 \mid jm \rangle = \hbar^2 j(j+1)\delta_{m'm}, \tag{4.54}$$

$$\langle jm' \mid J_+ \mid jm \rangle = \hbar\sqrt{(j-m)(j+m+1)}\,\delta_{m',m+1}, \tag{4.55}$$

$$\langle jm' \mid J_- \mid jm \rangle = \hbar\sqrt{(j+m)(j-m+1)}\,\delta_{m',m-1}. \tag{4.56}$$

Next, it follows from the definition of the ladder operators given in (4.23) that we have:

$$J_1 = \frac{1}{2}(J_+ + J_-), \qquad J_2 = \frac{1}{2i}(J_+ - J_-). \tag{4.57}$$

From these relations it follows that if we know matrix representations of J_+ and J_- then we can use them to determine the corresponding matrix representations of J_1 and J_2 (actually, this needs a "little" proof, but you should be able to convince yourself that it is true–operators and their matrix representations are different mathematical objects).

Now we can begin our construction of matrix representations of the angular momentum operators for different values of j.

$\mathbf{j = 0}$

This is an "easy" case. There is only one eigenvector, $|\ 00\rangle$, and it follows from (4.53) - (4.56) that all the relevant matrix elements are zero. Therefore the matrix representations of $\mathbf{J}^2$, J_1, J_2, and J_3 are all zero.

$\mathbf{j = \frac{1}{2}}$: The "spin $\frac{1}{2}$ representation"

In this case the relevant eigenvectors are $|\ \frac{1}{2}, \frac{1}{2}\rangle$ and $|\ \frac{1}{2}, -\frac{1}{2}\rangle$. The "general form" of the matrix representation is as follows:

$$\begin{pmatrix} \langle \frac{1}{2}, \frac{1}{2}\ |\ \ |\ \frac{1}{2}, \frac{1}{2}\rangle & \langle \frac{1}{2}, \frac{1}{2}\ |\ \ |\ \frac{1}{2}, -\frac{1}{2}\rangle \\ \\ \langle \frac{1}{2}, -\frac{1}{2}\ |\ \ |\ \frac{1}{2}, \frac{1}{2}\rangle & \langle \frac{1}{2}, -\frac{1}{2}\ |\ \ |\ \frac{1}{2}, -\frac{1}{2}\rangle \end{pmatrix}, \tag{4.58}$$

where by the phrase "general form" we mean that the matrix representation for a particular operator (e.g. J_3) is obtained by inserting the operator in the "gap" between the bra's and ket's in the general form, and using (4.53) - (4.56) to compute the associated matrix elements.

Therefore, the matrix representation for J_3 is given by:

$$\begin{pmatrix} \langle \frac{1}{2}, \frac{1}{2}\ |\ J_3\ |\ \frac{1}{2}, \frac{1}{2}\rangle & \langle \frac{1}{2}, \frac{1}{2}\ |\ J_3\ |\ \frac{1}{2}, -\frac{1}{2}\rangle \\ \\ \langle \frac{1}{2}, -\frac{1}{2}\ |\ J_3\ |\ \frac{1}{2}, \frac{1}{2}\rangle & \langle \frac{1}{2}, -\frac{1}{2}\ |\ J_3\ |\ \frac{1}{2}, -\frac{1}{2}\rangle \end{pmatrix} = \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ \\ 0 & -1 \end{pmatrix}. \tag{4.59}$$

The matrix representation for J_- is given by:

$$\begin{pmatrix} \langle \frac{1}{2}, \frac{1}{2}\ |\ J_-\ |\ \frac{1}{2}, \frac{1}{2}\rangle & \langle \frac{1}{2}, \frac{1}{2}\ |\ J_-\ |\ \frac{1}{2}, -\frac{1}{2}\rangle \\ \\ \langle \frac{1}{2}, -\frac{1}{2}\ |\ J_-\ |\ \frac{1}{2}, \frac{1}{2}\rangle & \langle \frac{1}{2}, -\frac{1}{2}\ |\ J_-\ |\ \frac{1}{2}, -\frac{1}{2}\rangle \end{pmatrix} = \hbar \begin{pmatrix} 0 & 0 \\ \\ 1 & 0 \end{pmatrix} \tag{4.60}$$

The matrix representation for J_+ is given by:

$$\begin{pmatrix} \langle \frac{1}{2}, \frac{1}{2}\ |\ J_+\ |\ \frac{1}{2}, \frac{1}{2}\rangle & \langle \frac{1}{2}, \frac{1}{2}\ |\ J_+\ |\ \frac{1}{2}, -\frac{1}{2}\rangle \\ \\ \langle \frac{1}{2}, -\frac{1}{2}\ |\ J_+\ |\ \frac{1}{2}, \frac{1}{2}\rangle & \langle \frac{1}{2}, -\frac{1}{2}\ |\ J_+\ |\ \frac{1}{2}, -\frac{1}{2}\rangle \end{pmatrix} = \hbar \begin{pmatrix} 0 & 1 \\ \\ 0 & 0 \end{pmatrix} \tag{4.61}$$

Using the matrix representations for J_- and J_+ with (4.57), the matrix representations for J_1 and J_2, respectively, are easily obtained and are found to be:

$$\frac{\hbar}{2} \begin{pmatrix} 0 & 1 \\ \\ 1 & 0 \end{pmatrix}, \qquad \frac{\hbar}{2} \begin{pmatrix} 0 & -i \\ \\ i & 0 \end{pmatrix} \tag{4.62}$$

The spin $\frac{1}{2}$ matrix representations for J_1, J_1, J_3 (neglecting the factor $\frac{\hbar}{2}$) occur frequently enough in many applications that they warrant a special name. They are referred to as the *Pauli spin matrices*, which are denoted as follows:

$$X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{4.63}$$

$\mathbf{j = 1}$: The "spin 1 representation"

We proceed exactly as in the case for $j = \frac{1}{2}$. In this case the relevant eigenvectors are $|\, 1, 1\rangle$, $|\, 1, 0\rangle$ and $|\, 1, -1\rangle$. The "general form" of the matrix representation is as follows:

$$\begin{pmatrix} \langle 1,1| \;\; |1,1\rangle & \langle 1,1| \;\; |1,0\rangle & \langle 1,1| \;\; |1,-1\rangle \\ \langle 1,0| \;\; |1,1\rangle & \langle 1,0| \;\; |1,0\rangle & \langle 1,0| \;\; |1,-1\rangle \\ \langle 1,-1| \;\; |1,1\rangle & \langle 1,-1| \;\; |1,0\rangle & \langle 1,-1| \;\; |1,-1\rangle \end{pmatrix} \tag{4.64}$$

where by the phrase "general form", as above, we mean that the matrix representation for a particular operator (e.g. J_3) is obtained by inserting the operator in the "gap" between the bra's and ket's in the general form, and using (4.53) - (4.56) to compute the associated matrix elements.

Therefore, the matrix representation for J_3 is given by:

$$\begin{pmatrix} \langle 1,1 \mid J_3 \mid 1,1\rangle & \langle 1,1 \mid J_3 \mid 1,0\rangle & \langle 1,1 \mid J_3 \mid 1,-1\rangle \\ \langle 1,0 \mid J_3 \mid 1,1\rangle & \langle 1,0 \mid J_3 \mid 1,0\rangle & \langle 1,0 \mid J_3 \mid 1,-1\rangle \\ \langle 1,-1 \mid J_3 \mid 1,1\rangle & \langle 1,-1 \mid J_3 \mid 1,0\rangle & \langle 1,-1 \mid J_3 \mid 1,-1\rangle \end{pmatrix} = \hbar \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} \tag{4.65}$$

The matrix representation for J_- is given by:

$$\begin{pmatrix} \langle 1,1 \mid J_- \mid 1,1\rangle & \langle 1,1 \mid J_- \mid 1,0\rangle & \langle 1,1 \mid J_- \mid 1,-1\rangle \\ \langle 1,0 \mid J_- \mid 1,1\rangle & \langle 1,0 \mid J_- \mid 1,0\rangle & \langle 1,0 \mid J_- \mid 1,-1\rangle \\ \langle 1,-1 \mid J_- \mid 1,1\rangle & \langle 1,-1 \mid J_- \mid 1,0\rangle & \langle 1,-1 \mid J_- \mid 1,-1\rangle \end{pmatrix} = \hbar \begin{pmatrix} 0 & 0 & 0 \\ \sqrt{2} & 0 & 0 \\ 0 & \sqrt{2} & 0 \end{pmatrix} \tag{4.66}$$

The matrix representation for J_+ is given by:

$$\begin{pmatrix} \langle 1,1 \mid J_+ \mid 1,1\rangle & \langle 1,1 \mid J_+ \mid 1,0\rangle & \langle 1,1 \mid J_+ \mid 1,-1\rangle \\ \langle 1,0 \mid J_+ \mid 1,1\rangle & \langle 1,0 \mid J_+ \mid 1,0\rangle & \langle 1,0 \mid J_+ \mid 1,-1\rangle \\ \langle 1,-1 \mid J_+ \mid 1,1\rangle & \langle 1,-1 \mid J_+ \mid 1,0\rangle & \langle 1,-1 \mid J_+ \mid 1,-1\rangle \end{pmatrix} = \hbar \begin{pmatrix} 0 & \sqrt{2} & 0 \\ 0 & 0 & \sqrt{2} \\ 0 & 0 & 0 \end{pmatrix} \tag{4.67}$$

Using the matrix representations for J_- and J_+ with (4.57), the matrix representations for J_1 and J_2, respectively, are easily obtained.

Orbital Angular Momentum

Recall the way in which we began our study of the quantum theory of angular momentum. We began with the *classical* expression of orbital angular momentum of a particle about a point, and then "quantized" this expression by associating position and momentum operators to the position and momentum coordinates (Definition 4.13). Then the commutation relations for position and momentum operators that had been considered earlier in the course were used to derive commutation relations for the different components of the (orbital) angular momentum (Proposition 2). We then made the "leap" of defining a "general" quantum mechanical angular momentum operator as any three Hermitian operators satisfying the commutation relations derived for the orbital angular momentum (Definition 4.18). The theory subsequently developed was solely a consequence of these three commutation relations for the three Hermitian operators. We justified this leap into abstraction by saying that there we would encounter different types of angular momentum and, quantum mechanically, their common feature is the commutation relations. There is a justification for this statement, but it involves a more deep consideration of symmetries and rotations than we will go into in this course. An excellent discussion can be found in the "Littlejohn Notes" or in Cohen-Tannoudji et al.

However, the important point we want to make now is the following. The allowed values of j and m that we obtained above follow solely from the commutation relations for J_1, J_2, and J_3 (convince your self of this). A particular "type" of angular momentum may have physical constraints that limit those values in some way. That is what we now want to address by returning to that particular type of angular momentum from which we started, orbital angular momentum. In particular, we have the following result.

Theorem 12. *For* orbital *angular momentum j and m must both be integers.*

Proof. We will give the standard "proof" of this result (and explain

why we put "proof" in quotes afterwards).

Our work in this part of the course with angular momentum has been completely algebraic in nature. That is, we have not considered wave functions as you did earlier in the course. A natural "next step" in our study of angular momentum would be to consider angular momentum wave functions, i.e. spatial dependence of angular momentum. This topic is discussed extremely well in the "Littlejohn Notes". We will not pursue this topic here. However, for this "proof" we will need to consider an eigenfunction of L_3 (not J_3 since we are considering orbital angular momentum).

A point in space written in spherical coordinates has the form:

$$\mathbf{x} = (r\sin\theta\cos\phi, r\sin\theta\sin\phi, r\cos\theta) \tag{4.68}$$

and using the chain rule we have:

$$\begin{aligned}\frac{\partial}{\partial\phi} &= \frac{\partial x_1}{\partial\phi}\frac{\partial}{\partial x_1} + \frac{\partial x_2}{\partial\phi}\frac{\partial}{\partial x_2} + \frac{\partial x_3}{\partial\phi}\frac{\partial}{\partial x_3}, \\ &= -x_2\frac{\partial}{\partial x_1} + x_1\frac{\partial}{\partial x_2}.\end{aligned} \tag{4.69}$$

From the last line of this expression (using $P_k = \frac{\hbar}{i}\frac{\partial}{\partial x_k}$, and associating the position operator X_k with the position coordinate x_k, $k = 1, 2$), we see that:

$$-i\hbar\frac{\partial}{\partial\phi} = X_1P_2 - P_1X_2 = L_3 \tag{4.70}$$

Now let $\psi_m(r, \theta, \phi)$ denote an eigenfunction of L_3. Then we have:

$$-i\hbar\frac{\partial\psi_m}{\partial\phi} = L_3\psi_m = m\hbar\psi_m \tag{4.71}$$

or

$$\frac{\partial\psi_m}{\partial\phi} = im\psi_m. \tag{4.72}$$

Integrating this expression from 0 to ϕ (leaving r and θ fixed) gives:

$$\psi_m(r, \theta, \phi) = e^{im\phi}\psi_m(r, \theta, 0). \tag{4.73}$$

Physically, we require the wavefunction to be a single valued function of the coordinates, i.e. at a given point in space, the wavefunction assumes only one value. In other words, we require:

$$\begin{aligned}\psi_m(r, \theta, \phi) &= \psi_m(r, \theta, \phi + 2\pi) = e^{im(\phi+2\pi)}\psi_m(r, \theta, 0), \\ &= e^{2\pi im}\left(e^{im\phi}\psi_m(r, \theta, 0)\right) = e^{2\pi im}\psi_m(r, \theta, \phi).\end{aligned} \tag{4.74}$$

Hence, the only way that:

$$\psi_m(r,\theta,\phi) = e^{2\pi i m}\psi_m(r,\theta,\phi), \tag{4.75}$$

is if m is an integer, which implies that j must be an integer (why?).

□

Now we need to explain why we put quotation marks around the word "proof". The result is certainly true, and there are no errors in our computations. The issue comes with the the statement "Physically, we require the wavefunction to be a single valued function of the coordinates". Why should this be true? The wavefunction is not a "physical field", such as pressure or density, where the value of such a field corresponds to a "physical observable, i.e. there is a unique value of the observable at each point in space. Recall our interpretation of the wavefunction as a *probability density*. Moreover, the wavefunction is "arbitrary" up to a phase factor, e.g a wavefunction ψ should have the same physical implications as $-\psi$. Clearly, a deeper analysis of this notion of "single valuedness" of the wave is required[2]. A resolution comes in a deeper understanding of the nature of the eigenfunctions of the Schroedinger equation (e.g. the Hilbert space of its solutions), which is beyond the scope of this course. However, a proof that the eigenvalues of the orbital angular momentum operator must take on integer values can be found in the paper:

[2] The book by Schumacher and Westmoreland has a nice section call "How not to think about ψ", (section 10.6) which I recommend.

D. M. Kaplan and F. Y. Wu [1971] On the eigenvalues of orbital angular momentum. *Chinese Journal of Physics*, **9**(1), 31-33.

The proof is surprisingly simple (at the level of the mathematics that we have already done in the course) and it uses the algebraic approach of operators that we have been developing, rather than wave mechanics, which was the focus of the first part of the course. However, as we have mentioned, a natural further development of angular momentum would be to develop angular momentum eigenfunctions that are appropriate solutions of the Schroedinger equation. This is pursued in the "Littlejohn Notes" as well as in Cohen-Tannoudji et al.

Spin, and the Stern-Gerlach Experiment

We have mentioned that there are types of angular momentum other than orbital angular momentum, and *spin* is probably the preeminent example. Despite its name, spin is an intrinsically quantum mechanical property with no classical analog. It was definitively demonstrated in the 1922 experiment of Otto Stern and Walther Gerlach (the "Stern-Gerlach experiment"). The background and history are well described in the following article:

B. Friedrich, D. Herschbach [2003] Stern and Gerlach: How a Bad Cigar Helped Reorient Atomic Physics. *Physics Today*, December 2003, 53-59.

The following two quotes taken from this article describe its importance in the history of physics, as well as its continuing broad impact.

The demonstration of space quantization, carried out in Frankfurt, Germany, in 1922 by Otto Stern and Walther Gerlach, ranks among the dozen or so canonical experiments that ushered in the heroic age of quantum physics.

Descendants of the Stern-Gerlach experiment (SGE) and its key concept of sorting quantum states via space quantization are legion. Among them are the prototypes for nuclear magnetic resonance, optical pumping, the laser, and atomic clocks, as well as incisive discoveries such as the Lamb shift and the anomalous increment in the magnetic moment of the electron, which launched quantum electrodynamics The means to probe nuclei, proteins, and galaxies; image bodies and brains; perform eye surgery, read music or data from compact discs; and scan bar codes on grocery packages or DNA base pairs in the human genome all stem from exploiting transitions between space-quantized quantum states.

For the rest of this course the general framework of the Stern-Gerlach experiment will play a a central role in several ways. Most notably, in the topic of "measurement" in quantum mechanics. Now we want to explain the essentials of this experiment (which is no substitute for a detailed study of the underlying physics–a good discussion can be found in Cohen-Tannoudji et al.).

Stern and Gerlach sent a beam of neutral atoms through a region of an inhomogeneous magnetic field. This means that the magnetic field had a nonzero spatial gradient, which was chosen to be the z direction. Stern and Gerlach used silver atoms in their experiment (Ag; Z=47). This was a very deliberate choice. We will briefly explain the reasons behind this choice (you don't need to know these details–the main point is at the end of our discussion). In the ground state a silver atom has one *valence electron* in a 5s subshell. The other 46 electrons fill all subshells for n=1, n=2, and n=3, and the 4d subshell. *All these closed shells and subshells contribute zero to the total angular momentum of the atom. So the properties measured in the Stern-Gerlach experiment are the properties of the valence electron*. This is very significant. In effect, for the properties under consideration, the Stern-Gerlach experiment is an experiment on single electrons, and the angular momentum of the electron in the 5s subshell corresponds to $l = 0$.

The following is "what happens" When a silver atom passes through the magnetic field it experiences a deflection (as a result of the magnetic field gradient) depending on the angular momentum of the particle. If the silver atom behaved as a classical particle, it is expected that the angular momentum vectors of the beam of particles would

be randomly distributed, and the beam would strike a detector in a "broad spot" centered around the axis of the beam (see Figure 4 for an illustration). This was not observed. The silver atoms did not behave as classical particles. The silver atoms had total (quantum) orbital angular momentum zero (with the electrons occupying the various shells and subhells as described above). Nevertheless, the beam struck the detector in two discrete locations, and it was a big mystery as to "why". (see Figure 4 for an illustration of what was observed). It could be argued that deflection of the beam of silver atoms due to the force generated by the inhomogeneous field was due to the interaction of the magnetic field with the "angular momentum" of the silver atoms, which was completely due to the angular momentum of the valence electron, which corresponded to $l = 0$ for the electron in the 5s subshell. The distribution of silver atoms on the detector could not be explained by classical angular momentum, and it could not be explained by quantum orbital angular momentum. So what was going on?

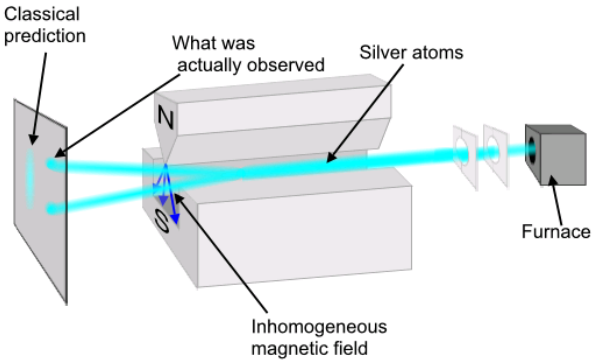


Figure 4.1: Schematic of the Stern-Gerlach experiment (figure from Wikipedia).

The Physics Today paper of Friedrich and Herschbach describes the efforts to come up with an explanation. Goudsmit and Uhlenbeck finally proposed the idea of an "intrinsic angular momentum" (or "spin") that provided an explanation of the results.

If you would like to read Goudsmit's on personal account of the discovery of electron spin go to the following URL:

http://www.lorentz.leidenuniv.nl/history/spin/goudsmit.html

A very complete account of the topic of spin is given here (and it is also not too difficult to read):

S. Tomonaga, T. Oka [1998] The story of Spin. University of Chicago Press

If you want to perform your own Stern-Gerlach experiment there is

a nice Java applet that will allow you to do that here:

http://www.if.ufrgs.br/~betz/quantum/SGPeng.htm

For the purpose of the rest of the course, the important facts are that the spin of an electron is described by a "two state quantum mechanical system"–spin "up" and spin "down". where up and down are defined by the direction of the gradient of the magnetic field. In this case the gradient is in the z direction. We take spin up as the positive z direction and spin down as the negative z direction (now we are in danger of beginning to encounter some conceptual difficulties by "confusing" the physical space with the abstract Hilbert space describing the "spin dynamics" of the electron–but we will address this later on). The "Stern-Gerlach device" sorts the beam of silver atoms into these two quantum "spin states". Next we want to describe such two state quantum systems in more detail, and that will take us into the final part of the course.

Problems

1. *Orbital angular momentum commutation relations.*

   Show that $[AB, C] = A[B, C] + [A, C]B$ for three operators A, B, C. Work out a similar expression for $[AB, CD]$. Hence show that the orbital angular momentum operators L_1, L_2, L_3, where $L_1 = X_2P_3 - X_3P_2$, $L_2 = X_3P_1 - X_1P_3$ and $L_3 = X_1P_2 - X_2P_1$, satisfy the angular momentum commutation relations

$$[L_j, L_k] = i\hbar \sum_{m=1}^{3} \epsilon_{jkm} L_m.$$

   (X_j and P_k satisfy the usual canonical commutation relations $[X_j, P_k] = i\hbar\delta_{jk}$.)

2. *Raising and lowering operators.*

   Consider three operators J_1, J_2, J_3 satisfying the angular momentum commutation relations. With the definitions $J_\pm = J_1 \pm iJ_2$ and $J^2 = J_1^2 + J_2^2 + J_3^2$, show that

$$\begin{aligned} [J^2, J_\pm] &= 0; \\ J_+J_- &= J^2 - J_3^2 + \hbar J_3; \\ J_-J_+ &= J^2 - J_3^2 - \hbar J_3; \\ [J_+, J_-] &= 2\hbar J_3; \\ [J_3, J_\pm] &= \pm\hbar J_\pm. \end{aligned}$$

3. *Pauli matrices.*

   The Pauli matrices are defined by

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}; \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}; \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

   (a) Show that

$$\sigma_x^2 = I; \quad \sigma_y^2 = I; \quad \sigma_z\sigma_x = i\sigma_y; \quad \sigma_z\sigma_y = -i\sigma_x;$$

   where I is the identity matrix on $\mathbb{C}^2$

   (b) Consider two 3-component vectors $\mathbf{A}$ and $\mathbf{B}$ with real entries. Let $\mathbf{A}.\sigma$ and $\mathbf{B}.\sigma$ be the matrices

$$\mathbf{A}.\sigma = A_x\sigma_x + A_y\sigma_y + A_z\sigma_z; \quad \mathbf{B}.\sigma \quad = \quad B_x\sigma_x + B_y\sigma_y + B_z\sigma_z.$$

Write down the matrices $\mathbf{A}.\sigma$ and $\mathbf{B}.\sigma$ explicitly, and hence, or otherwise show that

$$(\mathbf{A}.\sigma)\ (\mathbf{B}.\sigma) \quad = \quad (\mathbf{A}.\mathbf{B})\ I + i(\mathbf{A}\times\mathbf{B}).\sigma.$$

(c) Let $\mathbf{n}$ be a three component unit vector and let a be a real parameter. Show that

$$\exp\left(ia\mathbf{n}.\sigma\right) = I\cos a + i\mathbf{n}.\sigma\sin a$$

4. *The spin one representation.*

For the $j = 1$ representation of the angular momentum commutation relations, write down the action of the operators $J_\pm$ and J_3 on basis vectors which are simultaneous eigenvectors of J^2 and J_3.

Calculate the action of J_1 and J_2 in this case and hence show that the operators J_1, J_2, J_3 satisfy the angular momentum commutation relations $[J_j, J_k] = i\hbar\sum_{m=1}^{3}\epsilon_{jkm}J_m$.

Calculate the matrices of J_1, J_2, J_3 with respect to the basis you have used above, and show that these matrices also satisfy the angular momentum commutation relations.

5. *Matrix elements for angular momentum operators.*

Let J_1, J_2, J_3 satisfy angular momentum commutation relations, and let $|j,m\rangle$ be the usual (normalised) simultaneous eigenvectors of J^2 and J_3 with eigenvalues $\hbar^2 j(j+1)$ and $m\hbar$ respectively.

By considering $\langle j,m|\, J_+^2\,|j,m\rangle$ or otherwise, show that

$$\langle j,m|\, J_1^2\,|j,m\rangle = \langle j,m|\, J_2^2\,|j,m\rangle\,,$$

and find the value of this matrix element.

6. *A spin Hamiltonian.*

Consider the Hamiltonian

$$H = \frac{1}{2I}(J_1^2 + J_3^2)\ ,$$

where I is a constant.

(a) Show that

$$\langle 1, m|H|1, m\rangle = \frac{1}{4I}(2 + m^2)\hbar^2 \ ,$$

for $m = -1, 0, 1$.

(b) Show also that for $m \neq n$, $\langle 1, m|H|1, n\rangle = 0$ unless $(m, n) = (1, -1)$ or $(m, n) = (-1, 1)$. Calculate $\langle 1, 1|H|1, -1\rangle$.

(c) Deduce the three eigenvalues of H and the corresponding eigenstates.

7. Let J_1, J_2, J_3 be self-adjoint angular momentum operators obeying:

$$[J_1, J_2] = i\hbar J_3, \quad [J_2, J_3] = i\hbar J_1, \quad [J_3, J_1] = i\hbar J_2,$$

with the operators $\mathbf{J}^2$, J_+ and J_- defined by $\mathbf{J}^2 = J_1^2 + J_2^2 + J_3^2$ and $J_\pm = J_1 \pm iJ_2$. We denote a simultaneous eigenvector of $\mathbf{J}^2$ and J_3 by $\mid jm\rangle$, i.e.

$$\begin{aligned} \mathbf{J}^2 \mid jm\rangle &= \hbar^2 j(j+1) \mid jm\rangle, \\ J_3 \mid jm\rangle &= \hbar m \mid jm\rangle. \end{aligned}$$

Consider the Hamiltonian:

$$H = \mathbf{J^2} - J_3^2. \tag{4.76}$$

(a) Show that H is self-adjoint.

(b) Show that

$$\mathbf{J}^2 = J_+J_- + J_3^2 - \hbar J_3 = J_-J_+ + J_3^2 + \hbar J_3.$$

(c) Compute the eigenvalues of H.

(d) Are $\mid jm\rangle$ eigenvectors for H *and* $\mathbf{J}^2$? Justify your answer.

(e) Suppose the state space for H is C^2, with basis vectors $\left\{\mid \frac{1}{2}, \frac{1}{2}\rangle, \mid \frac{1}{2}, -\frac{1}{2}\rangle\right\}$. Compute the matrix representation of H with respect to this basis.

8. Let $\mathbf{J} = (J_1, J_2, J_3)$ where J_1, J_2, J_3 are self-adjoint angular momentum operators obeying:

$$[J_1, J_2] = i\hbar J_3, \quad [J_2, J_3] = i\hbar J_1, \quad [J_3, J_1] = i\hbar J_2,$$

with the operators $\mathbf{J}^2$, J_+ and J_- defined by $\mathbf{J}^2 = J_1^2 + J_2^2 + J_3^2$ and $J_\pm = J_1 \pm iJ_2$. We denote a simultaneous eigenvector of $\mathbf{J}^2$ and J_3 by $\mid jm\rangle$, i.e.

$$\begin{aligned} \mathbf{J}^2 \mid jm\rangle &= \hbar^2 j(j+1) \mid jm\rangle, \\ J_3 \mid jm\rangle &= \hbar m \mid jm\rangle. \end{aligned}$$

Consider the Hamiltonian:

$$H = \frac{1}{2I_1}\left(J_1^2 + J_2^2\right) + \frac{1}{2I_3}J_3^2,$$

where I_1 and I_3 are positive constants.

(a) Show that $\mathbf{J} \times \mathbf{J} = i\hbar\mathbf{J}$.

(b) Show that

$$\mathbf{J}^2 = J_+J_- + J_3^2 - \hbar J_3 = J_-J_+ + J_3^2 + \hbar J_3.$$

(c) Using the simultaneous eigenvectors of $\mathbf{J}^2$ and J_3, compute the corresponding eigenvalues of H.

(d) Suppose the state space for H is $\mathbb{C}^2$, with basis vectors $\left\{ \mid \frac{1}{2}, \frac{1}{2}\rangle, \mid \frac{1}{2}, -\frac{1}{2}\rangle \right\}$. Compute the matrix representation of H with respect to this basis.

(e) Compute the expectation value of J_3^2 in the state $\mid \frac{1}{2}, -\frac{1}{2}\rangle$.

5
The Postulates of Quantum Mechanics, Measurement, Composite Systems, Tensor Products, and Entanglement

In this part of the course we will study some of the aspects of quantum mechanics that are most at odds with our intuition from classical mechanics. These topics are of intense current interest and, potentially, will be at the heart of significant technological advances (the phrase "quantum engineering" is becoming more and more common). Consequently, there is intense research activity in these areas.

We begin by summarizing the "Postulates of Quantum Mechanics". For a more detailed discussion of the postulates see "Littlejohn's Notes 3".

The Postulates of Quantum Mechanics

States: Physical systems are associated with a Hilbert space. This Hilbert space is typically called the *state space*. In Dirac's notation, the vectors in this Hilbert space are called *kets*. A state is a *ray in Hilbert space*[1] Now you are probably asking, 'why a ray (and "what is a ray?") and not just a vector in the Hilbert space?" Recall that quantum states are arbitrary up to a normalization (we tend to work with vectors of length one) and a phase (which does not effect the "physics", i.e. expectation values and the modulus squared of wave functions). With this motivation, a ray in a Hilbert space is an equivalence class of vectors where every vector in the equivalence class differs by multiplication by a nonzero complex number.

[1] You will sometimes hear this referred to as a *pure state*–which begs the question of "what kind of state is not a pure state?". If you pursue more study of quantum mechanics you will undoubtedly some across the answer to this question.

Observables: An observable is a property of a physical system that can be measured (at least in principle). In quantum mechanics an observable is represented by a Hermitian, or self-adjoint, operator. You know that Hermitian operators have many "nice" properties, e.g. their eigenvalues are real and they have a complete set of eigenvectors. What is particularly relevant to us now is the *spec-*

tral decomposition of a Hermitian operator. We will restrict ourselves to the situation where the spectrum is discrete (since that is all that we will need for our purposes).

Consider a Hermitian operator A acting on a Hilbert space and let $\{\lambda_n\}$ denote its eigenvalues. Then A can be represented as:

$$A = \sum_n \lambda_n P_n, \tag{5.1}$$

where P_n is the orthogonal projection onto the space of eigenvectors with eigenvalue λ_n. The operators P_n satisfy:

$$\begin{aligned} P_n P_m &= \delta_{n,m} P_n, \\ P_n^\dagger &= P_n. \end{aligned} \tag{5.2}$$

If there are no degeneracies, then for every eigenvalue, λ_n, there is a unique eigenvector, ψ_n, and the projection operators are given by:

$$P_n = \mid \psi_n \rangle \langle \psi_n \mid . \tag{5.3}$$

Measurement: The main points in the measurement of observables in quantum mechanics are the following.

1. The numerical outcome of the measurement of an observable A is an eigenvalue of A.If the quantum state just prior to the measurement was $\mid \psi \rangle$, then the probability that the outcome of the measurement is λ_n is given by:

$$p(\lambda_n) = \parallel P_n \mid \psi \rangle \parallel^2 = \langle \psi \mid P_n \mid \psi \rangle. \tag{5.4}$$

2. Immediately after the measurement, the state of the quantum system is the projection of the original state into the space spanned by the eigenvectors corresponding to the eigenvalue that was just measured. In the case of no degeneracies, the state of the system after measurement is the eigenvector of A corresponding to the eigenvalue that was just measured.

   More precisely, if, as a result of the measurement of A on $\mid \psi \rangle$, the value λ_n is attained. Then immediately after the measurement the (normalized) quantum state becomes:

$$\frac{P_n \mid \psi \rangle}{(\langle \psi \mid P_n \mid \psi \rangle)^{\frac{1}{2}}} \tag{5.5}$$

Dynamics: Time evolution of a quantum state is governed by the Schrödinger equation.

The quantum theory of measurement seems very counterintuitive (at least, to someone whose intuition is developed almost entirely through classical mechanics). It is probabilistic in nature (given an observable, we can compute the probabilities for the different outcomes of measuring the observable). Probably the most strange thing about it is that once a measurement is made, the wavefunction of the system "collapses" to the eigenstate of the eigenvalue corresponding to the outcome of the measurement (and the phrase used for this is "collapse of the wavefunction"). Thus the very act of measuring a quantum system "changes it". Throughout the rest of this course we will focus almost exclusively on issues associated with "quantum measurement".

The Stern-Gerlach Experiment–Revisited

We return to the Stern-Gerlach experiment in order to study more deeply "quantum measurement". In particular, we will consider quantum particles having spin $\frac{1}{2}$, i.e. particles whose behavior is described by quantum mechanics, such as electrons.

The Pauli Spin Matrices and Dirac Notation–Review of Some Background

We begin by first developing some necessary mathematical tools.

The Complex Vector Space, $\mathbb{C}^2$

When a beam of electrons[2] is passed through a Stern-Gerlach apparatus the beam is split into two–half correspond to electrons whose spin is in the direction of the magnetic field gradient (spin $\frac{1}{2}$) and the other correspond to electrons whose spin is in the opposite direction of the magnetic field gradient (spin $-\frac{1}{2}$) . If we characterize the electrons solely by their spin, then this system provides an ideal example of a "two state system". The Hilbert space is two dimensional, and can be taken as $\mathbb{C}^2$, which we view as a (complex) linear vector space equipped with an inner product.

[2] Remember, Stern and Gerlach did not use a beam of electrons, but a beam of silver atoms. However, recall the discussion. The silver atoms were such that the measure properties were those of the single valence electron. We will abuse history, slightly, by referring to a "beam of electrons".

The Pauli Spin Matrices

Recall the Pauli spin matrices that we derived earlier:

$$J_1 = \frac{\hbar}{2}X, \quad J_2 = \frac{\hbar}{2}Y, \quad J_3 = \frac{\hbar}{2}Z, \tag{5.6}$$

where

$$X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (5.7)$$

We can now state the problem that we will address.

Statement of Problem. *We imagine a beam of spin $\frac{1}{2}$ quantum particles propagating in the y direction and passing through certain Stern-Gerlach devices. We will be considering measurements of J_3 and J_1. The relevant Hilbert space is $\mathbb{C}^2$, and J_3 and J_1 are Hermitian operators defined on this Hilbert space.*

We first need to develop some mathematical properties of J_3 and J_1 that will be necessary in applying the postulates of quantum mechanics for the purpose of measurement.

$$\boxed{J_3 = \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \frac{\hbar}{2} Z}$$

The eigenvalues of Z are 1, with corresponding eigenvector $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$, and -1, with corresponding eigenvector $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$. We will use Dirac's bra-ket notation in our calculations. Different notations are used in the literature for the ket vector corresponding to $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$. Some of these are collected below.

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} : \quad |\uparrow\rangle, |\, 0\rangle, |\, +\frac{1}{2}\rangle, |\, +z\rangle. \quad (5.8)$$

Similarly, notations for the ket vector corresponding to $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ are:

$$\begin{pmatrix} 0 \\ 1 \end{pmatrix} : \quad |\downarrow\rangle, |\, 1\rangle, |\, -\frac{1}{2}\rangle, |\, -z\rangle. \quad (5.9)$$

We will use the notation $|\uparrow\rangle$ for $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $|\downarrow\rangle$ for $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ in this section.
The projection operator onto the eigenvector $|\uparrow\rangle$ is given by:

$$|\uparrow\rangle\langle\uparrow| = \begin{pmatrix} 1 \\ 0 \end{pmatrix} (1 \;\; 0) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}. \quad (5.10)$$

Similarly, the projection operator onto the eigenvector $|\downarrow\rangle$ is given by:

$$|\downarrow\rangle\langle\downarrow| = \begin{pmatrix} 0 \\ 1 \end{pmatrix} (0 \;\; 1) = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}. \tag{5.11}$$

Consequently, it is easy to see that the spectral decomposition for J_3 is given by:

$$J_3 = \frac{\hbar}{2} |\uparrow\rangle\langle\uparrow| - \frac{\hbar}{2} |\downarrow\rangle\langle\downarrow| = \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{5.12}$$

$$\boxed{J_1 = \frac{\hbar}{2} \begin{pmatrix} 0 & 1 \\ & \\ 1 & 0 \end{pmatrix} = \frac{\hbar}{2} X}$$

The eigenvalues of X are 1, with corresponding eigenvector $\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, and -1, with corresponding eigenvector $\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$. Some notation for kets corresponding to $\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$, respectively, are given below:

$$\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} : \quad |\rightarrow\rangle, \, |+\rangle, \, |+x\rangle, \tag{5.13}$$

$$\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} : \quad |\leftarrow\rangle, \, |-\rangle, \, |-x\rangle. \tag{5.14}$$

The corresponding projection operators onto the respective eigenvectors are given by:

$$|\rightarrow\rangle\langle\rightarrow| = \frac{1}{2} \begin{pmatrix} 1 \\ 1 \end{pmatrix} (1 \;\; 1) = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \tag{5.15}$$

$$|\leftarrow\rangle\langle\leftarrow| = \frac{1}{2} \begin{pmatrix} 1 \\ -1 \end{pmatrix} (1 \;\; -1) = \frac{1}{2} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}, \tag{5.16}$$

and the spectral decomposition of J_1 has the following form:

$$J_1 = \frac{\hbar}{2} |\rightarrow\rangle\langle\rightarrow| - \frac{\hbar}{2} |\leftarrow\rangle\langle\leftarrow| = \frac{\hbar}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{5.17}$$

Change of Basis. We will need to relate the basis of eigenvectors of J_3 to the basis of eigenvectors of J_1:

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \leftrightarrow \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}. \tag{5.18}$$

The relations between the two bases of eigenvectors can be determined essentially by inspection:

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{1}{\sqrt{2}}\left(\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix} + \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}\right) \tag{5.19}$$

$$\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \frac{1}{\sqrt{2}}\left(\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix} - \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}\right) \tag{5.20}$$

or, in bra-ket notation:

$$|\uparrow\rangle = \frac{1}{\sqrt{2}}|\rightarrow\rangle + \frac{1}{\sqrt{2}}|\leftarrow\rangle \tag{5.21}$$

$$|\downarrow\rangle = \frac{1}{\sqrt{2}}|\rightarrow\rangle - \frac{1}{\sqrt{2}}|\leftarrow\rangle, \tag{5.22}$$

and

$$\begin{aligned} |\rightarrow\rangle &= \frac{1}{\sqrt{2}}|\uparrow\rangle + \frac{1}{\sqrt{2}}|\downarrow\rangle \\ |\leftarrow\rangle &= \frac{1}{\sqrt{2}}|\uparrow\rangle - \frac{1}{\sqrt{2}}|\downarrow\rangle. \end{aligned} \tag{5.23}$$

Four "Thought Experiments" on the Measurement of Spin

Now we will consider four "thought experiments" [3] that illustrate measuring an observable in a quantum system. The observables will be J_3 and J_1 (recall that we have proven that these are Hermitian operators) and the Hilbert space defining the system is $\mathbb{C}^2$. In the "thought experiments" by the phrase "Stern-Gerlach apparatus" we will mean an apparatus having an inhomogeneous magnetic field whose gradient is in a specific direction. For our purposes the magnetic field gradient will either be in the vertical (z) direction or the horizontal (x direction), and these directions are "local" to the particular Stern-Gerlach apparatus. A beam of spin $\frac{1}{2}$ particles (e.g. electrons) propagates in the y direction (relative to the Stern-Gerlack apparatus) and passes through the inhomogeneous magnetic field associated with the Stern-Gerlach apparatuses in each experiment.

Experiment 1:

This is the "standard" Stern-Gerlach experiment. A beam of spin $\frac{1}{2}$ particles passes through a Stern-Gerlach apparatus with the inhomogeneous magnetic field gradient in the z direction, see Fig. 5. We assume that the initial state of the beam of spin $\frac{1}{2}$ particles is given by:

$$|\psi\rangle = \alpha|\uparrow\rangle + \beta|\downarrow\rangle, \tag{5.24}$$

[3] The state vector $|\psi\rangle$ is a description of the beam of spin $\frac{1}{2}$ quantum particles. It is not the state vector for each individual particle in the beam (you might think about why I say this). The real question would be how does one realize a beam in this state in a "real experiment"? This is why we are using the phrase "thought experiment". It is mainly to give experience with the postulates of quantum mechanics, and quantum measurement. "Thought experiments" have played (and still play) an important role in quantum mechanics. We will see that later when we consider the Einstein-Podolsky-Rosen (EPR) paradox. It took quite a few years before the conditions of the "thought experiment" of EPR could be realized in the laboratory. However, the consequences of the thought experiment were so compelling that the need to design a "real experiment" was clear.

where normalization of $|\psi\rangle$ implies that $|\alpha|^2 + |\beta|^2 = 1$.

We wish to measure J_3. The possible values of J_3 that we can measure are its eigenvalues, $\pm\frac{\hbar}{2}$. Recalling (5.4), the probability that we measure $\frac{\hbar}{2}$ is given by:

$$\begin{aligned} p\left(\frac{\hbar}{2}\right) &= \langle\psi \mid (|\uparrow\rangle\langle\uparrow|) \mid \psi\rangle, \\ &= (\alpha^*\langle\uparrow| + \beta^*\langle\downarrow|)(|\uparrow\rangle\langle\uparrow|)(\alpha|\uparrow\rangle + \beta|\downarrow\rangle), \\ &= |\alpha|^2 . \end{aligned} \tag{5.25}$$

A similar calculation shows that the probability that we measure $-\frac{\hbar}{2}$ is given by:

$$p\left(-\frac{\hbar}{2}\right) = |\beta|^2 . \tag{5.26}$$

So, as expected, the Stern-Gerlach apparatus splits the beam into two beams–one beam consisting of "spin up" particles and the other beam consisting of "spin down" particles (where "up" and "down" are measured with respect to the vertical (z) direction).

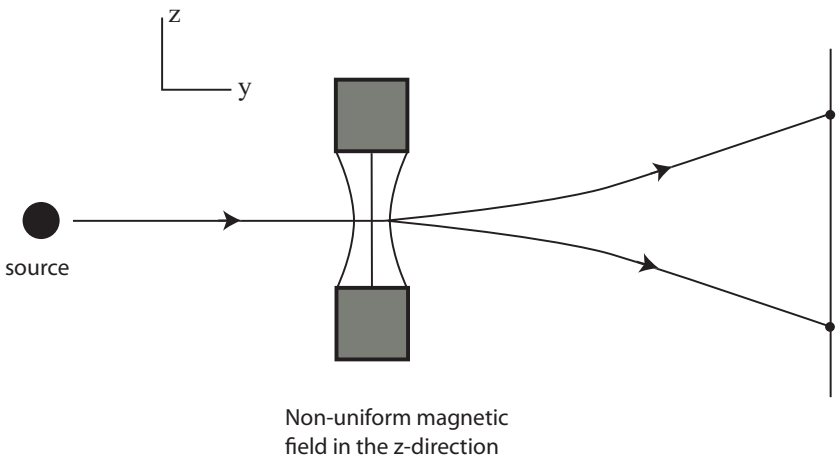


Figure 5.1: A beam of spin $\frac{1}{2}$ particles passing through a Stern-Gerlach apparatus, where the gradient of the magnetic field is in the z direction.

Experiment 2:

In this experiment the beam of spin $\frac{1}{2}$ particles passes through two consecutive Stern-Gerlach apparatuses, with each having their magnetic field gradient oriented in the z direction. As in the previous experiment, the initial state (before the beam passes through either Stern-Gerlach apparatus) is given by:

$$|\psi\rangle = \alpha|\uparrow\rangle + \beta|\downarrow\rangle, \quad |\alpha|^2 + |\beta|^2 = 1. \tag{5.27}$$

We assume that a measurement has been made immediately after the beam passes through the first Stern-Gerlach apparatus and only particles with spin J_3 eigenvalues of $\frac{\hbar}{2}$ are allowed to continue to the second Stern-Gerlach apparatus. Recalling (5.5), the state following the measurement that gives the eigenvalue $\frac{\hbar}{2}$ is given by:

$$\frac{|\uparrow\rangle\langle\uparrow|\,(\alpha\,|\uparrow\rangle+\beta\,|\downarrow\rangle)}{\sqrt{\langle\psi\mid(|\uparrow\rangle\langle\uparrow|)\mid\psi\rangle}}=|\uparrow\rangle \tag{5.28}$$

This is consistent with the measurement postulate of quantum mechanics. The eigenvalue $\frac{\hbar}{2}$ is measured and the state "collapses" to the eigenstate of corresponding to the eigenvalue $\frac{\hbar}{2}$. The beam of particles in state $|\uparrow\rangle$ then passes through the second Stern-Gerlach apparatus. Another measurement of J_3 on is state is carried out, and the probabilities that $\frac{\hbar}{2}$ and $-\frac{\hbar}{2}$ are measured are, respectively, given by:

$$\begin{aligned} p\left(\frac{\hbar}{2}\right) &= \langle\uparrow|\,(|\uparrow\rangle\langle\uparrow|)\,|\uparrow\rangle=1, \\ p\left(-\frac{\hbar}{2}\right) &= \langle\uparrow|\,(|\downarrow\rangle\langle\downarrow|)\,|\uparrow\rangle=0. \end{aligned} \tag{5.29}$$

Hence, we always get the outcome $\frac{\hbar}{2}$, as expected, since the state of the beam that enters the second Stern-Gerlach device is $|\uparrow\rangle$.

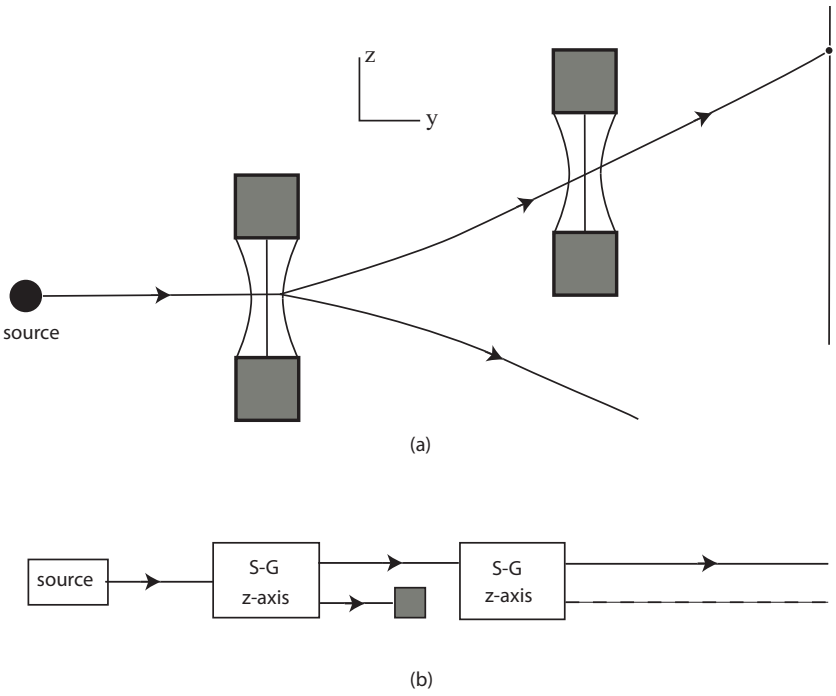


Figure 5.2: A beam of spin $\frac{1}{2}$ particles passes through two consecutive Stern-Gerlach devices, with each measuring J_3. a) After passing through the first Stern-Gerlach device only particles in the state $|\uparrow\rangle$ are allowed to pass through the second Stern-Gerlach device. b) A more "schematic" depiction of the experiment that is shown in a).

Experiment 3:

We consider a beam of spin $\frac{1}{2}$ particles in the initial state:

$$\psi = \alpha \left|\uparrow\right\rangle + \beta \left|\downarrow\right\rangle, \quad |\,\alpha\,|^2 + |\,\beta\,|^2 = 1. \tag{5.30}$$

The beam passes through a Stern-Gerlach apparatus where J_3 is measured, and particles where the outcome of the measurement is $\frac{\hbar}{2}$ are allowed to pass through a second Stern-Gerlach device where the magnetic field gradient is in the x direction. The particles passing through the second Stern-Gerlach device are in the state $\left|\uparrow\right\rangle$, and we wish to measure J_1. We know that the possible outcomes of this measurement are $\frac{\hbar}{2}$ and $-\frac{\hbar}{2}$ (the two eigenvalues of J_1). The probability that the outcome of the measurement is $\frac{\hbar}{2}$ is given by:

$$p\left(\frac{\hbar}{2}\right) = \langle\uparrow| \left(|\rightarrow\rangle\langle\rightarrow|\right) |\uparrow\rangle. \tag{5.31}$$

Now in order to calculate this quantity we need to express the state $\left|\uparrow\right\rangle$ in terms of the basis of J_1 given by $\left|\rightarrow\right\rangle$ and $\left|\leftarrow\right\rangle$. The formula relating the basis elements was previously given in (5.22), which we rewrite here:

$$\left|\uparrow\right\rangle = \frac{1}{\sqrt{2}} \left|\rightarrow\right\rangle + \frac{1}{\sqrt{2}} \left|\leftarrow\right\rangle. \tag{5.32}$$

Substituting this expression in (5.31) gives:

$$p\left(\frac{\hbar}{2}\right) = \left(\frac{1}{\sqrt{2}}\langle\rightarrow| + \frac{1}{\sqrt{2}}\langle\leftarrow|\right) \left(|\rightarrow\rangle\langle\rightarrow|\right) \left(\frac{1}{\sqrt{2}} \left|\rightarrow\right\rangle + \frac{1}{\sqrt{2}} \left|\leftarrow\right\rangle\right) = \frac{1}{2}. \tag{5.33}$$

Similarly, the probability that the outcome of the measurement is $-\frac{\hbar}{2}$ is given by:

$$p\left(-\frac{\hbar}{2}\right) = \langle\uparrow| \left(|\leftarrow\rangle\langle\leftarrow|\right) |\uparrow\rangle \tag{5.34}$$

Substituting the change of basis expression (5.32) into this expression gives:

$$p\left(-\frac{\hbar}{2}\right) = \left(\frac{1}{\sqrt{2}}\langle\rightarrow| + \frac{1}{\sqrt{2}}\langle\leftarrow|\right) \left(|\leftarrow\rangle\langle\leftarrow|\right) \left(\frac{1}{\sqrt{2}} \left|\rightarrow\right\rangle + \frac{1}{\sqrt{2}} \left|\leftarrow\right\rangle\right) = \frac{1}{2}. \tag{5.35}$$

Hence, both outcomes of the measurement of J_1 occur with equal probability, and the beam is split into two again.

Experiment 4:

We consider a beam of spin $\frac{1}{2}$ particles in the initial state:

$$\psi = \alpha \left|\uparrow\right\rangle + \beta \left|\downarrow\right\rangle, \quad |\,\alpha\,|^2 + |\,\beta\,|^2 = 1. \tag{5.36}$$

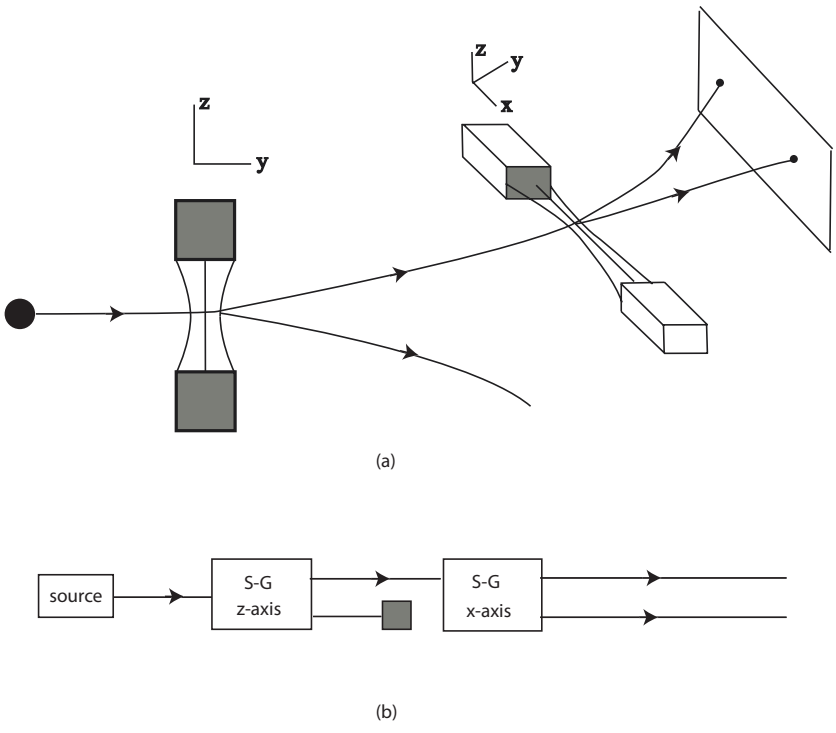


Figure 5.3: A beam of spin $\frac{1}{2}$ particles pass through two consecutive Stern-Gerlach devices. a) The first Stern-Gerlach device measures J_3, and after passing through the first Stern-Gerlach device only particles in the state $|\uparrow\rangle$ are allowed to pass through the second Stern-Gerlach device. The second Stern-Gerlach device measures J_1. b) A more "schematic" depiction of the experiment that is shown in a).

The beam passes through a Stern-Gerlach apparatus where J_3 is measured, and particles where the outcome of the measurement is $\frac{\hbar}{2}$ are allowed to pass through a second Stern-Gerlach device where the magnetic field gradient is in the x direction. Immediately before passing through the second Stern-Gerlach device the particles are in the state $|\uparrow\rangle$, and we wish to measure J_1. We saw from experiment 3 that these particles will be in state $|\rightarrow\rangle$ with probability $\frac{1}{2}$ and in state $|\leftarrow\rangle$ with probability $\frac{1}{2}$. The particles in the state $|\rightarrow\rangle$ (corresponding to a measurement outcome of $\frac{\hbar}{2}$) are allowed to continue to a third Stern-Gerlach apparatus where the gradient of the magnetic field is in the z direction. We measure J_3 when the particles have passed through this final Stern-Gerlach apparatus. The possible outcomes are the eigenvalues of J_3, $\pm\frac{\hbar}{2}$. The probability that the outcome of the measurement is $\frac{\hbar}{2}$ is given by:

$$p\left(\frac{\hbar}{2}\right) = \langle\rightarrow| \left(|\uparrow\rangle\langle\uparrow|\right) |\rightarrow\rangle \tag{5.37}$$

In order to calculate this quantity we need to express the state $|\rightarrow\rangle$ in terms of the basis of $\mathbb{C}^2$ given by the eigenvectors of J_3. Recall that this

was given in (5.23), which we rewrite here:

$$|\rightarrow\rangle = \frac{1}{\sqrt{2}}\,|\uparrow\rangle + \frac{1}{\sqrt{2}}\,|\downarrow\rangle \tag{5.38}$$

Then we have:

$$p\left(\frac{\hbar}{2}\right) = \left(\frac{1}{\sqrt{2}}\langle\uparrow| + \frac{1}{\sqrt{2}}\langle\downarrow|\right)(|\uparrow\rangle\langle\uparrow|)\left(\frac{1}{\sqrt{2}}\,|\uparrow\rangle + \frac{1}{\sqrt{2}}\,|\downarrow\rangle\right) = \frac{1}{2}. \tag{5.39}$$

The probability that the outcome of the measurement is $-\frac{\hbar}{2}$ is given by:

$$p\left(-\frac{\hbar}{2}\right) = \langle\rightarrow|\,(|\downarrow\rangle\langle\downarrow|)\,|\rightarrow\rangle \tag{5.40}$$

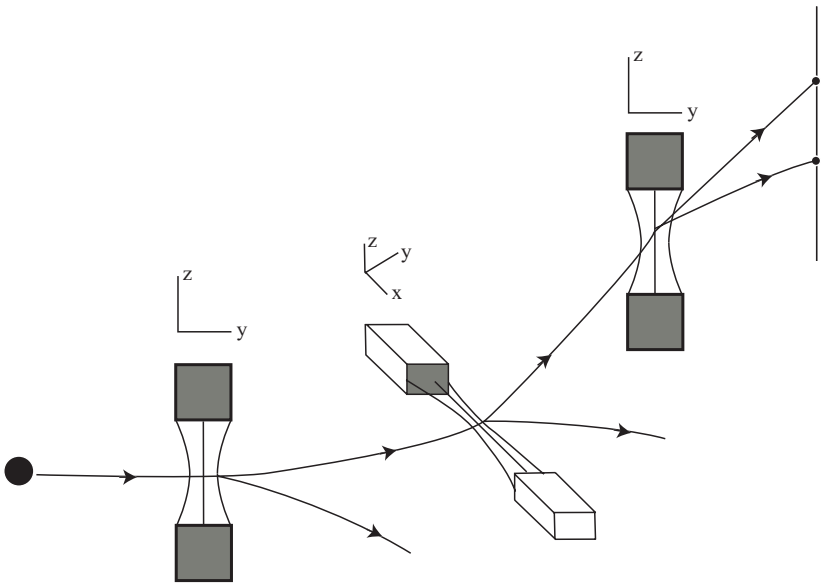
A similar calculation to that given above shows that:

$$p\left(-\frac{\hbar}{2}\right) = \frac{1}{2}.$$

Hence, the final Stern-Gerlach devices splits the beam into two, corresponding to "spin up" and "spin down", as measured in the vertical (z) direction. This is exactly the situation that occurred immediately after passing through the first Stern-Gerlach device, where we "threw away" the "spin down" part of the beam. Somehow, some "spin down" component of the beam has now been recovered.

Composite Systems, Tensor Products, and Entanglement

Motivation for considering the tensor product. To begin with, we consider classical mechanics. Imagine a point particle of mass m undergoing two dimensional motion as a result of a (net) force acting on the particle. In classical mechanics the phrase *configuration space* is used to describe the two coordinates that describe the position of the particle, which is two dimensional. Now consider n separate and identical (without loss of generality for the point we now wish to make) such systems that we view as a single *composite system*. The configuration space for composite systems is the *cartesian product* of the configuration spaces of each two dimensional system, and it is $2n$ dimensional. As we have described it thus far, the individual two dimensional systems do not interact with each other. However, we could consider them to be coupled in a variety of ways. The nature of the coupling (i.e. the coupling forces) can be determined by applying Newton's law's to determine the forces due to the coupling of the individual systems. These coupling forces can be expressed in terms of the configuration space coordinates of the cartesian product of the individual systems. So even if the individual two dimensional systems are coupled, the



(a)

(b)

Figure 5.4: A beam of spin $\frac{1}{2}$ particles pass through three consecutive Stern-Gerlach devices. a) The first Stern-Gerlach device measures J_3, and after passing through the first Stern-Gerlach device only particles in the state $|\uparrow\rangle$ are allowed to pass through the second Stern-Gerlach device. A measurement of J_1 is made on these particles, and only particles with outcome $\frac{\hbar}{2}$ are allowed to continue to the third Stern-Gerlach device. The third Stern-Gerlach device measures J_3. b) A more "schematic" depiction of the experiment that is shown in a).

configuration space for the n coupled systems is the cartesian product of the configuration spaces of the individual two dimensional systems, and it is $2n$ dimensional.

Now let's consider the quantum mechanical analog of this situation. We consider n quantum mechanical systems, each described by a two dimensional state space (Hilbert space). Now we consider the *composite system* of all n two dimensional systems. In quantum mechanics the state space describing the composite system is *not* the cartesian product of the two dimensional state spaces (which would have dimension $2n$), but the *tensor product* of the two dimensional state spaces, which has dimension 2^n. This is a somewhat bewildering difference between classical and quantum mechanics and arises as a result of our consideration of composite (or coupled, or interacting) systems. The first

point to take notice of is the difference in dimension for the composite system of n two dimensional classical mechanical systems versus the composite system of n two dimensional quantum mechanical systems. In classical mechanics the dimensions is $2n$ and in quantum mechanics the dimension is 2^n (more details of this will be provided when we consider tensor product spaces in detail below). Hence, as the number of two dimensional systems is increased, the dimension of the state space of the composite system grows linearly with n in classical mechanics, and exponentially with n in quantum mechanics. This is a HUGE difference between classical and quantum mechanics (and it is at the heart of why a quantum computer (assuming such a thing existed) could solve much more complex problems than a classical computer).

However, the real question is "why the difference?", i.e. why a cartesian product in classical mechanics and a tensor product in quantum mechanics? I don't have a terribly convincing reason for this, but I do believe that I have a reason that you might accept. In classical mechanics Newton's laws are our guide to writing down the "equations of motion" for a classical mechanical system. Consider the situation of n systems each exhibiting two-dimensional motion as described above. The solution of Newton's equations for the composite system is a time varying vector in the cartesian product of the state spaces of the individual two dimensional systems. In quantum mechanics the Schrödinger equation is our guide to writing down the equations of motion for a quantum mechanical system, and its solution is the wavefunction. For the composite system of n two dimensional systems the solution of Schrödinger's equation is a wave function, which is a function of the $2n$ configuration space variables for the composite system. Mathematically, this wavefunction can be represented as the tensor product of wavefunctions for the individual two dimensional systems. Hence, the Schrödinger equation description of "quantum reality" implies the tensor product description of composite systems. A more detailed discussion of this argument is given in Notes 17 of the "Littlejohn Notes". However, you will not need to worry about this question of "why" in what follows (unless you want to). We will begin our study of composite systems by considering Hilbert spaces that are tensor products and linear operators defined on Hilbert spaces that are tensor products.

Tensor Product State Spaces and Operators on Tensor Product State Spaces

We begin by defining the tensor product of two Hilbert spaces[4]. Finite dimensional Hilbert spaces will suffice for our needs (and the main

[4] Tensor products of a finite number of Hilbert spaces are straightforwardly defined, but they will not be required for our purposes

example that we will use repeatedly is $\mathbb{C}^2$).

Definition 31 (Tensor Product)**.** *Let $\mathcal{H}_A$ and $\mathcal{H}_B$ denote finite dimensional, complex linear vector spaces, each equipped with an inner product (i.e. finite dimensional Hilbert spaces)*[5]*. The tensor product of $\mathcal{H}_A$ and $\mathcal{H}_B$, denoted $\mathcal{H}_A \otimes \mathcal{H}_B$ is a finite dimensional, complex linear vector space, equipped with an inner product, consisting of i) all elements of the form $\mid a\rangle \otimes \mid b\rangle$, where $\mid a\rangle \in \mathcal{H}_A$ and $\mid b\rangle \in \mathcal{H}_B$ and ii) all possible (finite) linear combinations of elements $\mid a\rangle \otimes \mid b\rangle$, where $\mid a\rangle \in \mathcal{H}_A$ and $\mid b\rangle \in \mathcal{H}_B$. The operations on elements in $\mathcal{H}_A \otimes \mathcal{H}_B$ obey the following rules.*

1. $\lambda\left(\mid a\rangle \otimes \mid b\rangle\right) = \left(\lambda \mid a\rangle\right) \otimes \mid b\rangle = \mid a\rangle \otimes \left(\lambda \mid b\rangle\right), \quad \lambda \in \mathbb{C},$
2. $$\begin{aligned}&\left(\mid a_1\rangle + \mid a_2\rangle\right) \otimes \mid b\rangle = \mid a_1\rangle \otimes \mid b\rangle + \mid a_2\rangle \otimes \mid b\rangle,\\ &\mid a\rangle \otimes \left(\mid b_1\rangle + \mid b_2\rangle\right) = \mid a\rangle \otimes \mid b_1\rangle + \mid a\rangle \otimes \mid b_2\rangle\end{aligned}$$
3. *Let $\langle \cdot \mid \cdot \rangle_A$ denote the inner product on $\mathcal{H}_A$, $\langle \cdot \mid \cdot \rangle_B$ denote the inner product on $\mathcal{H}_B$, and $\langle \cdot \mid \cdot \rangle_{A\otimes B}$ denote the inner product on $\mathcal{H}_A \otimes \mathcal{H}_B$. Then for $\mid a\rangle \otimes \mid b\rangle$, $\mid a'\rangle \otimes \mid b'\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$, the inner product of these two vectors is given by:*

$$\left\langle \left(\mid a\rangle \otimes \mid b\rangle\right) \mid \left(\mid a'\rangle \otimes \mid b'\rangle\right) \right\rangle_{A\otimes B} \equiv \left\langle a \mid a' \right\rangle_A \left\langle b \mid b' \right\rangle_B$$

[5] It is worth commenting on the nature of the subscripts "A" and "B'. We will be discussing measurement, which requires a state space and an observable (i.e. Hermitian operator acting on that state space). Practically, measurements are carried out by people, and it is traditional to speak of "Alice" (A) and "Bob" (B) as two people carrying out measurements on a system. More people could be involved (e.g. C for "Charlie"), but we will not consider the situation of more than two people carrying out measurements.

There are some "mathematical holes" in our definition. We are *stating* that $\mathcal{H}_A \otimes \mathcal{H}_B$ is a finite dimensional, complex linear vector space. That statement requires a proof, which we will not give here. However, recall that a vector space must have a set of scalars associated with it (for us, these are the complex numbers) and a way of adding the objects that are elements of the vector space. Then addition of the objects and multiplication of the objects by scalars must satisfy certain axioms. It would be a useful exercise for you to prove that $\mathcal{H}_A \otimes \mathcal{H}_B$ satisfies the requirements to be a vector space[6] Moreover, an inner product on a vector space must satisfy certain requirements, and it would be a useful exercise for you to show that the inner product defined above satisfies these requirements.

[6] In order to show the existence of an additive identity element it would be useful to first show that $\mid a\rangle \in \mathcal{H}_A$, $\mid b\rangle \in \mathcal{H}_B$, we have $\mid \text{zero}\rangle_{\mathcal{H}_A} \otimes \mid b\rangle = \mid a\rangle \otimes \mid \text{zero}\rangle_{\mathcal{H}_B} = \mid \text{zero}\rangle_{\mathcal{H}_A \otimes \mathcal{H}_B}$, where $\mid \text{zero}\rangle_{\mathcal{H}_A}$ denotes the zero vector in $\mathcal{H}_A$, etc.

Next, we turn to the important issue of a basis for the tensor product space.

Proposition 5 (Orthonormal Basis for the Tensor Product Space)**.** *Let $\{\mid a_i\rangle\}$, $i = 1, \ldots, N_A$ be an orthonormal basis in $\mathcal{H}_A$ and let $\{\mid b_i\rangle\}$, $i = 1, \ldots, N_B$ be an orthonormal basis in $\mathcal{H}_B$. Then $\mid a_i\rangle \otimes \mid b_j\rangle$, $i = 1, \ldots, N_A$, $j = 1, \ldots N_B$ is an orthonormal basis in $\mathcal{H}_A \otimes \mathcal{H}_B$.*

Note that it follows immediately from this result that the dimension of $\mathcal{H}_A \otimes \mathcal{H}_B$ is $N_A N_B$. The proof of this result is left as an exercise (and it is one that you should really make an effort to do).

Example $\mathbb{C}^2 \otimes \mathbb{C}^2$. As an example we construct a basis for $\mathbb{C}^2 \otimes \mathbb{C}^2$. We first choose a basis for $\mathbb{C}^2$:

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \equiv |\, 0\rangle, \qquad \begin{pmatrix} 0 \\ 1 \end{pmatrix} \equiv |\, 1\rangle. \tag{5.41}$$

Recall from (5.7), (5.8) and (5.9) that these are the eigenvectors of the matrix Z. Following Proposition 5, an orthonormal basis for $\mathbb{C}^2 \otimes \mathbb{C}^2$ consists of the following four vectors (with a shorthand notation for each vector immediately to the right):

$$\begin{aligned} |\, 0\rangle \otimes |\, 0\rangle &\equiv |\, 0\rangle\, |\, 0\rangle, \\ |\, 0\rangle \otimes |\, 1\rangle &\equiv |\, 0\rangle\, |\, 1\rangle, \\ |\, 1\rangle \otimes |\, 0\rangle &\equiv |\, 1\rangle\, |\, 0\rangle, \\ |\, 1\rangle \otimes |\, 1\rangle &\equiv |\, 1\rangle\, |\, 1\rangle. \end{aligned} \tag{5.42}$$

Now we come to a very important idea. Let $|\, \psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$. It follows from Proposition 5 that $|\, \psi\rangle$ can be written as a linear combination of elements of $\mathcal{H}_A \otimes \mathcal{H}_B$ of the form $|\, a_i\rangle \otimes |\, b_j\rangle$. However, it is *not* true that *for every* $|\, \psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$ we can find $|\, \psi_A\rangle \in \mathcal{H}_A$, $|\, \psi_B\rangle \in \mathcal{H}_B$ such that:

$$|\, \psi\rangle = |\, \psi_A\rangle \otimes |\, \psi_B\rangle. \tag{5.43}$$

We will show that this is the case by considering a specific example in $\mathbb{C}^2 \otimes \mathbb{C}^2$ that *cannot* be represented in product form.

Consider the state:

$$|\, \phi\rangle = \frac{|\, 0\rangle\, |\, 0\rangle + |\, 1\rangle\, |\, 1\rangle}{\sqrt{2}} \tag{5.44}$$

We wish to find states $|\, \phi_A\rangle$ and $|\, \phi_B\rangle$ such that $|\, \phi\rangle = |\, \phi_A\rangle \otimes |\, \phi_B\rangle$. Letting

$$|\, \phi_A\rangle = a_0\, |\, 0\rangle + a_1\, |\, 1\rangle, \qquad |\, \phi_B\rangle = b_0\, |\, 0\rangle + b_1\, |\, 1\rangle$$

Then we have:

$$|\, \phi_A\rangle \otimes |\, \phi_B\rangle = (a_0\, |\, 0\rangle + a_1\, |\, 1\rangle) \otimes (b_0\, |\, 0\rangle + b_1\, |\, 1\rangle) = a_0 b_0\, |\, 0\rangle\, |\, 0\rangle + a_0 b_1\, |\, 0\rangle\, |\, 1\rangle + a_1 b_0\, |\, 1\rangle\, |\, 0\rangle + a_1 b_1\, |\, 1\rangle\, |\, 1\rangle. \tag{5.45}$$

Then if (5.44) = (5.45) we must have:

$$\begin{aligned} \frac{1}{\sqrt{2}} &= a_0 b_0 \\ 0 &= a_0 b_1 \\ 0 &= a_1 b_0 \\ \frac{1}{\sqrt{2}} &= a_1 b_1. \end{aligned} \tag{5.46}$$

If we examine the second equality in this list, $0 = a_0 b_1$ implies either $a_0 = 0$ or $b_1 = 0$. However, either of these conditions being satisfied is inconsistent with the first and the fourth inequalities in this list. So we cannot find states $\mid \phi_A\rangle$ and $\mid \phi_B\rangle$ such that $\mid \phi\rangle = \mid \phi_A\rangle \otimes \mid \phi_B\rangle$. This leads to the following definition.

Definition 32 (Entangled State)**.** *Any state $\mid \phi\rangle \in \mathcal{H}_\mathcal{A} \otimes \mathcal{H}_\mathcal{B}$ that cannot be written in the form $\mid \phi\rangle = \mid \phi_A\rangle \otimes \mid \phi_B\rangle$, for some $\mid \phi_A\rangle \in \mathcal{H}_A$, $\mid \phi_B\rangle \in \mathcal{H}_B$ is said to be an* entangled state*.*

The following definition should be clear.

Definition 33 (Product State)**.** *Any state $\mid \phi\rangle \in \mathcal{H}_\mathcal{A} \otimes \mathcal{H}_\mathcal{B}$ that can be written in the form $\mid \phi\rangle = \mid \phi_A\rangle \otimes \mid \phi_B\rangle$, for some $\mid \phi_A\rangle \in \mathcal{H}_A$, $\mid \phi_B\rangle \in \mathcal{H}_B$ is said to be a* product state*.*

Linear Operators on Tensor Product Spaces. Suppose $A : \mathcal{H}_A \to \mathcal{H}_A$ and $B : \mathcal{H}_B \to \mathcal{H}_B$ are linear operators. We construct a linear operator on $\mathcal{H}_A \otimes \mathcal{H}_B$ using A and B, which we will refer to as $A \otimes B$, as follows.

Recall Definition 5. It suffices to define a linear operator of a complex linear vector space by specifying its action of each basis element. Then the action of the linear map on a general vector follows by linearity since any vector can be expressed as a linear combination of basis vectors. We define:

$$(A \otimes B) \mid a_i\rangle \mid b_j\rangle \equiv A \mid a_i\rangle B \mid b_j\rangle, \quad i = 1, \ldots N_A,\ j = 1, \ldots N_B. \tag{5.47}$$

Now choose a general $\mid \psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$. Then we can write:

$$\mid \psi\rangle = \sum_{i,j} \alpha_{i,j} \mid a_i\rangle \mid b_j\rangle.$$

Then we have:

$$(A \otimes B) \mid \psi\rangle = \sum_{i,j} \alpha_{i,j} A \mid a_i\rangle B \mid b_j\rangle. \tag{5.48}$$

The following result describes the eigenvalues and eigenvectors of $A \otimes B$ in terms of the eigenvalues and eigenvectors of A and the eigenvalues and eigenvectors of B.

Proposition 6. *Suppose the linear operator A has eigenvalues λ_i and eigenvectors $\mid a_i\rangle$ and suppose the linear operator B has eigenvalues μ_i and eigenvectors $\mid b_i\rangle$. Then the linear operator $A \otimes B$ has eigenvalues $\lambda_i \mu_j$ and eigenvectors $|a_i\rangle \mid b_j\rangle$.*

Proof. Verification of this result involves a straightforward calculation:

$$\begin{aligned} A \otimes B \mid a_i\rangle \mid b_j\rangle &= A \mid a_i\rangle\, B \mid b_j\rangle, \\ &= \lambda_i \mid a_i\rangle\, \mu_j \mid b_j\rangle, \\ &= \lambda_i \mu_j \mid a_i\rangle \mid b_j\rangle. \end{aligned} \tag{5.49}$$

□

Now the natural question arises. If A and B have certain properties (e.g. if they are Hermitian), then does $A \otimes B$ have the same property? We have the following proposition.

Proposition 7.

$$(A \otimes B)^\dagger = A^\dagger \otimes B^\dagger.$$

Proof. The idea for the proof is that we show that this equality holds on basis vectors $\mid a_i\rangle \mid b_j\rangle$, and then it follows that it holds on an arbitrary vector since any vector can be expressed as a linear combination of basis vectors and the map is linear. We have:

$$\begin{aligned} \langle a_i \mid \langle b_j \mid A^\dagger \otimes B^\dagger \mid a_k\rangle \mid b_l\rangle &= \langle a_i \mid A^\dagger \mid a_k\rangle \langle b_j \mid B^\dagger \mid b_l\rangle, \\ &= \left(\langle a_k \mid A \mid a_i\rangle \langle b_l \mid B \mid b_j\rangle\right)^*, \\ &= \left(\langle a_k \mid \langle b_l \mid A \otimes B \mid a_i\rangle \mid b_j\rangle\right)^* \\ &= \langle a_i \mid \langle b_j \mid (A \otimes B)^\dagger \mid a_k\rangle \mid b_l\rangle \end{aligned} \tag{5.50}$$

□

This result implies that if A and B are Hermitian, then $A \otimes B$ is Hermitian. Using this result, together with Proposition 6 allows us to conclude that $A \otimes B$ has a spectral decomposition analogous to what we have derived earlier. In order to keep the discussion simple (and this is the most general case that we will need in the course), let us assume that $\mathcal{H}_A$ and $\mathcal{H}_B$ are finite dimensional, and A and B are nondegenerate. In this case, for A there is a unique eigenvector, $\mid a_i\rangle$, for every eigenvalue λ_i, $i = 1, \ldots N_A$, and for B there is a unique eigenvector, $\mid b_i\rangle$, for every eigenvalue μ_i, $i = 1, \ldots N_B$. Then we have:

$$\begin{aligned} A \otimes B &= \sum_{i,j} \lambda_i \mu_j \mid a_i\rangle \mid b_j\rangle \langle a_i \mid \langle b_j \mid, \\ &= \left(\sum_i \lambda_i \mid a_i\rangle \langle a_i \mid\right) \otimes \left(\sum_j \mu_j \mid b_j\rangle \langle b_j \mid\right). \end{aligned} \tag{5.51}$$

The following result will be useful for computations.

Proposition 8.

$$(A \otimes B)(C \otimes D) = AC \otimes BD.$$

Proof. The idea is the same as the idea for the proof of Proposition 7. We show that this equality holds on basis vectors $\mid a_i\rangle \mid b_j\rangle$, and then it follows that it holds on an arbitrary vector since any vector can be expressed as a linear combination of basis vectors and the map is linear. We have:

$$\begin{aligned}(A \otimes B)\,(C \otimes D) \mid a_i\rangle \mid b_j\rangle &= (A \otimes B)\left(C \mid a_i\rangle \otimes D \mid b_j\rangle\right), \\ &= AC \mid a_i\rangle\, BD \mid b_j\rangle, \\ &= (AC \otimes BD) \mid a_i\rangle \mid b_j\rangle. \end{aligned} \tag{5.52}$$

□

Example: The Pauli Spin Operators and Tensor Products Recall the matrices X and Z associated with the Pauli spin matrices defined on $\mathbb{C}^2$ given in (5.7). We denote the basis of $\mathbb{C}^2$ by $\mid 0\rangle$ and $\mid 1\rangle$ (recall (5.8) and (5.9)). Then we have:

$$\begin{aligned} X \mid 0\rangle &= \mid 1\rangle, \quad Z \mid 0\rangle = \mid 0\rangle, \\ X \mid 1\rangle &= \mid 0\rangle, \quad Z \mid 1\rangle = -\mid 1\rangle. \end{aligned} \tag{5.53}$$

We define the tensor product operator $X \otimes Z$ on $\mathbb{C}^2 \otimes \mathbb{C}^2$ by defining the action of the operator on the basis elements of $\mathbb{C}^2 \otimes \mathbb{C}^2$ as follows:

$$\begin{aligned} X \otimes Z \mid 0\rangle \mid 0\rangle &= X \mid 0\rangle Z \mid 0\rangle = \mid 1\rangle \mid 0\rangle, \\ X \otimes Z \mid 1\rangle \mid 1\rangle &= X \mid 1\rangle Z \mid 1\rangle = -\mid 0\rangle \mid 1\rangle, \\ X \otimes Z \mid 0\rangle \mid 1\rangle &= X \mid 0\rangle Z \mid 1\rangle = -\mid 1\rangle \mid 1\rangle, \\ X \otimes Z \mid 1\rangle \mid 0\rangle &= X \mid 1\rangle Z \mid 0\rangle = \mid 0\rangle \mid 0\rangle. \end{aligned}$$

Then an example of the action of $X \otimes Z$ on a general state in $\mathbb{C}^2 \otimes \mathbb{C}^2$ is the following:

$$X \otimes Z \left(\frac{\mid 0\rangle \mid 0\rangle + \mid 1\rangle \mid 1\rangle}{\sqrt{2}} \right) = \frac{\mid 1\rangle \mid 0\rangle - \mid 0\rangle \mid 1\rangle}{\sqrt{2}}. \tag{5.54}$$

At this point it is useful to note that $X \otimes Z \neq Z \otimes X$. In order to show this it suffices to show that inequality holds on only one element of $\mathbb{C}^2 \otimes \mathbb{C}^2$:

$$X \otimes Z \mid 0\rangle \mid 0\rangle = \mid 1\rangle \mid 0\rangle,$$

and

$$Z \otimes X \mid 0\rangle \mid 0\rangle = \mid 0\rangle \mid 1\rangle.$$

Example. The operator corresponding to doing nothing is the identity operator:

$$\mathbb{I} \mid \psi\rangle = \mid \psi\rangle, \quad \forall \mid \psi\rangle. \tag{5.55}$$

Then $X \otimes \mathbb{I}$ is a tensor product operator on $\mathbb{C}^2 \otimes \mathbb{C}^2$. Examples of the action of this operator on states in $\mathbb{C}^2 \otimes \mathbb{C}^2$ are the following:

$$X \otimes \mathbb{I} \mid 0\rangle \mid 0\rangle = \mid 1\rangle \mid 0\rangle. \tag{5.56}$$

$$X \otimes \mathbb{I} \left(\frac{\mid 0\rangle \mid 0\rangle + \mid 1\rangle \mid 1\rangle}{\sqrt{2}} \right) = \frac{\mid 1\rangle \mid 0\rangle + \mid 0\rangle \mid 1\rangle}{\sqrt{2}}. \tag{5.57}$$

Measurement: Observables as Tensor Product Operators on Tensor Product Spaces.

We now consider the issue of measurement, where the observable is the tensor product of two Hermitian operators on a Hilbert space that is the tensor product of the two Hilbert spaces on which the individual operators making up the tensor product of the two operators acts (you might want to think about this sentence a bit).

However, first let's consider Z acting on $\mathbb{C}^2$. We wish to measure Z in the state

$$\mid \psi\rangle = \frac{\mid 0\rangle + \mid 1\rangle}{\sqrt{2}} \in \mathbb{C}^2.$$

Recall from (5.12) that the spectral decomposition of Z is given by:

$$Z = \mid 0\rangle\langle 0 \mid - \mid 1\rangle\langle 1 \mid,$$

$P_0 \equiv \mid 0\rangle\langle 0 \mid$ is the projection operator onto the eigenstate corresponding to the eigenvalue $+1$ and $P_1 \equiv \mid 1\rangle\langle 1 \mid$ is the projection operator onto the eigenstate corresponding to the eigenvalue -1. The post-measurement state given that the outcome of the measurement is $+1$ is given by:

$$\begin{aligned} \frac{P_0 \mid \psi\rangle}{\parallel P_0 \mid \psi\rangle \parallel} &= \frac{\frac{1}{\sqrt{2}} \mid 0\rangle\langle 0 \mid (\mid 0\rangle + \mid 1\rangle)}{\parallel \frac{1}{\sqrt{2}} \mid 0\rangle\langle 0 \mid (\mid 0\rangle + \mid 1\rangle) \parallel}, \\ &= \frac{\frac{1}{\sqrt{2}} \mid 0\rangle}{\parallel \frac{1}{\sqrt{2}} \mid 0\rangle \parallel}, \\ &= \mid 0\rangle, \end{aligned} \tag{5.58}$$

and the probability that the outcome of the measurement is $+1$ is given by:

$$\begin{aligned}\langle\psi \mid P_0 \mid \psi\rangle &= \frac{1}{2}\left(\langle 0 \mid +\langle 1 \mid\right)\left(\mid 0\rangle\langle 0 \mid\right)\left(\mid 0\rangle+\mid 1\rangle\right), \\ &= \frac{1}{2}\left(\langle 0 \mid +\langle 1 \mid\right)\left(\mid 0\rangle\right) = \frac{1}{2}. \end{aligned} \tag{5.59}$$

Similarly, the post-measurement state, given that the outcome of the measurement is -1, is given by:

$$\frac{P_1 \mid \psi\rangle}{\parallel P_1 \mid \psi\rangle \parallel} =\mid 1\rangle, \tag{5.60}$$

and the probability that the outcome of the measurement is -1 is given by:

$$\langle\psi \mid P_1 \mid \psi\rangle = \frac{1}{2}. \tag{5.61}$$

Now let's consider the operator $Z \otimes \mathbb{I}$ on $\mathbb{C}^2 \otimes \mathbb{C}^2$, and we wish to measure $Z \otimes \mathbb{I}$ in the state,

$$\mid \psi\rangle = \frac{\mid 0\rangle \mid 0\rangle+ \mid 1\rangle \mid 1\rangle}{\sqrt{2}} \in \mathbb{C}^2 \otimes \mathbb{C}^2. \tag{5.62}$$

We know that Z has eigenvalues $+1$ and -1, and $\mathbb{I}$ has eigenvalues $+1$ and $+1$. Therefore by Proposition 6 $Z \otimes \mathbb{I}$ has eigenvalues $+1$, $+1$, -1, and -1. Moreover, from (5.51) we have the following spectral representation:

$$\begin{aligned} Z \otimes \mathbb{I} &= \left(\mid 0\rangle\langle 0 \mid - \mid 1\rangle\langle 1 \mid\right) \otimes \mathbb{I} \\ &= \mid 0\rangle\langle 0 \mid \otimes \mathbb{I} - \mid 1\rangle\langle 1 \mid \otimes \mathbb{I}, \end{aligned} \tag{5.63}$$

where $P_0 \equiv \mid 0\rangle\langle 0 \mid \otimes \mathbb{I}$ is the projection onto the space of eigenvectors corresponding to the eigenvalue $+1$ and $P_1 \equiv \mid 1\rangle\langle 1 \mid \otimes \mathbb{I}$ is the projection onto the space of eigenvectors corresponding to the eigenvalue -1.

Therefore, the post-measurement state given that the outcome of the measurement is $+1$ is given by:

$$\begin{aligned} \frac{P_0 \mid \psi\rangle}{\parallel P_0 \mid \psi\rangle \parallel} &= \frac{\left(\mid 0\rangle\langle 0 \mid \otimes \mathbb{I}\right)\frac{1}{\sqrt{2}}\left(\mid 0\rangle \mid 0\rangle+ \mid 1\rangle \mid 1\rangle\right)}{\parallel P_0 \mid \psi\rangle \parallel}, \\ &= \frac{\frac{1}{\sqrt{2}} \mid 0\rangle \mid 0\rangle}{\parallel \frac{1}{\sqrt{2}} \mid 0\rangle \mid 0\rangle \parallel} =\mid 0\rangle \mid 0\rangle, \end{aligned} \tag{5.64}$$

and the probability that the outcome of the measurement is $+1$ is given by:

$$\begin{aligned}\langle\psi \mid P_0 \mid \psi\rangle &= \langle\psi \mid (\mid 0\rangle\langle 0 \mid \otimes \mathbb{I})\frac{1}{\sqrt{2}}(\mid 0\rangle \mid 0\rangle + \mid 1\rangle \mid 1\rangle), \\ &= \frac{1}{2}(\langle 0 \mid \langle 0 \mid + \langle 1 \mid \langle 1 \mid) \mid 0\rangle \mid 0\rangle = \frac{1}{2}, \end{aligned} \tag{5.65}$$

Similar calculations can be carried out for the post-measurement state given that the outcome of the measurement is -1 and the probability that the outcome of the measurement is -1.

Non-Locality and Bell Inequalities

The Einstein-Podolsky-Rosen (EPR) Paradox

The thought experiment of Einstein, Podolsky, and Rosen has played a fundamental role in our understanding of quantum mechanics. The original reference is the following:

> A. Einstein, B. Podolsky, and N. Rosen, Can quantum-mechanical description of physical reality be considered complete? *Phys. Rev.*, **47**, 777 (1935).

We will first describe the thought experiment, and discuss what were believed to be "paradoxical features" of the outcome of the thought experiments afterwards.

Suppose we have a source that creates two quantum particles of spin $\frac{1}{2}$. One is sent to Alice and one is sent to Bob ("A" and "B"). The two particles that are sent to Alice and Bob are created by the source through some type of physical process. Consequently, the two particles form a composite system. Figure 5 shows a schematic diagram of the two particles created by the source, with one sent to Alice and one sent to Bob. Alice and Bob will each perform measurements on their particles. We assume that Alice and Bob are sufficiently spatially separated that no communication between them is possible in the time that it takes them to perform their measurements (even at the speed of light).

Figure 5.5: Schematic of the source creating two particles, one is sent to Alice and one is sent to Bob.

The Hilbert space for the composite system of two spin $\frac{1}{2}$ particles is $\mathbb{C}^2 \otimes \mathbb{C}^2$, and we assume that the two particles are in the following

state (you might recall this state from earlier in the material on tensor products):

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|0\rangle \otimes |0\rangle + |1\rangle \otimes |1\rangle. \tag{5.66}$$

Let us suppose that Alice measures Z, and Bob does not do anything. Then the measurement operator on $\mathbb{C}^2 \otimes \mathbb{C}^2$ is $Z \otimes \mathbb{I}$ (and we know that it is Hermitian). The spectral decomposition of this operator is given by:

$$Z \otimes \mathbb{I} = \underbrace{|0\rangle\langle 0| \otimes \mathbb{I}}_{P_0} - \underbrace{|1\rangle\langle 1| \otimes \mathbb{I}}_{P_1},$$

where P_0 is the projection onto the eigenspace corresponding to eigenvalues of $+1$ and P_1 is the projection onto the eigenspace corresponding to eigenvalues of -1.

Using the postulate of quantum mechanics concerned with measurement, if the outcome of Alice's measurement is $+1$ then the state collapses to:

$$\frac{P_0 |\psi\rangle}{\| P_0 |\psi\rangle \|} = \frac{(|0\rangle\langle 0| \otimes \mathbb{I}) \left(\frac{1}{\sqrt{2}}(|0\rangle \otimes |0\rangle + |1\rangle \otimes |1\rangle\right)}{\| (|0\rangle\langle 0| \otimes \mathbb{I}) \left(\frac{1}{\sqrt{2}}(|0\rangle \otimes |0\rangle + |1\rangle \otimes |1\rangle\right) \|} = |0\rangle \otimes |0\rangle \tag{5.67}$$

If the outcome of Alice's measurement is -1 then the state collapses to:

$$\frac{P_1 |\psi\rangle}{\| P_1 |\psi\rangle \|} = \frac{(|1\rangle\langle 1| \otimes \mathbb{I}) \left(\frac{1}{\sqrt{2}}(|0\rangle \otimes |0\rangle + |1\rangle \otimes |1\rangle\right)}{\| (|1\rangle\langle 1| \otimes \mathbb{I}) \left(\frac{1}{\sqrt{2}}(|0\rangle \otimes |0\rangle + |1\rangle \otimes |1\rangle\right) \|} = |1\rangle \otimes |1\rangle \tag{5.68}$$

It is important to realize that Alice's measurement has collapsed the wave function *for the composite system*.

After Alice has completed her measurement (and therefore collapsed the wave function) suppose that Bob measures Z. In this case the measurement operator for the composite system is given by:

$$\mathbb{I} \otimes Z = \underbrace{\mathbb{I} \otimes |0\rangle\langle 0|}_{P_0'} - \underbrace{\mathbb{I} \otimes |1\rangle\langle 1|}_{P_1'}, \tag{5.69}$$

where P_0' is the projection onto the eigenspace corresponding to $+1$ and P_1' is the projection onto the eigenspace corresponding to -1.

Now if Alice obtained $+1$ for her measurement of Z then the state collapses to $|0\rangle \otimes |0\rangle$ and the probabilities that Bob measures 1 or -1

are given by:

$$\begin{array}{rcccl} p\,(\text{Bob obtains } +1) & = & \langle 0 \mid \langle 0 \mid P_0' \mid 0\rangle \mid 0\rangle & = & 1, \\ p\,(\text{Bob obtains } -1) & = & \langle 0 \mid \langle 0 \mid P_1' \mid 0\rangle \mid 0\rangle & = & 0. \end{array} \tag{5.70}$$

If Alice obtained -1 for her measurement of Z then the state collapses to $\mid 1\rangle \mid 1\rangle$ and the probabilities that Bob measures 1 or -1 are given by:

$$\begin{array}{rcccl} p\,(\text{Bob obtains } +1) & = & \langle 1 \mid \langle 1 \mid P_0' \mid 1\rangle \mid 1\rangle & = & 0, \\ p\,(\text{Bob obtains } -1) & = & \langle 1 \mid \langle 1 \mid P_1' \mid 1\rangle \mid 1\rangle & = & 1. \end{array} \tag{5.71}$$

So Bob always obtains the same result as Alice.

One could ask if this would still hold if Alice measured X, and then Bob measured X? In this case the measurement operators are given by:

$$\begin{array}{rcl} X \otimes \mathbb{I} & = & \mid +\rangle\langle + \mid \otimes \mathbb{I} - \mid -\rangle\langle - \mid \otimes \mathbb{I}, \quad \text{Alice}, \\ \mathbb{I} \otimes X & = & \mathbb{I} \otimes \mid +\rangle\langle + \mid - \mathbb{I} \otimes \mid -\rangle\langle - \mid \otimes \mathbb{I}, \quad \text{Bob}. \end{array} \tag{5.72}$$

We need to express $\mid \psi\rangle$ in terms of the basis of eigenvectors of X. Recall from (5.22) we have:

$$\mid 0\rangle = \frac{\mid +\rangle + \mid -\rangle}{\sqrt{2}}, \quad \mid 1\rangle = \frac{\mid +\rangle - \mid -\rangle}{\sqrt{2}}, \tag{5.73}$$

and therefore

$$\begin{array}{rcl} \mid \psi\rangle & = & \dfrac{1}{\sqrt{2}} \left(\mid 0\rangle \mid 0\rangle + \mid 1\rangle \mid 1\rangle \right), \\ & = & \dfrac{1}{\sqrt{2}} \left(\dfrac{\mid +\rangle + \mid -\rangle}{\sqrt{2}} \right) \left(\dfrac{\mid +\rangle + \mid -\rangle}{\sqrt{2}} \right) + \dfrac{1}{\sqrt{2}} \left(\dfrac{\mid +\rangle - \mid -\rangle}{\sqrt{2}} \right) \left(\dfrac{\mid +\rangle - \mid -\rangle}{\sqrt{2}} \right), \\ & = & \dfrac{1}{\sqrt{2}} \left(\mid +\rangle \mid +\rangle + \mid -\rangle \mid -\rangle \right). \end{array} \tag{5.74}$$

It follows that we will obtain the same results for measuring X as we did for Z simply by realizing that interchanging the kets $\mid 0\rangle \leftrightarrow \mid +\rangle$ and $\mid 1\rangle \leftrightarrow \mid -\rangle$ gives same results. Therefore, if Alice first measures X and Bob measures X afterwards, he will obtain exactly the same results as Alice.

Note that $\mid \psi\rangle$ is entangled. Would the same conclusions hold if the two particles were not entangled?

In summary the postulate of measurement in quantum mechanics says that if Alice measures a quantity then the state collapses to the state corresponding to the eigenstate of the outcome of Alice's measurement. Therefore if Bob measures the same quantity he will obtain exactly the same value as Alice. The result of this experiment seems contrary to common experience in two ways.

Locality. The experiment has been arranged to that no communication is possible between Alice and Bob, after Alice makes her measurement, and before Bob makes his measurement. ("Faster than the speed of light" communications violate special relativity.) Yet, somehow, at the instant Alice makes her measurement Bob's wave function "collapses" to the eigenstate corresponding to the eigenvalue that Alice measured.

Reality. We believe that the properties of a particle should be an intrinsic characteristic of the particle and they should *not* depend on what measurements are made. Once Alice makes her measurement of Z, if Bob then measures Z he will always obtain the same value as Alice.

These issues troubled Einstein greatly. In letters to Niels Bohr he coined the famous phrase "spooky action at a distance"[7] to describe the results of this thought experiment. Einstein sought a way around these issues, and that led to "local hidden variable models", that we now describe.

[7] Letter from Einstein to Max Born, 3 March 1947; *The Born-Einstein Letters; Correspondence between Albert Einstein and Max and Hedwig Born from 1916 to 1955*, Walker, New York, 1971.

Local Hidden Variable Models

We give a brief description of the idea behind "hidden variable theories" Suppose that in addition to the wave function, the particles had a "hidden" list of "answers" to measurements, and all the measurement does is reveal this answer. This is shown schematically in Figure 5 for the observable Z.

If additional observables are measured, we just extend the list of hidden variables to include further measurement outcomes. This is illustrated schematically in Fig. 5 for the observables X and Z.

Suppose that the EPR source produces pairs of particles that carry the same hidden variables, as schematically shown in Fig. 5.

This would explain why Alice and Bob always get the same outcome, without any mysterious wave function collapse or instantaneous action at a distance, i.e. the particles carry hidden variables along with them- and the measurements are just revealing a "local" property of the particles. If the source sometimes produces pairs with one set of hidden variables, and sometimes another set (but always the same values a given pair of particles, i.e. the two particles that were produced at the same time by the source), this would also explain why sometimes $+1$ is measured and sometimes -1 is measured.

Einstein believed that there were hidden variables which, if we knew their values, would allow us to predict the outcomes of their measurements exactly. Therefore "local realism" would be restored. Effectively, this means that quantum mechanics would be a statistical theory, much like classical statistical mechanics describes the be-

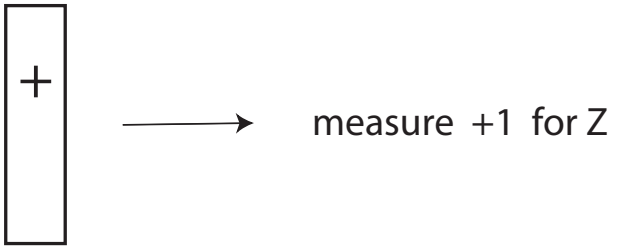
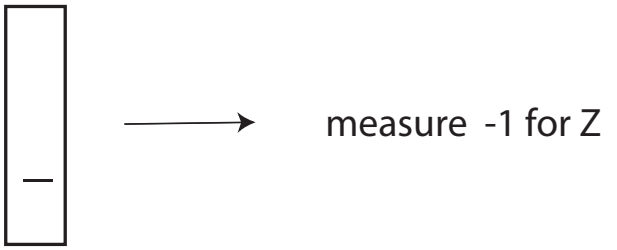


Figure 5.6: The vertical rectangles are a schematic representation of the particles. The symbol in the rectangle shows the sign of the value of Z which is revealed when Z is measured.



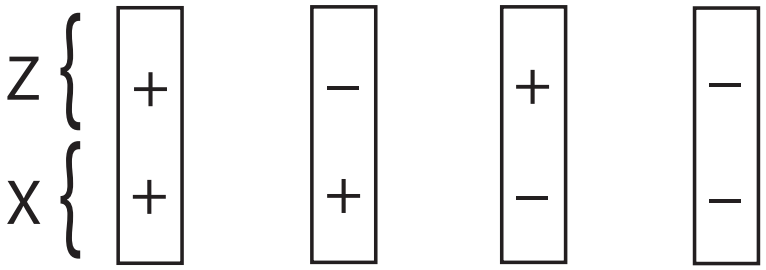


Figure 5.7: Particles contain the outcomes of measurement of X and Z.

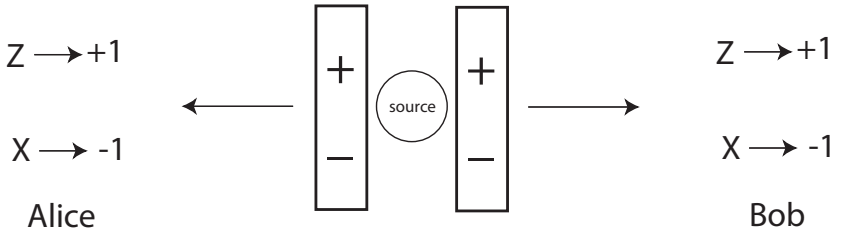


Figure 5.8: Schematic of the pairs of particles produced by the EPR source with each carrying the same "hidden variables", i.e. outcomes of the measurements of Z and X.

haviour of gasses probabilistically, even though the underlying classical mechanics is deterministic. However, Bohr believed that wavefunction collapse and probabilistic outcomes were a fundamental part of

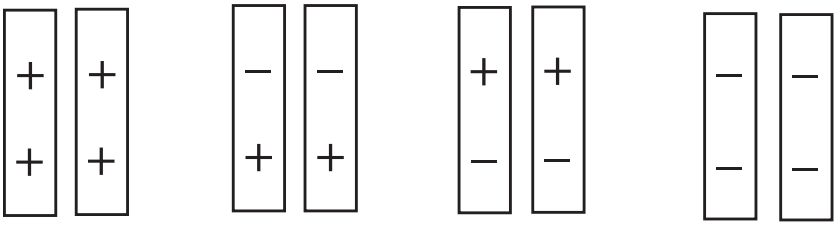


Figure 5.9: Pairs of particles produced by the EPR source with the same "hidden variables".

how nature works. Both of these explanations were consistent with the EPR thought experiment, and this is the way the situation remained for some years, until John Bell came up with a twist on the EPR experiment that allowed hidden variables models to be put to the test[8], and this is the next topic that we will consider.

[8] J. S. Bell [1964] On the Einstein Podolsky Rosen Paradox. *Physics*, **1**, 195-200.

Bell's Experiment (The Clauser-Horne-Shimony-Holt (CHSH) Version)

Rather than discuss Bell's experiment, we will consider a slightly different, but similar in spirit, experiment due to Clauser, Horne, Shimony and Holt[9]. They described another thought experiment that, in principle, could be realized in the laboratory (we will mention this a bit more later).

[9] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt [1969] Proposed experiment to test local hidden variable theories. *Phys. Rev. Lett*, **23**(15), 880-884.

- A source creates a pair of particles. One is sent to Alice, and the other to Bob. Note that we have not specified whether or not these are classical particles or quantum particles.
- Alice has a choice of *two* properties, denoted A_1 and A_2, that she can measure on her particle, and the possible outcome of either measurement is ± 1. More precisely, we denote the possible values of the property A_1 by $a_1 = \pm 1$ and the possible values of the property A_2 by $a_2 = \pm 1$. Similarly, Bob has a choice of measuring two properties, B_1 and B_2, that take possible values $b_1 = \pm 1$ and $b_2 = \pm 1$, respectively. We assume that what Alice decides to measure has no effect on what Bob decides to measure. This is the "locality assumption".
- Alice and Bob repeat the experiment many times (i.e. many particle pairs are sent to Alice and Bob from the source), and they choose at random which property that they will measure. They keep a record of both their measurement choices (i.e. the property they choose to measure) and the outcomes of the measurement.
- We denote the expected value of the product of $a_1 b_1$ $(= \pm 1)$ by $\mathbb{E}(a_1 b_1)$, the expected value of the product of $a_1 b_2$ $(= \pm 1)$ by $\mathbb{E}(a_1 b_2)$, the expected value of the product of $a_2 b_1$ $(= \pm 1)$ by $\mathbb{E}(a_2 b_1)$, and

the expected value of the product of $a_1 b_1$ ($= \pm 1$) by $\mathbb{E}(a_2 b_2)$. We assume that there is a joint probability distribution that governs the outcome of all measurements that Alice and Bob might perform, $P(a_1.a_2, b_1, b_2)$. This is the hypothesis of "reality". If the values of the outcomes are are known exactly, then the outcome of any measurement can be predicted with certainty–the measurement outcomes are described probabilistically because the values of the hidden variables are drawn from an ensemble of possible values.

Based on these assumptions, Clauser, Horne, Shimony and Holt (CHSH) derived an inequality that the experimental results must satisfy, under the assumptions above.

Theorem 13 (The CHSH Inequality)**.**

$$\mathbb{E}(a_1 b_1) + \mathbb{E}(a_1 b_2) + \mathbb{E}(a_2 b_1) - \mathbb{E}(a_2 b_2) \leq 2$$

Proof. Consider the quantity:

$$\begin{aligned} C &= a_1 b_1 + a_1 b_2 + a_2 b_1 - a_2 b_2, \\ &= a_1(b_1 + b_2) + a_2(b_1 - b_2). \end{aligned} \tag{5.75}$$

(Note that in writing (5.75) down we are assuming that values of all four observables can be specified simultaneously, and this is an assumption of the "hidden variable theories" and not consistent with quantum mechanics.) Since b_1, $b_2 = \pm 1$ we have

$$\begin{aligned} \text{either} \quad & b_1 + b_2 = \pm 2 \quad \text{and} \quad b_1 - b_2 = 0, \\ \text{or} \quad & b_1 + b_2 = 0 \quad \text{and} \quad b_1 - b_2 = \pm 2. \end{aligned} \tag{5.76}$$

And also a_1, $a_2 = \pm 1$. Using these relations in (5.75) gives:

$$\begin{aligned} C &= a_1(b_1 + b_2) + a_2(b_1 - b_2), \\ &= (\pm 1)\,(\pm 2), \\ &= \pm 2 \leq 2. \end{aligned} \tag{5.77}$$

Let $P(a_1, a_2, b_1, b_2)$ be the probability that the particles are in the state where property A_1 has the value a_1, property A_2 has the value a_2, property B_1 has the value b_1, and property B_2 has the value b_2. Then the expected value for the quantity C is given by:

$$\begin{aligned} \mathbb{E}(C) &= \sum_{a_1, a_2, b_1, b_2} \left(a_1(b_1 + b_2) + a_2(b_1 - b_2)\right) P(a_1, a_2, b_1, b_2), \\ &\leq \sum_{a_1, a_2, b_1, b_2} 2\, P(a_1, a_2, b_1, b_2), \quad \text{using} \quad (5.77). \\ &= 2. \end{aligned} \tag{5.78}$$

But we also have:

$$\begin{aligned}\mathbb{E}(C) &= \mathbb{E}\left(a_1b_1 + a_1b_2 + a_2b_1 - a_2b_2\right) \\ &= \mathbb{E}(a_1b_1) + \mathbb{E}(a_1b_2) + \mathbb{E}(a_2b_1) - \mathbb{E}(a_2b_2) \qquad (5.79)\end{aligned}$$

Combining this expression, with (5.78), gives:

$$\mathbb{E}(a_1b_1) + \mathbb{E}(a_1b_2) + \mathbb{E}(a_2b_1) - \mathbb{E}(a_2b_2) \leq 2. \qquad (5.80)$$

□

Now let's consider this experiment with quantum particles. Suppose the source emits pairs of quantum particles of spin $\frac{1}{2}$. For each pair emitted, one is sent to Alice and one is sent to Bob. Moreover, we assume that each emitted pair of particles is in the state:

$$|\,\psi\rangle = \frac{|\,0\rangle\,|\,0\rangle + |\,1\rangle\,|\,1\rangle}{\sqrt{2}} \qquad (5.81)$$

(Note: this state should be familiar to you.)

Alice chooses to measure $A_1 = X$ or $A_2 = Z$ and Bob chooses to measure $B_1 = \frac{X+Z}{\sqrt{2}}$ or $B_2 = \frac{X-Z}{\sqrt{2}}$. Recall that:

$$X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

and therefore

$$\frac{X+Z}{\sqrt{2}} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \quad \frac{X-Z}{\sqrt{2}} = \frac{1}{\sqrt{2}}\begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix}.$$

It is easy to verify that each of these four matrices has eigenvalues ± 1. Therefore the outcome of any measurement of these four quantities is ± 1.

We will need the following two calculations:

$$\begin{aligned}\langle\psi\,|\,Z\otimes Z\,|\,\psi\rangle &= \frac{1}{2}\left(\langle 0\,|\,\langle 0\,| + \langle 1\,|\,\langle 1\,|\right) Z\otimes Z\left(|\,0\rangle\,|\,0\rangle + |\,1\rangle\,|\,1\rangle\right), \\ &= \frac{1}{2}(1+1) = 1, \qquad (5.82)\end{aligned}$$

and

$$\langle\psi\,|\,X\otimes X\,|\,\psi\rangle = \frac{1}{2}\left(\langle 0\,|\,\langle 0\,| + \langle 1\,|\,\langle 1\,|\right) X\otimes X\left(|\,0\rangle\,|\,0\rangle + |\,1\rangle\,|\,1\rangle\right) = 1. \qquad (5.83)$$

Now we have:

$$
\begin{aligned}
&\mathbb{E}(a_1b_1) + \mathbb{E}(a_1b_2) + \mathbb{E}(a_2b_1) - \mathbb{E}(a_2b_2) \\
&= \langle\psi \mid X \otimes \frac{X+Z}{\sqrt{2}} \mid \psi\rangle + \langle\psi \mid X \otimes \frac{X-Z}{\sqrt{2}} \mid \psi\rangle \\
&+ \langle\psi \mid Z \otimes \frac{X+Z}{\sqrt{2}} \mid \psi\rangle - \langle\psi \mid Z \otimes \frac{X-Z}{\sqrt{2}} \mid \psi\rangle \\
&= \frac{1}{\sqrt{2}} \left(2\langle\psi \mid X \otimes X \mid \psi\rangle + 2\langle\psi \mid Z \otimes Z \mid \psi\rangle\right) \\
&= \frac{4}{\sqrt{2}} = 2\sqrt{2} > 2
\end{aligned} \tag{5.84}
$$

But this result contradicts Theorem 13. Therefore there must have been assumptions in the proof of this inequality that are not valid for quantum mechanics. Recall that there were two key assumptions–the locality assumption and the "realism" assumption. One, or both, of these assumptions are not consistent with quantum mechanics. However, violation of the CHSH inequality rules out the hidden variable model as an explanation for the results of the EPR experiment.

Not that $\mid \psi\rangle$ is entangled. Would the CHSH inequality be violated if $\mid \psi\rangle$ were not entangled?

The experiments that we have described have been "thought experiments". However, technology has developed to the point where they can be realized in a laboratory. The main difficulties have been creating the pairs of entangled particles and ensuring that the locality assumption is enforced. A summary of these experiments can be found here:

http://en.wikipedia.org/wiki/Bell_test_experiments

Problems

1. *The spin observable for an arbitrary direction.*

   The operator $J(\theta,\phi)$ corresponding to spin in the spatial direction (θ,ϕ) in spherical coordinates on the unit sphere in three dimensions is

   $$J(\theta,\phi) = \sin\theta\cos\phi J_x + \sin\theta\sin\phi J_y + \cos\theta J_z.$$

   Consider the spin-1/2 representation.

   (a) Show that the eigenvalues of $J(\theta,\phi)$ are $\pm\hbar/2$ and that the corresponding normalised eigenvectors may be taken to be

   $$|\theta,\phi\rangle = \cos\left(\frac{\theta}{2}\right)\left|\frac{1}{2}\ \ \frac{1}{2}\right\rangle + e^{i\phi}\sin\left(\frac{\theta}{2}\right)\left|\frac{1}{2}\ \ -\frac{1}{2}\right\rangle,$$

   and

   $$|\pi-\theta,\phi+\pi\rangle = \sin\left(\frac{\theta}{2}\right)\left|\frac{1}{2}\ \ \frac{1}{2}\right\rangle - e^{i\phi}\cos\left(\frac{\theta}{2}\right)\left|\frac{1}{2}\ \ -\frac{1}{2}\right\rangle,$$

   where $\left|\frac{1}{2}\ \ \pm\frac{1}{2}\right\rangle$ satisfy $J_z\left|\frac{1}{2}\ \ \pm\frac{1}{2}\right\rangle = \pm\frac{\hbar}{2}\left|\frac{1}{2}\ \ \pm\frac{1}{2}\right\rangle$.

   (b) Show that the pair of vectors $|\theta,\phi\rangle$ and $|\pi-\theta,\phi+\pi\rangle$ form an orthonormal basis for $\mathbb{C}^2$ (for fixed values of (θ,ϕ)).

   (c) Show also that the operator $J(\theta,\phi)$ may be written in terms of the projectors onto $|\theta,\phi\rangle$ and $|\pi-\theta,\phi+\pi\rangle$ as

   $$J(\theta,\phi) = \frac{\hbar}{2}|\theta,\phi\rangle\langle\theta,\phi| - \frac{\hbar}{2}|\pi-\theta,\phi+\pi\rangle\langle\pi-\theta,\phi+\pi|.$$

2. *Measurement of spin in arbitrary directions.*

   Consider a spin-1/2 particle in the state $\left|\frac{1}{2}\ \ \frac{1}{2}\right\rangle$. By writing the state in terms of the eigenstates of $J(\theta,\phi)$ (defined in question 1), or otherwise, calculate the probability that the eigenvalue $\hbar/2$ is found when $J(\theta,\phi)$ is measured.

3. *The uncertainty relations for spin.*

   (a) Derive the following uncertainty relation for spin from the commutation relations:

   $$\Delta_{|\psi\rangle}(J_x)\,\Delta_{|\psi\rangle}(J_y) \geq \frac{\hbar}{2}\left|E_{|\psi\rangle}(J_z)\right|. \tag{5.85}$$

(note that the right hand side of this equation depends on the state $|\psi\rangle$ unlike the case for the canonical commutation relations). Under what circumstances is there equality in (5.85)?

(b) Calculate the terms in the uncertainty relation above for the state $|\psi\rangle = \left|\frac{1}{2}\ \frac{1}{2}\right\rangle$, and confirm that the uncertainty relation is satisfied. Comment on your answer in the light of your answer to the last part of (a).

(c) Calculate $\Delta_{|\psi\rangle}(J_x)$ for the state

$$|\psi\rangle = \frac{1}{\sqrt{2}}\left(\left|\frac{1}{2}\ \frac{1}{2}\right\rangle + \left|\frac{1}{2}\ -\frac{1}{2}\right\rangle\right).$$

Is this value consistent with the uncertainty relation?

4. *Tensor product operations.*

Let us define the operators X, Y and Z on $\mathbb{C}^2$ by

$$\begin{aligned} X|1\rangle &= |2\rangle; & X|2\rangle &= |1\rangle \\ Y|1\rangle &= i|2\rangle; & Y|2\rangle &= -i|1\rangle \\ Z|1\rangle &= |1\rangle; & Z|2\rangle &= -|2\rangle. \end{aligned}$$

and let $\mathbf{I}$ denote the identity operator on $\mathbb{C}^2$.

(a) Show that the operator $X \otimes \mathbf{I}$ on $\mathbb{C}^2 \otimes \mathbb{C}^2$ is unitary. You may use, without proof, the facts that $(A \otimes B)^\dagger = A^\dagger \otimes B^\dagger$, and $(A \otimes B)(C \otimes D) = AC \otimes BD$, for operators A, B, C, D.

(b) Let $|\Psi\rangle$ be the state

$$|\Psi\rangle = \frac{1}{\sqrt{2}}(|1\rangle|2\rangle - |2\rangle|1\rangle)$$

on $\mathbb{C}^2 \otimes \mathbb{C}^2$. Calculate $|\Psi_X\rangle = X \otimes \mathbf{I}|\Psi\rangle$.

(c) Calculate also $|\Psi_Y\rangle = Y \otimes \mathbf{I}|\Psi\rangle$, and $|\Psi_Z\rangle = Z \otimes \mathbf{I}|\Psi\rangle$. Show that the four states $|\Psi\rangle, |\Psi_X\rangle, |\Psi_Y\rangle, |\Psi_Z\rangle$ form an orthonormal basis for $\mathbb{C}^2 \otimes \mathbb{C}^2$.

5. Suppose we have a source that creates two quantum particles of spin $\frac{1}{2}$. One is sent to Alice and one is sent to Bob ("A" and "B"). The two particles that are sent to Alice and Bob are created by the source through some type of physical process and they form a composite system. Figure 5 shows a schematic diagram of the two particles created by the source, with one sent to Alice and one sent to

Bob. Alice and Bob can each independently perform measurements on their particles. We assume that Alice and Bob are sufficiently spatially separated that no communication between them is possible in the time that it takes them to perform their measurements (even at the speed of light).

Figure 5.10: Schematic of the source creating two particles, one is sent to Alice and one is sent to Bob.

Let $\{|\,0\rangle, |\,1\rangle\}$ denote an orthonormal basis of $\mathbb{C}^2$ and let Z be the operator on $\mathbb{C}^2$ defined by:

$$Z\,|\,0\rangle = |\,0\rangle, \qquad Z\,|\,1\rangle = -1.$$

The Hilbert space for the composite system of two spin $\frac{1}{2}$ particles is $\mathbb{C}^2 \otimes \mathbb{C}^2$, and we assume that the two particles are in the following state:

$$|\,\psi\rangle = \frac{1}{\sqrt{2}}\left(|\,0\rangle\,|\,0\rangle + |\,1\rangle\,|\,1\rangle\right). \tag{5.86}$$

Let us suppose that Alice measures Z, and Bob does not do anything. Then the measurement operator on $\mathbb{C}^2 \otimes \mathbb{C}^2$ is $Z \otimes \mathbb{I}$.

(a) Show that $Z \otimes \mathbb{I}$ is self-adjoint.

(b) Compute the eigenvalues and eigenvectors of $Z \otimes \mathbb{I}$.

(c) Express $Z \otimes \mathbb{I}$ in spectral form in Dirac notation.

(d) Show that $|\,\psi\rangle$ is entangled.

(e) Suppose the outcome of Alice's measurement is $+1$. Determine the state of the system after this measurement.

(f) Suppose Alice measures $+1$. What is the probability that Bob measures $+1$ after Alice's measurement? What is the probability that Bob measures -1 after Alice's measurement?

Bibliography

M. Amaku, F. A. B. Coutinho, and F. M. Toyama. The normalization of wave functions of the continuous spectrum. *Revista Brasileira de Ensino de Física*, 42, 2020.

D. Bohm. *Quantum theory*. Courier Corporation, 1951.

C. Cohen-Tannoudji, B. Diu, and F. Laloe. *Quantum Mechanics*. Wiley-VCH, 1992.

P. A. M. Dirac. A new notation for quantum mechanics. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 35, pages 416–418. Cambridge University Press, 1939.

P. A. M. Dirac. *The principles of quantum mechanics*. Number 27. Oxford university press, 1981.

A. Furuta. One thing is certain: Heisenberg's uncertainty principle is not dead. *Scientific American*, 2012.

K. Hannabuss. *An introduction to quantum theory*, volume 1. Clarendon Press, 1997.

W. Heisenberg. *The physical principles of the quantum theory*. Courier Corporation, 1949.

W. Heisenberg. Über den anschaulichen inhalt der quantentheoretischen kinematik und mechanik. In *Original Scientific Papers Wissenschaftliche Originalarbeiten*, pages 478–504. Springer, 1985.

A. J. G. Hey, T. Hey, and P. Walters. *The new quantum universe*. Cambridge University Press, 2003.

E. H. Kennard. Zur quantenmechanik einfacher bewegungstypen. *Zeitschrift für Physik*, 44(4-5):326–352, 1927.

B. Mielnik and O. Rosas-Ortiz. Factorization: little or great algorithm? *Journal of Physics A: Mathematical and General*, 37(43):10007, 2004.

H. P. Robertson. The uncertainty principle. *Physical Review*, 34(1):163, 1929.

E. Schrödinger. An undulatory theory of the mechanics of atoms and molecules. *Physical review*, 28(6):1049, 1926.

B. Schumacher and M. Westmoreland. *Quantum processes systems, and information*. Cambridge University Press, 2010.

A D. Stone. *Einstein and the quantum: The quest of the valiant Swabian*. Princeton University Press, 2015.

D. J Tannor. *Introduction to quantum mechanics: a time-dependent perspective*. University Science Books, 2007.

H. Weyl. Gruppentheorie und quantenmechanik, hirzel, leipzig. *Theory of Groups and Quantum Mechanics, 2nd ed.(1931), transl. H. P. Robertson, Dover, NY (1950)*, pages 100–101, 1928.

Index