Quantum Mechanics Primer for P4331

E.H. Majzoub

June 8, 2022

Abstract

This is a short summary of some of the tools we will use and topics we will cover in P4331. The primary purpose of these notes is to refresh your knowledge of linear algebra and demonstrate how it is used in quantum mechanics. These notes are not a replacement of the material we will cover in class. Many of the topics touched on here will be generalized in the lectures and supplemented in class with specific examples and further explanation. You should view these notes as a toolbox to consult as needed throughout the course.

Sections 2 and 3 cover linear algebra and its applications to quantum mechanics. These are the two most important sections for you to read at the beginning of the course. Section 4 covers the Schrodinger equation and may be read when we reach that point in the course. It may help to bounce between sections 3 and 4 as you read them. The remaining sections cover more advanced topics we may not get to in class. Finally, there is an appendix on some of the basic mathematical tools necessary for working problems in the course. I suggest you read over these sections as we come to use these methods in class.

If you find any typos or other errors please let me know at majzoube@umsystem.edu.

Contents

1 Classical vs. Quantum Mechanics

Classical mechanics and quantum mechanics are not without similarities, but there are major differences in their underlying mathematical frameworks, and in their conceptual approaches. In classical mechanics we search for particle trajectories. In contrast, quantum mechanics introduces a fundamentally new idea into physics, that of the state function. Some of the differences and similarities are listed below.

Classical Mechanics

- Described by Newton's Laws. The second law $\vec{F} = m\vec{a} = m\frac{d^2}{dt^2}x(t)$, is a real, ordinary differential equation (ODE).
- The goal is to find all particle **trajectories** $x_i(t)$ for a collection of particles labeled by $i = \{1, 2, 3, \cdots\}$.
- The set $\{x_i(t)\}$ is a complete description of the system. The energy, E, angular momentum $\vec{L}$, and other quantities can be calculated from the $x_i(t)$.
- Classical E&M + mechanics fails at describing systems like the hydrogen atom.
- Mathematical structure: Lagrangian and Hamiltonian formulations. ODEs/PDEs.

Quantum Mechanics

- Described by Schrödinger's equation, $i\hbar\frac{\partial}{\partial t}\psi(x,t) = \hat{H}\psi(x,t)$, a *complex* partial differential equation (PDE).
- Goal is to find the **state function** $\psi(x,t)$, also known as a **state vector**.
- $\psi(x,t)$ does not describe a particle trajectory like $x(t)$ does in classical mechanics. Instead, $\psi(x,t)$ gives the state of the particle where the location of the particle may be found with a certain *probability*. That probability is $P(x, x+dx, t) = \psi^*(x,t)\psi(x,t)\,dx$.
- The state function is an abstract object and may not depend on position at all! It may describe some aspect of the system that has no classical analog, e.g. particle *spin*.
- Mathematical structure: **Linear Algebra** of the state vectors that satisfy Schrödinger's equation. This is the more abstract notion of **Hilbert spaces** and operators that act on the Hilbert space.

2 Basic Linear Algebra

We will use linear algebra almost every day in our course on quantum mechanics. It is *very important* to know the basics so that you are not lost after the first few days of class.

We begin with some basic definitions of **finite** vector spaces that we will generalize when we study quantum mechanics. Keep in mind that in quantum mechanics the "vectors" will be more abstract, but they will follow essentially the same rules. If you are feeling confused about the state vectors in quantum mechanics, I encourage you to return to basic linear algebra and think about vectors in 3-dimensions for reference.

2.1 Vector spaces over $\mathbb{R}$

A **vector space**, V, over the real numbers, is a set of objects called vectors $\{\vec{v}_1, \vec{v}_2, \vec{v}_3, \cdots\}$ together with a law of composition, addition "+", such that the following conditions hold:

1. V is **closed under addition**. If $\vec{v}_i, \vec{v}_j \in V$, then

$$(\vec{v}_i + \vec{v}_j) \in V,$$

   i.e., the sum of two vectors is also an element of the vector space.

2. V is **closed under scalar multiplication** by a number in the reals $r \in \mathbb{R}$, i.e.

$$r\vec{v}_i \in V.$$

3. V is **commutative** under addition

$$\vec{v}_i + \vec{v}_j = \vec{v}_j + \vec{v}_i.$$

4. V is **associative** under addition

$$(\vec{v}_i + \vec{v}_j) + \vec{v}_k = \vec{v}_i + (\vec{v}_j + \vec{v}_k)\,.$$

5. V is **distributive** over the reals

$$r(\vec{v}_i + \vec{v}_j) = r\vec{v}_i + r\vec{v}_j.$$

6. There exists a **null vector**, $\vec{0}$, such that

$$\vec{0} + \vec{v} = \vec{v} + \vec{0}.$$

7. There exists an **inverse element**, $-\vec{v}$, for each element $\vec{v} \in V$ such that

$$\vec{v} + (-\vec{v}) = \vec{0}.$$

An important generalization of the concept of a vector space is to include an *inner product*. In this case the vector space V becomes a **normed vector space** or an **inner product space**. The inner product is a law of composition $\vec{v}_1 \cdot \vec{v}_2$ denoted by a dot "$\cdot$" that gives a real number if the vector space is over $\mathbb{R}$. The inner product must satisfy the following properties:

1. The inner product of two vectors $\vec{v}$, and $\vec{w}$ is defined as

$$\vec{v} \cdot \vec{w} = v_1 w_1 + v_2 w_2 + \cdots + v_n w_n \tag{2.1}$$
$$= |\vec{v}||\vec{w}| \cos\theta \tag{2.2}$$

   where the v_i and w_i are the *components* of the vectors, and $|\vec{v}| = \sqrt{\vec{v} \cdot \vec{v}}$, and θ is the angle between the vectors. The inner product is also called the **overlap** between the vectors $\vec{v}$ and $\vec{w}$.

2. The inner or "dot" product of a vector with itself is positive definite

$$\vec{v} \cdot \vec{v} \geq 0.$$

3. The inner product satisfies the Cauchy-Schwarz inequality

$$\vec{v} \cdot \vec{w} \leq |\vec{v}||\vec{w}|. \tag{2.3}$$

2.1.1 Linear combination of vectors

A linear combination of vectors is a sum of the basis vectors with real coefficients. It is written

$$\vec{w} = c_1\vec{v}_1 + c_2\vec{v}_2 + c_3\vec{v}_3 + \cdots.$$

Example: An example of a linear combination of vectors in $\mathbb{R}^3$ is

$$\vec{w} = 2(1,0,0) + 3(0,1,0) + 5(0,0,1)$$

which evaluates to $\vec{w} = (2,3,5)$.

2.1.2 Basis sets

A vector space V has a **basis set** $\{\vec{e}_1, \vec{e}_2, \vec{e}_3, \cdots, \vec{e}_n\}$ that *spans* V. This means that for *any* vector $\vec{v} \in V$ one can write

$$\vec{v} = v_1\vec{e}_1 + v_2\vec{e}_2 + v_3\vec{e}_3 + \cdots + v_n\vec{e}_n. \tag{2.4}$$

The number of elements in the basis set is the **dimension** of V. An orthogonal basis set is one such that

$$\vec{e}_i \cdot \vec{e}_j = \delta_{ij}. \tag{2.5}$$

This means that the inner product of any two *different* basis set vectors is 0, and any two identical basis set vectors is 1. This also means the basis set vectors have unit length. Unit length basis vectors are sometimes denoted with hats, for example in n-space $\{\hat{e}_1, \hat{e}_2, \hat{e}_3, \ldots, \hat{e}_n\}$. We can find the components of a vector $\vec{v}$ by using the inner product of $\vec{v}$ with each the basis set vectors in turn. Note that

$$\begin{aligned}\hat{e}_i \cdot \vec{v} &= \hat{e}_i \cdot (v_1\vec{e}_1 + v_2\vec{e}_2 + v_3\vec{e}_3 + \cdots + v_n\vec{e}_n) \\ &= v_1\,(\hat{e}_i \cdot \hat{e}_1) + v_2\,(\hat{e}_i \cdot \hat{e}_2) + \cdots + v_n\,(\hat{e}_i \cdot \hat{e}_n) \\ &= v_i\end{aligned}$$

and the overlap of $\vec{v}$ with $\hat{e}_i$ is v_i, and is the **projection** of $\vec{v}$ along the direction $\hat{e}_i$ since $\vec{v} \cdot \hat{e}_i = |\vec{v}||\hat{e}_i| \cos\theta$, but $|\hat{e}_i| = 1$, and we find $\vec{v} \cdot \hat{e}_i = v \cos\theta = v_i$, where $v = |\vec{v}|$.

2.1.3 Linear independence

One central concept in vector spaces is that of linear independence of vectors. We say that two vectors $\vec{v}, \vec{w} \in V$ are **linearly independent** if

$$a\vec{v} + b\vec{w} = \vec{0} \tag{2.6}$$

implies that $a = b = 0$. In other words, if the *resultant vector* is zero, and $\vec{v}$ and $\vec{w}$ are *not* zero, and the only way to satisfy (2.6) is to put $a = b = 0$, then the vectors are linearly independent. To see this, let us specialize to $D = 3$ with the familiar basis set $\{\hat{x}, \hat{y}, \hat{z}\}$. Let us take the opposite view and suppose that a and b are *not* zero and determine the consequences. In components, we write

$$\begin{aligned}a\vec{v} + b\vec{w} &= a(v_x\hat{x} + v_y\hat{y}) + b(w_x\hat{x} + w_y\hat{y}) \\ &= (av_x + bw_x)\hat{x} + (av_y + bw_y)\hat{y} \\ &= \vec{0}\end{aligned}$$

From (2.6) we are assuming that the x- and y-components of the resultant vector are zero. Setting the x-component to zero we have $a = -bw_x/v_x$ and from setting the y-component to zero gives $a = -bw_y/v_y$. Setting these two expression equal gives

$$w_x/v_x = w_y/v_y, \tag{2.7}$$

i.e., the vectors are proportional. Note that (2.7) implies $w_x = cv_x$ and $w_y = cv_y$ for some constant c. The vectors are $\vec{v} = (v_x, v_y)$ and $\vec{w} = c(v_x, v_y)$ and they point in the same direction but are of different lengths. Clearly one can be subtracted from the other with a scaling constant to give a resultant vector of zero length. Therefore these vectors are **linearly dependent**.

One can also approach linear dependence or independence from the *inner product* of two vectors. In 3-dimensional Euclidean space we call this the dot product: $\vec{v} \cdot \vec{w} = v_x w_x + v_y w_y + v_z w_z$. In 2-dimensions if the dot product is zero, we have $v_x w_x = -v_y w_y$ or $w_x/w_y = -v_y/v_x$.

If we define the angle of the vector with respect to the x-axis with $\tan(\theta_w) = w_y/w_x$ we see that

$$\begin{aligned}\tan(\theta_w) &= w_y/w_x \\ &= -v_x/v_y \\ &= -(v_y/v_x)^{-1} \\ &= -(\tan(\theta_v))^{-1} \\ &= -\cos(\theta_v)/\sin(\theta_v),\end{aligned}$$

or we must have that

$$\sin(\theta_w)/\cos(\theta_w) = -\cos(\theta_v)/\sin(\theta_v)$$

or $\sin(\theta_w)\sin(\theta_v) + \cos(\theta_w)\cos(\theta_v) = 0$. From your understanding of trigonometry, this can be written $\cos(\theta_w - \theta_v) = 0$. The cosine is zero at $\pi/2$, $3\pi/2$, etc., so the vectors are *perpendicular*.

Example: Calculate the inner product of $\vec{a} = (2, 0)$ and $\vec{b} = (1, 5)$:

$$\vec{a} \cdot \vec{b} = (2)(1) + (0)(5) = 2.$$

Thus the vectors $\vec{a}$ and $\vec{b}$ are not linearly independent because they both have components along the x-axis.

2.1.4 Vector representations in real space

We can represent vectors in matrix format by considering each vector as a column of the components of the vector,

$$\vec{v} = v_1\hat{e}_1 + v_2\hat{e}_2 + \cdots v_n\hat{e}_n = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}. \tag{2.8}$$

We can form another vector called the **transpose**,

$$\vec{v}^T = (v_1\ v_2\ \cdots\ v_n)\,. \tag{2.9}$$

Using the rules of matrix multiplication we see that $\vec{w}^T\vec{v} = \vec{w} \cdot \vec{v}$.

$$\begin{aligned}\vec{w}^T\vec{v} &= (w_1\ w_2\ \cdots\ w_n)\begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} \\ &= w_1v_1 + w_2v_2 + \cdots + w_nv_n.\end{aligned} \tag{2.10}$$

Example: Compute the dot product of $\vec{v} = 2\hat{e}_1 + 3\hat{e}_2$ and $\vec{w} = -2\hat{e}_1 - 4\hat{e}_2$. Using the formula above,

$$\begin{aligned}\vec{w} \cdot \vec{v} &= (-2)(2) + (3)(-4) \\ &= -4 - 12 = -16.\end{aligned}$$

Example: Compute the magnitude squared of $\vec{w} = -2\hat{e}_1 - 4\hat{e}_2$.

$$\begin{aligned}|\vec{w}|^2 &= \vec{w} \cdot \vec{w} \\ &= (-2)(-2) + (-4)(-4) \\ &= 4 + 16 = 20.\end{aligned}$$

2.1.5 Normalizing a vector

A vector has a direction and a magnitude. In general the length of the vector is $|\vec{v}| = \sqrt{v_1^2 + v_2^2 + \cdots}$. It is often useful to make the vector a unit vector, or a vector of length 1. This is easy to do by taking the vector and simply dividing by its magnitude,

$$\hat{v} = \frac{\vec{v}}{|\vec{v}|}$$

where the unit vector $\hat{v}$ carries a "hat" to indicate it is of unit length.

2.2 Complex numbers

Ccomplex numbers are essential in quantum mechanics. We will cover the basics here, but if any of this is foreign to you be sure you master the following few sections.

Recall that a complex number z is represented by $z = x + iy$, where both x and y are real numbers, i.e., $x, y \in \mathbb{R}$, and i is defined by

$$i^2 = -1 \tag{2.11}$$

or $i = \sqrt{-1}$. Multiplying complex numbers is straightforward if you use (2.11). For example, take two complex numbers $z_1 = a + ib$, and $z_2 = c + id$. The **product of two complex** numbers $z_1 z_2$ is given by

$$\begin{aligned}(a + ib)(c + id) &= ac + iad + ibc + i^2 bd \\ &= (ac - bd) + i(ad + bc)\end{aligned}$$

where we used $i^2 = -1$ to get $-bd$ in the second line.

In the complex number system we define the **magnitude squared** of a complex number $z = a + ib$ as

$$|z|^2 = z^* z$$

where

$$z^* = a - ib$$

is the **complex conjugate** of the number $z = a + ib$. Complex conjugation is simply changing the sign of the i in the expression. Computing $|z|^2$ gives

$$\begin{aligned} |z|^2 &= (a - ib)(a + ib) \\ &= a^2 + iab - iab - i^2 b^2 \\ &= a^2 + b^2. \end{aligned}$$

Note that the magnitude squared of a complex number is always positive definite because both of the numbers a, and b, are squared.

For clarity, we can identify the real and imaginary parts of a complex number z as

$$z = z_R + iz_I$$

where z_R is the real part and z_I is the imaginary part of z.

Complex numbers can also be represented in exponential form by making use of the **Euler formula**

$$e^{ia} = \cos a + i \sin a.$$

Representing a complex number by $z = x + iy$, with the coordinates (x, y) in the complex plane, we calculate the distance from the origin using the Pythagorean theorem, $r^2 = x^2 + y^2$. Let the angle θ be from the x-axis to the radial line defined by the point (x, y). Then $x = r\cos\theta$ and $y = r\sin\theta$, and we can write

$$\begin{aligned} z &= x + iy \\ &= r\cos\theta + ir\sin\theta \\ &= re^{i\theta}. \end{aligned}$$

So we can represent the point $z = x + iy$ just as well with $z = re^{i\theta}$ with $r^2 = x^2 + y^2$. The coordinates (r, θ) are the usual polar coordinates on the plane. We will use the Euler formula frequently.

2.3 Vector spaces over $\mathbb{C}$

These are the vector spaces we will use throughout much of the course. In a complex vector space W, the components, or coefficients of a vector $\vec{w} \in W$ are numbers in $\mathbb{C}$ instead of $\mathbb{R}$. Explicitly,

$$\vec{w} = w_1\hat{e}_1 + w_2\hat{e}_2 + \cdots + w_n\hat{e}_n$$

where

$$w_j = (w_j)_R + i\,(w_j)_I\,.$$

The complex conjugate of w_j is

$$w_j^* = (w_j)_R - i\,(w_j)_I\,.$$

One generalizes the inner product from n-dimensional real space to an n-dimensional complex vector space by defining the inner product as

$$\vec{v} \cdot \vec{w} = v_1^* w_1 + v_2^* w_2 + \cdots + v_n^* w_n \tag{2.12}$$

where v_i^* is the complex conjugate of v_i. The inner product is defined with the complex conjugate of one of the vectors so that the inner product of a vector with itself is *positive definite*. In particular

$$\begin{aligned} \vec{w} \cdot \vec{w} &= w_1 w_1^* + w_2 w_2^* + \cdots + w_n w_n^* \\ &= |w_1|^2 + |w_2|^2 + \cdots + |w_n|^2. \end{aligned}$$

In particular the magnitude of a vector with complex coefficients is defined just as in the real vector space, $|\vec{w}|^2 = \vec{w} \cdot \vec{w}$, but the individual terms are the magnitude squared: $|\vec{w}|^2 = \sqrt{|w_1|^2 + |w_2|^2 + \cdots}$.

In matrix notation we can write the inner product as

$$\vec{v}^{\,\dagger} \vec{w} = \begin{pmatrix} v_1^* & v_2^* & \cdots & v_n^* \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix}, \tag{2.13}$$

where

$$\vec{v}^{\,\dagger} = \left(\vec{v}^{\,*}\right)^T. \tag{2.14}$$

The dagger represents the combination of the complex conjugate and transpose.

Example: Compute the vector dot product of $\vec{p} = p_1 \hat{e}_1 + p_2 \hat{e}_2$ and $\vec{q} = q_1 \hat{e}_1 + q_2 \hat{e}_2$, where $\{p_1, p_2, q_1, q_2\} \in \mathbb{C}$, and $p_1 = 2 + i$, $p_2 = 3 - 2i$, $q_1 = 1 - 4i$, and $q_2 = 2 - 2i$.

Using the form of the dot product given in (2.12), we write

$$\begin{aligned} \vec{p} \cdot \vec{q} &= p_1^* q_1 + p_2^* q_2 \\ &= (2 - i)(1 - 4i) + (3 + 2i)(2 - 2i) \\ &= 2 - 8i - i + i^2 4 + 6 - 6i + 4i - 4i^2 \\ &= 8 - 11i. \end{aligned}$$

Example: Compute the magnitude squared of $\vec{p} = p_1 \hat{e}_1 + p_2 \hat{e}_2$ with $p_1 = 2 + i$, and $p_2 = 3 - 2i$.

The magnitude squared of a vector is $|\vec{p}|^2 = \vec{p} \cdot \vec{p}$, so we calculate

$$\begin{aligned} \vec{p} \cdot \vec{p} &= p_1^* p_1 + p_2^* p_2 \\ &= (2 - i)(2 + i) + (3 + 2i)(3 - 2i) \\ &= 4 - i^2 + 9 - 4i^2 \\ &= 4 + 1 + 9 + 4 = 18. \end{aligned}$$

2.4 Switching basis sets

When we switch basis sets for a vector $\vec{v}$ in $\mathbb{R}^n$ it is the *basis set elements* and the *components* of the vector that that change. The vector itself is not changed – it is the same vector as seen by an observer with a different coordinate system. Given two basis sets denoted[†] $\{e_i\}$ and $\{e'_i\}$ we write this as

$$\vec{v} = v_1 e_1 + v_2 e_2 + \cdots + v_n e_n \tag{2.15}$$
$$= v'_1 e'_1 + v'_2 e'_2 + \cdots + v'_n e'_n. \tag{2.16}$$

Because these two expressions are equal, we can take the inner product of (2.15) with e_i to obtain the component v_i. But this must give the same answer as taking the inner product of (2.16) with e_i. We have

$$\begin{aligned} v_i &= e_i \cdot (v'_1 e'_1 + v'_2 e'_2 + \cdots + v'_n e'_n) \\ &= v'_1 e_i \cdot e'_1 + v'_2 e_i \cdot e'_2 + \cdots + v'_n e_i \cdot e'_n. \end{aligned}$$

If we do this for every basis vector e_i we can form the matrix equation

$$\begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} e_1 \cdot e'_1 & e_1 \cdot e'_2 & \cdots & e_1 \cdot e'_n \\ e_2 \cdot e'_1 & e_2 \cdot e'_2 & & e_2 \cdot e'_n \\ \vdots & & \ddots & \vdots \\ e_n \cdot e'_1 & e_n \cdot e'_2 & \cdots & e_n \cdot e'_n \end{pmatrix} \begin{pmatrix} v'_1 \\ v'_2 \\ \vdots \\ v'_n \end{pmatrix}. \tag{2.17}$$

This means that we can compute the components of a vector $\vec{v}$ in a different basis set by calculating the inner products of the basis set elements between the two and forming the matrix in (2.17). In matrix notation we can write

$$\vec{v} = M\vec{v}'$$

and if the matrix is invertible we can write

$$\vec{v}' = M^{-1}\vec{v}.$$

Example Let $\{e_1, e_2\} = \{\hat{x}, \hat{y}\}$, and $\{e'_1, e'_2\} = \{(\hat{x} + \hat{y})/\sqrt{2}, (\hat{x} - \hat{y})/\sqrt{2}\}$. Suppose that $\vec{v}' = (1/\sqrt{2}, 1/\sqrt{2})^T$. Then $M = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix}$ and $v = (1, 0)^T$.

2.5 Matrix multiplication

We have already introduced matrices, but they are so important that we remind ourselves of some of the basic properties.

[†]We drop the hats in this section to reduce clutter.

2.5.1 Matrix multiplication on a vector

We have seen in section 2.4 the action of a matrix on a vector. We will also encounter matrices in the context of rotations, for example. I expect that you are familiar with standard manipulations such as matrix multiplication and matrix on vector multiplication. For a matrix M multiplying a vector $\vec{v}$ from the left, we write $\vec{w} = M\vec{v}$. We write this in component notation as

$$w_i = \sum_k M_{ik} v_k. \tag{2.18}$$

Here $\vec{v}$ is a column vector, so the index k on matrix M must be the *column index*. Therefore i labels the rows of M. We define the **matrix inverse** M^{-1} from the equation $MM^{-1} = I$. If M is a 2×2 matrix we can write its inverse as

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

where the **determinant** of the 2×2 matrix above is given by $(ad - bc)$. In general the determinant of a matrix is defined as the *sum of all signed elementary products of a matrix*. A matrix whose rows are linearly *dependent* will have no inverse because the determinant will vanish.

2.5.2 Rotation matrices

Let us consider a rotation matrix operating on a vector. We denote the matrix as R. For orthogonal rotations we have $R^T = R^{-1}$, and $\det R = +1$ for proper rotations. We write

$$\vec{v}\,' = R\vec{v}. \tag{2.19}$$

We can use (2.18) and write the j-th component of $R\vec{v}$

$$(R\vec{v})_j = \sum_i R_{ji} v_i \tag{2.20}$$

or in matrix notation

$$R\vec{v} = \begin{pmatrix} R_{11} & \cdots & R_{1n} \\ \vdots & \ddots & \vdots \\ R_{n1} & \cdots & R_{nn} \end{pmatrix} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} v_1' \\ \vdots \\ v_n' \end{pmatrix}$$

and let the matrix R *operate on the components* of the vector in some basis set $\{\vec{e}_1, \cdots, \vec{e}_n\}$. This is one definition of the rotation matrix. Alternatively we can write

$$R\vec{v} = R\left(v_1\vec{e}_1 + v_2\vec{e}_2 + \cdots + v_n\vec{e}_n\right) \tag{2.21}$$

and let the matrix R *operate on the basis vectors* themselves. We have

$$R\vec{v} = v_1\left(R\vec{e}_1\right) + v_2\left(R\vec{e}_2\right) + \cdots v_n\left(R\vec{e}_n\right) \tag{2.22}$$

and the matrix R acts on the basis vectors.

Example: In two dimensions let R be the rotation matrix that rotates vectors by 90 degrees counterclockwise. Let this matrix act on the vector $(1, 0)$ that points along the x-direction:

$$R\hat{e}_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \hat{e}_2.$$

Thus, if we let R act *on the basis vectors themselves* then we could write instead that

$$R\vec{e_i} = \sum_j R_{ij}\vec{e}_j = \vec{e}_i\,' \tag{2.23}$$

where this equation says that R rotates the vector $\vec{e}_i$ into some linear combination of the other $\vec{e}_j$.[†] Then we have

$$\begin{aligned} R\vec{v} &= R\sum_i v_i\vec{e}_i = \sum_i v_i(R\vec{e}_i) \\ &= \sum_i v_i \left(\sum_j R_{ij}\vec{e}_j\right) \\ &= \sum_j \left(\sum_i R_{ij}v_i\right)\vec{e}_j \\ &= \sum_j (R^T\vec{v})_j\vec{e}_j \\ &= \sum_j v'_j\vec{e}_j. \end{aligned} \tag{2.24}$$

So the components of the new vector here are

$$\vec{v}\,' = R^T\vec{v}. \tag{2.25}$$

The difference between (2.19) and (2.25) is that we are rotating the vector itself in (2.19) and rotating the basis vectors in (2.25). We expect them to rotate in opposite directions.

2.5.3 Einstein summation convention

Albert Einstein introduced a summation convention so that one does not need to keep writing the summation symbol $\sum$. The rule is simply that any index that appears twice on the same side of an equation is summed over. Therefore we write (2.19) as

$$v'_i = R_{ij}v_j.$$

We will use this convention from now on.

[†]Note that dotting both sides of (2.23) with $\vec{e}_k$ gives $R_{ik} = \vec{e}_i\,' \cdot \vec{e}_k$.

2.5.4 Matrix multiplication

Multiplying two matrices is similar to matrix multiplication of vectors. Let's multiply the following 2×2 matrices

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} B = \begin{pmatrix} e & f \\ g & h \end{pmatrix}.$$

The product is

$$AB = \begin{pmatrix} ae + bg & af + bh \\ ce + dg & cf + dh \end{pmatrix}$$

where the first entry at the top left is given by vector multiplication of the first row vector in A, which is (a, b) and the column vector $(e, g)^T$. The entry on the top right is given by the vector multiplication of the row vector (a, b) and the column vector $(f, h)^T$. The second row is given by multiplying the second row vector in A.

In components we write

$$(AB)_{ij} = \sum_k A_{ik} B_{kj}.$$

So that

$$\begin{aligned} (AB)_{11} &= A_{11}B_{11} + A_{12}B_{21} \\ &= ae + bg. \end{aligned}$$

The rest of the entries are computed similarly.

Example: We compute the product of the following matrices. Calculate this yourself to make sure you understand it.

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 3 & 1 \\ 5 & 6 \end{pmatrix} = \begin{pmatrix} 13 & 13 \\ 29 & 27 \end{pmatrix}.$$

2.5.5 Matrix determinant

The *determinant* of a matrix is the *sum of all signed elementary products* of the matrix. This means the sum over all combinations of products of one element from each row (or column) with a sign depending on the even or odd permutation of the column (or row) ordering. This can be written very succinctly with the antisymmetric symbol. For a $n \times n$ matrix A, we write

$$\det(A) = \epsilon^{i_1 i_2 \cdots i_n} A^{1 i_1} A^{2 i_2} \cdots A^{N i_n}. \tag{2.26}$$

The symbol $\epsilon^{i_1 i_2 \cdots i_n}$ is the called the **antisymmetric symbol**. It is antisymmetric in its indices. For example, for two the two index symbol ϵ^{ij} for the values of $i, j \in \{1, 2\}$, the symbol takes the following values: $\epsilon^{12} = 1$, $\epsilon^{21} = -1$, with all other values being zero.

Example: For a two dimensional matrix the determinant is thus:

$$\begin{aligned}\det\begin{pmatrix} A^{11} & A^{12} \\ A^{21} & A^{22} \end{pmatrix} &= \epsilon^{12}A^{11}A^{22} + \epsilon^{21}A^{12}A^{21} \\ &= A^{11}A^{22} - A^{12}A^{21}.\end{aligned}$$

We can derive some useful formulas using the determinant, such as the formula $\det(I + \epsilon A) = 1 + \epsilon \operatorname{tr}(A)$, where $\epsilon^2 = 0$, for infinitesimal ϵ. We can write the matrix elements $(I + \epsilon A)_{ij} = \delta_{ij} + \epsilon A^{ij}$. Our formula for the determinant looks like

$$\det(I + \epsilon A) = \epsilon^{i_1 i_2 \cdots i_n}\left(\delta_{1i_1} + \epsilon A^{1i_1}\right)\left(\delta_{1i_2} + \epsilon A^{1i_2}\right)\cdots\left(\delta_{1i_n} + \epsilon A^{1i_n}\right).$$

All terms with powers of ϵ greater than one vanish. Trying this for a 3×3 matrix gives

$$\begin{aligned}&\epsilon^{i_1 i_2 i_3}(\delta_{1i_1} + \epsilon A^{1i_1})(\delta_{2i_2} + \epsilon A^{2i_2})(\delta_{3i_3} + \epsilon A^{3i_3}) \\ &= \epsilon^{i_1 i_2 i_3}\left(\delta_{1i_1}\delta_{2i_2}\delta_{3i_3} + \delta_{1i_1}\delta_{2i_2}\epsilon A^{3i_3} + \delta_{1i_1}\delta_{3i_3}\epsilon A^{2i_2} + \delta_{2i_2}\delta_{3i_3}\epsilon A^{1i_1}\right) \\ &= \epsilon^{i_1 i_2 i_3}\delta_{1i_1}\delta_{2i_2}\delta_{3i_3} + \epsilon^{i_1 i_2 i_3}\delta_{1i_1}\delta_{2i_2}\epsilon A^{3i_3} + \cdots \\ &= 1 + \epsilon A^{33} + \epsilon A^{22} + \epsilon A^{11}.\end{aligned}$$

2.6 Linear transformations on vectors

The transformation of vectors by **matrix operators** does not always mean a rotation or a basis change. The matrix could represent a projection operator, for example, or some other operator. We are interested in particular in **linear transformations**. Denote a matrix transformation operator as T. We are interested in operators that take our vector space V to itself, i.e. $T : V \to V$. Let $\vec{\alpha}, \vec{\beta} \in V$ be vectors in V. For a linear operator we have

$$T(a\vec{\alpha} + b\vec{\beta}) = a(T\vec{\alpha}) + b(T\vec{\beta}).$$

Following section 2.5 and writing $\vec{\alpha} = \alpha_i \vec{e}_i$, we have

$$T\vec{\alpha} = \alpha_i T\vec{e}_i$$

and we need to know how T transforms the basis vectors. For each basis vector we have $T\vec{e}_i = T_{ij}\vec{e}_j$ just as in (2.23) and so

$$\vec{\alpha}\,' = T^T\vec{\alpha}. \tag{2.27}$$

2.6.1 Basis change of a linear transformation T

We will now determine how the matrix representation of a linear transformation T changes if we change basis vectors. In (2.27) the vectors $\vec{\alpha}'$ and $\vec{\alpha}$ are simply two vectors in V. Under a basis change we have seen in (2.25) that $\vec{v} \to R^T\vec{v}$. Therefore (2.27) goes to

$$R^T\vec{\alpha}\,' = T^T R^T\vec{\alpha}.$$

For simplicity put $L = T^T$, and noting that $\left(R^T\right)^{-1} = R$, we have

$$R^T \vec{\alpha}\,' = LR^T \vec{\alpha}.$$

Multiplying both sides by $\left(R^T\right)^{-1}$ we find that

$$L \to RLR^T. \tag{2.28}$$

This is called a **similarity transform** and it is used very often in quantum mechanics.

2.7 Systems of linear equations

A system of linear equations can often be represented in matrix format. For example the equations

$$a_{11}x_1 + a_{12}x_2 = p_1$$
$$a_{21}x_1 + a_{22}x_2 = p_2$$

can be represented as

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} p_1 \\ p_2 \end{pmatrix}$$

or more compactly as

$$A\vec{x} = \vec{p}. \tag{2.29}$$

The solution to this set of equations can be done slowly using brute force substitution. A better approach is to use Gaussian elimination to reduce the matrix to something more manageable. One of the most important methods to solve the system is by finding the **matrix inverse** of A. For the matrix to have an inverse, the rows must be *linearly independent*. A quick way to check this is to calculate the *matrix determinant*. If $\det A \neq 0$, then an inverse can be found. If one can calculate the matrix inverse, the solution can be found by multiplying both sides of (2.29) by A^{-1} giving

$$\vec{x} = A^{-1}\vec{p}.$$

Finding the matrix inverse of small matrices, say 2×2 matrices, is possible by hand. A formula worth remembering is

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix},$$

noting that $(ad - bc) = \det \begin{pmatrix} a & b \\ c & d \end{pmatrix}$.

2.8 Eigenvalue equations

An **eigenvalue equation** is of the form

$$M\vec{x} = \alpha\vec{x} \tag{2.30}$$

where M is an $n \times n$ matrix, $\vec{x}$ is an n-component vector, and α is a number called the **eigenvalue**. This means that when M operates on $\vec{x}$ it gives back *the same vector* $\vec{x}$ only multiplied by a constant. This type of equation has enormous importance in quantum mechanics and we will solve eigenvalue problems repeatedly in different circumstances. The first step in extracting the information in an eigenvalue equation is to calculate the eigenvalues. This is done by writing (2.30) in the form

$$M\vec{x} = \lambda I\vec{x}$$

where I is the $n \times n$ identity matrix, and λ is a *variable*, not to be confused with the specific eigenvalue α in (2.30) above. Then we can move r.h.s. to the l.h.s. and write

$$(M - \lambda I)\,\vec{x} = \vec{0}. \tag{2.31}$$

Note that the r.h.s. of this equation is the zero vector and consider what this must mean for the l.h.s. of the equation to satisfy this equality. It could be solved by $\vec{x} = 0$, which is called the **trivial solution** and is of little interest to us. Otherwise the matrix $(M - \lambda I)$ must be something like "zero". In contrast to the system of equations in (2.29) where we demanded that the determinant *not* be zero, in the case of an eigenvalue equation we *require* the determinant of $(M - \lambda I)$ to be zero. Solving

$$\det(M - \lambda I) = 0 \tag{2.32}$$

will give a polynomial of degree n on the l.h.s. of (2.32), the solution of which will give a set of n values $\{\lambda_i\}$ that give a zero determinant. These are the eigenvalues of the matrix M.

Example Consider the equation

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\vec{x} = \lambda\vec{x}.$$

We require

$$\det\left[\begin{pmatrix} a & b \\ c & d \end{pmatrix} - \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}\right] = 0$$

or

$$\begin{aligned} \det\begin{pmatrix} a-\lambda & b \\ c & d-\lambda \end{pmatrix} &= 0 \\ (a-\lambda)(d-\lambda) - bc &= 0. \end{aligned}$$

We solve the quadratic in λ for the solutions in terms of a, b, c, d. □

Once the set of eigenvalues $\{\lambda_i\}$ is in hand, one then computes the i-th **eigenvector** of M by using one of the specific eigenvalues λ_i and solving

$$M\vec{v}_i = \lambda_i \vec{v}_i$$

where $\vec{v}_i$ is the i-th eigenvector and its components are undetermined. When solving for the components of an eigenvector, the components are variable so that they can be solved for. The eigenvector $\vec{v}_i$ is tied specifically to the eigenvalue λ_i.

Example: If $M = \begin{pmatrix} 3 & 2 \\ 3 & 4 \end{pmatrix}$ then we have from the example above $\lambda^2 - \lambda(a+d) + \det M = 0$. The solutions are

$$\lambda = \frac{(a+d) \pm \sqrt{(a+d)^2 - 4\det M}}{2}$$

or $\lambda = \frac{1}{2}(7 \pm \sqrt{(49-24)}) = (7 \pm 5)/2$ or $\lambda = \{1, 6\}$. Then one would solve the matrix equations corresponding to each eigenvalue,

$$\begin{pmatrix} 3 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 1 \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \qquad \text{and}$$

$$\begin{pmatrix} 3 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 6 \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

For the eigenvalue $\lambda = 1$ we multiply out the top matrix,

$$\begin{pmatrix} 3x_1 + 2x_2 \\ 3x_1 + 4x_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

There are two equations here,

$$3x_1 + 2x_2 = x_1, \qquad \text{and}$$
$$3x_1 + 4x_2 = x_2.$$

The first equation is solved yielding $x_2 = -x_1$, thus giving the eigenvector $(x_1, -x_1)^T$. Since x_1 can be anything we can set it to 1 giving the eigenvector $(1, -1)^T$ which can be normalized. Thus

$$\lambda = 1 \quad \longleftrightarrow \quad \vec{x} = \begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{pmatrix}.$$

Using the eigenvalue equation for $\lambda = 6$ we would write

$$3x_1 + 2x_2 = 6x_1$$
$$3x_1 + 4x_2 = 6x_2$$

and solve for (x_1, x_2) corresponding to the eigenvalue $\lambda_2 = 6$. We have for each of these equations

$$3x_1 = 2x_2$$

or the corresponding non-normalized vector is $(1, 3/2)^T$. The magnitude of this vector is $\sqrt{1+9/4} = \sqrt{13}/2$ and the normalized eigenvector is

$$\lambda = 6 \quad \longleftrightarrow \quad \vec{x} = \begin{pmatrix} 0.5547 \\ 0.8321 \end{pmatrix}. \quad \square$$

These are the simplest cases of matrix equations. Problems that are commonly encountered include things such as multiple identical eigenvalues that indicate some kind of degeneracy in the solution of the matrix equation. We will encounter these in the course and discuss the procedure to deal with them.

2.9 Matrix Diagonalization with Eigenvectors

Suppose we are given a matrix A we wish to diagonalize it. In order to do this there needs to be a basis set such that we can find the eigenvalues and eigenvectors using A. In other words we assume that for an $n \times n$ matrix A, there exists n vectors $\vec{v}^{(j)}$ such that $A\vec{v}^{(j)} = \lambda_j \vec{v}^{(j)}$. By component we have $A_{ik}\vec{v}_k^{(j)} = \lambda_j \vec{v}_i^{(j)}$. It is easiest to see this if you write out the matrix and vectors and look at the i^{th} row of A.

$$\begin{pmatrix} A_{11} & A_{12} & A_{13} & \dots & A_{1n} \\ & \ddots & & & \\ A_{i1} & A_{i2} & & & A_{in} \\ & & & & \\ A_{n1} & A_{n2} & & & A_{nn} \end{pmatrix} \begin{pmatrix} v_1^{(j)} \\ v_2^{(j)} \\ \vdots \\ \\ v_n^{(j)} \end{pmatrix} = \lambda_j \begin{pmatrix} v_1^{(j)} \\ v_2^{(j)} \\ \vdots \\ \\ v_n^{(j)} \end{pmatrix}. \tag{2.33}$$

Start by solving for the eigenvalues and eigenvectors in the usual way. Once this is completed, construct a matrix V out of the column eigenvectors then $V^T AV$ is diagonal. This is easy to see looking at $V^T(AV)$ which looks like:

$$\begin{pmatrix} v_1^{(1)} & \dots & & \\ & & & \\ v_1^{(i)} & v_2^{(i)} & \dots & v_n^{(i)} \\ & & & \\ & & \dots & v_n^{(n)} \end{pmatrix} \begin{pmatrix} \lambda_1 v_1^{(1)} & \dots & \lambda_i v_1^{(i)} & \\ & & \lambda_i v_2^{(i)} & \\ & & & \\ & & \lambda_i v_n^{(i)} & \dots \quad \lambda_n v_n^{(n)} \end{pmatrix} = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix} \tag{2.34}$$

and clearly we assume that we have orthogonality $\vec{v}^{(i)} \cdot \vec{v}^{(j)} = \delta_{ij}$. This result is general for matrices over the field $\mathbb{R}$. If the matrices are complex, we need to use the Hermitian conjugate,[†] which is the complex conjugate transpose, instead of just the transpose.

Example: Complex-valued matrix

[†]The Hermitian conjugate of a matrix is the complex conjugate of the transpose of the matrix, $H^\dagger = H^{t*}$, denoted with a dagger.

Consider a matrix $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ defined over $\mathbb{C}$. We need to diagonalize the matrix, so we calculate its eigenvalues and eigenvectors. Form $\det(A - \lambda I) = 0$. The matrix is $\begin{pmatrix} -\lambda & -1 \\ 1 & -\lambda \end{pmatrix}$, and the characteristic equation is $\lambda^2 + 1 = 0$, with solutions $\lambda_\pm = \pm i$. We now need to find the eigenvectors, keeping in mind that they are complex.

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} u_r + iu_i \\ v_r + iv_i \end{pmatrix} = \pm i \begin{pmatrix} u_r + iu_i \\ v_r + iv_i \end{pmatrix}. \tag{2.35}$$

For the positive eigenvalue, we have $-v_r - iv_i = iu_r - u_i$, or $u_i = v_r$,and $v_i = -u_r$. The second equation from the first eigenvalue is $u_r + iu_i = iv_r - v_i$, giving the same result. Using the eigenvalue $-i$ we find $u_r = v_i$, and $u_i = -v_r$. We have the following non-normalized eigenvectors: $\begin{pmatrix} 1+i \\ 1-i \end{pmatrix}$, and $\begin{pmatrix} 1+i \\ -1+i \end{pmatrix}$. Note that the inner product is $\langle w_1, w_2 \rangle = \sum w_{1i}^* w_{2i}$. If we form now the matrix V whose column vectors correspond to the two eigenvalues, we have

$$V = \begin{pmatrix} 1+i & 1+i \\ 1-i & -1+i \end{pmatrix}. \tag{2.36}$$

The Hermitian conjugate of V is $V^\dagger = V^{T*}$, the complex conjugate transpose.

$$V^\dagger = \begin{pmatrix} 1+i & 1-i \\ 1+i & -1+i \end{pmatrix}^* = \begin{pmatrix} 1-i & 1+i \\ 1-i & -1-i \end{pmatrix}. \tag{2.37}$$

Calculating $V^\dagger AV$ we find

$$AV = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1+i & 1+i \\ 1-i & -1+i \end{pmatrix} = \begin{pmatrix} -1+i & 1-i \\ 1+i & 1+i \end{pmatrix} \tag{2.38}$$

and

$$\begin{pmatrix} 1-i & 1+i \\ 1-i & -1-i \end{pmatrix} \begin{pmatrix} -1+i & 1-i \\ 1+i & 1+i \end{pmatrix} = \begin{pmatrix} 4i & 0 \\ 0 & -4i \end{pmatrix}. \tag{2.39}$$

The magnitude of each eigenvector is 2, so we get a factor of $1/2$ for each matrix above, and dividing the RHS by 4 we find $\Lambda = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$, and we have diagonalized the matrix. We can now use $V^{-1} e^\Lambda V = e^A$ where $\Lambda = V^\dagger AV$. Clearly $e^{t\Lambda} = \begin{pmatrix} e^{it} & 0 \\ 0 & e^{-it} \end{pmatrix}$, and we get

$$\frac{V^\dagger AV}{4} = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}. \tag{2.40}$$

2.10 Unitary and Hermitian Operators

The Hermitian conjugate of a matrix is defined as the complex conjugate of the transpose of the matrix

$$H^\dagger = (H^T)^* \qquad \text{(def. of Hermitian conjugate)}$$

and denoted with a dagger. A matrix is called **Hermitian** if it is equal to its Hermitian conjugate,

$$H^\dagger = H \qquad \text{(def. of a Hermitian matrix)}.$$

Hermitian matrices are important in quantum mechanics because their eigenvalues are *real* (have no imaginary component), and so they are associated with *observables*, things that can be measured.

A **unitary matrix**, often denoted U, satisfies the following property

$$U^\dagger U = I,$$

which implies that $U^\dagger = U^{-1}$, the Hermitian conjugate is the inverse of the matrix. This implies in particular that U preserves the length of vectors. Suppose a is a unit vector, $a^\dagger a = 1$. Then operate on a with a unitary operator U, giving $b = Ua$. By definition the Hermitian conjugate of b is $b^\dagger = a^\dagger U^\dagger$, and therefore

$$b^\dagger b = a^\dagger U^\dagger U a = a^\dagger a = 1$$

and b is also a unit vector. Unitarity is a very special property of operators in quantum mechanics because it preserves the total probability. This must be so, otherwise information would be lost if time evolution, for example, were not unitary.

Any unitary matrix can be written as the exponential of a Hermitian matrix, $U = e^{iH}$. We can show this by expanding the exponential,

$$e^{iH} \cong 1 + iH - \frac{H^2}{2} + \cdots$$

and using $(H^n)^\dagger = (H^\dagger)^n$, and $H^\dagger = H$. Then $(e^{iH})^\dagger = e^{-iH}$. Then

$$\begin{aligned} U^\dagger U &= e^{-iH} e^{iH} \\ &= 1. \end{aligned}$$

To obtain the last line we have to use the fact that H commutes with itself.

3 Linear algebra in quantum mechanics

Vector spaces in quantum mechanics have many similarities to those in finite-dimensional Euclidean real and complex vector spaces. In particular *the concepts are the same*. There are vector spaces that are spanned by a basis set that can be used to decompose vectors. There are inner products of vectors, and "overlap" between vectors can be quantified.

3.1 Dirac notation

We will use different notation for vectors in quantum mechanics, in particular what is called **Dirac notation** after the famous physicist Paul Dirac who invented it. Dirac notation is an especially powerful notation; learning and adopting this notation is well worth the effort. In quantum mechanics, the vector spaces are necessarily complex, and below we generalize the notation in section 2.3. We write a vector $\vec{v}$ as

$$\vec{v} \equiv |v\rangle$$

and its complex conjugate transpose as

$$\vec{v}^{\,\dagger} \equiv \langle v| \,.$$

The inner product of two vectors is written

$$\begin{aligned} \vec{v}^{\,\dagger}\vec{w} &= \langle v|w\rangle \\ &= v_1^* w_1 + v_2^* w_2 + \cdots + v_n^* w_n. \end{aligned} \tag{3.1}$$

This implies an important property of brackets in quantum mechanics,

$$\langle v|w\rangle = \langle w|v\rangle^*$$

which follows because $(v_1^* w_1 + \cdots) = (v_1 w_1^* + \cdots)^*$.

We can expand a vector over a basis set in the usual way. If the basis set is $\{|\alpha_1\rangle, |\alpha_2\rangle, \cdots, |\alpha_n\rangle\}$ then we would write

$$|v\rangle = v_1 |\alpha_1\rangle + v_2 |\alpha_2\rangle + \cdots + v_n |\alpha_n\rangle$$

In quantum mechanics the vector spaces *can* be rather simple like those finite-dimensional vector spaces we have those discussed above. However, they can also be considerably more general. Quantum mechanical vector spaces can be spanned an infinite basis set, for example, an infinite number of continuous complex functions that are orthogonal. To give an example of a space spanned by functions let $f(x)$ and $g(x)$ be two complex-valued functions in this space (not necessarily basis functions). Then the inner product is defined as

$$\langle f \,|\, g\rangle = \int dx\, f^*(x) g(x) \tag{3.2}$$

in complete analogy with (3.1), the difference being a discrete sum going into a integral over continuous functions. One can consider a continuous function of one variable, $f(x)$, as a column vector with an (uncountably) infinite number of components.

3.2 Switching basis sets for spin-1/2

In the following discussion we consider the Stern-Gerlach experiments introduced in class. We have established that $|z+\rangle$ and $|z-\rangle$ span the states of an SGz experiment. We also know that

$$|\langle x \pm |z\pm\rangle|^2 = 1/2 \tag{3.3}$$

We can use (3.3) along with $\langle x - |x-\rangle = 1$ and $\langle x - |x+\rangle = 0$ to obtain the basis vectors $|x\pm\rangle$in terms of $|z\pm\rangle$. A good guess based on (3.3) is to try[†]

$$|x+\rangle = \frac{1}{\sqrt{2}}|z+\rangle + \frac{e^{i\alpha}}{\sqrt{2}}|z-\rangle$$
$$|x-\rangle = \frac{1}{\sqrt{2}}|z+\rangle + \frac{e^{i\beta}}{\sqrt{2}}|z-\rangle .$$

The coefficients of each equation above satisfy $|c_1|^2 + |c_2|^2 = 1$, as is necessary. We also require $\langle x + |x-\rangle = 0$. This gives

$$\frac{1}{2} + \frac{e^{i(\beta-\alpha)}}{2} = 0$$

which is satisfied by $\beta - \alpha = \pi$. We can choose $\alpha = 0$ and $\beta = \pi$. This gives

$$|x\pm\rangle = \frac{1}{\sqrt{2}}|z+\rangle \pm \frac{1}{\sqrt{2}}|z-\rangle \tag{3.4}$$

There are three directions in space and we have taken care of two of them. Now we need another set of basis vectors along the y-direction. We require $|\langle z \pm |y\pm\rangle|^2 = |\langle x \pm |y\pm\rangle|^2 = 1/2$, as well as the usual orthonormality of $|y\pm\rangle$. Let us try

$$|y+\rangle = \frac{1}{\sqrt{2}}|z+\rangle + \frac{e^{i\beta}}{\sqrt{2}}|z-\rangle$$
$$|y-\rangle = \frac{1}{\sqrt{2}}|z+\rangle + \frac{e^{i\gamma}}{\sqrt{2}}|z-\rangle .$$

This is designed so that $|\langle z \pm |y\pm\rangle|^2 = 1/2$. The values of β and γ must give kets $|y\pm\rangle$ that are orthogonal to each other but also satisfy $|\langle x + |y+\rangle|^2 = 1/2$.

$$\begin{aligned}|\langle x + |y+\rangle|^2 &= \frac{1}{4}\left(1 + e^{i\beta}\right)\left(1 + e^{-i\beta}\right) \\ &= \frac{1}{4}\left(2 + 2\cos\beta\right).\end{aligned}$$

For this to be equal to $1/2$ we need the cosine term to be zero, requiring $\beta = \pi/2$. This gives

$$|y+\rangle = \frac{1}{\sqrt{2}}|z+\rangle + \frac{i}{\sqrt{2}}|z-\rangle .$$

[†]We could be very general and use $|x+\rangle = \frac{a}{\sqrt{2}}|z+\rangle + \frac{b}{\sqrt{2}}|z-\rangle$ with $|a|^2 + |b|^2 = 1$ and deduce that $a = 1$ and $b = e^{i\alpha}$.

The orthogonality of the $|y\pm\rangle$ requires

$$\langle y+|y-\rangle = \frac{1}{2}\left(1 - ie^{i\gamma}\right) = 0$$

which is satisfied if $e^{i\gamma} = -i$ or $\gamma = -\pi/2$. Finally we have

$$|y\pm\rangle = \frac{1}{\sqrt{2}}|z+\rangle \pm \frac{i}{\sqrt{2}}|z-\rangle\,. \tag{3.5}$$

3.3 Change of basis for matrix representations of operators

The matrix representation of the S_z operator is given by

$$S_z = \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

in the z-basis. We would like to express this operator in the x-basis. We could do this by brute force by writing

$$(S_{11})_{x\text{-basis}} = \langle z+|S_z|z+\rangle$$

and inserting 1 in the form $1 = |x+\rangle\langle x+| + |x-\rangle\langle x-|$ on the left and right side of S_z on the rhs. Computing the result for S_{11} gives

$$S_{11} = \frac{1}{2}\left(S'_{11} + S'_{22} + S'_{21} + S'_{12}\right)$$

where the primes indicate elements in the x-basis. Let's do this more generally by writing one basis as $\{|i\rangle\}$ and the other basis as $\{|i'\rangle\}$. Switching bases amounts to writing $|i\rangle = \sum_{i'}\langle i'|i\rangle\,|i'\rangle$. This gives the transformation of the components of an operator S

$$S_{ij} = \langle i|S|j\rangle = \sum_{i'j'}\langle i|i'\rangle\langle i'|S|j'\rangle\langle j'|j\rangle\,.$$

Let us define $\alpha_{ii'} = \langle i|i'\rangle$, and $S_{i'j'} = \langle i'|S|j'\rangle$, then we can write

$$S_{ij} = \sum_{i'j'}\alpha_{ii'}S_{i'j'}\alpha^*_{jj'}$$

and if we further take the transpose of $\alpha^*_{jj'}$ to make the matrix multiplication explicit we have

$$S_{ij} = \sum_{i'j'}\alpha_{ii'}S_{i'j'}\alpha^{*T}_{j'j}.$$

But the conjugate transpose is the Hermitian conjugate and we can then write

$$S = \alpha S'\alpha^\dagger$$

or

$$S' = \alpha^\dagger S\alpha.$$

In our case for S_z we have

$$\begin{aligned}\alpha &= \begin{pmatrix} \langle z+|x+\rangle & \langle z+|x-\rangle \\ \langle z-|x+\rangle & \langle z-|x-\rangle \end{pmatrix} \\ &= \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\end{aligned}$$

and

$$\alpha^\dagger = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

Then

$$\begin{aligned}(S_z)_{x\text{-basis}} &= \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \\ &= \frac{\hbar}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.\end{aligned}$$

Why would we bother to do this? Representing the operator S_z in the x-basis allows us to evaluate an expression like $S_z |\phi\rangle$ where $\phi = a_+ |x+\rangle + a_- |x-\rangle$ is represented in the x-basis. Just to check that this gives us the correct answer, we can calculate $S_z |z+\rangle$ where both are written in the x-basis. Given what we have above this would be

$$\begin{aligned}S_z |z+\rangle &= (S_z)_{x\text{-basis}} \frac{1}{\sqrt{2}}(|x+\rangle + |x-\rangle) \\ &= \frac{\hbar}{2}\frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \left[\begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right] \\ &= \frac{\hbar}{2}\frac{1}{\sqrt{2}} \left[\begin{pmatrix} 0 \\ 1 \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right] \\ &= \frac{\hbar}{2} |z+\rangle .\end{aligned}$$

We obtain the correct result. This approach is used to illustrate how a change of basis affects the matrix representation of operators, and that it is computationally useful to move back and forth between basis sets.

3.4 Observables and expectation values

Suppose we have an observable $\hat{A}$, e.g. the spin of an electron along the z-direction, $\hat{S}_z$ is an observable. Further suppose that there is a basis set $\{|a_1\rangle, |a_2\rangle, \cdots, |a_N\rangle\}$ in which

$$\hat{A} |a_i\rangle = \lambda_i |a_i\rangle$$

with eigenvalues $\{\lambda_1, \lambda_2, \cdots\}$. An arbitrary state in the basis $\{|a_1\rangle, |a_2\rangle, \cdots, |a_N\rangle\}$ is given by

$$|\psi\rangle = c_1 |a_1\rangle + c_2 |a_2\rangle + \cdots + c_N |a_N\rangle$$

with the coefficients $c_i \in \mathbb{C}$, the complex numbers.

We define the **expectation value** of an observable $\hat{A}$ in the state $|\psi\rangle$ as

$$\langle\psi|\hat{A}|\psi\rangle \equiv \langle\hat{A}\rangle_\psi .$$

For example, suppose we have only a 2-dimensional Hilbert space spanned by the two kets $\{|a_1\rangle, |a_2\rangle\}$. Then an arbitrary state can be written $|\psi\rangle = c_1|a_1\rangle + c_2|a_2\rangle$. The expectation value of $\hat{A}$ in this state is

$$\begin{aligned}\langle\hat{A}\rangle_\psi &= (c_1^*\langle a_1| + c_2^*\langle a_2|)\,\hat{A}\,(c_1|a_1\rangle + c_2|a_2\rangle)\\ &= (c_1^*\langle a_1| + c_2^*\langle a_2|)\,(c_1\lambda_1|a_1\rangle + c_2\lambda_2|a_2\rangle)\\ &= |c_1|^2\lambda_1 + |c_2|^2\lambda_2.\end{aligned}$$

The **uncertainty** or **variance** is defined by

$$\Delta A = \sqrt{\langle\hat{A}^2\rangle - \langle\hat{A}\rangle^2}.$$

This definition should remind you of the variance of a distribution you learned in statistical mechanics.

3.5 Momentum and the translation operator

For any smooth function $f(x)$ one can Taylor expand the function about the point x. Writing this out gives

$$f(x+a) = f(x) + af'(x) + \frac{a^2}{2!}f''(x) + \frac{a^2}{3!}f'''(x) + \cdots$$

but this can also be written as a summation

$$f(x+a) = \sum_{n=0}^{\infty}\frac{a^n}{n!}\left(\frac{\partial}{\partial x}\right)^n f(x).$$

The function on the r.h.s. does not depend on n and so can be pulled out to the right of the summation, leaving

$$f(x+a) = e^{a\frac{\partial}{\partial x}}f(x). \tag{3.6}$$

We can write the operator in exponential form, let us call it $t_x(a) = e^{a\frac{\partial}{\partial x}}$. Note in particular that $t_x(a)$ translates the function $f(x)$ by a in the direction of $-\hat{x}$. Take the function value $f(0)$ before translation; it is translated to $x = -a$. Translating to the right we would need $e^{-a\partial_x}$.

In QM the momentum operator is written as

$$\hat{p} = -i\hbar\frac{\partial}{\partial x}$$

where $\partial/\partial x$ may be written more simply as ∂_x. If we introduce $1 = -i\cdot i$ into the exponent of (3.6) we can write the new operator that translates to the right as

$$\begin{aligned}\mathscr{T}_x(a) &= e^{-a(-i\cdot i)\partial_x} = e^{-ia/\hbar(-i\hbar\partial_x)}\\ &= e^{-ia\frac{\hat{p}}{\hbar}}.\end{aligned}$$

The operator $\mathscr{T}_x(a)$ is the translation operator in QM.

3.6 Angular momentum

Suppose we have a wave function in 2D space, $\psi = \psi(x, y)$. Now suppose we perform a rotation in the plane (about the z-axis). From elementary geometry it is easy to show

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

and if $\theta = \epsilon \ll 1$, then $x' \approx x - \epsilon y$ and $y' \approx \epsilon x + y$. In other words

$$\begin{aligned} \delta x &= -\epsilon y \\ \delta y &= +\epsilon x \end{aligned}$$

are the changes in the coordinates. What does this mean for the change in the wave function $\psi(x, y)$ due to the transformation? We calculate

$$\begin{aligned} \delta\psi(x, y) &= \frac{\partial\psi}{\partial x}\delta x + \frac{\partial\psi}{\partial y}\delta y \\ &= \left(-\epsilon y\frac{\partial}{\partial x} + \epsilon x\frac{\partial}{\partial y}\right)\psi \end{aligned}$$

and since we know $\hat{p}_x = -i\partial_x$ and $\hat{p}_y = -i\partial_y$ we can write

$$\begin{aligned} \delta\psi(x, y) &= \epsilon\left(-iy\hat{p}_x + ix\hat{p}_y\right)\psi \\ &= \epsilon i\left(x\hat{p}_y - y\hat{p}_x\right)\psi. \end{aligned}$$

We define the angular momentum operator about the z-axis as

$$\hat{L}_z = \left(x\hat{p}_y - y\hat{p}_x\right). \tag{3.7}$$

Following this definition we construct $\hat{L}_x$ and $\hat{L}_y$ by cyclic iteration of $(x \to y \to z \to x)$, giving

$$\hat{L}_x = \left(y\hat{p}_z - z\hat{p}_y\right) \tag{3.8}$$

$$\hat{L}_y = \left(z\hat{p}_x - x\hat{p}_z\right). \tag{3.9}$$

It is straightforward to show that the commutators between the L's are

$$[\hat{L}_i, \hat{L}_j] = i\epsilon_{ijk}\hat{L}_k$$

for $i, j, k \in \{x, y, z\}$. The totally antisymmetric symbol ϵ_{ijk} is defined as $+1$ for $ijk = \{123, 312, 231\}$, -1 for $ijk = \{132, 213, 321\}$,and zero for any two indices equal to each other. These are the even and odd permutations of the numbers (123).

3.7 The state function, or "ray"

Suppose a general state function is given by the ket $|\psi\rangle$. The expectation value of an observable in this wave function is given by $\langle\hat{O}\rangle_\psi = \langle\psi|\hat{O}|\psi\rangle$ and for a normalized wave function we have $\langle\psi|\psi\rangle = 1$. If we multiply our wave function by an overall phase, a complex number that can be represented by $e^{i\alpha}$, then neither the wave function normalization or the expectation value of an observable will change since $e^{i\alpha}e^{-i\alpha} = 1$. Therefore, multiplication of any state vector $|\psi\rangle$ by an overall phase $e^{i\alpha}$, will give an equivalent description of the physical system. This degree of freedom of multiplying by an arbitrary phase gives rise to Weinberg's use of the term *ray* to describe the collection of state vectors with different values of the arbitrary phase angle α. A state in quantum mechanics is given by a ray in the Hilbert space.

3.8 Complex conjugation

Quantum mechanics necessarily deals with complex numbers, and our wave functions are in general complex, i.e., $\psi(x) = \psi_r(x) + i\psi_i(x)$, where the function can be written in terms of real and imaginary parts. If we complex conjugate, we obtain $\psi^*(x) = \psi_r(x) - i\psi_i(x)$. If we multiply two wave functions $\phi(x)$ and $\psi(x)$ we obtain

$$\begin{aligned}\phi(x)\psi(x) &= (\phi_r + i\phi_i)(\psi_r + i\psi_i) \\ &= (\phi_r\psi_r - \phi_i\psi_i) + i(\phi_r\psi_i + \phi_i\psi_r)\end{aligned}$$

and if we add two wave functions we obtain

$$\phi(x) + \psi(x) = (\phi_r + \psi_r) + i(\phi_i + \psi_i).$$

Note that the magnitude of $\phi(x) + \psi(x)$ is

$$(\phi + \psi)(\phi + \psi)^* = \phi\phi^* + \phi\psi^* + \psi\phi^* + \psi\psi^*.$$

3.9 The spin of the photon

Classical electricity and magnetism tells us that the photon has two polarization states. They can be considered linear, with vertical (y) and horizontal (x) polarization where the electric fields are perpendicular to each other along the direction of propagation. They can also be written in terms of circularly polarized light in which the electric field rotates either clockwise (R) or counter clockwise (L) about the axis of propagation.

Consider linear polarized light that comes out of a polarizer, say along the x'-direction. The electric field is written as $\vec{E} = E_0 e^{i(kz'-\omega t)}\hat{x}'$ for propagation along the z'-axis. If we now introduce another polarizer further down and still perpendicular to the z'-axis, oriented at an angle ϕ with respect to the first polarizer, then the electric field entering the polarizer can be written in the new (rotated) frame as $\vec{E} = E_0 e^{i(kz'-\omega t)}(\cos\phi\,\hat{x} + \sin\phi\,\hat{y})$. We are just writing $\hat{x}' = \cos\phi\,\hat{x} + \sin\phi\,\hat{y}$ as we would for the usual rotation of a vector in the plane.

If this polarizer were an x-polarizer, then the transmitted intensity will go as $\cos^2\phi$, and if this were a y-polarizer, the intensity would go as $\sin^2\phi$. Quantum mechanically we can then represent this experimental result as

$$|\langle x|x'\rangle|^2 = \cos^2\phi$$
$$|\langle x|y'\rangle|^2 = \sin^2\phi$$

giving the amplitudes

$$\langle x|x'\rangle = \cos\phi \qquad \langle y|x'\rangle = -\sin\phi$$
$$\langle x|y'\rangle = \sin\phi \qquad \langle y|y'\rangle = \cos\phi.$$

So we can write

$$|x'\rangle = \cos\phi\,|x\rangle + \sin\phi\,|y\rangle$$
$$|y'\rangle = -\sin\phi\,|x\rangle + \cos\phi\,|y\rangle$$

or

$$\begin{pmatrix} |x\rangle \\ |y\rangle \end{pmatrix} = \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix} \begin{pmatrix} |x'\rangle \\ |y'\rangle \end{pmatrix}.$$

If we define the right and left circularly polarized light as

$$|R\rangle = |x\rangle + i\,|y\rangle$$
$$|L\rangle = |x\rangle - i\,|y\rangle$$

then we can write these in terms of the basis $\{|x'\rangle, |y'\rangle\}$ as

$$|R\rangle = (\cos\phi\,|x'\rangle - \sin\phi\,|y'\rangle) + i(\sin\phi\,|x'\rangle + \cos\phi\,|y'\rangle)$$
$$|L\rangle = (\cos\phi\,|x'\rangle - \sin\phi\,|y'\rangle) - i(\sin\phi\,|x'\rangle + \cos\phi\,|y'\rangle).$$

Collecting terms we have

$$|R\rangle = e^{i\phi}\,|x'\rangle + ie^{i\phi}\,|y'\rangle = e^{i\phi}(|x'\rangle + i\,|y'\rangle)$$
$$|L\rangle = e^{-i\phi}\,|x'\rangle - ie^{-i\phi}\,|y'\rangle = e^{-i\phi}(|x'\rangle - i\,|y'\rangle).$$

or

$$|R\rangle = e^{i\phi}\,|R'\rangle$$
$$|L\rangle = e^{-i\phi}\,|L'\rangle.$$

But these equations say

$$|R'\rangle = e^{-i\phi}\,|R\rangle$$
$$|L'\rangle = e^{+i\phi}\,|L\rangle.$$

If our rotation operator is defined as $R(\phi) = e^{-i\hat{J}_z\phi/\hbar}$, then our result implies that we must have

$$\hat{J}_z\,|R\rangle = \hbar\,|R\rangle$$
$$\hat{J}_z\,|L\rangle = -\hbar\,|L\rangle$$

in order that $e^{-i\hat{J}_z\phi/\hbar}\,|R\rangle = e^{-i\phi}\,|R\rangle$ and similarly for the left circularly polarized light. The fact that we have $1\hbar$ on the rhs, as opposed to $1/2$ or 2 in front of the $\hbar$ implies that the spin of the photon is 1. There is no $S_z = 0$ state because the photon is massless.

4 The Schrödinger Equation

I encourage you to read other books on quantum mechanics in addition to the text we use in class. Other authors may present the same material in a way that you find easier to understand. Sometimes reading about the same physics but simply said a different way is all you need to understand a concept. The notation used in different quantum mechanical textbooks can vary widely. It is important to be aware of differences in notation, and you should get used to seeing different symbols for e.g. various functions, operators, and inner products in quantum mechanics.

Notation for Schrödinger's equation Schrödinger's wave equation is a partial differential equation in both space and time. In many of the cases we will investigate, the wave function can be separated (separation of variables) into two parts, once space, and one time. Some textbooks will use $\psi(x,t)$ for the whole wave function while others will use $\Psi(x,t)$ and reserve $\psi(x)$, or $\phi(x)$ for only the spatial part of the function. If we "separate" our wave function into two parts, where it is appropriate, we often write $\Psi(x,t) = \psi(x)\phi(t)$. Some use $\phi(t)$ for the time-dependent part. However, it is common in other textbooks to use $\phi(x)$ for the space dependent part and I will use this notation in several discussions below. Be careful when reading the class notes and other textbooks and be sure that you know which function is which for any particular discussion.

The non-relativistic Schrödinger Equation (SE) in its most general form is given by

$$i\hbar\partial_t\psi(x,t) = \hat{H}\psi(x,t) \tag{4.1}$$

where most often we will specialize to a Hamiltonian given by $\hat{H} = \frac{\hat{p}^2}{2m} + V(x)$. Here $\hat{p}^2 = \hat{p}_x^2 + \hat{p}_y^2 + \hat{p}_x^2$, and we leave the "potential function" $V(x)$ unspecified. The momentum operator can be written as $\hat{p}_j = -i\hbar\partial_j$, for $j = \{x, y, z\}$. We will derive this form of the momentum operator in class, but we note that $\frac{\hat{p_x}^2}{2m} = -\frac{\hbar^2}{2m}\partial_x^2$. Using these assumptions, the SE in one space dimension takes the form

$$i\hbar\frac{\partial}{\partial t}\psi(x,t) = \left[-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V(x)\right]\psi(x,t).$$

The time-*independent* Schrödinger equation is found by assuming that the Hamiltonian is independent of the time, i.e., $\hat{H} \neq \hat{H}(t)$. In this case it is permissible to separate the wave function and we write $\psi(x,t) = \chi(x)\phi(t)$. Note that here we are using Griffiths notation with $\phi(t)$ for the time dependence. Then the time derivative passes through the spatial function on the LHS, and vice versa on the RHS, and the SE goes to $i\hbar\chi\partial_t\phi = -\frac{\hbar^2}{2m}\chi''\phi + \phi V\chi$. Dividing through by ψ gives

$$i\hbar\frac{\phi'}{\phi} = -\frac{\hbar^2}{2m}\frac{\chi''}{\chi} + V(x). \tag{4.2}$$

The derivatives represent differentiation with respect to the variable of the function, either x or t. In this separated form, the left and right sides are linked by a common value even

as x and t vary, so that they must be equal to a constant. Call this "separation" constant E, and we can solve the left hand side immediately. Since $\frac{d\phi}{dt} = -\frac{iE}{\hbar}\phi(t)$, we have that $\phi(t) = \phi(0)e^{-iE/\hbar}$. We cannot solve the spatial differential equation unless we know $V(x)$. Note that we can write the RHS of (4.2) as

$$\hat{H}\chi(x) = E\chi(x), \tag{4.3}$$

known as the time-independent SE. We will spend a lot of time solving this equation with particular "potential" functions $V(x)$ in class.

Note that the SE looks very much like the diffusion equation $\partial_t c(x,t) = D\partial_x^2 c(x,t)$, where the diffusion constant D has units of $\mathrm{cm}^2\mathrm{s}^{-1}$ and $c(x,t)$ can be considered the concentration of solute for example. Also recall the wave equation $\partial_t^2 q(x,t) = v^2\partial_x^2 q(x,t)$, with propagation velocity v admits wave solutions such as $q(x,t) = q_0 e^{i(kx-\omega t)}$. The diffusion equation and the wave equation only differ in the power of the time derivative, but the behavior of the solutions to these two equations is dramatically different. The SE looks like the diffusion equation because there is only one power of the time derivative, but the complex number i in front of $\frac{\partial}{\partial t}$ in the SE changes the solutions from those of the diffusion equation. If we assume we have a free particle so that $\hat{H} = -\frac{\hbar^2}{2m}\partial_x^2$, then plugging in a trial solution for the SE using $\psi(x,t) = Ae^{i(kx-\omega t)}$ and taking the partial derivatives gives $\hbar\omega = \frac{\hbar^2 k^2}{2m}$. If we identify $E = \hbar\omega$ and $p = \hbar k$, then obtain $E = \frac{p^2}{2m}$, exactly as we expect for a free particle! The SE admits wave solutions. We will see that its resemblance to the diffusion equation is also very important.

4.1 The quantum mechanical commutator

A study of the Schrodinger equation shows that one can describe the position and momentum as operators $\hat{x}$ and $\hat{p}$ that satisfy the canonical commutation relation

$$[\hat{x}, \hat{p}] = i\hbar. \tag{4.4}$$

This is in stark contrast to classical mechanics where the phase space variables x and p commute. We would write classically that

$$[x, p] = 0. \tag{4.5}$$

Going from (4.5) to (4.4) can be described as *deforming* the relation (4.5); we have made the rhs nonzero. In fact we can consider taking ordinary phase space in which (4.5) holds and *impose* the condition (4.4). What happens when we do this? Clearly if the variables x and p no longer commute then they have a matrix representation. What do matrices operate on? A vector space. Therefore, imposing the condition that the phase space variables do not commute forces us into a vector space where our phase space variables become operators: $x \to \hat{x}$ and $y \to \hat{y}$. This essentially describes the process of deformation that takes us from classical mechanics to quantum mechanical Hilbert space.

4.2 Eigenfunction Expansions

In Dirac notation Schrödinger's equation is simply $i\hbar\partial_t|\psi(x,t)\rangle = \hat{H}|\psi(x,t)\rangle$, and we will use this notation whenever it is convenient. We will cover the specifics of Dirac notation in detail during the semester. In quantum mechanics, the solutions to the SE live in a mathematical linear vector space called "Hilbert space". Just as in Euclidean geometry in 3-space, where we expand an arbitrary vector over the basis of unit vectors pointing along the axes, $\vec{v} = v_x\hat{x}+v_y\hat{y}+v_z\hat{z}$, we can do the same in Hilbert space with functions, which are more general objects than are 3-space vectors. The wave function $|\psi(x,t)\rangle$ can be "expanded" in a "basis set" consisting of the functions $\{|\phi_n(x)\rangle\}$, so we write

$$|\psi(x,t)\rangle = \sum_n c_n(t)|\phi_n(x)\rangle. \tag{4.6}$$

Note that we are now using $\phi(x)$ as the spatial component. Equation (4.6) should remind you very much of a Fourier expansion. The basis states are often chosen so that they are orthogonal to each other and normalized so the inner product of a function with itself is unity. We write this condition mathematically in Dirac notation as

$$\langle\phi_n(x)|\phi_m(x)\rangle = \int dx\ \phi_n(x)\phi_m(x) = \delta_{nm}, \tag{4.7}$$

where δ_{nm} is the Kronecker delta.

Equation (4.6) is very general, but we often start with this equation and then go to a special case for some particular problem. Suppose we have a set of basis states that are energy eigenstates, so that $\hat{H}\phi_n(x) = E_n\phi_n(x)$. Then if we operate on a general wave function $\psi(x,t)$ with our time-independent Hamiltonian, we obtain $\hat{H}\sum_n c_n\phi_n = \sum_n c_nE_n\phi_n$. Therefore we can write the SE as $i\hbar\sum_n c_n'\phi_n = \sum_n c_nE_n\phi_n$, and because these sums must be equal term-for-term, we can easily pick off the time dependence of the c_n and obtain

$$\psi(x,t) = \sum_n c_n(0)e^{-iE_nt/\hbar}\phi_n(x). \tag{4.8}$$

This equation tells us that if we know the wave function at time $t = 0$, then we can find it for all time. In other words, if we know $\psi(x,0)$ then we can find $\psi(x,t)$. We will solve (4.8) in class in several situations.

4.3 Interpretation of the Wave Function

For a single particle, say an electron, we interpret the magnitude squared of the wave function to be the probability of finding the electron. Therefore, given a properly normalized wave function, the integral over all space of this probability must be one, $\int\psi^*(x)\psi(x)dx = 1$. In other words, the electron must be *somewhere*. From this interpretation we will calculate such things as fluxes and use them to understand scattering of particles off potentials, for example. We say that $\psi^*(x)\psi(x)$ is a *probability density*, and that $\psi(x)$ is the *amplitude* for a process. In this case it is the amplitude to find the particle.

An "observable" in quantum mechanics has specific mathematical restrictions that we will cover in class. In short, we calculate an observable in the following manner:

$$\langle O \rangle = \frac{\int dx\, \psi^*(x)\, \hat{O}\, \psi(x)}{\int dx\, \psi^*(x)\, \psi(x)}.$$

We say that we have calculated the "expectation value" of the observable O, and you can think of something like the particle position, or momentum for example.

4.4 Generalized Uncertainty Principle

The uncertainty principle can be derived from the variance of expectation values as follows. This presentation is taken from Gasiorowicz's text. If we define two Hermitian operators U, V, and form an arbitrary wave function $\phi(x) = (U + i\lambda V)\psi(x)$, we can use the fact that $\int dx\, \phi * \phi$ is positive definite to obtain the uncertainty principle in terms of the commutator of U and V. We have a function of the parameter λ.

$$\begin{aligned} I(\lambda) &= \int dx\, [(U + i\lambda V)\psi(x)]^* [(U + i\lambda V)\psi(x)] \\ &= \int dx\, [\psi(x)]^* \left[(U^+ - i\lambda V^+)(U + i\lambda V)\psi(x)\right] \\ &= \int dx\, [\psi(x)]^* [(U - i\lambda V)(U + i\lambda V)\psi(x)] \end{aligned}$$

So that $I(\lambda) = \langle U^2 \rangle + i\lambda\langle [U, V] \rangle + \lambda^2 \langle V^2 \rangle$. Minimizing this with respect to $\lambda$and noting that $I_{min} \geq 0$, one finds

$$\langle U^2 \rangle \langle V^2 \rangle \geq \frac{1}{4} \langle i[U, V] \rangle.$$

If we let $U = A - \langle A \rangle$, and similarly for V, then we see that $(\Delta A)^2 (\Delta B)^2 \geq \frac{1}{4}\langle i[A, B] \rangle$.

4.5 Time evolution operator

The time evolution operator is given by $U(t; \delta t) = (1 - \frac{i}{\hbar} H(t)\delta t)$ for an infinitesimal translation in time. If $H \neq H(t)$, then breaking up a finite time interval into a set of small intervals, $t = n\delta t$, it follows that $U(t, 0) = e^{-iHt/\hbar}$. If the Hamiltonian is time dependent then this expression needs to be modified. Let's set $\hbar = 1$ for the moment. Then

$$\begin{aligned} U(t; \delta t) &= (1 - iH(t)\delta t) \\ U(t + \delta t; \delta t) &= (1 - iH(t + \delta t)\delta t) \\ &\vdots \\ U(t + n\delta t; \delta t) &= (1 - iH(t + n\delta t)\delta t)\,. \end{aligned}$$

We evolve the system a finite time by incrementally going from $t' = 0$ to $t' = t$. Thus

$$\begin{aligned} U(t;0) &= (1 - iH(n\delta t)\delta t) \cdots (1 - iH(\delta t)\delta t)\,(1 - iH(0)\delta t) \\ &\approx e^{-iH(n\delta t)\delta t} \cdots e^{-iH(\delta t)\delta t} e^{-iH(0)\delta t} \\ &= \exp\left[-i\left[H(0) + H(\delta t) + H(2\delta t) + \cdots + H(n\delta t)\right]\delta t + O(\delta t^2)\right] \end{aligned}$$

Letting $t = n\delta t$ or $\delta t = t/n$ we have

$$\begin{aligned} U(t;0) &= \exp\left[-i\left[H(0) + H(t/n) + H(2t/n) + \cdots + H((n-1)t/n) + H(t)\right]\frac{t}{n}\right] \\ &= \exp\left[-i\sum_{j=0}^{n}\frac{t\Delta j}{n}H(\frac{j}{n}t)\right]. \end{aligned}$$

In the limit as $n \to \infty$ and with $\Delta j = 1$, we define the ratios $\frac{t\Delta j}{n} \equiv dt'$ and $\frac{tj}{n} = t'$ then we have

$$\sum_{j=0}^{n}\frac{t\Delta j}{n}H(\frac{j}{n}t) \to \int_0^t dt'\, H(t')$$

and so

$$U(t;0) = e^{-i/\hbar\int_0^t dt'\, H(t')}.$$

5 Classical vs. Quantum

5.1 Classical Mechanics is not linear

We have seen that in QM a linear combination of solutions is a solution to the Schrödinger equation. For contrast, let us consider classical mechanics in one dimension and show that it is not linear, even though the addition of two solutions to a linear ODE provide a solution to that ODE.

In one dimension let us write Newton's equation as $F = m\,d^2z/dt^2$ and consider a ball thrown up from the ground at $z = 0$ with velocity $\dot{z} = 1$. We have from Newton that

$$-mg = m\frac{d^2z}{dt^2}$$

which we can integrate and find

$$z(t) = -\frac{g}{2}t^2 + c_1 t + c_2$$

and use the boundary conditions (z and $\dot{z}$ at time $t = 0$) to determine that $c_1 = 1$,and $c_2 = 0$. So our solution is

$$z(t) = -\frac{g}{2}t^2 + t.$$

This is a parabola and gives $F = md^2z/dt^2 = -mg$ as we expect. The solution is unique given the boundary conditions. This means that any other boundary conditions would give a different path.

Suppose we wanted to try to add two solutions together. They are given by

$$z_1(t) = -\frac{g}{2}t^2 + c_{11}t + c_{12}$$
$$z_2(t) = -\frac{g}{2}t^2 + c_{21}t + c_{22}.$$

Then $z(t) = z_1(t) + z_2(t)$ gives

$$F = -mg - mg = -2mg. \quad \perp$$

This gives a contradiction, it is not correct. How about if we try to improve the result by mixing the two solutions so that they have a combined weight of 1? For example we could write

$$z(t) = \eta z_1(t) + (1 - \eta)z_2(t)$$

where $\eta \in [0, 1]$. Then when we calculate the force we find

$$F = -mg\eta - mg(1 - \eta) = -mg.$$

This looks great but it is still wrong. Why? In general if we choose a solution like this with arbitrary c_{ij} it will not satisfy the boundary conditions.† To see this let's look at the boundary conditions of the full additive "solution" $z(t) = z_1(t) + z_2(t)$ with arbitrary c_{ij}. Using the bounday conditions $z(0) = 0$ and $\dot{z}(0) = 0$ gives the equations

$$\begin{aligned} \eta c_{12} + (1-\eta)c_{22} &= 0 \\ \eta c_{11} + (1-\eta)c_{21} &= 1. \end{aligned}$$

These can be written in matrix form as

$$\begin{pmatrix} c_{11} & c_{21} \\ c_{12} & c_{22} \end{pmatrix} \begin{pmatrix} \eta \\ 1-\eta \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

with solution

$$\begin{pmatrix} \eta \\ 1-\eta \end{pmatrix} = \frac{1}{\det C} \begin{pmatrix} c_{22} & -c_{21} \\ -c_{12} & c_{11} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

where $\det C = c_{11}c_{22} - c_{12}c_{21}$. This gives

$$\begin{pmatrix} \eta \\ 1-\eta \end{pmatrix} = \frac{1}{\det C} \begin{pmatrix} c_{22} \\ -c_{12} \end{pmatrix}$$

or

$$\begin{aligned} \eta &= \frac{c_{22}}{c_{11}c_{22} - c_{12}c_{21}} \\ 1-\eta &= \frac{-c_{12}}{c_{11}c_{22} - c_{12}c_{21}} \end{aligned}$$

We can now see the result given that $\eta + 1 - \eta = 1$. We have

$$c_{11}c_{22} - c_{12}c_{21} = c_{22} - c_{12}.$$

But this equation allows for many solutions. To take one we can set $c_{11} = c_{21} = 1$, then we have that

$$c_{22} - c_{12} = \text{arbitrary}$$

and can be anything. Suppose we try $c_{22} - c_{12} = 1$, say, and let $c_{12} = 1$, then $c_{22} = 2$. This will satisfy $F = -mg$, but let's look at the trajectory itself. We'll have

$$z(t) = \eta\left(-\frac{g}{2m}t^2 + t + 1\right) + (1-\eta)\left(-\frac{g}{2m}t^2 + t + 2\right)$$

and we find

$$\begin{aligned} z(0) &= \eta + 2(1-\eta) = 2 - \eta \\ &\neq 0 \end{aligned}$$

†Well, obviously we could choose the solution to be

$$z(t) = \frac{1}{2}z_1 + \frac{1}{2}z_2$$

with both $z_1 = z_2 = -\frac{g}{2m}t^2 + t$, but this is just our original solution above. It is one very special case.

unless $\eta = 2$. But we stipulated that $\eta \in [0, 1]$. Choosing $\eta = 2$ will spoil the second derivative giving $F = -mg$.

So in general, adding solutions will not give a new solution. This illustrates that classical mechanical solutions are pinned down by the boundary conditions and there is a unique solution to a classical linear ODE with those boundary conditions. In quantum mechanics there are still boundary conditions to impose, but the linear superposition of state vectors also satisfies the Schrödinger equation.

6 Perturbation Theory

Perturbation theory falls into two main classes, those that are time-independent and those that are time-dependent. Time-independent perturbation theory is covered thoroughly and nearly identically in every standard textbook, so we skip it here and give an example instead. Below we will develop time-dependent perturbation theory.

6.1 Perturbing a quantum system

A perturbation to a quantum system can be represented as a component of the Hamiltonian. So for example given a quantum system with Hamiltonian H_0 the perturbed system has Hamiltonian $H_0 + V$.

Let us begin with a simple quantum mechanical system with known energy eigenstates, $\{|n\rangle\}$. A general solution is given by $|\Psi\rangle = \sum_n c_n(t)\,|n\rangle$ with $H_0\,|n\rangle = E_n\,|n\rangle$. The Schrödinger equation of the unperturbed system gives for the coefficients

$$i\partial_t c_n(t) = E_n c_n(t)$$

or $c_n(t) = c_n(0)e^{-iE_n t}$. Let us suppose that the perturbation is given by

$$V = \lambda\delta(t - t_0)\,|n\rangle\,\langle m|\,. \tag{6.1}$$

This is a perturbation that mixes part of the state $|m\rangle$ with $|n\rangle$ at the instant of time $t = t_0$. Solving the Schrödinger equation now for the coefficients gives a different result for $c_n'(t)$,

$$ic_n'(t) = E_n c_n(t) + \lambda c_m(t)\delta(t - t_0). \tag{6.2}$$

We expect the solution of (6.2) to go something like $e^{-iE_n t}$ but be modified suitably by the perturbation. We can solve this equation using Fourier transforms. Multiplying both sides by $e^{i\omega t}$ and integrating over t we have

$$i\int dt\, e^{i\omega t}c_n'(t) = E_n c_n(\omega) + \lambda c_m(t_0)e^{i\omega t_0}$$

where

$$c_n(\omega) \equiv \int dt\, c_n(t)e^{i\omega t}.$$

Integrating by parts on the l.h.s. gives

$$+\omega c_n(\omega) = E_n c_n(\omega) + \lambda c_m(t_0)e^{i\omega t_0}$$

or

$$c_n(\omega) = \lambda c_m(t_0)\frac{e^{i\omega t_0}}{\omega - E_n}. \tag{6.3}$$

In inverting the FT

$$c_n(t) = \int d\omega\, e^{-i\omega t}c_n(\omega)$$

we need to perform the integral

$$\int_{-\infty}^{\infty} d\omega\, e^{-i\omega t} \frac{e^{i\omega t_0}}{\omega - E_n}.$$

This integral is easy to do by shifting the integration variable. Let $y = \omega - E_n$ then $dy = d\omega$, the limits stay the same, and we have

$$\begin{aligned} c_n(t) &= \int dy\, \frac{e^{i(y+E_n)(t_0-t)}}{y} \\ &= e^{-iE_n(t-t_0)} \int dy\, \frac{e^{-iy(t-t_0)}}{y}. \end{aligned}$$

Let $I = \int dy\, e^{-iy\alpha}/y$ where $\alpha = t - t_0$. Then $dI/d\alpha = -i \int dy\, e^{iy\alpha} = -i\delta(\alpha)$. The integral of $dI/d\alpha$ gives $I(\alpha) = s - i\theta(\alpha)$ with some integration constant s, where $\theta(\alpha)$ is the Heaviside step function. Therefore we have

$$c_n(t) = e^{-iE_n(t-t_0)} \lambda c_m(t_0) \left[s - i\theta(t - t_0) \right].$$

For $t < t_0$ the solution must be $c_n(t < t_0) = c_n(0)e^{-iE_n t}$ and here the step function vanishes, so the constant s can be solved for,

$$c_n(0)e^{-iE_n t} = e^{-iE_n(t-t_0)} \lambda c_m(t_0) s$$

giving

$$s = c_n(0) \frac{e^{-iE_n t_0}}{\lambda c_m(t_0)}$$

and yielding the final solution for the coefficient $c_n(t)$

$$c_n(t) = c_n(0)e^{-iE_n t} - i\lambda c_m(t_0) e^{-iE_n(t-t_0)} \theta(t - t_0) \tag{6.4}$$

just as we expected.

So we see that the perturbation changes the amplitude that a measurement will find the system in state $|n\rangle$. This perturbation does not conserve the energy because it changes no other amplitudes, so if the amplitudes without the perturbation are normalized, $\sum_n |c_n|^2 = 1$, then they will not be following the perturbation. The amplitude to find the system in state $|n\rangle$ is $P_n = |c_n|^2$ with

$$c_n^* c_n = |c_n(0)|^2 + \lambda c_n(0) c_m(t_0) 2 \sin(E_n t_0) \theta(t - t_0) + O(\lambda^2). \tag{6.5}$$

6.2 Time-dependent perturbation theory

We now develop the general theory of time-dependent perturbation theory.[†] We begin with the Schrodinger equation

[†] In this section we set $\hbar = 1$.

$$i\partial_t |\psi\rangle = H |\psi\rangle$$

where $H = H_0 + H'$. Write $|\psi\rangle = \sum_n c_n(t) |\phi_n\rangle$ with $H_0 |\phi_n\rangle = E_n |\phi_n\rangle$. Without the perturbation we have

$$i \sum_n \dot{c}_n(t) |\phi_n\rangle = H_0 \sum_n c_n(t) |\phi_n\rangle$$

and taking the inner product with $\langle\phi_m|$ gives

$$i\dot{c_m}(t) = E_m c_m(t)$$

or $c_m(t) = c_m(0)e^{-iE_m t}$. Then we can write

$$|\psi\rangle = \sum_n c_n(0) e^{-iE_n t} |\phi_n\rangle$$

where there is no time dependence in $c_m(0)$ or $|\phi_n\rangle$. Define a new set of kets $|n\rangle$ with the simple exponential time dependence included,

$$|n\rangle = e^{-iE_n t} |\phi_n\rangle$$

and note that

$$\partial_t |n\rangle = -iE_n |n\rangle .$$

Then we can write the wave function generally as

$$|\psi\rangle = \sum_n b_n(t) |n\rangle$$

where any additional time dependence in the b's will be due to the time-dependent perturbation in the Hamiltonian. Writing out the Schrödinger equation we have

$$i\partial_t |\psi\rangle = i \sum_n \left(\dot{b}_n - iE_n b_n \right) |n\rangle = (H_0 + \lambda H') \sum_n b_n |n\rangle .$$

Dotting through with $\langle m|$ and noting that $\langle m|n\rangle = e^{-i(E_m - E_n)t}\delta_{mn}$ gives

$$\dot{b}_m - iE_m b_m = -iE_m b_m - i\lambda \sum_n b_n H'_{mn}$$

or

$$\dot{b}_m = -i\lambda \sum_n b_n H'_{mn}(t).$$

Integrating this expression we have

$$b_m(t) = -i\lambda \sum_n \int_{t_0}^{t} dt' \, b_n(t) H'_{mn}(t).$$

Recalling that $H'_{mn}(t) = \langle m|H'(t)|n\rangle$ we write

$$H'_{mn}(t) = e^{-i(E_n - E_m)t} \langle\phi_m|H'(t)|\phi_n\rangle$$

giving

$$b_m(t) = -i\lambda \sum_n \int_{t_0}^{t} dt'\, b_n(t) e^{-i(E_n - E_m)t} \langle \phi_m | H'(t) | \phi_n \rangle \,. \tag{6.6}$$

We can now write the perturbation expansion for the coefficients as

$$b_m(t) = b_m^0(t) + \lambda b_m^1(t) + \lambda^2 b_m^2(t) + \cdots$$

and inserting this into (6.6) gives the b_m^j at different orders

$$\begin{aligned} b_m^1(t) &= -i \sum_n \int_{t_0}^{t} dt'\, b_n^0(t') e^{-i(E_n - E_m)t'} \langle \phi_m | H'(t') | \phi_n \rangle \quad &: \lambda \\ &\vdots \\ b_m^{p+1}(t) &= -i \sum_n \int_{t_0}^{t} dt'\, b_n^p(t') e^{-i(E_n - E_m)t'} \langle \phi_m | H'(t') | \phi_n \rangle \quad &: \lambda^{p+1}. \end{aligned}$$

Once the $b_n^j(t)$ are computed to the desired order, the approximate wave function is determined.

7 Two-particle Systems and Quantum Chemistry

This section contains an elementary introduction to some of the concepts used in quantum chemistry (QC) and density functional theory (DFT). DFT methods for periodic solids rely on Fourier analysis, but the underlying concepts in QC and DFT are the same: constructing appropriately normalized initial wave functions and using the variational principle and self consistent field (SCF) methods to minimize the total energy. Perturbation and other methods are used to obtain information about excited states. Section 7.3 illustrates how to incorporate electron spin into the wave function. Section 7.7 describes the Coulomb and exchange integrals for a two-electron wave function, and defines the correlation energy.

7.1 Basis states for a two-spin system

We will work with spin-1/2 in this section, but the results are easily generalized. For a system with two spin-1/2 particles, say a proton and an electron, we can write the basis states of each individual particle in the z-basis as $\{|z+\rangle_1, |z-\rangle_1\}$ and $\{|z+\rangle_2, |z-\rangle_2\}$, where the subscripts refer to a specific particle. We will denote the basis states of the two-particle system as

$$\{|z+,z+\rangle, |z+,z-\rangle, |z-,z+\rangle, |z-,z-\rangle\}$$

with a column vector representing each state as

$$|z+,z+\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \qquad |z+,z-\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \tag{7.1}$$

$$|z-,z+\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \qquad |z-,z-\rangle = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}. \tag{7.2}$$

► *aside*: There are several different ways to represent the same basis vector $|z+,z+\rangle$, in particular it is common to see the following,

$$|z+,z+\rangle \equiv |z+\rangle_1 \otimes |z+\rangle_2 \equiv |z+\rangle_1 |z+\rangle_2,$$

where the symbol $\otimes$ is called the **tensor product**. The tensor product, also known as the direct product, is a notation to keep track of the different Hilbert spaces for each particle. In this case with two spin-1/2 particles the Hilbert spaces are identical, but one could also consider a spin-1/2 particle and a spin-3/2 particle as a two-particle system. Then the Hilbert spaces would not even have the same dimension and so on. Generally, the tensor product is used to form products of vector spaces. Hilbert spaces are vector spaces, and the product must be used whenever there is more than one particle in a 'system'. The notation $|z+\rangle_1 \otimes |z+\rangle_2$ emphasizes that $|z+\rangle_1$ is in a *different* vector space than $|z+\rangle_2$. One can also

write the basis kets in terms of the spin and spin projection $|s_1, m_1; s_2, m_2\rangle$, which for the state $|z+, z-\rangle$ is $|\frac{1}{2}, \frac{1}{2}; \frac{1}{2}, -\frac{1}{2}\rangle$.◀

Staying within the z-basis we can define kets such as $|x-, z+\rangle$, for example, using $|x-\rangle = (|z+\rangle - |z-\rangle)/\sqrt{2}$, giving

$$|x-, z+\rangle = \frac{1}{\sqrt{2}} \left(|z+, z+\rangle - |z-, z+\rangle \right).$$

The spin operator for each spin is

$$\hat{S}_1 = (\hat{S}_{1x}, \hat{S}_{1y}, \hat{S}_{1z})$$
$$\hat{S}_2 = (\hat{S}_{2x}, \hat{S}_{2y}, \hat{S}_{2z})$$

where they act on the basis kets as shown in the following examples,

$$\hat{S}_{2z} |x-, z+\rangle = +\frac{\hbar}{2} |x-, z+\rangle$$
$$\hat{S}_{2z} |x+, z-\rangle = -\frac{\hbar}{2} |x+, z-\rangle .$$

Most importantly, the spin operators $\hat{S}_{1,2}$ operate *independently.* This means in particular that their commutator vanishes,

$$[\hat{S}_1, \hat{S}_2] = 0.$$

We now apply the above formalism to the hydrogen atom. Consider a hydrogen atom and ignore spin altogether. Then the potential (interaction between the electron and proton) is $V(r) = -e^2/r$ and the ground state energy of the system is $E_0 = -13.6$ eV. This is not quite correct however, because the spin of the electron and the spin of the proton will slightly change the energy. There are two important corrections to the ground state energy E_0. The largest correction is called the **fine structure**, and is obtained by considering the **spin-orbit interaction** of the electron spin and the orbital motion that gives rise to a magnetic field generated by the proton as it goes around the electron (from the electron's point of view). The second correction is smaller, but still important, and is called the **hyperfine splitting**. The hyperfine splitting is due to the direct interaction of each particle's intrinsic spin (non-orbital). The Hamiltonian for the hyperfine splitting is given by

$$\hat{H}_1 = \frac{2A}{\hbar^2} \hat{S}_1 \cdot \hat{S}_2. \tag{7.3}$$
$$= \frac{2A}{\hbar^2} \left(\hat{S}_{1x}\hat{S}_{2x} + \hat{S}_{1y}\hat{S}_{2y} + \hat{S}_{1z}\hat{S}_{2z} \right) \tag{7.4}$$

We want to write this Hamiltonian in a more useful form, and we do this by making use of the raising and lowering operators defined by

$$\hat{S}_{1\pm} = \hat{S}_{1x} \pm i\hat{S}_{1y}$$
$$\hat{S}_{2\pm} = \hat{S}_{2x} \pm i\hat{S}_{2y}.$$

Now we calculate the following products

$$\begin{aligned}\hat{S}_{1+}\hat{S}_{2-} &= \left(\hat{S}_{1x} + i\hat{S}_{1y}\right)\left(\hat{S}_{2x} - i\hat{S}_{2y}\right) \\ &= \hat{S}_{1x}\hat{S}_{2x} + i\left(\hat{S}_{1y}\hat{S}_{2x} - \hat{S}_{1x}\hat{S}_{2y}\right) + \hat{S}_{1y}\hat{S}_{2y}\end{aligned}$$

and

$$\begin{aligned}\hat{S}_{1-}\hat{S}_{2+} &= \left(\hat{S}_{1x} - i\hat{S}_{1y}\right)\left(\hat{S}_{2x} + i\hat{S}_{2y}\right) \\ &= \hat{S}_{1x}\hat{S}_{2x} - i\left(\hat{S}_{1y}\hat{S}_{2x} - \hat{S}_{1x}\hat{S}_{2y}\right) + \hat{S}_{1y}\hat{S}_{2y},\end{aligned}$$

and add them together to obtain $\hat{S}_{1+}\hat{S}_{2-} + \hat{S}_{1-}\hat{S}_{2+} = 2\hat{S}_{1x}\hat{S}_{2x} + 2\hat{S}_{1y}\hat{S}_{2y}$. Then we can write

$$2\hat{S}_1 \cdot \hat{S}_2 = \hat{S}_{1+}\hat{S}_{2-} + \hat{S}_{1-}\hat{S}_{2+} + 2\hat{S}_{1z}\hat{S}_{2z}. \tag{7.5}$$

Using (7.5) we can calculate the matrix elements of $\hat{H}_1$. Calculating one finds $\langle 1|\hat{H}_1|1\rangle = A/2$, for example, and the matrix representation one finds is

$$\hat{H}_1 = \begin{pmatrix} A/2 & 0 & 0 & 0 \\ 0 & -A/2 & A & 0 \\ 0 & A & -A/2 & 0 \\ 0 & 0 & 0 & A/2 \end{pmatrix}. \tag{7.6}$$

Calculating the eigenvalues, and letting $\alpha = A/2$, requires

$$\begin{vmatrix} \alpha - \lambda & 0 & 0 & 0 \\ 0 & -\alpha - \lambda & 2\alpha & 0 \\ 0 & 2\alpha & -\alpha - \lambda & 0 \\ 0 & 0 & 0 & \alpha - \lambda \end{vmatrix} = 0$$

giving

$$(\alpha - \lambda)^2\left[(\alpha + \lambda)^2 - A^2\right] = 0.$$

The eigenvalues are $\lambda = \alpha$ twice from the first factor, and $\lambda = \{\alpha, -3\alpha\}$ from the second factor, or $\lambda = \{\alpha, \alpha, \alpha, -3\alpha\}$, indicating there is degeneracy (equal eigenvalues).

Now that we have the eigenvalues we need to find the eigenvectors, and we proceed using $\lambda = \alpha$. Somehow, this should give us 3 separate eigenvectors because this eigenvalue is triply degenerate. We proceed to calculate

$$\begin{pmatrix} \alpha & 0 & 0 & 0 \\ 0 & -\alpha & 2\alpha & 0 \\ 0 & 2\alpha & -\alpha & 0 \\ 0 & 0 & 0 & \alpha \end{pmatrix}\begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = \alpha\begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix}$$

giving the equations

$$\alpha a = \alpha a \qquad -\alpha b + 2\alpha c = \alpha b \qquad 2\alpha b - \alpha c = \alpha c \qquad \alpha d = \alpha d.$$

The first and last equations provide no information, leaving us with 2 equations in 4 unknowns, so there is no unique answer. The two middle equations reduce to $c = b$, leaving us with the following column vector

$$\begin{pmatrix} a \\ b \\ b \\ d \end{pmatrix}$$

as the eigenvector for $\lambda = \alpha$. The first and last entries can be anything we want, but the middle two must be the same. Since we need three independent eigenvectors we can make the choices

$$\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \qquad \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \qquad \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \tag{7.7}$$

where we chose the prefactor in the last one so that the three are orthonormal. For the remaining eigenvector we use the eigenvalue -3α, which gives the following equations

$$\alpha a = -3\alpha a \qquad -\alpha b + 2\alpha c = -3\alpha b \qquad 2\alpha b - 3\alpha c = \alpha c \qquad \alpha d = -3\alpha d.$$

The first and last imply that $a = d = 0$, and the middle two give $c = -b$, suggesting we choose the eigenvector

$$\frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix}, \tag{7.8}$$

which we note is orthogonal to the third eigenvector from the eigenvalue α. The eigenvectors (7.7) and (7.8) together form a complete set. They are in the z-basis, so for example, using (7.1) and (7.2)

$$\frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} = \frac{1}{\sqrt{2}} \left(|z+, z-\rangle + |z-, z+\rangle \right).$$

The three vectors (7.7) are called the **triplet states** because there are three of them, and the state (7.8) is called the **singlet state**. The energy between the triplet and singlet states (singlet state is lowest) is $4\alpha = 2A$. The energy required to transition from the singlet state to one of the triplet states is

$$2A = \hbar\omega = h\nu.$$

In hydrogen the experimental value for the frequency is $\nu_{\text{expt}} = 1420$ MHz, with a corresponding wavelength of $\lambda = 21$ cm, which is in the microwave region. This value has been measured to one part in 10^{13}, and is one of the most accurately measured values in experimental physics.

7.2 Addition of angular momentum

In a two-particle Hilbert space an arbitrary ket may be written as

$$|\Psi\rangle = |\psi_1\rangle \otimes |\psi_2\rangle \,.$$

Suppose an operator $\hat{A}$ acts only on the space in which $|\psi_1\rangle$ lives, then we can write the operator acting on the ket $|\Psi\rangle$ as

$$\left(\hat{A} \otimes 1\right) (|\psi_1\rangle \otimes |\psi_2\rangle) = \hat{A}\,|\psi_1\rangle \otimes |\psi_2\rangle \,,$$

and likewise for an operator $\hat{B}$ acting on the states in the space of kets for the second particle,

$$\left(1 \otimes \hat{B}\right) (|\psi_1\rangle \otimes |\psi_2\rangle) = |\psi_1\rangle \otimes \hat{B}\,|\psi_2\rangle \,.$$

Using these ideas we define the *total spin operator* for a two-particle system (using spin-1/2 specifically)

$$\hat{S} = \left(\hat{S}_1 \otimes 1\right) + \left(1 \otimes \hat{S}_2\right),$$

which is usually shortened to $\hat{S} = \hat{S}_1 + \hat{S}_2$. Calculating $\hat{S}^2$ we have

$$\begin{aligned} \hat{S}^2 &= \left(\hat{S}_1 + \hat{S}_2\right)^2 \\ &= \hat{S}_1^2 + 2\hat{S}_1 \cdot \hat{S}_2 + \hat{S}_2^2 \end{aligned}$$

and because $\hat{S}_1$ and $\hat{S}_2$ operate on different Hilbert spaces *all components of* $\hat{S}_1$ *, e.g.* $\hat{S}_{1y}$*, commute with the components of* $\hat{S}_2$, and vice versa.

$\hat{S}^2$ represents the total angular momentum squared of the two-particle system and its action on the kets $|s, m\rangle$ is

$$\begin{aligned} \hat{S}^2\,|s, m\rangle &= s(s+1)\hbar^2\,|s, m\rangle \\ \hat{S}_z &= m\hbar\,|s, m\rangle \end{aligned}$$

where s is the total spin, and m is the z-projection of the total spin. For example,

$$\begin{aligned} \hat{S}_1^2\,|z+, z+\rangle &= \hat{S}_1^2\,|z+\rangle \otimes |z+\rangle \\ &= \frac{1}{2}\left(\frac{1}{2}+1\right)\hbar^2\,|z+\rangle \otimes |z+\rangle \end{aligned}$$

and

$$\begin{aligned} \hat{S}_2^2\,|z+, z+\rangle &= |z+\rangle \otimes \hat{S}_2^2\,|z+\rangle \\ &= |z+\rangle \otimes \frac{1}{2}\left(\frac{1}{2}+1\right)\hbar^2\,|z+\rangle \,. \end{aligned}$$

We can now calculate

$$\begin{aligned}\hat{S}^2 |z+,z+\rangle &= \left(\hat{S}_1^2 + 2\hat{S}_1 \cdot \hat{S}_2 + \hat{S}_2^2\right) |z+,z+\rangle \\ &= \left[\frac{1}{2}\left(\frac{1}{2}+1\right)\hbar^2 + 2\hat{S}_1 \cdot \hat{S}_2 + \frac{1}{2}\left(\frac{1}{2}+1\right)\hbar^2\right] |z+,z+\rangle\end{aligned}$$

and using (7.5) for the middle term with only $2\hat{S}_{1z}\hat{S}_{2z}$ contributing gives $2\frac{1}{2}\frac{1}{2}$, for a total of

$$\hat{S}^2 |z+,z+\rangle = 2\hbar |z+,z+\rangle .$$

But this means that $s(s+1) = 2$ or $s = 1$, so the spin of the object $|z+,z+\rangle$ is 1. All of the triplet states (7.7) have spin 1 and the singlet state (7.8) has spin 0. Calculation of the z-projection of the spins is straightforward.

$$\begin{aligned}\hat{S}_z |z+,z+\rangle &= (\hat{S}_{1z} + \hat{S}_{1z}) |z+,z+\rangle \\ &= (\hbar/2 + \hbar/2) |z+,z+\rangle \\ &= \hbar |z+,z+\rangle\end{aligned}$$

From this calculation it is easy to see that $\hat{S}_z |z-,z-\rangle = -\hbar |z-,z-\rangle$, and $\hat{S}_z |z+,z-\rangle = 0$. We can now relabel the states in the form $|s,m_s\rangle$, giving

$$|1,1\rangle = |z+,z+\rangle \tag{7.9}$$

$$|1,0\rangle = \frac{1}{\sqrt{2}}\left(|z+,z-\rangle + |z-,z+\rangle\right) \tag{7.10}$$

$$|1,-1\rangle = |z-,z-\rangle \tag{7.11}$$

$$|0,0\rangle = \frac{1}{\sqrt{2}}\left(|z+,z-\rangle - |z-,z+\rangle\right) \tag{7.12}$$

where (7.9)-(7.11) are the triplet states and (7.12) is the singlet state.

We end this section with two examples. The first example shows how to switch bases. Suppose we want to write $|1,0\rangle$ in the $|x\pm\rangle$ basis. We would calculate as follows.

$$|1,0\rangle = \frac{1}{\sqrt{2}}\left(|z+\rangle \otimes |z-\rangle + |z-\rangle \otimes |z+\rangle\right)$$

Now substitute in $|x\pm\rangle = (|z+\rangle \pm |z-\rangle)/\sqrt{2}$, and compute

$$\begin{aligned}|1,0\rangle &= \frac{1}{\sqrt{2}}\frac{1}{2}\left[(+)+(-)\otimes(+)-(-) \quad + \quad (+)-(-)\otimes(+)+(-)\right] \\ &= \frac{1}{\sqrt{2}}\frac{1}{2}\left[(++)+(-+)-(+-)-(--) \quad + \quad (++)-(-+)+(+-)-(--)\right] \\ &= \frac{1}{\sqrt{2}}\left[(++)-(--)\right] = \frac{1}{\sqrt{2}}\left[|x+,x+\rangle - |x-,x-\rangle\right].\end{aligned}$$

For our second example we will define a **product state**. Consider the state

$$\begin{aligned}|\psi\rangle &= (|z+\rangle \otimes |z+\rangle) + (|z+\rangle \otimes |z-\rangle) \\ &= |z+\rangle_1 \otimes (|z+\rangle_2 \otimes |z-\rangle_2)\end{aligned}$$

which we were able to factor because the first ket in the Hilbert space for the first particle was the same in both terms. An example of a state that is not factorable, or not a product state is

$$|\phi\rangle = |z+, z-\rangle - |z-, z+\rangle\,,$$

and we say that particle 1 and 2 are **entangled**. The ket $|\phi\rangle$ is an entangled state. This state is an example of a **Bell pair**, after John S. Bell, and is frequently written as

$$|\psi\rangle = \frac{|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle}{\sqrt{2}}$$

for spin-1/2 particles.

7.3 Singlet and triplet states of He

Understanding triplet and singlet states requires the quantum principle of the *indistinguishably of identical particles*, in the sense that an operator that exchanges otherwise identical particles must leave the system indistinguishable. Define the exchange operator $\hat{P}_{12}$ such that when operating on a *valid* wave function of N indistinguishable particles $\psi(1, 2, 3, \ldots, N)$ we obtain

$$\hat{P}_{12}\psi(1, 2, 3, \ldots, N) = c\psi(2, 1, 3, \ldots, N)$$

where c is a complex number. The label number (n) refers to all space and spin components of particle n in the wave function. Operating twice with $\hat{P}_{12}$ gives

$$\hat{P}_{12}\hat{P}_{12}\psi(1, 2, \ldots) = c\hat{P}_{12}\psi(2, 1, \ldots) = c^2\psi(1, 2, \ldots).$$

Obviously, we should obtain the same wave function that we started with, therefore we require that $c^2 = 1$ or $c = \{+1, -1\}$. These are the eigenvalues of the exchange operator. If the wave function we are operating on is valid, then we must obtain either $\pm 1\psi$.

Now consider a two-spin system such as the He atom electronic wave function and label the electrons (1) and (2). For Fermions we require that the wave function be antisymmetric under exchange, requiring that the wave function for He be antisymmetric. We write the wave function as a product of space and spin parts, $\psi = \phi(x_1, x_2)\gamma(1, 2)$, where γ is the spin part of the wave function, and require that ψ be antisymmetric under $\hat{P}_{12}$. In other words we require that $\hat{P}_{12}\psi = \phi(x_2, x_1)\gamma(2, 1) = -\phi(x_1, x_2)\gamma(1, 2)$.

7.4 Acceptable wave functions

Definition of a valid wave function. We can use the definition of the exchange operator operating on either a spin or space wave function to determine what combinations are allowed.

If we find that $\hat{P}_{12}\phi \neq \pm\phi$ or $\hat{P}_{12}\gamma \neq \pm\gamma$, then this space or spin function is not *valid* and cannot be considered when constructing a wave function.

Consider any two-electron system. There are four enumerable spin configurations for the the electrons: $\alpha(1)\alpha(2)$, $\beta(1)\beta(2)$, $\alpha(1)\beta(2)$, and $\alpha(2)\beta(1)$. The first two spin functions have both electrons either spin up or spin down, respectively, while the last two have one of the electrons up and the other down. We can construct linear combinations of these spin functions to make a spin function that satisfies the validity requirement. Consider that $\hat{P}_{12}[\alpha(1)\beta(2)] = \alpha(2)\beta(1) \neq \pm\alpha(1)\beta(2)$, so this is not an allowed spin function. However if we form a linear combination of spin functions we can define $\gamma_\pm \equiv 2^{-\frac{1}{2}}[\alpha(1)\beta(2) \pm \alpha(2)\beta(1)]$, noting that $\hat{P}_{12}\gamma_\pm = \pm\gamma_\pm$.

The He atom has two electrons, and in the ground state, both are in $1s$-like orbitals, and must have opposite spin according to the Pauli principle. The space part for two electrons in $1s$ orbitals is $1s(1)1s(2)$, and on exchange goes to $1s(2)1s(1)$. Since both electrons are in $1s$ orbitals these two functions represent an indistinguishable state and this is a valid *symmetric* space function. Now we must add the electron spin to the wave function. Because the space part $1s(1)1s(2)$ is symmetric, the spin part must be antisymmetric. Note that the ground state wave function cannot be $\psi = 1s(1)1s(2)\alpha(1)\beta(2)$, because the spin part is not *valid*; it allows that the electrons are distinguishable in the spin part. Instead, a valid wave function is

$$\psi_g = 1s(1)1s(2)\gamma_- = \frac{1}{\sqrt{2}}1s(1)1s(2)[\alpha(1)\beta(2) - \alpha(2)\beta(1)]. \tag{7.13}$$

The total spin is $s = 0$, and the multiplicity is $(2s + 1) = (2 \cdot 0 + 1) = 1$. This is a singlet state because there is only one way to write the wave function.

7.5 First excited state of He

Now consider the first excited state of He, where one electron is promoted to a 2s state. Now the space function can only legitimately be a linear combination of $1s(1)2s(2)$ and $1s(2)2s(1)$, so that the exchange operator returns ±1, i.e., the space function must be either $\phi_\pm \equiv 2^{-\frac{1}{2}}[1s(1)2s(2) \pm 1s(2)2s(1)]$. If we take the space function ϕ_- that is antisymmetric on exchange, then we must take spin functions that are symmetric. There are three spin functions that are symmetric on exchange: $\alpha(1)\alpha(2)$, $\beta(1)\beta(2)$, and γ_+. Therefore we have three energy-degenerate wave functions for the first excited state of He. They are

$$\psi_1 = \phi_-[\alpha(1)\alpha(2)] \tag{7.14}$$

$$\psi_0 = \frac{1}{\sqrt{2}}\phi_-[\alpha(1)\beta(2) + \alpha(2)\beta(1)] \tag{7.15}$$

$$\psi_{-1} = \phi_-[\beta(1)\beta(2)]. \tag{7.16}$$

In this set of equations, the total spin s is 1, and the m_s values are 1, 0, and -1, respectively. Explicitly, the $s = 1$, $m_s = 0$ wave function for the He first excited state is given by

$$\psi_0 = \frac{1}{2}[1s(1)2s(2) - 1s(2)2s(1)][\alpha(1)\beta(2) + \alpha(2)\beta(1)]. \tag{7.17}$$

Recalling the addition of angular momentum rules, if $s_1 = s_2 = \frac{1}{2}$, then $s = |s_1+s_2|, \cdots, |s_1-s_2|$. The possibilities are $s = 1, 0$, with $m_{s=1} = \{1, 0, -1\}$ and $m_{s=0} = 0$. If the spin $s = s_1 + s_2 = 1$ for the state given in (7.15) and (7.17) is confusing, consider the result of the lowering operator $\hat{S}_- = \hat{S}_{1-} + \hat{S}_{2-}$ on the spin state $\alpha(1)\alpha(2)$. It mixes the up and down spins and gives a linear combination that is proportional to the spin part in (7.15). It lowers the m_s value from 1 to 0. Also, the total spin operator $\hat{S}^2 = (\hat{S}_1^2 + \hat{S}_2^2 + 2\hat{S}_1 \cdot \hat{S}_2)$ operating on (7.15) gives +2 for an eigenvalue, so that $s(s+1) = 2$ and $s = 1$. To see this it is necessary to recall that $\hat{S}_\pm = \hat{S}_x + i\hat{S}_y$, giving $\hat{S}_x = \frac{1}{2}(\hat{S}_+ + \hat{S}_-)$ and $\hat{S}_y = \frac{1}{2i}(\hat{S}_+ - \hat{S}_-)$. Writing out $\hat{S}^2$ gives

$$\hat{S}^2 = \hat{S}_1^2 + \hat{S}_2^2 + \hat{S}_{1+}\hat{S}_{2-} + \hat{S}_{1-}\hat{S}_{2+} + 2\hat{S}_{1z}\hat{S}_{2z}. \tag{7.18}$$

If we operate on $[\alpha(1)\beta(2)]$, and using $\hat{S}_{1-}\alpha(1) = \beta(1)$, and $\hat{S}_{1+}\beta(1) = \alpha(1)$, and so forth, we obtain

$$\hat{S}^2[\alpha(1)\beta(2)] = \left(\frac{3}{4} + \frac{3}{4}\right)[\alpha(1)\beta(2)] + [\beta(1)\alpha(2)] - 2\left(\frac{1}{2}\,\frac{1}{2}\right)[\alpha(1)\beta(2)] \tag{7.19}$$

$$= \left(\frac{3}{2} - \frac{1}{2}\right)[\alpha(1)\beta(2)] + [\beta(1)\alpha(2)] = [\alpha(1)\beta(2) + \beta(1)\alpha(2)]. \tag{7.20}$$

Likewise, we find that $\hat{S}^2[\alpha(2)\beta(1)] = [\alpha(1)\beta(2) + \alpha(2)\beta(1)]$.

7.6 Slater determinants

Considering the space and spin function separately for helium was possible in this example because of the small number of possibilities. In general, for larger systems, one constructs the antisymmetric wave function from *Slater determinants*. For the He ground state (7.13) it is given by the formula

$$\psi_g = \frac{1}{\sqrt{2}}\begin{vmatrix} 1s(1)\alpha(1) & 1s(1)\beta(1) \\ 1s(2)\alpha(2) & 1s(2)\beta(2) \end{vmatrix}. \tag{7.21}$$

Taking the determinant of a matrix gives the sum of all signed elementary products of its elements, which is to say it multiplies one distinct element from each row and does this for all possible permutations, and assigns a sign of +(-) if the permutation is even(odd).

7.7 Coulomb and Exchange Integrals

This subsection defines the *Coulomb* and *exchange* integrals for two electrons in a box.

7.7.1 Ground state wave function

Consider two electrons in a box and label them 1, and 2. Let $\{|\phi_n(x)\rangle\}$ be one-particle eigenfunctions for the uncoupled electron Hamiltonian. Write $\hat{H}_0 = \hat{H}_{1,0} + \hat{H}_{2,0}$, with $\hat{H}_{1,0}|\phi_n(x_1)\rangle = \epsilon_n|\phi_n(x_1)\rangle$ and $\hat{H}_{2,0}|\phi_n(x_2)\rangle = \epsilon_n|\phi_n(x_2)\rangle$. The Hamiltonian for the both electrons in the box can be written

$$\hat{H} = \hat{H}_0 + \hat{V}(x_1, x_2) = \frac{\hat{p}_1^2}{2m} + \frac{\hat{p}_2^2}{2m} + \hat{V}(|x_1 - x_2|), \tag{7.22}$$

where the potential only depends on the distance between the two electrons. If the potential $V(x_1, x_2)$ is weak it can be treated in perturbation theory to find the change in energy eigenvalues and eigenstates. Taking this approach, the zeroth-order ground-state wave function for this system would simply be (ignoring the potential)

$$\psi(x_1, x_2) = \phi_1(x_1)\phi_1(x_2)\frac{1}{\sqrt{2}}\left[\alpha(1)\beta(2) - \alpha(2)\beta(1)\right]. \tag{7.23}$$

The spin functions α and β denote spin up and down, respectively. The spin part of Eq. (7.23) is antisymmetric because the total wave function must be antisymmetric, and the space part of the wave function, as written, is symmetric.

7.8 Spin function normalization

Let us pause for a moment and examine the spin function normalization rules. The spin functions are eigenfunctions of the operator corresponding to the z-component of the angular momentum of the electron, $\hat{S}_{1z}$, and $\hat{S}_{2z}$. Because these operators are Hermitian, the eigenvectors of, e.g. $\hat{S}_{1z}$, are orthogonal. Explicitly, $\hat{S}_{1z}|\pm\rangle = \pm\frac{1}{2}|\pm\rangle$, so that $\langle +|-\rangle = 0$, and $\langle +|+\rangle = \langle -|-\rangle = 1$.

What we would like to do is have a rule for normalization that looks like the space normalization of a 1D wave function, $1 = \int dx\, \phi^*(x)\phi(x)$. We can work out the rules for spin by analogy with the rules for space. If the space were discretized, this integral would become a sum. In the case of spin there are only two possible states, up and down, and the sum has only two terms.

If we denote $\alpha = |+\rangle$ and $\beta = |-\rangle$, then $|+\rangle_1 = \alpha(1)$ and $|-\rangle_1 = \beta(1)$, et cetera. For space normalization we use the notation $\langle x|\phi\rangle = \phi(x)$ for the wave function, and integrate over x: $\int dx\, \phi^*(x)\phi(x) = 1$. We can write something similar for spin: $\langle m_s|+\rangle = \alpha_{m_s}$ and $\langle m_s|-\rangle = \beta_{m_s}$. For a normalized space basis state we would write $\langle\phi_n|\phi_n\rangle = 1$. We can insert a complete set of states and write

$$1 = \langle\phi_n|\phi_n\rangle = \left\langle \phi_n \middle| \left(\int dx\ |x\rangle\langle x|\right) \middle| \phi_n \right\rangle = \int dx\, \phi_n^*(x)\phi_n(x).$$

Likewise, with the spin states we can write

$$\langle +|+\rangle = \left\langle + \middle| \left(\sum_{m_s} |m_s\rangle\langle m_s|\right) \middle| + \right\rangle = 1.$$

The spin function normalization would then be written as

$$1 = \langle + \,|\, + \rangle = \sum_{m_{s1}=-\frac{1}{2}}^{\frac{1}{2}} \langle + \,|\, m_s \rangle \langle m_s \,|\, + \rangle = \sum_{m_{s1}=-\frac{1}{2}}^{\frac{1}{2}} \langle m_s \,|\, + \rangle^* \langle m_s \,|\, + \rangle \,.$$

This is written more compactly as

$$\sum_{m_s=-1/2}^{+1/2} \alpha^*_{m_s} \alpha_{m_s} = 1. \tag{7.24}$$

Because α and β are different basis functions we also have

$$\sum_{m_s=-1/2}^{+1/2} \beta^*_{m_s} \beta_{m_s} = 1. \tag{7.25}$$

The orthogonality of $|+\rangle$ and $|-\rangle$ can be written as

$$0 = \langle + \,|\, - \rangle = \left\langle + \left| \left(\sum_{m_s} |m_s\rangle \langle m_s| \right) \right| - \right\rangle$$

or

$$\sum_{m_s=-1/2}^{+1/2} \alpha^*_{m_s} \beta_{m_s} = 0. \tag{7.26}$$

Equations (7.24), (7.25), and (7.26) are the orthonormality conditions on spin. They imply that $\alpha_{m_s=+\frac{1}{2}} = 1$, and $\alpha_{m_s=-\frac{1}{2}} = 0$, and likewise for β.

Another way to see all of this is to consider a general spin state $(b_+ |+\rangle + b_- |-\rangle)$. Normalization of this state requires that

$$|b_+|^2 + |b_-|^2 = 1. \tag{7.27}$$

We can construct a wave function using this spin function. For simplicity look at

$$|\psi\rangle = |\phi_n\rangle \left(b_+ |+\rangle + b_- |-\rangle \right).$$

This wave function can be written as

$$\psi(x, m_s) = \langle x, m_s | \left\{ |\phi_n\rangle \left(b_+ |+\rangle + b_- |-\rangle \right) \right\} = \phi_n(x) \left(b_+ \alpha_{m_s} + b_- \beta_{m_s} \right).$$

Normalizing this wave function we write

$$\sum_{m_s} \int dx \, \psi^*(x, m_s) \psi(x, m_s) = 1.$$

The integral over the space part is $\int dx \, \phi^*_n(x) \phi_n(x) = 1$, and the sum over m_s, invoking the orthogonality relations, results in Eq. (7.27).

7.9 Ground state energy and first order correction

Now, going back to the wave function in Eq. (7.23), we see that it is antisymmetric and satisfies $\psi(x_2, x_1) = -\psi(x_1, x_2)$. We may now calculate $\left\langle \hat{H}_0 \right\rangle_\psi$, the expectation value of the Hamiltonian in the wave function ψ. we have

$$\left\langle \hat{H}_0 \right\rangle_\psi = \sum_{m_{s1}, m_{s2}} \int d^3x \; \psi^*(x_1, x_2) \hat{H}_0 \psi(x_1, x_2).$$

The unperturbed Hamiltonian only operates on the space part of the wave function and returns $\epsilon_1 + \epsilon_1 = 2\epsilon_1$. Carrying out the integration and summation gives 1, and the overall result is $\left\langle \hat{H}_0 \right\rangle_\psi = 2\epsilon_1$, as expected. In perturbation theory notation, we would write $E^{(0)}_{\text{gnd}} = 2\epsilon_1$. The first order change in the ground state energy due to the perturbation of the potential is

$$E^{(1)}_{\text{gnd}} = \left\langle \hat{V} \right\rangle_\psi .$$

Writing this equation out is cumbersome. Dropping the hats, and leaving off the limits of the integrals and sums, we can write

$$\begin{aligned} E^{(1)}_{\text{gnd}} \;\; = \;\; & \frac{1}{2} \sum_{s_1 s_2} \int dx_1 dx_2 \, \phi^*_1(x_1) \phi^*_1(x_2) \left[\alpha^*(1)\beta^*(2) - \alpha^*(2)\beta^*(1) \right] \times \\ & V \phi_1(x_1) \phi_1(x_2) \left[\alpha(1)\beta(2) - \alpha(2)\beta(1) \right]. \end{aligned}$$

This can be written more compactly by writing $\phi_{12} = \phi_1(x_2)$, and so on for the space functions, and $\alpha(1) = \alpha_1$ for the spin functions. Expanding and combining terms results in

$$\frac{1}{2} \sum_{s_1 s_2} \int dx_1 dx_2 \; \left(\phi^*_{11} \phi^*_{12} V \phi_{11} \phi_{12} \right) \left(\alpha^*_1 \beta^*_2 \alpha_1 \beta_2 - \alpha^*_1 \beta^*_2 \alpha_2 \beta_1 - \alpha^*_2 \beta^*_1 \alpha_1 \beta_2 + \alpha^*_2 \beta^*_1 \alpha_2 \beta_1 \right).$$

Labeling the four terms in brackets as $((1) + (2) + (3) + (4))$, and denoting $\alpha(1)\beta(2) \rightarrow \alpha(m_{s1})\beta(m_{s2})$, and invoking the orthogonality relations for the sums, it can be easily shown that terms 2 and 3 are zero, and terms 1 and 4 are each unity. For example, the first term will be $\sum_{s_1} \alpha^*_1 \alpha_1 \sum_{s_2} \beta^*_2 \beta_2$, while the second term will be $\sum_{s_1} \alpha^*_1 \beta_1 \sum_{s_2} \alpha_2 \beta^*_2$. This gives us finally

$$E^{(1)}_{\text{gnd}} = \int dx_1 dx_2 \; \left(\phi^*_{11} \phi^*_{12} V \phi_{11} \phi_{12} \right).$$

This is called a Coulomb integral if the potential is $\frac{e^2}{|x_1 - x_2|}$. This integral is a measure of the repulsion. From the viewpoint of a two spin system, the state we have described is called the singlet state.

7.10 First excited state for two electrons in a box

Let us now consider the first excited state of two electrons in a box. We must still have a wave function that is overall antisymmetric. If the spatial part of the wave function is

antisymmetric then the spin part can be symmetric. We can write the first excited state in this manner as

$$\left\{ \begin{array}{c} \psi_1 \\ \psi_0 \\ \psi_{-1} \end{array} \right\} = \frac{1}{\sqrt{2}} \left[\phi_1(x_1)\phi_2(x_2) - \phi_1(x_2)\phi_2(x_1)\right] \begin{cases} \alpha_1\alpha_2 \\ \frac{1}{\sqrt{2}}\left(\alpha_1\beta_2 + \alpha_2\beta_1\right) \\ \beta_1\beta_2 \end{cases} \tag{7.28}$$

where each of the three spin functions is an acceptable symmetric function. In this case it is easy to see that $\hat{H}_0\psi = \epsilon_1 + \epsilon_2$, giving the total kinetic energy as expected. Without the perturbation we have $E^{(0)}_{1\text{st}} = \epsilon_1 + \epsilon_2$. Now we want to calculate the first order change in the energy due to the perturbation V. Let us look at $E^{(1)}_{1\text{st}}$ for the first option, ψ_1. We have

$$\langle V \rangle_{\psi_2} = \frac{1}{2} \sum_{s_1 s_2} \int d^2x \, \left[\phi_{11}\phi_{22} - \phi_{12}\phi_{21}\right]^* \alpha_1^*\alpha_2^* V \left[\phi_{11}\phi_{22} - \phi_{12}\phi_{21}\right] \alpha_1\alpha_2.$$

Now the sum over the spin functions is just $\sum_{s_1} \alpha_1^*\alpha_1 \sum_{s_2} \alpha_2^*\alpha_2 = 1$, and we are left with the spatial integrals

$$\frac{1}{2} \int d^2x \, \left(\phi_{11}^*\phi_{22}^* V \phi_{11}\phi_{22} + \phi_{12}^*\phi_{21}^* V \phi_{12}\phi_{21} - \phi_{12}^*\phi_{21}^* V \phi_{11}\phi_{22} - \phi_{11}^*\phi_{22}^* V \phi_{12}\phi_{21}\right). \tag{7.29}$$

Look carefully at the second term in the kernel of the integral for the following discussion. The first two terms are integrals like

$$J = \int dx_1 dx_2 \, \phi_1^*(x_2)\phi_1(x_2) V \phi_2^*(x_1)\phi_2(x_1),$$

where the argument of ϕ_1 and ϕ_1^* is x_2, for example, and are of the Coulomb form. However, the last two integrals are of the form

$$K = \int dx_1 dx_2 \, \phi_1^*(x_2)\phi_1(x_1) V \phi_2^*(x_1)\phi_2(x_2),$$

where the argument of ϕ_1 and ϕ_1^* is x_1 and x_2, respectively. This is an example of an *exchange integral*. The first two terms of Eq. (7.29) give the same contribution, as do the last two terms. The three different choices of spin functions in Eq. (7.28) all sum to 1, so all three of the options for the first excited state give the same first order change to the energy in perturbation theory, $E^{(1)}_{1\text{st}} = J + K$.

In general, construction of antisymmetric wave functions can be done with the Slater determinant, and this can be done for a general many electron wave function. Constructing wave functions in this manner with a single Slater determinant (for closed shell molecular orbitals) is called the Hartree-Fock method. The exchange energies are important, and depend on electron correlation. Both Coulomb and exchange integrals contribute to what is called the "correlation energy"; it is defined to be the difference of the exact energy and the Hartree-Fock energy, $E_{\text{corr}} = E_{\text{exact}} - E_{\text{HF}}$. Terminology varies, but in general if one uses more than a single Slater determinant to more accurately incorporate the correlation energy the method is called "post-Hartree Fock".

8 Conclusions

This concludes our brief summary of the basic concepts you will need to follow the lectures. Please be sure you understand the contents of this document in full by the midpoint of the semester. It will be critical in moving forward with the more complicated material. Please see me if you have questions regarding anything in these notes.

A Mathematical Tools

A.1 Derivatives

I will assume you have mastered calculus and understand the concepts of limits, derivatives, Riemann integration, and so forth. It is important to know the difference between a total and partial derivative. A partial derivative acts explicitly only on the variable given, so $\frac{\partial}{\partial t}$ acts only on the variable t to the right of the operator. A total derivative can "get to" the variable when there are nested functions, for example $\frac{d}{dt}f(x(t),t) = \frac{\partial f}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial f}{\partial t}$. So if $f(x,t) = x^2 + 5t$, then $\frac{d}{dt}f = 2x\frac{\partial x}{\partial t} + 5$, but $\frac{\partial}{\partial t}f = 5$.

Shorthand notation for derivatives will be used extensively and will make our work much easier. In particular we will define $\partial_t = \frac{\partial}{\partial t}$, and likewise for spatial derivatives. We will also use the subscript notation with and sometimes without a comma to denote partial derivatives. All of the following are equivalent:

$$\frac{\partial}{\partial t}\phi(t) = \partial_t\phi(t) = \phi_{,t}(t) = \phi_t(t).$$

A.2 Integration by parts

We will frequently use the method of integration by parts, where an integrand contains a product of two functions that are integrated separately. Using the "d" notation from calculus we can write $d(ab) = b\,da + a\,db$, and integrating both sides obtain $ab = \int b\,da + \int a\,db$, so that $\int a\,db = ab - \int b\,da$. In more general terms we write:

$$\int_a^b g(x)\,f(x)\,dx = g(x)\,F(x)\Big|_a^b - \int_a^b F(x)\,g'(x)\,dx \tag{A.1}$$

where we have made the association $f(x)dx = db$, and where $F'(x) = f(x)$ or $F(x)$ is the anti-derivative of $f(x)$. As an example, we can integrate $x\cos(x)$ by integrating the cosine in the first term and differentiating the x in the second term: $\int dx\; x\cos(x) = x\sin(x) - \int dx\;\sin(x) = x\sin(x) + \cos(x)$.

A.3 Taylor's series

One of the most common approximations used in physics and mathematics is the Taylor expansion. The general form for the Taylor expansion for a function $f(x)$ about the point x is given by

$$f(x+h) = f(x) + h\,f'(x) + \frac{h^2}{2!}\,f''(x) + \frac{h^3}{3!}\,f^{(3)}(x) + \cdots \tag{A.2}$$

In physics we often make approximations where one quantity is small or large with respect to another. When we do, we frequently find expression such as $g(x) = (1 \pm x)^n$, where x is small. We can Taylor expand this expression about $g(0)$ to find $g(x) \approx 1 \pm nx$. Note that if $n = 1/2$ we have, to first order, $(1-x)^{-1/2} \approx 1 + \frac{x}{2}$. A surprising number of

physics formulas come from this simple approximation. A great example is the low velocity approximation of the energy of a particle of rest mass m_0. Relativistically, the particle has energy $E = m_0c^2/\sqrt{1-\frac{v^2}{c^2}}$, and using our approximation, if $v/c \ll 1$, we have

$$E = m_0c^2(1+\frac{v^2}{2c^2}) = m_0c^2 + \frac{1}{2}m_0v^2.$$

The sum of the rest energy and the non-relativistic kinetic energy $\frac{1}{2}m_0v^2$ that you are used to.

A.4 Fourier Transforms

I assume you have seen Fourier integrals and are aware of Fourier analysis, although we will spend significant time working this out in detail in class. Remember that differential equations are often approached by transforming the equation through an integral transform to something that can be solved algebraically. The standard Fourier transform pair is defined by the equations

$$g(k) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} dx\, e^{-ikx}\, f(x) \tag{A.3}$$

$$f(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} dk\, e^{+ikx}\, g(k). \tag{A.4}$$

Look at the integral for $f(x)$. This is a function of the variable x, not of k, which is "integrated out." The integral is a weighted sum of the exponential e^{+ikx} with weighting function $g(k)$. Likewise for the function $g(k)$, where the variable x is integrated out. Also note the symmetry in the definitions that differ only in the sign of the exponent.

A.5 Dirac delta function

The Dirac delta function $\delta(x)$ is actually not a formal function, but rather a distribution, and is defined by its behavior when integrating over it. In general, the delta function is infinitely peaked at $x = 0$ but has finite area. Mathematically this is written $\int_{-\infty}^{\infty}\delta(x)\,dx = 1$, while $\delta(x \neq 0) = 0$. It has the effect of picking out a function value when performing an integration in the following sense: $f(0) = \int_{-\epsilon}^{\epsilon} dx\,\delta(x)\,f(x)$, where ϵ is a very small number. Note that we only need to integrate over the region where the delta function is nonzero. The Dirac delta function is intimately related to the Heaviside step function defined by

$$\theta(x) = \begin{cases} 1 & x \geq 0 \\ 0 & x < 0 \end{cases}.$$

Using the usual definition of the derivative $\theta'(x) = \lim_{h\to 0}\frac{\theta(x+h)-\theta(x)}{h}$, which is indeterminate as $h \to 0$. Still, the derivative of the Heaviside step function furnishes one definition of the

Dirac delta function: $\delta(x) = \theta'(x)$. How? Consider the integral $\int f(x)\delta(x)dx$. Using our definition of the delta function from Heaviside we write $\int f(x)\theta'(x)dx$, and integrating by parts we find $f(x)\theta(x)|_{-\epsilon}^{\epsilon} - \int f'(x)\theta(x)dx = f(\epsilon) - \int_0^{\epsilon} \frac{df}{dx}dx = f(\epsilon) - (f(\epsilon) - f(0)) = f(0)$. In agreement with the calculation above.

A.6 Gaussian integrals

Gaussian integrals come up frequently in many areas in physics. There are a few tricks to calculating these integrals shown here that we will cover in class in detail. The simplest Gaussian integral is given by

$$I = \int_{-\infty}^{\infty} e^{-x^2}.$$

The trick to solving this integral is to write $I^2 = \int dx \int dy\, e^{-(x^2+y^2)}$ and change to polar coordinates, with $dx\,dy \to r\,dr\,d\theta$. The integral is easy and one finds that $I = \sqrt{\pi}$. A generalization of this integral is

$$J = \int_{-\infty}^{\infty} e^{-ax^2}$$

that can be solved easily by the substitution $u = a^{1/2}x$, giving $J = \sqrt{\pi/a}$. Note that we can consider $J = J(a)$, it is a function of the parameter a. The next step makes the integrand more complicated with the integral

$$K = \int_{-\infty}^{\infty} x\, e^{-ax^2}$$

but this integral is seen to vanish by symmetry; the integrand has odd symmetry about the origin. However the integral

$$L = \int_{-\infty}^{\infty} x^2 e^{-ax^2}$$

does not vanish, it has even symmetry about the origin. Its value can be determined by noting that $L = -J'(a)$, and therefore $L = \frac{\sqrt{\pi}}{2a^{3/2}}$. Most of the Gaussian integrals you encounter in this course can be solved by repeated differentiation under the integral sign of these prototype integrals.

A.7 Ordinary differential equations

The Schrödinger equation is a partial differential equation, but can often be simplified in specific problems to a set of ordinary differential equations (ODEs) in the time and space variables. You should be able to integrate the most simple ODEs that come up frequently. For example, these include equations such as $y' = \pm ay$ and $y'' = \pm ay$, with exponential and complex exponential solutions, respectively. In general, any *linear* ODE with constant coefficients

$$y^{(n)} + c_{n-1}y^{(n-1)} + \cdots + cy' + c_0 y = 0$$

can be solved with exponentials. In the example $y' = \pm ay$, we write $dy/dx = \pm ay$. This is a perfect differential, and we can write $\frac{dy}{y} = \pm a\, dx$. This can be immediately integrated to give $ln(y) = \pm ax$, or after exponentiating both sides, $y = y_0 e^{\pm ax}$. In our second order example, $y'' = \pm ay$, we can try the solution $y = e^{rx}$. Then $y' = re^{rx}$, and $y'' = r^2 e^{rx}$, and we have that $e^{rx} r^2 = \pm a e^{rx}$, or $r^2 = \pm a$. The resulting polynomial in r is called the indicial equation. Solving for r we find $r = \sqrt{a}$ or $r = i\sqrt{a}$, and the general solution is written $y = c_1\, e^{x\sqrt{a}} + c_2\, e^{ix\sqrt{a}}$. More sophisticated methods for solving ODEs will be covered in class.

A.8 Probability theory

Some basic facts about probability theory are very useful in the definition of expectation values in quantum mechanics. First, we define a probability when we know something about possible outcomes of an experiment, such as rolling dice or flipping a coin. We define from this a *sample space*, say heads or tails for coin flipping, and then demand that the sum of probabilities of the possible outcomes must sum to unity: $\sum_n P_n = 1$, where P_n is the probability for some event. This is called normalization of the probability. The average of an external function f over the various possibilities is then defined as

$$\bar{f} = \frac{\sum_n P_n f_n}{\sum_n P_n}. \tag{A.5}$$

The expression in (A.5) allows for non-normalized probabilities, and the denominator would be unity if they are normalized. We can generalize (A.5) for continuous variables by integrating:

$$\bar{f} = \frac{\int dx\, P(x)\, f(x)}{\int dx\, P(x)}.$$

A.9 Multidimensional Taylor Expansion

The Taylor expansion for one dimension is

$$f(x+h) = f(x) + h\, f'(x) + \frac{h^2}{2!}\, f''(x) + \frac{h^3}{3!}\, f^{(3)}(x) + O(h^4), \tag{A.6}$$

or more compactly as $f(x+h) = \sum_n^\infty \frac{h^n}{n!}\, f^{(n)}(x)$. Moving on to two dimensions, the function $f(x,y)$ can be expanded in both variables by performing the usual expansion (A.6) for one variable first, and then for the other.

$$\begin{aligned}
f(x+h, y+t) &= f(x, y+t) + h\, f_x(x, y+t) + \frac{h^2}{2!}\, f_{xx}(x, y+t) + \frac{h^3}{3!}\, f_{xxx}(x, y+t) + \dots \\
f(x, y+t) &= f(x,y) + t\, f_y(x,y) + \frac{t^2}{2!}\, f_{yy}(x,t) + \frac{t^3}{3!}\, f_{yyy}(x,y) + \dots \\
f_x(x, y+t) &= f_x(x,y) + t\, f_{xy}(x,y) + \frac{t^2}{2!}\, f_{xyy}(x,y) + \frac{t^3}{3!}\, f_{xyyy}(x,y) + \dots \\
f_{xx}(x, y+t) &= f_{xx}(x,y) + t\, f_{xxy}(x,y) + \frac{t^2}{2!}\, f_{xxyy}(x,y) + \frac{t^3}{3!}\, f_{xxyyy}(x,y) + \dots
\end{aligned}$$

Combining these equations we gives

$$f(x+h,y+t) = \left[f(x,y) + t\, f_y(x,y) + \frac{t^2}{2!}\, f_{yy}(x,t) + \frac{t^3}{3!}\, f_{yyy}(x,y) + \dots\right] +$$
$$h\left[f_x(x,y) + t\, f_{xy}(x,y) + \frac{t^2}{2!}\, f_{xyy}(x,y) + \frac{t^3}{3!}\, f_{xyyy}(x,y) + \dots\right] +$$
$$\frac{h^2}{2!}\left[f_{xx}(x,y) + t\, f_{xxy}(x,y) + \frac{t^2}{2!}\, f_{xxyy}(x,y) + \frac{t^3}{3!}\, f_{xxyyy}(x,y) + \dots\right] +$$
$$\frac{h^3}{3!} f_{xxx}(x,y+t) + \dots$$

Collecting terms gives

$$f(x+h,y+t) = f(x,y) + (h\partial_x + t\partial_y)f(x,y) + \frac{(h\partial_x + t\partial_y)^2}{2!}f(x,y) + \dots$$
$$= \sum_n^\infty \frac{(h\partial_x + t\partial_y)^n}{n!} f(x,y).$$

Index