Quantum Chemistry: Fundamentals and Methods

Chem 115A/215A Lecture Notes

Thomas P Fay

University of California, Los Angeles
thomaspfay@ucla.edu

Contents

Introduction

These lecture notes accompany the Chem 115A/215A Graduate Quantum Chemistry course at UCLA. This courser aims to provide an introduction to the "axiomatic" approach to quantum mechanics, where everything is built up from a few fundamental postulates together with the toolkit of linear algebra. In this course we aim to review the fundamentals of linear algebra and use this to build up a quantum mechanical framework. This framework is applied to a set of important basic systems that form the foundations of many other problems in quantum mechanics, namely the harmonic oscillator, angular momenta, and ultimately the hydrogen atom.

Assumed knowledge

A basic working knowledge of introductory undergraduate quantum mechanics, multivariable calculus and linear algebra is assumed including the topics

- Basics of (finite-dimensional) vectors and matrices. Matrix-vector multiplication, determinants, inverses, eigenvalues and eigenvectors.
- Multivariable calculus, partial differentiation, changes of variables, integration, solving simple differential equations.
- The time independent and time-dependent Schrödinger equation.
- Position-space wave functions, position and momentum operators.
- Wave functions for simple 1D systems including the particle in a box, the particle on a ring, and the harmonic oscillator.
- Wave functions for a particle in 2D and 3D box and the hydrogen atom.

If you are unfamiliar with any of these topics, I advise you to review your favourite undergraduate physical chemistry textbook, or one of many other resources available on these topics.

Accompanying Textbooks

Much of this course is based on Modern Quantum Mechanics by J.J. Sukurai and Jim Napolitano. Molecular Quantum Mechanics by Peter Atkins can also serve as a useful complementary text to parts of this course.

Acknowledgements

These course materials were prepared based on materials kindly provided to me by Benjamin Schwartz, Daniel Neuhauser, and David Manolpoulos.

1 Linear Algebra

1.1 Vector Spaces

A vector space V consists of vectors $|v\rangle$ and complex numbers. (We will use $|v\rangle, |u\rangle, |w\rangle$ etc. to denote vectors).

We can combine two vectors through **vector addition** to get a new vector, and we can combine a vector with a complex number through **scalar multiplication** to get a new vector.

Vector spaces have the following properties:

1. **Closure under addition:** For any two vectors $|v\rangle$ and $|u\rangle$

$$|u\rangle + |v\rangle = |w\rangle \text{ is a vector in the same space}$$

2. **Closure under scalar multiplication:** For any complex number a and vector $|v\rangle$

$$a\,|v\rangle = |w\rangle \text{ is a vector in the same space}$$

3. **Commutativity of vector addition:**

$$|u\rangle + |v\rangle = |v\rangle + |u\rangle$$

4. **Associativity of vector addition:**

$$|u\rangle + (|v\rangle + |w\rangle) = (|v\rangle + |u\rangle) + |w\rangle$$

5. **Zero vector:** There exists a zero vector 0_V such that for any $|v\rangle$

$$|v\rangle + 0_V = |v\rangle$$

6. **Subtraction:** For any vector $|v\rangle$ there exists an additive inverse $-|v\rangle$ in the same space such that

$$0_V = |v\rangle + (-|v\rangle) \equiv |v\rangle - |v\rangle$$

7. **Compatibility of scalar and complex-number multiplication:** For any vector $|v\rangle$ and complex numbers a and b:

$$a(b\,|v\rangle) = (ab)\,|v\rangle$$

8. **Identity of scalar multiplication:** For any vector

$$|v\rangle = 1\,|v\rangle$$

9. **Distributivity of scalar multiplication with vector additon:**

$$a(|v\rangle + |u\rangle) = a\,|v\rangle + a\,|u\rangle$$

10. **Distributivity of scalar multiplication with complex number additon:**

$$(a+b)\,|v\rangle = a\,|v\rangle + b\,|v\rangle$$

This is a convoluted way of saying that vector addition is compatible with multiplication and addition of complex numbers, and obeys the same basic rules of algebra that we are used to with numbers. This means we don't have to worry about using

1.2 Examples

Q: Why abstract in this way?

A: Because lots of different objects (in QM and beyond) have these properties, so we can use these properties to uncover general truths about many things.

1.2.1 Example 1: Complex numbers, $V = \mathbb{C}$

Checks:

1. If z and $w \in \mathbb{C}$ then $z + w \in \mathbb{C}$ ✓
2. If a and $w \in \mathbb{C}$ then $zw \in \mathbb{C}$ ✓
3. $z + w = w + z$ for complex numbers ✓
4. $z + (w + v) = (z + w) + v$ for complex numbers ✓
5. The zero vector is just the number 0: $0_V \equiv 0$ and $z + 0 = z$ ✓
6. $-z$ does exist in $\mathbb{C}$ ✓
7. Scalar multiplication is just standard complex number multiplication so $a(bz) = (ab)z$ ✓
8. $z = 1z$ ✓
9. Scalar multiplication is the same as complex number multiplication so $a(z + w) = az + aw$ ✓
10. Same as above $(a + b)z = az + bz$ ✓

1.2.2 Example 2: Complex numbers with $\mathrm{Re}[z] \geq 0$

V is the set of complex numbers $z = x + iy$ with $x, y \in \mathbb{R}$, and $x \geq 0$, with vector addition and scalar multiplication defined in the same way as complex number multiplication.

Properties 1, 3- - 5 and 7 – 10 all hold, from the above but if $x > 0$ then $-z = -x - iy$, but $(-x) \ngeq 0$. So property 6 does not hold for this vector space. Likewise for $a = -1$, $az = -x - iy$, which is also not a complex number with positive real part, so the set is not closed under scalar multiplication.

This means V as defined is **not** a vector space.

1.2.3 Example 3: Pairs of complex numbers $V = \mathbb{C}^2$

$$|v\rangle = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \quad v_1, v_2 \in \mathbb{C}, \quad |v\rangle \in \mathbb{C}^2 \leftarrow \text{pair of } \mathbb{C}\text{-numbers}$$

[Also we could consider triples $= \mathbb{C}^3$, etc. of complex numbers.]

with addition:

$$|v\rangle + |u\rangle = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} + \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \stackrel{\text{def}}{=} \begin{pmatrix} v_1 + u_1 \\ v_2 + u_2 \end{pmatrix}$$

and multiplication:

$$a\,|v\rangle = a \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \stackrel{\text{def}}{=} \begin{pmatrix} av_1 \\ av_2 \end{pmatrix}$$

Verify for yourself that this satisfies the requirements to be a vector space.

1.2.4 Example 3: Complex valued functions of a real number

$$|\psi\rangle = \psi(x) \quad \psi(x) \in \mathbb{C}, \quad \forall x \in \mathbb{R}$$

Let's check the requirements for this to be a vector space, with the following definitions of addition and scalar multiplication defined in the obvious way.

Addition:

$$|\psi\rangle + |\phi\rangle = \psi(x) + \phi(x)$$

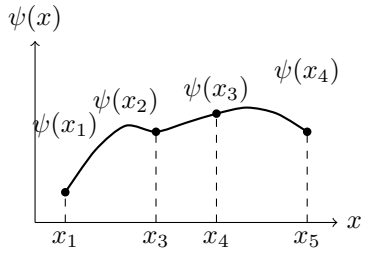
Scalar multiplication:

$$a\,|\psi\rangle = a\psi(x)$$

All properties satisfied. ✓

Note: Functions and "normal" vectors are not so different!

Pick a set of points in $\mathbb{R}$, $x_1, x_2, x_3, \ldots$



and stack them

$$|v\rangle = \begin{pmatrix} \psi(x_1) \\ \psi(x_2) \\ \psi(x_3) \\ \vdots \end{pmatrix}$$

"Just" make the sequence of x_n infinite (and upset a mathematician)!

Then the information in $|v\rangle$ fully describes $\psi(x)$.

1.3 Properties of vector spaces

We won't dive too deeply into proving lots of basic properties of vector spaces here formally. The utility of the abstract defining properties of vector spaces is that we can apply all of our basic algebra directly to vectors.

As a simple example of using these properties let's prove one very useful thing: $0\,|v\rangle = 0_V$.

Proof: We can write 0_V using property 6 and go from there

$$\begin{aligned}
0_V &= |v\rangle - |v\rangle \text{ from 6} \\
&= 1\,|v\rangle - |v\rangle \text{ from 8} \\
&= (0+1)\,|v\rangle - |v\rangle \text{ basic property of } \mathbb{C} \text{ numbers} a = a+0 = 0+a \\
&= (0\,|v\rangle + 1\,|v\rangle) - |v\rangle \text{ from 10} \\
&= (0\,|v\rangle + |v\rangle) - |v\rangle \text{ from 8} \\
&= 0\,|v\rangle + (|v\rangle) - |v\rangle) \text{ from 4} \\
&= 0\,|v\rangle + 0_V \text{ from 6} \\
&= 0\,|v\rangle \text{ from 5} \quad \square
\end{aligned}$$

We have just proved something very useful about vector spaces, that we didn't have to assume from the fundamental properties.

In a similar manner we could prove other useful things like $(-1)\,|v\rangle = -\,|v\rangle$.

Note: From here onwards we will shamelessly abuse notation and write $0_V \equiv 0$.

From the context it should always be obvious if we mean the zero of the vector space 0_V or the complex number 0.

1.4 Inner products

Given two vectors in a vector space, we can define an inner product as a function of two vectors that returns a complex number

$$\langle v|u\rangle = c$$

where $|v\rangle$, $|u\rangle$ are vectors and c is a complex number.

It is an inner product if this function has the following properties:

(i) **Conjugate Symmetry**

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

Note: we will use z^* to denote the complex conjugate of z i.e. for $z = x+iy$ with $x, y \in \mathbb{R}$, $z^* = x-iy$.

(ii) **Linearity:**

$$\langle u|av+bw\rangle = a\,\langle u|v\rangle + b\,\langle u|w\rangle$$

for $a, b \in \mathbb{C}$

(iii) **Positivity:**

$$\langle u|u\rangle > 0 \quad \text{for all } |u\rangle \neq 0$$
$$(\langle u|u\rangle \in \mathbb{R} \text{ by (i)})$$

From these basic properties we can prove interesting things.

(1) For $|u\rangle \neq 0$, if $|v\rangle = 0$ then $\langle u|v\rangle = 0$.

Proof:

$$\begin{aligned}
0 = 0 \cdot |w\rangle \quad &\text{for any } |w\rangle \neq 0 \\
\text{if } |v\rangle = 0 \Rightarrow |v\rangle = 0\,|w\rangle \quad &\text{for } |w\rangle \neq 0 \\
\text{By (ii):} \quad \langle u|v\rangle = \langle u|0\cdot w\rangle = 0 \cdot \langle u|w\rangle &= 0 \quad \square
\end{aligned}$$

(2) $\langle au + bv|w\rangle = a^* \langle u|w\rangle + b^* \langle v|w\rangle$
Note: we will sometimes denote $a|v\rangle + b|u\rangle \equiv |av + bu\rangle$
Proof:

$$\begin{aligned}\langle au + bv|w\rangle &= \langle w|au + bv\rangle^* \quad \text{(by (i))} \\ &= (\langle w|u\rangle a + \langle w|v\rangle b)^* \quad \text{(by (ii))} \\ &= a^* \langle w|u\rangle^* + b^* \langle w|v\rangle^* \quad \text{(basic property of } ^*) \\ &= a^* \langle u|w\rangle + b^* \langle v|w\rangle \quad \text{(by (i))} \quad \square\end{aligned}$$

1.4.1 Examples of inner products

(a) For $|v\rangle \in \mathbb{C}^n \quad |v\rangle = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$

$$\langle u|v\rangle = \sum_{k=1}^{n} u_k^* v_k$$

(b) For functions of $x \in \mathbb{R}$

$$\langle \psi|\phi\rangle = \int_{-\infty}^{+\infty} \psi(x)^* \phi(x)\, dx$$

Exercise: Verify properties (i)-(iii) for (a) and (b).

1.4.2 The Dual Space

The inner product allows us to define the "dual" of a vector $|u\rangle$:

$$\langle u| = (|u\rangle)^\dagger$$

where the "dagger" † means take the dual of $|u\rangle$.
The dual $\langle u|$ is defined as a function from V to $\mathbb{C}$

$$\langle u|\,(|v\rangle) = c$$

that yields the inner product with $|v\rangle$:

$$\langle u|\,(|v\rangle) = \langle u|v\rangle$$

So for example, for functions in 1D :

$$\langle \psi| \neq \psi(x)^*$$

$$\text{but} \quad \langle \psi| = \int_{-\infty}^{+\infty} \psi(x)^* \ \cdot \ dx$$

with another function in the integral implicitly, i.e.

$$\text{but} \quad \langle \psi|\,(|\phi\rangle) = \int_{-\infty}^{+\infty} \psi(x)^* \phi(x) dx$$

From the properties of the inner product:

$$\begin{aligned} \langle av| &= a^* \langle v| \quad \text{for } a \in \mathbb{C} \\ \langle v+u| &= \langle v| + \langle u| \end{aligned}$$

Often it's useful to treat $\langle v|$ and $|u\rangle$ separately when dealing with inner products.

1.4.3 Orthogonality and norms

We can also define the norm of a vector as

$$\| |v\rangle \| = \sqrt{\langle v|v\rangle}$$

From (iii) we know that if $\| |v\rangle \| = 0$ then $|v\rangle = 0$.

Now we can define some other useful things

- $|v\rangle$ and $|w\rangle$ are parallel if ($|v\rangle \neq 0$ and $|w\rangle \neq 0$)

$$\langle v|w\rangle = \| |v\rangle \| \cdot \| |w\rangle \| = \sqrt{\langle v|v\rangle \langle w|w\rangle}$$

- $|v\rangle$ and $|w\rangle$ are orthogonal if

$$\langle v|w\rangle = 0 \quad \text{and } |v\rangle \neq 0 \text{ and } |w\rangle \neq 0$$

- $|v\rangle$ is normalized if

$$\langle v|v\rangle = 1$$

Now we can show some interesting general properties

Triangle Inequality:

$$\| |v\rangle + |w\rangle \| \leq \| |u\rangle \| + \| |v\rangle \| \quad (TI)$$

and the **Cauchy-Schwarz Inequality:**

$$| \langle v|w\rangle |^2 \leq \langle v|v\rangle \cdot \langle w|w\rangle \quad (CS)$$

1.4.4 Proof of (CS):

If $|v\rangle = 0$ then (CS) holds. For $|v\rangle \neq 0$ define

$$|z\rangle = |w\rangle - \frac{\langle v|w\rangle}{\langle v|v\rangle} |v\rangle$$

Now consider $\langle v|z\rangle$

$$\begin{aligned} \langle v|z\rangle &= \langle v|w\rangle - \frac{\langle v|w\rangle}{\langle v|v\rangle} \langle v|v\rangle \\ &= \langle v|w\rangle - \langle v|w\rangle = 0 \end{aligned}$$

So $|v\rangle$ and $|z\rangle$ are orthogonal and

$$|w\rangle = \frac{\langle v|w\rangle}{\langle v|v\rangle}|v\rangle + |z\rangle$$

Now consider $\langle w|w\rangle$

$$\begin{aligned}
\langle w|w\rangle &= \left(\langle z| + \frac{\langle v|w\rangle^*}{\langle v|v\rangle}\langle v|\right)\left(|v\rangle\frac{\langle v|w\rangle}{\langle v|v\rangle} + |z\rangle\right) \\
&= \langle z|z\rangle + \frac{\langle v|w\rangle^*\langle v|w\rangle}{\langle v|v\rangle^2}\cdot\langle v|v\rangle \\
&+ \frac{\langle v|w\rangle^*}{\langle v|v\rangle}\cancelto{0}{\langle v|z\rangle} + \frac{\langle v|w\rangle}{\langle v|v\rangle}\cancelto{0}{\langle z|v\rangle} \\
&= \langle z|z\rangle + \frac{|\langle v|w\rangle|^2}{\langle v|v\rangle}
\end{aligned}$$

This gives

$$\langle w|w\rangle\langle v|v\rangle = \overbrace{\langle z|z\rangle\langle v|v\rangle}^{\geq 0} + |\langle v|w\rangle|^2 \geq |\langle v|w\rangle|^2 \quad \square$$

As an exercise prove the triangle inequality.

1.5 Orthonormal basis

An orthonormal basis of a vector space is a set of vectors $|n\rangle$ such that

(i) $\langle n|n\rangle = 1$

(ii) $\langle n|m\rangle = 0$ if $|n\rangle \neq |m\rangle$ i.e. $\langle n|m\rangle = \delta_{nm} \overset{\text{def}}{=} \begin{cases} 1 & n = m \\ 0 & n \neq m \end{cases}$

(iii) Any vector $|v\rangle$ can be written as

$$|v\rangle = \sum_n |n\rangle c_n \quad \text{where } c_n \in \mathbb{C}$$

Q: How do we find c_n?

A: Consider $\langle n|v\rangle$

$$\begin{aligned}
\langle n|v\rangle &= \langle n|\left(\sum_{n'}|n'\rangle c_{n'}\right) \\
&= \sum_{n'}\langle n|n'\rangle c_{n'} \\
&= \sum_{n'}\delta_{nn'}c_{n'} \\
&= \overbrace{\delta_{n,0}}^{0}\cdot c_1 + \overbrace{\delta_{n,1}}^{0}\cdot c_1 + \cdots + \overbrace{\delta_{n,n}}^{1}\cdot c_n + \overbrace{\delta_{n,n+1}}^{0}\cdot c_{n+1} + \cdots \\
&= c_n \cdot 1 = c_n
\end{aligned}$$

1.6 Linear independence

A set of vector $|v_n\rangle \neq 0$, with $n = 1, 2, \ldots, N$ is linearly independent if

$$\sum_{k=1}^{N} a_n |v_n\rangle = 0 \Rightarrow \text{all } a_n = 0$$

If this is not true for a set of vectors $|v_n\rangle$, that means there exists a vector in this set $|v_k\rangle$ with $a_k =\neq 0$

$$0 = a_k |v_k\rangle + \sum_{n \neq k} a_n |v_n\rangle$$

we can rearrange this to give

$$|v_k\rangle = \sum_{n \neq k} a'_n |v_n\rangle$$

where $a'_n = -a_n/a_k$. So if a set of vectors are linearly dependent (i.e. not linearly dependent), we can write at least one of the vectors as a linear combination of the others. This is an equivalent definition of linear independence/dependence.

1.7 Dimensionality

A vector space is finite-dimensional with dimension N if at most N vectors can be linearly independent

i.e.

$$\sum_{k=1}^{N} a_k |v_k\rangle = 0 \Rightarrow \text{all } a_k = 0$$

but

$$\sum_{n=1}^{N+1} a_k |v_k\rangle = 0 \nRightarrow \text{all } a_k = 0$$

Any vector space that is not finite is infinite.

Finite vector spaces are much easier to work with but unfortunately a lot of vector spaces we encounter in QM are infinite, but when doing numerical calculations we have to approximate them as finite. Almost all results from finite vector spaces transfer over to infinite dimensionality vector spaces, so we generally leave the more careful treatment of infinite vector spaces to the mathematicians.

1.8 Basis linear independence

The basis is linearly independent i.e. for any set of basis vectors $|n\rangle$

$$\sum_n a_n |n\rangle = 0 \quad \text{only if all } a_n = 0$$

Proof: Suppose $a_n \neq 0$

Define $|v\rangle = \sum_n a_n |n\rangle = 0$

From the above $a_n = \langle n|v\rangle$ but $\langle v|n\rangle = 0$ if $|v\rangle = 0$

So $a_n = 0$.

This is a contradiction of our assumption $a_n \neq 0$

So all $a_n = 0$. $\square$

1.9 Building a basis from linearly independent vectors

Suppose $\{|v_k\rangle\}$ are linearly indep. and the space is N-dimensional. Define $|1\rangle, |2\rangle, \ldots$ as

$$
\begin{aligned}
|1\rangle &= \frac{|v_1\rangle}{\| |v_1\rangle \|} \\
|2\rangle &= \frac{|v_2\rangle - |1\rangle \langle 1|v_2\rangle}{\| |v_2\rangle - |1\rangle \langle 1|v_2\rangle \|} \\
|3\rangle &= \frac{|v_3\rangle - |1\rangle \langle 1|v_3\rangle - |2\rangle \langle 2|v_3\rangle}{\| |v_3\rangle - |1\rangle \langle 1|v_3\rangle - |2\rangle \langle 2|v_3\rangle \|} \\
|4\rangle &= \ldots
\end{aligned}
$$

Clearly $\langle 1|1\rangle = 1$, $\langle 2|2\rangle = 1$, ...

Also:

$$\langle 1|2\rangle = \frac{\langle 1|v_2\rangle - \langle 1|1\rangle \langle 1|v_2\rangle}{\| |v_2\rangle - |1\rangle \langle 1|v_2\rangle \|} = \frac{\langle 1|v_2\rangle - \langle 1|v_2\rangle}{\| |v_2\rangle - |1\rangle \langle 1|v_2\rangle \|} = 0$$

and $\langle 1|3\rangle = 0$, $\langle 2|3\rangle = 0$, etc.

This is Gram-Schmidt orthogonalization and it will be useful later.

1.9.1 Examples:

For $|v\rangle \in \mathbb{C}^2$, one basis is

$$|1\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad |2\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

This is often called the *standard*, *natural* or *canonical* basis.

But also we could have

$$|1'\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad |2'\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

$$
\begin{aligned}
\langle 1'|1'\rangle &= \frac{1}{\sqrt{2}} \frac{1}{\sqrt{2}} (1^2 + 1^2) = 1 \\
\langle 2'|2'\rangle &= \frac{1}{\sqrt{2}} \frac{1}{\sqrt{2}} (1^2 + (-1)^2) = 1 \\
\langle 1'|2'\rangle &= \frac{1}{\sqrt{2}} \frac{1}{\sqrt{2}} (1^2 + 1 \cdot (-1)) = 0
\end{aligned}
$$

and for $|v\rangle = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$

$$|v\rangle = \left(\frac{1}{\sqrt{2}} (v_1 + v_2) \right) |1'\rangle + \left(\frac{1}{\sqrt{2}} (v_1 - v_2) \right) |2'\rangle$$

So any $|v\rangle = |1'\rangle c_1' + |2'\rangle c_2'$

In the homework you will find a basis for periodic functions $f(x) = f(x + 2\pi)$ where you'll find the basis is infinite!

1.10 Subspaces

A subspace U of a vector space V is a subset of vectors that forms a vector space. This means U is closed under addition and scalar multiplication in its own right.

The dimension of a subspace is defined as the maximum number of linearly independent vectors in the subspace, exactly the same as vector space dimensionality. The trivial, zero dimensional subspace is the subset of V just containing the zero vector 0_V.

Example 1: For $|v\rangle \in \mathbb{C}^2$ the set of vectors U such that

$$|u\rangle = \begin{pmatrix} z \\ 0 \end{pmatrix} \in U, \quad z \in \mathbb{C}$$

is closed under vector addition and scalar multiplication, so this is a subspace of $\mathbb{C}$. Its dimension is 1.

Example 2: The set of even complex valued functions of $x \in \mathbb{R}$, such that

$$f(x) = f(-x)$$

is closed under addition and scalar multiplication, and is a subset of all possible complex valued functions. So this is a subspace. Likewise the set of odd functions

$$f(x) = -f(x)$$

also forms a subspace.

1.11 Hilbert spaces

We've already basically introduced what a Hilbert space, $\mathcal{H}$, is without saying it explicitly.

A **Hilbert space** is a **vector space**, with an **inner product**, where an **orthonormal basis** can always we defined (even for infinite dimensional spaces). Hilbert spaces are said to be *complete*.

In Quantum Mechanics, Hilbert spaces play an essential role, because all quantum states are elements of a Hilbert space (more on that soon), and the evolution and properties of these states are governed by *linear operators* (more on that next time).

2 Linear Operators

2.1 Linear Operator Definition

We have introduced $\mathbb{C}$ vectors & inner products already. Now let's consider a special class of functions on vectors: linear operators.

A linear operator, $\hat{A}$, (denoted by the hat $\hat{\ }$) maps vectors to other vectors. It is "linear" if it satisfies

1. $\hat{A}(a\,|v\rangle) = a(\hat{A}\,|v\rangle)$ for $a \in \mathbb{C}$
2. $\hat{A}(|v\rangle + |w\rangle) = (\hat{A}\,|v\rangle) + (\hat{A}\,|w\rangle)$

This means $\hat{A} + \hat{B}$ is well defined: $(\hat{A} + \hat{B})\,|v\rangle = \hat{A}\,|v\rangle + \hat{B}\,|v\rangle$

And also $a\hat{A}$ is well defined: $(a\hat{A})\,|v\rangle = a(\hat{A}\,|v\rangle)$.

Note: We will sometimes write $|\hat{A}v\rangle = \hat{A}\,|v\rangle$.

This means that we can take linear combinations of operators e.g. $a\hat{A} + b\hat{B}$ and act them on linear combinations of vectors $c\,|v\rangle + d\,|u\rangle$, and expand and factor these expressions using our familiar rules of algebra. We will need to be careful about operator *multiplication* as we will see very shortly.

2.1.1 Examples

Example 1: For $|v\rangle \in \mathbb{C}^2$ & multiplying the first component by $s \in \mathbb{C}$ is linear

$$\hat{A}\,|v\rangle \stackrel{\text{def}}{=} \begin{pmatrix} sv_1 \\ v_2 \end{pmatrix}$$

$$s\begin{pmatrix} v_1 + v_2 \\ v_2 + u_2 \end{pmatrix} = \begin{pmatrix} sv_1 \\ v_2 \end{pmatrix} + \begin{pmatrix} su_1 \\ u_2 \end{pmatrix}$$

so $\hat{A}(|u\rangle + |v\rangle) = \hat{A}\,|u\rangle + \hat{A}\,|v\rangle$ ✓

and

$$\begin{pmatrix} sau_1 \\ au_2 \end{pmatrix} = a\begin{pmatrix} su_1 \\ u_2 \end{pmatrix}$$

so $\hat{A}(a\,|u\rangle) = a\hat{A}\,|u\rangle$ ✓

Note: In this example we can write

$$\hat{A}\,|v\rangle = \begin{pmatrix} s & 0 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$$

So we can write the linear operator acting on the vector as a matrix-vector multiplication. For finite dimensional vector spaces we can always do this, and there is a one-to-one correspondence between linear operators and matrices. The square matrix above is equivalent to the *matrix representation* of $\hat{A}$ in the canonical basis. We'll explain this fully below.

Example 2: For 1D functions

$$\hat{x}\psi(x) \stackrel{\text{def}}{=} x\psi(x)$$

$$\text{and} \quad \hat{D}\psi(x) \stackrel{\text{def}}{=} \frac{d}{dx}\psi(x)$$

are both linear

e.g.

$$\begin{aligned}\hat{D}\left(a\psi(x)+b\phi(x)\right) &= \frac{d}{dx}\left(a\psi(x)+b\phi(x)\right)\\ &= a\frac{d}{dx}\psi(x)+b\frac{d}{dx}\phi(x)\\ &= a\hat{D}\psi(x)+b\hat{D}\phi(x) \quad \checkmark\end{aligned}$$

2.2 Outer Products

The outer product (also called a **dyad**) of two vectors $|u\rangle$ and $|u\rangle$ is defined as

$$|v\rangle\langle u|$$

where we recall that $\langle u| = |u\rangle^\dagger$ is the *dual* of the vector $|u\rangle$. When this acts to a vector to the right, we get another vector

$$|v\rangle\langle u|\left(|w\rangle\right) = |v\rangle \overbrace{\langle u|w\rangle}^{\text{a complex number}} = \langle u|w\rangle|v\rangle$$

And because the inner product $\langle u|w\rangle$ is linear in its second argument i.e. $\langle u|a_1w_1+a_2w_2\rangle = a_1\langle u|w_1\rangle + a_2\langle u|w_2\rangle$, the dyad $|v\rangle\langle u|$ is a linear operator on vectors

$$|v\rangle\langle u|\left(|a_1w_1+a_2w_2\rangle\right) = a_1|v\rangle\langle u|w_1\rangle + a_2|v\rangle\langle u|w_2\rangle = a_1|v\rangle\langle u|\left(|w_1\rangle\right) + a_2|v\rangle\langle u|\left(|w_2\rangle\right)$$

We can also define the **projector** onto $|v\rangle$ as the special dyad

$$\hat{P}_v \stackrel{\text{def}}{=} |v\rangle\langle v|$$

2.3 The Identity Operator

The identity operator $\hat{1}$ is the linear operator that turns $|v\rangle$ into itself, i.e., it does nothing to $|v\rangle$

$$\hat{1}|v\rangle = |v\rangle$$

Before we saw that any $|v\rangle$ can be written in terms of a basis $\{|n\rangle\}$ as

$$|v\rangle = \sum_n |n\rangle c_n$$

where $c_n = \langle n|v\rangle$.

Therefore we can write

$$\begin{aligned}|v\rangle &= \sum_n |n\rangle\langle n|v\rangle\\ &= \sum_n \left(|n\rangle\langle n|\right)|v\rangle\end{aligned}$$

The projector $|n\rangle\langle n|$ turns a vector $|v\rangle$ into $|n\rangle\langle n|v\rangle = \langle n|v\rangle|n\rangle$ so this is an operator.

We also see that

$$|v\rangle = \left(\sum_n |n\rangle\langle n|\right)|v\rangle$$

so this means we can write the identity operator as

$$\hat{1} = \sum_n |n\rangle \langle n|$$

This is sometimes called a *resolution of the identity* because we are resolving the identity operator into a sum over a particualr set of vectors. Note we can choose *any* set of basis vectors $|n\rangle$ for our resolution of the identity.

2.4 Operator Equivalence

Two operators $\hat{A}$ and $\hat{B}$ are equivalent, i.e. $\hat{A} = \hat{B}$, if for all $|v\rangle$

$$\hat{A}|v\rangle = \hat{B}|v\rangle \quad \text{if and only if} \quad \hat{A} = \hat{B}$$

If we write out these two vectors in a basis we find

$$\hat{A}|v\rangle = \sum_n \hat{A}|n\rangle v_n$$

$$\hat{B}|v\rangle = \sum_n \hat{B}|n\rangle v_n$$

where $\langle n|v\rangle = v_n$. Two vectors are equivalent if and only if all there components in the particular basis are equivalent, so $\hat{A} = \hat{B}$ implies

$$\langle m|\hat{A}|v\rangle = \sum_n \langle m|\hat{A}|n\rangle v_n$$

$$\langle m|\hat{B}|v\rangle = \sum_n \langle m|\hat{B}|n\rangle v_n$$

This is true for all v_n, so $\hat{A} = \hat{B}$ implies

$$\langle m|\hat{A}|n\rangle = \langle m|\hat{B}|n\rangle$$

$\langle m|\hat{A}|n\rangle$ is called a *matrix element* of $\hat{A}$. Also if the matrix elements are equivalent then for any $|v\rangle$

$$\begin{aligned}
\hat{A}|v\rangle &= \sum_n \hat{1}\hat{A}|n\rangle v_n \\
&= \sum_{n,m} |m\rangle\langle m|\hat{A}|n\rangle v_n \\
&= \sum_{n,m} |m\rangle\langle m|\hat{B}|n\rangle v_n \quad \text{if } \langle m|\hat{A}|n\rangle = \langle m|\hat{B}|n\rangle \\
&= \sum_n \hat{1}\hat{B}|n\rangle v_n \\
&= \hat{B}|v\rangle
\end{aligned}$$

So equivalence of all matrix elements $\langle m|\hat{A}|n\rangle = \langle m|\hat{B}|n\rangle$ means $\hat{A} = \hat{B}$. So there is a one-to-one correspondence between *matrix elements in a particular basis* and the operators.

2.5 Operator Composition

We adopt the convention where operators to the right act first, i.e.

$$\hat{A}\hat{B}|v\rangle = \hat{A}\left(\hat{B}|v\rangle\right)$$

In general the order that linear operators are applied in matters!

$$\hat{A}\hat{B}\left|v\right\rangle \neq \hat{B}\hat{A}\left|v\right\rangle$$

If the order doesn't matter then we say operators **commute**, so $\hat{A}\hat{B} = \hat{B}\hat{A}$.

Let's consider the pair of operators $\hat{x}$ and $\hat{D}$ above. Do these commute? First let's find $\hat{x}\hat{D}$

$$\hat{x}\hat{D}\psi(x) = x\frac{d}{dx}\psi(x)$$

and now let's find $\hat{D}\hat{x}$

$$\begin{aligned}\hat{D}\hat{x}\psi(x) &= \hat{D}\left(x\psi(x)\right)\\ &= \psi(x) + x\frac{d}{dx}\psi(x)\\ &\neq \hat{x}\hat{D}\psi(x)\end{aligned}$$

So we see $\hat{x}$ and $\hat{D}$ don't commute.

We define the **commutator** as

$$[\hat{A}, \hat{B}] = \hat{A}\hat{B} - \hat{B}\hat{A}$$

e.g. for $\hat{A} = \hat{x}$ and $\hat{B} = \hat{D} = \frac{d}{dx}$

$$\begin{aligned}[\hat{x}, \hat{D}]\psi(x) &= x\frac{d}{dx}\psi(x) - \psi(x) - x\frac{d}{dx}\psi(x)\\ &= -\psi(x)\\ \Rightarrow [\hat{x}, \hat{D}] &= -\hat{1}\end{aligned}$$

2.6 Commutator Properties

Commutators have the following useful properties:

1. $[\hat{A}, \hat{B}] = -[\hat{B}, \hat{A}]$
2. $[\hat{A}, \hat{B} + \hat{C}] = [\hat{A}, \hat{B}] + [\hat{A}, \hat{C}]$
3. $[\hat{A}, \hat{B}\hat{C}] = [\hat{A}, \hat{B}]\hat{C} + \hat{B}[\hat{A}, \hat{C}]$
4. $[\hat{A}, b\hat{1}] = 0$

As an example, let's prove 1):

$$\begin{aligned}[\hat{A}, \hat{B}] &= \hat{A}\hat{B} - \hat{B}\hat{A}\\ &= -(\hat{B}\hat{A} - \hat{A}\hat{B})\\ &= -[\hat{B}, \hat{A}] \quad \square\end{aligned}$$

and 2):

Consider $[\hat{A}, \hat{B} + \hat{C}]$ acting on $\left|v\right\rangle$

$$\begin{aligned}[\hat{A}, \hat{B} + \hat{C}]\left|v\right\rangle &= \hat{A}(\hat{B} + \hat{C})\left|v\right\rangle - (\hat{B} + \hat{C})\hat{A}\left|v\right\rangle\\ &= \hat{A}\hat{B}\left|v\right\rangle + \hat{A}\hat{C}\left|v\right\rangle - \hat{B}\hat{A}\left|v\right\rangle - \hat{C}\hat{A}\left|v\right\rangle\\ &= \hat{A}\hat{B}\left|v\right\rangle - \hat{B}\hat{A}\left|v\right\rangle + \hat{A}\hat{C}\left|v\right\rangle - \hat{C}\hat{A}\left|v\right\rangle\\ &= [\hat{A}, \hat{B}]\left|v\right\rangle + [\hat{A}, \hat{C}]\left|v\right\rangle\end{aligned}$$

$\Rightarrow [\hat{A}, \hat{B} + \hat{C}] = [\hat{A}, \hat{B}] + [\hat{A}, \hat{C}] \quad \square$

2.7 Matrix Representation of Linear Operators

If a vector space has finite dimensionality, i.e., a finite basis, we can construct a matrix representation of an operator $\hat{A}$ in a given basis.

First, let us represent the vector $|v\rangle$ as a column vector $\underline{v}$ in basis $\{|n\rangle\}$.

If $|v\rangle = \sum_n |n\rangle\, v_n$ the column vector $\underline{v}$ is

$$\underline{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_N \end{pmatrix}$$

This connection between finite dimensional vectors and column vectors is one-to-one for a given choice of basis.

Now $\hat{A}\,|v\rangle =?$ Let's use the identity operator $\hat{1}$!

$$\begin{aligned}
\hat{A}\,|v\rangle &= \hat{A}\hat{1}\,|v\rangle \\
&= \hat{1}\hat{A}\hat{1}\,|v\rangle \\
&= \sum_{n=1}^{N} |n\rangle\,\langle n|\,\hat{A}\sum_{m=1}^{N} |m\rangle\,\langle m|v\rangle \\
&= \sum_{n=1}^{N} |n\rangle \sum_{m=1}^{N} \langle n|\hat{A}|m\rangle\, v_m \\
&= \sum_{n=1}^{N} |n\rangle \sum_{m=1}^{N} A_{n,m} v_m
\end{aligned}$$

So the column vector for $\hat{A}\,|v\rangle = |w\rangle$ in the $\{|n\rangle\}$ basis is:

$$\underline{w} = \begin{pmatrix} \sum_{m=1}^{N} A_{1,m} v_m \\ \sum_{m=1}^{N} A_{2,m} v_m \\ \vdots \\ \sum_{m=1}^{N} A_{N,m} v_m \end{pmatrix}$$

which we can write as

$$\underline{w} = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1N} \\ A_{21} & A_{22} & \cdots & A_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ A_{N1} & A_{N2} & \cdots & A_{N,N} \end{pmatrix} \cdot \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_N \end{pmatrix}$$

where matrix-vector multiplication is defined in the usual way as

$$\underline{w} = \underline{\underline{A}} \cdot \underline{v}$$

$$w_n = \sum_{m=1}^{N} A_{nm} v_m$$

The matrix elements of $\underline{\underline{A}}$ are

$$[\underline{\underline{A}}]_{n,m} = A_{n,m} = \langle n|\hat{A}|m\rangle$$

So overall for a finite dimensional space, $\hat{A}$ is fully defined by its action on basis vectors, in a given basis. This means there is a one-to-one correspondence between the matrix representation and the operator.

We can also define matrix multiplication as

$$\underline{\underline{A}} \cdot \underline{\underline{B}} = \underline{\underline{C}} \quad \text{if } \underline{\underline{A}} \text{ is } M \times N \text{ and } \underline{\underline{B}} \text{ is } N \times M'$$

$$C_{nm} = \sum_{p=1}^{N} A_{np} B_{pm}$$

It is fairly staightforward to show that if $\hat{C} = \hat{A}\hat{B}$ then the matrix representation of $\hat{C}$ in a given basis is

$$\underline{\underline{C}} = \underline{\underline{A}} \cdot \underline{\underline{B}}$$

and likewise for $\hat{D} = a\hat{A} + b\hat{B}$

$$\underline{\underline{D}} = a\underline{\underline{A}} + b\underline{\underline{B}}.$$

One important operator is the identity operator $\hat{1}$, its matrix elements in *any* basis are

$$\langle n|\hat{1}|m\rangle = \langle n|m\rangle = \delta_{n,m}$$

so the matrix representation is just the identity matrix, which is zero everywhere apart from the diagonal elements which are all 1

$$\underline{\underline{1}} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

Similarly the zero operator 0 has all zero matrix elements in any basis, so it its matrix representation is a matrix of zeros $\underline{\underline{0}}$.

For finite-dimensional complex vectors $|v\rangle = \underline{v} \in \mathbb{C}^N$, we can always write a linear operator $\hat{A}$ as

$$\hat{A}\,|v\rangle = \underline{\underline{A}} \cdot \underline{v}$$

where $\underline{\underline{A}}$ is the matrix representation of $\hat{A}$ in the canonical basis $|n\rangle = \underline{e}_n$ where the mth element of $\underline{e}_n$ is $[\underline{e}_n]_m = \delta_{nm}$.

Example 1: For example if we go back to the scaling example for $|v\rangle = \underline{v} \in \mathbb{C}^2$

$$\hat{A}\,|v\rangle = \begin{pmatrix} sv_1 \\ v_2 \end{pmatrix}$$

Pick a basis $\left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$ (this is the natural basis of $\mathbb{C}^2$) and the matrix representation of $\hat{A}$ in this basis is

$$\begin{aligned} A_{11} &= \langle 1|\hat{A}|1\rangle = \langle 1|1\rangle\, s = s \\ A_{12} &= \langle 1|\hat{A}|2\rangle = \langle 1|2\rangle = 0 \\ A_{21} &= \langle 2|\hat{A}|1\rangle = \langle 2|1\rangle\, s = 0 \\ A_{22} &= \langle 2|\hat{A}|2\rangle = \langle 2|2\rangle = 1 \end{aligned}$$

So:

$$\underline{\underline{A}} = \begin{pmatrix} s & 0 \\ 0 & 1 \end{pmatrix}$$

and we can also write

$$\hat{A}\,|v\rangle = \underline{\underline{A}} \cdot \underline{v} \tag{1}$$

2.8 Inverses

If $\hat{A}$ is invertible we can find another operator $\hat{A}^{-1}$ such that

$$\hat{A}^{-1}\hat{A} = \hat{A}\hat{A}^{-1} = \hat{1}$$

If $\hat{A}$ is not invertible, this means

$$\hat{A}\left|v\right\rangle = 0_V$$

for some $\left|v\right\rangle \neq 0_V$. (If it was invertible we could apply $\hat{A}^{-1}$ to the left and we'd find that $\left|v\right\rangle = 0_V$, in contraction of our assumption.)

Example 1

$$\hat{A}\left|v\right\rangle = \begin{pmatrix} s & 0 \\ 0 & 1 \end{pmatrix} \cdot \underline{v} \text{ on } \mathbb{C}^2 \text{ with } s \neq 0$$

$$\hat{A}^{-1}\left|u\right\rangle = \begin{pmatrix} \frac{1}{s} & 0 \\ 0 & 1 \end{pmatrix} \cdot \underline{v}$$

but if $s = 0$

$$\begin{pmatrix} s & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 0 \\ v_2 \end{pmatrix} \text{ but also } \begin{pmatrix} s & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} v_1' \\ v_2 \end{pmatrix} = \begin{pmatrix} 0 \\ v_2 \end{pmatrix}$$

$\begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$ and $\begin{pmatrix} v_1' \\ v_2 \end{pmatrix}$ both map to $\begin{pmatrix} 0 \\ v_2 \end{pmatrix}$ so there is no linear operator that can restore the original vector.

Not all linear operators are invertible. For example if $s = 0$ above.

Example 2 For square normalizable functions i.e. $\psi(x) \to 0$ as $|x| \to \infty$ if $\hat{D} = \frac{d}{dx}$

$$\hat{D}^{-1}\psi(x) = \int_{-\infty}^{x} \psi(x')dx'.$$

2.9 Adjoint Operators

The **adjoint** of $\hat{A}$ is denoted $\hat{A}^\dagger$. $\hat{A}^\dagger$ is the operator that satisfies

$$\langle v|\hat{A}|u\rangle = \left\langle \hat{A}^\dagger v \middle| u \right\rangle \quad \forall \left|u\right\rangle, \left|v\right\rangle$$

From the inner product properties we can also write this as

$$\langle v|\hat{A}|u\rangle = \left\langle u \middle| \hat{A}^\dagger v \right\rangle^*$$

Note we used $\dagger$ to denote the dual of a vector. We can define the dual of $\hat{A}\left|v\right\rangle$ to be

$$\left(\hat{A}\left|v\right\rangle\right)^\dagger = \left(\left|\hat{A}v\right\rangle\right)^\dagger = \left\langle v\right|\hat{A}^\dagger$$

Self-adjoint (also called **Hermitian**) operators are equal to their adjoint

$$\hat{A} = \hat{A}^\dagger$$

So:

$$\begin{aligned}\langle v|\hat{A}|u\rangle &= \left\langle \hat{A}v \middle| u \right\rangle \\ &= \left\langle u \middle| \hat{A}v \right\rangle^*\end{aligned}$$

Example: For $\hat{A}\,|v\rangle = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$ on $\mathbb{C}^2$, $\hat{A} = \hat{A}^\dagger$

From the above if $\hat{A}$ is Hermitian

$$\begin{aligned}\left\langle u \middle| \hat{A}v \right\rangle &= \langle v|\hat{A}|u\rangle^* \\ \langle u|\hat{A}|v\rangle &= (u_1^* u_2^*) \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = -iu_1^* v_2 + iu_2^* v_1\end{aligned}$$

$$\begin{aligned}\langle v|\hat{A}|u\rangle &= (v_1^* v_2^*) \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \\ &= +iu_1 v_2^* - iu_2 v_1^* \\ &= (-iu_1^* v_2 + iu_2^* v_1)^* \quad \checkmark\end{aligned}$$

2.9.1 Adjoint of products

The adjoint of the product $\hat{A}\hat{B}$ can be found as

$$\begin{aligned}\langle u|\hat{A}\hat{B}|v\rangle &= \langle \hat{A}^\dagger u|\hat{B}|v\rangle \quad \text{from definition of } \hat{A}^\dagger \\ &= \langle \hat{B}^\dagger \hat{A}^\dagger u||v\rangle \quad \text{from definition of } \hat{B}^\dagger\end{aligned}$$

$$\text{and} \quad \langle u|\hat{A}\hat{B}|v\rangle = \left\langle (\hat{A}\hat{B})^\dagger u \middle| v \right\rangle \quad \text{from definition of } (\hat{A}\hat{B})^\dagger$$

so we see that

$$(\hat{A}\hat{B})^\dagger = \hat{B}^\dagger \hat{A}^\dagger$$

2.9.2 Adjoint of products

Given a basis $\{|n\rangle\}$ we can write any $\hat{A}$ as

$$\hat{A} = \hat{1}\hat{A}\hat{1} = \sum_{nm} |n\rangle\ \langle n|\hat{A}|m\rangle \langle m|$$

and

$$\begin{aligned}\hat{A}^\dagger = \hat{1}\hat{A}^\dagger\hat{1} &= \sum_{nm} |n\rangle\ \langle n|\hat{A}^\dagger|m\rangle \langle m| \\ &= \sum_{nm} |n\rangle\ \langle m|\hat{A}|n\rangle^* \langle m|\end{aligned}$$

Therefore $\hat{A}$ is Hermitian if and only if its matrix elements satisfy

$$\langle n|\hat{A}|m\rangle = \langle m|\hat{A}|n\rangle^*$$

for all basis vectors $|n\rangle$ and $|m\rangle$. This means the matrix representation of $\hat{A}$ will always be **conjugate symmetric**

$$\underline{\underline{A}} = (\underline{\underline{A}}^T)^* = \underline{\underline{A}}^\dagger$$

where $\underline{\underline{A}}^T$ denotes the transpose of $\underline{\underline{A}}$ and $\underline{\underline{A}}^*$ denotes the complex conjugate of $\underline{\underline{A}}$, and $\underline{\underline{A}}^\dagger \stackrel{\text{def}}{=} (\underline{\underline{A}}^T)^*$.

Hermitian operators are very special in QM. But to understand why we'll need to define an eigenvector and eigenvalue.

2.10 Eigenvectors and eigenvalues

$|\lambda\rangle$ is an eigenvector of $\hat{A}$ with eigenvalue $\lambda \in \mathbb{C}$ if

$$\hat{A}\,|\lambda\rangle = |\lambda\rangle\,\lambda$$

and $|\lambda\rangle \neq 0$.

Hermitian operators always have **real-valued** eigenvalues.

Proof:

We know that $|\lambda\rangle \neq 0$

If $\hat{A}\,|\lambda\rangle = |\lambda\rangle\,\lambda$ then

$$\langle\lambda|\,\hat{A}\,|\lambda\rangle = \langle\lambda|\lambda\rangle\,\lambda$$

but also

$$\begin{aligned}
\langle\lambda|\,\hat{A}\,|\lambda\rangle &= \langle\lambda|\,\hat{A}^\dagger\,|\lambda\rangle^* \\
&= \langle\lambda|\,\hat{A}\,|\lambda\rangle^* \quad \because \hat{A} = \hat{A}^\dagger \\
&= \langle\lambda|\,\lambda\,|\lambda\rangle^* \\
&= \langle\lambda|\lambda\rangle\,\lambda^* \quad \because \langle\lambda|\lambda\rangle \in \mathbb{R}
\end{aligned}$$

$$\begin{aligned}
\Rightarrow \quad & \langle\lambda|\lambda\rangle\,\lambda = \langle\lambda|\lambda\rangle\,\lambda^* \\
\Rightarrow \quad & \lambda = \lambda^*
\end{aligned}$$

so $\lambda \in \mathbb{R}$ $\square$

For two eigenvectors $|\lambda_1\rangle$ and $|\lambda_2\rangle$, if $\lambda_1 \neq \lambda_2$ then

$$\langle\lambda_1|\lambda_2\rangle = 0$$

Proof:

$$\begin{aligned}
\langle\lambda_1|\,\hat{A}\,|\lambda_2\rangle &= \langle\lambda_1|\lambda_2\rangle\,\lambda_2 \\
&= \langle\lambda_2|\,\hat{A}\,|\lambda_1\rangle^* \quad \text{by hermiticity: } \hat{A} = \hat{A}^\dagger \\
&= \langle\lambda_2|\lambda_1\rangle^*\,\lambda_1^* \\
&= \langle\lambda_1|\lambda_2\rangle\,\lambda_1 \quad \because \lambda_1 = \lambda_1^*
\end{aligned}$$

So $\langle\lambda_1|\lambda_2\rangle\,(\lambda_1 - \lambda_2) = 0$ but $\lambda_1 \neq \lambda_2$

so $\langle\lambda_1|\lambda_2\rangle = 0$ $\square$

Note that if $|\lambda\rangle$ is an eigenvector of $\hat{A}$ then so is $c\,|\lambda\rangle$ for $c \neq 0$ so we can always normalize $|\lambda\rangle$ to give

$$|n\rangle = \frac{1}{\sqrt{\langle\lambda_n|\lambda_n\rangle}}\,|\lambda_n\rangle$$

and if $|\lambda_1\rangle$ and $|\lambda_2\rangle$ have the same eigenvalue, i.e. $\lambda_1 = \lambda_2$ and are linearly independent we can always use Gram-Schmidt orthogonalization to construct an orthonormal set of vectors with the same eigenvalue. The maximum number of linealry independent eigenvectors with the same eigenvalue is called the **degeneracy** of the eigenvalue.

This means that the orthonormal eigenvectors of $\hat{A}$ $(= \hat{A}^\dagger)$ form a basis i.e.

$$\hat{A}\,|\lambda_n\rangle = |\lambda_n\rangle\,\lambda_n$$

$$\hat{1} = \sum_n |\lambda_n\rangle\,\langle\lambda_n|$$

Note that we will encounter examples later where the eigenspectrum of $\hat{A}$ is continuous so

$$\hat{1} = \int_{-\infty}^{+\infty} d\lambda\,|\lambda\rangle\,\langle\lambda|$$

A very important consequence of this is that $\hat{A}$ can be resolved in terms of its eigenvectors and eigenvalues.

$$\begin{aligned}\hat{A} &= \hat{A}\hat{1}\\ &= \hat{A}\sum_n |\lambda_n\rangle\langle\lambda_n|\\ &= \sum_n \hat{A}\,|\lambda_n\rangle\langle\lambda_n|\\ &= \sum_n \lambda_n\,|\lambda_n\rangle\langle\lambda_n|\end{aligned}$$

This is sometimes called a *Spectral Decomposition*. This proves to be very useful in many contexts.

It's also fairly straightforward to see that

$$\hat{A}^k\,|\lambda_n\rangle = \lambda_n^k\,|\lambda_n\rangle$$

so following the same argument as above (you can fill in the steps I skip here yourself)

$$\begin{aligned}\hat{A}^k &= \hat{A}^k\sum_n |\lambda_n\rangle\langle\lambda_n|\\ &= \sum_n \lambda_n^k\,|\lambda_n\rangle\langle\lambda_n|\end{aligned}$$

2.11 Unitary operators

A unitary operator has the property that

$$\hat{U}^\dagger = \hat{U}^{-1}$$

In the problems you will show that eigenvalues of a unitary operator can all be written as

$$\lambda = e^{i\phi} \quad 0 \leq \phi < 2\pi$$

In QM Hermitian operators are connected to *measurements of observables* and unitary operators to *transformations*. They are very closely related (as we will see soon) and we can also always resolve the identity operator in terms of a unitary operator's eigenvectors, exactly as we can for a Hermitian operator.

2.11.1 Change of basis

Let's define an operator $\hat{S}$ that transforms a vector in one basis $|n\rangle$ to another basis $|n'\rangle$ preserving all the coefficients i.e. for any vector $|v\rangle = \sum_n c_n |n\rangle$

$$\hat{S}|v\rangle = \hat{S}\left(\sum_n c_n |n\rangle\right) = \sum_n c_n |n'\rangle$$

This operation is clearly linear. From this we know how $\hat{S}$ acts on basis vectors $|n\rangle$

$$\hat{S}|n\rangle = |n'\rangle$$

We can use a resolution of the identity to show

$$\hat{S} = \hat{S}\hat{1} = \sum_n \hat{S}|n\rangle\langle n| = \sum_n |n'\rangle\langle n|$$

From this and the fact that $\langle n'|m'\rangle = \delta_{n,m}$ and $\langle n|m\rangle = \delta_{n,m}$, we can show that

$$\hat{S}^\dagger\hat{S} = \hat{S}\hat{S}^\dagger = \hat{1}$$

So unitary operators represent a *change of basis*. We see therefore that unitary generalise the idea of a rotation to arbitrary dimensional complex vectors.

2.12 Functions of operators

For a function $f(x)$ with a well-defined Taylor series, we can define $f(\hat{A})$, by just replacing $x \to \hat{A}$ in the Taylor series

$$f(\hat{A}) = f(0) + f'(0)\hat{A} + \frac{1}{2!}f''(0)\hat{A}^2 + \frac{1}{2!}f'''(0)\hat{A}^3 + \cdots$$

A very commonly encountered example is the exponential function

$$\exp\left(\hat{A}\right) \equiv e^{\hat{A}} = \sum_{k=0}^{\infty} \frac{1}{k!}\hat{A}^k. \tag{2}$$

Now we have established most of the foundations of Linear Algebra that we'll need to understand quantum mechanics.

If $\hat{A}$ has a complete eigenbasis, i.e. if we can resolve the identity in terms of its eigenvectors $\hat{1} = \sum_n |\lambda_n\rangle\langle\lambda_n|$, we can use the Taylor expansion and $\hat{A}^k|\lambda_n\rangle = \lambda_n^k|\lambda_n\rangle$ to show that

$$\begin{aligned} f(\hat{A}) &= f(\hat{A})\hat{1} \\ &= \sum_n f(\lambda_n)|\lambda_n\rangle\langle\lambda_n| \end{aligned}$$

For example if $\hat{A}$ is hermitian or unitary we can do this.

2.13 Matrix representations of Inverses

In the finite dimensional case, if the matrix inverse exists, which corresponds to the inverse of a linear operator $\hat{A}$, it is given by

$$\underline{\underline{A}}^{-1} = \frac{1}{\det(\underline{\underline{A}})}\mathrm{adj}(\underline{\underline{A}})$$

$\underline{\underline{A}}^{-1}$ is not defined if and only if the determinant $\det(\underline{\underline{A}}) = 0$. $\mathrm{adj}(\underline{\underline{A}})$ is called the **adjugate matrix**. You are advised to consult your favourite reference on matrices for the full definition of this, because its quite cumbersome to write out.

For two dimensional matrices we have

$$\underline{\underline{A}} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

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

2.14 Matrix representations of Hermitian Operators

We previously showed there is an equivalence between matrix-vector multiplication and linear operators in finite dimensional complex vector spaces. What properties do the matrix representations of Hermitian operators have?

We know Hermitian operators satisfy

$$\langle n|\hat{A}|m\rangle = \langle m|\hat{A}|n\rangle^*$$

for all basis vectors $|n\rangle, |m\rangle$. We also recall that the matrix representation of $\hat{A}$ in a basis is

$$\underline{\underline{A}} = \begin{pmatrix} A_{1,1} & A_{1,2} & \cdots \\ A_{2,1} & A_{2,2} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}$$

with $A_{n,m} = [\underline{\underline{A}}]_{n,m} = \langle n|\hat{A}|m\rangle$ and the matrix representation *fully* defines a linear operator in finite dimensions. The Hermitian property above means that the matrix elements of Hermitian operators satsify

$$A_{n,m} = A^*_{m,n}$$

We also recall that the **transpose** of a matrix $\underline{\underline{A}}^T$ is

$$[\underline{\underline{A}}^T]_{n,m} = [\underline{\underline{A}}]_{m,n}$$

so the transpose is mirrored along the diagonal

$$\underline{\underline{A}}^T = \begin{pmatrix} A_{1,1} & A_{2,1} & \cdots \\ A_{1,2} & A_{2,2} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}$$

Likewise the complex conjugate of a matrix $\underline{\underline{A}}^*$ is just defined by taking the complex conjugate of all its matrix elements

$$[\underline{\underline{A}}^*]_{n,m} = [\underline{\underline{A}}]^*_{n,m}$$

so

$$\underline{\underline{A}}^* = \begin{pmatrix} A^*_{1,1} & A^*_{1,2} & \cdots \\ A^*_{2,1} & A^*_{2,2} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}$$

Returning the hermitian property above, $A_{n,m} = A^*_{m,n}$ means the matrix representation satisfies

$$\underline{\underline{A}} = (\underline{\underline{A}}^T)^* = (\underline{\underline{A}}^*)^T$$

where $(\underline{\underline{A}}^T)^*$ is the **conjugate-transpose** of the matrix $\underline{\underline{A}}$. This is also called the **hermitian conjugate** and we use the same † symbol as the hermitian conjugate of an operator to denote this

$$\underline{\underline{A}}^\dagger \stackrel{\text{def}}{=} (\underline{\underline{A}}^T)^*.$$

Because of the one-to-one mapping between operators and matrices in finite dimensional spaces, an operator is Hermitian/self-adjoint if and only if its matrix representation is Hermitian. So if we know the matrix representation of an operator in finite dimensions, then we can immediately tell if it's Hermitian by inspecting if its matrix elements. Note that the diagonal elements of the a Hermitian matrix satisfy

$$A_{n,n} = A^*_{n,n}$$

so these have to be real-valued.

Example: For $\mathbb{C}^2$, and operator $\hat{A}$ with a matrix representation $\underline{\underline{A}}$ in the canonical basis

$$\underline{\underline{A}} = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$$

in the canonical basis. $\underline{\underline{A}} = \underline{\underline{A}}^\dagger$ so $\hat{A} = \hat{A}^\dagger$.

But for $\hat{B}$ with matrix representation

$$\underline{\underline{B}} = \begin{pmatrix} i & 1 \\ 1 & 0 \end{pmatrix}$$

$B_{1,2} = B^*_{2,1}$, $B_{2,2} = B^*_{2,2}$ but $B^*_{1,1} = i^* = -i \neq B_{1,1}$ so this is not a Hermitian matrix and $\hat{B}$ is not a Hermitian operator.

2.15 Matrix Representations of Unitary Operators

In finite dimensions the one-to-one correspondence of matrices and operators means the matrix representation of unitary $\hat{U}$ satisfies

$$\underline{\underline{U}}^\dagger \cdot \underline{\underline{U}} = \underline{\underline{U}} \cdot \underline{\underline{U}}^\dagger = \underline{\underline{1}}$$

So if a matrix satisfies this, then the corresponding operator satisfies $\hat{U}^\dagger \hat{U} = \hat{U}\hat{U}^\dagger = \hat{1}$.

2.16 Hermitian Matrix Diagonalisation

For finite dimensional Hermitian matrices, we can find the eigenvalues by noting that

$$(\underline{\underline{A}} - \lambda \underline{\underline{1}})^{-1}$$

is not defined, which means

$$\det(\underline{\underline{A}} - \lambda \underline{\underline{1}}) = 0$$

Solving this equation gives the eigenvalues.

We know there exists a complete set of eigenvectors $\underline{u}_n$ that diagonalise the Hermitian matrix $\underline{\underline{A}}$.

$$\underline{\underline{A}} \cdot \underline{u}_n = \lambda_n \underline{u}_n$$

We can also choose these vectors to be orthonormal so $\underline{u}_n^\dagger \cdot \underline{u}_m = \delta_{n,m}$. This means we can put these vectors together into a matrix $\underline{\underline{U}}$

$$\underline{\underline{U}} = (\underline{u}_1 \ \underline{u}_2 \ \cdots \ \underline{u}_N)$$

From the definition of matrix multiplication we can show

$$\underline{\underline{U}}^\dagger \cdot \underline{\underline{U}} = \underline{\underline{1}}$$

and

$$\underline{\underline{U}} \cdot \underline{\underline{U}}^\dagger = \underline{\underline{1}}$$

so the matrix of orthonormal eigenvectors is unitary.

We also have

$$\underline{\underline{A}} \cdot \underline{\underline{U}} = (\lambda_1 \underline{u}_1 \ \lambda_2 \underline{u}_2 \ \cdots \ \lambda_N \underline{u}_N) = \underline{\underline{U}} \cdot \underline{\underline{\Lambda}}$$

where $\underline{\underline{\Lambda}}$ is a diagonal matrix of eigenvalues of $\underline{\underline{A}}$, so $[\underline{\underline{\Lambda}}]_{n,m} = \delta_{n,m}\lambda_n$.

Right multiplying this by $\underline{\underline{U}}^\dagger$ gives the eigendecomposition of $\underline{\underline{A}}$

$$\underline{\underline{A}} = \underline{\underline{U}} \cdot \underline{\underline{\Lambda}} \cdot \underline{\underline{U}}^\dagger$$

or left multiplying by $\underline{\underline{U}}^\dagger$ diagonalises $\underline{\underline{A}}$

$$\underline{\underline{U}}^\dagger \cdot \underline{\underline{A}} \cdot \underline{\underline{U}} = \underline{\underline{\Lambda}}.$$

3 Postulates of Quantum Mechanics

3.1 The Postulates

What is a postulate?

e.g. $\underline{F} = m\underline{a}$ - Everyone knows it gives a very good description of everyday objects but no-one really knows why!

These postulates for Quantum Mechanics give a reasonable starting point. We'll revisit them later though.

P1 The physical state of a quantum system is given by an "amplitude" $\psi(x) \in \mathbb{C}$. This takes any classical configuration x and associates it with a complex number $\psi(x)$. The set of allowed states form a Hilbert space $\mathcal{H}$.

E.g. For a particle in 1D

$$|\psi\rangle = \psi(x) \quad \text{for } x \in \mathbb{R}$$

For two particles in 3D

$$|\psi\rangle = \psi(\underline{r}_1, \underline{r}_2) \quad \text{for } \underline{r}_1, \underline{r}_2 \in \mathbb{R}^3 \quad (\underline{r}_i = (x_i, y_i, z_i))$$

Why? Allows superposition i.e. if $|\psi\rangle$ and $|\phi\rangle$ are states then $a|\psi\rangle + b|\phi\rangle$ is an allowed state.

P2 For every physical observable we have a Hermitian linear operator on $\mathcal{H}$.

$$\begin{array}{ccc} A & \longrightarrow & \hat{A} \\ \text{Observable} & & \text{Operator} \end{array}$$

E.g.

$$\text{position: } \hat{x} = x$$
$$\text{momentum: } \hat{p}_x = \frac{\hbar}{i}\frac{\partial}{\partial x}$$

Why? Hermitian operators have real eigenvalues and eigenvectors form a basis. We'll see why this is important soon.

P3 The average result of a measurement of $\hat{A}$ is

$$\langle A \rangle = \frac{\langle \psi | \hat{A} | \psi \rangle}{\langle \psi | \psi \rangle}$$

Why? Guaranteed to be real-valued (see below) and Born's rule, i.e. $|\psi(x)|^2 dx = \rho(x)dx$, is hidden in here. This postulate also means we can identify a classical expectation value of measurement of A, $\langle A \rangle_{\text{cl}}$, with this definition of the quantum expectation value, so $\langle A \rangle_{\text{cl}} = \langle A \rangle = \langle \psi | \hat{A} | \psi \rangle \,/\, \langle \psi | \psi \rangle$.

P4 Quantum states $|\psi(t)\rangle$ evolve in time according to the Schrödinger equation:

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

$\hat{H}$ is the energy operator, the Hamiltonian.

Why? (i) Recovers Planck's relation $E = \hbar\omega$ and the de Broglie relation $p = h/\lambda$. (ii) Conserves energy and more.

P5 Measurements of A can only take values a_n which are the eigenvalues of $\hat{A}$. If an observable is measured to have a value a_n then during the measurement the state is collapsed onto the eigenstate of $\hat{A}\ |a_n\rangle$

$$|\psi\rangle \longrightarrow |a_n\rangle \quad \text{if } a_n \text{ is measured.}$$

The probability of this measurement outcome is $P_n = |\langle a_n|\psi\rangle|^2 / \langle\psi|\psi\rangle$.

Why? We'll see later that the allowed measurement values, and the probability falls naturally out from (**P3**), but we need this postulate to define what happens *after* a measurement, i.e. the collapse.

Let's examine some of these postulates and check their sanity.

3.2 Average measurements

Average measurements of observables are real valued.

(P1) - (P3): Observables should always be measured to be real-valued. We've established that Hermitian operators $\hat{A} = \hat{A}^\dagger$ always have real eigenvalues and their eigenvectors form a complete basis. This means $\langle A\rangle$ in (P3) is always real.

Proof:

$$\langle A\rangle = \frac{\langle\psi|\hat{A}|\psi\rangle}{\langle\psi|\psi\rangle}, \quad \langle\psi|\psi\rangle \in \mathbb{R} \text{ by def.}$$

$$\begin{aligned}
\langle\psi|\hat{A}|\psi\rangle &= \langle\psi|\hat{1}\hat{A}|\psi\rangle \quad \text{choose: } \hat{1}\,|n\rangle = a_n\,|n\rangle \\
&= \langle\psi|\sum_n |n\rangle\,\langle n|\,\hat{A}|\psi\rangle \\
&= \sum_n \langle\psi|n\rangle\ \langle n|a_n|\psi\rangle \quad \because \langle n|\,\hat{A} = a_n\,\langle n| \\
&= \sum_n \langle\psi|n\rangle\,\langle n|\psi\rangle\,a_n \quad \because a_n \text{ is constant} \\
&= \sum_n |\langle\psi|n\rangle|^2 a_n \quad \because \langle\psi|n\rangle = \langle n|\psi\rangle^*
\end{aligned}$$

$a_n \in \mathbb{R}$, $|\langle n|\psi\rangle|^2 \in \mathbb{R} \Rightarrow \langle A\rangle \in \mathbb{R}$ $\square$

Note: In probability theory

$$\langle A\rangle_{\text{cl}} = \sum_n a_n P_n$$

P_n is the probability of measuring a_n.

So this implies $P_n \stackrel{?}{=} |\langle n|\psi\rangle|^2$? Do we really need (**P5**)?

We will revisit this shortly and show that more carefully $P_n = |\langle n|\psi\rangle|^2$ more rigorously.

3.3 Uncertainty in measurements

Eigenstates of $\hat{A}$ have zero uncertainty in the measurement of A.

We can define the uncertainty, or variance, of a measurement as

$$\sigma_A^2 = \left\langle (A - \langle A\rangle)^2 \right\rangle = \left\langle A^2 \right\rangle - \langle A\rangle^2$$

The only states where $\sigma_A^2 = \langle A^2 \rangle - \langle A \rangle^2 = 0$ are the eigenstates of $\hat{A}$.

Proof:

For $|\psi\rangle = \sum_n |n\rangle\, c_n$

$$\begin{aligned}\sigma_A^2 &= \langle\psi|(\hat{A} - \langle A\rangle)^2|\psi\rangle \\ &= \sum_n |c_n|^2 (a_n - \langle A\rangle)^2 = 0\end{aligned}$$

$|c_n|^2 \geq 0$ and $(a_n - \langle A\rangle)^2 \geq 0$

$\Rightarrow$ all $|c_n|^2(a_n - \langle A\rangle)^2 = 0$

$\Rightarrow$ either $|c_n|^2 = 0$ if $a_n - \langle A\rangle \neq 0$

or $|c_n|^2 > 0$ if $a_n - \langle A\rangle = 0$

$\sum_n |c_n|^2 = 1$ if $\langle\psi|\psi\rangle = 1$

So $|c_n| = 1$ for $a_n = \langle A\rangle$ i.e. $\langle A\rangle$ can only be a_n for $\sigma_A^2 = 0$. □

3.4 Commuting observables

Suppose $\hat{A}$, $\hat{B}$ are observable operators, i.e. Hermitian

Also suppose that they commute.

$$\hat{A}\hat{B} = \hat{B}\hat{A}$$

We will now show an important property of the eigenstates of commuting observables.

Now let us act on $|a_n\rangle$, where $\hat{A}\,|a_n\rangle = |a_n\rangle\, a_n$ with $\hat{A}\hat{B}$.

$$\begin{aligned}\hat{A}\hat{B}\,|a_n\rangle &= \hat{B}\hat{A}\,|a_n\rangle \quad \because \hat{A}\hat{B} = \hat{B}\hat{A} \\ &= \hat{B}(a_n\,|a_n\rangle) \quad \because \hat{A}\,|a_n\rangle = |a_n\rangle\, a_n \\ &= a_n(\hat{B}\,|a_n\rangle) \quad \because a_n \text{ is constant}\end{aligned}$$

This implies $|\phi\rangle = \hat{B}\,|a_n\rangle$ is an eigenstate of $\hat{A}$ as well.

If a_n is non-degenerate i.e. no other linearly independent $|a_m\rangle$ exist where $a_m = a_n$, this means

$$\hat{B}\,|a_n\rangle = b_n\,|a_n\rangle \quad \text{where } b_n \text{ is a const.}$$

So $|a_n\rangle$ is also an eigenstate of $\hat{B}$ with eigenvalue b_n.

Suppose however that a_n is degenerate and $\{\,|a_n, k\rangle : k = 1, \ldots, M\}$ is the set of orthonormal degenerate eigenstates.

$$\hat{A}\,|a_n, k\rangle = |a_n, k\rangle\, a_n$$

We know $\forall k$

$$\hat{B}\,|a_n, k\rangle = \sum_{j=1}^{M} B_{kj}^{(n)}\,|a_n, j\rangle$$

If we diagonalise $B_{jk}^{(n)}$, so finding the vectors $\underline{c}_m$, such that $\underline{\underline{B}}^{(n)}\underline{c}_m = b_m\underline{c}_m$ then

$$\hat{B}\sum_j c_{mj}\,|a_n, j\rangle = b_m \sum_j c_{mj}\,|a_n, j\rangle$$

$$\text{and } \hat{A}\sum_j c_{mj}\,|a_n, j\rangle = a_n \sum_j c_{mj}\,|a_n, j\rangle$$

So we can find a set of states which are **simultaneously** eigenstates of $\hat{A}$ and $\hat{B}$ if $[\hat{A}, \hat{B}] = 0$. These states have **no uncertainty** in the measurements of **both** A and B.

If $[\hat{A}, \hat{B}] \neq 0$ this is not true in general.

To show this we will find a general form of the uncertainty principle.

3.5 The Uncertainty Principle

Let us consider σ_A^2 and σ_B^2 for a general state

$$\sigma_A^2 = \left\langle A^2 \right\rangle - \langle A \rangle^2 = \left\langle (A - \langle A \rangle)^2 \right\rangle$$
$$\sigma_B^2 = \left\langle B^2 \right\rangle - \langle B \rangle^2 = \left\langle (B - \langle B \rangle)^2 \right\rangle$$

Define $\hat{a} = \hat{A} - \langle A \rangle$, $\hat{b} = \hat{B} - \langle B \rangle$ (both Hermitian operators)

$$\begin{aligned} \sigma_A^2 \sigma_B^2 = & \ \langle \psi | \hat{a}^2 | \psi \rangle \ \langle \psi | \hat{b}^2 | \psi \rangle \\ = & \ (\langle \hat{a}\psi | \hat{a}\psi \rangle)(\left\langle \hat{b}\psi \middle| \hat{b}\psi \right\rangle) \quad \because \hat{a}, \hat{b} \text{ are Hermitian} \end{aligned}$$

The Cauchy-Schwarz inequality tells us

$$\langle v | v \rangle \langle w | w \rangle \geq | \langle v | w \rangle |^2$$

for general Hilbert space vectors $|v\rangle$ and $|w\rangle$.

So setting $|v\rangle = \hat{a} |\psi\rangle$ and $|w\rangle = \hat{b} |\psi\rangle$

$$\begin{aligned} \sigma_A^2 \sigma_B^2 &= \langle \hat{a}\psi | \hat{a}\psi \rangle \left\langle \hat{b}\psi \middle| \hat{b}\psi \right\rangle \\ &\geq | \left\langle \hat{a}\psi \middle| \hat{b}\psi \right\rangle |^2 \quad \text{by Cauchy-Schwarz} \end{aligned}$$

Also for a general complex number $|z|^2 = \mathrm{Re}[z]^2 + \mathrm{Im}[z]^2$, so

$$| \left\langle \hat{a}\psi \middle| \hat{b}\psi \right\rangle |^2 = \mathrm{Re}[\left\langle \hat{a}\psi \middle| \hat{b}\psi \right\rangle]^2 + \mathrm{Im}[\left\langle \hat{a}\psi \middle| \hat{b}\psi \right\rangle]^2$$

For the real part we have

$$\begin{aligned} \mathrm{Re}[\left\langle \hat{a}\psi \middle| \hat{b}\psi \right\rangle] &= \frac{1}{2}(\left\langle \hat{a}\psi \middle| \hat{b}\psi \right\rangle + \left\langle \hat{a}\psi \middle| \hat{b}\psi \right\rangle^*) \\ &= \frac{1}{2}(\left\langle \hat{a}\psi \middle| \hat{b}\psi \right\rangle + \left\langle \hat{b}\psi \middle| \hat{a}\psi \right\rangle) \\ &= \frac{1}{2}(\ \langle \psi | \hat{a}\hat{b} | \psi \rangle + \ \langle \psi | \hat{b}\hat{a} | \psi \rangle) \\ &= \frac{1}{2}(\ \langle \psi | (\hat{a}\hat{b} + \hat{b}\hat{a}) | \psi \rangle) \end{aligned}$$

and for the imaginary part

$$\begin{aligned} \mathrm{Im}[\left\langle \hat{a}\psi \middle| \hat{b}\psi \right\rangle] &= \frac{1}{2i}(\left\langle \hat{a}\psi \middle| \hat{b}\psi \right\rangle - \left\langle \hat{a}\psi \middle| \hat{b}\psi \right\rangle^*) \\ &= \frac{1}{2i} \langle \psi | (\hat{a}\hat{b} - \hat{b}\hat{a}) | \psi \rangle \end{aligned}$$

So:

$$\begin{aligned}\sigma_A^2\sigma_B^2 &\geq |\left\langle \hat{a}\psi \middle| \hat{b}\psi \right\rangle|^2 \\ &\geq \mathrm{Im}[\left\langle \hat{a}\psi \middle| \hat{b}\psi \right\rangle]^2 \quad \because \mathrm{Re}[\left\langle \hat{a}\psi \middle| \hat{b}\psi \right\rangle]^2 \geq 0 \\ &= \left| \langle\psi| \frac{1}{2i}[\hat{a},\hat{b}] |\psi\rangle \right|^2\end{aligned}$$

Finally we can deduce that $[\hat{a},\hat{b}] = [\hat{A} - \langle A\rangle, \hat{B} - \langle B\rangle] = [\hat{A},\hat{B}]$.

So:

$$\sigma_A^2\sigma_B^2 \geq \left| \frac{1}{2i} \left\langle [\hat{A},\hat{B}] \right\rangle \right|^2 \quad \square$$

So if $[\hat{A},\hat{B}] \neq 0$ there will always be states where σ_A^2 and σ_B^2 are > 0.

Likewise if $\sigma_A^2 \to 0$ then $\sigma_B^2 \to +\infty$ if $[\hat{A},\hat{B}] \neq 0$

So increasing certainty in A decreases certainty in B!

3.6 Symmetries of particle in a 1D box

Let's look at a familiar example of the particle in a 1D box to see how commuting hermitian operators have simultaneous eigenstates.

The energy eigenstates for a particle in an infinite potential walled box between $0 \leq x \leq L$ are

$$|E_n\rangle = \psi_n(x) = \begin{cases} \sqrt{\frac{2}{L}} \sin\left(\frac{n\pi x}{L}\right) & 0 \leq x \leq L \\ 0 & \text{everywhere else} \end{cases}$$

for $n = 1, 2, 3, ...$ The energy eigenvalues are

$$E_n = \frac{\hbar^2}{2m}\left(\frac{\pi n}{L}\right)^2$$

The operator $\hat{\Pi}$ given by

$$\hat{\Pi}\psi(x) = \psi(L - x)$$

is a hermitian operator on functions restricted to be 0 at $x \leq 0$ and $x \geq L$, as can be confirmed with a simple integral substitution

$$\int_0^L \phi(x)^*\psi(L-x)\,\mathrm{d}x = \int_0^L \phi(L-x)^*\psi(x)\,\mathrm{d}x\,.$$

Furthermore $\hat{\Pi}^2 = \hat{1}$, so the eigenvalues of $\hat{\Pi}$ are $\lambda = \pm 1$. (Note that $\hat{P}i$ is also unitary.)

$$\hat{\Pi}^2|\lambda\rangle = \lambda^2|\lambda\rangle = |\lambda\rangle \implies \lambda^2 = 1$$

The particle in a 1D box hamiltonian commutes with this operator

$$\hat{\Pi}\frac{\partial^2}{\partial x^2}\psi(x) = \psi''(L-x) = \frac{\partial^2}{\partial(L-x)^2}\psi(L-x) = \frac{\partial^2}{\partial x^2}\hat{\Pi}\psi(x)$$

$$\hat{\Pi}(V(x)\psi(x)) = V(L-x)\psi(L-x) = V(x)\hat{\Pi}\psi(x)$$

So there must be simultaneous eigenstates of $\hat{H}$ and $\hat{\Pi}$. The energy eigenstates are non-degenerate, so from what we deduced above, these must be eigenstates of $\hat{\Pi}$ as well. It is striaghtforward to find their eigenvalues as

$$\hat{\Pi}\psi_n(x) = \sqrt{\frac{2}{L}}\sin\left(\frac{n\pi(L-x)}{L}\right) = \sqrt{\frac{2}{L}}\left[\sin(n\pi)\cos\left(\frac{n\pi x}{L}\right) - \cos(n\pi)\sin\left(\frac{n\pi x}{L}\right)\right] = (-1)^{n+1}\psi_n(x)$$

so the eigenvalues of $\hat{\Pi}$ for these states are $\lambda_n = (-1)^{n+1}$.

3.7 Probability densities

We will now consider the classical probability density for the measurement of A, which we'll denote $p(a)$. $p(a)\,\mathrm{d}a$ tells us the probability of that value of A will be measured to have a value between a and $a + \mathrm{d}a$. So the probability that A is between a_1 and a_2 is

$$\mathrm{Prob}(a_1 \leq A \leq a_2) = \int_{a_1}^{a_2} p(a)\,\mathrm{d}a$$

We know $p(a) \geq 0$, because probabilities are positive numbers. Likewise

$$\int_{-\infty}^{\infty} p(a)\,\mathrm{d}a = 1$$

and the (classical) moments of a probability density are

$$\langle A^n \rangle_{\mathrm{cl}} = \int_{-\infty}^{\infty} a^n p(a)\,\mathrm{d}a\,.$$

So if we know $p(a)$, we can full characterise the probability of any particular outcome for the measurement of A.

The probability density $p(a)$ is fully defined by its Fourier transform. This function $\mu(s)$ is called the **moment generating function**

$$\mu(s) = \int_{-\infty}^{\infty} p(a) e^{-ias}\,\mathrm{d}a = \left\langle e^{-isA} \right\rangle_{\mathrm{cl}}$$

This is because the Fourier transform is invertible

$$p(a) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \mu(s) e^{+ias}\,\mathrm{d}a$$

For using the series expansion for the exponential, we see that

$$\begin{aligned}
\mu(s) &= \int_{-\infty}^{\infty} p(a) \sum_{k=0}^{\infty} \frac{1}{k!}(-isa)^k\,\mathrm{d}a \\
&= \sum_{k=0}^{\infty} \frac{1}{k!}(-is)^k \int_{-\infty}^{\infty} p(a) a^k\,\mathrm{d}a \\
&= \sum_{k=0}^{\infty} \frac{1}{k!}(-is)^k \left\langle A^k \right\rangle_{\mathrm{cl}}
\end{aligned}$$

So if we know all the moments of a probability distribution $\langle A^n \rangle_{\mathrm{cl}}$, then we can construct the moment generating function $\mu(s)$, and if we know $\mu(s)$ we can construct $p(a)$ by using the inverse Fourier transform.

If A can only take discrete values a_n with a probability P_n then the moments are

$$\left\langle A^k \right\rangle_{\mathrm{cl}} = \sum_n a_n P_n$$

Putting this into the moment generating function we find

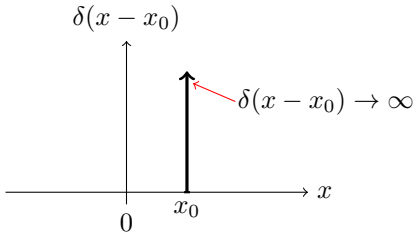
$$\begin{aligned}
\mu(s) &= \sum_{k=0}^{\infty} \frac{1}{k!}(-is)^k \sum_n P_n a_n^k \\
&= \sum_n P_n \left(\sum_{k=0}^{\infty} \frac{1}{k!}(-isa_n)^k \right) \\
&= \sum_n P_n e^{-isa_n}
\end{aligned}$$

From this we can recover the probability density using the inverse Fourier transform

$$\begin{aligned} p(a) &= \frac{1}{2\pi}\int_{-\infty}^{\infty}\sum_n P_n e^{-isa_n}\,\mathrm{d}a \\ &= \sum_n P_n \underbrace{\frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{is(a-a_n)}ds}_{\delta(a-a_n)} \\ &= \sum_n P_n \delta(a-a_n) \end{aligned}$$

Note: $\delta(x-x_0)$ is the"delta-function"

$\delta(x-x_0)$ is 0 everywhere apart from at x_0 where $\delta(x-x_0)\to\infty$



And

$$\int_{-\infty}^{+\infty}\delta(x-x_0)dx = 1$$

and

$$\int_{-\infty}^{+\infty}\delta(x-x_0)x^n dx = 0 \quad \text{for } n>0.$$

So

$$\int_{-\infty}^{+\infty}\delta(x-x_0)f(x)dx = f(x_0).$$

So these properties of the delta function mean this recovers the expected behaviour for the probabilities. $\mathrm{Prob}(\alpha_1 \le A \le \alpha_2)$ is zero if none of the discrete values a_n that A can take are in the range α_1 to α_2. If a single value a_n is in this range then $\mathrm{Prob}(\alpha_1 \le A \le \alpha_2) = P_n$, so also for a small positive $\delta\alpha$, with $\delta\alpha \to 0$, $\mathrm{Prob}(a_n - \delta\alpha \le A \le a_n + \delta\alpha) = P_n$.

Back the quantum mechanics...

Now let's go back to quantum mechanics. We want to show generally that $P_n = |\langle n|\psi\rangle|^2$.

Firstly we note that the quantum moments of measurement of A are well defined (show this)

$$\langle A^n\rangle = \frac{\sum_m |\langle m|\psi\rangle|^2 a_m^n}{\langle\psi|\psi\rangle}$$

Now let us define a function of $\hat{A}$ by its Taylor series of $f(a)$

$$f(a) = f(0) + f'(0)a + \frac{1}{2!}f''(0)a^2 + \frac{1}{3!}f'''(0)a^3 + \cdots$$

$$f(\hat{A}) \stackrel{\text{def}}{=} f(0)\hat{1} + f'(0)\hat{A} + \frac{1}{2!}f''(0)\hat{A}^2 + \frac{1}{3!}f'''(0)\hat{A}^3 + \cdots$$

So the moment generating *operator* $e^{-is\hat{A}}$ is well-defined, as is its expectation value,

$$\left\langle e^{-isA} \right\rangle = \frac{1}{\langle\psi|\psi\rangle} \langle\psi|e^{-is\hat{A}}|\psi\rangle$$

Let's assume $\langle\psi|\psi\rangle = 1$ without loss of generality. Using the definition of the moment generating operator this is

$$\begin{aligned}
\left\langle e^{-isA} \right\rangle &= \langle\psi|e^{-is\hat{A}}|\psi\rangle \\
&= \langle\psi|\sum_{k=0}^{\infty} \frac{1}{k!}(-is\hat{A})^k|\psi\rangle \\
&= \sum_{k=0}^{\infty} \frac{1}{k!}(-is)^k \langle\psi|\hat{A}^k|\psi\rangle \\
&= \sum_{k=0}^{\infty} \frac{1}{k!}(-is)^k \left\langle A^k \right\rangle
\end{aligned}$$

So this expectation value is well defined and we can identity this as the moment generating function using **P3** to equate the quantum and classical expectation values.

$$\mu(s) = \langle\psi|e^{-is\hat{A}}|\psi\rangle$$

Given that we have a well-defined moment generating function $\mu(s)$ we can re-write it in terms of the eigenstates of $\hat{A}$ and its eigenvalues as

$$\begin{aligned}
\mu(s) &= \langle\psi|e^{-is\hat{A}}|\psi\rangle \\
&= \sum_n \langle\psi|n\rangle \langle n|\psi\rangle e^{-isa_n} \\
&= \sum_n |\langle\psi|n\rangle|^2 e^{-isa_n}
\end{aligned}$$

So we can find $p(a)$ by applying the inverse Fourier transform to this

$$\begin{aligned}
p(a) &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{+isa} \mu(t) ds \\
&= \frac{1}{2\pi} \int_{-\infty}^{+\infty} \sum_n e^{isa} e^{-isa_n} |\langle n|\psi\rangle|^2 ds \\
&= \sum_n |\langle n|\psi\rangle|^2 \underbrace{\frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{is(a-a_n)} ds}_{\delta(a-a_n)} \\
&= \sum_n |\langle n|\psi\rangle|^2 \delta(a - a_n)
\end{aligned}$$

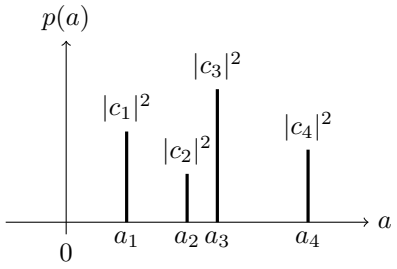
If A can only take discrete values, the the classical probability density is

$$p(a) = \sum_n P_n \delta(a - a_n)$$

So clearly $P_n = |\langle n|\psi\rangle|^2$. This is Born's rule.

Consequences:

Our expression for $p(a)$ shows $p(a) \neq 0$ only at eigenvalues of $\hat{A}$, so we only measure $A = a_n$



But if we measure $A = a_n$, the only states with $\sigma_A^2 = 0$ are the $\hat{A}\,|n\rangle = a_n\,|n\rangle$ states, so after the measurement we get $|\psi\rangle = |n\rangle$

This is consistent with (P5)!

Next-time: Continuous symmetries and time-evolution.

Final note: We could have cheated by doing the integral first

$$\begin{aligned} p(a) &= \frac{1}{2\pi}\int_{-\infty}^{+\infty} \left\langle e^{-isA}\right\rangle e^{isa} ds \\ &= \left\langle \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-is\hat{A}+isa} ds \right\rangle \\ &= \langle\psi| \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-is\hat{A}+isa} ds |\psi\rangle \\ &= \langle\psi|\delta(\hat{A}-a)|\psi\rangle \end{aligned}$$

This is okay to do but $\delta(x - x_0)$ doesn't have a Taylor series so our definition of $f(\hat{A})$ cannot be applied directly. It works because of the above, and so we can write

$$\delta(\hat{A}-a)\,|n\rangle = \delta(a_n - a)\,|n\rangle\,.$$

4 Time Evolution and Continuous Symmetries

4.1 Time Evolution

Last time we introduced the postulate that time evolution of a state $|\psi(t)\rangle$ is governed by

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

where $\hat{H}$ is the Hamiltonian (total energy) operator.

Suppose we know $|\psi(0)\rangle = |\psi(t=0)\rangle$. We can solve this using a resolution of the identity in the energy eigenbasis, $\hat{H}|n\rangle = |n\rangle E_n$.

The LHS of (TDSE) can be re-written as

$$\begin{aligned} i\hbar\frac{\partial}{\partial t}|\psi(t)\rangle &= i\hbar\frac{\partial}{\partial t}\hat{1}|\psi(t)\rangle \\ &= i\hbar\frac{\partial}{\partial t}\sum_n |n\rangle\langle n|\psi(t)\rangle \\ &= \sum_n |n\rangle\left(i\hbar\frac{\partial}{\partial t}\langle n|\psi(t)\rangle\right) \end{aligned}$$

And the RHS likewise can be found to be

$$\begin{aligned} \hat{H}|\psi(t)\rangle &= \hat{H}\hat{1}|\psi(t)\rangle \\ &= \sum_n \hat{H}|n\rangle\langle n|\psi(t)\rangle \\ &= \sum_n |n\rangle E_n c_n(t) \end{aligned}$$

Equating the LHS and RHS we find that

$$\sum_n |n\rangle\left(i\hbar\frac{\partial}{\partial t}c_n(t)\right) = \sum_n |n\rangle E_n c_n(t)$$

We equate coefficients of $|n\rangle$ to find an equation for $c_n(t)$

$$i\hbar\frac{\mathrm{d}}{\mathrm{d}t}c_n(t) = E_n c_n(t)$$

$$\begin{aligned} &\Rightarrow \quad \frac{1}{c_n(t)}\frac{\mathrm{d}}{\mathrm{d}t}c_n(t) = -i\frac{E_n}{\hbar} \\ &\Rightarrow \quad \frac{\mathrm{d}}{\mathrm{d}t}\ln(c_n(t)) = -i\frac{E_n}{\hbar} \\ &\Rightarrow \quad \int_0^t \frac{\mathrm{d}}{\mathrm{d}t'}\ln(c_n(t'))dt' = \int_0^t\left(-i\frac{E_n}{\hbar}\right)dt' \\ &\Rightarrow \quad \ln\left(\frac{c_n(t)}{c_n(0)}\right) = -i\frac{E_n t}{\hbar} \\ &\Rightarrow \quad c_n(t) = c_n(0)e^{-iE_n t/\hbar} \end{aligned}$$

So overall the full time-dependent quantum state $|\psi(t)\rangle$ can be written as

$$|\psi(t)\rangle = \sum_n e^{-iE_n t/\hbar}|n\rangle\langle n|\psi(0)\rangle$$

So this means if we can solve the energy eigenvalue equation

$$\hat{H}\,|n\rangle = E_n\,|n\rangle$$

then we can immediately obtain $|\psi(t)\rangle$ for any $|\psi(0)\rangle$.

4.2 Ehrenfest Theorem

We can also use the TDSE to obtain an equation for how $\langle A(t)\rangle = \langle\psi(t)|\hat{A}|\psi(t)\rangle$ changes with time. Here we assume that $\hat{A}$ itself is independent of time.

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\langle A(t)\rangle &= \frac{\mathrm{d}}{\mathrm{d}t}\langle\psi(t)|\hat{A}|\psi(t)\rangle \\
&= \left\langle \frac{\partial\psi(t)}{\partial t}\middle|\hat{A}\middle|\psi(t)\right\rangle + \left\langle\psi(t)\middle|\hat{A}\middle|\frac{\partial\psi(t)}{\partial t}\right\rangle \\
&= \left\langle -\frac{i}{\hbar}\hat{H}\psi(t)\middle|\hat{A}\middle|\psi(t)\right\rangle + \left\langle\psi(t)\middle|\hat{A}\middle|-\frac{i}{\hbar}\hat{H}\psi(t)\right\rangle \\
&= +\frac{i}{\hbar}\langle\psi(t)|\hat{H}\hat{A}|\psi(t)\rangle - \frac{i}{\hbar}\langle\psi(t)|\hat{A}\hat{H}|\psi(t)\rangle \\
&= \frac{i}{\hbar}\langle\psi(t)|(\hat{H}\hat{A}-\hat{A}\hat{H})|\psi(t)\rangle \\
&= \frac{i}{\hbar}\langle\psi(t)|[\hat{H},\hat{A}]|\psi(t)\rangle \\
&= \frac{i}{\hbar}\left\langle[\hat{H},\hat{A}](t)\right\rangle
\end{aligned}$$

This is called **Ehrenfest theorem**.

Example: For a particle in 1D with $\hat{H} = \frac{1}{2m}\hat{p}^2 + V(\hat{x})$ what is $\frac{\mathrm{d}}{\mathrm{d}t}\langle x(t)\rangle$?

$$\begin{aligned}
[\hat{H},\hat{x}] &= \left[\frac{\hat{p}^2}{2m},\hat{x}\right] + [V(\hat{x}),\hat{x}] \\
&= \frac{1}{2m}\left(\hat{p}[\hat{p},\hat{x}] + [\hat{p},\hat{x}]\hat{p}\right) \\
&= \frac{1}{2m}\left(\hat{p}\cdot(-i\hbar) + (-i\hbar)\hat{p}\right) \\
&= -i\hbar\frac{1}{m}\hat{p}
\end{aligned}$$

So putting this into Ehrenfest theorem we find

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\langle x(t)\rangle &= \frac{i}{\hbar}\left\langle\frac{-i\hbar}{m}p(t)\right\rangle \\
&= \frac{1}{m}\langle p(t)\rangle
\end{aligned}$$

which is what we expect from classical mechanics!

Note: A derivation of $\hat{p}_x = -i\hbar\frac{\partial}{\partial x}$ (?)

We could run the argument above in reverse. The only way $\frac{\mathrm{d}}{\mathrm{d}t}\langle x(t)\rangle = \frac{1}{m}\langle p(t)\rangle$ is for $[\hat{p},\hat{x}] = -i\hbar$, and the only operator that satisfies this is $\hat{p} = -i\hbar\frac{\partial}{\partial x}$.

4.3 Time-evolution operator

Going back to our solution for $|\psi(t)\rangle$, we can write this in a different form.

$$\begin{aligned}
|\psi(t)\rangle &= \sum_n |n\rangle \ \langle n|e^{-iE_n t/\hbar}|\psi(0)\rangle \\
&= \sum_n |n\rangle \ \langle n|e^{-i\hat{H}t/\hbar}|\psi(0)\rangle \\
&= e^{-i\hat{H}t/\hbar}\sum_n |n\rangle \langle n|\psi(0)\rangle \\
&= e^{-i\hat{H}t/\hbar}\hat{1}\,|\psi(0)\rangle \\
&= e^{-i\hat{H}t/\hbar}\,|\psi(0)\rangle
\end{aligned}$$

We call $e^{-i\hat{H}t/\hbar} = \hat{U}(t)$ the time-evolution operator.

It takes a state at a time t_0 and evolves it to $t_0 + t$:

$$|\psi(t_0 + t)\rangle = \hat{U}(t)\,|\psi(t_0)\rangle$$

$\hat{U}(t)$ has the following important properties (which follow from the fact that it is a function of the Hermitian operator $\hat{H}$)

$$\begin{aligned}
\hat{U}(0) &= \hat{1} \\
\hat{U}(t_1 + t_2) &= \hat{U}(t_1)\hat{U}(t_2) \\
\hat{U}(t)^\dagger &= \left(e^{-i\hat{H}t/\hbar}\right)^\dagger = e^{+i\hat{H}t/\hbar} = \hat{U}(-t) \\
\hat{U}(t)^\dagger &= \hat{U}(t)^{-1}
\end{aligned}$$

The last property follows from

$$\hat{U}(t)^\dagger\hat{U}(t) = e^{+i\hat{H}t/\hbar}e^{-i\hat{H}t/\hbar} = e^{i\hat{H}(t-t)/\hbar} = e^0 = \hat{1}$$

and likewise $\hat{U}(t)\hat{U}(t)^\dagger = \hat{1}$. So $\hat{U}(t)^\dagger = \hat{U}(t)^{-1}$.

This means $\hat{U}(t)$ is a unitary operator (remember those from our introduction to linear algebra?).

Using the time evolution operator, we can easily show things like $\langle\psi(t)|\phi(t)\rangle = \langle\psi(0)|\phi(0)\rangle$

Proof: From unitarity

$$\begin{aligned}
\langle\psi(t)|\phi(t)\rangle &= \langle\psi(0)|\hat{U}(t)^\dagger\hat{U}(t)|\phi(0)\rangle \\
&= \langle\psi(0)|\hat{1}|\phi(0)\rangle \\
&= \langle\psi(0)|\phi(0)\rangle \quad \square
\end{aligned}$$

4.4 Time-dependent operators

Previously we showed that states evolve according to

$$|\psi(t)\rangle = e^{-i\hat{H}t/\hbar}\,|\psi(0)\rangle$$

But states are never directly observable. We saw from Ehrenfest theorem that

$$\frac{d}{dt}\langle A\rangle = \frac{i}{\hbar}\left\langle[\hat{H}, \hat{A}]\right\rangle$$

We could also define the time-dependent (Heisenberg) operator $\hat{A}(t)$

$$\hat{A}(t) = \hat{U}(t)^\dagger \hat{A} \hat{U}(t) = e^{i\hat{H}t/\hbar} \hat{A} e^{-i\hat{H}t/\hbar}$$

The differential equation for the evolution of $\hat{A}$ is found to be

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\hat{A}(t) &= \frac{i}{\hbar}(\hat{H}e^{i\hat{H}t/\hbar}\hat{A}e^{-i\hat{H}t/\hbar} - e^{i\hat{H}t/\hbar}\hat{A}e^{-i\hat{H}t/\hbar}\hat{H}) \\ &= \frac{i}{\hbar}[\hat{H}, \hat{A}(t)]\end{aligned}$$

We can also write this equivalently as

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\hat{A}(t) &= \frac{i}{\hbar}(e^{i\hat{H}t/\hbar}\hat{H}\hat{A}e^{-i\hat{H}t/\hbar} - e^{i\hat{H}t/\hbar}\hat{A}\hat{H}e^{-i\hat{H}t/\hbar}) \\ &= \frac{i}{\hbar}[\hat{H}, \hat{A}](t) \\ &= \frac{i}{\hbar}[\hat{H}, \hat{A}(t)]\end{aligned}$$

where $[\hat{H}, \hat{A}](t) = \hat{U}(t)^\dagger [\hat{H}, \hat{A}] \hat{U}(t)$ has the same definition as $\hat{A}(t)$. The last line follows from the fact that $[\hat{U}(t), \hat{H}] = 0$

Using this we can write time-dependent expectation values in two different ways

$$\begin{aligned}\left\langle \psi(0)|\hat{A}(t)|\psi(0)\right\rangle &= \left\langle \psi(0)|U(t)^\dagger \hat{A} U(t)|\psi(0)\right\rangle \\ &= \left\langle \psi(t)|\hat{A}|\psi(t)\right\rangle\end{aligned}$$

So we can move time-dependence onto operators instead of states.

This can often simplify calculations as we will see with the harmonic oscillator.

Note: When we work with time-dependent states and time-independent operators, this is called the **Schrödinger picture**. When we work with time-independent states and time-dependent operators, this is called working in the **Heisenberg picture**. Both are equally valid descriptions of dynamics, but depending on the problem at hand, one may be more useful than the other.

4.5 Conserved Quantities

What if $[\hat{H}, \hat{A}] = 0$? Ehrenfest theorem tells us

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle A(t)\rangle = \frac{1}{i}\left\langle [\hat{H}, \hat{A}](t)\right\rangle = 0$$

$$\Rightarrow \quad \langle A(t)\rangle = \langle A(0)\rangle$$

So $\langle A(t)\rangle$ is a constant, i.e. it is a conserved quantity.

Equivalently, A is conserved if

$$[\hat{U}(t), \hat{A}] = 0$$

As can be shown by considering

$$\begin{aligned}\langle A(t)\rangle &= \langle \psi(0)|\hat{U}(t)^\dagger \hat{A} \hat{U}(t)|\psi(0)\rangle \\ &= \langle \psi(0)|\hat{U}(t)^\dagger \hat{U}(t)\hat{A}|\psi(0)\rangle \\ &= \langle \psi(0)|\hat{1}\hat{A}|\psi(0)\rangle \\ &= \langle A(0)\rangle\end{aligned}$$

Consequences:

$$[\hat{H}, \hat{1}] = 0 \Rightarrow \langle 1(t)\rangle = \langle \psi(t)|\psi(t)\rangle \text{ is conserved}$$
$$[\hat{H}, \hat{H}] = 0 \Rightarrow \langle H(t)\rangle \text{ is conserved, energy is conserved.}$$
$$[\hat{H}, \hat{H}^n] = 0 \Rightarrow \langle H^n(t)\rangle \text{, moments of energy distribution are conserved}$$
$$[\hat{H}, \delta(E-\hat{H})] \Rightarrow \left\langle \delta(E-\hat{H})(t)\right\rangle = \rho_E(E,t) \text{ the energy distribution is conserved.}$$

This is completely consistent with what we already knew, that $|c_n(t)|^2 = |c_n(0)|^2$ in our direct solution to the time-dependent Schrödinger equation.

4.6 The TDSE as a consequence of unitary dynamics

Let us consider a new way of arriving at the TDSE, without including it as a postulate.

First we assert that time evolution is generated by an operator $\hat{U}(t)$ – we don't know yet what $\hat{U}(t)$ is!

$$\hat{U}(t)\,|\psi(0)\rangle = |\psi(t)\rangle$$

Clearly $\hat{U}(0) = \hat{1}$.

Also $\hat{U}(t)$ should be norm conserving; this gives conservation of probability.

$$\langle\psi|\hat{U}(t)^\dagger\hat{U}(t)|\psi\rangle = \langle\psi|\psi\rangle$$

This implies $\langle\psi|\hat{U}(t)^\dagger\hat{U}(t)|\phi\rangle = \langle\psi|\phi\rangle$ (see problems) so $\hat{U}(t)^\dagger\hat{U}(t) = \hat{1}$. This doesn't quite show $\hat{U}(t)$ must be unitary because we also need $\hat{U}(t)\hat{U}(t)^\dagger = \hat{1}$. Let's show this now.

$\hat{U}(t)$ should also obey

$$\hat{U}(t_1+t_2) = \hat{U}(t_1)\hat{U}(t_2)$$

Evolving by $t_1 + t_2$ should give the same state as evolving by t_2 then t_1.

This implies

$$\begin{aligned} & \hat{U}(0) = \hat{1} = \hat{U}(-t)\hat{U}(t) = \hat{U}(+t)\hat{U}(-t) \\ \Rightarrow \quad & \hat{U}(t)^{-1} = \hat{U}(-t) \\ \Rightarrow \quad & \hat{U}(t)^{-1} = \hat{U}(t)^\dagger \quad \text{from norm-conservation.} \end{aligned}$$

This also incidentally proves that $\hat{U}(t)$ is invertible.

So simple physical arguments tell us $\hat{U}(t)$ is unitary.

Furthermore, consider $t = \delta t \to 0$, let us expand $\hat{U}(\delta t)$:

$$\hat{U}(\delta t) = \hat{U}(0) + \hat{U}'(0)\delta t + \mathcal{O}(\delta t^2)$$

We know $\hat{U}(\delta t)^\dagger\hat{U}(\delta t) = \hat{1}$

$$\begin{aligned} \hat{U}(\delta t)^\dagger\hat{U}(\delta t) &= (\hat{1} + \hat{U}'(0)^\dagger\delta t)(\hat{1} + \hat{U}'(0)\delta t) + \mathcal{O}(\delta t^2) \\ &= \hat{1} + (\hat{U}'(0)^\dagger + \hat{U}'(0))\delta t + \mathcal{O}(\delta t^2) \\ &= \hat{1} \end{aligned}$$

$$\Rightarrow \quad \hat{U}'(0)^\dagger = -\hat{U}'(0)$$

This means we can define $\hat{U}'(0) = -i\hat{\Omega}$, where $\hat{\Omega} = \hat{\Omega}^\dagger$

$$\hat{U}(\delta t) = \hat{1} - i\hat{\Omega}\delta t + \mathcal{O}(\delta t^2)$$

From $\hat{U}(t_1 + t_2) = \hat{U}(t_1)\hat{U}(t_2)$ we can see that

$$\begin{aligned}
\hat{U}(t) &= \lim_{N\to\infty} \hat{U}(t/N)^N \\
&= \lim_{N\to\infty} (\hat{1} - i\hat{\Omega}\frac{t}{N})^N \\
&= \lim_{N\to\infty} (e^{-i\hat{\Omega}(t/N)})^N \\
&= e^{-i\hat{\Omega}t}
\end{aligned}$$

So time evolution being unitary implies the existence of a Hermitian operator $\hat{\Omega}$ that "generates" $\hat{U}(t)$.

Clearly $[\hat{U}(t), \hat{\Omega}] = [e^{-i\hat{\Omega}t}, \hat{\Omega}] = 0$ (an operator commutes with a function of itself).

So $\langle\Omega\rangle$ is conserved for all quantum systems.

This implies $\hat{\Omega}$ is a function of $\hat{H}$ – the only universally conserved quantity.

$$\hat{\Omega} = f(\hat{H})$$

Finally, if we scale the energy of system by γ:

$$\hat{H}' = \gamma\hat{H}$$

Then dynamics look the same after re-scaling time by γ:

$$t' = \frac{1}{\gamma}t$$

This implies:

$$f(\gamma\hat{H})t/\gamma = f(\hat{H})t$$

Therefore $f(\hat{H})$ must be a linear function of $\hat{H}$

$$f(\hat{H}) = \frac{1}{\hbar}\hat{H}$$

where $\hbar$ is set by our choice of unit system for $\hat{H}$ and t.

As before take the derivative of $|\psi(t)\rangle$ and we get:

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

So we have derived the TDSE essentially just from conservation of probability and energy! This has some very deep connections to Noether's theorem and continuous symmetries in physics. We've effectively shown that a continuous symmetry of our system, the fact that the zero of time is arbitrary, so a unitary transformation must describe time-evolution. This implies a hermitian generator, which is also a conserved quantity for all systems, and which we therefore identify as the Hamiltonian.

4.7 Spatial Translations

Another continuous symmetry many systems possess is translation invariance. Translating a system in space from $x \to x + a$ should conserve the norm $\langle\psi|\psi\rangle$. So by the same arguments as for time evolution, the translation operator $\hat{T}(a)$ is unitary,

$$\hat{T}(a) = \hat{T}(a)^{-1}$$

Likewise

$$\hat{T}(a+b) = \hat{T}(a)\hat{T}(b)$$

If we translate by $+a$, that's the same as transforming $\psi(x)$ as

$$\hat{T}(a)\psi(x) = \psi(x-a)$$

This gives us an expression for $\hat{T}(a)$ as follows.

Consider $\hat{T}(\delta a)$ for small δa

$$\begin{aligned}\hat{T}(\delta a)\psi(x) &= \psi(x-\delta a)\\ &= \psi(x) - \delta a\frac{d}{dx}\psi(x) + \mathcal{O}(\delta a^2)\end{aligned}$$

As before

$$\hat{T}(a) = \lim_{N\to\infty}\hat{T}\left(\frac{a}{N}\right)^N$$

so

$$\begin{aligned}\hat{T}(a) &= \lim_{N\to\infty}\left(\hat{1} - \frac{d}{dx}\cdot\frac{a}{N}\right)^N\\ &= \lim_{N\to\infty}\left(e^{-\frac{a}{N}\frac{d}{dx}}\right)^N\\ &= e^{-a\frac{d}{dx}}\end{aligned}$$

So we have

$$\begin{aligned}\hat{T}(a) &= \exp\left(-a\frac{d}{dx}\right)\\ &= \exp\left(-ia\left(-i\frac{d}{dx}\right)\right)\\ &= \exp\left(-ia\hat{G}\right)\end{aligned}$$

$\hat{G} = \hat{G}^\dagger$ is Hermitian (show this) and it is called the generator for translations. $\hat{G}$ looks an awful lot like the momentum operator $\hat{p} = -i\hbar d/dx$. The key difference here is that previously we *defined* the momentum operator to be $\hat{p} = -i\hbar d/dx$, so this was treated as a postulate of how we formulate quantum mechanics. Our aim here is to uncover a deeper reason *why* $\hat{p} = -i\hbar d/dx$.

If a system is translationally invariant, so $V(x)$ is a constant, which we can take to be zero $V(x) = 0$, i.e. $\hat{H} = \frac{1}{2m}\hat{p}^2$, the energy of a state is invariant under translations. So this means for the free-particle

$$[\hat{H}, \hat{T}(a)] = 0$$

This means $[\hat{H}, \hat{G}] = 0$ so $\hat{G}$ is conserved. Likewise $[\hat{G}, \hat{p}] = 0$. So $\hat{G}$ must be a function of $\hat{p}$

$$\hat{G} = f(\hat{p})$$

Classical mechanics tells us that if we transform to a new coordinate system where $x' = \gamma x$, momentum transforms to $\frac{1}{\gamma}p = p'$. Back in quantum mechanics, in the transformed coordinate system, the wavefunction is

$$\psi'(x') = \frac{1}{\gamma}\psi\left(\frac{x'}{\gamma}\right)$$

In the new frame translation is given by

$$\hat{T}'(a') = \exp\left(-ia'\left(-i\frac{d}{dx'}\right)\right)$$

And this acts on states in the new coordinate system as

$$\hat{T}'(a')\psi'(x') = \psi'(x' - a')$$

A translation by a in the old frame should be equivalent to a translation by γa in the new frame, so $\hat{T}'(\gamma a) = \hat{T}(a)$ and therefore $\hat{G}' = \frac{1}{\gamma}\hat{G}$.

We can see this directly using the chain rule

$$\hat{G}' = -i\frac{d}{dx'} = -i\frac{d}{d(\gamma x)} = \frac{1}{\gamma}\left(-i\frac{d}{dx}\right) = \frac{1}{\gamma}\hat{G}$$

So $\hat{G}$ transforms as $\hat{p}$ under spatial re-scaling and we conclude $\hat{G} = \frac{1}{\hbar}\hat{p}$, where $\hbar$ again appears because of our choice of units.

We have now shown two important examples of continuous symmetry operations that give rise to conserved quantities in certain systems. We conclude that the Hermitian generator of spatial translations is momentum, much as the Hamiltonian is the generator for time translations.

4.8 Symmetries

In general if there is a conserved quantity $\hat{Q}$ for a given system satisfying $[\hat{H}, \hat{Q}]$, then we can construct a continuous set of unitary operators $\hat{u}(\tau) = \exp\left(-i\tau\hat{Q}\right)$ that are called **symmetries** of the system or equivalently symmetries of the Hamiltonian. Later we will consider more complex symmetries such as rotations.

5 Position and momentum representations

Thus far we have primarily worked with quantum states in the position representation, where we work with wave functions and operators that are defined in terms of a position x. It turns out that we can also equivalently work with **momentum space** wave functions and operators, where states and observable operators are instead defined in terms of momentum. In order to translate between position and momentum representations it is useful to first consider the eigenstates of the Hermitian operators $\hat{x}$ and $\hat{p}$.

5.1 Position Eigenstates

Let us now consider $\hat{x}$. What are its eigenstates? We need:

$$\hat{x}|x_0\rangle = |x_0\rangle x_0$$

$\hat{x}$ is Hermitian so $|x_0\rangle$ must exist. Write $|x_0\rangle = \varphi_{x_0}(x)$

$$|x_0\rangle = \varphi_{x_0}(x) \quad \text{and} \quad x\varphi_{x_0}(x) = x_0\varphi_{x_0}(x)$$

Let's Fourier transform this to a function of the wave-vector k

$$\begin{aligned}\tilde{\varphi}_{x_0}(k) &= \int_{-\infty}^{+\infty} dx\, x\, e^{-ikx}\varphi_{x_0}(x) \\ +i\frac{\partial}{\partial k}\tilde{\varphi}_{x_0}(k) &= \int_{-\infty}^{\infty} dx\, e^{-ikx}\, x\, \varphi_{x_0}(x) \quad \text{by differentiation under the integral}\end{aligned}$$

Using this in the eigenvalue equation we find

$$\begin{aligned}\Rightarrow \quad -i\frac{\partial}{\partial k}\tilde{\varphi}_{x_0}(k) &= x_0\tilde{\varphi}_{x_0}(k) \\ \Rightarrow \quad \tilde{\varphi}_{x_0}(k) &= \mathcal{N}_{x_0}e^{-ikx_0}\end{aligned}$$

$$\begin{aligned}\varphi_{x_0}(x) &= \frac{1}{2\pi}\int_{-\infty}^{+\infty} dk\, \mathcal{N}_{x_0}\, e^{+ikx}\, e^{-ikx_0} \\ &= \mathcal{N}_{x_0}\,\delta(x - x_0)\end{aligned}$$

We pick $\mathcal{N}_{x_0}$ so that $\langle x_1|x_0\rangle = \delta(x_1 - x_0)$

$$\begin{aligned}\langle x_1|x_0\rangle = \int_{-\infty}^{\infty} dx\, \varphi_{x_1}(x)^*\varphi_{x_0}(x) &= \mathcal{N}_{x_0}^2\delta(x - x_0) \\ &= \delta(x_1 - x_0) \\ \Rightarrow \quad \mathcal{N}_{x_0} &= 1\end{aligned}$$

Why normalize in this way? The states $|x\rangle$ form a basis and therefore we require the resolution of the identity to hold

$$\hat{1} = \int_{-\infty}^{\infty} dx\, |x\rangle\langle x|$$

Note that before we had $\hat{1} = \sum_n |n\rangle\langle n|$, but now we have a continuum of eigenstates, so we need to replace the sum with an integral $\sum_n \to \int$ and $|n\rangle\langle n| \to |x\rangle\langle x|\,\mathrm{d}x$. This can be shown by considering $\hat{1}|x_0\rangle$

$$\begin{aligned}
\hat{1}|x_0\rangle &= \int dx\,|x\rangle\langle x|x_0\rangle \\
&= \int dx\,|x\rangle\delta(x-x_0) \\
&= \int dx\,|x_0\rangle\delta(x-x_0) \\
&= |x_0\rangle \int dx\,\delta(x-x_0) \\
&= |x_0\rangle
\end{aligned}$$

We can also write

$$\begin{aligned}
\langle x|\psi\rangle &= \int dx'\,\varphi_{x_0}(x')\,\psi(x') \\
&= \int dx\,\delta(x-x')\psi(x') \\
&= \psi(x)
\end{aligned}$$

This might seem confusing because $\langle x|\psi\rangle$ is an amplitude like $\langle n|\psi\rangle$ i.e. just a complex number but $\psi(x)$ is a function...

Strictly speaking we shouldn't write $|\psi\rangle = \psi(x)$ but $|\psi\rangle = \psi$ where $\psi : \mathbb{R} \to \mathbb{C}$ is a function. ψ is a rule for mapping values in $\mathbb{R}$ onto values in $\mathbb{C}$. $\psi(x)$ is the value the rule produces at x. $|\psi\rangle$ can be be constructed in any basis $\langle n|\psi\rangle$, but it is often most convenient to use the position basis. The choice of basis just reflects how we choose to extract information from the state, but the state exists regardless of the choice of basis. Don't worry if this seems confusing; just avoid mixing $|\psi\rangle$ and $\psi(x)$ notation wherever possible, especially because we may need to deal with systems where there are more coordinates or discrete configuration spaces to worry about.

5.2 Momentum Eigenstates

An alternative representation to the position representation is the momentum representation. In order to construct this we need states $|p\rangle$ that are eigenstates of the momentum operator $\hat{p}$ such that

$$\hat{p}|p\rangle = |p\rangle p$$

Let's use position states to solve this, just to get used to how they behave. (We can always do it the usual way as well - set $\frac{d}{dx}\psi_p(x) = p\psi_p(x) \Rightarrow \psi_p(x) = \mathcal{N}e^{+ipx/\hbar}$.)

$$\langle x|\hat{p}|p\rangle = \langle x|p\rangle p$$

We know $|p\rangle$ is also an eigenstate of the translation operator,

$$\hat{T}(a)\,|p\rangle = e^{-iap/\hbar}\,|p\rangle\,.$$

When the translation operator acts on the position eigenstate $|x\rangle$ we get

$$\hat{T}(a)\,|x\rangle = |x+a\rangle$$

so we have

$$\langle x| = \langle x=0|\,\hat{T}(x)^\dagger = \langle x=0|\,\hat{T}(-x)$$

and therefore

$$\begin{aligned}\langle x|p\rangle &= \langle x=0|\hat{T}(-x)|p\rangle \\ &= \langle x=0|p\rangle\, e^{+ixp/\hbar}\end{aligned}$$

where $\langle x=0|p\rangle = \varphi_p(0)$ is determined by normalization, so we call it $\mathcal{N}$. Without solving a single differential equation we now know that

$$\varphi_p(x) = \langle x|p\rangle = \mathcal{N}_p e^{+ixp/\hbar}.$$

We use the normalization condition that $\hat{1} = \int \mathrm{d}p\, |p\rangle\langle p|$ much like for the position eigenstates, so $\langle p|p'\rangle = \delta(p-p')$

$$\begin{aligned}\langle p|p'\rangle = \delta(p-p') &= \mathcal{N}_p\mathcal{N}_{p'} \int dx\, e^{-ipx/\hbar} e^{+ip'x/\hbar} \\ &= \mathcal{N}_p\mathcal{N}_{p'}\hbar \int dx e^{+i(p'-p)x/\hbar} \\ &= \mathcal{N}_p^2 \hbar 2\pi\delta(p-p')\end{aligned}$$

So $\mathcal{N}_p = \frac{1}{\sqrt{2\pi\hbar}}$, and overall

$$\varphi_p(x) = \langle x|p\rangle = \frac{1}{\sqrt{2\pi\hbar}} e^{+ixp/\hbar}.$$

5.3 Momentum and position representations

From our knowledge linear algebra, we know we can write quantum states in terms of a given basis $|n\rangle$ using a resolution of the identity as

$$|\psi\rangle = \sum_n c_n\, |n\rangle$$

with $c_n = \langle n|\psi\rangle$. Likewise we can write matrix representations of operators in this basis as

$$\hat{A} = \sum_{n,m} |n\rangle\langle m|\, A_{n,m}$$

with $A_{n,m} = \langle n|\hat{A}|m\rangle$.

We can do exactly the same thing for position and momentum states

$$|\psi\rangle = \int \mathrm{d}x\, |x\rangle\, \psi(x)$$

$$|\psi\rangle = \int \mathrm{d}p\, |p\rangle\, \psi_p(p)$$

where **position space wave function** is $\psi(x) = \langle x|\psi\rangle$ and the **momentum space wave function** is $\psi_p(p) = \langle p|\psi\rangle$.

From these we can obtain the probability density for position

$$\rho(x) = \langle\psi|x\rangle\langle x|\psi\rangle = |\psi(x)|^2$$

and the probability density for momentum

$$\rho_p(p) = \langle\psi|p\rangle\langle p|\psi\rangle = |\psi_p(p)|^2.$$

This means observables like $\langle p^n \rangle$ are easy to obtain if we know $\psi_p(p)$

$$\langle p^n \rangle = \int \mathrm{d}p\, p^n |\psi_p(p)|^2.$$

Likewise we can construct momentum and position space represenations of operators. In position space we have

$$\hat{A} = \int \mathrm{d}x \int \mathrm{d}x'\, A(x,x')\, |x\rangle\langle x'|$$

with $A(x,x') = \langle x|\hat{A}|x'\rangle$ and in momentum space

$$\hat{A} = \int \mathrm{d}p \int \mathrm{d}p'\, A_p(p,p')\, |p\rangle\langle p'|$$

with $A_p(p,p') = \langle p|\hat{A}|p\rangle$

Example 1: The momentum operator can be written in position space as

$$\hat{p} = \int \mathrm{d}x \int \mathrm{d}x'\, i\hbar\delta'(x-x')\, |x\rangle\langle x'|$$

where $\delta'(x)$ is the derivative of the delta-function.

In momentum space this becomes

$$\hat{p} = \int \mathrm{d}p \int \mathrm{d}p'\, p\delta(p-p')\, |p\rangle\langle p'| = \int \mathrm{d}p\, p\, |p\rangle\langle p|\,.$$

Example 2: The translation operator $\hat{T}(a) = e^{-i\hat{p}a/\hbar}$ can be found in the momentum representation trivially by noting that $\hat{T}(a)$ is just a function of the momentum operator, so it is diagonal in the momentum eigenbasis

$$\hat{T}(a) = \int \mathrm{d}p \int \mathrm{d}p'\, e^{-ipa/\hbar}\delta(p-p')\, |p\rangle\langle p'| = \int \mathrm{d}p\, e^{-ipa/\hbar}\, |p\rangle\langle p|\,.$$

Using this we find in the position representation

$$\langle x|\hat{T}(a)|x'\rangle = \int \mathrm{d}p\, \langle x|p\rangle\, \langle p|x'\rangle\, e^{ipa/\hbar} = \frac{1}{2\pi\hbar}\int \mathrm{d}p\, e^{ipx/\hbar}e^{-ipx'/\hbar}e^{-ipa/\hbar} = \delta(x-(x'+a))$$

so the full operator can be written in the position represenation as

$$\hat{T}(a) = \int \mathrm{d}x \int \mathrm{d}x'\, |x\rangle\langle x'|\, \delta(x-(x'+a))$$

evaluating the integral over x' we find

$$\hat{T}(a) = \int \mathrm{d}x\, |x\rangle\langle x-a| = \int \mathrm{d}x\, |x+a\rangle\langle x|\,.$$

6 The Harmonic Oscillator

The harmonic oscillator serves as a useful approximate model for many complex systems we encounter in chemistry, e.g. vibrations of molecules, motion of atoms in solids, and even the electromagnetic field. Perhaps more importantly, the tools developed to find the eigenstates of the harmonic oscillator form the basis for understanding angular momentum, the structure of the hydrogen atom, and these tools also form the foundations of the *second quantisation*, which is the framework for tackling much more complicated many-body problems in quantum mechanics of molecules and condensed matter.

The Hamiltonian for the harmonic oscillator is

$$\hat{H} = \frac{1}{2m}\hat{p}^2 + \frac{1}{2}k\hat{x}^2$$

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

We will find the eigenstates of the harmonic oscillator solely using this property, and our toolkit from linear algebra and the fundamental postulates of quantum mechanics.

6.1 Creation and Annihilation Operators

Let us write the classical HO energy as

$$E = \hbar\omega a^* a$$

with

$$a = i\sqrt{\frac{1}{2\hbar\omega m}}p + \sqrt{\frac{k}{2\hbar\omega}}x \quad \text{and} \quad k = m\omega^2$$

Now let's write down the QM version which we will call the **ladder** operators (the reason for this name will become clear soon)

$$\hat{a} = \sqrt{\frac{k}{2\hbar\omega}}\hat{x} + i\sqrt{\frac{1}{2\hbar\omega m}}\hat{p}$$
$$\hat{a}^\dagger = \sqrt{\frac{k}{2\hbar\omega}}\hat{x} - i\sqrt{\frac{1}{2\hbar\omega m}}\hat{p}$$

And let's evaluate $\hbar\omega\hat{a}^\dagger\hat{a}$

$$\begin{aligned}
\hbar\omega\hat{a}^\dagger\hat{a} &= \hbar\omega\left[\frac{1}{2\hbar\omega m}\hat{p}^2 + \frac{k}{2\hbar\omega}\hat{x}^2 + i\sqrt{\frac{k}{m}}\frac{\hat{x}\hat{p} - \hat{p}\hat{x}}{2\hbar\omega}\right] \\
&= \frac{1}{2m}\hat{p}^2 + \frac{k}{2}\hat{x}^2 + \frac{i\omega}{2}\sqrt{\frac{k}{m}}[\hat{x}, \hat{p}] \\
&= \frac{1}{2m}\hat{p}^2 + \frac{k}{2}\hat{x}^2 + \frac{i\omega}{2} \\
&= \frac{1}{2m}\hat{p}^2 + \frac{k}{2}\hat{x}^2 - \frac{\hbar\omega}{2} \\
&= \hat{H} - \frac{\hbar\omega}{2}
\end{aligned}$$

Likewise $\hbar\omega\hat{a}\hat{a}^\dagger = \hat{H} + \hbar\omega/2$

So we can write the harmonic oscillator Hamiltonian in terms of these operators as

$$\begin{aligned}
\hat{H} &= \hbar\omega\left(\hat{a}^\dagger\hat{a} + \frac{1}{2}\right) \\
&= \hbar\omega\left(\hat{a}\hat{a}^\dagger - \frac{1}{2}\right)
\end{aligned}$$

This also shows $[\hat{a}, \hat{a}^\dagger] = 1$ but let's find this explicitly, just for fun

$$\begin{aligned}
[\hat{a}, \hat{a}^\dagger] &= \left[\sqrt{\frac{k}{2\hbar\omega}}\hat{x} + i\sqrt{\frac{1}{2\hbar\omega m}}\hat{p},\ \sqrt{\frac{k}{2\hbar\omega}}\hat{x} - i\sqrt{\frac{1}{2\hbar\omega m}}\hat{p}\right] \\
&= -\frac{i}{2}\sqrt{\frac{k}{m\hbar\omega}}\left([\hat{x}, \hat{p}] - [\hat{p}, \hat{x}]\right) \\
&= -\frac{i}{2\hbar}\left([\hat{x}, \hat{p}] + [\hat{x}, \hat{p}]\right) = -\frac{i}{2\hbar}(2i\hbar) = 1
\end{aligned}$$

6.2 Ladder operators and energy eigenstates

Now suppose $\hat{H}\,|E_n\rangle = E_n\,|E_n\rangle$ and let's consider

$$\begin{aligned}
\hat{H}\hat{a}\,|E_n\rangle &= \hbar\omega(\hat{a}^\dagger\hat{a} + \frac{1}{2})\hat{a}\,|E_n\rangle \\
&= \hbar\omega(\hat{a}^\dagger\hat{a}\hat{a} - [\hat{a}, \hat{a}^\dagger] + \frac{1}{2})\hat{a}\,|E_n\rangle \\
&= \hbar\omega(\hat{a}^\dagger - 1 + \frac{1}{2})\hat{a}\,|E_n\rangle \\
&= \hat{a}(\hbar\omega(\hat{a}^\dagger + \frac{1}{2}) - \hbar\omega)\,|E_n\rangle \\
&= \hat{a}(\hat{H} - \hbar\omega)\,|E_n\rangle \\
&= \hat{a}(E_n - \hbar\omega)\,|E_n\rangle \\
&= (E_n - \hbar\omega)\hat{a}\,|E_n\rangle
\end{aligned}$$

This shows that $\hat{a}\,|E_n\rangle$ is also an eigenstate of $\hat{H}$ with eigenvalues $E_n - \hbar\omega$.

Likewise we find

$$\hat{H}\hat{a}^\dagger\,|E_n\rangle = (E_n + \hbar\omega)\hat{a}^\dagger\,|E_n\rangle$$

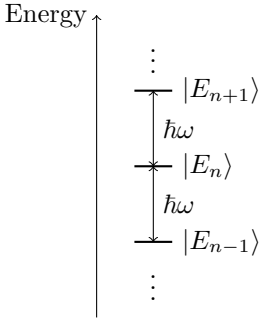
so $\hat{a}^\dagger\,|E_n\rangle$ is also an eigenstate of $\hat{H}$ with eigenvalues $E_n + \hbar\omega$.

So overall we have shown that

$$\begin{aligned}
\hat{a}\,|E_n\rangle &\propto |E_n - \hbar\omega\rangle \\
\hat{a}^\dagger\,|E_n\rangle &\propto |E_n + \hbar\omega\rangle
\end{aligned}$$

This means $\hat{a}$ and $\hat{a}^\dagger$ generate a ladder of states with

$$E_n - E_{n-1} = \hbar\omega :$$



This is why we call these operators ladder operators. We will also sometimes call $\hat{a}^\dagger$ the **raising** operator and $\hat{a}$ the **lowering** operator. Equivalently, and somewhat more dramatically, these are called the creation ($\hat{a}^\dagger$)

and annihilation ($\hat{a}$) operators (this naming convention comes from a connection to the second quantisation which we will explore in the second part of this course).

We will now deduce the absolute values of E_n.

Recall that $\hat{H} = \hbar\omega(\hat{a}^\dagger\hat{a} + \frac{1}{2})$. Assuming $|E_n\rangle$ is normalised, we have

$$E_n = \langle E_n|\hat{H}|E_n\rangle = \hbar\omega\left(\langle E_n|\hat{a}^\dagger\hat{a}|E_n\rangle + \frac{1}{2}\langle E_n|E_n\rangle\right) > 0$$

So this means the ladder of states must terminate at some $|E_0\rangle$ with a minimum $E_0 > 0$ i.e.

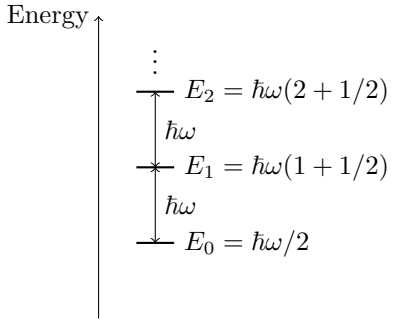
$$\hat{a}\,|E_0\rangle = 0$$

For this state we have (assume $\|\,|E_0\rangle\,\|^2 = 1$)

$$\begin{aligned} E_0 &= \langle E_0|\hat{H}|E_0\rangle \\ &= \hbar\omega\cancelto{0}{\langle E_0|\hat{a}^\dagger\hat{a}|E_0\rangle} + \frac{\hbar\omega}{2}\langle E_0|E_0\rangle \\ &= \frac{\hbar\omega}{2} \end{aligned}$$

Bound states in 1D are unique (as we will show below for the harmonic oscillator), so there is one unique ladder of states with

$$\begin{aligned} E_0 &= \frac{\hbar\omega}{2} & \hat{a}\,|E_0\rangle &= 0 \\ E_1 &= (1+\frac{1}{2})\hbar\omega & |E_1\rangle &\propto \hat{a}^\dagger\,|E_0\rangle \\ E_2 &= (2+\frac{1}{2})\hbar\omega & |E_2\rangle &\propto (\hat{a}^\dagger)^2\,|E_0\rangle \\ &\vdots & &\vdots \end{aligned}$$



6.3 Matrix elements of $\hat{a}$ and $\hat{a}^\dagger$

It turns out to be very useful to find the c_n^+ and c_n^- from before to fully define $|E_n\rangle$

$$\hat{a}\,|E_n\rangle = |E_{n-1}\rangle\,c_n^-$$

and

$$\begin{aligned} E_n &= \langle E_n|\hbar\omega(\hat{a}^\dagger\hat{a} + \frac{1}{2})|E_n\rangle \\ &= \hbar\omega(n + \frac{1}{2}) \end{aligned}$$

Also

$$E_n = \hbar\omega\langle E_{n-1}|E_{n-1}\rangle|c_n^-|^2 + \frac{1}{2}\langle E_n|E_n\rangle) = \hbar\omega(|c_n|^2 + \frac{1}{2})$$
$$\Rightarrow |c_n|^2 = n$$

Likewise

$$\begin{aligned} E_n &= \langle E_n|\hbar\omega(\hat{a}\hat{a}^\dagger - \frac{1}{2})|E_n\rangle \\ &= \hbar\omega(|c_n^+|^2 - \frac{1}{2}) \\ &= \hbar\omega(n + \frac{1}{2}) \end{aligned}$$

$|c_n^+|^2 = n + 1$

We can choose the phase of $|E_n\rangle$ so that $|c_n^+| = c_n^+$

So $c_n^+ = \sqrt{n+1}$ and $c_n^- = \sqrt{n}$

So we can now find all the matrix elements of $\hat{a}$ and $\hat{a}^\dagger$.

We now write $|n\rangle \equiv |E_n\rangle$

$$\langle n|\hat{a}|m\rangle = \langle n|\sqrt{m}|m+1\rangle = \sqrt{m}\delta_{n,m+1} = \sqrt{n-1}\delta_{n-1,m}$$
$$\langle n|\hat{a}^\dagger|m\rangle = \langle m|\hat{a}|n\rangle^* = \sqrt{n}\delta_{n,m+1} = \sqrt{m+1}\delta_{n,m+1}$$

And we can now find matrix elements of $\hat{x}$ and $\hat{p}$ as using

$$\hat{x} = \frac{1}{2}\sqrt{\frac{2\hbar\omega}{k}}(\hat{a} + \hat{a}^\dagger)$$
$$\hat{p} = \frac{1}{2i}\sqrt{2\hbar\omega m}(\hat{a} - \hat{a}^\dagger)$$

We can immediately deduce that the matrix elements for $\hat{x}$ and $\hat{p}$ are

$$\langle n|\hat{x}|m\rangle = \frac{1}{2}\sqrt{\frac{2\hbar\omega}{k}}(\langle n|\hat{a}|m\rangle + \langle n|\hat{a}^\dagger|m\rangle)$$
$$\langle n|\hat{p}|m\rangle = \frac{1}{2i}\sqrt{2\hbar\omega m}(\langle n|\hat{a}|m\rangle - \langle n|\hat{a}^\dagger|m\rangle)$$

One immediate consequence is:

$$\langle n|\hat{x}|n\rangle = 0 \quad \text{and} \quad \langle n|\hat{p}|n\rangle = 0$$

So eigenstates of the harmonic oscillator have an average position and momentum of zero.

We also see that $\langle m|\hat{x}|n\rangle$ is non-zero only if $m = n \pm 1$. This is exactly the selection rule for vibrational spectroscopy that you've probably seen before. The vibrational quantum number $n \equiv \nu$ can only change by 1 when infrared light is absorbed.

6.4 Constructing position space wave functions

We can find $\psi_0(x) = \langle x|0\rangle$ by noting

$$\hat{a}\psi_0(x) = 0$$

$$\Rightarrow \left(\sqrt{\frac{k}{2\hbar\omega}}x + i\sqrt{\frac{1}{2\hbar\omega m}}(-i\hbar)\frac{d}{dx}\right)\psi_0(x) = 0$$

$$\Rightarrow x\psi_0(x) + \overbrace{\frac{\hbar}{\sqrt{km}}}^{b^2}\frac{d}{dx}\psi_0(x) = 0$$

$$\Rightarrow \frac{d}{dx}\psi_0(x) = -\frac{x}{b^2}\psi_0(x)$$

$$\Rightarrow \frac{d}{dx}\ln\psi_0(x) = -\frac{x}{b^2}$$

$$\Rightarrow \int_{x_0}^{x}\ln\psi_0(x) = -\frac{1}{2}\left(\frac{x}{b}\right)^2$$

$$\Rightarrow \psi_0(x) = \psi_0(0)e^{-\frac{1}{2}(x/b)^2}$$

We can find $\psi_0(0)$ by normalising this state. Note that $\int_{-\infty}^{+\infty} dx|\psi(x)|^2 = 1$ and $\int_{-\infty}^{+\infty} e^{-ax^2}dx = \sqrt{\frac{\pi}{a}}$ so

$$\Rightarrow 1 = |\psi_0(0)|^2\int_{-\infty}^{+\infty} e^{-\frac{x^2}{b^2}}$$

$$\Rightarrow 1 = \sqrt{\frac{\pi}{(1/b^2)}}|\psi_0(0)|^2$$

$$\Rightarrow \psi_0(0) = \left(\frac{1}{\pi b^2}\right)^{1/4}$$

This gives the full normalized $\psi_0(x) = \left(\frac{b^2}{\pi}\right)^{1/4} e^{-\frac{1}{2}(x/b)^2}$

We can produce all $\psi_n(x)$ using

$$\psi_n(x) = \frac{1}{\prod_{m=1}^{n} c_m^+}(\hat{a}^\dagger)^n\psi_0(x)\cdot$$

$$= \frac{1}{\sqrt{n!}}(\hat{a}^\dagger)^n\psi_0(x)$$

Each factor of $\hat{a}^\dagger$ introduces a new power of x in front of the Gaussian function $\psi_0(x)$.

Recalling that the parity operator $\hat{\Pi}\psi(x) = \psi(-x)$, the operator $\hat{a}^\dagger$ is odd, i.e. $\hat{\Pi}\hat{a}^\dagger\hat{\Pi} = -\hat{a}^\dagger$, so each factor of $\hat{a}^\dagger$ also changes the overall parity of the HO eigenfunctions.

So we conclude that $\psi_n(x) = N_n H_n(x)\psi_0(x)$ where $H_n(x)$ is some polynomial in x of order n which is odd for odd n and even for even n.

Now we know the eigenstates for the HO we can consider dynamics of the HO next.

6.5 Basic Dynamics of the Harmonic Oscillator

Let's find general expressions for the time-dependent expectation values of position and momentum for the quantum harmonic oscillator.

Recall Ehrenfest theorem

$$\frac{d}{dt}\langle A\rangle = \frac{i}{\hbar}\langle[H, A]\rangle$$

Also recall that for $\hat{A} = \hat{x}$

$$\frac{d}{dt}\langle x\rangle = \frac{1}{m}\langle p\rangle$$

and $\hat{A} = \hat{p}$

$$\frac{d}{dt}\langle p\rangle = -\left\langle \frac{dV}{dx}\right\rangle$$

For the H.O. $\frac{dV}{dx} = kx$ so

$$\frac{d}{dt}\langle p\rangle = -k\langle x\rangle$$

Let's differentiate $\frac{d}{dt}\langle x\rangle$ again

$$\begin{aligned}\frac{d^2}{dt^2}\langle x\rangle &= \frac{d}{dt}\left(\frac{1}{m}\langle p\rangle\right)\\ &= \frac{1}{m}\frac{d}{dt}\langle p\rangle\\ &= -\frac{k}{m}\langle x\rangle\end{aligned}$$

Recall also that $\omega = \sqrt{\frac{k}{m}}$ so

$$\frac{d^2}{dt^2}\langle x\rangle = -\omega^2\langle x\rangle$$

We solve this with

$$\langle x\rangle = A\sin(\omega t) + B\cos(\omega t)$$

At $t = 0$, $\langle x(0)\rangle = B$ and

$$\begin{aligned}\langle p\rangle &= m\frac{d\langle x\rangle}{dt}\\ &= m(A\omega\cos(\omega t) - B\omega\sin(\omega t))\end{aligned}$$

So $\langle p(0)\rangle = m\omega A$

$$\begin{aligned}\langle x(t)\rangle &= \langle x(0)\rangle\cos(\omega t) + \frac{1}{m\omega}\langle p(0)\rangle\sin(\omega t)\\ \langle p(t)\rangle &= \langle p(0)\rangle\cos(\omega t) - m\omega\langle x(0)\rangle\sin(\omega t)\end{aligned}$$

This is just the classical equation of motion! (This is not true for general $V(x)$ though!)

6.6 Heisenberg equation of motion for the harmonic oscillator

An alternative approach is to consider the equation of motion for an operator $\hat{A}(t)$

$$\hat{A}(t) = e^{i\hat{H}t/\hbar}\hat{A}e^{-i\hat{H}t/\hbar}$$

Note that we can use the unitarity of the time-evolution operator to write this in a couple of equivalent forms

$$\begin{aligned}\frac{d}{dt}\hat{A}(t) &= \frac{i\hat{H}}{\hbar}e^{i\hat{H}t/\hbar}\hat{A}e^{-i\hat{H}t/\hbar} - \frac{i}{\hbar}e^{i\hat{H}t/\hbar}\hat{A}e^{-i\hat{H}t/\hbar}\hat{H}\\ &= \frac{i}{\hbar}[\hat{H}, \hat{A}(t)] = \frac{i}{\hbar}e^{i\hat{H}t/\hbar}[\hat{H}, \hat{A}]e^{-i\hat{H}t/\hbar}\end{aligned}$$

Now that us consider the dynamics of the annihilation operator. Set $\hat{A} = \hat{a}$ and we find

$$\frac{d}{dt}\hat{a}(t) = \frac{i}{\hbar}[\hat{H}, \hat{a}(t)]$$

$$= \frac{i}{\hbar} e^{i\hat{H}t/\hbar} \left(\hbar\omega(\hat{a}^\dagger \hat{a} + \frac{1}{2}), \hat{a} \right) e^{-i\hat{H}t/\hbar}$$
$$= i\omega e^{i\hat{H}t/\hbar} [\hat{a}^\dagger \hat{a}, \hat{a}] e^{-i\hat{H}t/\hbar}$$

Recall $[\hat{A}\hat{B}, \hat{C}] = \hat{A}[\hat{B}, \hat{C}] + [\hat{A}, \hat{C}]\hat{B}$

$$[\hat{a}^\dagger \hat{a}, \hat{a}] = \hat{a}^\dagger [\hat{a}, \hat{a}] + [\hat{a}^\dagger, \hat{a}]\hat{a}$$
$$= -\hat{a}$$

So:

$$\frac{d}{dt}\hat{a}(t) = i\omega \hat{U}(t)^\dagger (-\hat{a}) \hat{U}(t) = -i\omega \hat{a}(t)$$

In order to solve this we can consider the matrix elements $\langle n|\hat{a}(t)|m\rangle$.

$$\Rightarrow \tfrac{d}{dt}\langle n|\hat{a}(t)|m\rangle = -i\omega \langle n|\hat{a}(t)|m\rangle$$
$$\Rightarrow \langle n|\hat{a}(t)|m\rangle = e^{-i\omega t}\langle n|\hat{a}|m\rangle$$
$$\Rightarrow \hat{a}(t) = \textstyle\sum_n |n\rangle\langle n|\hat{a} \sum_m |m\rangle\langle m| e^{-i\omega t}$$
$$= e^{-i\omega t}\hat{a}$$

Also recall: $\hat{a}^\dagger(t) = \hat{a}(t)^\dagger = e^{+i\omega t}\hat{a}^\dagger$ and

$$\hat{x} = \frac{1}{2}\sqrt{\frac{2\hbar\omega}{k}}(\hat{a} + \hat{a}^\dagger)$$
$$\hat{p} = \frac{1}{2i}\sqrt{2\hbar\omega m}(\hat{a} - \hat{a}^\dagger)$$

So by taking the expectation value of the time-evolved $\hat{a}$ operator (and $\hat{a}^\dagger(t)$) we can find the time evolution of $\langle x(t)\rangle$ and $\langle p(t)\rangle$ directly. First note that

$$\langle a(t)\rangle = \langle a(0)\rangle e^{-i\omega t}$$
$$= \left(\sqrt{\frac{k}{2\hbar\omega}}\langle x(0)\rangle + i\sqrt{\frac{1}{2\hbar\omega m}}\langle p(0)\rangle \right) e^{-i\omega t}$$

But also $\langle a(t)\rangle = \left(\sqrt{\frac{k}{2\hbar\omega}}\langle x(t)\rangle + i\sqrt{\frac{1}{2\hbar\omega m}}\langle p(t)\rangle \right)$.

After some simple algebra we find this agrees with what we obtained from Ehrenfest theorem.

6.7 Coherent states

We will now explore the properties of a special class of states related to the Harmonic oscillator called coherent states.

The coherent state of the HO is denoted $|z\rangle$, where $z \in \mathbb{C}$. It is an eigenstate of $\hat{a}$:

$$\hat{a}\,|z\rangle = |z\rangle\, z \quad \text{and} \quad \langle z|z\rangle = 1$$

If this is true then

$$\begin{aligned}\langle x\rangle = \langle z|\,\hat{x}\,|z\rangle &= \frac{1}{2}\sqrt{\frac{2\hbar}{m\omega}}\left(\langle z|\,\hat{a}\,|z\rangle + \langle z|\,\hat{a}^\dagger\,|z\rangle\right)\\ &= \frac{1}{2}\sqrt{\frac{2\hbar}{m}}\left(\langle z|\,\hat{a}\,|z\rangle + \langle z|\,\hat{a}^*\,|z\rangle\right)\\ &= \frac{1}{2}\sqrt{\frac{2\hbar}{m\omega}}\left(z + z^*\right)\\ &= \sqrt{\frac{2\hbar}{m\omega}}\mathrm{Re}[z]\end{aligned}$$

Likewise for $\langle p\rangle$:

$$\langle p|p\rangle = \langle z|\,\hat{p}\,|z\rangle = \sqrt{2\hbar m\omega}\mathrm{Im}[z]$$

Let's find a ket $|z\rangle$ in the $|n\rangle$ basis:

$$\begin{aligned}\hat{a}\,|z\rangle &= \sum_{n=0}^{\infty} c_n\,|n\rangle\, z\\ \hat{a}\,|z\rangle &= \sum_{n=0}^{\infty} c_n\hat{a}\,|n\rangle\\ &= \sum_{n=0}^{\infty} c_n\sqrt{n}\,|n-1\rangle \quad n' = n-1, n = n'+1\\ &= \sum_{n=0}^{\infty} c_{n+1}\sqrt{n+1}\,|n'\rangle \quad \text{arbitrary index } n \text{ or } n'\end{aligned}$$

So $c_n = \sqrt{n+1}c_{n+1}$

$$\Rightarrow c_{n+1} = \frac{z}{\sqrt{n+1}}c_n$$

$$\Rightarrow c_n = \frac{z^n}{\sqrt{n!}}c_0$$

$$1 = \sum_{n=0}^{\infty}|c_n|^2 = \sum_{n=0}^{\infty}\frac{|z|^{2n}}{n!}c_0^2 = e^{|z|^2}c_0^2$$

$$c_0 = e^{-|z|^2/2}$$

So:

$$\begin{aligned}|z\rangle &= \sum_{n=0}^{\infty}|n\rangle\left(\frac{z^n}{\sqrt{n!}}e^{-|z|^2/2}\right)\\ &= \sum_{n=0}^{\infty}\frac{z^n}{\sqrt{n!}}\,|n\rangle\, e^{-|z|^2/2}\end{aligned}$$

And $\sqrt{n!}\,|n\rangle = (\hat{a}^\dagger)^n\,|0\rangle$. So we can also write this as:

$$\begin{aligned}|z\rangle &= e^{-|z|^2/2}\sum_{n=0}^{\infty}\frac{(z\hat{a}^\dagger)^n}{n!}\,|0\rangle\\ &= e^{-|z|^2/2}e^{z\hat{a}^\dagger}\,|0\rangle\end{aligned}$$

6.7.1 Time evolution

How do these states evolve? Consider $\hat{a}(t)\left|z\right\rangle z$:

$$e^{-i\hat{H}t/\hbar}\hat{a}\left|z\right\rangle = e^{-i\hat{H}t/\hbar}\left|z\right\rangle z$$

$$\Rightarrow e^{-i\hat{H}t/\hbar}\hat{a}e^{i\hat{H}t/\hbar}e^{-i\hat{H}t/\hbar}\hat{a}\left|z\right\rangle = \left|z(t)\right\rangle z$$
$$e^{i\omega t}\hat{a}\left\langle z(t)|z(t)\right\rangle = \left|z(t)\right\rangle z$$

So:

$$\hat{a}\left|z(t)\right\rangle = \left|z(t)\right\rangle ze^{-i\omega t}$$

which means $\left|z(t)\right\rangle$ is a coherent state with $z(t) = ze^{-i\omega t}$ up to a phase factor.

From this and $\hat{a}^\dagger(t) = e^{i\omega t}\hat{a}^\dagger$ we can show:

$$e^{-i\hat{H}t/\hbar}\hat{a}\left|z\right\rangle = \left|ze^{-i\omega t}\right\rangle e^{-i\omega t/2}$$

So coherent states evolve into other coherent states, with $\langle x\rangle$ and $\langle p\rangle$ obeying the classical equation of motion. In this sense coherent states evolve in as close a way to classical as possible.

6.7.2 Average energy

The energy of a coherent state is:

$$\begin{aligned}\left\langle z\right|\hat{H}\left|z\right\rangle &= \left\langle z\right|(\hat{a}^\dagger\hat{a} + \frac{1}{2})\left|z\right\rangle = \hbar\omega\left(\left\langle z\right|\hat{a}^\dagger\hat{a}\left|z\right\rangle + \frac{1}{2}\right)\\ &= \hbar\omega\left(\left\langle z\right|\hat{z}^* z\left|z\right\rangle + \frac{1}{2}\right)\\ &= \hbar\omega\left(\left|z\right|^2 + \frac{1}{2}\right)\\ &= \frac{\left\langle p\right\rangle^2}{2m} + \frac{1}{2}m\omega^2\left\langle x\right\rangle^2 + \frac{\hbar\omega}{2}\end{aligned}$$

As with the dynamics, we see that the energy is as close to classical as possible, with the minimal amount of zero-point energy present. As the energy of the coherent state increases, the zero-point energy term becomes negligible and these states behave essential as a perfectly classical particle would.

6.7.3 Coherent states, translations and boosts

Above we found that

$$\left|z\right\rangle = e^{-\left|z\right|^2/2}e^{z\hat{a}^\dagger}\left|0\right\rangle \tag{3}$$

We can also re-write this by noting $\hat{a}^n\left|0\right\rangle = 0$

$$\left|z\right\rangle = e^{-\left|z\right|^2/2}e^{z\hat{a}^\dagger}e^{-z^*\hat{a}}\left|0\right\rangle \tag{4}$$

We can then use a result called the Baker-Campell-Hausdorff formula to obtain

$$e^{\hat{A}+\hat{B}} = e^{\hat{A}}e^{\hat{B}}e^{-\frac{1}{2}[\hat{A},\hat{B}]}\ldots$$

where other factors depend on commutators more than two $\hat{A}$ and $\hat{B}$ operators e.g. $[\hat{A},[\hat{A},\hat{B}]]$.

Using this we obtain

$$|z\rangle = e^{z\hat{a}^\dagger - z^*\hat{a}}\,|0\rangle \tag{5}$$

and we see that $\hat{D}(z) = e^{z\hat{a}^\dagger - z^*\hat{a}}$ is a unitary operator. This is called the **displacement operator**.

When z is purely real, this reduces to the translation operator $e^{-i\langle x\rangle\hat{p}/\hbar}$ and likewise when it is purely imaginary it reduces to the momentum boost operator $e^{i\langle p\rangle\hat{x}/\hbar}$. So we see that the displacement operator generalises the translation and momentum boost operators. In fact we can also write this as

$$\hat{D}(z) = e^{i\phi(z)}e^{-i\langle x\rangle\hat{p}/\hbar}e^{i\langle p\rangle\hat{x}/\hbar}$$

so the displacement operator both translates and boosts the momentum of a quantum state. For coherent states the reference state is simply the harmonic oscillator ground state.

7 Angular Momentum

Angular momentum is a vast topic covering:

- Spherical potentials, and atomic structure
- Molecular rotations
- Spectroscopy including NMR & EPR.

In this section we will cover the fundamentals of angular momentum that underpins the quantum description of these systems.

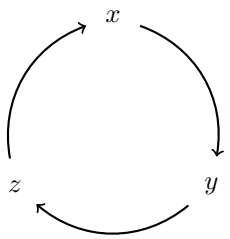
7.1 Defining orbital angular momentum

The classical orbital angular momentum for a particle is

$$\vec{\ell} = \vec{r} \times \vec{p}$$

It is a vector $\vec{q}$ with components

$$\begin{aligned} \ell_x &= yp_z - zp_y \\ \ell_y &= zp_x - xp_z \\ \ell_z &= xp_y - yp_x \end{aligned}$$



Note the cyclic symmetry: $x \to y \to z \to x$

This can be written concisely as

$$\ell_\alpha = \sum_{\beta\gamma} \varepsilon_{\alpha\beta\gamma} r_\beta p_\gamma$$

where $\alpha, \beta, \gamma \in \{x, y, z\}$ and $\varepsilon_{\alpha\beta\gamma}$ is the Levi-Civita symbol

$$\varepsilon_{xyz} = +1, \quad \varepsilon_{zxy} = -1 \quad \text{etc.}$$

It satisfies:

$$\varepsilon_{\alpha\alpha\beta} = 0, \quad \varepsilon_{\alpha\beta\alpha} = 0, \quad \varepsilon_{\beta\alpha\alpha} = 0$$

and

$$\begin{aligned} \varepsilon_{\alpha\beta\gamma} &= -\varepsilon_{\gamma\beta\alpha} \\ \varepsilon_{\alpha\beta\gamma} &= \varepsilon_{\beta\gamma\alpha} = \varepsilon_{\gamma\alpha\beta} \end{aligned}$$

We define the QM angular momentum operator as:

$$\hat{L} = \hat{r} \times \hat{p}$$

So

$$\hat{L}_\alpha = \sum_{\beta\gamma} \varepsilon_{\alpha\beta\gamma} \hat{r}_\beta \hat{p}_\gamma$$

i.e.

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

Note $[\hat{x}, \hat{p}_y] = 0$, $[\hat{z}, \hat{p}_x] = 0$ etc. So the order of $\hat{r}_\alpha$ and $\hat{p}_\alpha$ does not matter in the definition of $\hat{L}$.

In the exercises you will show using $[\hat{x}_\alpha, \hat{p}_\beta] = i\hbar$ that

$$[\hat{L}_\alpha, \hat{L}_\beta] = \sum_\gamma i\hbar\varepsilon_{\alpha\beta\gamma}\hat{L}_\gamma$$

i.e.

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

and

$$[\hat{L}^2, \hat{L}_\alpha] = 0$$

Consequences: $\hat{L}^2$ and $\hat{L}_\alpha$ can be well-defined simultaneously for a QM state, but not two different $\hat{L}_\alpha$ and $\hat{L}_\beta$.

7.2 Angular Momentum Eigenstates

From $[\hat{L}^2, \hat{L}_z] = 0$ we know a state $|l, m\rangle$ exists such that $|l, m\rangle$ is a simultaneous eigenstate of $\hat{L}^2$ & $\hat{L}_z$.

$$\hat{L}^2|l, m\rangle = f_l|l, m\rangle$$
$$\hat{L}_z|l, m\rangle = g_m|l, m\rangle$$

7.3 Ladder operators

As with the QHO we can define operators:

$$\hat{L}_+ = \hat{L}_x + i\hat{L}_y$$
$$\hat{L}_- = \hat{L}_x - i\hat{L}_y = \hat{L}_+^\dagger$$

And we find the following useful commutators:

$$\begin{aligned}[\hat{L}_+, \hat{L}_-] &= [\hat{L}_x + i\hat{L}_y, \hat{L}_x - i\hat{L}_y] \\ &= -i[\hat{L}_x, \hat{L}_y] + i[\hat{L}_y, \hat{L}_x] \\ &= -i(i\hbar\hat{L}_z) + i(-i\hbar\hat{L}_z) \\ &= 2\hbar\hat{L}_z\end{aligned}$$

$$\begin{aligned}[\hat{L}_+, \hat{L}_z] &= [\hat{L}_x + i\hat{L}_y, \hat{L}_z] \\ &= [\hat{L}_x, \hat{L}_z] + i[\hat{L}_y, \hat{L}_z] \\ &= -i\hbar\hat{L}_y + i \cdot i\hbar\hat{L}_x \\ &= -\hbar(\hat{L}_x + i\hat{L}_y) \\ &= -\hbar\hat{L}_+\end{aligned}$$

$$\begin{aligned}[\hat{L}_-, \hat{L}_z] &= -i\hbar\hat{L}_y - i(i\hbar\hat{L}_x) \\ &= \hbar\hat{L}_-\end{aligned}$$

$$\begin{aligned}[\hat{L}^2, \hat{L}_\pm] &= [\hat{L}^2, \hat{L}_x \pm i\hat{L}_y] \\ &= [\hat{L}^2, \hat{L}_x] \pm i[\hat{L}^2, \hat{L}_y] \\ &= 0\end{aligned}$$

So we can deduce that the states $|\ell, m\rangle$ behave as follows when $\hat{L}_\pm$ act on them:

$$\begin{aligned}\hat{L}^2\hat{L}_\pm|\ell, m\rangle &= \hat{L}_\pm\left[\hat{L}^2|\ell, m\rangle\right] \qquad \because [\hat{L^2}, \hat{L}_\pm] = 0 \\ &= \hat{L}_\pm\left(f_\ell|\ell, m\rangle\right) \\ &= f_\ell\left(\hat{L}_\pm|\ell, m\rangle\right)\end{aligned}$$

So $\hat{L}_\pm|\ell, m\rangle$ is also an eigenstate of $\hat{L}^2$.

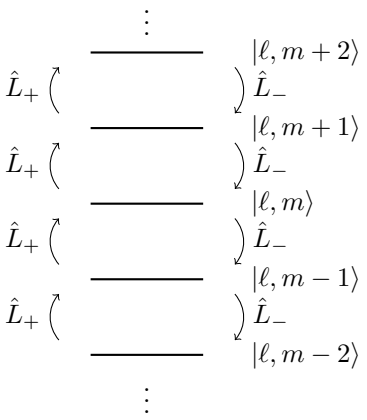
Also $\left[\hat{L}_z, \hat{L}_\pm\right] = \pm\hbar\hat{L}_\pm$, so

$$\begin{aligned}\hat{L}_z\hat{L}_\pm|\ell, m\rangle &= \left(\left[\hat{L}_z, \hat{L}_\pm\right] + \hat{L}_\pm\hat{L}_z\right)|\ell, m\rangle \\ &= \left(\pm\hbar\hat{L}_\pm + \hat{L}_\pm\hat{L}_z\right)|\ell, m\rangle \\ &= \left(\pm\hbar\hat{L}_\pm + \hat{L}_\pm g_m\right)|\ell, m\rangle \\ &= (g_m \pm \hbar)\left(\hat{L}_\pm|\ell, m\rangle\right)\end{aligned}$$

So $\hat{L}_\pm|\ell, m\rangle \propto |\ell, m \pm 1\rangle$ where $g_{m\pm1} = g_m \pm \hbar$.

(This looks familiar from the QHO $\hat{a}$ and $\hat{a}^\dagger$!)

$\hat{L}_+$ and $\hat{L}_-$ form a ladder of states which are eigenstates of $\hat{L}^2$ and $\hat{L}_z$ with $\hat{L}_z$ eigenvalues separated by $\pm\hbar$.



Note that the label "m" was arbitrary, so we can set

$$\hbar m = g_m$$

Question: Is this ladder infinite?

Answer: No! Let's see why...

We can write $\hat{L}^2$ as:

$$\begin{aligned}\hat{L}^2 &= \hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2 \\ &= \hat{L}_x^2 + \hat{L}_+\hat{L}_- - i\left[\hat{L}_x, \hat{L}_y\right] + i\left[\hat{L}_x, \hat{L}_y\right] \\ &= \hat{L}_z^2 + \hat{L}_+\hat{L}_- - \hbar\hat{L}_z\end{aligned}$$

And we also know

$$\begin{aligned}\langle \ell m | \hat{L}^2 | \ell m \rangle &= \langle \ell m | \left(\hat{L}_z^2 - \hat{L}_+\hat{L}_- + \hbar\hat{L}_z\right) | \ell m \rangle \\ &= \hbar^2(m^2 - m) + \langle \ell m | \hat{L}_-^\dagger \hat{L}_- | \ell m \rangle \\ &= f_\ell \geq 0\end{aligned}$$

So

$$\begin{aligned}f_\ell - \hbar^2 m(m-1) &= \langle \ell m | \hat{L}_-^\dagger \hat{L}_- | \ell m \rangle \\ &= \| \hat{L}_- | \ell m \rangle \|^2 \\ &\geq 0\end{aligned}$$

So $\hat{L}_- | \ell m \rangle$ must terminate for some $m = m_{\min} \leq 0$

Likewise

$$\hat{L}^2 = \hat{L}_z^2 + \hat{L}_-\hat{L}_+ + \hbar\hat{L}_z$$

and

$$f_\ell = \hbar^2 m(m+1) + \langle \ell m | \hat{L}_+^\dagger \hat{L}_+ | \ell m \rangle$$

$$\Rightarrow f_\ell - \hbar^2 m(m+1) \geq 0$$

For a given f_ℓ and $m \geq 0$, m cannot be increased arbitrarily, and satisfies this.

So we have

$$\begin{aligned}f_\ell - \hbar^2 m(m-1) &\geq 0 \\ f_\ell - \hbar^2 m(m+1) &\geq 0\end{aligned}$$

$\forall | \ell m \rangle$ that exist.

This means

$$\begin{aligned}\hat{L}_+ | \ell, m_{\max} \rangle &= 0 \\ \hat{L}_- | \ell, m_{\min} \rangle &= 0\end{aligned}$$

But also

$$\begin{aligned}f_\ell &= \hbar^2 m_{\max}(m_{\max}+1) + \| \hat{L}_+ | \ell, m_{\max} \rangle \|^2 \\ f_\ell &= \hbar^2 m_{\min}(m_{\min}-1) + \| \hat{L}_- | \ell, m_{\min} \rangle \|^2\end{aligned}$$

$$\Rightarrow m_{\min} = -m_{\max} \equiv \ell$$

and

$$f_\ell = \hbar^2 \ell(\ell+1)$$

And therefore for a given ℓ (i.e. f_ℓ), m takes values

$$m = -\ell, -\ell+1, \ldots, \ell-1, \ell$$

This implies either ℓ is an a positive integer, or half-integer.

For an orbital angular momentum only $\ell = 0, 1, 2, \ldots$ is allowed, we'll see why shortly. But half-integer ℓ is allowed by the angular momentum operator rules. This is true for intrinsic degrees of freedom, namely spin.

Let us conclude by finding the ladder operators' matrix elements. We have shown that:

$$\hat{L}_+ |\ell, m\rangle = c^+_{\ell m} |\ell, m+1\rangle$$

$$\hat{L}_- |\ell, m\rangle = c^-_{\ell m} |\ell, m-1\rangle$$

And also

$$f_\ell - \hbar^2 m(m \pm 1) = \| \hat{L}_\pm |\ell m\rangle \|^2$$

$$\Rightarrow \quad \hbar^2 \ell(\ell+1) - \hbar^2 m(m \pm 1) = |c^\pm_{\ell m}|^2$$

We can choose the phase of $|\ell m\rangle$ such that

$$c^+_{\ell m} = \hbar \sqrt{\ell(\ell+1) - m(m+1)}$$

And lastly

$$\begin{aligned} \langle \ell, m | \hat{L}_+ |\ell' m'\rangle &= c^+_{\ell\ell' m'} \delta_{\ell\ell'} \delta_{m,m'+1} \\ &= \langle \ell' m' | \hat{L}_- |\ell m\rangle^\dagger \\ &= c^*_{\ell' m'} \delta_{\ell\ell'} \delta_{m,m'+1} \end{aligned}$$

$$\Rightarrow \quad c^-_{\ell m} = \hbar \sqrt{\ell(\ell+1) - m'(m'-1)}$$

And lastly we can find matrix elements of $\hat{L}_x$ and $\hat{L}_y$

$$\begin{aligned} \langle \ell m | \hat{L}_x |\ell' m'\rangle &= \frac{1}{2} \left(\langle \ell m | \hat{L}_+ |\ell' m'\rangle + \langle \ell m | \hat{L}_- |\ell' m'\rangle \right) \\ &= \frac{\hbar}{2} \Big(\sqrt{\ell(\ell+1) - m'(m'+1)} \delta_{\ell\ell'} \delta_{m,m'+1} \\ &\quad + \sqrt{\ell(\ell+1) - m'(m'-1)} \delta_{\ell\ell'} \delta_{m,m'-1} \Big) \end{aligned}$$

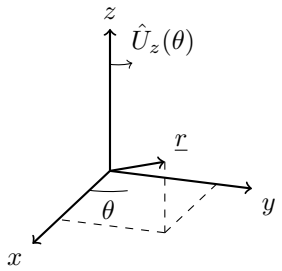
And for $\hat{L}_y$

$$\langle \ell m | \hat{L}_y | \ell' m' \rangle = \frac{\hbar}{2i} \Big(\sqrt{\ell(\ell+1) - m'(m'+1)} \delta_{\ell\ell'} \delta_{m,m'+1} - \sqrt{\ell(\ell+1) - m'(m'-1)} \delta_{\ell\ell'} \delta_{m,m'-1} \Big)$$

7.4 Angular momentum and rotations

Previously we saw continuous symmetry operations take the form of unitary operators on quantum states. Rotations about x, y, z axes must also have unitary representation with corresponding generators, i.e.

$$\hat{U}_x(\theta) = e^{-i\hat{J}_x\theta}$$
$$\hat{U}_y(\theta) = e^{-i\hat{J}_y\theta}$$
$$\hat{U}_z(\theta) = e^{-i\hat{J}_z\theta}$$



For a vector observable:

$$\langle \hat{\underline{A}} \rangle = \langle \psi | \hat{\underline{A}} | \psi \rangle$$

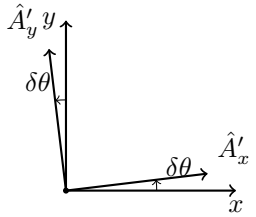
$$\langle \underline{\underline{R}}_\alpha \hat{\underline{A}} \rangle = \langle \psi | \hat{U}_\alpha^\dagger \hat{\underline{A}} \hat{U}_\alpha | \psi \rangle \quad \forall | \psi \rangle \in \mathcal{H}$$

So

$$\underline{\underline{R}}_\alpha \ \hat{\underline{A}} = \hat{U}_\alpha^\dagger \hat{\underline{A}} \hat{U}_\alpha$$

For a small rotation angle about z

$$(\underline{\underline{R}} \ \hat{\underline{A}})_z = \hat{A}_z$$
$$(\underline{\underline{R}} \ \hat{\underline{A}})_x = \hat{A}_x + \delta\theta \hat{A}_y + O(\delta\theta^2)$$
$$(\underline{\underline{R}} \ \hat{\underline{A}})_y = \hat{A}_y + \delta\theta \hat{A}_x + O(\delta\theta^2)$$



And for $\hat{U}_z(\delta\theta)$ we have

$$\hat{U}_z(\delta\theta) = \hat{1} - i\hat{J}_z\delta\theta + O(\delta\theta^2)$$

So

$$i[\hat{J}_z, \hat{A}_z] = 0 \quad \Rightarrow [\hat{A}_z, \hat{A}_z] = 0$$
$$i[\hat{J}_z, \hat{A}_x] = -\hat{A}_y \quad \Rightarrow [\hat{J}_z, \hat{A}_x] = i\hat{A}_y$$
$$i[\hat{J}_z, \hat{A}_y] = +\hat{A}_x \quad \Rightarrow [\hat{J}_z, \hat{A}_y] = i\hat{A}_x$$

And more generally

$$[\hat{J}_\alpha, \hat{A}_\beta] = i\sum_\gamma \epsilon_{\alpha\beta\gamma}\hat{A}_\gamma$$

An infinitesimal rotation about $\underline{n}$ is

$$\hat{U}_n(\delta\theta) = \hat{U}_x(n_x\delta\theta)\hat{U}_y(n_y\delta\theta)\hat{U}_z(n_z\delta\theta)$$

So

$$\hat{U}_n(\delta\theta) = \exp\Big(-i\underline{n}\cdot\underline{\hat{J}}\delta\theta\Big)$$

By the fundamental rules for how rotations compose, if we rotate by θ about α, and then $\delta\theta$ about $\underline{n}$, this is equivalent to rotating first by α and then by $\delta\theta$ about $\underline{\underline{R}}_\alpha^{-1}(\theta)\underline{n}$, i.e. we have

$$\hat{U}_n(\delta\theta)\hat{U}_\alpha(\theta) = \hat{U}_\alpha(\theta)\hat{U}_{n'}(\delta\theta)$$

$$\underline{n}' = \underline{\underline{R}}_\alpha^{-1}(\theta)\underline{n}$$

which is equivalent to $\underline{\hat{J}}$ transforming as

$$\hat{U}_\alpha(\theta)^\dagger\underline{\hat{J}}\hat{U}_\alpha(\theta) = \underline{\underline{R}}_\alpha(\theta)\underline{\hat{J}}$$

An alternative way of showing this is to consider a small rotation about an axis α, it is fairly striaghforward to show that

$$\underline{\underline{R}}_\alpha(\delta\theta_\alpha) = \underline{\underline{1}} + \delta\theta_\alpha\underline{\underline{e}}_\alpha^\times + \mathcal{O}(\delta\theta_\alpha^2)$$

where $\underline{\underline{v}}^\times\underline{u} = \underline{v}\times\underline{u}$, so $\underline{\underline{v}}^\times$ is an anti-symmetric matrix of the form

$$\underline{\underline{v}}^\times = \begin{pmatrix} 0 & -v_z & v_y \\ v_z & 0 & -v_x \\ -v_y & v_x & 0 \end{pmatrix} = -\sum_{\alpha\beta\gamma}\epsilon_{\alpha\beta\gamma}\underline{e}_\alpha\underline{e}_\beta^T v_\gamma$$

We can use the triple cross product rule $\underline{a}\times(\underline{b}\times\underline{c}) = \underline{b}\underline{a}^T\underline{c} - (\underline{a}^T\underline{b})\underline{c}$ to show that for x and y rotations

$$\underline{\underline{R}}_\alpha(\delta\theta_\alpha)\underline{\underline{R}}_\beta(\delta\theta_\alpha) = \underline{\underline{1}} + \delta\theta_\alpha\underline{\underline{e}}_\alpha^\times + \delta\theta_\beta\underline{\underline{e}}_\beta^\times + \delta\theta_\alpha\delta\theta_\beta(\underline{e}_\beta\underline{e}_\alpha^T - (\underline{e}_\alpha^T\underline{e}_\beta))$$

so

$$\underline{\underline{R}}_x(\delta\theta_x)\underline{\underline{R}}_y(\delta\theta_y) - \underline{\underline{R}}_y(\delta\theta_y)\underline{\underline{R}}_x(\delta\theta_x) = \delta\theta_x\delta\theta_y\underline{\underline{e}}_z^\times = \underline{\underline{R}}_z(\delta\theta_x\delta\theta_y) - \underline{\underline{1}}$$

The quantum rotation operators should obey this composition law so

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

we can repeat this argument for the other commutators to show that $\hat{J}_\alpha$ transforms as a vector.

This means

$$[\hat{J}_\alpha, \hat{J}_\beta] = i\sum_\gamma \epsilon_{\alpha\beta\gamma}\hat{J}_\gamma$$

So $\underline{\hat{J}}$ transforms as a vector under rotations as well.

The only operators satisfying this are angular momentum operators $\times\frac{1}{\hbar}$. So let us replace $\hat{J}_\alpha \to \hat{J}_\alpha/\hbar$ above.

Now let us define

$$\hat{J}^2\,|j,m\rangle = \hbar^2 j(j+1)\,|j,m\rangle$$
$$\hat{J}_z\,|j,m\rangle = \hbar m\,|j,m\rangle$$

For a given value of j, $\hat{1} = \sum_{m=-j}^{j} |j,m\rangle \langle j,m|$

Let's consider $\hat{U}_z(\theta)$ in the $|j,m\rangle$ basis

$$\hat{U}_z(\theta) = \hat{1} e^{-i\hat{J}_z\theta/\hbar} \hat{1}$$

$$= \sum_{m,m'} |j,m\rangle \langle j,m| \, e^{-i\hat{J}_z\theta/\hbar} \, |j,m'\rangle \langle j,m'|$$

$$= \sum_{m,m'} e^{-i\theta m'} \, |j,m\rangle \langle j,m'| \, \delta_{m,m'}$$

$$= \sum_{m=-j}^{j} e^{-i\theta m} \, |j,m\rangle \langle j,m|$$

Let $m = -j + n$ for $n = 0, 1, \ldots, 2j$

so we can write $\hat{U}_z(\theta)$ as

$$\hat{U}_z(\theta) = e^{+ij\theta} \sum_{n=0}^{2j} e^{-in\theta} \, |j,-j+n\rangle \langle j,-j+n|$$

If $\theta = 2\pi$ this is a 360° rotation. $\hat{U}_z(2\pi)$ is

$$\hat{U}_z(2\pi) = e^{2\pi ij\theta} \sum_{n=0}^{2j} e^{-2\pi n} \, |j,-j+n\rangle \langle j,-j+n|$$

$$= e^{2\pi ij} \hat{1}$$

So for integer j

$$\hat{U}_z(2\pi) = \hat{1}$$

But for half-integer j

$$\hat{U}_z(2\pi) = -\hat{1}$$

We note that for any state we can define

$$|\psi_j\rangle = \sum_{m=-j}^{j} |j,m\rangle \langle j,m|\psi\rangle \tag{6}$$

$$|\psi\rangle = \sum_{j \in \text{integer and half-integer}} |\psi_j\rangle \tag{7}$$

For position-space systems we need

$$\langle \underline{r}| \, \psi\rangle = \left\langle \underline{\underline{R}}_z(2\pi)\underline{r} \right| \psi\rangle$$
$$= \langle \underline{r}| \, \hat{U}_z^\dagger(2\pi) \, |\psi\rangle$$
$$= \sum_j \langle \underline{r}| \, \psi_j\rangle (-1)^{2j}$$

So in position space only j = integer is allowed.

If j = half-integer, the basic rules of rotations and angular momenta still apply, but in 3D position space this is forbidden, so j = half-integer is only allowed for other "intrinsic" degrees of freedom e.g. the electron spin or proton nuclear spin.

7.4.1 Euler-Rodriguez Formula

If we rotate about z then

$$\underline{\underline{R}}_z \underline{v} = \underline{e}_x(\cos(\theta)v_x - \sin(\theta)v_y) + \underline{e}_y(\cos(\theta)v_y + \sin(\theta)v_x) + v_z\underline{e}_z$$

which we can re-write as

$$\underline{\underline{R}}_z \underline{v} = \cos(\theta)(\underline{\underline{1}} - \underline{e}_z\underline{e}_z^T)\underline{v} + \sin(\theta)\underline{e}_z \times \underline{v} + \underline{e}_z\underline{e}_z^T\underline{v}$$

Since our choice of x, y and z axes are arbitrary, assuming they have the correct handedness, this should also apply to rotation about an arbitrary unit vector $\underline{n}$

$$\underline{\underline{R}}_n \underline{v} = \cos(\theta)(\underline{\underline{1}} - \underline{n}\underline{n}^T)\underline{v} + \sin(\theta)\underline{n} \times \underline{v} + \underline{n}\underline{n}^T\underline{v}$$

Taking the small θ limit of this gives the formula for small rotations we used above.

7.5 Coupled Angular Momenta

We previously saw that for an angular momentum operator $\hat{J}$ whose components satisfy:

$$[\hat{J}_\alpha, \hat{J}_\beta] = i\sum_\gamma \epsilon_{\alpha\beta\gamma}\hbar\hat{J}_\gamma \quad \text{(C)}$$

Its eigenvalue $\hat{J}^2$ and z (or x or y) component can be simultaneously known because

$$[\hat{J}^2, \hat{J}_\alpha] = 0$$

And the simultaneous eigenstates satisfy:

$$\hat{J}^2 |j, m_j\rangle = \hbar^2 j(j+1) |j, m_j\rangle$$
$$\hat{J}_z |j, m_j\rangle = \hbar m_j |j, m_j\rangle$$

For $j = 0, \frac{1}{2}, 1, \frac{3}{2}, \ldots$, $m_j = -j, -j+1, \ldots, j-1, j$.

We note that any set of operators satisfying (C) will have these eigenstates.

Suppose now we have two angular momenta $\underline{\hat{S}}$ and $\underline{\hat{L}}$:

$$[\hat{L}_\alpha, \hat{L}_\beta] = \sum_\gamma i\epsilon_{\alpha\beta\gamma}\hbar\hat{L}_\gamma$$
$$[\hat{S}_\alpha, \hat{S}_\beta] = \sum_\gamma i\epsilon_{\alpha\beta\gamma}\hbar\hat{S}_\gamma$$

And suppose these are independent so

$$[\hat{L}_\alpha, \hat{S}_\beta] = 0$$

Let us define the total angular momentum:

$$\underline{\hat{J}} = \underline{\hat{S}} + \underline{\hat{L}}$$
$$\hat{J}_\alpha = \hat{S}_\alpha + \hat{L}_\alpha$$

Is this also an angular momentum? If it satisfies (C) it is! Let's check this.

$$\begin{aligned}
[\hat{J}_\alpha, \hat{J}_\beta] &= [\hat{S}_\alpha, \hat{S}_\beta] + [\hat{S}_\alpha, \hat{L}_\beta] + [\hat{L}_\alpha, \hat{S}_\beta] + [\hat{L}_\alpha, \hat{L}_\beta] \\
&= \sum_\gamma i\epsilon_{\alpha\beta\gamma}\hbar\hat{S}_\gamma + \sum_\gamma i\epsilon_{\alpha\beta\gamma}\hbar\hat{L}_\gamma \\
&= \sum_\gamma i\epsilon_{\alpha\beta\gamma}\hbar(\hat{S}_\gamma + \hat{L}_\gamma) \\
&= \sum_\gamma i\epsilon_{\alpha\beta\gamma}\hbar\hat{J}_\gamma \quad \checkmark
\end{aligned}$$

This is an angular momentum.

We can also check $[\hat{J}_\alpha, \hat{L}^2]$:

$$\begin{aligned}
[\hat{J}_\alpha, \hat{L}^2] &= [\hat{S}_\alpha, \hat{L}^2] + [\hat{L}_\alpha, \hat{L}^2] \\
&\quad\uparrow \qquad\qquad \uparrow \\
0 &\because [\hat{S}_\alpha, \hat{L}_\beta] = 0 \qquad 0 \because \hat{L}_\alpha \text{ is an angular momentum} \\
&= 0
\end{aligned}$$

And likewise $[\hat{J}_\alpha, \hat{S}^2] = 0$. As an exercise show $[\hat{J}^2, \hat{L}^2] = [\hat{J}^2, \hat{S}^2] = 0$

This means we can simultaneously define $\hat{J}^2, \hat{J}_z, \hat{L}^2, \hat{S}^2$.

Can we also define $\hat{L}_z$? Clearly $[\hat{J}_z, \hat{L}_z] = 0$ and $[\hat{L}^2, \hat{L}_z] = 0$ and $[\hat{S}^2, \hat{L}_z] = 0$. So we only need to check $[\hat{J}^2, \hat{L}_z]$:

$$\begin{aligned}
[\hat{J}^2, \hat{L}_z] &= [\hat{L}^2, \hat{L}_z] + [\hat{S}^2, \hat{L}_z] + \sum_\alpha [2\hat{L}_\alpha\hat{S}_\alpha, \hat{L}_z] \\
&= \sum_\alpha 2\hat{S}_\alpha[\hat{L}_\alpha, \hat{L}_z] \\
&= -2i\hbar\hat{S}_x\hat{L}_y + 2i\hbar\hat{S}_y\hat{L}_x \\
&\neq 0
\end{aligned}$$

And likewise $[\hat{J}^2, \hat{S}_z] = -[\hat{J}^2, \hat{L}_z] \neq 0$. So we cannot define $\hat{L}_z$ & $\hat{S}_z$ if $\hat{J}^2$ is well defined.

This all means for a given eigenvalue of $\hat{L}^2$ and $\hat{S}^2$, we can either define $\hat{J}^2$ and $\hat{J}_z$ or $\hat{L}_z$ and $\hat{S}_z$, but not all four of them.

The $\hat{J}^2, \hat{J}_z$ eigenstates are called the **coupled representation**.

$$\begin{aligned}
\hat{J}^2 \left|j, m_j, \ell, s\right\rangle &= \hbar^2 j(j+1) \left|j, m_j, \ell, s\right\rangle \\
\hat{J}_z \left|j, m_j, \ell, s\right\rangle &= \hbar m_j \left|j, m_j, \ell, s\right\rangle \\
\hat{L}^2 \left|j, m_j, \ell, s\right\rangle &= \hbar^2 \ell(\ell+1) \left|j, m_j, \ell, s\right\rangle \\
\hat{S}^2 \left|j, m_j, \ell, s\right\rangle &= \hbar^2 s(s+1) \left|j, m_j, \ell, s\right\rangle
\end{aligned}$$

and the $\hat{L}_z, \hat{S}_z$ eigenstates are called the **uncoupled representation**.

$$\begin{aligned}
\hat{L}^2 \left|\ell, m_\ell; s, m_s\right\rangle &= \hbar^2 \ell(\ell+1) \left|\ell, m_\ell; s, m_s\right\rangle \\
\hat{S}^2 \left|\ell, m_\ell; s, m_s\right\rangle &= \hbar^2 s(s+1) \left|\ell, m_\ell; s, m_s\right\rangle \\
\hat{L}_z \left|\ell, m_\ell; s, m_s\right\rangle &= \hbar m_\ell \left|\ell, m_\ell; s, m_s\right\rangle \\
\hat{S}_z \left|\ell, m_\ell; s, m_s\right\rangle &= \hbar m_s \left|\ell, m_\ell; s, m_s\right\rangle
\end{aligned}$$

In what follows we will sometimes drop the explicit ℓ and s quantum numbers and use the simplified notation:

$$|j, m_j, \ell, s\rangle \equiv |j, m_j\rangle$$
$$|\ell, m_\ell; s, m_s\rangle \equiv |m_\ell; m_s\rangle = |m_\ell\rangle\,|m_s\rangle$$

Q: Can j take any value?

A: No! Let's see why.

We know the states $\{|j, m_j\rangle\}$ and $\{|m_\ell; m_s\rangle\}$ both form a basis for a given ℓ and s. This means there can only be $(2\ell+1)(2s+1)$ total coupled states (because this is the total number of uncoupled states), so there must be a maximum j.

We also know $|j, m_j\rangle$ and $|m_\ell; m_s\rangle$ are both eigenstates of $\hat{J}_z$:

$$\begin{aligned}\hat{J}_z\,|j, m_j\rangle &= \hbar m_j\,|j, m_j\rangle \\ \hat{J}_z\,|m_\ell; m_s\rangle &= (\hat{L}_z + \hat{S}_z)\,|m_\ell; m_s\rangle \\ &= \hbar(m_\ell + m_s)\,|m_\ell, m_s\rangle\end{aligned}$$

This means that

$$|j, m_j\rangle = \sum_{\substack{m_\ell, m_s \\ m_\ell + m_s = m_j}} C^{j,m}_{\ell, m_\ell, s, m_s}\,|m_\ell, m_s\rangle$$

where $C^{j,m}_{\ell, m_\ell, s, m_s} = 0$ unless $m_\ell + m_s = m_j$.

But the max value m_ℓ and m_s can take are ℓ and s and likewise m_j's max value is j, so the maximum value j can take is:

$$j_{\max} = \ell + s$$

Also j can only take integer steps from $j_{\max}$, because $m_\ell = \ell - \frac{1}{2}$ and $m_s = s - \frac{1}{2}$ states do not exist. This means:

$$j = \ell + s, \ell + s - 1, \ldots, j_{\min}$$

There is only one set of states with $j = \ell + s - 1$, because there are 2 states with $m_j = \ell + s - 1$:

$$|\ell - 1; s\rangle \quad , \quad |\ell; s - 1\rangle$$

where 1 state in this subspace is generated from total $j = \ell + s$, and $m_j = \ell + s - 1$:

$$\hat{J}_-\,|\ell + s, \ell + s\rangle \propto |\ell + s, \ell + s - 1\rangle \propto c^-_{\ell, m_\ell}\,|\ell - 1; s\rangle + c^-_{s, m_s}\,|\ell; s - 1\rangle$$

So the only other state orthogonal to $|\ell + s, \ell + s - 1\rangle$ is $|\ell + s - 1, \ell + s - 1\rangle$.

We can repeat this argument for every m_j to show that $|j, m_j\rangle$ is unique.

The last question is: what is $j_{\min}$?

Let's take $\ell + s =$ integer. The $m_j = 0$ subspace is spanned by $|j, 0\rangle$ for states $j = j_{\min}, j_{\min} + 1, \ldots, \ell + s - 1, \ell + s$.

But also the $m_j = 0$ subspace is spanned by the uncoupled states (take $\ell \geq s$):

$$|-s, s\rangle\,, |-s + 1, s - 1\rangle\,, \ldots, |s - 1, -s + 1\rangle\,, |s, -s\rangle$$

There are $2s + 1$ of these states. Likewise there are $\ell + s - j_{\min} + 1$ coupled states. This means:

$$\ell + s - j_{\min} + 1 = 2s + 1$$

$$\Rightarrow \quad j_{\min} = \ell - s$$

The same argument applies to the $|j, \pm\frac{1}{2}\rangle$ subspaces for $\ell + s =$ half integer. Overall we find:

$$j = |\ell - s|, |\ell - s| + 1, \ldots, \ell + s - 1, \ell + s.$$

This fully defines that set of allowed coupled angular momentum states.

7.6 Application of the coupled representation: Coupled spins

The Hamiltonian for two identical spins in a magnetic field is:

$$\hat{H} = \gamma B \hat{S}_{1z} + \gamma B \hat{S}_{2z} + \frac{2J}{\hbar} \hat{S}_1 \cdot \hat{S}_2$$

Q: What are the eigenstates & eigenvalues?

A: Let's rewrite this as:

$$\hat{S}_1 \cdot \hat{S}_2 = \frac{1}{2}(\hat{S}^2 - \hat{S}_1^2 - \hat{S}_2^2) \quad \text{with } \hat{S} = \hat{S}_1 + \hat{S}_2$$

So:

$$\begin{aligned} \hat{H} &= \gamma B(\hat{S}_{1z} + \hat{S}_{2z}) - \frac{J}{\hbar}(S(S+1) - S_1(S_1+1) - S_2(S_2+1)) \\ &= \gamma B \hat{S}_z - \frac{J}{\hbar}(\hat{S}^2 - \hat{S}_1^2 - \hat{S}_2^2) \end{aligned}$$

for $S = |\frac{1}{2} - \frac{1}{2}|, \ldots, \frac{1}{2} + \frac{1}{2}$, $m_S = -S, \ldots, +S$

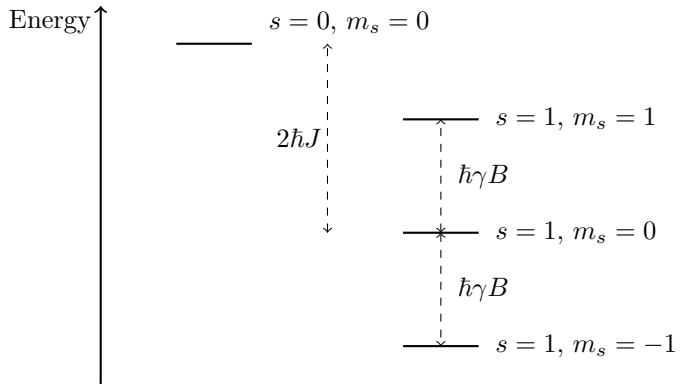
We have shown that $\hat{S}_1^2, \hat{S}_2^2, \hat{S}^2$ and $\hat{S}_z$ all commute and the simultaneous eigenstates are:

$$|s, m_s s_1, s_2\rangle$$

The energy eigenvalues are:

$$E_{s,m_s,s_1,s_2} = (\hbar\gamma B m_s - \hbar J(s(s+1) - s_1(s_1+1) - s_2(s_2+1)))$$

For $s_1 = \frac{1}{2}$ and $s_2 = \frac{1}{2}$ the energy levels are given by $s = |\frac{1}{2} - \frac{1}{2}|, \ldots, \frac{1}{2} + \frac{1}{2} = 0, 1$, $m_s = -s, \ldots, +s$. These energy levels are sketched below.



7.7 Application of the coupled representation: Atomic structure

We have seen that if a system's Hamiltonian is rotationally symmetric, then it commutes with the total angular momentum operator, and the energy eigenstates can also be set to be eigenstates of $\hat{L}^2$ and $\hat{L}_z$.

For a single electron in an atom, its orbitals (atomic wave functions without spin) will be eigenfunctions of $\hat{\ell}^2$ and $\hat{\ell}_z$ (with $\hat{\ell}_\alpha$ denoting the orbital angular momentum operator for a single electron) with corresponding quantum numbers ℓ and m. With spin added, the Hamiltonian is also invariant to rotations with $\hat{j}_\alpha = \hat{\ell}_\alpha + \hat{s}_\alpha$, so the eigenstates will also be eigenstates of $\hat{j}$, the total angular momentum of the electron (orbital plus spin). The electron spin quantum number is $s = 1/2$, which is a fundamental property of the electron, so the full eigenstates including spin will be eigenstates of $\hat{j}^2$ and $\hat{j}_z$. The magnetic field generated by orbital motion of the nucleus relative to the electron, $\underline{B}_\ell \propto (+Z)\underline{\ell}$, interacts with the electron spin magnetic moment, $\underline{\mu}_s \propto (-e)\underline{s}$, with a Hamiltonian given by

$$\hat{H}_{\text{soc}} = \xi(\hat{r})\hat{\underline{\ell}} \cdot \hat{\underline{s}}$$

where $\xi(\hat{r}) \propto Z/r^3$. Applying perturbation theory (more on this later) the spin-orbit coupling energy at lowest order is

$$E_{\text{soc}} = \zeta \left\langle \hat{\underline{\ell}} \cdot \hat{\underline{s}} \right\rangle = \frac{\zeta}{2} \left\langle \hat{j}^2 - \hat{s}^2 - \hat{\ell}^2 \right\rangle$$

with $\zeta \propto Z^4$.

In N-electron atoms, electron-repulsion means that for a given set of single-electron orbitals, with well-defined ℓ_i being occupied, to first order electron orbital and spin angular momenta couple separately, so we get a total orbital $\hat{\underline{L}}$ and spin $\hat{\underline{S}}$ angular momentum

$$\hat{\underline{L}} = \hat{\underline{\ell}}_1 + \hat{\underline{\ell}}_2 + \cdots \hat{\underline{\ell}}_N$$

$$\hat{\underline{S}} = \hat{\underline{s}}_1 + \hat{\underline{s}}_2 + \cdots \hat{\underline{s}}_N$$

This gives eigenstates for the N-electron atom with well-defined L and S quantum numbers. When the same sub-shell is occupied, only certain combinations of L and S are allowed on exchange symmetry grounds. The overall electronic wave function must be anti-symemtric with respect to swapping any two electrons. It turns out that the maximum L or S is symmetric, and the symmetry/anti-symmetry alternates as L or S is lowered. Anti-symmetric total orbital wavefunctions will be more stable, due to reduced repulsion, and larger total L, for a given symmetry, will be lower in energy. The Hamiltonian (including spin-orbit interactions) is invariant to total rotations by $\hat{J}_\alpha = \hat{L}_\alpha + \hat{S}_\alpha$, so these angular momenta couple into a total $\hat{J}_\alpha$, and the eigenstates are eigenstates of $\hat{J}^2$ and $\hat{J}_z$ as well as $\hat{L}^2$ and $\hat{S}^2$.

Due to spin-orbit coupling, there will be an effective spin-orbit interaction for a given orbital occupancy, L and S of the form

$$\hat{H}^{\text{eff}}_{\text{soc}} = \zeta \hat{\underline{L}} \cdot \hat{\underline{S}} = \frac{\zeta}{2}(\hat{J}^2 - \hat{L}^2 - \hat{S}^2)$$

In the hydrogenic atom $\zeta > 0$ but for multi-electron atoms, depending on the orbital occupancy ζ can be positive or negative. If a sub-shell is more than half filled, e.g. p^4 or d^7, then $\zeta < 0$, otherwise it is positive like in the one-electron atom.

7.8 Coupled States and Clebsch-Gordon Coefficients

We saw before that

$$|j, m_j, \ell, s\rangle = \sum_{m_\ell, m_s} C^{j,m_j}_{\ell,m_\ell,s,m_s} |\ell, m_\ell; s, m_s\rangle$$

The coefficients $C^{j,m_j}_{\ell,m_\ell,s,m_s} = \langle \ell, m_\ell; s, m_s | j, m_j, \ell, s\rangle \equiv \langle \ell, m_\ell; s, m_s | j, m_j\rangle$ are called the Clebsch-Gordon coefficients. Unfortunately there is no strong consensus on notation for the Clebsch-Gordon coefficients but throughout these notes we will use this convention. Noting that

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

We find

$$\begin{aligned}\hat{J}_\pm |j, m_j, \ell, s\rangle &= \sum_{m_\ell, m_s} C^{j,m_j\pm1}_{\ell,m_\ell,s,m_s} \hbar\sqrt{j(j+1) - m_j(m_j \pm 1)}\, |\ell, m_\ell; s, m_s\rangle \\ &= \sum_{m_\ell, m_s} \hbar C^{j,m_j}_{\ell,m_\ell,s,m_s} \times \Big(\hbar\sqrt{\ell(\ell+1) - m_\ell(m_\ell \pm 1)}\, |\ell, m_\ell \pm 1; s, m_s\rangle \\ &\quad + \hbar\sqrt{s(s+1) - m_s(m_s \pm 1)}\, |\ell, m_\ell; s, m_s \pm 1\rangle\Big)\end{aligned}$$

Therefore by comparing coefficients of $|\ell, m_\ell; s, m_s\rangle$

$$\begin{aligned}C^{j,m_j\pm1}_{\ell,m_\ell,s,m_s}\sqrt{j(j+1) - m_j(m_j \pm 1)} &= C^{j,m_j}_{\ell,m_\ell\mp1,s,m_s}\sqrt{\ell(\ell+1) - m_\ell(m_\ell \mp 1)} \\ &\quad + C^{j,m_j}_{\ell,m_\ell,s,m_s\mp1}\sqrt{s(s+1) - m_s(m_s \mp 1)}\end{aligned}$$

From this recursion we can construct arbitrary coupled states.

We can also obtain some orthogonality relations:

$$\begin{aligned}\langle j', m'_j, \ell, s | j, m_j, \ell, s\rangle &= \delta_{j,j'}\delta_{m_j,m'_j} \\ &= \sum_{m_\ell, m_s} C^{j',m'_j}_{\ell,m_\ell,s,m_s} C^{j,m_j}_{\ell,m_\ell,s,m_s}\end{aligned}$$

And likewise

$$\begin{aligned}\langle \ell, m'_\ell, s, m'_s | \ell, m_\ell, s, m_s\rangle &= \delta_{m_\ell,m'_\ell}\delta_{m_s,m'_s} \\ &= \sum_j \sum_{m_j} C^{j,m_j}_{\ell,m_\ell,s,m_s} C^{j,m_j}_{\ell,m'_\ell,s,m'_s}\end{aligned}$$

7.9 Matrix Elements using Wigner D-matrices

Suppose we want to find matrix elements of the form

$$\langle j', m' | \hat{O} | j, m\rangle$$

Is there an easy way of simplifying this problem?

Let's start by defining the Wigner D-matrices.

$$\hat{U}(R)\, |j, m\rangle = \sum_{m'} |j, m'\rangle\, \mathcal{D}^{(j)}_{m',m}(R)$$

We know the rotation operator $\hat{U}(R) = e^{-i\theta \underline{n}\cdot\underline{\hat{J}}/\hbar}$ commutes with $\hat{J}^2$.

Motivated by this let's assume we can also decompose $\hat{O}$ as

$$\hat{O} = \sum_{k=0}^{\infty} \sum_{q=-k}^{k} \hat{T}^{(k)}_q$$

where

$$\hat{U}^\dagger(R)\hat{T}^{(k)}_q\hat{U}(R) = \sum_{q'=-k}^{k} \mathcal{D}^{(k)*}_{q,q'}(R)\hat{T}^{(k)}_{q'}$$

For example, if $\hat{O}$ is a scalar

$$\hat{U}^\dagger(R)\hat{O}\hat{U}(R) = \hat{O}$$

$\mathcal{D}^{(0)}_{0,0}(R) = 1$, so $\hat{T}^{(0)}_0 = \hat{O}$ and all other $\hat{T}^{(k)}_q = 0$

If $\hat{O} = \hat{V}_\alpha$ is a component of a vector

$$\hat{U}^\dagger(R)\hat{V}_\alpha\hat{U}(R) = \sum_\beta R_{\alpha\beta}\hat{V}_\beta$$

So the set-of vector operators span a 3-dimensional space of operators, so

$$\hat{V}_\alpha = \sum_{q=-1}^{1} \hat{T}^{(1)}_q(V_\alpha)$$

We note that $\mathcal{D}^{(j)}_{m',m}(R)$ is given by

$$\mathcal{D}^{(j)}_{m',m}(R) = \langle j, m'|\hat{U}(R)|j, m\rangle$$

So for a small rotation we have

$$\begin{aligned} i\delta\theta\left[\underline{n}\cdot\underline{\hat{J}}, \hat{T}^{(k)}_q\right] &= +i\sum_{q'}\delta\theta T_{q'}\langle k, q'|\underline{n}\cdot\underline{\hat{J}}|k, q\rangle \\ \Rightarrow \left[\underline{n}\cdot\underline{\hat{J}}, \hat{T}^{(k)}_q\right] &= \sum_{q'}\hat{T}^{(k)}_{q'}\langle k, q'|\underline{n}\cdot\underline{\hat{J}}|k, q\rangle \end{aligned}$$

With $\underline{n} = \underline{e}_z$ we have

$$\begin{aligned} \left[\hat{J}_z, \hat{T}^{(k)}_q\right] &= \sum_{q'}\hat{T}^{(k)}_{q'}\langle k, q'|k, q\rangle \\ &= \hat{T}^{(k)}_q\hbar q \end{aligned}$$

and $\underline{n} = \underline{e}_x \pm i\underline{e}_y$

$$\begin{aligned} \left[\hat{J}_\pm, \hat{T}^{(k)}_q\right] &= \sum_{q'}\hat{T}^{(k)}_{q'}\langle k, q'|\hat{J}_\pm|k, q\rangle \\ &= \hat{T}^{(k)}_{q\pm1}\hbar\sqrt{k(k+1) - q(q\pm1)} \end{aligned}$$

Using this we can show some useful things e.g.

$$\langle j', m'|\hat{T}^{(k)}_q|j, m\rangle = \langle j', m+q|\hat{T}^{(k)}_0|j, m\rangle\delta_{m',m+q}$$

To prove this let's consider

$$\begin{aligned} 0 &= \langle j', m'|\left[\hat{J}_z, \hat{T}^{(k)}_q\right] - \hbar q\hat{T}^{(k)}_q|j, m\rangle \\ &= \hbar(m' - m - q)\langle j', m'|\hat{T}^{(k)}_q|j, m\rangle \\ \Rightarrow \langle j', m'|\hat{T}^{(k)}_q|j, m\rangle &= 0 \quad \text{unless } m' = m + q \end{aligned}$$

This turns out to be a special case of the Wigner-Eckart Theorem.

7.10 Properties of Wigner D-Matrices

before proceeding it is useful to derive the following property of the D-matrices. We start by considering how the state $|\ell, m_\ell; s, m_s\rangle$ rotates

$$\hat{U}(R)\,|\ell, m_\ell; s, m_s\rangle = \sum_{m'_\ell=-\ell}^{\ell} \sum_{m'_s=-s}^{s} \mathcal{D}^{(\ell)}_{m'_\ell,m_\ell}(R)\mathcal{D}^{(s)}_{m'_s,m_s}(R)\,|\ell, m'_\ell; s, m'_s\rangle$$

we note however that

$$|\ell, m_\ell; s, m_s\rangle = \sum_{j=|\ell-s|}^{\ell+s} \sum_{m_j=-j}^{j} C^{j,m_j}_{\ell,m_\ell,s,m_s}\,|j, m_j\rangle$$

and therefore the rotated state is

$$\hat{U}(R)\,|\ell, m_\ell; s, m_s\rangle = \sum_{j=|\ell-s|}^{\ell+s} \sum_{m_j=-j}^{j} \sum_{m'_j=-j}^{j} C^{j,m_j}_{\ell,m_\ell,s,m_s}\,\left|j, m'_j\right\rangle \mathcal{D}^{(j)}_{m'_j,m_j}(R)$$

By inserting the expansion of $\left|j, m'_j\right\rangle$ in terms of the uncoupled states, and equating the first and third expressions we find

$$\mathcal{D}^{(\ell)}_{m'_\ell,m_\ell}(R)\mathcal{D}^{(s)}_{m'_s,m_s}(R) = \sum_{j=|\ell-s|}^{\ell+s} \sum_{m_j=-j}^{j} \sum_{m'_j=-j}^{j} C^{j,m_j}_{\ell,m_\ell,s,m_s} C^{j,m'_j}_{\ell,m'_\ell,s,m'_s} \mathcal{D}^{(j)}_{m'_j,m_j}(R)$$

We also note that

$$\begin{aligned}
\mathcal{D}^{(j)}_{m',m}(R) &= \langle j, m'|\hat{U}(R)|j, m\rangle \\
&= \langle j, m|\hat{U}(R)^\dagger|j, m'\rangle^* \\
&= \langle j, m|\hat{U}(R^{-1})|j, m'\rangle^* \\
&= \mathcal{D}^{(j)}_{m,m'}(R^{-1})^*
\end{aligned}$$

This means

$$\hat{U}(R)\hat{T}^{(k)}_q\hat{U}(R)^\dagger = \sum_{q'=-k}^{k} \mathcal{D}^{(k)*}_{q,q'}(R^{-1})\hat{T}^{(k)}_{q'} = \sum_{q'=-k}^{k} \mathcal{D}^{(k)}_{q',q}(R)\hat{T}^{(k)}_{q'}$$

7.11 Wigner-Eckart Theorem

Let us consider how $\hat{T}^{(k)}_q\,|j, m\rangle$ transforms under rotations

$$\hat{U}(R)\hat{T}^{(k)}_q\,|j, m\rangle = \sum_{q'=-k}^{k} \sum_{n=-j}^{j} \mathcal{D}^{(k)}_{q',q}(R)\mathcal{D}^{(k)}_{n,m}(R)\hat{T}^{(k)}_{q'}\,|j, n\rangle \tag{8}$$

This immediately tells us that $\hat{T}^{(k)}_q\,|j, m\rangle$ has an expansion of the form

$$\hat{T}^{(k)}_q\,|j, m\rangle = \sum_{j'=|j-k|}^{j+k} \sum_{m'=-j'}^{j'} |j', m'\rangle\, T^{j',m'}_{k,q,j,m}$$

so when rotating this we find

$$\hat{U}(R)\hat{T}^{(k)}_q\,|j, m\rangle = \sum_{n'} \sum_{j'=|j-k|}^{j+k} \sum_{m'=-j'}^{j'} |j', n'\rangle\, \mathcal{D}^{(j')}_{n',m'} T^{j',m'}_{k,q,j,m}$$

If we instead insert the expression above for a product of two D-matrices we find into the first equation

$$\hat{U}(R)\hat{T}^{(k)}_q \left|j,m\right\rangle = \sum_{j',n',m'} \mathcal{D}^{(j')}_{n',m'}(R) \sum_{q'=-k}^{k} \sum_{n=-j}^{j} C^{j',m'}_{k,q,j,m} C^{j',n'}_{k,q',j,n} \hat{T}^{(k)}_{q'} \left|j,n\right\rangle$$

If we take the inner product with $\langle j', n''|$ for the two different equations immediately above for $\hat{U}(R)\hat{T}^{(k)}_q \left|j,m\right\rangle$ and equate them we get

$$\sum_{j',n',m'} \mathcal{D}^{(j')}_{n',m'}(R) \sum_{q'=-k}^{k} \sum_{n=-j}^{j} C^{j',m'}_{k,q,j,m} C^{j',n'}_{k,q',j,n} \left\langle j', n''\right| \hat{T}^{(k)}_{q'} \left|j,n\right\rangle = \sum_{j'm'n'} \delta_{n'',n'} \mathcal{D}^{(j')}_{n',m'}(R) T^{j',m'}_{k,q,j,m}$$

Equating the coefficients of $\mathcal{D}^{(j')}_{n',m'}(R)$ we obtain

$$C^{j',m'}_{k,q,j,m} \sum_{q'=-k}^{k} \sum_{n=-j}^{j} C^{j',n'}_{k,q',j,n} \left\langle j', n''\right| \hat{T}^{(k)}_{q'} \left|j,n\right\rangle = \delta_{n'',n'} T^{j',m'}_{k,q,j,m}$$

$$\Rightarrow C^{j',m'}_{k,q,j,m} \underbrace{\sum_{q'=-k}^{k} \sum_{n=-j}^{j} C^{j',n'}_{k,q',j,n} T^{j',n''}_{k,q,j,n}}_{\text{independent of } m,m',q} = \delta_{n'',n'} T^{j',m'}_{k,q,j,m}$$

The term inside the sum depends only on j, j', n', n''. The other factors on the LHS and RHS are independent of n' and n''. Dividing by $C^{j',m'}_{k,q,j,m}$ tells us that the ratio $T^{j',m'}_{k,q,j,m}/C^{j',m'}_{k,q,j,m}$ is independent of m, m', q, so this ratio must be proportional to a constant that only terms of j, j' and k, which we denote $\langle j||\hat{T}^{(k)}||j\rangle / \sqrt{2j+1}$. Overall we find that

$$T^{j',m'}_{k,q,j,m} = C^{j',m'}_{k,q,j,m} \frac{\langle j'\|\hat{T}^{(k)}\|j\rangle}{\sqrt{2j'+1}} \tag{9}$$

The constant is ratio to be divided by $\sqrt{2j'+1}$ by convention. We can check our deduction for consistency with the above straightforwardly.

The Wigner-Eckart theorem has big consequences e.g. atomic selection rules! It is also very useful in simplifying the calculation of many different matrix elements by immediately telling us that only a small handful of matrix elements of the form

$$T^{j',m'}_{k,q,j,m} = \langle j', m'|\hat{T}^{(k)}_q|j,m\rangle$$

are non-zero, and the non-zero terms are determined by which Clebsch-Gordon coefficients are non-zero.

An alternative proof

As we showed above general, the matrix element $\langle j', m'|\hat{T}^{(k)}_q|j,m\rangle$ is given by

$$\langle j', m'|\hat{T}^{(k)}_q|j,m\rangle = C^{j',m'_j}_{j,m_j,k,q} \frac{\langle j'\|\hat{T}^{(k)}\|j\rangle}{\sqrt{2j'+1}}$$

where $\langle j'\|\hat{T}^{(k)}\|j\rangle$ is called a reduced matrix element independent of m, m' and q and $C^{j',m'_j}_{j,m_j,k,q}$ is a Clebsch-Gordon coefficient $\langle j, m_j; k, q|j', m'\rangle$.

To prove this another way let's consider

$$\begin{aligned}
\langle j', m'| \left[\hat{J}_\pm, \hat{T}^{(k)}_q\right] \left|j,m\right\rangle &= \langle j', m'|\hat{T}^{(k)}_q|j,m\rangle \hbar\sqrt{k(k+1) - q(q\pm 1)} \\
&= \hbar\sqrt{j'(j'+1) - m'(m'\pm 1)} \langle j', m' \pm 1|\hat{T}^{(k)}_q|j,m\rangle \\
&- \hbar\sqrt{j(j+1) - m(m\pm 1)} \langle j', m'|\hat{T}^{(k)}_q|j,m\pm 1\rangle
\end{aligned}$$

We define $\langle j', m'|T_q^{(k)}|j, m\rangle = T^{k,q}_{j',m',j,m}$

Therefore:

$$\sqrt{j'(j'+1) - m'(m'\pm 1)}T^{k,j'}_{j,m\pm 1,q,m} = \sqrt{j(j+1) - m(m\pm 1)}T^{k,j'}_{j,m,q,m\pm 1} + \sqrt{k(k+1) - q(q\pm 1)}T^{k,j'}_{j,m,q\pm 1,m}$$

This is the same recursion as the CG coefficients, so

$$T^{k,j'}_{j,m,q,m} = \text{constant} \times C^{j',m'_j}_{j,m_j,k,q}$$

The constant is defined as $\frac{\langle j'\|\hat{T}^{(k)}\|j\rangle}{\sqrt{2j'+1}}$.

Projection Theorem Let's now consider a vector operator $\hat{V}_q = \hat{T}_q^{(1)}$.

How can we find $\langle j, m'|\hat{V}_q|j, m\rangle$?

Consider $\hat{V}_q$ is given by

$$\hat{V}_{\pm 1} = \mp\frac{1}{\sqrt{2}}(\hat{V}_x \pm i\hat{V}_y)$$

$\hat{V}_0 = \hat{V}_z$

The expectation value of this is

$$\begin{aligned}\langle j, m|\underline{\hat{V}}\cdot\underline{\hat{J}}|j, m\rangle &= \langle j, m|\hat{V}_0|j, m\rangle\hbar m \\ &- \frac{\hbar}{\sqrt{2}}\langle j, m|\hat{V}_{+1}|j, m-1\rangle\sqrt{j(j+1) - m(m-1)} \\ &+ \frac{\hbar}{\sqrt{2}}\langle j, m|\hat{V}_{-1}|j, m+1\rangle\sqrt{j(j+1) - m(m+1)} \\ &= a_{j,m}\langle j\|\hat{V}\|j\rangle\end{aligned}$$

Also by WET

$$\langle j, m|\underline{\hat{V}}\cdot\underline{\hat{J}}|j, m\rangle = C^{0,j}_{j,m,0,m}\frac{\langle j\|\underline{\hat{V}}\cdot\underline{\hat{J}}\|j\rangle}{\sqrt{2j+1}}$$

$C^{0,j}_{j,m,0,m} = \langle 0, 0; j, m|j, m, 0, j\rangle = 1.$

So $\Rightarrow a_{j,m} = a_j$, and setting $\hat{V} = \hat{J}$.

$$\langle j, m|\hat{J}^2|j, m\rangle = a_j\langle j\|\hat{J}\|j\rangle$$

This means

$$\langle j\|\hat{V}\|j\rangle = \langle j, m|\underline{\hat{V}}\cdot\underline{\hat{J}}|j, m\rangle\frac{\langle j\|\hat{J}\|j\rangle}{\langle j, m|\hat{J}^2|j, m\rangle}$$

Also WET tells us

$$C^{1,j}_{j,m,q,m'} = \frac{\langle j, m|\hat{V}_q|j, m'\rangle\cdot\sqrt{2j+1}}{\langle j\|\hat{V}\|j\rangle}$$

which is true for $\hat{V} = \hat{J}$ so

$$C^{1,j}_{j,m,q,m'} = \frac{\langle j, m|\hat{J}_q|j, m'\rangle\cdot\sqrt{2j+1}}{\langle j\|\hat{J}\|j\rangle}$$

Finally putting this back into WET for $\hat{V}$ we find

$$\begin{aligned}\langle j, m'|\hat{V}_q|j, m\rangle &= \frac{\langle j, m'|\hat{J}_q|j, m\rangle \cdot \sqrt{2j+1}}{\langle j\|\hat{J}\|j\rangle} \frac{\langle j, m|\underline{\hat{V}} \cdot \underline{\hat{J}}|j, m\rangle \langle j\|\hat{J}\|j\rangle}{\sqrt{2j+1}\langle j, m|\hat{J}^2|j, m\rangle} \\ &= \frac{\langle j, m|\underline{\hat{V}} \cdot \underline{\hat{J}}|j, m\rangle}{\hbar^2 j(j+1)} \langle j, m'|\hat{J}_q|j, m\rangle.\end{aligned}$$

This result finds use in deriving the Landé g-factor for the energy levels of atoms in magnetic fields.

8 The matrix formulation of quantum mechanics

Thus far we have used techniques from linear algebra, operators and differential equations to solve QM problems.

We will now explore how to formulate some of these problems instead as matrix problems.

8.1 Matrices and Vectors

As a reminder, vectors are arrays of complex numbers

$$\underline{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_N \end{pmatrix}$$

and matrices are 2D-arrays

$$\underline{\underline{A}} = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1N} \\ A_{21} & A_{22} & & \vdots \\ \vdots & & \ddots & \vdots \\ A_{N1} & \cdots & & A_{NN} \end{pmatrix} = \begin{pmatrix} \underline{A}^T_{r,1} \\ \underline{A}^T_{r,2} \\ \vdots \\ \underline{A}^T_{r,N} \end{pmatrix} = \begin{pmatrix} \underline{A}_{c,1} & \underline{A}_{c,2} & \cdots & \underline{A}_{c,N} \end{pmatrix}$$

Matrix-matrix multiplication is defined as

$$\underline{\underline{C}} = \underline{\underline{A}} \cdot \underline{\underline{B}}$$

$$C_{nm} = \sum_k A_{nk} B_{km} = \underline{A}^T_{r,n} \cdot \underline{B}_{c,m}$$

Matrix-vector multiplication is a special case of this

$$\underline{u} = \underline{\underline{A}}\underline{v}$$

$$u_n = \sum_k A_{nk} v_k = \underline{A}^T_{r,n} \cdot \underline{v}$$

Matrix-multiplication is a linear operation, i.e.

$$\underline{\underline{A}}(a\underline{v} + b\underline{u}) = a(\underline{\underline{A}}\underline{v}) + b(\underline{\underline{A}}\underline{u})$$

And we can define the inner product of two vectors as

$$\underline{u}^\dagger \cdot \underline{v} = \sum_n u_n^* v_n$$

8.2 Mapping QM to Matrix Equations

Suppose we have a basis $\{|n\rangle\}$ for a Hilbert space and we know the expansion of $|\psi\rangle$ and $|\phi\rangle$ in this basis

$$|\phi\rangle = \sum_n |n\rangle\, a_n \qquad a_n = \langle n|\phi\rangle$$

$$|\psi\rangle = \sum_n |n\rangle\, b_n \qquad b_n = \langle n|\psi\rangle$$

The inner product is

$$\begin{aligned}\langle\phi|\psi\rangle &= \langle\phi|\hat{1}|\psi\rangle \\ &= \sum_n \langle\phi|n\rangle\langle n|\psi\rangle \\ &= \sum_n a_n^* b_n\end{aligned}$$

If we define vectors $\underline{a}$ and $\underline{b}$, we have

$$\langle\phi|\psi\rangle = \underline{a}^\dagger \cdot \underline{b}$$

So we can map inner products to vector inner products in a particular basis. Likewise we can find actions of linear operators as

$$|\phi\rangle = \hat{A}\,|\psi\rangle$$

$$\begin{aligned}a_n &= \langle n|\phi\rangle \\ &= \langle n|\hat{A}|\psi\rangle \\ &= \langle n|\hat{A}\hat{1}|\psi\rangle \\ &= \sum_m \langle n|\hat{A}|m\rangle\langle m|\psi\rangle \\ &= \sum_m \langle n|\hat{A}|m\rangle\, b_m\end{aligned}$$

$\Rightarrow \quad \underline{a} = \underline{\underline{A}} \cdot \underline{b} \quad$ where $[\underline{\underline{A}}]_{nm} = \langle n|\hat{A}|m\rangle$

[Remember we often referred to $\langle n|\hat{A}|m\rangle$ as a matrix element]

This means eigenvalue/eigenvector problems in QM map onto eigenvalue decompositions of matrices.

$$\hat{A}\,|\psi\rangle = \lambda\,|\psi\rangle$$

is equivalent to

$$\underline{\underline{A}}\underline{b} = \lambda\underline{b}$$

So if we know $\langle n|\hat{A}|m\rangle$ in a particular basis, we can solve many problems in QM (at least numerically).

8.3 Example: Spin-$\frac{1}{2}$

We can choose the S_z projection eigenstates as a basis for a spin-$\frac{1}{2}$ system, $\{|s, m_s\rangle\}$

$$\begin{aligned}|1\rangle \equiv |\alpha\rangle &\equiv \left|\frac{1}{2}, +\frac{1}{2}\right\rangle \\ |2\rangle \equiv |\beta\rangle &\equiv \left|\frac{1}{2}, -\frac{1}{2}\right\rangle\end{aligned}$$

In general the matrix representation of $\hat{A}$ is

$$\underline{\underline{A}} = \begin{pmatrix}\langle\alpha|\hat{A}|\alpha\rangle & \langle\alpha|\hat{A}|\beta\rangle \\ \langle\beta|\hat{A}|\alpha\rangle & \langle\beta|\hat{A}|\beta\rangle\end{pmatrix}$$

The matrix representation of $\hat{1}$ in this basis is

$$\hat{1} = \begin{pmatrix}1 & 0 \\ 0 & 1\end{pmatrix}$$

For $\hat{S}_z$ we have

$$\left\langle \frac{1}{2}, m_s \middle| \hat{S}_z \middle| \frac{1}{2}, m_s' \right\rangle = \hbar m_s \delta_{m_s m_s'}$$

So

$$\underline{\underline{S}}_z = \begin{pmatrix} \frac{\hbar}{2} & 0 \\ 0 & -\frac{\hbar}{2} \end{pmatrix}$$

For $\hat{S}_+$ we have

$$\left\langle \frac{1}{2}, m_s \middle| \hat{S}_+ \middle| \frac{1}{2}, m_s' \right\rangle = \hbar \sqrt{\frac{3}{4} - m_s(m_s+1)} \delta_{m_s, m_s'+1}$$

$\Rightarrow \quad \langle\alpha|\hat{S}_+|\alpha\rangle = \langle\beta|\hat{S}_+|\beta\rangle = \langle\beta|\hat{S}_+|\alpha\rangle = 0$ and

$$\langle\alpha|\hat{S}_+|\beta\rangle = \hbar\sqrt{\frac{3}{4} - (-\frac{1}{2})(\frac{1}{2})} = \hbar$$

So

$$\underline{\underline{S}}_+ = \begin{pmatrix} 0 & \hbar \\ 0 & 0 \end{pmatrix}$$

Likewise $\hat{S}_-$ is given by

$$\underline{\underline{S}}_- = \begin{pmatrix} 0 & 0 \\ \hbar & 0 \end{pmatrix}$$

And $\hat{S}_x = \frac{1}{2}(\hat{S}_+ + \hat{S}_-)$ and $\hat{S}_y = \frac{1}{2i}(\hat{S}_+ - \hat{S}_-)$

$$\underline{\underline{S}}_x = \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad\qquad \underline{\underline{S}}_y = \begin{pmatrix} 0 & -i\frac{\hbar}{2} \\ +i\frac{\hbar}{2} & 0 \end{pmatrix}$$

8.4 Example: Two Spin-$\frac{1}{2}$

For a system with 2 spin-$\frac{1}{2}$ angular momenta, $\hat{S}_1$ and $\hat{S}_2$, we can define a basis

$$\begin{aligned} |1\rangle &= |\alpha_1\alpha_2\rangle \equiv |\alpha\alpha\rangle \\ |2\rangle &= |\alpha_1\beta_2\rangle \equiv |\alpha\beta\rangle \\ |3\rangle &= |\beta_1\alpha_2\rangle \equiv |\beta\alpha\rangle \\ |4\rangle &= |\beta_1\beta_2\rangle \equiv |\beta\beta\rangle \end{aligned}$$

With $\sigma_i = \alpha_i$ or β_i, we know if $\hat{A}_1$ only acts on spin 1

$$\langle\sigma_1\sigma_2|\hat{A}_1|\sigma_1'\sigma_2'\rangle = \langle\sigma_1|\hat{A}_1|\sigma_1'\rangle\langle\sigma_2|\sigma_2'\rangle = \langle\sigma_1|\hat{A}_1|\sigma_1'\rangle\,\delta_{\sigma_2\sigma_2'}$$

and likewise for $\hat{A}_2$ that only acts on 2

$$\langle\sigma_1\sigma_2|\hat{A}_2|\sigma_1'\sigma_2'\rangle = \delta_{\sigma_1\sigma_1'}\langle\sigma_2|\hat{A}_2|\sigma_2'\rangle$$

For example, $\hat{S}_{1z}$ is given in this basis by:

$$\underline{\underline{S}}_{1z} = \frac{\hbar}{2}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

And $\hat{S}_{1x}$:

$$\underline{\underline{S}}_{1x} = \frac{\hbar}{2}\begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

And $\hat{S}_{2z}$:

$$\underline{\underline{S}}_{2z} = \frac{\hbar}{2}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

And $\hat{S}_{2x}$:

$$\underline{\underline{S}}_{2x} = \frac{\hbar}{2}\begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

And $\hat{S}_{1z}\hat{S}_{2z} = \hat{C}$

$$\begin{aligned} \underline{\underline{C}} &= \frac{\hbar^2}{4}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \\ &= \underline{\underline{S}}_{1z} \cdot \underline{\underline{S}}_{2z} \end{aligned}$$

You'll notice in general that $\hat{A}_1$ is represented as

$$\hat{A}_1 = \begin{pmatrix} A_{11} & 0 & A_{12} & 0 \\ 0 & A_{11} & 0 & A_{12} \\ A_{21} & 0 & A_{22} & 0 \\ 0 & A_{21} & 0 & A_{22} \end{pmatrix} = \begin{pmatrix} A_{11}\underline{\underline{1}} & A_{12}\underline{\underline{1}} \\ A_{21}\underline{\underline{1}} & A_{22}\underline{\underline{1}} \end{pmatrix}$$

where $A_{nm} = \langle n|\hat{A}_1|m\rangle$ are the matrix elements of $\hat{A}_1$ in the α, β basis without $\underline{\hat{S}}_2$.

Likewise for $\hat{B}_2$ that only acts on $\underline{\hat{S}}_2$:

$$\hat{B}_2 = \begin{pmatrix} B_{11} & B_{12} & 0 & 0 \\ B_{21} & B_{22} & 0 & 0 \\ 0 & 0 & B_{11} & B_{12} \\ 0 & 0 & B_{21} & B_{22} \end{pmatrix} = \begin{pmatrix} \underline{\underline{B}} & \underline{\underline{0}} \\ \underline{\underline{0}} & \underline{\underline{B}} \end{pmatrix}$$

And generally for $\hat{C} = \hat{A}_1\hat{B}_2 = \hat{B}_2\hat{A}_1$

$$\underline{\underline{C}} = \underline{\underline{A}} \otimes \underline{\underline{B}} = \underline{\underline{A}}_1 \cdot \underline{\underline{B}}_2 = \begin{pmatrix} A_{11}\underline{\underline{B}} & A_{12}\underline{\underline{B}} \\ A_{21}\underline{\underline{B}} & A_{22}\underline{\underline{B}} \end{pmatrix}$$

This is called a Kronecker product and it's a consequence of ordering the composite system basis functions as

$$\begin{aligned} |1\rangle &\equiv |1,1\rangle \\ |2\rangle &\equiv |1,2\rangle \\ |3\rangle &\equiv |2,1\rangle \\ |4\rangle &\equiv |2,2\rangle \end{aligned}$$

In the next section of these notes we sketch how to apply these principles to more general multi-particle/composite systems.

8.5 General Composite Systems

Suppose we have multiple independent degrees of freedom in a QM system, e.g. x and y positions of a particle or two different angular momenta $\hat{L}$ and $\hat{S}$. The full Hilbert space of the composite system is a direct product of the two subsystem Hilbert spaces:

$$\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$$

We assume $\mathcal{H}_1$ has a dimension of N_1 and $\mathcal{H}_2$ has a dimension of N_2. Given a basis of $\mathcal{H}_1$, $\{|n_1\rangle\}$, and for $\mathcal{H}_2$, $\{|n_2\rangle\}$, we can construct a basis for $\mathcal{H}$ as:

$$|n_1, n_2\rangle \equiv |n_1\rangle \otimes |n_2\rangle \equiv |n_1\rangle\,|n_2\rangle$$

An operator acting on $\mathcal{H}_1$ has the following effect on the basis states:

$$\begin{aligned}\hat{A}_1\,|n_1, n_2\rangle &= (\hat{A}_1\,|n_1\rangle) \otimes |n_2\rangle \\ &= \sum_{m_1} A_{1,m_1,n_1}\,|m_1\rangle \otimes |n_2\rangle\end{aligned}$$

and for $\mathcal{H}_2$ we have:

$$\hat{A}_2\,|n_1, n_2\rangle = |n_1\rangle \otimes \sum_{m_2} A_{2,m_2,n_2}\,|m_2\rangle$$

The quantum state $|\psi\rangle \in \mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$ can be written in terms of a 2-index tensor (a matrix) as:

$$|\psi\rangle = \sum_{n_1,n_2} a_{n_1,n_2}\,|n_1, n_2\rangle$$

For practical calculations, we need to flatten this out into a vector:

$$|\psi\rangle = \sum_{n} a_n\,|n\rangle$$

We do this by defining a total composite index n, for $n_1 = 1, 2, \ldots, N_1$, $n_2 = 1, \ldots, N_2$, as:

$$n = (N_2 - 1)n_1 + n_2$$

So $|n\rangle = |n_1, n_2\rangle$, and n runs from 1 to $N = N_1 N_2$, which is the total dimensionality of the composite system.

This means $\underline{a}$ is given by:

$$\underline{a} = \begin{pmatrix} a_{11} \\ a_{12} \\ \vdots \\ a_{1N} \\ \vdots \\ a_{21} \\ \vdots \\ a_{NN} \end{pmatrix} \qquad a_n = a_{n_1,n_2}$$

Given this, how can we represent $\hat{A}_1$ and $\hat{A}_2$ as matrices?

We know:

$$\langle n_1, n_2|\hat{A}_1|m_1, m_2\rangle = A_{1,n_1,m_1}\delta_{n_2,m_2}$$

We can also define the Kronecker product of two matrices as:

$$\underline{\underline{A}} \otimes \underline{\underline{B}} = \begin{pmatrix} A_{11}\underline{\underline{B}} & A_{12}\underline{\underline{B}} & \cdots & A_{1N_1}\underline{\underline{B}} \\ A_{21}\underline{\underline{B}} & A_{22}\underline{\underline{B}} & \cdots & A_{2N_1}\underline{\underline{B}} \\ \vdots & \vdots & \ddots & \vdots \\ A_{N_1,1}\underline{\underline{B}} & A_{N_1,2}\underline{\underline{B}} & \cdots & A_{N_1,N_1}\underline{\underline{B}} \end{pmatrix}$$

Now if $A_{1,n_1,m_1} = \delta_{n_1,p_1}\delta_{m_1,q_1}$:

$$\begin{aligned} \hat{A}_1 |\psi\rangle &= \sum_{n_2} |p_1\rangle a_{q_1,n_2} \\ &\equiv \sum_{n} |n\rangle b_n \end{aligned}$$

So $\underline{b}$ is given by:

$$\underline{b} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ a_{q_1,1} \\ a_{q_1,2} \\ \vdots \\ a_{q_1,N_2} \\ \vdots \\ 0 \\ \vdots \\ 0 \end{pmatrix} \begin{matrix} \\ \\ \\ \leftarrow n = (N_2-1)p_1 + 1 \\ \\ \vdots \\ \leftarrow n = (N_2-1)p_1 + N_2 \\ \\ \\ \\ \end{matrix}$$

Which is equivalent to:

$$\underline{b} = (\underline{\underline{A}}_1 \otimes \underline{\underline{1}}_2)\underline{a}$$

$$\underline{\underline{A}}_1 \otimes \underline{\underline{1}}_2 = \begin{pmatrix} \underline{\underline{0}}_2 & \cdots & & \underline{\underline{0}}_2 \\ \vdots & \ddots & \underline{\underline{1}}_2 & \vdots \\ & & \underline{\underline{0}}_2 & \\ \underline{\underline{0}}_2 & \cdots & & \underline{\underline{0}}_2 \end{pmatrix}$$

It is straightforward ot generalise this to arbitrary $\hat{A}_1$ by decomposing $\hat{A}_1 = \sum_{p_1,q_1} A_{1,p_1,q_2} |p_1\rangle\langle q_2|$ and using linearity. So in general $\hat{A}_1 \longrightarrow \underline{\underline{A}}_1 \otimes \underline{\underline{1}}_2$.

Likewise for $\hat{A}_2 = |p_2\rangle \langle q_2|$:

$$\begin{aligned}\hat{A}_2 |\psi\rangle &= \sum_{n_1} |n_1, p_2\rangle \, a_{n_1,q_2} \\ &= \sum_n |n\rangle \, b_n\end{aligned}$$

$$\underline{b} = \begin{pmatrix} 0 \\ \vdots \\ a_{1,q_2} \\ \vdots \\ 0 \\ \vdots \\ a_{N_2 q_2} \\ \vdots \\ 0 \end{pmatrix} \begin{matrix} \\ \\ \leftarrow n = p_2 \\ \\ \\ \\ \leftarrow n = (N_1 - 1)N_2 + p_2 \\ \\ \end{matrix}$$

which is equivalent to:

$$\underline{b} = (\underline{\underline{1}}_1 \otimes \underline{\underline{A}}_2) \cdot \underline{a}$$

$$\underline{\underline{1}}_1 \otimes \underline{\underline{A}}_2 = \begin{pmatrix} \underline{\underline{A}}_2 & \underline{\underline{0}} & \cdots & \underline{\underline{0}} \\ \underline{\underline{0}} & \underline{\underline{A}}_2 & \cdots & \underline{\underline{0}} \\ \vdots & \vdots & \ddots & \vdots \\ \underline{\underline{0}} & \underline{\underline{0}} & \cdots & \underline{\underline{A}}_2 \end{pmatrix}$$

Again by decomposing an arbitrary $\hat{A}_2$ as we did for $\hat{A}_1$ into $\hat{A}_2 = \sum_{p_2,q_2} A_{2,p_2,q_2} \, |p_2\rangle\langle q_2|$ and using linearity we see that this expression applies to general $\hat{A}_2$.

For a general operator $\hat{C} = \hat{A}_1 \hat{B}_2$, its matrix representation is given by:

$$\begin{aligned}\underline{\underline{C}} &= \underline{\underline{A}}_1 \cdot \underline{\underline{B}}_2 \\ &= (\underline{\underline{A}}_1 \otimes \underline{\underline{1}}_2)(\underline{\underline{1}}_1 \otimes \underline{\underline{B}}_2) \\ &= \underline{\underline{A}}_1 \otimes \underline{\underline{B}}_2\end{aligned}$$

We can generalise this for N-degrees of freedom by iteratively applying the argument above.

$$\begin{gathered}\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \cdots \otimes \mathcal{H}_N \\ \hat{A}_1 \hat{A}_2 \cdots \hat{A}_N \longrightarrow \underline{\underline{A}}_1 \otimes \underline{\underline{A}}_2 \otimes \cdots \otimes \underline{\underline{A}}_N.\end{gathered}$$

9 The Hydrogen Atom

9.1 The Hamiltonian

The Hamiltonian for an electron in a Coulomb potential is

$$\hat{H} = \hat{T} + V = -\frac{Ze^2}{4\pi\epsilon_0 r} = -\frac{\hbar^2}{2m_e}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right) - \frac{Ze^2}{4\pi\epsilon_0 r}$$

where $r = \sqrt{x^2 + y^2 + z^2}$ and Z is the nuclear charge of the potential.

In the exercises you will show that

$$\hat{H} = E_h \hat{H}' = E_h\left(\hat{T}' - \frac{Z}{\rho}\right)$$

$$\hat{H}' = -\frac{1}{2}\left(\frac{\partial^2}{\partial x'^2} + \frac{\partial^2}{\partial y'^2} + \frac{\partial^2}{\partial z'^2}\right) - \frac{Z}{\rho}$$

where $x = a_0 x'$, $y = a_0 y'$, $z = a_0 z'$. This is a conversion to atomic units. E_h is called the Hartree, and it has a value of about 27.2 eV or 2625.5 kJ/mol. a_0 is called the Bohr, and it's about 0.529 Å.

$$E_h a_0^2 = \frac{\hbar^2}{2m_e} \qquad E_h = \frac{e^2}{4\pi\epsilon_0}$$

$$\Rightarrow a_0 = \frac{4\pi\hbar^2\epsilon_0}{e^2 m_e} \qquad E_n = \frac{\hbar^2}{m_e a_0^2}$$

Hereafter we will now drop the primes and work in atomic units.

9.2 Spherical polar coordinates

We can transform this to spherical polar coordinates

$$\begin{aligned} x &= r\cos\phi\sin\theta \\ y &= r\sin\phi\sin\theta \\ z &= r\cos\theta \end{aligned}$$

Using rules from multivariable calculus e.g.

$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial r}\frac{\partial r}{\partial x} + \frac{\partial f}{\partial \theta}\frac{\partial \theta}{\partial x} + \frac{\partial f}{\partial \phi}\frac{\partial \phi}{\partial x}$$

and $\frac{\partial r}{\partial x} = \left(\frac{\partial x}{\partial r}\right)^{-1}$ etc.

We end up finding:

$$\hat{T} = -\frac{1}{2}\left[\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{1}{r^2}\hat{\Lambda}_\Omega\right]$$

$$\hat{\Lambda}_\Omega = \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}$$

You will also show that

$$-\hat{\Lambda}_\Omega = \hat{L}^2$$

So

$$\hat{T} = \hat{T}_r + \frac{1}{r^2}\hat{L}^2$$

We can then also show that

$$[\hat{L}_\alpha, \hat{r}] = [\hat{L}_\alpha, \hat{T}_r] = 0$$

So the eigenstates of $\hat{H}$ can be chosen to be simultaneous eigenstates of $\hat{H}$, $\hat{L}^2$ and $\hat{L}_z$.

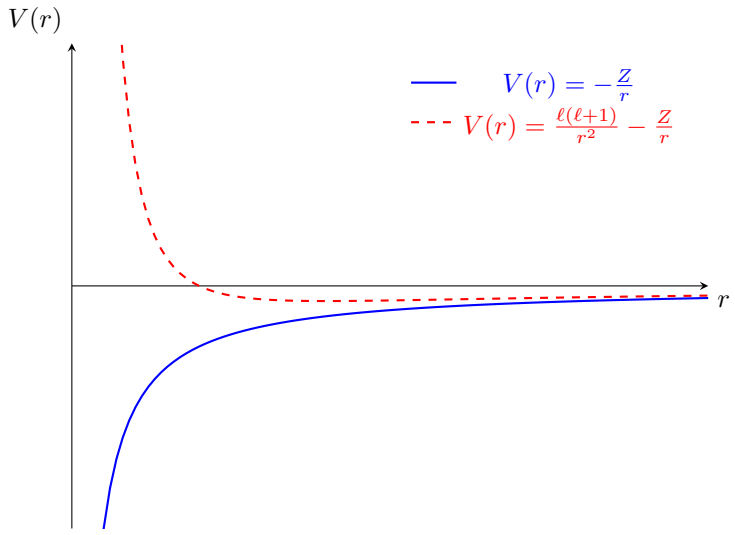
Such an eigenstate is denoted $|E, \ell, m\rangle$ and we find

$$\begin{aligned}
\hat{H}|E,\ell,m\rangle &= \left(\hat{T}_r + \frac{1}{r^2}\hat{L}^2 - \frac{Z}{r}\right)|E,\ell,m\rangle \\
&= \left(\hat{T}_r + \frac{\ell(\ell+1)}{r^2} - \frac{Z}{r}\right)|E,\ell,m\rangle \\
&\equiv H_\ell|E,\ell,m\rangle
\end{aligned}$$

Note we drop $\hbar^2$ from the $\hat{L}^2$ eigenvalue because in atomic units $\hat{p}_\alpha = -i\frac{\partial}{\partial x_\alpha}$, and the angular momentum commutators are $[\hat{L}_x, \hat{L}_y] = i\hat{L}_z$ etc..

We see there is a new effective potential $V_\ell(r)$ which depends on ℓ:

$$V_\ell(r) = \frac{\ell(\ell+1)}{r^2} - \frac{Z}{r}$$



The $\ell(\ell+1)/r^2$ term creates a barrier called the centrifugal barrier.

We can use a ladder operator technique to solve the effective eigenvalue problem above.

Because $[\hat{L}_\alpha, \hat{r}] = 0$ and $\hat{L}_\alpha$ is independent of r, we can separate the r part $|E, \ell\rangle \equiv R_\ell(r)$ and the angular part $|\ell, m\rangle \equiv Y_\ell^m(\theta, \phi)$ which is just a function of the spherical polar angles. These functions can be found explicitly using the angular momentum ladder operators.

$$|E, \ell, m\rangle = |E, \ell\rangle \otimes |\ell, m\rangle$$

Because $\hat{H}_\ell$ only acts on r, we can just work with the radial equation

$$\hat{H}_\ell |E, \ell\rangle = E|E, \ell\rangle.$$

[Write this out using $\Psi_{E,\ell,m}(r, \theta, \phi) = R_{E,\ell}(r)Y_{\ell m}(\theta, \phi)$ if you are unconvinced.]

9.3 Hydrogenic Atom Ladder Operators

Let us define $\hat{p}_r$ as

$$\hat{p}_r = -i\left(\frac{1}{r} + \frac{\partial}{\partial r}\right)$$

Noting $[-i\frac{\partial}{\partial r}, 1/r]] = +i\frac{1}{r^2}$, we find

$$\begin{aligned}
\frac{1}{2}\hat{p}_r^2 &= -\frac{1}{2}\left(\frac{\partial^2}{\partial r^2} + \frac{1}{r^2} + \frac{2}{r}\frac{\partial}{\partial r} - \frac{1}{r^2}\right) \\
&= -\frac{1}{2}\left(\frac{\partial^2}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r}\right) \\
&= \hat{T}_r
\end{aligned}$$

And let us define the ladder operators $\hat{A}_\ell$:

$$\hat{A}_\ell = \frac{1}{\sqrt{2}}\left(i\hat{p}_r + \frac{(\ell+1)}{r} - \frac{Z}{(\ell+1)}\right)$$

$$\hat{A}_\ell^\dagger = \frac{1}{\sqrt{2}}\left(-i\hat{p}_r + \frac{(\ell+1)}{r} - \frac{Z}{(\ell+1)}\right)$$

We find that $\hat{A}_\ell^\dagger \hat{A}_\ell$ is

$$\begin{aligned}
\hat{A}_\ell^\dagger \hat{A}_\ell &= \frac{1}{2}\hat{p}_r^2 + \frac{1}{2}\left[\frac{(\ell+1)}{r} - \frac{Z}{(\ell+1)}\right]^2 + i\frac{1}{2}\left[\hat{p}_r, -\frac{(\ell+1)}{r} + \frac{Z}{(\ell+1)}\right] \\
&= \frac{1}{2}\hat{p}_r^2 + \frac{((\ell+1))^2}{2r^2} - \frac{Z}{r} + \frac{Z^2}{2((\ell+1))^2} + \frac{i}{2}(-i)\frac{(\ell+1)}{r^2} \\
&= \frac{1}{2}\hat{p}_r^2 + \frac{(\ell+1)^2 + (\ell+1)}{2r^2} - \frac{Z}{r} + \frac{Z^2}{2((\ell+1))^2} \\
&= \frac{1}{2}\hat{p}_r^2 + \frac{(\ell+1)(\ell+2)}{2r^2} - \frac{Z}{r} + \frac{Z^2}{2((\ell+1))^2} \\
&= \hat{H}_{\ell+1} + \frac{Z^2}{2((\ell+1))^2}
\end{aligned}$$

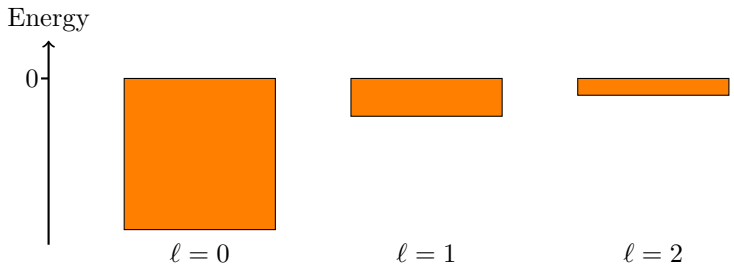
Likewise we find:

$$\hat{A}_\ell \hat{A}_\ell^\dagger = \frac{1}{2}\hat{p}_r^2 + \frac{(\ell+1)^2 - (\ell+1)}{2r^2} - \frac{Z}{r} + \frac{Z^2}{2((\ell+1))^2}$$
$$= \hat{H}_\ell + \frac{Z^2}{2((\ell+1))^2}$$

This shows that the eigenvalues for the given ℓ are bounded below by

$$E \geq \frac{Z^2}{2(\ell+1)^2}$$

i.e. the state $|E, \ell\rangle$ can only exist if $E \geq \frac{Z^2}{2(\ell+1)^2}$. The range of allowed E for a given ℓ is shown as the orange blocks below.



Now consider

$$\hat{A}_\ell \hat{H}_{\ell+1} |E, \ell\rangle = E(\hat{A}_\ell |E, \ell+1\rangle)$$

But also we can write the left-hand side as

$$\hat{A}_\ell \hat{H}_{\ell+1} |E, \ell+1\rangle = \hat{A}_\ell \left(\hat{A}_\ell^\dagger \hat{A}_\ell - \frac{Z^2}{2((\ell+1))^2} \right) |E, \ell+1\rangle$$
$$= \left(\hat{A}_\ell \hat{A}_\ell^\dagger - \frac{Z^2}{2((\ell+1))^2} \right) \hat{A}_\ell |E, \ell+1\rangle$$
$$= \hat{H}_\ell \left(\hat{A}_\ell |E, \ell+1\rangle \right)$$

So $\hat{A}_\ell |E, \ell+1\rangle \propto |E, \ell\rangle$, a state with the same energy but ℓ incremented down by 1.

Likewise we find that

$$\hat{A}_\ell^\dagger \hat{H}_\ell |E, \ell\rangle = \hat{H}_{\ell+1} \left(\hat{A}_\ell^\dagger |E, \ell\rangle \right)$$

so $\hat{A}_\ell^\dagger |E, \ell\rangle \propto |E, \ell+1\rangle$, a state with the same energy but ℓ incremented up by 1.

We cannot keep applying $\hat{A}_\ell^\dagger$ indefinitely because the eigenvalues of $\hat{H}_\ell$ are bounded below by $-Z^2/2(\ell+1)^2$. This means for a given energy $E < 0$ there is an $\ell_{\max}$ such that the ladder of states terminates

$$\hat{A}_{\ell_{\max}} |E, \ell_{\max}\rangle = 0$$

For this state we have

$$\langle H_{\ell_{\max}} \rangle = E = \langle E, \ell_{\max} | \hat{A}_{\ell_{\max}}^+ \hat{A}_{\ell_{\max}} |E, \ell_{\max}\rangle - \frac{Z^2}{2(\ell_{\max}+1)^2}$$

$$\Rightarrow E = -\frac{Z^2}{2(\ell_{\max}+1)^2}$$

We set $n = \ell_{\max} + 1$, so the energy eigenvalues are

$$E_n = -\frac{Z^2}{2n^2}, \quad \ell < n, \quad n = 0, 1, 2, \ldots$$

We now have the full set of energy eigenvalues, and a way of constructign all eigenstates of the hydrogenic atom, without ever solving a differential equation directly. We also see that the degeneracy of states with different ℓ arises from the $\frac{1}{r}$ potential.

We know degeneracies are rarely accidental in QM. Instead it implies there is some operator $\hat{A}$ such that

$$[\hat{H}, \hat{A}] = 0.$$

The Runge-Lenz Vector

Here we briefly go over the "hidden" symmetry in the Hydrogenic atom.

We start with a definition of the Runge-Lenz vector

$$\underline{\hat{A}} = \frac{1}{2}\left(\underline{\hat{p}} \times \underline{\hat{L}} - \underline{\hat{L}} \times \underline{\hat{p}}\right) - Z\frac{\hat{\underline{r}}}{\hat{r}}$$

It turns out that

$$[\hat{H}, \hat{A}_\alpha] = 0$$

And

$$[\hat{A}_\alpha, \hat{L}_\beta] = i\sum_\gamma \epsilon_{\alpha\beta\gamma}\hat{A}_\gamma$$

$$[\hat{A}_\alpha, \hat{A}_\beta] = -2\hat{H}\left(i\sum_\gamma \epsilon_{\alpha\beta\gamma}\hat{L}_\gamma\right)$$

Defining $\hat{D}_\alpha = \frac{1}{\sqrt{-2\hat{H}}}$, we find

$$[\hat{H}, \hat{L}_\alpha] = 0, \quad [\hat{H}, \hat{D}_\alpha] = 0$$

$$[\hat{L}_\alpha, \hat{L}_\beta] = [\hat{D}_\alpha, \hat{D}_\beta] = i\sum_\gamma \epsilon_{\alpha\beta\gamma}\hat{L}_\gamma$$

$$[\hat{D}_\alpha, \hat{L}_\beta] = i\sum_\gamma \epsilon_{\alpha\beta\gamma}\hat{D}_\gamma$$

It also turns out that

$$\hat{D}^2 + \hat{L}^2 = -\frac{Z^2}{2\hat{H}} + 1$$

Defining $\underline{\hat{B}}_\pm = \frac{1}{2}\left(\underline{\hat{L}} \pm \underline{\hat{D}}\right)$. Noting also that

$$\underline{\hat{A}} \cdot \hat{L} = 0 \quad \underline{\hat{L}} \cdot \underline{\hat{A}} = 0 \Rightarrow \underline{\hat{D}} \cdot \underline{\hat{L}} = \underline{\hat{L}} \cdot \underline{\hat{D}} = 0$$

$$4(\hat{B}_+^2 + \hat{B}_-^2) = \frac{1}{4}\left(\hat{D}^2 + \hat{L}^2\right) = -\frac{Z^2}{2\hat{H}} + 1$$

And also $[\hat{B}_{\pm\alpha}, \hat{B}_{\pm\beta}] = i\sum_\gamma \epsilon_{\alpha\beta\gamma}\hat{B}_{\pm\gamma}$

$$[\hat{B}_{+\alpha}, \hat{B}_{-\beta}] = 0$$

This has the same structure as the angular momentum operator, so we can find a set of states

$$|b, m_+, m_-\rangle$$

Such that

$$\hat{B}_\pm^2|b, m_+, m_-\rangle = m_\pm|b, m_+, m_-\rangle$$

$$\hat{B}_\pm^2|b, m_+, m_-\rangle = b(b+1)|b, m_+, m_-\rangle$$

Also

$$\hat{H} = -\frac{Z^2}{2}\left(4\hat{B}_+^2 - 1\right)^{-1}$$

Denoting $b = \frac{n-1}{2}$ for $n \in \mathbb{N}$, $n \geq 1$

$$4b(b+1) = 4\frac{(n-1)}{2}\frac{(n+1)}{2} = n^2 - 1$$

$$\Rightarrow \hat{H}|b, m_+, m_-\rangle = -\frac{Z^2}{2}\frac{1}{n^2}|b, m_+, m_-\rangle$$

And we arrive at the same eigenvalues with a degeneracy of n^2. The unitary operator

$$\hat{U} = \exp\left[-i\underline{a}_+ \cdot \underline{\hat{B}}_+ - i\underline{a}_- \cdot \underline{\hat{B}}_-\right]$$

commutes with $\hat{H}$ and contains the rotational symmetry with $\underline{a}_+ = \underline{a}_- = \theta\underline{n}$ plus the additional hidden symmetry of $\hat{H}$.

A trick to separate radial and angular terms

Note that the radial distance $\hat{r}$ operator is a rank 0 tensor operator, i.e. a scalar, so it automatically commutes with $\hat{L}_\alpha$. We can write $\hat{L}^2$ as

$$\begin{aligned}\frac{1}{\hat{r}^2}\hat{L^2} &= -\underline{\hat{p}} \times \frac{\underline{\hat{r}}}{\hat{r}} \cdot \frac{\underline{\hat{r}}}{\hat{r}} \times \underline{\hat{p}} \\ &= -\sum_{\alpha\beta} \hat{p}_\alpha \hat{r}_\beta \frac{1}{\hat{r}^2} \hat{r}_\alpha \hat{p}_\beta + \sum_{\alpha\beta} \hat{p}_\alpha \hat{r}_\beta \frac{1}{\hat{r}^2} \hat{r}_\beta \hat{p}_\alpha \\ &= -(\underline{\hat{p}} \cdot \frac{\underline{\hat{r}}}{\hat{r}})(\frac{\underline{\hat{r}}}{\hat{r}} \cdot \underline{\hat{p}}) + \sum_\alpha \hat{p}_\alpha \frac{\hat{r}^2}{\hat{r}^2} \hat{p}_\alpha\end{aligned}$$

The left-hand term has to be hermitian so we can write it as

$$\frac{\underline{\hat{r}}}{\hat{r}} \cdot \underline{\hat{p}} = -i\frac{\partial}{\partial r}$$

so

$$\underline{\hat{p}} \cdot \frac{\underline{\hat{r}}}{\hat{r}} = \frac{\underline{\hat{r}}}{\hat{r}} \cdot \underline{\hat{p}} + \sum_\alpha [\hat{p}_\alpha, \hat{r}_\alpha/\hat{r}]$$

the commutator is

$$\sum_\alpha [\hat{p}_\alpha, \hat{r}_\alpha/\hat{r}] = \sum_\alpha [\hat{p}_\alpha, \hat{r}_\alpha]/\hat{r} + \sum_\alpha \hat{r}_\alpha [\hat{p}_\alpha, 1/\hat{r}] = -3i/r + i/r = -2i/r$$

so

$$\underline{\hat{p}} \cdot \frac{\underline{\hat{r}}}{\hat{r}} = \frac{\underline{\hat{r}}}{\hat{r}} \cdot \underline{\hat{p}} - 2i/\hat{r}$$

so

$$(\underline{\hat{p}} \cdot \frac{\underline{\hat{r}}}{\hat{r}})(\frac{\underline{\hat{r}}}{\hat{r}} \cdot \underline{\hat{p}}) = (\frac{\underline{\hat{r}}}{\hat{r}} \cdot \underline{\hat{p}} - 2i/\hat{r})(\frac{\underline{\hat{r}}}{\hat{r}} \cdot \underline{\hat{p}})$$

Finally, because $\underline{\hat{r}} \cdot \underline{\hat{p}}$ is a scalar operator (it is rotationally invariant) it must only depend on r and derivatives with respect to r. It's fairly straightforward to show that

$$-i\frac{\partial}{\partial r} = -i\sum_\alpha \frac{r_\alpha}{r}\frac{\partial}{\partial r_\alpha} = \frac{\underline{\hat{r}}}{\hat{r}} \cdot \underline{\hat{p}}$$

Putting this together we find

$$(\underline{\hat{p}} \cdot \frac{\underline{\hat{r}}}{\hat{r}})(\frac{\underline{\hat{r}}}{\hat{r}} \cdot \underline{\hat{p}}) = (-i)^2(\frac{\partial}{\partial r} + \frac{2}{r})\frac{\partial}{\partial r} = 2\hat{T}_r = 2\hat{p}_r^2$$

so

$$\begin{aligned}\frac{1}{\hat{r}^2}\hat{L^2} &= -2\hat{p}_r^2 + 2\hat{T} \\ \implies \hat{T} &= \frac{1}{2}\hat{p}_r^2 + \frac{\hat{L}^2}{2\hat{r}^2}\end{aligned}$$

Finding explicit radial functions

The radial functions (often called **spherical harmonics**) can be found by first noting the following expressions for the angular momentum operators in spherical polar coordinates (and atomic units)

$$\begin{aligned}\hat{L}_x &= i\left(\sin\phi\frac{\partial}{\partial\theta} + \cot\theta\cos\phi\frac{\partial}{\partial\phi}\right) \\ \hat{L}_y &= i\left(-\cos\phi\frac{\partial}{\partial\theta} + \cot\theta\sin\phi\frac{\partial}{\partial\phi}\right) \\ \hat{L}_z &= -i\frac{\partial}{\partial\phi}\end{aligned}$$

Noting that the angular momentum eigenstates are eigenstates of $\hat{L}_z$, we immediately find that

$$Y_\ell^m(\theta,\phi) = \frac{1}{\sqrt{2\pi}} e^{im\phi} y_{\ell,m}(\theta)$$

The function of θ can be found by constructing $\hat{L}_\pm = \hat{L}_x \pm i\hat{L}_y$

$$\begin{aligned}\hat{L}_\pm &= i\left((\sin\phi \mp i\cos\phi)\frac{\partial}{\partial\theta} + \cot\theta(\cos\phi \pm i\sin\phi)\frac{\partial}{\partial\phi}\right)\\ &= i\left(\mp ie^{\pm i\phi}\frac{\partial}{\partial\theta} + \cot\theta e^{\pm i\phi}\frac{\partial}{\partial\phi}\right)\end{aligned}$$

The functions $y_{\ell,+\ell}(\theta)$ can be found using $\hat{L}_+ Y_\ell^m(\theta,\phi) = 0$

$$\begin{aligned}\hat{L}_+ Y_\ell^m(\theta,\phi) &= i\left(-ie^{+i\phi}\frac{\partial}{\partial\theta}y_{\ell,\ell}(\theta) + \cot\theta e^{+i\phi}(i\ell)y_{\ell,\ell}(\theta)\right)\frac{e^{i\ell\phi}}{\sqrt{2\pi}} = 0\\ \implies 0 &= \frac{\mathrm{d}}{\mathrm{d}\theta}y_{\ell,\ell}(\theta) - \ell\cot\theta\, y_{\ell,\ell}(\theta)\\ \implies \ln y_{\ell,\ell}(\theta) &= \ell\ln(\sin\theta) + c\\ \implies y_{\ell,\ell}(\theta) &= N_{\ell,\ell}(\sin\theta)^\ell\end{aligned}$$

Integrated from 0 to π with the spherical polar Jacobian (integral weight) factor $\sin\theta$ we have

$$1 = N_{\ell,\ell}^2\int_0^\pi (\sin\theta)^{2\ell}\sin\theta\,\mathrm{d}\theta$$

The integral can be evaluated as

$$1 = N_{\ell,\ell}^2\left(2\frac{(2\ell)!!}{(2\ell+1)!!}\right) \implies N_{\ell,\ell} = (-1)^\ell\sqrt{\frac{(2\ell+1)!!}{2(2\ell)!!}}$$

The apparently random factor of $(-1)^\ell$ is chosen by convention in quantum mechanics. The final result is

$$Y_\ell^\ell(\theta,\phi) = \frac{N_{\ell,\ell}}{\sqrt{2\pi}}(\sin\theta)^\ell e^{i\ell\phi}$$

Successive applications of $\hat{L}_-$ generates the rest of the angular functions. Clearly the ϕ part will be $e^{im\phi}$, and the remaining term is a function of θ. These turn out to be associate Legendre "polynomials" of $\cos\theta$ (they're not strictly polynomials because they often involve square roots) denoted $P_\ell^m(\cos\theta)$. So in general

$$Y_\ell^m(\theta,\phi) = \mathcal{N}_{\ell,m}P_\ell^m(\cos\theta)e^{im\phi}.$$

It should be noted that these have parity (the sign change on inversion of all coordinates $\underline{r} \to -\underline{r}$) of $(-1)^\ell$.

10 Techniques of approximation

Most problems we want to solve in QM are untractable analytically. This means we have to use techniques to approximate solutions. Approximation also often gives us useful physical insight into problems.

10.1 Variational Theorem

The variational theorem is a very powerful result in QM. A lot of quantum chemistry is based on this.

Statement:

For any $|\psi\rangle$ and any Hamiltonian $\hat{H}$ which has a ground state energy eigenvalue E_0, we have

$$\frac{\langle\psi|\,\hat{H}\,|\psi\rangle}{\langle\psi|\psi\rangle} \geq E_0$$

Proof: Consider $\langle\psi|\,\hat{H}\,|\psi\rangle$ and insert an identity operator in the $\hat{H}$ eigenbasis

$$\begin{aligned}
\hat{H}\,|n\rangle &= E_n\,|n\rangle \\
\langle\psi|\,\hat{H}\,|\psi\rangle &= \langle\psi|\,\hat{H}\hat{1}\,|\psi\rangle \\
&= \sum_n \langle\psi|\,\hat{H}\,|n\rangle\,\langle n|\psi\rangle \\
&= \sum_n |\langle n|\psi\rangle|^2 E_n
\end{aligned}$$

Likewise:

$$\langle\psi|\psi\rangle = \sum_n |\langle n|\psi\rangle|^2 \geq 0$$

Now $E_n \geq E_0 \Rightarrow E_n - E_0 \geq 0$, so

$$\begin{aligned}
\langle\psi|\,\hat{H}\,|\psi\rangle &= \sum_n |\langle n|\psi\rangle|^2[(E_n - E_0) + E_0] \\
&= E_0 \sum_n |\langle n|\psi\rangle|^2 + \sum_n |\langle n|\psi\rangle|^2 (E_n - E_0) \\
&\geq E_0 \sum_n |\langle n|\psi\rangle|^2 \quad \because (E_n - E_0) \geq 0 \text{ and } |\langle n|\psi\rangle|^2 \geq 0
\end{aligned}$$

Combining this with $\langle\psi|\psi\rangle \geq 0$ we find

$$\frac{\langle\psi|\,\hat{H}\,|\psi\rangle}{\langle\psi|\psi\rangle} \geq \frac{E_0 \sum_n |\langle n|\psi\rangle|^2}{\sum_n |\langle n|\psi\rangle|^2} = E_0 \quad \square$$

Consequences: If we have a trial wavefunction which depends on parameters $(\alpha_1, \alpha_2, \ldots) = \underline{\alpha}$, then minimizing

$$\langle E(\underline{\alpha})\rangle = \frac{\langle\psi(\underline{\alpha})|\,\hat{H}\,|\psi(\underline{\alpha})\rangle}{\langle\psi(\underline{\alpha})|\psi(\underline{\alpha})\rangle} \geq E_0$$

So minimizing $\langle E(\underline{\alpha})\rangle$ will give the best possible estimate of E_0.

Example: LCAO Theory

Suppose we ignore e-e repulsion and consider only electrons in a molecule. The Hamiltonian is

$$\begin{aligned}\hat{H} &= \hat{T} + \sum_A \frac{Z_A}{|\hat{\underline{r}} - \underline{R}_A|} \\ &= \hat{T} + \sum_A \hat{V}_A\end{aligned}$$

Close to atom A: $\hat{H} \approx \hat{T} + \hat{V}_A = \hat{H}_A$. So the atomic orbital solutions are good guesses for the exact eigenstates.

We take as a guess

$$|\psi(c)\rangle = \sum_p c_p |\chi_p\rangle$$

where $c_p \in \mathbb{R}$.

$$\begin{aligned}\langle\psi| \hat{H} |\psi\rangle &= \sum_{pq} c_p c_q \langle\chi_p| \hat{H} |\chi_q\rangle \\ &= \sum_{pq} c_p c_q H_{pq} \\ \langle\psi|\psi\rangle &= \sum_{pq} c_p c_q \langle\chi_p|\chi_q\rangle \\ &= \sum_{pq} c_p c_q S_{pq}\end{aligned}$$

Note: $\langle\chi_p|\chi_q\rangle \neq \delta_{pq}$!

We also take $\chi_p(\underline{r}) \in \mathbb{R}$ so $S_{pq} = S_{qp}$, $H_{pq} = H_{qp}$

Applying the variational theorem we find

$$\begin{aligned}0 = \frac{\partial}{\partial c_m}\langle E\rangle &= \frac{1}{\langle\psi|\psi\rangle}\sum_{pq}\frac{\partial}{\partial c_m} c_p c_q H_{pq} - \frac{\langle E\rangle}{\langle\psi|\psi\rangle}\sum_{pq}\frac{\partial}{\partial c_m} c_p c_q S_{pq} \\ &= \frac{1}{\langle\psi|\psi\rangle}\sum_p (c_p H_{pm} + H_{mp} c_p) - \frac{\langle E\rangle}{\langle\psi|\psi\rangle}\sum_p (c_p S_{pm} + S_{mp} c_p) \\ \Rightarrow \quad &\sum_p H_{mp} c_p = \langle E\rangle \sum_p S_{mp} c_p \quad \text{for all } m\end{aligned}$$

This is satisfied by the generalized eigenvalue problem eigenvectors and eigenvalues:

$$\underline{\underline{H}}\ \underline{c}_i = \lambda_i \underline{\underline{S}}\ \underline{c}_i$$

with $\langle E\rangle = \lambda_i$

Note: $\underline{\underline{H}}\underline{c}_i = \lambda_i \underline{\underline{S}}\underline{c}_i$ has the same properties as a normal Hamiltonian eigenvalue problem.

Define:

$$\tilde{\underline{c}}_i = \underline{\underline{S}}^{-\frac{1}{2}} \underline{c}_i \qquad\qquad \underline{\underline{\tilde{H}}} = \underline{\underline{S}}^{-\frac{1}{2}} \underline{\underline{H}}\, \underline{\underline{S}}^{-\frac{1}{2}}$$

And we find:

$$\begin{aligned} & \underline{\underline{\tilde{H}}}\, \underline{\underline{S}}^{-\frac{1}{2}} \underline{c}_i = \lambda_i \underline{\underline{S}}^{-\frac{1}{2}} \underline{c}_i \\ \Rightarrow \quad & \underline{\underline{S}}^{-\frac{1}{2}} \underline{\underline{\tilde{H}}}\, \underline{\underline{S}}^{-\frac{1}{2}} \tilde{\underline{c}}_i = \lambda_i \tilde{\underline{c}}_i \\ \Rightarrow \quad & \underline{\underline{\tilde{H}}}\, \tilde{\underline{c}}_i = \lambda_i \tilde{\underline{c}}_i \end{aligned}$$

So a generalized eigenvalue problem is just a regular eigenvalue problem with a transformation. These are sometimes called the **Secular equations**.

Example: H_2^+.

Let's consider a single electron with two protons, i.e. the simplest possible molecule H_2^+. We know the ground state should resemble a hydrogen 1s orbital close to each of the two protons, A and B. So we expand the wave function as a linear combination of $|1s_A\rangle$ and $|1s_B\rangle$. The secular equations for the two 1s orbitals are

$$\begin{pmatrix} \alpha & \beta \\ \beta & \alpha \end{pmatrix} \begin{pmatrix} c_A \\ c_B \end{pmatrix} = E \begin{pmatrix} 1 & S \\ S & 1 \end{pmatrix} \begin{pmatrix} c_A \\ c_B \end{pmatrix}$$

where $\alpha = \langle 1s_A|\hat{H}|1s_A\rangle = \langle 1s_B|\hat{H}|1s_B\rangle$, $\beta = \langle 1s_A|\hat{H}|1s_B\rangle = \langle 1s_B|\hat{H}|1s_A\rangle$ and $S = \langle 1s_A|1s_B\rangle = \langle 1s_B|1s_A\rangle$. Symmetry of the H_2^+ molecule means that $\langle 1s_A|\hat{H}|1s_A\rangle = \langle 1s_B|\hat{H}|1s_B\rangle$. Also note that $\alpha < 0$ and $\beta < 0$ and $0 < S < 1$.

The Secular equations are solved by noting that $\underline{\underline{H}} - E\underline{\underline{S}}$ is not invertible for the non-trivial $\underline{c}$ solutions, just like in a normal eigenvalue problem. This means the determinant of $\underline{\underline{H}} - E\underline{\underline{S}}$ is zero for the eigenvalues E.

$$\begin{aligned} 0 &= \det\left(\underline{\underline{H}} - E\underline{\underline{S}}\right) \\ &= \begin{vmatrix} \alpha - E & \beta - SE \\ \beta - SE & \alpha - E \end{vmatrix} \\ &= (\alpha - E)^2 - (\beta - SE)^2 \\ &= \alpha^2 - \beta^2 - 2(\alpha + S\beta)E + (1 - S^2)E^2 \\ &= (\alpha - \beta - (1 - S)E)(\alpha + \beta - (1 + S)E) \end{aligned}$$

so the eigenvalues are

$$E_\pm = \frac{\alpha \pm \beta}{1 \pm S}$$

Solving the secular equations with these eigenvalues gives

$$\underline{c}_\pm = \frac{1}{\sqrt{2(1 \pm S)}} \begin{pmatrix} 1 \\ \pm 1 \end{pmatrix}.$$

Because $\alpha < 0$ and $\beta < 0$ and $0 < S < 1$, the lowest energy solution is E_+. This corresponds to a σ bonding molecular orbital.

Note that as the separation between the nuclei R increases, the diagonal Hamiltonian matrix elements just become the bare hydrogen atommatrix elements, which are just the 1s energies, so $\alpha \to E_{1s}$. Likewise the overlap vanishes, so $\beta \to 0$ and $S \to 0$ as $R \to \infty$. This means overall $E_\pm \to E_{1s}$, which is the exact ground state energy in this limit.

10.2 Perturbation theory

Often we can divide a QM problem into an easily solved part $\hat{H}_0$, and a remainder, $\lambda\hat{V}$, where λ controls the strength of the remaining part.

$$\hat{H} = \hat{H}_0 + \lambda\hat{V}$$

We assume we can find the eigenstates of $\hat{H}_0$ easily, so we know $\{\left|n^{(0)}\right\rangle\}$.

$$\hat{H}_0 \left|n^{(0)}\right\rangle = E_n^{(0)} \left|n^{(0)}\right\rangle$$

Because $\{\left|n^{(0)}\right\rangle\}$ is complete, we can write the full $|n\rangle$ as

$$|n\rangle = \sum_m c_{n,m} \left|m^{(0)}\right\rangle$$

Normalization is arbitrary so we take (for ease later)

$$\left\langle n^{(0)}\middle|n\right\rangle = 1$$

So

$$|n\rangle = \left|n^{(0)}\right\rangle + \sum_{m\neq n} c_{n,m} \left|m^{(0)}\right\rangle$$

We also know $|n\rangle$ has some expansion in λ

$$|n\rangle = \left|n^{(0)}\right\rangle + \lambda \left|n^{(1)}\right\rangle + \lambda^2 \left|n^{(2)}\right\rangle + \cdots$$

And this satisfies $\hat{H}|n\rangle = E_n|n\rangle$. This means that every term in the expansion must satisfy this equation.

$$\begin{aligned}
& & (\hat{H}_0 + \lambda\hat{V})|n\rangle &= E_n|n\rangle \\
\Rightarrow\quad & \hat{H}_0\left(\left|n^{(0)}\right\rangle + \lambda\left|n^{(1)}\right\rangle + \cdots\right) + \lambda\hat{V}\left(\left|n^{(0)}\right\rangle + \lambda\left|n^{(1)}\right\rangle + \cdots\right) &&= \hat{H}_0\left|n^{(0)}\right\rangle \\
& && + \lambda(\hat{H}_0\left|n^{(1)}\right\rangle + \hat{V}\left|n^{(0)}\right\rangle) \\
& && + \lambda^2(\hat{H}_0\left|n^{(2)}\right\rangle + \hat{V}\left|n^{(1)}\right\rangle) \\
& && + \cdots \\
& && = E_n\left(\left|n^{(0)}\right\rangle + \lambda\left|n^{(1)}\right\rangle + \cdots\right)
\end{aligned}$$

Likewise the energy can be expanded as

$$E_n = E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \cdots$$

With this $E_n|n\rangle$ is

$$
\begin{aligned}
E_n \left|n\right\rangle = & E_n^{(0)} \left|n^{(0)}\right\rangle + \lambda(E_n^{(0)} \left|n^{(1)}\right\rangle + E_n^{(1)} \left|n^{(0)}\right\rangle) \\
& + \lambda^2(E_n^{(0)} \left|n^{(2)}\right\rangle + E_n^{(1)} \left|n^{(1)}\right\rangle + E_n^{(2)} \left|n^{(0)}\right\rangle) \\
& + \cdots
\end{aligned}
$$

So we find by comparing coefficients:

$$
\begin{aligned}
&\lambda^0 : E_n^{(0)} \left|n^{(0)}\right\rangle = \hat{H}_0 \left|n^{(0)}\right\rangle \\
&\lambda^1 : E_n^{(1)} \left|n^{(0)}\right\rangle + E_n^{(0)} \left|n^{(1)}\right\rangle = \hat{H}_0 \left|n^{(1)}\right\rangle + \hat{V} \left|n^{(0)}\right\rangle \\
&\lambda^2 : E_n^{(2)} \left|n^{(0)}\right\rangle + E_n^{(1)} \left|n^{(1)}\right\rangle + E_n^{(0)} \left|n^{(2)}\right\rangle = \hat{H}_0 \left|n^{(2)}\right\rangle + \hat{V} \left|n^{(1)}\right\rangle \\
&\vdots \\
&\lambda^p : \sum_{k=0}^{p} E_n^{(p-k)} \left|n^{(k)}\right\rangle = \hat{H}_0 \left|n^{(p)}\right\rangle + \hat{V} \left|n^{(p-1)}\right\rangle
\end{aligned}
$$

We note that

$$
\left\langle n^{(0)} \middle| n \right\rangle = 1 \Rightarrow \left\langle n^{(0)} \middle| n^{(k)} \right\rangle = 0 \quad \text{for} \quad k > 0
$$

So we can produce the equations above with $\left\langle n^{(0)}\right|$ for $p > 0$

$$
\begin{aligned}
\left\langle n^{(0)}\right| \textstyle\sum_{k=0}^{p} E_n^{(p-k)} \left|n^{(k)}\right\rangle &= \left\langle n^{(0)}\right| (\hat{H}_0 \left|n^{(p)}\right\rangle + \hat{V} \left|n^{(p-1)}\right\rangle) \\
\Rightarrow \left\langle n^{(0)} \middle| n^{(0)} \right\rangle E_n^{(p)} &= E_n^{(0)} \left\langle n^{(0)} \middle| n^{(p)} \right\rangle + \left\langle n^{(0)}\right| \hat{V} \left|n^{(p-1)}\right\rangle \\
\Rightarrow E_n^{(p)} &= \left\langle n^{(0)}\right| \hat{V} \left|n^{(p-1)}\right\rangle
\end{aligned}
$$

So $E_n^{(p)}$ depends on $\left|n^{(p-1)}\right\rangle$. How do we find $\left|n^{(p-1)}\right\rangle$?

Consider the λ^1 equation above. Let's rearrange this to

$$
(E_n^{(0)} - \hat{H}_0) \left|n^{(1)}\right\rangle = -E_n^{(1)} \left|n^{(0)}\right\rangle + \hat{V} \left|n^{(0)}\right\rangle
$$

Project this onto $\left\langle m^{(0)}\right|$, $m \neq n$, and we find

$$
\begin{aligned}
\left\langle m^{(0)} \right| (E_n^{(0)} - \hat{H}_0) \left|n^{(1)}\right\rangle &= -E_n^{(1)} \left\langle m^{(0)} \middle| n^{(0)} \right\rangle + \left\langle m^{(0)} \right| \hat{V} \left|n^{(0)}\right\rangle \\
\Rightarrow (E_n^{(0)} - E_m^{(0)}) \left\langle m^{(0)} \middle| n^{(1)} \right\rangle &= \left\langle m^{(0)} \right| \hat{V} \left|n^{(0)}\right\rangle
\end{aligned}
$$

If $E_m^{(0)} \neq E_n^{(0)}$ for all m.

$$
\Rightarrow \left|n^{(1)}\right\rangle = \sum_{m \neq n} \left|m^{(0)}\right\rangle \frac{\left\langle m^{(0)}\right| \hat{V} \left|n^{(0)}\right\rangle}{E_n^{(0)} - E_m^{(0)}}
$$

We can repeat this for the order p equation

$$
\begin{aligned}
(E_n^{(0)} - \hat{H}_0) \left|n^{(p)}\right\rangle &= -\sum_{k=0}^{p-1} E_n^{(p-k)} \left|n^{(k)}\right\rangle + \hat{V} \left|n^{(p-1)}\right\rangle \\
&= -\sum_{k=0}^{p-1} \left|n^{(k)}\right\rangle \langle n^{(0)}|\hat{V}|n^{(k-1)}\rangle + \hat{V} \left|n^{(p-1)}\right\rangle
\end{aligned}
$$

$$\left\langle m^{(0)}\middle|n^{(p)}\right\rangle = \frac{\left\langle m^{(0)}\right|\hat{V}\left|n^{(p-1)}\right\rangle}{E_n^{(0)}-E_m^{(0)}} - \sum_{k=0}^{p-1}\frac{\left\langle m^{(0)}\middle|n^{(k)}\right\rangle\left\langle n^{(0)}\right|\hat{V}\left|n^{(p-k-1)}\right\rangle}{E_n^{(0)}-E_m^{(0)}}$$

In practice we rarely go beyond 2nd order in perturbation theory.

$$E_n^{(1)} = \left\langle n^{(0)}\right|\hat{V}\left|n^{(0)}\right\rangle$$

$$\left|n^{(1)}\right\rangle = \sum_{m\neq n}\left|m^{(0)}\right\rangle\frac{\left\langle m^{(0)}\right|\hat{V}\left|n^{(0)}\right\rangle}{E_n^{(0)}-E_m^{(0)}}$$

$$E_n^{(2)} = \left\langle n^{(0)}\right|\hat{V}\left|n^{(1)}\right\rangle = \sum_{m\neq n}\frac{\left|\left\langle m^{(0)}\right|\hat{V}\left|n^{(0)}\right\rangle\right|^2}{E_n^{(0)}-E_m^{(0)}}$$

Note that if $E_m^{(0)} > E_n^{(0)}$ for m then $E_n^{(2)} \leq 0$, so 2nd order corrections to the ground state always reduce the energy.

10.2.1 Example: The second order Stark effect

When an electric field is applied to a hydrogen atom along the z direction, the electron feels a new Hamiltonian (in atomic units where $e = 1$).

$$\hat{V} = -\mathcal{E}\hat{z}$$

For the ground level, $n = 1$, i.e. when the electron is in the 1s orbital, we can use perturbation theory to evaluate the effect of the electric field on the electron.

By inversion symmetry of the 1s orbital wave-function (or alternatively by recognising that $\hat{V}$ is a rank 1 spherical tensor operator and using Wigner-Eckart theorem) for the 1s wave function $|1\mathrm{s}\rangle \equiv |n=1,\ell=0,m_\ell=0\rangle$ the first order perturbation energy is zero

$$E_{1\mathrm{s}}^{(1)} = -\mathcal{E}\,\langle 1\mathrm{s}|\hat{z}|1\mathrm{s}\rangle = 0$$

The second order energy is non-zero however

$$E_{1\mathrm{s}}^{(2)} = -\mathcal{E}^2\sum_{n\geq 2}\frac{|\,\langle 1\mathrm{s}|\hat{z}|n,\ell=1,m_\ell=0\rangle\,|^2}{E_n-E_1}$$

The fact that $\hat{V} = \hat{r}\hat{T}_0^{(1)}$ means that only $\ell = 1$ and $m_\ell = 0$ terms are non-zero in the second order energy. [Note that the sum over "$n \geq 2$" includes both the bound states with $E_n < 0$ and continuum states with $E \geq 0$.]

We can also work out first order properties using the first-order wave-function. For example the dipole moment $\hat{\mu}_z = -\hat{z}$ is given by

$$\begin{aligned}\langle\mu_z\rangle^{(1)} &= \langle 1\mathrm{s}|\hat{\mu}_z|\psi_{1\mathrm{s}}^{(1)}\rangle + \langle\psi_{1\mathrm{s}}^{(1)}|\hat{\mu}_z|1\mathrm{s}\rangle \\ &= \mathcal{E}\left(2\sum_{n\geq 2}\frac{|\,\langle 1\mathrm{s}|\hat{z}|n,\ell=1,m_\ell=0\rangle\,|^2}{E_n-E_1}\right)\end{aligned}$$

From this we can identify the polarisability α of the electron in the hydrogen atom from $\langle\mu_z\rangle = \alpha\mathcal{E}$

$$\alpha = 2\sum_{n\geq 2}\frac{|\,\langle 1\mathrm{s}|\hat{z}|n,\ell=1,m_\ell=0\rangle\,|^2}{E_n-E_1}.$$

The perturbation of energy levels of a system by an external electric field is called the **Stark Effect**. Because in this case the lowest order non-zero perturbation arises at second order in the electric field, this is called a second-order stark effect.

10.3 An alternative derivation

An alternative approach to derive perturbation theory goes as follows. This is based on **projection operators**. An operator $\hat{P}$ is a projection operator if it is **idempotent**

$$\hat{P}^2 = \hat{P}$$

If $\hat{P}$ is a projection operator it is straightforward to show that $\hat{Q} = \hat{1} - \hat{P}$ is a projection operator as well and

$$\hat{P}\hat{Q} = \hat{Q}\hat{P} = 0$$

Let us take the energy eigenvalue equation divided into a reference and perturbation term as above

$$\hat{H}_0 |n\rangle + \lambda \hat{V} |n\rangle = E_n |n\rangle$$

Now let us define the projection operator $\hat{P} = |n^{(0)}\rangle\langle n^{(0)}|$. It's straightforward to verify that $[\hat{P}, \hat{H}_0] = 0$ and $[\hat{Q}, \hat{H}_0] = 0$. Now let's project the eigenvalue equation with $\hat{Q}$

$$\hat{H}_0\hat{Q} |n\rangle + \lambda \hat{Q}\hat{V} |n\rangle = E_n \hat{Q} |n\rangle$$

We can rearrange this to find

$$\hat{Q} |n\rangle = \lambda (\hat{H}_0 - E_n)^{-1} \hat{Q}\hat{V} |n\rangle$$

Assuming E_n does not coincide with any of the eigenvalues of $\hat{H}_0$ the inverse is well-defined. Now using $\hat{1} = \hat{P} + \hat{Q}$ we can write $|n\rangle$ as

$$\begin{aligned} |n\rangle &= \hat{P} |n\rangle + \hat{Q} |n\rangle \\ &= |n^{(0)}\rangle + \lambda \hat{Q} (\hat{H}_0 - E_n)^{-1} \hat{Q}\hat{V} |n\rangle \end{aligned}$$

where we used the normalisation $\langle n^{(0)}|n\rangle = 1$. We can rearrange this to find

$$\begin{aligned} (\hat{1} - \lambda \hat{Q} (\hat{H}_0 - E_n)^{-1} \hat{Q}\hat{V}) |n\rangle &= |n^{(0)}\rangle \\ \implies |n\rangle &= (\hat{1} - \lambda \hat{Q} (\hat{H}_0 - E_n)^{-1} \hat{Q}\hat{V})^{-1} \left|n^{(0)}\right\rangle \end{aligned}$$

We can expand the inverse using a Taylor series $(\hat{1} - \lambda \hat{A})^{-1} = \sum_{k=1}^{\infty} \lambda^k \hat{A}^k$ which immediately gives an expansion for $|n\rangle$.

$$|n\rangle = |n^{(0)}\rangle + \lambda \hat{Q} (\hat{H}_0 - E_n)^{-1} \hat{Q}\hat{V} |n^{(0)}\rangle + \lambda^2 \hat{Q} (\hat{H}_0 - E_n)^{-1} \hat{Q}\hat{V}\hat{Q} (\hat{H}_0 - E_n)^{-1} \hat{Q}\hat{V} |n^{(0)}\rangle + \cdots$$

We're not quite done yet though because E_n is itself a function of λ.

In order to obtain the energy as a function of λ, we rearrange the original eigenvalue equation and project with $\left\langle n^{(0)}\right|$ to give

$$\begin{aligned} (E_n - \hat{H}_0) |n\rangle &= \lambda \hat{V} |n\rangle \\ E_n - E_n^{(0)} &= \lambda \langle n^{(0)}|\hat{V}|n\rangle \end{aligned}$$

We can then insert what we had above for $|n\rangle$ to obtain

$$E_n = E_n^{(0)} + \lambda \langle n^{(0)}| \hat{V} \left(|n^{(0)}\rangle + \lambda \hat{Q} (\hat{H}_0 - E_n)^{-1} \hat{Q}\hat{V} |n^{(0)}\rangle + \lambda^2 \hat{Q} (\hat{H}_0 - E_n)^{-1} \hat{Q}\hat{V}\hat{Q} (\hat{H}_0 - E_n)^{-1} \hat{Q}\hat{V} |n^{(0)}\rangle + \cdots \right)$$

By inserting the same expression for E_n into this equation, we can find the same perturbation series for E_n as above. With this we can then also find the perturbation series for $|n\rangle$ using the expressions above.

10.4 Degenerate Perturbation theory

The above equations assumed $E_n^{(0)}$ is non-degenerate, so $(E_m^{(0)} - E_n^{(0)})^{-1}$ is well defined. What if this is not the case?

Suppose there is some set of states $\{\left|n^{(0)}, k\right\rangle\}$ which all have the eigenvalue $E_n^{(0)}$ for H_0.

We choose $\left\langle n^{(0)}, k\middle| n, k\right\rangle = 1$ but $\left\langle n^{(0)}, \ell\middle| n, k\right\rangle \neq 0$ for $k \neq \ell$.

Consider the 1st equation we had before

$$E_{n,k}^{(1)} \left|n^{(0)}, k\right\rangle + E_n^{(0)} \left|n^{(1)}, k\right\rangle = \hat{H}_0 \left|n^{(1)}, k\right\rangle + \hat{V} \left|n^{(0)}, k\right\rangle$$

Project this with $\left\langle n^{(0)}, \ell\right|$ for $\ell \neq k$

$$\begin{aligned}
\left(H_0 - E_n^{(0)}\right) \left|n^{(1)}, k\right\rangle &= \left(-E_{n,k}^{(1)} + \hat{V}\right) \left|n^{(0)}, k\right\rangle \\
\langle n^{(0)}, \ell| \left(\hat{H}_0 - E_n^{(0)}\right) |n^{(1)}, k\rangle &= \langle n^{(0)}, \ell| \left(-E_{n,k}^{(1)} + \hat{V}\right) |n^{(0)}, k\rangle \\
\Rightarrow \left(E_n^{(0)} - E_n^{(0)}\right) \left\langle n^{(0)}, \ell\middle| n^{(1)}, k\right\rangle &= \langle n^{(0)}, \ell| \hat{V} |n^{(0)}, k\rangle \\
\Rightarrow \langle n^{(0)}, \ell| \hat{V} |n^{(0)}, k\rangle &= 0 \quad \text{for } \ell \neq k.
\end{aligned}$$

So $\left|n^{(0)}, k\right\rangle$ must diagonalize $\hat{V}$ in the $E_n^{(0)}$ subspace for perturbation theory to work. Projecting with $\left\langle n^{(0)}, k\right|$ gives ht same result as before

$$E_{nk}^{(1)} = \langle n^{(0)}, k| \hat{V} |n^{(0)}, k\rangle$$

as long as $\left|n^{(0)}, k\right\rangle$ diagonalises $\hat{V}$ in the $E_n^{(0)}$ subspace.

Projecting the 1st equation with $\left\langle m^{(0)}\right|$ gives the same expression for $\left\langle m^{(0)}\middle| n^{(1)}, k\right\rangle$ as before

$$\left\langle m^{(0)}\middle| n^{(1)}, k\right\rangle = \frac{\langle m^{(0)}| \hat{V} |n^{(0)}, k\rangle}{E_n^{(0)} - E_m^{(0)}}$$

But we need to find $\left\langle n^{(0)}, \ell\middle| n^{(1)}, k\right\rangle$. Consider the 2nd equation

$$E_{nk}^{(2)} (E_n^{(0)} - \hat{H}_0) \left|n^{(1)}, k\right\rangle = (\hat{V} - E_{nk}^{(1)}) \left|n^{(1)}, k\right\rangle - E_{nk}^{(2)} \left|n^{(0)}, k\right\rangle$$

Projecting with $\left\langle n^{(0)}, \ell\right|$ gives

$$0 = \langle n^{(0)}, \ell| \hat{V} |n^{(1)}, k\rangle - E_{nk}^{(1)} \left\langle n^{(0)}, \ell\middle| n^{(1)}, k\right\rangle$$

From the above we have

$$\left|n^{(1)}, k\right\rangle = \sum_{m \neq n} \left|m^{(0)}\right\rangle \frac{\langle m^{(0)}| \hat{V} |n^{(0)}, k\rangle}{E_m^{(0)} - E_n^{(0)}} + \sum_{\ell \neq k} C_\ell^{(1)} \left|n^{(0)}, \ell\right\rangle$$

So

$$\begin{aligned}
0 = &\sum_{m \neq n} \frac{\langle m^{(0)}| \hat{V} |n^{(0)}, k\rangle \ \langle n^{(0)}, \ell| \hat{V} |m^{(0)}\rangle}{E_m^{(0)} - E_n^{(0)}} \\
&+ \sum_{\ell' \neq k} \langle n^{(0)}, \ell| \hat{V} |n^{(0)}, \ell'\rangle \, C_{\ell'}^{(1)} - E_{nk}^{(1)} C_\ell^{(1)}
\end{aligned}$$

$$\Rightarrow C_\ell = \frac{1}{E^{(1)}_{nk} - E^{(1)}_{n\ell}} \sum_{m \neq n} \frac{\langle m^{(0)}|\hat{V}|n^{(0)}, k\rangle \ \langle n^{(0)}, \ell|\hat{V}|m^{(0)}\rangle}{E^{(0)}_m - E^{(0)}_n}$$

And projection with $\langle n^{(0)}, k|$ gives

$$\begin{aligned} E^{(2)}_{nk} &= \langle n^{(0)}, k|\hat{V}|n^{(1)}, k\rangle \\ &= \sum_{m \neq n} \frac{|\langle n^{(0)}, k|\hat{V}|n^{(0)}, k\rangle|^2}{E^{(0)}_n - E^{(0)}_m} \end{aligned}$$

Noting the contribution from $\langle n^{(0)}, \ell|n^{(1)}, k\rangle$ is zero because $\langle n^{(0)}, k|\hat{V}|n^{(0)}, \ell\rangle = 0$.

10.5 The projection operator approach

Projection operators make degenerate perturbation theory a lot easier. We state by defining $\hat{P}$ as the projection onto the entire degenerate subspace with eigenvalue E_n

$$\hat{P} = \sum_k |n^{(0)}, k\rangle\langle n^{(0)}, k|$$

Using this, the derivation above follows the same way to obtain an equation for $\hat{P}\,|n, k\rangle$

$$|n, k\rangle = (\hat{1} - \lambda\hat{Q}(\hat{H}_0 - E_n)^{-1}\hat{Q}\hat{V})^{-1}\hat{P}\,|n^{(0)}, k\rangle$$

We also require that

$$\begin{aligned} \langle n^{(0)}, \ell|\,(E_{n,k} - \hat{H}_0)\,|n, k\rangle &= \lambda\,\langle n^{(0)}, \ell|\,\hat{V}\,|n, k\rangle \\ (E_{n,k} - E^{(0)}_n)\,\langle n^{(0)}, \ell|n, k\rangle &= \lambda\,\langle n^{(0)}, \ell|\hat{V}|n, k\rangle \end{aligned}$$

Inserting the above expression for $|n, k\rangle$ we obtain

$$\begin{aligned} (E_{n,k} - E^{(0)}_n)\,\langle n^{(0)}, \ell|n, k\rangle &= \lambda\,\langle n^{(0)}, \ell|\,\hat{V}(\hat{1} - \lambda\hat{Q}(\hat{H}_0 - E_n)^{-1}\hat{Q}\hat{V})^{-1}\hat{P}\,|n^{(0)}, k\rangle & (10) \\ &= \lambda\,\langle n^{(0)}, \ell|\,\hat{V}\hat{P}\,|n, k\rangle + \lambda^2\,\langle n^{(0)}, \ell|\,\hat{V}\hat{Q}(\hat{H}_0 - E_n)^{-1}\hat{Q}\hat{V}\hat{P}\,|n, k\rangle + \cdots & (11) \end{aligned}$$

Now expanding $\langle n^{(0)}, \ell|n, k\rangle$ and $E_{n,k}$ as a series in λ we can find the perturbation series for the energies easily. This gives us the same condition as above that $\langle n^{(0)}, \ell|\hat{V}|n^{(0)}, k\rangle = (E^{(1)}_{n,k} - E^{(0)}_n)\delta_{\ell,k}$, so we must diagonalise $\hat{V}$ in the degenerate subspace to find the first order energies. This procedure also gives a transparent way to find a perturbation series if $\langle n^{(0)}, \ell|\hat{V}|n^{(0)}, k\rangle = V_n\delta_{\ell,k}$, i.e. if there is still degeneracy left after diagonalising $\hat{V}$ in the degenerate subspace, i.e. if degeneracy is only lifted at higher orders in λ.

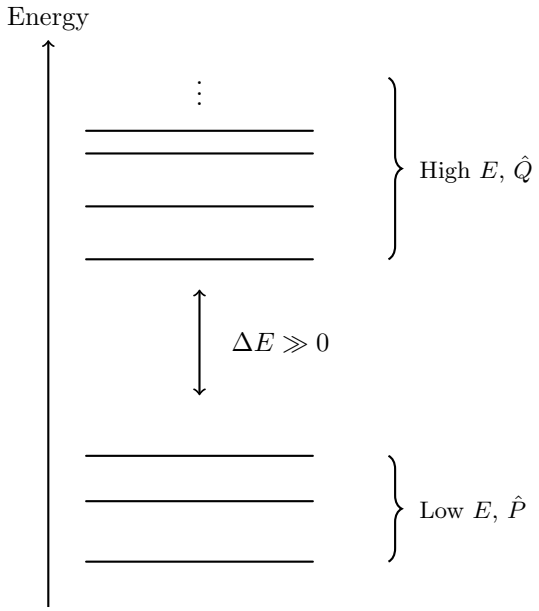
10.6 Effective Hamiltonian theory

The projection operator formalism also gives an alternative approach to approximating the eigenvalues of a complex Hamiltonian.

Let's assume as before that we can divide the Hamiltonian into a reference part $\hat{H}_0$ and a perturbation part $\lambda\hat{V}$. Suppose $\hat{H}_0$ has a set of states in a small spectral window, for example the ground state and set of low-lying excited states. We'll denote these $|p^{(0)}\rangle$ for $p \in \mathcal{P}$, and we'll denote the projection operator $\hat{P}$ as the projection onto these states

$$\hat{P} = \sum_{p \in \mathcal{P}} |p^{(0)}\rangle\langle p^{(0)}|$$

From now on we will use indices p, p', p'' etc. to reference exclusively to these states. The remaining state will be denoted $|q^{(0)}\rangle$ and likewise $\hat{Q} = \hat{1} - \hat{P}$ will be a projection operator onto these states.



Our aim is to find an approximate expression for the low energy states of the Hamiltonian, i.e. eigenstates with energies close to those of $|p^{(0)}\rangle$. For these states

$$\|\hat{P}\,|\psi\rangle\,\| \approx \|\hat{P}\,|\psi\rangle\,\| \iff \|\hat{Q}\,|\psi\rangle\,\| \approx 0 \tag{12}$$

Let's project the eigenvalue equation as before

$$\begin{aligned} E\hat{P}\,|\psi\rangle &= \hat{H}_0\hat{P}\,|\psi\rangle + \lambda\hat{P}\hat{V}\,|\psi\rangle \\ E\hat{Q}\,|\psi\rangle &= \hat{H}_0\hat{Q}\,|\psi\rangle + \lambda\hat{Q}\hat{V}\,|\psi\rangle \end{aligned}$$

Inserting $\hat{1} = \hat{P} + \hat{Q}$ into the last term of the second equation we obtain

$$E\hat{Q}\,|\psi\rangle = \hat{H}_0\hat{Q}\,|\psi\rangle + \lambda\hat{Q}\hat{V}(\hat{P} + \hat{Q})\,|\psi\rangle$$

and we can solve this for $\hat{Q}\,|\psi\rangle$, which we are asserting should be small

$$\hat{Q}\,|\psi\rangle = \lambda(E - \hat{H}_0\hat{Q} + \lambda\hat{Q}\hat{V}\hat{Q})^{-1}\hat{Q}\hat{V}\hat{P}\,|\psi\rangle$$

Now inserting $\hat{1} = \hat{P} + \hat{Q}$ into the last term of the $\hat{P}$ projected equation above we find

$$E\hat{P}\,|\psi\rangle = \hat{H}_0\hat{P}\,|\psi\rangle + \lambda\hat{P}\hat{V}\hat{P}\,|\psi\rangle + \lambda\hat{P}\hat{V}\hat{Q}\,|\psi\rangle$$

and we can eliminate $\hat{Q}\,|\psi\rangle$ from this equation using

$$E\hat{P}\,|\psi\rangle = \hat{H}_0\hat{P}\,|\psi\rangle + \lambda\hat{P}\hat{V}\hat{P}\,|\psi\rangle + \lambda^2\hat{P}\hat{V}(E - \hat{H}_0\hat{Q} + \lambda\hat{Q}\hat{V}\hat{Q})^{-1}\hat{Q}\hat{V}\hat{P}\,|\psi\rangle$$

From this equation we can do many things. Perhaps the simplest thing is to do is to truncate the equation at first order in the perturbation λ. This reduces to the **first order effective Hamiltonian** equation

$$\begin{aligned} E\hat{P}\,|\psi\rangle &= \hat{H}^{(1)}_{\text{eff}}\hat{P}\,|\psi\rangle \\ \hat{H}^{(1)}_{\text{eff}} &= \hat{H}_0\hat{P} + \lambda\hat{P}\hat{V}\hat{P} \end{aligned}$$

The energies from this equation are equivalent to those in first order degenerate perturbation theory, apart from we do not assume the set of states $|p^{(0)}\rangle$ are strictly degenerate. This is sometimes also called **quasi-degenerate** perturbation theory.

If we want to include second order effects we need to account for the second term. This is formally a function of energy, so the full (all orders in λ) $\hat{H}_{\text{eff}}(E)$ is a function of the energy eigenvalues E, so the equation is no longer a simple linear eigenvalues equation. Note that in condensed matter physics, and some quantum chemistry literature, the $\hat{\Sigma}(E) = \hat{H}_{\text{eff}}(E) - \hat{H}_{\text{eff}}^{(1)}$ is sometimes referred to as a **self-energy**.

In order to simplify this, we can approximate E as some characteristic value for the $\hat{P}$ space E_0. With this equation the second order effective Hamiltonian is given by

$$\hat{H}_{\text{eff}}^{(2)} = \hat{H}_{\text{eff}}^{(1)} + \lambda^2 \hat{P}\hat{V}\hat{Q}(E_0 - \hat{H}_0\hat{Q})^{-1}\hat{Q}\hat{V}\hat{P}$$

Alternatively we can avoid replacing E with a constant E_0, and instead iteratively solve the non-linear equation. This gives a result known as **Brilllouin-Wigner** perturbation theory. Note another approach can be taken to more rigorously obtain a pertubative linear eigenvalue equation known as the **Schrieffer-Wolff** transform.