

PH20013/60 Quantum Mechanics

Section I: Matrix Mechanics

Purpose of this part of the unit:

- 1) Constructing a mathematical framework that will be our new language of quantum mechanics
- 2) Moving away from the explicit focus on the wave function
- 3) Establishing the principles of quantum mechanics



Werner Heisenberg



Pascual Jordan



Max Born

Our first aim is to understand the following sentence:

“In quantum mechanics, the state of the particle is represented by a vector $|\psi(t)\rangle$ in a Hilbert space.”

In order to do that, we must familiarize ourselves with the concept of **linear vector spaces**. Let's start from linearity:

For two functions f and g and a scalar α , a linear operator \hat{O} satisfies:

- $\hat{O}(f + g) = \hat{O}f + \hat{O}g$,
- $\hat{O}(\alpha f) = \alpha \hat{O}f$.

Exercise: Many familiar things are linear operators; persuade yourself that all the following examples are:

- Matrix \mathbb{M} acting on vectors \vec{u} and \vec{v}
- Function of position $O(x)$ acting on functions $f(x)$ and $g(x)$
- Derivative $\frac{d}{dx}$ acting on functions $f(x)$ and $g(x)$

We want to build a framework capable of capturing all of these examples at once and will do so based on vectors and matrices. You are already familiar with the notion of vectors. However, here we introduce the so called Dirac notation, in which a vector is no longer denoted with an arrow: \vec{v} . Instead, we will write it as,

$$|v\rangle.$$

For reasons that will become clear soon, a vector written down using this notation is often also known as a “ket”.

We will call a linear vector space a set of objects on which our linear operators can act (we will call the objects vectors but they could also be for example functions like in the earlier exercise). An important point is that the space is closed (the result of any allowed operation is another vector and element of the space).

The set of vectors $\{|i\rangle\}$ is said to be linearly independent if

$$\sum_{i=1}^n a_i |i\rangle = |0\rangle$$

is only true for all $a_i = 0$. If the set of vectors is not linearly independent, we say they are linearly dependent.

A vector space has dimension n if it can accommodate a maximum of n linearly independent vectors.

Exercise: Consider three 2×2 matrices:

$$|1\rangle = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, |2\rangle = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, |3\rangle = \begin{bmatrix} -2 & -1 \\ 0 & -2 \end{bmatrix}.$$

Are they linearly independent?

Exercise: What is the dimension of the vector space of all 2×2 matrices?

Any vector $|v\rangle$ in an n -dimensional space can be written as a linear combination of n linearly independent vectors $|1\rangle, \dots, |n\rangle$.

(If there were a vector $|v\rangle$ for which this were not possible, it would join the given set of vectors and form a set of $n + 1$ linearly independent vectors, which is not possible in an n -dimensional space in which only n linearly independent vectors can be chosen at once).

Because a set of n linearly independent vectors in an n -dimensional space can be used to describe any vector in this space, we call such a set a basis. Notice that, using the definitions so far, we have for any $|v\rangle$,

$$|v\rangle = \sum_{i=1}^n v_i |i\rangle,$$

where the vectors $|i\rangle$ form a basis. The coefficients of expansion, v_i , of a vector in terms of a linearly independent basis $\{|i\rangle\}$ are called the components of the vector in that basis. Such an expansion is unique.

Note that, given a basis the components are unique but if we change the basis, the components will change.

(*) Exercise: Prove that expansion of a vector in a given basis is unique.

Since the vector $|v\rangle$ is uniquely specified by its components in a given basis, we can, in this basis, write it as a column vector

$$|v\rangle \rightarrow \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}.$$

We used the arrow instead of equality to emphasize that the vector $|v\rangle$ (which does not depend on the choice of the basis) is represented by this column vector only in the given basis. Normally, we will use the equality sign, but you should remember that a vector will be represented by different columns of coefficients in different bases.

For two vectors, $|v\rangle$ and $|u\rangle$, we define the inner (or scalar) product and denote it by $\langle v|u\rangle$. It is a number, generally complex, dependent on the two vectors and obeys the following axioms:

- $\langle v|u\rangle = \langle u|v\rangle^*$ (skew-symmetry),
- $\langle v|v\rangle \geq 0$, 0 iff $|v\rangle = 0$ (positive semidefiniteness),
- $\langle v|(\alpha|u\rangle + \beta|z\rangle) \equiv \langle v|\alpha u + \beta z\rangle = \alpha\langle v|u\rangle + \beta\langle v|z\rangle$ (linearity in ket).

Exercise: Show that the following are consequences of the statements on the top of the page:

- $\langle v|v\rangle$ is real,
- $\langle \alpha u + \beta z|v\rangle = \alpha^*\langle u|v\rangle + \beta^*\langle z|v\rangle$ (antilinearity)
- Two vectors are orthogonal (or perpendicular) if their inner product vanishes.
- We will refer to $\sqrt{\langle v|v\rangle} = |v|$ as the norm or length of the vector. A normalized vector has unit norm.
- A set of basis vectors all of unit norm, which are pairwise orthogonal is called an orthonormal basis.

We will refer to $\langle \dots |$ as a “bra”. This allows us to refer to an inner product $\langle v|u\rangle$ as a “bra-ket”.



Paul Dirac (1902 - 1984)

Given a linearly independent basis we can form linear combinations of the basis vectors to obtain an orthonormal basis. This is always possible and the procedure is known as the Gram-Schmidt orthonormalization process:

We start with a non-orthonormal basis of n vectors.

1. Rescale the first vector to normalize it.
2. Subtract from the second vector its projection on the first (the result is perpendicular to the first vector).
3. Normalize the result of the above point.
4. Subtract from the third vector its projections on the first and second vectors. Normalize the remainder.
5. Continue until the full orthonormal basis is produced.

In an orthonormal basis,

$$\langle i | j \rangle = \delta_{ij} \equiv \begin{cases} 1 & \text{for } i = j \\ 0 & \text{for } i \neq j \end{cases}.$$

Exercise: Use the Gram-Schmidt procedure to transform

$$|v_1\rangle = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, |v_2\rangle = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, |v_3\rangle = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix},$$

into an orthonormal basis.

For convenience, we will usually want to expand vectors in an orthonormal basis (and in the rest of this section we assume that such a basis has already been found).

Let us take

$$|v\rangle = \sum_{i=1}^n v_i |i\rangle,$$
$$|w\rangle = \sum_{j=1}^n w_j |j\rangle.$$

Then,

$$\langle v|w\rangle = \sum_i \sum_j v_i^* w_j \langle i|j\rangle$$

Remembering that

$$|v\rangle \rightarrow \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix},$$

we propose that,

$$\langle v| = [v_1^* \quad v_2^* \quad \dots \quad v_n^*],$$

and

$$\langle v|w\rangle = [v_1^* \quad v_2^* \quad \dots \quad v_n^*] \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} = \sum_i \sum_j v_i^* w_j \langle i|j\rangle,$$

as intended. In an orthonormal basis,

$$\langle v|w\rangle = \sum_i \sum_j v_i^* w_j \langle i|j\rangle = \sum_i \sum_j v_i^* w_j \delta_{ij} = \sum_i v_i^* w_i.$$

Note that in an orthonormal basis,

$$|v\rangle = \sum_{i=1}^n v_i |i\rangle,$$

so that,

$$\langle j|v\rangle = \langle j|\sum_{i=1}^n v_i |i\rangle = \sum_{i=1}^n v_i \langle j|i\rangle = v_j,$$

i.e., to find the j -th component of a vector we take a dot product with the j -th unit vector. Using this result,

$$|v\rangle = \sum_{i=1}^n |i\rangle \langle i|v\rangle.$$

Vector spaces associated to quantum mechanics are called Hilbert spaces, in honour of the mathematician David Hilbert. These are formally finite or infinite dimensional vector spaces with inner product and some additional technical properties to do with convergence and continuity. While we will discuss generalization of finite dimensional vector spaces to infinite dimensions, we will not overly concern ourselves with the additional requirements (and pretty much take them for granted).



David Hilbert (1862 - 1943)

We have developed the idea of a linear vector space (which might contain all sorts of objects other than the traditional vectors, including matrices and functions) and will now focus on deciphering the following statement:

“The independent variables x and p are represented by Hermitian operators \hat{X} and \hat{P} with the following matrix elements in the eigenbasis of \hat{X} :

$$\langle x | \hat{X} | x' \rangle = x \delta(x - x'),$$

$$\langle x | \hat{P} | x' \rangle = -i\hbar \frac{d}{dx} \delta(x - x')."$$

In order to do so, we need to learn more about **linear operators** and later generalize the notion of a vector space to infinite dimensions.

Linear operators are instructions for transforming any given vector $|v\rangle$ into another, $|v'\rangle$,

$$\hat{O}|v\rangle = |v'\rangle.$$

Operators can also act on bras:

$$\langle v|\hat{O} = \langle v''|.$$

The important feature of linear operators is that once their action on the basis vectors is determined, their action on any other vector is determined. If

$$\hat{\Omega}|i\rangle = |i'\rangle,$$

then

$$\hat{\Omega}|v\rangle = \sum_i v_i \hat{\Omega}|i\rangle = \sum_i v_i |i'\rangle.$$

The projections

$$\langle j|i'\rangle = \langle j|\hat{\Omega}|i\rangle \equiv \Omega_{ji}$$

are the matrix elements of the operator $\hat{\Omega}$ in the basis $\{|i\rangle\}$. If

$$\hat{\Omega}|v\rangle = |v'\rangle,$$

then the components of the transformed $|v'\rangle$ are expressable in terms of Ω_{ji} and components of $|v\rangle$:

$$\begin{aligned} v'_i &= \langle i|v'\rangle = \langle i|\hat{\Omega}|v\rangle = \langle i|\hat{\Omega}\left(\sum_j v_j |j\rangle\right) = \sum_j v_j \langle i|\hat{\Omega}|j\rangle \\ &= \sum_j v_j \Omega_{ij}. \end{aligned}$$

This equation can be cast in a matrix form

$$\begin{bmatrix} v'_1 \\ v'_2 \\ \vdots \\ v'_n \end{bmatrix} = \begin{bmatrix} \langle 1|\hat{\Omega}|1\rangle & \langle 1|\hat{\Omega}|2\rangle & \dots & \langle 1|\hat{\Omega}|n\rangle \\ \langle 2|\hat{\Omega}|1\rangle & \ddots & \dots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \langle n|\hat{\Omega}|1\rangle & \langle n|\hat{\Omega}|2\rangle & \dots & \langle n|\hat{\Omega}|n\rangle \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}.$$

Let's consider some examples:

1) Identity operator \hat{I} :

$$I_{ij} = \langle i|\hat{I}|j\rangle = \langle i|j\rangle = \delta_{ij}$$

2) Projection operators:

$$|v\rangle = \sum_{i=1}^n |i\rangle \langle i|v\rangle \Rightarrow \hat{I} = \sum_{i=1}^n |i\rangle \langle i|,$$

$$\mathbb{P}_i = |i\rangle \langle i| \quad \text{so that} \quad \hat{I} = \sum_{i=1}^n |i\rangle \langle i| = \sum_{i=1}^n \mathbb{P}_i.$$

Consider

$$\mathbb{P}_i |v\rangle = |i\rangle \langle i|v\rangle = |i\rangle v_i.$$

\mathbb{P}_i projects out the component of any ket $|v\rangle$ along the direction $|i\rangle$ and hence the name projection operator.

Note that

$$\langle v|\mathbb{P}_i = \langle v|i\rangle \langle i| = v_i^* \langle i|,$$

$$\mathbb{P}_i \mathbb{P}_j = |i\rangle \langle i|j\rangle \langle j| = \delta_{ij} \mathbb{P}_j,$$

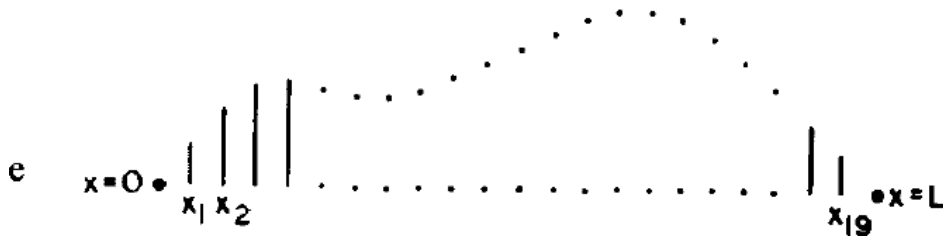
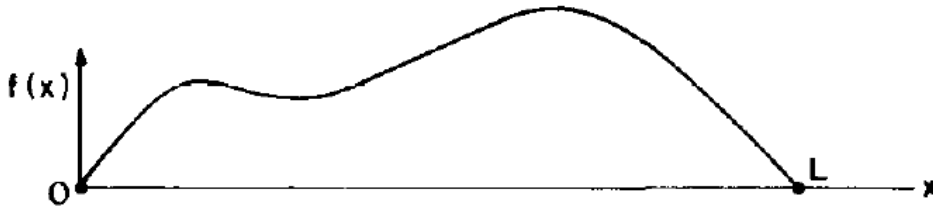
in particular

$$\mathbb{P}_i^2 = \mathbb{P}_i \mathbb{P}_i = \mathbb{P}_i.$$

We have

$$(\mathbb{P}_i)_{kl} = \langle k|i\rangle \langle i|l\rangle = \delta_{ki} \delta_{il} = \delta_{kl} \delta_{li}.$$

3) Derivative operator $\hat{D} = \frac{d}{dx}$:



$$|f\rangle \leftrightarrow \begin{bmatrix} f(x_0) \\ \vdots \\ f(x_i) \\ \vdots \\ f(x_N) \end{bmatrix} \leftarrow i\text{-th place}$$

Recall:

$$\frac{df}{dx} = \lim_{\Delta x \rightarrow 0} \frac{f(x + \Delta x) - f(x)}{\Delta x}.$$

However, we should always look for the most symmetric expression possible:

$$\frac{df}{dx} = \lim_{\Delta x \rightarrow 0} \frac{f(x + \Delta x) - f(x - \Delta x)}{2\Delta x}.$$

This gives

$$\frac{df}{dx} \rightarrow \hat{D}f$$

$$= \begin{bmatrix} \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \\ \dots & -\frac{1}{2\Delta x} & 0 & \frac{1}{2\Delta x} & 0 & 0 & \dots \\ \dots & 0 & -\frac{1}{2\Delta x} & 0 & \frac{1}{2\Delta x} & 0 & \dots \\ \dots & 0 & 0 & -\frac{1}{2\Delta x} & 0 & \frac{1}{2\Delta x} & \dots \\ \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} \begin{bmatrix} \vdots \\ f(x_{n-1}) \\ f(x_n) \\ f(x_{n+1}) \\ \vdots \end{bmatrix}.$$

Clearly, this matrix stops being sensible as $\Delta x \rightarrow 0$, as the off-diagonal elements diverge. However, this simple picture of turning any linear operator – even a derivative – into a matrix is central to our quantum mechanics language.

For a product of operators,

$$\begin{aligned}
 (\Omega\Lambda)_{ij} &= \langle i | \hat{\Omega} \hat{\Lambda} | j \rangle = \langle i | \hat{\Omega} I \hat{\Lambda} | j \rangle = \sum_k \langle i | \hat{\Omega} | k \rangle \langle k | \hat{\Lambda} | j \rangle \\
 &= \sum_k \Omega_{ik} \Lambda_{kj},
 \end{aligned}$$

what means that matrix representing a product of operators is the product of the matrices representing the factors.

For an operator Ω we define its adjoint $\hat{\Omega}^\dagger$ so that:

$$\hat{\Omega} | v \rangle = | \hat{\Omega} v \rangle \implies \langle \hat{\Omega} v | = \langle v | \hat{\Omega}^\dagger.$$

We have

$$(\Omega^\dagger)_{ij} = \langle i | \hat{\Omega}^\dagger | j \rangle = \langle \hat{\Omega} i | j \rangle = \langle j | \hat{\Omega} i \rangle^* = \langle j | \hat{\Omega} | i \rangle^* = \Omega_{ji}^*.$$

In a given basis, the adjoint operation is the same as taking the transpose conjugate. Note that

$$(\hat{\Omega} \hat{\Lambda})^\dagger = \hat{\Lambda}^\dagger \hat{\Omega}^\dagger.$$

Exercise: Prove the relation above.

For two operators, their order is important. We have

$$\hat{\Lambda}\hat{\Omega}|v\rangle = \hat{\Lambda}(\hat{\Omega}|v\rangle) = \hat{\Lambda}|\hat{\Omega}v\rangle,$$

where $|\hat{\Omega}v\rangle$ is the ket obtained by acting with $\hat{\Omega}$ on $|v\rangle$.

In general,

$$\hat{\Omega}\hat{\Lambda} - \hat{\Lambda}\hat{\Omega} \equiv [\hat{\Omega}, \hat{\Lambda}] \neq 0.$$

We call $[\hat{\Omega}, \hat{\Lambda}]$ the commutator of $\hat{\Omega}$ and $\hat{\Lambda}$. In turn, we call

$$\{\hat{\Omega}, \hat{\Lambda}\} \equiv \hat{\Omega}\hat{\Lambda} + \hat{\Lambda}\hat{\Omega}$$

the anticommutator of $\hat{\Omega}$ and $\hat{\Lambda}$.

Exercise: Show that:

$$\begin{aligned} [\hat{\Omega}, \hat{\Lambda}\hat{\Theta}] &= \hat{\Lambda}[\hat{\Omega}, \hat{\Theta}] + [\hat{\Omega}, \hat{\Lambda}]\hat{\Theta}, \\ [\hat{\Lambda}\hat{\Omega}, \hat{\Theta}] &= \hat{\Lambda}[\hat{\Omega}, \hat{\Theta}] + [\hat{\Lambda}, \hat{\Theta}]\hat{\Omega}. \end{aligned}$$

The inverse of an operator, $\hat{\Omega}^{-1}$, satisfies

$$\hat{\Omega}^{-1}\hat{\Omega} = \hat{\Omega}\hat{\Omega}^{-1} = \hat{I}.$$

Not every operator has an inverse. Also, one needs to inverse the order of operators when inverting a product of operators,

$$(\hat{\Omega}\hat{\Lambda})^{-1} = \hat{\Lambda}^{-1}\hat{\Omega}^{-1}$$

for only then we have

$$(\hat{\Omega}\hat{\Lambda})^{-1}(\hat{\Omega}\hat{\Lambda}) = \hat{\Lambda}^{-1}\hat{\Omega}^{-1}\hat{\Omega}\hat{\Lambda} = \hat{\Lambda}^{-1}\hat{\Lambda} = \hat{I}.$$

Eigenvalue problem - reminder

$$\hat{\Omega}|v\rangle = \omega|v\rangle$$

The equation above is an eigenvalue equation, $|v\rangle$ is an eigenket (eigenvector) of $\hat{\Omega}$ with eigenvalue ω .

We can rewrite this equation as

$$(\hat{\Omega} - \hat{I}\omega)|v\rangle = |0\rangle \Rightarrow |v\rangle = (\hat{\Omega} - \hat{I}\omega)^{-1}|0\rangle.$$

The right-hand side does not give zero only if the inverse operator does not exist. This means that

$$\det(\hat{\Omega} - \hat{I}\omega) = 0,$$

which determines the eigenvalues ω through the characteristic equation. Importantly, while the matrix form of $\hat{\Omega}$ depends on the choice of the basis, the eigenvalues are basis independent.

Exercise: Determine the eigenvalues and eigenvectors (in Cartesian co-ordinates) of the operator $\hat{R}\left(\frac{\pi}{2}\hat{\mathbf{i}}\right)$ (counterclockwise rotation by 90 degrees around the x -axis).

Back to linear operators

An operator $\hat{\Omega}$ is Hermitian if $\hat{\Omega}^\dagger = \hat{\Omega}$.

An operator $\hat{\Omega}$ is anti-Hermitian if $\hat{\Omega}^\dagger = -\hat{\Omega}$.

An operator \hat{U} is unitary if $\hat{U}\hat{U}^\dagger = \hat{I}$. Consequently, $\hat{U}^\dagger\hat{U} = \hat{I}$.

- Hermitian operators have real eigenvalues (and so can correspond to quantities that can be measured).
- Their eigenvectors are orthogonal.
- In the basis of its orthonormal eigenvectors, a Hermitian operator is diagonal and has its eigenvalues as its diagonal entries.
- The eigenvalues of a unitary operator are complex numbers of unit modulus.
- Their eigenvectors are orthogonal.
- Unitary operators preserve the inner product between the vectors they act on.

Exercise: Prove the properties of Hermitian operators listed above.

Exercise: Prove the properties of unitary operators listed above.

If one treats the columns of an $n \times n$ unitary matrix as components of n vectors, these vectors are orthonormal. In the same way, the rows may be interpreted as components of n orthonormal vectors.

Active and passive transformations and diagonalization

Suppose we subject all the vectors $|v\rangle$ in a space to a unitary transformation $|v\rangle \rightarrow \hat{U}|v\rangle$. Under this transformation, the matrix elements of any operator $\hat{\Omega}$ are modified as follows

$$\langle v' | \hat{\Omega} | v \rangle \rightarrow \langle \hat{U}v' | \hat{\Omega} | \hat{U}v \rangle = \langle v' | \hat{U}^\dagger \hat{\Omega} \hat{U} | v \rangle.$$

It is clear that the same change would be effected if we left the vectors alone and subjected all operators to the change $\hat{\Omega} \rightarrow \hat{U}^\dagger \hat{\Omega} \hat{U}$.

The first case is called an active transformation and the second a passive transformation (the terms are in reference to the vectors which are changed in the first transformation but not the second).

Importantly, the eigenvalues, trace and determinant of a matrix do not change under a unitary transformation.

Reminder: The trace of a matrix is defined to be the sum of its diagonal elements:

$$\text{Tr } \hat{\Omega} = \sum_i \Omega_{ii}.$$

Exercise: Show that the eigenvalues, trace and determinant of a matrix do not change under a unitary transformation. You may use the following properties of the trace/determinant:

$$\begin{aligned} \det(\hat{A}\hat{B}) &= \det(\hat{A}) \det(\hat{B}), \\ \text{Tr}(\hat{A}\hat{B}) &= \text{Tr}(\hat{B}\hat{A}). \end{aligned}$$

Consider a Hermitian operator $\hat{\Omega}$ represented as a matrix in some orthonormal basis $|1\rangle, \dots, |n\rangle$. If we trade this basis for the eigenbasis $|\omega_1\rangle, \dots, |\omega_n\rangle$, the matrix representing $\hat{\Omega}$ will become diagonal. The operator \hat{U} inducing the change of basis,

$$|\omega_i\rangle = \hat{U}|i\rangle,$$

is clearly unitary for it “rotates” one orthonormal basis into another. Hence, every Hermitian matrix can be diagonalized by a unitary change of basis. Equivalently, in terms of passive transformations, there exists a unitary matrix \hat{U} (built out of the eigenvectors of $\hat{\Omega}$) such that $\hat{U}^\dagger \hat{\Omega} \hat{U}$ is diagonal. Thus, the problem of finding a basis that diagonalizes $\hat{\Omega}$ is equivalent to solving its eigenvalue problem.

If $\hat{\Omega}$ and $\hat{\Lambda}$ are two commuting Hermitian operators, there exists a basis of common eigenvectors that diagonalizes them both (that is, the two operators share the same eigenvectors).

Exercise: Show that the following Hermitian matrices can be simultaneously diagonalized. Find the eigenvectors common to both and verify that under a unitary transformation to this basis both matrices are diagonalized.

$$\hat{\Omega} = \begin{bmatrix} 5 & 1 \\ 1 & 5 \end{bmatrix}, \quad \hat{\Lambda} = \begin{bmatrix} -1 & -1 \\ -1 & -1 \end{bmatrix}.$$

Let us consider a function of (Hermitian) operators that can be written as a power series,

$$f(\hat{\Omega}) = \sum_{n=0}^{\infty} a_n \hat{\Omega}^n,$$

for example,

$$e^{\hat{\Omega}} = \sum_{n=0}^{\infty} \frac{\hat{\Omega}^n}{n!}.$$

By going to the eigenbasis of $\hat{\Omega}$, we can readily perform the sum:

$$\hat{\Omega} = \begin{bmatrix} \omega_1 & & & \\ & \omega_2 & & \\ & & \ddots & \\ & & & \omega_n \end{bmatrix},$$

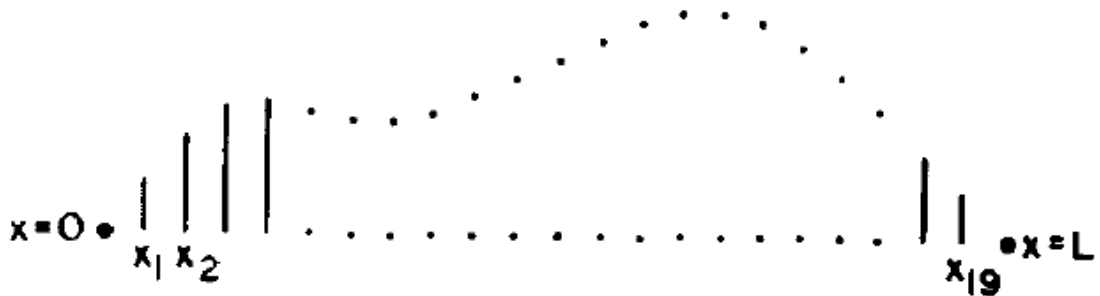
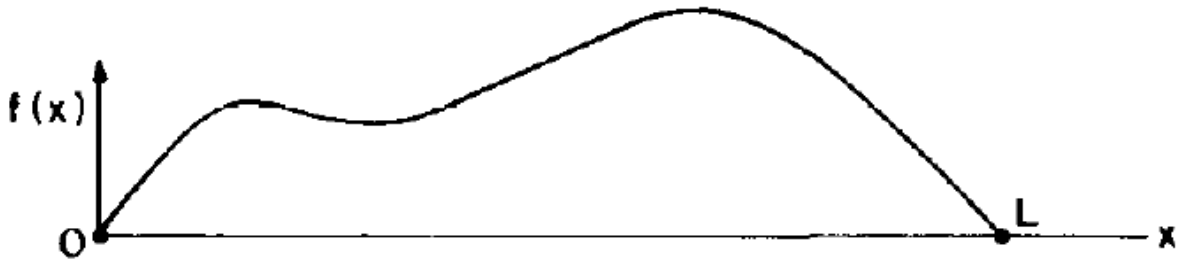
$$e^{\hat{\Omega}} = \begin{bmatrix} \sum_{m=0}^{\infty} \frac{\omega_1^m}{m!} & & & \\ & \sum_{m=0}^{\infty} \frac{\omega_2^m}{m!} & & \\ & & \ddots & \\ & & & \sum_{m=0}^{\infty} \frac{\omega_n^m}{m!} \end{bmatrix}.$$

Because each sum converges to the familiar limit e^{ω_n} , the operator $e^{\hat{\Omega}}$ is well defined in this basis (and therefore in any other).

Exercise: If the operator \hat{H} is Hermitian, show that $e^{i\hat{H}}$ is unitary.

Generalization to Infinite Dimensions – position space

In all the previous discussions, the dimensionality of the vector space n was unspecified but assumed to be a finite number. The last piece of the mathematical puzzle we need is to generalize these concepts to infinite dimensions. This will also allow us to make a connection to the wave function you are already familiar with.



$$|x_i\rangle \leftrightarrow \begin{bmatrix} 0 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \leftarrow i\text{-th place}$$

$$\langle x_i | x_j \rangle = \delta_{ij}$$

$$\sum_{i=1}^n |x_i\rangle \langle x_i| = I$$

Imagine a space containing n (the number of points along the string) mutually perpendicular axes, one for each point x_i . Along each axis is a unit vector $|x_i\rangle$. The function $f_n(x)$ is represented by a vector whose projection along the i -th direction is $f_n(x_i)$:

$$|f_n\rangle = \sum_{i=1}^n f_n(x_i) |x_i\rangle.$$

We define inner product

$$\langle f_n | g_n \rangle = \sum_{i=1}^n f_n^*(x_i) g_n(x_i) \Delta, \quad \Delta = \frac{L}{n+1},$$

(with the last factor necessary to obtain a smooth limit as n tends to infinity). In the limit, we get an integral

$$\begin{aligned} \langle f | g \rangle &= \int_a^b f^*(x) g(x) dx, \\ \langle f | f \rangle &= \int_a^b |f(x)|^2 dx, \end{aligned}$$

where we already introduced the limits a and b which depend on the particular physical problem (e.g., length of the box, volume of the material, all of 3D space; the last two require a simple generalization of the integral to more than one dimension). In this continuous limit, expansion of a state into the position basis takes the form

$$|g\rangle = \int_a^b dx' g(x') |x'\rangle.$$

What are the basis vectors in this space and how are they normalized? Each point x gets a basis vector and we can require

$$\langle x|x'\rangle = 0, \quad x \neq x'.$$

Should we require $\langle x|x\rangle = 1$? Let us look at the completeness relation

$$\sum_{i=1}^n |x_i\rangle \langle x_i| = I \rightarrow \int_a^b |x'\rangle \langle x'| dx' = I.$$

Multiplying by a ket $|f\rangle$ from the right and bra $\langle x|$ on the left,

$$\int_a^b \langle x|x'\rangle \langle x'|f\rangle dx' = \langle x|I|f\rangle = \langle x|f\rangle.$$

Note that $\langle x|f\rangle = f(x)$, $\langle x'|f\rangle = f(x')$. Also, we know that $\langle x|x'\rangle$ vanishes if $x \neq x'$ so that we can restrict the integral to an infinitesimal vicinity of x ,

$$\int_{x-\varepsilon}^{x+\varepsilon} \langle x|x'\rangle f(x') dx' = f(x).$$

In this infinitesimal region, for any smooth f , $f(x')$ can be approximated by $f(x)$ so that

$$f(x) \int_{x-\varepsilon}^{x+\varepsilon} \langle x|x'\rangle dx' = f(x) \rightarrow \int_{x-\varepsilon}^{x+\varepsilon} \langle x|x'\rangle dx' = 1.$$

Note that $\langle x|x' \rangle$ cannot be finite at $x = x'$ for then the integral over an infinitesimal region would be infinitesimally small. The bra-ket should be infinite in such a way so that the integral is unity! This is handled by the Dirac delta “function”

$$\delta(x - x') = 0, \quad x \neq x',$$

$$\int_a^b \delta(x - x') dx' = 1, \quad a < x < b.$$

We have the normalization condition

$$\langle x|x' \rangle = \delta(x - x').$$

Useful relations for the Dirac delta function

$$\int \delta(x - x') f(x') dx' = f(x)$$

$$\delta(x - x') = \lim_{\varepsilon \rightarrow 0} \frac{1}{2\sqrt{\pi\varepsilon}} e^{-(x-x')^2/(4\varepsilon)}$$

$$\delta(-x) = \delta(x)$$

$$\delta(\alpha x) = \frac{1}{|\alpha|} \delta(x)$$

$$\delta(x - x') = \frac{1}{2\pi} \int_{-\infty}^{\infty} dk e^{ik(x'-x)}$$

Let us now consider operators. Position is real – there should exist an associated Hermitian position operator, \hat{X} . What are its matrix elements in the position basis? We have

$$\begin{aligned}\hat{X}|x\rangle &= x|x\rangle, \\ \langle x'|\hat{X}|x\rangle &= x\langle x'|x\rangle = x\delta(x' - x).\end{aligned}$$

What about other operators? Here, we shall assume that the operators we are interested in are local, that is, they are diagonal in the position space,

$$\langle x|\hat{O}|x'\rangle = O_{xx'} = O_x\delta(x - x'),$$

where O_x is the matrix element or expression of the operator \hat{O} in position space (expression for \hat{O} in terms of functions of position and its derivatives).

Consider now some general orthonormal basis, $\{|f_i\rangle\}$. Each basis vector can be expanded in the position basis,

$$|f_i\rangle = \int_a^b dx' f_i|x'\rangle.$$

This means that the matrix element, O_{ij} , in the $\{|f_i\rangle\}$ basis, is

$$\begin{aligned}O_{ij} &= \langle f_i|\hat{O}|f_j\rangle = \int dx \int dx' \langle f_i|x\rangle \langle x|\hat{O}|x'\rangle \langle x'|f_j\rangle \\ &= \int dx \int dx' f_i^*(x) \langle x|\hat{O}|x'\rangle f_j(x') \\ &= \int dx \int dx' f_i^*(x) O_x \delta(x - x') f_j(x') \\ &= \int dx f_i^*(x) O_x f_j(x).\end{aligned}$$

Bases other than the position one can be (more) useful at times. For example, let's look at the derivative operator,

$$\hat{D} \rightarrow D_x = \frac{d}{dx}.$$

Unfortunately, such an operator is anti-Hermitian (**exercise**). To make it Hermitian, we choose,

$$\hat{K} = -i\hat{D}$$

What are the eigenvalues of \hat{K} ? Using its eigenstates,

$$\hat{K}|k\rangle = k|k\rangle \rightarrow \langle x|\hat{K}|k\rangle = k\langle x|k\rangle.$$

We have $\langle x|k\rangle = k(x)$, function different for every k . At the same time,

$$\begin{aligned} \langle x|\hat{K}|k\rangle &= \int \langle x|\hat{K}|x'\rangle \langle x'|k\rangle dx' = \\ &= -i \int \langle x|\hat{D}|x'\rangle k(x') dx' = -i \int dx' \frac{d}{dx} \delta(x - x') k(x') = \\ &= -i \frac{d}{dx} k(x). \end{aligned}$$

$$-i \frac{d}{dx} k(x) = k k(x) \rightarrow k(x) = A e^{ikx}$$

This is a plane wave! With this, we adjust our notation,

$$k(x) \equiv \langle x|k\rangle \rightarrow \psi_k(x),$$

to denote a plane wave with wave vector k .

The constant A is fixed by normalization:

$$\begin{aligned} A &= \frac{1}{(2\pi)^{1/2}} \rightarrow \langle k|k'\rangle = \int \langle k|x'\rangle \langle x'|k'\rangle dx' \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} dx' e^{-i(k-k')x'} = \delta(k - k'). \end{aligned}$$

Since \hat{K} is Hermitian, its eigenvectors form an orthonormal basis set – how does a ket $|f\rangle$ look like in this basis [in contrast to the \hat{X} basis in which $\langle x|f\rangle = f(x)$]? We find the expansion coefficients by computing:

$$f(k) = \langle k|f\rangle = \int_{-\infty}^{\infty} \langle k|x\rangle \langle x|f\rangle dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dx e^{-ikx} f(x).$$

Note that in the \hat{X} basis,

$$f(x) = \langle x|f\rangle = \int_{-\infty}^{\infty} \langle x|k\rangle \langle k|f\rangle dk = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dk e^{ikx} f(k).$$

The above relations show that the Fourier transform which you will learn about soon provides the passage from one complete basis, $\{|x\rangle\}$ (position), to another, $\{|k\rangle\}$ (wave vector). Given the scaling $\hbar k = p$, the latter is commonly referred to as the momentum space.

Exercise: Using the relation between the position and momentum spaces, show that knowing exactly the position of the particle would imply complete lack of information about its momentum (and vice versa). When might it be better to use one of these bases rather than the other?

We have established that in the position basis,

$$\langle x|\hat{X}|x'\rangle = x'\langle x|x'\rangle = x'\delta(x - x'),$$

$$\langle x|\hat{K}|x'\rangle = -i\frac{d}{dx}\delta(x - x')$$

(remember that the appearance of the Delta functions means the above expressions are only meaningful under an integral over position).

The matrix elements of K are trivial in the $|k\rangle$ basis,

$$\langle k|\hat{K}|k'\rangle = k'\langle k|k'\rangle = k'\delta(k - k').$$

What is $\langle k|\hat{X}|k'\rangle$?

$$\begin{aligned}\langle k|\hat{X}|k'\rangle &= \int \langle k|\hat{X}|x\rangle\langle x|k'\rangle dx = \int x\langle k|x\rangle\langle x|k'\rangle dx \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} dx e^{-ikx} x e^{ik'x} = i \frac{d}{dk} \left(\frac{1}{2\pi} \int_{-\infty}^{\infty} dx e^{i(k'-k)x} \right) \\ &= i \frac{d}{dk} \delta(k - k').\end{aligned}$$

In summary, in the \hat{X} basis, \hat{X} acts as x and \hat{K} as $-id/dx$ [on the functions $f(x)$], while in the \hat{K} basis, \hat{K} acts like k and \hat{X} like id/dk [on the functions $f(k)$]. Operators with such an interrelationship are said to be conjugate to each other.

Conjugate operators \hat{X} and \hat{K} do not commute:

$$\hat{X}|f\rangle \rightarrow xf(x), \quad \hat{K}|f\rangle \rightarrow -i\frac{df(x)}{dx},$$

$$\hat{X}\hat{K}|f\rangle \rightarrow -ix\frac{df(x)}{dx}, \quad \hat{K}\hat{X}|f\rangle \rightarrow -i\frac{d}{dx}[xf(x)],$$

$$[\hat{X}, \hat{K}]|f\rangle = -ix\frac{df(x)}{dx} + ix\frac{df(x)}{dx} + if(x) = if(x) \rightarrow i\hat{I}|f\rangle.$$

Since $|f\rangle$ is an arbitrary ket, we have

$$[\hat{X}, \hat{K}] = i\hat{I}$$

(irrespective of the basis in which we would perform the calculation).

We define the momentum operator $\hat{P} = \hbar\hat{K}$ (in the position basis in one dimension, $\hat{P} = -i\hbar\frac{d}{dx}$). This means that $[\hat{X}, \hat{P}] = i\hbar\hat{I} = i\hbar$.

Exercise: Show that the result for $[\hat{X}, \hat{K}]$ is the same if we compute it in the \hat{K} basis.

Exercise: Show that in more than one dimension, only position and momentum operators along the same axis do not commute.

We are now ready to formulate our foundations of quantum mechanics...

Principles of quantum mechanics

- I. The state of the particle is represented by a vector $|\psi(t)\rangle$ in a Hilbert space.
- II. The independent variables x and p are represented by Hermitian operators \hat{X} and \hat{P} with the following matrix elements in the eigenbasis of \hat{X} :

$$\langle x|\hat{X}|x'\rangle = x\delta(x - x'),$$

$$\langle x|\hat{P}|x'\rangle = -i\hbar \frac{d}{dx} \delta(x - x').$$

The operators corresponding to dependent variables $\omega(x, p)$ are given Hermitian operators $\hat{\Omega}(\hat{X}, \hat{P}) = \omega(x \rightarrow \hat{X}, p \rightarrow \hat{P})$ ($\hat{\Omega}$ is the same function of \hat{X} and \hat{P} as ω is of x and p – so called first quantization).

- III. If the particle is in a state $|\psi\rangle$, measurement of the variable corresponding to $\hat{\Omega}$ will yield one of the eigenvalues ω with probability $P(\omega) = |\langle \omega|\psi\rangle|^2$ (for a normalized ket $|\psi\rangle$). The state of the system will change from $|\psi\rangle$ to $|\omega\rangle$ as a result of the measurement.
- IV. The state vector $|\psi(t)\rangle$ obeys the Schrödinger equation

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

Where \hat{H} is the quantum Hamiltonian operator (total energy operator) obtained from its classical counterpart by taking $x \rightarrow \hat{X}, p \rightarrow \hat{P}$.

Complications

1. The recipe is ambiguous (classically, x and p commute).

Solution: Write a symmetric sum (for example, $\hat{\Omega} = (\hat{X}\hat{P} + \hat{P}\hat{X})/2$).

2. The operator $\hat{\Omega}$ is degenerate.

Solution: Let us say $\omega_1 = \omega_2 = \omega$. What is the probability? We have to include all eigenstates (ask for probability of ω_1 or ω_2),

$$P(\omega) = |\langle\omega, 1|\psi\rangle|^2 + |\langle\omega, 2|\psi\rangle|^2.$$

3. The eigenvalue spectrum of $\hat{\Omega}$ is continuous.

Solution: Generalize the expansion to

$$|\psi\rangle = \int |\omega\rangle \langle\omega|\psi\rangle d\omega.$$

$P(\omega)$ then becomes probability density (probability of measuring value between ω and $\omega + d\omega$).

4. The quantum variable $\hat{\Omega}$ has no classical counterpart (e.g., spin).

Solution: Use your intuition and hope for the best.

Comments and reminders

1. The expectation value is the mean value obtained across many measurements

$$\begin{aligned}\langle \hat{\Omega} \rangle &= \sum_i P(\omega_i) \omega_i = \sum_i \langle \psi | \omega_i \rangle \langle \omega_i | \psi \rangle \omega_i \\ &= \sum_i \langle \psi | \hat{\Omega} | \omega_i \rangle \langle \omega_i | \psi \rangle = \langle \psi | \hat{\Omega} \hat{I} | \psi \rangle = \langle \psi | \hat{\Omega} | \psi \rangle\end{aligned}$$

2. The uncertainty is given by the standard deviation

$$\begin{aligned}\Delta\Omega &= \left\langle (\hat{\Omega} - \langle \hat{\Omega} \rangle)^2 \right\rangle^{1/2} = \left\langle \hat{\Omega}^2 - 2\hat{\Omega}\langle \hat{\Omega} \rangle + \langle \hat{\Omega} \rangle^2 \right\rangle^{1/2} \\ &= \sqrt{\langle \hat{\Omega}^2 \rangle - \langle 2\hat{\Omega}\langle \hat{\Omega} \rangle \rangle + \langle \langle \hat{\Omega} \rangle^2 \rangle} = \sqrt{\langle \hat{\Omega}^2 \rangle - \langle \hat{\Omega} \rangle^2}.\end{aligned}$$

3. If operators $\hat{\Omega}$ and $\hat{\Lambda}$ commute, their corresponding variables are said to be compatible. There exists a complete basis which is an eigenbasis of both $\hat{\Omega}$ and $\hat{\Lambda}$.

Time evolution

Let us investigate time evolution of systems for which the Hamiltonian \hat{H} does not explicitly depend on time. For given $t = t_0 = 0$, we solve the time-independent Schrödinger equation,

$$\hat{H}|\varepsilon\rangle = \varepsilon|\varepsilon\rangle$$

(find the eigenkets of \hat{H}).

We have

$$|\psi(t)\rangle = \sum |\varepsilon\rangle \langle \varepsilon | \psi(t) \rangle \equiv \sum a_\varepsilon(t) |\varepsilon\rangle.$$

From Schrödinger equation,

$$\begin{aligned} 0 &= \left(i\hbar \frac{\partial}{\partial t} - \hat{H} \right) |\psi(t)\rangle = \sum \left(i\hbar \frac{da_\varepsilon}{dt} - \varepsilon a_\varepsilon \right) |\varepsilon\rangle \Rightarrow i\hbar \frac{da_\varepsilon}{dt} \\ &= \varepsilon a_\varepsilon \end{aligned}$$

(because $|\varepsilon\rangle$ are linearly independent). This means that

$$a_\varepsilon(t) = a_\varepsilon(0) e^{-i\varepsilon t/\hbar}.$$

Hence,

$$\langle \varepsilon | \psi(t) \rangle \equiv a_\varepsilon(t) = a_\varepsilon(0) e^{-i\varepsilon t/\hbar} = \langle \varepsilon | \psi(0) \rangle e^{-i\varepsilon t/\hbar}$$

so that

$$|\psi(t)\rangle = \sum |\varepsilon\rangle \langle \varepsilon | \psi(0) \rangle e^{-i\varepsilon t/\hbar} \equiv \hat{U} |\psi(0)\rangle$$

where

$$\hat{U} = \sum |\varepsilon\rangle \langle \varepsilon | e^{-i\varepsilon t/\hbar}.$$

The states $|\varepsilon(t)\rangle = |\varepsilon(0)\rangle e^{-i\varepsilon t/\hbar}$ are called stationary states (the probability distribution is time-independent in such a state).

Notice that the operator \hat{U} can be rewritten as

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

Since \hat{H} is Hermitian, $\hat{U}(t)$ is unitary. We can therefore think of the time evolution of a ket $|\psi\rangle$ as a “rotation” in Hilbert space. One immediate consequence is that the norm $\langle\psi(t)|\psi(t)\rangle$ is invariant,

$$\langle\psi(t)|\psi(t)\rangle = \langle\psi(0)|\hat{U}^\dagger(t)\hat{U}(t)|\psi(0)\rangle = \langle\psi(0)|\psi(0)\rangle.$$

Another consequence is that we could abandon the fixed basis we were using and adopt one that also rotates at the same rate as the state vectors. In such a basis the vectors would appear frozen, but the operators which were constant matrices in the fixed basis, would now appear to be time-dependent. Any physical entity, such as a matrix element, would, however, come out the same as before since $\langle\psi|\hat{\Omega}|\psi\rangle$, which is the dot product of $\langle\psi|$ and $|\hat{\Omega}\psi\rangle$, is invariant under rotations. This view of quantum mechanics is called the Heisenberg picture, while the one we have been using is called the Schrödinger picture.

Exercise: Consider Hermitian operators $\hat{\Gamma}$ and $\hat{\Omega}$ in a Hilbert space \mathcal{H} .

- Show that if their commutator $[\hat{\Omega}, \hat{\Gamma}]$ is written in the form $[\hat{\Omega}, \hat{\Gamma}] = i\hat{\Lambda}$, then $\hat{\Lambda}$ is Hermitian. Show also that their anticommutator $\{\hat{\Omega}, \hat{\Gamma}\}$ is Hermitian.
- Show that the product of two operators can always be written in terms of their commutator and anticommutator.
- Consider the product of squared uncertainties, $(\Delta\Omega)^2(\Delta\Gamma)^2$, where

$$(\Delta\Omega) = \left[\langle (\hat{\Omega} - \langle \hat{\Omega} \rangle)^2 \rangle \right]^{1/2},$$

and the angular brackets denote the expectation value, $\langle \Theta \rangle = \langle \psi | \hat{\Theta} | \psi \rangle$, with a normalized state vector $|\psi\rangle$. Demonstrate that

$$\begin{aligned} & (\Delta\Omega)^2(\Delta\Gamma)^2 \\ &= \langle (\hat{\Omega} - \langle \hat{\Omega} \rangle) \psi | (\hat{\Omega} - \langle \hat{\Omega} \rangle) \psi \rangle \langle (\hat{\Gamma} - \langle \hat{\Gamma} \rangle) \psi | (\hat{\Gamma} - \langle \hat{\Gamma} \rangle) \psi \rangle. \end{aligned}$$

- With the help of the result in c) and the Schwartz inequality,

$$|v||w| \geq |\langle v|w \rangle|,$$

for vectors $|v\rangle$ and $|w\rangle$, show that

$$(\Delta\Omega)^2(\Delta\Gamma)^2 \geq |\langle (\hat{\Omega} - \langle \hat{\Omega} \rangle)(\hat{\Gamma} - \langle \hat{\Gamma} \rangle) \rangle|^2.$$

- Now, using the result of b), show that

$$\begin{aligned} & (\Delta\Omega)^2(\Delta\Gamma)^2 \\ & \geq \frac{1}{4} \left| \langle \psi | \{ \hat{\Omega} - \langle \hat{\Omega} \rangle, \hat{\Gamma} - \langle \hat{\Gamma} \rangle \} + [\hat{\Omega} - \langle \hat{\Omega} \rangle, \hat{\Gamma} - \langle \hat{\Gamma} \rangle] | \psi \rangle \right|^2. \end{aligned}$$

f) Demonstrate that $[\hat{\Omega} - \langle \hat{\Omega} \rangle, \hat{\Gamma} - \langle \hat{\Gamma} \rangle] = [\hat{\Omega}, \hat{\Gamma}]$.

g) Demonstrate that

$$\langle \{\hat{\Omega} - \langle \hat{\Omega} \rangle, \hat{\Gamma} - \langle \hat{\Gamma} \rangle\} \rangle = \langle \{\hat{\Omega}, \hat{\Gamma}\} \rangle - 2\langle \hat{\Gamma} \rangle \langle \hat{\Omega} \rangle.$$

h) Prove the following form of the uncertainty principle for the operators $\hat{\Gamma}$ and $\hat{\Omega}$,

$$(\Delta\Omega)^2(\Delta\Gamma)^2 \geq \frac{1}{4} (\langle \{\hat{\Gamma}, \hat{\Omega}\} \rangle - 2\langle \hat{\Gamma} \rangle \langle \hat{\Omega} \rangle)^2 + \frac{1}{4} \langle \hat{\Lambda} \rangle^2.$$

i) Show that for position and momentum, the result in h) can be used to obtain the usual form of the Heisenberg uncertainty principle.

PH20013/60 Quantum Mechanics

Section II: Angular momentum and Spin

In this section we will discuss quantization of angular momentum. Apart from the importance of this topic to solutions of Schrodinger equation in 3D, our analysis will ultimately lead us to a quantum mechanical creature with no counterpart in classical world – spin. Our discussion will provide a practice ground for the use of the framework developed in Section I.

Angular momentum

In classical mechanics, the angular momentum of a particle at position \vec{r} and with momentum \vec{p} is

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

When moving from classical mechanics to quantum mechanics, we can simply substitute vectors \vec{r} and \vec{p} with operators \hat{r} and \hat{p} to obtain a new operator \hat{L} ,

$$\hat{L} = \hat{r} \times \hat{p},$$

Which we expand in the same way as a cross product of two vectors.

$$\begin{vmatrix} \vec{x} & \vec{y} & \vec{z} \\ \hat{x} & \hat{y} & \hat{z} \\ \hat{p}_x & \hat{p}_y & \hat{p}_z \end{vmatrix} \rightarrow \begin{aligned} \hat{L}_x &= \hat{y}\hat{p}_z - \hat{z}\hat{p}_y \\ \hat{L}_y &= \hat{z}\hat{p}_x - \hat{x}\hat{p}_z \\ \hat{L}_z &= \hat{x}\hat{p}_y - \hat{y}\hat{p}_x \end{aligned}$$

We would like to know all that can be known about angular momentum of a particle. A sensible first question is: how much are we allowed to know (can we know all three components of \hat{L} at once)? We know that to answer this question, it is enough to compute certain commutators.

Example

Find the commutator $[\hat{L}_x, \hat{L}_y]$.

$$\begin{aligned}
 [\hat{L}_x, \hat{L}_y] &= \hat{L}_x \hat{L}_y - \hat{L}_y \hat{L}_x \\
 &= (\hat{y} \hat{p}_z - \hat{z} \hat{p}_y)(\hat{z} \hat{p}_x - \hat{x} \hat{p}_z) - (\hat{z} \hat{p}_x - \hat{x} \hat{p}_z)(\hat{y} \hat{p}_z - \hat{z} \hat{p}_y) \\
 &= \hat{y} \hat{p}_z \hat{z} \hat{p}_x - \hat{y} \hat{p}_z \hat{x} \hat{p}_z - \hat{z} \hat{p}_y \hat{z} \hat{p}_x + \hat{z} \hat{p}_y \hat{x} \hat{p}_z - \hat{z} \hat{p}_x \hat{y} \hat{p}_z + \hat{z} \hat{p}_x \hat{z} \hat{p}_y \\
 &\quad + \hat{x} \hat{p}_z \hat{y} \hat{p}_z - \hat{x} \hat{p}_z \hat{z} \hat{p}_y
 \end{aligned}$$

We have established that only position and momentum operators along the same axis do not commute (that is, $[\hat{x}, \hat{p}_x] = i\hbar$ but $[\hat{y}, \hat{p}_x] = 0$). Hence, the only non-commuting operators are \hat{z} and \hat{p}_z in terms 1,4,5 and 8. We can swap order of all the other operators without any consequences. After cancelling terms 2 and 7 and 3 and 6,

$$\begin{aligned}
 [\hat{L}_x, \hat{L}_y] &= \hat{y} \hat{p}_z \hat{z} \hat{p}_x + \hat{z} \hat{p}_y \hat{x} \hat{p}_z - \hat{z} \hat{p}_x \hat{y} \hat{p}_z - \hat{x} \hat{p}_z \hat{z} \hat{p}_y \\
 &= \hat{y} \hat{p}_x \hat{p}_z \hat{z} + \hat{x} \hat{p}_y \hat{z} \hat{p}_z - \hat{y} \hat{p}_x \hat{z} \hat{p}_z - \hat{x} \hat{p}_y \hat{p}_z \hat{z} \\
 &= (\hat{y} \hat{p}_x - \hat{x} \hat{p}_y) \hat{p}_z \hat{z} - (\hat{y} \hat{p}_x - \hat{x} \hat{p}_y) \hat{z} \hat{p}_z = -\hat{L}_z [\hat{p}_z, \hat{z}] = i\hbar \hat{L}_z
 \end{aligned}$$

After similar calculations, we find

$$\left. \begin{aligned}
 [\hat{L}_x, \hat{L}_y] &= i\hbar \hat{L}_z \\
 [\hat{L}_y, \hat{L}_z] &= i\hbar \hat{L}_x \\
 [\hat{L}_z, \hat{L}_x] &= i\hbar \hat{L}_y
 \end{aligned} \right\} \begin{array}{l} \text{See Problem 1} \\ \text{Problem Set 3} \end{array}$$

Usually, we cannot know all three (or even two) components of angular momentum exactly (we will come back to this in the problem class). What about the magnitude of angular momentum, $\hat{L}^2 = \hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2$?

$$\begin{aligned}
 [\hat{L}^2, \hat{L}_x] &= [\hat{L}_x^2, \hat{L}_x] + [\hat{L}_y^2, \hat{L}_x] + [\hat{L}_z^2, \hat{L}_x] \\
 &= \hat{L}_y [\hat{L}_y, \hat{L}_x] + [\hat{L}_y, \hat{L}_x] \hat{L}_y + \hat{L}_z [\hat{L}_z, \hat{L}_x] + [\hat{L}_z, \hat{L}_x] \hat{L}_z \\
 &= \hat{L}_y (-i\hbar \hat{L}_z) + (-i\hbar \hat{L}_z) \hat{L}_y + \hat{L}_z i\hbar \hat{L}_y + i\hbar \hat{L}_y \hat{L}_z = 0
 \end{aligned}$$

We have

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

so that it is possible to know the magnitude of angular momentum and one of its components exactly.

Mathematically, this means we can look for states which are eigenstates of both \hat{L}^2 and one of the components of \hat{L} , conventionally \hat{L}_z .

The following operators will help us in this analysis

$$\begin{aligned}\hat{L}_+ &= \hat{L}_x + i\hat{L}_y \\ \hat{L}_- &= \hat{L}_x - i\hat{L}_y\end{aligned}$$

Exercise: Show that operators \hat{L}_+ and \hat{L}_- are not Hermitian. What are the physical consequences of this fact?

Let us investigate the properties of those two operators

$$\begin{aligned}\hat{L}_+\hat{L}_- &= (\hat{L}_x + i\hat{L}_y)(\hat{L}_x - i\hat{L}_y) = \hat{L}_x^2 + \hat{L}_y^2 - i\hat{L}_x\hat{L}_y + i\hat{L}_y\hat{L}_x \\ &= \hat{L}^2 - \hat{L}_z^2 - i[\hat{L}_x, \hat{L}_y] = \hat{L}^2 - \hat{L}_z^2 + \hbar\hat{L}_z\end{aligned}$$

Similarly,

$$\hat{L}_-\hat{L}_+ = \hat{L}^2 - \hat{L}_z^2 - \hbar\hat{L}_z$$

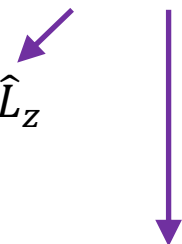
so that

$$[\hat{L}_+, \hat{L}_-] = \hat{L}_+\hat{L}_- - \hat{L}_-\hat{L}_+ = 2\hbar\hat{L}_z$$

Also,

$$\begin{aligned}[\hat{L}_z, \hat{L}_+] &= [\hat{L}_z, \hat{L}_x] + i[\hat{L}_z, \hat{L}_y] = i\hbar\hat{L}_y + i(-i\hbar\hat{L}_x) = \hbar\hat{L}_+, \\ [\hat{L}_z, \hat{L}_-] &= [\hat{L}_z, \hat{L}_x] - i[\hat{L}_z, \hat{L}_y] = i\hbar\hat{L}_y - i(-i\hbar\hat{L}_x) = -\hbar\hat{L}_-.\end{aligned}$$

See Problem 2
Problem Set 3



After the introductory maths, we shall now find the eigenvalues of \hat{L}^2 and \hat{L}_z . We know these two operators commute so have the same set of eigenstates,

$$\hat{L}^2|\phi_n\rangle = \alpha_n|\phi_n\rangle, \quad \hat{L}_z|\phi_n\rangle = \beta_n|\phi_n\rangle.$$

Because \hat{L}^2 corresponds to squared length of the angular momentum “vector” and \hat{L}_z to only one of its three components, we also expect that

$$\alpha_n \geq \beta_n^2$$

Let's start with

$$\hat{L}_+\hat{L}_z|\phi_n\rangle = \hat{L}_+\beta_n|\phi_n\rangle = \beta_n\hat{L}_+|\phi_n\rangle,$$

whereas switching the order

$$\begin{aligned} \hat{L}_z\hat{L}_+|\phi_n\rangle &= ([\hat{L}_z, \hat{L}_+] + \hat{L}_+\hat{L}_z)|\phi_n\rangle = (\hbar\hat{L}_+ + \hat{L}_+\hat{L}_z)|\phi_n\rangle \\ &= (\hbar\hat{L}_+ + \beta_n\hat{L}_+)|\phi_n\rangle = (\hbar + \beta_n)\hat{L}_+|\phi_n\rangle. \end{aligned}$$

Hence,

$$\hat{L}_z(\hat{L}_+|\phi_n\rangle) = (\beta_n + \hbar)(\hat{L}_+|\phi_n\rangle)$$

what means that if $|\phi_n\rangle$ is an eigenstate of \hat{L}_z with eigenvalue β_n , then $\hat{L}_+|\phi_n\rangle$ is also an eigenstate of \hat{L}_z with eigenvalue $(\beta_n + \hbar)$. Because the set of $\{\phi_n\}$ forms a complete basis set, the eigenstate $(\hat{L}_+|\phi_n\rangle)$ must belong to this set. The operator transforms one eigenstate into another with the latter corresponding to eigenvalue increased by \hbar as compared to the former.

Similarly, we can show that

$$\hat{L}_z(\hat{L}_-|\phi_n\rangle) = (\beta_n - \hbar)(\hat{L}_-|\phi_n\rangle). \quad \leftarrow \begin{array}{l} \text{See Problem 3} \\ \text{Problem Set 3} \end{array}$$

If $|\phi_n\rangle$ is an eigenstate of \hat{L}_z with eigenvalue β_n , then $\hat{L}_-|\phi_n\rangle$ is also an eigenstate of \hat{L}_z with eigenvalue $(\beta_n - \hbar)$.

Let's now investigate combinations of \hat{L}^2 and \hat{L}_+/\hat{L}_- ,

$$\hat{L}_+ \hat{L}^2 |\phi_n\rangle = \hat{L}_+ \alpha_n |\phi_n\rangle = \alpha_n \hat{L}_+ |\phi_n\rangle.$$

Because $[\hat{L}^2, \hat{L}_x] = [\hat{L}^2, \hat{L}_y] = 0$, we must have $[\hat{L}^2, \hat{L}_+] = [\hat{L}^2, \hat{L}_-] = 0$, so

$$\hat{L}^2 \hat{L}_+ |\phi_n\rangle = \hat{L}_+ \hat{L}^2 |\phi_n\rangle = \alpha_n \hat{L}_+ |\phi_n\rangle.$$

If $|\phi_n\rangle$ is an eigenstate of \hat{L}^2 with eigenvalue α_n , then $\hat{L}_+ |\phi_n\rangle$ is also an eigenstate of \hat{L}^2 with the same eigenvalue.

Similarly,

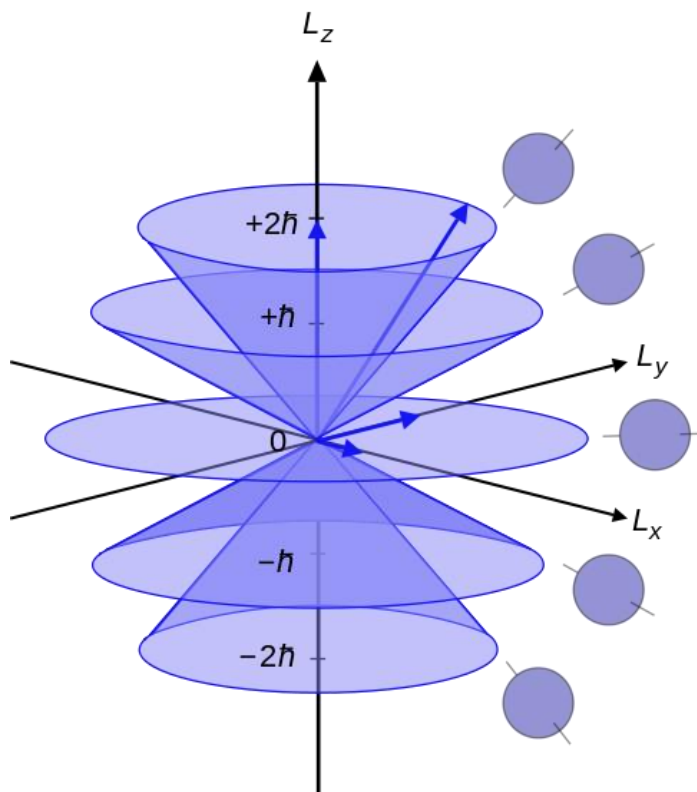
$$\hat{L}^2 (\hat{L}_- |\phi_n\rangle) = \alpha_n (\hat{L}_- |\phi_n\rangle).$$

If $|\phi_n\rangle$ is an eigenstate of \hat{L}^2 with eigenvalue α_n , then $\hat{L}_- |\phi_n\rangle$ is also an eigenstate of \hat{L}^2 with the same eigenvalue.

What does all this maths actually mean?

Consider a state $|\phi_n\rangle$ which corresponds to the value α_n of the magnitude of angular momentum, \hat{L}^2 , and the value β_n of its z-component \hat{L}_z . When we act with \hat{L}_+ or \hat{L}_- on this state, we “jump” to another eigenstate, $|\phi_{n'}\rangle$ which corresponds to the same eigenvalue α_n of \hat{L}^2 but $(\beta_n \pm \hbar)$ of \hat{L}_z . We conclude that

- 1) there are groups of eigenstates corresponding to a single value of α_n (length of the angular momentum “vector”) but different values of β_n (its projection on the z-axis);
- 2) the z-component of the angular momentum can only change in jumps of \hbar .



Within our “vector” model we just imagined, it makes sense to assume that

- 1) For a given value of α_n , $\alpha_n \geq \beta_n$ ² (projection on the z-axis can be at most as long as the length of the arrow). This means that for a given α_n , there are corresponding minimum and maximum values of β_n (minimum and maximum z-components of angular momentum corresponding to the “vector arrow” pointing “down” or “up”). We shall refer to them as β_n^{\min} and β_n^{\max} .
- 2) The eigenvalues β_n^{\min} and β_n^{\max} correspond to eigenstates we shall denote $|\phi_n^{\min}\rangle$ and $|\phi_n^{\max}\rangle$.

Question:

What happens when we act with \hat{L}_+ on $|\phi_n^{\max}\rangle$?

$$\hat{L}_+|\phi_n^{\max}\rangle = 0$$

Hence, $\hat{L}_-\hat{L}_+|\phi_n^{\max}\rangle = \hat{L}_-0 = 0$.

However, we have derived before the result

$$\hat{L}_-\hat{L}_+ = \hat{L}^2 - \hat{L}_z^2 - \hbar\hat{L}_z.$$

Hence,

$$\hat{L}_-\hat{L}_+|\phi_n^{\max}\rangle = (\hat{L}^2 - \hat{L}_z^2 - \hbar\hat{L}_z)|\phi_n^{\max}\rangle = 0$$

$$(\hat{L}^2 - \hat{L}_z^2 - \hbar\hat{L}_z)|\phi_n^{\max}\rangle = [\alpha_n - (\beta_n^{\max})^2 - \hbar\beta_n^{\max}]|\phi_n^{\max}\rangle = 0$$

$$\text{Thus, } \alpha_n - (\beta_n^{\max})^2 - \hbar\beta_n^{\max} = 0 \Rightarrow \beta_n^{\max}(\beta_n^{\max} + \hbar) = \alpha_n$$

In a similar way, starting from $\hat{L}_-|\phi_n^{\min}\rangle = 0$ and investigating $\hat{L}_+\hat{L}_-|\phi_n^{\min}\rangle$, we get

$$\beta_n^{\min}(\beta_n^{\min} - \hbar) = \alpha_n$$

See Problem 4
Problem Set 3



Notice that the two conditions can be reconciled if

$$\beta_n^{\min} = -\beta_n^{\max}.$$

The operators \hat{L}_+ and \hat{L}_- are called ladder operators as they move the eigenvalues β_n on a ladder from β_n^{\min} to β_n^{\max} . It is a ladder because only constant steps of $\pm\hbar$ are allowed. The total “length” of the ladder must be equal to integer number of hops, n ,

$$\beta_n^{\max} - \beta_n^{\min} = n\hbar \Rightarrow \beta_n^{\max} = -\beta_n^{\min} = \frac{n}{2}\hbar = \ell\hbar$$

Using this result, we get

$$\alpha_n = \beta_n^{\max}(\beta_n^{\max} + \hbar) = \ell(\ell + 1)\hbar^2,$$

where ℓ can be an integer (ladder of the type $\dots, -\hbar, 0, \hbar, \dots$) or half-integer (ladder of the type $\dots, -\frac{\hbar}{2}, \frac{\hbar}{2}, \dots$). For orbital angular momentum (what we have discussed so far), only the integer solutions matter. However, for more general angular momentum, the half-integer solutions are also relevant.

For each value of ℓ , eigenvalues of \hat{L}_z can be written as $\beta_n = m_l\hbar$, where m_l varies in integer steps from $-\ell$ to $+\ell$.

We see that our initial index n counting the eigenstates $|\phi_n\rangle$ should rather be replaced with two “good” quantum numbers for this problem, ℓ and m_l . A unique eigenstate can be denoted simply as $|\phi_n\rangle \equiv |\phi_{\ell,m}\rangle \equiv |\ell, m_l\rangle$. In this notation,

$$\hat{L}^2|\ell, m_l\rangle = \ell(\ell + 1)\hbar^2|\ell, m_l\rangle,$$

$$\hat{L}_z|\ell, m_l\rangle = m_l\hbar|\ell, m_l\rangle.$$

Also, we know that

$$\hat{L}_+|\ell, m_l\rangle \propto |\ell, m_l + 1\rangle, \quad \hat{L}_-|\ell, m_l\rangle \propto |\ell, m_l - 1\rangle,$$

that is, the ladder operators transform a state $|\ell, m_l\rangle$ into one with the quantum number m_l increased or decreased by one. In general, they also generate a scalar prefactor which we have not determined.

Side comment: because of other contexts, you may also hear \hat{L}_+ and \hat{L}_- (or similar ladder operators) referred to as “creation” and “annihilation” operators.

Using Dirac notation and operator algebra, we have “quantized” 3D space under the condition of spherical symmetry. We have done that without writing down the Schrodinger equation once and there was no need to derive/guess solutions of some differential equation. Our solution is very general and includes, for example, the hydrogen atom problem (or any other atom, for that matter). The details of the particle (charge) and the potential (caused by a proton) are not important.

However, we are missing some details because of two points:

- 1) quantization of the wave function along the distance away from the centre (dependence on the radius which is not described by angular momentum);
- 2) spin of the electron has purely quantum mechanical origin and we cannot obtain it simply by quantizing the classical equations for angular momentum.

We will discuss point 2) in the next part and then return to 1).

Finally, the drawback of our method is that we cannot say anything about the form of the eigenstates $|\phi_n\rangle \equiv |\ell, m\rangle$ in, let's say, position representation (their functional form as functions of position) – for this, we need to solve the right differential equations. To do this, we can start from

$$\hat{L} = \hat{r} \times \hat{p} = -i\hbar \vec{r} \times \vec{\nabla},$$

where, using spherical coordinates,

$$\vec{\nabla} = \vec{e}_r \frac{\partial}{\partial r} + \vec{e}_\theta \frac{1}{r} \frac{\partial}{\partial \theta} + \vec{e}_\varphi \frac{1}{r \sin \theta} \frac{\partial}{\partial \varphi},$$

with the radius r , in-plane angle φ and z-axis angle θ and \vec{e}_i denoting the unit vector for the i -th coordinate.

Because they are concerned with angular momentum, neither \hat{L} nor \hat{L}_z depend on r ,

$$\hat{L} = -i\hbar \left(\vec{e}_\varphi \frac{\partial}{\partial \theta} - \vec{e}_\theta \frac{1}{\sin \theta} \frac{\partial}{\partial \varphi} \right),$$

$$\hat{L}_z = -i\hbar \frac{\partial}{\partial \varphi},$$

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

Our eigenproblems for \hat{L}^2 and \hat{L}_z become differential equations,

$$\hat{L}^2 Y_\ell^{m_l}(\varphi, \theta) = \ell(\ell + 1)\hbar^2 Y_\ell^{m_l}(\varphi, \theta),$$

$$\hat{L}_z Y_\ell^{m_l}(\varphi, \theta) = m_l \hbar Y_\ell^{m_l}(\varphi, \theta),$$

where $Y_\ell^{m_l}(\varphi, \theta)$, called the spherical harmonics, correspond to $|\ell, m_l\rangle$. Their exact functional form can be obtained by solving the differential equations explicitly. Here, I provide some examples:

$$Y_0^0(\varphi, \theta) = \sqrt{\frac{1}{4\pi}},$$

$$Y_1^0(\varphi, \theta) = \sqrt{\frac{3}{4\pi}} \cos \theta,$$

$$Y_1^{\pm 1}(\varphi, \theta) = \mp \sqrt{\frac{3}{8\pi}} \sin \theta e^{\pm i\varphi}.$$

At this point...

You should be able to solve problems 1 to 4 from Problem Set 3.

Final comments:

1. Remember that there was nothing special in the z -component of angular momentum! We could have just as well performed the analysis by choosing the operators \hat{L}^2 and, say, \hat{L}_x . Apart from minor changes in the maths (like the definition of \hat{L}_+ and \hat{L}_-), nothing would really change. Hence, our main conclusion about the possible eigenvalues for \hat{L}_z has to apply to \hat{L}_x and \hat{L}_y as well: components of angular momentum are quantised in steps of \hbar ; for given orbital quantum number ℓ we can only obtain x -, y - or z -component of angular momentum equal to $m_l \hbar$, $m_l = -\ell, \dots, \ell$ (note that for different components we need different eigenfunctions).
2. In this section, we frequently imagined quantum angular momentum as a vector. This picture can only be stretched so far. For example, notice that the magnitude of this “vector” is greater than the maximum z -component, $\sqrt{\ell(\ell + 1)} > \ell$ (except for trivial $\ell = 0$). Evidently, we cannot get the angular momentum to point perfectly along the z direction. Why can't we just pick our axes so that z points along whatever the direction of angular momentum is? To do this, we need to know all three components simultaneously and the uncertainty principle says it is impossible. Can we hope that at least once in a while, by sheer luck, the z -axis will be pointing along the direction of angular momentum? No – it's not simply that we don't know the components of angular momentum, the particle cannot have a determinate angular momentum vector any more than it can simultaneously have a determinate position and momentum. If \hat{L}_z has a well-defined value, then \hat{L}_x and \hat{L}_y do not.

Summary

1. Different components of the angular momentum do not commute:

$$\begin{aligned}[\hat{L}_x, \hat{L}_y] &= i\hbar\hat{L}_z \\ [\hat{L}_y, \hat{L}_z] &= i\hbar\hat{L}_x \\ [\hat{L}_z, \hat{L}_x] &= i\hbar\hat{L}_y\end{aligned}$$

2. However, any component of angular momentum commutes with \hat{L}^2 .
3. By purely algebraic means, we determined the eigenvalues of \hat{L}^2 and (because of a convention) \hat{L}_z .
4. Angular momentum is quantized:
 - a) Angular quantum number ℓ determines the squared magnitude of angular momentum, $\ell(\ell + 1)\hbar^2$.
 - b) For given ℓ , eigenvalues for the x -, y - and z -components of angular momentum are equal to $m_l\hbar$, $m_l = -\ell, \dots, \ell$.
5. The eigenfunctions of the \hat{L}_z operator are the spherical harmonics $Y_\ell^{m_l}(\varphi, \theta)$.

Spin

In classical mechanics, a rigid object admits two kinds of angular momentum: orbital ($\vec{L} = \vec{r} \times \vec{p}$), associated with the motion of the centre of mass, and spin ($\vec{S} = I\vec{\omega}$), associated with motion about the centre of mass. This distinction is largely a matter of convenience because the latter is nothing but the sum of the “orbital” momenta of smaller masses (pieces of the larger body). In quantum mechanics, the distinction is fundamental. Here, we will focus on the case of spin $\frac{1}{2}\hbar$ (electron’s case; the requirement for this and not other value follows from the Stern-Gerlach experiment).



Wolfgang Pauli (1900 – 1958)

- Proposed a new quantum number for electrons and formulated the Pauli exclusion principle (1924)
- Introduced Pauli matrices and solved the relativistic spin-1/2 problem (1927)
- Nobel Prize in Physics (1945) for “decisive contribution through his discovery of a new law of Nature, the exclusion principle or Pauli principle”

Question:

If the electron were a classical solid sphere with radius

$$r_c = \frac{e^2}{4\pi\epsilon_0 mc^2}$$

(the so-called classical electron radius, obtained by assuming the electron's mass is attributable to energy stored in its electric field via the Einstein formula $E = mc^2$), and its angular momentum is $\frac{\hbar}{2}$, then how fast (in m/s) would a point on the “equator” be moving? Does this model make sense?

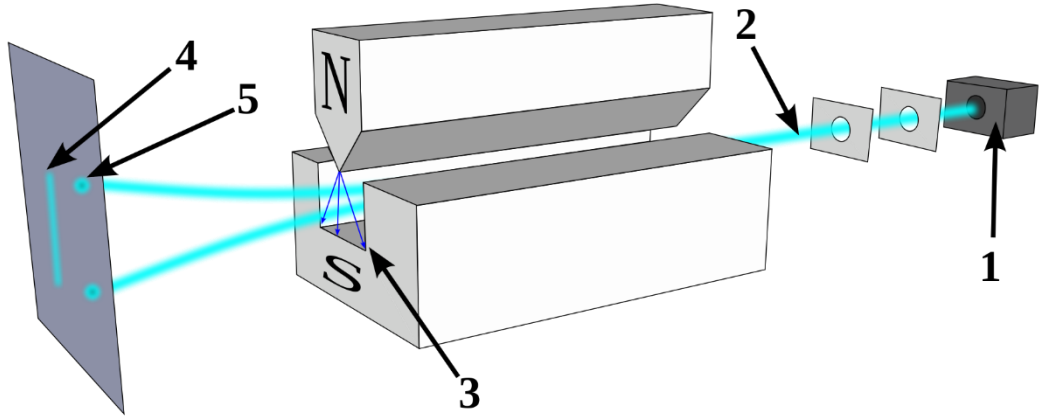


Figure: Stern–Gerlach experiment: silver atoms travelling through an inhomogeneous magnetic field, deflected up or down depending on their spin; (1) furnace, (2) beam of silver atoms, (3) inhomogeneous magnetic field, (4) classically expected result, (5) observed result.

Consider neutral (to avoid the Lorentz force) atoms passed through a setup like in the figure. Classically, given the Hamiltonian of a magnetic moment $\vec{\mu}$ in a magnetic field \vec{B} ,

$$\mathcal{H} = -\vec{\mu} \cdot \vec{B},$$

any inhomogeneity of the magnetic field will lead to a force,

$$\vec{F} = -\nabla\mathcal{H} = \nabla(\vec{\mu} \cdot \vec{B}) = \nabla(\vec{\mu} \cdot \nabla)\vec{B}$$

(I used here some vector identities, the fact that $\vec{\mu}$ does not depend on position and that $\nabla \times \vec{B} = 0$). The x - and y -components of the force vanish on average because of the precession of the spin (will come back to this). Hence,

$$\vec{F} = \mu_z \frac{\partial B_z}{\partial z} \vec{k}.$$

Classically, since μ_z is continuous, we expect a continuous trace (4). The actual experiment reveals a series of discrete dots (5) – hence, μ_z is discrete.

To test electron spin, one can use a beam of hydrogen atoms.

Following our earlier discussion of angular momentum, we postulate existence of spin operators $\hat{S}_x, \hat{S}_y, \hat{S}_z$ corresponding to its components along the three axes and obeying the commutation relations

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

It follows like before that the eigenvectors of $\hat{S}^2 = \hat{S}_x^2 + \hat{S}_y^2 + \hat{S}_z^2$ and \hat{S}_z satisfy

$$\hat{S}^2|s, m\rangle = s(s+1)\hbar^2|s, m\rangle,$$

$$\hat{S}_z|s, m\rangle = m\hbar|s, m\rangle,$$

where we now use symbol s (for spin) instead of ℓ . For $s = \frac{1}{2}$, there are only two eigenstates,

$$1) \quad s = \frac{1}{2}, m = \frac{1}{2} \rightarrow \left| \frac{1}{2}, \frac{1}{2} \right\rangle \equiv \left| \frac{1}{2} \right\rangle \equiv |\uparrow\rangle,$$

$$2) \quad s = \frac{1}{2}, m = -\frac{1}{2} \rightarrow \left| \frac{1}{2}, -\frac{1}{2} \right\rangle \equiv \left| -\frac{1}{2} \right\rangle \equiv |\downarrow\rangle.$$

A general spin state of a spin-1/2 particle is a linear combination of these two eigenstates,

$$|\chi\rangle = a|\uparrow\rangle + b|\downarrow\rangle.$$

In matrix representation, we associate a two-element column vector (spinor) with each of the eigenstates,

$$|\uparrow\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad |\downarrow\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix},$$

so that $|\chi\rangle = a \begin{bmatrix} 1 \\ 0 \end{bmatrix} + b \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix}.$

In this representation, operators $\hat{S}_x, \hat{S}_y, \hat{S}_z, \hat{S}^2$ must be 2×2 matrices. Notice that, as spin is an inherent property of a particle, no simple wave mechanistic description is possible (no dependence on position – no wave function to write down).

We will now find the matrix forms of operators $\hat{S}_x, \hat{S}_y, \hat{S}_z, \hat{S}^2$ and discuss their eigenvalues and eigenvectors. In doing this, we will apply some of the basic “maths” rules in Dirac notation and see how simple manipulation provides physical insight.

Starting from \hat{S}^2 , we can write it as a general 2×2 matrix

$$\hat{S}^2 = \begin{bmatrix} c & d \\ e & f \end{bmatrix},$$

With c, d, e , and f to be determined. We require (see earlier)

$$\hat{S}^2|\uparrow\rangle = \begin{bmatrix} c & d \\ e & f \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = s(s+1)\hbar^2|\uparrow\rangle = \frac{3}{4}\hbar^2 \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

what gives

$$\begin{bmatrix} c \\ e \end{bmatrix} = \frac{3}{4}\hbar^2 \begin{bmatrix} 1 \\ 0 \end{bmatrix} \Rightarrow c = \frac{3}{4}\hbar^2 \text{ and } e = 0.$$

Similarly,

$$\hat{S}^2|\downarrow\rangle = \begin{bmatrix} c & d \\ e & f \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = s(s+1)\hbar^2|\downarrow\rangle = \frac{3}{4}\hbar^2 \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

gives

$$\begin{bmatrix} d \\ f \end{bmatrix} = \frac{3}{4}\hbar^2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} \Rightarrow d = 0 \text{ and } f = \frac{3}{4}\hbar^2.$$

Hence,

$$\hat{S}^2 = \frac{3}{4}\hbar^2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

We find \hat{S}_z similarly by requiring

$$\left. \begin{aligned} \hat{S}_z|\uparrow\rangle &= \frac{\hbar}{2}|\uparrow\rangle \Rightarrow \begin{bmatrix} c & d \\ e & f \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \frac{\hbar}{2} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \\ \hat{S}_z|\downarrow\rangle &= -\frac{\hbar}{2}|\downarrow\rangle \Rightarrow \begin{bmatrix} c & d \\ e & f \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = -\frac{\hbar}{2} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \end{aligned} \right\} \Rightarrow \hat{S}_z = \frac{\hbar}{2} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

How can we find \hat{S}_x and \hat{S}_y ? Recall that for angular momentum in general,

$$\hat{L}_+|\ell, m\rangle \propto |\ell, m+1\rangle, \quad \hat{L}_-|\ell, m\rangle \propto |\ell, m-1\rangle,$$

where $\hat{L}_\pm = \hat{L}_x \pm i\hat{L}_y$. Moving to \hat{S}_+ and \hat{S}_- operators,

$$\hat{S}_+|\downarrow\rangle = z|\uparrow\rangle \quad \text{and} \quad \hat{S}_-|\uparrow\rangle = \tilde{z}|\downarrow\rangle, \quad (+)$$

where the complex numbers z and \tilde{z} are currently unknown. We also have the equalities describing the end of our “ladder”

$$\hat{S}_+|\uparrow\rangle = 0 \quad \text{and} \quad \hat{S}_-|\downarrow\rangle = 0. \quad (++)$$

Writing

$$\hat{S}_+ = \begin{bmatrix} c & d \\ e & f \end{bmatrix}, \quad \hat{S}_- = \begin{bmatrix} \tilde{c} & \tilde{d} \\ \tilde{e} & \tilde{f} \end{bmatrix},$$

we use $(++)$ to get $c = e = \tilde{d} = \tilde{f} = 0$. To use $(+)$, we need to determine z and \tilde{z} first. Let us look at

$$\begin{aligned} \langle \downarrow | \hat{S}_- \hat{S}_+ | \downarrow \rangle &= \langle (\hat{S}_-)^{\dagger} \downarrow | \hat{S}_+ \downarrow \rangle = \langle (\hat{S}_x - i\hat{S}_y)^{\dagger} \downarrow | \hat{S}_+ \downarrow \rangle \\ &= \langle (\hat{S}_x + i\hat{S}_y) \downarrow | \hat{S}_+ \downarrow \rangle = \langle \hat{S}_+ \downarrow | \hat{S}_+ \downarrow \rangle = z^* \langle \uparrow | z | \uparrow \rangle = |z|^2. \end{aligned}$$

At the same time, recalling $\hat{L}_- \hat{L}_+ = \hat{L}^2 - \hat{L}_z^2 - \hbar \hat{L}_z$,

$$\begin{aligned} \langle \downarrow | \hat{S}_- \hat{S}_+ | \downarrow \rangle &= \langle \downarrow | \hat{S}^2 - \hat{S}_z^2 - \hbar \hat{S}_z | \downarrow \rangle \\ &= \left\langle \downarrow \left| \frac{3}{4} \hbar^2 - \frac{1}{4} \hbar^2 - \hbar \left(-\frac{1}{2} \hbar \right) \right| \downarrow \right\rangle = \hbar^2 \langle \downarrow | \downarrow \rangle = \hbar^2. \end{aligned}$$

Hence, $z = \hbar$. Similarly, investigating $\langle \uparrow | \hat{S}_+ \hat{S}_- | \uparrow \rangle$ yields $\tilde{z} = \hbar$.

We can now use (+) to find $f = \tilde{c} = 0$ and $d = \tilde{e} = \hbar$ so that

$$\hat{S}_+ = \hbar \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad \hat{S}_- = \hbar \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}.$$

Because we have

$$\left. \begin{aligned} \hat{S}_+ &= \hat{S}_x + i\hat{S}_y \\ \hat{S}_- &= \hat{S}_x - i\hat{S}_y \end{aligned} \right\} \Rightarrow \begin{cases} \hat{S}_x = \frac{1}{2}(\hat{S}_+ + \hat{S}_-) \\ \hat{S}_y = \frac{1}{2i}(\hat{S}_+ - \hat{S}_-) \end{cases},$$

we obtain

$$\hat{S}_x = \frac{\hbar}{2} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \hat{S}_y = \frac{\hbar}{2} \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}.$$

The matrices

$$\sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \sigma_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad \sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix},$$

are known as the Pauli matrices.

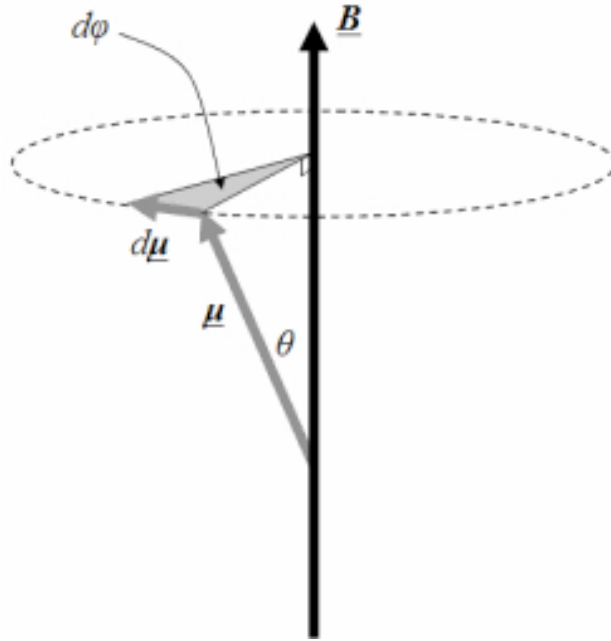
Question:

Can you determine eigenvalues and eigenvectors of \hat{S}_x and \hat{S}_y ?

At this point...

You should be able to solve problems 5 to 7 from Problem Set 3.

Consider magnetized object spinning about centre of mass, with angular momentum \vec{L} and magnetic moment $\vec{\mu} = \gamma \vec{L}$, with γ the gyromagnetic ratio.



An external magnetic field \vec{B} causes a torque,

$$\frac{d\vec{L}}{dt} = \vec{\tau} = \vec{\mu} \times \vec{B} = \gamma \vec{L} \times \vec{B}.$$

With $\vec{B} = B\vec{k}$, we have

$$\left. \begin{aligned} \frac{dL_x}{dt} &= \gamma L_y B \\ \frac{dL_y}{dt} &= -\gamma L_x B \end{aligned} \right\} \rightarrow \frac{d(L_x + iL_y)}{dt} = -i\gamma B(L_x + iL_y),$$

as well as $\frac{dL_z}{dt} = 0 \rightarrow L_z = \text{const.}$ Recognizing $L_+ = L_x + iL_y$

$$\frac{dL_+}{dt} = -i\gamma B L_+ \rightarrow L_+ = L_+^0 e^{-i\gamma B t}.$$

This means that the angular momentum \vec{L} precesses about magnetic field direction with angular velocity $\vec{\omega}_0 = -\gamma \vec{B}$, independent of the angle between \vec{L} ($\vec{\mu}$) and \vec{B} .

What about a quantum particle? Here, we focus on an electron.

First: For a general state $\alpha|\uparrow\rangle + \beta|\downarrow\rangle$, how do α and β relate to the orientation of the spin?

Let us assume that spin is pointing along the unit vector $\vec{n} = (\cos \varphi \sin \theta, \sin \varphi \sin \theta, \cos \theta)$, i.e., in direction (φ, θ) . Spin must be an eigenstate of $\vec{n} \cdot \hat{\sigma}$ with eigenvalue 1:

$$\begin{bmatrix} n_z & n_x - in_y \\ n_x + in_y & -n_z \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix},$$

and $|\alpha|^2 + |\beta|^2 = 1$. This gives

$$\begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} e^{-i\varphi/2} \cos \theta/2 \\ e^{i\varphi/2} \sin \theta/2 \end{bmatrix}.$$

Note that under a 2π rotation,

$$\begin{bmatrix} \alpha \\ \beta \end{bmatrix} \rightarrow - \begin{bmatrix} \alpha \\ \beta \end{bmatrix}$$

To return the spin to the starting point, revolution by 4π is necessary (for spin 1, rotation by 2π is needed and for spin 2, only π !).

Second: Consider a small current loop consisting of an electron in uniform circular motion. Quasiclassically, the electron's orbital angular momentum \vec{L} is related to the magnetic moment $\vec{\mu}$ of the loop,

$$\vec{\mu} = -\frac{e}{2m_e}\vec{L} = \gamma\vec{L},$$

where e and m_e are electron charge and mass, respectively. This suggests that there may be a similar relationship between magnetic moment and spin angular momentum. We can write

$$\hat{\mu} = -\frac{ge}{2m_e}\hat{S} \equiv \gamma\hat{S},$$

Where g is the Landé g -factor. Classically, we would expect $g = 1$ – instead, $g = 2.0023192$. The fact that the g -factor is (almost) twice that expected from classical physics is only explicable using relativistic quantum mechanics. The small corrections to the relativistic result $g = 2$ come from quantum field theory.

By analogy with the classical expression, we write the quantum Hamiltonian,

$$\hat{H} = -\hat{\vec{\mu}} \cdot \hat{\vec{B}},$$

where I have for added emphasis explicitly marked the quantities as both operators and vectors (that is, with three components) – just like with \hat{L} , we will not always do so. Given our geometry, we have

$$\hat{H} = -\hat{\vec{\mu}} \cdot \vec{B} = -\gamma\hat{S} \cdot \vec{B} = -\gamma\frac{\hbar}{2}\sigma_z B.$$

The dynamics can be inferred by using the time evolution operator,

$$\begin{aligned}\hat{U}(t) &= e^{-i\hat{H}t/\hbar} = \exp\left[-i\frac{\gamma}{2}\sigma_z Bt\right], \\ |\psi(t)\rangle &= \hat{U}|\psi(0)\rangle \rightarrow \begin{bmatrix} \alpha(t) \\ \beta(t) \end{bmatrix} = \begin{bmatrix} e^{-i\omega_0 t/2} & 0 \\ 0 & e^{i\omega_0 t/2} \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} \\ &= \begin{bmatrix} e^{-i(\varphi+\omega_0 t)/2} \cos \theta/2 \\ e^{i(\varphi+\omega_0 t)/2} \sin \theta/2 \end{bmatrix},\end{aligned}$$

where we defined $\vec{\omega}_0 = -\gamma\vec{B} = -g\omega_c\vec{k}$, with $\omega_c = \frac{eB}{2m_e}$ the cyclotron frequency. This means that spin precesses with frequency $\vec{\omega}_0$ (the exercise below will confirm this).

Exercise: Find the expectation values of the spin components for an electron in an external magnetic field along the z-axis. How do these compare to the results of a single measurement?

We will explore the analogy between the dynamics of a quantum classical magnetic moment a little further. Suppose we view the precession in a frame that is rotating at a frequency $\vec{\omega}$ parallel to our $\vec{B} = B\vec{k}$. In this rotating frame, the precession frequency is

$$\vec{\omega}_r = \vec{\omega}_0 - \vec{\omega} = -\gamma(B + \omega/\gamma)\vec{k}$$

(if the frame rotates exactly at precession frequency, $\vec{\omega} = \vec{\omega}_0$, spins pointing in any direction will remain at rest in that frame, as expected). The effective field is now $(B + \omega/\gamma)\vec{k}$. Suppose we then add a small additional component of the magnetic field which (in laboratory frame) is rotating with angular frequency $\vec{\omega}$ in the xy -plane,

$$\vec{B} = B\vec{k} + B_\omega(\cos(\omega t)\vec{i} - \sin(\omega t)\vec{j}).$$

The effective magnetic field in a frame rotating with the same frequency ω , is $\vec{B} = (B + \omega/\gamma)\vec{k} + B_\omega\vec{i}$. If we tune ω , so that it matches the precession frequency in the original magnetic field, $\omega = -\gamma B$, in the rotating frame the magnetic moment will only see the small field in the x -direction. Spin will therefore precess about the x -direction at slow angular frequency γB_ω - so called (para)magnetic resonance. This general principle applied to nuclear magnetic moments (the resonance frequency depends on the nucleus) underpins nuclear magnetic resonance (NMR) and magnetic resonance imaging (MRI).

Question: What happens when the in-plane field is removed?

Summary

1. Electron spin cannot be explained in terms of classical angular momentum of electron.
2. The x -, y - and z -components of electron spin can be described using 2×2 (Pauli) matrices

$$\hat{S}_x = \frac{\hbar}{2} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \hat{S}_y = \frac{\hbar}{2} \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \hat{S}_z = \frac{\hbar}{2} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

3. It is common to use the notation $|\uparrow\rangle$ and $|\downarrow\rangle$ for the eigenstates of \hat{S}_z corresponding to the eigenvalues $\frac{\hbar}{2}$ and $-\frac{\hbar}{2}$, respectively.
4. In the presence of an external magnetic field, spins undergo Larmor precession around the direction of the applied field. Application of an additional oscillating field can be used to select a new precession axis.

PH20013/60 Quantum Mechanics

Section III:

A few (nonexaminable) words on quantum weirdness

In this section, we shall discuss two examples which apply many elements from the previous sections and showcase some of the inherently novel questions physicists had to (and to some extent still have to) face after discovering/learning quantum mechanics.

Case Study I: Two electron spins

In the previous section, we have constructed a framework to describe electron spin. How do we go about describing two electrons? Here, we will skip the technical details and introduce an intuitive generalization of our single spin case. Consider the eigenstates of \hat{S}_z : clearly, we can list four cases

- 1) $|\uparrow_1\rangle|\uparrow_2\rangle \equiv |\uparrow\uparrow\rangle,$
- 2) $|\uparrow_1\rangle|\downarrow_2\rangle \equiv |\uparrow\downarrow\rangle,$
- 3) $|\downarrow_1\rangle|\uparrow_2\rangle \equiv |\downarrow\uparrow\rangle,$
- 4) $|\downarrow_1\rangle|\downarrow_2\rangle \equiv |\downarrow\downarrow\rangle,$

where the indices number the two electrons and correspond to the ordering of the arrows in the “two-electron” kets.

On the other hand, imagine we see the two electrons as a single “particle”: what is its spin s ? Without much rigour, we will argue that two cases are possible:

- 1) $s = s_1 + s_2 = \frac{1}{2} + \frac{1}{2} = 1$ (“vectors” add up),
- 2) $s = s_1 - s_2 = \frac{1}{2} - \frac{1}{2} = 0$ (“vectors” cancel out).

Let's assume that for our two-electron case, spin is the only form of angular momentum to concern ourselves with (these two electrons do not orbit around a nucleus). Then, we interpret spin s as total angular momentum. Following our earlier discussions of angular momentum and spin, we conclude that

- For the state $s = 0$, we can assign quantum numbers $s = 0$ and $m_s = 0$, leading to one eigenstate $|s = 0, m_s = 0\rangle$, for which

$$\hat{S}^2|0,0\rangle = 0|0,0\rangle; \quad \hat{S}_z|0,0\rangle = 0|0,0\rangle.$$

- For the state $s = 1$, we can assign quantum numbers $s = 1$ and $m_s = -1, 0, 1$, leading to three eigenstates $|1, -1\rangle$, $|1, 0\rangle$ and $|1, 1\rangle$ for which

- 1) $\hat{S}^2|1, -1\rangle = 2\hbar^2|1, -1\rangle; \quad \hat{S}_z|1, -1\rangle = -\hbar|1, -1\rangle,$
- 2) $\hat{S}^2|1, 0\rangle = 2\hbar^2|1, 0\rangle; \quad \hat{S}_z|1, 0\rangle = 0|1, 0\rangle,$
- 3) $\hat{S}^2|1, 1\rangle = 2\hbar^2|1, 1\rangle; \quad \hat{S}_z|1, 1\rangle = \hbar|1, 1\rangle.$

Intuition suggests that we can associate two of the eigenstates of $s = 1$ with the two-electron states,

$|\uparrow\uparrow\rangle \equiv |1, 1\rangle$, (largest z -component pointing up)

$|\downarrow\downarrow\rangle \equiv |1, -1\rangle$. (largest z -component pointing down)

However, the other two states ($|\uparrow\downarrow\rangle$ and $|\downarrow\uparrow\rangle$) are similar – why should one correspond to $s = 0$ and the other to $s = 1$?

Recall that,

$$\hat{S}_+ |1, -1\rangle \propto |1, 0\rangle.$$

Hence, $\hat{S}_+ |\downarrow\downarrow\rangle \propto |1, 0\rangle$, which is the state we are trying to identify. At the same time, for this 2-electron particle, $\hat{S}_+ = \hat{S}_+^{(1)} + \hat{S}_+^{(2)}$, where the superscripts refer to the first or the second electron. So,

$$\begin{aligned}\hat{S}_+ |1, -1\rangle &= (\hat{S}_+^{(1)} + \hat{S}_+^{(2)}) |\downarrow\downarrow\rangle = \hat{S}_+^{(1)} |\downarrow_1\rangle |\downarrow_2\rangle + \hat{S}_+^{(2)} |\downarrow_1\rangle |\downarrow_2\rangle \\ &= \hbar |\uparrow_1\rangle |\downarrow_2\rangle + \hbar |\downarrow_1\rangle |\uparrow_2\rangle = \hbar [|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle].\end{aligned}$$

(operators for a specific particle only act on the ket describing this particle).

Hence, the state $|1, 0\rangle$ is a linear combination of $|\downarrow\uparrow\rangle$ and $|\uparrow\downarrow\rangle$,

$$|1, 0\rangle = \frac{1}{\sqrt{2}} [|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle]$$

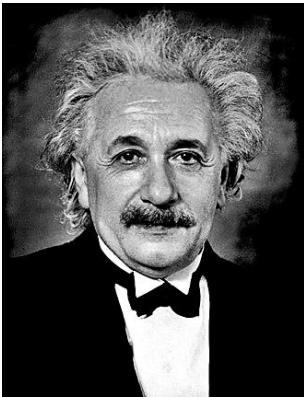
(the prefactor is due to normalization). This means that in a state with $s = 1$ (equivalent to $\ell = 1$) and $m_s = 0$, spins of individual electrons are not pointing in a specific directions – the state is a superposition of two configurations in which spins point opposite. We can find the state $|0, 0\rangle$ by looking for another linear combination that is orthogonal to $|1, 0\rangle$,

$$|0, 0\rangle = \frac{1}{\sqrt{2}} [|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle].$$

The state with angular momentum zero is also a superposition of spins pointing both up and down!

The $|0,0\rangle$ state is called a singlet (single state with $s = 0$) while the other three states are referred to as a triplet.

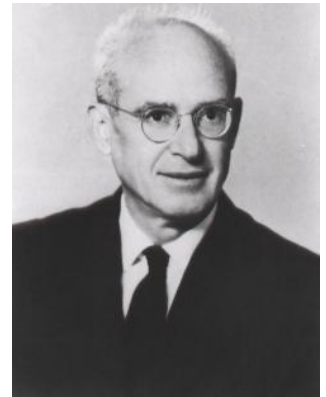
The two-electron system is a quantum version of Schrödinger's cat as in the states $|1,0\rangle$ and $|0,0\rangle$ electrons individually do not have a definite z-component of spin. We will employ this “catness” in the following discussion, known as the Einstein-Podolsky-Rosen (EPR) paradox first discussed in 1935.



A. Einstein



B. Podolsky

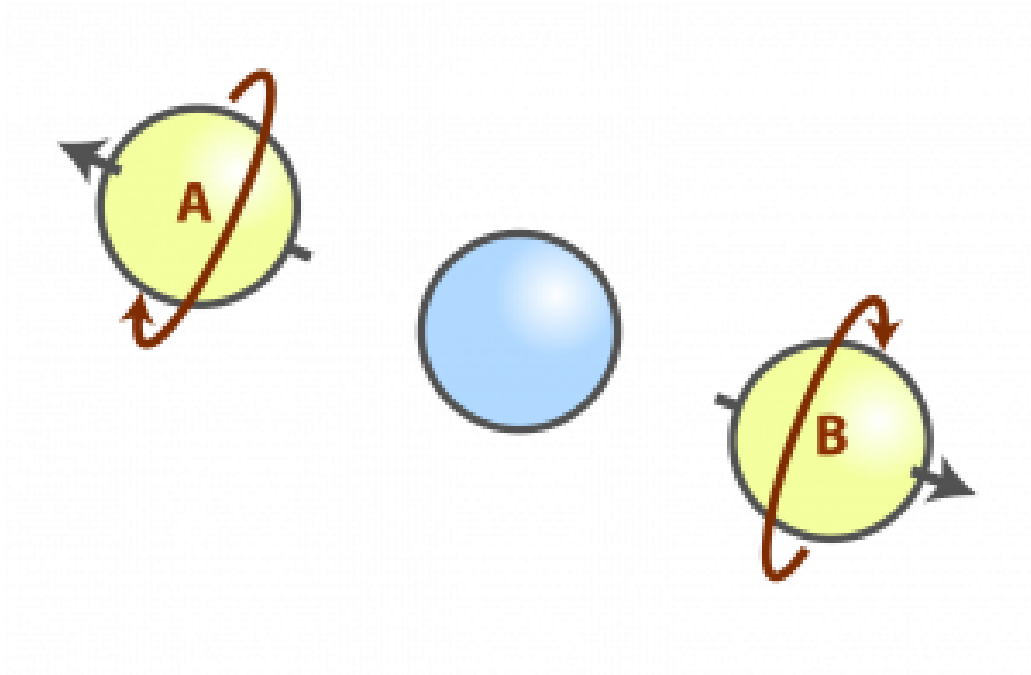


N. Rosen

Consider the decay of the neutral pi meson into an electron and a positron,

$$\pi^0 \rightarrow e^- + e^+.$$

Assuming the pi meson was at rest, the electron and positron fly off in opposite directions. Because π^0 has spin zero, conservation of angular momentum requires that the electron and positron are in the singlet spin state.



$$|0,0\rangle = \frac{1}{\sqrt{2}} [|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle] \equiv \frac{1}{\sqrt{2}} [|\uparrow_1\rangle|\downarrow_2\rangle - |\downarrow_1\rangle|\uparrow_2\rangle]$$

The consequence of this is that, if the particle 1, let's say the electron, is found to have spin up, particle 2, the positron, must have spin down (the first in the bracket above). Of course, if the electron has spin down, then the positron must have spin up. Quantum mechanics can't tell us which combination will be the result of a measurement in any particular decay event, but it does say we get each combination half the time (on average) and that the spins of the two particles are correlated (we also say this is an example of an **entangled state**).

Now, imagine we let the two particles fly a large distance apart (for our gedankenexperiment it could even be light years apart). Then, we measure the spin of the electron. Let's say we get spin up – we immediately know that someone far away will get spin down as a result of their measurement. With our measurement, we collapsed the wave function of the electron as well as, at the same moment in time, the positron which is far away. Einstein, Podolsky and Rosen considered such “spooky action-at-a-distance” (Einstein's words) preposterous.

EPR argument rests on the assumption that no influence can propagate faster than speed of light (the principle of locality). You might be tempted to propose that the collapse of the wave function is not instantaneous but “travels” at some finite velocity. However, this would lead to violations of angular momentum conservation, for if we measured the spin of the positron before the news of the collapse had reached it, there would be a 50/50 chance of finding both particles with spin up. The experiments are unequivocal: no such violation occurs – the correlation of the spins is perfect and the collapse of the wave function instantaneous.

Einstein, Podolsky and Rosen did not doubt that quantum mechanics is correct – they only claimed that it must be incomplete; that the wave function is not the whole story and some other quantity, so called “hidden variable”, is needed in addition to Ψ to characterize the system fully.

However, in 1964 John Bell stated what is known as Bell' theorem:

No physical theory of local hidden variables can ever reproduce all of the predictions of quantum mechanics.

In Bell's words: "In a theory in which parameters are added to quantum mechanics to determine the results of individual measurements, without changing the statistical predictions, there must be a mechanism whereby the setting of one measuring device can influence the reading of another instrument, however remote. Moreover, the signal involved must propagate instantaneously, so that such a theory could not be Lorentz invariant."



John Stewart Bell (1928 – 1990)

- Stated what is known as Bell's theorem (1964)
- Supported what is known as de Broglie-Bohm's interpretation of QM (nonlocal hidden variable theory)

Most importantly, Bell's theorem (in the form of what is known as Bell's inequalities) can be tested experimentally. This was first done in 1972 and every few years more stringent tests (accounting for various loopholes either due to experimental imperfections or theoretical models) support the theory of quantum theory as we discussed it and not the hypothesis that some hidden variables are missing from our description.

Note that Bell's theorem still allows nonlocal hidden variable theories and did not in any way close the discussions about more philosophical implications of quantum mechanics.

On Moodle you can find a selection of articles discussing interpretations of quantum mechanics and/or Bell's theorem.

Case Study II: Wave function symmetry

Consider the Schrödinger equation for two particles,

$$i\hbar \frac{\partial}{\partial t} |\psi\rangle = \left[\frac{\hat{P}_1^2}{2m_1} + \frac{\hat{P}_2^2}{2m_2} + \hat{V}(\hat{X}_1, \hat{X}_2) \right] |\psi\rangle = H|\psi\rangle.$$

In general, $\hat{V}(\hat{X}_1, \hat{X}_2) \neq \hat{V}(\hat{X}_1) + \hat{V}(\hat{X}_2)$. However, if $\hat{V}(\hat{X}_1, \hat{X}_2) = \hat{V}(\hat{X}_1) + \hat{V}(\hat{X}_2)$, then the Hamiltonian is a sum of the Hamiltonians of the two particles,

$\hat{H} = \hat{H}_1 + \hat{H}_2$. We say that \hat{H} is separable. In such a case, the particles are independent, and the total energy is a sum of the energies of the two particles, $E = E_1 + E_2$.

Mathematically, the solution can simply be a product of the single-particle states,

$$\psi(\vec{r}_1, \vec{r}_2) = \psi_a(\vec{r}_1)\psi_b(\vec{r}_2),$$

where a and b emphasize different wave functions for the particles. Using kets, I will write this as

$$|\psi\rangle = |\psi_a\rangle_1 |\psi_b\rangle_2$$

where the subscripts 1 and 2 emphasize that the kets refer to different particles.

A solution like that above means that particle 1 is in a state $|\psi_a\rangle$ (wave function ψ_a) and particle 2 in a state $|\psi_b\rangle$ (wave function ψ_b). This means they are distinguishable (we know which particle is in what state and hence can tell them apart). This, however, is not what happens in quantum mechanics: quantum particles are indistinguishable. This requires a state which is noncommittal as to which particle is in which state.

It turns out that there are two ways to build linear combinations that do not distinguish between particles:

$$|\psi_{\pm}\rangle = \frac{1}{\sqrt{2}}(|\psi_a\rangle_1|\psi_b\rangle_2 \pm |\psi_b\rangle_1|\psi_a\rangle_2).$$

Particles called bosons (particles with integer spin) occupy the state symmetric under exchange of the particles,

$$\begin{aligned} |\psi_{\text{boson}}\rangle &= \frac{1}{\sqrt{2}}(|\psi_a\rangle_1|\psi_b\rangle_2 + |\psi_b\rangle_1|\psi_a\rangle_2) \\ &= \frac{1}{\sqrt{2}}(|\psi_b\rangle_1|\psi_a\rangle_2 + |\psi_a\rangle_1|\psi_b\rangle_2). \end{aligned}$$

In contrast, fermions (particles with half-integer spin) occupy the state anti-symmetric under particle exchange,

$$\begin{aligned} |\psi_{\text{fermion}}\rangle &= \frac{1}{\sqrt{2}}(|\psi_a\rangle_1|\psi_b\rangle_2 - |\psi_b\rangle_1|\psi_a\rangle_2) \\ &= -\frac{1}{\sqrt{2}}(|\psi_b\rangle_1|\psi_a\rangle_2 - |\psi_a\rangle_1|\psi_b\rangle_2). \end{aligned}$$

The minus sign in front is inconsequential here because quantum states related by a global phase lead to the same physics (any $|\dots|^2$ operation gets rid of the phase).

Note that it follows that two fermions cannot occupy the same state because for $|\psi_b\rangle_i = |\psi_a\rangle_i \equiv |\psi\rangle_i$,

$$|\psi_{\text{fermion}}\rangle = \frac{1}{\sqrt{2}}(|\psi\rangle_1|\psi\rangle_2 - |\psi\rangle_1|\psi\rangle_2) = 0.$$

This is the famous Pauli exclusion principle.

What are the consequences of such a requirement for the symmetry of the quantum state? Let's assume that the states $|\psi_a\rangle$ and $|\psi_b\rangle$ are orthogonal and compute the expectation value of the square of the separation between the two particles,

$$\langle (x_1 - x_2)^2 \rangle = \langle x_1^2 \rangle + \langle x_2^2 \rangle - 2\langle x_1 x_2 \rangle.$$

For distinguishable particles we have (only bras, kets and operators with the same particle index can come together!)

$$\begin{aligned} \langle x_1^2 \rangle &= {}_1\langle\psi_a| {}_2\langle\psi_b|\hat{X}_1^2|\psi_a\rangle_1|\psi_b\rangle_2 = {}_1\langle\psi_a|\hat{X}_1^2|\psi_a\rangle_1 {}_2\langle\psi_b|\psi_b\rangle_2 \\ &= {}_1\langle\psi_a|\hat{X}_1^2|\psi_a\rangle_1 = \langle x^2 \rangle_a, \end{aligned}$$

$$\begin{aligned} \langle x_2^2 \rangle &= {}_1\langle\psi_a| {}_2\langle\psi_b|\hat{X}_2^2|\psi_a\rangle_1|\psi_b\rangle_2 = {}_1\langle\psi_a|\psi_a\rangle_1 {}_2\langle\psi_b|\hat{X}_2^2|\psi_b\rangle_2 \\ &= {}_2\langle\psi_b|\hat{X}_2^2|\psi_b\rangle_2 = \langle x^2 \rangle_b, \end{aligned}$$

$$\begin{aligned} \langle x_1 x_2 \rangle &= {}_1\langle\psi_a| {}_2\langle\psi_b|\hat{X}_1\hat{X}_2|\psi_a\rangle_1|\psi_b\rangle_2 \\ &= {}_1\langle\psi_a|\hat{X}_1|\psi_a\rangle_1 {}_2\langle\psi_b|\hat{X}_2|\psi_b\rangle_2 = \langle x \rangle_a \langle x \rangle_b. \end{aligned}$$

For indistinguishable particles,

$$\begin{aligned} &\langle x_1^2 \rangle \\ &= \frac{1}{\sqrt{2}} \left({}_1\langle\psi_a| {}_2\langle\psi_b| \pm {}_1\langle\psi_b| {}_2\langle\psi_a| \right) \hat{X}_1^2 \frac{1}{\sqrt{2}} (|\psi_a\rangle_1 |\psi_b\rangle_2 \\ &\quad \pm |\psi_b\rangle_1 |\psi_a\rangle_2) \\ &= \frac{1}{2} \left({}_1\langle\psi_a| {}_2\langle\psi_b|\hat{X}_1^2|\psi_a\rangle_1|\psi_b\rangle_2 \pm {}_1\langle\psi_a| {}_2\langle\psi_b|\hat{X}_1^2|\psi_b\rangle_1|\psi_a\rangle_2 \right. \\ &\quad \left. \pm {}_1\langle\psi_b| {}_2\langle\psi_a|\hat{X}_1^2|\psi_a\rangle_1|\psi_b\rangle_2 + {}_1\langle\psi_b| {}_2\langle\psi_a|\hat{X}_1^2|\psi_b\rangle_1|\psi_a\rangle_2 \right) \\ &= \frac{1}{2} \left({}_1\langle\psi_a|\hat{X}_1^2|\psi_a\rangle_1 {}_2\langle\psi_b|\psi_b\rangle_2 \pm {}_1\langle\psi_a|\hat{X}_1^2|\psi_b\rangle_1 {}_2\langle\psi_b|\psi_a\rangle_2 \right. \\ &\quad \left. \pm {}_1\langle\psi_b|\hat{X}_1^2|\psi_a\rangle_1 {}_2\langle\psi_a|\psi_b\rangle_2 + {}_1\langle\psi_b|\hat{X}_1^2|\psi_b\rangle_1 {}_2\langle\psi_a|\psi_a\rangle_2 \right) \\ &= \frac{1}{2} (\langle x^2 \rangle_a + \langle x^2 \rangle_b). \end{aligned}$$

Similarly,

$$\langle x_2^2 \rangle = \frac{1}{2} (\langle x^2 \rangle_a + \langle x^2 \rangle_b).$$

Finally,

$$\begin{aligned} & \langle x_1 x_2 \rangle \\ &= \frac{1}{2} \left({}_1\langle \psi_a | {}_2\langle \psi_b | \hat{X}_1 \hat{X}_2 | \psi_a \rangle_1 | \psi_b \rangle_2 \pm {}_1\langle \psi_a | {}_2\langle \psi_b | \hat{X}_1 \hat{X}_2 | \psi_b \rangle_1 | \psi_a \rangle_2 \right. \\ & \quad \left. \pm {}_1\langle \psi_b | {}_2\langle \psi_a | \hat{X}_1 \hat{X}_2 | \psi_a \rangle_1 | \psi_b \rangle_2 + {}_1\langle \psi_b | {}_2\langle \psi_a | \hat{X}_1 \hat{X}_2 | \psi_b \rangle_1 | \psi_a \rangle_2 \right) \\ &= \frac{1}{2} \left({}_1\langle \psi_a | \hat{X}_1 | \psi_a \rangle_1 {}_2\langle \psi_b | \hat{X}_2 | \psi_b \rangle_2 \pm {}_1\langle \psi_a | \hat{X}_1 | \psi_b \rangle_1 {}_2\langle \psi_b | \hat{X}_2 | \psi_a \rangle_2 \right. \\ & \quad \left. \pm {}_1\langle \psi_b | \hat{X}_1 | \psi_a \rangle_1 {}_2\langle \psi_a | \hat{X}_2 | \psi_b \rangle_2 + {}_1\langle \psi_b | \hat{X}_1 | \psi_b \rangle_1 {}_2\langle \psi_a | \hat{X}_2 | \psi_a \rangle_2 \right) \\ &= \frac{1}{2} (\langle x \rangle_a \langle x \rangle_b \pm \langle x \rangle_{ab} \langle x \rangle_{ba} \pm \langle x \rangle_{ba} \langle x \rangle_{ab} + \langle x \rangle_b \langle x \rangle_a) \\ &= \langle x \rangle_a \langle x \rangle_b \pm \langle x \rangle_{ab} \langle x \rangle_{ba}. \end{aligned}$$

Altogether, we for distinguishable particles,

$$\langle x^2 \rangle_a + \langle x^2 \rangle_b - 2\langle x \rangle_a \langle x \rangle_b.$$

For indistinguishable ones, we have

$$\langle x^2 \rangle_a + \langle x^2 \rangle_b - 2\langle x \rangle_a \langle x \rangle_b \mp 2\langle x \rangle_{ab} \langle x \rangle_{ba}.$$

The result is smaller for bosons (take the minus sign) and larger for fermions (take the plus sign). This means that bosons like to be on average closer together and fermions prefer to be on average further away from each other. We call this an exchange force (although it is not really a force) or exchange energy. This purely geometrical consequence of symmetrization is a strictly quantum mechanical phenomenon with no classical counterpart.

As a final remark, notice that when discussing spin, we did not comment on any other aspect of the quantum state. The conclusions relating the symmetry of the quantum state of bosons/fermions are general and we can ask how to marry the quantum state $|\psi\rangle$ (the so called spatial part which gives the familiar wave function, $\psi(x) = \langle x|\psi\rangle$) with the spin state $|s\rangle$. The short answer which will be expanded on in your future quantum mechanics units is that the overall quantum state is a product of $|\psi\rangle$ and $|s\rangle$. For two electrons, it is this overall product wave function that has to be antisymmetric with respect to exchange of the two particles. This means that if the electrons are in the spin singlet state which is antisymmetric under exchange (check this!), then the spatial part must be symmetric under exchange. Conversely, if the electrons are in the spin triplet state which is symmetric under exchange, the spatial part must be antisymmetric. For example, we can construct a state $|\Psi\rangle$ with identical spatial parts if they are accompanied by antisymmetric spin singlet,

$$|\Psi\rangle = \frac{1}{\sqrt{2}} |\psi_a\rangle_1 |\psi_a\rangle_2 [|\uparrow_1\rangle |\downarrow_2\rangle - |\downarrow_1\rangle |\uparrow_2\rangle].$$

PH20013/60 Quantum Mechanics

Section IV: Hydrogen problem and atomic orbitals

In this section we will discuss quantum mechanics in three dimensions, with a particular focus on the hydrogen problem and atomic orbitals.

Quantum mechanics in 3D

Before jumping into more complicated things, let's ask a question: how would we solve a **one-particle** quantum mechanical problem in three dimensions?

Let's start with a 3D Schrodinger equation,

$$i\hbar \frac{\partial}{\partial t} |\psi(t)\rangle = \left(\frac{\hat{P}^2}{2m} + \hat{V} \right) |\psi(t)\rangle$$

In the position basis, this gives a somewhat familiar

$$i\hbar \frac{\partial}{\partial t} \psi(\vec{r}, t) = \left(\frac{p^2}{2m} + V(\vec{r}) \right) \psi(\vec{r}, t).$$

Using the form of the (now explicitly 3D) momentum operator and recalling that the time evolution of an energy eigenstate is of the form $\langle \varepsilon | \psi(t) \rangle = \langle \varepsilon | \psi(0) \rangle e^{-i\varepsilon t/\hbar}$, (equivalent to saying that the wave function is separable, $\psi(\vec{r}, t) = \psi(\vec{r})\phi(t)$ with $\phi(t) = e^{-i\varepsilon t/\hbar}$), we have

$$\left(-\frac{\hbar^2}{2m} \nabla^2 + V(\vec{r}) \right) \psi(\vec{r}) = \varepsilon \psi(\vec{r}),$$

which is nothing but an eigenproblem for the total energy operator \hat{H} written down in a particular basis.

Example: 3D “box”

The particle is confined to a 3D box with dimensions $L_x \times L_y \times L_z$. The potential energy is

- $V(\vec{r}) = 0$ if $-\frac{L_x}{2} \leq x \leq \frac{L_x}{2}$, $-\frac{L_y}{2} \leq y \leq \frac{L_y}{2}$, $-\frac{L_z}{2} \leq z \leq \frac{L_z}{2}$,
- $V(\vec{r}) = \infty$ elsewhere.

This problem is separable in Cartesian co-ordinates:

- $V(\vec{r}) = V_x(x) + V_y(y) + V_z(z)$,
- $\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$,
- $\psi(\vec{r}) = \phi_x(x)\phi_y(y)\phi_z(z)$.

We have,

$$\left(-\frac{\hbar^2}{2m} \frac{1}{\phi_x} \frac{d^2 \phi_x}{dx^2} + V_x(x) \right) + \left(-\frac{\hbar^2}{2m} \frac{1}{\phi_y} \frac{d^2 \phi_y}{dy^2} + V_y(y) \right) + \left(-\frac{\hbar^2}{2m} \frac{1}{\phi_z} \frac{d^2 \phi_z}{dz^2} + V_z(z) \right) = \varepsilon$$

Each of the brackets is a function of only one variable but the whole equation holds at every point in space.

Therefore, each bracket must be a constant (say, ε_x , ε_y and ε_z) and the sum of these constants must be ε .

For the 3D box, we get a 1D infinite potential well along each of the three dimensions, for example,

$$\phi_x \rightarrow \phi_{x,n_x} = \begin{cases} \sqrt{\frac{2}{L_x}} \sin \frac{n_x \pi x}{L_x}, & \text{for } n_x \text{ even,} \\ \sqrt{\frac{2}{L_x}} \cos \frac{n_x \pi x}{L_x}, & \text{for } n_x \text{ odd,} \end{cases}$$
$$\varepsilon_x \rightarrow \varepsilon_{x,n_x} = \frac{\hbar^2 \pi^2 n_x^2}{2m L_x^2}.$$

Overall,

$\psi(\vec{r}) \rightarrow \psi_{n_x, n_y, n_z}(x, y, z) = \phi_{x, n_x}(x) \phi_{y, n_y}(y) \phi_{z, n_z}(z)$, (zero outside the box) and $\varepsilon = \varepsilon_{x, n_x} + \varepsilon_{y, n_y} + \varepsilon_{z, n_z}$.

The numbers n_x , n_y and n_z that characterize the solutions and energies are the *quantum numbers*.

If the potential is only a function of distance from the origin, we can use spherical co-ordinates (r, φ, θ) . The ∇^2 operator is somewhat complicated and we have

$$-\frac{\hbar^2}{2m} \left[\frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial \psi}{\partial r} \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial \psi}{\partial \theta} \right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2 \psi}{\partial \varphi^2} \right] + V(r)\psi = \varepsilon\psi.$$

We can look for a wave function in the form $\psi(r, \varphi, \theta) = R(r)Y(\varphi, \theta)$ and this gives

$$\left[-\frac{\hbar^2}{2m} \frac{1}{R} \frac{d}{dr} \left(r^2 \frac{dR}{dr} \right) + r^2 V(r) - r^2 \varepsilon \right] - \frac{\hbar^2}{2m} \left[\frac{1}{Y} \frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial Y}{\partial \theta} \right) + \frac{1}{Y} \frac{1}{\sin^2 \theta} \frac{\partial^2 Y}{\partial \varphi^2} \right] = 0.$$

Again, each bracket has to be a constant as they depend on different variables. Moreover, recall that

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

We have already studied its eigenproblem without dealing with the differential equation!

Because we already know the eigenvalues of \hat{L}^2 , we can write

$$-\frac{\hbar^2}{2m} \frac{1}{r^2} \frac{1}{R} \frac{d}{dr} \left(r^2 \frac{dR}{dr} \right) + V(r) + \frac{\hbar^2 l(l+1)}{2mr^2} = \varepsilon$$

Details of the potential were not important so far – if V only depends on r , the angular solutions are always spherical harmonics!

Notice that our solutions depend on l but not on the magnetic quantum number m_l (because no magnetic field is present). Solutions with a given l value (which fixes the total angular momentum) but different m_l values (i.e., different L_z components) are degenerate.

In the case of a hydrogen atom (m is now the mass of an electron and we assume that the nucleus is heavy and does not move)), the potential is

$$V(r) = -\frac{e^2}{4\pi\epsilon_0 r}.$$

Changing variables,

$$R(r) = \frac{\chi(r)}{r}, x = \frac{2r}{\beta a_0} \text{ with } a_0 = \frac{4\pi\epsilon_0 \hbar^2}{me^2}, \varepsilon = -\frac{\hbar^2}{2ma_0^2\beta^2},$$

with β a constant that controls both the radial scale of solutions (through the relation between x and r) and energy ε ,

$$\frac{d^2\chi}{dr^2} + \left(\frac{\beta}{x} - \frac{1}{4} - \frac{l(l+1)}{x^2} \right) \chi = 0.$$

Note that we now require that $\chi(r=0) = 0$ (or R will blow up to infinity at the origin).

Exercise: Show that

$$R(r) = Ae^{-r/a_0}$$

is a solution of the radial part of hydrogen atom equation for $l = 0$ and some β . Determine the energy of this solution and the normalization constant, A .

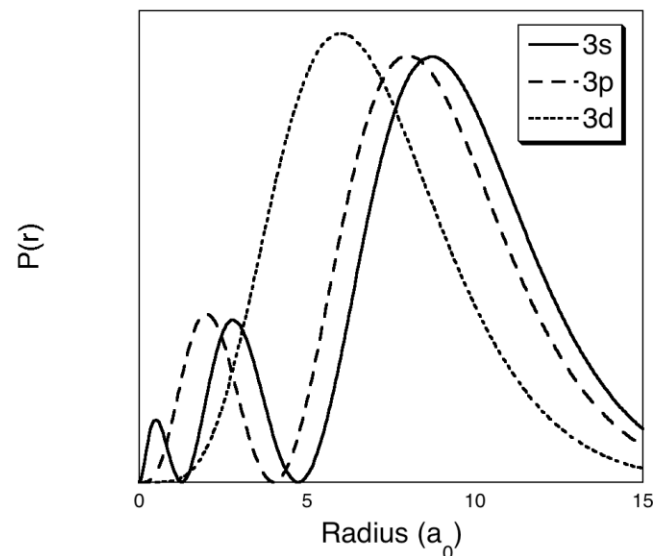
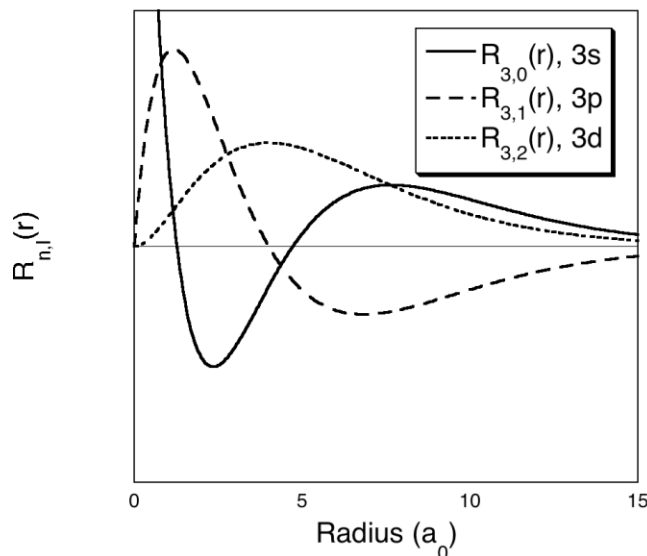
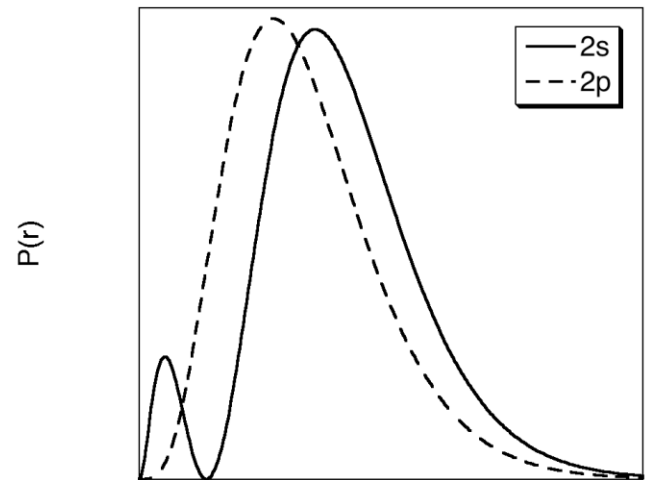
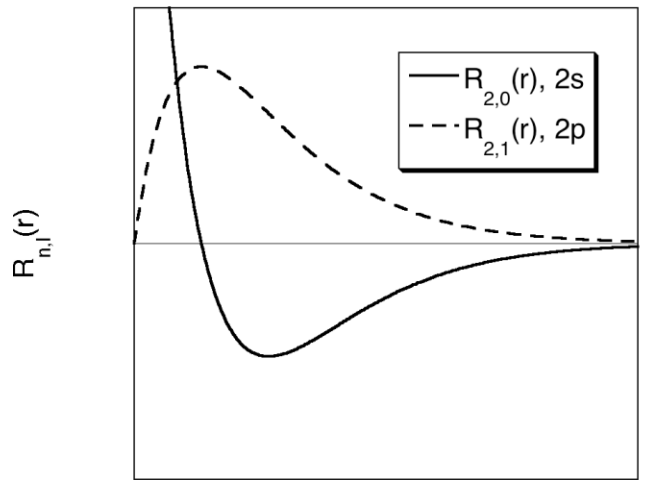
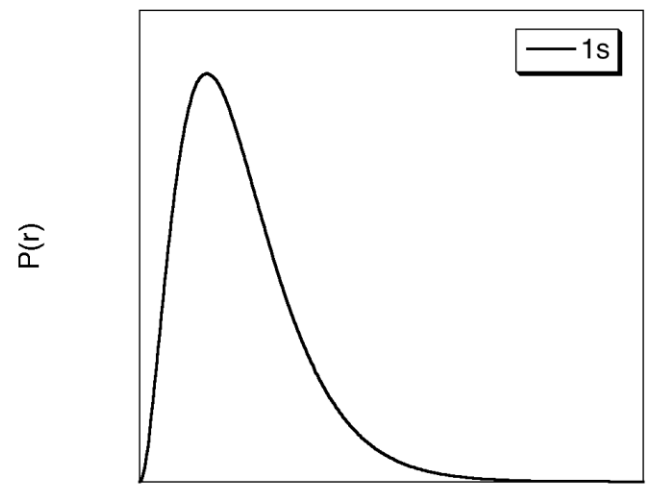
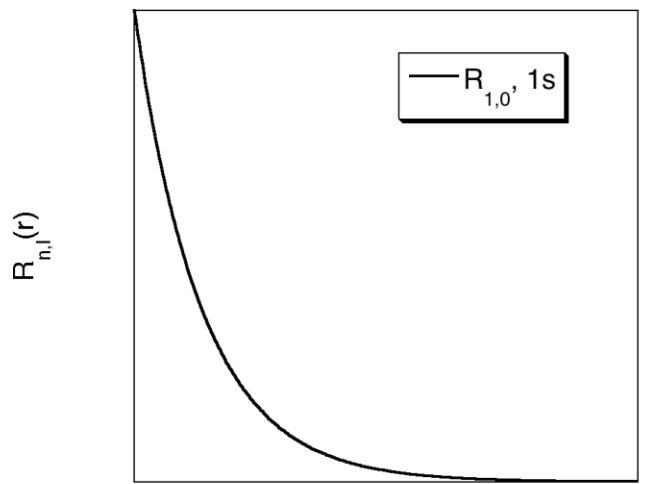
The radial equation can be solved exactly for any value of the quantum number l . It turns out that solutions that stay finite between $r = 0$ and $r = \infty$ can be labelled by an integer n where $n > l$. Radial states can therefore be labelled by $R_{n,l}(r)$. Overall, we have

$$\langle \vec{r} | nlm_l \rangle = \psi_{nlm_l}(r, \varphi, \theta) = R_{n,l}(r)Y_\ell^{m_l}(\varphi, \theta),$$

$$\varepsilon_n = -\frac{\hbar^2}{2ma_0^2n^2} = -\frac{m}{2\hbar^2} \left(\frac{e^2}{4\pi\epsilon_0} \right)^2 \frac{1}{n^2} = -\frac{\varepsilon_1}{n^2}.$$

By convention, $l = 0,1,2,3,4, \dots$ is often denoted with letters s, p, d, f, g, \dots , respectively.

n	l	m_l	Name	$\psi_{nlm_l}(r, \varphi, \theta)$
1	0	0	1s $ 1,0,0\rangle$	$\frac{1}{\sqrt{\pi}a_0^{3/2}} e^{-r/a_0}$
2	0	0	2s $ 2,0,0\rangle$	$\frac{1}{4\sqrt{2\pi}a_0^{3/2}} \left[2 - \frac{r}{a_0} \right] e^{-r/(2a_0)}$
2	1	0	2p $ 2,1,0\rangle$	$\frac{1}{4\sqrt{2\pi}a_0^{3/2}} \frac{r}{a_0} e^{-r/(2a_0)} \cos \theta$
2	1	± 1	2p $ 2,1,\pm 1\rangle$	$\frac{1}{8\sqrt{\pi}a_0^{3/2}} \frac{r}{a_0} e^{-r/(2a_0)} \sin \theta e^{\pm i\varphi}$



Radial wave function, $R(r)$, and radial probability distributions, $P(r) = |R(r)|^2 4\pi r^2$, for hydrogen orbitals. Note the energy for 1s is -13.6 eV, for 2s and 2p is -3.4 eV, and 3s, 3p and 3d is -1.5 eV. The energy operator has no dependence on the angular momentum.

- Wave functions have $(n - l - 1)$ radial zeros.
- $P(r)$ plots show that, for a given l , an electron is found further from the nucleus as n increases.
- $P(r)$ plots show that, for a given n , an electron is more likely to be found near the nucleus for small l (compare for example the $2s$ and $2p$ orbitals).
- The energies of orbitals are negative – they are bound states.
- The ground state, the lowest energy state, has $n = 1$.
- The states with $n = 2, 3, 4, \dots$ are excited states.
- Zero energy corresponds to $n = \infty$.
- Higher energies (> 0) are unbound and form a continuum of allowed energies: no quantization.
- An energy of $+13.6$ eV is required to excite an electron from the ground $1s$ state to infinity (escape!). This is called the ionization energy.

Finally, while the orbital wave functions are in general complex, it is possible to build straightforward linear combinations that are purely real and hence easier to visualize. For example,

$$|2p_x\rangle = \frac{1}{\sqrt{2}} (|2,1,1\rangle + |2,1,-1\rangle) \rightarrow \langle x|2p_x\rangle \equiv \psi_{2p_x}(r, \varphi, \theta)$$

$$= \frac{1}{4\sqrt{\pi}} \left(\frac{1}{a_0}\right)^{\frac{3}{2}} \frac{r}{a_0} e^{-\frac{r}{2a_0}} \sin \theta \cos \varphi ,$$

$$|2p_y\rangle = \frac{1}{\sqrt{2}i} (|2,1,1\rangle - |2,1,-1\rangle) \rightarrow \langle x|2p_y\rangle \equiv \psi_{2p_y}(r, \varphi, \theta)$$

$$= \frac{1}{4\sqrt{\pi}} \left(\frac{1}{a_0}\right)^{\frac{3}{2}} \frac{r}{a_0} e^{-\frac{r}{2a_0}} \sin \theta \sin \varphi .$$

While these have the same shape as $\psi_{210}(r, \varphi, \theta) \equiv \psi_{2p_z}(r, \varphi, \theta)$, they are not the eigenstates of \hat{L}_z (because they mix states with different m_l).

Exercise: Consider a single electron in a hydrogen atom. For the given value n of the principal quantum number,

- What are the possible values of l ?
- How many different quantum states are available with the same n (spin excluded)?
- How many different quantum states are available with the same n (spin included)?

Electronic states of atoms with many electrons

Notice that little would change in the hydrogen equations if we decided to describe potential due to a nucleus with more protons than hydrogen. However, adding more electrons introduces repulsion between them. Even for only two electrons, no exact solution is known. We can make a simplifying assumption that electrons feel the average effect of repulsion from other electrons, leading to an effective potential for each electron that is still a function of the radial distance, r , only. Because of the spherical symmetry, spherical harmonics are still useful as eigenfunctions of angular momentum and quantum numbers l and m_l remain good quantum numbers. Treating electrons as independent particles in the field of “average” charge leads again to the quantization of the radial distance and appearance of the principal quantum number n . Repulsion between electrons reduces the nuclear charge of protons felt by an electron – we introduce an effective nuclear charge $Z_{\text{eff}}(n, l)$ which depends on the two quantum numbers n and l that define the shape of an orbital. Because of this electronic screening, the energies of orbitals are renormalized,

$$\varepsilon_n = -\frac{Z_{\text{eff}}^2 \hbar^2}{2ma_0^2} \frac{1}{n^2} = \frac{\varepsilon_1 Z_{\text{eff}}^2}{n^2}.$$

Side comment: recalling previous section, we can put two electrons into the same state $|n, l, m_l\rangle$, if the spin part is a spin singlet,

$$|\Psi\rangle = \frac{1}{\sqrt{2}} |n, l, m_l\rangle_1 |n, l, m_l\rangle_2 [|\uparrow_1\rangle|\downarrow_2\rangle - |\downarrow_1\rangle|\uparrow_2\rangle].$$

This means that two electrons paired in the same $|n, l, m_l\rangle$ orbital have spin $s = 0$!

Hydrogenic orbitals (Z_{eff} is the effective nuclear charge)

n	l	m_l	name	$\psi_{nlm_l}(r, \varphi, \theta)$
1	0	0	1s $ 1,0,0\rangle$	$\frac{1}{\sqrt{\pi}} \left(\frac{Z_{\text{eff}}}{a_0} \right)^{\frac{3}{2}} e^{-Z_{\text{eff}}r/a_0}$
2	0	0	2s $ 2,0,0\rangle$	$\frac{1}{4\sqrt{2\pi}} \left(\frac{Z_{\text{eff}}}{a_0} \right)^{\frac{3}{2}} \left[2 - \frac{Z_{\text{eff}}r}{a_0} \right] e^{-Z_{\text{eff}}r/(2a_0)}$
2	1	0	2p $ 2,1,0\rangle$	$\frac{1}{4\sqrt{2\pi}} \left(\frac{Z_{\text{eff}}}{a_0} \right)^{\frac{3}{2}} \frac{Z_{\text{eff}}r}{a_0} e^{-Z_{\text{eff}}r/(2a_0)} \cos \theta$
2	1	± 1	2p $ 2,1, \pm 1\rangle$	$\frac{1}{8\sqrt{\pi}} \left(\frac{Z_{\text{eff}}}{a_0} \right)^{\frac{3}{2}} \frac{Z_{\text{eff}}r}{a_0} e^{-Z_{\text{eff}}r/(2a_0)} \sin \theta e^{\pm i\varphi}$

Effective Nuclear Charges

	H																	He
Z	1																	2
1s	1.000																	1.688
	Li	Be											B	C	N	O	F	Ne
Z	3	4											5	6	7	8	9	10
1s	2.691	3.685											4.680	5.673	6.665	7.658	8.650	9.642
2s	1.279	1.912											2.576	3.217	3.847	4.492	5.128	5.758
2p													2.421	3.136	3.834	4.453	5.100	5.758
	Na	Mg											Al	Si	P	S	Cl	Ar
Z	11	12											13	14	15	16	17	18
1s	10.626	11.609											12.591	13.575	14.558	15.541	16.524	17.508
2s	6.571	7.392											8.214	9.020	9.825	10.629	11.430	12.230
2p	6.802	7.826											8.963	9.945	10.961	11.977	12.993	14.008
3s	2.507	3.308											4.117	4.903	5.642	6.367	7.068	7.757
3p													4.066	4.285	4.886	5.482	6.116	6.764
	K	Ca	Sc	Ti	V	Cr	Mn	Fe	Co	Ni	Cu	Zn	Ga	Ge	As	Se	Br	Kr
Z	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36
1s	18.490	19.473	20.457	21.441	22.426	23.414	24.396	25.381	26.367	27.353	28.339	29.325	30.309	31.294	32.278	33.262	34.247	35.232
2s	13.006	13.776	14.574	15.377	16.181	16.984	17.794	18.599	19.405	20.213	21.020	21.828	22.599	23.365	24.127	24.888	25.643	26.398
2p	15.027	16.041	17.055	18.065	19.073	20.075	21.084	22.089	23.092	24.095	25.097	26.098	27.091	28.082	29.074	30.065	31.056	32.047
3s	8.680	9.602	10.340	11.033	11.709	12.368	13.018	13.676	14.322	14.961	15.594	16.219	16.996	17.790	18.596	19.403	20.219	21.033
3p	7.726	8.658	9.406	10.104	10.785	11.466	12.109	12.778	13.435	14.085	14.731	15.369	16.204	17.014	17.850	18.705	19.571	20.434
4s	3.495	4.398	4.632	4.817	4.981	5.133	5.283	5.434	5.576	5.711	5.842	5.965	7.067	8.044	8.944	9.758	10.553	11.316
3d			7.120	8.141	8.983	9.757	10.528	11.180	11.855	12.530	13.201	13.878	15.093	16.251	17.378	18.477	19.559	20.626
4p													6.222	6.780	7.449	8.287	9.028	9.338
	Rb	Sr	Y	Zr	Nb	Mo	Tc	Ru	Rh	Pd	Ag	Cd	In	Sn	Sb	Te	I	Xe
Z	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54
1s	36.208	37.191	38.176	39.159	40.142	41.126	42.109	43.092	44.076	45.059	46.042	47.026	48.010	48.992	49.974	50.957	51.939	52.922
2s	27.157	27.902	28.622	29.374	30.125	30.877	31.628	32.380	33.155	33.883	34.634	35.386	36.124	36.859	37.595	38.331	39.067	39.803
2p	33.039	34.030	35.003	35.993	36.982	37.972	38.941	39.951	40.940	41.930	42.919	43.909	44.898	45.885	46.873	47.860	48.847	49.835
3s	21.843	22.664	23.552	24.362	25.172	25.982	26.792	27.601	28.439	29.221	30.031	30.841	31.631	32.420	33.209	33.998	34.787	35.576
3p	21.303	22.168	23.093	23.846	24.616	25.474	26.384	27.221	28.154	29.020	29.809	30.692	31.521	32.353	33.184	34.009	34.841	35.668
4s	12.388	13.444	14.264	14.902	15.283	16.096	17.198	17.656	18.582	18.986	19.865	20.869	21.761	22.658	23.544	24.408	25.297	26.173
3d	21.679	22.726	25.397	25.567	26.247	27.228	28.353	29.359	30.405	31.451	32.540	33.607	34.678	35.742	36.800	37.839	38.901	39.947
4p	10.881	11.932	12.746	13.460	14.084	14.977	15.811	16.435	17.140	17.723	18.562	19.411	20.369	21.265	22.181	23.122	24.030	24.957
5s	4.985	6.071	6.256	6.446	5.921	6.106	7.227	6.485	6.640 (empty)	6.756	8.192	9.512	10.629	11.617	12.538	13.404	14.218	
4d			15.958	13.072	11.238	11.392	12.882	12.813	13.442	13.618	14.763	15.877	16.942	17.970	18.974	19.960	20.934	21.893
5p													8.470	9.102	9.995	10.809	11.612	12.425

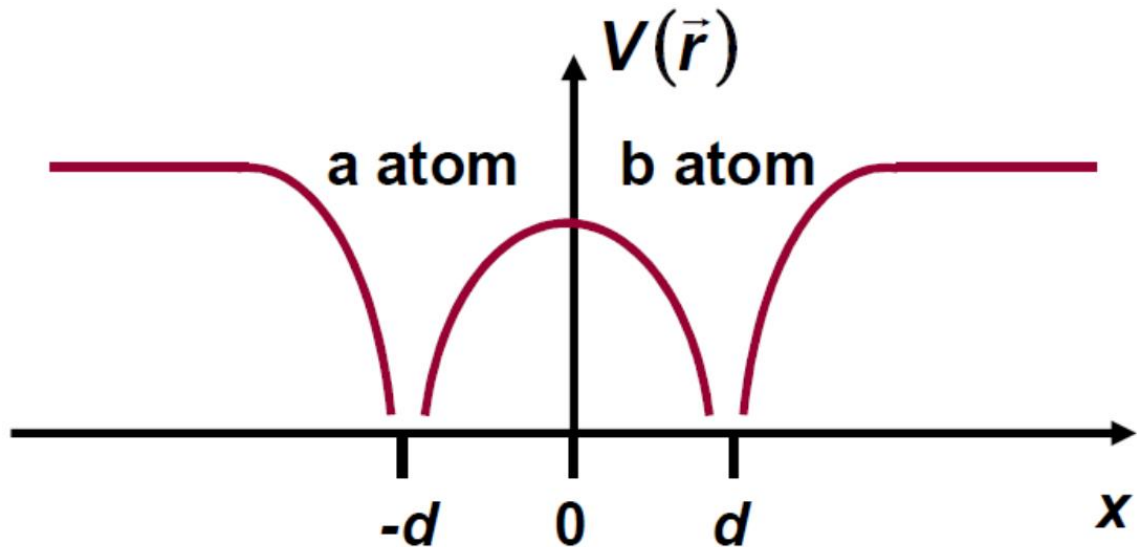
[illegible]

Atomic Number	Name	Symbol	Electron Configuration
1	Hydrogen	H	$1s^1$
2	Helium	He	$1s^2 = [\text{He}]$
3	Lithium	Li	$[\text{He}]2s^1$
4	Beryllium	Be	$[\text{He}]2s^2$
5	Boron	B	$[\text{He}]2s^22p^1$
6	Carbon	C	$[\text{He}]2s^22p^2$
7	Nitrogen	N	$[\text{He}]2s^22p^3$
8	Oxygen	O	$[\text{He}]2s^22p^4$
9	Fluorine	F	$[\text{He}]2s^22p^5$
10	Neon	Ne	$[\text{He}]2s^22p^6 = [\text{Ne}]$
11	Sodium	Na	$[\text{Ne}]3s^1$
12	Magnesium	Mg	$[\text{Ne}]3s^2$
13	Aluminium	Al	$[\text{Ne}]3s^23p^1$
14	Silicon	Si	$[\text{Ne}]3s^23p^2$
15	Phosphorous	P	$[\text{Ne}]3s^23p^3$
16	Sulphur	S	$[\text{Ne}]3s^23p^4$
17	Chlorine	Cl	$[\text{Ne}]3s^23p^5$
18	Argon	Ar	$[\text{Ne}]3s^23p^6 = [\text{Ar}]$
19	Potassium	K	$[\text{Ar}]4s^1$
20	Calcium	Ca	$[\text{Ar}]4s^2$
21	Scandium	Sc	$[\text{Ar}]3d^14s^2$
22	Titanium	Ti	$[\text{Ar}]3d^24s^2$
23	Vanadium	V	$[\text{Ar}]3d^34s^2$
24	Chromium	Cr	$[\text{Ar}]3d^54s^1$: anomaly!
25	Manganese	Mn	$[\text{Ar}]3d^54s^2$
26	Iron	Fe	$[\text{Ar}]3d^64s^2$
27	Cobalt	Co	$[\text{Ar}]3d^74s^2$
28	Nickel	Ni	$[\text{Ar}]3d^84s^2$
29	Copper	Cu	$[\text{Ar}]3d^{10}4s^1$: anomaly!
30	Zinc	Zn	$[\text{Ar}]3d^{10}4s^2$
31	Gallium	Ga	$[\text{Ar}]3d^{10}4s^24p^1$

Electron configuration of the first 4 rows of elements.

From hydrogen atom to hydrogen molecule

Let's build the simplest model of atomic bonding. Consider two hydrogen atoms sitting at a close(ish) distance $2d$ from each other. How could we describe our problem?



1. We assume (again) that the nuclei are too heavy to move.
2. We consider only one electron (to avoid questions about Coulomb repulsion between electrons).

In the position basis, the Hamiltonian for the electron is then the Schrodinger equation,

$$\hat{H}(\vec{r}) = -\frac{\hbar^2}{2m}\nabla^2 + V(\vec{r} - d\vec{i}) + V(\vec{r} + d\vec{i}).$$

Recall that we know the electronic ground states for each atom separately, the $|1s\rangle = |\phi_{1s}\rangle \equiv |nlm_l\rangle = |100\rangle$ state, in which we expect the electron to be if the atoms are far away from each other.

The simplest guess we can make is that the quantum state of the electron in the molecule, $|\psi\rangle$, is made up of a linear combination of the orbitals of isolated atoms. This means that,

$$|\psi\rangle = c_1|\phi_{1s}(\vec{r} - d\vec{t})\rangle + c_2|\phi_{1s}(\vec{r} + d\vec{t})\rangle \\ \equiv c_1|\phi_{1s}(a)\rangle + c_2|\phi_{1s}(b)\rangle,$$

where a/b is a shorthand to remind us which atom we mean. An important point resulting from above is that we have decided that only the $|\phi_{1s}\rangle$ states are important – that is, they provide a sufficient basis for this problem. Starting from the general,

$$\hat{H}|\psi\rangle = \varepsilon|\psi\rangle,$$

we close it first with $\langle\phi_{1s}(a)|$, so that on the left-hand side,

$$\langle\phi_{1s}(a)|\hat{H}(c_1|\phi_{1s}(a)\rangle + c_2|\phi_{1s}(b)\rangle) \\ = c_1\langle\phi_{1s}(a)|\hat{H}|\phi_{1s}(a)\rangle + c_2\langle\phi_{1s}(a)|\hat{H}|\phi_{1s}(b)\rangle.$$

The $|\phi_{1s}\rangle$ are eigenstates of the single hydrogen atom problem but not of our Hamiltonian \hat{H} here, because this accounts for both atoms together. Hence, there is nothing else we can do. On the right-hand side,

$$\langle\phi_{1s}(a)|\varepsilon(c_1|\phi_{1s}(a)\rangle + c_2|\phi_{1s}(b)\rangle) \\ = \varepsilon c_1\langle\phi_{1s}(a)|\phi_{1s}(a)\rangle + \varepsilon c_2\langle\phi_{1s}(a)|\phi_{1s}(b)\rangle.$$

We have, $\langle\phi_{1s}(a)|\phi_{1s}(a)\rangle = 1$. In principle, $\langle\phi_{1s}(a)|\phi_{1s}(b)\rangle \neq 0$, but, to a good approximation (and for simplicity here), $\langle\phi_{1s}(a)|\phi_{1s}(b)\rangle \approx 0$. We have,

$$c_1\langle\phi_{1s}(a)|\hat{H}|\phi_{1s}(b)\rangle + c_2\langle\phi_{1s}(a)|\hat{H}|\phi_{1s}(b)\rangle = \varepsilon c_1.$$

We now go back to our eigenproblem and close with $\langle \phi_{1s}(b) |$. On the left-hand side, we have

$$\begin{aligned} & \langle \phi_{1s}(b) | \hat{H} (c_1 |\phi_{1s}(a)\rangle + c_2 |\phi_{1s}(b)\rangle) \\ &= c_1 \langle \phi_{1s}(b) | \hat{H} | \phi_{1s}(a) \rangle + c_2 \langle \phi_{1s}(b) | \hat{H} | \phi_{1s}(b) \rangle. \end{aligned}$$

On the right-hand side,

$$\begin{aligned} & \langle \phi_{1s}(b) | \varepsilon (c_1 |\phi_{1s}(a)\rangle + c_2 |\phi_{1s}(b)\rangle) \\ &= \varepsilon c_1 \langle \phi_{1s}(b) | \phi_{1s}(a) \rangle + \varepsilon c_2 \langle \phi_{1s}(b) | \phi_{1s}(b) \rangle \approx \varepsilon c_2. \end{aligned}$$

Putting everything together, the two equations give us the following matrix expression,

$$\begin{bmatrix} \langle \phi_{1s}(a) | \hat{H} | \phi_{1s}(a) \rangle & \langle \phi_{1s}(a) | \hat{H} | \phi_{1s}(b) \rangle \\ \langle \phi_{1s}(b) | \hat{H} | \phi_{1s}(a) \rangle & \langle \phi_{1s}(b) | \hat{H} | \phi_{1s}(b) \rangle \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \varepsilon \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}.$$

We have cast our problem as an eigenproblem in the orthonormal (following our approximations) basis $\{|\phi_{1s}(a)\rangle, |\phi_{1s}(b)\rangle\}$. The coefficients c_1 and c_2 of the eigenstate will tell us how the states of the isolated atoms mix together to produce the state for the molecule as a whole.

We will now make further approximations:

1. We say that $\langle \phi_{1s}(a) | \hat{H} | \phi_{1s}(a) \rangle = \langle \phi_{1s}(b) | \hat{H} | \phi_{1s}(b) \rangle \approx \varepsilon_{1s}$ (when the electron is in a pure 1s state around one of the atoms, its energy is essentially what it would be if this atom was isolated).
2. We say that $\langle \phi_{1s}(a) | \hat{H} | \phi_{1s}(b) \rangle = -V_{ss\sigma}$ - a real ($V_{ss\sigma} > 0$) parameter that expresses “coupling” between the two 1s states on different atoms (the subscript is a convention that keeps information about the angular momentum symmetry of the coupling).

With these points, the Hamiltonian takes the form:

$$\begin{bmatrix} \varepsilon_{1s} & -V_{ss\sigma} \\ -V_{ss\sigma} & \varepsilon_{1s} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \varepsilon \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}.$$

The eigenvalues are:

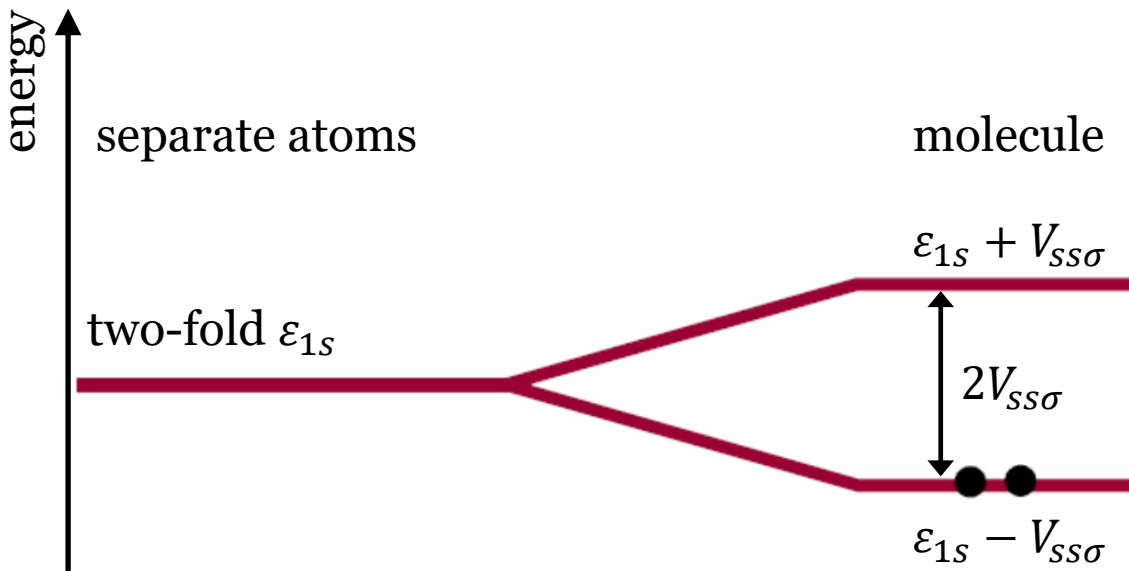
$$\varepsilon_B = \varepsilon_{1s} - V_{ss\sigma}, \quad \varepsilon_A = \varepsilon_{1s} + V_{ss\sigma}$$

with eigenstates

$$|\varepsilon_B\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \rightarrow |\psi\rangle = \frac{1}{\sqrt{2}} (|\phi_{1s}(a)\rangle + |\phi_{1s}(b)\rangle),$$

$$|\varepsilon_A\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \rightarrow |\psi\rangle = \frac{1}{\sqrt{2}} (|\phi_{1s}(a)\rangle - |\phi_{1s}(b)\rangle).$$

The $|\varepsilon_B\rangle$ is the bonding state and $|\varepsilon_A\rangle$ the antibonding state. The electron will occupy the bonding state and lower its energy – formation of the molecule is energetically beneficial (as we might have guessed).



(*) Exercise: How would the Hamiltonian matrix, eigenvalues and eigenvectors change if we relaxed one of the approximations and put $\langle \phi_{1s}(a) | \phi_{1s}(b) \rangle = s \neq 0$, where s is a complex number?

Back to atoms – addition of angular momenta

Recall our discussion of two electron spins. We effectively argued that two points of view are possible:

- Two particles with $s = \frac{1}{2}$ spin each. The basis consists of products of the single-particle bases, $\{|\uparrow\uparrow\rangle, |\uparrow\downarrow\rangle, |\downarrow\uparrow\rangle, |\downarrow\downarrow\rangle\}$ (basis states are eigenstates of single-particle z -component of spin operators).
- A composite system of two particles with the total angular momentum of $s = 1$. The basis consists of states that are eigenstates of the total spin and z -component of spin operators for the composite system, $\left\{\frac{1}{\sqrt{2}}[|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle], |\uparrow\uparrow\rangle, \frac{1}{\sqrt{2}}[|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle], |\downarrow\downarrow\rangle\right\}$.

Note that the two bases are not equivalent. This is because addition of angular momenta (including spins) presents some subtleties.

The angular momentum operator of a composite system is a sum of the operators of its parts. In this case

$$\hat{S} = \hat{S}^{(1)} + \hat{S}^{(2)},$$

where the superscripts refer to the first or second particle and the spin operators are in 3D – this means they have three components, x , y and z . Notice that for the z -component, addition like this is straightforward,

$$\hat{S}_z = \hat{S}_z^{(1)} + \hat{S}_z^{(2)},$$

which means that for two electrons we have the options $m_s = -1, 0, 1$ which exhaust all the combinations of two $s_z^{(i)} = \pm\frac{1}{2}$. For the total spin, the situation is more complicated because we are trying to add two “quantum vectors”.

We will go back to our short discussion in the previous section and check the two-electron case again. Focusing on the eigenstates of \hat{S}_z for each spin, we can list four different cases,

- 1) $|\uparrow_1\rangle|\uparrow_2\rangle \equiv |\uparrow\uparrow\rangle,$
- 2) $|\uparrow_1\rangle|\downarrow_2\rangle \equiv |\uparrow\downarrow\rangle,$
- 3) $|\downarrow_1\rangle|\uparrow_2\rangle \equiv |\downarrow\uparrow\rangle,$
- 4) $|\downarrow_1\rangle|\downarrow_2\rangle \equiv |\downarrow\downarrow\rangle,$

where the indices number the two electrons and correspond to the ordering of the arrows in the “two-electron” kets. Again, intuition suggests (can you check this mathematically?) that we can associate two of the four states above with (total) $m_s = -1$ and $m_s = +1$ – the maximum possible z-component of the combined spin,

$|\uparrow\uparrow\rangle$ - largest z-component pointing up, $m_s = +1$,

$|\downarrow\downarrow\rangle$ - largest z-component pointing down, $m_s = -1$.

We know that acting with an operator \hat{S}_+/\hat{S}_- on an eigenstate of \hat{S}_z will increase/decrease the value of while leaving the total length of angular momentum unchanged. Here, we mean operators for the two-electron system which are a sum of the single-particle operators,

$$\hat{S}_+ = \hat{S}_+^{(1)} + \hat{S}_+^{(2)}, \quad \hat{S}_- = \hat{S}_-^{(1)} + \hat{S}_-^{(2)}.$$

In the previous section, we acted with \hat{S}_+ on $|\downarrow\downarrow\rangle$,

$$\begin{aligned} \hat{S}_+|\downarrow\downarrow\rangle &= (\hat{S}_+^{(1)} + \hat{S}_+^{(2)})|\downarrow\downarrow\rangle = \hat{S}_+^{(1)}|\downarrow_1\rangle|\downarrow_2\rangle + \hat{S}_+^{(2)}|\downarrow_1\rangle|\downarrow_2\rangle \\ &= \hbar|\uparrow_1\rangle|\downarrow_2\rangle + \hbar|\downarrow_1\rangle|\uparrow_2\rangle = \hbar[|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle]. \end{aligned}$$

Similarly, we can see that

$$\begin{aligned} \hat{S}_-|\uparrow\uparrow\rangle &= (\hat{S}_-^{(1)} + \hat{S}_-^{(2)})|\uparrow\uparrow\rangle = \hat{S}_-^{(1)}|\uparrow_1\rangle|\uparrow_2\rangle + \hat{S}_-^{(2)}|\uparrow_1\rangle|\uparrow_2\rangle \\ &= \hbar|\downarrow_1\rangle|\uparrow_2\rangle + \hbar|\uparrow_1\rangle|\downarrow_2\rangle = \hbar[|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle]. \end{aligned}$$

Clearly, the three states, (i) $|\downarrow\downarrow\rangle$, (ii) $\frac{1}{\sqrt{2}}[|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle]$ (where I fixed the normalization constant) and (iii) $|\uparrow\uparrow\rangle$ create a ladder. Notice that, because the single-particle spins cannot be further increased/decreased, the ladder does not contain any further states. We can also confirm that, as expected,

$$\begin{aligned}
 & \hat{S}_z \frac{1}{\sqrt{2}} [|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle] \\
 &= \hat{S}_z^{(1)} \frac{1}{\sqrt{2}} [|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle] + \hat{S}_z^{(2)} \frac{1}{\sqrt{2}} [|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle] \\
 &= \frac{1}{\sqrt{2}} \left[\left(+\frac{\hbar}{2} \right) |\uparrow\downarrow\rangle + \left(-\frac{\hbar}{2} \right) |\downarrow\uparrow\rangle \right] + \frac{1}{\sqrt{2}} \left[\left(-\frac{\hbar}{2} \right) |\uparrow\downarrow\rangle + \left(+\frac{\hbar}{2} \right) |\downarrow\uparrow\rangle \right] \\
 &= 0 \times \frac{1}{\sqrt{2}} [|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle].
 \end{aligned}$$

This is all well and good, but at the beginning we have listed four orthogonal single-particle states – our vector space must be 4-dimensional and we are missing something.

Notice that $\frac{1}{\sqrt{2}}[|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle]$ is a symmetric combination of $|\uparrow\downarrow\rangle$ and $|\downarrow\uparrow\rangle$. In previous sections, we have already seen that an antisymmetric combination is orthogonal to the symmetric one. Hence, our fourth state is,

$$\frac{1}{\sqrt{2}} [|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle].$$

Exercise: Using the properties of the single-particle states, confirm that the singlet/triplet states are orthogonal and normalized.

Notice that

$$\begin{aligned}
& \hat{S}_z \frac{1}{\sqrt{2}} [|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle] \\
&= \hat{S}_z^{(1)} \frac{1}{\sqrt{2}} [|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle] + \hat{S}_z^{(2)} \frac{1}{\sqrt{2}} [|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle] \\
&= \frac{1}{\sqrt{2}} \left[\left(+\frac{\hbar}{2} \right) |\uparrow\downarrow\rangle - \left(-\frac{\hbar}{2} \right) |\downarrow\uparrow\rangle \right] + \frac{1}{\sqrt{2}} \left[\left(-\frac{\hbar}{2} \right) |\uparrow\downarrow\rangle - \left(+\frac{\hbar}{2} \right) |\downarrow\uparrow\rangle \right] \\
&= 0 \times \frac{1}{\sqrt{2}} [|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle].
\end{aligned}$$

All four of our states are the eigenstates of total \hat{S}_z . I claim that the three states making up their own ladder correspond to the total spin of 1 and the fourth state corresponds to the total spin of 0. If this is the case, we expect the corresponding total \hat{S}^2 eigenvalues of $2\hbar^2$ and $0\hbar^2$ (remember, for angular momentum l , eigenvalues of \hat{L}^2 are $l(l+1)\hbar^2$). To perform the calculations using only the single-particle operators, notice that,

$$\begin{aligned}
\hat{S}^2 &= (\hat{S}^{(1)} + \hat{S}^{(2)}) \cdot (\hat{S}^{(1)} + \hat{S}^{(2)}) \\
&= (\hat{S}^{(1)})^2 + (\hat{S}^{(2)})^2 + 2\hat{S}^{(1)} \cdot \hat{S}^{(2)}
\end{aligned}$$

(single-particle operators $\hat{S}^{(1)}$ and $\hat{S}^{(2)}$ act only on the part of the state with “their” particle and so commute with each other). The squared terms are straightforward, and we have

$$\begin{aligned}
& \left((\hat{S}^{(1)})^2 + (\hat{S}^{(2)})^2 \right) |\downarrow\downarrow\rangle = (\hat{S}^{(1)})^2 |\downarrow\downarrow\rangle + (\hat{S}^{(2)})^2 |\downarrow\downarrow\rangle \\
&= \frac{1}{2} \left(\frac{1}{2} + 1 \right) \hbar^2 |\downarrow\downarrow\rangle + \frac{1}{2} \left(\frac{1}{2} + 1 \right) \hbar^2 |\downarrow\downarrow\rangle = \frac{3}{2} \hbar^2 |\downarrow\downarrow\rangle.
\end{aligned}$$

The same result holds for $|\uparrow\uparrow\rangle$.

For the other two states,

$$\begin{aligned}
& \left((\hat{S}^{(1)})^2 + (\hat{S}^{(2)})^2 \right) \frac{1}{\sqrt{2}} [|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle] \\
&= \frac{1}{\sqrt{2}} (\hat{S}^{(1)})^2 |\uparrow\downarrow\rangle + \frac{1}{\sqrt{2}} (\hat{S}^{(1)})^2 |\downarrow\uparrow\rangle + \frac{1}{\sqrt{2}} (\hat{S}^{(2)})^2 |\uparrow\downarrow\rangle \\
&+ \frac{1}{\sqrt{2}} (\hat{S}^{(2)})^2 |\downarrow\uparrow\rangle = 2 \frac{3}{4} \hbar^2 \frac{1}{\sqrt{2}} [|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle] \\
&= \frac{3}{2} \hbar^2 \frac{1}{\sqrt{2}} [|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle],
\end{aligned}$$

and,

$$\left((\hat{S}^{(1)})^2 + (\hat{S}^{(2)})^2 \right) \frac{1}{\sqrt{2}} [|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle] = \frac{3}{2} \hbar^2 \frac{1}{\sqrt{2}} [|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle].$$

To handle the mixed term, $2\hat{S}^{(1)} \cdot \hat{S}^{(2)}$, we can use our knowledge of the single-electron spin operators in terms of the Pauli matrices. We have

$$\begin{aligned}
\hat{S}_x |\uparrow\rangle &= \frac{\hbar}{2} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \frac{\hbar}{2} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \frac{\hbar}{2} |\downarrow\rangle, \\
\hat{S}_x |\downarrow\rangle &= \frac{\hbar}{2} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \frac{\hbar}{2} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \frac{\hbar}{2} |\uparrow\rangle, \\
\hat{S}_y |\uparrow\rangle &= \frac{\hbar}{2} \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = i \frac{\hbar}{2} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = i \frac{\hbar}{2} |\downarrow\rangle, \\
\hat{S}_y |\downarrow\rangle &= \frac{\hbar}{2} \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = -i \frac{\hbar}{2} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = -i \frac{\hbar}{2} |\uparrow\rangle.
\end{aligned}$$

Using these relations, we get for example for the $|\downarrow\downarrow\rangle$ state,

$$\begin{aligned}
 2\hat{S}^{(1)} \cdot \hat{S}^{(2)}|\downarrow\downarrow\rangle &= 2\left(\hat{S}_x^{(1)}\hat{S}_x^{(2)} + \hat{S}_y^{(1)}\hat{S}_y^{(2)} + \hat{S}_z^{(1)}\hat{S}_z^{(2)}\right)|\downarrow\downarrow\rangle \\
 &= 2\frac{\hbar}{2}\frac{\hbar}{2}|\uparrow\uparrow\rangle + 2\left(-i\frac{\hbar}{2}\right)\left(-i\frac{\hbar}{2}\right)|\uparrow\uparrow\rangle + 2\left(-\frac{\hbar}{2}\right)\left(-\frac{\hbar}{2}\right)|\downarrow\downarrow\rangle \\
 &= \frac{\hbar^2}{2}|\downarrow\downarrow\rangle.
 \end{aligned}$$

Similarly, for $|\uparrow\uparrow\rangle$, $2\hat{S}^{(1)} \cdot \hat{S}^{(2)}|\uparrow\uparrow\rangle = \frac{\hbar^2}{2}|\downarrow\downarrow\rangle$. For the remaining two states,

$$\begin{aligned}
 2\hat{S}^{(1)} \cdot \hat{S}^{(2)}\frac{1}{\sqrt{2}}[|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle] &= \frac{\hbar^2}{2}\frac{1}{\sqrt{2}}[|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle], \\
 2\hat{S}^{(1)} \cdot \hat{S}^{(2)}\frac{1}{\sqrt{2}}[|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle] &= -\frac{3}{2}\hbar^2\frac{1}{\sqrt{2}}[|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle].
 \end{aligned}$$

Ultimately,

$$\begin{aligned}
 \hat{S}^2|\uparrow\uparrow\rangle &= \left(\frac{3}{2}\hbar^2 + \frac{1}{2}\hbar^2\right)|\uparrow\uparrow\rangle = 2\hbar^2|\uparrow\uparrow\rangle, \\
 \hat{S}^2|\downarrow\downarrow\rangle &= \left(\frac{3}{2}\hbar^2 + \frac{1}{2}\hbar^2\right)|\downarrow\downarrow\rangle = 2\hbar^2|\downarrow\downarrow\rangle, \\
 \hat{S}^2\frac{1}{\sqrt{2}}[|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle] &= 2\hbar^2\frac{1}{\sqrt{2}}[|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle], \\
 \hat{S}^2\frac{1}{\sqrt{2}}[|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle] &= 0\hbar^2\frac{1}{\sqrt{2}}[|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle].
 \end{aligned}$$

This shows that these four states are the eigenstates of \hat{S}_z and \hat{S}^2 for total spin $s = 1$.

Because of that, we can rename them the way we have done that in the previous section,

$$\begin{aligned} |\uparrow\uparrow\rangle &= |1,1\rangle, \\ \frac{1}{\sqrt{2}} [|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle] &= |1,0\rangle, \\ |\downarrow\downarrow\rangle &= |1,-1\rangle, \\ \frac{1}{\sqrt{2}} [|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle] &= |0,0\rangle, \end{aligned}$$

where the labels now denote the quantum numbers s and m_s of the total, two-particle system.

We have rigorously showed that a composite system with two separate sources of angular momentum (in our case, spins $s = \frac{1}{2}$) can be viewed as a system with a single total angular momentum (here, total spin $s = 1$). While we discussed the simplest possible case, our conclusions are not restricted to electrons or half-integer spin.

If one combines angular momentum l_1 with l_2 , the possible outcomes are all angular momenta from $(l_1 + l_2)$ to $(l_1 - l_2)$ (or $(l_2 - l_1)$ if $l_2 > l_1$).

- Combining a particle of spin $\frac{3}{2}$ with a particle of spin 2 gives total spin of $\frac{7}{2}, \frac{5}{2}, \frac{3}{2}$ or $\frac{1}{2}$, depending on the configuration of the two spins.
- If a hydrogen atom is in the state $|nlm_l\rangle$, the net angular momentum of the electron (spin plus orbital) is $l + \frac{1}{2}$ or $l - \frac{1}{2}$. If we now account for the proton (with spin $s = \frac{1}{2}$), the angular momentum quantum number of the atom is $l + 1, l$ or $l - 1$ (and l can be achieved in two distinct ways, depending on whether the electron is in the $l + \frac{1}{2}$ or $l - \frac{1}{2}$ configuration).

Ultimately, whether we view the system as an assembly of several parts (like two separate electron spins) or a one entity (total spin), depends on our choice and these options simply provide different bases.

When considering the total angular momentum of electrons (orbital + spin), it is traditional to use the total angular momentum quantum number j (and the corresponding total angular momentum operator \hat{J}).

Exercise: An electron is in an orbital with angular momentum $l = 1$.

- How many states are there with different quantum numbers?
- Now consider the total angular momentum j of this electron. What are the possible states and how many are there?

Does it matter which of our two points of view (for an electron, separate \hat{L} and \hat{S} vs total angular momentum \hat{J}) we adopt?

As is often the case, in some situations it does. We will look at one example which requires a detour into classical equations and then recalling some of the points from the first section...

Spin-Orbit Coupling

Consider an electron in an orbit in the xy -plane around the nucleus: from the electron's point of view, the proton is circling around it! Such a current loop generates a magnetic field that exerts a torque on the spin of the electron (with a tendency to align electron's magnetic moment with the direction of the applied field). The Hamiltonian of a magnetic moment $\hat{\mu}$ in a magnetic field \hat{B} , as before,

$$\hat{H} = -\hat{\mu} \cdot \hat{B}.$$

Also as before,

$$\hat{\mu} = -\frac{ge}{2m}\hat{S}$$

What is the magnetic field that the electron feels? Using the Biot-Savart law, at the centre of a loop with radius r ,

$$\vec{B} = \frac{\mu_0 I}{2r} \vec{k} = \frac{\mu_0 Z_{\text{eff}} e}{2rT} \vec{k},$$

where \vec{k} is the unit vector along the z -axis. The current is $I = Z_{\text{eff}} e / T$, with T the period of the orbit and Z_{eff} the (effective) nuclear charge. On the other hand, the orbital angular momentum of the electron (in the rest frame of the nucleus), is

$$\vec{L} = \vec{r} \times \vec{p} = rmv\vec{k} = rm \frac{2\pi r}{T} \vec{k} = \frac{2\pi m r^2}{T} \vec{k}.$$

Putting the expressions for \vec{B} and \vec{L} together (both are along the same direction),

$$\vec{B} = \frac{1}{4\pi\epsilon_0} \frac{Z_{\text{eff}} e}{mc^2 r^3} \vec{L},$$

where the speed of light c connects μ_0 and ϵ_0 .

Hence, the Hamiltonian is,

$$\begin{aligned}\hat{H} &= -\hat{\mu} \cdot \hat{B} = -\left(-\frac{ge}{2m}\hat{S}\right) \cdot \frac{1}{4\pi\epsilon_0} \frac{Z_{\text{eff}}e}{mc^2r^3} \hat{L} \\ &= \left(\frac{e^2}{4\pi\epsilon_0}\right) \frac{Z_{\text{eff}}g}{m^2c^2r^3} \hat{S} \cdot \hat{L}.\end{aligned}$$

This result is not fully correct, because electron's frame is not an inertial system – it accelerates as the electron orbits the nucleus. Proper calculation requiring more complicated electrodynamics adds the Thomas precession factor of $\frac{1}{2}$ which cancels the g -factor $g \approx 2$,

$$\hat{H}_{\text{so}} = \left(\frac{e^2}{4\pi\epsilon_0}\right) \frac{Z_{\text{eff}}}{m^2c^2r^3} \hat{S} \cdot \hat{L}.$$

This is so the called spin-orbit interaction (or “coupling”), hence the subscript ‘so’.

Matrix mechanics reminder

Consider a Hamiltonian \hat{H} that does not depend on time and an operator \hat{O} in the Schrodinger picture. In the Heisenberg picture, the transformed operator is

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

Notice that if $\hat{O} = \hat{H}$, then $\hat{H}_H = \hat{H}$, because \hat{H} commutes with itself and the exponentials can be multiplied by each other to give identity. This means that a Hamiltonian which does not depend on time looks the same in the Heisenberg and Schrodinger pictures.

Let us now assume that \hat{O} in the Schrodinger picture does not explicitly depend on time. Then,

$$\begin{aligned} \frac{d}{dt} \hat{O}_H &= \frac{d}{dt} (e^{i\hat{H}t/\hbar} \hat{O} e^{-i\hat{H}t/\hbar}) \\ &= i \frac{\hat{H}}{\hbar} e^{i\hat{H}t/\hbar} \hat{O} e^{-i\hat{H}t/\hbar} + e^{i\hat{H}t/\hbar} \hat{O} \left(-i \frac{\hat{H}}{\hbar} \right) e^{-i\hat{H}t/\hbar} \\ &= \frac{i}{\hbar} \hat{H} e^{i\hat{H}t/\hbar} \hat{O} e^{-i\hat{H}t/\hbar} - \frac{i}{\hbar} e^{i\hat{H}t/\hbar} \hat{O} e^{-i\hat{H}t/\hbar} \hat{H} = \frac{i}{\hbar} [\hat{H}, \hat{O}_H], \end{aligned}$$

which is the Heisenberg equation of motion. Notice that, if \hat{O}_H does not change in time, $[\hat{H}, \hat{O}_H] = 0$. This means that the observable associated with \hat{O}_H is conserved. If this is the case, we also have

$$\begin{aligned} [\hat{H}, \hat{O}_H] &= \hat{H} e^{i\hat{H}t/\hbar} \hat{O} e^{-i\hat{H}t/\hbar} - e^{i\hat{H}t/\hbar} \hat{O} e^{-i\hat{H}t/\hbar} \hat{H} \\ &= e^{i\hat{H}t/\hbar} \hat{H} \hat{O} e^{-i\hat{H}t/\hbar} - e^{i\hat{H}t/\hbar} \hat{O} \hat{H} e^{-i\hat{H}t/\hbar} = e^{i\hat{H}t/\hbar} [\hat{H}, \hat{O}] e^{-i\hat{H}t/\hbar} = 0. \end{aligned}$$

Because the exponential cannot be zero, we must have $[\hat{H}, \hat{O}] = 0$. If an observable is conserved, commutator of the associated operator with the Hamiltonian is zero in both the Schrodinger and Heisenberg pictures.

(*) Exercise: Consider a time-independent Hamiltonian of an electron in a generic (three-dimensional) form,

$$\hat{H} = \frac{\hat{p}^2}{2m} + V(\vec{r}).$$

Construct an argument that for a spherical potential that only depends on the radial distance and does not depend on spin, $V(\vec{r}) = V(r)$, both the angular momentum and spin of the electron are conserved.

For a Hamiltonian including only a spherically symmetric potential and kinetic energy, as you might expect, both the angular momentum and spin are conserved (that is, all of their components are conserved which also implies conservation of magnitudes). However, if we take into account the spin-orbit interaction,

$$\hat{H}_{\text{so}} = \alpha(r)\hat{S} \cdot \hat{L},$$

this is no longer the case. For example, consider,

$$\begin{aligned} [\hat{H}_{\text{tot}}, \hat{L}_x] &= [\hat{H} + \hat{H}_{\text{so}}, \hat{L}_x] = [\hat{H}, \hat{L}_x] + [\hat{H}_{\text{so}}, \hat{L}_x] = [\hat{H}_{\text{so}}, \hat{L}_x] \\ &= [\alpha(r)\hat{S} \cdot \hat{L}, \hat{L}_x]. \end{aligned}$$

Recall that in spherical coordinates, components of angular momentum do not depend on the radial distance and so $\alpha(r)$ must commute with any \hat{L}_i (it commutes with any \hat{S}_i because these do not depend on position at all). Also, $[\hat{S}_i, \hat{L}_j] = 0$. This means that,

$$\begin{aligned} [\hat{H}_{\text{so}}, \hat{L}_x] &= \alpha(r)[\hat{S} \cdot \hat{L}, \hat{L}_x] \\ &= \alpha(r)[\hat{S}_x \hat{L}_x, \hat{L}_x] + \alpha(r)[\hat{S}_y \hat{L}_y, \hat{L}_x] + \alpha(r)[\hat{S}_z \hat{L}_z, \hat{L}_x] = \\ &= \alpha(r)\hat{S}_y [\hat{L}_y, \hat{L}_x] + \alpha(r)[\hat{S}_y, \hat{L}_x] \hat{L}_y + \alpha(r)\hat{S}_z [\hat{L}_z, \hat{L}_x] \\ &+ \alpha(r)[\hat{S}_z, \hat{L}_x] \hat{L}_z = \alpha(r)\hat{S}_y (-i\hbar \hat{L}_z) + \alpha(r)\hat{S}_z (i\hbar \hat{L}_y) \\ &= i\hbar \alpha(r)(\hat{L}_y \hat{S}_z - \hat{L}_z \hat{S}_y) = i\hbar \alpha(r)(\hat{L} \times \hat{S})_x. \end{aligned}$$

Ultimately, looking at other components,

$$[\hat{H}_{\text{so}}, \hat{L}] = i\hbar\alpha(r)\hat{L} \times \hat{S}.$$

I invite you check by yourselves that, in contrast,

$$[\hat{H}_{\text{so}}, \hat{S}] = i\hbar\alpha(r)\hat{S} \times \hat{L},$$

so that for the total angular momentum $\hat{J} = \hat{L} + \hat{S}$,

$$[\hat{H}_{\text{so}}, \hat{J}] = [\hat{H}_{\text{so}}, \hat{L} + \hat{S}] = [\hat{H}_{\text{so}}, \hat{L}] + [\hat{H}_{\text{so}}, \hat{S}] = 0.$$

This means that

- components of \hat{L} and \hat{S} are not conserved (they change in time);
- The eigenstates of \hat{L}_z and \hat{S}_z are not the eigenstates of the total Hamiltonian.

Hence, m_l and m_s are no longer good quantum numbers (good labels to keep track of). However, components of the total angular momentum \hat{J} are conserved so m_j (keeping track of the z-component of the total angular momentum) is a good quantum number.

At the same time, notice that,

$$\begin{aligned} [\hat{H}_{\text{tot}}, \hat{L}^2] &= [\hat{H} + \hat{H}_{\text{so}}, \hat{L}^2] = [\hat{H}, \hat{L}^2] + [\hat{H}_{\text{so}}, \hat{L}^2] = [\hat{H}_{\text{so}}, \hat{L}^2] \\ &= [\alpha(r)\hat{S} \cdot \hat{L}, \hat{L}^2] \\ &= \alpha(r)[\hat{S}_x\hat{L}_x, \hat{L}^2] + \alpha(r)[\hat{S}_y\hat{L}_y, \hat{L}^2] + \alpha(r)[\hat{S}_z\hat{L}_z, \hat{L}^2]. \end{aligned}$$

Again, $[\hat{S}_i, \hat{L}_j] = 0$. Also, recall that $[\hat{L}^2, \hat{L}_i] = 0$. Hence, $[\hat{H}_{\text{so}}, \hat{L}^2] = 0$ and \hat{L}^2 is conserved in the presence of spin-orbit coupling. Similarly, you can check that $[\hat{H}_{\text{so}}, \hat{S}^2] = 0$. This also means that $[\hat{H}_{\text{so}}, \hat{J}^2] = 0$.

To summarize, if we take spin-orbit coupling into account, then l, s, j and m_j are good quantum numbers. However, m_l and m_s are not.

How important is spin-orbit coupling?

$$\hat{H}_{\text{so}} = \left(\frac{e^2}{4\pi\epsilon_0} \right) \frac{Z_{\text{eff}}}{m^2 c^2 r^3} \hat{S} \cdot \hat{L}$$

Let's estimate the energy scale of \hat{H}_{so} for hydrogen:

- $Z_{\text{eff}} = 1$
- $|\hat{S}| \sim \hbar$,
- $|\hat{L}| \sim \hbar$,
- $r \sim a_0 = \frac{4\pi\epsilon_0 \hbar^2}{me^2}$.

$$\varepsilon_{\text{so}} \sim \left(\frac{e^2}{4\pi\epsilon_0} \right) \frac{1}{m^2 c^2} \hbar^2 \left(\frac{me^2}{4\pi\epsilon_0 \hbar^2} \right)^3 = \left(\frac{e^2}{4\pi\epsilon_0 \hbar c} \right)^4 mc^2$$

We have the rest mass energy of an electron, $mc^2 = 0.51$ MeV, which looks large. However, $\frac{e^2}{4\pi\epsilon_0 \hbar c} = \alpha$, where $\alpha = \frac{1}{137}$ is the fine structure constant. Altogether, $\varepsilon_{\text{so}} \sim 1$ meV, which is small compared to the hydrogen level energies $\varepsilon_n = -\frac{\varepsilon_1}{n^2} = -\frac{13.6}{n^2}$ eV.

The smallness of spin-orbit coupling and the need to abandon the hydrogen quantum numbers m_l and m_s and work with the total angular momentum \hat{J} lead us to the question of approximations in quantum mechanics and the next section...

PH20013/60 Quantum Mechanics

Section V: Approximation methods (and some atomic physics)

In this section, we will discuss some options we can resort to when a quantum mechanical problem cannot be solved exactly (or if that requires an effort greater than we want to make). We will then apply them to learn more about the spectrum of the hydrogen atom.

Variational Principle

The list of known quantum mechanical problems that can be solved exactly is shorter than you might expect.

Therefore, it is a very important question to ask: what should we do if we encounter a complicated problem?

The first method we discuss is the variational method. It is based on two notions:

- For many situations, a sufficient “solution” to a quantum problem is the knowledge of its ground state (the state of the lowest energy) and ground state energy, rather than the whole eigensystem.
- Physicists’ intuition might be helpful in “guessing” what the ground state might be, thus avoiding some of the mathematical complications.

The idea behind the method is that the ground state $|\psi_0\rangle$ is the one which corresponds to the lowest possible energy (ground state energy ε_0). This is because everything in nature has a tendency to lower its energy – if a different state/wave function yielded a lower energy, it would become the new ground state. Hence, we could test any guess for the quantum state $|\psi_0\rangle$ and the lower energy we get for this state, the closer we are to the true ground state. Let’s see what this means more mathematically.

For a state $|\psi\rangle$, we have its expansion into the eigenbasis $\{|\varepsilon\rangle\}$ of \hat{H} ,

$$|\psi\rangle = \sum_n \langle \varepsilon_n | \psi \rangle |\varepsilon_n\rangle \equiv \sum_n a_n |\varepsilon_n\rangle.$$

All the eigenstates are normalized and so is $|\psi\rangle$, so that

$$\begin{aligned} 1 = \langle \psi | \psi \rangle &= \sum_m a_m^* \langle \varepsilon_m | \sum_n a_n |\varepsilon_n\rangle = \sum_{m,n} a_m^* a_n \langle \varepsilon_m | \varepsilon_n \rangle \\ &= \sum_{m,n} a_m^* a_n \delta_{mn} = \sum_n a_n^* a_n = \sum_n |a_n|^2. \end{aligned}$$

Computing the expectation value of \hat{H} on $|\psi\rangle$,

$$\begin{aligned} \langle \psi | \hat{H} | \psi \rangle &= \sum_m a_m^* \langle \varepsilon_m | \hat{H} \sum_n a_n |\varepsilon_n\rangle = \sum_{m,n} a_m^* a_n \langle \varepsilon_m | \hat{H} | \varepsilon_n \rangle \\ &= \sum_{m,n} a_m^* a_n \langle \varepsilon_m | \varepsilon_n \rangle \varepsilon_n = \sum_{m,n} \varepsilon_n a_m^* a_n \delta_{mn} = \sum_n \varepsilon_n |a_n|^2. \end{aligned}$$

Let's subtract the ground state energy ε_0 from this expectation value,

$$\begin{aligned} \langle \psi | \hat{H} | \psi \rangle - \varepsilon_0 &= \sum_{n=0}^{+\infty} \varepsilon_n |a_n|^2 - 1 \times \varepsilon_0 = \sum_{n=0}^{+\infty} \varepsilon_n |a_n|^2 - \sum_{n=0}^{+\infty} |a_n|^2 \varepsilon_0 \\ &= \sum_{n=0}^{+\infty} |a_n|^2 (\varepsilon_n - \varepsilon_0) = \sum_{n=1}^{+\infty} |a_n|^2 (\varepsilon_n - \varepsilon_0). \end{aligned}$$

Because $\varepsilon_n > \varepsilon_0$ for $n \geq 1$, we must have $\langle \psi | \hat{H} | \psi \rangle - \varepsilon_0 \geq 0$ and so for any trial state $|\psi\rangle$,

$$\langle \psi | \hat{H} | \psi \rangle \geq \varepsilon_0.$$

Exercise: Consider the quantum harmonic oscillator in one dimension,

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

Determine the upper bound on the ground state energy using a trial wave function $\psi(x) = A e^{-bx^2}$, where A is a normalization constant you need to establish, and b is a real parameter. For what value of b is the energy minimal? You might find the following integrals useful:

$$\int_{-\infty}^{+\infty} dx e^{-bx^2} = \sqrt{\frac{\pi}{b}}, \quad \int_{-\infty}^{+\infty} dx x^2 e^{-bx^2} = \sqrt{\frac{\pi}{4b^3}}.$$

A good choice of the trial wave function form is essential for the success of the variational method. Normally, it requires a combination of many functions to obtain an optimal result.

Non-degenerate Perturbation Theory

The second approximation method is based on the idea that for a problem that is a little different from one we can solve exactly, we can use the exact solutions as the starting point and figure out what corrections are needed.

Consider a Hamiltonian

$$\hat{H} = \hat{H}_0 + \hat{H}_1$$

where \hat{H}_0 is a Hamiltonian for which we can solve the eigenproblem and \hat{H}_1 is a small addition. One refers to \hat{H}_0 as the unperturbed Hamiltonian and \hat{H}_1 as the perturbing Hamiltonian or perturbation.

We assume that to every eigenket $|\varepsilon_n^0\rangle \equiv |n^{(0)}\rangle$ of \hat{H}_0 with eigenvalue $\varepsilon_n^{(0)}$ there is an eigenket $|n\rangle$ of \hat{H} with eigenvalue ε_n (states and eigenvalues are continuously deformed from the unperturbed into perturbed ones as the perturbation is “turned on” from zero to its full strength). We then assume that the eigenkets and eigenvalues of \hat{H} can be expanded in a perturbation series,

$$\begin{aligned} |n\rangle &= |n^{(0)}\rangle + |n^{(1)}\rangle + |n^{(2)}\rangle + \dots, \\ \varepsilon_n &= \varepsilon_n^{(0)} + \varepsilon_n^{(1)} + \varepsilon_n^{(2)} + \dots, \end{aligned}$$

the superscript (k) on each term gives the power of the matrix element of \hat{H}_1 that it is expected to come out proportional to. A term with superscript (k) is called a *kth-order term* (for example, a product like $\varepsilon_n^{(k)} |n^{(q)}\rangle$ is a term of order $k + q$). We hope that as the order increases, the terms get systematically smaller.

To find the term in expansions we proposed, we start with the eigenvalue equation,

$$\hat{H}|n\rangle = \varepsilon_n|n\rangle.$$

Using our expansions,

$$\begin{aligned} & (\hat{H}_0 + \hat{H}_1)(|n^{(0)}\rangle + |n^{(1)}\rangle + \dots) \\ &= \left(\varepsilon_n^{(0)} + \varepsilon_n^{(1)} + \dots \right) (|n^{(0)}\rangle + |n^{(1)}\rangle + \dots). \end{aligned}$$

Let us pick out zeroth-order terms:

$$\hat{H}_0|n^{(0)}\rangle = \varepsilon_n^{(0)}|n^{(0)}\rangle.$$

By our initial assumption, this eigenproblem can be solved.

Let's move on to first order terms:

$$\hat{H}_0|n^{(1)}\rangle + \hat{H}_1|n^{(0)}\rangle = \varepsilon_n^{(0)}|n^{(1)}\rangle + \varepsilon_n^{(1)}|n^{(0)}\rangle. \quad (*)$$

We close on the left with the known unperturbed state $\langle n^{(0)}|$, to get

$$\langle n^{(0)}|\hat{H}_0|n^{(1)}\rangle + \langle n^{(0)}|\hat{H}_1|n^{(0)}\rangle = \varepsilon_n^{(0)}\langle n^{(0)}|n^{(1)}\rangle + \varepsilon_n^{(1)}\langle n^{(0)}|n^{(0)}\rangle.$$

Notice that, because \hat{H}_0 is Hermitian, $\langle n^{(0)}|\hat{H}_0 = \langle n^{(0)}|\varepsilon_n^{(0)}$.

Also, $\langle n^{(0)}|n^{(0)}\rangle = 1$. We get,

$$\langle n^{(0)}|\hat{H}_1|n^{(0)}\rangle = \varepsilon_n^{(1)},$$

that is, the first-order change in the energy is the expectation value of the perturbation in the unperturbed state.

Let us go back to (*) and dot with another unperturbed state, $\langle m^{(0)}|$, $m \neq n$. We get

$$\langle m^{(0)}|\hat{H}_0|n^{(1)}\rangle + \langle m^{(0)}|\hat{H}_1|n^{(0)}\rangle = \varepsilon_n^{(0)}\langle m^{(0)}|n^{(1)}\rangle + \varepsilon_n^{(1)}\langle m^{(0)}|n^{(0)}\rangle.$$

This time, we have $\langle m^{(0)}|\hat{H}_0 = \langle m^{(0)}|\varepsilon_m^{(0)}$ and $\langle m^{(0)}|n^{(0)}\rangle = 0$.

This gives,

$$\varepsilon_m^{(0)}\langle m^{(0)}|n^{(1)}\rangle + \langle m^{(0)}|\hat{H}_1|n^{(0)}\rangle = \varepsilon_n^{(0)}\langle m^{(0)}|n^{(1)}\rangle,$$

$$\langle m^{(0)}|n^{(1)}\rangle = \frac{\langle m^{(0)}|\hat{H}_1|n^{(0)}\rangle}{\varepsilon_n^{(0)} - \varepsilon_m^{(0)}}.$$

The above inner product represents the coefficients of the correction $|n^{(1)}\rangle$ to the eigenstate in the basis of \hat{H}_0 , except the component $\langle n^{(0)}|n^{(1)}\rangle$ because we chose $m \neq n$. It can be shown that the normalization requirement results in $\langle n^{(0)}|n^{(1)}\rangle = 0$ so that, to the first order,

$$|n\rangle = |n^{(0)}\rangle + |n^{(1)}\rangle = |n^{(0)}\rangle + \sum_{m \neq n} \frac{\langle m^{(0)}|\hat{H}_1|n^{(0)}\rangle}{\varepsilon_n^{(0)} - \varepsilon_m^{(0)}} |m^{(0)}\rangle.$$

(*) Exercise: Show that, keeping terms only to the first order, the normalization condition for $|n\rangle$ implies $\langle n^{(0)}|n^{(1)}\rangle = 0$.

One can proceed in a similar fashion to derive corrections to the eigenvalues and eigenstates of higher orders. We will stop here. However, notice that our expansion for $|n\rangle$ is only safe if the eigenvalues are non-degenerate, that is, $\varepsilon_n^{(0)} - \varepsilon_m^{(0)} \neq 0$ for any $m \neq n$.

Exercise: Consider the one-dimensional infinite potential well,

$$V(x) = \begin{cases} 0 & \text{for } |x| < \frac{L}{2}, \\ \infty & \text{for } |x| > \frac{L}{2}, \end{cases}$$

with energies

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

and eigenstates

$$\psi_n(x) = \begin{cases} \sqrt{\frac{2}{L}} \sin \frac{n\pi x}{L}, & \text{for } n \text{ even,} \\ \sqrt{\frac{2}{L}} \cos \frac{n\pi x}{L}, & \text{for } n \text{ odd.} \end{cases}$$

- Suppose we perturb the system by simply raising the “floor” of the well by a constant amount, V_0 (so that $V(x) = V_0$ for $|x| < \frac{L}{2}$). Find the first-order correction to the energies.
- What if only half of the floor is raised, $V(x) = V_0$ for $-\frac{L}{2} < x < 0$?
- Finally, what if the perturbation is a linear function, $V(x) = \frac{2V_0}{L}x$ for $|x| < \frac{L}{2}$, instead?

Degenerate Perturbation Theory

The perturbation approach we have derived so far is bound to break down miserably if the unperturbed Hamiltonian \hat{H}_0 is degenerate. This is because of the divergence in the expression for the first-order correction to the eigenstate,

$$|n^{(1)}\rangle = \sum_{m \neq n} \frac{\langle m^{(0)} | \hat{H}_1 | n^{(0)} \rangle}{\varepsilon_n^{(0)} - \varepsilon_m^{(0)}} |m^{(0)}\rangle.$$

This is because the p eigenvectors corresponding to the p -fold degenerate eigenvalue define a p -dimensional vector space such that any linear combination of these vectors is also an eigenstate of \hat{H}_0 . This means there are in fact infinitely many choices of a basis that are all good enough to serve as eigenkets of \hat{H}_0 . However, once the perturbation is applied, typically at least some, if not all, of the degeneracy associated with the eigenvalues $\varepsilon_n^{(0)} = \varepsilon_m^{(0)} = \dots$, is lifted as these eigenvalues are modified by the perturbation. Once the eigenvalues stop being degenerate, it is no longer true that any linear combination of the corresponding eigenkets is still an eigenket. Hence, only some of the originally infinitely many options (as few as one if the degeneracy has been lifted completely) provide a good set of eigenvectors for the total Hamiltonian $\hat{H} = \hat{H}_0 + \hat{H}_1$. If we are lucky to have made a good choice at the beginning, then the matrix element $\langle m^{(0)} | \hat{H}_1 | n^{(0)} \rangle = 0$ which prevents the divergence in $|n^{(1)}\rangle$. If not, we can do nothing. Ideally, we want to develop a method that avoids the need to be lucky, especially given that plenty of applications of perturbation theory involve degenerate states.

We will consider here the simplest example of two-fold degeneracy. Suppose that,

$$\begin{aligned}\hat{H} &= \hat{H}_0 + \hat{H}_1 \\ \hat{H}_0|\psi_{0,a}\rangle &= \varepsilon^{(0)}|\psi_{0,a}\rangle, \quad \hat{H}_0|\psi_{0,b}\rangle = \varepsilon^{(0)}|\psi_{0,b}\rangle, \\ \langle\psi_{0,a}|\psi_{0,b}\rangle &= 0,\end{aligned}$$

with both of the states normalized.

Exercise: Persuade yourself that any linear combination of $|\psi_{0,a}\rangle$ and $|\psi_{0,b}\rangle$ is still an eigenstate of \hat{H}_0 with the same eigenvalue, $\varepsilon^{(0)}$.

We want to solve the eigenproblem,

$$\hat{H}|\psi\rangle = \varepsilon|\psi\rangle,$$

at energies close to the degenerate $\varepsilon^{(0)}$, with the unperturbed state being a linear combination of two eigenstates, $|\psi_{0,a}\rangle$ and $|\psi_{0,b}\rangle$, $|\psi^{(0)}\rangle = \alpha|\psi_{0,a}\rangle + \beta|\psi_{0,b}\rangle$, which might or might not turn out to be a particularly lucky choice to describe the perturbation.

As before, the energy and eigenstate can be expanded,

$$\begin{aligned}|\psi\rangle &= |\psi^{(0)}\rangle + |\psi^{(1)}\rangle + |\psi^{(2)}\rangle + \dots, \\ \varepsilon &= \varepsilon^{(0)} + \varepsilon^{(1)} + \varepsilon^{(2)} + \dots,\end{aligned}$$

and plugging these back into the eigenproblem,

$$\begin{aligned}(\hat{H}_0 + \hat{H}_1)(|\psi^{(0)}\rangle + |\psi^{(1)}\rangle + \dots) \\ = (\varepsilon^{(0)} + \varepsilon^{(1)} + \dots)(|\psi^{(0)}\rangle + |\psi^{(1)}\rangle + \dots).\end{aligned}$$

Again, we need to keep track of the order of the terms,

$$\begin{aligned} & \hat{H}_0|\psi^{(0)}\rangle + (\hat{H}_0|\psi^{(1)}\rangle + \hat{H}_1|\psi^{(0)}\rangle) + \dots \\ &= \varepsilon^{(0)}|\psi^{(0)}\rangle + (\varepsilon^{(0)}|\psi^{(1)}\rangle + \varepsilon^{(1)}|\psi^{(0)}\rangle) + \dots. \end{aligned}$$

The 0th order term cancels out because $\hat{H}_0|\psi\rangle = \varepsilon^{(0)}|\psi\rangle$. At the first order, we have,

$$\hat{H}_0|\psi^{(1)}\rangle + \hat{H}_1|\psi^{(0)}\rangle = \varepsilon^{(0)}|\psi^{(1)}\rangle + \varepsilon^{(1)}|\psi^{(0)}\rangle. \quad (*)$$

Taking the inner product with $\langle\psi_{0,a}|$,

$$\begin{aligned} & \langle\psi_{0,a}|\hat{H}_0|\psi^{(1)}\rangle + \langle\psi_{0,a}|\hat{H}_1|\psi^{(0)}\rangle \\ &= \varepsilon^{(0)}\langle\psi_{0,a}|\psi^{(1)}\rangle + \varepsilon^{(1)}\langle\psi_{0,a}|\psi^{(0)}\rangle. \end{aligned}$$

Because \hat{H}_0 is Hermitian,

$$\begin{aligned} \langle\psi_{0,a}|\hat{H}_0|\psi^{(1)}\rangle &= \langle\psi_{0,a}|\hat{H}_0^\dagger|\psi^{(1)}\rangle = (\hat{H}_0|\psi_{0,a}\rangle)^\dagger|\psi^{(1)}\rangle \\ &= \varepsilon^{(0)}\langle\psi_{0,a}|\psi^{(1)}\rangle, \end{aligned}$$

so that the first terms on the left and right cancel each other. We are left with

$$\langle\psi_{0,a}|\hat{H}_1|\psi^{(0)}\rangle = \varepsilon^{(1)}\langle\psi_{0,a}|\psi^{(0)}\rangle.$$

We now plug in $|\psi^{(0)}\rangle = \alpha|\psi_{0,a}\rangle + \beta|\psi_{0,b}\rangle$,

$$\langle\psi_{0,a}|\hat{H}_1(\alpha|\psi_{0,a}\rangle + \beta|\psi_{0,b}\rangle) = \varepsilon^{(1)}\langle\psi_{0,a} | (\alpha|\psi_{0,a}\rangle + \beta|\psi_{0,b}\rangle).$$

Because $\langle\psi_{0,a}|\psi_{0,b}\rangle = 0$, we get

$$\alpha\langle\psi_{0,a}|\hat{H}_1|\psi_{0,a}\rangle + \beta\langle\psi_{0,a}|\hat{H}_1|\psi_{0,b}\rangle = \alpha\varepsilon^{(1)}.$$

We now go back to (*) and take the inner product with $\langle\psi_{0,b}|$.

In a manner similar to before, we obtain

$$\langle \psi_{0,b} | \hat{H}_1 | \psi^{(0)} \rangle = \varepsilon^{(1)} \langle \psi_{0,b} | \psi^{(0)} \rangle.$$

Using $|\psi^{(0)}\rangle = \alpha|\psi_{0,a}\rangle + \beta|\psi_{0,b}\rangle$ again,

$$\langle \psi_{0,b} | \hat{H}_1 (\alpha|\psi_{0,a}\rangle + \beta|\psi_{0,b}\rangle) = \varepsilon^{(1)} \langle \psi_{0,b} | (\alpha|\psi_{0,a}\rangle + \beta|\psi_{0,b}\rangle).$$

so that

$$\alpha \langle \psi_{0,b} | \hat{H}_1 | \psi_{0,a} \rangle + \beta \langle \psi_{0,b} | \hat{H}_1 | \psi_{0,b} \rangle = \beta \varepsilon^{(1)}.$$

Let's look at both the equations together,

$$\begin{aligned} \alpha \langle \psi_{0,a} | \hat{H}_1 | \psi_{0,a} \rangle + \beta \langle \psi_{0,a} | \hat{H}_1 | \psi_{0,b} \rangle &= \alpha \varepsilon^{(1)}, \\ \alpha \langle \psi_{0,b} | \hat{H}_1 | \psi_{0,a} \rangle + \beta \langle \psi_{0,b} | \hat{H}_1 | \psi_{0,b} \rangle &= \beta \varepsilon^{(1)}. \end{aligned}$$

Recall that $|\psi_{0,a}\rangle$ and $|\psi_{0,b}\rangle$ provide one of possible choices for a basis in a two-dimensional (sub)space corresponding to the two-fold degenerate eigenvalue $\varepsilon^{(0)}$. The equations above describe relations for the matrix element of the perturbation, \hat{H}_1 , in this basis. We can rewrite both equations using a matrix,

$$\begin{bmatrix} \langle \psi_{0,a} | \hat{H}_1 | \psi_{0,a} \rangle & \langle \psi_{0,a} | \hat{H}_1 | \psi_{0,b} \rangle \\ \langle \psi_{0,b} | \hat{H}_1 | \psi_{0,a} \rangle & \langle \psi_{0,b} | \hat{H}_1 | \psi_{0,b} \rangle \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \varepsilon^{(1)} \begin{bmatrix} \alpha \\ \beta \end{bmatrix},$$

which casts the energy correction $\varepsilon^{(1)}$ as an eigenvalue of an eigenproblem of \hat{H}_1 . The size of the matrix is ultimately prescribed by the degeneracy of $\varepsilon^{(0)}$. We can compute all the matrix elements because the basis vectors are the eigenstates of the unperturbed Hamiltonian \hat{H}_1 .

$$\begin{bmatrix} \langle \psi_{0,a} | \hat{H}_1 | \psi_{0,a} \rangle & \langle \psi_{0,a} | \hat{H}_1 | \psi_{0,b} \rangle \\ \langle \psi_{0,b} | \hat{H}_1 | \psi_{0,a} \rangle & \langle \psi_{0,b} | \hat{H}_1 | \psi_{0,b} \rangle \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \varepsilon^{(1)} \begin{bmatrix} \alpha \\ \beta \end{bmatrix},$$

It is normal to introduce a shorthand for the matrix elements just like we have done it before ($\Omega_{ij} = \langle i | \hat{\Omega} | j \rangle$),

$$\langle \psi_{0,i} | \hat{H}_1 | \psi_{0,j} \rangle = H_{1,ij},$$

so that our eigenproblem is given by

$$\begin{bmatrix} H_{1,aa} & H_{1,ab} \\ H_{1,ba} & H_{1,bb} \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \varepsilon^{(1)} \begin{bmatrix} \alpha \\ \beta \end{bmatrix}.$$

The eigenvalues of the above matrix give the first-order corrections to the energy and the corresponding eigenvectors tell us the coefficients α and β that determine the “good” states that are the eigenstates after taking the perturbation into account. Notice that we must have $H_{1,ab} = H_{1,ba}^*$ because as part of the total Hamiltonian \hat{H} , \hat{H}_1 must be Hermitian. We often call the off-diagonal element the coupling between states $|\psi_{0,a}\rangle$ and $|\psi_{0,b}\rangle$. If this off-diagonal element $H_{1,ab} = H_{1,ba}^* = 0$, the two eigenvectors are $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$. The perturbation matrix is diagonal with the eigenvalues,

$$\begin{aligned} \varepsilon_1^{(1)} &= \hat{H}_{1,aa} = \langle \psi_{0,a} | \hat{H}_1 | \psi_{0,a} \rangle, \\ \varepsilon_2^{(1)} &= \hat{H}_{1,bb} = \langle \psi_{0,b} | \hat{H}_1 | \psi_{0,b} \rangle. \end{aligned}$$

which agrees with the result obtained from non-degenerate perturbation theory. This is the case of us being lucky and choosing the “good” basis from the very start. However, if $H_{1,ab} = H_{1,ba}^* \neq 0$, we have to diagonalize the perturbation matrix.

The determinant of interest is

$$\begin{vmatrix} H_{1,aa} - \varepsilon^{(1)} & H_{1,ab} \\ H_{1,ab}^* & H_{1,bb} - \varepsilon^{(1)} \end{vmatrix} = 0,$$

where we accounted for $H_{1,ba} = H_{1,ab}^*$. This gives,

$$(H_{1,aa} - \varepsilon^{(1)})(H_{1,bb} - \varepsilon^{(1)}) - |H_{1,ab}|^2 = 0,$$

with solutions

$$\varepsilon_{\pm}^{(1)} = \frac{1}{2} \left(H_{1,aa} + H_{1,bb} \pm \sqrt{(H_{1,aa} - H_{1,bb})^2 + 4|H_{1,ab}|^2} \right).$$

Let's say that $H_{1,aa} = H_{1,bb} = 0$. Then,

$$\varepsilon_{\pm}^{(1)} = \pm |H_{1,ab}|,$$

and accounting for the initial energy of the levels,

$$\varepsilon = \varepsilon^{(0)} + \varepsilon^{(1)} = \varepsilon^{(0)} \pm |H_{1,ab}|.$$

This illustrates generic behaviour of degenerate states under perturbation: coupling between states introduced by the perturbation leads to their repulsion (one level moves up and one level moves down).

Finally, notice that we started from the total Hamiltonian, $\hat{H} = \hat{H}_0 + \hat{H}_1$. Our discussion showed that, in the case of a degeneracy, we can use a basis of eigenvectors associated with the degenerate eigenvalue (just choose/make them orthogonal) to describe the perturbation. These eigenvectors are already eigenstates of \hat{H}_0 , so that the full solution,

$$\varepsilon = \varepsilon^{(0)} + \varepsilon^{(1)},$$

is equivalent (check this!) to diagonalizing the full Hamiltonian,

$$\begin{bmatrix} \langle \psi_{0,a} | \hat{H} | \psi_{0,a} \rangle & \langle \psi_{0,a} | \hat{H} | \psi_{0,b} \rangle \\ \langle \psi_{0,b} | \hat{H} | \psi_{0,a} \rangle & \langle \psi_{0,b} | \hat{H} | \psi_{0,b} \rangle \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \varepsilon \begin{bmatrix} \alpha \\ \beta \end{bmatrix}$$

$$\begin{bmatrix} \langle \psi_{0,a} | \hat{H}_0 + \hat{H}_1 | \psi_{0,a} \rangle & \langle \psi_{0,a} | \hat{H}_0 + \hat{H}_1 | \psi_{0,b} \rangle \\ \langle \psi_{0,b} | \hat{H}_0 + \hat{H}_1 | \psi_{0,a} \rangle & \langle \psi_{0,b} | \hat{H}_0 + \hat{H}_1 | \psi_{0,b} \rangle \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \varepsilon \begin{bmatrix} \alpha \\ \beta \end{bmatrix}$$

$$\begin{bmatrix} \varepsilon_a^{(0)} + H_{1,aa} & H_{1,ab} \\ H_{1,ba} & \varepsilon_b^{(0)} + H_{1,bb} \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \varepsilon \begin{bmatrix} \alpha \\ \beta \end{bmatrix},$$

where $\varepsilon_i^{(0)}$ is the unperturbed energy of state $|\psi_{0,i}\rangle$.

1. Compare the expression above with our discussion of the hydrogen molecule. In this picture, we can view the bonding as “perturbation due to coupling” of the previously degenerate states of the two isolated atoms.
2. Framing the complete problem like above allows to generalize degenerate perturbation theory to the “almost degenerate” case (see one of the **Exercises**).

Summary

1. Variational principle allows to place an upper bound on the ground state energy. The better the guess of the trial wave function, the tighter the bound.
2. For non-degenerate perturbation theory, the first-order correction to the energy $\varepsilon_n^{(0)}$ of the unperturbed state $|n^{(0)}\rangle$ due to the perturbation \hat{H}_1 is,

$$\langle n^{(0)} | \hat{H}_1 | n^{(0)} \rangle = \varepsilon_n^{(1)}.$$

3. In the case of a degenerate level, the expression from point 2 might fail. Instead, one should diagonalize the perturbation (or the full Hamiltonian) in the basis of the eigenstates associated with the degenerate eigenvalue (there are many choices of such a basis but any will do as long as the states are orthogonal).

Let's go back to spin-orbit coupling,

$$\hat{H}_{\text{so}} = \left(\frac{e^2}{4\pi\epsilon_0} \right) \frac{Z_{\text{eff}}}{m^2 c^2 r^3} \hat{S} \cdot \hat{L}$$

On the one hand, the electronic states are highly degenerate: for every n , there are $2n^2$ states that have the same energy, $\epsilon_n = -\frac{\hbar^2}{2ma_0^2} \frac{1}{n^2}$. Should we use degenerate perturbation theory?

We have shown that most of the quantum numbers remain good quantum numbers after taking \hat{H}_{so} into account. The basis of hydrogen-like states is “good enough” and it is enough to use the first-order non-degenerate perturbation theory result, $\epsilon_n^{(1)} = \langle n^{(0)} | \hat{H}_1 | n^{(0)} \rangle$. In our context, this means that the energy correction is

$$\epsilon_{\text{so}}^{(1)} = \langle nlm_l | \hat{H}_{\text{so}} | nlm_l \rangle = \left(\frac{e^2}{4\pi\epsilon_0} \right) \frac{Z_{\text{eff}}}{m^2 c^2} \langle nlm_l | \frac{1}{r^3} \hat{S} \cdot \hat{L} | nlm_l \rangle.$$

Notice that, because $\hat{J} = \hat{L} + \hat{S}$,

$$\hat{J}^2 = (\hat{L} + \hat{S}) \cdot (\hat{L} + \hat{S}) = \hat{L}^2 + \hat{S}^2 + 2\hat{S} \cdot \hat{L},$$

and we have

$$\hat{S} \cdot \hat{L} = \frac{1}{2} (\hat{J}^2 - \hat{L}^2 - \hat{S}^2).$$

Therefore, the eigenvalues of $\hat{S} \cdot \hat{L}$ are

$\frac{\hbar^2}{2} (j(j+1) - l(l+1) - s(s+1))$. Here, we have $s = \frac{1}{2}$, so

$$\begin{aligned} \epsilon_{\text{so}}^{(1)} &= \left(\frac{e^2}{4\pi\epsilon_0} \right) \frac{Z}{m^2 c^2} \frac{\hbar^2 \left(j(j+1) - l(l+1) - \frac{3}{4} \right)}{2} \langle nlm_l | \frac{1}{r^3} | nlm_l \rangle. \end{aligned}$$

It can be shown that,

$$\langle nlm_l | \frac{1}{r^3} | nlm_l \rangle = \frac{Z_{\text{eff}}^3}{l(l + \frac{1}{2})(l + 1)n^3 a_0^3},$$

so that

$$\varepsilon_{\text{so}}^{(1)} = \left(\frac{e^2}{4\pi\epsilon_0} \right) \frac{Z_{\text{eff}}^4}{m^2 c^2} \frac{\hbar^2 \left(j(j + 1) - l(l + 1) - \frac{3}{4} \right)}{2l(l + \frac{1}{2})(l + 1)n^3 a_0^3}.$$

Given that $\varepsilon_n = -\frac{m}{2\hbar^2} \left(\frac{e^2}{4\pi\epsilon_0} \right)^2 \frac{1}{n^2} = -\frac{\varepsilon_1}{n^2}$, we have

$$\varepsilon_{\text{so}}^{(1)} = \frac{(\varepsilon_n)^2 Z_{\text{eff}}^4}{mc^2} \frac{n \left(j(j + 1) - l(l + 1) - \frac{3}{4} \right)}{l(l + \frac{1}{2})(l + 1)}.$$

As you will learn next year, the ultimate origin of the spin-orbit coupling lies in relativistic effects (notice the prefactor which compares the energy ε_n to the rest energy of an electron, mc^2). It turns out that another effect provides a correction to energy of similar order. Consider the kinetic energy, $\hat{T} = \hat{P}^2/(2m)$, as the low-velocity quantum limit of the classical relativistic expression,

$$T = \sqrt{p^2 c^2 + m^2 c^4} - mc^2$$

(total relativistic energy minus the rest energy – all that remains must be due to motion). If we take $p \ll mc$ and expand in powers of (p/mc) ,

$$\begin{aligned} T &= mc^2 \left[\sqrt{1 + \left(\frac{p}{mc} \right)^2} - 1 \right] \\ &= mc^2 \left[1 + \frac{1}{2} \left(\frac{p}{mc} \right)^2 - \frac{1}{8} \left(\frac{p}{mc} \right)^4 + \dots - 1 \right] \approx \frac{1}{2} \frac{p^2}{m} - \frac{1}{8} \frac{p^4}{m^3 c^2}. \end{aligned}$$

The first order correction to the “basic” non-relativistic kinetic energy is,

$$\hat{T}_{\text{rel}} = -\frac{\hat{p}^4}{8m^3c^2}.$$

For a hydrogen-like atom, in first-order perturbation theory,

$$\varepsilon_{\text{rel}}^{(1)} = \langle nlm_l | -\frac{\hat{p}^4}{8m^3c^2} | nlm_l \rangle = -\frac{1}{8m^3c^2} \langle nlm_l | \hat{p}^4 | nlm_l \rangle.$$

The required expectation value can be calculated to give,

$$\varepsilon_{\text{rel}}^{(1)} = -\frac{(\varepsilon_n)^2 Z_{\text{eff}}^4}{mc^2} \left(\frac{2n}{l + \frac{1}{2}} - \frac{3}{2} \right).$$

Adding this to the spin-orbit correction gives (after relating quantum number l to j , $j = l \pm \frac{1}{2}$),

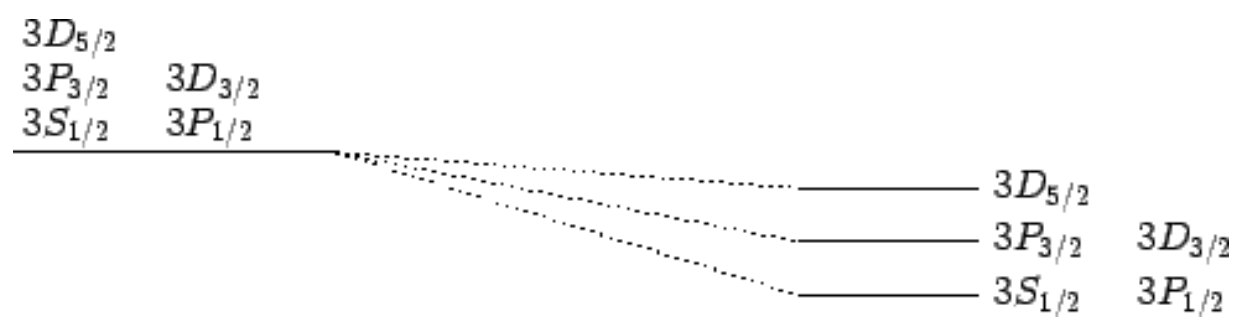
$$\varepsilon_{n,j}^{(1)} = \varepsilon_{\text{so}}^{(1)} + \varepsilon_{\text{rel}}^{(1)} = \frac{(\varepsilon_n)^2 Z_{\text{eff}}^4}{2mc^2} \left(3 - \frac{4n}{j + \frac{1}{2}} \right).$$

- The energy now depends on both the principal quantum number n and the total angular momentum quantum number j .
- Note that our corrections are proportional to Z_{eff}^4 . While they are small for the hydrogen atom, $Z_{\text{eff}} = 1$, $\varepsilon^{(1)} \sim 1$ meV, the energy scale clearly grows very quickly with increasing Z_{eff} .

Let us focus on hydrogen. Together with the energies we determined at the beginning of Section 4, we have the grand result for the energy levels of hydrogen,

$$\varepsilon_{n,j} = -\frac{13.6 \text{ eV}}{n^2} \left[1 + \frac{\alpha^2}{n^2} \left(\frac{n}{j + \frac{1}{2}} - 3 \right) \right].$$

- The “basic” Schrodinger equation, despite providing more quantum numbers, gives energy levels in agreement with the simple Bohr model.
- We have managed to obtain the first-order corrections to the “Bohr” result (so called “fine structure”). They can be measured experimentally which was first done in 1887 by Michelson and Morley, predating the quantum mechanical explanation. The correctness of our calculations contributes to the confirmation of the validity of quantum mechanics.



unperturbed

+ fine structure

- It is conventional to refer to the energy eigenstates which are also simultaneous eigenstates of \hat{J}^2 as nL_j states, where n is the principal quantum number, $L = (S, P, D, F, \dots)$ as $l = (0, 1, 2, 3, \dots)$, and j is the total angular momentum quantum number.
- For $n = 1$, in the absence of fine structure, there are two (because of spin!) degenerate $1S_{1/2}$ states. The fine structure induced energy-shifts of these two states are the same. Hence, fine structure does not break their degeneracy.
- For $n = 2$, in the absence of fine structure, there are two $2S_{1/2}$ states, two $2P_{1/2}$ states, and four $2P_{3/2}$ states, all of which are degenerate. The fine structure-induced energy shifts of the $2S_{1/2}$ and $2P_{1/2}$ states are the same, but different from the energy shift of the $2P_{3/2}$ states. Hence, fine structure does not break the degeneracy of the $2S_{1/2}$ and $2P_{1/2}$ states but does break the degeneracy of these states relative to the $2P_{3/2}$ states.
- For $n = 3$, in the absence of fine structure, there are two $3S_{1/2}$ states, two $3P_{1/2}$ states, four $3P_{3/2}$ states, four $3D_{3/2}$ states, and six $3D_{5/2}$ states, all of which are degenerate. Fine structure breaks these states into three groups: (i) the $3S_{1/2}$ and $3P_{1/2}$ states, (ii) the $3P_{3/2}$ and $3D_{3/2}$ states, and (iii) the $3D_{5/2}$ states.

Exercise: Determine

- the energy splitting between the $2P_{1/2}$ and $2P_{3/2}$ states of hydrogen;
- the ratio of the fine structure correction to the energy difference between the $1S_{1/2}$ and $2P_{3/2}$ states to the unperturbed energy difference between these states.