CHEM552: Advanced Quantum Chemistry

Jussi Eloranta (Eucalyptus Hall 2025)
E-mail: Jussi.Eloranta@csun.edu

(Updated: May 7, 2018)

Chapter 1: The foundations of quantum mechanics



Niels Bohr (1885 - 1962; Nobel price 1922):
"Anyone who is not shocked by quantum theory has not understood it"



Enrico Fermi (1901 - 1954; Nobel price 1932):
"You cannot understand quantum mechanics but you can get used to it"



Richard Feynman (1918-1988; Nobel price 1965):
"I think I can safely say that nobody understands quantum mechanics"

Operators in quantum mechanics

Observable: Any dynamical variable of the system that can be measured. In classical mechanics these are represented by functions whereas in quantum mechanics they are operators.

Operator: Is a symbol that defines a set of mathematical operations that act on a given function. Examples of operators are: multiplication by a constant, differentiation, etc. General operators are denoted by Ω whereas H is reserved to represent the Hamiltonian operator, which yields the total energy of the system. Sometimes operators are represented with a "hat" above them, $\hat{H}$.

$$\textit{The Schrödinger equation: } H\psi = E\psi.$$

This is a linear eigenvalue problem, where H contains typically partial derivatives with respect to the spatial coordinates.

1.1 Linear operators

All operators in quantum mechanics are linear. A linear operator is defined as follows:

$$\Omega(f+g) = \Omega f + \Omega g \ (f \text{ and } g \text{ are functions}) \tag{1.1}$$
$$\Omega(af) = a\Omega f \ (a \text{ is a constant and } f \text{ is a function})$$

Example 1.1 Are the following operators linear?

a) Multiplication by a constant C.

b) Differentiation.

c) Integration.

d) Taking natural logarithm.

Solution.

a) $C(f+g) = Cf + Cg$ and $C(af) = Caf = aCf$. This fulfills Eq. (1.1).

b) Let $f = f(x)$. $\frac{d}{dx}(f+g) = \frac{df}{dx} + \frac{dg}{dx}$ and $\frac{d}{dx}(af) = a\frac{df}{dx}$. This is operator is linear.

c) $\int (f+g)dx = \int f dx + \int g dx$ and $\int af dx = a\int f dx$.

d) $\ln(f+g) \neq \ln(f) + \ln(g) = \ln(fg)$. This operator is not linear.

1.2 Eigenfunctions and eigenvalues

When operators operate on a given function, the outcome is another function. For example, differentiation of $\sin(x)$ gives $\cos(x)$. In some special cases the outcome of an operation is the same function multiplied by a constant. These functions are called eigenfunctions of the given operator Ω. If f is an eigenfunction of Ω, we have:

$$\Omega f = \omega f \tag{1.2}$$

where ω is a constant and is called the eigenvalue of Ω.

Example 1.2 Is the function $f(x) = \cos(3x + 5)$ an eigenfunction of the operator $\Omega = d^2/dx^2$ and, if so, what is the corresponding eigenvalue?

Solution. By operating on the function we get:

$$\underbrace{\frac{d^2}{dx^2}}_{=\Omega}\underbrace{\cos(3x+5)}_{=f} = \frac{d}{dx}\left(-3\sin(3x+5)\right) = \underbrace{-9}_{=\omega}\underbrace{\cos(3x+5)}_{=f}$$

Thus this has the right form required in Eq. (1.2) and f is an eigenfunction of operator Ω. The corresponding eigenvalue is -9. Note that eigenfunctions and eigenvalues go together in pairs. There are many possible (eigenfunction, eigenvalue) pairs for a given operator.

Any well behaved function can be expressed as a linear combination of eigenfunctions of an operator ($\Omega f_n = \omega_n f_n$):

$$g = \sum_{i=1}^{\infty} c_n f_n \tag{1.3}$$

where c_n are coefficients that are specific to function g. The advantage of this expansion is that we know exactly how Ω operates on each term:

$$\Omega g = \Omega \sum_{i=1}^{\infty} c_n f_n = \sum_{i=1}^{\infty} c_n \Omega f_n = \sum_{i=1}^{\infty} c_n \omega_n f_n$$

When many functions have the same eigenvalue, these eigenfunctions are said to be *degenerate*. Let $f_1, f_2, ..., f_k$ be all eigenfunctions of Ω so that they have the same eigenvalue ω, then we have:

$$\Omega f_n = \omega f_n, \text{ with } n = 1, 2, ..., k \tag{1.4}$$

Any linear combination of these functions is also an eigenfunction of Ω. Let g be a linear combination of f_n's, then we have:

$$\Omega g = \Omega \sum_{n=1}^{k} c_n f_n = \sum_{n=1}^{k} c_n \Omega f_n = \sum_{n=1}^{k} c_n \omega f_n = \omega \sum_{n=1}^{k} c_n f_n = \omega g$$

This has the form of an eigenvalue equation: $\Omega g = \omega g$.

Example. Show that any linear combination of e^{2ix} and e^{-2ix} is an eigenfunction of the operator d^2/dx^2.

Solution.

$$\frac{d^2}{dx^2}e^{\pm 2ix} = \pm 2i\frac{d}{dx}e^{\pm 2ix} = -4e^{\pm 2ix}$$

Operation on any linear combination gives then:

$$\underbrace{\frac{d^2}{dx^2}}_{=\Omega}\underbrace{\left(ae^{2ix} + be^{-2ix}\right)}_{=g} = \underbrace{-4}_{=\omega}\underbrace{\left(ae^{2ix} + be^{-2ix}\right)}_{=g}$$

A set of functions $g_1, g_2, ..., g_k$ are said to be *linearly independent* if it is not possible to find constants $c_1, c_2, ..., c_k$ such that

$$\sum_{i=1}^{k} c_i g_i = 0$$

when exlcuding the trivial solution $c_1 = c_2 = ... = c_k = 0$. The dimension of the set, k, gives the number of possible linearly independent functions that can be constructed from the functions. For example, from three $2p$ orbitals, it is possible to construct three different linearly independent functions.

1.3 Representations

The most common representation is the *position representation*. In this representation the position (x) and momentum (p_x) operators are given by:

$$x \rightarrow x\times \text{ and } p_x \rightarrow \frac{\hbar}{i}\frac{\partial}{\partial x} \tag{1.5}$$

where $\hbar = \frac{h}{2\pi}$. An alternative representation is the *momentum representation*:

$$x \rightarrow -\frac{\hbar}{i}\frac{\partial}{\partial x} \text{ and } p_x \rightarrow p_x\times \tag{1.6}$$

The two representations are related to each other via the Fourier duality (see Quantum Chemistry I lecture notes):

$$\psi(p_x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\psi(x)e^{p_x x/\hbar}dx$$

$$\text{or } \psi(k) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\psi(x)e^{ikx}dx \text{ with } p_x = \hbar k$$

Example 1.2a Derive the position representation form of the operator p_x.

Solution. We begin by calculating the expectation value of p_x:

$$\langle p_x \rangle = \langle \hbar k \rangle = \hbar \langle \psi(k) | k | \psi(k) \rangle = \hbar \int_{-\infty}^{\infty} \psi^*(k) k \psi(k) dk$$

$$= \frac{\hbar}{2\pi} \int_{-\infty}^{\infty} \psi^*(x') \int_{-\infty}^{\infty} e^{ikx'} k \underbrace{\int_{-\infty}^{\infty} \psi(x) e^{-ikx} dx}_{\text{integration by parts}} \, dk dx'$$

$$= -\frac{i\hbar}{2\pi} \int_{-\infty}^{\infty} \psi^*(x') \int_{-\infty}^{\infty} e^{ikx'} \int_{-\infty}^{\infty} \frac{d\psi(x)}{dx} e^{-ikx} dx dk dx'$$

Next we will use the following result for Driac delta measure (δ):

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

Recall that δ is defined as:

$$\delta(x) = \begin{cases} \infty \text{ when } x = 0 \\ 0 \text{ when } x \neq 0 \end{cases} \text{ and } \int_{-\infty}^{\infty} \delta(x) dx = 1$$

Now we can continue working with the expectation value:

$$... = -i\hbar \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \psi^*(x') \frac{d\psi(x)}{dx} \delta(x' - x) dx dx'$$

$$= \int_{-\infty}^{\infty} \psi^*(x) \left(-i\hbar \frac{d}{dx} \right) \psi(x) dx = \int_{-\infty}^{\infty} \psi^*(x) p_x \psi(x) dx$$

This gives us the momentum operator:

$$p_x = -i\hbar \frac{d}{dx}$$

1.4 Commutation and non-commutation

When two operators operate successively on a given function, it is important to note the order in which they operate. In general, $(AB)g \neq (BA)g$ where A and B are operators and g is a function. Recall that the same holds for matrices where the order of multiplication may not usually be changed. A *commutator* gives the difference between operators AB and BA:

$$[A, B] = AB - BA \tag{1.7}$$

If $[A, B] = 0$ the operators commute and their operation order may be changed without further consideration. In general operators do not commute. When calculating commutators, one has to include a function that $(AB - BA)$ operates on and then in the end "cancel" out the function.

Example 1.3 Evaluate the commutator $[x, p_x]$ in the position representation.

Solution. Choosen an arbitrary function f. Note that we don't need to know its actual form. The commutator can be now evaluated as:

$$[x, p_x] = (xp_x - p_x x)\, f = x \times \frac{\hbar}{i}\frac{\partial f}{\partial x} - \frac{\hbar}{i}\frac{\partial (xf)}{\partial x}$$

$$= x \times \frac{\hbar}{i}\frac{\partial f}{\partial x} - \frac{\hbar}{i} f - x \times \frac{\hbar}{i}\frac{\partial f}{\partial x} = i\hbar f$$

$$\Rightarrow [x, p_x] = i\hbar$$

1.5 The construction of operators

Operators for other observables can be constructed from the operators for position (x) and momentum (p_x). The kinetic energy operator T can be constructed by using the classical analogy $T = \frac{p^2}{2m}$ where m is the mass of the particle. Here we have considered only one dimension and in this case the kinetic energy operator becomes:

$$T = \frac{p_x^2}{2m} = \frac{1}{2m}\left(\frac{\hbar}{i}\frac{d}{dx}\right)^2 = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2} \tag{1.8}$$

In three dimensions T is given by:

$$T = -\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right) \equiv -\frac{\hbar^2}{2m}\nabla^2 \equiv -\frac{\hbar^2}{2m}\Delta \tag{1.9}$$

The operator for potential energy is simply given by multiplication by a potential function $V(x)$. In three dimensions it typically depends on the distance between particles ($r = \sqrt{x^2 + y^2 + z^2}$). For example, in the position representation the Coulomb potential energy of an electron (e^-) in the field of a nucleus of atomic number Z is:

$$V(r) = -\frac{Ze^2}{4\pi\epsilon_0 r}\times \tag{1.10}$$

where r is the distance from the nucleus to the electron. Note that often the multiplication sign is omitted when specifying operators.

The operator for the total energy (called the Hamiltonian) is given as the sum of kinetic and potential parts:

$$H = T + V \tag{1.11}$$

For example, for particle moving in one dimension subject to potential V, H takes the following form:

$$H = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + V(x) \tag{1.12}$$

For an electron with mass m_e in hydrogen atom, the Hamiltonian takes the form:

$$H = -\frac{\hbar^2}{2m_e}\nabla^2 - \frac{e^2}{4\pi\epsilon_0 r} \tag{1.13}$$

The general prescription for constructing operators in the position representation is:

1. Write the classical expression for the observable in terms of position coordinates and the linear momentum.
2. Replace x by multiplication by x and replace p_x by $\frac{\hbar}{i}\frac{\partial}{\partial x}$ (likewise for other coordinates).

1.6 Integrals over operators

In our calculations we will frequently have to evaluate integrals of the following form:

$$I = \int f_m^* \Omega f_n d\tau \tag{1.14}$$

where '*' denotes a complex conjugate and $d\tau$ is the *volume element* (for example, in 3-D Cartesian coordinates $dxdydz$). The integral is take over the whole space, for 3-D Cartesian coodrinates: $x : -\infty \to \infty$, $y : -\infty \to \infty$, $z : -\infty \to \infty$.

When $\Omega = 1$, Eq. (1.14) turns into an *overlap integral*:

$$S = \int f_m^* f_n d\tau \tag{1.15}$$

When f_m and f_n are *orthogonal*, $S = 0$. If S is close to 1, the two functions are nearly "parallel". A special case of Eq. (1.15) with $m = n$ is the *normalization integral*:

$$\int f_m^* f_m d\tau \equiv \int |f_m|^2 \, d\tau = 1 \tag{1.16}$$

Given that $\int |g|^2 \, d\tau$ has a finite value, it is possible to multiply g by a constant and make it normalized.

Example 1.4 Function $f(x) = \sin(n\pi/L)$ between $x = 0$ and $x = L$ and zero elsewhere. Find the normalized form of this function ("normalize it").

Solution. Since the given function is non-zero only between $[0, L]$ we can restrict the integration in this interval:

$$\int_0^L |f(x)|^2\, dx = \int_0^L f^*(x) f(x) dx = \int_0^L \sin^2(\pi x/L)\, dx = \frac{1}{2} L$$

where we have used the following result from tablebook: $\int \sin^2(ax) dx = \frac{1}{2a}(ax - \sin(ax)\cos(ax))$. Here we can see that $f(x)$ must be divided by $\sqrt{\frac{1}{2}L}$ to make it normalized. Thus defining $g(x) = \sqrt{\frac{2}{L}} \times f(x) = \sqrt{\frac{2}{L}} \sin(\pi x/L)$ gives a normalized function.

A set of functions f_n that are normalized and mutually orthogonal are said to be *orthonormal*:

$$\int f_m^* f_n d\tau = \delta_{mn} \tag{1.17}$$

where δ_{mn} is the Kronecker delta, which is 1 if $m = n$ and 0 otherwise.

1.7 Dirac bracket notation

Since we will be mostly dealing with integrals such as in Eqs. (1.15) and (1.16), it is convenient to use short notation ("*Dirac bracket notation*"):

$$\langle m \left|\Omega\right| n \rangle = \int f_m^* \Omega f_n d\tau \tag{1.18}$$

The symbol $|n\rangle$ is called a *ket* and denotes the state described by function f_n. Symbol $\langle m|$ is *bra* and denotes the state f_m with complex conjugation. When they are combined, they will give $\langle m|n\rangle$ ("bracket") or with an operator between them $\langle m \left|\Omega\right| n\rangle$. The orthonormaization condition (Eq. (1.17) can be written using the Dirac notation as:

$$\langle m|n\rangle = \int f_m^* f_n d\tau \tag{1.19}$$

$$\langle m|n\rangle = \delta_{mn} \tag{1.20}$$

Complex conjugation of an overlap integral can be evaluated as follows:

$$\langle m|n\rangle = \int f_m^* f_n d\tau = \int (f_n^* f_m)^* d\tau = \left(\int f_n^* f_m d\tau \right)^* = \langle n|m\rangle^*$$

1.8 Hermitian operators

An operator is *hermitian* if it satisfies the following relation:

$$\int f_m^* \Omega f_n d\tau = \left(\int f_n^* \Omega f_m d\tau \right)^* \tag{1.21}$$

or alternatively:

$$\int f_m^* \Omega f_n d\tau = \int (\Omega f_m)^* f_n d\tau$$

By using the Dirac notation, we can write Eq. (1.21):

$$\langle m | \Omega | n \rangle = \langle n | \Omega | m \rangle^* \tag{1.22}$$

Example 1.5 Show that both position (x) and momentum (p_x) operators are hermitian.

Solution. Consider operator x first (note that x is real):

$$\int f_m^* x f_n d\tau = \int f_n x f_m^* d\tau = \left(\int f_n^* x f_m d\tau \right)^*$$

For momentum we have $p_x = \frac{\hbar}{i} \frac{d}{dx}$ and then integration by parts:

$$\int_a^b u'v = \Big/_a^b uv - \int_a^b uv'$$

gives:

$$\int_{-\infty}^{\infty} f_m^* p_x f_n dx = \int_{-\infty}^{\infty} f_m^* \frac{\hbar}{i} \frac{d}{dx} f_n dx = \frac{\hbar}{i} \Big/_{-\infty}^{\infty} f_m^* f_n - \frac{\hbar}{i} \int_{-\infty}^{\infty} f_n \frac{d}{dx} f_m^* dx = \ldots$$

Since both f_m and f_n must approach zero when x approaches ∞, we can simplify:

$$\ldots = -\frac{\hbar}{i} \int_{-\infty}^{\infty} f_n \frac{d}{dx} f_m^* dx = \left(\int_{-\infty}^{\infty} f_n^* \frac{\hbar}{i} \frac{d}{dx} f_m dx \right)^*$$

Thus we have shown that Eq. (1.21) holds and p_x is hermitean.

There are a number of important properties that hold for hermitian operators:

1. *The eigenvalues of hermitean operators are real.*

   Proof. Consider an eigenvalue equation: $\Omega |\omega\rangle = \omega |\omega\rangle$ and multiply it from the left by $\langle\omega|$: $\langle\omega |\Omega| \omega\rangle = \omega \langle\omega|\omega\rangle = \omega$ where $\langle\omega|\omega\rangle = 1$ (normalization). Complex conjugating both sides: $\langle\omega |\Omega| \omega\rangle^* = \omega^*$ By hermiticity we have $\langle\omega |\Omega| \omega\rangle^* = \langle\omega |\Omega| \omega\rangle$. The two above equations now yield $\omega = \omega^*$. This implies that ω is real.

2. Eigenfunctions corresponding to *different* eigenvalues of an hermitian operator are orthogonal:

$\langle f_m | f_n \rangle = \delta_{mn}$ where f_m and f_n belong to different eigenvalues (non-degenerate)

Proof. Choose two different eigenfunctions $|\omega\rangle$ and $|\omega'\rangle$ that satisfy:

$$\Omega \, |\omega\rangle = \omega \, |\omega\rangle \;\text{ and }\; \Omega \, |\omega'\rangle = \omega' \, |\omega'\rangle$$

Multiplication side by side by ω and ω' gives:

$$\langle \omega' | \, \Omega \, |\omega\rangle = \omega \, \langle \omega' | \omega \rangle \;\text{ and }\; \langle \omega | \, \Omega \, |\omega'\rangle = \omega' \, \langle \omega | \omega' \rangle$$

Taking complex conjugate of both sides of the 2nd relation above and subtracting it from the first we get:

$$\langle \omega' | \, \Omega \, |\omega\rangle - \langle \omega | \, \Omega \, |\omega'\rangle^* = \omega \, \langle \omega' | \omega \rangle - \omega' \, \langle \omega | \omega' \rangle^*$$

Since Ω is Hermitian, the left side of the above expression is zero. Since $\langle \omega | \omega \rangle$ is real and $\langle \omega' | \omega \rangle = \langle \omega | \omega' \rangle$ we have:

$$(\omega - \omega') \, \langle \omega' | \omega \rangle = 0$$

Since we have non-degenerate situation, $\omega \neq \omega'$ and hence $\langle \omega' | \omega \rangle = 0$. For example, eigenfunctions of "Harmonic oscillator" are orthogonal. Note that this result does not apply to degenerate states.

1.9 States and wavefunctions

The word postulate basically means assumption. If we make certain assumptions then quantum mechanics is a full theory in a sense that everything can be derived from them.

Postulate #1. The state of a system is fully described by a function $\Psi(r_1, r_2, ..., r_n, t)$ ("wavefunction") where n is the number of particles in the system, r_i are their spatial coordinates and t is the time. In addition to space and time, sometimes wavefunctions depend also on other degrees of freedom such as spin.

Wavefunctions are often expressed as a linear combination of eigenfunctions and gien such representation the associated quantum numbers may be used in specifying the wavefunction uniquely. If the wavefunction is known then all properties of the system may be obtained from it. For example, the Hamiltonian operator will give the energy of the system (as either expectation value or eigenvalue as we will see later).

1.10 The fundamental prescription

Postulate #2. Observables in quantum mechanics are represented by hermitian operators chosen to satisfy the following commutation relations:

$$\left[q, p_{q'}\right] = i\hbar\delta_{qq'} \text{ with } \left[q, q'\right] = 0 \text{ and } \left[p_q, p_{q'}\right] = 0$$

where q and q' each denote one of the coordinates x, y, z and p_q and $p_{q'}$ the corresponding linear momenta. Note that since these operators are hermitian their eigenvalues are real and that x, y, z commute with each other as well as p_x, p_y, p_z commute with each other. Previously we used the Fourier duality to derive the exact form for the momentum operator.

Example. Show that the operators p_x and p_y fulfill the above requirement.

Solution. First we note that $p_x = \frac{\hbar}{i}\frac{\partial}{\partial x}$ and $p_y = \frac{\hbar}{i}\frac{\partial}{\partial y}$. Trivially the constants can be exchanged in the commutator and we only need to worry about the derivatives. For any well behaved function we have: $\frac{\partial^2}{\partial x \partial y} f(x, y) = \frac{\partial^2}{\partial y \partial x} f(x, y)$ and hence the operators commute. In general, if two operators depend on different variables, they commute.

1.11 The outcome of measurements

The following postulate establishes the link between the wavefunctions & operators and experimental observations:

Postulate #3. When a system is described by a wavefunction $\psi(r)$ (note time independent wavefunction) we obtain the *expectation value* of operator Ω as:

$$\langle \Omega \rangle = \frac{\int \psi^* \Omega \psi d\tau}{\int \psi^* \psi d\tau} = \frac{\langle \psi | \, \Omega \, | \psi \rangle}{\langle \psi | \psi \rangle} \tag{1.23}$$

If the wavefunction ψ is normalized:

$$\langle \Omega \rangle = \int \psi^* \Omega \psi d\tau = \langle \psi | \, \Omega \, | \psi \rangle \tag{1.24}$$

In general we assume that the we wavefunctions that we work with are normalized. Note that calculation of expectation value of an eigenfunction gives the corresponding eigenvalue ($\Omega \, |\psi\rangle = \omega \, |\psi\rangle$):

$$\langle \Omega \rangle = \int \psi^* \Omega \psi d\tau = \int \psi^* \omega \psi d\tau = \omega \int \psi^* \psi d\tau = \omega \tag{1.25}$$

Recall that a general wavefunction can be always represented as a linear combination of eigenfunctions:

$$\psi = \sum_{n=1}^{\infty} c_n \psi_n \text{ where } \Omega \psi_n = \omega_n \psi_n$$

The expectation value of this wavefunction can be calculated as:

$$\langle \Omega \rangle = \sum_{n=1}^{\infty} c_n^* c_n \omega_n \int \psi_n^* \psi_n d\tau = \sum_{n=1}^{\infty} c_n^* c_n \omega_n = \sum_{n=1}^{\infty} |c_n|^2 \omega_n \tag{1.26}$$

where we have used the fact that $\int \psi_n^* \psi_m d\tau = \delta_{mn}$. The expectation value is therefore a weighted average of over the eigen states of Ω that contribute to ψ. This can be cast into an additional postulate:

#3' When ψ is an eigenfunction of the operator Ω, the determination of the corresponding property Ω always yields one result, namely the corresponding eigenvalue ω. The expectation value will simply be the eigenvalue ω. When ψ is not an eigenfunction of Ω, a single measurement of the property yields a single outcome which is one of the eigenvalues of Ω, and the probability that a particular eigenvalue ω_n is measured is equal to $|c_n|^2$, where c_n is the coefficient of the eigenfunction ψ_n in the expansion of the wavefunction.

In practice, the above postulate means that multiple measurements on the system produce a distribution of outcomes and the distribution is determined by the coefficients, $|c_n|^2$.

Example 1.6 An operator A has eigenfunctions $f_1, f_2, ..., f_n$ with corresponding eigenvalues $a_1, a_2, ..., a_n$. The state of the system is described by a normalized wavefunction ψ given by:

$$\psi = \frac{1}{2} f_1 - \left(\frac{3}{8}\right)^{1/2} f_2 + \left(\frac{3i}{8}\right)^{1/2} f_3$$

What will be the outcome of measuring the observable A?

Solution. We need to evaluate the expectation value of ψ:

$$\langle A \rangle = \left\langle \left(\frac{1}{2} f_1 - \left(\frac{3}{8}\right)^{1/2} f_2 + \left(\frac{3i}{8}\right)^{1/2} f_3 \right) \middle| A \middle| \left(\frac{1}{2} f_1 - \left(\frac{3}{8}\right)^{1/2} f_2 + \left(\frac{3i}{8}\right)^{1/2} f_3 \right) \right\rangle$$

$$= \frac{1}{4} \langle f_1 | A | f_1 \rangle + \frac{3}{8} \langle f_2 | A | f_2 \rangle + \frac{3}{8} \langle f_3 | A | f_3 \rangle$$

$$= \frac{1}{4} a_1 + \frac{3}{8} a_2 + \frac{3}{8} a_3$$

1.12 The interpretation of the wavefunction

The wavefunction is given the *probability interpretation* (also called the *Born interpretation*):

Postulate #4. For normalized wavefunctions $\psi(r)$, the probability that a particle will be found in the volume element $d\tau$ at the point r is equal to $|\psi(r)|^2\, d\tau$.

In one dimension the volume element $d\tau$ is dx and in three dimensions $dxdydz$ (Cartesian coodinates) or $r^2 \sin(\theta) dr d\theta d\phi$ (spherical coordinates; $r : 0 \to \infty, \theta : 0 \to \pi, \phi : 0 \to 2\pi$). According to the Born interpretation $|\psi(r)|^2\, d\tau$ gives the probability of the particle to occur within the volume element $d\tau$. The *wavefunction itself has no direct physical* meaning since experiments can only determine quantities that depend on the square of the wavefunction. Note that the wavefunction may be a complex valued function whereas $|\psi(r)|^2$ is always real and non-negative.

Note that any wavefunction that is square integrable (i.e. $\psi \in L^2$) can be normalized and given the probability interpretation. This also implies that $|\psi(r)| \to 0$ when $|r| \to \infty$. This will be especially useful result when applying the Green's formula:

$$\int_\Omega \frac{\partial v}{\partial x_j} w dx + \int_\Omega v \frac{\partial w}{\partial x_j} dx = \int_{\partial\Omega} v w n_j ds$$

where v and w are functions that depend on variables $x_1, x_2, ..., x_j, ..., x_N$ and n_j are the components of the outward normal to $\partial\Omega$. If v and w are such that they appraoch zero at the boudary, the left hand side term is zero.

1.13 The equation for the wavefunction

The last postulate concerns the dynamical evolution of the wavefunction.

Postulate #5. The wavefunction $\Psi(r_1, r_2, ..., t)$ evolves in time according with time:

$$i\hbar \frac{\partial \Psi}{\partial t} = H\Psi \tag{1.27}$$

This partial differential equation is called the **Schrödinger equation**, which was first proposed by Erwin Schrödinger (1926). The operator H above is the Hamiltonian operator, which gives "the total energy $\times$ ψ" after operating on ψ. Note that the time-independent Schrödinger equation can be derived from Eq. (1.27). This will be shown in the next section.

In one dimension the time-dependent Schrödinger equation can be written as:

$$i\hbar \frac{\partial \Psi}{\partial t} = -\frac{\hbar^2}{2m} \frac{\partial^2 \Psi}{\partial x^2} + V(x)\Psi \tag{1.28}$$

1.14 The separation of the Schrödinger equation

When the Hamiltonian operator H does not depend on time, Eq. (1.27) can be separated into two equations: 1) The time-independent Schrödinger equation and 2) an auxiliary equation defining the phase evolution of the time-dependent wavefunction. Since most often $H = T + V$ ("kinetic + potential"), it is sufficient to have the potential function V time-independent. In the following we will demonstrate how the separation can be carried out in one dimension:

$$H\Psi = -\frac{\hbar^2}{2m}\frac{\partial^2\Psi}{\partial x^2} + V(x)\Psi = i\hbar\frac{\partial\Psi}{\partial t}$$

Apart from the wavefunction itself the left hand side depends only on the spatial coordinate x and the right hand side only on time t. In such case one can write the solution in the following form: $\Psi(x,t) = \psi(x)\theta(t)$ where ψ and θ are function that depend only on x and t respectively. After this substitution, we get:

$$-\frac{\hbar^2}{2m}\theta\frac{d^2\psi}{dx^2} + V(x)\psi\theta = i\hbar\psi\frac{d\theta}{dt}$$

By dividing both sides by $\psi\theta$, we have:

$$-\frac{\hbar^2}{2m}\frac{1}{\psi}\frac{d^2\psi}{dx^2} + V(x) = i\hbar\frac{1}{\theta}\frac{d\theta}{dt}$$

The left hand side depends only on ψ and the right hand side only on θ. Since they can be varied independently and the equation still must hold, both sides of the equation must be constant. This constant we denote by E and call it "energy":

$$-\frac{\hbar^2}{2m}\frac{1}{\psi}\frac{d^2\psi}{dx^2} + V(x) = E$$

By multiplying both sides by ψ we finally have:

$$-\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2} + V(x)\psi = E\psi \tag{1.29}$$

For short we can write this as $H\psi = E\psi$ (*the time-independent Schrödinger equation*). In a similar way we obtain another equation for θ:

$$i\hbar\frac{d\theta}{dt} = E\theta$$

which has a solution:

$$\theta \propto e^{-iEt/\hbar} \tag{1.30}$$

This results in phase spinning in the complex plane with frequency dictated by E:

$$\Psi(x,t) = \psi(x)e^{-iEt/\hbar} \tag{1.31}$$

The higher the energy, the faster the phase spinning. It is helpful to remeber the Euler's relation: $e^{ix} = \cos(x) + \sin(x)$ with the parity rules: $\sin(-x) = -\sin(x)$ and $\cos(-x) = \cos(x)$. Thus the overall wavefunction Ψ will have both ψ and θ dependent parts in it with the first giving the spatial behavior and the latter the temporal behavior. Note that $|\Psi|^2$ does not depend on time and hence these states are called *stationary*.

1.15 Simultaneous observables

If a system is in one of the eigenstates of operator A then is it possible to simultaneously determine another property which is expressed by operator B? For example, if we know the momentum of the particle exactly then is it possible to measure the position exactly? It turns out that some times it is possible to measure both A and B at the same time and some times not.

Next we will prove the following result:

Property #3. Two operators A and B have precisely defined observables $\Leftrightarrow$ $[A, B] = 0$ (i.e. the operators must commute).

Proof. "$\Rightarrow$" First we note that in order to precisely define the outcome from both A and B, they must have share the same eigenfunctions. Thus: $A|\psi\rangle = a|\psi\rangle$ and $B|\psi\rangle = b|\psi\rangle$. Thus we can write:

$$AB|\psi\rangle = Ab|\psi\rangle = bA|\psi\rangle = ba|\psi\rangle = ab|\psi\rangle = aB|\psi\rangle = Ba|\psi\rangle = BA|\psi\rangle$$

"$\Leftarrow$" We need to show that given that $A|\psi\rangle = a|\psi\rangle$ and $[A, B] = 0$, we have $B|\psi\rangle = b|\psi\rangle$. Because we have $A|\psi\rangle = a|\psi\rangle$, we can write:

$$BA|\psi\rangle = Ba|\psi\rangle = aB|\psi\rangle$$

Because A and B commute, the first term can also be written as $AB|\psi\rangle$ and hence:

$$A(B|\psi\rangle) = a(B|\psi\rangle)$$

This has the same form as the eigenvalue equation for A and therefore $B|\psi\rangle$ must be proportional to $|\psi\rangle$. We denote this proportionality constant by b and then we get the result we were looking for: $B|\psi\rangle = b|\psi\rangle$.

In order to determine if two observables can be determined simultaneously with arbitrarily high precision, one must inspect the commutator between the corresponding operators.

Example. Is it possible to determine both position x and momentum p_x (i.e. momentum along the x-axis) simultaneously? How about x and p_y?

Solution. We have already calculated the commutator $[x, p_x]$ in Example 1.3 and noticed that it gives a non-zero result. Hence operators p_x and x cannot be determined simultaneously with arbitrarily high precision. On the other hand x and p_y commute and they can be determined simultaneously with arbitrarily high precision.

Pairs of observables that cannot be determined simultaneously are said to be *complementary*.

1.16 The uncertainty principle

As we saw, if two operators do not commute, it is not possible to specify their eigenvalues of the operators simultaneously. However, it is possible to give up precision in one of the observables to acquire greater precision in the other. For example, if we have unertainty of Δx in position x and Δp_x in momentum p_x, we can show that the following relation holds:

$$\Delta x \Delta p_x \geq \frac{1}{2}\hbar \tag{1.32}$$

This states that if Δx increases (i.e. greater uncertainty) then we can have smaller Δp_x (i.e. greater accuracy in momentum). This result was first presented by Heisenberg (1927). In general, for operators A and B with uncertainties ΔA and ΔB, respectively, we have:

$$\Delta A \Delta B \geq \frac{1}{2}\left|\langle [A, B] \rangle\right| \tag{1.33}$$

where the uncertainties of A (or B) are defined as:

$$\Delta A = \left\{ \left\langle A^2 \right\rangle - \langle A \rangle^2 \right\}^{1/2} \tag{1.34}$$

Proof. Let A and B be operators and choose a wavefunction ψ that is not necessarily an eigenfunction of A or B. We will optimize the following non-negative integral with respect to scaling factor α to yield the minimum combined error:

$$I = \int \psi^* \left|(\alpha\delta A + i\delta B)\right|^2 \psi d\tau$$

The scaling factor α acts to reduce the error in A while the whole integral will give the combined error of both A and B. Note that the contribution of δB is included in the imaginary part as we want to be sure not have cancellation of the δA and δB contributions by sign. By squaring the whole integrand, we ensure that we get contributions from both errors added up as positive numbers.

To simplify the calculation, we define the expectation values of A and B as:

$$\langle A \rangle = \langle \psi \left| A \right| \psi \rangle \text{ and } \langle B \rangle = \langle \psi \left| B \right| \psi \rangle$$

and furthermore deviation of each operator around its expectation value by:

$$\delta A = A - \langle A \rangle \text{ and } \delta B = B - \langle B \rangle$$

A direct calculation gives the following result (*):

$$[\delta A, \delta B] = [A - \langle A \rangle, B - \langle B \rangle] = [A, B] \equiv iC$$

Next we rewrite I as follows:

$$I = \int \psi^* \left(\alpha\delta A + i\delta B\right)^* \left(\alpha\delta A + i\delta B\right) \psi d\tau$$

$$= \int \psi^* \left(\alpha \delta A - i \delta B\right) \left(\alpha \delta A + i \delta B\right) \psi d\tau$$

In the last step we used the fact that the operators are hermitian. In the Dirac notation this can be written as:

$$I = \langle \psi | \left(\alpha \delta A - i \delta B\right) \left(\alpha \delta A + i \delta B\right) | \psi \rangle$$

This can be expanded as follows (see the result marked with (*) above for substituting in C):

$$I = \alpha^2 \left\langle (\delta A)^2 \right\rangle + \left\langle (\delta B)^2 \right\rangle + i\alpha \left\langle \delta A \delta B - \delta B \delta A \right\rangle = \alpha^2 \left\langle (\delta A)^2 \right\rangle + \left\langle (\delta B)^2 \right\rangle + \alpha \left\langle C \right\rangle$$

Since we want to minimize I with respect to α, we reorganize the above expression:

$$I = \left\langle (\delta A)^2 \right\rangle \left(\alpha + \frac{\langle C \rangle}{2 \left\langle (\delta A)^2 \right\rangle} \right)^2 + \left\langle (\delta B)^2 \right\rangle - \frac{\langle C \rangle^2}{4 \left\langle (\delta A)^2 \right\rangle}$$

Clearly the minimum value for I is reached when α is chosen such that the first term above is zero. At this point I takes the value:

$$I = \left\langle (\delta B)^2 \right\rangle - \frac{\langle C \rangle^2}{4 \left\langle (\delta A)^2 \right\rangle} \geq 0$$

This can be rearranged as:

$$\left\langle (\delta A)^2 \right\rangle \left\langle (\delta B)^2 \right\rangle \geq \frac{1}{4} \langle C \rangle^2$$

The left side of the equation can be simplified by using:

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

$$= \left\langle A^2 \right\rangle - 2 \langle A \rangle \langle A \rangle + \langle A \rangle^2 = \left\langle A^2 \right\rangle - \langle A \rangle^2 = \Delta A^2$$

By doing the same operation on B and substituting in ΔA and ΔB we arrive at:

$$\Delta A^2 \Delta B^2 \geq \frac{1}{4} \langle C \rangle^2$$

Taking square root of both sides yields the uncertainty principle (recall that $[A, B] = iC$):

$$\Delta A \Delta B \geq \frac{1}{2} |\langle C \rangle| = \frac{1}{2} |\langle [A, B] \rangle|$$

1.17 Consequences of the uncertainty principle

It is instructive to see how Eq. (1.33) applies to position and momentum.

Example 1.8 Consider a particle prepared in a state given by wavefunction $\psi = Ne^{-x^2/2\Gamma}$ (Gaussian function) where $N = (\pi\Gamma)^{-1/4}$. Evaluate Δx and Δp_x and confirm that the uncertainty principle is satisfied.

Solution. We must calculate the following expectation values for Eq. (1.33): $\langle x \rangle$, $\langle x^2 \rangle$, $\langle p_x \rangle$ and $\langle p_x^2 \rangle$.

1. $\langle x \rangle = 0$ because x is an antisymmetric function with respect to origin and the square of the given Gaussian function is symmetric. Product of symmetric and antisymmetric functions is always antisymmetric. Integration of antisymmetric function gives zero.
2. $\langle p_x \rangle = 0$ because differentiation of symmetric function gives antisymmetric function. When this is multilpied by the symmetric wavefunction, the result is antisymmetric function. Hence the integral is zero.
3. The following integrals from tablebook will be useful for the remaining integrals: $\int_{-\infty}^{\infty} e^{-ax^2} dx = \left(\frac{\pi}{a}\right)^{1/2}$ and $\int_{-\infty}^{\infty} x^2 e^{-ax^2} dx = \frac{1}{2a}\left(\frac{\pi}{a}\right)^{1/2}$. For $\langle x^2 \rangle$ this gives:

$$\langle x^2 \rangle = N^2 \int_{-\infty}^{\infty} x^2 e^{-x^2/\Gamma} dx = \frac{1}{2}\Gamma$$

4. For $\langle p_x^2 \rangle$ we have:

$$\langle p_x^2 \rangle = N^2 \int_{-\infty}^{\infty} \exp\left(-\frac{x^2}{2\Gamma}\right)\left(-\hbar^2 \frac{d^2}{dx^2}\right)\exp\left(-\frac{x^2}{2\Gamma}\right) dx$$

$$= \hbar^2 N^2 \left\{ \frac{1}{\Gamma}\int_{-\infty}^{\infty} e^{-x^2/\Gamma} dx - \frac{1}{\Gamma}\int_{-\infty}^{\infty} x^2 e^{-x^2/\Gamma} dx \right\} = \frac{\hbar^2}{2\Gamma}$$

Now it follows that $\Delta x = \sqrt{\langle x^2 \rangle - \langle x \rangle^2} = \sqrt{\Gamma/2}$ and $\Delta p = \sqrt{\langle p^2 \rangle - \langle p \rangle^2} = \sqrt{\frac{\hbar^2}{2\Gamma}}$. This gives $\Delta x \Delta p = \sqrt{\Gamma/2} \times \sqrt{\frac{\hbar^2}{2\Gamma}} = \frac{\hbar}{2}$. Thus for this Gaussian wavefunction appears to be "optimal" in a sense that it gives the best accuracy for the uncertainty principle.

This problem can also be solved using the Fourier dualism between the position and momentum spaces. Exercise: Show that by Fourier transforming $\psi(x)$ into $\psi(k)$ one gets another Gaussian. Then take its width as Δp, the width of the original Gaussian as Δx, and finally calculate $\Delta x \Delta p$.

1.18 The uncertainty in energy and time

Often the uncertainty between energy and time is expressed as:

$$\Delta E \Delta t \geq \hbar$$

However, time is not an observable in nonrelativistic quantum mechanics but just a parameter with no corresponding operator. For this reason Eq. (1.33) cannot be applied. We will see later what the meaning of this "uncertainty" relation is.

Note that this will have important implications to spectroscopy and especially broadening of resonances. For example, consider a simple UV/VIS absorption experiment. When the molecule is promoted to the excited state, there could be some external perturbation that disturbs the excited state and hence "shortens" its lifetime. This would result in line broadening in UV/VIS spectrum.

1.19 Time-evolution and conservation laws

In addition to providing information about simultaneous exact measurement of observables, a commutator between two operators also plays an important role in determining the time-evolution of the expectation values. When H is the Hamiltonian operator and operator Ω corresponding to some observable *does not depend on time*:

$$\frac{d\langle\Omega\rangle}{dt} = \frac{i}{\hbar}\langle[H,\Omega]\rangle \tag{1.35}$$

It is important to notice that when $[H,\Omega] = 0$, the expectation value does not depend on time.

Proof. Differentiation of Ω with respect to time gives:

$$\frac{d\langle\Omega\rangle}{dt} = \frac{d}{dt}\langle\Psi|\Omega|\Psi\rangle = \int\left(\frac{\partial\Psi^*}{\partial t}\right)\Omega\Psi d\tau + \int\Psi^*\Omega\left(\frac{\partial\Psi}{\partial t}\right)d\tau$$

Note that Ω does not depend on time whereas Ψ and Ψ^* do. Next we apply the time-dependent Schrödinger equation (Eq. (1.27)):

$$\int\Psi^*\Omega\left(\frac{\partial\Psi}{\partial t}\right)d\tau = \int\Psi^*\Omega\left(\frac{1}{i\hbar}\right)H\Psi d\tau = \frac{1}{i\hbar}\int\Psi^*\Omega H\Psi d\tau$$

For the other therm we have (note that H is hermitiean, see Eq. (1.21)):

$$\int \left(\frac{\partial \Psi^*}{\partial t}\right) \Omega \Psi d\tau = -\int \left(\frac{1}{i\hbar}\right) (H\Psi)^* \, \Omega \Psi d\tau = -\frac{1}{i\hbar} \int \Psi^* H \Omega \Psi d\tau$$

By combining these expressions we get the final result:

$$\frac{d\langle \Omega \rangle}{dt} = -\frac{1}{i\hbar} \left(\langle H\Omega \rangle - \langle \Omega H \rangle \right) = \frac{i}{\hbar} \langle [H, \Omega] \rangle$$

Example. Calculate the expectation value of linear momentum as a function of time for a particle in one-dimensional system. The total Hamiltonian is $H = T + V$.

Solution. The commutator between H and p_x is:

$$[H, p_x] = \left[-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + V, \frac{\hbar}{i} \frac{d}{dx} \right] = \frac{\hbar}{i} \left[V, \frac{d}{dx} \right]$$

To work out the remaining commutator, we need to writen the wavefunction that we operate on:

$$[H, p_x] = \frac{\hbar}{i} \left\{ V \frac{d\psi}{dx} - \frac{d(V\psi)}{dx} \right\} = \frac{\hbar}{i} \left\{ V \frac{d\psi}{dx} - V \frac{d\psi}{dx} - \frac{dV}{dx} \psi \right\} = -\frac{\hbar}{i} \frac{dV}{dx} \psi$$

This holds for all ψ and hence:

$$[H, p_x] = -\frac{\hbar}{i}\frac{dV}{dx} \tag{1.36}$$

Eq. (1.35) can now be written as:

$$\frac{d}{dt}\langle p_x \rangle = \frac{i}{\hbar}\langle [H, p_x] \rangle = -\left\langle \frac{dV}{dx} \right\rangle \tag{1.37}$$

Here we note that force is given by $F = -dV/dx$ and we can rewrite the above equation as:

$$\frac{d}{dt}\langle p_x \rangle = \langle F \rangle \tag{1.38}$$

This states that the rate of change of the expectation value of the linear momentum is equal to the expectation value of the force. In a similar way one can show that:

$$\frac{d}{dt}\langle x \rangle = \frac{\langle p_x \rangle}{m} \tag{1.39}$$

Eqs. (1.38) and (1.39) consitute so called *Ehrenfest's theorem*. This states that classical mechanics deals with expectation values (i.e. quantum mechanical averages).

1.20 Matrices in quantum mechanics: Matrix elements

If a complete set of functions is specified ("basis set"), then we can write operators in matrix form, provided that the number of functions in the set is finite. In case of infinite basis set, the corresponding matrices would have infinite dimensions. Recall that matrix multiplication follows the rule:

$$P_{rc} = \sum_s M_{rs} N_{sc} \tag{1.40}$$

and especially note that in general $MN \neq NM$. In other words, they do not commute.

Heisenberg formulated his version of quantum mechanics by using matrices rather than operators. The two formulations are equivalent with the exception on infinite basis set expansions where the matrix formulation becomes problematic. In many cases, we will use integrals of the form $\langle m \left| \omega \right| n \rangle$, which we will often denote just by Ω_{mn}. When m and n are members of the specified basis set, Ω_{mn} is called a *matrix element* of operator Ω.

We will also often encounter sums of the following form:

$$\sum_s \langle r \left| A \right| s \rangle \langle s \left| B \right| c \rangle$$

By using the matrix notation, we can write this as:

$$\sum_s A_{rs} B_{sc} = (AB)_{rc} = \langle r \left| AB \right| c \rangle$$

where $(AB)_{rc}$ denotes the matrix element of AB, which corresponds to operator AB. In the last step we have used the *completeness relation* (also known as the *closure relation*), which states that:

$$\sum_s |s\rangle \langle s| = 1$$

Example 1.9 Use the completeness relation to show that the expectation value of the square of an hermitian operator is non-negative.

Solution. We calculate the expectation value:

$$\left\langle \omega \left| \Omega^2 \right| \omega \right\rangle = \langle \omega \left| \Omega\Omega \right| \omega \rangle = \sum_s \langle \omega \left| \Omega \right| s \rangle \langle s \left| \Omega \right| \omega \rangle = \sum_s \langle \omega \left| \Omega \right| s \rangle \langle \omega \left| \Omega \right| s \rangle^*$$

$$= \sum_s \underbrace{\left| \langle \omega \left| \Omega \right| s \rangle \right|^2}_{\geq 0} \geq 0$$

1.21 The diagonalization of the Hamiltonian

The time-independent Schrödinger equation ($H\psi = E\psi$) can be written it matrix form (given a suitable basis set expansion; not eigenfunctions of H). Consider first

$$\underline{H\left|\psi\right\rangle} = H\sum_n c_n \left|n\right\rangle = E\sum_n c_n \left|n\right\rangle = \underline{E\left|\psi\right\rangle}$$

If this is multiplied by $\langle m|$ side by side, we get:

$$\sum_n c_n \left\langle m\right|H\left|n\right\rangle = E\sum_n c_n \left\langle m|n\right\rangle$$

By denoting matrix $H_{mn} = \left\langle m\right|H\left|n\right\rangle$ and vector $c = c_m$ we have:

$$Hc = Ec \text{ or } \sum_n H_{mn}c_n = Ec_m \text{ for each } m \tag{1.41}$$

This is the matrix form of the Schrödinger equation and it is extremely useful when considering a system with only few basis functions or numerical solution to Schrödinger equation. If one can find a basis set such that H becomes a diagonal matrix then we have:

$$H_{mm}c_m = Ec_m \tag{1.42}$$

This states that each diagonal element H_{mm} is equal to E. If this holds for all m, then we have all eigenvalues of H arranged on the diagonal of matrix H. Note that we have used H for both the operator and the matrix which is somewhat confusing.

Example. Diagonalizing matrices using the Maxima program (you may also consider wxmaxima, which is graphical user interface to maxima). Consider the following matrix:

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \tag{1.43}$$

To diagonalize this matrix with maxima, enter the following commands (%i corresponds to input and %o to output):

```
(%i1) m:matrix([1, 2, 3], [4, 5, 6], [7, 8, 9]);
                                  [ 1  2  3 ]
                                  [         ]
(%o1)                             [ 4  5  6 ]
                                  [         ]
                                  [ 7  8  9 ]
(%i2) eigenvalues(m);
                  3 sqrt(33) - 15  3 sqrt(33) + 15
(%o3)        [[- ---------------, ---------------, 0], [1, 1, 1]]
                         2                2
```

The first three numbers are the eigenvalues and the following vector ($[1, 1, 1]$) states that the degeneracy factor of each of these eigenvalues is one.

```
(%i4) eigenvectors(m);
           3 sqrt(33) - 15  3 sqrt(33) + 15
(%o4) [[[- ---------------, ---------------, 0], [1, 1, 1]],
                   2                2
      3 sqrt(33) - 19    3 sqrt(3) sqrt(11) - 11
[1, - ---------------, - -----------------------],
            16                       8
    3 sqrt(33) + 19  3 sqrt(3) sqrt(11) + 11
[1, ---------------, -----------------------], [1, - 2, 1]]
          16                    8
```

The first vector in the output is the eigenvalues followed by the degeneracies (just like with the eigenvalues command). The three vectors after these are the corresponding eigenvectors. These could be converted into a wavefunction by multiplying the vector components by the corresponding basis functions. Also note that in this case Maxima was able to find exact solution rather than approximate one.

Chapter 2: Linear motion and the harmonic oscillator

The characteristics of acceptable wavefunctions

The Born interpretation implies that wavefunctions must be square integrable and should be made to satisfy the normalization condition:

$$\int \psi^* \psi d\tau = 1 \tag{2.1}$$

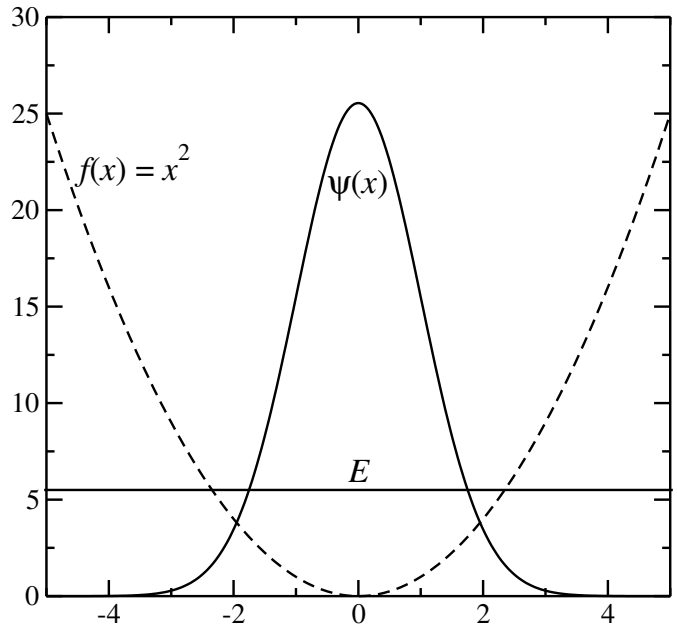
This implies that the wavefunction cannot become infinte over a finte region in space. It can, however, be infinite in some inifinitesimal region in space (zero measure set), which would not contribute to the value of the integral. Furthermore, ψ must be a function (i.e. it may only have a single value at a given point in space) and have continuous first derivative except in some infinitesimal regions in space.

2.1 & 2.2 The curvature of the wavefunction & Qualitative solutions

The Schrödinger equation can be used to obtain curvature of ψ:

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

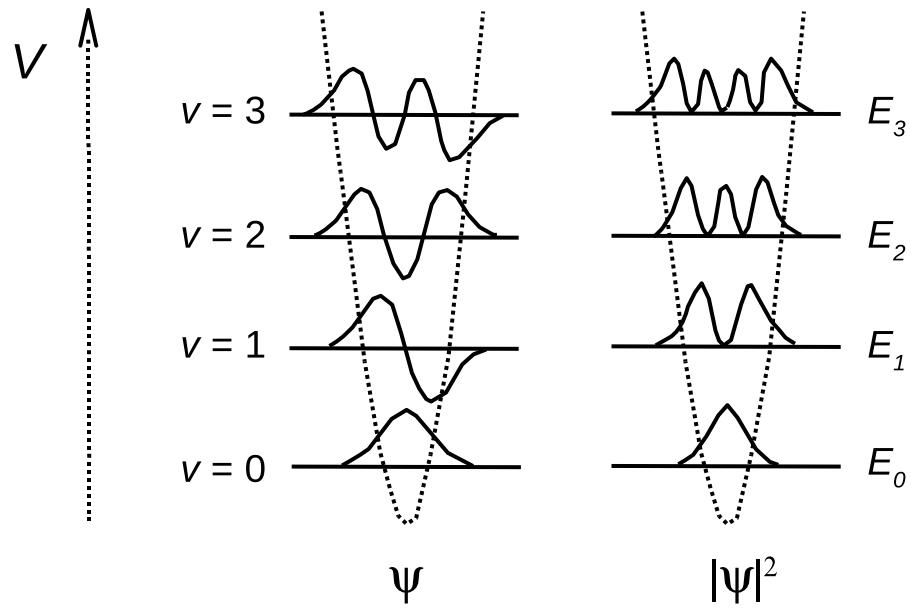
If $(V - E)\psi < 0$ the curvature goes like $\cap$ or $(V - E)\psi > 0$ then $\cup$. Consider, for example, the following:



The curvature is large where the amplitude of ψ is large. The solutions must be such that they tend to zero far away from the origin.

2.3 The emergence of quantization

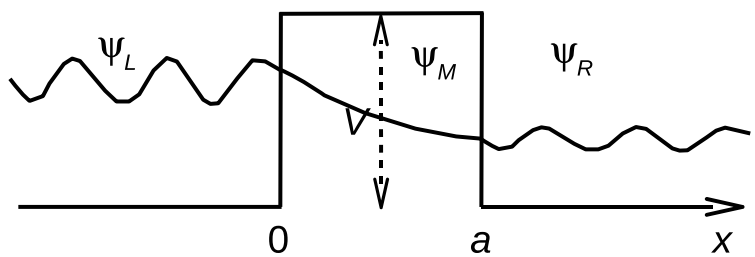
For systems where the external potential V varies sufficiently strongly as a function of the spatial coordinates, the energy may be *quantitized*. A similar quantitization may also follow from the *boundary conditions* such as the *Dirichlet* ($\psi(r) = 0$ at the boundary; corresponds to infinite potential at boundary), *Neumann* ($\partial\psi/\partial r = 0$ at the boundary; corresponds to constant and continuous potential at the boundary), or *cyclic boundary* condition (e.g. molecular rotation). Note that without these conditions the energy will not be quantitized (i.e. continuous spectrum).



Note that if the potential well has finite depth, it is possible to have both discrete eigenstates at low energies and then at higher energies a continuum of states.

2.4 Penetration into non-classical regions

On previous slide we have seen that the probability amplitude $|\psi(r)|^2$ asymptotically approaces zero when we go far away from the origin. Thus it is necessary that at some point the potential energy will be larger than the energy of the state that the particle is on. This region is forbidden classically as the total energy of the particle is less than the potential energy. It is often said that the particle *penetrates* or *tunnels* into the potential barrier. Most often when referring to quantum mechanical tunneling, we are discussing non-classical extent of the particle probability amplitude on the opposite side of a potential barrier:



2.5 Energy and momentum of translational motion

The easiest type of motion to consider is that of a completely free particle (i.e. no external potential) travelling in unbounded one dimensional region. Since there is no potential, $V(x) \equiv 0$ and the total Hamiltonian is then:

$$H = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2} \tag{2.2}$$

and the time-independent Schrödinger equation is then:

$$-\frac{\hbar^2}{2m}\frac{d^2\psi(x)}{dx^2} = E\psi(x) \tag{2.3}$$

The general solutions to this equation are:

$$\psi(x) = Ae^{ikx} + Be^{-ikx} \text{ and } k = \left(\frac{2mE}{\hbar^2}\right)^{1/2} \tag{2.4}$$

Recall that $e^{ikx} \equiv \cos(kx) + i\sin(kx)$. Note that these functions cannot be normalized over the unbounded region as $|\psi(x)|^2 = 1$. If the region would be finite then it would be possible to carry out the normalization. In any case, the probability amplitude of ψ states that there will be equal probability to find the particle anywhere in the region.

Since we can choose any value for k in Eq. (2.4), there is no quantitization in this system. If one operates on Eq. (2.4) with the Hamiltonian of Eq. (2.2) or the momentum operator in Eq. (1.5), we get:

$$p = k\hbar \text{ and } E = \frac{k^2\hbar^2}{2m} \tag{2.5}$$

Note that sometimes the notation $\lambda \equiv \frac{2\pi}{k}$ is used above. In this case we would have:

$$p = \frac{2\pi}{\lambda} \times \hbar = \frac{h}{\lambda} \tag{2.6}$$

This is called the *de Broglie relation*.

2.6 The significance of the coefficients

So far we have not discussed the meaning of the coefficients A and B in the solution $\psi(x) = A\exp(ikx) + B\exp(-ikx)$. Their meaning can be seen by operating on ψ by the momentum operator and choosing $B = 0$ and $A = 0$ in turn:

$$p\psi = \frac{\hbar}{i}\frac{d\psi}{dx} = \frac{\hbar}{i}\frac{d\left(Ae^{ikx}\right)}{dx} = k\hbar Ae^{ikx} = k\hbar\psi \tag{2.7}$$
$$p\psi = \frac{\hbar}{i}\frac{d\psi}{dx} = \frac{\hbar}{i}\frac{d\left(Be^{-ikx}\right)}{dx} = -k\hbar Be^{-ikx} = -k\hbar\psi$$

The two eigenvalues $k\hbar$ and $-h\hbar$ correspond to states having linear momentum along positive and negative x-axis, respectively. The significance of the coefficients A and B depend on how the state of the particle was prepared. If $B = 0$ then the particle will travel along the positive x-axis direction.

Note that neither $\cos(kx)$ nor $\sin(kx)$ are eigenfunctions of the momentum operator. For example:

$$p\psi = \frac{\hbar}{i}\frac{d\psi}{dx} = \frac{\hbar}{i}\frac{d(\cos(kx))}{dx} = ik\hbar\sin(kx)$$

A similar calculation can be carried out for $\cos(kx)$. For $\psi(x) = \sin(kx)$ we can write $\psi(x) = \frac{1}{2}Ce^{ikx} + \frac{1}{2}Ce^{-ikx}$ and we see that both $\sin(kx)$ and $\cos(kx)$ can be expressed as a linear combination of the eigenfunctions with equal weights.

2.7 The flux density

The *flux density*, which indicates the quantum motion present in a given wavefunction, is defined as (defined about x-axis below):

$$J_x = \frac{1}{2m}\left(\Psi^* p_x \Psi + \Psi p_x^* \Psi^*\right) = \frac{\hbar}{2mi}\left(\Psi^* \frac{d}{dx}\Psi - \Psi \frac{d}{dx}\Psi^*\right) \tag{2.8}$$

where, in the position representation, $p_x = \frac{\hbar}{i} d/dx$ and $p_x^* = -\frac{\hbar}{i} d/dx$. If Ψ corresponds to a stationary state (i.e. just a simple phase evolution corresponding to the eigenvalue), we can simplify the above expression as:

$$J_x = \frac{1}{2m}\left(\psi^* p_x \psi + \psi p_x^* \psi^*\right) = \frac{\hbar}{2mi}\left(\psi^* \frac{d}{dx}\psi - \psi \frac{d}{dx}\psi^*\right) \tag{2.9}$$

The general expression in 3-D is given by:

$$\vec{J} = \frac{1}{2mi}\left(\Psi^* \vec{\nabla}\Psi - \Psi \vec{\nabla}\Psi^*\right)$$

If the probability of finding particle in $dxdydz$ is $|\Psi(x,y,z)|^2\, dxdydz$ then the probability flux of particle passing from/into this volume element is given by $\vec{J}(x,y,z)$. Note that this has both magnitude and direction.

Example 2.1 Calculate the flux density for a system that is described by $\psi(x) = Ae^{ikx}$.

Solution. Evaluate $J_x\psi$ as follows:

$$J_x = \frac{1}{2m}\left\{\left(Ae^{ikx}\right)^*\left(\frac{\hbar}{i}\frac{d}{dx}\right)\left(Ae^{ikx}\right) + \left(Ae^{ikx}\right)\left(\frac{\hbar}{i}\frac{d}{dx}\right)^*\left(Ae^{ikx}\right)^*\right\}$$

$$= \frac{1}{2m}\left\{\left(A^*e^{-ikx}\right)\left(\frac{\hbar}{i}\frac{d}{dx}\right)\left(Ae^{ikx}\right) - \left(Ae^{ikx}\right)\left(\frac{\hbar}{i}\frac{d}{dx}\right)\left(A^*e^{-ikx}\right)\right\}$$

$$= \frac{\hbar\left|A\right|^2}{2mi}\left\{\left(e^{-ikx}\right)(ik)\left(e^{ikx}\right) - \left(e^{ikx}\right)(-ik)\left(e^{-ikx}\right)\right\} = \frac{k\hbar\left|A\right|^2}{m}$$

Note that $\pm k\hbar/m$ is the classical velocity of the particle (from $p = mv$), so that the flux density above is the velocity multiplied by the probability density that the particle is in that state.

2.8 Wavepackets

A *wavepacket* is a special wavefunction representing a particle that is spatially localized. Note that spatial localization necessarily introduces uncertainty in the momentum. As the name implies, wavepackets are constructed from multiple "waves", which here refer to a superposition of momentum eigenfunctions (e.g., e^{ikx} and e^{-ikx}). Such a wavepacket can be stationary (i.e., it does not move but may broaden as a function of time; *dispersion*) or it can be given momentum so that it can move in a given direction. In this sense, wavepackets move through space in a similar way as classical particles. A wavepacket can be formed as a superposition of functions $\Psi_k(x,t)$:

$$\Psi_k(x,t) = Ae^{ikx}e^{-iE_kt/\hbar} \tag{2.10}$$

The time dependent phase factor is dictated by the time-dependent Schrödinger equation. Instead of summing $\Psi_k(x,t)$, we take a continuum of states and integrate over all momenta to form the wavepacket:

$$\Psi(x,t) = \int_{-\infty}^{\infty} g(k)\Psi_k(x,t)dk \tag{2.11}$$

where $g(k)$ is called the *shape function*. Because of the interference between different $\Psi_k(x,t)$, the overall wavepacket appears localized. The time-dependent part in Eq. (2.10) makes the wavepacket "walk".

Example. Gaussian wavepacket in harmonic potential.

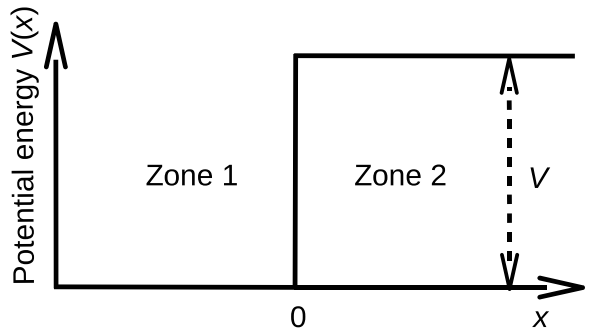
Click here to start the movie

Example. Gaussian wavepacket propagation through two slits ("the two slit experiment").

Click here to start the movie

2.9 Penetration into and through barriers: An infinitely thick potential wall

In the following we will consider how a particle in freely translational motion will react to a potential barrier. First we consider an infinitely thick potential barrier as shown below.



We take the particle to approach from the left (Zone 1) and proceed to the right toward Zone 2. According to classical physics, the particle would be able to go over the barrier if it has enough kinetic energy ($T \geq V$). If this is not the case, the particle would bounce back and continue to the left. Quantum mechanics predicts something very different as it allows particle to overcome the barrier even if it does not have the sufficient kinetic energy.

To solve the problem, we have to write the Schrödinger equation for the problem. Since we have two different zones with different potential energy, we can write two different equations:

$$\begin{aligned} \text{Zone 1 } (x<0)\text{: } H &= -\frac{\hbar^2}{2m}\frac{d^2}{dx^2} \\ \text{Zone 2 } (x\geq 0)\text{: } H &= -\frac{\hbar^2}{2m}\frac{d^2}{dx^2}+V \end{aligned} \tag{2.12}$$

In zone 1 the eigenfunctions correspond to those of a free particle (i.e., e^{ikx} and e^{-ikx}). In Zone 2 we have to replace E by $E-V$ to account for the change in potential. According to Eq. (2.4) the solutions in these to zones are:

$$\begin{aligned} &\text{Zone 1: } \psi(x) = Ae^{ikx}+Be^{-ikx} \text{ with } k\hbar=\sqrt{2mE} \\ &\text{Zone 2: } \psi(x) = A'e^{ik'x}+B'e^{-ik'x} \text{ with } k'\hbar=\sqrt{2m\left(E-V\right)} \end{aligned} \tag{2.13}$$

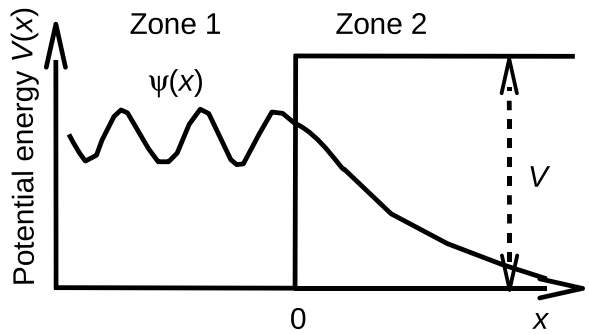
If $E<V$ then k' must be imaginary so we will denote $k'=i\kappa$ to simplfy notation. Then in Zone 2 we have:

$$\text{Zone 2: } \psi(x) = A'e^{-\kappa x}+B'e^{\kappa x} \text{ where } \kappa\hbar=\sqrt{2m\left(V-E\right)} \tag{2.14}$$

The above wavefunction is a mixture of decaying and increasing exponentials and thus the wavefunction does not oscillate in Zone 2 (as opposed to Zone 1). Because the potential barrier is infinitely wide, we can safely exclude the increasing exponential in the solution since we expect the amplitude of the wavefunction to decay as we go deeper into the barrier:

$$\text{Zone 2: } \psi(x) = A' e^{-\kappa x}$$

We can now give a qualitative picture of the solution:

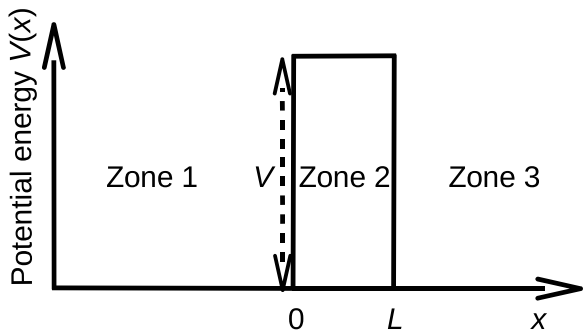


The decay reate is determined by κ and κ^{-1} is called the *penetration depth*. As can be seen in Eq. (2.14), the penetration depth decreases when the mass increases or the potential barrier height V increases.

Macroscopic particles have such high mass that in practice they will not be found in classically forbidden regions. However, for light particles, such electrons or protons, this effect is often significant.

2.10 A barrier of finite width

Next we will consider a barrier that has a finite width:



Here the potential function $V(x)$ is given by

$$\begin{aligned} \text{Zone 1 } (x < 0): &\qquad V(x) = 0 \\ \text{Zone 2 } (0 \leq x < L): &\qquad V(x) = V \\ \text{Zone 3 } (x \geq 0): &\qquad V(x) = 0 \end{aligned} \tag{2.15}$$

Following our previous calculation, we can writen down the solution to Schrödinger equation as:

$$\begin{aligned} &\text{Zone 1: } \psi(x) = Ae^{ikx} + Be^{-ikx} \text{ with } k\hbar = \sqrt{2mE} \\ &\text{Zone 2: } \psi(x) = A'e^{ik'x} + B'e^{-ik'x} \text{ with } k'\hbar = \sqrt{2m\left(E - V\right)} \\ &\text{Zone 3: } \psi(x) = A''e^{ikx} + B''e^{-ikx} \text{ with } k\hbar = \sqrt{2mE} \end{aligned} \tag{2.16}$$

Components with positive momentum (k) are called *incoming waves* and with negative momentum *outgoing waves*. In Zone 1 the coefficient A gives the initial weight of the incoming wave and B the outgoing wave. We will divide the solution into three different cases:

1. Case $E < V$. In this case the particle does not have enough energy to classically cross the barrier. According to quantum mehcanics, it is possible that the particle will be able to cross the barrier. In Zone 2 Eq. (2.14) gives the wavefunction as $\psi = A'e^{-\kappa x} + B'e^{\kappa x}$. To connect Zones 1, 2 and 3, we require that the wavefunction and its derivative are continuous at the boundaries (i.e., when $x = 0$ and $x = L$). The wavefunction continuity gives:

$$\begin{aligned} &\text{At } x = 0\text{: } A + B = A' + B' \\ &\text{At } x = L\text{: } A'e^{-\kappa L} + B'e^{\kappa L} = A''e^{ikL} + B''e^{-ikL} \end{aligned} \tag{2.17}$$

Continuity for the first derivative gives:

$$\begin{aligned} &\text{At } x = 0\text{: } ikA - ikB = -\kappa A' + \kappa B' \\ &\text{At } x = L\text{: } -\kappa A'e^{-\kappa L} + \kappa B'e^{\kappa L} = ikA''e^{ikL} - ikB''e^{-ikL} \end{aligned} \tag{2.18}$$

1. (continued) So far we have four equations and six unknowns (the coefficients A, B, A', B', A'', B''). Normalization condition for the wavefunction gives one more equation. We can also state that in Zone 3 we should only have forward propagating wave, which means that $B'' = 0$. In zone 1 we can identify B as the amplitude of the reflected wave. Then the *reflection probability* R and *transmission probability* T are given by (see Example 2.1):

$$R = \frac{J_x(-k)}{J_x(k)} = \frac{(k\hbar/m)\,|B|^2}{(k\hbar/m)\,|A|^2} = \frac{|B|^2}{|A|^2} \tag{2.19}$$

$$T = \frac{|A''|^2}{|A|^2}$$

Note that we have not written down the solution for A, B, A', B', A'', B'' since the expressions are quite complicated. The full expression for T is:

$$T = \frac{1}{1 + \left(e^{\kappa L} - e^{-\kappa L}\right)^2 / \left(16\,(E/V)\,(1 - E/V)\right)} \tag{2.20}$$

$$R = 1 - T$$

where $\kappa = \sqrt{2mV\,(1 - e/V)}/\hbar$. Since we have $E < V$, T is often called the *tunneling probability*.

2. $\underline{E > V}$. Classcally, the particle now has sufficient energy to overcome the potential barrier. One might expect to get $T = 1$ and $R = 0$ but this is not the case. To get the result, one can replace κ by $-ik'$:

$$T = \frac{1}{1 + \left(\sin^2\left(k'L\right)\right) / \left(4\left(E/V\right)\left(E/V - 1\right)\right)} \tag{2.21}$$
$$R = 1 - T$$

where $\hbar k' = \sqrt{2mV\left(E/V - 1\right)}$. Note that T reaches unity whenever $\sin(k'L) = 0$, which can happen at:

$$k' = \frac{n\pi}{L} \text{ where } n = 1, 2, 3, ... \tag{2.22}$$

Furthermore, T has minima at:

$$k' = \frac{n\pi}{2L} \text{ where } n = 1, 3, ...$$

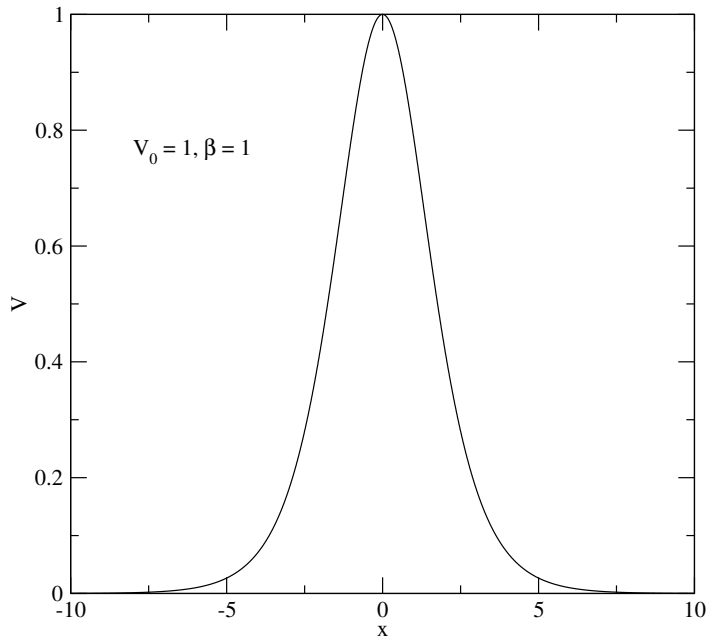
At the high energy limit ($E >> V$), T approaches one (i.e. the classical behavior). Note that even when $E > V$ there is a possibility that the particle is reflected back. This phenomenom is called *antitunneling* or *non-classical reflection*.

2.11 The Eckart potential barrier

The rectangular shape potential that we have considered previously is not very realistic in many practical problems. Unfortunately, the algebra becomes very tedious even with simple potential functions. One potential function with known solutions is the *Eckart potential*:

$$V(x) = \frac{4V_0 e^{\beta x}}{\left(1 + e^{\beta x}\right)^2} \tag{2.23}$$

where V_0 and β are constants with dimensions of energy and inverse length, respectively. This function is plotted below.



The Schrödinger equation corresponding to the potential in Eq. (2.23) can be solved analytically. However, the calculation is quite complicated and we just note the end result for the transmission coefficient:

$$T = \frac{\cosh\left(4\pi\sqrt{2mE}/(\hbar\beta)\right) - 1}{\cosh\left(4\pi\sqrt{2mE}/(\hbar\beta)\right) + \cosh\left(2\pi\left|8mV_0 - (\hbar\beta)^2\right|^{1/2}/(\hbar\beta)\right)} \tag{2.24}$$

When $E << V_0$ we have $T \approx 0$. When $E \to V_0$ then T approaches 1 and the classical limit is recovered when $E >> V_0$. More details of the Eckart potential problem can be found from *Phys. Rev.* 35, 1303 (1930).

2.12 Particle in a box: Solutions of the one dimensional problem

Particle in one dimensional box is confined by a "box potential", which is essentially an infinitely deep square well. The Hamiltonian for this problem is given by:

$$H = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + V(x) \text{ where } V(x) = \begin{cases} 0, & \text{for } 0 \leq x \leq L \\ \infty, & \text{otherwise} \end{cases} \tag{2.25}$$

Because the potential energy at the walls (i.e. when $x = 0$ or $x = L$), the particle cannot penetrate the wall at these points. Thus we can replace the potential term with a *Dirichlet boundary condition* where we require the wave function to be zero at $x = 0$ and $x = L$. Now Eq. (2.25) can be written simply as:

$$H = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2} \tag{2.26}$$

but we require that ψ in $H\psi = E\psi$ goes to zero at points $x = 0$ and $x = L$. Note that the previous equation looks like there is no external potential and one might think that there is no quantitization of energy. However, in this case the boundary conditions enforce the quantitization. The general solution to the eigenvalue problem $H\psi = E\psi$ with H given in Eq. (2.26) is:

$$\psi(x) = C\cos(kx) + D\sin(kx) \text{ with } k\hbar = \sqrt{2mE}$$

This has to be zero when $x = 0$: $\psi(0) = C = 0$. Thus $C = 0$ and we have $\psi(x) = D\sin(kx)$ so far. At $x = L$ we have: $\psi(L) = D\sin(kL) = 0$. This can be satisfied when:

$$k = \frac{n\pi}{L} \text{ where } n = 1, 2, ... \tag{2.27}$$

Note that the value $n = 0$ is excluded because it would yield $\psi(x) = 0$ for all x. The integer n is called a *quantum number* and can be used to label states of the system as well as to calculate the energy:

$$E_n = \frac{k^2\hbar^2}{2m} = \frac{n^2\hbar^2\pi^2}{2mL^2} = \frac{n^2h^2}{8mL^2} \text{ with } n = 1, 2, ... \tag{2.28}$$

The important outcome of this calculation is to see that the energy is quantitized, which means that it can only take certain values (given by Eq. (2.28). We still need to determine the constant D in the wavefunction, which will be given by the normalization condition:

$$\int_0^L \psi^*(x)\psi(x)dx = D^2 \int_0^L \sin^2\left(\frac{n\pi x}{L}\right) dx = \frac{1}{2}LD^2$$

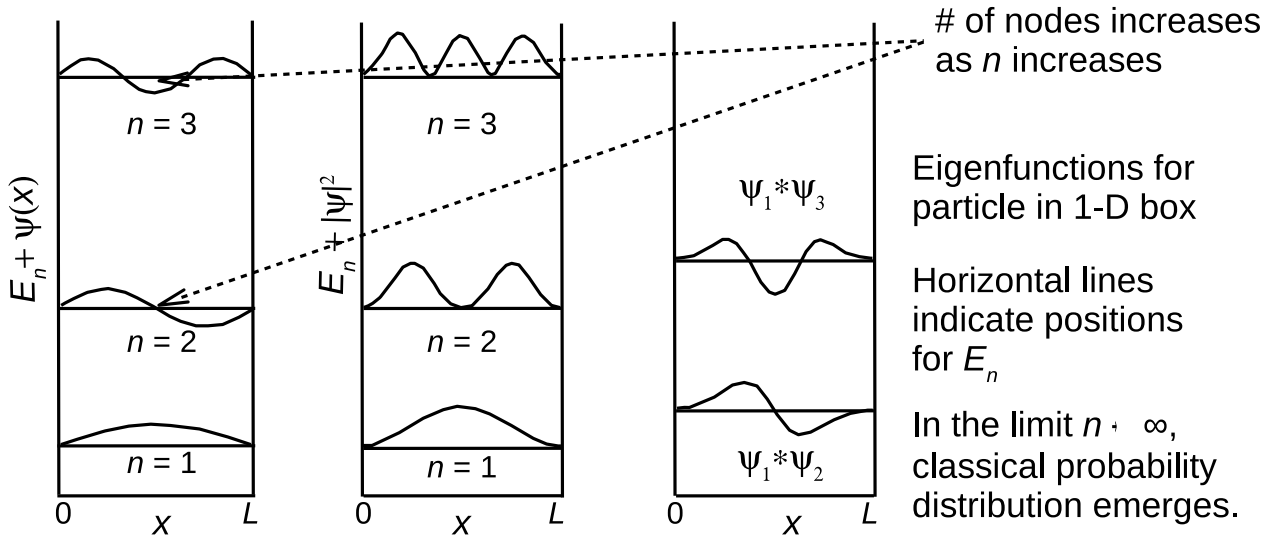
which gives $D = \sqrt{2/L}$.

The complete wavefunction can now be written as:

$$\psi(x) = \sqrt{\frac{2}{L}} \sin\left(\frac{n\pi x}{L}\right) \tag{2.29}$$

with the energies given by:

$$E_n = \frac{n^2 h^2}{8mL^2} \text{ where } n = 1, 2, ...$$



Energy levels and wavefunctions for a particle in one dimensional box.

2.13 Particle in a box: Features of the solution

By looking at the solutions, we can immediately note the following:

- The lowest energy the particle can have is reached when $n = 1$. *This energy is non-zero*: $E_1 = h^2/(8mL^2)$. This irremovable energy is called the *zero-point energy*. In classical physics, where $h = 0$, there is no zero-point energy. This is a direct consequence of the uncertainty present in the momentum. Since $\Delta p \neq 0$, we have $\langle p^2 \rangle \neq 0$ and the average kinetic energy $T \propto \langle p^2 \rangle$ is not zero.
- The energy separation between the neighboring states decreases as a function of the box size. A larger box gives more freedom to the particle and hence the energy separation (and the zero-point energy) will be smaller:

$$E_{n+1} - E_n = \left\{ (n+1)^2 - n^2 \right\} \frac{h^2}{8mL^2} = (2n+1) \frac{h^2}{8mL^2} \tag{2.30}$$

  If the width of the box would approach macroscopic dimensions, the energy levels would be very close to each other and energy would appear as a continuous function. The same holds for the particle mass m.

2.14 The two-dimensional square well ("2D box")

The Hamiltonian for a particle in two-dimensional square is given by:

$$H = -\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right) \tag{2.31}$$

with Dirichlet boundary conditions $\psi(L_1, y) = 0$ and $\psi(x, L_2) = 0$. Inside the well the Schrödinger equation is:

$$\frac{\partial^2\psi}{\partial x^2} + \frac{\partial^2\psi}{\partial y^2} = -\frac{2mE}{\hbar^2}\psi \tag{2.32}$$

Since the Hamiltonian is a sum of terms each depending on x and y separately, we can write the solution as a product: $\psi(x, y) = X(x)Y(y)$. Inserting this into Eq. (2.32):

$$Y(y)\frac{d^2X(x)}{dx^2} + X(x)\frac{d^2Y(y)}{dy^2} = -\frac{2mE}{\hbar^2}X(x)Y(y)$$

Division of both sides by $X(x)Y(y)$ yields:

$$\frac{1}{X(x)}\frac{d^2X(x)}{dx^2} + \frac{1}{Y(y)}\frac{d^2Y(y)}{dy^2} = -\frac{2mE}{\hbar^2}$$

Next we divide E into two contributions $E = E^X + E^Y$ and separate the above equation as:

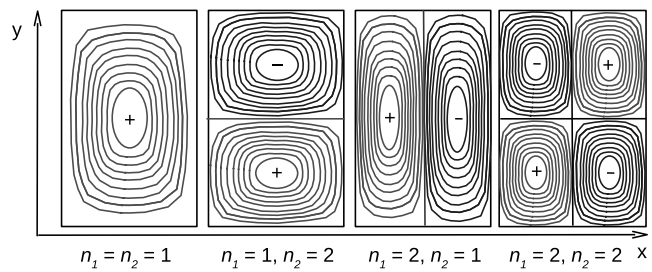
$$\frac{d^2X(x)}{dx^2} = -\frac{2mE^X}{\hbar^2}X \text{ and } \frac{d^2Y}{dy^2} = -\frac{2mE^Y}{\hbar^2}Y$$

Both equations have the same form as the one-dimensional particle in a box. Hence we form the overall wavefunction as a product of the one dimensional problem solutions:

$$\psi_{n_1,n_2}(x,y) = \frac{2}{\sqrt{L_1 L_2}} \sin\left(\frac{n_1\pi x}{L_1}\right) \sin\left(\frac{n_2\pi y}{L_2}\right) \tag{2.33}$$

$$E_{n_1,n_2} = \frac{h^2}{8m}\left(\frac{n_1^2}{L_1^2} + \frac{n_2^2}{L_2^2}\right) \text{ with } n_1 = 1, 2, ... \text{ and } n_2 = 1, 2, ...$$

Note that in order to define a state, one must now use two independent quantum numbers n_1 and n_2.



Contour plots of the lowest eigenfunctions for the two-dimensional particle in a box problem.

2.15 Degeneracy

In two-dimensions, when $L_1 = L_2 = L$, the state energies are given by:

$$E_{n_1,n_2} = \frac{h^2}{8mL^2}\left(n_1^2 + n_2^2\right) \tag{2.34}$$

For example, states with ($n_1 = 1$, $n_2 = 2$ or $|1,2\rangle$) and ($n_1 = 2$, $n_2 = 1$ or $|2,1\rangle$) (i.e. reversed order) will give exactly the same energy. We say that these two states are *degenerate*. In general, the reversal of n_1 and n_2 will result in another state that still has the same energy. Note that the wavefunctions corresponding to these states are different. Following the previous example we have:

$$\psi_{1,2}(x,y) = \frac{2}{L}\sin\left(\frac{\pi x}{L}\right)\sin\left(\frac{2\pi y}{L}\right) \text{ and } \psi_{2,1}(x,y) = \frac{2}{L}\sin\left(\frac{2\pi x}{L}\right)\sin\left(\frac{\pi y}{L}\right)$$

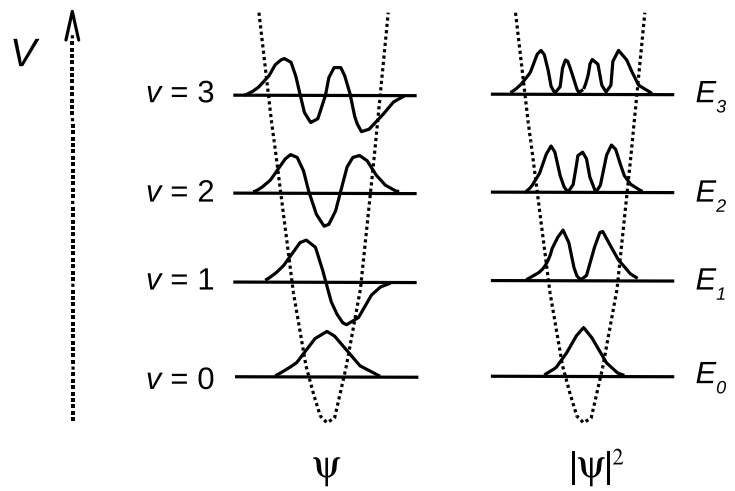
In quantum mechanics degeneracy may arise in two different ways: 1) due to symmetry present in the potential or 2) by accident. The latter is called *accidental degeneracy*. To demonstrate this, consider the particle in a two-dimensional box problem. Choose $L_1 = L$ and $L_2 = 2L$ then there will be accidental degeneracy between states $|1,4\rangle$ and $|2,2\rangle$. In this case there is *hidden symmetry* in the system that is reflect by the fact that the box sides are multiples of L.

2.16 The harmonic oscillator

The most important potential in chemistry and physics is the *harmonic oscillator*. The name originates from the fact that this potential can also be used to describe oscillations in a spring or the movement of a pendulum. In physics it can be used to approximate the potential that atoms feel when they are trapped in a lattice, or in chemistry, when a chemical bond is approximated as a spring. The potential function has the following shape:

$$V(x) = \frac{1}{2}kx^2 \tag{2.35}$$

where k is the *force constant* that defines the steepness of the potential function. Just like in the particle in one-dimensional box problem, this potential has quantitized levels with the associated energies and eigenfunctions. This is demonstrated below:



Harmonic oscillator potential, quantitized energies and eigenfunctions.

The Hamiltonian for the harmonic oscillator problem is:

$$H = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + \frac{1}{2}kx^2 \tag{2.36}$$

and the corresponding Schrödinger equation is then:

$$-\frac{\hbar^2}{2m}\frac{d^2\psi(x)}{dx^2} + \frac{1}{2}kx^2\psi(x) = E\psi(x) \tag{2.37}$$

Often, instead of using k, the equation is written in terms of ω, which is related to k by $\omega = \sqrt{k/m}$ (m is the particle mass). The easiest way to solve Eq. (2.37) is to use so called *rising* (or *creation*) and *lowering* (or *annihilation*) operators. The idea is to express the original Hamiltonian as a product of two operators that are easier to deal with. The following choice can be used:

$$a = \sqrt{\frac{m\omega}{2\hbar}}\left(x + i\frac{p}{m\omega}\right)$$

$$a^+ = \sqrt{\frac{m\omega}{2\hbar}}\left(x - i\frac{p}{m\omega}\right)$$

When these operators are multiplied, we get:

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

The Hamiltonian written in terms of ω is:

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

Comparison of the two previous expressions yields the following:

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

The operators a and a^+ do not commute. This can be seen by calculating aa^+ and comparing the result with the previous calculation of a^+a:

$$aa^+ = \frac{m\omega}{2\hbar}\left(x + i\frac{p}{m\omega}\right)\left(x - i\frac{p}{m\omega}\right) = \frac{1}{\hbar\omega}\left(\frac{p^2}{2m} + \frac{m\omega}{2}x^2\right) + \frac{1}{2}$$

Thus we get $\left[a, a^+\right] = 1$. Next we will consider eigenvalues of H:

$$H\psi_n = \hbar\omega\left(a^+a + \frac{1}{2}\right)\psi_n = \lambda_n\psi_n$$

The eigenvalue λ_n is positive since:

$$\left\langle \psi_n \left| \hbar\omega\left(a^+a + \frac{1}{2}\right)\right| \psi_n \right\rangle = \hbar\omega\left\langle \psi_n \left| a^+a \right| \psi_n \right\rangle + \frac{1}{2}\hbar\omega$$

$$= \hbar\omega\left(\left\langle \psi_n \left| a^+a \right| \psi_n \right\rangle + \frac{1}{2}\right) = \hbar\omega\left\langle a\psi_n | a\psi_n \right\rangle + \frac{\hbar\omega}{2} = \hbar\omega\int |a\psi_n|^2\, d\tau + \frac{\hbar\omega}{2} \geq \frac{\hbar\omega}{2}$$

Consider operation of a^+ on ψ (here $H\psi_n = \lambda_n \psi_n$):

$$H\left(a^+\psi_n\right) = \hbar\omega\left(a^+a + \frac{1}{2}\right)a^+\psi_n = \hbar\omega\left(\underbrace{a^+aa^+}_{=a^+(1+a^+a)} + \frac{1}{2}a^+\right)\psi_n$$

$$= \hbar\omega\left(a^+(1+a^+a) + \frac{1}{2}a^+\right)\psi_n = \hbar\omega a^+\left(1 + a^+a + \frac{1}{2}\right)\psi_n$$

$$= a^+\left(\hbar\omega\left(a^+a + \frac{1}{2}\right) + \hbar\omega\right)\psi_n = a^+\left(\lambda_n + \hbar\omega\right)\psi_n = \left(\lambda_n + \hbar\omega\right)a^+\psi_n$$

This states that the operator a^+ generated ψ_{n+1} from ψ_n, which has one unit ($\hbar\omega$) higher energy. In a similar way, one can show that:

$$H\left(a\psi_n\right) = \left(\lambda_n - \hbar\omega\right)a\psi_n$$

This immediately tells us that the energy levels in harmonic oscillator are evenly spaced and that their separation is $\hbar\omega$. Furthermore, since all λ_n's are positive, there must be a lower bound λ_0.

If we take the lower limit for $\lambda_0 = \frac{\hbar\omega}{2}$ (our previous estimate) then we can write the original problem as:

$$-\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2} + \frac{m\omega^2}{2}x^2\psi = \underbrace{\lambda_0}_{=\hbar\omega/2}\psi$$

which appears to be satisfied by the following Gaussian function (direct calculation):

$$\psi_0(x) \propto \exp\left(-\frac{m\omega x^2}{2\hbar}\right)$$

Then we have clearly found the ground state solution to the harmonic oscillator problem. The rising operator can now be used to generate the excited states:

$$\lambda_n = \left(n + \frac{1}{2}\right)\hbar\omega \text{ where } \omega = \sqrt{k/m} \text{ and } n = 0, 1, 2... \tag{2.38}$$

To find the eigenfunctions, we must operate on ψ_0 by a^+:

$$\psi_n = \left(a^+\right)^n \psi_0 = \left(\frac{m\omega}{2\hbar}\right)^{n/2}\left(x - i\frac{p}{m\omega}\right)^n \psi_0$$

The rising operator essentially generates Hermite polynomials $H_n\left(\sqrt{\frac{m\omega}{\hbar}}x\right)$ in front of the Gaussian function. Thus we can writen the complete wavefunction as (note that in the textbook, $n = v$):

$$\psi_n(x) = N_n H_n(\alpha x)\, e^{-\alpha^2 x^2/2} \quad \alpha = \left(\frac{mk}{\hbar^2}\right)^{1/4} \quad N_n = \sqrt{\frac{\alpha}{2^n n!\sqrt{\pi}}} \tag{2.39}$$

Exercise. Show that N_n above produces a normalized wavefunction.

For the two lowest stats the corresponding wavefunctions are:

$$\psi_0(x) = \sqrt{\frac{\alpha}{\sqrt{\pi}}} e^{-\alpha^2 x^2/2} \quad \psi_1(x) = \sqrt{\frac{2\alpha^3}{\sqrt{\pi}}} x e^{-\alpha^2 x^2/2}$$

Some of the first Hermite polynomials are listed below:

n	$H_n(x)$
0	1
1	$2x$
2	$4x^2 - 2$
3	$8x^3 - 12x$
4	$16x^4 - 48x^2 + 12$
5	$32x^5 - 160x^3 + 120x$

For example, the following Maxima program can be used to generate Hermite polynomials:

```
load(orthopoly);
hermite(0, x);
hermite(1, x);
hermite(2, x);
hermite(3, x);
/* etc. */
```

The following realtions are useful when working with Hermite polynomials:

$$H_v''(y) - 2yH_v'(y) + 2vH_v(y) = 0 \text{ (characteristic equation)}$$

$$H_{v+1}(y) = 2yH_v(y) - 2vH_{v-1}(y) \text{ (recursion relation)}$$

$$\int_{-\infty}^{\infty} H_{v'}(y)H_v(y)e^{-y^2}\,dy = \begin{cases} 0, & \text{if } v' \neq v \\ \sqrt{\pi}2^v v!, & \text{if } v' = v \end{cases}$$

2.17 Properties of the solutions

The following table summarizes properties of the harmonic oscillator:

Energies:	$E_v = \left(v + \frac{1}{2}\right)\hbar\omega,\ \omega = \sqrt{k/m}$
Wavefunctions:	$\psi_v(x) = N_v H_v(\alpha x) e^{\alpha^2 x^2/2}$
	$\alpha = \left(\frac{mk}{\hbar^2}\right)^{1/4} \quad N_v = \sqrt{\frac{\alpha}{2^v v! \sqrt{\pi}}}$
Matrix elements:	$\langle v+1 \left\vert x \right\vert v\rangle = \sqrt{\frac{\hbar}{2m\omega}}\sqrt{v+1},\ \langle v-1 \left\vert x \right\vert v\rangle = \sqrt{\frac{\hbar}{2m\omega}}\sqrt{v}$
	$\langle v+1 \left\vert p_x \right\vert v\rangle = i\sqrt{\frac{\hbar m\omega}{2}}\sqrt{v+1},\ \langle v-1 \left\vert p_x \right\vert v\rangle = -i\sqrt{\frac{\hbar m\omega}{2}}\sqrt{v}$
Virial theorem:	$\langle E_K\rangle = \langle E_P\rangle$ for all v

As we noted previously, the spacing between the energy levels is given by:

$$E_{v+1} - E_v = \hbar\omega \tag{2.40}$$

Just like for particle in a box, we can see that the system approaches classical behavior if $\omega \to 0$ (i.e. $m \to \infty$ or $k \to 0$). The *virial theorem* describes the balance between the kinetic and potential energies as follows:

$$2\langle E_K\rangle = s\langle E_P\rangle \tag{2.41}$$

$$\langle E_{tot}\rangle = \langle E_K\rangle + \langle E_P\rangle = (1 + 2/s)\langle E_K\rangle \tag{2.42}$$

As indicated above in the table, for harmonic oscillator $s = 2$ (calculation not shown).

2.18 The classical limit

We consider two classical limits:

1. As the quantum number $v \to \infty$, the maxima in $|\psi|^2$ tends to the left and right sides of the potential. These maxima correspond to *classical turning points* where the system spends most of its time. Note that still on the average, the system will be found at the minimum of the potential.
2. If the wavefunction is localized in space, we have a *wavepacket*. Apart from the width, this type of wavefunction gives a well defined location.

Example. Show that whatever superposition of harmonic oscillator is used to construct a wavepacket, it is localized at the same place at the times $0, T, 2T, ...$ where T is the classical period of the oscillator.

Solution. The classical period is $T = 2\pi/\omega$. We need to form a time-dependent wavepacket by superimposing the $\Psi(x,t)$ for the oscillator, and then evaluate it at $t = nT$, with $n = 0, 1, 2, ...$ The wavepacket has the following form:

$$\Psi(x,t) = \sum_v c_v \Psi(x,t) = \sum_v c_v \psi_v(x) e^{-E_v t/\hbar} = \sum_v c_v \psi_v(x) e^{-i(v+1/2)\omega t}$$

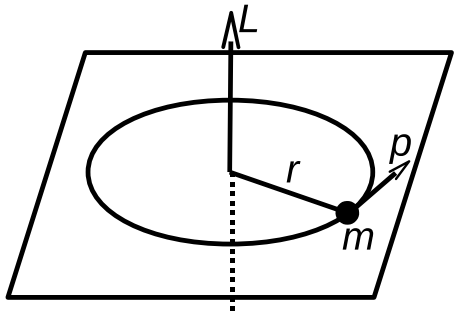
At time nT this becomes:

$$\Psi(x,nT) = \sum_v \psi_v(x) e^{-2\pi n i(v+1/2)} = \sum_v c_v \psi_v(x)\,(-1)^n = (-1)^n\,\Psi(x,0)$$

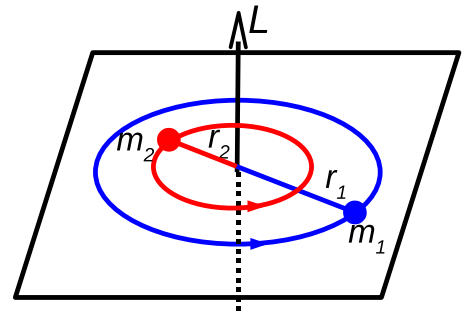
Chapter 3: Rotational motion and the hydrogen atom

3.1 Particle on a ring: The Hamiltonian and the Schrödinger equation

First we consider quantum mechanics of a particle travelling on a circular ring as shown below. Note the analogy to molecular rotation.



Rotation about a fixed point



Rotation of diatomic molecule around the center of mass

Particle mass is denoted by m, distance from the origin by $\vec{r}$ and the classical *angular momentum* associated with the rotational motion by $\vec{l} = \vec{r} \times \vec{p}$. Classically the *moment of inertia* is given by $I = m\,|r|^2$.

The particle of mass m is confined to circular orbit in the xy-plane. The potential energy is taken to be zero but we require periodic boundary copndition (i.e. the wavefunction must return to the same value after 360°rotation). The Hamiltonian is given by:

$$H = -\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right) \tag{3.1}$$

Note that here we keep the distance between the particle and the center of rotation constant. Hence we can use the polar coorinates to simlify our calculations: $x = r\cos(\phi)$ and $y = r\sin(\phi)$ where $\phi : 0 \to 2\pi$. The Laplacian ∇^2 in polar coordinates is:

$$\nabla^2 = \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right) = \frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\frac{\partial^2}{\partial \phi^2} \tag{3.2}$$

In our problem r is a constant and hence the r dependent derivatives in the Hamiltonian are zero:

$$H = -\frac{\hbar^2}{2mr^2}\frac{d^2}{d\phi^2} = -\frac{\hbar^2}{2I}\frac{d^2}{d\phi^2} \tag{3.3}$$

The wavefunction depends only on the angle ϕ. We denote this angle dependent wavefunction by $\Phi = \Phi(\phi)$. The Schrödinger equation is now:

$$\frac{d^2\Phi}{d\phi^2} = -\frac{2IE}{\hbar^2}\Phi \tag{3.4}$$

The general solutions to this equation are:

$$\Phi = Ae^{im_l\phi} + Be^{-im_l\phi} \tag{3.5}$$

The quantity m_l is a dimensionless number. Φ must fulfills *cyclic boundary condition*, which states that $\Phi(\phi) = \Phi(\phi + 2\pi)$. This implies that:

$$\underbrace{Ae^{im_l(\phi+2\pi)} + Be^{-im_l(\phi+2\pi)}}_{=\Phi(\phi+2\pi)} = Ae^{im_l\phi}e^{2\pi im_l} + Be^{-im_l\phi}e^{-2\pi im_l} = \underbrace{Ae^{im_l\phi} + Be^{-im_l\phi}}_{=\Phi(\phi)}$$

We can immediately see that for the above to hold, we must have $e^{2\pi im_l} = 1$. This means that $m_l = 0, \pm 1, \pm 2, ...$ It now follows from Eqs. (3.4) and (3.5) that:

$$E_{m_l} = \frac{m_l^2\hbar^2}{2I} \text{ with } m_l = 0, \pm 1, \pm 2, ... \tag{3.6}$$

3.2 The angular momentum

On the previous slide, the positive and negative values of m_l correspond to different directions of rotation. This is similar what we had for linear momentum with k and $-k$. To confirm this interpretation, we need to evaluate the z component of angular momentum l. The classical expression for angular momentum is:

$$\vec{l} = \vec{r} \times \vec{p} = \begin{vmatrix} i & j & k \\ x & y & z \\ p_x & p_y & p_z \end{vmatrix} \tag{3.7}$$

where i, j, k are orthogonal unit vectors along the x, y, z axis, respctively. After expanding the determinant (see your CHEM352 notes), we get the z component as:

$$L_z = xp_y - yp_x \tag{3.8}$$

Substitution of the quantum mechanical momentum operators gives:

$$l_z = x\left(\frac{\hbar}{i}\frac{\partial}{\partial y}\right) - y\left(\frac{\hbar}{i}\frac{\partial}{\partial x}\right)$$

In polar coordinates (see your CHEM352 notes) this becomes:

$$l_z = \frac{\hbar}{i}\frac{\partial}{\partial \phi} \tag{3.9}$$

Next we operate on the eigenfunctions we obtained earlier by l_z:

$$l_z \Phi_{m_l} = \frac{\hbar}{i} \frac{\partial}{\partial \phi} A e^{i m_l \phi} = m_l \hbar A e^{i m_l \phi} = m_l \hbar \Phi_{m_l} \tag{3.10}$$

This shows that the z component of the angular momentum is positive if m_l is positive and vice versa for negative values. Thus, when m_l is positive, the rotation about z axis proceeds counter clockwise and when m_l is negative, the rotation proceeds clockwise.

The remaining task is to normalize the wavefunctions (i.e. determine the constant A; note that B is not needed since we allowed m_l to absorb the sign). The normalization integral is:

$$\int_0^{2\pi} \Phi^* \Phi d\phi = |A|^2 \int_0^{2\pi} e^{-i m_l \phi} e^{i m_l \phi} d\phi = |A|^2 \int_0^{2\pi} f\phi = 2\pi |A|^2$$

$$\Rightarrow A = \frac{1}{\sqrt{2\pi}} \text{ (chosen to be real)}$$

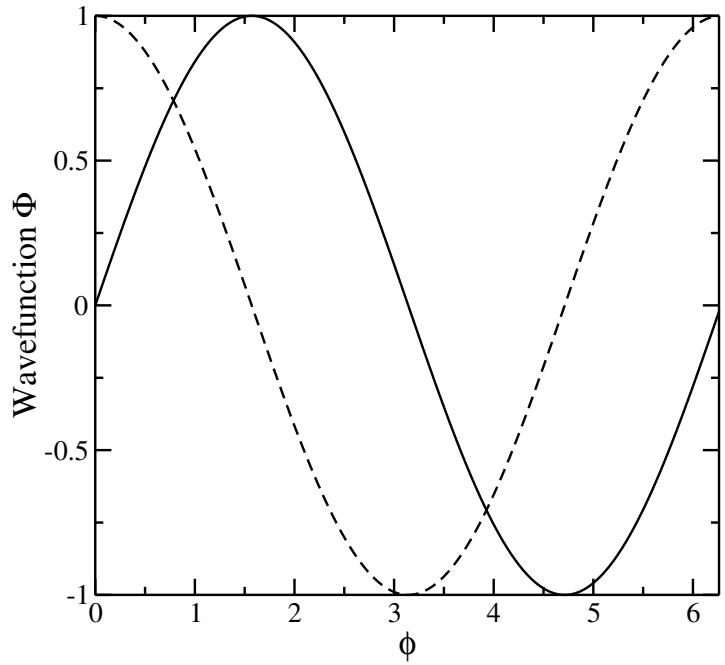
Excercise. Show that the eigenfunctions Φ_{m_l} form an orthogonal set of functions.

3.3 The shapes of the wavefunctions

To understand the shapes of the previous rotational wavefunctions, we expand the complex exponential term:

$$\Psi_{m_l} = \sqrt{\frac{1}{2\pi}} e^{i m_l \phi} = \sqrt{\frac{1}{2\pi}} \left(\cos(m_l \phi) + i \sin(m_l \phi)\right) \tag{3.11}$$

This wavefunction is complex when $m_l \neq 0$. The real and imaginary parts are shown below.



Note that eigenstates with $m_l \neq 0$ are doubly degenerate as the energy depends on m_l^2. Furthermore, there is no zero-point energy as the energy for $m_l = 0$ is zero. For all states the probability density is independent of the angle ϕ:

$$|\Phi_{m_l}|^2 = \sqrt{\frac{1}{2\pi}} e^{-im_l\phi} \sqrt{\frac{1}{2\pi}} e^{im_l\phi} = \frac{1}{2\pi} \tag{3.12}$$

Later we will find the following result useful:

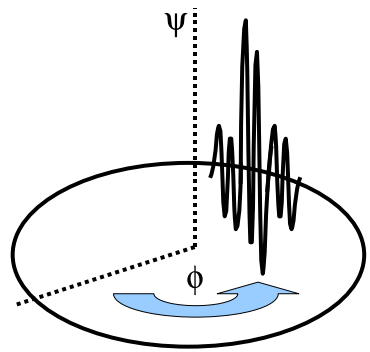
$$\Phi_{m_l}(\phi + \pi) = \sqrt{\frac{1}{2\pi}} e^{im_l(\phi+\pi)} = \sqrt{\frac{1}{2\pi}} e^{im_l\phi} \left(e^{i\pi}\right)^{m_l} = (-1)^{m_l}\, \Phi_{m_l}(\phi) \tag{3.13}$$

3.4 The classical limit

It is possible to preapre a wavepacket from the rotational eigenfunctions. In this each angular momentum component is written as:

$$\Psi_{m_l} = \frac{1}{\sqrt{2\pi}} e^{-m_l \phi - i m_l^2 \hbar t/(2I)} \tag{3.15}$$

By adding these functions weighted by the shape function (see Eq. (2.11)) we can obtain the rotational wavepacket. An example of a rotational wavepacket is shown below.

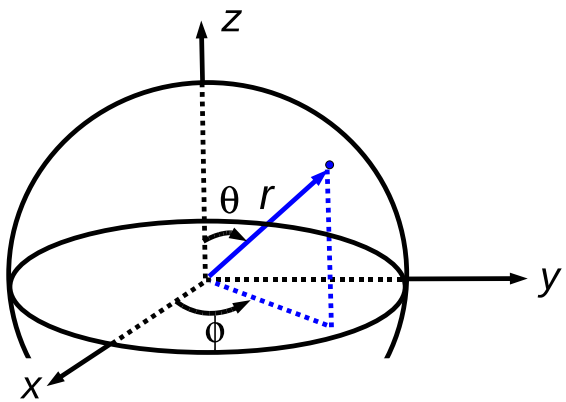


3.5 Particle on a sphere: The Schrödinger equation and its solution

Next we consider a particle moving on a spherical shell (rather than in a circle), which means that we will need to consider two angular variables, i.e., spherical coordinates with r fixed. Such system can be, for example, a diatomic molecule that rotates about its center of mass. The following calculation will also be useful when we solve the hydrogen atom problem. Here the potential energy is also zero and we need to consider cyclic boundary conditions for the angular variables. The Hamiltonian is then simply:

$$H = -\frac{\hbar^2}{2m}\nabla^2 \tag{3.16}$$

This is most convenient to express in spherical coordinates:



$$\begin{aligned} x &= r\sin(\theta)\cos(\phi) \\ y &= r\sin(\theta)\sin(\phi) \\ z &= r\cos(\theta) \end{aligned} \tag{3.17}$$

$$\theta \in [0,\pi]\,,\ \phi \in [0,2\pi] \ \text{ and } \ r \in [0,\infty]$$

The laplacian can be written in spherical coordinates in various different equivalent forms (see a mathematics tablebook):

$$\nabla^2 = \frac{1}{r}\frac{\partial^2}{\partial r^2} r + \frac{1}{r^2}\Lambda^2 \tag{3.18}$$

$$\nabla^2 = \frac{1}{r^2}\frac{\partial}{\partial r} r^2 \frac{\partial}{\partial r} + \frac{1}{r^2}\Lambda^2$$

$$\nabla^2 = \frac{\partial^2}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\Lambda^2$$

The *legendrian* Λ^2 is the angular part of the laplacian:

$$\Lambda^2 = \frac{1}{\sin^2(\theta)}\frac{\partial^2}{\partial \phi^2} + \frac{1}{\sin(\theta)}\frac{\partial}{\partial \theta}\sin(\theta)\frac{\partial}{\partial \theta} \tag{3.19}$$

If the particle is confined to a spherical shell then r is constant in the laplacian and then the hamiltonian reduces to:

$$H = -\frac{\hbar^2}{2mr^2}\Lambda^2 \tag{3.20}$$

Since the moment of inertia $I = mr^2$ the Schrödinger equation now reads:

$$\Lambda^2 \psi(\theta, \phi) = -\left(\frac{2IE}{\hbar^2}\right) \psi(\theta, \phi) \tag{3.21}$$

By multiplying the Schrödinger Eq. (3.21) by $\sin^2(\theta)$, we can separate θ and ϕ parts. Thus the solution must be $\psi(\theta, \phi) = \Theta(\theta)\Phi(\phi)$. For Θ the equation becomes:

$$\frac{d^2\Phi}{d\phi^2} = \text{constant} \times \Phi$$

This equation with cyclic boundary conditions we have already seen (particle confined to a ring). The eigenfunctions were given in Eq. (3.11) and are labelled by the quantum number m_l. Solution for Θ is more complicated and we skip the derivation. The combined $\Phi\Theta$ functions are called the *spherical harmonics* $Y_{l,m_l}(\theta, \phi)$. They satisfy the required eigenvalue equation:

$$\Lambda^2 Y_{l,m_l} = -l(l+1) Y_{l,m_l} \tag{3.22}$$

where the quantum numbers l and m_l may have the following values:

$$l = 0, 1, 2, ... \text{ and } m_l = -l, -l+1, ..., 0, ..., l-1, l$$

Spherical harmonics are composed of two factors:

$$Y_{l,m_l}(\theta,\phi) = \Theta_{l,m_l}(\theta)\Phi_{m_l}(\phi) \tag{3.23}$$

where Φ are solutions for the particle in a ring problem and Θ are called *associated Legendre functions*. Some spherical harmonics are listed below.

$$Y_{0,0} = \frac{1}{2\sqrt{\pi}},\ ,\ Y_{1,0} = \sqrt{\frac{3}{4\pi}}\cos(\theta),\ Y_{1,1} = -\sqrt{\frac{3}{8\pi}}\sin(\theta)e^{i\phi}$$

$$Y_{1,-1} = \sqrt{\frac{3}{8\pi}}\sin(\theta)e^{-i\phi},\ Y_{2,0} = \sqrt{\frac{5}{16\pi}}(3\cos^2(\theta)-1),\ Y_{2,1} = -\sqrt{\frac{15}{8\pi}}\sin(\theta)\cos(\theta)e^{i\phi}$$

$$Y_{2,-1} = \sqrt{\frac{15}{8\pi}}\sin(\theta)\cos(\theta)e^{-i\phi},\ Y_{2,2} = \sqrt{\frac{15}{32\pi}}\sin^2(\theta)e^{2i\phi},\ Y_{2,-2} = \sqrt{\frac{15}{32\pi}}\sin^2(\theta)e^{-2i\phi}$$

ETC.

Comparison of Eqs. (3.21) and (3.22) shows that (note degeneracy with respect to m_l):

$$E_{l,m_l} = l(l+1)\frac{\hbar^2}{2I} \tag{3.24}$$

Example. Show that the spherical harmonic $Y_{1,0}$ is a solution of Eq. (3.22).

Solution. We substitute $Y_{1,0}$ into Eq. (3.22):

$$\Lambda^2 Y_{1,0} = \frac{1}{\sin(\theta)} \frac{\partial}{\partial\theta} \sin(\theta) \frac{\partial}{\partial\theta} N\cos(\theta) = -N \frac{1}{\sin(\theta)} \frac{d}{d\theta} \sin^2(\theta)$$

$$= -2N \frac{1}{\sin(\theta)} \sin(\theta)\cos(\theta) = -2Y_{1,0}$$

where here N is the normalization factor for $Y_{1,0}$. The following Maxima program can be used to evaluate spherical harmonics. Maxima follows the Condon-Shortley convention but may have a different overall sign than in the previous table.

```
load(orthopoly);
/* spherical_harmonic(l, m, theta, phi) */
spherical_harmonic(0, 0, theta, phi); /* l = 0, m = 0 */
spherical_harmonic(1, 1, theta, phi);       /* l = 1, m = 1 */
```

3.6 The angular momentum of the particle

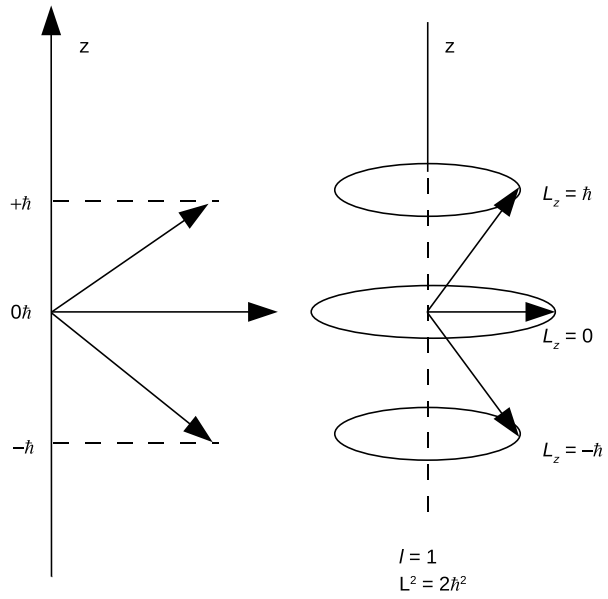
The quantum numbers l and m_l are directly related to the magnitude and projected value of angular momentum. To see this, we first use expressions from classical physics. The rotational energy is given by $E = \frac{1}{2}I\omega^2 = \frac{l^2}{2m}$ where I is the moment of inertia, ω is the angular momentum and l is the magnitude of the angular momentum. If one compares this expression with Eq. (3.24) we get:

$$|l| = \sqrt{l(l+1)}\hbar \text{ or } l^2 = l(l+1)\hbar^2 \tag{3.25}$$

where $|l|$ denotes the magintude of angular momentum. This is why l is called the *angular momentum quantum number* as it is directly related to the magnitude of angular momentum. The spherical harmonics are also eigenfunctions of l_z:

$$l_z Y_{l,m_l} = \frac{\hbar}{i}\left(\Theta_{l,m_l}\frac{e^{im_l\phi}}{\sqrt{2\pi}}\right) = m_l\hbar Y_{l,m_l} \tag{3.26}$$

which shows that m_l specifies the projection of angular momentum on the z-axis. Because m_l is restricted to only certain values, l_z must also obey this restriction (*space quantitization*). The term space quantitization follows from the vector model of angular momentum:

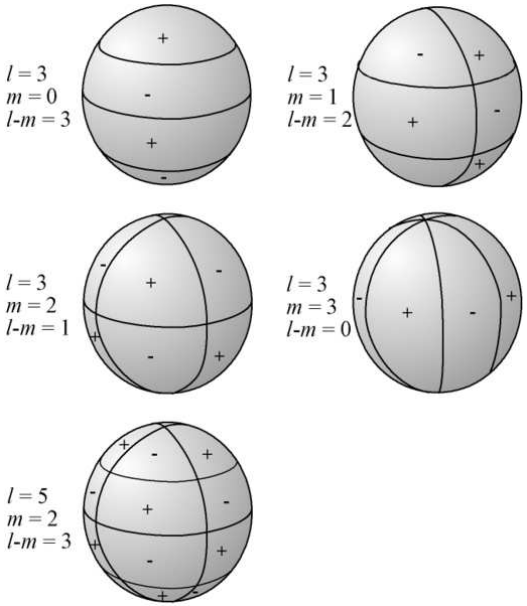


The angle between the z-axis and the $\vec{l}$ is given by:

$$\cos(\theta) = \frac{m_l}{\sqrt{l(l+1)}} \qquad (3.27)$$

Since l_x, l_y and l_z do not commute, we cannot specify l_x and l_y exactly. This uncertainty is drawn on the left with a circle extending around the z-axis.

3.7 The properties of the solutions



Representation of Y_{l,m_l} on a unit sphere. Y_{l,m_l} is equal to 0 (i.e. *angular node*) along m_l great circles passing through the poles, and along $l - m_l$ circles of equal latitude. The function changes sign each time it crosses one of these lines. Note that it is possible to create rotational wavepackets, which have a well defined orientation.

3.8 The rigid rotor

Now we will apply the previous theory to describe rotation of diatomic molecule where we assume that the bond length is fixed ("rigid rotor"). The masses of the nuclei are represented by m_1 and m_2 and their separation is fixed at distance R. The total Hamiltonian for this probelm is:

$$H = -\frac{\hbar^2}{2m_1}\nabla_1^2 - \frac{\hbar^2}{2m_2}\nabla_2^2 \tag{3.28}$$

where ∇_i^2 differentiates with respect to the coordinates of particle i. This can be transformed as (see Atkins & Friedman, Molecular Quantum Mechanics, Further information 4):

$$\frac{1}{m_1}\nabla_1^2 + \frac{1}{m_2}\nabla_2^2 = \frac{1}{m}\nabla_{\mathrm{cm}}^2 + \frac{1}{\mu}\nabla^2$$

The subscript cm refers to the variables describing the center of mass translational movement of the molecule, the Laplacian without subscript refers to the internal coordinates of the molecule (vibration and rotation), $m = m_1 + m_2$ and:

$$\frac{1}{\mu} = \frac{1}{m_1} + \frac{1}{m_2} \text{ or } \mu = \frac{m_1 m_2}{m_1 + m_2} \tag{3.29}$$

where μ is called the *reduced mass*.

Now we can rewrite the Schrödinger equation as:

$$-\frac{\hbar^2}{2m}\nabla^2_{\mathrm{cm}}\Psi - \frac{\hbar^2}{2\mu}\nabla^2\Psi = E_{\mathrm{total}}\Psi \tag{3.30}$$

Since the Hamiltonia is a sum of two terms depending on different variables, it is possible to separate this equation as:

$$-\frac{\hbar^2}{2m}\nabla^2_{\mathrm{cm}}\psi_{\mathrm{cm}} = E_{\mathrm{cm}}\psi_{\mathrm{cm}} \tag{3.31}$$

$$-\frac{\hbar^2}{2\mu}\nabla^2\psi = E\psi$$

with $E_{\mathrm{total}} = E_{\mathrm{cm}} + E$. The first equation is related to free translation of the molecule where as the second equation describes both rotation and vibration of the (diatomic) molecule. The 2nd equation can be solved in spherical coordinates and simplified with the rigid rotor assumption (R is constant). Thus the radial derivatives in the Laplacian are zero and only the angular part is retained:

$$-\frac{\hbar^2}{2\mu R^2}\Lambda^2\psi = E\psi \tag{3.32}$$

To simplify notation we write $I = \mu R^2$. This is identical to Eq. (3.21). The solutions can be labelled by two quantum numbers l and m_l, which in the context of molecular rotation, are usually written as J and M_J.

The eigenfunctions for the rotating diatomic molecule are hence spherical harmonics Y_{J,M_J} and the rotational energy levels are given by:

$$E_{J,M_J} = J(J+1)\frac{\hbar^2}{2I} \tag{3.33}$$

where $J = 0, 1, 2, ...$ and $M_J = 0, \pm 1, \pm 2, ..., \pm J$. Because the energy is independent of M_J, each rotational level is $2J + 1$ times degenerate.

Example. What are the reduced mass and moment of inertia of $H^{35}Cl$? The equilibrium internuclear distance R_e is 127.5 pm (1.275 Å). What are the values of L, L_z and E for the state with $J = 1$? The atomic masses are: $m_H = 1.673470 \times 10^{-27}$ kg and $m_{Cl} = 5.806496 \times 10^{-26}$ kg.

Solution. First we calculate the reduced mass (Eq. (3.29)):

$$\mu = \frac{m_H m_{^{35}Cl}}{m_H + m_{^{35}Cl}} = \frac{(1.673470 \times 10^{-27}\ \text{kg})(5.806496 \times 10^{-26}\ \text{kg})}{(1.673470 \times 10^{-27}\ \text{kg}) + (5.806496 \times 10^{-26}\ \text{kg})}$$

$$= 1.62665 \times 10^{-27}\ \text{kg}$$

Next, $I = \mu R^2$ gives the moment of inertia:

$$I = \mu R_e^2 = (1.626 \times 10^{-27}\ \text{kg})(127.5 \times 10^{-12}\ \text{m})^2 = 2.644 \times 10^{-47}\ \text{kg m}^2$$

L is given by Eq. (3.25):

$$L = \sqrt{J(J+1)}\hbar = \sqrt{2}\left(1.054 \times 10^{-34}\ \text{Js}\right) = 1.491 \times 10^{-34}\ \text{Js}$$

L_z is given by Eq. (3.26):

$$L_z = -\hbar, 0, \hbar \text{ (three possible values)}$$

Energy of the $J = 1$ level is given by Eq. (3.33) (or Eq. (3.24)):

$$E = \frac{\hbar^2}{2I} J(J+1) = \frac{\hbar^2}{I} = 4.206 \times 10^{-22}\ \text{J} = 21\ \text{cm}^{-1}$$

This rotational spacing can be, for example, observed in gas phase infrared spectrum of HCl.

3.9 The Schrödinger equation for hydrogenic atoms

In the following we will obtain analytic solution to a problem that consists of one positively charged nucleus and one electron ("hydrogenic atom"). The nuclear charge is denoted by Z, which, for example, would be one for hydrogen atom. The solution will be ontained in spherical coordinates where we can separate the equation into radial and angular parts. The angular part we have already solved previously. The Hamiltonian for the two-particle electron-nucleus system is:

$$H = -\frac{\hbar^2}{2m_e}\nabla_e^2 - \frac{\hbar^2}{2m_N}\nabla_N^2 - \frac{Ze^2}{4\pi\epsilon_0 r} \tag{3.34}$$

where m_e is the electron mass, m_N is the mass of the nucleus, and ∇_e^2 and ∇_N^2 are the Laplacians with respect to electron and nuclear coordinates, respectively. The quantity ϵ_0 in the Coulomb potential term is a constant (*vacuum permittivity*). This Hamiltonian is very close to our previous rotational Hamiltonian apart from the Coulomb potential (attraction between + and − charges). Again, we can convert to center-of-mass and relative coordinates, the Halmitonian becomes:

$$H = -\frac{\hbar^2}{2m}\nabla_{\mathrm{cm}}^2 - \frac{\hbar^2}{2\mu}\nabla^2 - \frac{Ze^2}{4\pi\epsilon_0 r} \tag{3.35}$$

where $m = m_e + m_N$ and the reduced mass μ is given by Eq. (3.29).

This can be separated into the translational and relative motion parts. The latter can be written as:

$$-\frac{\hbar^2}{2\mu}\nabla^2\psi - \frac{Ze^2}{4\pi\epsilon_0 r}\psi = E\psi \tag{3.36}$$

Unlike the rigid rotor problem, the distance between the electron and the nucleus is not constant and therefore we need to include the radial part in our Laplacian. The Schrödinger equation now becomes:

$$\frac{1}{r}\frac{\partial^2}{\partial r^2}r\psi + \frac{1}{r^2}\Lambda^2\psi + \frac{Ze^2\mu}{2\pi\epsilon_0\hbar^2 r}\psi = -\left(\frac{2\mu E}{\hbar^2}\right)\psi \tag{3.37}$$

3.10 The separation of the relative coordinates

To separate Eq. (3.37) into radial and angular parts, we write a product trial solution:

$$\psi(r,\theta,\phi) = R(r)Y(\theta,\phi) \tag{3.38}$$

where Y is spherical harmonic function. When this is substituted into the Schrödinger equation and use $\Lambda^2 Y = -l(l+1)Y$, we get:

$$\frac{1}{r}\frac{\partial^2}{\partial r^2} rRY - \frac{l(l+1)}{r^2}RY + \left(\frac{Ze^2\mu}{2\pi\epsilon_0\hbar^2 r}\right)RY = -\left(\frac{2\mu E}{\hbar^2}\right)RY$$

Both sides of the equation can be divided by Y, which leaves us with the *radical Schrödinger equation*:

$$\frac{1}{r}\frac{d^2(rR)}{dr^2} + \left(\frac{Ze^2\mu}{2\pi\epsilon_0\hbar^2 r} - \frac{l(l+1)}{r^2}\right)R = -\left(\frac{2\mu E}{\hbar^2}\right)R$$

This can be simplified by $u = rR$:

$$\frac{d^2u}{dr^2} + \left(\frac{2\mu}{\hbar^2}\right)\left(\frac{Ze^2}{4\pi\epsilon_0 r} - \frac{l(l+1)\hbar^2}{2\mu r^2}\right)u = -\left(\frac{2\mu E}{\hbar^2}\right)u \tag{3.39}$$

This is a one dimensional Schrödinger equation with an *effective potential* V_{eff}:

$$V_{\text{eff}} = -\frac{Ze^2}{4\pi\epsilon_0 r} + \frac{l(l+1)\hbar^2}{2\mu r^2} \tag{3.40}$$

3.11 The radial Schrödinger equation

The first term appearing in the effective potential (Eq. (3.40)) is the Coulomb interaction between the electron and the nucleus. The second term corresponds to a centrifugal force that impels he electron away from the nucleus. When $l = 0$ (i.e. spherical s-orbital), the centrifugal term is zero. When $l > 0$ the angular momentum drags the electron farther away from the nucleus. States with $l = 0$ and $l > 0$ have very different behavior near the nucleus: For $l = 0$ the effective potential tends toward $-\infty$ whereas for $l > 0$ it approaches $+\infty$. This implies that the $l > 0$ states have a node at the nucleus whereas $l = 0$ do not.

The solutions to the radical Schrödinger equation are related to the *associated Laguerre polynomials* (derivation not shown). Each Laguerre polynomial is labelled by two numbers n and l where $n = 1, 2, ...$ and $l = 0, 1, ..., n-1$. The overall solution R is a product of the normalization factor, the associated Laguerre polynomial L and the exponential factor:

$$R_{n,l}(r) = -\left\{\left(\frac{2Z}{na}\right)^3 \frac{(n-l-1)!}{2n\left((n+l)!\right)^3}\right\} \rho^l L_{n+l}^{2l+1}(\rho) e^{-\rho/2}$$

where $\rho = \frac{2Zr}{na_0}$ and a_0 is the Bohr radius, which defined as:

$$a_0 = \frac{4\pi\epsilon_0\hbar^2}{m_e e^2} \approx 52.9 \text{ pm} \tag{3.43}$$

Orbital	n	l	R_{nl}
1s	1	0	$2\left(\frac{Z}{a_0}\right)^{3/2} e^{-\rho/2}$
2s	2	0	$\frac{1}{2\sqrt{2}}\left(\frac{Z}{a_0}\right)^{3/2} (2-\rho)e^{-\rho/2}$
2p	2	1	$\frac{1}{2\sqrt{6}}\left(\frac{Z}{a_0}\right)^{3/2} \rho e^{-\rho/2}$
3s	3	0	$\frac{1}{9\sqrt{3}}\left(\frac{Z}{a_0}\right)^{3/2} (6-6\rho-\rho^2)e^{-\rho/2}$
3p	3	1	$\frac{1}{9\sqrt{6}}\left(\frac{Z}{a_0}\right)^{3/2} (4-\rho)\rho e^{-\rho/2}$
3d	3	2	$\frac{1}{9\sqrt{30}}\left(\frac{Z}{a_0}\right)^{3/2} \rho^2 e^{-\rho/2}$

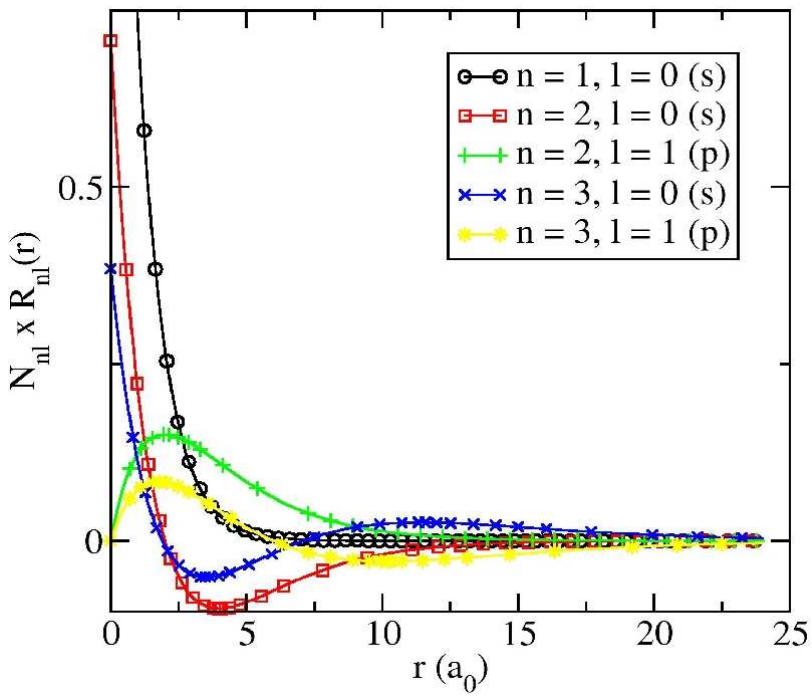
Table: Examples of the radial wavefunctions for hydrogenlike atoms.

- When $l = 0$ the radical wavefunction has non-zero value at $r = 0$. When $l > 0$, the radial wavefunction is zero at $r = 0$.
- Each radical wavefunction has $n - l - 1$ nodes in it (we exclude the zero at $r = 0$ for $l > 0$).

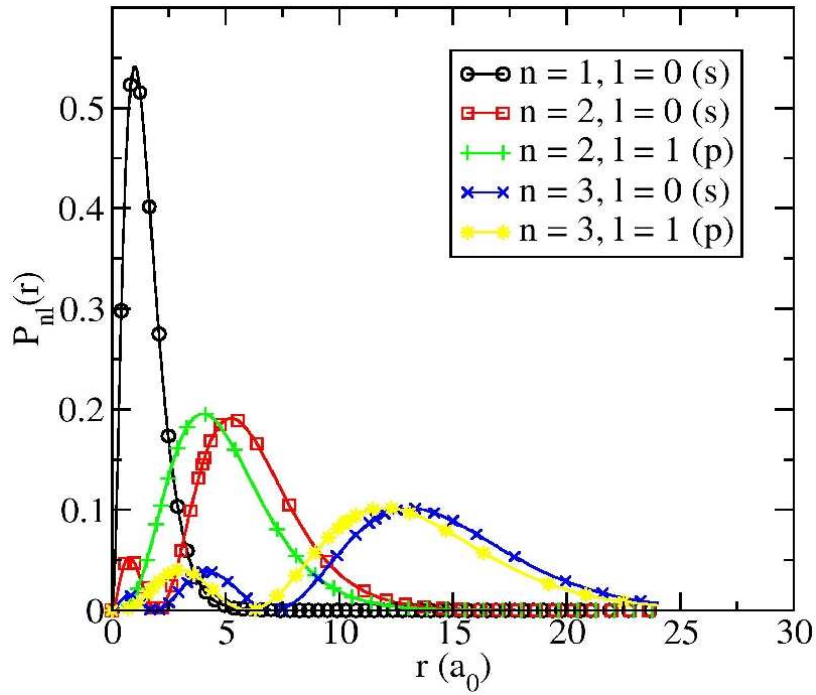
Inserting the radial wavefunctions into Eq. (3.39) gives the energy:

$$E_n = -\left(\frac{Z^2 \mu e^4}{32\pi^2 \epsilon_0^2 \hbar^2}\right) \frac{1}{n^2} \tag{3.44}$$

Quantum #	Values	E contrib.	Meaning
n	$1, 2, ...$	YES	The principal quantum numer
l	$l < n$	NO	Total angular momentum $(s, p, ...)$
m_l	$-l, -l+1, ..., 0, l-1, l$	NO	Projection of angular momentum.



Radial wavefunctions.



Radial probabilities ($\times r^2$; see next section).

The following Maxima program can be used to plot te radial wavefunctions on the previous page (if wxMaxima is used, replace plot2d with wxplot2d):

```
/* Plot normalized radial wavefunctions for hydrogen atom */
load(orthopoly);
Z:1; /* Nuclear charge 1 (hydrogen) */
a0:1;/* Output in the units of Bohr */
/* Definition of the radial wavefunction */
R(r,n,l):= (
  rho:2*Z*r/(n*a0),  /* definition of rho */
  norm:sqrt((2*Z/(n*a0))^3 * (n - l - 1)! / (2*n*((n + l)!))),
  norm * rho^l * exp(-rho/2) * gen_laguerre(n-l-1,2*l+1,rho)
  );
/* Plot the radial wavefunction (range from 0 to 24) */
plot2d(
  [R(r,1,0),R(r,2,0),R(r,2,1),R(r,3,0),R(r,3,1)],
  [r,0,24],
  [
    legend,"n = 1, l = 0(s)", "n = 2, l = 0(s)",
    "n = 2, l = 1(p)", "n = 3, l = 0(s)", "n = 3, l = 1(p)"
  ],
  [xlabel,"r(a0)"],
  [ylabel,"Nnl x Rnl(r)"]
  );
```

3.12 Probabilities and the radial distribution function

The complete wavefunction for hydrogenic atoms is:

$$\psi_{n,l,m_l} = R_{n,l} Y_{l,m_l}$$

where $R_{n,l}$ are the associated Laguerre functions and the Y_{l,m_l} are the spherical harmonics. In spherical coordinates the volume element for integration is $d\tau = r^2 \sin(\theta) d\theta d\phi dr$. We need to consider integratetion of $\left|\psi_{n,l,m_l}\right|^2 d\tau$ over a given region in space. If we want to calculate the *radial distribution function* $P(r)$, we have to integrate over the angular variables:

$$P(r)dr = \int_0^{\pi} \int_0^{2\pi} R_{n,l}^2 \left|Y_{l,m_l}\right|^2 r^2 \sin(\theta) d\theta d\phi dr \tag{3.45}$$

Since spherical harmonics are normalized in a sense that

$$\int_0^{\pi} \int_0^{2\pi} \left|Y_{l,m_l}\right|^2 \sin(\theta) d\theta d\phi = 1$$

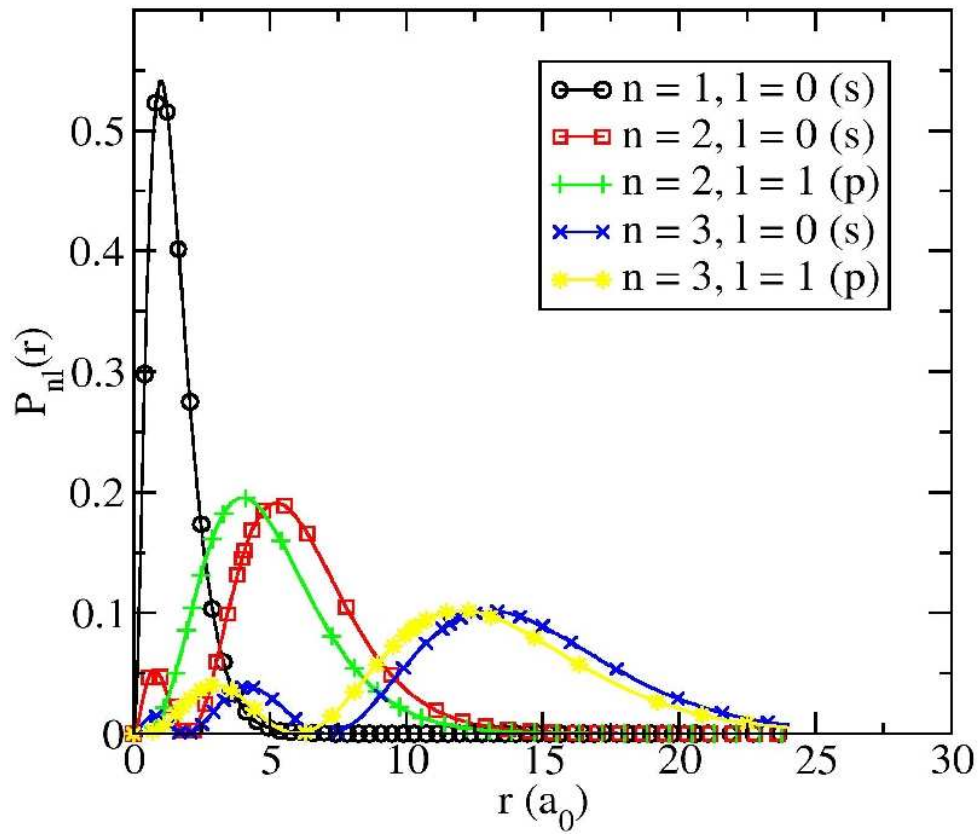
We then can calculate the probability for the electron to be found between r and $r + dr$ as:

$$P(r)dr = R_{n,l}^2 r^2 dr \tag{3.46}$$

For 1s orbital ($n = 1$ and $l = 0$) we get the following expression:

$$P(r) = 4\left(\frac{Z}{a_0}\right)^3 r^2 e^{-2Zr/a_0}$$

We plotted $P(r)$ earlier, which shows the radial behavior of the wavefunctions:



Note that all wavefunctions now approach zero when $r \to 0$ or $r \to \infty$. For $1s$ wavefunction, it is easy to see that the maximum occurs at:

$$r_{\max} = \frac{a_0}{Z} \tag{3.47}$$

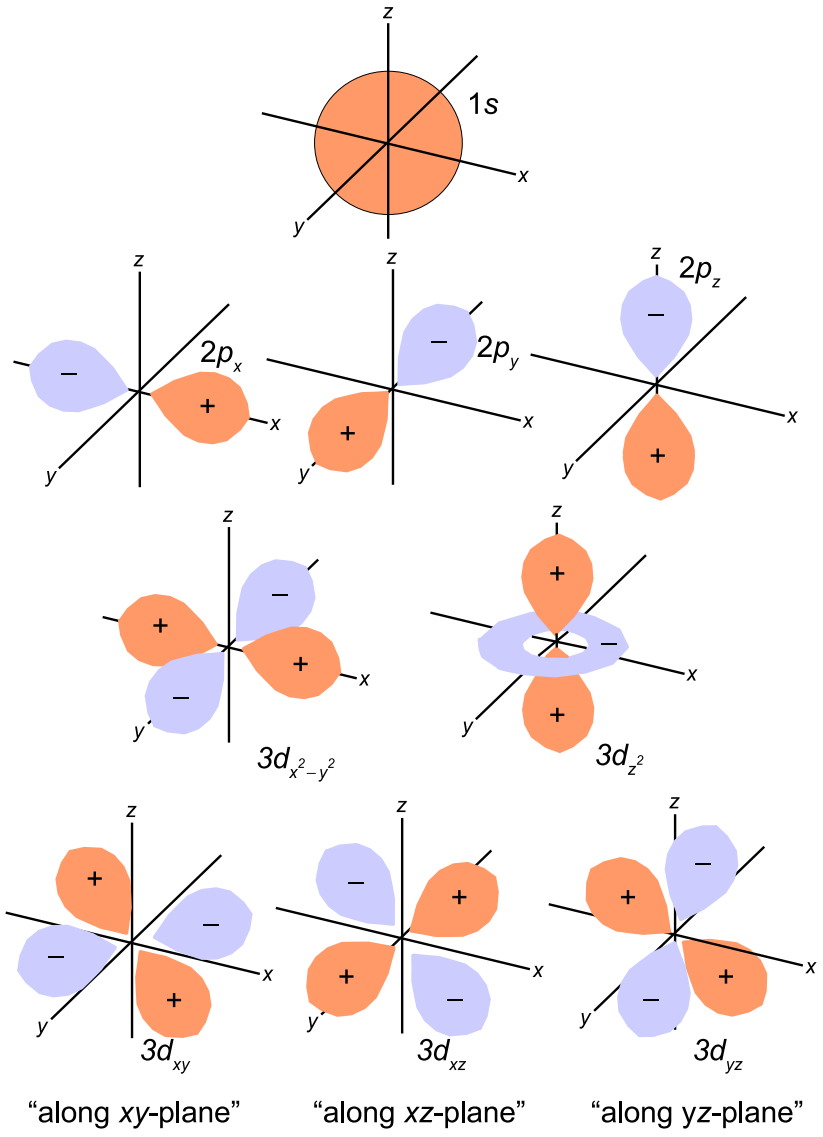
by solving for r in $dP(r)/dr = 0$. For hydrogen atom this gives $r_{\max} = a_0$. Note that the most probable radius decreases when the nuclear charge Z increases.

3.13 Atomic orbitals

One electron wavefunctions are called *atomic orbitals*. The atomic orbitals are labelled as s-orbital when $l = 0$, p-orbital when $l = 1$, etc. When electron is described by the wavefunction ψ_{n,l,m_l} we say that the electron *occupies* the orbital. An electron that occupies s-orbital is called an s-electron and similarly for the other orbitals.

There are two common ways to represent an orbital graphically: 1) by denoting the electron density by shading or 2) by drawing *boundary surface*. In the latter case the boundary surface is often chosen in such way that 90% of the electron density resides inside the surface. For real orbitals, the sign is often denoted by writing $+$ or $-$ on the relevant lobes of the orbitals. Note that this representation only gives information about the shape and the radical extent of the orbital. Examples of atomic orbitals are shown on the next page.

For p-orbitals ($l = 1$), we have three different m_l values allowed ($m_l = +1, 0, -1$). These are often labelled as p_{+1}, p_0, and p_{-1}. The orbital with $m_l = 0$ is also called a p_z*-orbital*. The other two orbitals p_{+1} and p_{-1} are complex, and their maximum amplitude lies in the xy-plane.



Atomic orbital boundary surfaces.

The wavefunctions for p-orbitals are given by:

$$p_0 = p_z = \sqrt{\frac{3}{4\pi}} R_{n,l}(r)\cos(\theta) \tag{3.48}$$
$$p_{+1} = -\sqrt{\frac{3}{8\pi}} R_{n,l}(r)\sin(\theta)e^{i\phi}$$
$$p_{-1} = \sqrt{\frac{3}{8\pi}} R_{n,l}(r)\sin(\theta)e^{-i\phi}$$

Since p-orbitals are degenerate, linear combinations of the above functions are also eigenfunctions. We can write the *Cartesian p-orbitals* as:

$$p_x = \frac{1}{\sqrt{2}}(p_{-1} - p_{+1}) = \sqrt{\frac{3}{4\pi}} R_{n,l}(r)\sin(\theta)\cos(\phi) \tag{3.49}$$
$$p_y = \frac{i}{\sqrt{2}}(p_{-1} + p_{+1}) = \sqrt{\frac{3}{4\pi}} R_{n,l}(r)\sin(\theta)\sin(\phi)$$

When there are well-defined x and y axes (such non-linear molecules), it is more approriate to use the Cartesian forms. The p_x, p_y, and p_z orbitals are aligned along the correponding axes. Note that they change sign depending on passing through the origin along the axis.

d-orbitals are five fold degenerate ($n > 2$). All orbitals, except $m_l = 0$, are complex and correspond to definite states of orbital angular momentum around the z-axis. It is possible to express these orbitals in their Cartesian forms as follows:

$$d_{z^2} = d_0 = \sqrt{\frac{5}{16\pi}} R_{n,2}(r) \left(3\cos^2(\theta) - 1\right) = \sqrt{\frac{5}{16\pi}} R_{n,2}(r) \left(3z^2 - r^2\right)/r^2$$

$$d_{x^2-y^2} = \frac{1}{\sqrt{2}} \left(d_{+2} + d_{-2}\right) = \sqrt{\frac{15}{16\pi}} R_{n,2}(r) \left(x^2 - y^2\right)/r^2$$

$$d_{xy} = \frac{1}{i\sqrt{2}} \left(d_{+2} - d_{-2}\right) = \sqrt{\frac{15}{4\pi}} R_{n,2}(r) xy/r^2$$

$$d_{yz} = \frac{1}{i\sqrt{2}} \left(d_{+1} + d_{-1}\right) = -\sqrt{\frac{15}{4\pi}} R_{n,2}(r) yz/r^2$$

$$d_{zx} = \frac{1}{\sqrt{2}} \left(d_{+1} - d_{-1}\right) = -\sqrt{\frac{15}{4\pi}} R_(n,2)(r) zx/r^2$$

Given the analytic forms for the orbitals, it is now possible to calculate all their properties. For example, it is possible to obtain the mean radius for a given orbital by using r as the operator (or its inverse $1/r$):

$$\langle r \rangle_{n,l,m_l} = \frac{n^2 a_0}{Z}\left(1 + \frac{1}{2}\left(1 - \frac{l(l+1)}{n^2}\right)\right) \tag{3.50}$$

$$\left\langle \frac{1}{r} \right\rangle_{n,l,m_l} = \frac{Z}{a_0 n^2}$$

Note that the mean radius for ns orbital is larger than for np orbital.

3.14 The degeneracy of hydrogenic atoms

The energies of hydrogen atom orbitals depdend only on the principal quantum number n. This means that each shell n has the following degeneracy factor:

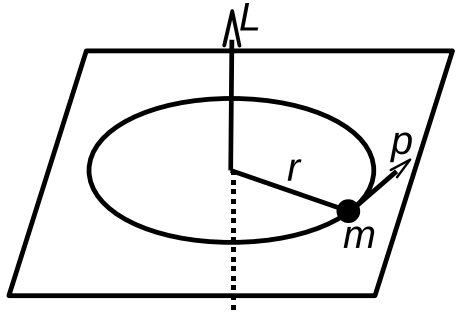
$$g_n = \sum_{l=0}^{n-1} (2l+1) = n^2 \tag{3.52}$$

Here all ns, np, nd, ... levels have the same energy. Each $np_{x,y,z}$, etc. are clearly related to each other by symmetry but the degeneracy between different l orbitals is due to hidden symmetry. In fact, in many atom molecules this symmetry is broken and only the m_l components belonging to the given l level remain degenerate. The degeneracy between the m_l levels can be removed by applying electric or magnetic field on the system.

Chapter 4: Angular momentum

4.1 The angular momentum and their commutation relations

In classical mechanics, the *angular momentum*, l, of a particle travelling with a linear momentum p at a position r on its path is defined as the vector product $l = r \times p$:



The direction of the vector l is given by the right-hand screw rule. It points perpendicular to the plane of rotation.

If the position $r = xi + yj + zk$ and momentum $p = p_x i + p_y j + p_z k$ are multiplied out, we get:

$$l = r \times p = (yp_z - zp_y)i + (zp_x - xp_z)j + (xp_y - yp_x)k \tag{4.2}$$

where the components of l can be identified:

$$l_x = yp_z - zp_y,\ l_y = zp_x - xp_z,\ l_z = xp_y - yp_x \tag{4.3}$$

The magnitude of the angular momentum is obtained through:

$$l^2 = l_x^2 + l_y^2 + l_z^2 \tag{4.4}$$

Transition to quantum mechanics can be carried out by replacing the classical observables by the corresponding operators:

$$l_x = \frac{\hbar}{i}\left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}\right),\ l_y = \frac{\hbar}{i}\left(z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z}\right),\ l_z = \frac{\hbar}{i}\left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right) \tag{4.5}$$

The commutation relations between different Cartesian components of l can be calculated as follows:

$$\begin{aligned}[l_x, l_y] = [yp_z - zp_y, zp_x - xp_z] = [yp_z, zp_x] - [yp_z, xp_z] - [zp_y, zp_x] + [zp_y, xp_z] \\ y\,[p_z, z]\,p_x - 0 - 0 + xp_y\,[z, p_z] = i\hbar(-yp_x + xp_y) = i\hbar l_z\end{aligned} \tag{4.6}$$

Note that above, for example, y and p_x commute since they depend on different coordinates. In a similar way we can get the other commutators:

$$[l_x, l_y] = i\hbar l_z,\ \ [l_y, l_z] = i\hbar l_x,\ \ [l_z, l_x] = i\hbar l_y \tag{4.7}$$

For l^2 we have, since $\left[l_z^2, l_z\right] = 0$:

$$\left[l^2, l_z\right] = \left[l_x^2 + l_y^2 + l_z^2, l_z\right] = \left[l_x^2, l_z\right] + \left[l_y^2, l_z\right]$$

For the other two commutators above we have:

$$\left[l_x^2, l_z\right] = L_x l_x l_z - l_z l_x l_x = l_x l_x l_z - l_x l_z l_x + l_x l_z l_x - l_z l_x l_x$$
$$= l_x \left[l_x, l_z\right] + \left[l_x, l_z\right] l_x = -i\hbar \left(l_x l_y + l_y l_x\right)$$

In a similar way we have:

$$\left[l_y^2, l_z\right] = i\hbar \left(l_x l_y + l_y l_x\right)$$

Since these are of equal magniture but opposite signs, they cancel out and we get:

$$\left[l^2, l_q\right] = 0 \text{ where } q = x, y, z \tag{4.8}$$

An observable is called an angular momentum if it follows the above commutation rules. It turns out that, for example, electron spin follows these rules and hence we say that it has angular momentum.

4.2 Angular momentum observables

Earlier we saw that observables are complementary and restricted by the uncertainty relation if their operators do not commute. We just showed that l_x, l_y, l_z do not commute and hence are subject to the uncertainty. However, l^2 commutes with l_x, l_y, l_z and allows for accurate simultaneous determination. For example, one can determine l^2 and l_z exactly but then l_x and l_y could not be. For this reason we used the "cone of uncertainty" in our earlier vector model for angular momentum. The commutation relations can be written in a compact form:

$$l \times l = i\hbar l \tag{4.9}$$

which can be seen by expanding the left hand side cross product and using the previous commutation relations to identify the vector components.

4.3 The shift operators

In the context of the harmonic oscillator, we have already seen shift operators. Here we define these operators for angular momentum where they can be used to change the state (i.e. rise or lower). The rising operator l_+ and lowering operator l_- are defined as follows:

$$l_+ = l_x + il_y \text{ and } l_- = l_x - il_y \tag{4.10}$$

These can be inverted:

$$l_x = \frac{l_+ + l_-}{2} \text{ and } l_y = \frac{l_+ - l_-}{2i} \tag{4.11}$$

The following commutation relations will be useful:

$$[l_z, l_+] = \hbar l_+, \ [l_z, l_-] = -\hbar l_-, \ [l_+, l_-] = 2\hbar l_z \tag{4.12}$$

$$\left[l^2, l_\pm\right] = 0 \tag{4.13}$$

Excercise. Show that the above commutation relations hold.

4.4 The effect of the shift operators

In the following we will explore the properties of angular momentum (without reference to our earlier work with spherical harmonics). We use two quantum numbers λ and m_l to denote the simultaneous eigenvalues of l^2 and l_z. The eigenstates are now labelled as $|\lambda, m_l\rangle$. Here m_l is defined as:

$$l_z |\lambda, m_l\rangle = m_l \hbar |\lambda, m_l\rangle \tag{4.14}$$

This definition is correct since m_l is just a numerical factor multiplying $\hbar$ (which itself has the right units for angular momentum). Since l_z is hermitean, m_l must be real but at this point we don't know that it can only take discrete values.

Since l_z and l^2 commute, $|\lambda, m_l\rangle$ is also an eigenstate of l^2. The eigenvalue of l^2 we take to depend on both λ and m_l as follows:

$$l^2 |\lambda, m_l\rangle = f(\lambda, m_l)\hbar^2 |\lambda, m_l\rangle \tag{4.15}$$

where f is an unknown function (that we know to be $\lambda(\lambda + 1)$ based on our earlier work based on spherical harmonics). Since l^2 is hermitean, f must be a real valued function.

Show that the function f is non-negative.

Proof. First $l^2 = l_x^2 + l_y^2 + l_z^2 \Rightarrow l^2 - l_z^2 = l_x^2 + l_y^2$ and then:

$$\langle \lambda, m_l | \, l^2 - l_z^2 \, | \lambda, m_l \rangle = \langle \lambda, m_l | \, l_x^2 + l_y^2 \, | \lambda, m_l \rangle$$

$$= \langle \lambda, m_l | \, l_x^2 \, | \lambda, m_l \rangle + \langle \lambda, m_l | \, l_y^2 \, | \lambda, m_l \rangle \geq 0$$

where the last inequality follows from the fact that squares of hermitean operators give non-negative expectation values (Exercise: expand $|\lambda, m_l\rangle$ in eigenfunctions of l_x or l_y and inspect the terms in the resulting sum). We can also write:

$$\langle \lambda, m_l | \, (f(\lambda, m_l) - m_l^2)\hbar^2 \, | \lambda, m_l \rangle = \langle \lambda, m_l | \, l^2 - l_z^2 \, | \lambda, m_l \rangle \geq 0$$

since $|\lambda, m_l\rangle$ are eigenfunctions of both l^2 and l_z. This means that:

$$f(\lambda, m_l) \geq m_l^2 \geq 0 \tag{4.16}$$

where the last inequality follows from the hermiticity of l_z.

Next we will see how the shift operators operate on a given eigenstate. First we note that since l_+ and l_- are defined in terms of l_x and l_y, which do not have $|\lambda, m_l\rangle$ as their eigenfunctions, and one expects that a new state will be created after their operation.

1. Apply l^2 on function $l_+ |\lambda, m_l\rangle$ and show that it is an eigenfunction and that there is no change in the eigenvalue (i.e. function f value).

$$l^2 l_+ |\lambda, m_l\rangle = l_+ l^2 |\lambda, m_l\rangle = l_+ f(\lambda, m_l)\hbar^2 |\lambda, m_l\rangle = f(\lambda, m_l)\hbar^2 l_+ |\lambda, m_l\rangle$$

   where we have used the fact that $\left[l^2, l_+\right] = 0$ (Eq. (4.13)).

2. Next we apply l_z on function $l_+ |\lambda, m_l\rangle$ and show that it is an eigenfunction and that the eigenvalue increases by $\hbar$:

$$l_z l_+ |\lambda, m_l\rangle = (l_+ l_z + [l_z, l_+]) |\lambda, m_l\rangle = (l_+ l_z + \hbar l_+) |\lambda, m_l\rangle$$

$$= (l_+ m_l \hbar + \hbar l_+) |\lambda, m_l\rangle = (m_l + 1)\hbar l_+ |\lambda, m_l\rangle$$

   This is an eigenvalue equation with $l_+ |\lambda, m_l\rangle$ as the eigenfunction for l_z and $(m_l + 1)\hbar$ as the eigenvalue.

Thus the state $l_+ |\lambda, m_l\rangle$ is proportional to the state $|\lambda, m_l + 1\rangle$ and we can write:

$$\begin{aligned} l_+ |\lambda, m_l\rangle &= c_+ (\lambda, m_l) \hbar |\lambda, m_l + 1\rangle \\ l_- |\lambda, m_l\rangle &= c_- (\lambda, m_l) \hbar |\lambda, m_l - 1\rangle \end{aligned} \tag{4.17}$$

where c_+ and c_- are dimensionless constants. The 2nd line follows from analogous consideration for l_-. Clearly l_+ generates the next higher state in terms of m_l and l_- does the opposite. This is why they are called shift operators.

4.5 The eigenvalues of angular momentum

From our previous calculation it is apparent that l_+ and l_- step m_l by ± 1. We have established that there must be an upper limit for m_l since $f(\lambda, m_l) \geq m_l^2$. The maximum value for m_l^2 we denote by l. If we operate on a state with $m_l = l$, we generate nothing, because there is no state with a larger value of m_l:

$$l_+ |\lambda, l\rangle = 0$$

This relation will give us the unknown function f. When we apply l_- on this, we get:

$$l_- l_+ |\lambda, l\rangle = 0$$

On the other hand, the product $l_- l_+$ can be expanded as follows:

$$\begin{aligned} l_- l_+ &= (l_x - i l_y)(l_x + i l_y) = l_x^2 + l_y^2 + i l_x l_y - i l_y l_x = l_x^2 + l_y^2 + i [l_x, l_y] \\ &= l^2 - l_z^2 + i (i\hbar l_z) \end{aligned} \tag{4.18}$$

Thus we can write the previous equation as:

$$l_- l_+ |\lambda, l\rangle = (l^2 - l_z^2 - \hbar l_z) |\lambda, l\rangle = 0$$

Rearranging the above gives now:

$$l^2 |\lambda, m\rangle = \left(l_z^2 + \hbar l_z\right) |\lambda, l\rangle = \left(l^2 + l\right) \hbar^2 |\lambda, l\rangle$$

Now it follows that $f(\lambda, l) = l(l+1)$. Since l_- does not change the value of l^2, all states $|\lambda, l\rangle, |\lambda, l-1\rangle, ...$ have the same eigenvalue of l^2. Therefore:

$$f(\lambda, m_l) = l(l+1), \text{ for } m_l = l, l-1, ...$$

Using similar reasoning, it can be shown that the lower limit for m_l is $-l$. Thus overall we have:

$$f(\lambda, m_l) = l(l+1), \text{ for } m_l = l, l-1, ..., 0, ..., -l+1, -l$$

By comparing the role of λ in the original eigenvalue equation and the l above, we can identify that they are the same: $\lambda = l$ (i.e. the maximum value of $|m_l|$). Now we have:

$$f(l, m_l) = l(l+1) \text{ for } m_l = l, l-1, ..., -l+1, -l \tag{4.19}$$

and also:

$$l^2 |l, m_l\rangle = l(l+1)\hbar^2 |l, m_l\rangle \tag{4.20}$$

and we already knew that:

$$l_z |l, m_l\rangle = m_l \hbar |l, m_l\rangle \tag{4.21}$$

Given a value for l then m_l can vary between $-l$ and l taking unit steps. For example, $l = 2$ will give $m_l = +2, +1, 0, -1, -2$. l can also take a value of half-integer. For example, in this case $l = 3/2$ and $m_l = +3/2, 1/2, -1/2, -3/2$.

Overall, based on the hermiticity of the angular momentum operators and their commutation relations, we have shown that:

- The magnitude of the angular momentum is confined to the values $\sqrt{l(l+1)}\hbar$ with $l = 0, 1, 2, ...$
- The component on an arbitrary z-axis is limited to the $2l + 1$ values $m_l\hbar$ with $m_l = l, l - 1, ..., -l$.

Note that the value of l earlier in the context of rotation (cyclic boundary condition) was fixed to an integer whereas here it may also take half-integer values. It turns out that spin (electron or nuclear) will follow the angular momentum laws with integra or half-integral value of l. For systems with cyclic boundary conditions we will use notation l and m_l, for general discussion on angular momentum j and m_j and s/m_s for half-integral systems. The general notation would therefore be:

$$j^2 \left|j, m_j\right\rangle = j(j+1)\hbar^2 \left|j, m_j\right\rangle \text{ and } j_z \left|j, m_j\right\rangle = m_j\hbar \left|j, m_j\right\rangle \tag{4.22}$$

with $m_j = j, j - 1, ..., -j$.

4.6 The matrix elements of the angular momentum

We still have not worked out the function $c_\pm(j, m_j)$:

$$j_\pm \left|j, m_j\right\rangle = c_\pm \left(j, m_j\right) \hbar \left|j, m_j \pm 1\right\rangle \tag{4.23}$$

Because these functions form an orthonormal set, multiplication from the left gives:

$$\left\langle j, m_j \pm 1\right| j_\pm \left|j, m_j\right\rangle = c_\pm(j, m_j)\hbar \tag{4.24}$$

So we basically need to know the above matrix elements in order to know $c_\pm$. These matrix elements play important role in calculating transition intensities in magnetic resonance spectroscopy.

Calculation of matrix elements. Eqs. (4.18) and (4.23) give two different ways to evaluate the operation of $j_- j_+$:

$$j_- j_+ \left|j, m_j\right\rangle = \left(j^2 - j_z^2 - \hbar j_z\right) \left|j, m_j\right\rangle = \left(j(j+1) - m_j(m_j+1)\right) \hbar^2 \left|j, m_j\right\rangle$$

$$j_- j_+ \left|j, m_j\right\rangle = j_- c_+(j, m_j)\hbar \left|j, m_j + 1\right\rangle = c_+(j, m_j)c_-(j, m_j + 1)\hbar^2 \left|j, m_j\right\rangle$$

Comparison of these two expressions yields:

$$c_+(j,m_j)c_-(j,m_j+1) = j(j+1) - m_j(m_j+1) \tag{4.25}$$

We need one more equation to obtain the $c_\pm$. Consider the following matrix element:

$$\langle j,m_j|\, j_-\, |j,m_j+1\rangle = c_-(j,m_j+1)\hbar$$

This can be manipulated as follows (j_x and j_y are hermitean):

$$\begin{aligned}
\langle j,m_j|\, j_-\, |j,m_j+1\rangle &= \langle j,m_j|\, j_x - ij_y\, |j,m_j+1\rangle \\
&= \langle j,m_j|\, j_x\, |j,m_j+1\rangle - i\,\langle j,m_j|\, j_y\, |j,m_j+1\rangle \\
&= \langle j,m_j+1|\, j_x\, |j,m_j\rangle^* - i\,\langle j,m_j+1|\, j_y\, |j,m_j\rangle^* \\
&= (\langle j,m_j+1|\, j_x\, |j,m_j\rangle + i\,\langle j,m_j+1|\, j_y\, |j,m_j\rangle)^* \\
&= \langle j,m_j+1|\, j_+\, |j,m_j\rangle^*
\end{aligned}$$

We have just shown that j_- and j_+ are each other's *hermitean conjugates*:

$$\langle j,m_j|\, j_-\, |j,m_j+1\rangle = \langle j,m_j+1|\, j_+\, |j,m_j\rangle^* \tag{4.26}$$

In general, operators A and B are each other's hermietan conjugates if:

$$\langle a|\, A\, |b\rangle = \langle b|\, B\, |a\rangle^* \tag{4.27}$$

By combining Eqs. (4.23) and (4.26) we get the following relation between c_- and c_+:

$$c_j(j, m_j + 1) = c_+^*(j, m_j) \tag{4.28}$$

Eq. (4.25) now gives:

$$|c_+(j, m_j)|^2 = j(j+1) - m_j(m_j+1)$$

Next we choose c_+ to be real and positive:

$$c_+(j, m_j) = \sqrt{j(j+1) - m_j(m_j+1)} = c_-(j, m_j) \tag{4.29}$$

Example. Evaluate the matrix elements (a) $\langle j, m_j + 1|\, j_x\, |j, m_j\rangle$, (b) $\langle j, m_j + 2|\, j_x\, |j, m_j\rangle$, and (c) $\langle j, m_j + 2|\, j_x^2\, |j, m_j\rangle$.

Solution. Express j_x in terms of the shift operators and use the previous results.

(a)

$$\langle j, m_j + 1 | \, j_x \, | j, m_j \rangle = \frac{1}{2} \langle j, m_j + 1 | \, j_+ + j_- \, | j, m_j \rangle$$

$$= \frac{1}{2} \langle j, m_j + 1 | \, j_+ \, | j, m_j \rangle + \frac{1}{2} \underbrace{\langle j, m_j + 1 | \, j_- \, | j, m_j \rangle}_{=0}$$

$$= \frac{1}{2} c_+(j, m_j)\hbar$$

(b)

$$\langle j, m_j + 2 | \, j_x \, | j, m_j \rangle = 0$$

(the rising and lowering operators can not produce $|j, m_j + 2\rangle$ state from $|j, m_j\rangle$)

(c)

$$\langle j, m_j + 2 | \, j_x^2 \, | j, m_j \rangle = \frac{1}{4} \langle j, m_j + 2 | \, j_+^2 + j_-^2 + j_+ j_- + j_- j_+ \, | j, m_j \rangle$$

$$= \frac{1}{4} \langle j, m_j + 2 | \, j_+^2 \, | j, m_j \rangle = \frac{1}{4} c_+(j, m_j + 1) c_+(j, m_j) \hbar^2$$

$$= \frac{1}{4} \sqrt{j(j+1) - (m_j + 1)(m_j + 2)} \sqrt{j(j+1) - m_j(m_j + 1)} \hbar^2$$

4.7 The angular momentum eigenfunctions

In the following we will consider *orbital angular momentum* explicitly. We had concluded earlier that we needed to solve a 2nd order differential equation, which we just stated to be spherical harmonics (without the actual calculation). Here we show that there is an easier way to solve this problem where we just need to consider a first order differential equation.

We start by looking for the wavefunction $|l, l\rangle$ (i.e. $m_l = l$). Once we have found this wavefunction, we can apply the lowering operator l_- on it to get the other wavefunctions corresponding to different m_l values. We need to solve the equation:

$$l_+ |l, l\rangle = 0$$

To solve this equation, we need to express the Cartesian angular momentum operators in spherical coordinates:

$$\begin{aligned} l_x &= -\frac{\hbar}{i}\left(\sin(\phi)\frac{\partial}{\partial\theta} + \cot(\theta)\cos(\phi)\frac{\partial}{\partial\phi}\right) \\ l_y &= \frac{\hbar}{i}\left(\cos(\phi)\frac{\partial}{\partial\theta} - \cot(\theta)\sin(\phi)\frac{\partial}{\partial\phi}\right) \\ l_z &= \frac{\hbar}{i}\frac{\partial}{\partial\phi} \end{aligned} \tag{4.30}$$

where $\cot(\theta) = \frac{1}{\tan(\theta)}$.

Now we can write the lowering and rising operators using the above Cartesian forms:

$$l_+ = \hbar e^{i\phi}\left(\frac{\partial}{\partial\theta} + i\cot(\theta)\frac{\partial}{\partial\phi}\right) \tag{4.31}$$
$$l_- = -\hbar e^{-i\phi}\left(\frac{\partial}{\partial\theta} - i\cot(\theta)\frac{\partial}{\partial\phi}\right)$$

Using these representations we can now rewrite our $l_+ |l,l\rangle = 0$ equation as:

$$\hbar e^{i\phi}\left(\frac{\partial}{\partial\theta} + i\cot(\theta)\frac{\partial}{\partial\phi}\right)\psi_{l,l}(\theta,\phi) = 0$$

This partial differential equation can be separated by writing $\psi(\theta,\phi) = \Theta(\theta)\Phi(\phi)$. After substituting, differentiating and dividing appropriately, we get:

$$\frac{\tan(\theta)}{\Theta}\frac{d\Theta}{d\theta} = -\frac{i}{\Phi}\frac{d\Phi}{d\phi}$$

Since the LHS depends on θ only and the RHS on ϕ only so they must be equal to some constant, which we denote by c. Therefore we can write two equations:

$$\tan(\theta)\frac{d\Theta}{d\theta} = c\Theta \text{ and } \frac{d\Phi}{d\phi} = ic\Phi$$

These two equations can be integrated directly to give:

$$\Theta \propto \sin^c(\theta) \text{ and } \Phi \propto e^{ic\phi}$$

The constant c can be found to be equal to l by requiring that $l_z\psi_{l,l} = l\hbar\psi_{l,l}$. Thus we can write the complete solution as:

$$\psi_{l,l} = N\sin^l(\theta)e^{il\phi} \tag{4.32}$$

where N is a normalization constant. This can be verified to correspond to the spherical harmonics we listed earlier. To get the rest of the eigenfunctions, one now needs to apply l_- repeatedly on Eq. (4.32).

Example. Construct the wavefunction for the state $|l, l-1\rangle$.

Solution. We know that $l_- |l,l\rangle = c_-\hbar |l, l-1\rangle$ and the position representation for l_- we can proceed:

$$l_-\psi_{l,l} = -\hbar e^{-i\phi}\left(\frac{\partial}{\partial\theta} - i\cot(\theta)\frac{\partial}{\partial\phi}\right) N\sin^l(\theta)e^{il\phi}$$

$$= -N\hbar e^{-i\phi}\left(l\sin^{l-1}(\theta)\cos(\theta) - i(il)\cot(\theta)\sin^l(\theta)\right)e^{il\phi}$$

$$= -2Nl\hbar\sin^{l-1}(\theta)\cos(\theta)e^{i(l-1)\phi}$$

On the other hand we also know that:

$$l_- |l, l\rangle = \sqrt{l(l+1) - l(l-1)}\hbar |l, l-1\rangle = \sqrt{2l}\hbar |l, l-1\rangle$$

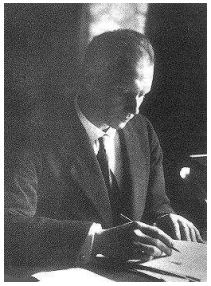
Therefore we can now extract $|l, l-1\rangle$:

$$\psi_{l,l-1} = -\sqrt{2l} N \sin^{l-1}(\theta) \cos(\theta) e^{i(l-1)\phi}$$

Note that the normalization constant is universal in a sense that one N will provide normalization for all $|l, m_l\rangle$:

$$N = \frac{1}{2^l l!} \sqrt{\frac{(2l+1)!}{4\pi}}$$

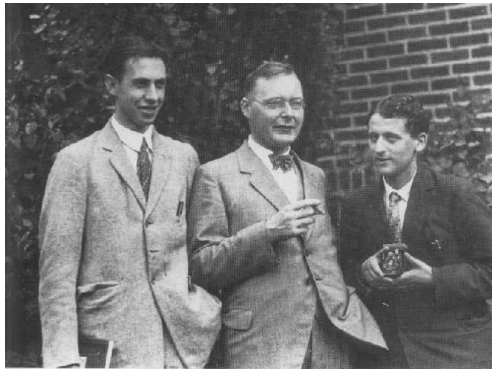
4.8 Spin



Walter Gerlach
(1889 - 1979)

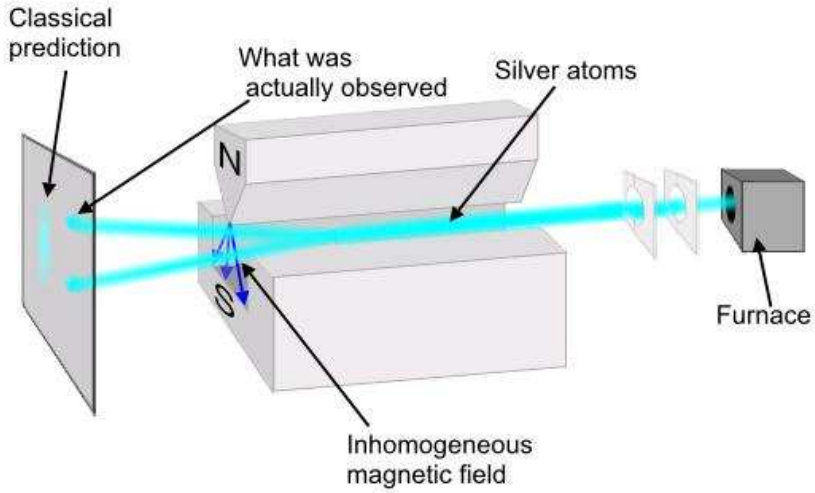


Otto Stern (1889 - 1979),
Nobel price 1943



L: George Uhlenbeck (1900 - 1988),
M: Hendrik Kramers (1894 - 1952),
R: Samuel Goudsmit (1902 - 1978).

Stern-Gerlach experiment:



Note that silver atoms have one unpaired electron.

The electron appears to have an intrinsic magnetic moment, which originates from ele ctron spin.

Based on the Stern-Gerlach experiment it appears that electrons have intrinsic angular momentum, which is called the *spin*. The quantum number for electron spin is $s = 1/2$ and there are two states associated with it $m_s = +1/2$ (often denoted by α or $\uparrow$) and $m_s = -1/2$ (often denoted by β or $\downarrow$). Note that spin does not naturally arise from non-relativistic quantum mechanics but requires the relativistic theory developed by Paul Dirac. Fortunately, by using the general laws of angular momentum, it is possible to incorporte it into Schrödinger equation.

If we denote the $\left|\frac{1}{2}, +\frac{1}{2}\right\rangle$ by α and $\left|\frac{1}{2}, -\frac{1}{2}\right\rangle$ by β then we can apply the previously developed general expressions to yield:

$$s_z\alpha = +\frac{1}{2}\hbar\alpha \text{ and } s_z\beta = -\frac{1}{2}\hbar\beta \text{ and } s^2\alpha = \frac{3}{4}\hbar^2\alpha \text{ and } s^2\beta = \frac{3}{4}\hbar^2\beta \tag{4.33}$$

and the effects of the shift operators are:

$$s_+\alpha = 0 \text{ and } s_+\beta = \hbar\alpha \text{ and } s_-\alpha = \hbar\beta \text{ and } s_-\beta = 0 \tag{4.34}$$

The matrix elements of s_- and s_+ are:

$$\langle\alpha| s_+ |\beta\rangle = \langle\beta| s_- |\alpha\rangle = \hbar \tag{4.35}$$

These spin eigenfunctions can also be represented as vectors:

$$\alpha = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \text{ and } \beta = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{4.36}$$

The spin operators can then be written as matrices (note that there is a finite number of possible eigenstates). For example, consider the s_z operator:

$$s_z\alpha = +\frac{1}{2}\hbar \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = +\frac{1}{2}\hbar \begin{pmatrix} 1 \\ 0 \end{pmatrix} = +\frac{1}{2}\hbar\alpha$$

The same calculation can be carried out for β. For s_x we can write:

$$s_x\alpha = +\frac{1}{2}\hbar \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = +\frac{1}{2}\hbar \begin{pmatrix} 0 \\ 1 \end{pmatrix} = +\frac{1}{2}\hbar\beta$$

All spin $s = \frac{1}{2}$ operators can thus be represented by matrices as follows:

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{4.37}$$

$$\sigma_+ = \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix} \quad \sigma_- = \begin{pmatrix} 0 & 0 \\ 2 & 0 \end{pmatrix}$$

with the following relation:

$$s_q = \frac{1}{2}\hbar\sigma_q \text{ where } q = x, y, z, +, - \tag{4.38}$$

These are called the *Pauli spin matrices*.

Example. Show that the Pauli matrices correctly follow the commutation relation $[s_x, s_y] = i\hbar s_z$.

Solution. First we calculate the commutator between σ_x and σ_y:

$$[\sigma_x, \sigma_y] = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} - \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

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

$$= 2 \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} = 2i \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = 2i\sigma_z$$

Eq. (4.38) now states that:

$$[s_x, s_y] = \frac{1}{4}\hbar^2 [\sigma_x, \sigma_y] = \frac{1}{4}\hbar^2 (2i\sigma_z) = i\hbar s_z$$

4.9 The specification of coupled angular momenta

In the following we will consider two sources of angular momentum, which are denoted by j_1 and j_2. These can arise from one particle in the form of its spin and orbital angular momenta or from two different particles. We need to find a way to express the total angular momentum j that considers contribution of both j_1 and j_2.

We denote the angular momentum quantum numbers for particle 1 by j_1 and m_{j_1} and for particle 2 by j_2 and m_{j_2}. In order to represent the overall wavefunction as $|j_1, m_{j_1}; j_2, m_{j_2}\rangle$ we must make sure that all operators j_1^2, j_2^2, j_{1z}, j_{2z} commute. In fact, each Cartesian component of j_1 and j_2 commute:

$$\left[j_{1q}, j_{2q'}\right] = 0 \text{ where } q = x, y, z \text{ and } q' = x, y, z \tag{4.39}$$

This can be seen simply by noting that each source of angular momentum have their own variables. Operators depending on different variables commute trivially. Since each Cartesian component commutes and $j_q^2 = j_{qx}^2 + j_{qy}^2 + j_{qz}^2$, we have the necessary commutation relations for expressing the wavefunction as $|j_1, m_{j_1}; j_2, m_{j_2}\rangle$.

Next we will check if the *total angular momentum* $j = j_1 + j_2$ could also be specified. First we need to verify that j is angular momentum by using Eq. (4.7):

$$\begin{aligned}
&[j_x, j_y] = [j_{1x} + j_{2x}, j_{1y} + j_{2y}] \\
&= [j_{1x}, j_{1y}] + [j_{2x}, j_{2y}] + [j_{1x}, j_{2y}] + [j_{2x}, j_{1y}] \\
&= i\hbar j_{1z} + i\hbar j_{2z} + 0 + 0 = i\hbar j_z
\end{aligned} \tag{4.40}$$

The other commutation relations required by Eq. (4.7) can be obtained from the above calculation by cyclic permutation of the coordinate labels. Thus we conclude that j is angular momentum. (Exercise: Show that $j_1 - j_2$ is not angular momentum). Now we can also get the magnitude of j as $\sqrt{j(j+1)}\hbar$ with j integral or half-integral and the z component given by $m_j\hbar$ with $m_j = j, j-1, ..., -j$.

Can we specify j if we have specified j_1 and j_2? To answer this question we need to consider the commutation relation between these operators:

$$\begin{aligned}
&[j^2, j_1^2] = [j_x^2 + j_y^2 + j_z^2, j_{1x}^2 + j_{1y}^2 + j_{1z}^2] \\
&= \left[(j_{1x} + j_{2x})^2 + (j_{1y} + j_{2y})^2 + (j_{1z} + j_{2z})^2, j_{1x}^2 + j_{1y}^2 + j_{1z}^2\right] \\
&= [j_{1x}^2, j_{1x}^2] + [j_{1x}j_{2x}, j_{1x}^2] + ... = 0
\end{aligned} \tag{4.41}$$

because $[j_{1q}, j_{2q}] = 0$, $[j_{1q}, j_{1q}] = 0$, and $[j_{2q}, j_{2q}] = 0$.

This means that we can simultaneously specify the eigenvalues of j_1^2, j_2^2, and j^2. The operator j also commutes with $j_z = j_{1z} + j_{2z}$. This allows us to specify the our eigenstates as $|j_1, j_2; j, m_j\rangle$. Can we proceed even further by also specifying m_{j1} and m_{j2}? To answer this question we inspect the commutator between j^2 and j_{1z} as an example:

$$\begin{aligned}
&\left[j_{1z}, j^2\right] = \left[j_{1z}, j_x^2\right] + \left[j_{1z}, j_y^2\right] + \left[j_{1z}, j_z^2\right] \qquad (4.42)\\
&= \left[j_{1z}, (j_{1x} + j_{2x})^2\right] + \left[j_{1z}, (j_{1y} + j_{2y})^2\right] + \left[j_{1z}, (j_{1z} + j_{2z})^2\right]\\
&= \left[j_{1z}, j_{1x}^2 + 2j_{1x}j_{2x}\right] + \left[j_{1z}, j_{1y}^2 + 2j_{1y}j_{2y}\right]\\
&= \left[j_{1z}, j_{1x}^2 + j_{1y}^2\right] + 2\left[j_{1z}, j_{1x}\right]j_{2x} + 2\left[j_{1z}, j_{1y}\right]j_{2y}\\
&= \left[j_{1z}, j_1^2 - j_{1z}^2\right] + 2i\hbar j_{1y}j_{2x} - 2i\hbar j_{1x}j_{2y}\\
&= 2i\hbar(j_{1y}j_{2x} - j_{1x}j_{2y}) \neq 0
\end{aligned}$$

Because this commutator is not zero, we cannot specify j_{1z} and j^2 simultaneously. The same calculation can also be carried out for $[j_{2z}, j]$. Thus we cannot specify m_{j1} or m_{j2} in this case. Thus we have the following choices:

Representation	Wavefunction	Interpretation
Coupled representation	$\|j_1, j_2; j, m_j\rangle$	Idividual m_{j1} and m_{j2} unspecified.
Uncoupled representation	$\|j_1, m_{j1}; j_2, m_{j2}\rangle$	Total angular momentum unspecified.

Later we will see that depending on the situation, one representation might be easier to use than the other. Both representations are complete in a sense that either one can be used to describe the system.

4.10 The permitted values of the total angular momentum

Consider the coupled representation $|j_1, j_2; j, m_j\rangle$. What are the allowed values for j and m_j? For example, if we have one p-electron, which has $l = 1$ and $s = 1/2$, then it does not seem possible to have total angular momentum to be larger than $3/2$. The allowed values for m_j follow from $j_z = j_{1z} + j_{2z}$:

$$m_j = m_{j_1} + m_{j_2} \tag{4.43}$$

This states that the z-axis projections are just summed to give the total projected value. The possible values for j are given by the Glebsch-Gordan (GG) series:

$$j = j_1 + j_2, j_1 + j_2 - 1, ..., |j_1 - j_2| \tag{4.44}$$

To justify this results we first notice that it gives the same number of states as the uncoupled representation. For example, when $j_1 = j_2 = 1$, the uncoupled representation has states: $|1, 1; 1, 1\rangle$, $|1, 0; 1, 1\rangle$, $|1, -1; 1, 1\rangle$, ... ($3 \times 3 = 9$ states). The coupled representation, according to the above prescription, gives: $|1, 1; 2, 2\rangle$, $|1, 1; 2, 1\rangle$, $|1, 1; 2, 0\rangle$, ... ($(2 \times 2 + 1) + (2 \times 1 + 1) + (2 \times 0 + 1) = 9$ states).

To justify the first term in the GG series consider the state with the maximum amount of angular momentum about the quantitization axis: $m_{j_1} = j_1$ and $m_{j_2} = j_2$. Eq. (4.43) gives then the following maximum value for $m_j = m_{j_1} + m_{j_2} = j_1 + j_2$. Since the maximum value for m_j is j by definition, we have $j = j_1 + j_2$.

The second term in the GG series ($j = j_1 + j_2 - 1$) can be justified as follows. Consider two uncoupled states with 1) $m_{j_1} = j_1 - 1$ & $m_{j_2} = j_2$ and 2)$m_j = j_1$ & $m_{j_2} = j_2 - 1$. These two states contribute to $j = j_1 + j_2$ with $m_j = j - 1$ but also to $j = j_1 + j_2 - 1$ with $m_j = j$. One can continue this process to account for all the states that are shown in the GG series. Overall the procedure must be repeated until the right number of states are identified.

Example. What angular momentum can arise from two sources, which have $j_1 = \frac{1}{2}$ and $j_2 = \frac{3}{2}$? Express the eigenstates in both coupled and uncoupled representations.

Solution. To find out the possible values for the total angular momentum j, we need to write the GG series (Eq. (4.44)):

$$j = 2, 1$$

Eigenstates in the uncoupled representation are:

$$\left|\frac{1}{2}, +\frac{1}{2}; \frac{3}{2}, +\frac{3}{2}\right\rangle, \left|\frac{1}{2}, +\frac{1}{2}; \frac{3}{2}, +\frac{1}{2}\right\rangle, \left|\frac{1}{2}, +\frac{1}{2}; \frac{3}{2}, -\frac{1}{2}\right\rangle, \left|\frac{1}{2}, +\frac{1}{2}; \frac{3}{2}, -\frac{3}{2}\right\rangle$$

$$\left|\frac{1}{2}, -\frac{1}{2}; \frac{3}{2}, +\frac{3}{2}\right\rangle, \left|\frac{1}{2}, -\frac{1}{2}; \frac{3}{2}, +\frac{1}{2}\right\rangle, \left|\frac{1}{2}, -\frac{1}{2}; \frac{3}{2}, -\frac{1}{2}\right\rangle, \left|\frac{1}{2}, -\frac{1}{2}; \frac{3}{2}, -\frac{3}{2}\right\rangle$$

In the coupled representation the eigenstates are:

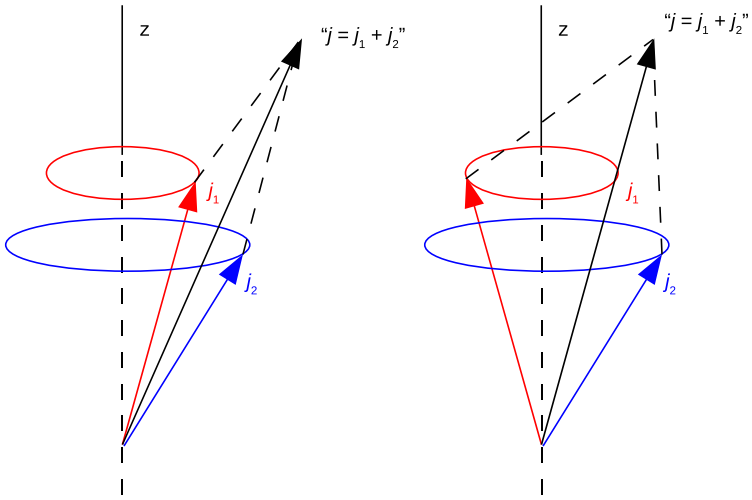
$$\left|\frac{1}{2},\frac{3}{2};2,+2\right\rangle,\ \left|\frac{1}{2},\frac{3}{2};2,+1\right\rangle,\ \left|\frac{1}{2},\frac{3}{2};2,0\right\rangle,\ \left|\frac{1}{2},\frac{3}{2};2,-1\right\rangle,\ \left|\frac{1}{2},\frac{3}{2};2,-2\right\rangle$$

$$\left|\frac{1}{2},\frac{3}{2};1,+1\right\rangle,\ \left|\frac{1}{2},\frac{3}{2};1,0\right\rangle,\ \left|\frac{1}{2},\frac{3}{2};1,-1\right\rangle$$

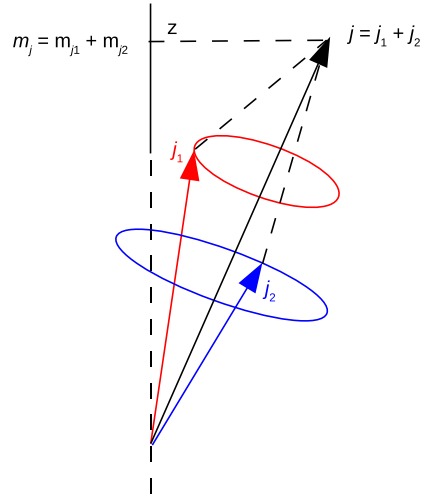
4.11 The vector model of coupled angular momenta

The *vector model* is an attempt to represent pictorially the features of coupled angular momenta. These vector diagrams express the following information:

- The length of the vector representing the total angular momentum is $\sqrt{j(j+1)}$ with j being one of the permitted values in the Glebsch-Gordan series.
- The above vector must lie at an indeterminate angle on a cone about the z-axis. This is because when j_z is specified, j_x and j_y are subject to the uncerntainty principle.
- The lengths of the contributing angular momenta j_1 and j_2 are given by $\sqrt{j_1(j_1+1)}$ and $\sqrt{j_2(j_2+1)}$, respectively. These lengths have definite values even when j is specified.
- In the *coupled picture* the total angular momentum (j) and the projection on the z-axis (m_j) are well defined. The values of m_{j_1} and m_{j_2} are indefinite but their sum $m_{j_1} + m_{j_2} = m_j$ has a definite value.
- In the *uncoupled picture* j is indefinite but j_1 and j_2 along with their projections m_{j_1} and m_{j_2} all have well defined values.



Uncoupled picture: j is not well defined.



Coupled picture: j is well defined but j_1 and j_2 are not.

Example. Consider two particles with $j_1 = 1/2$ and $j_2 = 1/2$. These can be, for example, electrons or protons. Write the states in both uncoupled and coupled representations.

Solution. In the uncoupled representation we have $\left|\frac{1}{2}, +\frac{1}{2}; \frac{1}{2}, +\frac{1}{2}\right\rangle$, $\left|\frac{1}{2}, -\frac{1}{2}; \frac{1}{2}, +\frac{1}{2}\right\rangle$, $\left|\frac{1}{2}, +\frac{1}{2}; \frac{1}{2}, -\frac{1}{2}\right\rangle$, and $\left|\frac{1}{2}, -\frac{1}{2}; \frac{1}{2}, -\frac{1}{2}\right\rangle$.

In the coupled representations (using the Gelbsch-Gordan series): $\left|\frac{1}{2}, \frac{1}{2}; 1, +1\right\rangle$, $\left|\frac{1}{2}, \frac{1}{2}; 1, 0\right\rangle$, $\left|\frac{1}{2}, \frac{1}{2}; 1, -1\right\rangle$ and $\left|\frac{1}{2}, \frac{1}{2}; 0, 0\right\rangle$. The first three terms beloning to $j = 1$ is called *triplet state* as it consists of three different m_j levels. The last term is called *singlet state* since there is only one m_j level associated with it.

4.12 The relation between the schemes

The state $|j_1, j_2; j, m_j\rangle$ consists of all values of m_{j_1} and m_{j_2} such that $m_j = m_{j_1} + m_{j_2}$. Thus it should be possible to express the coupled state as a linear combination of the uncoupled states $|j_1, m_{j_1}; j_2, m_{j_2}\rangle$:

$$|j_1, j_2; j, m_j\rangle = \sum_{m_{j_1}, m_{j_2}} C(m_{j_1}, m_{j_2})\, |j_1, m_{j_1}; j_2, m_{j_2}\rangle \tag{4.45}$$

The coefficients $C(m_{j_1}, m_{j_2})$ are called *vector coupling coefficients* (or *Glebsch-Gordan coefficients*, *Wiegner coefficients*, and in a slightly modified form *3j-symbols*).

Let us consider the two spin-half particles in the previous example. The relation ship between the two representations is then:

$$|1, +1\rangle = \left|\frac{1}{2}, +\frac{1}{2}; \frac{1}{2}, +\frac{1}{2}\right\rangle$$

$$|1, 0\rangle = \frac{1}{\sqrt{2}}\left|\frac{1}{2}, +\frac{1}{2}; \frac{1}{2}, -\frac{1}{2}\right\rangle + \frac{1}{\sqrt{2}}\left|\frac{1}{2}, -\frac{1}{2}; \frac{1}{2}, - + \frac{1}{2}\right\rangle$$

$$|1, -1\rangle = \left|\frac{1}{2}, -\frac{1}{2}; \frac{1}{2}, -\frac{1}{2}\right\rangle$$

$$|0, 0\rangle = \frac{1}{\sqrt{2}}\left|\frac{1}{2}, +\frac{1}{2}; \frac{1}{2}, -\frac{1}{2}\right\rangle - \frac{1}{\sqrt{2}}\left|\frac{1}{2}, -\frac{1}{2}; \frac{1}{2}, - + \frac{1}{2}\right\rangle$$

Note that in the triplet state we have a possibilty for the spins to be parallel ($\uparrow\uparrow$ or $\downarrow\downarrow$) where as in the singlet state we can only have opposite spins ($\uparrow\downarrow$ or $\downarrow\uparrow$). The $m_j = 0$ components differ only by their relative sign. For the triplet state the plus sign signifies that the spins are precessing in with the same phase whereas in the singlet state they are out of phase.

Vector coupling coefficients can be derived based on the idea of the following example. In general, however, it is much easier to use numerical tables that list the coefficients. If we multiply Eq. (4.45) from the left by $\left\langle j_1, m'_{j_1}; j_2, m'_{j_2} \right|$ and notice that the only term that survives is $m_{j_1} = m'_{j_1}$ and $m_{j_2} = m'_{j_2}$ (orthogonality):

$$\left\langle j_1, m'_{j_1}; j_2, m'_{j_2} \right| \, j_1, j_2; j, m_j \rangle = C(m'_{j_1}, m'_{j_2}) \tag{4.46}$$

In the following we specifically consider $j_1 = \frac{1}{2}$ and $j_2 = \frac{1}{2}$. The state with maximum m_j must arise from the two angular momenta having the same m_j:

$$\left| \frac{1}{2}, +\frac{1}{2}; \frac{1}{2}, +\frac{1}{2} \right\rangle \tag{4.47}$$

By operating on this state with j_- we get (see Eqs. (4.23) and (4.29)):

$$j_- \, |j, m_j\rangle = \sqrt{j(j+1) - m_j(m_j - 1)}\hbar \, |j, m_j - 1\rangle \tag{4.48}$$

and thus: $j_- \, |1, +1\rangle = \sqrt{2}\hbar \, |1, 0\rangle$

On the other hand we can also express j_- using the individual lowering operators: $j_- = j_{1-} + j_{2-}$. Operation with this on $|1, +1\rangle$ gives:

$$j_- |j, m_j\rangle = (j_{1-} + j_{2-})\, \alpha_1 \alpha_2 = \hbar\, (\alpha_1 \beta_2 + \beta_1 \alpha_2)$$

Combining the two results above now gives:

$$|1, 0\rangle = \frac{1}{\sqrt{2}} (\alpha_1 \beta_2 + \beta_1 \alpha_2) \tag{4.49}$$

One could proceed by applying the above procude once more to obtain $|1, -1\rangle$ or just by stating that both angular momenta must have the same m_j:

$$|1, -1\rangle = \beta_1 \beta_2 \tag{4.50}$$

Subce j^2 is hermitean operator, its eigenfunctions are orthogonal. Changing the sign in Eq. (4.49) will give this function:

$$|0, 0\rangle \frac{1}{\sqrt{2}} (\alpha_1 \beta_2 - \beta_1 \alpha_2) \tag{4.51}$$

The same procedure can be repeated for any values of j_1 and j_2. The above results can be compiled into a table of vector coupling coefficients.

m_{j_1}	m_{j_2}	$\lvert 1,+1\rangle$	$\lvert 1,0\rangle$	$\lvert 0,0\rangle$	$\lvert 1,-1\rangle$
$+\frac{1}{2}$	$+\frac{1}{2}$	1	0	0	0
$+\frac{1}{2}$	$-\frac{1}{2}$	0	$1/\sqrt{2}$	$1/\sqrt{2}$	0
$-\frac{1}{2}$	$+\frac{1}{2}$	0	$1/\sqrt{2}$	$-1/\sqrt{2}$	0
$-\frac{1}{2}$	$-\frac{1}{2}$	0	0	0	1

4.13 The coupling of several angular momenta

If we need to couple more than two angular momenta then we essentially couple two at a time and then proceed coupling this to the rest. For example, if we have two electrons ($s_1 = 1/2$ and $s_2 = 1/2$) each with unit orbital angular momentum $l_1 = 1$ and $l_2 = 1$. We could first couple the spin and orbital angular momenta for each electron and then couple the resulting angular momenta or we could first couple the electron spins and orbital angular momenta separately and then couple the total spin and orbital angular momenta. For example the spin angular momenta couple to give the singlet ($S = 0$) and triplet ($S = 1$) states and the orbital angular momentum yields $L = 2, 1, 0$. Then coupling these to each other will give (for $S = 1$): $J = 3, 2, 1$; $J = 2, 1, 0$; $J = 1$ and for ($S = 0$): $J = 2$, $J = 1$, $J = 0$.

Chapter 5: Group theory

5.1 Symmetry operations and elements

Molecules in their equilibrium geometries often exhibit a certain degree of symmetry. For example, a benzene molecule is symmetric with respect to rotations around the axis perpendicular to the molecular plane. The concept of symmetry can be applied in quantum mechanics to simplify the underlying calculations. For example, in chemistry, symmetry can be used to predict optical activities of molecules as well as their dipole moments. Especially, in spectroscopy symmetry is a powerful tool for predicting optically allowed transitions.

Symmetry element: A symmetry element is a geometrical entity, which acts as a center of symmetry. It can be a plane, a line or a point.

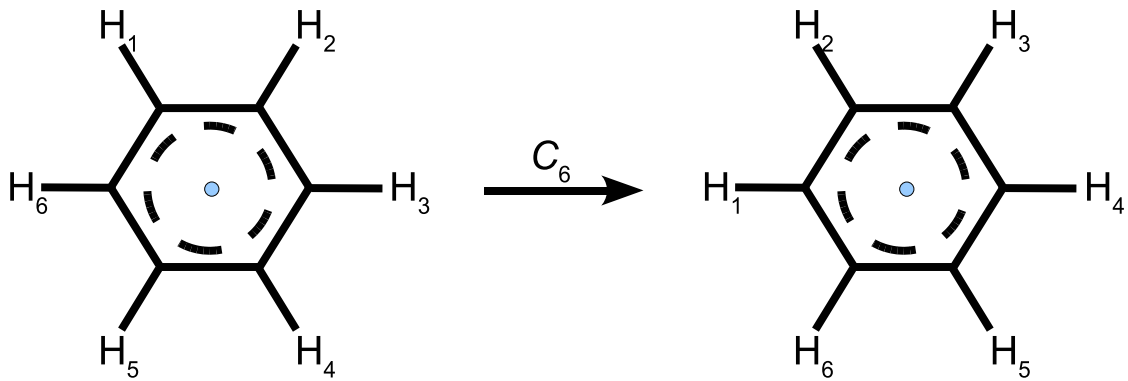
Symmetry operation: Action that leaves an object looking the same after it has been carried out is called a symmetry operation. Typical symmetry operations include rotations, reflections and inversions. The corresponding symmetry element defines the reference point for the symmetry operation. In quantum mechanics symmetry operations appear as operators, which can operate on a given wavefunction.

Point group: A collection of symmetry operations defines the overall symmetry for the molecule. When these operations form a mathematical group, they are called a point group. As we will see later, molecules can be classified in terms of point groups.

Symmetry operation	Symmetry element	Operation
i	Center of symmetry (point)	Projection through the center of symmetry to the equal distance on the opposite side.
C_n	Proper rotation axis (line)	Counterclockwise rotation about the axis by $2\pi/n$, where n is an integer.
σ	Mirror plane (plane)	Reflection across the plane of symmetry.
S_n	Improper rotation axis (line)	Counterclockwise rotation about the axis by $2\pi/n$ followed by a reflection across the plane perpendicular to the rotation axis.
E	Identity element	This operation leaves the object unchanged.

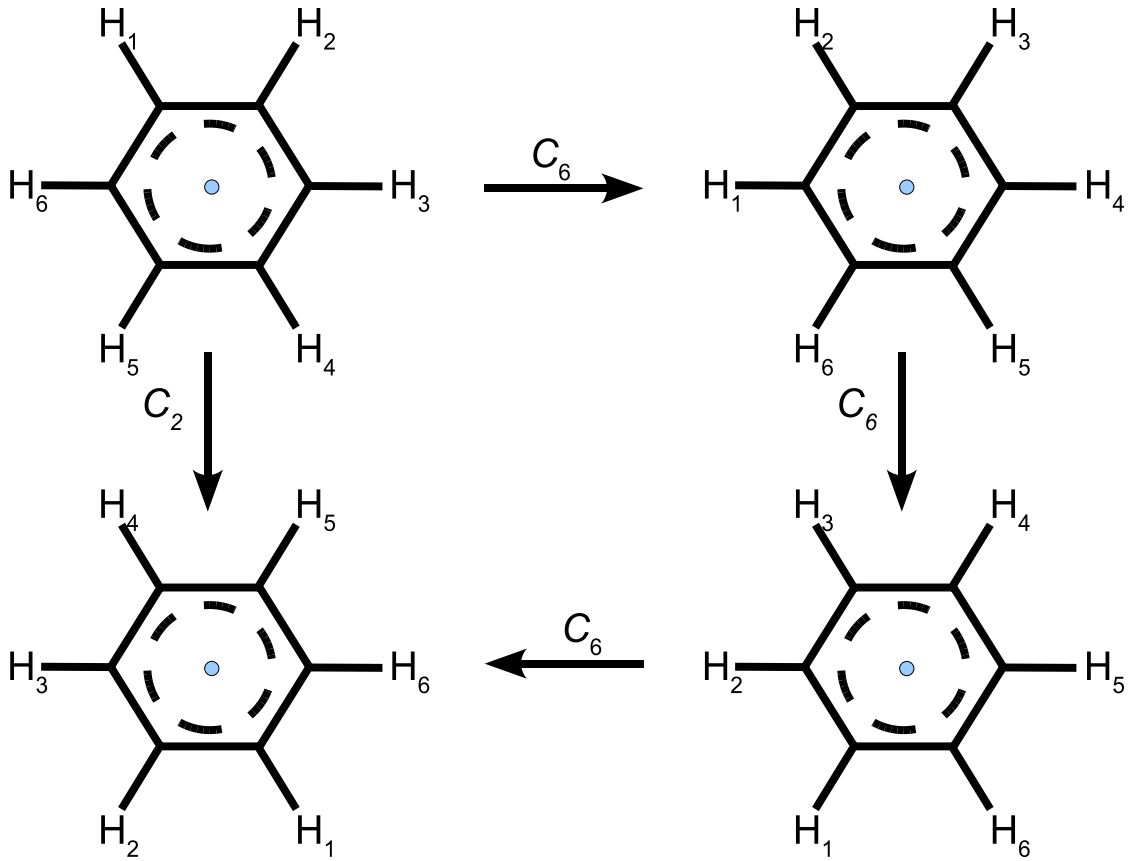
Rotation. The rotation operation is denoted by C_n, where the (counterclockwise) rotation angle is given by $2\pi/n$ in radians. Thus a C_1 operation rotates a given object by 360°, which effectively does nothing to the object. Here n is called the **order of rotation** and the corresponding symmetry element is called an n-fold rotation axis. Often notation C_n^+ is used to denote clockwise and C_n^- counterclockwise rotations.

Consider a planar benzene molecule as an example (note that both C and H nuclei are transformed):



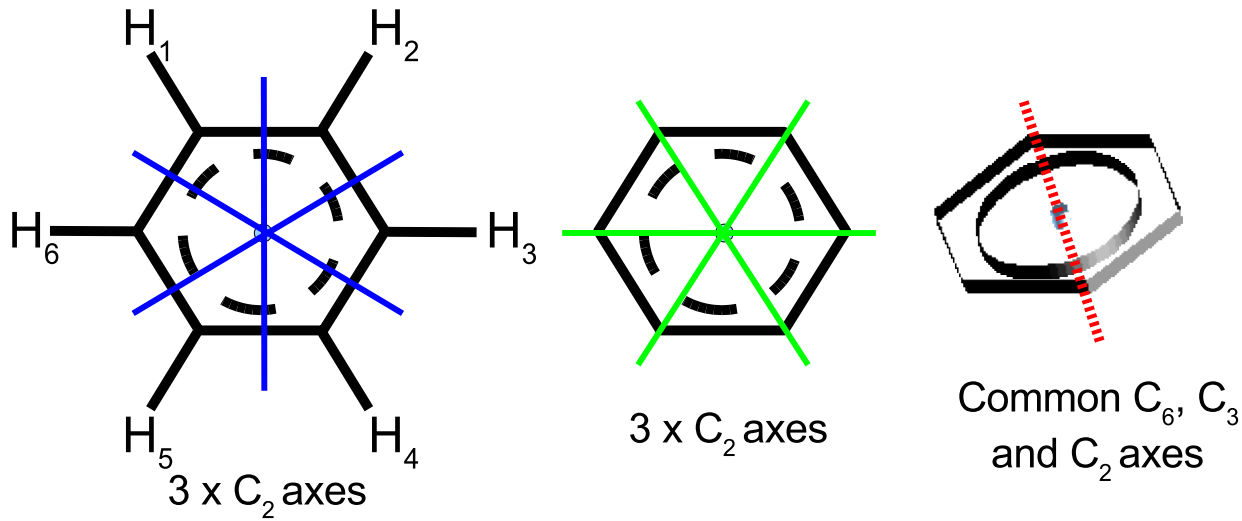
The symmetry element is indicated in the middle (line pointing out of plane).

Rotations can be combined to yield other rotation operations. For example, for benzene $C_6^3 = C_2$:



Demonstration of $C_6^3 = C_2$.

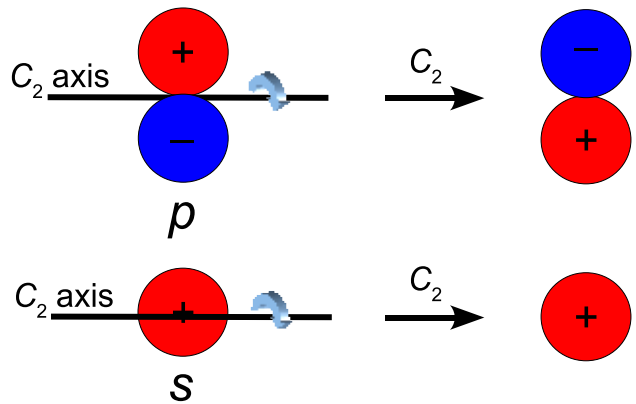
A molecule may have many different rotation symmetry axes. For example, benzene has a number of different possible C_n with various symmetry elements. Consider the C_6 symmetry element going through the center of the molecule and being perpendicular to the plane of the molecule. As shown previously, both C_6 and C_2 have collinear symmetry axes. In addition, C_3 also has the same symmetry axis. Furthermore, there are six other C_2 symmetry axes. These axes are indicated below.



Various C_6, C_3 and C_2 symmetry axes in benzene.

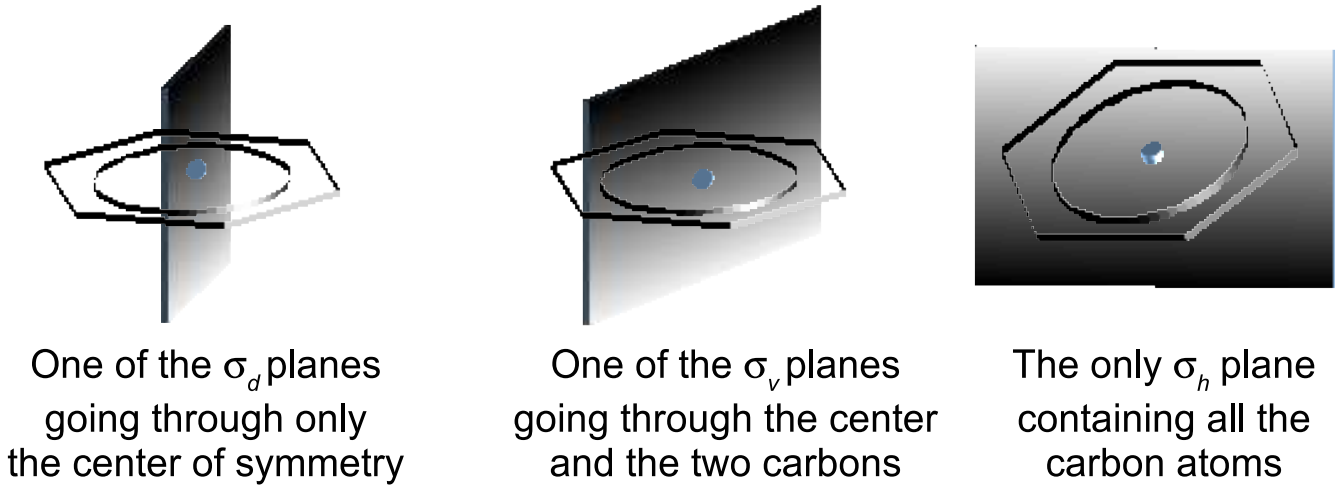
Note that there are three different kinds of C_2 axes and in this case we distinguish between them by adding primes to them (e.g. C_2, C_2', C_2''). The **principal axis** of rotation is the C_n axis with the highest n. For benzene this is C_6.

Symmetry operations can be performed on any object defined over the molecule. For example, a C_2 operation on a s and p_z orbitals can visualized as follows:



Operation of C_2 on s and p orbitals.

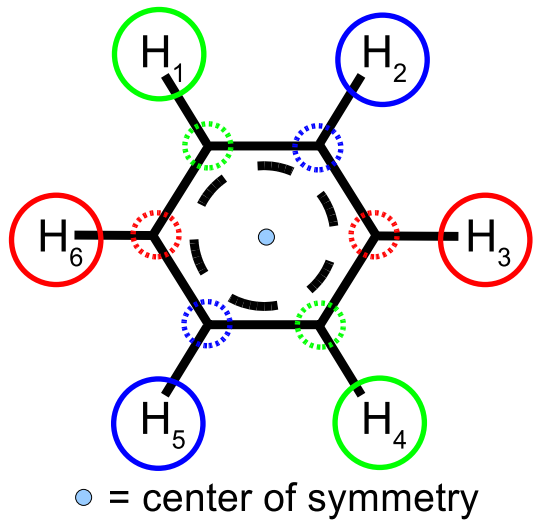
Reflection. The reflection operation is denoted by σ and the corresponding symmetry element is called a mirror plane. Given a symmetry plane, the σ operation reflects each point to the opposite side of the plane. For example, some of the σ symmetry elements in benzene are shown below.



Some of the σ symmetry elements in benzene.

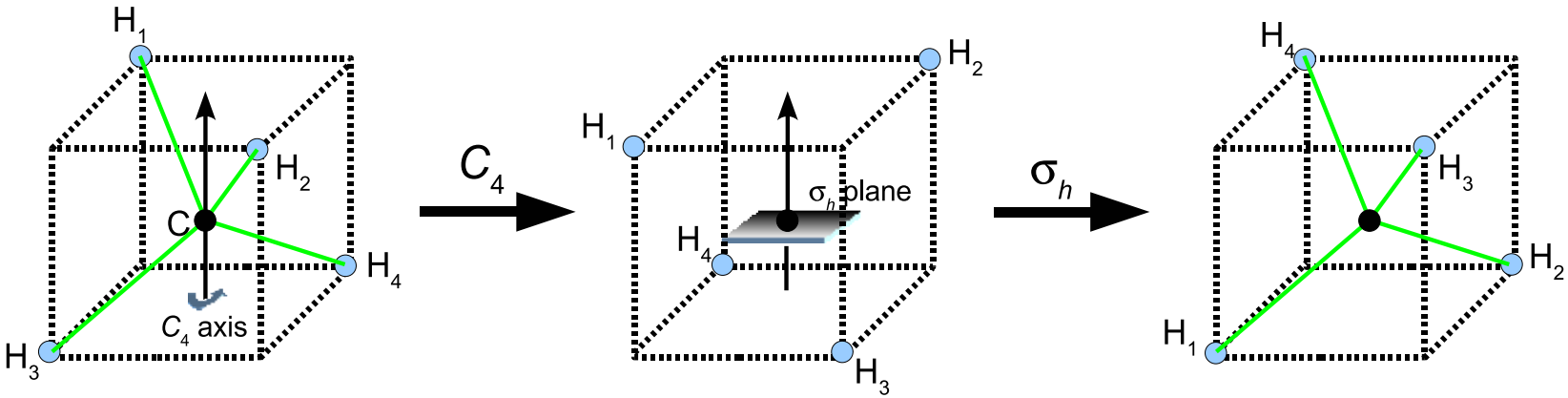
σ_d denotes a plane, which bisects the angle between the two C_2 axes and lies parallel to the principal axis. The σ_v plane includes the protons and the principal axis. The σ_h is perpendicular to the principal axis. Note that two successive reflections $\sigma\sigma$ bring the molecule back to its original configuration (corresponding to an E operation).

Inversion. The inversion operation is denoted by i and it is expressed relative to the central point (i.e. the symmetry element) in the molecule through which all the symmetry elements pass. This point is also called the **center of symmetry**. If the center point is located at origin $(0, 0, 0)$, the inversion operation changes coordinates as $(x, y, z) \rightarrow (-x, -y, -z)$. Molecules with inversion symmetry are called **centrosymmetric**. Note that there are obviously molecules which do not fulfill this requirement. Application of i operation twice (e.g. $i^2 = ii$) corresponds to the identity operation E.



Atoms related to each other via the inversion symmetry in benzene.

Rotation-reflection. Rotation-reflection operation is denoted by S_n. It consists of two different operations: C_n and σ_h, which are executed in sequence. Note that a molecule may not necessary possess a proper symmetry with respect to these individual operations but may still have the overall S_n symmetry. For example, benzene has S_6 symmetry as well as C_6 and σ_h whereas a tetrahedral CH_4 has S_4 but not C_4 or σ_h alone:



S_4 symmetry operation in methane. Note that the symmetry is temporarily lost after C_4.

It can be shown that $(S_4)^2 = C_2$ and $(S_4)^4 = E$.

5.2 The classification of molecules

To classify a molecule according to its symmetry, we list all its symmetry operations and the assign a label (*point group*) that is based on the list of those operations. The word point refers to the fact that we will only consider the operations corresponding to symmetry elements that intersect in at least one point. To classify crystals, we would also need to consider translational symmetry, which would require classification according to their *space group*. In the following we will consider the *Schoenflies* notation (rather than the *Hermann-Mauguin*):

1. The groups C_1, C_s, and C_i consist of the identity operation alone (C_1), the identity and a reflection (C_s), and the identity and an inversion (C_i).
2. The groups C_n. These groups have both identity and n-fold rotation.
3. The groups C_{nv}. In addition to the operations of C_n, this includes also n vertical reflections. An important special case is $C_{\infty v}$ to which, for example, diatomic molecules belong.
4. The groups C_{nh}. In addition to C_n operations, these groups also contain a horizontal reflection together with possible additional operations that these imply.
5. The groups D_n. In addition to the operations in C_n, these groups possess n two-fold rotations perpendicular to the n-fold (principal) axis. Also, operations that the previous imply are included.

6. The groups D_{nh}. In addition to operations in D_n, these groups also include the horizontal reflection (and other possible operations that the presence of these operations imply). An example of D_{nh} molecule is a diatomic homonuclear molecule.

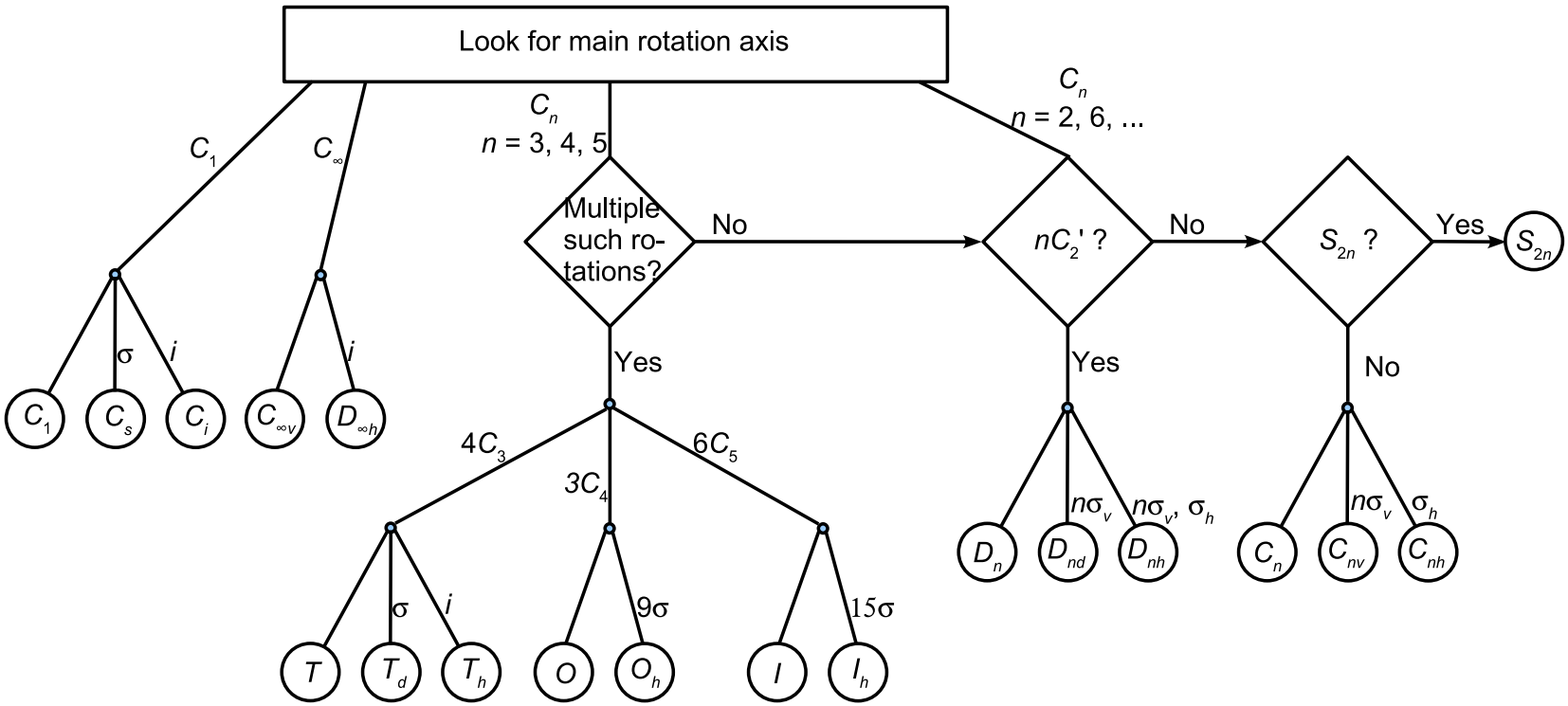
7. The groups D_{nd}. These groups contain the operations of the groups D_n and n dihedral reflections (and other possible operations that the presence of these operations imply).

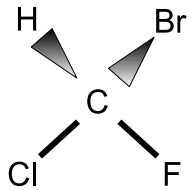
8. The groups S_n (n even). These groups contain the identity and an n-fold improper rotation (and other possible operations that the presence of these operations imply). Note that when n is odd, these are identical to C_{nh}. Also S_2 is identical to C_i.

9. The cubic and icosahedral groups. These groups contain more than one n-fold rotation with $n \geq 3$. The cubic groups are labeled as T (for tetrahedral), O for octahedral, and I for icosahedral. The group T_d is the group of the regular tetrahedron. T_h is a tetrahedral group with inversion. The regular octahedron group is called O_h and, if it lacks inversion, it is just O. In a smilar way, I_h is icosahedron and just I if it lacks inversion.

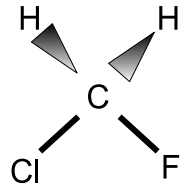
10. The full rotation group R_3. This group consists of all rotations through any angle and in any orientation. It is the symmetry group of the sphere. For example, atoms belong to this group but no molecule does. The properties of this group are related to the properties of angular momentum.

Determination of the point group for a given molecule can be a tedious task. Therefore it is helpful to use the following flowchart for determining the point group:

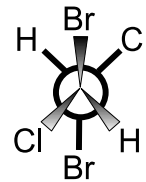




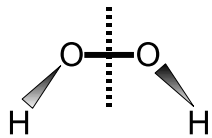
C_1: E



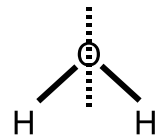
C_s: E, σ



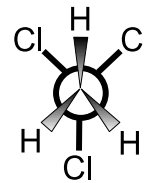
C_i: E, i



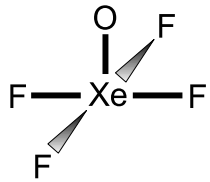
C_2: E, C_2



C_{2v}: E, C_2,
$\sigma_v(xz)$, $\sigma_v'(yz)$



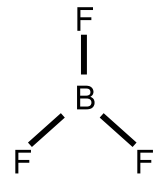
C_{3v}: E, $2C_3$, $3\sigma_v$



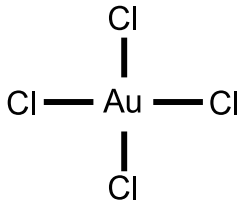
C_{4v}: E, $2C_4$, C_2,
$2\sigma_v$, $2\sigma_d$



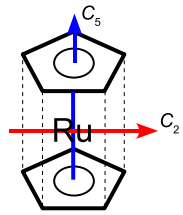
$C_{\infty v}$: E, $2C_\infty$, $\infty\sigma_v$



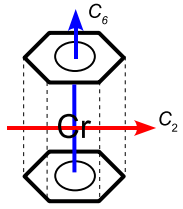
D_{3h}: E, $2C_3$, $3C_2$,
σ_h, $2S_3$, $3\sigma_v$



D_{4h}: E, $2C_4$, C_2, $2C_2'$, $2C_2''$, i, $2S_4$, σ_h, $2\sigma_v$, $2\sigma_d$



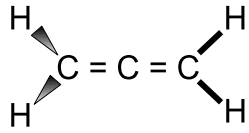
D_{5h}: E, $2C_5$, $2C_5^2$, $5C_2$, σ_h, $2S_5$, $2S_5^2$, $5\sigma_v$



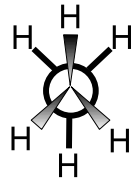
D_{6h}: E, $2C_6$, $2C_3$, C_2, $3C_2'$, $3C_2''$, i, $2S_3$, $2S_6$, σ_h, $3\sigma_d$, $3\sigma_v$



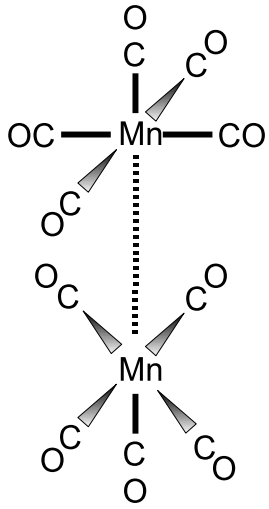
$D_{\infty h}$: E, $2C_\infty$, $\infty\sigma_v$, i, $2S_\infty$, ∞C_2



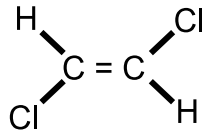
D_{2d}: E, $2S_4$, C_2, $2C_2'$, $2\sigma_d$



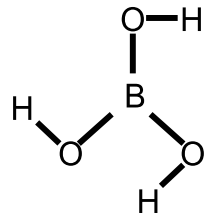
D_{3d}: E, $2C_3$, $3C_2$, i, $2S_6$, $3\sigma_d$



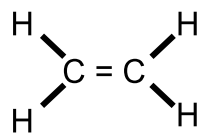
D_{4d}: E, $2S_8$, $2C_4$, $2S_8^3$, C_2, $4C_2'$, $4\sigma_d$



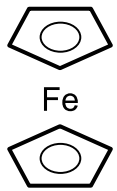
C_{2h}: E, C_2, i, σ_h



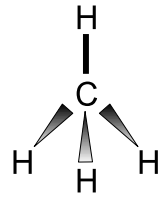
C_{3h}: E, C_3, C_3^2, σ_h, S_3, S_3^5



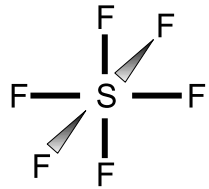
D_{2h}: E, $C_2(z)$, $C_2(y)$, $C_2(x)$, i, σ_{xy}, σ_{xz}, σ_{yz}



D_{5d}: E, $2C_5$, $2C_5^2$, $5C_2$, i, $2S_{10}^3$, $2S_{10}$, $5\sigma_d$



T_d: E, $8C_3$, $3C_2$, $6S_4$, $6\sigma_d$



O_h: E, $8C_3$, $6C_2$, $6C_4$, $3C_2$, i, $6S_4$, $8S_6$, $3\sigma_h$, $6\sigma_d$

5.3 The calculus of symmetry

We will proceed in two stages: 1) explore the properties of the symmetry operators and 2) associate symmetry operations with their matrix representations.

First we define how we "multiply" symmetry operations. We will take this to follow the same definition that we had for operators (symmetry operations are, in fact, operators). For symmetry operations S and R, we define the product SR in such way that we apply R first and then on the outcome of this, we will apply S. If SR and RS give the same overall operation, we say that S and R *commute*. The product SR always corresponds to some other single symmetry operation. For example, $\sigma_h C_2 = i$. In general, for all symmetry operations of an object, we have:

$$RS = T \tag{5.1}$$

where R, S, T are the symmetry operations. Symmetry operations are *associative*, which means that $(RS)T = R(ST)$ for all symmetry operations R, S, T. The general features of symmetry operations can be summarized:

1. The identity operation is a symmetry operation.
2. Symmetry operations combine in accord with the associative law of multiplication.

3. If R and S are symmetry operations, then RS is also a symmetry operation. Especially, when R and S are the same, we have $RR = R^2$, which is then a symmetry operation as well.
4. The inverse of each symmetry operation is also a symmetry operation. The inverse operation is defined through $RR^{-1} = R^{-1}R = E$.

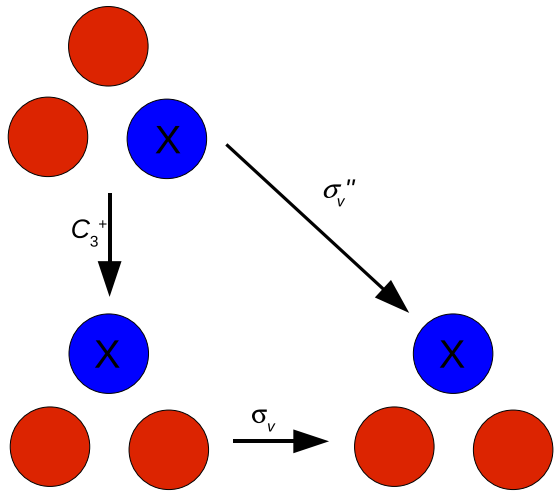
A mathematical *group* is defined as follows:

1. The identity element is part of the group.
2. The elements multiply associatively.
3. If R and S are members of the group, then RS is also a member of the group.
4. The inverse of each element is a member of the group.

Since the symmetry operations form a mathematical group, the theory dealing with the symmetry operations is called *group theory*. We saw examples of these groups earlier (C_{2v}, D_{2h}, etc.).

5.4 Group multiplication table

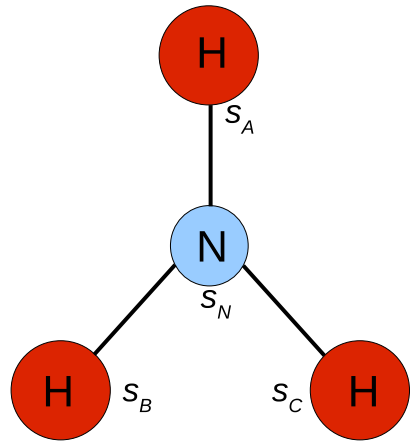
A table showing the results from multiplying the symmetry operations is called *group multiplication table*. The number of elements in a group is called the *order* of the group. For example, C_{3v} has the order 6. Note that the symmetry operations do not necessarily commute. If all of the symmetry operations in a group commute, the group is called *Abelian*. Given an object, one can construct the group multiplication table (C_{3v} below):



First:	E	C_3^+	C_3^-	σ_v	σ_v'	σ_v''
E	E	C_3^+	C_3^-	σ_v	σ_v'	σ_v''
C_3^+	C_3^+	C_3^-	E	σ_v'	σ_v''	σ_v
C_3^-	C_3^-	E	C_3^+	σ_v''	σ_v	σ_v'
σ_v	σ_v	σ_v''	σ_v'	E	C_3^-	C_3^+
σ_v'	σ_v'	σ_v	σ_v''	C_3^+	E	C_3^-
σ_v''	σ_v''	σ_v'	σ_v	C_3^-	C_3^+	E

5.5 Matrix representations

Once we define an entity defined over the object in vector notation (*basis*), we can write down the *matrix representations* for symmetry operations. There is one matrix representation for each symmetry operation in the group.



Nitrogen out of plane

NH_3 molecule (C_{3v})

We define the basis to consist of the s-orbitals located at each atom in NH_3. These are denoted by S_N, S_A, S_B, and S_C. The *dimension* of this basis is 4. We express the basis functions as a vector: (S_N, S_A, S_B, S_C). In general for a basis with dimension d, we would write $\vec{f} = (f_1, f_2, ..., f_d)$. Consider first one of the σ_v operations, which gives $\sigma_v(S_N, S_A, S_B, S_C) = (S_N, S_A, S_C, S_B)$.

The above σ_v operation can be represented by the following matrix:

$$\sigma_v(S_N, S_A, S_B, S_C) = (S_N, S_A, S_B, S_C) \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \tag{5.2}$$

The matrix representation of σ_v in the given basis is denoted by $D(\sigma_v)$. This is always a square matrix with the dimensions given by the length of the basis vector. Following the matrix muliplication rule, we can write:

$$Rf_i = \sum_j f_j D_{ji}(R) \tag{5.3}$$

where $D_{ij}(R)$ is the matrix element of representative $D(R)$ of the operation R. The other matrix representatives can be found in the same way. Note that for the identity operation we always gave $D(E)$ to be an identity matrix (i.e. ones on the diagonal with zeros everywhere else).

Example. Find the matrix representative for the operation C_3^+ in the group C_{3v} for the s-orbital basis used above.

Solution. Remember that the + sign implies clockwise rotation around the principal axis. This gives $C_3^+(S_N, S_A, S_B, S_C) = (S_N, S_B, S_C, S_A)$. The corresponding matrix representation is then:

$$C_3^+(S_N, S_A, S_B, S_C) = (S_N, S_A, S_B, S_C) \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} = (S_N, S_B, S_C, S_A)$$

The complete set of matrix representatives for this case are given below:

$D(E)$
$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$
$\chi(E) = 4$

$D(C_3^+)$
$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$
$\chi(C_3^+) = 1$

$D(C_3^-)$
$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$
$\chi(C_3^-) = 1$

$D(\sigma_v)$
$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$
$\chi(\sigma_v) = 2$

$D(\sigma_v')$
$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$
$\chi(\sigma_v') = 2$

$D(\sigma_v'')$
$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$
$\chi(\sigma_v'') = 2$

The above matrix representatives follow the group multiplication table. For example,

$$D(\sigma_v)D(C_3^+) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} = D(\sigma_v'')$$

In general we have for all members of the group:

$$\text{if } RS = T, \text{ then } D(R)D(S) = D(T) \tag{5.4}$$

Proof. Consider two elements R and S which multiply to give the element T. From Eq. (5.3) it follows that for the general basis f:

$$RSf_i = R\sum_j f_j D_{ji}(S) = \sum_{j,k} f_k D_{kj}(R) D_{ji}(S)$$

The last part is the same as a matrix product and hence we can write:

$$RSf_i = \sum_k f_k \left\{D(R)D(S)\right\}_{ki}$$

where $\{D(R)D(S)\}$ refers to the element in row k and column i of the product matrix. We also know that $RS = T$ which implies that:

$$RSf_i = Tf_i = \sum_k f_k \left\{D(T)\right\}_{ki}$$

By comparing the two equations above, we conclude that

$$\left\{D(R)D(S)\right\}_{ki} = \left\{D(T)\right\}_{ki}$$

This holds for all elements and therefore

$$D(R)D(S) = D(T)$$

5.6 The properties of matrix representations

Two matrix representations are called *similar* if the representatives for the two bases are related by the *similarity transformation*:

$$D(R) = cD'(R)c^{-1} \tag{5.5}$$
$$D'(R) = c^{-1}D(R)c \text{ (the inverse relation)}$$

where c is a (non-singular) matrix formed by the coefficients that relate the two bases. The inverse relation is obtained by multiplying from the left by c^{-1} and from the right by c.

Proof. Denote the old basis by $f(f_1, f_2, ..., f_n)$ and the new basis by $f'(f'_1, f'_2, ..., f'_n)$. The relationship between these two bases is obtained by taking a linear combination:

$$f'_i = \sum_j f_j c_{ji}$$

where the coefficients c_{ij} are constants that define linear combination. The right hand side can be written as a matrix product:

$$f' = fc$$

where c is the matrix consisting of coefficients c_{ij}. Denote the matrix representation of operation R by $D(R)$ (basis f) and $D'(R)$ (basis f') then:

$$Rf = fD(R) \text{ and } Rf' = f'D'(R)$$

We can now substitute $f' = fc$ into the last equation:

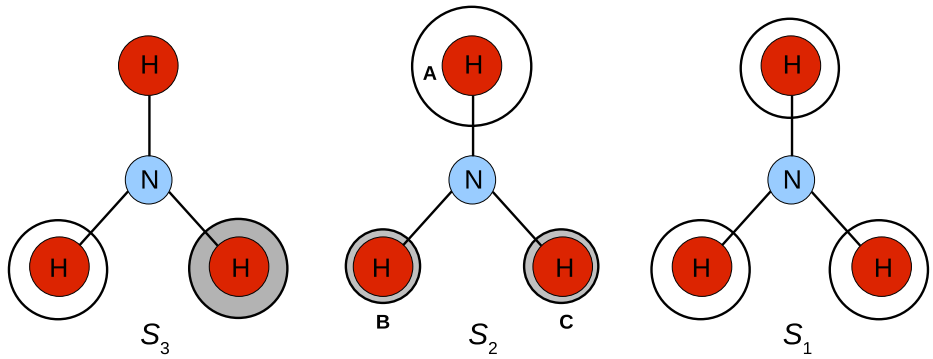
$$Rfc = fcD'(R)$$

Then we multiply from the right by c^{-1} and note that $cc^{-1} = 1$:

$$Rf = fcD'(R)c^{-1}$$

Comparison of this with $Rf = fD(R)$ gives: $D(R) = cD'(R)c^{-1}$.

Example. Recall the s-orbital basis that we used for NH_3: (S_N, S_A, S_B, S_C). We can define a similar basis as follows: (S_N, S_1, S_2, S_3) where $S_1 = S_A + S_B + S_C$, $S_2 = 2S_A - S_B - S_C$, and $S_3 = S_B - S_C$. This gives so called symmetry adapted basis functions as shown below:



The matrix representation of C_{3v} in (S_N, S_1, S_2, S_3) basis is given by:

$D(E)$
$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$
$\chi(E) = 4$

$D(C_3^+)$
$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1/2 & -1/2 \\ 0 & 0 & 1/2 & -1/2 \end{pmatrix}$$
$\chi(C_3^+) = 1$

$D(C_3^-)$
$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1/2 & 1/2 \\ 0 & 0 & -1/2 & -1/2 \end{pmatrix}$$
$\chi(C_3^-) = 1$

$D(\sigma_v)$
$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$
$\chi(\sigma_v) = 2$

$D(\sigma_v')$
$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1/2 & 1/2 \\ 0 & 0 & 3/2 & 1/2 \end{pmatrix}$$
$\chi(\sigma_v') = 2$

$D(\sigma_v'')$
$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1/2 & -1/2 \\ 0 & 0 & -3/2 & 1/2 \end{pmatrix}$$
$\chi(\sigma_v'') = 2$

Example. The matrix representation of C_3^+ in C_{3v} for the s-orbital basis (not the above symmetry adapted basis) was given previously. Construct the matrix for carrying out the similarity transformation between this basis and the symmetry adapted basis (above). Evaluate C_3^+ in the new basis (the result is given above).

Solution. The relationship between the two bases is: $S_N = S_N, S_1 = S_A + S_B + S_C, S_2 = 2S_A - S_B - S_C, S_3 = S_B - S_C$
The transformation between the bases can be represeted as:

$$(s_N, S_1, S_2, S_3) = (S_N, S_A, S_B, S_C) \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 1 & -1 & 1 \\ 0 & 1 & -1 & -1 \end{pmatrix}$$

This is the matrix c. Next we need to invert c:

$$c^{-1} = \frac{1}{6} \begin{pmatrix} 6 & 0 & 0 & 0 \\ 0 & 2 & 2 & 2 \\ 0 & 2 & -1 & -1 \\ 0 & 0 & 3 & -3 \end{pmatrix}$$

This allows us to calculate C_3^+ in the new basis:

$$D'(C_3^+) = c^{-1} D(C_3^+) c = \ldots = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1/2 & -1/2 \\ 0 & 0 & 1/2 & 1/2 \end{pmatrix}$$

5.7 The characters of representations

The diagonal sum of matrix elements is called the *character of the matrix*:

$$\chi(R) = \sum_i D_{ii}(R) \tag{5.6}$$

The two different matrix representations using the two different bases previously all had the same characters for a given operation. In mathematics this operation is called trace of the matrix and is denoted by tr. For example:

$$\chi(R) = trD(R) \tag{5.7}$$

It turns out that similarity transforms preserve the matrix trace (invariance) and this is the reason why the two previous cases results in identical matrix traces. In proving this results, we will use the following general results for matrices:

$$trABC = trCAB = trBCA \tag{5.8}$$

Proof of matrix trace invariance. First, we express the trace as a diagonal sum:

$$trABC = \sum_i (ABC)_{ii}$$

Next we expand the matrix product using matrix multiplication:

$$trABC = \sum_{ijk} A_{ij} B_{jk} C_{ki}$$

Matrix elements are just plain numbers and they can be multiplied in any order. Thus we recast the above as:

$$trABC = \sum_{ijk} B_{jk} C_{ki} A_{ij} = \sum_{j} (BCA)_{jj} = trBCA$$

This result can be used to establish the invariance of the trace under a similarity transformation ($D(R) = cD'(R)c^{-1}$):

$$\chi(R) = trD(R) = trcD'(R)c^{-1} = trD'(R)c^{-1}c = trD'(R) = \chi'(R)$$

This shows that the traces in both representations are equal.

5.8 Characters and classes

Another important observation from the two previous examples is that the characters of the two rotations are the same as well as the characters of the three reflections are the same. This can be used to divide symmetry operations into *classes*. Two symmetry operations R and R' belong to the same class if there exists another symmetry operation S of the group such that:

$$R' = S^{-1}RS \tag{5.9}$$

The elements R and R' are said to be *conjugate*. In practice this means that the symmetry operations R and R' are the same kind of operations but performed with respect to symmetry elements that are related by a symmetry operation.

Example. Show that the symmetry operations C_3^+ and C_3^- are conjugate in the group C_{3v}.

Solution. We need to find a symmetry transformation (S and S^{-1}) such that Eq. (5.9) holds ($R = C_3^+$ and $R' = C_3^-$). If we choose $S = \sigma_v$ then also $S^{-1} = \sigma_v$. Group multiplication now gives:

$$\sigma_v^{-1} C_3^+ \sigma_v = \sigma_v \left(C_3^+ \sigma_v \right) = \sigma_v \sigma_v' = C_3^-$$

$$\Rightarrow C_3^+ \text{ and } C_3^- \text{ are conjugate (belong to the same class)}$$

Next we will show that symmetry operations in the same class have the same character (χ):

Proof. We will use Eq. (5.8) ($trABC = trCAB = trBCA$) and Eq. (5.9) ($R' = S^{-1}RS$):

$$\chi(R') = trD(R') = trD^{-1}(S)D(R)D(S) = trD(R)D(S)D^{-1}(S) = trD(R) = \chi(R)$$

Although it is true that all members of the samel class have the same character in a given representation, the characters of different classes may be the same as one another.

5.9 Irreducible representations

Recall the matrix representation of C_{3v} in the basis (S_N, S_A, S_B, S_C):

$$D(E) \quad \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \chi(E) = 4$$

$$D(C_3^+) \quad \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \quad \chi(C_3^+) = 1$$

$$D(C_3^-) \quad \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{pmatrix} \quad \chi(C_3^-) = 1$$

$$D(\sigma_v) \quad \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \quad \chi(\sigma_v) = 2$$

$$D(\sigma_v') \quad \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \chi(\sigma_v') = 2$$

$$D(\sigma_v'') \quad \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \quad \chi(\sigma_v'') = 2$$

All these matrices appear to be *block-diagonal form*:

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & X & X & X \\ 0 & X & X & X \\ 0 & X & X & X \end{pmatrix}$$

This means that we can break our original four-dimensional basis into two: one consisting of S_N alone and the other three-dimensional basis:

$$\begin{array}{ccc} E & C_3^+ & C_3^- \\ (1) & (1) & (1) \\ \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} & \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} & \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \end{array}$$

$$\begin{array}{ccc} \sigma_v & \sigma_v' & \sigma_v'' \\ (1) & (1) & (1) \\ \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} & \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} & \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \end{array}$$

The first row in each case indicates the one-dimensional representation spanned by S_N and the 3×3 matrices form the three-dimensional representation spanned by the basis (S_A, S_B, S_C).

The separation of the representation into sets of matrices of lower dimension (as done above) is called the *reduction* of the representation. In this case we write:

$$D^{(4)} = D^{(3)} \oplus D^{(1)} \tag{5.10}$$

where S_N is the basis for $D^{(1)}$.

Above we say that the four-dimensional representation has been reduced to a *direct sum* of a three-dimensional and a one-dimensional representation. Note that we are not simply just adding the matrices but also changing the dimensions. The one-dimensional representation above is called *unfaithful representation* (here 1×1 matrices with all the same element 1).

Can we reduce $D^{(3)}$ further? If we use the symmetry adapted basis (as we had earlier):

$$D(E) \quad \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \chi(E) = 4$$

$$D(C_3^+) \quad \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1/2 & -1/2 \\ 0 & 0 & 1/2 & 1/2 \end{pmatrix} \quad \chi(C_3^+) = 1$$

$$D(C_3^-) \quad \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1/2 & 1/2 \\ 0 & 0 & -1/2 & -1/2 \end{pmatrix} \quad \chi(C_3^-) = 1$$

$$D(\sigma_v) \quad \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \quad \chi(\sigma_v) = 2$$

$$D(\sigma_v') \quad \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1/2 & 1/2 \\ 0 & 0 & 3/2 & 1/2 \end{pmatrix} \quad \chi(\sigma_v') = 2$$

$$D(\sigma_v'') \quad \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1/2 & -1/2 \\ 0 & 0 & -3/2 & 1/2 \end{pmatrix} \quad \chi(\sigma_v'') = 2$$

This looks like these matrices resemble more diagonal form with the off diagonal block smaller:

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & X & X \\ 0 & 0 & X & X \end{pmatrix}$$

This corresponds to the following reduction:

$$D^{(4)} = D^{(1)} \oplus D^{(1)} \oplus D^{(2)}$$

The two one-dimensional representations are unfaithful and identical to the single one-dimensional representation introduced earlier. We just reduced the $D^{(3)}$ term:

$$D^{(3)} = D^{(1)} \oplus D^{(2)}$$

Here $S_1 = S_A + S_B + S_C$ is a basis for $D^{(1)}$. Note that both basis (i.e. S_N and S_1) for the two $D^{(1)}$'s have the "same symmetry", which can be seen in the figure in Sec. 5.6. The same symmetry here is taken to mean that they act as a basis of the same matrix representation.

The next question is obviously if the remaining two-dimensional representation $D^{(2)}$ can be reduced into two one-dimensional representations? We will see later that there is no similarity transformation that would reduce this representation further and as such we say that it is an *irreducible representation* (or "irrep" for short). Another example of irreducible representation were the unfaithful representations $D^{(1)}$.

Each irreducible representation is labeled by *symmetry species*. The symmetry species is ascribed on the basis of the list of characters of the representation. For unfaithful representations (1×1 matrices) the trace for each matrix corresponding to a given symmetry operation is always one. Thus we have $(1, 1, 1, 1, 1, 1)$ in C_{3v} for $(E,\, C_3^+,\, C_3^-,\, \sigma_v,\, \sigma_v',\, \sigma_v'')$. This vector consisting of the characters is labeled by A_1. The two-dimensional irreducible representation has character $(2, -1, -1, 0, 0, 0)$ (label E). These can be identified from below:

$$
\begin{array}{ccc}
D(E) & D(C_3^+) & D(C_3^-) \\
(1) & (1) & (1) \\
(1) & (1) & (1) \\
\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} & \begin{pmatrix} -1/2 & -1/2 \\ 1/2 & -1/2 \end{pmatrix} & \begin{pmatrix} -1/2 & 1/2 \\ -1/2 & -1/2 \end{pmatrix}
\end{array}
$$

$$
\begin{array}{ccc}
D(\sigma_v) & D(\sigma_v') & D(\sigma_v'') \\
(1) & (1) & (1) \\
(1) & (1) & (1) \\
\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} & \begin{pmatrix} -1/2 & 1/2 \\ 3/2 & 1/2 \end{pmatrix} & \begin{pmatrix} -1/2 & -1/2 \\ -3/2 & 1/2 \end{pmatrix}
\end{array}
$$

We have two A_1 (S_N and S_1) and doubly degenerate E (S_2 and S_3).

In general the letter A and B are used for the one-dimensional irreducible representations, E is used for two-dimensional, and T for three-dimensional. A general irreducible representation is denoted by Γ. If a particular set of functions is a basis for an irreducible representation Γ, we say that the basis *spans* that irreducible representation. The complete list of characters of all possible irreducible representations of a group is called a *character table*. There are only finite number of irreducible representations.

The following three tasks are left:

- Determine which symmetry species of irreducible representation may occur in a group and establish their characters.
- Determine to what direct sum of irreducible representations an arbitrary matrix representation can be reduced to.
- Construct the linear combinations of members of an arbitrary basis that span a particular irreducible representation.

5.10 The great and little orthogonality theorems

Consider a group of order h and let $D^{(l)}(R)$ be the representative of the operation R in a d_l-dimensional irreducible representation of symmetry species $\Gamma^{(l)}$ of the group. The *great orthogonality theorem* states that:

$$\sum_R D^{(l)}_{ij}(R)^* D^{(l')}_{i'j'}(R) = \frac{h}{d_l}\delta_{ll'}\delta_{ii'}\delta_{jj'} \tag{5.11}$$

where * denotes complex conjugation. Here the matrices will be real and the complex conjugation will have no effect. Note that the indices i, j and i', j' refer to the elements within the matrix. The summation runs over all possible symmetry operations R in the group.

Example. Consider C_{3v} with $h = 6$ and choose a) $l = \mathrm{A}_1$, $l' = \mathrm{A}_1$, $i = j = i' = j' = 1$ and b) $l = \mathrm{E}$, $l' = \mathrm{E}$, $i = i' = 3$, $j = 4$, $j' = 3$.

Solution. a) For A_1 we have $d_l = 1$ with matrices $(1), (1), (1), (1), (1), (1)$. Here the indices i, j, i', j' point to this only element in these matrices. The sum in Eq. (5.11) now reads ($R = E, C_3^+, C_3^-, \sigma_v, \sigma_v', \sigma_v''$):

$$\sum_R D^{(\mathrm{A}_1)}_{11}(R)^* D^{(\mathrm{A}_1)}_{11}(R) = 1\times 1 + 1\times 1 + 1\times 1 + 1\times 1 + 1\times 1 + 1\times 1 = 6$$

The right-hand-side in Eq. (5.11) gives also $6/1 = 6$.

b) Here we get the left-hand-side sum of Eq. (5.11) to be:

$$\sum_R D^{(\mathrm{E})}_{34}(R)^* D^{(\mathrm{E})}_{33}(R) = D^{(\mathrm{E})}_{34}(E)^* D^{(\mathrm{E})}_{33}(E) + D^{(\mathrm{E})}_{34}(C_3^+)^* D^{(\mathrm{E})}_{33}(C_3^+) + \ldots$$

$$= 0 \times 1 + 0 \times 0 + 1 \times 0 + 1 \times 0 + 0 \times 0 + 0 \times 1 = 0$$

The right-hand-side of Eq. (5.11) is also zero.

The *little orthogonality theorem* states that:

$$\sum_R \chi^{(l)}(R)^* \chi^{(l')}(R) = h\delta_{ll'} \tag{5.12}$$

Proof. This result can be derived directly from the great orthogonality theorem by setting $j = i$ and $j' = i'$ and summing over the diagonal elements:

$$\sum_{i,i'} \sum_R D^{(l)}_{ii}(R)^* D^{(l')}_{i'i'}(R) = \sum_R \left\{ \sum_i D^{(l)}_{ii}(R)^* \right\} \left\{ \sum_{i'} D^{(l')}_{i'i'}(R) \right\}$$

$$= \sum_R \chi^{(l)}(R)^* \chi^{(l')}(R)$$

The right-hand-side of Eq. (5.11), with the same modifications as above, becomes:

$$\sum_{i,i'} \left(\frac{h}{d_l} \right) \delta_{ll'} \delta_{ii'} \delta_{ii'} = \frac{h}{d_l} \delta_{ll'} \sum_{i} \delta_{ii} = h \delta_{ll'}$$

The little orthogonality theorem can be expressed in a simpler form by using the fact that all operations of the same class have the same character. If the number of symmetry operations in a class c is $g(c)$ then:

$$\sum_{c} g(c) \chi^{(l)}(c)^* \chi^{(l')}(c) = h \delta_{ll'} \tag{5.13}$$

For example, for C_{3v} we have $g(C_3) = 2$ and $g(\sigma_v) = 3$. With $l = l'$ this becomes:

$$\sum_{c} g(c) \left| \chi^{(l)}(c) \right|^2 = h \tag{5.14}$$

By denoting $\sqrt{g(c)} \chi_c^{(l)}$ as a component $v_c^{(l)}$ of a vector $v^{(l)}$ we can write the above equation as:

$$\sum_{c} v_c^{(l)*} v_c^{(l')} = \vec{v}^{(l)*} \cdot \vec{v}^{(l')} = h \delta_{ll'} \tag{5.15}$$

This states that the two vectors are orthogonal unless $l = l'$. This also gives that the number of symmetry species is equal to the number of classes.

The vector notation can also be applied to the great orthogonality theorem. We denote $v^{(l,i,j)} = D^{(l)}_{ij}$ and then:

$$v^{(l,i,j)*} \cdot v^{(l',i',j')} = \frac{h}{d_l}\delta_{ll'}\delta_{ii'}\delta_{jj'} \tag{5.16}$$

Since the summation in Eq. (5.11) is carried out over all h operations of the group, the vectors v must be h dimensional. The total number of vectors of a given irreducible representation is d_l^2 because the labels i and j can take d_l values in a $d_l \times d_l$ matrix. The number of these orthogonal vectors clearly cannot exceed the dimension h. They can, in fact, be shown to be equal:

$$\sum_i d_l^2 = h \tag{5.17}$$

Example. Derive the C_{3v} character table.

Solution. The order of the group is $h = 6$ and there are three different classes of symmetry operations $(E, 2C_3, 3\sigma_v)$. We have already found A_1 and E previously. The dimensionality d of the missing irrep can be found using Eq. (5.17):

$$1^2 + 2^2 + d^2 = 6 \Rightarrow d = 1$$

This new irrep we will call A_2. With $l = A_2$, Eq. (5.14) gives:

$$\left(\chi^{(A_2)}(E)\right)^2 + 2\left(\chi^{(A_2)}(C_3)\right)^2 + 3\left(\chi^{(A_2)}(\sigma_v)\right)^2 = h = 6$$

With $l = \mathrm{A}_2$ and $l' = \mathrm{A}_1$ we get (Eq. (5.13)):

$$\chi^{(\mathrm{A}_2)}(E)\chi^{(\mathrm{A}_1)}(E) + 2\chi^{(\mathrm{A}_2)}(C_3)\chi^{(\mathrm{A}_1)}(C_3) + 3\chi^{(\mathrm{A}_2)}(\sigma_v)\chi^{(\mathrm{A}_1)}(\sigma_v) = 0$$

and with $l = \mathrm{A}_2$ and $l' = \mathrm{E}$:

$$\chi^{(\mathrm{A}_2)}(E)\chi^{(\mathrm{E})}(E) + 2\chi^{(\mathrm{A}_2)}(C_3)\chi^{(\mathrm{E})}(C_3) + 3\chi^{(\mathrm{A}_2)}(\sigma_v)\chi^{(\mathrm{E})}(\sigma_v) = 0$$

Substitution of the previously found characters for A_1 and E into the above two equations then gives:

$$\chi^{(\mathrm{A}_2)}(E) + 2\chi^{(\mathrm{A}_2)}(C_3) + 3\chi^{(\mathrm{A}_2)}(\sigma_v) = 0$$

$$2\chi^{(\mathrm{A}_2)}(E) - 2\chi^{(\mathrm{A}_2)}(C_3) = 0$$

We have three equations that can now determine three the three unknown characters for A_2. By solving the three equations, we get $\chi^{(\mathrm{A}_2)}(E) = 1$, $\chi^{(\mathrm{A}_2)}(C_3) = 1$, and $\chi^{(\mathrm{A}_2)}(\sigma_v) = -1$.

Table: Character table for C_1 point group (Abelian, possibly chiral).

C_1	E	Modes	Operators
A	1	$R_x, R_y, R_z, T_x, T_y, T_z$	$x, y, z, ...$

Table: Character table for $C_s = C_h$ point group (Abelian, achiral).

C_s	E	σ_h	Modes	Operators
A'	1	1	R_z, T_x, T_y	x, y, x^2, y^2, z^2, xy
A''	1	-1	R_x, R_y, T_z	z, yz, xz

Table: Character table for $C_i = S_2$ point group (Abelian, achiral).

C_i	E	i	Modes	Operators
A_g	1	1	R_x, R_y, R_z	$x^2, y^2, z^2, xy, xz, yz$
A_u	1	-1	T_x, T_y, T_z	x, y, z

Table: Character table for C_{2v} point group (Abelian, achiral).

C_{2v}	E	C_2	$\sigma_v(xz)$	$\sigma'_v(yz)$	Modes	Operators
A_1	1	1	1	1	T_z	z, x^2, y^2, z^2
A_2	1	1	-1	-1	R_z	xy
B_1	1	-1	1	-1	T_x, R_y	x, xz
B_2	1	-1	-1	1	T_y, R_x	y, yz

Table: Character table for C_{3v} point group (non-Abelian, achiral).

C_{3v}	E	$2C_3$	$3\sigma_v$	Modes	Operators
A_1	1	1	1	T_z	z, x^2+y^2, z^2
A_2	1	1	-1	R_z	
E	2	-1	0	T_x, T_y, R_x, R_y	$x, y, x^2-y^2, xy, xz, yz$

Table: Character table for C_{4v} point group (non-Abelian, achiral).

C_{4v}	E	C_2	$2C_4$	$2\sigma_v$	$2\sigma_d$	Modes	Operators
A_1	1	1	1	1	1	T_z	z, x^2+y^2, z^2
A_2	1	1	1	-1	-1	R_z	
B_1	1	1	-1	1	-1		x^2-y^2
B_2	1	1	-1	-1	1		xy
E	2	-2	0	0	0	T_x, T_y, R_x, R_y	x, y, xz, yz

Table: Character table for C_{5v} point group (non-Abelian, achiral, $\alpha = 2\pi/5$).

C_{5v}	E	$2C_5$	$2C_5^2$	$5\sigma_v$	Modes	Operators
A_1	1	1	1	1	T_z	z, x^2+y^2, z^2
A_2	1	1	1	-1	R_z	
E_1	2	$2\cos(\alpha)$	$2\cos(2\alpha)$	0	R_x, R_y, T_x, T_y	x, y, xz, yz
E_2	2	$2\cos(2\alpha)$	$2\cos(\alpha)$	0		xy, x^2-y^2

Table: Character table for C_{6v} point group (non-Abelian, achiral).

C_{6v}	E	C_2	$2C_3$	$2C_6$	$3\sigma_d$	$3\sigma_v$	Modes	Operators
A_1	1	1	1	1	1	1	T_z	z, x^2+y^2, z^2
A_2	1	1	1	1	-1	-1	R_z	
B_1	1	-1	1	-1	-1	1		
B_2	1	-1	1	-1	1	-1		
E_1	2	-2	-1	1	0	0	R_x, R_y, T_x, T_y	x, y, xz, yz
E_2	2	2	-1	-1	0	0		xy, x^2-y^2

Table: Character table for $C_{\infty v}$ point group (non-Abelian, achiral). When $\phi = \pi$ only one member in C_ϕ.

$C_{\infty v}$	E	$2C_\phi$	...	$\infty\sigma_v$	Modes	Operators
$A_1 = \Sigma^+$	1	1	...	1	T_z	z, x^2+y^2, z^2
$A_2 = \Sigma^-$	1	1	...	-1	R_z	
$E_1 = \Pi$	2	$2\cos(\phi)$	...	0	T_x, T_y, R_x, R_y	x, y, xz, yz
$E_2 = \Delta$	2	$2\cos(2\phi)$	...	0		x^2-y^2, xy
$E_3 = \Phi$	2	$2\cos(3\phi)$	...	0		
...	...	...	...	...		

Table: Character table for D_2 point group (Abelian, possibly chiral).

D_2	E	$C_2(z)$	$C_2(y)$	$C_2(x)$	Modes	Operators
A_1	1	1	1	1		x^2, y^2, z^2
B_1	1	1	-1	-1	R_z, T_z	z, xy
B_2	1	-1	1	-1	R_y, T_y	y, xz
B_3	1	-1	-1	1	R_x, T_x	x, yz

Table: Character table for D_{2h} point group (Abelian, achiral).

D_{2h}	E	$C_2(z)$	$C_2(y)$	$C_2(x)$	i	$\sigma(xy)$	$\sigma(xz)$	$\sigma(yz)$	Modes	Operators
A_g	1	1	1	1	1	1	1	1		x^2, y^2, z^2
B_{1g}	1	1	−1	−1	1	1	−1	−1	R_z	xy
B_{2g}	1	−1	1	−1	1	−1	1	−1	R_y	xz
B_{3g}	1	−1	−1	1	1	−1	−1	1	R_x	yz
A_u	1	1	1	1	−1	−1	−1	−1		
B_{1u}	1	1	−1	−1	−1	−1	1	1	T_z	z
B_{2u}	1	−1	1	−1	−1	1	−1	1	T_y	y
B_{3u}	1	−1	−1	1	−1	1	1	−1	T_x	x

Table: Character table for D_{2d} point group (non-Abelian, achiral).

D_{2d}	E	$2S_4$	C_2	$2C_2'$	$2\sigma_d$	Modes	Operators
A_1	1	1	1	1	1		x^2+y^2, z^2
A_2	1	1	1	−1	−1	R_z	
B_1	1	−1	1	1	−1		x^2-y^2
B_2	1	−1	1	−1	1	T_z	z, xy
E	2	0	−2	0	0	T_x, T_y, R_x, R_y	x, y, xz, yz

Table: Character table for D_3 point group (non-Abelian, possibly chiral).

D_3	E	$2C_3$	$3C_2'$	Modes	Operators
A_1	1	1	1		z^2, x^2+y^2
A_2	1	1	-1	R_z, T_z	z
E	2	-1	0	R_x, R_y, T_x, T_y	$x, y, xz, yz, xy, x^2-y^2$

Table: Character table for D_{3h} point group (non-Abelian, achiral).

D_{3h}	E	σ_h	$2C_3$	$2S_3$	$3C_2'$	$3\sigma_v$	Modes	Operators
A_1'	1	1	1	1	1	1		x^2+y^2, z^2
A_2'	1	1	1	1	-1	-1	R_z	
A_1''	1	-1	1	-1	1	-1		
A_2''	1	-1	1	-1	-1	1	T_z	z
E'	2	2	-1	-1	0	0	T_x, T_y	x, y, x^2-y^2, xy
E''	2	-2	-1	1	0	0	R_x, R_y	xz, yz

Table: Character table for D_4 point group (non-Abelian, possibly chiral).

D_4	E	C_2	$2C_4$	$2C_2'$	$2C_2''$	Modes	Operators
A_1	1	1	1	1	1		z^2, x^2+y^2
A_2	1	1	1	-1	-1	R_z, T_z	z
B_1	1	1	-1	1	-1		x^2-y^2
B_2	1	1	-1	-1	1		xy
E	2	-2	0	0	0	R_x, R_y, T_x, T_y	x, y, xz, yz

Table: Character table for D_{6h} point group (non-Abelian, achiral).

D_{6h}	E	$2C_6$	$2C_3$	C_2	$3C_2'$	$3C_2''$	i	$2S_3$	$2S_6$	σ_h	$3\sigma_d$	$3\sigma_v$	Modes	Operators
A_{1g}	1	1	1	1	1	1	1	1	1	1	1	1		x^2+y^2, z^2
A_{2g}	1	1	1	1	-1	-1	1	1	1	1	-1	-1	R_z	
B_{1g}	1	-1	1	-1	1	-1	1	-1	1	-1	1	-1		
B_{2g}	1	-1	1	-1	-1	1	1	-1	1	-1	-1	1		
E_{1g}	2	1	-1	-2	0	0	2	1	-1	-2	0	0	R_x, R_y	xz, yz
E_{2g}	2	-1	-1	2	0	0	2	-1	-1	2	0	0		x^2-y^2, xy
A_{1u}	1	1	1	1	1	1	-1	-1	-1	-1	-1	-1		
A_{2u}	1	1	1	1	-1	-1	-1	-1	-1	-1	1	1	T_z	z
B_{1u}	1	-1	1	-1	1	-1	-1	1	-1	1	-1	1		
B_{2u}	1	-1	1	-1	-1	1	-1	1	-1	1	1	-1		
E_{1u}	2	1	-1	-2	0	0	-2	-1	1	2	0	0	T_x, T_y	x, y
E_{2u}	2	-1	-1	2	0	0	-2	1	1	-2	0	0		

Table: Character table for $D_{\infty h}$ point group (non-Abelian, achiral).

$D_{\infty h}$	E	$2C_\phi$	...	$\infty\sigma_v$	i	$2S_\phi$	...	$\infty C_2'$	Modes	Operators
$A_{1g} = \Sigma_g^+$	1	1	...	1	1	1	...	1		$x^2 + y^2, z^2$
$A_{1u} = \Sigma_u^+$	1	1	...	1	−1	−1	...	−1	T_z	z
$A_{2g} = \Sigma_g^-$	1	1	...	−1	1	1	...	−1	R_z	
$A_{2u} = \Sigma_u^-$	1	1	...	−1	−1	−1	...	1		
$E_{1g} = \Pi_g$	2	$2\cos(\phi)$	...	0	2	$-2\cos(\phi)$	...	0	R_x, R_y	xz, yz
$E_{1u} = \Pi_u$	2	$2\cos(\phi)$	...	0	−2	$2\cos(\phi)$	...	0	T_x, T_y	x, y
$E_{2g} = \Delta_g$	2	$2\cos(2\phi)$	...	0	2	$2\cos(2\phi)$	...	0		$x^2 - y^2, xy$
$E_{2u} = \Delta_u$	2	$2\cos(2\phi)$	...	0	−2	$-2\cos(2\phi)$	...	0		

Table: Character table for T_d point group (non-Abelian, achiral).

T_d	E	$8C_3$	$3C_2$	$6S_4$	$6\sigma_d$	Modes	Operators
A_1	1	1	1	1	1		$x^2 + y^2 + z^2$
A_2	1	1	1	−1	−1		
E	2	−1	2	0	0		$2z^2 - x^2 - y^2, x^2 - y^2$
T_1	3	0	−1	1	−1	R_x, R_y, R_z	
T_2	3	0	−1	−1	1	T_x, T_y, T_z	x, y, z, xy, xz, yz

Table: Character table for O_h point group (non-Abelian, achiral).

O_h	E	$8C_3$	$3C_2$	$6C_4$	$6C_2'$	i	$8S_6$	$3\sigma_h$	$6S_4$	$6\sigma_d$	Modes	Operators
A_{1g}	1	1	1	1	1	1	1	1	1	1		$x^2+y^2+z^2$
A_{2g}	1	1	1	−1	−1	1	1	1	−1	−1		
E_g	2	−1	2	0	0	2	−1	2	0	0		$2z^2-x^2-y^2, x^2-y^2$
T_{1g}	3	0	−1	1	−1	3	0	−1	1	−1	R_x, R_y, R_z	
T_{2g}	3	0	−1	−1	1	3	0	−1	−1	1		xy, xz, yz
A_{1u}	1	1	1	1	1	−1	−1	−1	−1	−1		
A_{2u}	1	1	1	−1	−1	−1	−1	−1	1	1		
E_u	2	−1	2	0	0	−2	1	−2	0	0		
T_{1u}	3	0	−1	1	−1	−3	0	1	−1	1	T_x, T_y, T_z	x, y, z
T_{2u}	3	0	−1	−1	1	−3	0	1	1	−1		

Table: Character table for I point group (non-Abelian, possibly chiral, $\alpha = 2\pi/5$).

I	E	$12C_5$	$12C_5^2$	$20C_3$	$15C_2$	Modes	Operators
A	1	1	1	1	1		$x^2+y^2+z^2$
T_1	3	$-2\cos(2\alpha)$	$-2\cos(\alpha)$	0	−1	$R_x, R_y, R_z, T_x, T_y, T_z$	x, y, z
T_2	3	$-2\cos(\alpha)$	$-2\cos(2\alpha)$	0	−1		
G	4	−1	−1	1	0		
H	5	0	0	−1	1		$2z^2-x^2-y^2, x^2-y^2,$ xy, yz, xz

5.11 The reduction of representations

Next we will find out how to determine the irreducible representations that a given set of basis functions span. Most often we are interested in finding the irrep symmetry species (e.g. A_1, ...) and the character rather (χ) than the actual irrep matrix representation. We have seen that a representation can be expressed as a direct sum of irreducible representations:

$$D(R) = D^{(\Gamma^{(1)})}(R) \oplus D^{(\Gamma^{(2)})}(R) \oplus ... \tag{5.18}$$

by finding a similarity transformation that *simultaneously* converts the matrix representatives to block-diagonal form. We can use simplified notation for this as:

$$\Gamma = \sum_l a_l \Gamma^{(l)} \tag{5.19}$$

where a_l is the number of times the irreducible representation of symmetry species $\Gamma^{(l)}$ appears in the *direct sum*. For example, the reduction of the previously considered s-orbital basis would be written as $\Gamma = 2A_1 + E$.

First we will find the coefficients a_l. Since the matrix trace operation is not affected by a similarity transformation, we can relate the overall matrix character to taking characters of the irreps:

$$\chi(R) = \sum_l a_l \chi^{(l)}(R) \tag{5.20}$$

Next we multiply Eq. (5.20) by $\chi^{(l')}(R)^*$ and sum over R:

$$\sum_R \chi^{(l')}(R)^* \chi(R) = \sum_R \sum_l a_l \chi^{(l')}(R)^* \chi^{(l)}(R) = h \sum_l a_l \delta_{ll'} = h a_{l'}$$

where the 2nd last step follows from the little orthogonality theorem (Eq. (5.12)). Thus the coefficients a_l are given by the following relation:

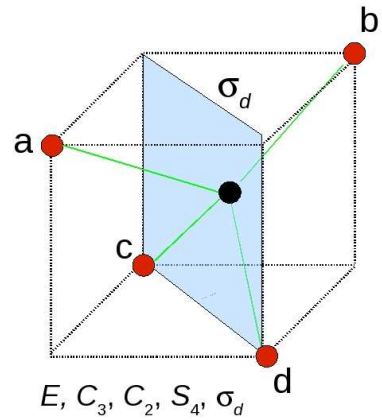
$$a_l = \frac{1}{h} \sum_R \chi^{(l)}(R)^* \chi(R) \tag{5.21}$$

Because characters for operations in the same class are identical, we can simplify the sum to go over the classes:

$$a_l = \frac{1}{h} \sum_c g(c) \chi^{(l)}(c)^* \chi(c) \tag{5.22}$$

Sometimes a simple approach works for finding a_l (Eq. (5.20)). Consider again the s-orbital example (NH_3). The characters for the three different classes are 4, 1, and 2. By inspecting the C_{3v} character table, we see that $A_1 = (1, 1, 1)$ and $E = (2, -1, 0)$. By taking $2 \times A_1 + E = 2(1, 1, 1) + (2, -1, 0) = (4, 1, 2)$. Thus we can immediately conclude that $\Gamma = 2A_1 + E$.

Example. What symmetry species do the four hydrogen 1s-orbitals of methane span?



Solution. Methane belongs to T_d point group (character table given earlier, $h = 24$). The basis is four-dimensional (H_a, H_b, H_c, H_d). We need to use Eq. (5.22). First we obtain the g-factors from T_d character table: $g(E) = 1, g(C_3) = 8, g(C_2) = 3, g(S_4) = 6, g(\sigma_d) = 6$. To get $\chi(c)$'s we use the following method:

We determine the number (N) of members left in their original localtion after the application of each symmetry operation: a 1 occurs in the diagonal of the representative in each case, and so, the character is the sum of 1 taken N times. Note that if the member of the basis moves, a zero appears along the diagonal which makes no contribution to the character. Only one operation from each class need to be considered because the characters are the same for all members of a class.

The number of unchanged basis members for the operations $E, C_3, C_2, S_4, \sigma_d$ are 4, 1, 0, 0, and 2, respectively. For example, for C_3 operation only one H atom basis function remains in its original place and hence we have 1. For E nothing moves, so we have trivially all four H atom basis functions that remain in their original places. The characters $\chi^{(l)}(c)$ follow directly from the T_d character table. For example, for A_1 we have $E : 1, C_3 : 1, C_2 : 1, \sigma_d : 1, S_4 : 1$ and for E we have $E : 2, C_3 : -1, C_2 : 2, \sigma_d : 0, S_4 : 0$. Overall we then get (Eq. (5.22)):

$$a(\mathrm{A}_1) = \frac{1}{24}\left(1 \times (4 \times 1) + 8 \times (1 \times 1) + 3 \times (0 \times 1) + 6 \times (0 \times 1) + 6 \times (2 \times 1)\right) = 1$$

$$a(\mathrm{A}_2) = \frac{1}{24}\left(1 \times (4 \times 1) + 8 \times (1 \times 1) + 3 \times (0 \times 1) + 6 \times (0 \times -1) + 6 \times (2 \times -1)\right) = 0$$

$$a(\mathrm{E}) = \frac{1}{24}\left(1 \times (4 \times 2) + 8 \times (1 \times -1) + 3 \times (0 \times 2) + 6 \times (0 \times 0) + 6 \times (2 \times 0)\right) = 0$$

$$a(\mathrm{T}_1) = \frac{1}{24}\left(1 \times (4 \times 3) + 8 \times (1 \times 0) + 3 \times (0 \times -1) + 6 \times (0 \times 1) + 6 \times (2 \times -1)\right) = 0$$

$$a(\mathrm{T}_2) = \frac{1}{24}\left(1 \times (4 \times 3) + 8 \times (1 \times 0) + 3 \times (0 \times -1) + 6 \times (0 \times -1) + 6 \times (2 \times 1)\right) = 1$$

Hence the four s-orbitals span $\Gamma = \mathrm{A}_1 + \mathrm{T}_2$.

5.12 Symmetry-adapted bases

Now we will find the linear combinations of the basis functions that span an irreducible representation of a given symmetry species. For example, we saw previously the s-orbital example for NH_3 where we were able to obtain better block-diagonal form when choosing the basis set appropriately. Now we will find a general method for doing this (i.e. finding a *symmetry-adapted basis*). The resulting basis set is called *symmetry-adapted linear combinations*.

We define a *projection operator* as follows:

$$P_{ij}^{(l)} = \frac{d_l}{h} \sum_R D_{ij}^{(l)}(R)^* R \tag{5.23}$$

This operator consists of a mixture of symmetry operations R of the group, which are weighted by the value of the matrix elements D_{ij} of the representation. The effect of this operator is:

$$P_{ij}^{(l)} f_{j'}^{(l')} = f_i^{(l)} \delta_{ll'} \delta_{jj'} \tag{5.24}$$

where $f^{(l')} = \left(f_1^{(l')}, f_2^{(l')}, ..., f_d^{(l')}\right)$ represents a basis for a $d_{l'}$-dimensional irrep $D^{(l')}$ of symmetry species $\Gamma^{(l')}$ of a group of order h.

Proof. We can express any operation of the group as (Eq. (5.3)):

$$Rf_{j'}^{(l')} = \sum_{i'} f_{i'}^{(l')} D_{i'j'}^{(l')}(R)$$

Next we multiply the complex conjugate of $D_{ij}^{(l)}(R)$, sum over the symmetry operations R, and apply the great orthogonality theorem (Eq. (5.11)):

$$\sum_R D_{ij}^{(l)}(R)^* Rf_{j'}^{(l')} = \sum_R \sum_{i'} D_{ij}^{(l)}(R)^* f_{i'}^{(l')} D_{i'j'}^{(l')}(R)$$

$$= \sum_{i'} f_{i'}^{(l')} \left\{ \sum_R D_{ij}^{(l)}(R)^* D_{i'j'}^{(l')}(R) \right\}$$

$$= \sum_{i'} f_{i'}^{(l')} \left(\frac{h}{d_{l'}} \right) \delta_{ll'} \delta_{ii'} \delta_{jj'}$$

$$= \left(\frac{h}{d_{l'}} \right) \delta_{ll'} \delta_{jj'} f_i^{(l')} = \left(\frac{h}{d_l} \right) \delta_{ll'} \delta_{jj'} f_i^{(l)}$$

This corresponds to Eqs. (5.23) and (5.24).

What does P do?

- Suppose that $l \neq l'$. In this case acting on basis function $f_{j'}^{(l')}$ with $P_{ij}^{(l)}$ gives zero. This is because $f_{j'}^{(l')}$ is not a member of the basis set that spans $\Gamma^{(l)}$.
- Suppose that $l = l'$ but $j \neq j'$. Now the basis function is not at the right location ($j \neq j'$) and the projection operator gives zero.
- Suppose that $l = l'$ and $j = j'$. In this case the projection operator moves the basis function at j (of the set that spans $\Gamma^{(l)}$) to new location i. Thus we say that P *projects* a member from one location to another location. If we know one only one member of a basis of a representation, we can project all the other members out of it.

A special case $l = l'$ and $i = j$ gives the original basis function if $i = j'$ or zero if $i \neq j'$:

$$P_{ii}^{(l)} f_{j'}^{(l)} = f_i^{(l)} \delta_{ij'} \tag{5.25}$$

Consider an arbitrary linearly independent basis set $f = (f_1, f_2, ..., f_d)$. This could be, for example, the s-orbital basis set we considered previously in NH_3. Denote the symmetry-adapted basis by f'. The two basis can be related to each other by linear combination (the coefficients have been absorbed into the $f_{j'}^{(l')}$):

$$f_j = \sum_{l',j'} f_{j'}^{(l')} \tag{5.26}$$

where l' runs over the irreducible representations and j' over the basis vectors indices in $f^{(l')}$. Operation by $P_{ii}^{(l)}$ on Eq. (5.26) gives:

$$P_{ii}^{(l)} f_j = \sum_{l',j'} P_{ii}^{(l)} f_{j'}^{(l')} = \sum_{l',j'} \delta_{ll'} \delta_{ij'} f_{j'}^{(l')} = f_i^{(l)} \tag{5.27}$$

Thus we generated from f_j the ith basis function of the symmetry-adapted basis $f^{(l)}$. Once we have found this, we can use $P_{ji}^{(l)}$ to get the jth member of $f^{(l)}$ (Eq. (5.24)):

$$P_{ji}^{(l)} f_i^{(l)} = f_j^{(l)}$$

The problem with the above approach is that we would need to know the elements of all the representatives of the irreducible representation (construction of P in Eq. (5.23)). Usually we only know that characters of the irreps (i.e. the sum over the diagonal elements). This information is also useful as can be seen by defining another projection operator $p^{(l)}$ as:

$$p^{(l)} = \sum_i P_{ii}^{(l)} = \frac{d_l}{h} \sum_{i,R} D_{ii}^{(l)}(R)^* R \tag{5.28}$$

This can be rewritten in terms of the charaters:

$$p^{(l)} = \frac{d_l}{h} \sum_R \chi^{(l)}(R)^* R \tag{5.29}$$

The above projection operator can be constructed just by using the information available in the character table. It generates a sum of the members of a basis spanning an irreducible representation:

$$p^{(l)} f_j = \sum_i P_{ii}^{(l)} f_j = \sum_i f_i^{(l)} \tag{5.30}$$

For one-dimensional irreps the sum poses no problems since there is only one member of the basis set. However, for higher-dimensional irreps the sum will contain more than one term and can cause problems. Usually, however, we are usually interested in low-dimensional irreps so this does not cause problems in practice.

Example. Construct the symmetry-adapted basis for the group C_{3v} using the s-orbital basis (see the previous NH_3 example).

Solution. We have already found out that the s-orbital basis spans $2A_1 + E$. We can then use Eq. (5.30) to construct the symmetry-adapted basis by projection. We will use the following method:

1. Prepare a table headed by the basis and show in the columns the effect of the operations. A given column is then headed by f_j and an entry in the table shows Rf_j.

Original set:	s_N	s_A	s_B	s_C
E	s_N	s_A	s_B	s_C
C_3^+	s_N	s_B	s_C	s_A
C_3^-	s_N	s_C	s_A	s_B
σ_v	s_N	s_A	s_C	s_B
σ_v'	s_N	s_B	s_A	s_C
σ_v''	s_N	s_C	s_B	s_A

2. Multiply each member of the column by the character of the corresponding operation. This step produces $\chi(R)Rf_j$ at each location.
3. Add the entries within each column. This produces $\sum_R \chi(R)Rf_i$ for a given f_j (see also Eq. (5.29)).
4. Multiply by dimension/order. This produces $p^{(l)}f_j$.

For C_{3v} we have $h = 6$. For the irreducible representation of symmetry species A_1, $d = 1$ and $\chi(R) = 1$ for all R. The first column now gives:

$$\frac{1}{6}\left(s_N + s_N + s_N + s_N + s_N + s_N\right) = s_N$$

The second column gives:

$$\frac{1}{6}\left(s_A + s_B + s_C + s_A + s_B + s_C\right) = \frac{1}{3}\left(s_A + s_B + s_C\right)$$

The remaining two columns give exactly the same outcome as above. For E, $d = 2$ and for the six symmetry operations, the characters are $\chi = (2, -1, -1, 0, 0, 0)$. The first column gives:

$$\frac{2}{6}\left(2s_N - s_N - s_N + 0 + 0 + 0\right) = 0$$

The second column gives:

$$\frac{2}{6}\left(2s_A - s_B - s_C + 0 + 0 + 0\right) = \frac{1}{3}\left(2s_A - s_B - s_C\right)$$

The remaining two columns give:

$$\frac{1}{3}\left(2s_B - s_C - s_A\right) \text{ and } \frac{1}{3}\left(2s_C - s_A - s_B\right)$$

These three functions are linearly dependent. One can form a linear combination from the last two that is orthogonal to the first:

$$\frac{1}{3}\left(2s_B - s_C - s_A\right) - \frac{1}{3}\left(2s_C - s_A - s_B\right) = s_B - s_C$$

Thus we have found four symmetry-adapted basis functions:

$$s_1 = s_N,\; s_2 = \frac{1}{3}\left(s_A + s_B + s_C\right),\; s_3 = \frac{1}{3}\left(2s_A - s_B - s_C\right),\; s_4 = s_A - s_C$$

Summary of the notation used:

R	Symmetry operation (e.g., C_3, E, etc.)
$D^{(l)}(R)$	Matrix representation for symmetry operation R in irreducible representation l (e.g., A_1, etc.).
$D_{ij}(R)$	Matrix element (i,j) of the matrix representation of R.
f	Basis functions in a form of a vector $\vec{f} = (f_1, f_2, ..., f_n)$ (e.g., $f = (S_N, S_A, S_B, S_C)$)
f_i	ith element in the basis function vector.
$f^{(l)}$	Basis functions (vector) for irreudcible representation l.
$\chi^{(l)}(R)$	Character of symmetry operation (matrix) R in irrep l.
$g(c)$	The number of symmetry operations in class c (e.g., $c = C_3$; C_3^+, C_3^-, etc.).
h	The order of the group. I.e. the total number of symetry operations in the group.
d	Dimension of the basis set.
d_l	Dimension of the basis in irreducible representation l.
c & c^{-1}	Similarity transformation matrices.

5.13 The symmetry properties of functions

Consider again NH_3 molecule (C_{3v}) and the three p-orbitals located on the nitrogen atom (denoted by p_x, p_y, p_z). We expect this to reduce into one irrep spanned by p_z and a two-dimensional irreducible representation spanned by p_x, p_y. Earlier we have found that (Sec. 3.13) the Cartesian p-orbitals can be written:

$$p_x = xf(r),\ p_y = yf(r),\ p_z = zf(r)$$

where r is the distance from the nitrogen nucleus. All operations in the point group do not operate on r so we can ignore $f(r)$ and see how (x, y, z) transform.

The effect of σ_v on (x, y, z) is:

$$\sigma_v(x,y,z) = (-x,y,z) = (x,y,z)\overbrace{\begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}}^{=D(\sigma_v)}$$

Under the rotation C_3^+ we get:

$$C_3^+(x,y,z) = \left(-\frac{1}{2}x + \frac{1}{2}\sqrt{3}y, -\frac{1}{2}\sqrt{3}x - \frac{1}{2}y, z\right) = (x,y,z)\begin{pmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} & 0 \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$D(E) \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \chi(E) = 3$$

$$D(C_3^+) \quad \begin{pmatrix} -1/2 & -\sqrt{3}/2 & 0 \\ \sqrt{3}/2 & -1/2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \chi(C_3^+) = 0$$

$$D(C_3^-) \quad \begin{pmatrix} -1/2 & \sqrt{3}/2 & 0 \\ -\sqrt{3}/2 & -1/2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \chi(C_3^-) = 0$$

$$D(\sigma_v) \quad \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \chi(\sigma_v) = 1$$

$$D(\sigma_v') \quad \begin{pmatrix} 1/2 & -\sqrt{3}/2 & 0 \\ -\sqrt{3}/2 & -1/2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \chi(\sigma_v') = 1$$

$$D(\sigma_v'') \quad \begin{pmatrix} 1/2 & \sqrt{3}/2 & 0 \\ \sqrt{3}/2 & -1/2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \chi(\sigma_v'') = 1$$

The characters of E, $2C_3$, and σ_v in the basis (x, y, z) are 3, 0, and 1. From the C_{3v} character table we can see that $\mathrm{A}_1 = (1, 1, 1)$, $\mathrm{A}_2 = (1, 1, -1)$, and $E = (2, -1, 0)$. Taking $\Gamma = \mathrm{A}_1 + \mathrm{E}$ gives $(3, 0, 1)$. The function z is basis for A_1 and the two p_x, p_y orbitals span E. Formally, one should block-diagonalize the above representations as we did earlier to prove this. Character tables usually list the symmetry species of irreps spanned by x, y, z because they are very important. We could continue this with d-orbitals where the Cartesian forms span $\Gamma = \mathrm{A}_1 + 2\mathrm{E}$ in C_{3v}. The Cartesian d-orbitals are:

$$d_{xy} = xyf(r),\ d_{yz} = yzf(r),\ d_{zx} = zxf(r)$$

$$d_{x^2-y^2} = \left(x^2 - y^2\right) f(r),\ d_{z^2} = \left(3z^2 - r^2\right) f(r)$$

5.14 The decomposition of direct-product bases

As we just saw with d-orbitals, they correspond to quadratic forms of the Cartesian components x, y, z. For example, $d_{xy} = xyf(r)$. Can we find the the symmetry species of these quadratic forms just be knowning the symmetry species for individual x, y, z? In more general terms, if we know the symmetry species spanned by a basis $(f_1, f_2, ...)$, can we state the symmetry species spanned by their products, such as $(f_1^2, f_2^2, ...)$? It turns out that this information is carried by the character tables.

If $f_i^{(l)}$ is a member of a basis for an irreducible representation of symmetry species $\Gamma^{(i)}$ of dimension d_l, and $f_{i'}^{(l')}$ is a member of a basis for an irreducible representation of symmetry species $\Gamma^{(l')}$ of dimension $d_{l'}$, then the products also form a basis for a representation, which is called a *direct-product representation* with dimension $d_l \times d_{l'}$.

Proof. Let R be a symmetry operation of a group with its matrix representation denoted by $D(R)$. Two basis functions transform as follows:

$$Rf_i^{(l)} = \sum_j f_j^{(l)} D_{ji}^{(l)}(R) \text{ and } Rf_{i'}^{(l')} = \sum_{j'} f_{j'}^{(l')} D_{j'i'}^{(l')}(R)$$

When these are multiplied, we get:

$$\left(Rf_i^{(l)}\right)\left(Rf_{i'}^{(l')}\right) = \sum_{j,j'} f_j^{(l)} f_{j'}^{(l')} D_{ji}^{(l)}(R) D_{j'i'}^{(l')}(R)$$

This is appears to be a linear combination of products $f_j^{(l)} f_{j'}^{(l')}$ with the expansion coefficients given by $D_{ji}^{(l)}(R) D_{j'i'}^{(l')}(R)$ and hence forms a basis. The sums run to d_l and $d_{l'}$ and hence the dimensionality of the expansion is $d_l \times d_{l'}$. The matrix representatitve of R in the direct-product basis is $D_{ji}^{(l)}(R) D_{j'i'}^{(l')}(R)$ with i, i' and j, j' label the columns and rows of the matrix, respectively.

Next we calculate the character of R for its matrix representation in the direct-product basis. The diagonal elements now correspond to $i = i'$ and $j = j'$. and therefore the character is given by:

$$\chi(R) = \sum_{i,i'} D_{ii}^{(l)}(R) D_{i'i'}^{(l')}(R) = \left\{\sum_i D_{ii}^{(l)}(R)\right\}\left\{\sum_{i'} D_{i'i'}^{(l')}(R)\right\} \tag{5.31}$$

$$= \chi^{(l)}(R)\chi^{(l')}(R)$$

This simply states that the characters of the operations in the direct-product basis are the products of the corresponding characters for the original bases.

Example. Determine the symmetry species of the irreducible representations spanned by (a) the quadratic forms x^2, y^2, z^2 and (b) the basis (xz, yz) in the group C_{3v}.

Solution. (a) First we use Eq. (5.31) to find out the characters of each symmetry operation in C_{3v}. We saw previously that the basis (x, y, z) spans a representation with characters 3, 0, 1 (corresponding to $E, 2C_3, 3\sigma_v$). The direct-product basis composed of x^2, y^2, z^2 therefore spans a representation with characters 9, 0, 1. Earlier we developed a simple method to find out the corresponding symmetry elements by looking for a sum of them that reproduces the right characters. Here $2\mathrm{A}_1 + \mathrm{A}_2 + 3\mathrm{E}$ yields the required character 9, 0, 1.

(b) The basis (xz, yz) is the direct product of the bases z and (x, y) which span A_1 and E, respectively. From the C_{3v} character table we get:

$$(1\ 1\ 1) \times (2\ -1\ 0) = (2\ -1\ 0)$$

This corresponds clearly to E itself. Therefore, (xz, yz) is a basis for E. This is often denoted by writing $\mathrm{A}_1 \times \mathrm{E} = \mathrm{E}$. In a similar way the direct product of (x, y) with itself, consists of the basis (x^2, xy, yx, y^2) which spans $\mathrm{E} \times \mathrm{E} = \mathrm{A}_1 + \mathrm{A}_2 + \mathrm{E}$.

Tables of decompositions of direct products like the ones above are called *direct-product tables*. Note that the direct-product $\Gamma^{(l)} \times \Gamma^{(l')}$ contains only A_1 when $l = l'$. Direct-product tables are listed in the following pages.

Table: Direct product table for C_1.

C_1	A
A	A

Table: Direct product table for C_s.

C_s	A'	A''
A'	A'	A''
A''	A''	A'

Table: Direct product table for C_i.

C_i	A_g	A_u
A_g	A_g	A_u
A_u	A_u	A_g

Table: Direct product table for C_{2v}.

C_{2v}	A_1	A_2	B_1	B_2
A_1	A_1	A_2	B_1	B_2
A_2	A_2	A_1	B_2	B_1
B_1	B_1	B_2	A_1	A_2
B_2	B_2	B_1	A_2	A_1

Table: Direct product table for C_{3v} and D_3.

C_{3v}	A_1	A_2	E
A_1	A_1	A_2	E
A_2	A_2	A_1	E
E	E	E	$A_1 + A_2 + E$

Table: Direct product table for C_{4v}, D_{2d} and D_4.

C_{4v}	A_1	A_2	B_1	B_2	E
A_1	A_1	A_2	B_1	B_2	E
A_2	A_2	A_1	B_2	B_1	E
B_1	B_1	B_2	A_1	A_2	E
B_2	B_2	B_1	A_2	A_1	E
E	E	E	E	E	$A_1 + A_2 + B_1 + B_2$

Table: Direct product table for C_{5v}.

C_{5v}	A_1	A_2	E_1	E_2
A_1	A_1	A_2	E_1	E_2
A_2	A_2	A_1	E_2	E_1
E_1	E_1	E_1	$A_1 + A_2 + E_2$	$E_1 + E_2$
E_2	E_2	E_2	$E_1 + E_2$	$A_1 + A_2 + E_2$

Table: Direct product table for C_{6v} and D_{6h}. For D_{6h}: $g \times g = g$, $g \times u = u$, $u \times g = u$, $u \times u = g$.

C_{6v}	A_1	A_2	B_1	B_2	E_1	E_2
A_1	A_1	A_2	B_1	B_2	E_1	E_2
A_2	A_2	A_1	B_2	B_1	E_1	E_2
B_1	B_1	B_2	A_1	A_2	E_2	E_1
B_2	B_2	B_1	A_2	A_1	E_2	E_1
E_1	E_1	E_1	E_2	E_2	$A_1 + A_2 + E_2$	$B_1 + B_2 + E_1$
E_2	E_2	E_2	E_1	E_1	$B_1 + B_2 + E_1$	$A_1 + A_2 + E_2$

Table: Direct product table for D_2 and D_{2h}. For D_{2h}: $g \times g = g$, $g \times u = u$, $u \times g = u$, $u \times u = g$.

D_2	A	B_1	B_2	B_3
A	A	B_1	B_2	B_3
B_1	B_1	A	B_3	B_2
B_2	B_2	B_3	A	B_1
B_3	B_3	B_2	B_1	A

Table: Direct product table for D_{3h}.

C_{3h}	A_1'	A_2'	E'	A_1''	A_2''	E''
A_1'	A_1'	A_2'	E'	A_1''	A_2''	E''
A_2'	A_2'	A_1'	E'	A_2''	A_1''	E''
E'	E'	E'	$A_1' + A_2' + E'$	E''	E''	$A_1'' + A_2'' + E''$
A_1''	A_1''	A_2''	E''	A_1'	A_2'	E'
A_2''	A_2''	A_1''	E''	A_2'	A_1'	E'
E''	E''	E''	$A_1'' + A_2'' + E''$	E'	E'	$A_1' + A_2' + E'$

Table: Direct product table for T_d and O_h. For O_h: $g \times g = g$, $g \times u = u$, $u \times g = u$, $u \times u = g$.

T_d	A_1	A_2	E	T_1	T_2
A_1	A_1	A_2	E	T_1	T_2
A_2	A_2	A_1	E	T_2	T_1
E	E	E	$A_1 + A_2 + E$	$T_1 + T_2$	$T_1 + T_2$
T_1	T_1	T_2	$T_1 + T_2$	$A_1 + E + T_1 + T_2$	$A_2 + E + T_1 + T_2$
T_2	T_2	T_1	$T_1 + T_2$	$A_2 + E + T_1 + T_2$	$A_1 + E + T_1 + T_2$

Table: Direct product table for I_h with $g \times g = g$, $g \times u = u$, $u \times g = u$, $u \times u = g$.

I_h	A	T_1	T_2	G	H
A	A	T_1	T_2	G	H
T_1	T_1	$A + T_1 + H$	$G + H$	$T_2 + G + H$	$T_1 + T_2 + G + H$
T_2	T_2	$G + H$	$A + T_2 + H$	$T_1 + G + H$	$T_1 + T_2 + G + H$
G	G	$T_2 + G + H$	$T_1 + G + H$	$A + T_1 + T_2 + G + H$	$T_1 + T_2 + G + 2H$
H	H	$T_1 + T_2 + G + H$	$T_1 + T_2 + G + H$	$T_1 + T_2 + G + 2H$	$A + T_1 + T_2 + 2G + 2H$

We define the *symmetrized direct product* as follows:

$$f_{ij}^{(+)} = \frac{1}{2}\left(f_i^{(l)} f_j^{(l)} + f_j^{(l)} f_i^{(l)}\right) \tag{5.32}$$

and the *antisymmetrized direct product*:

$$f_{ij}^{(-)} = \frac{1}{2}\left(f_i^{(l)} f_j^{(l)} - f_j^{(l)} f_i^{(l)}\right) \tag{5.33}$$

For example, for xy and yx the symmetrized direct product gives xy whereas the antisymmetrized case gives identically zero. In this case, the irreducible representations spanned by the antisymmetrized direct product should be discarded from the decomposition. The characters corresponding to the symmetrized and antisymmetrized direct products are (derivation not shown):

$$\chi^{+}(R) = \frac{1}{2}\left(\chi^{(l)}(R)^2 + \chi^{(l)}(R^2)\right) \text{ and } \chi^{-}(R) = \frac{1}{2}\left(\chi^{(l)}(R)^2 - \chi^{(l)}(R^2)\right) \tag{5.34}$$

In the direct-product tables the symmetry species of the antisymmetrized product is denoted by $[\Gamma]$. In the present case, we would denote:

$$\mathrm{E} \times \mathrm{E} = \mathrm{A}_1 + [\mathrm{A}_2] + \mathrm{E}$$

So now we would know that (x^2, xy, y^2) spans $\mathrm{A}_1 + \mathrm{E}$.

5.15 Direct-product groups

Smaller groups can be extended into larger groups by "glueing" the character tables of smaller groups together. Consider two groups G and G' both of order h. The elemments in these groups are denoted by $R_1, R_2, ..., R_h$ and $R'_1, R'_2, ..., R'_h$. Let the two group satisfy:

1. The only element in common is the identity.
2. The elements of group G commute with the elements of group G' (i.e. $RR' = R'R$).

For example, C_s and C_{3v} fulfill these rulse. The product RR' with each element of G and G' form a *direct-product group*:

$$G'' = G \otimes G' \tag{5.35}$$

Here G'' is also a group:

- The identity element is clearly $E'' = EE'$.
- Associativity can be derived from $(RS)T = R(ST)$, $(R'S')T' = R'(S'T')$ and the commutability requirement.
- $R'' = R'R$ and $T'' = T'T$, their product $R''T'' = R'RT'T = R'T'RT$. Now $R'T' \in G'$ and $RT \in G$ and hence the result belongs to G''.
- $R''R''^{-1} = RR'R^{-1}R'^{-1} = RR^{-1}R'R'^{-1} = EE' = E'' \Rightarrow$ Inverse exists and is in G''.

The order of the direct-product group is hh'. For example, for $C_s \otimes C_{3v}$ this is $2 \times 6 = 12$. The direct-product group identified by constructing its elements and the character table can be constructed from the character tables of the component groups. Let $(f_1, f_2, ...)$ be a basis for an irreducible representation of G and $(f'_1, f'_2, ...)$ be a basis for an irreducible representation G'. We can then write:

$$Rf_i = \sum_j f_j D_{ij}(R) \text{ and } R'f'_r = \sum_s f'_s D_{sr}(R') \tag{5.36}$$

Then the effect of RR' on the direct-product basis is:

$$RR' f_i f'_r = (Rf_i)(R'f'_r) = \sum_{j,s} f_j f'_s D_{ji}(R) D_{sr}(R')$$

The character of the operation RR' is now the sum of the diagonal elements:

$$\chi(RR') = \sum_{ir} D_{ii}(R) D_{rr}(R') = \chi(R)\chi(R') \tag{5.37}$$

This means that the character table of the direct-product group can be written down simply by multiplying together the appropriate characters of the two contributing groups.

Example. Construct the direct-product group $C_s \otimes C_{3v}$, identify it, and build its character table from the constituent groups.

Solution. The groups C_s and C_{3v} have, respectively, two and three classes of symmetry operations, so the direct-product group has $2 \times 3 = 6$ classes. Then it will also have six symmetry species of irreducible representations. The classes of C_s are (E, σ_h) and C_{3v} are $(E, 2C_3, 3\sigma_v)$ (from character tables). Consider first the identity operation E in C_s. When this is multiplied by the classes of C_{3v} then the same classes are reproduced $(E, 2C_3, \sigma_v)$. Multiplication of σ_h in C_s and E in C_{3v} gives again σ_h. The remaining combinations require more thinking, but essentially we get the following: $C_3\sigma_h = S_3$ and $\sigma_v\sigma_h = C_2$. The results are shown in the table below.

	E $= EE$	σ_h $= E\sigma_h$	$2C_3$ $= E2C_3$	$2S_3$ $= \sigma_h 2C_3$	$3\sigma_v$ $= E3\sigma_v$	$3C_2$ $= \sigma_h 3\sigma_v$
A_1' $(= A_1A')$	1	1	1	1	1	1
A_1'' $(= A_2A'')$	1	-1	1	-1	-1	1
A_2' $(= A_2A')$	1	1	1	1	-1	-1
A_2'' $(= A_1A''$	1	-1	1	-1	1	-1
E' $(= EA')$	2	2	-1	-1	0	0
E'' $(= EA'')$	2	-2	-1	1	0	0

Based on the character tables this corresponds to D_{3h}, so $C_s \otimes C_{3v} = D_{3h}$.

5.16 Vanishing integrals

Group theory can be used to decide wether integrals produce zero or non-zero values. For example, integration of even function from $-a$ to a will give a non-zero value whereas the same integration of an odd function would always give zero. In C_s point group, even functions g belong to A$'$ ($Eg = g$ and $\sigma_h g = g$) and odd functions f to A$''$ ($Ef = f$ and $\sigma_h f = -f$). If the integrand is not a basis for the totally symmetric irrep (here A$'$) then the integral is necessarily zero. If the integrand is a basis for the totally symmetric irrep then the integral may be non-zero.

If one considers f^2 or g^2 then it is easy to see that these are a basis for A$'$ and hence the corresponding integral may be non-zero. Furthermore, fg is necessarily zero because it is A$''$. This is consistent with the discussion above: f^2 : A$'' \times$ A$'' =$ A$'$, g^2 : A$' \times$ A$' =$ A$'$, and fg : A$'' \times$ A$' =$ A$''$. This demonstrates that the basis functions that span irreducible representations of different symmetry species are *orthogonal*.

If $f_i^{(l)}$ is the ith member of a basis that spans the irreducible representation of symmetry species $\Gamma^{(l)}$ of a group, and $f_j^{(l')}$ is the jth member of a basis that spans the irreducible representation of symmetry species $\Gamma^{(l')}$ of the same group, then for a symmetric range of integration:

$$\int f_i^{(l)*} f_j^{(l')} d\tau \propto \delta_{ll'}\delta_{ij} \tag{5.38}$$

The above result can be proven by applying the great orthogonality theorem (proof not shown). The above integral may be non-zero only when $l = l'$ and $i = j$. An important practical result can now be stated:

An integral $\int f^{(l)*} f^{(l')} d\tau$ *over a symmetric range is necessarily zero unless the integrand is a basis for the totally symmetric irreducible representation of the group. This will happen only if* $\Gamma^{(l)} = \Gamma^{(l')}$.

If three functions are multiplied inside the integral (symmetric integration range):

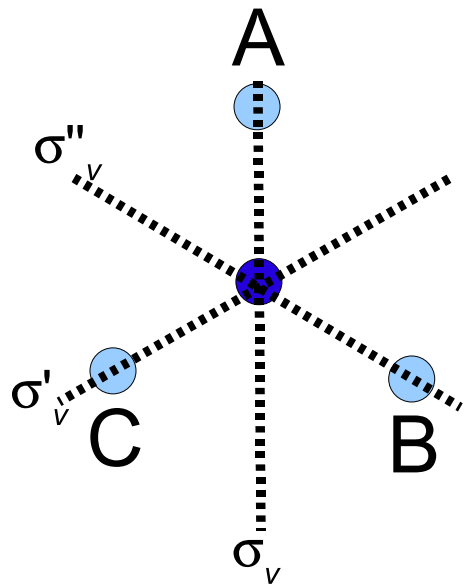
$$I = \int f^{(l)*} f^{(l')} f^{(l'')} d\tau \tag{5.39}$$

then this is necessarily zero unless the integrand is a basis for the totally symmetric irrep (such as A_1). To determine this, one has to evaluate the direct product $\Gamma^{(l)} \times \Gamma^{(l')} \times \Gamma^{(l'')}$. In quantum mechanics this is often enountered in the following form (matrix elements):

$$\langle a| \Gamma |b\rangle = \int \psi_a^* \Gamma \psi_b d\tau$$

Note that sometimes it is also possible to assign irreps to operators based on how they transform the function they operate on.

Example. Consider NH_3 molecule (C_{3v}) with just the atomic s orbitals on the hydrogens as a basis set. Note that we do not consider any functions on the nitrogen as we will try to see which of its atomic orbitals would have the right symmetry to form MOs with the hydrogen atom orbitals (AO). The hydrogen AOs should be combined to have the proper symmetry within C_{3v}. Such orbitals are called symmetry adapted linear combinations (SALCs). Label the hydrogen AOs as s_A, s_B, s_C.



The C_{3v} axis is perpendicular to the plane of the paper and goes through the nitrogen atom.

First we construct the matrix representations for the symmetry operations in C_{3v}. We go over the whole procedure (eventhough we did some of this work already earlier). The symmetry operations have the following effect on the hydrogen AOs:

	s_A	s_B	s_C
E	s_A	s_B	s_C.
C_3^-	s_C	s_A	s_B.
C_3^+	s_B	s_C	s_A.
σ_v	s_A	s_C	s_B.
σ_v'	s_B	s_A	s_C.
σ_v''	s_C	s_B	s_A.

Thus the matrix representatives can be written:

$$(A,B,C) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} A \\ B \\ C \end{pmatrix} \Rightarrow D(E) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \text{(with Tr} = 3)$$

$$(C,A,B) = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} A \\ B \\ C \end{pmatrix} \Rightarrow D(C_3^-) = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \text{(with Tr} = 0)$$

$$(B,C,A) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} A \\ B \\ C \end{pmatrix} \Rightarrow D(C_3^+) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \text{ (with Tr} = 0)$$

$$(A,C,B) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} A \\ B \\ C \end{pmatrix} \Rightarrow D(\sigma_v) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \text{ (with Tr} = 1)$$

$$(B,A,C) = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} A \\ B \\ C \end{pmatrix} \Rightarrow D(\sigma_v') = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \text{ (with Tr} = 1)$$

$$(C,B,A) = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} A \\ B \\ C \end{pmatrix} \Rightarrow D(\sigma_v'') = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \text{ (with Tr} = 1)$$

Note that the matrix trace operation is invariant under similarity transformations (i.e. multiplication by rotation matrices). Thus if we "rotate" our basis set in such a way that we choose it to be some linear combination of our present basis functions, the matrix character is unaffected by this choice.

To summarize the matrix characters:

$$\begin{array}{ccc} E & C_3^+ & \sigma_v \\ 3 & 0 & 1 \end{array}$$

Next we could proceed in finding the irreps for the matrix representatives but there is a shortcut we can take. Since the matrix character is invariant with respect to basis set rotations, we can just find the irreps that sum up to give the above characters. If we sum A_1 ($(1,1,1)$ from the character table) and E ($(2,-1,0)$ from the character table) we get:

$$A_1 + E = (1,1,1) + (2,-1,0) = (3,0,1).$$

This means that the three s orbitals may form SALCs with A_1 and E symmetries within C_{3v}. Note that E is doubly degenerate and that we have a consistent number of orbitals (three AOs giving three SALCs). This approach tells us only the symmetries of the orbitals but does not give explicit expressions for them. The expressions could be ontained by finding the diagonal matrix representations but this would involve essentially diagonalization of matrices which can be rather laborous. Instead we use the following rules for constructing the SALCs (the result will not be properly normalized):

1. Construct a table showing the effect of each operation on each orbital of the original basis (this was done already on page 241).
2. To generate the combination of a specified symmetry species, take each column in turn and:
   - i Multiply each member of the column by the character of the corresponding operation.
   - ii Add together all the orbitals in each column with the factors determined in (i).
   - iii Multiply the sum by "dimension of the symmetry element / order of the group" The dimension of the symmetry element is the character under operation E ($E = 2$, $T = 3$, ..., $A/B = 1$) and the order of the group is the total number of characters; for C_{3v} this is 6.

The first SALC with A_1 symmetry can now found to be (the s_A column multiplied by A_1 characters (1, 1, 1, 1, 1, 1); the total number of symmetry operations is 6 in C_{3v}) (dimension = 1):

$$\psi_{A_1} = \frac{1}{6}\left(s_A + s_B + s_C + s_A + s_B + s_C\right) = \frac{1}{3}\left(s_A + s_B + s_C\right)$$

From our previous consideration we know that we are still missing two orbitals, which belong to degenerate E. The same method with each column of the table (page 241) and E characters (2, −1, −1, 0, 0, 0) gives (dimension = 2):

$\psi_E = \frac{2}{6}(2s_A - s_C - s_B)$, $\psi'_E = \frac{2}{6}(2s_B - s_A - s_C)$, $\psi''_E = \frac{2}{6}(2s_C - s_B - s_A)$

We know that we should only have two orbitals in E but the above gives us three orbitals. It turns out that any one of these three expressions can be written as a sum of the other two (i.e. they are linearly dependent). The difference of the second and third equations gives:

$$\psi_E = s_B - s_C$$

which is orthogonal to the second equation. Thus the required two orthogonal SALCs are:

$$\psi_E = s_B - s_C \text{ and } \psi'_E = \frac{2}{3}(2s_B - s_A - s_C)$$

The remaining question is that which of these SALCs may have non-zero overlap with the AOs of the nitrogen atom? Recall that a non-zero overlap leads to formation of MOs. The nitrogen atom has s, p_x, p_y and p_z valence AOs, which may overlap with the SALCs. The s orbital is clearly A_1 since it is spherically symmetric. By inspecting the character table, one can see labels x, y and z on the "Operator" column. In addition to just operators, it also tells us the symmetries of the p orbitals. Thus both p_x and p_y belong to E and p_z belongs to A_1. Recall that for overlap to occur, the multiplication of orbital symmetries must give A_1. To check for this:

SALC	N AO	N AO symmetry	Overlap integral
ψ_{A_1}	s	A_1	$A_1 \times A_1 = A_1$ (**overlap**)
ψ_{A_1}	p_x	E	$A_1 \times E = E$ (no overlap)
ψ_{A_1}	p_y	E	$A_1 \times E = E$ (no overlap)
ψ_{A_1}	p_z	A_1	$A_1 \times A_1 = A_1$ (**overlap**)
ψ_E	s	A_1	$E \times A_1 = E$ (no overlap)
ψ_E	p_x	E	$E \times E = A_1$ (**overlap**)
ψ_E	p_y	E	$E \times E = A_1$ (**overlap**)
ψ_E	p_z	A_1	$E \times A_1 = E$ (no overlap)
ψ'_E	s	A_1	$E \times A_1 = E$ (no overlap)
ψ'_E	p_x	E	$E \times E = A_1$ (**overlap**)
ψ'_E	p_y	E	$E \times E = A_1$ (**overlap**)
ψ'_E	p_z	A_1	$E \times A_1 = E$ (no overlap)

Following the LCAO method, we would therefore construct three linear combinations, which form the final molecular orbitals:

LC1: $c_1\psi_{A_1} + c_2 s + c_3 p_z$ (with overall symmetry A_1)
LC2: $c_4\psi_E + c_5 p_x + c_6 p_y$ (with overall symmetry E)
LC3: $c_7\psi'_E + c_8 p_x + c_9 p_y$ (with overall symmetry E)

Example. Do the following integrals (a) $\left\langle d_{xy}|z|d_{x^2-y^2}\right\rangle$ and (b) $\left\langle d_{xy}|l_z|d_{x^2-y^2}\right\rangle$ vanish in C_{4v} molecule?

Solution. We need to check if $\Gamma^{(l)} \times \Gamma^{(l')} \times \Gamma^{(l'')}$ evaluates to A_1. (a) Based on the C_{4v} character table, d_{xy} spans B_2, $d_{x^2-y^2}$ B_1 and z A_1. The direct-product table then gives:

$$\mathrm{B}_2 \times \mathrm{A}_1 \times \mathrm{B}_1 = \mathrm{B}_2 \times \mathrm{B}_1 = \mathrm{A}_2$$

Because this is not A_1, this integral is necessarily zero. (b) l_z transforms as rotation about the z-axis (R_z) and hence it spans A_2. Direct-product evaluation now gives:

$$\mathrm{B}_2 \times \mathrm{A}_2 \times \mathrm{B}_1 = \mathrm{B}_2 \times \mathrm{B}_2 = \mathrm{A}_1$$

Since this is equal to A_1, the integral may have non-zero value.

5.17 Symmetry and degeneracy

As we have noted earlier, in quantum mechanics symmetry and degeneracy are often related. First we note that the hamiltonian H of a given system is invariant under every symmetry operation R of the relevant point group:

$$(RH) = H \tag{5.40}$$

This states that the total energy of the system does not change under a symmetry operation. For example, in the harmonic oscillator problem the kinetic energy is proportional to d^2/dx^2 and the potential energy is proportional to x^2 Both operators are invariant to inversion about the x-axis (i.e. $x \rightarrow -x$). Furthermore, H is also invariant under a similarity transformation of the group and we can write:

$$RHR^{-1} = H$$

Multiplication from the right by R then gives $RH = HR$, which means that R and H commute.

Eigenfunctions that are related by symmetry transformations of the system are degenerate.

Proof. Let ψ_i be an eigenfunction of H with eigenvalue E (i.e. $H\psi_i = E\psi_i$). Multiply the eigenvalue equation from the left by R and we obtain $RH\psi_i = ER\psi_i$. Then we insert $R^{-1}R$ (which is equal to the identity matrix) to get:

$$RHR^{-1}R\psi_i = ER\psi_i$$

From the invariance of H we then have:

$$HR\psi_i = ER\psi_i$$

Therefore, ψ_i and $R\psi_i$ correspond to the same energy E and are then degenerate.

The degree of degeneracy of a set of functions is equal to the dimension of the irreducible representation they span.

Proof. Consider a member ψ_j of a basis for an irreducible representation of dimension d of the point group in question. Suppose that the eigenvalue (i.e. energy) is E. Previously we saw that the projection operator P_{ij} (Eq. (5.23)) can be used to generate all other members of the basis if one basis member is given. Since P_{ij} is a linear combination of the symmetry operations in the group, it commutes with the hamiltonian. Hence:

$$H\psi_j = E\psi_j$$

$$\Rightarrow P_{ij}H\psi_j = HP_{ij}\psi_j = H\psi_i \text{ and } P_{ij}E\psi_j = EP_{ij}\psi_j = E\psi_i$$

$$\Rightarrow H\psi_i = E\psi_i$$

One can generate all d members of that belong to the degenerate set by applying all different possible P_{ij} operations. The dimension is always given by the character of the identity operation, $\chi(E)$.

Chapter 6: Techniques of approximation

6.1 Time-independent perturbation theory (2 levels)

In *time-independent perturbation theory* we divide the original hamiltonian H into two parts, dominant $H^{(0)}$ and less contributing $H^{(1)}$:

$$H = H^{(0)} + H^{(1)} \tag{6.1}$$

where $H^{(1)}$ is refered to as *peturbation*. Our aim is to approximately generate the wavefunction and energy of the perturbed system by using the solution of the unperturbed system.

First we consider a simple two-level system (only two eigenstates). We assume that these two eigenstates of $H^{(0)}$ are known and we denote them by $|1\rangle$ ($\psi_1^{(0)}$) and $|2\rangle$ ($\psi_2^{(0)}$). These states and functions form a complete orthonormal basis. The eigenstates have the corresponding energies $E_1^{(0)}$ and $E_2^{(0)}$. Thus both states fulfill the time-independent Schrödinger equation:

$$H^{(0)}\psi_m^{(0)} = E_m^{(0)}\psi_m^{(0)} \text{ where } m = 1, 2$$

Since we hope to solve the full Schrödinger equation:

$$H\psi = E\psi \tag{6.2}$$

we do so by using a linear combination trial function:

$$\psi = a_1\psi_1^{(0)} + a_2\psi_2^{(0)} \tag{6.3}$$

To find constants a_m, we insert the linear combination into Eq. (6.2):

$$a_1(H-E)\,|1\rangle + a_2(H-E)\,|2\rangle = 0$$

Multiplication of this equation by $\langle 1|$ and $\langle 2|$ in turn, gives two equations:

$$\begin{aligned} a_1\,(H_{11}-E) + a_2 H_{12} &= 0 \\ a_1 H_{21} + a_2\,(H_{22}-E) &= 0 \end{aligned} \tag{6.4}$$

where we have used notation $H_{mn} = \langle m|H|n\rangle$. The condition for the existence of non-trivial solution is (see your linear algebra notes):

$$\begin{vmatrix} H_{11}-E & H_{12} \\ H_{21} & H_{22}-E \end{vmatrix} = 0$$

At this point it should be noted that the coefficients a_1 and a_2 have disappeared and only the unknown E remains. Expansion of this determinant gives:

$$(H_{11}-E)\,(H_{22}-E) - H_{12}H_{21} = 0$$

$$\Rightarrow E^2 - (H_{11}+H_{22})\,E + H_{11}H_{22} - H_{12}H_{21} = 0$$

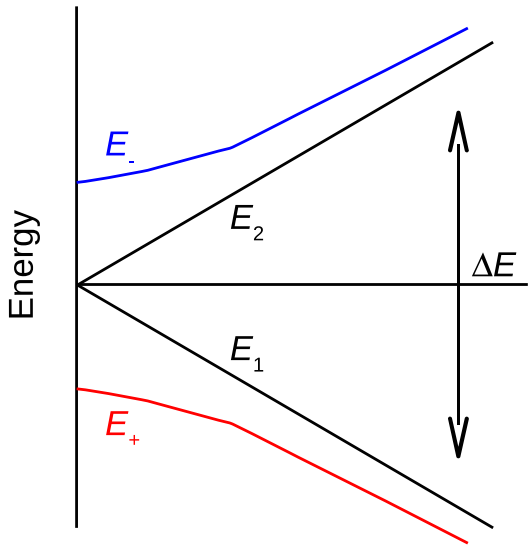
The above quadratic equation has the following two roots:

$$E_{\pm} = \frac{1}{2}\left(H_{11} + H_{22}\right) \pm \frac{1}{2}\sqrt{\left(H_{11} - H_{22}\right)^2 + 4H_{12}H_{21}} \tag{6.5}$$

In a special case, when the diagonal matrix elements of the perturbation are zero (i.e. $H^{(1)}_{mm} = 0$), we can write $H_{mm} = E^{(0)}_m$ and the above expression simplifies to:

$$E_{\pm} = \frac{1}{2}\left(E^{(0)}_1 - E^{(0)}_2\right) \pm \frac{1}{2}\sqrt{\left(E^{(0)}_1 - E^{(0)}_2\right)^2 + 4\epsilon^2} \tag{6.6}$$

where $\epsilon^2 = H^{(1)}_{12}H^{(1)}_{21}$. Since $H^{(1)}$ is hermitian, we can write $\epsilon^2 = \left|H^{(1)}_{12}\right|^2$. Note that when perturbation is absent, $\epsilon = 0$ and therefore $E_+ = E^{(0)}_1$ and $E_- = E^{(0)}_2$.



- The lower level E_+ is lowered from the unperturbed case and E_- raised.
- The perturbation is preventing the energy levels from crossing (*non-crossing rule*).
- The perturbation is greater at small ΔE values. For small ΔE, $E_+ - E_- \approx 2\epsilon$.

To summarize the effect of perturbation:

1. When a perturbation is applied, the lower level moves down in energy and the upper level moves up.
2. The closer the unperturbed states are, the greater the effect of a perturbation.
3. The stronger the perturbation, the greater the effect on the energies of the levels.

The effect of perturbation can be seen in more detail by considering the case of a weak perturbation relative to ΔE: $\epsilon^2 << \left(E_1^{(0)} - E_2^{(0)}\right)^2$. Eqn. (6.6) can now be expanded in series according to $\sqrt{1+x} = 1 + \frac{1}{2}x + ...$:

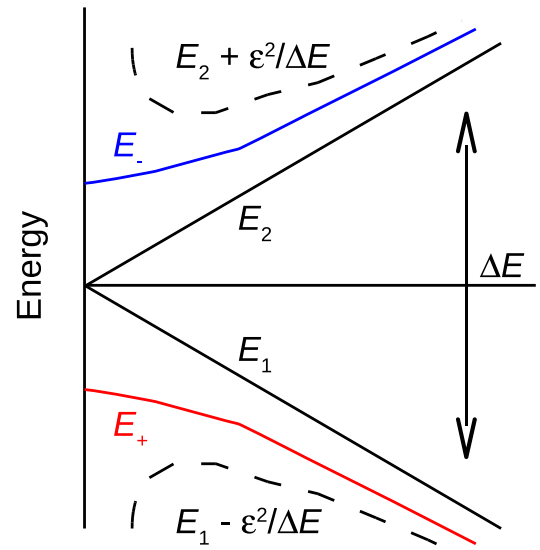
$$E_\pm = \frac{1}{2}\left(E_1^{(0)} + E_2^{(0)}\right) \pm \frac{1}{2}\left(E_1^{(0)} - E_2^{(0)}\right)\sqrt{1 + \frac{4\epsilon^2}{\left(E_1^{(0)} - E_2^{(0)}\right)^2}}$$

$$= \frac{1}{2}\left(E_1^{(0)} + E_2^{(0)}\right) \pm \frac{1}{2}\left(E_1^{(0)} - E_2^{(0)}\right)\left\{1 + \frac{2\epsilon^2}{\left(E_1^{(0)} - E_2^{(0)}\right)^2} + ...\right\}$$

When retaining the first two terms in the expansion of the square root we get (correct to 2nd order in ϵ):

$$E_+ \approx E_1^{(0)} - \frac{\epsilon^2}{\Delta E^{(0)}} \text{ and } E_- \approx E_2^{(0)} + \frac{\epsilon^2}{\Delta E^{(0)}}$$

where $\Delta E^{(0)} = E_2^{(0)} - E_1^{(0)}$. This expansion is valid when $\left(2\epsilon/\Delta E^{(0)}\right)^2 << 1$:



In the limit $\left(2\epsilon/\Delta E^{(0)}\right)^2 \to 0$, the exact solution and the approximate solutions become identical.

To get the wavefunctions (i.e. the coefficients c_1 and c_2), we need to substitute E_+ and E_- back into the original linear equations. Subsitution of E_+ and E_- in turn into Eqn. (6.4) gives:

$$\psi_+ = \psi_1^0 \cos(\zeta) + \psi_2^{(0)} \sin(\zeta) \text{ and } \psi_- = -\psi_1^{(0)} \sin(\zeta) + \psi_2^{(0)} \cos(\zeta) \tag{6.7}$$

with ($H_{12}^{(1)} = H_{21}^{(1)}$ is taken real here) gives:

$$\tan(2\zeta) = \frac{2\left|H_{12}^{(1)}\right|}{E_1^{(0)} - E_2^{(0)}} \tag{6.8}$$

If the system is degenerate (i.e. $E_1^{(0)} = E_2^{(0)}$) the above gives $\tan(2\zeta) \to \infty$ which corresponds to $\zeta \to \pi/4$. This yields the following wavefunctions:

$$\psi_+ = \frac{1}{\sqrt{2}}\left(\psi_1^{(0)} + \psi_2^{(0)}\right) \text{ and } \psi_- = \frac{1}{\sqrt{2}}\left(\psi_2^{(0)} - \psi_1^{(0)}\right) \tag{6.9}$$

This is a 50%-50% mixture of the original states. When the perturbation is small (i.e., ζ is small) compared to the energy level separation, we can take $\sin(\zeta) \approx \zeta$, $\cos(\zeta) \approx 1$, and $\tan(2\zeta) \approx 2\zeta = -\left|H_{12}^{(1)}\right| / \Delta E$. Eqn. (6.7) can then be written as:

$$\psi_+ \approx \psi_1^{(0)} - \frac{\left|H_{12}^{(1)}\right|}{\Delta E^{(0)}}\psi_2^{(0)} \text{ and } \psi_- \approx \psi_2^{(0)} + \frac{\left|H_{12}^{(1)}\right|}{\Delta E^{(0)}}\psi_1^{(0)} \tag{6.10}$$

6.2 Many-level systems

Previously we dealt with only two levels but how does the above theory generalize when there are many levels present in the system? Note that it might be possible that two levels in the set are degenerate and special precautions must be taken in such cases. In this section we will take the levels to be non-degenerate but we will return to the degerenrate case later.

Let's assume that we know the eigenvalues $E_n^{(0)}$ and eigenfunctions $|n\rangle$ (orthonormal) of the hamiltonian $H^{(0)}$, which differs from H to a small extent:

$$H^{(0)}\,|n\rangle = E_n^{(0)}\,|n\rangle \;\text{ where } n = 0, 1, 2, ... \tag{6.11}$$

In the following we will be calculating the perturbed form of state $|0\rangle$ (with energy $E_0^{(0)}$), which may no longer be the ground state of the system. The hamiltonian for the perturbed system is written as:

$$H = H^{(0)} + \lambda H^{(1)} + \lambda^2 H^{(2)} + ... \tag{6.12}$$

The parameter λ will act only as a "counter" for the order of the perturbation. In the end, the powers of λ will allow us to identify first, second, etc. order terms. The perturbed wavefunction is written as:

$$\psi_0 = \psi_0^{(0)} + \lambda\psi_0^{(1)} + \lambda^2\psi_0^{(2)} + ... \tag{6.13}$$

The energy can also be written in a similar way:

$$E_0 = E_0^{(0)} + \lambda E_0^{(1)} + \lambda^2 E_0^{(2)} + ... \tag{6.14}$$

where $E_0^{(1)}$ is the *first order correction* to the energy and $E_0^{(2)}$ is the *second order correction* etc. The overall equation that we are trying to solve is:

$$H\psi = E\psi \tag{6.15}$$

Next we insert the previous forms of H, ψ, and E into the above equation:

$$\lambda^0 \left(H^{(0)}\psi_0^{(0)} - E_0^{(0)}\psi_0^{(0)} \right)$$

$$+\lambda^1 \left(H^{(0)}\psi_0^{(1)} + H^{(1)}\psi_0^{(0)} - E_0^{(0)}\psi_0^{(1)} - E_0^{(1)}\psi_0^{(0)} \right)$$

$$+\lambda^2 \left(H^{(0)}\psi_0^{(2)} + H^{(1)}\psi_0^{(1)} + H^{(2)}\psi_0^{(0)} - E_0^{(0)}\psi_0^{(2)} - E_0^{(1)}\psi_0^{(1)} - E_0^{(2)}\psi_0^{(0)} \right)$$

$$+\text{...higher order terms...} = 0$$

Since λ is an arbitrary parameter, the terms in front of each λ must each be zero:

$$\begin{aligned} H^{(0)}\psi_0^{(0)} &= E_0^{(0)}\psi_0^{(0)} \\ \left(H^{(0)} - E_0^{(0)} \right) \psi_0^{(1)} &= \left(E_0^{(1)} - H^{(1)} \right) \psi_0^{(0)} \\ \left(H^{(0)} - E_0^{(0)} \right) \psi_0^{(2)} &= \left(E_0^{(2)} - H^{(2)} \right) \psi_0^{(0)} + \left(E_0^{(1)} - H^{(1)} \right) \psi_0^{(1)} \end{aligned} \tag{6.16}$$

and also for the higher order terms...

6.3 The first-order correction to the energy

The first equation in Eq. (6.16) is assumed to be known. The 2nd equation will give us the first order correction. Since H_0 is hermitian and its eigenfunctions for a complete basis, we can write the first order wavefunction as a linear combination:

$$\psi_0^{(1)} = \sum_n a_n \psi_n^{(0)} \tag{6.17}$$

where the sum is over all states of the system (might include continuum states). When this is inserted into the 2nd equation in Eq. (6.15), we get:

$$\sum_n a_n \left(H^{(0)} - E_0^{(0)}\right)|n\rangle = \sum_n a_n \left(E_n^{(0)} - E_0^{(0)}\right)|n\rangle = \left(E_0^{(1)} - H^{(1)}\right)|0\rangle \tag{6.18}$$

Multiplication of this equation from the left by $\langle 0|$ gives:

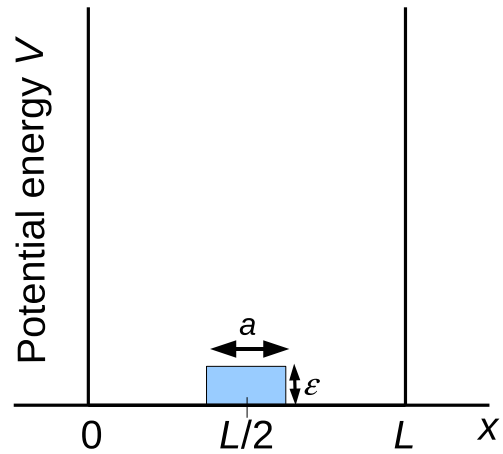
$$\sum_n a_n \left(E_n^{(0)} - E_0^{(0)}\right)\langle 0|n\rangle = \langle 0|\left(E_0^{(1)} - H^{(1)}\right)|0\rangle = E_0^{(1)} - \langle 0|\, H^{(1)}\,|0\rangle$$

The left hand side of the above equation is zero (orthogonality and $n = 0$ is zero otherwise). Therefore we can write:

$$E_0^{(1)} = \langle 0|\, H^{(1)}\,|0\rangle = H_{00}^{(1)} \tag{6.19}$$

When the diagonal elements of $\langle 0|\, H^{(1)}\,|0\rangle$ are zero, there is no 1st order correction.

Example. A small step is introduced into the one-dimensional square-well problem as shown below. Calculate the first order correction to the energy of a particle confined to the well when $a = L/10$ and (a) $n = 1$ and (b) $n = 2$.



Solution. We will use Eq. (6.19) by using the following hamiltonian:

$$H^{(1)} = \begin{cases} \epsilon, \text{ if } \frac{1}{2}(L-a) \leq x \leq \frac{1}{2}(L+a) \\ 0, \text{ if } x \text{ is outside this region} \end{cases}$$

The wavefunctions for the unperturbed system are given by Eq. (2.29). One would expect that the effect for $n = 2$ would be smaller than for $n = 1$ because the former has a node in this region. Integration then gives:

$$E_n^{(1)} = \frac{2\epsilon}{L} \int_{(L-a)/2}^{L+a)2} \sin^2\left(\frac{n\pi x}{L}\right) dx = \epsilon\left(\frac{a}{L} - \frac{(-1)^n}{n\pi}\sin\left(\frac{n\pi a}{L}\right)\right)$$

where we have used the following intergral from a tablebook (or integration by parts):

$$\int \sin^2(kx)dx = \frac{1}{2}x - \frac{1}{4k}\sin(2kx) + C$$

Now with $a = L/10$ and (a) $n = 1$, $E^{(1)} = 0.1984\epsilon$, and (b) $n = 2$, $E^{(1)} = 0.0065\epsilon$. This is just as we expected. When $\epsilon > 0$, the effect of the perturbation is to increase the energy.

6.4 The first-order correction to the wavefunction

To find out the first-order corrected wavefunction, we multiply Eq. (6.18) from the left by $\langle k|$ ($k \neq 0$):

$$\sum_n a_n \langle k| \left(E_n^{(0)} - E_0^{(0)}\right) |n\rangle = \langle k| \left(E_0^{(1)} - H^{(1)}\right) |0\rangle$$

Since the states are orthonormal, the above expression simplifies to:

$$a_k \left(E_k^{(0)} - E_0^{(0)}\right) = E_0^{(1)} \langle k|0\rangle - \langle k| H^{(1)} |0\rangle = - \langle k| H^{(1)} |0\rangle$$

Since the state $|0\rangle$ is non-degenerate, the differences $E_k^{(0)} - E_0^{(0)}$ are all non-zero for $k \neq 0$. Therefore we can carry out the division:

$$a_k = \frac{H_{k0}^{(1)}}{E_0^{(0)} - E_k^{(0)}} \tag{6.20}$$

where $H_{k0}^{(1)} = \langle k| H^{(1)} |0\rangle$. Then we can write the first-order wavefunction as:

$$\psi_0^{(1)} = \sum_k{}' \left(\frac{H_{k0}^{(1)}}{E_0^{(0)} - E_k^{(0)}}\right) \psi_k^{(0)} \tag{6.21}$$

$$\text{and } \psi_0 \approx \psi_0^{(0)} + \psi_0^{(1)}$$

where the prime means that $k = 0$ should be excluded from the summation. Note that both the matrix element $H^{(1)}_{k0}$ and the energy separation between the states 0 and k affect the extent to which the other states are mixed in. For example, states that are far away energetically from the state 0 are not expected to be mixed in by a weak perturbation.

6.5 The second-order correction to the energy

The 2nd order correction can be extracted from Eq. (6.16) the same way as the first order correction. The 2nd order correction to the wavefunction is written as a linear combination:

$$\psi_0^{(2)} = \sum_n b_n \psi_n^{(0)} \tag{6.22}$$

When this is substituted into the 3rd line in Eq. (6.16), we get:

$$\sum_n b_n \left(E_n^{(0)} - E_0^{(0)}\right) |n\rangle = \left(E_0^{(2)} - H^{(2)}\right) |0\rangle + \sum_n a_n \left(E_0^{(1)} - H^{(1)}\right) |n\rangle$$

Multiplication from the left by $\langle 0|$ gives:

$$\sum_n b_n \left(E_n^{(0)} - E_0^{(0)}\right) \langle 0|n\rangle = \langle 0| \left(E_0^{(2)} - H^{(2)}\right) |0\rangle + \sum_n a_n \langle 0| \left(E_0^{(1)} - H^{(1)}\right) |n\rangle$$

$$= E_0^{(2)} - \langle 0| H^{(2)} |0\rangle + \sum_n a_n \langle 0| \left(E_0^{(1)} - H^{(1)}\right) |n\rangle$$

$$= E_0^{(2)} - \langle 0| H^{(2)} |0\rangle + a_0 \left(E_0^{(1)} - \langle 0| H^{(1)} |0\rangle\right) + \sum_n{}' a_n \langle 0| \left(E_0^{(1)} - H^{(1)}\right) |n\rangle$$

Again the left hand side above is zero. Also the third term on the right is zero. In the last term on the right, the part $E_0^{(1)} \langle 0|n\rangle$ is also zero since the summation excludes $n = 0$ (the prime). Therefore we can get:

$$E_0^{(2)} = \langle 0| H^{(2)} |0\rangle + \sum_n{}' a_n \langle 0| H^{(1)} |n\rangle$$

The cofficients from a_k can be obtained from Eq. (6.20):

$$E_0^{(2)} = H_{00}^{(2)} + \sum_n{}' \frac{H_{0n}^{(1)} H_{n0}^{(1)}}{E_0^{(0)} - E_n^{(0)}} \tag{6.23}$$

By hermiticity we have $H_{0n}^{(1)} H_{n0}^{(1)} = H_{0n}^{(1)} H_{0n}^{(1)*} = \left|H_{0n}^{(1)}\right|^2$. If $E_n^{(1)} > E_0^{(1)}$ for all n (i.e. 0 is the ground state), the 2nd order correction lowers the energy.

Notes:

- This process can be carried out to any order and hence provides a systematic way to improve the approximation. However, it is usually not necessary to go beyond the 2nd order expansion.
- To know the energy correct to order $2n + 1$ in the perturbation, it is sufficient to know the wavefunctions only to nth order in the perturbation.
- The perturbation series is usually just assumed to converge. A more formal requirements for convergence were developed by Rellich and Kato.

6.6 Comments on the perturbation expressions

A common problem the 2nd order perturbation formula (Eq. (6.23)) is that it requires us to know all the excited states that occur in the sum. Furthermore, if there is a continuum set of states, all of these should be included.

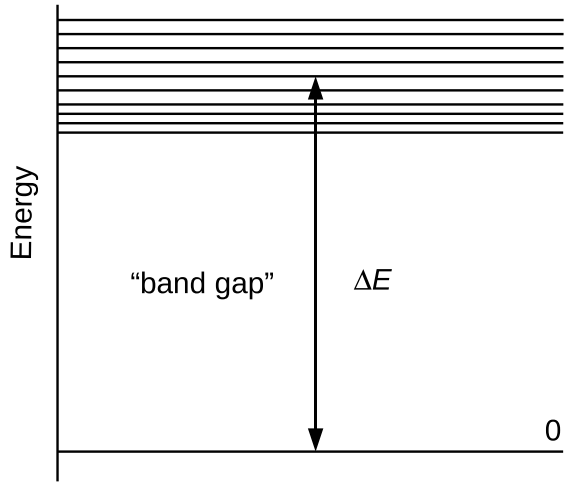
Since the energy difference between the states appears in the denominator in Eq. (6.23), it is apparent that only energetically nearby states contribute significantly to the sum. If we are calculating the ground state energy, the continuum set of states is usually far away energetically and hene these states can often be safely ignored. One should note, however, that eventhough states high in energy do not contribute much, there may be many of them and can lead to large overall contribution.

An additional simplification can be obtained by considering the symmetry properties of the matrix elements $H_{0n}^{(1)}$. By using group theory one can state that these matrix elements may be non-zero only if the direct product of the irreps corresponding to the two wavefunctions and the operator give the totally symmetric irrep:

$$\Gamma^{(0)} \times \Gamma^{(pert)} \times \Gamma^{(n)} = ... = \mathrm{A}_1 ?$$

6.7 The clousure approximation

If there is a significant band gap between the ground state and the excited states (which are located very close in energy), we can write the 2nd order correction to energy approximately as:



$$E_0^{(2)} \approx H_{00}^{(2)} - \frac{1}{\Delta E} {\sum_n}' H_{0n}^{(1)} H_{n0}^{(1)}$$

This is almost the same form as a matrix product:

$$\sum_n A_{rn} B_{nc} = (AB)_{rc}$$

To complete the first expression above in terms of a matrix product, we need to subtract the additional terms that should not be present:

$$E_0^{(2)} \approx H_{00}^{(2)} - \frac{1}{\Delta E} \sum_n H_{0n}^{(1)} H_{n0}^{(1)} + \frac{1}{\Delta E} H_{00}^{(1)} H_{00}^{(1)} \tag{6.24}$$

$$\approx H_{00}^{(2)} - \frac{1}{\Delta E} \left(H^{(1)} H^{(1)} \right)_{00} + \frac{1}{\Delta E} H_{00}^{(1)} H_{00}^{(1)}$$

Another way to see the above, is the use the completeness relation: $\sum_n |n\rangle \langle n| = 1$.

The important point is that the energy correction is now expressed only in terms of integrals over the ground state of the system and we don't need the excited state wavefunctions. Note, however, that this worked only because we approximated the energy differences by one constant ΔE. The above result is called the *closure approximation*.

The closure approximation for the 2nd order energy can be expressed as follows (derivation not shown):

$$E_0^{(2)} \approx H_{00}^{(2)} - \frac{\epsilon^2}{\Delta E} \tag{6.25}$$
$$\text{where } \epsilon^2 = \langle 0| H^{(1)2} |0\rangle - \langle 0| H^{(1)} |0\rangle^2$$

Most often the energy level structure can deviate from the picture shown on the previous slide. In this case the closure approximation is not good. If one defines ΔE as follows:

$$\Delta E = \frac{H_{00}^{(1)2} - \left(H^{(1)2}\right)_{00}}{\sum_n{}' H_{0n}^{(1)} H_{n0}^{(1)} / \left(E_0^{(0)} - E_n^{(0)}\right)} \tag{6.26}$$

then the closure approximation becomes exact. However, this requires again the knowledge of the excited states!

6.8 Perturbation theory for degenerate states

One important assumption in the previous derivation of time-independent perturbation theory was that we did not allow degeneracy between $|0\rangle$ and some other level. This is also obvious from the 1st and 2nd order energy expressions (see Eqs. (6.21) and (6.23)), which would produce infinity in case of such degeneracy. Another problem is that when the energy levels are close, even a small perturbation can introduce very large changes (i.e., the effect of perturbation is not small). Remember that any linear combination of wavefunctions belonging to a degenerate set is still an eigenfunction. This gives us extra freedom to choose our wavefunction. Our task is to work around the possible divisions by zero and see if we can find an optimum way of chooing starting combinations of the wavefunctions.

Let us assume that the level of interest is r-fold degenerate and that the states corresponding to the energy $E_0^{(0)}$ are $|0,l\rangle$, with $l = 1, 2, ..., r$, and the corresponding wavefunctions are $\psi_{0,l}^{(0)}$. All of the r states satisfy:

$$H^{(0)} |0,l\rangle = E_0^{(0)} |0,l\rangle \tag{6.27}$$

We form a linear combination of the degenerate states:

$$\phi_{0,i}^{(0)} = \sum_{l=1}^{r} c_{il} \psi_{0,l}^{(0)} \tag{6.28}$$

When the perturbation is applied, the state $\phi_{0,i}^{(0)}$ distorts into ψ_i and its energy changes from $E_0^{(0)}$ to E_i. The index i is required now because the perturbation may be such that the degeneracy is lifted. As previously, we write the perturbed wavefunction and energy in terms of λ:

$$\psi_i = \phi_{0,i}^{(0)} + \lambda\psi_{0,i}^{(1)} + ...$$

$$E_i = E_0^{(0)} + \lambda E_{0,i}^{(1)}$$

Substitution of these into $H\psi_i = E_i\psi_i$ and collection of powers of λ as done previously gives (up to 1st order shown):

$$\begin{aligned} H^{(0)}\phi_{0,i}^{(0)} &= E_0^{(0)}\phi_{0,i}^{(0)} \\ \left(H^{(0)} - E_0^{(0)}\right)\psi_{0,i}^{(1)} &= \left(E_{0,i}^{(1)} - H^{(1)}\right)\phi_{0,i}^{(0)} \end{aligned} \tag{6.29}$$

Next we express the first order corrected wavefunction $\psi_{0,i}^{(1)}$ as a linear combination of the unperturbed states. We divide the summation into two parts: degenerate and non-degenerate:

$$\psi_{0,i}^{(1)} = \sum_l a_l\psi_{0,l}^{(0)} + \sum_n{}' a_n\psi_n\psi_n^{(0)}$$

Insertion of this into the 2nd equation of Eq. (6.29) gives:

$$\sum_l a_l \left(H^{(0)} - E_0^{(0)}\right)|0,l\rangle + \sum_n{}' a_n \left(H^{(0)} - E_0^{(0)}\right)|n\rangle = \sum_l c_{il}\left(E_{0,i}^{(1)} - H^{(1)}\right)|0,l\rangle$$

Above we can replace $H^{(0)}|0,l\rangle$ by $E_0^{(0)}|0,l\rangle$ and $H^{(0)}|n\rangle$ by $E_n^{(0)}|n\rangle$:

$$\sum_l a_l \left(E_0^{(0)} - E_0^{(0)}\right)|0,l\rangle + \sum_n{}' a_n \left(E_n^{(0)} - E_0^{(0)}\right)|n\rangle = \sum_l c_{il}\left(E_{0,i}^{(1)} - H^{(1)}\right)|0,l\rangle$$

The first term above is clearly zero. Multiplication from the left by $\langle 0,k|$ gives zero on the left because the states $|n\rangle$ are orthogonal to the states $|0,k\rangle$, and we have:

$$\sum_l c_{il}\left(E_{0,i}^{(1)}\langle 0,k|0,l\rangle - \langle 0,k|H^{(1)}|0,l\rangle\right) = 0$$

The degenerate wavefunctions above are not necessary orthogonal and we introduce the following overlap integral:

$$S_{kl} = \langle 0,k|0,l\rangle \tag{6.30}$$

If the degenerate wavefunctions are orthogonal, we would have $S_{kl} = \delta_{kl}$. We also denote the hamiltonian matrix element as:

$$H_{kl}^{(1)} = \langle 0,k|\, H^{(1)}\, |0,l\rangle \tag{6.31}$$

Thus we arrive at the *secular equation*:

$$\sum_l c_{il}\left(E_{0,j}^{(1)} S_{kl} - H_{kl}^{(1)}\right) = 0 \tag{6.32}$$

There is one equation for each value of i ($i = 1, 2, ..., r$). This can also be expressed by using the *secular determinant*:

$$det\left|H_{kl}^{(1)} - E_{0,i}^{(1)} S_{kl}\right| = 0 \tag{6.33}$$

The solution of this equation will give the first-order energy correction $E_{0,i}^{(1)}$. Note that it also specifies the optimum linear combination of the degenerate functions to use for any subsequent perturbation distortion.

Example. What is the first-order correction to the energies of a doubly degenerate pair of orthonormal states?

Solution. The degenerate pair is orthogonal, which means that $S_{12} = \delta_{12} = 0$. The secular determinant then reads:

$$\begin{vmatrix} H_{11}^{(1)} - E_{0,i}^{(1)} & H_{12}^{(1)} \\ H_{21}^{(1)} & H_{22}^{(1)} - E_{0,i}^{(1)} \end{vmatrix} = 0$$

This determinant expands to:

$$\left(H_{11}^{(1)} - E_{0,i}^{(1)}\right)\left(H_{22}^{(1)} - E_{0,i}^{(1)}\right) - H_{12}^{(1)} H_{12}^{(1)} = 0$$

This gives the following quadratic equation for the energy:

$$E_{0,i}^{(1)} - \left(H_{11}^{(1)} + H_{22}^{(1)}\right) E_{0,i}^{(1)} + \left(H_{11}^{(1)} H_{22}^{(1)} - H_{12}^{(1)} H_{21}^{(1)}\right) = 0$$

The roots are the same as we obtained earlier for the two-level problem:

$$E_{0,i}^{(1)} = \frac{1}{2}\left(H_{11}^{(1)} + H_{22}^{(1)}\right) \pm \left(\left(H_{11}^{(1)} + H_{22}^{(1)}\right)^2 - 4\left(H_{11}^{(1)} H_{22}^{(1)} - H_{12}^{(1)} H_{21}^{(1)}\right)\right)^{1/2}$$

6.9 Variation theory

Variation theory provides a way to asses how good a given approximate wavefunction (denoted by ψ_{trial}) is. This theorem can be used as a guide in improving ψ_{trial}. It turns out that this becomes essentially an optimization problem with respect to the degrees of freedom present in ψ_{trial}. This method was developed by Lord Rayleigh.

Suppose that the system is described by a hamiltonian H and that the *lowest eigenvalue* is denoted by E_0. The *Rayleigh ratio* E is defined as:

$$E = \frac{\int \psi^*_{trial} H \psi_{trial} d\tau}{\int \psi^*_{trial} \psi_{trial} d\tau} \tag{6.34}$$

The *variation theorem* then states that:

$$E \geq E_0 \text{ for any } \psi_{trial} \tag{6.35}$$

Note that this result applies only to the ground state solution.

Proof. The trial function can be written as a linear combination of the true (but unknown) eigenfunctions of the hamiltonian. These eigenfunctions form a complete set since the hamiltonian is hermitian.

$$\psi_{trial} = \sum_n c_n \psi_n \text{ where } H\psi_n = E_n \psi_n$$

Next consider the following integral I:

$$I = \int \psi^*_{trial} \left(H - E_0\right) \psi_{trial} d\tau = \sum_{n,n'} c^*_n c_{n'} \int \psi^*_n \left(H - E_0\right) \psi_{n'} d\tau$$

$$= \sum_{n,n'} c^*_n c_{n'} \left(E_{n'} - E_0\right) \int \psi^*_n \psi_{n'} d\tau = \sum_n c^*_n c_n \left(E_n - E_0\right) \geq 0$$

where the last inequality follows from $E_n \geq E_0$ and $|c_n|^2 \geq 0$. Thus we have:

$$\int \psi^*_{trial} \left(H - E_0\right) \psi_{trial} d\tau \geq 0$$

$$\Rightarrow E \geq E_0$$

This means that the trial wavefunction ψ_{trial} giving the lowest value for E can be considered to be the best estimate to the true ground state eigenfunction. Also an approximate ψ_{trial} will immediately give you an upper limit esimate for the ground state energy. Often ψ_{trial} is expressed in terms of a few variational parameters and optimization of them for the lowest value of E will give the best possible ψ_{trial}.

Example. Use the trial wavefunction $\psi_{trial} = e^{-kr}$ to find an upper limit estimate for hydrogen ground state energy.

Solution. The hamiltonian for hydrogen atom is:

$$H = -\frac{\hbar^2}{2\mu}\nabla^2 - \frac{e^2}{4\pi\epsilon_0 r}$$

Since our trail wavefunction is independent of the angular variables (in spherical coordinates), we can rewrite our hamiltonian in terms of r only:

$$H = \frac{1}{r}\frac{d^2}{dr^2}(r\psi) - \frac{e^2}{4\pi\epsilon_0 r}$$

Next we need to evaluate the required integrals for Eq. (6.34):

$$\int \psi^*_{trial}\psi_{trial}d\tau = \overbrace{\int_0^{} 2\pi d\phi}^{=2\pi} \overbrace{\int_0^{\pi} \sin(\theta)d\theta}^{=2} \overbrace{\int_0^{\infty} e^{-2kr}r^2 dr}^{=1/(4k^3)} = \frac{\pi}{k^3}$$

$$\int \psi^*_{trial}\left(\frac{1}{r}\right)\psi_{trial}d\tau = \int_0^{2\pi} d\phi \int_0^{\pi} \sin(\theta)d\theta \overbrace{\int_0^{\infty} e^{-2kr}r dr}^{=1/(4k^2)} = \frac{\pi}{k^2}$$

$$\int \psi^*_{trial}\nabla^2\psi_{trial}d\tau = \int \psi^*_{trial}\left(\frac{1}{r}\frac{d^2}{dr^2}\right)re^{-kr}d\tau$$

$$= \int \psi^*_{trial} \left(k^2 - \frac{2k}{r} \right) \psi_{trial} d\tau = k^2 \int \psi^*_{trial} \psi_{trial} d\tau - 2k \int \psi^*_{trial} \left(\frac{1}{r} \right) \psi_{trial} d\tau$$

$$= \frac{\pi}{k} - \frac{2\pi}{k} = -\frac{\pi}{k}$$

Then we can write down the expectation value of H as:

$$\int \psi^*_{trial} H \psi_{trial} d\tau = \frac{\pi\hbar^2}{2\mu k} - \frac{e^2}{4\epsilon_0 k^2}$$

and the Rayleigh ratio is then:

$$E(k) = \frac{\left(\pi\hbar^2/(2\mu k)\right) - \left(e^2/(4\epsilon_0 k^2)\right)}{\pi/k^3} = \frac{k^2\hbar^2}{2\mu} - \frac{e^2 k}{4\pi\epsilon_0}$$

To find the minimum value for $E(k)$ we set $dE(k)/dk = 0$ zero:

$$\frac{dE(k)}{dk} = \frac{k\hbar^2}{\mu} - \frac{e^2}{4\pi\epsilon_0} = 0$$

This gives $k = \frac{e^2\mu}{4\pi\epsilon_0\hbar^2}$. The optimum value is then $E = -\frac{e^4\mu}{32\pi^2\epsilon_0^2\hbar^2}$.
Note that this value happens to be the true ground state energy because our trial wavefunction was such that it can become the exact ground state eigenfunction.

6.10 The Rayleigh-Ritz method

When the trial function ψ_{trial} is represented as a linear combination of some given basis functions, we obtain a modified Rayleigh variational method, which is called the *Rayleigh-Ritz* method. The coefficients appearing in the linear combination are optimize to yield the smalles value for the Rayleigh ratio (Eq. (6.34)). The trial function is written as:

$$\psi_{trial} = \sum_i c_i \psi_i \tag{6.36}$$

where only the coefficients c_i are varied but not the basis functions ψ_i. In the following we will assume that the coefficients c_i are real. The Rayleigh ratio is then:

$$E = \frac{\int \psi^*_{trial} H \psi_{trial} d\tau}{\int \psi^*_{trial} \psi_{trial} d\tau} = \frac{\sum\limits_{i,j} \int \psi_i H \psi_j d\tau}{\sum\limits_{i,j} c_i c_j \int \psi_i \psi_j d\tau} = \frac{\sum\limits_{i,j} c_i c_j H_{ij}}{\sum\limits_{i,j} c_i c_j S_{ij}} \tag{6.37}$$

To find the minimum value for this ratio, we differentiate with respect to each c_i and require that these derivatives are zero:

$$\frac{\partial E}{\partial c_k} = \frac{\sum\limits_j c_j H_{kj} + \sum\limits_i c_i H_{ik}}{\sum\limits_{i,j} c_i c_j S_{ij}} - \frac{\left(\sum\limits_j c_j S_{kj} + \sum\limits_i c_i S_{ik} \right) \sum\limits_{i,j} c_i c_j H_{ij}}{\left(\sum\limits_{i,j} c_i c_j S_{ij} \right)^2}$$

$$= \frac{\sum_j c_j \left(H_{kj} - ES_{kj}\right)}{\sum_{i,j} c_i c_j S_{ij}} + \frac{\sum_i c_i \left(H_{ik} - ES_{ik}\right)}{\sum_{i,j} c_i c_j S_{ij}} = 0$$

This equation is staisfied if the numerators are zero and hence we end up with the secular equations:

$$\sum_i \left(H_{ik} - ES_{ik}\right) c_i = 0 \tag{6.38}$$

In the matrix form this is essentially $Ac = 0$ where the matrix A has elements above shown in parentheses and c is a vector consisting of the c_i's. For such equation to have a non-trivial solution, the determinant corresponding to matrix A must be zero:

$$\det \left| H_{ik} - ES_{ik} \right| = 0 \tag{6.39}$$

The determinant, which does not have the coefficients c_i in it anymore, expands as a polynomial in terms of E. If the original matrix was $n \times n$, then the resulting polynomial will have degree n. The lowest energy from the roots should be selected as the approximate value for the energy. Note that it may not be trivial to find the roots for high order polynomials. Once the energy E is known, one can go back to Eq. (6.38) and solve for the coefficients c_i.

Notes:

- The variation principle leads to an *upper bound* for the energy.
- It is also possible to determine an upper bound for the first excited state by formulating a trial function that is orthogonal to the ground state function. In principle this procedure can be continued through out the rest of the excited states as well (may not be very practical).
- There are also methods for finding the *lower bound* for the energy. Thus one could sandwitch the value between the limits.
- The wavefunction is optimized in terms of energy and there is no guarantee that the resulting wavefunction would be able to describe any other property of the system well.

6.11 The Hellmann-Feynman theorem

Consider a system that is described by a hamiltonian that depends parametrically on parameter P. This might be a coordinate of an atom in a molecule or the strength of the electric field to which the molecule is exposed. The exact solution (both energy E and the wavefunction ψ) to the Schrödinger equation then also depends on P. In the following we will derive a result that will tell us how the energy E of the system varies as a function of P (the *Hellmann-Feynman theorem*):

$$\frac{dE}{dP} = \left\langle \frac{\partial H}{\partial P} \right\rangle \tag{6.40}$$

Proof. Let the exact wavefunction to be normalized to one for all values of P. Then we can write:

$$E(P) = \int \psi^*(P) H(P) \psi(P) d\tau$$

The derivative of E with respect to P is then:

$$\begin{aligned} \frac{dE}{dP} &= \int \left(\frac{\partial \psi^*}{\partial P}\right) H\psi d\tau + \int \psi^* \left(\frac{\partial H}{\partial P}\right) \psi d\tau + \int \psi^* H \left(\frac{\partial \psi}{\partial P}\right) d\tau \\ &= E \int \left(\frac{\partial \psi^*}{\partial P}\right) \psi d\tau + \int \psi^* \left(\frac{\partial H}{\partial P}\right) \psi d\tau + E \int \psi^* \left(\frac{\partial \psi}{\partial P}\right) d\tau \end{aligned}$$

$$= E\frac{d}{dP}\int \psi^* \psi d\tau + \int \psi^* \left(\frac{\partial H}{\partial P}\right) \psi d\tau$$

$$= \int \psi^* \left(\frac{\partial H}{\partial P}\right) \psi d\tau = \left\langle \frac{\partial H}{\partial P} \right\rangle$$

The Hellmann-Feynman theorem is particularly useful result when the hamiltonian depends on P in a simple way. For example, $H = H^{(0)} + Px$ would give $\partial H/\partial P = x$, which simply means that $dE/dP = \langle x \rangle$. The main disadvantage of the Hellmann-Feynman theorem is that the wavefunction must be the exact solution to the Schrödinger equation. It can be shown that this requirement can be relaxed so that it sufficient for the wavefunction to be variationally optimized. This can be very useful for computational chemistry methods and, for example, the Carr-Parrinello *ab initio* based molecular dynamics is largely based on this result. In this case the forces acting on the nuclei are calculated using the Hellmann-Feynman theorem. In other cases it may be possible to use the perturbation theory to calculate the response of the system on changing P.

6.12 Time-dependent perturbation theory

Most perturbations that are considered in physical chemistry are time-dependent. Even stationary perturbations must be turned on and there is some initial period when the perturbation is essentially time dependent. All spectroscopic methods require time-dependent treatment since they use either the oscillating electric or magnetic field components of photons passing through the sample. Sometimes the response of the molecule to a time-dependent perturbation is so fast that a time-independent situation is observed fairly quickly. In these cases time-independent treatment would be sufficient. Note, howevere, that there are many kinds of perturbations that never "settle down" and hence a true time-dependent treatment is required (e.g., spectroscopy). In the following, we will first consider a two-level system (i.e. ground and excited state) and later extend the theory to describe many levels.

The total hamiltonian of the system is written as a sum of time independent and time dependent parts:

$$H = H^{(0)} + H^{(1)}(t) \tag{6.41}$$

A typical example of a time-dependent perturbation is one that oscillates at an angular frequency ω:

$$H^{(1)}(t) = 2H^{(1)} \cos(\omega t) \tag{6.42}$$

where $H^{(1)}$ is a time-dependent perturbation operator. We have to start with the time-dependent Schrödinger equation:

$$H\Psi = i\hbar\frac{\partial\Psi}{\partial t} \tag{6.43}$$

We consider two levels, which have energies $E_1^{(0)}$ and $E_2^{(0)}$ with the corresponding eigenfunctions $\psi_1^{(0)}$ and $\psi_2^{(0)}$:

$$H^{(0)}\psi_n^{(0)} = E_n^{(0)}\psi_n^{(0)} \tag{6.44}$$

with their time-dependent phase factors that satisfy the time-dependent Schrödinger equation:

$$\Psi_n^{(0)}(t) = \psi_n^{(0)} e^{-iE_n^{(0)}t/\hbar} \tag{6.45}$$

In the presence of the perturbation $H^{(1)}(t)$, the state of the system is expressed as a linear combination of the basis functions:

$$\Psi(t) = a_1(t)\Psi_1^{(0)}(t) + a_2(t)\Psi_2^{(0)}(t) \tag{6.46}$$

Here the coefficients are also time-dependent since the composition of the state may evolve with time. The overall time-dependency arises then from both the osillatory behavior in the basis functions as well as the coefficients a_i. The probability for the system to be in state i $(i = 1, 2)$ is given by $|a_i|^2$.

Substitution of Eq. (6.46) into the time-dependent Schrödinger equation (Eq. (6.45)) gives:

$$H\Psi = a_1 H^{(0)}\Psi_1^{(0)} + a_1 H^{(1)}(t)\Psi_1^{(0)} + a_2 H^{(0)}\Psi_2^{(0)} + a_2 H^{(1)}(t)\Psi_2^{(0)}$$

$$= i\hbar \frac{\partial}{\partial t}\left(a_1 \Psi_1^{(0)} + a_2 \Psi_2^{(0)}\right)$$

$$= i\hbar a_1 \frac{\partial \Psi_1^{(0)}}{\partial t} + i\hbar \Psi_1^{(0)} \frac{da_1}{dt} + i\hbar a_2 \frac{\partial \Psi_2^{(0)}}{\partial t} + i\hbar \Psi_2^{(0)} \frac{da_2}{dt}$$

Each basis function satisfies:

$$H^{(0)} \Psi_n^{(0)} = i\hbar \frac{\partial \Psi_n^{(0)}}{\partial t}$$

which allows us to rewrite the previous equation as ($\dot{a}_i \equiv da_i/dt$):

$$a_1 H^{(1)}(t) \Psi_1^{(0)} + a_2 H^{(1)}(t) \Psi_2^{(0)} = i\hbar \dot{a}_1 \Psi_1^{(0)} + i\hbar \dot{a}_2 \Psi_2^{(0)}$$

Next we put in the explicit forms for $\Psi_i^{(0)}$'s and use the Dirac notation for $\psi_i^{(0)}$:

$$a_1 H^{(1)}(t) \left|1\right\rangle e^{-iE_1^{(0)} t/\hbar} + a_2 H_{12}^{(1)}(t) \left|2\right\rangle e^{-iE_2^{(0)} t/\hbar}$$

$$= i\hbar \dot{a}_1 \left|1\right\rangle e^{-iE_1^{(0)} t/\hbar} + i\hbar \dot{a}_2 \left|2\right\rangle e^{-iE_2^{(0)} t/\hbar}$$

Next we multiply by $\langle 1|$ from the left and use the orthonormality of the states:

$$a_1 H_{11}^{(1)}(t) e^{-iE_1^{(0)} t/\hbar} + a_2 H_{12}^{(1)}(t) e^{-iE_2^{(0)} t/\hbar} = i\hbar \dot{a}_1 e^{-iE_1^{(0)} t/\hbar}$$

where we have denoted $H^{(1)}_{ij}(t) = \langle i \left| H^{(1)}(t) \right| j \rangle$. We can simplify the expression further by setting $\hbar\omega_{21} = E^{(0)}_2 - E^{(0)}_1$:

$$a_1 H^{(1)}_{11}(t) + a_2 H^{(1)}_{12}(t) e^{-\omega_{21} t} = i\hbar \dot{a}_1 \tag{6.47}$$

Often the diagonal elements of the perturbation are zero (i.e. $H_{ii} = 0$, $i = 1, 2$). This yields (+ similar treatment for a_2 when multiplied by $\langle 2|$):

$$\begin{aligned} \dot{a}_1 &= \frac{1}{i\hbar} a_2 H^{(1)}_{12}(t) e^{-i\omega_{21} t} \\ \dot{a}_2 &= \frac{1}{i\hbar} a_2 H^{(1)}_{21}(t) e^{-i\omega_{21} t} \end{aligned} \tag{6.48}$$

Case 1: No perturbation ($H_{12}(t) = H_{21}(t) = 0$). Eq. (6.48) gives that $\dot{a}_1 = \dot{a}_2 = 0$. Thus the coefficients $a_i(t)$ are constant and the only time evolution present in the wavefunction Ψ is due to the exponential phase factors.

Case 2: Constant perturbation applied at $t = 0$, $H^{(1)}_{12} = \hbar V$ and by hermiticity $H^{(1)}_{21} = \hbar V^*$. The equations become then:

$$\dot{a}_1 = -iV a_2 e^{-i\omega_{21} t} \text{ and } \dot{a}_2 = -iV^* a_1 e^{i\omega_{21} t} \tag{6.49}$$

To solve the coupled differential equations, we differentiate $\dot{a}_2$ and then use the expression for $\dot{a}_1$:

$$\ddot{a}_2 = -iV^*\dot{a}_1 e^{i\omega_{21}t} + \omega_{21}V^* a_1 e^{i\omega_{21}t} = -|V|^2 a_2 + i\omega_{21}\dot{a}_2 \tag{6.50}$$

The corresponding expression for $\ddot{a}_1$ can be obtained by exchanging the roles of a_1 and a_2. The general solution to this equation is:

$$a_2(t) = \left(Ae^{i\Omega t} + Be^{-i\Omega t}\right) e^{i\omega_{21}t/2} \text{ where } \Omega = \frac{1}{2}\left(\omega_{21}^2 + 4|V|^2\right)^{1/2} \tag{6.51}$$

where the constants A and B are determined by the initial conditions. One can obtain a similar expression for a_2. In the initial state $a_1(0) = 1$ and $a_2(0) = 0$, which allows us to determine A and B (i.e. substitute these into the two equations for a_1 and another for a_2 and solve for A and B). Then we obtain the following expressions:

$$a_1(t) = \left(\cos(\Omega t) + \frac{i\omega_{21}}{2\Omega}\sin(\Omega t)\right) e^{-i\omega_{21}t/2} \text{ and } a_2(t) = -\frac{i|V|}{\Omega}\sin(\Omega t)e^{0\omega_{21}t/2} \tag{6.52}$$

This is an exact solution for the two level problem (i.e. no approximations).

6.13 The Rabi formula

We are interested in finding the system in one of the two states as a function of time. The probabilities for the two states are given by $P_1(t) = |a_1(t)|^2$ and $P_2(t) = |a_2(t)|^2$. Initially we have $P_2(0) = 0$. We can now write down the *Rabi formula* by using Eq. (6.52):

$$P_2(t) = \left(\frac{4|V|^2}{\omega_{21}^2 + 4|V|^2} \right) \sin^2 \left(\frac{1}{2} \left(\omega_{21}^2 + 4|V|^2 \right)^{1/2} t \right) \tag{6.53}$$

where the probability for the system to be in state 1 is given by $P_2(t) = 1 - P_1(t)$.

Case 1: Degenerate states. In this case $\omega_{21} = 0$ and Eq. (6.53) becomes:

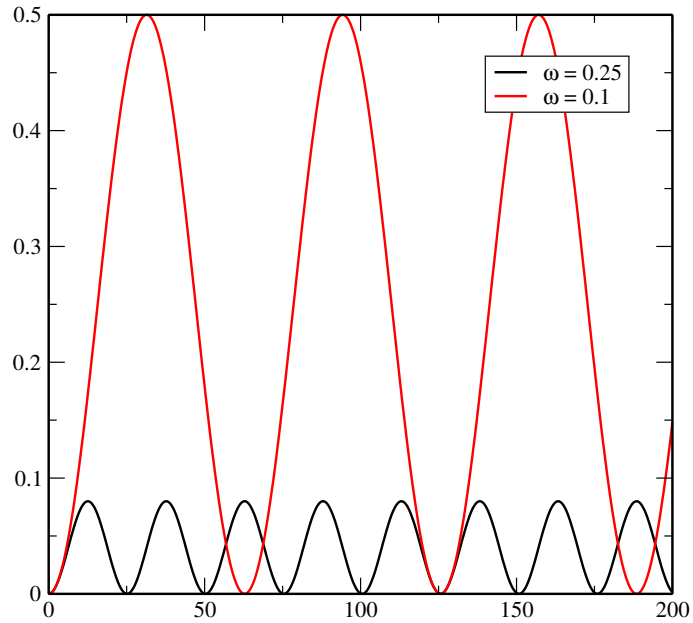
$$P_2(t) = \sin^2 (|V| t) \tag{6.54}$$

This shows that the population oscillates (*Rabi oscillations*) between the two states. Because the oscillation period depends on $|V|$, we see that a strong perturbation drives the system more rapidly than a weak perturbation. Regardless of the strength of the perturbation, it is possible to achieve condition where the system is purely in state 1 or state 2.

Case 2: $\omega_{21}^2 >> 4|V|^2$, which means that the energy levels are widely separated compared to the strength of the perturbation. In this case we can approximately remove $4|V|^2$ from the denominator and the argument of sin in Eq. (6.53).

$$P_2(t) \approx \left(\frac{2\,|V|}{\omega_{21}}\right)^2 \sin^2\left(\frac{1}{2}\omega_{21}t\right) \tag{6.55}$$

$P_2(t)$ with two different values of ω_{21} are plotted below.



Note that this can be used to prepare population inversion even in a two-level system! This is, for example, used in NMR spectroscopy.

6.14 Many level systems

The previous two-level system shows that even very simple systems lead to very complicated differential equations. For n-level system one would have to solve an n-order differential equation, which is very difficult. Furthermore even for a two-level system the resulting equations can be solved only for relatively simple perturbations. In the following we discuss Dirac's alternative approach, which is called the *variation of constants*.

As before the hamiltonian is denoted by $H = H^{(0)} + H^{(1)}(t)$ and the eigenfunctions of $H^{(0)}$ by $|n\rangle$ or $\psi_n^{(0)}$:

$$\Psi_n^{(0)}(t) = \psi_n^{(0)} e^{-iE_n^{(0)}t/\hbar} \quad H^{(0)}\Psi_n^{(0)} = i\hbar\frac{\partial \Psi_n^{(0)}}{\partial t}$$

The perturbed time-dependent wavefunction Ψ is expressed as a linear combination of the time-dependent unperturbed states:

$$\Psi(t) = \sum_n a_n(t)\Psi_n^{(0)}(t) = \sum_n a_n(t)\psi_n^{(0)} e^{-iE_n^{(0)}t/\hbar} \text{ and } H\Psi = i\hbar\frac{\partial \Psi}{\partial t} \qquad (6.56)$$

Our task is to set up differential equations for the coefficients $a_n(t)$. To do this we substitute Ψ into the time-dependent Schrödinger equation:

$$\overbrace{H\Psi}^{**} = \sum_n a_n(t) \overbrace{H^{(0)}\Psi_n^{(0)}(t)}^{*} + \sum_n a_n(t) H^{(1)}(t)\Psi_n^{(0)}(t)$$

$$\underbrace{i\hbar\frac{\partial\Psi}{\partial t}}_{**} = \sum_n a_n(t) \underbrace{i\hbar\frac{\partial\Psi_n^{(0)}}{\partial t}}_{*} + i\hbar\sum_n \dot{a}_n(t)\Psi_n^{(0)}(t)$$

The terms indicated by * are equal, likewise terms marked with ** are equal. So we are left with:

$$\sum_n a_n(t) H^{(1)}(t)\Psi_n^{(0)}(t) = i\hbar\sum_n \dot{a}_n(t)\Psi_n^{(0)}(t)$$

Substituting in the explicit expression for $\Psi_n^{(0)}$ gives:

$$\sum_n a_n(t) H^{(1)}(t)\,|n\rangle\, e^{-iE_n^{(0)}t/\hbar} = i\hbar\sum_n \dot{a}_n(t)\,|n\rangle\, e^{-iE_n^{(0)}t/\hbar}$$

To extract one of the $\dot{a}_k(t)$ on the right hand side, we multiply by $\langle k|$ from the left:

$$\sum_n a_n(t) \left\langle k \left| H^{(1)} \right| n \right\rangle e^{-iE_n^{(0)}t/\hbar} = i\hbar\dot{a}_k(t) e^{-iE_k^{(0)}t/\hbar}$$

To simplify the expression, we use notation $H^{(1)}_{kn}(t) = \left\langle k \left| H^{(1)}(t) \right| n \right\rangle$ and $\hbar\omega_{kn} = E^{(0)}_k - E^{(0)}_n$:

$$\dot{a}_k(t) = \frac{1}{i\hbar} \sum_n a_n(t) H^{(1)}_{kn}(t) e^{-\omega_{kn} t} \tag{6.57}$$

The above equation is still exact (no approximations have been made yet). From this point on, we will proceed in an approximate manner. First we integrate Eq. (6.57):

$$a_k(t) - a_k(0) = \frac{1}{i\hbar} \sum_n \int_0^t a_n(t') H^{(1)}_{kn}(t') e^{i\omega_{kn} t'} \, dt' \tag{6.58}$$

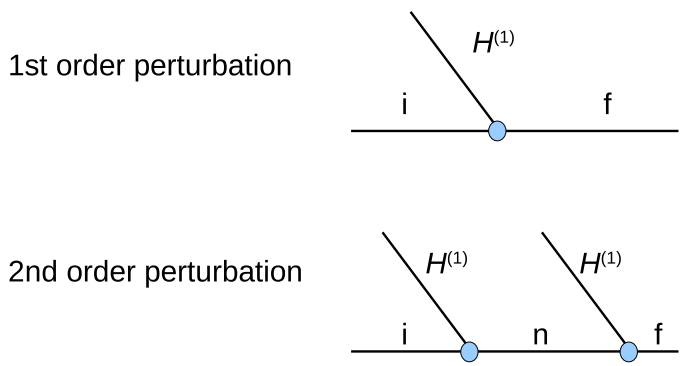
The problem with this expression is that the expression for $a_k(t)$ depends on all other coefficients $a_n(t)$ inclusing $a_k(t)$ itself. The other coefficients $a_n(t)$ are unknown and are determined by similar expression to Eq. (6.58). Therefore to solve Eq. (6.58), we need to know all the coefficients before hand! In the following we make an approximation that the perturbation is weak and applied for a short period of time, which implies that the coefficients should approximately remain close to their initial values. If the system is initially in state $|i\rangle$ then all coefficients other than $a_i(t)$ remain close to zero throughout the period. A state $|f\rangle$, which is initially zero, will now be given by:

$$a_f(t) \approx \frac{1}{i\hbar} \int_0^t a_i(t') H_{fi}^{(1)}(t') e^{i\omega_{fi}t'} dt'$$

This expression is derived from Eq. (6.58) with all coefficients equal to zero except $a_i(t)$, which remains close to one. The latter fact gives:

$$a_f(t) \approx \frac{1}{i\hbar} \int_0^t H_{fi}^{(1)}(t') e^{i\omega_{fi}t'} dt' \tag{6.59}$$

This expression gives the population of an initially unoccupied state $|f\rangle$ as a function of time. This ignores the possibility that the state $|f\rangle$ could be populated via an indirect route (i.e. through some other state). Another way to put this is to say that the perturbation acts only once. This is called the *first order time-dependent perturbation theory*. The interaction of the perturbation can be expressed by using the *Feynman diagrams* as shown below.



Feynman diagrams for 1st and 2nd order perturbation

6.15 The effect of a slowly switched constant perturbation

Here we apply Eq. (6.59) to a perturbation that rises slowly from zero to a steady final value:

$$H^{(1)}(t) = \begin{cases} 0 & \text{for } t < 0 \\ H^{(1)}(1 - e^{-kt}) & \text{for } t \geq 0 \end{cases} \tag{6.60}$$

where $H^{(1)}$ is a time-indepndent operator and, for slow switching, k is small and positive constant. The coefficient of an initially occupied state is given by Eq. (6.59):

$$a_f(t) = \frac{1}{i\hbar} H^{(1)}_{fi} \int_0^t \left(1 - e^{-kt'}\right) e^{i\omega_{fi}t'} \, dt' \tag{6.61}$$

$$= \frac{1}{i\hbar} H^{(1)}_{fi} \left(\frac{e^{i\omega_{fi}t} - 1}{i\omega_{fi}} + \frac{e^{-(k - i\omega_{fi})t} - 1}{k - i\omega_{fi}} \right)$$

At long times ($t >> 1/k$ and slow switching $k^2 << \omega_{fi}^2$) we can write this approximately as (compare to Eq. (6.20) – time-independent formula):

$$\left|a_f(t)\right|^2 = \frac{\left|H^{(1)}_{fi}\right|^2}{\hbar^2 \omega_{fi}^2} = \frac{\left|H^{(1)}_{fi}\right|^2}{\left(E_f^{(0)} - E_i^{(0)}\right)^2} \tag{6.62}$$

6.16 The effect of an oscillating perturbation

Next we consider an oscillating perturbation that could correspond, for example, to the oscillating electric field from light (e.g., electromagnetic radiation from a light source in a spectrometer or sun light). Once we know how to deal with oscillating perturbations, we will be able to work with all types of perturbations because they can always be expressed as a superposition of the oscillating waves (i.e. Fourier expansion).

First we consider transitions between states $|i\rangle$ and $|f\rangle$. A perturbation oscillating with an angular frequency $\omega = 2\pi\nu$, which is turned on at $t = 0$, is:

$$H^{(1)}(t) = 2H^{(1)}\cos(\omega t) = H^{(1)}\left(e^{i\omega t} + e^{-i\omega t}\right) \text{ for } t \geq 0 \tag{6.63}$$

When this is inserted into Eq. (6.59) (the 1st order expression for $a_f(t)$), we get:

$$a_f(t) = \frac{1}{\hbar i} H_{fi}^{(1)} \int_0^t \left(e^{i\omega t'} + e^{-i\omega t'}\right) e^{i\omega_{fi} t'}\, dt' \tag{6.64}$$

$$= \frac{1}{i\hbar} H_{fi}^{(1)} \left(\frac{e^{i(\omega_{fi}+\omega)t} - 1}{i(\omega_{fi} + \omega)} + \frac{e^{i(\omega_{fi}-\omega)t} - 1}{i(\omega_{fi} - \omega)} \right)$$

To simplify the above expression, we note that in electronic spectroscopy the frequencies ω and ω_{fi} are of the order of 10^{15} s^{-1} and in NMR the lowest frequencies are still higher than 10^6 s^{-1}. The both numerators are close to one (the norm of $\exp(-i...)$ is one), the denominator in the first term is of the order of the frequencies, so we can estimate an upper limit for it as 10^{-6} (and 10^{-15} in electronic spectroscopy), and the denominator in the second term may become close to zero. This implies that the 2nd term dominates and the first term can be approximately ignored. This suggests that it is sufficient to write the original hamiltonian of Eq. (6.63) as:

$$H^{(1)}(t) = 2H^{(1)}\cos(\omega t) = H^{(1)}e^{-i\omega t}$$

This is called the *rotating wave approximation* (RWA) and it is commonly applied in spectroscopy. By just retaining the 2nd term in Eq. (6.64) we can write the population of the final state as ($P_f(t) = \left|a_f(t)\right|^2$):

$$P_f(t) = \frac{4\left|H_{fi}^{(1)}\right|^2}{\hbar^2\left(\omega_{fi}-\omega\right)^2}\sin^2\left(\frac{1}{2}\left(\omega_{fi}-\omega\right)t\right) \tag{6.65}$$

Taking the amplitude of the perturbation to be a constant, $\left|H_{12}^{(1)}\right|^2 = \hbar^2\left|V_{fi}\right|^2$ (contains transition matrix element), gives:

$$P_f(t) = \frac{4\left|V_{fi}\right|^2}{\left(\omega_{fi} - \omega\right)^2} \sin^2\left(\frac{1}{2}\left(\omega_{fi} - \omega\right)t\right) \tag{6.66}$$

This expression is essentially the same that was obtained earlier for a two-level system with constant perturbation (cf. Eq. (6.55)). The main difference is that the energy level difference ω_{21} is now replaced by the difference between the frequncy of the oscillating perturbation and the energy level separation $(\omega_{fi} - \omega)$. This can be interpreted as an effective shift in energy caused by the oscillating perturbation.

According to Eq. (6.66), the time-dependence of the probability of being in state $|f\rangle$ depends on the frequency shift $\omega_{fi} - \omega$. When the offset is zero, the field and the system are said to be in *resonance* and the population of state $|f\rangle$ increases rapidly in time. To obtain a better understanding how this happens, consider Eq. (6.66) when $\omega \to \omega_{fi}$. First we note that sin is linear when x is small:

$$\sin(x) \approx x \text{ when } x \approx 0$$

And then:

$$\lim_{\omega \to \omega_{fi}} P_f(t) = \left|V_{fi}\right|^2 t^2 \tag{6.67}$$

This indicates that the final state population increases quadratically with time when on resonance. Note that we must remember that we have implied the first order approximation ($\left|V_{fi}\right|^2 t^2 << 1$; as the population transfer must be small).

6.17 Transition rates to continuum states

If a continuum of states exists in the neighborhood of the final state, we can still use Eq. (6.66) to calculate the transition probability to one member of these states. To account for all of the states, one has to integrate over the continuum states. Before this can be done, one needs to specify the *density of states* $\rho(E)$, which tells us the number of states available at energy E. In differential form $\rho(E)dE$ gives the number of states between E and $E+dE$. The total transition probability can then be obtained by integrating over the continuum:

$$P(t) = \int P_f(t)\rho(E)dE \tag{6.68}$$

where the integration is carried over the energy range corresponding to the continuum set of states in question. To evaluate this integral, we first express the transition frequency ω_{fi} in terms of the energy E by writing $\omega_{fi} = E/\hbar$:

$$P(t) = \int 4\left|V_{fi}\right|^2 \frac{\sin^2\left(\frac{1}{2}\left(E/\hbar-\omega\right)t\right)}{\left(E/\hbar-\omega\right)^2}\rho(E)dE$$

Since the incident frequency is going to excite a fairly narrow band in the continuum, we take the density of states to be independent of E. We also take the matrix element $\left|V_{fi}\right|$ to be constant (even though they vary with E). Then we can approximately write:

$$P(t) = \left|V_{fi}\right|^2 \rho(E_{fi}) \int \frac{\sin^2\left(\frac{1}{2}\left(E/\hbar - \omega\right)t\right)}{\left(E/\hbar - \omega\right)^2} dE$$

The integrand peaks at fairly narrow region in terms of E and hence we can safely extend the limits of integration. To simplify the integral we also change the integration variable $x = \frac{1}{2}\left(E/\hbar - \omega\right)t \rightarrow dE = (2\hbar/t)dx$. The integral reads now:

$$P(t) = \frac{2\hbar}{t} \left|V_{fi}\right|^2 \rho(E_{fi}) t^2 \int_{-\infty}^{\infty} \frac{\sin^2(x)}{x^2} dx$$

From a mathematics table book we can find the following integral:

$$\int_{-\infty}^{\infty} \frac{\sin^2(x)}{x^2} dx = \pi$$

This allows us the evaluate the integral:

$$P(t) = 2\pi\hbar t \left|V_f i\right|^2 \rho(E_{fi}) \tag{6.69}$$

Note that the power of t appears as one rather than t^2 (Eq. (6.67)). This is because linewidth of the peak decreases as $1/t$ (i.e. the line gets narrower) and the overall area is then $t^2 \times 1/t = t$.

The *transition rate* W is the rate of change of probability of being in an initially empty state:

$$W = \frac{dP}{dt} \tag{6.70}$$

The intensities of spectral lines are proportional to this because they depend on the rate of energy transfer between the system and the electromagnetic field. Therefore we can combine Eqs. (6.69) and (6.70):

$$W = 2\pi\hbar \left|V_{fi}\right|^2 \rho(E_{fi}) \tag{6.71}$$

This result is called *Fermi's golden rule*, which can be used to calculate transition rates that are based on the density of states $\rho(E_{fi})$ and the transition matrix element V_{fi}.

6.18 The Einstein transition probabilities

Fermi's golden rule is able to account for absorption of photons by matter but it cannot explain why emission from an excited state to the ground state would occur. Transitions should only occur if the system is subject to electromagnetic radiation. From pratical experience we know that once a molecule is promoted to an excited state, it will fluoresce even when no electromagnetic radiation is present. Clearly we are missing something in our description of this phenomenom.

Einstein considered a collection of atoms that were in thermal equilibrium with the electromagnetic field at a temperature T. If the perturbation due to the incident electromagnetic radiation takes the form $V_{fi} = -\vec{\mu} \cdot \vec{\varepsilon}$ where $\vec{\mu} = \sum q_i r_i$ is the electric dipole moment and $\vec{\varepsilon}$ is the electric field. Thus $\left|V_{fi}\right|^2 \propto |\vec{\varepsilon}|^2$. Note that the oscillatory term (i.e. $\cos(\omega t)$) is present in $H^{(1)}(t)$ but was not written as part of V_{fi} (or $H^{(1)}_{fi}$ in Eq. (6.42)). Intensity of light is also proportional to $|\vec{\varepsilon}|^2$ and hence $\left|V_{fi}\right|^2 \propto I$.

The relationship between intensity I (J/m^2) and the energy E (J) of the electromagnetic field is given by:

$$dE = I(\nu) A \Delta t d\nu \tag{6.72}$$

where $I(\nu)$ is the intensity of the electromagnetic field at frequency ν, A is the incident area, Δt is the passage time through the sample so that Δtc gives the length of the sample (c is the speed of light). The volume containing the electromagnetic radiation is then given by ΔtcA. This allows us to calculate the energy density:

$$\rho_{rad}(\nu)d\nu = \frac{dE}{Ac\Delta t} = \frac{I(c)}{c}d\nu \Rightarrow \rho_{rad}(\nu) = \frac{I(\nu)}{c} \tag{6.73}$$
$$\text{Since } \left|V_{fi}\right|^2 \propto I(\nu) \Rightarrow \left|V_{fi}\right|^2 \propto \rho_{rad}(\nu)$$

We also have $E_{fi} = h\nu_{fi}$ and we can use Fermi's golden rule (Eq. (6.71)) to write the transition rate from $|i\rangle$ to $|f\rangle$ (note the difference: ρ vs. ρ_{rad} and $H_{fi} = \hbar V_{fi}$):

$$W_{f\leftarrow i} = 2\pi\hbar \underbrace{\left|V_{fi}\right|^2}_{\propto\rho_{rad}} \rho(E_{fi}) \equiv B_{fi}\rho_{rad}(E_{fi}) \tag{6.74}$$

where the proportionality constant B_{fi} is called the *Einstein coefficient of stimulated absorption*. It is also possible to have the reverse process (i.e. light induced emission from $|f\rangle \rightarrow |i\rangle$; *stimulated emission*):

$$W_{i\leftarrow f} = B_{if}\rho_{rad}(E_{if}) = B_{if}\rho_{rad}(E_{fi}) \tag{6.75}$$

Since V_{if} is hermitian, $\left|V_{if}\right|^2 = V_{if}V_{if}^* = V_{fi}^*V_{fi} = \left|V_{fi}\right|^2$. This implies that also $B_{if} = B_{fi} \equiv B$.

Next we will show the connection between B_{fi} and the *transition dipole moment* $\vec{\mu}_{fi}$:

$$\vec{\mu}_{fi} = \int \psi_f^*\vec{\mu}\psi_i d\tau \tag{6.76}$$

Transition dipole moment and the Einstein coefficient B:

A molecule is exposed to light with its electric field vector lying along the z-axis ($\varepsilon(t) = 2\varepsilon\cos(\omega t)$) and oscillating at a frequency $\omega = 2\pi\nu$. The perturbation is then:

$$H^{(1)}(t) = 2H^{(1)}_{fi}\cos(\omega t) = -2\mu_z\varepsilon\cos(\omega t)$$

where μ_z is the z-component of the transition diple moment. Since $H^{(1)}_{fi} = \hbar V_{fi}$ we have $V_{fi} = -\mu_z\varepsilon/\hbar$. The transition rate from an initial state $|i\rangle$ to a continuum of final states $|f\rangle$ due to a perturbation of this form is given by Eq. (6.71):

$$W_{i\to f} = 2\pi\hbar\left|V_{fi}\right|^2\rho(E_{fi}) = \frac{2\pi}{\hbar}\mu^2_{fi,z}\varepsilon^2\rho(E_{fi})$$

where $\rho(E_{fi})$ is the density of the continuum states at an energy $E_{fi} = \hbar\omega_{fi}$ and ω_{fi} is the transition angular frequency. In liquid or gaseous phases molecules are able to rotate freely, which randomizes the direction. Then we would replace $\mu^2_{fi,z}$ by the average over all directions $\frac{1}{3}\left|\mu_{fi}\right|^2$.

The energy of a classical electromagnetic field is:

$$E = \frac{1}{2}\int\left(\epsilon_0\left\langle\varepsilon^2\right\rangle + \mu_0\left\langle \mathrm{h}^2\right\rangle\right)d\tau$$

wherer $\langle \varepsilon^2 \rangle$ and $\langle \mathrm{h}^2 \rangle$ are the time-averages of the squared field strengths (ε = electric, h = magnetic). In the present case the period is $2\pi/\omega$ and we average over that:

$$\langle \varepsilon^2 \rangle = \frac{4\varepsilon^2}{2\pi/\omega} \int_0^{2\pi/\omega} \cos^2(\omega t) dt = 2\varepsilon^2$$

From electromagnetic theory (Maxwell's equations): $\mu_0 \langle \mathrm{h}^2 \rangle = \epsilon_0 \langle \varepsilon^2 \rangle$ where μ_0 is the vacuum permeability and ϵ_0 is the permittivity. Therefore, for a field in a region of volume V, we have:

$$E = 2\epsilon_0 \varepsilon^2 V \text{ or } \varepsilon^2 = \frac{E}{2\epsilon_0 V}$$

To obtain the transition rate for non-monochromatic light, we need to consider a distribution of frequencies. This can done by definining the density of radiation states $\rho'_{rad}(E)$, where $\rho'_{rad}(E)dE$ gives the number of waves with photon energies in the range E to $E + dE$. Integration over energies now gives:

$$W_{i\to f} = \frac{1}{6\epsilon_0\hbar^2}\left|\mu_{fi}\right|^2\left(\frac{E_{fi}\rho'_{rad}(E_{fi})}{V}\right)$$

This integration is similar to what was used to get Eq. (6.71) but now instead of integrating over density of final states, we integrate over denisty of non-monochromatic electromagnetic radiation. The term $E_{fi}\rho'_{rad}(E_{fi})$ gives the energy density of radiation states (of monochromatic light) and therefore we can write:

$$W_{i\to f} = \frac{1}{6\epsilon_0\hbar^2}\left|\mu_{fi}\right|^2\rho_{rad}(E_{fi})$$

By comparing this with Eq. (6.75) we can identify the Einstein coefficient B:

$$B_{if} = \frac{\left|\mu_{fi}\right|^2}{6\epsilon_0\hbar^2} \tag{6.77}$$

The transition probabilities above refer to individual atoms. For a collection of atoms the rates must be multiplied by the number of atoms in each state. For a system in thermal equilibrium when there is no net energy transfer between the system and the field:

$$N_i W_{f\leftarrow i} = N_f W_{f\to i}$$

Since the two transition rates are equal, we conclude that the populations of the two levels are equal. This would be in apparent conflict with the Boltzmann distribution:

$$\frac{N_f}{N_i} = e^{-E_{fi}/(kT)}$$

Therefore there must be an additional process that contributes to emission:

$$W_{f\to i}^{spont} = A_{fi} \tag{6.78}$$

where A_{fi} is called the *Einstein coefficient of spontaneous emission.* The total rate of emission is then:

$$W_{f\to i} = A_{fi} + B_{fi}\rho_{rad}(E_{fi}) \tag{6.79}$$

If we accept the Boltzmann distribution then the above expression can be rearranged as:

$$\rho_{rad}(E_{fi}) = \frac{A_{fi}/B_{fi}}{(B_{if}/B_{fi})e^{E_{fi}/(kT)} - 1}$$

On the other hand Planck distribution gives:

$$\rho_{rad}(E_{fi}) = \frac{8\pi h\nu_{fi}^3/c^3}{e^{E_{fi}/(kT)} - 1} \tag{6.80}$$

Comparison of the above two equations yields that $B_{fi} = B_{if}$ as well as the relationship between B and A coefficients:

$$A_{fi} = \frac{8\pi h\nu_{fi}^3}{c^3} B_{fi} \tag{6.81}$$

It is important to note that the spontaneous emission increases as the third power of the transition frequency. Thus this process becomes more and more important at high frequencies. This predicts that it is very difficult to make lasers that operate at X-ray frequencies (and this is indeed the case!) because it is very difficult to maintain excited state populations. The spontaneous emission can be viewed to originate from the zero-point fluctuations present in the electromagnetic field (even when photons are not present). These cause random perturbations that act on atoms/molecules and cause them to emit.

6.19 Lifetime and energy uncertainty

If a system is stationary (i.e., $|\Psi|^2$ is time independent), it is described by one of its eigenfunctions with the phase part dictated by the state energy (i.e., $\Psi_i = \psi e^{-iE_i t\hbar}$). If the wavefunction is not stationary, it means that the system is evolving towards another states and then its energy is imprecise. Another way to think about this is to say that energy and time are Fourier pairs and subject to uncertainty. However, this is not a well defined statement since time is not an observable in non-relativistic quantum mechanics. Supppose that the probability of finding the system in a particular excited state i decays exponentially with time-constant τ:

$$|\Psi_i|^2 = |\psi_i|^2 e^{-t/\tau} \tag{6.82}$$

The exponential decay form is approriate for many types of processes (including spontaneous emission). The amplitude of the state i has then the following form:

$$\Psi_i = \psi_i e^{-iE_i t/\hbar - t/(2\tau)} \tag{6.83}$$

The amplitude of this function decays in an oscillatory manner. What is the energy of this state? We can see this by analyzing the frequency components by using Fourier transformation:

$$e^{-iE_i t/\hbar - t/(2\tau)} = \int_{-\infty}^{\infty} g(E') e^{-iE' t/\hbar} dE'$$

Fourier transforming this gives the following form for $g(E')$:

$$G(E') = \frac{\hbar/(2\pi\tau)}{(E - E')^2 + (\hbar/(2\tau))^2} \tag{6.84}$$

This expression is a superposition of many energies and therefore we have *lifetime broadening* of the state. In spectroscopy, the lineshape that follows from this expression is called *Lorentzian lineshape*. It is the most common lineshape in gas and liquid phase experiments where inhomogeneities of the sample are negligible. To get a qualitative result we note that the width at half-height is $\hbar/(2\tau)$ which gives:

$$\tau\Delta E \approx \frac{1}{2}\hbar \tag{6.85}$$

Note that this looks similar to the uncertainty principle between position and momentum. This is because position and momentum are also Fourier pairs and subject to uncertainty. Remember that if you have a narrow signal, its widht will be large in the corresponding Fourier space (and the other way around).

Chapter 7: Atomic structure

7.1 Energy levels of hydrogen atom

The energy level structure of hydrogen atom can be studied experimentally by fluorescence spectroscopy where the position of the emission lines corresponds to the energy differences between the electronic states. Based on Eq. (3.44), the transition energies are (cm^{-1}; see also the undergraduate quantum chemistry lecture notes):

$$\Delta\tilde{\nu}_{n_1,n_2} = R_H \left(\frac{1}{n_1^2} - \frac{1}{n_2^2} \right) \tag{7.1}$$

where $R_H = \frac{m_e e^4}{4\pi c (4\pi\epsilon_0)^2 \hbar^3}$ is the Rydberg constant (e is the electron charge, ϵ_0 is the vacuum permittivity, m_e is the electron mass, c is the speed of light). Note that this is independent of the orbital and spin quantum numbers.

While the experiments show that transitions are observed between all combinations n_1 and n_2, it appears that the expected degeneracy due to orbital and spin degrees of freedom does not contribute to the line intensities to the full extent. For example, transitions between s states ($l = 0$) are not observed. To account for this observation, we have to evaluate the transition moments between the states, which determine the overall line intensities (selection rules).

7.2 Transition probabilities and selection rules

Based on Eqs. (6.76), (6.77), (6.78), and (6.81), the fluorescence intensity is determined by the transition dipole moment between states f and i:

$$\vec{\mu}_{fi} = \langle \psi \, | \vec{\mu} | \, \psi \rangle = \int \psi^* \hat{\mu} \psi d\tau$$

where the electric dipole operatior $\vec{\mu} = -e\vec{r}$ ($\vec{r}$ is the position vector). Large values of $|\vec{\mu}|$ result in intense transitions between the states. Selection rules state whether this integral is zero (forbidden transition) or non-zero (allowed transition). For atoms these rules can be written in terms of changes in the orbital angular momentum quantum numbers (l and m).

Properties of spherical harmonics:

$$\int_0^{\pi} \int_0^{2\pi} Y_{l'}^{m'*}(\theta,\phi) Y_l^m(\theta,\phi) \sin(\theta) d\theta d\phi = \delta_{l,l'} \delta_{m,m'} \tag{7.2}$$

$$\int_0^{\pi} \int_0^{2\pi} Y_{l''}^{m''*}(\theta,\phi) Y_{l'}^{m'}(\theta,\phi) Y_l^m(\theta,\phi) \sin(\theta) d\theta d\phi = 0 \tag{7.3}$$

unless $m'' = m + m'$ and a triangle with sides l, l', l'' can be formed:
$l''^2 = l^2 + l'^2 - 2ll' \cos(\alpha)$ with some value of α ($|\cos(\alpha)| \leq 1$)

$$Y_l^{m*} = (-1)^m Y_l^{-m} \text{ (Condon-Shortley)} \tag{7.4}$$

In terms of symmetry, the vector components of $\vec{\mu}$ behave the same way as p-orbitals (in spherical harmonic basis). Therefore $l' = 1$ and $m' = +1, 0, -1$ in Eq. (7.3).

Selection rule $\Delta l = 0, \pm 1$: The triangle condition in Eq. (7.3) gives:

$$l''^2 = l^2 + l'^2 + 2ll'\alpha'$$

where all quantum numbers must be positive whole integers, $l' = 1$, and $\alpha' = \cos(\alpha)$ is a free parameter restricted to values between -1 and $+1$. This becomes then:

$$l''^2 = l^2 + 1 + 2l\alpha'$$

l	l''	l''^2	$l^2 + 1 + 2l\alpha'$	α'
0	0	0	1	No solution
0	1	1	1	Any value of α' OK
0	2	4	1	No solution
		ETC.		
1	1	1	$1 + 1 + 2\alpha'$	No solution
1	2	4	$1 + 1 + 2\alpha'$	$\alpha' = 1$
1	3	9	$1 + 1 + 2\alpha'$	No solution
		ETC.		
2	3	9	$4 + 1 + 4\alpha'$	$\alpha' = 1$

The equation is therefore satisfied when $\Delta l = l'' - l = \pm 1$ (e.g., $s - p$ transition).

Selection rule $\Delta m = 0, \pm 1$: The second condition in Eq. (7.3) with $m' = -1, 0, +1$ leads to $m'' = m + 1$, $m'' = m - 1$, or $m'' = m$. For short this is $\Delta m = m'' - m = 0, \pm 1$.

These results can also be derived using group theory as follows. Although atoms have infinitely high symmetry, for the following example we can restrict to D_{2h}. In this point group: s orbitals belong to A_g, p_x to B_{3u}, p_y to B_{2u}, and p_z to B_{1u}. The transition dipole operators share these same irreps.

$\langle s|\hat{\mu}|p_x\rangle$:
$x : A_g \times B_{3u} \times B_{3u} = A_g \times A_g = A_g$ (possibly non-zero; allowed transition!)
$y : A_g \times B_{2u} \times B_{3u} = B_{2u} \times B_{3u} = B_{1g}$ (zero)
$z : A_g \times B_{1u} \times B_{3u} = B_{1u} \times B_{3u} = B_{2g}$ (zero)

$\langle s|\hat{\mu}|p_y\rangle$:
$x : A_g \times B_{3u} \times B_{2u} = B_{3u} \times B_{2u} = B_{1g}$ (zero)
$y : A_g \times B_{2u} \times B_{2u} = A_g \times A_g = A_g$ (possibly non-zero; allowed transition!)
$z : A_g \times B_{1u} \times B_{2u} = B_{1u} \times B_{2u} = B_{3g}$ (zero)

$\langle s|\hat{\mu}|p_z\rangle$:
$x : A_g \times B_{3u} \times B_{1u} = B_{3u} \times B_{1u} = B_{2g}$ (zero)
$y : A_g \times B_{2u} \times B_{1u} = B_{2u} \times B_{1u} = B_{3g}$ (zero)
$z : A_g \times B_{1u} \times B_{1u} = A_g \times A_g = A_g$ (possibly non-zero; allowed transition!)

Therefore $s - p$ transitions are electric dipole allowed (optical transition).

However, $s - s$ transitions are not allowed because (e.g., $1s - 2s$):

$x : A_g \times B_{3u} \times A_g = B_{3u}$ (zero)
$y : A_g \times B_{2u} \times A_g = B_{2u}$ (zero)
$z : A_g \times B_{1u} \times A_g = B_{1u}$ (zero)

Remember that electron spin does not change in optical transitions (optical spectroscopy). In magnetic resonance spectroscopy, transitions between either electron or nuclear spin states are observed.

Review the atomic and molecular term symbol sections from the undergraduate quantum notes.

7.3 Schrödinger equation for ground state helium

The Hamiltonian for helium atom (2 electrons and nuclear charge +2) is:

$$\hat{H} = -\frac{\hbar^2}{2m_e}\left(\nabla_1^2 + \nabla_2^2\right) - \frac{2e^2}{4\pi\epsilon_0 r_1} - \frac{2e^2}{4\pi\epsilon_0 r_2} + \frac{e^2}{4\pi\epsilon_0 r_{12}} \tag{7.5}$$

where r_1 and r_2 are the distances from the nucleus to electrons 1 and 2, respectively, and r_{12} is the distance between the two electrons. The Schrödinger equation is then:

$$\hat{H}\psi(r_1, r_2) = E\psi(r_1, r_2) \tag{7.6}$$

where ψ is an eigenfunction and E is the corresponding eigenvalue. Note that there are many possible (ψ, E) pairs that can satisfy this equation. The overall dimensionality of this problem is 6 (i.e., the x, y, z coordinates for both electrons).

Egil Hylleraas (Norwegian theoretical physicist; 1898 – 1965) spent almost his whole life in trying to solve this problem analytically. He got very close but was not able to find the exact solution. Up to this day, no analytical solutions to this equation have been found. In practice, solutions to systems with two or more electrons must be obtained numerically.

Simple perturbation theory approach for ground state helium

We partition the Hamiltonian into two hydrogenic terms ($\hat{H}^{(0)}$) plus perturbation ($\hat{H}^{(1)}$):

$$\begin{aligned} \hat{H} &= \hat{H}^{(0)} + \hat{H}^{(1)} \\ \hat{H}^{(0)} &= \hat{H}_1 + \hat{H}_2 \text{ with } \hat{H}_i = -\frac{\hbar^2}{2m_e}\nabla_i^2 - \frac{2e^2}{4\pi\epsilon_0 r_i} \text{ and } i = 1, 2 \\ \hat{H}^{(1)} &= \frac{e^2}{4\pi\epsilon_0 r_{12}} \text{ (perturbation)} \end{aligned} \tag{7.7}$$

Next we recall from the undergraduate quantum chemistry course that when $\hat{H}$ can be expressed as a sum of independent terms, the solution will be a product of the eigenfunctions of these terms:

$$\hat{H} = \hat{H}_1 + \hat{H}_2 \text{ (independent)} \Rightarrow \psi(r_1, r_2) = \phi_1(r_1) \times \phi_2(r_2)$$

where ϕ_i are eigenfunctions of $\hat{H}_i$.

For helium atom, the eigenfunctions of $\hat{H}^{(0)}$ can be then written as:

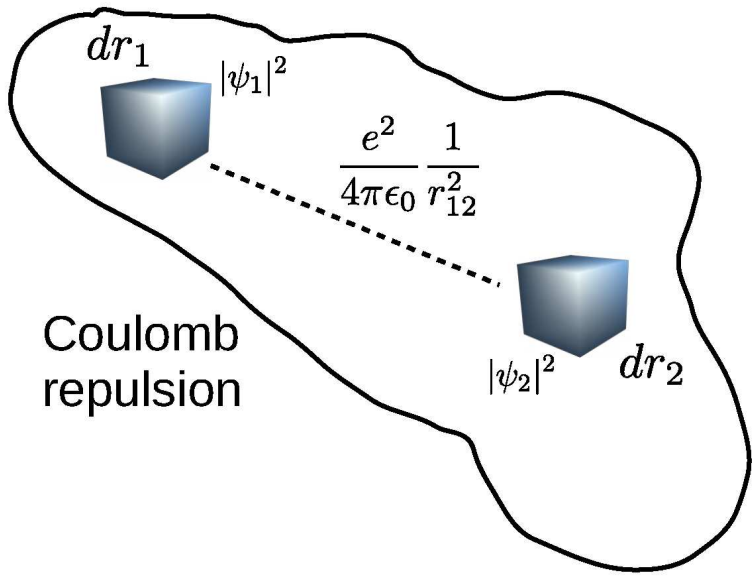
$$\psi(r_1, r_2) = \psi_{n_1,l_1,m_{l,1}}(r_1) \times \psi_{n_2,l_2,m_{l,2}}(r_2) \tag{7.8}$$

where the ψ's are hydrogenic orbitals with the specified quantum numbers (Eq. (3.38)).

The first order correction to energy (see Eq. (6.19)) is:

$$
\begin{aligned}
E^{(1)} &= \left\langle \psi(r_1, r_2) | \hat{H}^{(1)} | \psi(r_1, r_2) \right\rangle \qquad (7.9)\\
&= \left\langle \psi_{n_1,l_1,m_{l,1}}(r_1)\psi_{n_2,l_2,m_{l,2}}(r_2) | \hat{H}^{(1)} | \psi_{n_1,l_1,m_{l,1}}(r_1)\psi_{n_2,l_2,m_{l,2}}(r_2) \right\rangle \\
&= \frac{e^2}{4\pi\epsilon_0} \int \left|\psi_{n_1,l_1,m_{l,1}}(r_1)\right|^2 \times \frac{1}{r_{12}} \times \left|\psi_{n_2,l_2,m_{l,2}}(r_2)\right|^2 dr_1 dr_2
\end{aligned}
$$

This integral is caled the *Coulomb integral* and is often denoted by symbol J.



To get numerical values for the energy of ground state He atom, explicit forms for the wavefunctions ψ would be needed.

7.4 Excited states of helium atom

In ground state helium atom ($1s2$ configuration) there was only one possible configuration. However, excited states (e.g., $1s^1 2s^2$) have two possible configurations that are degenerate. Denote the two orbitals by a and b and include the electron index by a number (1 or 2) and then the two configurations can be written as $a(1)b(2)$ (state 1) and $a(2)b(1)$ (state 2). Since the two configurations are degenerate, we have to use the degenerate version of perturbation theory (see Eq. (6.33) and the example below that).

The required matrix elements (2x2 determinant) are:

$$H_{11} = \left\langle a(1)b(2)| \underbrace{H_1 + H_2}_{H^{(0)}} + \underbrace{\frac{e^2}{4\pi\epsilon_0 r_{12}}}_{H^{(1)}} |a(1)b(2) \right\rangle = E_a + E_b + J \quad (7.10)$$

$$H_{22} = \left\langle a(2)b(1)|H_1 + H_2 + \frac{e^2}{4\pi\epsilon_0 r_{12}}|a(2)b(1) \right\rangle = H_{11}$$

$$H_{12} = H_{21} = \left\langle a(1)b(2)|H_1 + H_2 + \frac{e^2}{4\pi\epsilon_0 r_{12}}|a(2)b(1) \right\rangle$$

$$= \left\langle a(1)b(2)|\frac{e^2}{4\pi\epsilon_0 r_{12}}|a(2)b(1) \right\rangle \equiv K$$

Note that atomic orbitals a and b are orthgonal (e.g., $\langle a(1)|b(1)\rangle = 0$).

The last integral in Eq. (7.10), which is denoted by K, is called the *exchange integral*. The values of both J and K are positive. The determinant corresponding to Eq. (6.33) is then:

$$\begin{vmatrix} H_{11} - ES_{11} & H_{12} - ES_{12} \\ H_{21} - ES_{21} & H_{22} - ES_{22} \end{vmatrix} = \begin{vmatrix} H_{11} - E & H_{12} \\ H_{21} & H_{22} - E \end{vmatrix} \tag{7.11}$$
$$= \begin{vmatrix} E_a + E_b + J - E & K \\ K & E_a + E_b + J - E \end{vmatrix} = 0$$

The overlap matrices are: $S_{11} = S_{22} = 1$ (normalization) and $S_{12} = S_{21} = 0$ (orthogonality). The two solutions to this determnant equation are:

$$E = E_a + E_b + J \pm K \tag{7.12}$$

where

$$J = \frac{e^2}{4\pi\epsilon_0} \left\langle a(1)b(2) \left| \frac{1}{r_{12}} \right| a(1)b(2) \right\rangle \tag{7.13}$$
$$K = \frac{e^2}{4\pi\epsilon_0} \left\langle a(1)b(2) | \frac{1}{r_{12}} | a(2)b(1) \right\rangle$$

Note that K involves the "change of indices" and hence is called the exchange.

The corresponding eigen functions can be solved from the determinant equation by substituting in the two energy values. The two eigen functions are:

$$\psi_+(1,2) = \frac{1}{\sqrt{2}}\left(a(1)b(2) + b(1)a(2)\right) \text{ (symmetric)} \tag{7.14}$$
$$\psi_-(1,2) = \frac{1}{\sqrt{2}}\left(a(1)b(2) - b(1)a(2)\right) \text{ (anti-symmetric)}$$

where symmetric and anti-symmetric refer to the exchange of the electron indices. The anti-symmetric eigen function $(-)$ changes its sign when the indices 1 and 2 are exchanged whereas the symmetric $(+)$ remains unchanged. Remember that electrons are fermions and their wavefunction should be anti-symmetric with respect to swapping the indices.

Two electrons can exist as electronic singlet ($S = 0$; "opposite spins") or triplet state ($S = 1$; "parallel spins"). The possible spin functions are:

$$\sigma_-(1,2) = \frac{1}{\sqrt{2}}\left(\alpha(1)\beta(2) - \beta(1)\alpha(2)\right) \text{ (anti-symmetric)} \tag{7.15}$$
$$\sigma_+^{+1}(1,2) = \alpha(1)\alpha(2) \text{ (symmetric)}$$
$$\sigma_+^{0}(1,2) = \frac{1}{\sqrt{2}}\left(\alpha(1)\beta(2) + \beta(1)\alpha(2)\right) \text{ (symmetric)}$$
$$\sigma_+^{-1}(1,2) = \beta(1)\beta(2) \text{ (symmetric)}$$

In order to get a wavefunction that is anti-symmetric overall, we have to connect symmetric and anti-symmetric functions together. The four possible eigen functions (with fermionic symmetry) are then:

$$\Psi_{0,0}(1,2) = \frac{1}{\sqrt{2}}\left(a(1)b(2) + b(1)a(2)\right) \times \frac{1}{\sqrt{2}}\left(\alpha(1)\beta(2) - \beta(1)\alpha(2)\right) \quad (7.16)$$

$$\Psi_{1,+1}(1,2) = \frac{1}{\sqrt{2}}\left(a(1)b(2) - b(1)a(2)\right) \times \alpha(1)\alpha(2)$$

$$\Psi_{1,0}(1,2) = \frac{1}{\sqrt{2}}\left(a(1)b(2) - b(1)a(2)\right) \times \frac{1}{\sqrt{2}}\left(\alpha(1)\beta(2) + \beta(1)\alpha(2)\right)$$

$$\Psi_{1,-1}(1,2) = \frac{1}{\sqrt{2}}\left(a(1)b(2) - b(1)a(2)\right) \times \beta(1)\beta(2)$$

where the subscripts indicate the value of S and M_S. The first line corresponds to the singlet state and the last three lines to the triplet state. Based on Eq. (7.12), these states are separated in energy by $2K$.

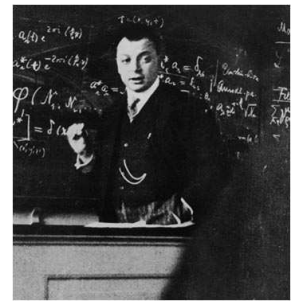
The singlet and triplet states show distinctly different behavior when the two electrons approach each other (i.e., $|r_1 - r_2| \to 0$). The probability density increases for the singlet state (*Fermi heap*) whereas it decreases for the triplet state (*Fermi hole*).

7.5 Pauli exclusion principle

To see where the two spin functions used in the previous section came from, we consider two electrons as an example. The spin functions can be written as products:

$$\alpha(1)\alpha(2),\ \beta(1)\alpha(2),\ \alpha(1)\beta(2) \text{ or } \beta(1)\beta(2) \qquad (7.17)$$

where, for example, $\alpha(1)$ indicates that electron 1 has spin α. Electrons are indistinguishable and therefore it makes no sense to try to use the two middle terms alone but combine them as:



Wolfgang Pauli (1900 - 1958), Austrian physicist, Nobel prize 1945.

$$\alpha(1)\alpha(2) \text{ (symmetric if 1 and 2 are exchanged)} \qquad (7.18)$$

$$\beta(1)\beta(2) \text{ (symmetric)} \qquad (7.19)$$

$$\frac{1}{\sqrt{2}}\left(\alpha(1)\beta(2) + \alpha(2)\beta(1)\right) \text{ (symmetric)} \qquad (7.20)$$

$$\frac{1}{\sqrt{2}}\left(\alpha(1)\beta(2) - \alpha(2)\beta(1)\right) \text{ (antisymmetric)} \qquad (7.21)$$

In 1926 Pauli showed that the wavefunction must be antisymmetric with respect to exchange of the electron indices. This applies also for systems with more than two electrons. An outcome of this principle is that no two electrons can have exactly the same quantum numbers: n, l, m_l, s, m_s. In non-relativistic quantum mechanics this is a postulate.

In order to construct a complete wavefunction for He atom, we need to tag on the spin part to the wavefunction:

$$\psi = 1s(1)1s(2) \times \frac{1}{\sqrt{2}} \left(\alpha(1)\beta(2) - \alpha(2)\beta(1)\right) \tag{7.22}$$



John Slater (1900 - 1976), American physicist

A generalization of this result for many-electron systems was proposed by Slater in 1929:

$$\psi(r_1, ..., r_n) = \frac{1}{\sqrt{n!}} \begin{vmatrix} 1s(1)\alpha(1) & 1s(1)\beta(1) & 2s(1)\alpha(1) & ... \\ 1s(2)\alpha(2) & 1s(2)\beta(2) & 2s(2)\alpha(2) & ... \\ 1s(3)\alpha(3) & ... & ... & ... \\ ... & ... & ... & ... \end{vmatrix} \tag{7.23}$$

For a ground state He atom this can be written:

$$\psi(r_1, r_2) = \frac{1}{\sqrt{2}} \begin{vmatrix} 1s(1)\alpha(1) & 1s(1)\beta(1) \\ 1s(2)\alpha(2) & 1s(2)\beta(2) \end{vmatrix} \tag{7.24}$$

The Slater determinant automatically ensures that the total wavefunction is antisymmetric. Note that from the mathematical point of view the antisymmetry requirement restricts the allowed solutions to the Schrödinger equation. The lowest energy solutions are typically symmetric and the antisymmetric states correspond to higher energy. However, one must be careful with terminology here because only the antisymmetric states exist for electrons (Fermions) and as such they are the lowest energy (ground state) solutions for them. Particles that have symmetric wavefunctions are called Bosons (for example, ^{4}He atoms).

In general, particles with half-integral spin ($s = \frac{1}{2}, \frac{3}{2}, ...$) are Fermions (Fermi-Dirac statistics) and particles with integral spin ($s = 0, 1, 2, ...$) are Bosons (Bose-Einstein statistics). Note that electron spin enters the Hamiltonian only when external fields are present or when spin-orbit interaction is included (will be discussed later).

Example. Show that the Slater determinant in Eq. (7.23) for the ground state helium atom is an eigenfunction of the total spin operator $\hat{S}_{z,tot} = \hat{S}_{z_1} + \hat{S}_{z_2}$, where 1 and 2 refer to the two electrons.

Solution. First we recall how $\hat{S}_z$ operates on electron spin as follows:

$$\hat{S}_z|\alpha\rangle = +\frac{\hbar}{2}|\alpha\rangle \text{ and } \hat{S}_z|\beta\rangle = -\frac{\hbar}{2}|\beta\rangle$$

Next, we expand the Slater determinant in Eq. (7.23):

$$\psi = \overbrace{1s(1)1s(2)}^{\text{symmetric}} \times \overbrace{\frac{1}{\sqrt{2}}\left(\alpha(1)\beta(2) - \alpha(2)\beta(1)\right)}^{\text{antisymmetric}}$$

Operate on this by $\hat{S}_{z_1}$ and $\hat{S}_{z_2}$. They operate only on the spin-part and on the corresponding electron only:

$$\hat{S}_{z_1}|\psi\rangle = |1s(1)1s(2)\rangle \times \frac{1}{\sqrt{2}}\left(\frac{\hbar}{2}|\alpha(1)\beta(2)\rangle + \frac{\hbar}{2}|\beta(1)\alpha(2)\rangle\right)$$

$$\hat{S}_{z_2}|\psi\rangle = |1s(1)1s(2)\rangle \times \frac{1}{\sqrt{2}}\left(-\frac{\hbar}{2}|\alpha(1)\beta(2)\rangle - \frac{\hbar}{2}|\beta(1)\alpha(2)\rangle\right)$$

$$\Rightarrow\ \hat{S}_{z,tot}|\psi\rangle = \left(\hat{S}_{z_1} + \hat{S}_{z_2}\right)|\psi\rangle = \hat{S}_{z_1}|\psi\rangle + \hat{S}_{z_2}|\psi\rangle = 0$$

Note that the two terms are equal in magnitude but have opposite signs and they cancel. Thus the eigenvalue of the z-component of the total spin is zero. It can also be shown that $S^2 = 0$. This kind of electronic configuration is called a *singlet state* (i.e. the two electrons have opposite spins).

Previously we had both electrons on $1s$ orbital with opposite spins. If the electrons reside on two different orbitals, for example, $1s$ and $2s$, we would have an excited helium atom. Such state can be created experimentally by a suitable high-energy process (laser induced break-down etc.). The spatial part of the wavefunction is $\psi = 1s(1)2s(2)$. It could as well be $\psi = 2s(1)1s(2)$ as we cannot distinguish the electrons from each other. Obviously we must form a linear combination of these so that both electrons appear identical (two possibilities):

$$\psi_{sym} = \frac{1}{\sqrt{2}}\left(1s(1)2s(2) + 1s(2)2s(1)\right)\ \text{(symmetric)} \tag{7.25}$$

$$\psi_{asym} = \frac{1}{\sqrt{2}}\left(1s(1)2s(2) - 1s(2)2s(1)\right)\ \text{(antisymmetric)} \tag{7.26}$$

Note that these two states may have different energies.

Next, we consider adding the spin part to these wavefunctions. Because the electrons are on two different orbitals, we have the following four possibilities:

$$\left.\begin{array}{c} \phi_1 = \alpha(1)\alpha(2) \text{ (symmetric)} \\ \phi_2 = \beta(1)\beta(2) \text{ (symmetric)} \\ \phi_3 = \frac{1}{\sqrt{2}}\left(\alpha(1)\beta(2) + \beta(1)\alpha(2)\right) \text{ (symmetric)} \end{array}\right\} \text{Triplet state}$$

$$\left.\phi_4 = \frac{1}{\sqrt{2}}\left(\alpha(1)\beta(2) - \beta(1)\alpha(2)\right)\right\} \text{Singlet state}$$

Before we can combine the spatial and spin wavefunctions, we must consider the symmetries of these functions. Remember that the total wavefunction must be antisymmetric. Thus the allowed combinations are: symmetric (spatial) $\times$ antisymmetric (spin) or antisymmetric (spatial) $\times$ symmetric (spin). The total wavefunction for the triplet state is therefore:

$$\psi_1 = \frac{1}{\sqrt{2}}\left(1s(1)2s(2) - 2s(1)1s(2)\right)\alpha(1)\alpha(2) \tag{7.27}$$

$$\psi_2 = \frac{1}{2}\left(1s(1)2s(2) - 2s(1)1s(2)\right)\left(\alpha(1)\beta(2) + \beta(1)\alpha(2)\right) \tag{7.28}$$

$$\psi_3 = \frac{1}{\sqrt{2}}\left(1s(1)2s(2) - 2s(1)1s(2)\right)\beta(1)\beta(2) \tag{7.29}$$

For the singlet state we have:

$$\psi_4 = \frac{1}{2}\left(1s(1)2s(2) + 2s(1)1s(2)\right)\left(\alpha(1)\beta(2) - \alpha(2)\beta(1)\right) \tag{7.30}$$

Singlet and triplet states have been named after the number of spin degenerate levels they posses. The total spin $\hat{S}^2$ and $\hat{S}_z$ operators for these states yield:

$$\text{Triplet: } \begin{cases} \hat{S}_z|\psi_1\rangle = +\hbar|\psi_1\rangle,\ \hat{S}_z|\psi_2\rangle = 0|\psi_2\rangle,\ \hat{S}_z|\psi_3\rangle = -\hbar|\psi_3\rangle \\ \hat{S}^2|\psi_i\rangle = 2\hbar^2|\psi_i\rangle \text{ where } i = 1,2,3. \end{cases} \tag{7.31}$$

$$\text{Singlet: } \begin{cases} \hat{S}_z|\psi_4\rangle = 0|\psi_4\rangle \\ \hat{S}^2|\psi_4\rangle = 0|\psi_4\rangle \end{cases}$$

For helium atom, the Pauli principle relevant for the excited states but, for example, for lithium the principle is important for the ground state. The Slater determinant for Li can be written as (expansion of the determinant looks already quite complicated):

$$\psi = \frac{1}{\sqrt{6}}\begin{vmatrix} 1s(1)\alpha(1) & 1s(1)\beta(1) & 2s(1)\alpha(1) \\ 1s(2)\alpha(2) & 1s(2)\beta(2) & 2s(2)\alpha(2) \\ 1s(3)\alpha(3) & 1s(3)\beta(3) & 2s(3)\alpha(3) \end{vmatrix} \tag{7.32}$$

Note that the last column could have been labeled as β as well (degeneracy).

Chapter 8: Molecular structure

8.1 Hydrogen molecule-ion, H_2^+

Recall the Born-Oppenheimer approximation from the undergraduate quantum chemistry course (electrons follow nuclear motion adiabatically):

$$\Psi(R, r) = \phi(R) \times \psi(r) \tag{8.1}$$

where $\Psi(R, r)$ is the total wavefunction that depends on both nuclear and electronic coordinates R and r, respectively, $\phi(R)$ is the nuclear wavefunction, and $\psi(r)$ is the electronic wavefunction. Note that this approximation can fail when the electronic states become degenerate (non-adiabatic transitions).

We will now quickly review how molecular orbitals come out from the Schrödinger equation by considering a simple example. The electronic Hamiltonian for H_2^+ is:

$$\hat{H} = -\frac{\hbar^2}{2m_e}\nabla^2 - \frac{e^2}{4\pi\epsilon_0 r_A} - \frac{e^2}{4\pi\epsilon_0 r_B} + \frac{e^2}{4\pi\epsilon_0 R} \tag{8.2}$$

where r_A is the distance between the electron and nucleus A, r_B is the same quantity for nucleus B, and R is the distance between the two nuclei (A and B).

While this one electron problem could be solved exactly (in elliptical coordinates), we will seek for an approximate solution by expressing the solution as a linear combination of suitable basis functions:

$$\psi(r) = \sum_{i=1}^{\infty} c_i \phi_i(r) \tag{8.3}$$

Here we choose a simple (finite) basis set that consists of only two hydrogen atom orbitals centered on each atom:

$$\psi(r) = c_A\phi_A(r) + c_B\phi_B(r) \tag{8.4}$$

This can be inserted into Eq. (6.33), which gives the following determinant equation:

$$\begin{vmatrix} \alpha - E & \beta - ES \\ \beta - ES & \alpha - E \end{vmatrix} = 0 \tag{8.5}$$

where the overlap matrix $S = \langle\phi_A|\phi_B\rangle = 0$, $\alpha = \left\langle \phi_A \left| \hat{H} \right| \phi_A \right\rangle$, $\beta = \left\langle \phi_A \left| \hat{H} \right| \phi_B \right\rangle$, and E is the energy of the molecule. Just like the example below Eq. (6.33), this gives the following roots for E (eigenvalues):

$$E_+ = \frac{\alpha + \beta}{1 + S} \tag{8.6}$$
$$E_- = \frac{\alpha - \beta}{1 - S}$$

By inserting these values of E in turn into Eq. (6.32), one can obtain the corresponding eigenvectors (orthogonal):

$$E^+ : c_A = c_B = \frac{1}{\sqrt{2(1+S)}} \tag{8.7}$$
$$E^- : c_A = -c_B = \frac{1}{\sqrt{2(1-S)}}$$

where the latter equalities follow from normalization: $\langle\psi|\psi\rangle = 1$.

The quantities α (Coulomb integral) and β (resonance integral) are (see Eq. (8.2)):

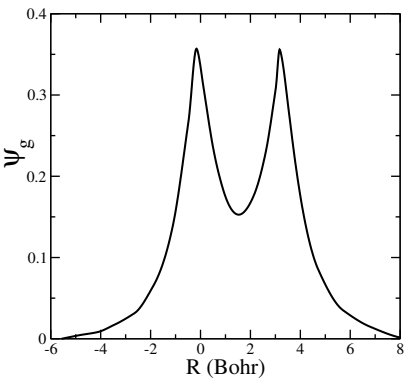
$$\alpha = \left\langle \phi_A \left| \hat{H} \right| \phi_A \right\rangle = E_{1s} - \frac{e^2}{4\pi\epsilon_0} \left\langle \phi_A \left| \frac{1}{r_B} \right| \phi_A \right\rangle + \frac{e^2}{4\pi\epsilon_0 R} \tag{8.8}$$

$$\beta = \left\langle \phi_A \left| \hat{H} \right| \phi_B \right\rangle = E_{1s} \langle \phi_A | \phi_B \rangle - \frac{e^2}{4\pi\epsilon_0} \left\langle \phi_A \left| \frac{1}{r_B} \right| \phi_B \right\rangle + \frac{e^2}{4\pi\epsilon_0 R} \langle \phi_A | \phi_B \rangle$$

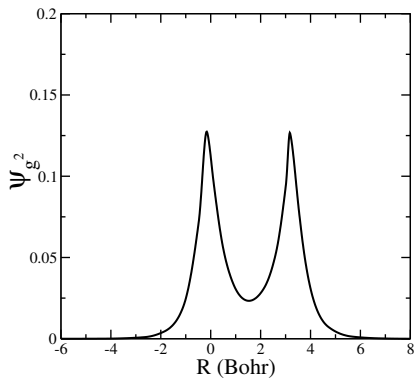
The eigenfunctions corresponding to Eq. (8.7) are:

$$1\sigma_g \equiv \psi_+(r) = \frac{1}{\sqrt{2(1+S)}} \left(\phi_A(r) + \phi_B(r)\right) \tag{8.9}$$

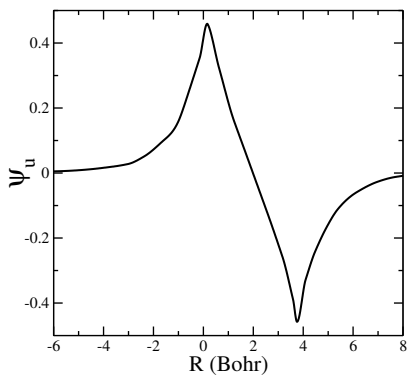
$$1\sigma_u \equiv \psi_-(r) = \frac{1}{\sqrt{2(1-S)}} \left(\phi_A(r) - \phi_B(r)\right)$$



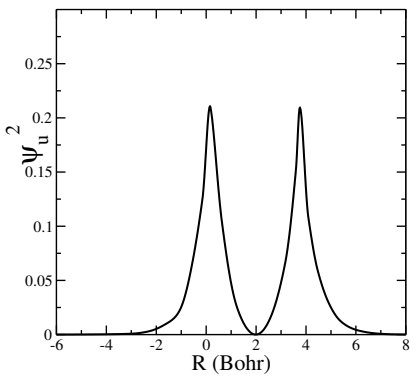
Bonding orbital (ψ_+)



Bonding orbital (ψ_+^2)



Antibonding orbital (ψ_-)



Antibonding orbital (ψ_-^2)

Note that the antibonding orbital has zero electron density between the nuclei.

8.2 Homonuclear diatomic molecules

The previous treatment demonstrates demonstrates the formation of molecular orbitals (MOs). For homonuclear diatomic molecules, the same energy atomic orbitals combine as follows:

$1s \pm 1s$	$1\sigma_g$ (bonding) and $1\sigma_u$ (antibonding).
$2s \pm 2s$	$2\sigma_g$ (bonding) and $2\sigma_u$ (antibonding).
$2p_{x,y} \pm 2p_{x,y}$	$1\pi_u$ (bonding) and $1\pi_g$ (antibonding).
$2p_z \pm 2p_z$	$3\sigma_g$ (bonding) and $3\sigma_u$ (antibonding).

If the electron-electron interaction in the Hamiltonian (H_2 has two electrons) is ignored, we can write the electronic configuration as $(1\sigma_g)^2$ (ground state), etc. Such electronic configurations can be written for homonuclear diatomics:

N_2	$1\sigma_g^2 1\sigma_u^2 1\pi_u^4 2\sigma_g^2$	Term sybmol: $^1\Sigma_g$
O_2	$1\sigma_g^2 1\sigma_u^2 2\sigma_g^2 1\pi_u^4 1\pi_g^2$	Term sybmol: $^3\Sigma_g$

An anti-symmetric total wavefunction can then be written using the Slater apprach discussed earlier:

$$\psi_{MO}^{(1\sigma_g)^2} = \frac{1}{\sqrt{2}} \begin{vmatrix} 1\sigma_g(1)\alpha(1) & 1\sigma_g(1)\beta(1) \\ 1\sigma_g(2)\alpha(2) & 1\sigma_g(2)\beta(2) \end{vmatrix} \tag{8.10}$$

Review term symbols from the undergraduate quantum notes.

8.3 Configuration interaction

In order to improve the approximate wavefunction constructed from a single Slater determinant, it is possible to use a linear combination of Slater determinants (each with their own electronic configuration & weight coefficient) as the wavefunction. This is called the *method of configuration interaction* (CI). Consider H_2 molecule- with the following electronic configurations: $1\sigma_g^2$, $1\sigma_g^1\sigma_u^1$, and $1\sigma_u^2$. The wavefunctions with the corresponding spin parts are:

$$\begin{aligned}
\Psi_1(^1\Sigma_g) &= 1\sigma_g(1)1\sigma_g(2) \times \phi_4 \\
\Psi_2(^1\Sigma_u) &= \frac{1}{\sqrt{2}}\left(1\sigma_g(1)1\sigma_u(2) + 1\sigma_g(2)1\sigma_u(1)\right) \times \phi_4 \\
\Psi_3(^1\Sigma_g) &= 1\sigma_u(1)1\sigma_u(2) \times \phi_4 \\
\Psi_4(^3\Sigma_u) &= \frac{1}{\sqrt{2}}\left(1\sigma_g(1)1\sigma_u(2) - 1\sigma_g(2)1\sigma_u(1)\right) \times \phi_{1,2,3}
\end{aligned} \tag{8.11}$$

where ϕ_1, ϕ_2, and ϕ_3 are the triplet spin parts (see Eqs. (7.27), (7.28), and (7.29)) and ϕ_4 is the singlet spin part (see Eq. (7.30)). Only the same symmetry states may mix and therefore the CI wavefunction for the ground state H_2 would be written as:

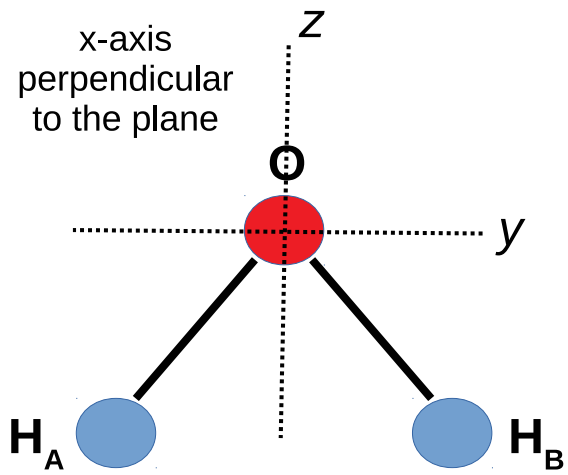
$$\Psi = c_1\Psi_1 + c_3\Psi_3 \tag{8.12}$$

where c_1 and c_3 are variational parameters to be determined. This wavefunction is more flexible than just Ψ_1.

8.4 Polyatomic molecules

Calculations of more complex molecules can also employ the linear combination of atomic orbitals (LCAO) technique as we have used in the previous sections. The resulting molecular orbitals can be classified in terms of the irreps of the point group in which the molecule belongs to. This was covered to some extent in the undergraduate quantum chemistry notes (review the material), but we will demonstrate the concept for water molecule below as we will need this for the computational chemistry part.

H_2O belongs to C_{2v} point group, which consists of A_1 (fully symmetric), A_2, B_1, and B_2 irreps. The two hydrogen atoms are labeled by A and B. The atomic basis functions are: $H(1s_A)$, $H(1s_B)$, $O(2s)$, $O(2p_x)$, $O(2p_y)$, and $O(2p_z)$. The orientation of the molecule is shown below.



C_{2v}	E	C_2	$\sigma_v(xz)$	$\sigma_v'(yz)$
A_1	1	1	1	1
A_2	1	1	-1	-1
B_1	1	-1	1	-1
B_2	1	-1	-1	1

Let us first consider the effect of the symmetry operations on these atomic orbitals:

C_{2v}	O($2s$)	O($2p_x$)	O($2p_y$)	O($2p_z$)	H($1s_A$)	H($1s_B$)
E	O($2s$)	O($2p_x$)	O($2p_y$)	O($2p_z$)	H($1s_A$)	H($1s_B$)
C_2	O($2s$)	$-$O($2p_x$)	$-$O($2p_y$)	O($2p_z$)	H($1s_B$)	H($1s_A$)
σ_v	O($2s$)	O($2p_x$)	$-$O($2p_y$)	O($2p_z$)	H($1s_B$)	H($1s_A$)
σ_v'	O($2s$)	$-$O($2p_x$)	O($2p_y$)	O($2p_z$)	H($1s_A$)	H($1s_B$)

We can then construct the matrix representations for the symmetry operations. As an example, this is shown for C_2 below (order of basis functions as shown above):

$$D(C_2) = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & \mathbf{0} & \mathbf{1} \\ 0 & 0 & 0 & 0 & \mathbf{1} & \mathbf{0} \end{pmatrix}$$

This matrix is diagonal with the exception of the block at the lower right (in boldface). The other matrix representations can be constructed similarily and the diagonals are:

Orbital	E	C_2	σ_v	σ_v'	Irrep
$\mathrm{O}(2s)$	$+1$	$+1$	$+1$	$+1$	A_1
$\mathrm{O}(2p_x)$	$+1$	-1	$+1$	-1	B_1
$\mathrm{O}(2p_y)$	$+1$	-1	-1	$+1$	B_2
$\mathrm{O}(2p_z)$	$+1$	$+1$	$+1$	$+1$	A_1

The 2x2 block on the previous page (in boldface) is not diagonal but can be diagonalized by rotating the original basis set $\mathrm{H}(1s_A)$, $\mathrm{H}(1s_B)$ as follows:

$$\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} \mathrm{H}_{1s_A} \\ \mathrm{H}_{1s_B} \end{pmatrix} = \begin{pmatrix} \mathrm{H}_{1s_A} + \mathrm{H}_{1s_B} \\ \mathrm{H}_{1s_A} - \mathrm{H}_{1s_B} \end{pmatrix}$$

Thus the new basis functions are $\mathrm{H}_{1s_A} + \mathrm{H}_{1s_B}$ and $\mathrm{H}_{1s_A} - \mathrm{H}_{1s_B}$, which belong to A_1 and B_2 irreps, respectively. Only basis functions with the same symmetry can group together to form the molecular orbitals (SALCs):

$$\begin{aligned} \psi(A_1) &= c_1\mathrm{O}(2s) + c_2\mathrm{O}(2p_z) + c_3\left(\mathrm{H}(1s_A) + \mathrm{H}(1s_B)\right) \\ \psi(B_1) &= \mathrm{O}(2p_x) \\ \psi(B_2) &= c_4\mathrm{O}(2p_y) + c_5\left(\mathrm{H}(1s_A) - \mathrm{H}(1s_B)\right) \end{aligned}$$

If one would calculate the electronic structure of water molecule then the resulting electronic configuration could be expressed as: $1a_1^2 1b_2^2 2a_1^2 1b_1^2$ and the overall symmetry would then be ${}^1A_1$ (direct product). The convention for writing the electronic configuration uses lower case letters whereas the overall symmetry is in upper case.

Chapter 9: The calculation of electronic structure

9.1 Mathematical definition of the problem

Electronic structure calculations employ the time-independent Schrödinger equation:

$$\hat{H}\psi(r;R) = E(R)\psi(r;R) \tag{9.1}$$

where $\psi(r;R)$ is the wavefunction (r represents the electronic coodinates and R the fixed molecular geometry), $E(R)$ is the electronic energy of the system at a given nuclear configuration R, and $\hat{H}$ is the electronic Hamiltonian:

$$\hat{H} = -\frac{\hbar^2}{2m_e}\sum_{i=1}^{n}\nabla_i^2 - \sum_{i=1}^{n}\sum_{j=1}^{N}\frac{Z_j e^2}{4\pi\epsilon_0 r_{ij}} + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{e^2}{4\pi\epsilon_0 r_{ij}} \tag{9.2}$$

Here n is the number of electrons and N the number of nuclei in the molecule. Note that this Hamiltonian is missing the nuclear - nuclear repulsion term, but this can be added to the total energy at the end of the calculation as the nuclei are assumed to be stationary. In terms of mathematics, Eq. (9.1) classifies as an operator eigevalue problem. For practical numerical calculations, this equation is always discretized such that it becomes a matrix eigenvalue problem. In addition to the direct calculation of the energy, the equilibrium molecular geometry can be obtained by minimizing the energy in terms of R (geometry optimization).

There are two categories of methods for calculating the electronic structure: 1) *ab initio* where the only input to the calculation is Eq. (9.1) and a basis set to represent the wavefunction and 2) *semiempirical* where a simplified form of the electronic Hamiltonian is used with possible additional parameters included from experimental data. *Ab initio* provide a clear strategy for improving the accuracy of the calculation whereas the situation is the opposite for most semiempirical methods.

9.2 The Hartree-Fock method

As we have seen before, the main difficulty in solving the electronic Schrödinger equation arises from the electron - electron interaction term in the Hamiltonian. In the Hartree-Fock (HF) method this will be included in an average fashion such that the correlated motion between the electrons is neglegted. Since the HF wavefunction is not going to be an exact solution of the electronic Schrödinger equation, we have to calculate the corresponding energy by Eq. (6.34):

$$E = \frac{\left\langle \psi \left| \hat{H} \right| \psi \right\rangle}{\langle \psi | \psi \rangle} = \left\langle \psi \left| \hat{H} \right| \psi \right\rangle \text{ (if } \psi \text{ normalized)} \tag{9.3}$$

The wavefunction in the HF method is taken to be a single Slater determinant (see Eq. (7.23)), which accounts for the proper antisymmetry for fermions:

$$\psi(r_1, ..., r_n) = \frac{1}{\sqrt{n!}} \begin{vmatrix} \phi_1(1) & \phi_2(1) & \phi_3(1) & ... \\ \phi_1(2) & \phi_2(2) & \phi_3(2) & ... \\ \phi_1(3) & ... & ... & ... \\ ... & ... & ... & ... \end{vmatrix} \tag{9.4}$$

where ϕ_i's are the molecular orbitals, n is the number of electrons, and the orbitals are taken to be orthonormalized: $\langle \phi_i | \phi_j \rangle = \delta_{ij}$ (Kronecker delta). More advanced methods include electron correlation by including more than one Slater determinant in the wavefunction, each with their own variational coefficient.

9.3 The energy of a single Slater determinant

We will first express the Slater determinant (Φ) in terms of the anti-symmetrizer operator $\hat{A}$ and the simple product of orbitals ϕ_i (Π):

$$\Phi = \hat{A}\left(\phi_1(1)\phi_2(2)...\phi_n(n)\right) = \hat{A}\Pi \tag{9.5}$$

The antisymmetrizer operator is defined as:

$$\hat{A} = \frac{1}{\sqrt{n!}}\sum_{p=0}^{n-1}(-1)^p\,\hat{P} = \frac{1}{\sqrt{n!}}\left(\hat{1} - \sum_{i,j}\hat{P}_{i,j} + \sum_{i,j,k}\hat{P}_{i,j,k} - ...\right) \tag{9.6}$$

where $\hat{1}$ is the identity operator, $\hat{P}_{i,j}$ generates all possible permutations of two electron coordinates, $\hat{P}_{i,j,k}$ all possible permutations of three electron coordinates, etc. The following results for $\hat{A}$ can be shown to hold:

$$\hat{A}\hat{H} = \hat{H}\hat{A} \text{ (commutes with the Hamiltonian)} \tag{9.7}$$

$$\hat{A}\hat{A} = \sqrt{n!}\hat{A} \tag{9.8}$$

The total Hamiltonian can be broken down as follows:

$$\hat{H} = \hat{T}_e + \hat{V}_{ne} + \hat{V}_{ee} + \hat{V}_{nn} \tag{9.9}$$

where $\hat{T}_e$ is the electron kinetic energy, $\hat{V}_{ne}$ is the nuclear-electron attraction, $\hat{V}_{ee}$ is the electron-electron repulsion, and $\hat{V}_{nn}$ is the nuclear-nuclear repulsion (constant for given molecular geometry).

The component operators above are defined as (N is the number of nuclei in the molecule):

$$\hat{T}_e = -\sum_{i=1}^{n} \frac{1}{2}\nabla_i^2 \tag{9.10}$$

$$\hat{V}_{ne} = -\sum_{i=1}^{n}\sum_{j=1}^{N} \frac{Z_j}{|R_j - r_i|}$$

$$\hat{V}_{ee} = \sum_{i<j}^{n} \frac{1}{|r_i - r_j|}$$

$$V_{nn} = \sum_{i<j}^{N} \frac{Z_i Z_j}{|R_i - R_j|}$$

Since the nuclear - nuclear repulsion term V_{nn} does not depend on the electron coordinates, it can be added to the total energy after the calculation. The remaining electronic part ($\hat{H}_e$) can be written in terms of the one and two electron operators:

$$\hat{H}_e = \sum_{i=1}^{n} \hat{h}_i + \sum_{i<j}^{n} \hat{g}_{ij} + V_{nn} \tag{9.11}$$

$$\hat{h}_i = -\frac{1}{2}\nabla_i^2 - \sum_{j=1}^{N} \frac{Z_j}{|R_j - r_i|} \text{ and } \hat{g}_{ij} = \frac{1}{|r_i - r_j|}$$

The energy of a Slater determinant can be writen in terms of the permutation operator P as follows:

$$E = \left\langle \Phi \left| \hat{H} \right| \Phi \right\rangle = \left\langle \hat{A}\Pi \left| \hat{H} \right| \hat{A}\Pi \right\rangle = \sqrt{n!} \left\langle \Pi \left| \hat{H} \right| \hat{A}\Pi \right\rangle = \sum_{p} (-1)^p \left\langle \Pi \left| \hat{H} \right| \hat{P}\Pi \right\rangle \tag{9.12}$$

For the one-electron $\hat{h}_1$ ($i = 1$; other indices evaluated the same way) and the above becomes:

$$\left\langle \Pi \left| \hat{h}_1 \right| \Phi \right\rangle = \left\langle \phi_1(1)\phi_1(2)...\phi_n(n) \left| \hat{h}_1 \right| \phi_1(1)\phi_2(2)...\phi_n(n) \right\rangle \tag{9.13}$$
$$\left\langle \phi_1(1) \left| \hat{h}_1 \right| \phi_1(1) \right\rangle \langle \phi_2(2)|\phi_2(2)\rangle \langle \phi_3(3)|\phi_3(3)\rangle ... \langle \phi_n(n)|\phi_n(n)\rangle$$

Note that only the identity operator in the permutation operator of Eq. (9.6) can give a non-zero contribution. All actual permutation operators involve change in the electron indices and therefore they include terms such as $\langle \phi_2(2)|\phi_1(2)\rangle$ etc., which are zero based on orthogonality. Consider, for example, the following two-electron permutation operator:

$$\left\langle \Pi \left| \hat{h}_1 \right| \hat{P}_{12}\Pi \right\rangle = \left\langle \phi_1(1)\phi_2(2)...\phi_n(n) \left| \hat{h}_1 \right| \phi_2(1)\phi_1(2)...\phi_n(n) \right\rangle \tag{9.14}$$
$$= \left\langle \phi_1(1) \left| \hat{h}_1 \right| \phi_2(1) \right\rangle \langle \phi_2(2)|\phi_1(2)\rangle ... \langle \phi_n(n)|\phi_n(n)\rangle$$

For the two-electron operator (g_{ij}), both the identity and $\hat{P}_{ij}$ operators can give non-zero contributions. However, three and higher permutation operators create at least one overlap integral between two different orbitals (orthogonality) and hence these must be zero. The identity operator yields the following contribution (*Coulomb integral*):

$$\begin{aligned}\langle \Pi \left|\hat{g}_{12}\right| \Pi\rangle &= \langle \phi_1(1)\phi_2(2)...\phi_n(n) \left|\hat{g}_{12}\right| \phi_1(1)\phi_2(2)...\phi_n(n)\rangle \\ &= \langle \phi_1(1)\phi_2(2) \left|\hat{g}_{12}\right| \phi_1(1)\phi_2(2)\rangle \langle \phi_3(3)|\phi_3(3)\rangle ... \langle \phi_n(n)|\phi_n(n)\rangle \\ &= \langle \phi_1(1)\phi_2(2) \left|\hat{g}_{12}\right| \phi_1(1)\phi_2(2)\rangle = J_{12}\end{aligned} \tag{9.15}$$

This represents the classical Coulomb interaction between the two electrons on these orbitals. The two-electron permutation operator yields (*Exchange integral*):

$$\begin{aligned}\left\langle \Pi \left|\hat{g}_{12}\right| \hat{P}_{12}\Pi \right\rangle &= \langle \phi_1(1)\phi_2(2)...\phi_n(n) \left|\hat{g}_{12}\right| \phi_2(1)\phi_1(2)...\phi_n(n)\rangle \\ &= \langle \phi_1(1)\phi_2(2) \left|\hat{g}_{12}\right| \phi_2(1)\phi_1(2)\rangle \langle \phi_3(3)|\phi_3(3)\rangle ... \langle \phi_n(n)|\phi_n(n)\rangle \\ &= \langle \phi_1(1)\phi_2(2) \left|\hat{g}_{12}\right| \phi_2(1)\phi_1(2)\rangle = K_{12}\end{aligned} \tag{9.16}$$

The above quantities can be arranged in matrix form (vector h_i & matrices J and K) according to electron indices. The total energy corresponding to a given Slater determinant can then be obtained from:

$$\begin{aligned}E &= \sum_{i=1}^{n} h_i + \sum_{i=1}^{n}\sum_{j>i}^{n} (J_{ij} - K_{ij}) + V_{nn} \\ &= \sum_{i=1}^{n} h_i + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} (J_{ij} - K_{ij}) + V_{nn}\end{aligned} \tag{9.17}$$

Note that the minus sign in front of K_{ij} above originates from the $(-1)^p$ term in Eq. (9.6). This can also be expressed in terms of Coulomb and Exchange operators:

$$E = \sum_{i=1}^{n} \left\langle \phi_i \left| \hat{h}_i \right| \phi_i \right\rangle + \frac{1}{2} \sum_{i,j}^{n} \left(\left\langle \phi_j \left| \hat{J}_i \right| \phi_j \right\rangle - \left\langle \phi_j \left| \hat{K}_i \right| \phi_j \right\rangle \right) + V_{nn} \tag{9.18}$$

The operators $\hat{J}_i$ and $\hat{K}_i$ are defined as:

$$\begin{aligned} \hat{J}_i |\phi_j(2)> &= \langle \phi_i(1) \left| \hat{g}_{12} \right| \phi_i(1) \rangle \, |\phi_j(2)> \\ \hat{K}_i |\phi_j(2)> &= \langle \phi_i(1) \left| \hat{g}_{12} \right| \phi_j(1) \rangle \, |\phi_i(2)> \end{aligned} \tag{9.19}$$

Given the molecular orbitals, this expression yields the electronic energy for the atom/molecule.

9.4 Optimization of the orbitals

Our next task is to find the optimal orbitals that minimize the energy given by Eq. (9.19). We will use a technique called *calculus of variations* to obtain the minimal energy condition for the orbitals. To maintain normalization of the orbitals as well as their orthogonality, we introduce Lagrange multipliers λ_{ij} and write the Langrangian function as:

$$L = E - \sum_{i,j}^{n} \lambda_{ij} \left(\langle \phi_i | \phi_j \rangle - \delta_{ij} \right) \tag{9.20}$$

where δ_{ij} is the Kronecker delta. The variation of L given by:

$$\delta L = \delta E - \sum_{i,j}^{n} \lambda_{ij} \left(\langle \delta\phi_i | \phi_j \rangle + \langle \phi_i | \delta\phi_j \rangle \right) = 0 \tag{9.21}$$

which we have set to zero to find the extremum point with respect to variations in the orbitals (similar to setting the normal first derivative to zero). The variation in energy is given by:

$$\begin{aligned} \delta E = & \sum_{i}^{n} \left(\left\langle \delta\phi_i \left| \hat{h}_i \right| \phi_i \right\rangle + \left\langle \phi_i \left| \hat{h}_i \right| \delta\phi_i \right\rangle \right) \\ & + \frac{1}{2} \sum_{i,j}^{n} \left(\left\langle \delta\phi_i \left| \hat{J}_j - \hat{K}_j \right| \phi_i \right\rangle + \left\langle \phi_i \left| \hat{J}_j - \hat{K}_j \right| \delta\phi_i \right\rangle \right. \\ & \left. + \left\langle \delta\phi_j \left| \hat{J}_i - \hat{K}_i \right| \phi_j \right\rangle + \left\langle \phi_j \left| \hat{J}_i - \hat{K}_i \right| \delta\phi_j \right\rangle \right) \end{aligned} \tag{9.22}$$

Since the summation on the 2nd and 3rd lines above runs symmetrically over i and j, the 3rd and 5th terms and 4th and 6th terms are equal. This cancels the factor $1/2$ in front of the summation:

$$\delta E = \sum_{i}^{n} \left\langle \delta\phi_i \left| \hat{h}_i \right| \phi_i \right\rangle + \left\langle \phi_i \left| \hat{h}_i \right| \delta\phi_i \right\rangle \tag{9.23}$$
$$+ \sum_{i,j}^{n} \left(\left\langle \delta\phi_i \left| \hat{J}_j - \hat{K}_j \right| \phi_i \right\rangle + \left\langle \phi_i \left| \hat{J}_j - \hat{K}_j \right| \delta\phi_i \right\rangle \right)$$

Next we introduce the *Fock operator* $\hat{F}$:

$$\hat{F}_i = \hat{h}_i + \sum_{j}^{n} \left(\hat{J}_j - \hat{K}_j \right) \tag{9.24}$$

which allows to writen Eq. (9.23) as:

$$\delta E = \sum_{i}^{n} \left(\left\langle \delta\phi_i \left| \hat{F}_i \right| \phi_i \right\rangle + \left\langle \phi_i \left| \hat{F}_i \right| \delta\phi_i \right\rangle \right) \tag{9.25}$$

Therefore we can now writen variation in the total Lagrangian, Eq. (9.21), as:

$$\delta L = \sum_{i}^{n} \left(\left\langle \delta\phi_i \left| \hat{F}_i \right| \phi_i \right\rangle + \left\langle \phi_i \left| \hat{F}_i \right| \delta\phi_i \right\rangle \right) - \sum_{i,j}^{n} \lambda_{i,j} \left(\langle \delta\phi_i | \phi_j \rangle + \langle \phi_i | \delta\phi_j \rangle \right) \tag{9.26}$$

Recall that the variational principle states that the optimal orbitals are the ones that make $\delta L = 0$. This should hold for variations with respect to all $\langle \delta\phi_i |$ or $| \delta\phi_i \rangle$. Choosing $\langle \delta\phi_i |$ in Eq. (9.26) yields (functional derivative of calculus of variations):

$$\frac{\delta L}{\delta \phi_i^*} = \hat{F}_i \, |\phi_i\rangle - \sum_j^n \lambda_{i,j} \, |\phi_j\rangle = 0 \tag{9.27}$$

Therefore the final set of Hartree-Fock equations is:

$$\hat{F}_i \phi_i = \sum_j^n \lambda_{i,j} \phi_j \tag{9.28}$$

where matrix $\lambda_{i,j}$ is Hermitian. The diagonal elements of λ represent orbital energies whereas the purpose of the off-diagonals is to keep the orbitals orthogonal to each other. Given a suitable (external) procedure for enforcing orbital orthogonality, the off-diagonals of λ become zero and Eq. (9.28) becomes:

$$\hat{F}_i \phi_i' = \epsilon_i \phi_i' \tag{9.29}$$

where the prime signifies that the original basis set has been rotated such that λ became diagonal. Note that there is one equation per electron (or orbital) and all these equations are coupled (by $\hat{J}$ and $\hat{K}$). This means that the equations must be solved iteratively with a given initial guess (e.g., from semi-empirical *ab inito* methods). The Hartree-Fock iterative procedure is called the *self-consistent field method* (SCF).

The total energy can be calculated from the marix elements of the Fock operator as follows:

$$E = \sum_{i}^{n} \epsilon_i - \frac{1}{2} \sum_{i}^{n} (J_{i,j} - K_{i,j}) + V_{nn} \tag{9.30}$$

where the Fock operator expectation value is:

$$\epsilon_i = \left\langle \phi_i \left| \hat{F}_i \right| \phi_i \right\rangle = h_i + \sum_{j}^{n} (J_{i,j} - K_{i,j}) \tag{9.31}$$

Note: *The total energy is not simply a sum of the orbital energies!* Hartree-Fock method considers the electron-electron interaction only on average and therefore it ignores electron correlation (i.e., uses a single Slater determinant as wavefunction).

It can be shown that the Hartree-Fock orbital energies represent (approximately) the ionization energies for the electrons on those orbitals. This is known as the *Koopman's theorem*. It assumes that all the other molecular orbitals are frozen during the ionization process. For HOMO this gives the first ionization energy and for LUMO the electron affinity. However, in the latter case the frozen orbital approximation may no longer be valid.

9.5 Basis set approximation

Every numerical calculation requires a representation for the orbitals. For small molecules this could consist of a grid where the function values are specified at each grid point. While this type of representation is very flexible, it also requires large amounts of memory and computational time. For this reason, almost all electronic structure calculations employ either gaussian or planewave basis function representations. The former are typically used by quantum chemists whereas the latter by computational physicists. Although Gaussian functions do not have exactly the correct behavior at the nucleus (vs. the Slater-type orbital), they can be grouped together to get the correct behavior approximately. The main advantage of using Gaussian functions is that the integrals required in the calculations are known analytically. This is much faster and more accurate than employing numerical integration techniques.

We express molecular orbitals (MO) ϕ_i as a linear combination of basis functions (χ_α):

$$\phi_i = \sum_{\alpha}^{m} c_{\alpha,i}\chi_\alpha \tag{9.32}$$

Inserting this into Eq. (9.29) gives:

$$\hat{F}_i \sum_{\alpha}^{m} c_{\alpha,i}\chi_\alpha = \epsilon_i \sum_{\alpha}^{m} c_{\alpha,i}\chi_\alpha \tag{9.33}$$

Multiplying the above equation by each basis function χ_β in turn yields the Roothaan-Hall equations for a closed-shell system in atomic orbital basis (M equations collected into matrix notation):

$$FC = SC\epsilon \tag{9.34}$$

where the Fock matrix elements are $F_{\alpha,\beta} = \left\langle \chi_\alpha \left| \hat{F} \right| \chi_\beta \right\rangle$ and the overlap matrix elements $S_{\alpha,\beta} = \langle \chi_\alpha | \chi_\beta \rangle$. Eq. (9.34) is called *generalized matrix eigenvalue problem*. Here C is the eigenvector matrix and ϵ is a diagonal matrix containing the eigenvalues.

Before expanding the Fock matrix element, we define the two-electron integral notation ($\hat{g}$ is the two-electron operator):

$$\langle \alpha\gamma \left| \hat{g} \right| \beta\delta \rangle \equiv \langle \alpha(1)\gamma(2) \left| \hat{g} \right| \beta(1)\delta(2) \rangle \tag{9.35}$$

The Fock matrix element can be expanded as follows:

$$\begin{aligned} \left\langle \chi_\alpha \left| \hat{F} \right| \chi_\beta \right\rangle &= \left\langle \chi_\alpha \left| \hat{h} \right| \chi_\beta \right\rangle + \sum_{j}^{occ.MO} \left\langle \chi_\alpha \left| \hat{J}_j - \hat{K}_j \right| \chi_\beta \right\rangle \\ &= \left\langle \chi_\alpha \left| \hat{h} \right| \chi_\beta \right\rangle + \sum_{j}^{occ.MO} \left(\langle \chi_\alpha \phi_j \left| \hat{g} \right| \chi_\beta \phi_j \rangle - \langle \chi_\alpha \phi_j \left| \hat{g} \right| \phi_j \chi_\beta \rangle \right) \end{aligned} \tag{9.36}$$

We can further substitute in the expansion coefficients from the MO linear combination into the above:

$$
\begin{aligned}
... &= \left\langle \chi_\alpha \left| \hat{h} \right| \chi_\beta \right\rangle + \sum_{j}^{occ.MO} \sum_{\gamma}^{AO} \sum_{\delta}^{AO} c_{\gamma,j} c_{\delta,j} \left(\left\langle \chi_\alpha \chi_\gamma \left| \hat{g} \right| \chi_\beta \chi_\delta \right\rangle - \left\langle \chi_\alpha \chi_\gamma \left| \hat{g} \right| \chi_\delta \chi_\beta \right\rangle \right) \\
&= \left\langle \chi_\alpha \left| \hat{h} \right| \chi_\beta \right\rangle + \sum_{\gamma}^{AO} \sum_{\delta}^{AO} D_{\gamma,\delta} \left(\left\langle \chi_\alpha \chi_\gamma \left| \hat{g} \right| \chi_\beta \chi_\delta \right\rangle - \left\langle \chi_\alpha \chi_\gamma \left| \hat{g} \right| \chi_\delta \chi_\beta \right\rangle \right)
\end{aligned}
\tag{9.37}
$$

where in the last line the the density matrix $D_{\gamma,\delta}$ was introduced:

$$
D_{\gamma,\delta} = \sum_{j}^{occ.MO} c_{\gamma,j} c_{\delta,j} \tag{9.38}
$$

The total energy can also be written in terms of the density matrix as follows:

$$
E = \sum_{\alpha,\beta}^{m} D_{\alpha,\beta} h_{\alpha,\beta} + \frac{1}{2} \sum_{\alpha,\beta,\gamma,\delta}^{m} D_{\alpha,\beta} D_{\gamma,\delta} \left(\left\langle \chi_\alpha \chi_\gamma \left| \hat{g} \right| \chi_\beta \chi_\delta \right\rangle - \left\langle \chi_\alpha \chi_\gamma \left| \hat{g} \right| \chi_\delta \chi_\beta \right\rangle \right) + V_{nn} \tag{9.39}
$$

where m is the number of basis functions. Given the number of occupied orbitals n, the number of unoccupied orbitals is then $m - n$ (also called virtual orbitals).

Note: The two-electron integral notation in Eq. (9.35) is called the *physicist's notation*. Another notation known as the *Mulliken* or *chemist's notation* has a different order of the electron indices:

$$\langle \alpha\beta \,|\hat{g}|\, \gamma\delta \rangle \equiv \langle \alpha(1)\beta(1) \,|\hat{g}|\, \gamma(2)\delta(2) \rangle \tag{9.40}$$

Sometimes the two-electron operator $\hat{g}$ is dropped when writing denoting the integrals. Also <> and () with electron indices inside are often used to express these integrals (but double check how the notation is defined exactly before doing anything!)

To construct the Fock matrix in Eq. (9.36), integrals over all pairs of basis functions and the one-electron operator $\hat{h}$ are needed. There are a total of m^2 such *one-electron integrals* (or core integrals). The two-electron part involves integration over all combinations of four basis functions, which means that there are m^4 of such *two-electron integrals*. In the traditional Hartree-Fock calculation these integrals are precomputed and stored on disk and retrieved whenever needed. In so-called direct SCF method these integrals are calculated on the fly whenever needed and this approach avoids the (usually large) disk space requirement. Although this increases the computational demand, it can be compensated by the relatively slow transfer rate of disk storage devices.

For the two-electron integrals, the basis functions can be located at 1 (all on the same atom), 2, 3 or 4 (all on different atoms). The calculation of the three and four center two-electron integrals can become computationally very demanding and, for this reason, one must be very careful in choosing a suitable form for the basis functions. Gaussian functions have the important property that a product of two gaussians centered on two different nuclei (positions R_A and R_B with exponents α and β) becomes another gaussian that has its center between R_A and R_B. Given the Gaussian functions:

$$G_A(r) = \left(\frac{2\alpha}{\pi}\right)^{3/4} e^{-\alpha(r-R_A)^2} \tag{9.41}$$

$$G_B(r) = \left(\frac{2\alpha}{\pi}\right)^{3/4} e^{-\beta(r-R_B)^2}$$

Their product is given by:

$$G_A(r)G_B(r) = Ke^{-\gamma(r-R_c)^2} \tag{9.42}$$

where $\gamma = \alpha + \beta$, $R_c = (\alpha R_A - \beta R_B) / (\alpha + \beta)$, and the normalization constant is given by:

$$K = \left(\frac{2}{\pi}\right)^{3/2} (\alpha\beta)^{3/4} e^{-\frac{\alpha\beta}{\alpha+\beta}(R_A-R_B)^2}$$

So, integration of products of Gaussians can be reduced to integration of single Gaussian functions. Such integrals can be evaluated in one dimension.

As the number of basis functions is increased, the calculation becomes more and more accurate (but also takes more longer to complete on a computer). In the limit of infinitely large basis set (which cannot be reached in practice), we would obtain so-called *Hartree-Fock limit*, which is the best energy that can be obtained within this theory. Note that this is only approximation and certainly not the solution to the Schrödinger equation (with the exception of hydrogen atom with only one electron). Deviation from the Hartree-Fock limit is called the *basis set truncation error*.

All modern electronic structure programs use Gaussian-type orbitals (GTOs), which are constructed as:

$$\chi_{i,j,k}(r - r_c) = (x - x_c)^i (y - y_c)^j (z - z_c)^k e^{-\alpha |r - r_c|^2} \tag{9.43}$$

where $r_c = (x_c, y_c, z_c)$ is the center of the basis function (usually centered at the nuclei), $r = (x, y, z)$ is the electron coordinate, α is a positive exponent, and i, j, k are non-negative integers. The "angular momentum" of the GTO is determined by the sum $i + j + k$ as: $0 = s$-type Gaussian, $1 = p$-type Gaussian, $2 = d$-type Gaussian, etc. Note that the sum can be equal to zero only in one way (only one s Gaussian) whereas a p Gaussian can be obtained by having $i = 1$, $j = 1$, or $k = 1$. These can be thought to correspond to functions that try to mimic p_x, p_y, p_z -like orbitals. Unfortunately, for d GTO there are six such functions rather than five and one of the functions if often eliminated from the basis set (the sixth one behaves similarly to an s GTO).

As mentioned earlier, Gaussians do not have the desired behavior near the nuclei. For this reason, primitive GTOs are often contracted to form a more flexible function such that they share the same variational coefficient. For example, to form such contracted function χ, we group several GTOs together with their own expansion coefficients:

$$\chi = \sum_i d_i g_i \tag{9.44}$$

where g_i's are the primitive GTOs and d_i's are the contraction coefficients. This scheme allows to construct more flexible basis functions without increasing the computational cost significantly. Note that, in principle, one coule include all the primitive GTOs in the calculation directly with their own variational parameters, but this could make the calculation computationally very expensive.

We conclude that a basis set definition includes the following information:

- GTO type (s, p, d, etc.), position (usually on nuclei), and exponent.
- If contracted term, contraction coefficients for each GTO in the contraction.

Note that specifying a GTO, for example, with angular momentum of p, creates three GTOs: $i = 1, j = 0, k = 0$; $i = 0, j = 1, k = 0$; $i = 0, j = 0, k = 1$. Therefore, high angular momentum GTOs are more expensive to include from the computational point of view.

An example of basis set definition for hydrogen atom ("STO-3G") in Molpro program format:

```
!  STO-3G  EMSL  Basis Set Exchange Library   4/30/18 3:30 PM
! Elements                                References
! --------                                ----------
!  H - Ne: W.J. Hehre, R.F. Stewart and J.A. Pople, J. Chem. Phys. 2657 (1969).
! Na - Ar: W.J. Hehre, R. Ditchfield, R.F. Stewart, J.A. Pople,
!          J. Chem. Phys.  2769 (1970).
! K,Ca - : W.J. Pietro, B.A. Levy, W.J. Hehre and R.F. Stewart,
! Ga - Kr: J. Am. Chem. Soc. 19, 2225 (1980).
! Sc - Zn: W.J. Pietro and W.J. Hehre, J. Comp. Chem. 4, 241 (1983) + Gaussian.
!  Y - Cd: W.J. Pietro and W.J. Hehre, J. Comp. Chem. 4, 241 (1983). + Gaussian
!
!
! HYDROGEN       (3s) -> [1s]
! HYDROGEN       (3s)->[1s]
basis={
s, H , 3.42525091, 0.62391373, 0.16885540
c, 1.3, 0.15432897, 0.53532814, 0.44463454
}
```

(Source: https://bse.pnl.gov/bse/portal)

Three s GTOs with exponents 3.42525091, 0.62391373, and 0.16885540 were grouped into single contraction with coefficients 0.15432897, 0.53532814, and 0.44463454.

Common description of basis sets:

- *Minimial basis set*: one function is used to represent each of the orbitals of elementary valence bond theory. For H and He: one function (for $1s$ orbital); five functions for Li - Ne ($1s$, $2s$, and three $2p$'s); nine functions for Na to Ar, etc.
- *Double-zeta (DZ) basis set:* Each basis function in a minimal basis set is replaced by two functions.
- *Triple-zeta (TZ) basis set:* Each basis function in a minimal basis set is replaced by three functions.
- *Split-valence (SV) basis set:* Compromise between DZ and TZ where each valence atomic orbital is represented by two basis functions while each innter shell atomic orbital is represented by a single basis function.

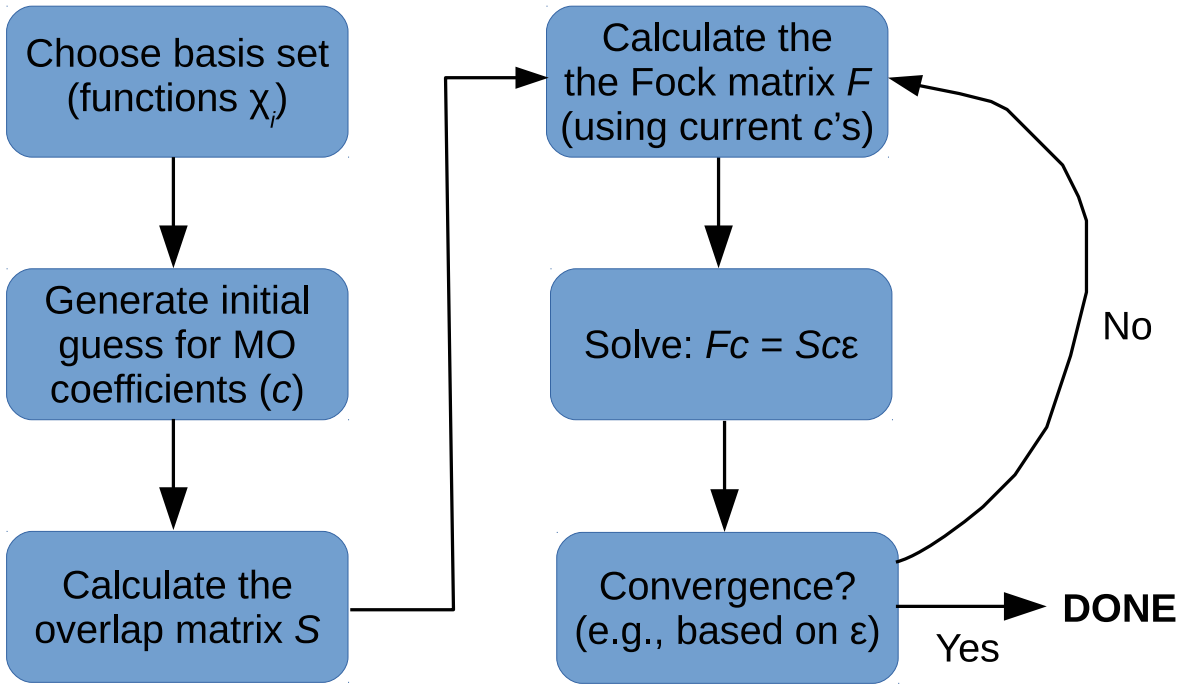
Since atomic orbitals are disturbed by adjacent atom that it forms the chemical bond with, polarization and diffuse functions may be added to improve the accuracy. Polarization can be accounted for by including GTOs with higher angular momentum. Diffuse functions (with small values of exponents α) are often required to account for long-range effects (e.g., hydrogen bonding).
A simplest basis set series is called STO-XG where X is a number 3, 4, 5, 6. This number describes how many GTOs were used to approximate a single Slater-type orbital (STO). In the previous example, STO-3G employed a contraction scheme $(3s) \rightarrow [1s]$. In general, this notation is used to describe contractions used in a basis set (e.g., $(9s5p)/[3s2p]$). There are many acronyms for basis sets: 3-21G, av5z, etc.

Note: When computing weak interactions (e.g., van der Waals interaction), the Hartree-Fock method is not sufficient and additional problems may also arise from the incompleteness of the basis set. This is so-called *basis set superposition error*, which can contribute significantly to binding between binding energy. Such error can be eliminated to some extent by applying *counterpoise correction*, which is outside the scope of this course.

Accuracy of basis sets using the Hartree-Fock method (energies in Hartree; 1 Hartree = 2625.5 kJ/mol).

Basis set	H_2	N_2	CH_4	NH_3	H_2O
STO-3G	−1.117	−107.496	−39.727	−55.454	−74.963
4-31G	−1.127	−108.754	−40.140	−56.102	−75.907
6-31G*	−1.127	−108.942	−40.195	−56.184	−76.011
6-31G**	−1.131	−108.942	−40.202	−56.195	−76.023
HF limit	−1.134	−108.997	−40.225	−56.225	−76.065

Note that optimize molecular geometries converge usually much faster with respect to the basis set than the energies.



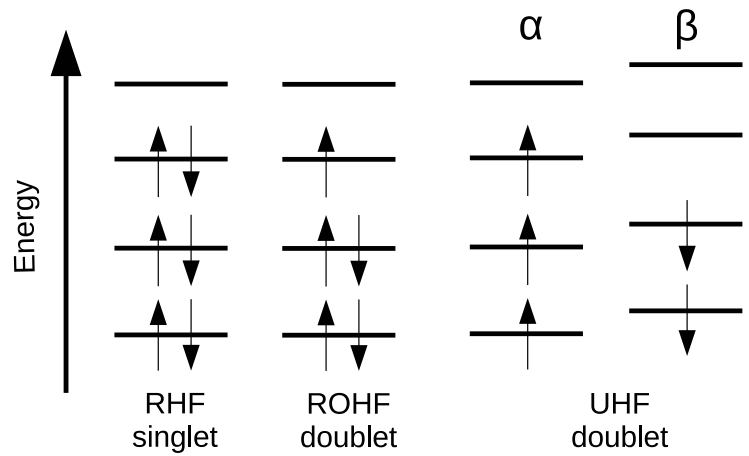
Flow chart of Hartree-Fock calculation.

9.6 Restricted and unrestricted Hartree-Fock

When the Slater determinant is written in terms of the spin orbitals (i.e., the spatial part times the spin part), there is no restriction on the overall spin of the wavefunction. This means that the expectation value of the $\hat{S}^2$ operator may not necessarily correspond exactly to that originally sought. For an electron pair with alpha and beta orbitals, their spatial parts may now also become different from each other. This is called *unrestricted Hartree-Fock* (UHF). For example, for an electronic doublet state molecule, the inner spin orbitals will have different energies (*spin polarization*). Note that for closed shell systems, all alpha/beta orbital pairs remain degenerate. When the orbital degeneracy is enforced (e.g., closed shell species), the method is called *restricted Hartree-Fock* (RHF) method. Since RHF places additional constraints on the wavefunction, UHF will always yield lower energy than RHF. UHF calculations often suffer from *spin contamination*, which means that higher spin states are (artificially) mixed into the solution: singlet may have triplet mixed in, doublet may have quarted mixed in, etc. The degree of spin contamination can be judged by comparing the desired value of $\hat{S}^2$ with the one obtained from the calculation.

If the alpha/beta pairs are restricted to have the same spatial parts for an open shell species, this approach is termed *restricted open shell Hartree-Fock* (ROHF). Note, however, that the orbital energies from this method are not uniquely determined and therefore, for example, Koopman's theorem cannot be used to estimate ionization energies from such wavefunction.

Example electron configurations from the different methods are shown below.



Electronic configurations for RHF, ROHF, and UHF.

Note also that most post-Hartree-Fock methods (to include electron correlation) can typically only employ ROHF orbitals.

9.7 Convergence of the SCF procedure

The self-consistent field (SCF) method used to solve the Hartree-Fock problem is *not guaranteed to convege*. Although this is often the case for molecular geometries near the equilibrium and small basis sets, highly distorted geometries (e.g., transition states) or large basis sets with diffuse basis functions often experience convergence problems. Furthermore, metal complexes where several states with nearly the same energy are present are usually even more difficult to get to converge. Below is a quick overview of various methods that can help in SCF convergence (stability and speed).

- **Extrapolation**: Make the convergence faster by extrapolating previous Fock matrices to generate a better Fock matrix than the one calculated directly from the current density matrix. Often the last three Fock matrices are used in this procedure.
- **Damping**: A common reason for slow convergence (or even divergence) is the oscillatory behavior of the subsequent density matrices (D_n, D_{n+1}, D_{n+2}, ...). In this case, D_n and D_{n+2} can be similar but D_{n+1} very different from the two (i.e., oscillation). To alleviate this problem, the current density matrix can include information from the past density matrix as: $D_{n+1}^{new} = \alpha D_n + (1-\alpha)\, D_{n+1}$ where the weighting factor α can be chosen as a constant or varied dynamically during the calculation.

- **Level shifting**: In this method the virtual orbitals are artificially raised in energy such that they do not interfere with the occupied orbitals. It can be shown that that the total energy is guaranteed to decrease when such level shifting is applied (i.e., avoids oscillatory behavior). However, this will also reduce the convergece speed.
- **Direct Inversion in the Iterative Subspace (DIIS)**: This method is very effective in both enforcing convergence and reducing the number of iterations required to reach convergence. While the exact details of the method are out of the scope of this lecture, it employs information from the past Fock and density matrices to minimize an a priori known error estimate function.
- **Use molecular symmetry!** Specifying the exact electronic configuration is always better than relying on the calculation trying to establish it. This involves specifying the point group for the molecule, irreducible representations for the occupied orbitals, irreducible representation for the overall wavefunction, and specifying the total electron spin (multiplicity). Note that group theory allows to eliminate many of the one- and two-electron integrals need not to be calculated. This may involve forming the proper symmetry adapted functions from the basis (SALCs). The direct product between the irreps must yield the fully symmetric irrep of the point group in question for the integral to yield non-zero value.
- **Use a good initial guess!** If the initial guess orbitals for the SCF procedure are far off from the actual solution, SCF will converge slowly or may not converge at all.

Note: The Hartree-Fock procedure ensures that the the variation of energy with respect to the orbitals is zero. In multi variable calculus, this corresponds to requirement that the gradient vector is zero. However, this only guarantees that the solution is an extremum point, which means that in reality it can be a minimum, maximum, or a saddle point. In order to ensure that the obtained Hartree-Fock solution corresponds to a minimum,

9.8 Existing electronic structure codes

The most common electronic structure codes are:

- Gaussian (commercial): One of the oldest and most widely used electronic structure programs. See http://gaussian.com/
- Gamess (free): Electronic structure code developed by Gordon et al. See http://www.msg.ameslab.gov/gamess/
- Molpro (commercial): Very efficient code for highly correlated methods (small molecules). See https://www.molpro.net/
- Aces-III (free): Similar to molpro (focus on parallel execution). See http://www.qtp.ufl.edu/ACES/
- Psi4 (free): Similar to molpro. See http://www.psicode.org/
- Spartan (commercial): Electronic structure calculations made simple (nice graphical user interface). See https://www.wavefun.com/

Note that most of these programs are "text based", which means that they require a text file (ASCII) as input. This input file describes the point group of the molecule, number of electrons (or charge), the number of nuclei, the positions of nuclei, their types, etc. For large molecules, it is cumbersome to specify the molecular geometry by "hand". This done usually either by specifying the x, y, z coordinates for each atom or so-called Z-matrix, which constructs the molecule based on bond lengths and angles. The output from these programs is also an ASCII file that provides the calculated quantities. Other graphical programs are often employed to construct and visualize the results (e.g., molden). Some commercial programs, such as Gaussian and Spartan, have graphical user interfaces.

9.9 Hartree-Fock examples (Molpro)

The following example will compute the energy of H_2 molecule at the HF/VDZ level (here HF = Hartree-Fock and VDZ denotes the basis set).

```
***,H2
basis=vdz;
geometry={
angstrom;
h1;
h2, h1, 0.74;
}
hf
```

The first *** line is a comment. The second line specifies the basis set as vdz (this is also known as cc-pVDZ; see the basis set order form for available basis sets). There is a large number of basis sets that are builtin Molpro.The third line specifies the molecular geometry (Z-matrix format). In this simple case we have only two atoms and the Z-matrix description just says that there are two hydrogen atoms, h1 and h2, which are separated by 0.74 Å. If this is placed in input file called h2.inp then the output will be placed in h2.out .

Single point Hartree-Fock calculation of water using Cartesian coordinates (x, y, z).

```
geomtyp=xyz
geometry={
3            ! number of atoms
This is an example of geometry input for water with an XYZ file
O ,0.0000000000,0.0000000000,-0.1302052882
H ,1.4891244004,0.0000000000, 1.0332262019
H,-1.4891244004,0.0000000000, 1.0332262019
}
hf
```

The first line specifies that the molecular geometry will be given in Cartesian coordinates. The geometry-block (defined by the curly braces), specifies: number of molecules; comment line; atom type, x-coord, y-coord, z-coord. The last hf indicates that this will be Hartree-Fock calculation.

Single point Hartree-Fock calculation of water using Z-matrix ("internal coordinates").

```
***,h2o                    !A title
r=1.85,theta=104           !set geometry parameters
geometry={O;               !z-matrix geometry input
          H1,O,r;
          H2,O,r,H1,theta}
hf                         !closed-shell scf
```

The first line is comment. The second line defines two variables "r" and "theta", which are set to the given values. These variables are used in the Z-matrix specification below. The geometry block holds the Z-matrix: The first O defines that we start with the oxygen atom; the second H1 says that H1 is located at distance "r" from the oxygen; the third line says that H2 is located at distance "r" from the first oxygen and it forms angle "theta" between H2 - H1 - O. The program determined the molecular symmetry automatically (C_{2v}), which can be seen by inspecting the output file. Also, no basis set was defined and the program assumes vdz basis set in this case (see the output file to confirm).

Note: You can invoke geometry optimization by adding command optg after hf in the above example.

Inspecting the output file (.out) for the previous water example. The basis set used for each atom:

```
 Library entry O      S cc-pVDZ                 selected for orbital group  1
 Library entry O      P cc-pVDZ                 selected for orbital group  1
 Library entry O      D cc-pVDZ                 selected for orbital group  1
 Library entry H      S cc-pVDZ                 selected for orbital group  2
 Library entry H      P cc-pVDZ                 selected for orbital group  2
```

Information on molecular geometry:

```
Molecule type: Asymmetric top,     Order of principal axis = 2
Symmetry elements: X,Y
Rotational constants: 273.1674388 777.0700186 421.2525104 GHz (calculated with

Point group  C2v

ATOMIC COORDINATES
NR  ATOM    CHARGE       X              Y              Z
  1  O       8.00    0.000000000    0.000000000   -0.127449274
  2  H1      1.00    0.000000000    1.457819894    1.011524455
  3  H2      1.00    0.000000000   -1.457819894    1.011524455
```

Overview of the Hartree-Fock procedure:

```
1PROGRAM * RHF-SCF (CLOSED SHELL)       Authors: W. Meyer, H.-J. Werner

 NUMBER OF ELECTRONS:       5+    5-
 CONVERGENCE THRESHOLDS:    1.00E-05 (Density)    1.00E-07 (Energy)
 MAX. NUMBER OF ITERATIONS:       60
 INTERPOLATION TYPE:            DIIS
 INTERPOLATION STEPS:              2 (START)      1 (STEP)
 LEVEL SHIFTS:                  0.00 (CLOSED)  0.00 (OPEN)

 Orbital guess generated from atomic densities.
 Full valence occupancy:    4   1   2   0

 Molecular orbital dump at record        2100.2

 Initial alpha occupancy:   3   1   1   0
 Initial beta  occupancy:   3   1   1   0
 Wave function symmetry:    1
```

The numbers above (3 1 1 0 etc.) represent irreducible representations in C_{2v}: $1 = A_1$, $2 = B_1$, $3 = B_2$, $4 = A_2$.

The Hartree-Fock iterations:

```
ITER DDIFF     GRAD      ENERGY       2-EL.EN.            DIPOLE MOMENTS DIIS O
1    0.000D+00 0.000D+00 -75.98489162 77.641371 0.00000 0.00000 1.17879  0    s
2    0.000D+00 0.378D-01 -76.01535790 74.243294 0.00000 0.00000 0.75877  1    d
3    0.781D-01 0.204D-01 -76.02463622 75.624937 0.00000 0.00000 0.83922  2    d
4    0.263D-01 0.224D-02 -76.02486900 75.527940 0.00000 0.00000 0.82525  3    d
5    0.481D-02 0.627D-03 -76.02488924 75.540341 0.00000 0.00000 0.82098  4    d
6    0.158D-02 0.982D-04 -76.02488988 75.540224 0.00000 0.00000 0.82025  5    d
7    0.253D-03 0.139D-04 -76.02488990 75.540168 0.00000 0.00000 0.82012  6    d
8    0.432D-04 0.113D-05 -76.02488990 75.540160 0.00000 0.00000 0.82012  7    d
9    0.303D-05 0.134D-06 -76.02488990 75.540157 0.00000 0.00000 0.82012  0    o

Final occupancy:   3   1   1   0

!RHF STATE 1.1 Energy               -76.024889896305
Nuclear energy                      8.99162654
One-electron energy                 -122.78659484
Two-electron energy                 37.77007840
Virial quotient                     -1.00152916
!RHF STATE 1.1 Dipole moment 0.00000000 0.00000000 0.82011576
Dipole moment /Debye                0.00000000 0.00000000 2.08438982
```

The orbital information:

```
Orbital energies:
   1.1         2.1         3.1         4.1         5.1
-20.554518  -1.326350  -0.563601  0.181733  0.836590

   1.2         2.2         3.2
-0.491854   1.199673   1.668166

   1.3         2.3         3.3
-0.686895  0.252984   0.772437

   1.4         2.4
1.472460     3.315890

HOMO      1.2    -0.491854 =  -13.3840eV
LUMO      4.1     0.181733 =    4.9452eV
LUMO-HOMO         0.673587 =   18.3292eV
```

How to run molpro on Linux?

```
% mg h2o.inp           (or any other text editor)
% molpro h2o.inp
% more h2o.out         (to view the output)
```

Molden program (free) can be used to construct Z-matrices and visualize the output (e.g., orbitals). To get molden compatible output from molpro, include the following line at the end of the molpro input file:

```
put,molden,h2o.molden;
```

Then use molden on this file:

```
% molden h2o.molden
```