Advanced Quantum Mechanics for Chemists: Chem. 115ab/215ab Notes

Daniel Neuhauser

Department of Chemistry and Biochemistry,
UCLA

April 3, 2020

Abstract

This set of notes is for a two-quarter graduate-course on quantum mechanics. It assumes you have had the standard one-quarter physical chemistry QM course for chemistry undergrads that covers wavefunction based quantum mechanics, including: the Schrödinger equation in 1D and 3D, coordinate-based solution to Harmonic Oscillators (H.O.), and the hydrogen atom and spherical harmonics.

You also need to know linear-algebra (for real vectors; linear-algebra for complex operators is developed here) and have some minimal exposure to complex numbers. Also, knowledge of Lagrange multipliers is assumed.

The notes first develop the mathematics of quantum mechanics, including delta-functions and the bra-ket notation (Chapters 1-3). Then this language is used to develop quantum mechanics in bra-ket notation (Chapters 4-9,11) including operator-based solutions to H.O., the Heisenberg picture, angular momenta and spin, and perturbation theory and the variational theorem. In Chapter 10 we will insert more math: Functional derivatives and Lagrange multipliers, needed for the variational theorem and later for DFT.

Latter chapters (Chapters 12-19) develop the many-electron problem, including Slater-determinants, Hartree-Fock theory, DFT, MP2 and 2nd-quantization, with an epilogue on other modern approaches.

Later topics (Chapter XXX) first cover the Born-Oppenheimer and non-adiabatic interaction,

Then (Chapters XXX-XXX) we spend several chapters on time-dependent perturbation theory, using both the Fermi Golden rule and the correlation-function picture. These chapters use first a classical description of the electromagnetic interaction, with the following chapter developing the interaction with a quantum electromagnetic field.

Path integrals follow in Chapter XXX

Several advanced topics then follow. XXX

Contents

1 Math: General

1.1 Complex numbers

Made from two real numbers:

$$z = x + iy, \;\; i^2 = -1, \tag{1.1}$$

and the complex-conjugate of a complex number is defined as:

$$z^* = x - iy. \tag{1.2}$$

Properties:

$$zz' = (x+iy)(x'+iy') = xx'+i^2yy'+i(xy'+x'y) = (xx'-yy')+i(xy'+x'y) \tag{1.3}$$

and

$$(zz')^* = z^*z'^* \tag{1.4}$$

and the conjugate of a conjugate brings us back to the original number

$$(z^*)^* = z. \tag{1.5}$$

The product of a complex number with its conjugate will be a real-positive #, the absolute value squared:

$$z^*z = (x+iy)(x-iy) = x^2 + y^2 \equiv |z|^2. \tag{1.6}$$

Note that this gives a prescription for the inverse of a complex #:

$$\frac{1}{z} = \frac{z^*}{|z|^2}, \tag{1.7}$$

i.e.,

$$\frac{1}{x+iy} = \frac{x-iy}{x^2+y^2}. \tag{1.8}$$

Finally, recall that each complex number can also be written in terms of its absolute magnitude and a phase factor

$$z = re^{i\theta} = r\left(\cos\theta + i\sin\theta\right) \tag{1.9}$$

where

$$r = |z| = \sqrt{x^2+y^2} \tag{1.10}$$

and we get θ from the relation above, $\cos\theta = \Re(z)/r = x/r$ (Here: $\Re$(z) refers to the real-part of z).

Derivatives of complex #'s:

Similar to real numbers:

$$\frac{dz}{dz} = 1, \quad \frac{dz^2}{dz} = 2z, \ etc. \tag{1.11}$$

but note that a complex-number and its derivative should be viewed in differentiation as two independent quantities:

$$\frac{\partial z^*}{\partial z} = 0. \tag{1.12}$$

Example: differentiating $|z|^2$

$$\frac{\partial\left(|z|^2\right)}{\partial z} = \frac{\partial\left(zz^*\right)}{\partial z} = z^*, \tag{1.13}$$

while

$$\frac{\partial\left(zz^*\right)}{\partial z^*} = z. \tag{1.14}$$

Residue Theorem

An important property of complex functions is the residue theorem – we will not need it but you should read about it if you want to advance beyond what we teach.

1.2 Delta functions

A delta function (or formally a Dirac delta function) $\delta(x-x')$ is defined so that **for any function** $f(x)$ the integral over the delta function gives **the value of the function at the point** x' :

$$\int_{-\infty}^{\infty} \delta(x-x')f(x)dx = f(x'). \tag{1.15}$$

or equivalently we can shift x' to 0, i.e., a delta function simply needs to fulfill, for any f,

$$\int_{-\infty}^{\infty} \delta(x)f(x)dx = f(0). \tag{1.16}$$

As a corollary (take $f(x) = 1$):

$$\int_{-\infty}^{\infty} \delta(x-x')dx' = 1, \tag{1.17}$$

and generally we can write

$$f(x)\delta(x-x') = f(x')\delta(x). \tag{1.18}$$

To prove the last eq., integrate both sides w.r.t. x.

Examples of delta functions:

The simplest one is

$$\delta(x-x') = \frac{1}{\sqrt{\pi}\varepsilon} e^{-\frac{(x-x')^2}{\varepsilon^2}}, \quad \varepsilon \to 0. \tag{1.19}$$

Indeed, this function becomes narrower and narrower as $\varepsilon \to 0$, and its integral is 1:

$$\int \frac{1}{\sqrt{\pi}\varepsilon} e^{-\frac{(x-x')^2}{\varepsilon^2}} dx' = \frac{1}{\sqrt{\pi}} \int e^{-y^2} dy = 1 \tag{1.20}$$

where we defined $y = (x-x')/\varepsilon$, and the last equality is the Gaussian integral, which you may have seen a proof of in other classes.

Since this function is becoming narrower and narrower, and its integral is 1, we can label it as a bona fide delta-function.

> Insert: generally for any bounded function $g(x)$ such that $\int_{-\infty}^{\infty} g(x)dx = 1$,
>
> $$\delta(x) = \frac{1}{\varepsilon} g\left(\frac{x}{\varepsilon}\right), \quad \varepsilon \to 0. \tag{1.21}$$
>
> This form means that we squeeze the "g" function in x, and stretch its height, so its area is conserved; the limit of this procedure is the delta function.
>
> This procedure even applies to "crazy" oscillating functions. For example, since $\frac{1}{\pi}\int \frac{\sin(x)}{x} = 1$, it follows that
>
> $$\delta(x) = \frac{1}{\pi\varepsilon} \frac{\sin\left(\frac{x}{\varepsilon}\right)}{\frac{x}{\varepsilon}} = \frac{1}{\pi} \frac{\sin\left(\frac{x}{\varepsilon}\right)}{x}, \qquad \varepsilon \to 0. \tag{1.22}$$
>
> Note that this delta-function does not become smaller and smaller at a given x when ε is made smaller and smaller, but it just oscillates faster and faster – unless x is 0 or close to it. This is enough to make it a proper delta function, i.e., fulfilling $\int_{-\infty}^{\infty} \delta(x) f(x) dx = f(0)$ for any normal function f. (I.e., at regions away from 0 it oscillates so much so that an integral of it associated with those none-near-0 regions will vanish due to cancellation of positive and negative contributions from neighboring points.)

A related delta-function is (we won't prove):

$$\int_{-\infty}^{\infty} e^{ipx} dp = 2\pi\delta(x). \tag{1.23}$$

Note that it "makes sense"; for x=0, the integral is infinity; for $x \neq 0$, the oscillations lead to a vanishing integral.

Delta function as limit of Discretization

Another delta function emerges when we consider the limit of "discretization" of integrals, or of converting functions to lists (a function $f(x)$ represented by an infinite list of values $f(x_0), f(x_0+dx), f(x_0+2dx), ...$ and see plot below). Then:

$$\delta(x-x') = \frac{\delta_{xx'}}{dx}, \tag{1.24}$$

where the Kronecker delta $\delta_{xx'}$ is 1 when its arguments are equal, and 0 otherwise.

Note that this definition gives the desired fundamental relation of delta functions when we use it for discretized integrals (reversing here x and x'):

$$\int f(x)\delta(x-x')dx' \simeq \sum_{x'} f(x')\frac{\delta_{xx'}}{dx}dx = f(x). \tag{1.25}$$

3D

The extension of delta functions to 3D is simple,

$$\delta(\boldsymbol{r}-\boldsymbol{r'}) = \delta(x-x')\delta(y-y')\delta(z-z'), \tag{1.26}$$

so

$$\int \delta(\boldsymbol{r}-\boldsymbol{r'})f(\boldsymbol{r})d\boldsymbol{r'} = f(\boldsymbol{r}), \tag{1.27}$$

where $d\boldsymbol{r} = dx * dy * dz = d^3r$.

There are other properties of delta functions, many books discuss them.

Physical interpretation of Dirac Delta function

Physically, think of a Dirac delta function as if it is **the number density of a classical point particle.** I.e., it is 0 everywhere except for the position of the particle (where it is infinite), but its volume integral gives 1.

Of course, physically we think of delta functions not as infinitely thin and tall, but as almost-infinitely thin and tall.

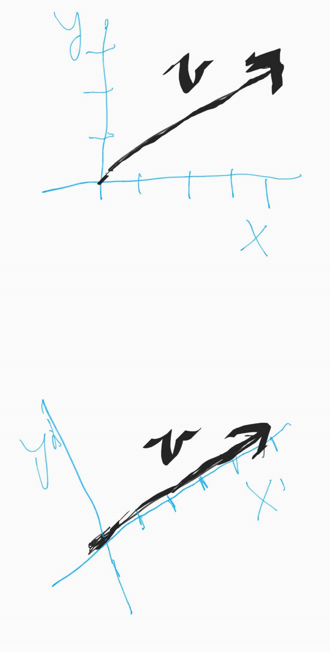
1.3 Vectors

I presume you had linear algebra, but let's recall a few basic facts and then extend them.

A vector, designated for the moment as $\boldsymbol{v}$, is an "arrow" with a specific length in a given number of dimensions. For example, in the figure below we plot a vector in 2D.

- In one coordinate system the vector has the coefficients, say, $\begin{pmatrix} 3 \\ 4 \end{pmatrix}$,meaning it is $\boldsymbol{v} = 3\hat{x} + 5\hat{\boldsymbol{y}}$, where we defined the unit vectors in that coordinate system
- If we rotate the coordinate system so the x axis lies along the direction of the vector, then the coordinates of the vector will be $\begin{pmatrix} 5 \\ 0 \end{pmatrix}$. The vector did not change, but its list-of-coefficients did. Note that in practice we often call this list of coefficients also a vector, apologies in advance for the possible confusion.

Interesting vectors in physics could be, for example, in 3D (directions in space) or 4D (space-time). However, they could also be **infinite-dimensional,** as the section below shows.



1.4 Functions as vectors

Say we are in 1D, and to simplify say we only consider a finite region of space, say $0 \leq x \leq 4$. Say we plot a function $f(x)$ where x is in this ranges

If we want to store this function, one simple way is to **discretize** space. For example say we discretize it to 401 points, spaced by $dx = 0.01$. I.e.,

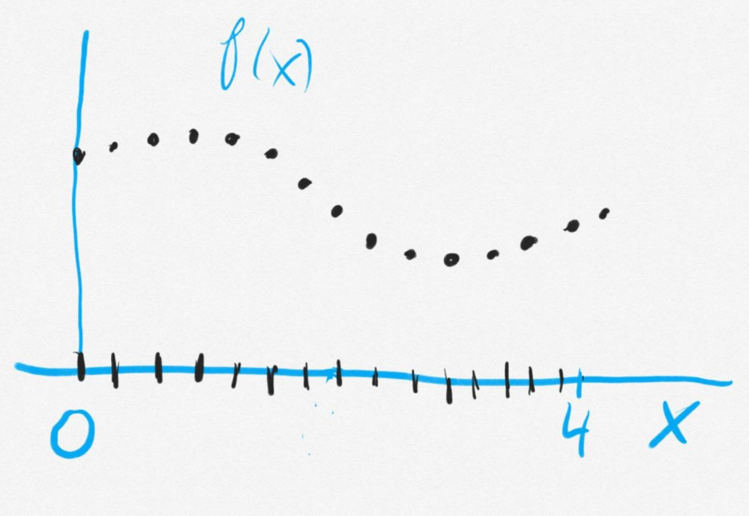
$$x = 0, 0.01, 0.02, ..., 3.99, 4.00$$

Now we can characterize the function by giving its list of values, i.e.,

$$\begin{pmatrix} f(0) \\ f(0.01) \\ f(0.02) \\ .. \\ .. \\ f(3.98) \\ f(3.99) \\ f(4.00) \end{pmatrix}. \tag{1.28}$$

Thus, we replaced the continuous function $f(x)$ by a vector (or list-of-coefficients if you want to be fancy) of length 401. Of course, if we discretized it further and further, and/or extended space to be from $x = -\infty$ to $x = \infty$ (rather than from 0 to 4), then we will need an infinitely long vector to represent f.

See also the figure below, which has a much larger grid spacing for drawing purposes.



We'll use these concepts in the next section.

2 Bra-kets, operators

2.1 Bra-kets

Denote

$$\int f^*(\boldsymbol{r})g(\boldsymbol{r})d\boldsymbol{r} = \langle f|g\rangle = \langle bra|ket\rangle. \tag{2.1}$$

The beauty of the bra-ket notation is that it applies no matter what the space is – it could be a 3D $\boldsymbol{r}$, or a single dimension x, or a vector denoting the position of multiple electrons, etc., or even the overlap could be a summation over two spin states rather than a continuous integral.

Similarly, given a linear operator A, we define

$$\langle f|A|g\rangle = \langle f|Ag\rangle. \tag{2.2}$$

i.e., the bra-ket "sandwich" of an operator A is obtained by first acting with that operator on the "ket" (here the function g), and then overlapping the result with the bra. Occasionally we'll write it explicitly as

$$\langle f|A|g\rangle = \langle f|A|g\rangle = \int f^*(\boldsymbol{r})(Ag)(\boldsymbol{r})d\boldsymbol{r}. \tag{2.3}$$

Properties of bra-kets:

- First

  $$\langle f|g\ \rangle = \langle g|f\rangle^*. \tag{2.4}$$

  Proof:

  $$\langle f|g\ \rangle^* = \left(\int f^*(\boldsymbol{r})g(\boldsymbol{r})d\boldsymbol{r}\right)^* = \int f(\boldsymbol{r})g^*(\boldsymbol{r})\,d\boldsymbol{r} = \langle g|f\rangle. \tag{2.5}$$

- Note also that if c denotes a complex number, then the function cf, when evaluated at a given point $\boldsymbol{r}$, is simply $cf(\boldsymbol{r})$. Therefore

  $$|cf\rangle = c|f\rangle. \tag{2.6}$$

- Similarly, since

  $$(cf(\boldsymbol{r}))^* = c^* f^*(\boldsymbol{r}), \tag{2.7}$$

  it follows that

  $$\langle cf| = c^*\langle f|. \tag{2.8}$$

- Finally, the norm-squared of a function is its bra-ket overlap, i.e.,

  $$\text{norm-squared}(u) \equiv \int |u(\boldsymbol{r})|^2 d\boldsymbol{r} = \int u^*(\boldsymbol{r})u(\boldsymbol{r})d\boldsymbol{r} = \langle u|u\rangle \tag{2.9}$$

  then

  $$\text{norm}(u) \equiv \sqrt{\langle u|u\rangle}.$$

Physical reasons to use the bra-ket notation

The bra-ket notation emphasizes that the "vectors" and/or "functions" (which are the same thing as we saw earlier) are abstract quantities.

Specifically, later we'll see that there are different ways to represent kets, i.e., to represent functions. (For example, we can represent functions by their Fourier coefficients, as we'll see later, or by other coefficients, etc.)

Therefore, we generalize (without actually overly changing the content) the notion that a physical system is represented by a Schrödinger wavefunction by saying that **any state of a physical system is represented by a complex ket**. **That complex ket has a norm of** 1. This means, as we'll see later, that the overall probability to find the particle (or particles if the system has several of those) **anywhere** in space is 1.

Ket-bra are operators

Another curious aspect is that **ket-bra combinations are operators.** I.e., given a bra $\langle g|$ and a ket $|f\rangle$,then $|f\rangle\langle g|$ is an operator! i.e., we can write

$$A = |f\rangle\langle g|$$

The proof is that an operator is defined as something which given a ket (i.e., a function and/or vector) returns another ket. So given any ket, $|\Psi\rangle$,the operator takes this ket and produces indeed another ket

$$A|\Psi\rangle = |f\rangle\langle g|\Psi\rangle = c|f\rangle, \tag{2.10}$$

where $c = \langle g|\Psi\rangle$. So:

- A bra dotted with a ket is a number, $\langle g|f\rangle$
- A ket first then a bra gives an operator, $|f\rangle\langle g|$.

If this looks confusing, recall that in linear algebra you learned that a column vector $\boldsymbol{u^T}$ (i.e., the transpose of a row vector), times a row vector $\boldsymbol{v}$, is a number, $\boldsymbol{u^T v}$, while a row vector times a column vector gives a matrix, $\boldsymbol{vu^T}$. Exchange the words row-vector with ket, column-vector with bra, and matrix with operator and you get what we just learned.

2.2 Adjoints of operators and Hermitian operators

Adjoint of an operator:

For any linear operator A, its adjoint $A^+$is **defined** as the operator that fulfills the following requirement: for every two functions f and g:

$$\langle f|A^+|g\rangle = \langle Af|g\rangle. \tag{2.11}$$

Explicitly, the adjoint is the operator that when working on a function on the right (the ket) and dotted by a function on the left (the bra), gives the same result as if the original operator acted on the left function. Note that

$$\langle f|A^+|g\rangle = \langle g|A|f\rangle^*. \tag{2.12}$$

Adjoint properties:

Adjoint of products:

$$(AB)^+ = B^+A^+. \tag{2.13}$$

Proof: (a little nauseating). First, recall the definition of adjoint, that says that for every two functions f, g,

$$\langle f|(AB)^+|g\rangle = \langle ABf|g\rangle \tag{2.14}$$

Now define an intermediate function $\eta = Bf$, so

$$\langle ABf|g\rangle = \langle A\eta|g\rangle = \langle \eta|A^+|g\rangle = \langle Bf|A^+g\rangle = \langle f|B^+|A^+g\rangle = \langle f|B^+A^+|g\rangle. \tag{2.15}$$

Combining the last two equations gives $\langle f|(AB)^+|g\rangle = \langle f|B^+A^+|g\rangle$, and since this is true for every f, g we get $(AB)^+ = B^+A^+$, as stipulated.

Hermitian operators

An operator W is **Hermitian if for every two functions f and g it fulfills:**

$$\langle f|W|g\rangle = \langle g|W|f\rangle^*. \tag{2.16}$$

Put differently: W **is Hermitian if it is equal to its adjoint**:

$$W = W^+ \qquad \text{(Hermitian } W\text{)}. \tag{2.17}$$

We'll see later that the Hamiltonian H and momentum p are Hermitian operators.

Hermitian operators are the natural extension to the complex plane of the concept of real symmetric operators, and share similar properties.

Property I. The first and most important property is that Hermitian operators have **real eigenvalues**; i.e., if

$$W|u\rangle = w|u\rangle, \tag{2.18}$$

(so w, $|u\rangle$ are an eigenvalue and its associated eigenvector) **then w is real, $w = w^*$.**

Proof: recall that

$$\langle u|u\rangle = \langle u|u\rangle^*, \langle u|u\rangle > 0,$$

and note that if W is a Hermitian operator, its **expectation value is real,** i.e., $\langle u|W|u\rangle = \langle u|W|u\rangle^*$ (it follows from the definition of a Hermitian operator, replace f and g by the same function u). Therefore

$$0 = \langle u|W|u\rangle - \langle u|W|u\rangle^* = \langle u|w|u\rangle - \langle u|w|u\rangle^* \\ = w\langle u|u\rangle - (w\langle u|u\rangle)^* = (w - w^*)\langle u|u\rangle. \quad (2.19)$$

so $w = w^*$, Q.E.D.

Measurements prelude: Note that all measured properties in QM are real. We'll see that measurement in Quantum Mechanics amount to expectation values (and more precisely to eigenvalues) of linear operators. This will imply that operators that can be measured are Hermitian, since those operators have real eigenvalues (we wont prove it, but it works both ways, i.e., if an operator has real eigenvalues it is Hermitian, not just the opposite). More on this later.

Property II. two eigenvectors are **automatically orthogonal if their associated eigenvalues are different,** i.e.,

$$\langle u|u'\rangle = 0 \quad \text{if } w \neq w'. \quad (2.20)$$

Proof: write the difference of matrix element of W, recalling the meaning of Hermitian operators:

$$0 = \langle u|W|u'\rangle - \langle u'|W|u\rangle^* = w'\langle u|u'\rangle - (w\langle u'|u\rangle)^* = \\ w'\langle u|u'\rangle - w^*\langle u'|u\rangle^* = (w' - w^*)\langle u|u'\rangle, \quad (2.21)$$

i.e., since w is real as proved earlier,

$$0 = (w' - w)\langle u|u'\rangle \quad (2.22)$$

so if $w - w' \neq 0$, then $\langle u|u'\rangle = 0$, *Q.E.D.*

Property III. if two eigenvectors have the same eigenvalues (what we label as **"degenerate states"**), **we can still make them (by linear combination) orthogonal**.

Example:

Consider an electron is hydrogen atom in n=2 states.

> Side note about notation. Instead of calling them as usual $\psi_{2s}, \psi_{2p_x}, \ldots$,we will refer to the kets, $|\Psi_{2s}\rangle, |\Psi_{2p_x}\rangle$ etc. Further, we often do a further shortcut and just label the kets by the eigenvector, i.e., write $|2s\rangle, |2p_x\rangle, \ldots$. This can be a very confusing notation, since in $|2s\rangle$,e.g., the "2" does not refer to multiplying the state by 2, it just refers to it being an eigenstate associated with n=2 (and $\ell = m = 0$).

Next, note that we can construct different non-orthogonal degenerate (same energy-eigenvalue) states. For example in hydrogen one *sp* state, labeled $\frac{1}{\sqrt{2}}(|2s\rangle + |2p_z\rangle)$, is not orthogonal to $|2p_z\rangle$, and they have the same energy eigenvalue ($E_{n=2} = -\frac{13.6}{2^2}$eV).

Of course we can make orthogonal combinations from these two states, e.g., $\frac{1}{\sqrt{2}}(|2s\rangle \pm |2p_z\rangle)$, or if we rather prefer $0.8|2s\rangle + 0.6|2p_z\rangle$ and $-0.6|2s\rangle + 0.8|2p_z\rangle$, or any other orthogonal linear combination.

Hermiticity and commutators:

Note that even if A, B are both Hermitian, the product is not necessarily Hermitian. I.e., if $A = A^+, B = B^+$, then

$$(AB)^+ = B^+A^+ = BA = AB + (BA - AB) = AB + [B, A], \tag{2.23}$$

where we defined the **commutator:**

$$[B, A] = BA - AB. \tag{2.24}$$

Thus, only if the Hermitian operators A and B commute (i.e., their commutator is zero) then their product is Hermitian.

Commutators and their properties:

First, the most important one:

$$[p, x] = -i\hbar. \tag{2.25}$$

Proof: for any function f

$$[p, x]f = pxf - xpf = -i\hbar \left(\frac{d}{dx}(xf) - x\frac{d}{dx}(f) \right) = -i\hbar \left(x\frac{df}{dx} + f\frac{dx}{dx} - x\frac{df}{dx} \right) = -i\hbar f. \tag{2.26}$$

Q.E.D.

Commutators of products:

$$[AB, C] = A[B, C] + [A, C]B. \tag{2.27}$$

Proof: add and subtract the same term

$$[AB, C] = ABC - CAB = ABC - ACB + ACB - CAB = A[B, C] + [A, C]B. \tag{2.28}$$

A commutator of Hermitian operators, times i, is a Hermitian operator:

$$C = i[A, B], \;\; A = A^+, B = B^+ \rightarrow C = C^+. \tag{2.29}$$

Proof:

$$\begin{aligned} C^+ = -i\left([A, B]\right)^+ &= -i(AB - BA)^+ = -i(B^+A^+ - A^+B^+) \\ &= -i(BA - AB) = i(AB - BA) = C. \end{aligned} \tag{2.30}$$

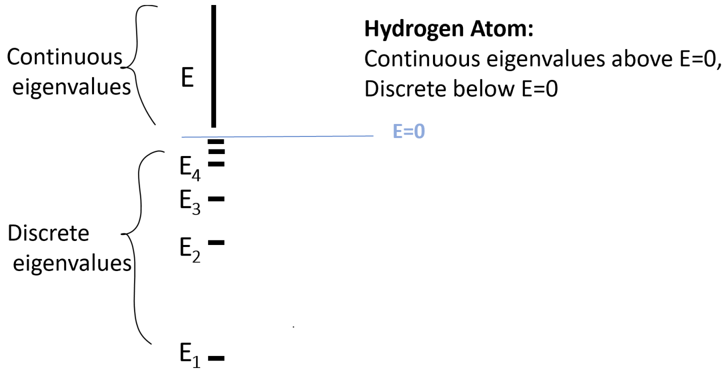
Put differently, the commutator itself (of Hermitian operators), without the extra i, is anti-Hermitian: $([A, B])^+ = -[A, B]$.

Energy levels of Hermitian operators: discrete and continuous.

Some Hermitian operators have discrete spectra – examples are the Hamiltonian of a particle in a box, or the Hamiltonian of a harmonic oscillator.

Some have continuous spectra; an example is the Hamiltonian operator for a free particle (i.e., the kinetic energy).

And some have a mixed spectra – discrete eigenvalues in some regions, continuous in others. This includes, of course, the Hamiltonian of hydrogen, or of crystals with bands and in-gap states. For example, for a hydrogen atom the discrete eigenstates would be the 1s, 2s, 2p, 3s, 3p, 3d, etc. states, and the continuous states would have $E > 0$ and would extend over all space. See picture:



Orthogonality and orthonormality:

Let's use w to denote a eigenvalue of a Hermitian operator W(usually the Hamiltonian, but could be others), and let $|u\rangle$ designate the associated eigenstate, so $W|u_w\rangle = w|u_w\rangle$. Similarly, w' will denote a possibly different eigenvalue, associated with an eigenstate $|u_{w'}\rangle$.

We proved $\langle u_w|u_{w'}\rangle = 0$ if $w \neq w'$. Now let's first specialize to discrete eigenvalues, labeled w_n, and label the eigenvector then as $|u_n\rangle$ (instead of the more cumbersome $|u_{w_n}\rangle$.

An eigenvector associated with a discrete eigenvalue will be normalizable (we wont prove that), so that we can multiply it by a constant to make its norm one, i.e., $\langle u_n|u_n\rangle = 1$. We also explained above that we can make each eigenvector orthogonal to each other.

So ignoring degenerate eigenvalues, then what we said amounts to being able to making the eigenvectors **orthonormal**:

$$\langle u_n|u_m\rangle = \delta_{n,m}. \tag{2.31}$$

What about continuous range of eigenvalue (also ignoring for now the possibility that there is degeneracy)? It turns out that then we can normalize the eigenvectors, but differently, so they fulfill:

$$\langle u_w|u_{w'}\rangle = \delta(w - w'), \tag{2.32}$$

i.e., their overlap is zero if the eigenvalues are different, and if they are the same eigenvector its norm-squared is the "proper" infinity delta function.

Of course, for a mixed spectra case (i.e., some eigenvalues are discretely distributed, some continuous), an eigenvector with a discrete eigenvalue and another eigenvector with a continuous eigenvalue must be orthogonal to each other (zero overlap), since their eigenvalues are different. I.e., if w_n is a discrete eigenvalue and w in a continuous range, then the associated eigenvectors fulfill $\langle u_n | u_w \rangle = 0$.

2.3 Completeness

Note: we are generally very flexible with notation, modifying it as we go.

Hermitian operators have complete basis sets

Recall what a complete basis set it. It is a basis which any vector (associated with the same "space") can be expanded in. For example, the unit vectors along the x and y directions, $\hat{\boldsymbol{x}}$ and $\hat{\boldsymbol{y}}$, can be used to expand any 2-D vector as $\boldsymbol{v} = v_x \hat{\boldsymbol{x}} + v_y \hat{\boldsymbol{y}}$. Similarity, for any "dimensionality", we can expand a vector in a complete basis.

Now given a Hermitian operator, W, I claim without proof that the **the set of its eigenstates is a complete basis.** This means that any function $|\Psi\rangle$ can be expanded in terms of the eigenstates.

$$|\Psi\rangle = \sum_n c_n |u_n\rangle + \int c(w) |u_w\rangle \tag{2.33}$$

where $|u_n\rangle$ are the discrete-spectrum eigenstates of W while $|u_w\rangle$ are the continuous spectrum eigenstates.

Extracting the coefficients

Let's specialize to the continuum case. Our results will be easily generalizable to the discrete and mixed-spectra cases. Then, we are claiming that any function (i.e., "ket") $|\Psi\rangle$ can be written as:

$$|\Psi\rangle = \int c(w) |u_w\rangle dw. \tag{2.34}$$

Claim: you can get the coefficient by a dot product with the eigenvector:

$$c(w) = \langle u_w | \Psi \rangle. \tag{2.35}$$

Proof (changing dummy indices in the integral from w to w'):

$$\begin{aligned} \langle u_w | \Psi \rangle = \langle u_w | \int c(w') |u_{w'}\rangle dw' &= \int c(w') \langle u_w | u_{w'} \rangle dw' \\ &= \int c(w') \delta(w - w') dw' = c(w). \quad Q.E.D. \end{aligned} \tag{2.36}$$

Note that in the proof we "inserted" the bra into the integral, that's allowed.

Norm-Squared

Claim:

$$\langle\Psi|\Psi\rangle = \sum_n |c_n|^2 + \int |c(w)|^2 dw. \tag{2.37}$$

Proof:

$$\langle\Psi|\Psi\rangle = \left(\sum_n c_n^* \langle u_n| + \int c^*(w)\langle u_w| dw\right)\left(\sum_{n'} c_{n'}|u_{n'}\rangle + \int c(w')|u_{w'}\rangle dw'\right), \tag{2.38}$$

which gives four terms which we symbolically write $\langle\Psi|\Psi\rangle = I_{dd}+I_{cc}+I_{cd}+I_{dc}$, where "d" stands for discrete (sum), and "c" is continuous (integral).

The two cross terms vanish since $\langle u_n|u_{w'}\rangle = 0$; for example

$$I_{dc} = \int \sum_{nw} c_n^* c(w)\langle u_n|u_w\rangle dw = \int \sum_n c_n^* c(w)\cdot 0 \cdot dw = 0. \tag{2.39}$$

So we're left with the $I_{dd} + I_{cc}$ terms, which give

$$\begin{aligned}\langle\Psi|\Psi\rangle &= \sum_{nn'} c_n^* c_{n'}\langle u_n|u_{n'}\rangle + \int\int c^*(w)c(w')\langle u_w|u_{w'}\rangle dwdw' \\ &= \sum_{nn'} c_n^* c_{n'}\delta_{nn'} + \int\int c^*(w)c(w')\delta(w-w')dw\,dw' \\ &= \sum_n |c_n|^2 + \int |c(w)|^2 dw. \quad Q.E.D.\end{aligned} \tag{2.40}$$

Note: $|c_n|^2$ is the probability for Ψ "to be" in a discrete state $|u_n\rangle$, while $|c(w)|^2 dw$ is the probability for Ψ "to be" in a state with energy in the range $[w, w+dw]$, i.e., $|c(w)|^2$ is the **probability density for Ψ at** w. We usually call either c_n or $c(w)$ the "amplitude" (or probability amplitude).

Bra-kets and completeness

Completeness looks nice in bra-kets (and horrendous without it)

For a general discrete-continuous basis we claim:

$$1 \equiv \text{ identity operator } = \text{(claim)} = \sum_n |u_n\rangle\langle u_n| + \int |u_w\rangle\langle u_w| dw \tag{2.41}$$

Proof: apply both sides of the eq. from above:

$$\begin{aligned}|\Psi\rangle = \text{(claim)} = \left(\sum_n |u_n\rangle\langle u_n|\right)|\Psi\rangle + \int |u_w\rangle\langle u_w|\Psi\rangle \\ = \sum_n |u_n\rangle\langle u_n|\Psi\rangle + \int |u_w\rangle\langle u_w|\Psi\rangle dw \quad (2.42)\end{aligned}$$

i.e., we get exactly the eq. from above, so the claim is valid!

$$|\Psi\rangle = \sum_n c_n|n\rangle + \int c(w)|u_w\rangle dw.\ Q.E.D. \quad (2.43)$$

with $c_n = \langle u_n|\Psi\rangle$ and $c(w) = \langle u_w|\Psi\rangle$, as needed.

3 Matrices and operators

Overall goal here: from $H|\Psi\rangle = E|\Psi\rangle$ to matrices. (Lots of this is a linear-algebra refresher, with the twist that now we consider complex operators/matrices, symmetric matrices are replaced by Hermitian ones, etc.)

3.1 Matrices from operators

Take a complete basis (discrete):

$$\sum_m |u_m\rangle\langle u_m| = 1. \tag{3.1}$$

Use this to turn

$$C = AB \ \text{(operators)} \quad \leftrightarrow C = AB \ \text{(matrices)}$$

by defining

$$C_{nm} = \langle u_n|C|u_m\rangle. \tag{3.2}$$

Of course, the choice of basis determines the matrices, while the operators are abstract, i.e., independent of basis set (i.e., independent of "representation").

Also, we're usually using the same symbol for the matrix and operator, the meaning should be clear by the context.

To show the operator-matrix relation we need to show:

$$C_{nm} = \sum_j A_{nj}B_{jm}. \tag{3.3}$$

Proof:

$$C_{nm} = \langle u_n|C|u_m\rangle = \langle u_n|AB|u_m\rangle = \sum_j \langle u_n|A|u_j\rangle\langle u_j|B|u_m\rangle = \sum_j A_{nj}B_{jm}. \tag{3.4}$$

Q.E.D.

Also: what's the matrix of the adjoint of an operator?

$$\left(A^+\right)_{nj} = \langle u_n|A^+|u_j\rangle = \langle u_j|A|u_n\rangle^* = A^*_{jn} \tag{3.5}$$

i.e., the adjoint matrix is the complex conjugate of the transpose of the original matrix.

3.2 Operators from matrices

We previously saw how to determine a matrix from an operator. The other direction is quite interesting.

To start, recall that we learned that a **ket-bra is an operator.** Thus, take a generic ket-bra: $|f\rangle\langle g|$. Given a general function $|\psi\rangle$ it operates on it as:

$$\textit{Operator on } |\psi\rangle = |f\rangle\langle g|\psi\rangle. \tag{3.6}$$

Then, given a matrix A_{ij}, and an associated basis $|u_n\rangle$, we define the operator A as

$$A = \sum_{ij} |u_i\rangle A_{ij} \langle u_j|. \tag{3.7}$$

Of course, the matrix elements of the operator A are indeed A_{ij} :

$$\langle u_i|A|u_j\rangle = A_{ij}. \tag{3.8}$$

3.3 Diagonalization

Given a vector and its eigenvectors and eigenvalues,

$$B|v_n\rangle = \beta_n |v_n\rangle, \tag{3.9}$$

then dot with $\langle u_j|$ and insert a complete basis $\sum_m |u_m\rangle\langle u_m|$ to get

$$\sum_m \langle u_j|B|u_m\rangle\langle u_m|v_n\rangle = \beta_n \langle u_j|v_n\rangle, \tag{3.10}$$

i.e.,

$$\sum_m B_{jm} V_{mn} = \beta_n V_{jn}, \tag{3.11}$$

where $B_{jm} = \langle u_j|B|u_m\rangle$, and

$$V_{mn} \equiv \langle u_m|v_n\rangle. \tag{3.12}$$

This is a matrix equation,

$$BV = V\beta, \tag{3.13}$$

where here $\beta_{ij} \equiv \beta_i \delta_{ij}$ refers to the diagonal eigenvalue matrix.

Note that if the two basis sets are orthonormal (both the $|u_j\rangle$ basis used to define the matrices and the basis $|v_n\rangle$ of eigenfunctions) then the eigenvector matrix, V_{nm}, is **unitary,** i.e., fulfills

$$V^+ V = 1. \tag{3.14}$$

Proof:

$$\begin{aligned}\left(V^+ V\right)_{nm} &= \sum_j (V^+)_{nj} V_{jm} = \sum_j V^*_{jn} V_{jm} = \sum_j \langle u_j|v_n\rangle^* \langle u_j|v_m\rangle \\ &= \sum_j \langle v_n|u_j\rangle\langle u_j|v_m\rangle = \langle v_n|v_m\rangle = \delta_{nm}. \quad \textit{Q.E.D.}\end{aligned} \tag{3.15}$$

3.4 Functions of matrices and operators

You have presumably learned in your undergrad math that a general symmetric matrix B can be diagonalized with an orthogonal matrix, R, satisfying $RR^T = 1$, as $B = R\beta R^T$ where β is diagonal (if you didn't, then get some resources from the TA).

A general (complex) Hermitian matrix can be similarly diagonalized as

$$A = V\lambda V^+ \tag{3.16}$$

or equivalently

$$AV = V\lambda \tag{3.17}$$

where V is **unitary,** i.e., its product with its adjoint is the identity matrix I

$$VV^+ = V^+V = I, \tag{3.18}$$

and as we saw earlier (when we talked about operators) the diagonal eigenvalue matrix, λ, is **real.**

Surprisingly, this leads to a simple prescription for functions of matrices/operators. Specifically, start from A^2

$$A^2 = \left(V\lambda V^+\right)\left(V\lambda V^+\right) = V\lambda\lambda V^+ = V\lambda^2 V^+. \tag{3.19}$$

Similarly, for any polynomial

$$A^n = \left(V\lambda V^+\right)\left(V\lambda V^+\right)...\left(V\lambda V^+\right) = V\lambda^n V^+. \tag{3.20}$$

But since every "reasonable" function $g(A)$ can be written as a polynomial

$$g(A) = g_0 I + g_1 A + g_2 A^2 + ... \tag{3.21}$$

we get the general relation

$$g(A) = Vg(\lambda)V^+. \tag{3.22}$$

Proof:

$$\begin{aligned} g(A) = g_0 I + g_1 A + g_2 A^2 + ... &= g_0 VV^+ + g_1 V\lambda V^+ + g_2 V\lambda^2 V^+ + ... \\ &= V\left(g_0 + g_1\lambda + g_2\lambda^2 + ...\right)V^+ = Vg(\lambda)V^+. \end{aligned} \tag{3.23}$$

So any function of an operator (or a matrix) can be replaced by the same function for the eigenvalues!

For example,

$$\frac{1}{2+A} = V \begin{pmatrix} \frac{1}{2+\lambda_1} & & & \\ & \frac{1}{2+\lambda_2} & & \\ & & ... & \\ & & & ... \end{pmatrix} V^{+}. \tag{3.24}$$

Another example: the propagation operator in quantum mechanics: $U = e^{-itH}$ (we'll see this operator a lot soon). Writing $H = V\varepsilon V^{+}$, we get

$$e^{-itH} = V \begin{pmatrix} e^{-it\varepsilon_1} & & & \\ & e^{-it\varepsilon_2} & & \\ & & ... & \\ & & & ... \end{pmatrix} V^{+}. \tag{3.25}$$

4 QM with bra-kets

4.1 The Schrödinger Equation, Hamiltonian and momentum operators

Consider a Hamiltonian (in this section we'll use 1d, the results are general):

$$H = \frac{p^2}{2M} + V, \tag{4.1}$$

where **now V means potential, not volume!** In general $V = V(x,t)$, while

$$p = -i\hbar \frac{d}{dx}. \tag{4.2}$$

The Schrödinger Eq. is then:

$$i\hbar \frac{\partial \Psi}{\partial t} = H\Psi, \tag{4.3}$$

i.e.,

$$i\hbar \frac{\partial |\Psi(t)\rangle}{\partial t} = H|\Psi(t)\rangle. \tag{4.4}$$

> **Henceforth we will assume we work in units where $\hbar = 1$.**
> So the Schrödinger equation will be
>
> $$i\frac{\partial |\Psi(t)\rangle}{\partial t} = H|\Psi(t)\rangle, \tag{4.5}$$
>
> and the momentum operator will be
>
> $$p = -i\frac{d}{dx}. \tag{4.6}$$

If $V(x,t) = V(x)$, i.e., the potential is time-independent, then we can find stationary solutions, labeled by the eigenvalue of H, i.e., if

$$|\Psi(t=0)\rangle = |u_E\rangle \tag{4.7}$$

where

$$H|u_E\rangle = E|u_E\rangle \tag{4.8}$$

then

$$|\Psi(t)\rangle = e^{-iEt}|u_E\rangle. \tag{4.9}$$

More generally, the direct solution of the Schrodinger equation for any initial state is

$$|\Psi(t)\rangle = e^{-iHt}|\Psi(t=0)\rangle. \tag{4.10}$$

Proof: differentiate the RHS:

$$i\frac{\partial|\Psi(t)\rangle}{\partial t} = i\frac{\partial e^{-iHt}|\Psi(t=0)\rangle.}{\partial t} = i(-i)He^{-iHt}|\Psi(t=0)\rangle = H|\Psi(t)\rangle, \;\; Q.E.D. \tag{4.11}$$

The Hamiltonian is Hermitian

Claim: H is Hermitian. Proof – using integration by parts:

$$\begin{aligned}\langle f|H|g\rangle = \int f^*(x)\left(-\frac{1}{2M}\frac{d^2}{dx^2} + V(x)\right)g(x)dx = \\ -\frac{1}{2M}\left(f^*(x)\frac{dg}{dx}\bigg|_{-\infty}^{\infty} - \int\frac{df^*(x)}{dx}\frac{dg}{dx}dx\right) + \int f^*(x)g(x)V(x)dx.\end{aligned} \tag{4.12}$$

The boundary term vanishes (assuming our functions decay at far distances) so

$$\langle f|H|g\rangle = \frac{1}{2M}\int\frac{df^*(x)}{dx}\frac{dg}{dx}dx + \int f^*(x)g(x)V(x)dx. \tag{4.13}$$

Clearly, if we exchange f and g and complex conjugate we'll get the same result, so H is Hermitian, $Q.E.D.$

Note: Since H is Hermitian, its eigenvalues (E) are real.

Propagation:

We can analytically propagate a w.f. if we know the eigenstates. Given $\Psi(x, t = 0)$, i.e., given $|\Psi(t=0)\rangle$, then expand the wavefunction in terms of the eigenstates of H, which are labeled by the eigenvalues and fulfill $H|u_n\rangle = E_n|u_n\rangle$ (assuming a discrete spectrum, i.e., only discrete eigenvalues)

$$|\Psi(t=0)\rangle = \sum_n c_n|u_n\rangle, \quad c_n = \langle u_n|\Psi(t=0)\rangle \tag{4.14}$$

so

$$|\Psi(t)\rangle = e^{-iHt}|\Psi(t=0)\rangle = e^{-iHt}\sum_n c_n|u_n\rangle = \sum_n c_n e^{-iE_nt}|u_n\rangle, \tag{4.15}$$

so, for 1D, e.g.,

$$\Psi(x,t) = \sum_n c_n e^{-iE_nt}u_n(x). \tag{4.16}$$

Note that here $|c_n|^2$ is the **probability** to be in a state $|u_n\rangle$ with energy E_n and it is unchanged with time (since $|c_n|^2 = |e^{-iE_nt}c_n|^2$). If the Hamiltonian has also a continuous spectrum, then $|c(E)|^2$ (for an energy E in the continuous range) the **probability density** for the initial state to be at energy E, and will also be unchanged with time.

Another Hermitian operator: p

p is Hermitian; proof – again by integration by parts:

$$\begin{aligned}\langle g|p|f\rangle^* &= \left(\int g^*(x)\left(-i\frac{df}{dx}\right)dx\right)^* = +i\int g(x)\frac{df^*}{dx}dx \\ &= i\left(g(x)f^*(x)\Big|_{-\infty}^{\infty} - \int f^*(x)\frac{dg}{dx}dx\right) = 0 - i\int f^*(x)\frac{dg}{dx}dx \\ &= \langle f|p|g\rangle. \qquad Q.E.D. \qquad (4.17)\end{aligned}$$

Eigenfunctions of p : here we ran into a **labeling nightmare**, since usually we use p for both the operator, and the eigenvalues. To avoid confusion, **we'll denote the operator below as** $\hat{p}$. **Note that the hat is meant here to designate an operator,** not as a unit vector like it does often in vector analysis. But I caution you that I will often later revert to using p (rather than $\hat{p}$) as the operator; this should eventually be clear by context.

Further, we will often denote the eigenfunctions (ket) associated with an eigenvalue p as $|p\rangle$ (rather than $|u_p\rangle$). But when we denote it as a function of x, we'll call the eigenfunctions $u_p(x)$.

Anyway, the eigenfunction equation is here then

$$\hat{p}|p\rangle = p|p\rangle, \tag{4.18}$$

or

$$-i\frac{d}{dx}u_p(x) = p \cdot u_p(x). \tag{4.19}$$

The solution is

$$u_p(x) = \frac{e^{-ipx}}{\sqrt{2\pi}}, \tag{4.20}$$

where the constant factor "normalization factor" ensures the proper normalization; i.e.,

$$\langle p|p'\rangle = \frac{1}{2\pi}\int e^{+ipx}e^{-ip'x}dx = \frac{1}{2\pi}\int e^{-i(p'-p)x}dx, \tag{4.21}$$

and using what we said about delta functions in a previous chapter, this implies

$$\langle p|p'\rangle = \delta(p-p') \quad Q.E.D. \tag{4.22}$$

The spectrum of p is continuous (only, no discrete eigenvalues).

Representation with eigenfunctions of p and Fourier transforms:

A general Ψ can be represented with eigenfunctions of p:

$$|\Psi\rangle = \int \phi(p)|p\rangle dp, \tag{4.23}$$

i.e.,

$$\Psi(x) = \int \phi(p)u_p(x)dp, \tag{4.24}$$

i.e.,

$$\Psi(x) = \frac{1}{\sqrt{2\pi}} \int \phi(p)e^{+ipx}dx. \tag{4.25}$$

Therefore, dotting the ket equation above with a bra $\langle p|$ yields (to prove, change the integration variable from p to p' above before "dotting" with the bra):

$$\phi(p) = \langle p|\Psi\rangle, \tag{4.26}$$

i.e.,

$$\phi(p) = \frac{1}{\sqrt{2\pi}} \int e^{-ipx}\Psi(x)dx. \tag{4.27}$$

The relations between $\Psi(x)$ and $\phi(p)$ called **a Fourier transform and inverse Fourier transform**. If you have not seen that, then you should consult a mathematics book. (They are akin to **Fourier sums**, which you most likely have seen, but are applicable to any function, not just periodic functions as needed for Fourier sums.)

Interpretation:

Just like

$$|\Psi(x)|^2 dx = \text{probability to find the particle in the momentum range } [x, x+dx] \tag{4.28}$$

we have

$$|\phi(p)|^2 dp = \text{probability to find the particle in the momentum range } [p, p+dp]. \tag{4.29}$$

4.2 x-Basis set

This part can be highly confusing.

Given a wavefunction $\Psi(x)$ we can act on it by x. Therefore, as we know, we can think of x as an operator (and sometimes when we do, we'll call it $\hat{\boldsymbol{x}}$). Thus, $\hat{\boldsymbol{x}}\Psi(x) = x\Psi(x)$, etc.

But if $\hat{\boldsymbol{x}}$ is an operator, it has eigenfunctions, which we'll label for now as $|x_0\rangle$ and in wavefunction form as $u_{x_0}(x)$. Those fulfill

$$\hat{\boldsymbol{x}}|x_0\rangle = x_0|x_0\rangle \tag{4.30}$$

i.e,

$$x u_{x_0}(x) = x_0 u_{x_0}(x) \tag{4.31}$$

i.e.,

$$(x - x_0)\, u_{x_0}(x) = 0 \tag{4.32}$$

Note what this means: for $x \neq x_0$, $u_{x_0}(x)$ needs to vanish, but since u_{x_0} does not have a 0 norm, this means that it must be infinity at $x = x_0$. This means that $u_{x_0}(x)$ must be a delta-functions,

$$u_{x_0}(x) = \delta(x - x_0). \tag{4.33}$$

(More precisely it needs to be proportional to a delta function, but the proportionality constant turns out to be 1.) Thus

$$\langle x|x_0\rangle = \delta(x - x_0). \tag{4.34}$$

Note that the completeness relation becomes (where we replace "x_0" by "x" as an integration variable)

$$I = \int |x\rangle\langle x| dx. \tag{4.35}$$

This gives us a clearer angle to look at wavefunctions and their value.

> (I reemphasize – $|x\rangle$ is a state that is localized at x,i.e., $\langle x'|x\rangle = \delta(x - x').$)

Specifically, **we need to think of $|\Psi\rangle$ as an abstract vector**. Such a vector can be written in terms of the "x" states as

$$|\Psi\rangle = \int |x\rangle\langle x|\Psi\rangle dx, \tag{4.36}$$

so $\langle x|\Psi\rangle$ (i.e., $\Psi(x)$), the amplitude for Ψ to be at a point x, is **the coordinate of Ψ along the "direction"** associated with the value x.

> Note: in the earlier example where we discretized space, there would be 401 values of x, and $\langle x|\Psi\rangle$ would be a "vector", i.e., list of numbers, with 401 numbers.

Of course, $|\langle x|\Psi\rangle|^2$ is the probability density to be at x.)

Or we can resolve the abstract vector Ψ in terms of the momentum eigenstates,

$$|\Psi\rangle = \int |p\rangle\langle p|\Psi\rangle dp. \tag{4.37}$$

The overlap of two functions has of course a similar form no matter which basis we calculate. Specifically, using completeness

$$\langle\Psi_1|\Psi_2\rangle = \int \langle\Psi_1|x\rangle\langle x|\Psi_2\rangle dx = \int \Psi_1^*(x)\Psi_2(x)dx, \tag{4.38}$$

and also

$$\langle\Psi_1|\Psi_2\rangle = \int \langle\Psi_1|p\rangle\langle p|\Psi_2\rangle dp = \int \Psi_1^*(p)\Psi_2(p)dp. \tag{4.39}$$

Fourier-relation revisited

We now get the Fourier-relation we got above (Eqs. (4.25),(4.27)) in a slightly more satisfying abstract manner, by writing

$$\begin{aligned}\Psi(p) = \langle p|\Psi\rangle = \int \langle p|x\rangle\langle x|\Psi\rangle dx \\ = \int \langle x|p\rangle^* \langle x|\Psi\rangle dx = \frac{1}{\sqrt{2\pi}}\int e^{-ipx}\Psi(x)dx,\end{aligned} \tag{4.40}$$

and similarly (exchange x and p)

$$\Psi(x) \quad = \quad \langle x|\Psi\rangle \quad = \quad \int \langle x|p\rangle\langle p|\Psi\rangle dp \quad = \quad \frac{1}{\sqrt{2\pi}}\int e^{ipx}\Psi(p)dp, \tag{4.41}$$

so that there's no need for a new symbol, $\phi(p)$, we just we $\Psi(p)$.

> **Be careful:** what does, e.g., $\Psi(1.38)$ mean?
>
> **Answer**: it depends on what does 1.38 refer to! For example, it could refer to momentum, position, etc.
>
> So the proper formula would be $\Psi(x = 1.38)$, or $\Psi(p = 1.38)$, etc.
>
> This is the same as realizing that a vector in 3D is well-defined, but its components along the "x" or "y" or "z" direction depend on the orientation of the coordinate system, so that if we rotate the coordinate system the components of the vector change, while the vector does not.
>
> In the same spirit, $|\Psi(1.38)\rangle$ is not meaningful unless we specify what it refers to. $|\Psi(x = 1.38)\rangle$ refers to a state that is an eigenstate of the x-operator, i.e., $\langle x'|\Psi(x = 1.38)\rangle = \delta(x - x')$,
> while $|\Psi(p = 1.38)\rangle$ refers to a plane wave, with position amplitude:
> $\langle x'|\Psi(p = 1.38)\rangle = \frac{1}{\sqrt{2\pi}}e^{ix'p} = \frac{1}{\sqrt{2\pi}}e^{ix'\cdot 1.38}$

4.3 Derivatives and Schrödinger equation in the bra-ket picture

The crucial pre-step in going from the abstract time-independent $\boxed{H|\Psi\rangle = E|\Psi\rangle}$ to the coordinate-dependent Schrödinger equation is

$$\langle x|p|\Psi\rangle = -i\frac{d}{dx}\langle x|\Psi\rangle. \tag{4.42}$$

Proof: (putting hat-operator subscript occasionally; also, there's an element of circularity in this proof, but it is still instructive).

First, note that

$$\langle x|\hat{\boldsymbol{p}}|\Psi\rangle = \int \langle x|\hat{\boldsymbol{p}}|p\rangle\langle p|\Psi\rangle dp. \tag{4.43}$$

But since, by definition of eigenstates, $\hat{\boldsymbol{p}}|p\rangle = p|p\rangle$, we write:

$$\langle x|\hat{\boldsymbol{p}}|\Psi\rangle = \int p\langle x|p\rangle\langle p|\Psi\rangle dp \tag{4.44}$$

and recalling $\langle x|p\rangle = \frac{1}{\sqrt{2\pi}}e^{ixp}$ so $p\langle x|p\rangle = -i\frac{d}{dx}\langle x|p\rangle$, we get that

$$\langle x|\hat{\boldsymbol{p}}|\Psi\rangle = -i\frac{d}{dx}\int \langle x|p\rangle \langle p|\Psi\rangle \, dp \tag{4.45}$$

i.e.,

$$\boxed{\langle x|\hat{\boldsymbol{p}}|\Psi\rangle = -i\frac{d}{dx}|\Psi\rangle}. \quad Q.E.D. \tag{4.46}$$

Based then on the previous equation, we see immediately:

$$\begin{aligned} E\langle x|\Psi\rangle = \langle x|H|\Psi\rangle = \langle x|\frac{\hat{\boldsymbol{p}}^{\boldsymbol{2}}}{2M} + V(\hat{\boldsymbol{x}})|\Psi\rangle = \frac{1}{2M}\left(-i\frac{d}{dx}\right)^2 \langle x|\Psi\rangle + \langle x|V(x)|\Psi\rangle \\ = \left(-\frac{1}{2M}\frac{d^2}{dx^2} + V(x)\right)\langle x|\Psi\rangle, \end{aligned} \tag{4.47}$$

so $\Psi(x) \equiv \langle x|\Psi\rangle$ fulfills indeed the Schrödinger equation, as we stipulated earlier on:

$$E\Psi(x) = \left(-\frac{1}{2M}\frac{d^2}{dx^2} + V(x)\right)\Psi(x) \tag{4.48}$$

4.4 Projection operators

Given a general Ψ, and a general discrete basis set $|n\rangle$,we write

$$|\Psi\rangle = \sum_n |n\rangle\langle n|\Psi\rangle = \sum_n P_n|\Psi\rangle \tag{4.49}$$

where we defined the projection operator

$$P_n = |n\rangle\langle n|. \tag{4.50}$$

Note that

$$\sum_n P_n = 1 \tag{4.51}$$

$$P_n P_m = |n\rangle\langle n|m\rangle\langle m| = \delta_{nm}|n\rangle\langle m| = \delta_{nm}P_n, \tag{4.52}$$

and therefore $P_n^2 = P_n$.

The states we defined above ($P_n = |n\rangle\langle n|$) are called projection operators to a specific state. We can also add such single-state projection operator in bunches, and such bunches will still be called projection operators. E.g., define $B_1 = P_1 + P_2 + P_3$, $B_2 = P_4 + ... + P_{10}$,..., $B_j = \sum_{\text{few } n} P_n, ...$; then we'll still have

$$\sum_j B_j = 1, \quad B_j B_k = \delta_{jk} B_j. \tag{4.53}$$

(Exercise: prove for the definition here that $B_1 B_1 = B_1$,and $B_1 B_2 = 0$).

Any set of operators that fulfills Eq. 4.53 is a set of projection operators, and each of the operators that fulfills this relation can be shown to be made of sum of several single-set projection operators.

Also, this discussion carries with some slight modifications to continuous basis sets.

4.5 The density operator: single electron

The projection operator to a specific position has a special name, $\hat{n}(\boldsymbol{r})$, the (spatial) density operator:

$$\hat{n}(\boldsymbol{r}) = |\boldsymbol{r}\rangle\langle\boldsymbol{r}|. \tag{4.54}$$

For a single electron, the expectation value of $\hat{n}(\boldsymbol{r})$ is just the electron density at the point $\boldsymbol{r}$

$$\langle\Psi|\hat{n}(\boldsymbol{r})|\Psi\rangle = \langle\Psi|\boldsymbol{r}\rangle\langle\boldsymbol{r}|\Psi\rangle = |\Psi(\boldsymbol{r})|^2. \tag{4.55}$$

Note that, as it stands, this operator is **a projection operator.** It projects a wavefunction to the state associated with the point $|\boldsymbol{r}\rangle$, i.e., what's $|\hat{n}(\boldsymbol{r})\Psi\rangle$?

$$|\hat{n}(\boldsymbol{r})\Psi\rangle = \hat{n}(\boldsymbol{r})|\Psi\rangle = |\boldsymbol{r}\rangle\langle\boldsymbol{r}|\Psi\rangle = \Psi(\boldsymbol{r})|\boldsymbol{r}\rangle, \tag{4.56}$$

i.e., it becomes a ket localized at $\boldsymbol{r}$, with an amplitude $\Psi(\boldsymbol{r})$. Therefore, when we ask what's the amplitude of $\hat{n}(\boldsymbol{r})|\Psi\rangle$ at an arbitrary point $\boldsymbol{r'}$, we are going to get a delta function:

$$\langle\boldsymbol{r'}|\hat{n}(\boldsymbol{r})\Psi\rangle = \Psi(\boldsymbol{r})\langle\boldsymbol{r'}|\boldsymbol{r}\rangle = \delta(\boldsymbol{r} - \boldsymbol{r'})\Psi(\boldsymbol{r}). \tag{4.57}$$

The density operator would be very important later when we consider **multi-electron systems (there it will not be a projection operator anymore).**

4.6 Measurements

Preliminaries

First recall: Any operator A can be written in terms of projection to its eigenstates, $|n\rangle$ (that fulfill $A|n\rangle = a_n|n\rangle$)

$$A = \sum_n a_n |n\rangle\langle n| = \sum_n a_n P_n. \tag{4.58}$$

Proof: it is enough to prove this by applying to the complete basis of eigenstates.

Thus, apply both sides on an arbitrary state $|m\rangle$.

The LHS gives:$A|m\rangle = a_m|m\rangle$, while the RHS gives

$$\sum_n a_n P_n |m\rangle = \sum_n a_n |n\rangle\langle n|m\rangle = \sum_n a_n \delta_{nm} |n\rangle = a_m|m\rangle, \tag{4.59}$$

i.e., the same – Q.E.D.!

Example: the Hamiltonian can be written, when the eigenvalues are discrete

$$H = \sum_n E_n |n\rangle\langle n|, \tag{4.60}$$

or, when the spectrum is continuous

$$H = \int |E\rangle\langle E| dE, \tag{4.61}$$

or even a mixture in a mixed continuous-discrete case.

Thus, the average energy, for example, is

$$\langle\Psi|H|\Psi\rangle = \sum_n \langle\Psi|n\rangle E_n \langle n|\Psi\rangle = \sum_n |\langle\Psi|n\rangle|^2 E_n \tag{4.62}$$

Measurements: Overview

A measuring device can be denoted as a Hermitian operator; call it A. Given a general normalized wavefunction Ψ (i.e., $\langle\Psi|\Psi\rangle = 1$) the **measurement will choose a specific eigenfunction of** A with a probability given by the overlap squared: i.e.,

$$|\Psi\rangle = \sum_a c(a)|u_a\rangle, \quad A|u_a\rangle = a|u_a\rangle \to$$

$$\text{Post-measurement, } \Psi \text{ is replaced by } \textbf{one} \text{ of the } |u_a\rangle, \text{ with probability } |c(a)|^2 = |\langle u_a|\Psi\rangle|^2. \tag{4.63}$$

Another way of saying this: Let's say the operator is the Hamiltonian. Then, measuring the "operator" (e.g., measuring the energy) means that given a normalized $|\Psi\rangle$ the **measurement device forces the wavefunction to a randomly chosen n'th energy-resolved state and that the probability to end up at the n'th state is** $|\langle\Psi|n\rangle|^2$**.** I.e.,

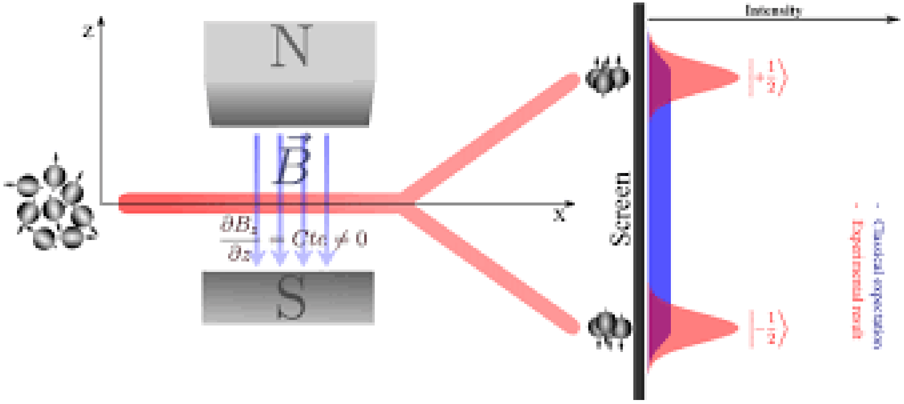
$$|\Psi\rangle \to P_n|\Psi\rangle = |n\rangle\langle n|\Psi\rangle = (final\ state) * (amplitude\ to\ reach\ it). \tag{4.64}$$

Example: the Stern Gerlach experiment.

Given a beam of atoms going along the y-direction. We'll choose atoms with a single s-shell valence electron (e.g., Ag). The beam passes through an inhomogeneous magnetic field along, say, the z-direction, the beam will be split into two.

The reason the beam is split into two is that if an electron has a spin up it will be repelled from the high field region, and if down then it will be attracted to it. Thus atoms with a single electron don't experience the same energy at different spins. They need to "choose" which direction they go through depending on the spin value.

Put differently, the Stern-Gerlach device measures the **spin** along this direction, S_z.



Note the **subtle difference** from the first-year chemistry interpretation of the Stern-Gerlach experiment. In your first-year course **you were taught that the spin are initially up and down and the beam splits it. But that's wrong!**

The correct interpretation is that the **spin is initially in a random mixture state**, i.e., it could be in **any linear-combination of "up" and "down"** states (which corresponds to being "along" a different axis – see the spin pointing along different axis in the LHS of the picture above. **Going through the measurement device** (the inhomogeneous magnet) **forces the spin to be either "up" or "down",** even though before it goes through the measurement device it isn't purely up or down.

Measurements relate to commutators

Assertion:

> If and only if $[A, B] = 0$, then the eigenvalues of A and of B can be measured simultaneously.

(Example: hydrogen atom, $A = H$, $B = L^2$, the eigenstates $\phi_{nlm}(\boldsymbol{r})$ are simultaneous eigenstates of H and L^2)

Let's prove one side, i.e., if we can find a complete set of vectors $|\Psi_{ab}\rangle$ that are eigenstates of A and of B, then A and B commute. Proof: apply AB and BA on all members of the set:

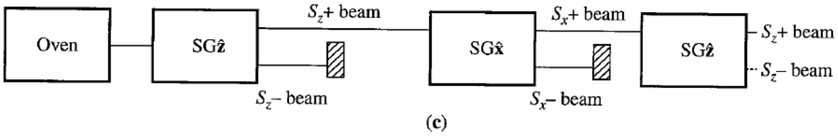
$$\begin{aligned} AB|\Psi_{ab}\rangle &= Ab|\Psi_{ab}\rangle = ab|\Psi_{ab}\rangle \\ BA|\Psi_{ab}\rangle &= Ba|\Psi_{ab}\rangle = ba|\Psi_{ab}\rangle = AB|\Psi_{ab}\rangle \end{aligned} \tag{4.65}$$

and since $|\Psi_{ab}\rangle$ are a complete set, this proves $AB = BA$, Q.E.D.

More details in Gasiorowicz.

Stern Gerlach Experiment: Details

We can now understand a more sophisticated version of the Stern Gerlach experiment (see Sakurai, 1.1, for more details; the figure is taken from there).



Sequential Stern-Gerlach experiments.

Say we run the silver-atom beam (which goes along the "y-axis") through three consecutive (rather than just one) magnets:

- The first is the same as in the original one, i.e., has an inhomogeneous field in the **z-direction**. After the atoms go through it, half will go in one direction (spin up along z) and half in the other (spin down along z). **But now we'll block the half of the beam associated with spin down along z.**
- We then ran the remaining beam through a rotated magnet, which has an inhomogeneous field in the x-direction. Therefore, half the atoms will go in one direction (those that have, or more precisely that **the magnet measured to have**, spin up along the **x-direction**), and half, those that have "spin-x down" will go in the other direction. We block again the latter beam.

- We now run the remaining beam (i.e., of atoms that were first measured to have spin up along z, and then were measured to have spin up along x) through an identical copy of the first magnet, i.e., measuring the spin along z.
  - Naively we would have expected that the atoms will still have spin up along z, since we measured them earlier to have spin up along z.
  - But instead they split to two equal beams. I.e., **the measurement of the spin along x (in the 2nd magnet) erased the memory of the measurement of the spin along z (in the 1st magnet).** In the mathematical language we used before, **the spin along z and along x are not commuting operators, so they cannot be measured simultaneously.**

We will verify our conclusions later when we learn about spins and Pauli matrices.

4.7 The evolution of expectation values.

We'll now conclude the chapter with a prelude to the next, harmonic oscillator part, by discussing the evolution of expectation values.

Consider a general operator A, either time-dependent or not. Its expectation value will generally be time-dependent:

$$\langle A\rangle = \langle\Psi(t)|A|\Psi(t)\rangle = \int \Psi^*(x,t)\,(A\Psi)(x,t)\,dx. \tag{4.66}$$

Then (omitting where convenient the explicit mentioning of the dependence on t):

$$\begin{aligned}\frac{d\langle A\rangle}{dt} &= \left\langle \frac{\partial}{\partial t}\Psi(t)\,|A|\,\Psi(t)\right\rangle + \left\langle \Psi(t)\left|\frac{\partial}{\partial t}A\right|\Psi(t)\right\rangle + \left\langle \Psi(t)\,|A|\,\frac{\partial}{\partial t}\Psi(t)\right\rangle \\ &= \langle -iH\Psi\,|A|\,\Psi(t)\rangle + \left\langle \Psi(t)\left|\frac{\partial A}{\partial t}\right|\Psi(t)\right\rangle + \langle\Psi(t)\,|A|\,(-i)H\Psi(t)\rangle \end{aligned}\tag{4.67}$$

and since H is Hermitian, we can rotate $\langle H\Psi|A|\Psi\rangle = \langle\Psi|HA|\Psi\rangle$, so (also note the sign change of the $-i$ factor when we take it out of the bra)

$$\frac{d\langle A\rangle}{dt} = +i\,\langle\Psi\,|HA|\,\Psi\rangle - i\,\langle\Psi\,|AH|\,\Psi\rangle + \left\langle\frac{\partial A}{\partial t}\right\rangle = +i\,\langle\Psi\,|[H,A]|\,\Psi\rangle + \left\langle\frac{\partial A}{\partial t}\right\rangle, \tag{4.68}$$

i.e.,

$$\boxed{\frac{d\langle A\rangle}{dt} = +i\,\langle[H,A]\rangle + \left\langle\frac{\partial A}{\partial t}\right\rangle.} \tag{4.69}$$

Example: the evolution of $\langle x\rangle, \langle p\rangle$ (Ehrenfest relations)

Consider the time dependence of $\langle x\rangle, \langle p\rangle$. The two operators x and p are time-independent, i.e., do not explicitly dependent on time (an example of an operator that depends on t would be, e.g., $A = tx + 5t^2p$).

Therefore $\frac{\partial x}{\partial t} = \frac{\partial p}{\partial t} = 0$, so inserting to the boxed relation above, we get

$$\frac{d\left\langle x\right\rangle}{dt} = i\left\langle [H,x]\right\rangle. \tag{4.70}$$

But

$$\begin{aligned} i[H,x] = i\left[\frac{p^2}{2M} + V(x), x\right] = \frac{i}{2M}\left[p^2,x\right] + i\left[V(x),x\right] = \frac{i}{2M}\left(p[p,x] + [p,x]p\right) + 0 \\ = \frac{i}{2M}\left(p(-i) + (-i)p\right) = (i)(-i)\frac{p}{M} = \frac{p}{M}, \end{aligned} \tag{4.71}$$

i.e.,

$$\boxed{\frac{d\langle x\rangle}{dt} = \frac{\langle p\rangle}{M}.} \tag{4.72}$$

This is a very enticing relation: no matter what the Hamiltonian and the initial wavefunction, t**he average velocity** (derivative w.r.t time of the average position) **is the average momentum divided by the mass. Just like classical mechanics!**

Similarly:

$$\frac{d\langle p\rangle}{dt} = i\langle [H,p]\rangle = i\left\langle\left[\frac{p^2}{2M}, p\right]\right\rangle + i\left\langle [V,p]\right\rangle = 0 + i\left\langle [V,p]\right\rangle, \tag{4.73}$$

Let's evaluate the action of the last commutator on any wavefunction:

$$i\left[V,p\right]\Psi(x) = i\left(V(x)\left(-i\right)\frac{\partial\Psi}{\partial x} - (-i)\frac{\partial V\Psi}{\partial x}\right) = V(x)\frac{\partial\Psi}{\partial x} - \frac{\partial V\Psi}{\partial x} = \left(-\frac{\partial V}{\partial x}\right)\Psi \tag{4.74}$$

i.e., the commutator is the same as $-\frac{\partial V}{\partial x}$! Therefore:

$$\boxed{\frac{d\langle p\rangle}{dt} = -\left\langle\frac{\partial V}{\partial x}\right\rangle.} \tag{4.75}$$

Again, this is a very classical-like equation.

> Note however that this expression will generally not be exactly the same as classical mechanics (the exception is Harmonic oscillator, as discussed below). In contrast, the classical mechanics expression would have been

$$\text{WRONG (Classical-like):} \qquad \frac{d\langle p\rangle}{dt} = -\frac{\partial V\left(\langle x\rangle\right)}{\partial\left\langle x\right\rangle} \tag{4.76}$$

i.e., in this classical-like expression the potential would have been evaluated only at the average position. In QM the derivative of the potential is evaluated at a range of position values, and is integrated over.

Ehrenfest relations for Harmonic Oscillator

An important special case where the Ehrenfest relations can be "closed" is H.O.(=Harmonic Oscillator):

$$V = \frac{1}{2}M\omega^2 x^2 \tag{4.77}$$

where

$$\frac{\partial V}{\partial x} = M\omega^2 x \tag{4.78}$$

so the relations become together:

$$\frac{d\langle x\rangle}{dt} = \frac{\langle p\rangle}{M},$$

$$\frac{d\langle p\rangle}{dt} = -M\omega^2\langle x\rangle, \tag{4.79}$$

i.e., just like a classical H.O.!

5 Harmonic Oscillator – algebraic approach

5.1 Raising and Lowering Operators

Recall that the (1-d) H.O. Hamiltonian is

$$H = \frac{p^2}{2M} + \frac{1}{2}M\omega^2 x^2. \tag{5.1}$$

First, simplify the notation. Define the natural width (definition justified later) of the H.O. as $\sigma = \sqrt{\hbar/(M\omega)}$. Then we redefine the momentum and coordinate:

$$\bar{p} = \frac{\sigma p}{\hbar} \qquad \bar{x} = \frac{x}{\sigma}. \tag{5.2}$$

Note that

$$[\bar{p}, \bar{x}] = -i. \tag{5.3}$$

Then, by substitution, we can easily show that

$$H = \hbar\omega\left(\frac{\bar{p}^2}{2} + \frac{\bar{x}^2}{2}\right). \tag{5.4}$$

Henceforth we'll again set $\hbar = 1$. Further, we'll remove the "bars" from x and p, but recall that from now on x and p refer to the scaled coordinates, not the original one.

We now define the lowering operator (name justified later) as

$$a = \frac{(x + ip)}{\sqrt{2}}, \tag{5.5}$$

while its conjugate is the raising operator

$$a^+ = \frac{(x - ip)}{\sqrt{2}}. \tag{5.6}$$

Now comes the key relation:

$$[a, a^+] = 1. \tag{5.7}$$

Proof:

$$[a, a^+] = \frac{1}{2}[x + ip, x - ip] = \frac{i}{2}[p, x] - \frac{i}{2}[x, p] = \frac{i}{2}(-i) - \frac{i}{2}(+i) = 1 \tag{5.8}$$

Claim:

$$\frac{p^2}{2} + \frac{x^2}{2} = a^+ a + \frac{1}{2}. \tag{5.9}$$

Proof:

$$\begin{aligned} 2a^+a = (x - ip)\,(x + ip) = \bar{x}^2 - ipx + ixp + p^2 &= x^2 - i[p,x] + \bar{p}^2 \\ &= x^2 - i(-i) + p^2 = x^2 + p^2 - 1. \end{aligned} \tag{5.10}$$

Therefore:

$$H = \omega\left(a^+a + \frac{1}{2}\right). \tag{5.11}$$

Corollaries:

$$[H,a] = \omega\left[a^+a, a\right] = \omega\left(a^+\left[a,a\right] + \left[a^+,a\right]a\right) = 0 + \omega(-1)a \tag{5.12}$$

i.e.,

$$[H,a] = -\omega a \tag{5.13}$$

and we can similarly show

$$\left[H, a^+\right] = +\omega a^+ \tag{5.14}$$

Claim: a^+ is a raising operator.

In other words, given an eigenstate $|\phi_E\rangle$ of H, fulfilling

$$H|\phi_E\rangle = E|\phi_E\rangle \tag{5.15}$$

(i.e., $(H\phi_E)(x) = E\phi_E(x)$), then we can find **a new eigenstate, by acting with a^+, and the new eigenstate is associated with a higher energy,** $E + \omega$, (i.e., $E + \hbar\omega$ if we were to keep $\hbar$).

Proof: Let's act with H on the new state, and use the commutation relation:

$$Ha^+|\phi_E\rangle = \left(a^+H + \left[H, a^+\right]\right)|\phi_E\rangle = a^+E|\phi_E\rangle + \omega|\phi_E\rangle$$

i.e.,

$$Ha^+|\phi_E\rangle = (E + \omega)\,a^+|\phi_E\rangle, \tag{5.16}$$

So the state $a^+|\phi_E\rangle$ is an eigenstate of H with an energy higher than that of ϕ_E by ω. (Hence: $a^+$is a raising operator.)

a is a lowering operator.

Similarly, $a|\phi_E\rangle$ is also an eigenstate of H, with energy $E - \omega$.

This last part is troubling, since it ostensibly means that the spectrum could have infinitely low energy (start with one eigenvalue E, and then lower it successively ("downward ladder") by applying a successively on the associated eigenvector to get $E - \omega$, $E - 2\omega$, $E - 3\omega, ...$

But this stands in contradiction to the fact, that we'll prove below, that the energy cannot be negative. Specifically, for any normalized wavefunction Ψ the expectation value of a^+a is never negative:

$$\langle\Psi|a^+a|\Psi\rangle = \langle\eta|\eta\rangle \geq 0, \;\; |\eta\rangle \equiv a|\Psi\rangle \tag{5.17}$$

and therefore ($H = \omega(a^+a + \frac{1}{2})$)

$$E_n = \langle n|H|n\rangle \geq \frac{\omega}{2} \tag{5.18}$$

Therefore, there must be a state, which we call $|\phi_{E_0}\rangle$ or even more simply $|0\rangle$, such that the "downward ladder" terminates, i.e.,

$$a|0\rangle = 0 \tag{5.19}$$

Then,

$$E_0 = \langle 0|H|0\rangle = \omega\langle 0|a^+a|0\rangle + \frac{\omega}{2} = \frac{\omega}{2}. \tag{5.20}$$

Finding the eigenstates using the raising operator.

Once we determined that the zero order state has $E_0 = \frac{\omega}{2}$,we can get the higher energy states by successive application of the raising operator. I.e., the 1st excited state (which we can label as $|\phi_{E_1}\rangle$ or simply as $|1\rangle$– we'll use the latter notation henceforth) will be obtained by applying a single raising operator:

$$E_1 = \left(1 + \frac{1}{2}\right)\omega \tag{5.21}$$

$$|1\rangle = c_1 a^+|0\rangle \tag{5.22}$$

(to understand why - recall that we proved that applying a^+ on a state raises its energy by ω).

The normalization factor c_1 will be obtained below.

We can similarly write:

$$E_2 = \left(2 + \frac{1}{2}\right)\omega \tag{5.23}$$

$$|2\rangle \propto a^+|1\rangle \tag{5.24}$$

i.e.,

$$|2\rangle = c_2\left(a^+\right)^2|0\rangle \tag{5.25}$$

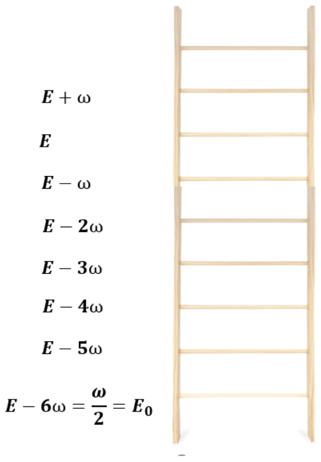
and therefore, climbing up the ladder by applying successively a^+, we get

$$E_n = \left(n + \frac{1}{2}\right)\omega, \quad n = 0, 1, 2, ... \tag{5.26}$$

$$|n\rangle = c_n\left(a^+\right)^n|0\rangle \tag{5.27}$$

Example:

The figure below shows an example – start from an arbitrary energy level; by lowering successively, applying a each time, we eventually will need to reach state the state with E_0,below which applying a yields a vanishing state. (I.e., $a|0\rangle \equiv a|u_0\rangle = 0$). In this case this was achieved after 6 states in the ladder. Since $E_0 = \frac{\omega}{2}$, this means that our original state was state $|6\rangle$ and had $E = (6 + \frac{1}{2})\omega$.



Note: we don't even know yet how $|0\rangle$looks like in coordinate space! Interestingly, we have not yet even determined how $u_{E_n}(x) \equiv \langle x|n\rangle$ (henceforth labeled $u_n(x)$) looks like! We'll determine it shortly, but the beauty of the algebraic approach is that we don't even need to know it to get lots of information.

5.2 Normalization

I will write down the normalization constant, and then prove it by induction:

Claim:

$$c_n = \frac{1}{\sqrt{n!}} \tag{5.28}$$

Proof – by induction: **First**, take the starting case, $n = 0$. Then by definition

$$\langle 0|0\rangle = 1 \tag{5.29}$$

and since

$$c_0 = \frac{1}{\sqrt{0!}} = 1$$

our claim is correct for $n = 0$.

Next, assume that the claim is correct for $n-1$,i.e., that

$$c_{n-1} = \frac{1}{\sqrt{(n-1)!}} \tag{5.30}$$

i.e., that

$$\langle n-1|n-1\rangle = \frac{1}{(n-1)!}\langle 0|a^{(n-1)}\left(a^{+}\right)^{n-1}|0\rangle = 1$$

and try to prove then that

$$\langle n|n\rangle = \frac{1}{n!}\langle 0|a^{n}\left(a^{+}\right)^{n}|0\rangle = 1. \tag{5.31}$$

Proof: The "big-picture" in the proof is that we want to use commutation relation to end up with expressions where a acts directly on $|0\rangle$, since that's 0. So let's start

First, separate

$$a^{n}\left(a^{+}\right)^{n} = a^{n-1}a\left(a^{+}\right)^{n} \tag{5.32}$$

and concentrate on the "brown" part, writing it as

$$a\left(a^{+}\right)^{n} = \left(a^{+}\right)^{n}a + \left[a,\left(a^{+}\right)^{n}\right]. \tag{5.33}$$

Let's evaluate the commutator. For this I need the insert:

Insert: what's the commutator, $[A, B^n]$ for any two operators A, B?

Answer: We can write our desired commutator as

$$[A, B^n] = ABB^{n-1} - BB^{n-1}A \tag{5.34}$$

and therefore, adding and subtracting a term:

$$[A, B^n] = ABB^{n-1} - BAB^{n-1} + BAB^{n-1} - BB^{n-1}A = [A, B]\,B^{n-1} + B\left[A, B^{n-1}\right] \tag{5.35}$$

Now we can repeat the same process one more time on the "green" commutator to get

$$\begin{aligned}[A, B^n] &= [A, B]\,B^{n-1} + B\left([A, B]\,B^{n-2} + B\left[A, B^{n-2}\right]\right)\\ &= [A, B]\,B^{n-1} + B\,[A, B]\,B^{n-2} + B^2\left[A, B^{n-2}\right]\end{aligned} \tag{5.36}$$

and we can repeat the procedure until we'll end up at our desired answer:

$$\boxed{[A,B^n] = [A,B]\,B^{n-1} + B\,[A,B]\,B^{n-2} + B^2\,[A,B]\,B^{n-3} + ... + B^{n-1}\,[A,B]} \tag{5.37}$$

In our case, A and B are a and a^+. The commutator of them is extremely simple

$$[A,B] = \left[a, a^+\right] = 1 \tag{5.38}$$

and therefore the last two equations give:

$$\left[a, \left(a^+\right)^n\right] = 1 \cdot \left(a^+\right)^{n-1} + a^+ \cdot 1 \cdot \left(a^+\right)^{n-2} + ... + \left(a^+\right)^{n-1} \cdot 1 \tag{5.39}$$

so

$$\boxed{\left[a, \left(a^+\right)^n\right] = n\left(a^+\right)^{n-1}} \tag{5.40}$$

and therefore

$$a\left(a^+\right)^n|0\rangle = \left(a^+\right)^n a|0\rangle + \left[a, \left(a^+\right)^n\right]|0\rangle = 0 + n\left(a^+\right)^{n-1}|0\rangle, \tag{5.41}$$

and therefore, plugging to our original equations

$$\langle n|n\rangle = \frac{1}{n!}\langle 0|a^n\left(a^+\right)^n|0\rangle = \frac{1}{n!}\langle 0|a^{n-1}a\left(a^+\right)^n|0\rangle = \frac{1}{n!}\langle 0|a^{n-1}n\left(a^+\right)^{n-1}|0\rangle = \frac{1}{(n-1)!}\langle 0|a^{n-1}\left(a^+\right)^{n-1}|0\rangle. \tag{5.42}$$

where we used $\frac{n}{n!} = \frac{1}{(n-1)!}$. But by our induction assumption, the last expression is just $\langle n-1|n-1\rangle$, i.e., equals 1; therefore

$$\langle n|n\rangle = \langle n-1|n-1\rangle = 1. \qquad Q.E.D. \tag{5.43}$$

5.3 Example for operators →matrices: H.O.

We can continue with the algebraic approach and derive matrix elements for the operators. In an H.O.:

$$H_{nm} = \langle n|H|m\rangle = E_n\delta_{nm} = \left(n + \frac{1}{2}\right)\omega, \tag{5.44}$$

i.e., (numbering the columns here $n = 0, 1, 2....$), the matrix of H in the H.O. eigenfunctions basis is (of course) diagonal

$$H=\frac{\omega}{2}\begin{pmatrix}1 & & & & \\ & 3 & & & \\ & & 5 & & \\ & & & ... & \\ & & & & ...\end{pmatrix}. \tag{5.45}$$

Further, let's discuss the matrix of the a, a^+ operators:

$$(a^+)_{mn}=\langle m|a^+|n\rangle, \tag{5.46}$$

But it turns out that

$$a^+|n\rangle=\sqrt{n+1}|n+1\rangle. \tag{5.47}$$

Proof: $|n\rangle=\frac{1}{\sqrt{n!}}\left(a^+\right)^n|0\rangle$, so apply

$$a^+|n\rangle=a^+\frac{\left(a^+\right)^n}{\sqrt{n!}}|0\rangle=\sqrt{n+1}\frac{\left(a^+\right)^{n+1}}{\sqrt{(n+1)!}}|0\rangle=\sqrt{n+1}|n{+}1\rangle\ \ Q.E.D. \tag{5.48}$$

so therefore

$$\left(a^+\right)_{mn}=\langle m|a^+|n\rangle=\sqrt{n+1}\langle m|n+1\rangle=\sqrt{n+1}\delta_{m,n+1}. \tag{5.49}$$

Also, there's no need to recalculate the matrix elements of a, since $\langle m|a|n\rangle=\langle n|a^+|m\rangle^*$.

In matrix form: (numbering the columns $n=0,1,2....$) a^+will be a non-diagonal, "upper-only" matrix:

$$a^+=\begin{pmatrix}0 & 0 & & & \\ \sqrt{1} & 0 & 0 & & \\ & \sqrt{2} & 0 & 0 & \\ & & \sqrt{3} & 0 & ... \\ & & & ... & ...\end{pmatrix}, \tag{5.50}$$

and the opposite for a :

$$a=\begin{pmatrix}0 & \sqrt{1} & & & \\ 0 & 0 & \sqrt{2} & & \\ & 0 & 0 & \sqrt{3} & \\ & & 0 & 0 & ... \\ & & & ... & ...\end{pmatrix}. \tag{5.51}$$

We can multiply the two infinite matrices to show that indeed the three matrices above indeed obey $\omega\left(a^+a+\frac{1}{2}\right)=H$, just like the operators do.

The matrices for x,p follow similarly:

$$x_{mn}=\langle m|x|n\rangle=\left(\langle m|a|n\rangle+\langle m|a^+|n\rangle\right) \tag{5.52}$$

i.e., in matrix form:

$$x = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & \sqrt{1} & & & \\ \sqrt{1} & 0 & \sqrt{2} & & \\ & \sqrt{2} & 0 & \sqrt{3} & \\ & & \sqrt{3} & 0 & ... \\ & & & ... & ... \end{pmatrix}. \tag{5.53}$$

Similarly,

$$p = \frac{-i}{\sqrt{2}} \begin{pmatrix} 0 & \sqrt{1} & & & \\ -\sqrt{1} & 0 & \sqrt{2} & & \\ & -\sqrt{2} & 0 & \sqrt{3} & \\ & & -\sqrt{3} & 0 & ... \\ & & & ... & ... \end{pmatrix}. \tag{5.54}$$

We can again show easily that the two infinite matrices fulfill $[p, x] = -i$.

> But: this is only true when we consider the infinite matrices; if we truncate at a finite size (say, only 100 functions, for a 100*100 matrix), it will not be true

5.4 Heisenberg Picture: Generally and for H.O.

The H.O. example is a good point to introduce a new concept: the Heisenberg and Schrödinger pictures.

The Schrödinger picture:

Our picture of time-dependent wavefunctions is called the Schrödinger picture. In that picture wavefunctions change in time, and most operators do not.

In the Schrödinger picture, the Schrödinger wavefunction fulfills

$$\frac{\partial|\Psi(t)\rangle}{\partial t} = -iH|\Psi(t)\rangle,$$

so then, for a time-independent Hamiltonian,

$$|\Psi(t)\rangle = e^{-iHt}|\Psi(0)\rangle. \tag{5.55}$$

In 115b/215b we'll generalize this solution to a time-dependent Hamiltonian.

The expectation value of any time-independent operator B is then

$$\langle B(t)\rangle = \langle\Psi(t)|B|\Psi(t)\rangle = \langle e^{-iHt}\Psi|B|e^{-iHt}\Psi\rangle = \langle\Psi(0)|e^{iHt}Be^{-iHt}|\Psi(0)\rangle. \tag{5.56}$$

The key is the realization that the only measurable property are matrix elements, not operators by themselves nor wavefunctions by themselves.

The Heisenberg Picture

In the Heisenberg picture, t**he wavefunction is a fixed object**, i.e., $|\Psi(0)\rangle$. Instead, what changes are operators, i.e., given a time-independent operator (in the Schrodinger picture) which we call B, then in the Heisenberg picture t**he operator is time dependent**

$$\text{Heiseberg picture: } B(t) \equiv e^{iHt} B e^{-iHt}, \tag{5.57}$$

From the previous two equations, we learn that the measurable property, matrix elements of

$$\langle B(t)\rangle \quad = \quad \langle\Psi(0)|B(t)|\Psi(0)\rangle \quad = \quad \langle\Psi(0)|e^{iHt} B e^{-iHt}|\Psi(0)\rangle. \tag{5.58}$$

and therefore $\langle B(t)\rangle$ is the same in both pictures! Whichever picture is used depends on convenience; for most of our discussion we'll continue using the Schroeidnger one, but you should be aware of the Heisenberg picture since it is a better starting point for approximations, necessary in physical chemistry. It is especially important for NMR.

Time-dependent equation for the evolution of operators in the Heisenberg picture.

From the definition, we see that

$$\frac{dB(t)}{dt} = \frac{d}{dt}\left(e^{iHt} B e^{-iHt}\right) = iHe^{iHt} B e^{-iHt} - e^{jHt} B e^{-iHt} iH = i\left[H, e^{iHt} B e^{-iHt}\right], \tag{5.59}$$

i.e., we reach the Heisenberg equation of motion, equivalent to the Schrödinger equation:

$$\boxed{\text{Heisenberg picture operator-evolution equation: } \frac{dB(t)}{dt} = i\left[H, B(t)\right]} \tag{5.60}$$

Example for the Heisenberg relation: Harmonic Oscillator

Preface: Products are preserved when we go from the Schroeidnger picture to the Heisenberg one, i.e., in general

$$\text{If} \quad C = AB \quad \text{then} \quad C(t) = A(t)B(t). \tag{5.61}$$

Proof:

$$C(t) = e^{+iHt} AB e^{-iHt} = e^{+iHt} A e^{-iHt} e^{+iHt} B e^{-iHt} = A(t)B(t), \quad Q.E.D. \tag{5.62}$$

As a corollary, commutation are also preserved, i.e.,

$$\textbf{If}\quad D=[A,B]\quad\textbf{then}\quad D(t)=[A(t),B(t)]\,. \tag{5.63}$$

Next, look at the H.O. Hamiltonian,

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

What we just learned is that

$$H(t)=\omega\left(a(t)a^{+}(t)+\frac{1}{2}\right) \tag{5.64}$$

But since H is time-independent and commutes with itself it would not be modified by the Heisenberg time-evolution

$$H(t)\equiv e^{+\frac{i}{\hbar}Ht}He^{-\frac{i}{\hbar}Ht}=He^{+\frac{i}{\hbar}Ht}e^{-\frac{i}{\hbar}Ht}=H, \tag{5.65}$$

so from the above we can write,

$$H=\omega\left(a(t)a^{+}(t)+\frac{1}{2}\right). \tag{5.66}$$

We also need to recall from above that $[H,a]=-\omega a$, and therefore

$$[H,a(t)]=-\omega a(t).$$

Plugging to the Heisenberg equation of motion we get:

$$\frac{da(t)}{dt}=-i[H,a(t)] \tag{5.67}$$

i.e.,

$$\boxed{\frac{da(t)}{dt}=-i\omega a(t)} \tag{5.68}$$

Note that this equation has a simple solution:

$$a(t)=e^{-i\omega t}a, \tag{5.69}$$

and conjugating gives

$$a^{+}(t)=e^{i\omega t}a^{+}. \tag{5.70}$$

The nice thing is that from these equations we can get directly the evolution of the x, p operators. Specifically, since $a=(x+ip)/\sqrt{2}$, $a^{+}=(x-ip)/\sqrt{2}$, it follows that $x=\frac{a+a^{+}}{\sqrt{2}}$, so

$$\bar{x}(t)=\frac{a(t)+a^{+}(t)}{\sqrt{2}}=\frac{e^{-i\omega t}a+e^{i\omega t}a^{+}}{\sqrt{2}}=\frac{e^{-i\omega t}\left(x+ip\right)+e^{i\omega t}\left(x-ip\right)}{2}, \tag{5.71}$$

i.e.,

$$x(t) = x\cos\omega t\ + \frac{p}{\omega}\sin\omega t, \tag{5.72}$$

and similarly we can show

$$p(t) = p\cos\omega t\ - \omega x\sin\omega t. \tag{5.73}$$

These are just the classical equations! (for scaled coordinates).

5.5 The Schrödinger wavefunctions for the H.O.

Finally, we can ask: what are the H.O. functions in space? I.e., what is $\langle x|n\rangle$?

Answer: Recall that

$$p = -i\frac{d}{dx}$$

so since

$$0 = a|0\rangle = (x+ip)\,|0\rangle = \left(x+\frac{d}{dx}\right)|0\rangle \tag{5.74}$$

so the ground-state wavefunction in scaled coordinates, $\phi_0(x) \equiv \langle x|0\rangle$ would fulfill

$$\langle x|\left(x+\frac{d}{dx}\right)|0\rangle = 0 \tag{5.75}$$

i.e.,

$$\left(x+\frac{d}{dx}\right)\phi_0(x) = 0 \tag{5.76}$$

and the solution is

$$\phi_0(x) \propto e^{-\frac{x^2}{2}} \tag{5.77}$$

and therefore when we go back to non-scaled coordinates (and recall) and normalize properly:

$$\text{Original unscaled coordinates: } \phi_0(x) = \frac{1}{\pi^{\frac{1}{4}}\sqrt{\sigma}}e^{-\frac{x^2}{2\sigma^2}} \tag{5.78}$$

where the normalization constant results from requiring $\int|\phi_0(x)|^2dx = 1$.

Similarly (we're casual on the symbols, and go back to scaled coordinates):

$$\phi_1(x) \propto \langle x|a^+\phi_0\rangle \propto \left(x-\frac{d}{dx}\right)e^{-\frac{x^2}{2}} = 2xe^{-\frac{x^2}{2}} \tag{5.79}$$

i.e.,

$$\text{Original unscaled coordinates: } \phi_1(x) = C_1xe^{-\frac{x^2}{2\sigma^2}} \tag{5.80}$$

where C_1 is a normalization constant. In general, since $a^+ \propto x - \frac{d}{dx}$, we get

$$\phi_n(x) = const * \left(x-\frac{d}{dx}\right)^n e^{-\frac{x^2}{2}}. \tag{5.81}$$

6 Angular Momentum: General

Recall classical mechanics (the definition carries over to QM):

$$\boldsymbol{L} = \boldsymbol{r} \times \boldsymbol{p}. \tag{6.1}$$

The same definition carries over to QM. For example,

$$L_x = yp_z - zp_y, \tag{6.2}$$

and the other components are defined similarly (circle $x \to y \to z \to x$).

Note that the components of $\boldsymbol{L}$ are Hermitian, e.g.,

$$L_x^+ = (yp_z)^+ - (zp_y)^+ = p_z^+ y^+ - p_y^+ z^+ = p_z y - p_y z = yp_z - zp_y = L_x, \tag{6.3}$$

etc.

Let's look at the commutators:

$$[L_x, L_y] = [yp_z - zp_y, zp_x - xp_z]\,. \tag{6.4}$$

Of all the operators in this commutator, only z and p_z don't commute, so

$$[L_x, L_y] = y\,[p_z, z]\,p_x + p_y\,[z, p_z]\,x = -i\hbar\,(yp_x - xp_y) = i\hbar L_z. \tag{6.5}$$

Similarly, (circle $x \to y \to z \to x$) :

$$[L_y, L_z] = i\hbar L_{x,}, \quad [L_z, L_x] = i\hbar L_{y,}, \tag{6.6}$$

Claim:

$$\boldsymbol{L} = i\hbar \boldsymbol{L} \times \boldsymbol{L}. \tag{6.7}$$

Proof:

$$(L \times L)_x = L_y L_z - L_z L_x = [L_y, L_z] = i\hbar L_x, \tag{6.8}$$

etc. for the other components.

Claim:

$$\left[L_z, L^2\right] = 0 \tag{6.9}$$

Proof:

$$\begin{aligned}\left[L_z, L^2\right] &= \left[L_z, L_x^2 + L_y^2 + L_z^2\right] = \left[L_z, L_x^2\right] + \left[L_z, L_y^2\right] + 0 \\ &= L_x\,[L_z, L_x] + [L_z, L_x]\,L_x + L_y\,[L_z, L_y] + [L_z, L_y]\,L_y \\ &= i\hbar\,(L_x L_y + L_y L_x) - i\hbar\,(L_x L_y + L_y L_x) = 0\end{aligned} \tag{6.10}$$

and the same for the other components. Thus, we can take common eigenfunctions of L_z and L^2 (since they commute).

Side-note: we mentioned that operators that don't commute can't share eigenfunctions. Let's see here an example – L_x and L_z. Say they did, i.e., there was an eigenfunction

$$L_x|u\rangle = l_x|u\rangle \tag{6.11}$$

$$L_z|u\rangle = l_z|u\rangle. \tag{6.12}$$

Thus:

$$\left[L_x, L_z\right]|u\rangle = -i\hbar L_y|u\rangle, \tag{6.13}$$

$$\left(l_x l_z - l_z l_x\right)|u\rangle = -i\hbar L_y|u\rangle \tag{6.14}$$

i.e.,

$$0 = L_y|u\rangle, \tag{6.15}$$

and therefore

$$\begin{aligned} l_x|u\rangle = L_x|u\rangle = -\frac{i}{\hbar}\left[L_y, L_z\right]|u\rangle = -\frac{i}{\hbar}\left(L_y L_z - L_z L_y\right)|u\rangle = \\ -\frac{i}{\hbar}\left(L_y l_z - l_z \cdot 0\right)|u\rangle = -\frac{i}{\hbar} l_z L_y|u\rangle = 0 \end{aligned} \tag{6.16}$$

i.e., $l_x = 0$; similarly it follows that $l_z = 0$. Thus: it is an S state. Any other state cannot be an eigenstate of both L_x and L_z.

End of side note.

So we'll designate states by the eigenvalues of L^2 and L_z :

$$L^2|l, m\rangle = \hbar^2 l(l+1)|l, m\rangle, \tag{6.17}$$

$$L_z|l, m\rangle = \hbar m|l, m\rangle. \tag{6.18}$$

We know from undergrad QM that these $|l, m\rangle$ state are spherical harmonics, and that l, m are both integers; we'll prove it now with operator methods.

The proof is quite similar to that of H.O. Again we define raising and lowering operators (names justified later):

$$L_+ = L_x + iL_y \tag{6.19}$$

$$L_- = L_x - iL_y. \tag{6.20}$$

Since L_x, L_y, L_z are all Hermitian, L_+ and $L_-$are adjoints of each other: $(L_+)^+ = L_-$.

We then form a few important commutators involving these raising/lowering operators:

$$[L_+, L_-] = [L_x + iL_y, L_x - iL_y] = i\,[L_y, L_x] - i\,[L_x, L_y] = -2i(i\hbar)L_z = 2\hbar L_z \tag{6.21}$$

while

$$[L_z, L_+] = [L_z, L_x + iL_y] = [L_z, L_x] + i\,[L_z, L_y] = i\hbar L_y + i(-i\hbar)L_x = \hbar L_+ \tag{6.22}$$

and similarly

$$[L_z, L_-] = -\hbar L_-. \tag{6.23}$$

The last two eqs. are reminiscent of $[H, a^+] = \hbar\omega a^+$ in the H.O. case, and that's why these operators got the same title (raising/lowering). Another important relation is

$$\left[L_+, L^2\right] = 0. \tag{6.24}$$

Proof: I leave it up to you, it is too painstaking.

Claim: (alluded to above): L_+ raises the eigenvalue of L_z by $\hbar$, i.e., raises m by 1. (To simplify the discussion below, **henceforth $\hbar = 1$.)** Specifically, given an eigenstate

$$L_z|l,m\rangle = m|l,m\rangle, \tag{6.25}$$

we note that

$$L_zL_+|l,m\rangle = (L_+L_z + [L_z, L_+])\,|l,m\rangle = (L_+m + L_+)\,|l,m\rangle = (m+1)L_+|l,m\rangle, \tag{6.26}$$

so $L_+|l,m\rangle$ is indeed an eigenstate of L_z with an eigenvalue $m+1$.

Similarly, $L_-|l,m\rangle$ is an eigenstate of L_z with an eigenvalue $m-1$. So we need to limit the "ladder", otherwise $|m|$ would go to infinity.

In a familiar type of proof (following the H.O. example), we first define $|\eta\rangle = L_-|l,m\rangle$. Then

$$0 \le \langle\eta|\eta\rangle = \langle l,m|L_+L_-|l,m\rangle. \tag{6.27}$$

For a similar reason, the expectation value of L_x^2,L_y^2, L_z^2 are non-negative, and therefore the expectation value of L^2, i.e.,

$$l(l+1) \ge 0.$$

which means $l > 0$ or $l < -1$. And we can always choose $l > 0$ (if $l < -1$, replace it by $-(l+1)$; they both give the same $l(l+1)$.)

In addition:

$$L_+L_- = (L_x+iL_y)(L_x-iL_y) = L_x^2+L_y^2-i[L_x,L_y] = L^2-L_z^2+L_z \qquad Q.E.D. \tag{6.28}$$

Similarly

$$L_-L_+ = L^2 - L_z^2 - L_z. \tag{6.29}$$

Putting what we learned together we learn that

$$0 \leq \langle l,m|L_+L_-|l,m\rangle = \langle l,m|L^2 - L_z^2 + L_z|l,m\rangle \tag{6.30}$$

i.e.,

$$0 \leq l(l+1) - m^2 - m = l(l+1) - m(m+1) \tag{6.31}$$

Similarly, replacing L_+L_- by L_-L_+ above, we get

$$0 \leq l(l+1) - m(m-1), \tag{6.32}$$

and together both equations give

$$-l \leq m \leq l. \tag{6.33}$$

Since m varies by integers, this implies that

$$2l = \text{integer}, \tag{6.34}$$

so m it at least half-integer. Examples:

$$l=3,\ \ m=-3,-2,...,3$$

$$l=2.5,\ \ m=-2.5,-1.5,...,2.5.$$

Expectation values

Classically a vector is aligned, so if it is aligned along z then the x and y components vanish. But this is not true in QM. Even if a state has $m_z = l$, still $\langle L_x^2 + L_y^2\rangle > 0$ (except for S-states). Specifically:

$$\langle l,l|L_x^2+L_y^2|l,l\rangle = \langle l,l|\boldsymbol{L^2} - L_z^2|l,l\rangle = l(l+1) - l^2 = l \tag{6.35}$$

Interestingly, as we raise l we approach the classical limit. Specifically, the relative uncertainty in the most-aligned state decreases, where we define this relative uncertainty as

$$\text{Uncertainty in alignment} = \frac{\langle l,l|L_x^2+L_y^2|l,l\rangle}{\langle l,l|\boldsymbol{L^2}|l,l\rangle} = \frac{l}{l(l+1)} = \frac{1}{(l+1)} \to 0 \text{ as } l \!\to \infty. \tag{6.36}$$

Spherical Harmonics

We can now get the angular momentum eigenfunctions, the spherical harmonics which you know (for integer l; for half integer we'll need spin, later).

Specifically, using the concept of solid angle:

$$y_{lm}(\theta,\phi) = \langle\theta,\phi|l,m\rangle \equiv \langle\Omega|l,m\rangle. \tag{6.37}$$

We need also the solid-angle integration, using $d\Omega = \sin(\theta)d\theta d\phi$

$$1 = \int d\Omega|\theta,\phi\rangle\langle\theta,\phi| = \int_0^{2\pi} d\phi \int_0^{\pi} \sin(\theta)d\theta|\theta,\phi\rangle\langle\theta,\phi|. \tag{6.38}$$

This integration weight implies that the bra-ket overlap must be

$$\langle\theta,\phi|\theta',\phi'\rangle = \frac{1}{\sin(\theta)}\delta(\theta-\theta')\delta(\phi-\phi') \tag{6.39}$$

(Proof: apply the eq. before the last to $|\theta',\phi'\rangle$.)

Also, using the definition of the angles in terms of x, y, z and the associated derivatives w.r.t. ϕ,θ, it can be shown (as was done presumably in undergrad QM) that

$$L_z = -i\frac{\partial}{\partial\phi}. \tag{6.40}$$

This immediately implies that

$$-i\frac{\partial}{\partial\phi}y_{lm}(\theta,\phi) = \langle\theta,\phi|L_z|l,m\rangle = m\langle\theta,\phi|l,m\rangle = my_{lm}(\theta,\phi), \tag{6.41}$$

so

$$y_{lm}(\theta,\phi) = y_{lm}(\theta,0)e^{im\phi}. \tag{6.42}$$

In addition, we get the theta dependence of y_{lm} by first considering $y_{l,m=l}$ and then applying the lowering operators successively to lower m. It is straightforward to show that

$$L_\pm = e^{i\phi}\left(\pm\frac{\partial}{\partial\theta} + i\cot(\theta)\frac{\partial}{\partial\phi}\right), \tag{6.43}$$

so noticing that raising the highest m orbital leads to a vanishing state, we get

$$0 = \langle\theta,\phi|L_+|l,m=l\rangle = e^{i\phi}\left(\frac{\partial}{\partial\theta} + i\cot(\theta)\frac{\partial}{\partial\phi}\right)y_{ll}(\theta,\phi) = e^{i\phi}\left(\frac{\partial}{\partial\theta} + i\cot(\theta)(il)\right)y_{ll}(\theta,\phi) \tag{6.44}$$

where we used $y_{ll}(\theta,\phi) = e^{il\phi}y_{ll}(\theta,0)$. Thus,

$$\left(\frac{\partial}{\partial\theta} - l\cot(\theta)\right)y_{ll}(\theta,0) = 0, \tag{6.45}$$

i.e.,

$$y_{ll}(\theta, \phi = 0) = \text{const.} * (\sin(\theta))^l \,. \tag{6.46}$$

The other y_{lm} follow by applying successively L_-,i.e.,

$$y_{lm}(\theta, \phi) \propto (L_-)^{(l-m)} y_{ll}(\theta, \phi). \tag{6.47}$$

See Gasiorowicz Chap. 7 for details.

Another example for operators → matrices: Angular momentum

$$\langle l, m | \boldsymbol{L^2} | l', m' \rangle = l(l+1)\delta_{ll'}\delta_{m'm} \tag{6.48}$$

$$\langle l, m | L_z | l', m' \rangle = m\delta_{ll'}\delta_{m'm} \,. \tag{6.49}$$

We will now form specific matrices, **each for a different l,** that represent the angular momentum components for that specific l. Thus,

$$L_+ | l, m \rangle = c | l, m+1 \rangle \tag{6.50}$$

where $c = c^+(l, m)$ is a function of l, m, and is obtained by noting (repeating the previous derivation)

$$|c|^2 = |c|^2 \langle l, m+1 | l, m+1 \rangle = \langle l, m | L_- L_+ | l, m \rangle = \\ \langle l, m | \left(L^2 - L_z^2 - L_z \right) | l, m \rangle = l(l+1) - m(m+1) \tag{6.51}$$

i.e., (we can make c real and positive):

$$\langle l, m' | L_+ | l, m \rangle = \delta_{ll'}\delta_{m'm+1}\sqrt{l(l+1) - m(m+1)} \tag{6.52}$$

Example: $l = 1$

Let's consider the matrices for $l = 1$. Historically we label the indices in the matrix in the order $m = 1, 0, -1$ (top to bottom)

Then:

$$L_z = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} \tag{6.53}$$

$$L_+ = \begin{pmatrix} 0 & \sqrt{2} & 0 \\ 0 & 0 & \sqrt{2} \\ 0 & 0 & 0 \end{pmatrix} \tag{6.54}$$

$$L_- = (L_+)^+ = \begin{pmatrix} 0 & 0 & 0 \\ \sqrt{2} & 0 & 0 \\ 0 & \sqrt{2} & 0 \end{pmatrix} \tag{6.55}$$

so

$$L_x = \frac{1}{2}(L_+ + L_-) = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \tag{6.56}$$

while

$$L_y = -\frac{i}{2}(L_+ - L_-) = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix}. \tag{6.57}$$

(Note that L_y is Hermitian!)

Note that the matrices of the angular momentum operators obey the same relation as the operators themselves. It is straightforward to check that the three matrices defined above fulfill $[L_x, L_y] = iL_z$, etc.

I reemphasize that the matrices we defined above are only for $l = 1$, i.e., involve the subspace of three spherical harmonics with $l = 1$ and different m. In principle we should have labeled these matrices with an $l = 1$ superscript, but it would be too cluttered.

Of course, for each different l we'll have three such matrices that are of the order $(2l+1) \times (2l+1)$, and fulfill the commutation relations, $[L_x, L_y] = iL_z$, etc.

Lie-Algebras

In mathematics, a finite group of operators $I_1, I_{2,}, ..., I_M$ such that their commutators are linear combination of each other are called **Lie-algebras,** i.e., they fulfill

$$[I_i, I_j] = i\sum_{k=1}^{M} C_{i,j}^{k} I_k \tag{6.58}$$

The angular momentum operators are Lie-algebras ($M = 3$, $C_{xy}^{z} = 1$, etc.)

Symmetry

Lie-algebras are important for symmetry operations. In the context of angular momentum, we can show that upon a rotation by an angle γ around any axis denoted by its direction $\hat{c}$ (here, unlike the rest of the notes, a "hat" means a unit vector, not an operator), then a wavefunction changes as

$$|\psi\rangle \to R(\gamma, \hat{\boldsymbol{c}})|\psi\rangle = e^{i\gamma L\cdot\hat{c}}|\psi\rangle \tag{6.59}$$

where R is the rotation operator, and $\boldsymbol{L}\cdot\hat{\boldsymbol{c}} = c_x L_x + c_y L_y + c_z L_z$.

To motivate the proof of the last equation (without fully proving it), start from an infinitesimally small amount we'll call δ_z. When we rotate the wavefunction by, say, δ_z =0.01 radian around the z-axis, than whatever value it has at ϕ is the value it has earlier at $\phi - 0.01$ (see class). Thus

$$R\psi(\theta,\phi) = \langle\theta,\phi|R(\delta_z,\hat{z})\psi\rangle = \psi(\theta,\phi-\delta_z) \tag{6.60}$$

But (in casual notation)

$$\begin{aligned}\psi(\theta,\phi-\delta_z) = \psi(\theta,\phi) - \delta_z\cdot\frac{\partial\psi}{\partial\phi} &= \left(1-\delta_z\frac{\partial}{\partial\phi}\right)\psi \\ &= \left(1+\delta_z\cdot i^2\frac{\partial}{\partial\phi}\right)\psi = (1+i\delta_z\cdot L_z)\,\psi \simeq e^{i\delta_z\cdot L_z}\psi \end{aligned} \tag{6.61}$$

This was for a tiny rotation around one axis (the z axis). Turns out that for a tiny rotation by δ_x around the x-axis the rotation operator is, of course, $e^{i\delta_x\cdot L_x}|\psi\rangle$, and analogously for the y-axis.

We can then combine infinitesimal rotation around the x, y, z axis to make an infinitesimal rotation around our desired $\hat{\boldsymbol{c}}$ axis. When we do it, the crucial step it to prove that combining two small rotations is the same as one small rotation around the combined axis. e.g.,

$$e^{i\delta_z L_z}e^{i\delta L_z} = e^{i\delta_z L_z + i\delta_x L_x} + O(\delta^2). \tag{6.62}$$

To first order the proof is trivial.

Once we combine the rotations we rotate a little bit (by $\delta\gamma$) around some average axis, i.e., we write

$$R(\delta\gamma,\hat{\boldsymbol{c}}) = e^{i\delta\gamma L\cdot\hat{c}}, \tag{6.63}$$

we can do the tiny rotations around the same axis many times (see class), to end up at

$$R(\gamma,\hat{\boldsymbol{c}}) = R(\delta\gamma,\hat{\boldsymbol{c}})R(\delta\gamma,\hat{\boldsymbol{c}})...R(\delta\gamma,\hat{\boldsymbol{c}}) = e^{i\delta\gamma L\cdot\hat{c}}e^{i\delta\gamma L\cdot\hat{c}}...e^{i\delta\gamma L\cdot\hat{c}} = e^{i\gamma L\cdot\hat{c}} \quad Q.E.D. \tag{6.64}$$

Representations

The spherical harmonics are, for each l, a finite set of functions such that any rotation mixes them up among themselves (at the same l and different m).

For example, take $l = 2$. Rotating by 90 degrees around the y-axis turns the five d-spherical Harmonics into a combination of the original five (can you work it out?).

Such sets are called representations of the rotation group. We will next see, when we talk about spin, that there are also such representations of which are half-integer (have l which is half-integer).

7 Spin

Spin (or more precisely spin half) is angular momenta with $l = \frac{1}{2}$, so two m values ($m = \pm\frac{1}{2}$).

Single-stage Stern-Gerlach experiment

To remind you, the original Stern Gerlach experiment proved that an electron has a property with two values. Specifically, an inhomogeneous field is applied along, say, the z axis, and a beam of neutral Ag atoms with one valence s-electron (with zero "regular" spatial momentum) is split by the field – half go down, repelled from the field, and half go up.

This is consistent with the fact that in classical mechanics an object with an angular momentum of $\boldsymbol{L}$ will interact with a magnetic field $\boldsymbol{B}$ with an interaction potential

$$-\, const. * (\boldsymbol{L} \cdot \boldsymbol{B}) \tag{7.1}$$

where the constant is proportional to the inverse mass (of the valence s-electron in our case). Thus, if the angular momentum is along the magnetic field, the interaction is negative so the electron (and therefore the Ag atom) would be attracted to the field and will turn up to regions where the field is stronger, and if the spin is opposite to the field the electron will be repelled.

But we just said that the valence electron in Ag is an s-electron, with zero spatial angular momentum; so this property must be a different angular momentum (not related to the spatial L) which we call spin, $\boldsymbol{S}$.

Multi-stage Stern Gerlach experiment (following Sakurai, Modern QM)

The multi stage S.G. experiment proved that different components of the spin cannot be specified simultaneously, i.e., the spin along z and along y (or x) cannot be specified independently.

Three stages:

- First an inhomogeneous field along the z axis; the beam is split to two, one with $S_z = \frac{\hbar}{2}$, the other with $S_z = -\frac{\hbar}{2}$ (we'll occasionally use $\hbar$ explicitly, before setting it back to $\hbar = 1$). We then block the beam with $S_z = -\frac{\hbar}{2}$.
- Then we pass the remaining Ag atoms (that have $S_z = \frac{\hbar}{2}$) through a region with an inhomogeneous field pointing along the y axis; half the atoms bend in the positive y axis, half in the negative, showing that we have two beams with $S_y = \pm\frac{\hbar}{2}$. We block the beam with $S_y = -\frac{\hbar}{2}$, so only $S_y = \frac{\hbar}{2}$ atoms remain.
- We now pass the atoms through a third stage, similar to the first, i.e., splitting a beam according to S_z. Now IF the z-component and the y-component of the spin could have been specified together, then as a result

> of the 2nd stage the valence electron would have simultaneously $S_z = \frac{\hbar}{2}$ (from the first stage) and $S_y = \frac{\hbar}{2}$ (from the 2nd stage). So in that case all atoms would have gone up (along the positive z-axis).

But instead, after the third stage we get the same result as after the first, i.e., the beam splits into two! Half with spin up and half with spin down! Thus, specifying S_y in the 2nd stage ERASED the memory of specifying $S_z$in the first stage.

So to conclude: the components of the spin do not commute. **We can't specify S_y and S_z simultaneously.**

This is consistent with the fact that spin is an angular momentum, so it fulfills:

$$\boldsymbol{S} \times \boldsymbol{S} = i\hbar \boldsymbol{S} \tag{7.2}$$

i.e.,

$$[S_x, S_y] = i\hbar S_z, \tag{7.3}$$

etc. for the other components.

Spin half and Pauli spin matrices

A spin-half (i.e., $l = \frac{1}{2}$) ket cannot be represented with spatial functions (one s, three p's, five d's, etc.). Instead we need to use the matrix representation (analogous to the 3*3 matrix rep. for $l = 1$). Here, it will be 2*2 matrices.

Replacing L by S, and numbering the matrix again top to bottom, i.e., the first column is $m = \frac{1}{2}$ and the 2nd column is $m = -\frac{1}{2}$, gives

$$S_z = \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{7.4}$$

Further, applying the general rules on what L_+ and L_- should be for $l = 2$, and therefore what L_x and L_y should be, we get that

$$S_+ = \hbar\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad S_- = (S_+)^+ = \hbar\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \tag{7.5}$$

(the "1" in S_+ is $\sqrt{l(l+1) - m(m+1)}$ for $m = -\frac{1}{2}$, i.e, $\sqrt{\frac{1}{2}\cdot\frac{3}{2} - \left(-\frac{1}{2}\right)\frac{1}{2}}{=}1$). Therefore:

$$S_x = \frac{S_+ + S_-}{2} = \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{7.6}$$

$$S_y = \frac{S_+ - S_-}{2i} = \frac{\hbar}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}. \tag{7.7}$$

For simplicity, we abbreviate

$$\boldsymbol{S} = \frac{\hbar}{2}\boldsymbol{\sigma}, \tag{7.8}$$

where we defined the Pauli spin matrix,

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{7.9}$$

$$\sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \tag{7.10}$$

$$\sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{7.11}$$

Note that the components of S and the Pauli spin matrices are all Hermitian (prove it for σ_y!).

Due to the factor of two between **S** and the Pauli spin matrices (in addition to $\hbar$), the latter fulfill

$$\boldsymbol{\sigma} \times \boldsymbol{\sigma} = \mathbf{2}i\boldsymbol{\sigma} \tag{7.12}$$

i.e.,

$$[\sigma_x, \sigma_y] = 2i\sigma_z, \tag{7.13}$$

etc.

Interestingly (and this is a particular feature to $l = \frac{1}{2}$) the Pauli spin matrices anticommute:

$$\sigma_x\sigma_y = -\sigma_y\sigma_x \tag{7.14}$$

etc. for the other two pairs. (Prove it!)

Spin-half eigenstates

The eigenstates of any Pauli matrix are length-two vectors, labeled spinors. The eigenstates of σ_z are labeled

$$\chi^+ = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{7.15}$$

$$\chi^- = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{7.16}$$

> Side note: in chemistry (NOT IN Gasiorowicz) these are often labeled instead, in casual writing, as a spin-up vector
>
> $$|\alpha\rangle = \chi^+,$$
>
> (the components of $|\alpha\rangle$ when measured along the z axis are 1 and 0), and similarly the spin-down
>
> $$|\beta\rangle = \chi^-.$$
>
> We'll use that notation later, but not in the next page or two.
>
> End of note)

It is trivial to note that $\sigma_z\chi^{\pm} = \pm\chi^{\pm}$, so $S_z\chi^{\pm} = \pm\frac{\hbar}{2}\chi^{\pm}$.

We can similarly ask what are the spin eigenstates along different directions; e.g., eigenstates of S_x or S_y, or even of a general vector. Specifically, given a normalized length-3 vector

$$\hat{\boldsymbol{n}} = (n_x, n_y, n_z), \quad \hat{n}\cdot\hat{n} = 1, \tag{7.17}$$

we can define the "Pauli spin matrix along the direction $\hat{\boldsymbol{n}}$"

$$\sigma_{\hat{\boldsymbol{n}}} = \hat{\boldsymbol{n}}\cdot\boldsymbol{\sigma} = n_x\sigma_x + n_y\sigma_y + n_z\sigma_z = \begin{pmatrix} n_z & n_x - in_y \\ n_x + in_y & -n_z \end{pmatrix} \tag{7.18}$$

and each of its eigenstates is a 2-component spinor denoted as

$$\begin{pmatrix} \gamma_+ \\ \gamma_- \end{pmatrix}. \tag{7.19}$$

I use $\gamma_{\pm}$ instead of $\alpha^{\pm}$ like Gasiorowicz, since α is used in chemistry for spin up. The eigenstates (the spinors) are normalized, $|\gamma_+|^2 + |\gamma_-|^2 = 1$.

Example Eigenstates of spin pointing in the x-y plane

Let's take an example, the eigenstates of a spin aligned along a vector in the x-y plane, at an angle ϕ to the x axis:

$$\hat{\boldsymbol{n}} = (\cos\phi, \sin\phi, 0) \tag{7.20}$$

so

$$\hat{\boldsymbol{n}}\cdot\boldsymbol{\sigma} = n_x\sigma_x + n_y\sigma_y = \begin{pmatrix} n_z & n_x - in_y \\ n_x + in_y & -n_z \end{pmatrix} = \begin{pmatrix} 0 & \cos\phi - i\sin\phi \\ \cos\phi + i\sin\phi & 0 \end{pmatrix} = \begin{pmatrix} 0 & \exp(-i\phi) \\ \exp(i\phi) & 0 \end{pmatrix}. \tag{7.21}$$

The normalized eigenstates are labeled as $\begin{pmatrix} u \\ v \end{pmatrix}$ and fulfill

$$\begin{pmatrix} 0 & \exp(-i\phi) \\ \exp(i\phi) & 0 \end{pmatrix}\begin{pmatrix} u \\ v \end{pmatrix} = \lambda\begin{pmatrix} u \\ v \end{pmatrix}, \tag{7.22}$$

so

$$\exp(-i\phi)v = \lambda u \tag{7.23}$$

$$\exp(i\phi)u = \lambda v. \tag{7.24}$$

Multiply the last two eqs. together to get

$$uv = \lambda^2 uv. \tag{7.25}$$

We know that neither u nor v are 0, (if u was 0 then by the eqs. above v=0 so the vector was 0, but it needs to have a unit norm). Therefore, we can divide the eq. above by uv, to get that the eigenvalue of $\hat{\boldsymbol{n}} \cdot \boldsymbol{\sigma}$ is

$$\lambda = \pm 1. \tag{7.26}$$

So the eigenvalue of the spin ($\hat{\boldsymbol{n}} \cdot \boldsymbol{S}$) along this axis is $\pm\frac{\hbar}{2}$. Note that this conforms to what we said about spin: the projection of the spin along any axis is always up or down, and has the same value regardless of the axis.

What about the eigenvectors themselves? First, since $|u| = |v|$, and the norm-squared is 1, it follows that

$$|u| = |v| = \frac{1}{\sqrt{2}}. \tag{7.27}$$

If $\lambda = 1$,then

$$\exp(-i\phi)v = u, \tag{7.28}$$

and we can write the up-eigenvector then

$$|\eta^{+}(\phi)\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} \exp\left(-\frac{i}{2}\phi\right) \\ \exp\left(+\frac{i}{2}\phi\right) \end{pmatrix}. \tag{7.29}$$

Similarly, the down-eigenvector is easily shown to be

$$|\eta^{-}(\phi)\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} \exp\left(-\frac{i}{2}\phi\right) \\ -\exp\left(+\frac{i}{2}\phi\right) \end{pmatrix}. \tag{7.30}$$

Note that

$$\langle\eta^{-}(\phi)|\eta^{+}(\phi)\rangle = \frac{1}{2}\left(\exp\left(+\frac{i}{2}\phi\right)\exp\left(-\frac{i}{2}\phi\right) - \exp\left(+\frac{i}{2}\phi\right)\exp\left(-\frac{i}{2}\phi\right)\right) = 0, \tag{7.31}$$

as it should.

Rotating by 360 degrees.

Interestingly: **if we rotate the spin vectors by 360 degrees,** for example by slowly rotating the magnetic field by an angle ϕ that starts at 0 and ends at 2π, **the spin vector change sign**! i.e.,

$$|\eta^{+}(\phi = 0)\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} \exp(i0) \\ \exp(i0) \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}. \tag{7.32}$$

while

$$|\eta^{+}(\phi = 2\pi)\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} \exp\left(-\frac{i}{2}2\pi\right) \\ \exp\left(+\frac{i}{2}2\pi\right) \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ -1 \end{pmatrix} = -|\eta^{+}(\phi = 0)\rangle. \tag{7.33}$$

This has profound implications in scattering in chemistry; we won't get into there, but beware.

Expectation value for a general spin direction

Back to a spin vector along a general direction (not just in the x-y plane). A given spin vector, $|\gamma\rangle = \begin{pmatrix} \gamma_+ \\ \gamma_- \end{pmatrix}$ has the following expectation values (henceforth we're back to $\hbar = 1$, so $\boldsymbol{S} = \boldsymbol{\frac{\sigma}{2}}$) :

$$\langle\gamma|S_z|\gamma\rangle = \frac{1}{2} \begin{pmatrix} \gamma_+^* & \gamma_-^* \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \gamma_+ \\ \gamma_- \end{pmatrix} = \frac{|\gamma_+|^2 - |\gamma_-|^2}{2}. \tag{7.34}$$

Similarly for S_x, S_y (see the book).

Measurement along two axis at an angle ϕ

More importantly, let's see what happens when we start with a spin up along a given direction, and then measure the spin along a different axis at an angle ϕ. What's the probability to end again up (along the new axis?)

Let's say the first axis was along x, and the next at an angle ϕ along the x-y plane, as we just studied. Then we start at $|\eta^+(\phi = 0)\rangle$

$$|\eta^+(\phi = 0)\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \tag{7.35}$$

so the amplitude to transfer to $|\eta^+(\phi)\rangle$ is $\langle\eta^+(\phi)|\eta^+(\phi = 0)\rangle$, and the probability is then

$$P^\phi = |\langle\eta^+(\phi)|\eta^+(\phi = 0)\rangle|^2 = \frac{1}{4} \left| \exp\left(-\frac{i}{2}\phi\right) + \exp\left(\frac{i}{2}\phi\right) \right|^2 = \cos^2\left(\frac{\phi}{2}\right) \tag{7.36}$$

8 Addition of two angular momenta

Classically: angular momenta add. Moon around the earth: $\boldsymbol{L_1}$, earth around the sun $\boldsymbol{L_2}$, so moon around the sun is $\boldsymbol{L_1} + \boldsymbol{L_2}$.

Quantum addition:

$$\boldsymbol{J} = \boldsymbol{L_1} + \boldsymbol{L_2}, \tag{8.1}$$

where $\boldsymbol{L_1}, \boldsymbol{L_2}$ are angular momenta, either spatially (of a particle) and/or spin. I.e., we can add the spin and the spatial angular momentum for the same particle, spins for different particles, spatial angular momentum for different particles, or other combinations.

Note that the different angular momenta are presumed to commute,

$$[L_{1,i}, L_{2,j}] = 0 \quad i, j = x, y, z. \tag{8.2}$$

> So a side note: different angular momenta commute, but an angular momenta "does not commute with itself" which is a fancy way of saying that an angular momenta x, y, z components don't commute with each other. That's a confusing point which we need to get used to.

With this it is easy to prove that the **sum of two angular momenta is also an angular momenta,** i.e., it fulfills

$$\boldsymbol{J} \times \boldsymbol{J} = i\boldsymbol{J}. \tag{8.3}$$

Proof:

$$\begin{aligned}\boldsymbol{J} \times \boldsymbol{J} = (\boldsymbol{L_1} + \boldsymbol{L_2}) \times (\boldsymbol{L_1} + \boldsymbol{L_2}) = \\ \boldsymbol{L_1} \times \boldsymbol{L_1} + \boldsymbol{L_2} \times \boldsymbol{L_2} + \boldsymbol{L_1} \times \boldsymbol{L_2} + \boldsymbol{L_2} \times \boldsymbol{L_1}.\end{aligned} \tag{8.4}$$

The brown-colored terms cancel each other since all components in $\boldsymbol{L_1}$ and $\boldsymbol{L_2}$ commute. Thus

$$\boldsymbol{J} \times \boldsymbol{J} = i\,\boldsymbol{L_1} + i\,\boldsymbol{L_2} = i\boldsymbol{J} \qquad Q.E.D. \tag{8.5}$$

This equation is completely equivalent to

$$[J_x, J_y] = iJ_z, \tag{8.6}$$

etc.

Triangle rule

Since $\boldsymbol{J} \times \boldsymbol{J} = i\boldsymbol{J}$, the proof that we had before on angular momenta is still valid, i.e., we can choose states with a definite integer value for $J^2 = \hbar^2 j(j+1)$, $J_z = m\hbar$ (I reinserted just this once $\hbar$). Just as before, the allowed range for m

would be $-j, -j+1, ..., j$. But what about the allowed values for l, if l_1 and l_2 are known?

Classically, when we add two vectors, $J = \boldsymbol{L_1} + \boldsymbol{L_2}$, then the length of the vector sum ranges between the difference and sum of the length of each, i.e. (see class):

$$\text{Classical triangle rule:} \quad \Big| |\boldsymbol{L}_1| - |\boldsymbol{L}_2| \Big| \leq |\boldsymbol{L_1} + \boldsymbol{L_2}| \leq |\boldsymbol{L_1}| + |\boldsymbol{L_2}|. \tag{8.7}$$

Quantally, the same rule turns out to be true, though we won't prove it:

$$\text{Quantum triangle rule:} \quad \Big| l_1 - l_2 \Big| \leq j \leq l_1 + l_2. \tag{8.8}$$

Example: addition of 2 spins $\boldsymbol{S_1} + \boldsymbol{S_2}$

Use $|\alpha_1\rangle$ to denote a spin up along the z axis for electron 1 ($m_1 = +\frac{\hbar}{2}$), while spin-down for electron 1 is denoted as $|\beta_1\rangle$. So there are four possible states:

- $|\alpha_1\alpha_2\rangle$, abbreviated as $|\alpha\alpha\rangle$(the first letter refers to the 1st electron, etc.)
- $|\alpha\beta\rangle$
- $|\beta\alpha\rangle$
- $|\beta\beta\rangle$

From the triangle rule, the total spin for the two electrons (abbreviated here as $\boldsymbol{S}$ rather than $\boldsymbol{J}$) will have the associated values between $s = \frac{1}{2} + \frac{1}{2} = 1$ and $s = \frac{1}{2} - \frac{1}{2} = 0$, i.e., it will be 0 or 1.

The $|s = 1, m = 1\rangle$ state

Let's find the states. Start with: $|s = 1, m = 1\rangle$. Since $S_z = S_{1z} + S_{2z}$, the only way for a state to have an eigenvalue of 1 is for the two eigenvalues to align, i.e.,

$$|s = 1, m = 1\rangle = |\alpha\alpha\rangle \tag{8.9}$$

Note our cursory use of bra-kets; on the left we refer to the eigenvalues of the combined spin and its projection along z, on the right we refer to the projection of the spin of each of the electrons on the z-axis.

We could proceed without proving the previous. equation (the RHS is the only available state for use on the LHS!), but for good measure (and practice) we will prove it. First apply S_z,

$$S_z|\alpha\alpha\rangle = (S_{z1} + S_{z2})\,|\alpha\alpha\rangle = \left(\frac{1}{2} + \frac{1}{2}\right)|\alpha\alpha\rangle = 1 * |\alpha\alpha\rangle. \tag{8.10}$$

So this works. Now apply $\boldsymbol{S^2}$,

$$\boldsymbol{S^2}|\alpha\alpha\rangle = \left(\boldsymbol{S_1^2} + \boldsymbol{S_2^2} + \boldsymbol{2S_1 \cdot S_2}\right)|\alpha\alpha\rangle = \left(\frac{1}{2} * \frac{3}{2} + \frac{1}{2} * \frac{3}{2} + 2S_1 \cdot S_2\right)|\alpha\alpha\rangle. \tag{8.11}$$

But generally, for two commuting angular momentum operators

$$2\,\boldsymbol{L_1 \cdot L_2} = (L_{1+}L_{2-} + L_{1-}L_{2+}) + 2L_{1z}L_{2z}, \tag{8.12}$$

Proof:

$$\begin{aligned} &L_{1+}L_{2-} + L_{1-}L_{2+} = (L_{1x} + iL_{1y})(L_{2x} - iL_{2y}) + (1 \leftrightarrow 2) = \\ &\quad L_{1x}L_{2x} + L_{1y}L_{2y} + i(L_{1y}L_{2x} - L_{2y}L_{1x}) + (1 \leftrightarrow 2) = 2L_{1x}L_{2x} + 2L_{1y}L_{2y} \end{aligned}$$

as needed. So

$$\boldsymbol{S^2}|\alpha\alpha\rangle = \left(\frac{3}{2} + S_{1+}S_{2-} + S_{1-}S_{2+} + 2S_{1z}S_{2z}\right)|\alpha\alpha\rangle. \tag{8.13}$$

But $S_{1+}|\alpha\rangle = 0$, $S_{1z}|\alpha\rangle = \frac{1}{2}|\alpha\rangle$, so

$$\boldsymbol{S^2}|\alpha\alpha\rangle = \left(\frac{3}{2} + 0 + 0 + 2 * \frac{1}{2} * \frac{1}{2}\right)|\alpha\alpha\rangle = 2 * 1|\alpha\alpha\rangle, \tag{8.14}$$

as expected for an $s = 1$ state (since s*(s+1)=1*(1+1)=2).

Thus we proved that $|\alpha\alpha\rangle$ is an eigenstate of $\boldsymbol{S^2}$ and of S_z, so it must be the $|s = 1, m = 1\rangle$ state.

Similarly, $|\beta\beta\rangle = |s = 1, m = -1\rangle$

Lowering operator to get $|s = 1, m = 0\rangle$

Next, apply a lowering operator:

$$\begin{aligned} |s = 1, m = 0\rangle &= const. * L_{-}|s = 1, m = 1\rangle = const. * (L_{1-} + L_{2-})|\alpha\alpha\rangle \\ &= const. * (L_{1-}|\alpha\alpha\rangle + L_{2-}\alpha\alpha\rangle) = const. * (|\beta\alpha\rangle + |\alpha\beta\rangle), \end{aligned} \tag{8.15}$$

and when we normalize properly

$$|s = 1, m = 0\rangle = \frac{|\beta\alpha\rangle + |\alpha\beta\rangle}{\sqrt{2}}. \tag{8.16}$$

Note that the three **triplet states** ($s = 1$, $m = -1, 0, 1$) are completely **symmetric** w.r.t. exchanging 1 and 2, as you probably know already.

Singlet

The singlet state, $|s = 0, m = 0\rangle$, is constructed to be orthogonal to $|s = 1, m = 0\rangle$, so it must be the antisymmetric combination

$$|s = 0, m = 0\rangle = \frac{|\beta\alpha\rangle - |\alpha\beta\rangle}{\sqrt{2}}. \tag{8.17}$$

Application: Spin-Spin Interaction (NMR, EPR, etc.)

Take an electron-nucleus, nucleus-nucleus, etc., interaction and assume that for some reason it has a spin-spin interaction component:

$$V(r) = V_0(r) + V_I(r)\boldsymbol{S_1} \cdot \boldsymbol{S_2}. \tag{8.18}$$

But as we saw

$$\boldsymbol{S_1} \cdot \boldsymbol{S_2} = \frac{1}{2}\left(\boldsymbol{S^2} - \boldsymbol{S_1^2} - \boldsymbol{S_2^2}\right). \tag{8.19}$$

When applying on individual spins, assume they are spin-half, i.e., $s_1 = s_2 = \frac{1}{2}$ (either an electron or a nucleon – proton or neutron – or a nucleus with net-spin 1/2).

$$S_1^2 = S_2^2 = \frac{1}{2} \times \frac{3}{2} = \frac{3}{4},$$

So $\boldsymbol{S_1} \cdot \boldsymbol{S_2}$ is diagonal (i.e., conserves) the total momentum and its z-projection, so

$$\boldsymbol{S_1} \cdot \boldsymbol{S_2}|s,m\rangle = \frac{1}{2}\left(\boldsymbol{S^2} - \boldsymbol{S_1^2} - \boldsymbol{S_2^2}\right)|s,m\rangle = \left(\frac{\boldsymbol{S^2}}{2} - \frac{3}{4}\right)|s,m\rangle = \left(\frac{s(s+1)}{2} - \frac{3}{4}\right)|s,m\rangle \tag{8.20}$$

The prefactor is then $\frac{s(s+1)}{2} - \frac{3}{4} = \frac{1*2}{2} - \frac{3}{4} = \frac{1}{4}$ for triplet states (s=1), and $-\frac{3}{4}$ for singlets. Thus the total interaction is

$$V(r) = V_0(r) + \frac{1}{4}V_I(r) \qquad s = 1$$

$$V(r) = V_0(r) - \frac{3}{4}V_I(r) \qquad s = 0 \tag{8.21}$$

Example: deuterium, p-n interaction; since $V_I(r) < 0$ (attraction) we find that the singlet would be unbound (its energy is too positive!)

The Einstein-Podolsky-Rose "paradox"

A very important and puzzling phenomena. Should be really the beginning of a few-weeks discussion on measurement and entanglement if we had time.

Prepare two electrons in a spin singlet state and then send them very far away from each other (while preserving their state). That's experimentally feasible. One will be on the left, the other on the right. They are in an "entangled" state, $\frac{|\beta\alpha\rangle - |\alpha\beta\rangle}{\sqrt{2}}$.

Now measure the spin on the left electron, i.e., measure $S_{left,z}$. If the result is up ($+\frac{1}{2}$, i.e., $|\alpha\rangle$) then it means you projected from the full w.f. the second component, i.e., (using colors)

$$\frac{|\beta\alpha\rangle - |\alpha\beta\rangle}{\sqrt{2}} \to |\alpha\beta\rangle \tag{8.22}$$

(and vice-versa for measuring $S_{left,z}$ to be $|\beta\rangle$.) But the eq. above means that, immediately, i.e., **faster than the speed of light,** the "information" that the left electron was measured to be spin-up "propagated" to the right-electron, which will be **for sure** in **a spin-down state!**

That was presented as a paradox by Einstein, Podolsky and Rosen in the 1930's, but it turns out to be experimentally measurable and true. The (not fully satisfying, for me at least) "answer" to the paradox is that QM is a non-local theory; the wavefunction "extends" over space, and affecting it in one place immediately affects it all over space.

Now back to "examples" and "chemistry".

Next example: Spin-Orbit interaction

Origin of the S.O. effect: the nuclei and electrons create an overall Coulomb electric potential, $V_c(r)$, and therefore an electric field,

$$\boldsymbol{E} = -\frac{1}{e}\nabla V_c(r) \approx -\frac{1}{e}\frac{\boldsymbol{r}}{r}\frac{dV_c(r)}{dr}. \tag{8.23}$$

where we used the fact that $V_c(r)$ is approximately spherical, so the electric field points radially.

The electric field when coupled with the motion of the electrons (velocity $\boldsymbol{v}$) creates a magnetic field

$$\boldsymbol{B} = -\frac{\boldsymbol{v}\times\boldsymbol{E}}{c} = -\frac{\boldsymbol{p}\times\boldsymbol{E}}{em_ec} = \frac{1}{em_ecr}\boldsymbol{p}\times\boldsymbol{r}\frac{dV_c(r)}{dr} = \frac{1}{em_ecr}\frac{dV_c(r)}{dr}\boldsymbol{L}. \tag{8.24}$$

Thus, as we learned in classical mechanics, an angular momentum creates a magnetic field. This magnetic field interacts with the dipole-moment that the electron spin creates, which turns out to be

$$\mu = g\frac{e\boldsymbol{S}}{2m_ec}, \tag{8.25}$$

where we denote explicitly the electron mass and velocity of light. The eq. above is classical in origin (each angular momentum, classical or quantum, creates such a dipole moment) except for the factor of $g \simeq 2$ due to relativity. So together the S.O. term in the Hamiltonian is

$$H_{S.O.} = f(r)\,\boldsymbol{S}\cdot\boldsymbol{L} \tag{8.26}$$

where

$$f(r) \approx \frac{1}{m_e^2cr}\frac{dV_c(r)}{dr}. \tag{8.27}$$

The spin-orbit interaction and combining L and S

Since $H_{S.O.} = f(r)\,\boldsymbol{S}\cdot\boldsymbol{L}$, we need to find its eigenstates. We know by now that when when combining spatial eigenstates with definite l and an electron spin (with $s=\frac{1}{2}$) then $\boldsymbol{S}\cdot\boldsymbol{L}$ will be diagonal when applied on eigenstates of $\boldsymbol{J^2}$ (and J_z) where $\boldsymbol{J}=\boldsymbol{L}+\boldsymbol{S}$ since

$$\boldsymbol{L}\cdot\boldsymbol{S}=\frac{1}{2}(\boldsymbol{J^2}-\boldsymbol{L^2}-\boldsymbol{S^2}), \tag{8.28}$$

so

$$\boldsymbol{L}\cdot\boldsymbol{S}|j,m\rangle=\frac{1}{2}\left(j(j+1)-l(l+1)-\frac{3}{4}\right)|j,m\rangle \tag{8.29}$$

where $\frac{3}{4}=s(s+1)$, and "m" refers to the eigenvalue of the total angular momentum, J_z (m_l and m_s refer to the eigenvalues of L_z and S_z, respectfully).

The triangle rule gives

$$j=l-\frac{1}{2},l+\frac{1}{2} \tag{8.30}$$

where we assume $l>0$ (the $l=0$ case is trivial – zero spin-orbit interaction). So therefore:

$$\text{for } j{=}l{+}\frac{1}{2}:\quad \boldsymbol{L}\cdot\boldsymbol{S}|j,m\rangle=\frac{1}{2}\left(\left(l+\frac{1}{2}\right)\left(l+\frac{3}{2}\right)-l(l+1)-\frac{3}{4}\right)|j,m\rangle=\frac{l}{2}|j,m\rangle \tag{8.31}$$

$$\text{for } j{=}l{-}\frac{1}{2}:\quad \boldsymbol{L}\cdot\boldsymbol{S}|j,m\rangle=\frac{1}{2}\left(\left(l-\frac{1}{2}\right)\left(l+\frac{1}{2}\right)-l(l+1)-\frac{3}{4}\right)|j,m\rangle=-\frac{l+1}{2}|j,m\rangle \tag{8.32}$$

So the splitting (difference in interaction) is linear in l.

Forming $|j,m\rangle$ eigenstates

We now need to ask: how can we combine the $(2l{+}1)$*2 different states, which we'll label as $|m_l m_s\rangle$ (omitting the l,s labels, and putting spatial angular momenta before the spin one). We wont get to the glory details, just a quick overview.

Combining $m_l=+l$ with a spin-up ($|\alpha\rangle$, i.e., $m_s=+\frac{1}{2}$) must give the higher-j state:

$$|j=l+\frac{1}{2},m=l+\frac{1}{2}\rangle=|m_l=l,m_s=\frac{1}{2}\rangle=|m_l=l,\alpha\rangle \tag{8.33}$$

Again note how the labels in the kets on the RHS and LHS refer to different things! This is made clearer when we write, in casual notation

$$\Psi^l_{j=l+\frac{1}{2},m=l+\frac{1}{2}}=y_{l,l}(\theta,\phi)|\alpha\rangle. \tag{8.34}$$

Now to a general $|j, m\rangle$state. We can reach an eigenvalue "m" of J_z in two ways. Combining $m_l = m - \frac{1}{2}$ with a spin-up, or $m_l = m + \frac{1}{2}$ with spin down. Thus:

$$|j, m\rangle = c^j_{m,\alpha}|m - \frac{1}{2}, \alpha\rangle + c^j_{m,\beta}|m + \frac{1}{2}, \beta\rangle \tag{8.35}$$

where Gasiorowicz derives the explicit form of $c^j_{m,\alpha}$ and $c^j_{m,\beta}$ (that depend of course also on the value of l), but we won't do it.

General addition of two angular momenta: Clebsch Gordan coefficients.

We added so far two spin halves, or a spin-half with a general angular momentum.

Now add two angular general momenta, i.e., $\boldsymbol{J} = \boldsymbol{L_1} + \boldsymbol{L_2}$. We know that given eigenstates $|y_{l_1 m_1}\rangle$ (labeled $|l_1, m_1\rangle$) of $(\boldsymbol{L_1})^2$and $L_{1,z}$, and analogous eigenstates labeled $|l_2, m_2\rangle$ of $(\boldsymbol{L_2})^2$and L_{2z}, we can form eigenstates of $\boldsymbol{J}^2, J_z, (\boldsymbol{L_1})^2, (\boldsymbol{L_2})^2$, labeled $|j, m, l_1, l_2\rangle$ from them; formally, we can write

$$|j, m, l_1 l_2\rangle = \sum_{m_1 m_2} C^{jm}_{l_1 m_1 l_2 m_2} |l_1, m_1, l_2, m_2\rangle. \tag{8.36}$$

Often this is abbreviated by omitting the $l_1 l_2$ from the LHS ket, i.e.,

$$|j, m\rangle = \sum_{m_1 m_2} C^{jm}_{l_1 m_1 l_2 m_2} |l_1, m_1, l_2, m_2\rangle. \tag{8.37}$$

The coefficients in the expansion are called Clebsch Gordan coefficients, and they are tabulated in books and on the web (e.g., use Wolfram Alpha). Formally, by dotting with the appropriate bra, we see that

$$\langle l_1 m_1 l_2 m_2 | j, m\rangle = C^{jm}_{l_1 m_1 l_2 m_2} \tag{8.38}$$

Some properties of the Clebsch Gordan coefficients

First, since $J_z = L_{1z} + L_{2z}$, we need to have $m = m_1 + m_2$, i.e.

$$C^{jm}_{l_1 m_1 l_2 m_2} = 0 \text{ if } \quad m_1 + m_2 \neq m \tag{8.39}$$

Similarly, the triangle rule states that

$$C^{jm}_{l_1 m_1 l_2 m_2} = 0 \text{ if } \quad j < |l_1 - l_2| \quad \text{or} \quad j > l_1 + l_2 \tag{8.40}$$

Next, there's one (actually two) CG coefficients we always know. Specifically, as we did earlier with spins, if we our two angular vectors are maximally aligned states, *i.e.*, if we consider the state $|l_1, m_1 = l_1, l_2, m_2 = l_2\rangle$ then the answer must be the combined system at the maximum angular momentum and most aligned, $|j = l_1 + l_2, m = j\rangle$, i.e.

$$|j = l_1 + l_2, m = j\rangle = |l_1, m_1 = l_1, l_2, m_2 = l_2\rangle. \tag{8.41}$$

This implies that

$$C^{j=l_1+l_2, m=m_1+m_2}_{l_1, m_1=l_1, l_2, m_2=l_2} = 1, \tag{8.42}$$

and if we reverse the sign of m_1, m_2, m in the eq. above we'll get 1 again of course.

The CG coefficients. fulfill sum rules; i.e., since $\langle j', m'|j, m\rangle = \delta_{jj'}\delta_{mm'}$ we get by dotting Eq. (8.37) with itself:

$$\sum_{m_1 m_2} C^{j,m}_{l_1 m_1 l_2 m_2} C^{j'm'}_{l_1 m_1 l_2 m_2} = \langle j', m'|j, m\rangle = \delta_{jj'}\delta_{mm'}. \tag{8.43}$$

It is easy to see that we can further generalize this relation to include the case of different l_1 and different l_2 in the two coefficients.

Finally, let's recall the simplest example, addition of two spins, in terms of CG coefficients. Specfcially, we learned that, for example, $|s = 1, m = 0\rangle = \frac{1}{\sqrt{2}}\left(|\alpha\beta\rangle + |\beta\alpha\rangle\right)$. In terms of CG coefficients this says

$$C^{1,0}_{\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, -\frac{1}{2}} = C^{1,0}_{\frac{1}{2}, -\frac{1}{2}, \frac{1}{2}, \frac{1}{2}} = \frac{1}{\sqrt{2}} \tag{8.44}$$

and you can similarly work out the coefficients for the other s=1 states and for the s=0 state.

9 Perturbation Theory

Crucial for chemistry – correlation is the difference between the mean-field and the actual correlated electron Hamiltonian. That's later. Now general theory and some atomic applications.

Non-degenerate perturbation theory

Given a solvable Hamiltonian, H_0, with easily solved eigenstates and eigenvalues

$$H_0|\phi_n\rangle = \varepsilon_n^0|\phi_n\rangle, \tag{9.1}$$

add a perturbation

$$H = H_0 + \lambda H_1, \tag{9.2}$$

with a small ("λ") strength; the new eigenstates and eigenvalues, which depend on the perturbation strength, satisfy

$$(H_0 + \lambda H_1)\,|\psi_n\rangle = \varepsilon_n|\psi_n\rangle. \tag{9.3}$$

Expand

$$|\psi_n\rangle = N(\lambda)\left(|\phi_n\rangle + \sum_{k\neq n} C_{nk}(\lambda)|\phi_k\rangle\right). \tag{9.4}$$

Expand order-by-order in λ (**the "1", "2", etc. on the C's and ε's are superscripts, not powers**).

$$C_{nk}(\lambda) = \lambda C_{nk}^1 + \lambda^2 C_{nk}^2 + ... \tag{9.5}$$

$$\varepsilon_n(\lambda) = \varepsilon_n^0 + \lambda\varepsilon_n^1 + \lambda^2\varepsilon_n^2 + ... \tag{9.6}$$

So

$$(H - \varepsilon_n)\psi_n = 0 \tag{9.7}$$

becomes (we divide $N(\lambda)$ out):

$$\left(H_0 + \lambda H_1 - \varepsilon_n^0 - \lambda\varepsilon_n^1 - \lambda^2\varepsilon_n^2 - ...\right)\left(|\phi_n\rangle + \lambda\sum_{k\neq n} C_{nk}^1|\phi_k\rangle + \lambda^2\sum_{k\neq n} C_{nk}^2|\phi_k\rangle + ...\right) = 0 \tag{9.8}$$

We'll derive the terms one by one; once we find them, the energy modification etc. would be found by setting $\lambda = 1$ (which is the as assuming that we replace λH_1 by H_1, i.e., absorb λ in H_1).

zeroth order

To zeroth order we get our undisturbed eq.,

$$\text{zeroth-order in } \lambda\text{:} \quad H_0|\phi_n\rangle = \varepsilon_n^0|\phi_n\rangle. \tag{9.9}$$

1st order

Now collect the terms linearly proportional to λ:

$$\left(\lambda H_1 - \lambda\varepsilon_n^1\right)|\phi_n\rangle + \left(H_0 - \varepsilon_n^0\right)\left(\lambda\sum_{k\neq n} C_{nk}^1|\phi_k\rangle\right) = 0 \tag{9.10}$$

i.e.,

$$\sum_{k\neq n} C_{nk}^1\left(\varepsilon_k^0 - \varepsilon_n^0\right)|\phi_k\rangle = -\left(H_1 - \varepsilon_n^1\right)|\phi_n\rangle. \tag{9.11}$$

1st order energy correction

Now first dot product with $\langle\phi_n|$ to get

$$-\langle\phi_n|\left(H_1 - \varepsilon_n^1\right)|\phi_n\rangle = \sum_{k\neq n} C_{nk}^1\left(\varepsilon_k^0 - \varepsilon_n^0\right)\langle\phi_n|\phi_k\rangle = \sum_{k\neq n} C_{nk}^1\left(\varepsilon_k^0 - \varepsilon_n^0\right) * 0 = 0, \tag{9.12}$$

i.e., the **first order correction to the energy is**

$$\varepsilon_n^1 = \langle\phi_n|H_1|\phi_n\rangle. \tag{9.13}$$

1st order correction to the eigenvectors

Next dot product Eq. (9.11) with $\langle\phi_j|$ $(j\neq n)$ to get

$$C_{nj}^1\left(\varepsilon_j^0 - \varepsilon_n^0\right)| = -\langle\phi_j|\left(H_1 - \varepsilon_n^1\right)|\phi_n\rangle = -\langle\phi_j|H_1|\phi_n\rangle, \tag{9.14}$$

i.e., the **first order correction to the eigenstates is**

$$C_{nj}^1 = \frac{\langle\phi_j|H_1|\phi_n\rangle}{\varepsilon_n^0 - \varepsilon_j^0} \quad (n\neq j). \tag{9.15}$$

Note that distant energy state contribute less than nearby states — we know it from first-year chem. where in M.O.s we mix only states with similar energy.

Energy denominator

We should pause here to **think about the energy denominators,** $(\frac{1}{\varepsilon_n^0-\varepsilon_j^0})$. They are prevalent also in 2nd-order and any-order perturbation theory. Beyond the dry math, why are they there physically?

To better appreciate, we'll need time-dependent perturbation theory, developed in 115b/215b. But already here we can state the key aspect (my reasoning is extremely heuristic) in two different ways:

- Extremely simple: starting from a state 'n', we in a sense violate conservation of energy by "jumping" to state j. Put differently, the perturbation H_1 polarizes the system. But we can only do it for a "short time" – and the length of time is $\hbar$/(Energy Difference), i.e., $\frac{1}{\varepsilon_n^0-\varepsilon_j^0}$.

- Somewhat more rigorous: The undisturbed eigenstates $|\phi_n\rangle$ and $|\phi_j\rangle$ oscillate in time with phases $\exp(-i\varepsilon_n^0 t)$ and $\exp(-i\varepsilon_j^0 t)$. Adding a perturbation is in a certain sense "kicking" the n'th state to the j'th state (with kicking-amplitude $\langle\phi_j|H_1|\phi_n\rangle$). But once we start "kicking" we have interference between the kicked-state amplitude (at level "j") and the original part (at level "n"), i.e., dephasing, since the two states evolve differently, with a relative phase $\exp\left(-i(\varepsilon_j^0 - \varepsilon_n^0)t\right)$. Thus, "kicked" amplitudes at different times interfere constructively or destructively, and the **integral of the dephased contribution amplitudes will be**

$$\langle\phi_j|H_1|\phi_n\rangle \int_0^\infty \exp\left(-i(\varepsilon_j^0 - \varepsilon_n^0)t\right) dt = \frac{\langle\phi_j|H_1|\phi_n\rangle}{i(\varepsilon_j^0 - \varepsilon_n^0)} \tag{9.16}$$

Now back to the formulae, this time for:

2nd order perturbation theory

The 2nd order terms in Eq. (9.8) are

$$\lambda^2\left(H_0 - \varepsilon_n^0\right)\sum_{k\neq n} C_{nk}^2|\phi_k\rangle + \left(\lambda H_1 - \lambda\varepsilon_n^1\right)\lambda\sum_{k\neq n} C_{nk}^1|\phi_k\rangle - \lambda^2\varepsilon_n^2|\phi_n\rangle = 0. \tag{9.17}$$

Dot with $\langle\phi_n|$ to get (and notice that the 1st term involving $|\phi_k\rangle,\ k\neq n$, drops out, as does the ε_n^1 term):

$$\sum_{k\neq n} C_{nk}^1\langle\phi_n|H_1|\phi_k\rangle = \varepsilon_n^2, \tag{9.18}$$

i.e., based on the previous equation for C_{nk}^1,

$$\varepsilon_n^2 = \sum_{j\neq n}\frac{\langle\phi_j|H_1|\phi_n\rangle}{\varepsilon_n^0 - \varepsilon_j^0}\langle\phi_n|H_1|\phi_j\rangle, \tag{9.19}$$

i.e.,

$$\varepsilon_n^2 = \sum_{j\neq n}\frac{|\langle\phi_j|H_1|\phi_n\rangle|^2}{\varepsilon_n^0 - \varepsilon_j^0}. \tag{9.20}$$

Note that if n is the ground state the denominator would always be negative, so the correction ε_n^2 would then be negative. I.e., any perturbation that has a vanishing first-order contribution would lower the ground-state energy!

Similarly, if n is an excited state, any higher energy state j gives a negative contribution ($\varepsilon_n^0 - \varepsilon_j^0 < 0$), i.e., "pushes" the energy down, and any lower-energy state "pushes" the energy up.

We know this from first-year MO diagrams. There, the 0'th order states are separate atoms, and the perturbation is the difference between the full molecular Hamiltonian and this 0th order Hamiltonian. Then we know that when we "combine" two states, j and n (in this case belonging to two separate atoms), the low one is "pushed" lower and the high one is pushed higher. See diagram in class.

2nd order perturbation theory and vdW interaction

Another way to look at the 2nd order perturbation theory is the vdW picture, which we saw already when we talked about C^1_{nj} and the energy denominators. Starting from a state “n”, the system can fluctuate to a state “j”, for a short time (1/(energy difference)) and then go back to state “n”. Going back and forth has a squared amplitude of $|\langle\phi_j|H_1|\phi_n\rangle|^2$, and is multiplied by the short time.

Sum rules and 2nd order perturbation theory example: Stark shift

Apply on hydrogen an electric field, assumed to be along z. The perturbation due to the field is then

$$H_1 = -e\,\boldsymbol{E}\boldsymbol{\cdot}\,\boldsymbol{r} = -fz \quad (\boldsymbol{E} \parallel \boldsymbol{z},\ f = eE). \tag{9.21}$$

The first order perturbation correction vanishes for any state due to symmetry

$$\varepsilon^1_n \equiv \varepsilon^1_{n=1} = -f\langle n|z|n\rangle = \int z|\psi_n(\boldsymbol{r})|^2 d\boldsymbol{r} = 0 \tag{9.22}$$

since the states are all symmetric or antisymmetric, so the “z” due to the electric field times the w.f. squared will be antisymmetric in z.

We’ll now concentrate on the 2nd order change in energy of the ground state. We’ll switch notation, defining the shift of the ground-state energy as

$$-\,\delta\varepsilon \equiv -\varepsilon^2_{n=1} = f^2\sum_{j\neq 1}\frac{|\langle\phi_j|z|\phi_1\rangle|^2}{\varepsilon^0_j-\varepsilon^0_1}. \tag{9.23}$$

This summation can be approximated, once we realize that

- First, most of the contribution will be from bound state (i.e., ignore the continuum states
- Also, the energy differences from the ground state are in a narrow range ($13.6 * \frac{3}{4}$ eVfor n=1 to n=2; 13.6eV for n=1 to very high n).

We therefore substitute, using

$$\frac{1}{\varepsilon^0_j-\varepsilon^0_1} \leq \frac{1}{\varepsilon^0_2-\varepsilon^0_1} = \frac{1}{\varepsilon^0_2-\varepsilon^0_1} \tag{9.24}$$

getting

$$-\,\delta\varepsilon \leq f^2\sum_{j\neq 1}\frac{|\langle\phi_j|z|\phi_1\rangle|^2}{\varepsilon^0_2-\varepsilon^0_1} = \frac{f^2}{\varepsilon^0_2-\varepsilon^0_1}\sum_j|\langle\phi_j|z|\phi_1\rangle|^2 \tag{9.25}$$

where in the last summation we extended the summation also to j=1, since $\langle\phi_j|z|\phi_1\rangle = 0$

We now introduce the technique of **sum rules**

$$\sum_j |\langle\phi_j|z|\phi_1\rangle|^2 = \sum_j \langle\phi_1|z|\phi_j\rangle\langle\phi_j|z|\phi_1\rangle = \langle\phi_1|z{*}1{*}z|\phi_1\rangle = \langle\phi_1|z^2|\phi_1\rangle = \frac{\langle\phi_1|r^2|\phi_1\rangle}{3} \tag{9.26}$$

which is easily calculable; but we won't even worry about calculating it, since by dimensional analysis $\langle\phi_1|r^2|\phi_1\rangle$ should be, for hydrogen, a number not too far from 1 times $(a_0)^2$.

All together then:

$$|\delta\varepsilon| \leq \frac{(eE)^2}{\varepsilon_2^0 - \varepsilon_1^0}\frac{\langle\phi_1|r^2|\phi_1\rangle}{3}. \tag{9.27}$$

Breakdown of perturbation theory

If the field strength is too big, then we can't trust 2nd order perturbation theory; in fact, if it is really strong the perturbation expansion fails altogether. Physically, at very strong field strengths the electron would be ripped out from the atom. See picture in the book and in class.

Degenerate Perturbation Theory

Happens if

$$\langle\phi_j|H_1|\phi_n\rangle \neq 0, \quad \varepsilon_n^0 = \varepsilon_j^0 \tag{9.28}$$

Then perturbation theory explodes since the $\langle\phi_j|H_1|\phi_n\rangle/(\varepsilon_j^0 - \varepsilon_n^0)$ term (and others like it, with the same denominator) explode.

The solution is very simple: take the starting subset of states that are degenerate-energy with $|\phi_n\rangle$ (and of course $|\phi_n\rangle$ itself) and **rotate** it by redefining a new orthogonal basis which is a linear combination of states of the **same energy**

$$|\tilde{\phi}_n\rangle = \sum_{l,\varepsilon_l^0=\varepsilon_n^0} a_{ln}|\phi_l\rangle, \tag{9.29}$$

so that in this subspace (of rotated states with the same energy as ε_n^0) the perturbation is diagonal

$$\langle\tilde{\phi}_m|H|\tilde{\phi}_l\rangle = \delta_{ml}\langle\tilde{\phi}_m|H|\tilde{\phi}_n\rangle. \tag{9.30}$$

An example will clarify.

2nd order degenerate perturbation theory: Stark shift for $n = 2$ states in hydrogen

Recall $H_1 = -eEz$. What's the energy perturbation for n=2 states? There are 4 such states of course; we can ignore $|2p_x\rangle$ and $|2p_y\rangle$, and concentrate of $|2s\rangle, |2p_z\rangle$.

Here, the problem is that

$$c \equiv \langle 2s|z|2p_z\rangle \neq 0 \tag{9.31}$$

So in this $|2s\rangle, |2p_z\rangle$ subspace, the perturbation potential matrix (we use the same symbol for it as the operator) has the form

$$H_1 = \begin{pmatrix} 0 & -eEc \\ -eEc & 0 \end{pmatrix}. \tag{9.32}$$

Clearly what we need is to define two sp-states,

$$|left\rangle \equiv \frac{|2s\rangle - |2p_z\rangle}{\sqrt{2}}, \tag{9.33}$$

and

$$|right\rangle \equiv \frac{|2s\rangle + |2p_z\rangle}{\sqrt{2}}. \tag{9.34}$$

so in the rotated basis

$$\tilde{H}_1 = \begin{pmatrix} eEc & 0 \\ 0 & -eEc \end{pmatrix}, \tag{9.35}$$

and since there's no interaction between different states (between Left and Right here) we can remove these interaction in the perturbation expansion.

After a two-line calculation, we see then (**prove it!**) that the Stark shift 2nd-order perturbation energy corrections would be

$$\varepsilon_{n=2} = -\frac{Ze^2R}{4} + eEc \times \begin{cases} 1 & (left) \\ 0 & p_x, p_y \\ -1 & (right). \end{cases} \tag{9.36}$$

Atomic examples of perturbation theory

In the remainder of the chapter we'll briefly consider perturbation theory as applied to atoms.

The most important interactions in an atom is, of course:

- H_0 : kinetic $\frac{p^2}{2m_e}$+nuclear+ $e-e$ interaction on a mean-field level (the interaction with the mean-field of all other electrons, we'll see later).

Beyond that, the other interactions are.

- e-e interaction beyond the mean-field level – here we'll only include a little discussion on atomic vdW interactions and continue in latter chapters.
- Spin orbit interaction: $\propto \boldsymbol{L}\cdot\boldsymbol{S}$
- Interaction with magnetic field: $\propto \boldsymbol{B}\cdot(\boldsymbol{L}+g\boldsymbol{S})$
- Relativistic corrections to the kinetic energy.

Let's tackle them one-by-one, starting with the last term:

Relativistic corrections to the kinetic energy.

The true K.E. in relativity is ("c" is the velocity of light here)

$$\sqrt{p^2c^2+m_e^2c^4}-m_ec^2=m_ec^2\left(\sqrt{1+\frac{p^2}{m_e^2c^2}}-1\right)\simeq$$

$$m_ec^2\left((1+\frac{p^2}{2m_e^2c^2}-\frac{p^4}{4m_e^4c^4}-1\right)=\frac{p^2}{2m_e}-\frac{p^4}{4m_e^3c^2} \quad (9.37)$$

where we used the Taylor expansion expression, $\sqrt{1+a}\simeq 1+\frac{a}{2}-\frac{a^2}{4}+....$ Thus, the correction to the orbital energy is proportional to $\langle\phi_n|p^4|\phi_n\rangle$.

Spin-orbit interactions:

Recall $H_{S.O.}=f(r)\boldsymbol{L}\cdot\boldsymbol{S}$. We recall that this requires combination of the y_{lm} and spin states, so the S.O. interaction is diagonal in a state labeled as $\Psi_{n,l,j,m}$ where "n" is the usual # of angular nodes, and as before "l" labels the spatial angular momentum, "j" labels the total angular momentum, and "m" is the eigenvalue of J_z. Then the orbital perturbation is:

$$\Delta_{nljm}=\langle\Phi_{nljm}|f(r)\boldsymbol{L}\cdot\boldsymbol{S}|\Phi_{nljm}\rangle, \quad (9.38)$$

where (casual notation):

$$\Phi_{nljm}=R_{nl}(r)\Psi^l_{jm} \quad (9.39)$$

so

$$\delta\varepsilon^{S.O.}_{nljm}=b_{nl}\langle j,m|\boldsymbol{L}\cdot\boldsymbol{S}|j,m\rangle, \quad (9.40)$$

where

$$b_{nl}\equiv\langle R_{nl}|f|R_{nl}\rangle=\int|R_{nl}(r)|^2f(r)r^2dr. \quad (9.41)$$

The b_{nl} term turns out to be positive.

Note that bra-kets refer above to a different space than the previous equation! (radial here, full 3-D a few equations above).

Also, remember that the combined angular momentum state $|j,m\rangle$ depends also on l.

We calculated earlier $\boldsymbol{L}\cdot\boldsymbol{S}|j,m\rangle$ so we recall the result

$$\langle j,m|\boldsymbol{L}\cdot\boldsymbol{S}|j,m\rangle=\frac{1}{2}\langle j,m|\boldsymbol{J}^2-\boldsymbol{L}^2-\boldsymbol{S}^2|j,m\rangle=\frac{1}{2}\left(j(j+1)-l(l+1)-\frac{3}{4}\right) \quad (9.42)$$

so the final result is:

$$\delta\varepsilon^{S.O.}_{nljm}=\frac{b_{nl}}{2}\times\begin{cases}l & \left(j=l+\frac{1}{2}\right)\\ -(l+1) & \left(j=l-\frac{1}{2}\right).\end{cases} \quad (9.43)$$

The splitting between the $j = l \pm \frac{1}{2}$ levels is essentially proportional to l.

Example: Yellow line in Na. The valence electron is in 3s; the six $3p$ states (3 spatially*2spin) are split due to the S.O. interaction to two levels, denoted as $3p_{\frac{3}{2}}$, i.e., $n = 3, l = 1, j = \frac{3}{2}$ (of course with four m-states), and $3p_{\frac{1}{2}}$ (with two m states); the latter level is lower in energy, by the formula above.

And indeed experimentally the $3s \to 3p$ transition has two peaks, the lower-frequency $3s \to 3p_{\frac{1}{2}}$ (5896 Angstrom) and the higher-frequency $3s \to 3p_{\frac{3}{2}}$ (5890 Angstrom). The frequency splitting between these levels matches (when multiplied by $\hbar$) the calculated 1st order perturbation theory term.

> Question for thought: would a classical mechanics treatment of $\boldsymbol{L}$ give a similar result for the expectation value$\langle \boldsymbol{L} \cdot \boldsymbol{S} \rangle$?

Zeeman effect: interaction with magnetic field.

We talked about the interaction with $\boldsymbol{B}$, with an angular momentum, which is (for a field pointing along the z-axis)

$$H_B = \mu_B \boldsymbol{B} \cdot (\boldsymbol{L} + g\boldsymbol{S}) \simeq \mu_B \boldsymbol{B} \cdot (\boldsymbol{L} + 2S) = \mu_B \boldsymbol{B} \cdot (\boldsymbol{J}+\boldsymbol{S}) = \mu_B \left(J_z + S_z \right), \tag{9.44}$$

since the g-factor for electron (the enhancement of the interaction with magnetic field for spin relative to angular momentum) is about $g \simeq 2$. Here,

$$\mu_B = \text{Bohr Magneton}=\frac{e}{2m_e c}.$$

As long as the magnetic field is not too strong compared to the S.O. interaction we should use as our basis the $|j, m\rangle$ states (for which $H_{S.O.}$ is diagonal) as our basis, and calculate the expectation value of H_B in that basis (the opposite case, very strong magnetic fields, is discussed later):

$$\frac{1}{\mu_B} \cdot \delta\varepsilon^B_{nljm} = \langle j, m|H_B|j, m\rangle = \langle j, m|J_z + S_z|j, m\rangle = \\ \langle j, m|m + S_z|j, m\rangle = (m + \langle j, m|S_z|j, m\rangle). \tag{9.45}$$

The matrix element $\langle j, m|S_z|j, m\rangle$ can be calculated from the expressions we had of $|j, m\rangle = c^j_{m,\alpha}|m_l = m - \frac{1}{2}, \alpha\rangle + c^j_{m,\beta}|m_l = m + \frac{1}{2}, \beta\rangle$. But at any rate, it is not going to be large, i.e.,

$$-\frac{1}{2} \leq \langle j, m|S_z|j, m\rangle \leq \frac{1}{2} \tag{9.46}$$

so we can roughly ignore it,

$$\delta\varepsilon^B_{nljm} \simeq \mu_B B m. \tag{9.47}$$

Thus, the m-levels split linearly (approximately).

Finally, if the field is extremely strong then the S.O. can be ignored; the best basis will be one where H_B is diagonal, i.e., the $|y_{lm_l}, m_s\rangle$basis (denoted earlier more simply as $|m, m_s\rangle$or better yet write it as $|l, m, m_s\rangle$. Then

$$\delta\varepsilon^B_{nlm_lm_s} = \langle l, m_l, m_s|H_B|l, m_l, m_s\rangle = \mu_B B\langle l, m_l, m_s|L_z{+}gS_z|l, m_l, m_s\rangle = \mu_B B\,(m_l + gm_s) \tag{9.48}$$

The book has some figures on the transition between the $H_B \ll H_{SO}$ and $H_{SO} \ll H_B$ cases.

Following Sakurai, we thus give a table comparing the weak-B-field and strong one case.

- Weak field: **dominant interaction:** $H_{SO.}$ **Almost conserved operators:** $\boldsymbol{J^2}$ (*i.e.,*$\boldsymbol{L}\cdot\boldsymbol{S}$). **Not good:** L_z, S_z.
- Strong field: **dominant interaction:** $H_{B.}$ **Almost conserved operators:** L_z, S_z.. **Not good:** $\boldsymbol{J^2}$ (*i.e.,* $\boldsymbol{L}\cdot\boldsymbol{S}$).
- Both cases, always good operators: $\boldsymbol{L^2}$**,** $\boldsymbol{S^2}$**,**J_z

Example: vdW interaction energy of two distant hydrogen atoms

Take two distant hydrogens, i.e., one electron around one proton and the 2nd electron around the other one. Ignore antisymmetry and exchange, i.e., treat the electrons as a distinct particles. Label the (vector) distance of electron 1 from proton 1 as $\boldsymbol{r_1}$, and the (vector) distance of electron 2 from proton 2 as $\boldsymbol{r_2}$. Further:

- The vector distance between protons is denoted $\boldsymbol{R}$,
- so the vector distance between the two electrons is $\boldsymbol{r_1} - \boldsymbol{R} + \boldsymbol{r_2}$ (see picture in book/class), and
- the distance between electron 1 and proton 2 is $|\boldsymbol{r_1} - \boldsymbol{R}|$,
- and the distance electron 2 and proton 1 is $|\boldsymbol{r_2} + \boldsymbol{R}|$.

The Hamiltonian is made from a zero-order part, and the "perturbation"

$$H = H_0 + H_1 \tag{9.49}$$

where the 0-order part has each electron around its own nucleus

$$H_0 = H_{01} + H_{02}, \tag{9.50}$$

where H_{01} is the 0-th order Hamiltonian for the 1st electron,

$$H_{01} = \frac{(\boldsymbol{p_1})^2}{2m_e} - \frac{e^2}{r_1}. \tag{9.51}$$

H_{02} is analogous.

The perturbation is made from four terms: ion-ion interaction, the electron-electron repulsion, electron-1 attraction with proton-2, and the electron-2 proton-1 attraction

$$H_1 = +\frac{e^2}{R} + \frac{e^2}{|\boldsymbol{r_1} + \boldsymbol{R} - \boldsymbol{r_2}|} - \frac{e^2}{|\boldsymbol{r_1} - \boldsymbol{R}|} - \frac{e^2}{|\boldsymbol{r_2} + \boldsymbol{R}|} \tag{9.52}$$

The 0-order 2-elecrron wavefunction is (ignoring spin)

$$\Psi(\boldsymbol{r_1}, \boldsymbol{r_2}) = \phi_{1s}(r_1)\phi_{1s}(r_2). \tag{9.53}$$

Let's calculate the matrix element at the limit of large-R; then we can show (a long derivation which we will skip) that

$$H_1 = e^2 \frac{\boldsymbol{r_1} \cdot \boldsymbol{r_2} - \boldsymbol{3}\left(\boldsymbol{r_1} \cdot \boldsymbol{\hat{R}}\right)\left(\boldsymbol{r_2} \cdot \boldsymbol{\hat{R}}\right)}{R^3} \tag{9.54}$$

where $\boldsymbol{\hat{R}} = \frac{\boldsymbol{R}}{R}$. Thus, the R-dependence of the matrix element is R^{-3}, and when we square it we get the vdW R^{-6}dependence.

The matrix element can be calculated analytically, but it is quite painstaking.

10 Functional derivatives and Lagrange Multipliers

In this part we'll cover two mathematically important concepts we have not discussed yet: funcitonal derivatives and Lagrange multipliers (you probably have seen the latter).

Functional derivatives

Let's again treat integrals as discrete summation, and ask what's the derivative of an integral w.r.t. the value of a function at a particular point. The simplest such expression will be

$$\frac{\partial}{\partial f(x)}\int f(x')dx' = \frac{\partial}{\partial f(x)}\sum_{x'} f(x')dx = \sum_{x'}\frac{\partial f(x')}{\partial f(x)}dx = \sum_{x'}\delta_{xx'}dx = dx, \tag{10.1}$$

where we used

$$\frac{\partial f(x')}{\partial f(x)} = \delta_{xx'}. \tag{10.2}$$

To get rid of the resulting "dx" we define then the functional derivative as

$$\frac{\delta}{\delta f(x)} = \frac{1}{dx}\frac{\partial}{\partial f(x)}. \tag{10.3}$$

and this refers, of course, to the limit of tiny dx. Therefore

$$\frac{\delta}{\delta f(x)}\int f(x')dx' = 1. \tag{10.4}$$

Formally, since $\frac{\partial}{\partial f(x)} f(x') = \delta_{xx'}$, we get

$$\frac{\delta f(x')}{\delta f(x)} = \frac{\delta_{xx'}}{dx} = \delta(x-x'). \tag{10.5}$$

The generalizations to 3D are straightforward.

First Example: the classical electrostatic repulsion energy due to a charge density $n(\boldsymbol{r})$ will be (we'll see it later too):

$$E_{ee} = \frac{1}{2}\int \frac{n(\boldsymbol{r})n(\boldsymbol{r'})}{|\,\boldsymbol{r}-\boldsymbol{r'}|}d\boldsymbol{r}d\boldsymbol{r'}. \tag{10.6}$$

Note that we generally use atomic units (no $4\pi\epsilon_0$, the electron charge is unity $e=1$, and $\hbar = 1$ – although often we'll keep $\hbar$ explicitly).

Upon differentiation we get (the factor of 2 is due to two equal terms; also note that before taking the derivative I am changing the integration variable):

$$\frac{\delta E_{ee}}{\delta n(\boldsymbol{r})} = \frac{1}{2}\frac{\delta}{\delta n(\boldsymbol{r})}\int \frac{n(\boldsymbol{r''})n(\boldsymbol{r'})}{|\,\boldsymbol{r''}-\boldsymbol{r'}|}d\boldsymbol{r''}d\boldsymbol{r'} = \frac{1}{2}*2*\int \frac{\delta n(\boldsymbol{r''})}{\delta n(\boldsymbol{r})}\frac{n(\boldsymbol{r'})}{|\,\boldsymbol{r''}-\boldsymbol{r'}|}d\boldsymbol{r''}d\boldsymbol{r'}$$
$$= \int \delta(\boldsymbol{r}-\boldsymbol{r''})\frac{n(\boldsymbol{r'})}{|\,\boldsymbol{r''}-\boldsymbol{r'}|}d\boldsymbol{r''}d\boldsymbol{r'}, \quad (10.7)$$

i.e.,

$$\frac{\delta}{\delta n(\boldsymbol{r})}E_{ee} = \int \frac{n(\boldsymbol{r'})}{|\,\boldsymbol{r}-\boldsymbol{r'}|}d\boldsymbol{r'}. \quad (10.8)$$

Another example: the kinetic energy associated with a wavefunction $\Psi(\boldsymbol{r})$ of an electron with mass m_e is

$$T = \frac{\hbar^2}{2m_e}\int \nabla\Psi^*(\boldsymbol{r})\cdot\nabla\Psi(\boldsymbol{r})d\boldsymbol{r}. \quad (10.9)$$

Differentiating w.r.t. the complex-conjugate of the w.f. gives:

$$\frac{\delta T}{\delta\Psi^*(\boldsymbol{r})} = \frac{\hbar^2}{2m_e}\frac{\delta T}{\delta\Psi^*(\boldsymbol{r})}\int \nabla\Psi^*(\boldsymbol{r'})\cdot\nabla\Psi(\boldsymbol{r'})d\boldsymbol{r'}$$
$$= \frac{\hbar^2}{2m_e}\int \nabla\left(\frac{\delta T}{\delta\Psi^*(\boldsymbol{r})}\Psi^*(\boldsymbol{r'})\right)\cdot\nabla\Psi(\boldsymbol{r'})d\boldsymbol{r'}, \quad (10.10)$$

so (we don't do a derivative of Ψ w.r.t. Ψ^*!)

$$\frac{\delta T}{\delta\Psi^*(\boldsymbol{r})} = \frac{\hbar^2}{2m_e}\int \nabla\left(\delta(\boldsymbol{r}-\boldsymbol{r'})\right)\cdot\nabla\Psi(\boldsymbol{r'})d\boldsymbol{r'}. \quad (10.11)$$

Let's integrate by parts (where the non-integral term is taken over a distant surface)

$$\frac{\delta T}{\delta\Psi^*(\boldsymbol{r})} = \frac{\hbar^2}{2m_e}\left(\delta(\boldsymbol{r}-\boldsymbol{r'})\nabla\Psi(\boldsymbol{r'})|_{\text{surface},|\boldsymbol{r}|\,\rightarrow\,\infty} - \int \delta(\boldsymbol{r}-\boldsymbol{r'})\nabla^2\Psi(\boldsymbol{r'})d\boldsymbol{r'}\right). \quad (10.12)$$

Note that the surface term vanishes since we assume that the w.f. is bounded; so integrate the delta function in the integral to get

$$\frac{\delta T}{\delta\Psi^*(\boldsymbol{r})} = -\frac{\hbar^2}{2m_e}\nabla^2\Psi(\boldsymbol{r}), \quad (10.13)$$

and this is just the action of the kinetic energy operator on the wavefunction!

Lagrange Multipliers

To extremize an expression while holding some constraint fixed, subtract the constraint times a Lagrange multiplier.

E.g., define (again using the same symbol T, but for a different expression; this is actually the same kinetic energy as above, but for a specific case, a uniform electron gas – we'll see later in the course the relation):

$$T = \frac{3}{5} c \int n(\boldsymbol{r})^{\frac{5}{3}} d\boldsymbol{r}, \tag{10.14}$$

where c is some numerical constant we'll see later.

Maximize T w.r.t. $n(\boldsymbol{r})$, keeping the total number of electrons (the integral over the density) fixed: $\int n(\boldsymbol{r})d\boldsymbol{r} = N$, by using a Lagrange multiplier. Thus, define D, a modified functional of the density, including the Lagrange multiplier:

$$D = T - \mu \left(\int n(\boldsymbol{r}) d\boldsymbol{r} - N \right) \tag{10.15}$$

and then find the optimal density by functional differentiation

$$0 = \frac{\delta D}{\delta n(\boldsymbol{r})} = \frac{\delta \left(T - \mu \left(\int n(\boldsymbol{r}) d\boldsymbol{r} - N \right) \right)}{\delta n(\boldsymbol{r})}. \tag{10.16}$$

Here, this gives

$$0 = \frac{\delta}{\delta n(\boldsymbol{r})} \left(\frac{3}{5} c \int n(\boldsymbol{r'})^{\frac{5}{3}} d\boldsymbol{r'} - \mu \left(\int n(\boldsymbol{r}') d\boldsymbol{r'} - N \right) \right) = c n(\boldsymbol{r})^{\frac{2}{3}} - \mu, \tag{10.17}$$

i.e., $n(\boldsymbol{r}) = (\mu/c)^{3/2}$=const. So the density is uniform indeed. Put differently, starting from a uniform density expression, we get back a uniform density. So the uniform density expression is <u>stable.</u>

11 Variational Principle

Simpler approach

First, simpler approach (that you probably saw). Assume we're given a set of eigenfunctions $|k\rangle$, $H|k\rangle = \varepsilon_k|k\rangle$, numbered from k=0. Any wavefunction $|\eta\rangle$can be expanded|

$$|\eta\rangle = \sum_k c_k|k\rangle, \tag{11.1}$$

so that

$$\langle H\rangle \equiv \frac{\langle\eta|H|\eta\rangle}{\langle\eta|\eta\rangle} = \frac{\sum_k |c_k|^2\varepsilon_k}{\sum_k |c_k|^2} \geq \frac{\sum_k |c_k|^2\varepsilon_0}{\sum_k |c_k|^2} = \varepsilon_0 \tag{11.2}$$

i.e., the normalized expectation value of the Hamiltonian with any function is larger or equal than ε_0.

Note that if $|\eta\rangle$ is close to the ground state, i.e., if

$$c_n \sim O(\lambda), \quad n \neq 0 \tag{11.3}$$

(where $O(\lambda)$ means that it is proportional to λ to the 1st order, which is a small number), then $|c_0|^2 = 1 - O(\lambda^2)$. Proof: $\sum_k |c_k|^2 = 1$, so

$$|c_0|^2 = 1 - \sum_k |c_k|^2 = 1 - O(\lambda^2). \tag{11.4}$$

Therefore, in that case:

$$\langle H\rangle \equiv \frac{\sum_k |c_k|^2\varepsilon_k}{1} = |c_0|^2\varepsilon_0 + \sum_{k\neq 0} |c_k|^2\varepsilon_k = \left(1 - O\left(\lambda^2\right)\right)\varepsilon_0 + O(\lambda^2) = \varepsilon_0 + O(\lambda^2). \tag{11.5}$$

i.e., **the error in the energy is of second order in the deviation of the w.f. from the ground state.**

Put differently, even a crummy w.f. gives a good estimate to the ground-state energy.

Variational principle: with a Lagrange variable.

Instead of $\frac{\langle\Psi|H|\Psi\rangle}{\langle\Psi|\Psi\rangle} \geq \varepsilon_0$, minimize $\langle\Psi|H|\Psi\rangle$ under the constraint $\langle\Psi|\Psi\rangle - 1 = 0$, i.e., extremize the **functional** of the w.f.,

$$D \equiv D[\Psi^*, \Psi, \mu] = \langle\Psi|H|\Psi\rangle - \mu\left(\langle\Psi|\Psi\rangle - 1\right). \tag{11.6}$$

Note that we need to extremize both w.r.t. μ and Ψ^* (or w.r.t. the parameters in Ψ^* if it is restricted).

Extremizing w.r.t. μ always gives:

$$0 = \frac{\partial D}{\partial \mu} = \frac{\partial \mu}{\partial \mu}\left(\langle\Psi|\Psi\rangle - 1\right) = \langle\Psi|\Psi\rangle - 1. \tag{11.7}$$

so Ψ needs to be normalized indeed. Further, if Ψ (and therefore Ψ^*) is unrestricted (except for the overall normalization) we'll get back the Schrödinger equation; for example, if our space is 3-d for a single particle then

$$0 = \frac{\delta D}{\delta \Psi^*(\boldsymbol{r})} = \langle \boldsymbol{r}|H\Psi\rangle - \mu\langle \boldsymbol{r}|\Psi\rangle, \tag{11.8}$$

i.e.,

$$H\Psi = \mu\Psi \tag{11.9}$$

so in that case μ, the Lagrange multiplier, has the interpretation that it is the energy of the w.f.

Note that if the w.f. is restricted (as we'll see later in Hartree-Fock and DFT) then μ is not the total energy. In fact, in DFT it turns out to be the Fermi energy (the HOMO or LUMO or middle of both, depending on the situation).

Parameter dependence

Formally, we can think of the variational principle with a "restricted" w.f., as relating to a w.f. that depends on a parameter set, $\boldsymbol{\lambda}$ (i.e., $\lambda_1, \lambda_2, ...$; the number of parameters can be finite or infinite and they can even be continuous). Also, here the parameters are not necessarily "small", unlike λ in the previous section. Then $\psi = \psi(\boldsymbol{r};\boldsymbol{\lambda})$, so we'll denote its ket as $|\psi(\boldsymbol{\lambda})\rangle$, so

$$D = \langle\psi(\boldsymbol{\lambda})|H|\psi(\lambda)\rangle - \mu\left(\langle\psi(\boldsymbol{\lambda})|\psi(\boldsymbol{\lambda})\rangle - 1\right) \tag{11.10}$$

and the best set of $\boldsymbol{\lambda}$ is found by extremizing

$$\frac{\partial D}{\partial \boldsymbol{\lambda}} = 0 \rightarrow \text{Best } \lambda \tag{11.11}$$

No nodes for ground-states

We can use the minimum energy principle to prove that the ground-state w.f. cannot have nodes, i.e., its sign is fixed (all positive, or multiply by -1, to get all negative). Let's prove for a single-particle in 1 dimension.

Let's do a proof by contradiction.

Say a 1D w.f. $\psi(x)$ did have a node at x_0, such that at x close to but below x_0 it would be positive, and for $x > x_0$ (but nearby) it would be negative.

And assume for simplicity that $b \equiv \frac{d\psi}{dx}|_{x=x_0} \neq 0$.

We recall that

$$\langle\psi|H|\psi\rangle = \langle\psi|T|\psi\rangle + \langle\psi|V|\psi\rangle = \frac{1}{2M}\int \left|\frac{d\psi(x)}{dx}\right|^2 dx + \int V(x)\left|\psi(x)\right|^2 dx. \tag{11.12}$$

Then take a small region (of size labeled as $2\delta x$) around x_0, i.e., $[x_0 - \delta x, x_0 + \delta x]$. The size is assumed to be so small that the w.f. behaves linearly there, $\psi(x) \simeq \psi(x_0) + \frac{d\psi}{dx}|_{x=x_0} \cdot (x - x_0) = 0 + b \cdot (x - x_0)$, so $\max(|\psi(x)|)$ in that region is $b \cdot \delta x$.

Then:

- The wavefunction is tiny in this $[x_0 - \delta x, x_0 + \delta x]$ region (maximum absolute size: $b \cdot \delta x$), and therefore the contribution of the potential energy times the w.f. squared to that region would be very small,

$$\left| \int_{x_0-\delta x}^{x_0+\delta x} V(x)|\psi(x)|^2 dx \right| \leq V(x_0) \cdot |b\delta x|^2 \cdot (2\delta x) \sim O(\delta x^3) \qquad (11.13)$$

- while the kinetic energy contribution would be much larger (and positive), since in this region $\frac{d\psi(x)}{dx} \sim \frac{d\psi(x)}{dx}|_{x=x_0} = b$ is non-zero

$$\frac{\hbar^2}{2m_e} \int_{x_0-\delta x}^{x_0+\delta x} \left| \frac{d\psi(x)}{dx} \right|^2 dx \simeq \frac{\hbar^2}{2m_e} \cdot |b|^2 \cdot (2\delta x) \sim O(\delta x). \qquad (11.14)$$

So we're going to construct a new w.f., that has a much smaller kinetic energy (and possibly slightly higher potential energy, but that's not going to matter). We'll do it in two stages.

1. Construct a continuous but not smooth w.f. that's "peaked" (see figure in class)

$$\psi'(x) = \begin{cases} \psi(x) & (x \leq x_0) \\ -\psi(x) & (x \geq x_0) \end{cases} \qquad (11.15)$$

2. This w.f. is then "smoothed" $\psi'(x) \rightarrow \psi''(x)$, where the "sharp corner" at $x = x_0$ is smoothed out over a region of size 2δ. This reduces the K.E. by $O(\delta)$, and potentially increases the P.E. by $O(\delta^3)$. The w.f. also needed to be renormalized, but that's an $O(\delta^3)$ effect.

So the main effect is that the K.E. was reduced, and therefore ψ'' has a lower associated energy than ψ, in contradiction to our initial assumption that ψ (with the nodes) has the lowest energy.

This proof carries over to the many-particle case, but not for fermions since those have to change sign upon particle exchange.

12 Many-electron Hamiltonian and Wavefunction: General

We now enter the last part of the course, dealing with the Coulombic many body system, i.e., the energy of many electron systems. Here we'll study the Hartree-Fock and DFT approaches, and also learn the language of more sophisticated approaches. At least initially we'll follow along Baer's notes (with abbreviations and modifications), and for the latter part also read Schatz and Ratner's Advanced Q. M. for Chem., especially Chapter 6.

Wavefunction Notation:

Starting this chapter, upper case Greek letters (essentially only Ψ) would be reserved to many-electron wavefunctions; lower case Greek letters (ψ, η, ϕ) would be reserved to one-electron wavefunctions.

The many-electron Coulomb Hamiltonian

The starting point is the N-electron Coulomb Hamiltonian

$$\hat{H} = \hat{H}_{1B} + \hat{U} \tag{12.1}$$

where

$$\hat{H}_{1B} = \hat{T} + \hat{V} \tag{12.2}$$

(don't confuse V for potential with V for volume...) and where the kinetic energy is

$$\hat{T} = -\frac{\hbar^2}{2m_e}\left(\frac{\partial^2}{\partial \boldsymbol{r}_1^2} + \frac{\partial^2}{\partial \boldsymbol{r}_1^2} + ... + \frac{\partial^2}{\partial \boldsymbol{r}_N^2}\right), \tag{12.3}$$

while the electron-nuclear potential energy is

$$V(\boldsymbol{r}_1, ..., \boldsymbol{r}_N) = v\left(\boldsymbol{r}_1\right) + ...v\left(\boldsymbol{r}_N\right), \tag{12.4}$$

where in atoms (recall e=1 for us):

$$v(\boldsymbol{r}) = -\sum_{j=1}^{N_{\mathrm{NUC}}} \frac{Z}{|\boldsymbol{r} - \boldsymbol{R}_j|}. \tag{12.5}$$

In addition, the 2-body interaction is

$$U = \frac{1}{2}\sum_{i \neq j} u(\boldsymbol{r}_i, \boldsymbol{r}_j) \tag{12.6}$$

$$u(\boldsymbol{r}_i, \boldsymbol{r}_j) = \frac{1}{|\boldsymbol{r}_i - \boldsymbol{r}_j|}, \tag{12.7}$$

and we'll ignore spin-orbit interactions.

We'll label V as a **one-body** term, and U is a **2-body** term. We're generally lax about adding the "hat" for operators, we'll do it where things are otherwise unclear.

The $i \equiv \boldsymbol{r}_i, s_i$ notation

We will often label coordinates and w.f. very strange (but economical – which eventually makes life easier!) notation:

$$\Psi(1,2,...,N) \equiv \Psi(\boldsymbol{r}_1, s_1, \boldsymbol{r}_2, s_2...., \boldsymbol{r}_N, s_N), \tag{12.8}$$

where "1" as coordinate is abbreviation for the spatial coordinate of the first electron and its spin, i.e.,

$$"1" \equiv \boldsymbol{r_1,} s_1 \tag{12.9}$$

"2" means similarly "$\boldsymbol{r_2,} s_2$", etc.

Pauli principle

The Pauli principle states that any electrons w.f. must satisfy

$$\Psi(....,i,....,j,...) = -\Psi(....,j,....,i,...), \tag{12.10}$$

i.e.,

$$\Psi(....,\boldsymbol{r}_i, s_i,....,\boldsymbol{r}_j, s_j,...) = -\Psi(....,\boldsymbol{r}_j, s_j,....,\boldsymbol{r}_i, s_i,...), \tag{12.11}$$

i.e., if we replace particle i and j, including BOTH their positions and their spins, the w.f. changes sign.

The expectation value of the Hamiltonian

The density operator: many-electrons case

In the early chapters we defined the single-electron density operator, $\hat{n}(\boldsymbol{r})$, and we got the result $\langle \boldsymbol{r'} | \hat{n}(\boldsymbol{r}) | \psi_{1e} \rangle = \psi_{1e}(\boldsymbol{r}) \delta(\boldsymbol{r} - \boldsymbol{r'})$, where the "1e" subscript refers to a single-electron wavefunction. We now extend it to a multi-electron case. Specifically,

$$\hat{n}(\boldsymbol{r})\Psi(1,2,..,N) = \sum_{i=1}^{N} \delta(\boldsymbol{r} - \boldsymbol{r}_i)\Psi(1,2,..,N), \tag{12.12}$$

while a more formal way is to write:

$$\hat{n}(\boldsymbol{r}) = \sum_{i=1}^{N} |\boldsymbol{r}(electron\, i)\rangle\langle \boldsymbol{r}(electron\, i))|, \tag{12.13}$$

where "$|\boldsymbol{r}(electron\, i)\rangle$" implies a ket for electron i only.

Note that the density operator acts on the full w.f.

The expectation value of the density operator is nothing more than the **one-electron spatial density** (which we usually abbreviate simply as the **electron-density or just "density"**)

$$n(\boldsymbol{r}) \equiv \langle\Psi|\hat{n}(\boldsymbol{r})|\Psi\rangle = \int |\Psi(,2,..,i,.,N)|^2 \sum_{i=1}^{N} \delta(\boldsymbol{r}-\boldsymbol{r}_i) d1d2d3....dN, \qquad (12.14)$$

and since the w.f. is completely antisymmetric, the w.f. squared is completely symmetric, so each "delta" in the sum would give the same contribution. We therefore replace the N delta functions by a single delta function over the 1st coordinate, multiplied by N.

$$n(\boldsymbol{r}) = N \int \delta(\boldsymbol{r}-\boldsymbol{r}_1)\, |\Psi(1,2,..,i,.,N)|^2\, d1d2d3....dN. \qquad (12.15)$$

Further, we'll expand "$d1$" to explicitly denote the integration over position and spin, $d1 = d\boldsymbol{r}\, ds_1$ and expand "1" in the wavefunction to be $\boldsymbol{r_1,}s_1$, so

$$n(\boldsymbol{r}) = N \int \delta(\boldsymbol{r}-\boldsymbol{r}_1)\, |\Psi(\boldsymbol{r}_1, s_1, 2,..,i,.,N)|^2\, d\boldsymbol{r}_1 ds_1 d2d3....dN, \qquad (12.16)$$

i.e., the final expression is (replacing s_1 by s)

$$n(\boldsymbol{r}) = N \int |\Psi(\boldsymbol{r}, s, 2,..,i,.,N)|^2\, ds\, d2d3....dN. \qquad (12.17)$$

The spatial density $n(\boldsymbol{r})$ is the probability to find ANY of the N electrons at $\boldsymbol{r}$ regardless of the positions of the other electrons.

Note that since $\langle\Psi|\Psi\rangle = 1$, we get the normalization constant.

$$\int n(\boldsymbol{r}) d\boldsymbol{r} = N \qquad (12.18)$$

i.e., the total electron density represents N electrons, as expected.

One-body potential in terms of the density operator

We can further write the one-body potential in terms of the density operator

$$\hat{V} = \int v(\boldsymbol{r}) \hat{n}(\boldsymbol{r}) d\boldsymbol{r}. \qquad (12.19)$$

This eq. would prove invaluable in DFT. To prove it, act with both sides on a general w.f. The L.H.S. gives

$$\langle \boldsymbol{r}_1, \boldsymbol{r}_2, ..., \boldsymbol{r}_N|\hat{V}|\Psi\rangle = (v(\boldsymbol{r}_1) + v(\boldsymbol{r}_2) + ...v(\boldsymbol{r}_N))\, \Psi(\boldsymbol{r}_1, \boldsymbol{r}_2, ..., \boldsymbol{r}_N), \qquad (12.20)$$

while the R.H.S. gives, once we recall what's $\hat{n}(\boldsymbol{r})|\Psi\rangle$ is (we color some parts for clarity):

$$\langle \boldsymbol{r}_1, \boldsymbol{r}_2, ..., \boldsymbol{r}_N| \int v(\boldsymbol{r})\hat{n}(\boldsymbol{r})d\boldsymbol{r}|\Psi\rangle = \int v(\boldsymbol{r})\langle \boldsymbol{r}_1, \boldsymbol{r}_2, ..., \boldsymbol{r}_N|\hat{n}(\boldsymbol{r})|\Psi\rangle d\boldsymbol{r} =$$
$$\int v(\boldsymbol{r}) \left(\delta(\boldsymbol{r}-\boldsymbol{r}_1) + \delta(\boldsymbol{r}-\boldsymbol{r}_2) + ...\delta(\boldsymbol{r}-\boldsymbol{r}_2)\right) \Psi(\boldsymbol{r}_1, \boldsymbol{r}_2, ..., \boldsymbol{r}_N)d\boldsymbol{r}$$
$$= \left(v(\boldsymbol{r}_1) + v(\boldsymbol{r}_2) + ...v(\boldsymbol{r}_N)\right) \Psi(\boldsymbol{r}_1, \boldsymbol{r}_2, ..., \boldsymbol{r}_N), \quad (12.21)$$

Q.E.D.

The one-body density matrix

For the kinetic energy later, and also for the exchange (also later), we need to define now the **spatial one-body density matrix** (or just "density matrix" below), in analogous steps to the definition of the density above.

Formally the one-body density-matrix operator is defined

$$\hat{\rho}(\boldsymbol{r}, \boldsymbol{r'}) = \sum_{i=1}^{N} |\boldsymbol{r}(electron\ i)\rangle\langle \boldsymbol{r'}(electron\ i))| \quad (12.22)$$

and in practice we can show as before that, defining the one-body density matrix (not operator!) as

$$\rho(\boldsymbol{r}, \boldsymbol{r'}) \equiv \langle\Psi|\hat{\rho}(\boldsymbol{r}, \boldsymbol{r'})|\Psi\rangle, \quad (12.23)$$

we get after similar manipulations that

$$\rho(\boldsymbol{r}, \boldsymbol{r'}) \equiv N \int \Psi(\boldsymbol{r}, s, 2, .., i, ., N)\Psi^*(\boldsymbol{r'}, s, 2, .., i, ., N)ds_1 d2d3....dN. \quad (12.24)$$

Note that

$$\rho(\boldsymbol{r}, \boldsymbol{r}) \equiv \rho\left(\boldsymbol{r}, \boldsymbol{r'} = \boldsymbol{r}\right) = n(\boldsymbol{r}). \quad (12.25)$$

One-body potential expectation value and electron density

Since H involves only 1-body and 2-body terms, the expectation value of a specific wavefunction Ψ is:

$$\langle\Psi|H|\Psi\rangle = \langle\Psi|\hat{T}|\Psi\rangle + \langle\Psi|\hat{V}|\Psi\rangle + \langle\Psi|U|\Psi\rangle \quad (12.26)$$

where we saw that $\hat{V} = \int v(\boldsymbol{r})\hat{n}(\boldsymbol{r})d\boldsymbol{r}$, so the one-body expectation value is simply:

$$\langle\hat{V}\rangle \equiv \langle\Psi|\hat{V}|\Psi\rangle = \int v(\boldsymbol{r})n(\boldsymbol{r})d\boldsymbol{r}. \quad (12.27)$$

i.e., the expectation of the potential becomes an **integral over the total electron density.**

Kinetic energy expectation value and one-body density matrix

It is straightforward to show, using similar steps to above, that the kinetic energy is related to the behavior of the density matrix "near the diagonal"

$$\langle\Psi|\hat{T}|\Psi\rangle = \frac{\hbar^2}{2m_e}\int \frac{\partial}{\partial \boldsymbol{r}}\frac{\partial}{\partial \boldsymbol{r'}}\rho(\boldsymbol{r'},\boldsymbol{r})\bigg|_{\boldsymbol{r'}=\boldsymbol{r}} d\boldsymbol{r}. \tag{12.28}$$

2-body density and the electron-electron interaction

Finally, we define the 2-body density. The associated operator is:

$$\hat{Y}(\boldsymbol{r},\boldsymbol{r}') = \hat{n}(\boldsymbol{r})\hat{n}(\boldsymbol{r'}) - \delta(\boldsymbol{r}-\boldsymbol{r'})\hat{n}(\boldsymbol{r}) \tag{12.29}$$

i.e., the operator of localizing any electron at one position and any other electron at another position (the "minus delta" at the end guarantees that we wont measure the same electron twice).

Similar derivations to the ones we did lead to:

$$Y(\boldsymbol{r},\boldsymbol{r'}) = N(N-1)\int |\Psi(\boldsymbol{r},s,\boldsymbol{r'},s',3,4,...,N)|^2\, ds\, ds'\, d3....dN \tag{12.30}$$

Note that $Y(\boldsymbol{r},r')$ is the probability of finding ANY electron at $\boldsymbol{r}_1$ with ANY OTHER electron at $\boldsymbol{r}_2$.

Using the same techniques as above we can show that

$$\langle\Psi|U|\Psi\rangle = \frac{1}{2}\int \frac{Y(\boldsymbol{r},\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r'}|} d\boldsymbol{r}d\boldsymbol{r'}. \tag{12.31}$$

Warning: don't confuse $\rho(\boldsymbol{r},\boldsymbol{r}')$ and $Y(\boldsymbol{r},\boldsymbol{r}')$.

Summary: overall expectation value $\langle H\rangle$

$$\langle\Psi|H|\Psi\rangle = \frac{\hbar^2}{2m_e}\int \frac{\partial}{\partial \boldsymbol{r}}\frac{\partial}{\partial \boldsymbol{r'}}\rho(\boldsymbol{r'},\boldsymbol{r})\bigg|_{\boldsymbol{r'}=\boldsymbol{r}} d\boldsymbol{r} + \int v(\boldsymbol{r})n(\boldsymbol{r})d\boldsymbol{r} + \frac{1}{2}\int \frac{Y(\boldsymbol{r},\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r'}|} d\boldsymbol{r}d\boldsymbol{r'}. \tag{12.32}$$

Solving for the lowest energy:

There are many approaches to find the lowest energy or a good approximation to it. We'll talk mostly about two in the next few lectures; qualitatively (and somewhat inaccurately):

- One method would be to take the eqs. above and insert an approximate w.f. for which all these quantities are reasonable, and then look for the best such approximate function. The one we will use is a Slater determinant, and the lowest-energy Slater determinant is called the Hartree-Fock solution.

- Another would be to approximate the terms in the Hamiltonian – this is the regime of DFT, density functional theory.
- Other techniques would be perturbation theory, and further more sophisticated methods – we'll briefly comment on these at the end.

13 Slater Determinants

Slater determinants are "simple w.f." that satisfy the Pauli antisymmetry relation and are therefore the basic ingredients in electronic structure. Getting used to them involves lots of notation. So to warm up, let's consider first instead the case of:

Warmup: Non-identical particles and Hartree products.

Say for example we want to talk about a system of different particles, e.g., hydrogen (electron+proton), when we now consider the quantum nature of the proton. The simplest quantum function of the two coordinates, **r** for the electron and **R** for the proton, will be a product (labeled usually a Hartree product)

$$\phi(\boldsymbol{r})\zeta(\boldsymbol{R}), \tag{13.1}$$

where ϕ and ζ are some functions of the electron and proton coordinates. Further, if we have a complete basis of the electrons, $\phi_i(\boldsymbol{r})$ and a similar complete basis of the protons $\zeta_j(\boldsymbol{R})$ then we can make a complete basis of the two-particle system:

$$\Psi_{ij}(\boldsymbol{r},\boldsymbol{R}) = \phi_i(\boldsymbol{r})\zeta_j(\boldsymbol{R}), \tag{13.2}$$

which we could write in a bra-ket notation as

$$|\Psi_{ij}\rangle = |\phi_i\zeta_j\rangle.$$

The true wavefunction for such a system could be written as

$$|\Psi\rangle = \sum_{ij} c_{ij}|\Psi_{ij}\rangle. \tag{13.3}$$

Such an expansion is called a CI expansion (CI means Configuration Interaction). Note that it may not be very economical, since in each $|\Psi_{ij}\rangle$ the particles move independently (while in practice we expect the electrons to move with the protons). Therefore, we may need may ij terms, i.e., many products $\phi_i(\boldsymbol{r})\zeta_j(\boldsymbol{R})$, but at least it is easy to understand.

So let's move to electrons:

Fermions: 2-electron wavefunctions.

Specific examples for 2-electrons w.f.

Ground-state Helium

Take Helium, assign each electron to 1s, and put the electrons in spin-zero combination; then (casual notation)

$$\Psi = \psi_{1s}(\boldsymbol{r}_1)\psi_{1s}(\boldsymbol{r}_2)\frac{|\alpha\beta\rangle - |\beta\alpha\rangle}{\sqrt{2}}, \tag{13.4}$$

(This is of course just an approximation to the true wavefunction, which has extra terms due to correlation).

In a more "formally correct" and tedious notation (which we'll avoid generally), it will be

$$\Psi(\boldsymbol{r}_1, s_1, \boldsymbol{r}_2, s_2) = \psi_{1s}(\boldsymbol{r}_1)\psi_{1s}(\boldsymbol{r}_2)\frac{\delta_{s_1,\frac{1}{2}}\delta_{s_2,-\frac{1}{2}} - \delta_{s_1,-\frac{1}{2}}\delta_{s_2,\frac{1}{2}}}{\sqrt{2}}. \tag{13.5}$$

Excited Helium

Another example: H_e^*, where one electron (with the "down" spin) is excited to $2p_y$:

$$\Psi = \frac{\psi_{1s}(\boldsymbol{r}_1)\psi_{2p_y}(\boldsymbol{r}_2)|\alpha\beta\rangle - \psi_{2p_y}(\boldsymbol{r}_1)\psi_{1s}(\boldsymbol{r}_2)|\beta\alpha\rangle}{\sqrt{2}} \tag{13.6}$$

2-electron w.f.: Antisymmetrization

Since electrons are fermions, their wavefunction must be antisymmetric w.r.t. particle exchange. Say we want to take one electron in an orbital "η" and another in an orbital "ζ". We can make a 2-electron wavefunction from them. We will label this function in a seemingly very strange (but economical – which eventually makes life easier!) notation:

$$\Psi(1,2) = \frac{1}{\sqrt{2}}\left(\eta(1)\zeta(2) - \zeta(1)\eta(2)\right) \tag{13.7}$$

where "1" as coordinate is abbreviation for the spatial coordinate of the first electron and its spin, i.e.,

$$\eta(1) \equiv \eta(\boldsymbol{r_1,}s_1) \tag{13.8}$$

and "2" means similarly "$\boldsymbol{r_2,}s_2$". Note the antisymmetrization above, which results in the desired $\Psi(1,2) = -\Psi(2,1)$, i.e., if we exchange at the same time both the spins and the spatial coordinates of the two electrons, the w.f. reverses sign.

Revisiting Helium Let's consider again the Slater determinant that represents two electrons in Helium. Say each is in a 1s-type orbital denoted as $\phi_0(\boldsymbol{r})$, but one is in a spin up (α) and the other is down (β). We write (in a fairly relaxed non-rigiorus notation, as usual):

$$\eta(1) = \phi_0(\boldsymbol{r_1})|\alpha\rangle$$

$$\bar{\eta}(1) = \phi_0(\boldsymbol{r_1})|\beta\rangle$$

So

$$\Psi(1,2) = \frac{1}{\sqrt{2}}\left(\eta(1)\bar{\eta}(2) - \bar{\eta}(1)\eta(2)\right) = \\ \frac{1}{\sqrt{2}}\left(\phi_0(\boldsymbol{r_1})|\alpha\rangle\phi_0(\boldsymbol{r_2})|\beta\rangle - \phi_0(\boldsymbol{r_1})|\beta\rangle\phi_0(\boldsymbol{r_1})|\alpha\rangle\right) \quad (13.9)$$

i.e.,

$$\Psi(1,2) = \phi_0(\boldsymbol{r_1})\phi_0(\boldsymbol{r_2})\frac{|\alpha\rangle|\beta\rangle - |\beta\rangle|\alpha\rangle}{\sqrt{2}} \quad (13.10)$$

which is most often written as

$$\Psi(1,2) = \phi_0(\boldsymbol{r_1})\phi_0(\boldsymbol{r_2})\frac{|\alpha\beta - \beta\alpha\rangle}{\sqrt{2}} \quad (13.11)$$

or even more compactly:

$$|\Psi\rangle = \frac{1}{\sqrt{2}}|\eta\bar{\eta} - \bar{\eta}\eta\rangle. \quad (13.12)$$

(Of course, the true ground-state w.f. of Helium is a CI expansion with more configurations, as in Eq. (13.3), but this Slater determinant is the dominant part of the true ground-state wavefunction).

An arbitrary Slater determinant for excited-state Helium As another example, consider a complete basis for the electrons (for the combined spatial and spin coordinate), which we'll denote as η_1,η_2,η_3.... . Let's pick two orbitals for the electrons which are denoted as $|\eta_8\rangle$ and $|\eta_{11}\rangle$ (why 8 and 11 and not just 1 and 2? I want you to remember that we can use any orbitals; the electrons are denoted as 1, 2,... but we can use any orbitals we want). Then

$$|\Psi_{8,11}\rangle = \frac{1}{\sqrt{2}}|\eta_8\eta_{11} - \eta_{11}\eta_8\rangle \quad (13.13)$$

i.e.,

$$\Psi_{8,11}(1,2) = \frac{1}{\sqrt{2}}\left(\eta_8(1)\eta_{11}(2) - \eta_{11}(1)\eta_8(2)\right). \quad (13.14)$$

Time for yet another "simplification": we'll often denote this antisymmetric product of two basis-set members as "simply"

$$|8,11\rangle \equiv |\Psi_{8,11}\rangle = \frac{1}{\sqrt{2}}|\eta_8\eta_{11} - \eta_{11}\eta_8\rangle. \quad (13.15)$$

This last notation, $|8,11\rangle$ assumes that we have a specific basis in mind. **But it can be confusing (does it mean antisymmetrized w.f., like here; or does it not?), and the precise meaning depends on the context,** so we'll avoid it for now.

Slater Determinants: example

Since

$$\det\begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc \tag{13.16}$$

the expression above can be rewritten as

$$\Psi_{8,11}(1,2) = \frac{1}{\sqrt{2}}\det\begin{pmatrix} \eta_8(1) & \eta_{11}(1) \\ \eta_8(2) & \eta_{11}(2) \end{pmatrix}, \tag{13.17}$$

and we often abbreviate it as $|\Psi_{8,11}\rangle = \det(\eta_8, \eta_{11})$ — this is a definition of the RHS, and the 1/sqrt(2) factor is absorbed into the definition.

This generalizes to more electrons; e.g., a 3-electron Slater determinant made from three basis functions $\eta_6, \eta_8, \eta_{11}$, will be

$$\Psi_{6,8,11}(1,2,3) = \frac{1}{\sqrt{3!}}\det\begin{pmatrix} \eta_6(1) & \eta_8(1) & \eta_{11}(1) \\ \eta_6(2) & \eta_8(2) & \eta_{11}(2) \\ \eta_6(3) & \eta_8(3) & \eta_{11}(3) \end{pmatrix}, \tag{13.18}$$

or simply $|6,8,11\rangle$ or $|\Psi_{6,8,11}\rangle$ or $\det(\eta_6, \eta_8, \eta_{11})$.

Note that the determinant <u>does not change when we take linear combination of orbitals</u>. This is very important below.

General and closed-shell determinants

A general Slater determinant made from N orbitals (N is the number of electrons here) that will be numbered $\eta_1, \eta_2, ..., \eta_N$ would be denoted as:

$$|1,...,N\rangle \text{ or } |\Psi_{1,2,..,N}\rangle \text{ or } |\det(\eta_1\eta_2...\eta_N),$$

and the meaning would be

$$\Psi_{1,....,N}(1,2,...,N) = \frac{1}{\sqrt{N!}}\det\begin{pmatrix} \eta_1(1) & ... & .... & \eta_N(1) \\ ... & ... & ... & ... \\ ... & ... & & ... \\ \eta_1(N) & & & \eta_N(N) \end{pmatrix}. \tag{13.19}$$

Could be highly confusing, practice makes perfect.

Closed shell-determinants are ones where for every up-spin orbital there's a similar down-spin one; in that case, denoting the spatial orbitals by ϕ_n, we get:

$$\eta_1(1) = \phi_1(\boldsymbol{r_1})|\alpha\rangle$$

$$\eta_2(1) = \phi_1(\boldsymbol{r_1})|\beta\rangle$$

$$\eta_3(1) = \phi_2(\boldsymbol{r_1})|\alpha\rangle$$

$$\eta_3(1) = \phi_2(\boldsymbol{r_1})|\beta\rangle$$

$$...$$

$$\eta_{N-1}(1) = \phi_{\frac{N}{2}}(\boldsymbol{r_1})|\alpha\rangle$$

$$\eta_N(1) = \phi_{\frac{N}{2}}(\boldsymbol{r_1})|\beta\rangle.$$

It is OK to consider only orthogonal orbitals

Recall you learned in linear algebra that the value of a determinant does not change when we linearly combine a column with another; for example

$$\det(\eta_1, \eta_2, \eta_3, ..., \eta_N) = \det(\eta_1 + c\eta_3, \eta_2, \eta_3, ..., \eta_N).$$

Therefore, we can always do linear transformation of the states, till they are orthogonal:

$$\langle \eta_j | \eta_k \rangle = \delta_{jk}. \tag{13.20}$$

Note that if each state has a definite spin, i.e., is either $|\alpha\rangle$ or $|\beta\rangle$ (the usual case but there are exceptions) then the spatial orbitals within each spin are orthogonal. For closed-shell determinants as defined above the orthogonality condition becomes simply:

$$\langle \phi_j | \phi_k \rangle = \int \phi_j^*(\boldsymbol{r})\, \phi_k(\boldsymbol{r}) d\boldsymbol{r} = \delta_{jk}. \tag{13.21}$$

One-body density and density matrix for Slater determinants

One-body density for a Slater-Determinant

For a Slater determinant, the density is very simple

$$n(\boldsymbol{r}) = \sum_{i=1}^{N} \sum_{s} |\eta(\boldsymbol{r}, s)|^2 \tag{13.22}$$

(here we explicitly sum over spins, which is what we really mean when we write that we integrate over them). In lieu of a full blown proof, let's show this for a 2-electron system, with a Slater determinant made from 2 orthogonal orbitals, η and ζ.

$$\Psi(1,2) = \frac{\eta(1)\zeta(2) - \zeta(2)\eta(1)}{\sqrt{2}}, \tag{13.23}$$

so

$$n(\boldsymbol{r}) = N\int |\Psi(1,2,..,i,.,N)|^2\, ds_1 d2 = \frac{2}{2}\sum_{s_1}\int |\eta(1)\zeta(2)-\zeta(1)\eta(2)|^2\, d2 =$$

$$= \sum_{s_1}\int \left(\eta^*(\boldsymbol{r}_1,s_1)\zeta^*(2)-\zeta^*(\boldsymbol{r}_1,s_1)\eta^*(2)\right)\left(\eta(\boldsymbol{r}_1,s_1)\zeta(2)-\zeta(\boldsymbol{r}_1,s_1)\eta(2)\right) d2. \tag{13.24}$$

where the $\frac{2}{2}$ term is first because N=2, and second because the Slater determinant squared gives a factor of $\frac{1}{N!}$.

The cross term vanish, since $\int \zeta^*(2)\eta(2)d2$=0, so we're left with the diagonal terms

$$n(\boldsymbol{r}) = \sum_{s_1} |\eta(\boldsymbol{r}_1,s_1)|^2 \int |\zeta(2)|^2\, d2 + \ (\eta \leftrightarrow \zeta). \tag{13.25}$$

But

$$\int |\zeta(2)|^2\, d2 = 1, \tag{13.26}$$

so, removing the "1" subscripts:

$$n(\boldsymbol{r}) = \sum_{s}\left(|\eta(\boldsymbol{r},s)|^2 + |\zeta(\boldsymbol{r},s)|^2\right), \tag{13.27}$$

as stipulated.

Closed-shell density

For a closed-shell Slater-determinant, where each electron-up has a corresponding electron down at the same spatial orbital, the density becomes

$$n(\boldsymbol{r}) = 2\sum_{i=1}^{\frac{N}{2}} |\phi_i(\boldsymbol{r})|^2. \tag{13.28}$$

One-body density matrix for a closed-shell Slater-Determinant

The closed-shell Slater determinant density matrix is similarly shown to be

$$\rho(\boldsymbol{r},\boldsymbol{r'}) = 2\sum_{i=1}^{\frac{N}{2}} \phi_i(\boldsymbol{r})\phi_i^*(\boldsymbol{r'}). \tag{13.29}$$

Interestingly, this is (except for the factor of 2) just the form of a **projection** operator to the space of the occupied spatial orbitals (i.e., the $\phi_1,..,\phi_{\frac{N}{2}}$). We talked about projection operators earlier, so I remind you:

Formally, given a **general one-electron spatial wavefunction** $|\psi\rangle$ (in the language of the beginning of the course), then ρ projects from ψ the ground-state components

$$\frac{\rho}{2}|\psi\rangle = \sum_{n=1}^{\frac{N}{2}} c_n|\phi_n\rangle, \quad c_n = \langle\phi_n|\psi\rangle \tag{13.30}$$

Also, it is easy to see that formally the **density matrix from the Slater-determinant** fulfills:

$$\left(\frac{\rho}{2}\right)^2 = \frac{\rho}{2} \tag{13.31}$$

where the product here is in a matrix sense, i.e.,

$$\int \rho(\boldsymbol{r},\boldsymbol{r'})\rho(\boldsymbol{r'},\boldsymbol{r''})d\boldsymbol{r'} = 2\rho(\boldsymbol{r},\boldsymbol{r''}). \tag{13.32}$$

Two-body density for Slater determinants

It is painstaking, but we can show that the two-body density for a general Slater determinant $\det(\eta_1, ..., \eta_N)$ has the form

$$Y(\boldsymbol{r},\boldsymbol{r'}) = \sum_{s,s'}\sum_{i,j} \eta_i^*(\boldsymbol{r},s)\eta_j^*(\boldsymbol{r'},s')\left(\eta_i(\boldsymbol{r},s)\eta_j(\boldsymbol{r'},s') - \eta_j(\boldsymbol{r},s)\eta_i(\boldsymbol{r'},s')\right). \tag{13.33}$$

Note that we don't need to explicitly restrict the terms to $i \neq j$ since the $i = j$ term automatically vanishes.

For a closed shell determinant, the summation over spins is simple and gives

$$Y(\boldsymbol{r},\boldsymbol{r'}) = \sum_{i,j=1}^{\frac{N}{2}} \phi_i^*(\boldsymbol{r})\phi_j^*(\boldsymbol{r'})\left(4\phi_i(\boldsymbol{r})\phi_j(\boldsymbol{r'}) - 2\phi_j(\boldsymbol{r})\phi_i(\boldsymbol{r'})\right). \tag{13.34}$$

The factor of "+4" above comes from the summation over 2 values of " s " times 2 values of " s' ". The factor of -2 (and not -4) derives from the cross-term in Eq. (13.33), where only the $s = s'$ term contributes.

The "+4" term in the two body density is easily seen to be just the density squared, while the "-2" term relates to the density matrix, so (**prove this!) for a closed-shell Slater-determinant, the two-body density is built from the one-body density matrix:**

$$Y(\boldsymbol{r},\boldsymbol{r'}) = n(\boldsymbol{r})n(\boldsymbol{r'}) - \frac{1}{2}|\rho(\boldsymbol{r},\boldsymbol{r'})|^2. \tag{13.35}$$

This is not a general property and relates only to Slater determinants.

Fermi hole

Y is quite interesting at the diagonal:

$$Y(\boldsymbol{r}, \boldsymbol{r'} = \boldsymbol{r}) = n(\boldsymbol{r})^2 - \frac{1}{2}|\rho(\boldsymbol{r}, \boldsymbol{r})|^2 = \frac{n(\boldsymbol{r})^2}{2} \text{ (Closed-Shell Slater Determinant)} \tag{13.36}$$

Contrast this with the true 2-body density matrix, which, for electrons (or any strongly **repulsive** particles) must vanish

$$Y(\boldsymbol{r}, \boldsymbol{r'} = \boldsymbol{r}) = 0 \text{ (True ground-state wavefunciton),} \tag{13.37}$$

since two repulsive electrons can't be at the same place.

So the Slater two-body density is not vanishing at it should; physically this is since in Slater determinant a spin-up electron and a spin-down electron can occupy the same point in space.

BUT at least the Slater determinant two-body density is reduced from the density-density product (i.e., it is not $n(\boldsymbol{r})^2$, only $\frac{n(\boldsymbol{r})^2}{2}$). This reduction leads to better ground-state energies for Slater determinants (i.e., fermions) than in case of bosons where the w.f. is completely symmetric.

The reduction of $Y(\boldsymbol{r}, \boldsymbol{r'} = \boldsymbol{r})$ in a Slater-determinant is (or more precisely: is related to something called) the **Fermi hole.** The (stronger) reduction in the true ground-state two-body density is associated with a **Coulomb hole.**

14 Hartree-Fock theory

We will ask: what's the Slater determinant $|\Psi\rangle$ (for simplicity, closed-shell one) with the lowest $\langle H\rangle$? So first:

$\langle H\rangle$

Recall:

$$\langle H\rangle = \langle\Psi|\hat{T}|\Psi\rangle + \langle\Psi|\hat{V}|\Psi\rangle + \langle\Psi|\hat{U}|\Psi\rangle \equiv E_{1B} + E_{ee} \tag{14.1}$$

where

$$E_{1B} = E_{Kin} + E_{eN} = \langle H_{1B}\rangle \tag{14.2}$$

and we recall $H_{1B} = \hat{T} + \hat{V}$. From above, Eq. (12.32) – and we're back to $\hbar = 1$:

$$E_{Kin} = \frac{1}{2m_e}\int \frac{\partial}{\partial \boldsymbol{r}}\frac{\partial}{\partial \boldsymbol{r'}}\rho(\boldsymbol{r'},\boldsymbol{r})\bigg|_{\boldsymbol{r'}=\boldsymbol{r}} d\boldsymbol{r} \tag{14.3}$$

$$E_{eN} = \int n(\boldsymbol{r})u(\boldsymbol{r})d\boldsymbol{r} \tag{14.4}$$

$$E_{ee} = \frac{1}{2}\int Y(\boldsymbol{r},\boldsymbol{r'})u(\boldsymbol{r},\boldsymbol{r'})dr = \frac{1}{2}\cdot\int \frac{n(\boldsymbol{r})n(\boldsymbol{r'})}{|\boldsymbol{r}-\boldsymbol{r'}|}d\boldsymbol{r}d\boldsymbol{r'} - \frac{1}{4}\int \frac{|\rho(\boldsymbol{r},\boldsymbol{r'})|^2}{|\boldsymbol{r}-\boldsymbol{r'}|}d\boldsymbol{r}d\boldsymbol{r'}. \tag{14.5}$$

Plug in the expressions for n, ρ, D for Slater determinant in terms of the occupied eigenstates as derived in the previous chapter in terms of the orthogonal orbitals and get (with factors of 2,4 due to spin, as mentioned):

$$E_{Kin} = \frac{2}{2m_e}\sum_{i=1}^{N}\int \nabla\phi_i^*(\boldsymbol{r})\nabla\phi_i(\boldsymbol{r})d\boldsymbol{r} = 2\sum_i\langle\phi_i| - \frac{\nabla^2}{2m_e}|\phi_i\rangle \tag{14.6}$$

$$E_{eN} = 2\cdot\sum_{i=1}^{N}\int |\phi_i(\boldsymbol{r})|^2 v(\boldsymbol{r})d\boldsymbol{r} = 2\sum_i\langle\phi_i|v|\phi_i\rangle \tag{14.7}$$

so:

$$E_{1B} \equiv E_{Kin} + E_{eN} = 2\sum_i\langle\phi_i|\hat{h}|\phi_i\rangle \tag{14.8}$$

where

$$\hat{h} \equiv -\frac{\nabla^2}{2m_e} + V_{eN}. \tag{14.9}$$

Further, you can also show from our expressions that

$$E_{ee} = E_H + K, \tag{14.10}$$

where the Coulomb direct energy (so called Hartree term, or classical electrostatic term) is

$$E_H = \frac{1}{2}\int \frac{n(\boldsymbol{r})n(\boldsymbol{r'})}{|\boldsymbol{r}-\boldsymbol{r}'|} d\boldsymbol{r} dr = 2\sum_{ij}\int \frac{|\phi_i(\boldsymbol{r})|^2|\phi_j(\boldsymbol{r}')|^2}{|\boldsymbol{r}-\boldsymbol{r}'|} d\boldsymbol{r} d\boldsymbol{r}', \tag{14.11}$$

while the exchange energy is (bad symbol "K" used historically, don't confuse with K.E.)

$$K = -\frac{1}{4}\int \frac{|\rho(\boldsymbol{r.r'})|^2}{|\boldsymbol{r}-\boldsymbol{r}'|} d\boldsymbol{r} d\boldsymbol{r'} = -\sum_{ij}\int \frac{\phi_i^*(\boldsymbol{r})\phi_j^*(\boldsymbol{r'})\phi_i(\boldsymbol{r'})\phi_j(\boldsymbol{r})}{|\boldsymbol{r}-\boldsymbol{r}'|} d\boldsymbol{r} d\boldsymbol{r}'. \tag{14.12}$$

Minimizing $\langle H \rangle$

We need to minimize $\langle H \rangle$ w.r.t. the spatial orbitals, subject to the constraint $\langle \phi_i | \phi_j \rangle = \delta_{ij}$, i.e., extremize the "Lagrangian"

$$L = \frac{1}{2}\langle H \rangle - C_L \tag{14.13}$$

where the $\frac{1}{2}$ is for convenience (removing the spin factor), C_L are the constraints times Lagrange multipliers,

$$C_L = \sum_{i,j=1}^{N} \lambda_{ij} \left(\langle \phi_i | \phi_j \rangle - \delta_{ij} \right) \tag{14.14}$$

and λ_{ij} are a set of Lagrange multipliers.

Thus, we need find the set of $\phi_i(\boldsymbol{r})$ that fulfill:

$$\frac{\delta L}{\delta \phi_i^*(\boldsymbol{r})} = 0. \tag{14.15}$$

The Fock operator

From above we find

$$\frac{1}{2}\frac{\delta E_{1B}}{\delta \phi_i^*(\boldsymbol{r})} = \hat{h}\phi_i(\boldsymbol{r}), \tag{14.16}$$

$$\frac{1}{2}\frac{\delta E_H}{\delta \phi_i^*(\boldsymbol{r})} = v_H(\boldsymbol{r})\phi(\boldsymbol{r}), \tag{14.17}$$

where we defined the Hartree potential, i.e., the classical Coulombic potential due to the density

$$v_H(\boldsymbol{r}) = \int \frac{n(\boldsymbol{r'})}{|\boldsymbol{r}-\boldsymbol{r'}|} d\boldsymbol{r'}, \tag{14.18}$$

and

$$\frac{1}{2}\frac{\delta K}{\delta \phi_i^*(\boldsymbol{r})} = \langle \boldsymbol{r}|\hat{k}\phi_i\rangle \tag{14.19}$$

where we defined the Fock exchange operator

$$\langle \boldsymbol{r}|\hat{k}\phi_i\rangle = \sum_j \phi_j(\boldsymbol{r}) \int \frac{\phi_j^*(\boldsymbol{r'})\phi_i(\boldsymbol{r'})}{|\boldsymbol{r}-\boldsymbol{r'}|} d\boldsymbol{r'} = \int \frac{\rho(\boldsymbol{r},\boldsymbol{r'})\phi_i(\boldsymbol{r'})}{|\boldsymbol{r}-\boldsymbol{r'}|} d\boldsymbol{r'}, \tag{14.20}$$

In addition, differentiating the constraint term leads to

$$\frac{\delta C_L}{\delta \phi_i^*(\boldsymbol{r})} = \sum_{j=1}^{N} \lambda_{ij}\phi_j(\boldsymbol{r}) \tag{14.21}$$

Putting it all together converts $\frac{\delta L}{\delta \phi_i^*(\boldsymbol{r})} = 0$ to

$$\hat{f}\phi_i = \sum_{j=1}^{N} \lambda_{ij}\phi_j(\boldsymbol{r}) \tag{14.22}$$

where the Fock operator

$$\hat{f} = \hat{h} + v_H + \hat{k}, \tag{14.23}$$

is a one-body operator, made from:

- combining the underlying one-body operator $\hat{h}$ (kinetic+ e-N attraction) with:
- the one-body Hartree and exchange terms that result from the "folding" of the two-body Coulomb interaction and the Hartree-Fock two-body density.

Rotating the orbitals

Now comes a trick. We have freedom to "rotate" the orbitals, i.e., to take orthogonal linear combination of them, $\tilde{\phi}_i = \sum_j B_{ij}\phi_j$ where $B^+B = I$ – since that will not change the Slater determinant, and therefore will not change the Fock operator. So we'll rotate them till the Fock equation is diagonal, i.e., till $\hat{f}\tilde{\phi}_i = \varepsilon_i\tilde{\phi}_i$, (question for you: what should B fulfill?).

We can then remove the tilde and get

$$\hat{f}\phi_i = \varepsilon_i\phi_i. \tag{14.24}$$

Thus, we have a new interpretation of the Hartree-Fock procedure: the process of minimizing the total energy w.r.t. a set of orthogonal orbital is completely equivalent to the process of finding the lowest $\frac{N}{2}$ eigenstates of the Fock operator.

Hartree-Fock energy

Given the Fock orbitals, the Hartree-Fock energy is then (and all the bra-kets henceforth except the LHS on the present equation refer to a single-electron expectation values), as we wrote:

$$\langle H\rangle = E_{1B} + E_H + K, \tag{14.25}$$

where all summations in this subsection are for $i = 1, ..., \frac{N}{2}$. On the other hand

$$\varepsilon_i = \langle\phi_i|\hat{f}|\phi_i\rangle = \langle\phi_i|\hat{h}|\phi_i\rangle + \langle\phi_i|v_H(\boldsymbol{r})|\phi_i\rangle + \langle\phi_i|\hat{k}|\phi_i\rangle, \tag{14.26}$$

and it is not too hard to prove (**do it on your own!)** that

$$2\sum_i\langle\phi_i|\hat{h}|\phi_i\rangle = E_{1B} \equiv E_{Kin} + E_{eN} \tag{14.27}$$

where we define the one-body energy; and further

$$2\sum_i\langle\phi_i|v_H(\boldsymbol{r})|\phi_i\rangle = 2E_H, \tag{14.28}$$

$$2\sum_i\langle\phi_i|\hat{k}|\phi\rangle = 2K, \tag{14.29}$$

and therefore

$$2\sum_i\varepsilon_i = E_{1B}+2(E_H+K) = E_{1B}+2E_{ee} = \langle H\rangle+E_{ee} = \langle H\rangle+E_H+K, \tag{14.30}$$

i.e.,

$$\langle H\rangle \quad = \quad 2\sum_i\varepsilon_i \quad - \quad E_{ee} \quad = \quad 2\sum_{i=1}^{\frac{N}{2}}\varepsilon_i \quad - \quad E_H \quad - \quad K. \tag{14.31}$$

The interpretation of this eq. is very simple. Each orbital energy ε_i double-counts the Coulomb interaction (in the direct and in the exchange terms).

For example, if i=2 and j=5, then ε_2 will contain the interaction of the 2nd-electron with the 5th. On the other hand, ε_5 counts this interaction again, this time as the interaction of the 5th-electron with the 2nd one. So in short, the sum over all orbital energies (times 2, for spin) double-counts the Coulomb interaction, so we need to subtract half of it, i.e., to subtract E_{ee}.

SCF procedure

Finally, note that the Hartree Fock equations need to be solved self-consistently, since the Fock operator depends on the eigenstates, i.e., $\hat{f} = \hat{f}[\phi]$. Thus we need to solve a procedure schematically noted as:

1. Choose a random set of $\frac{N}{2}$ functions $\phi_i(\boldsymbol{r})$;
2. Then, repetitively:

2.i Calculate the density and Fock operator from these $\phi_i(\boldsymbol{r})$

2.i Find the lowest $\frac{N}{2}$ eigenstates of $\hat{f}$, $\phi_i(\boldsymbol{r})$,$i = 1, .., \frac{N}{2}$.

Basis-sets solution of the Hartree-Fock equation

Basis-sets

While it is possible to do the SCF procedure above with actual dense orbitals defined over space, in chemistry it is quite common to use basis set methods (and these are important also for MP2, and higher-level correlated approaches).

Specfcially, we write the unknown eigenstates in terms of an orthogonal set of fixed functions

$$\phi_i(\boldsymbol{r}) = \sum_{l=1}^{M} C_{li}\chi_l(\boldsymbol{r}). \tag{14.32}$$

where $M > \frac{N}{2}$ is the overall basis-set size, and the $\chi_j(\boldsymbol{r})$ are in turn LCAO, i.e., linear-combination of atomic orbitals

$$\chi_l(\boldsymbol{r}) = \sum_{k=1}^{M} d_{kl} g_l(\boldsymbol{r}), \tag{14.33}$$

Here, d_{kl} are linear coefficients constructed to make the χ_l orbitals orthogonal.

Each $g_l(\boldsymbol{r})$ is associated with a specific atom, and a specific nlm, and is obtained by placing one or several Gaussians. This is because some of the integrals we'll show shortly turn out to be analytically solvable for Gaussian basis functions.

For example, a fairly simple basis for LiH would be made from 7 functions, two on the hydrogen (representing 1s and 2s) and five on the lithium (representing 1s, 2s, and 2px, 2py, 2pz)

$$g_1(\boldsymbol{r}) = c_{1a}e^{-\zeta_{1a}(\boldsymbol{r}-\boldsymbol{R_H})^2} + c_{1b}e^{-\zeta_{1b}(\boldsymbol{r}-\boldsymbol{R_H})^2}$$

$$g_2(\boldsymbol{r}) = c_2e^{-\zeta_2(\boldsymbol{r}-\boldsymbol{R_H})^2}$$

$$g_3(\boldsymbol{r}) = c_{3a}e^{-\zeta_{3a}(\boldsymbol{r}-\boldsymbol{R_{Li}})^2} + c_{3b}e^{-\zeta_{3b}(\boldsymbol{r}-\boldsymbol{R_{Li}})^2}$$

$$g_4(\boldsymbol{r}) = c_4e^{-\zeta_4(\boldsymbol{r}-\boldsymbol{R_{Li}})^2}$$

$$g_{5x}(\boldsymbol{r}) = c_5\left(x - X_{Li}\right)e^{-\zeta_5(\boldsymbol{r}-\boldsymbol{R_{Li}})^2}$$

$$g_{5y}(\boldsymbol{r}) = c_5\left(y - Y_{Li}\right)e^{-\zeta_5(\boldsymbol{r}-\boldsymbol{R_{Li}})^2}$$

$$g_{5z}(\boldsymbol{r}) = c_5\left(z - Z_{Li}\right)e^{-\zeta_5(\boldsymbol{r}-\boldsymbol{R_{Li}})^2}$$

In practice, the values of the coefficients (the c's and ζ's) and the number of Gaussians in each basis functions is optimized based on studies of small molecules. Better basis sets have more functions and more associated work.

The basis sets have names, e.g., STO-3G, 6-31G, and 6-31G**, we won't go over these.

Hartree-Fock eqs. with basis-sets

Given $\hat{f}\phi_i = \varepsilon_i\phi_i$, we turn it to a basis-set eq. involving the orthogonal basis χ_l (the LCAO). Note that even though in the derivations we treat the χ_l as a complete set, they are NOT A COMPLETE basis, so the results would need to be converged with basis (i.e., increase M till convergence).

Dot-product with χ_l to get

$$\langle\chi_l|\hat{f}|\phi_i\rangle = \langle\chi_l|\phi_i\rangle\varepsilon_i. \tag{14.34}$$

Now insert a "resolution of the identity", i.e., $\sum_{m=1}^{M}|\chi_m\rangle\langle\chi_m| \simeq I$ (not exactly 1 since M is not infinitely large) to get

$$\sum_{m=1}^{M}\langle\chi_l|\hat{f}|\chi_m\rangle\langle\chi_m|\phi_i\rangle = \langle\chi_l|\phi_i\rangle\varepsilon_i \tag{14.35}$$

i.e., a matrix relation

$$\sum_m f_{lm}C_{mi} = C_{li}\varepsilon_i \tag{14.36}$$

or in short

$$fC = C\varepsilon \tag{14.37}$$

where $f_{lm} \equiv \langle\chi_l|\hat{f}|\chi_m\rangle$, and $C_{li} = \langle\chi_l|\phi_i\rangle$. In short, we need to find the lowest $\frac{N}{2}$ eigenvalues of the matrix f_{lm}

Further, to make the matrix f_{lm} we note, from the eq. we derived, $\hat{f} = \hat{h} + v_H(\boldsymbol{r}) + \hat{k}$, that

$$f_{lm} = h_{lm} + (v_H)_{lm} + \left(\hat{k}\right)_{lm} \tag{14.38}$$

where

$$h_{lm} = \langle\chi_l| - \frac{\nabla^2}{2m_e} + v(\boldsymbol{r})|\chi_m\rangle = \langle\chi_l| - \frac{1}{2m_e}\nabla^2 - \sum_{q=1}^{N_{NUC}}\frac{Z_q}{|\,\boldsymbol{r} - \boldsymbol{R_q}|}|\chi_m\rangle \tag{14.39}$$

is evaluated analytically, since the matrix element of $1/r$ between two Gaussians is analytic, i.e.,

$$\langle g_m|\hat{h}|g_l\rangle = \left(t_{ml} - \sum_{q=1}^{N_{NUC}} Z_q G_{ml}^q\right) \tag{14.40}$$

and therefore:

$$t_{ml} \equiv -\frac{1}{2m_e}\langle g_m|\nabla^2|g_l\rangle \tag{14.41}$$

$$G^q_{ml} \equiv \langle g_m|\frac{1}{|\,\boldsymbol{r}-\boldsymbol{R_q}|}|g_l\rangle \tag{14.42}$$

are both evaluated analytically, so

$$h_{ml} = \sum_{m',l'=1}^{M} d_{mm'}d_{ll'}\langle g_{m'}|\hat{h}|g_{l'}\rangle \tag{14.43}$$

is calculated readily.

In addition, $(v_H)_{lm}$ is also calculated readily (but expensively) from

$$(v_H)_{lm} = \langle\chi_l|v_H|\chi_m\rangle = 2\sum_{i=1}^{\frac{N}{2}}\int \chi_l^*(\boldsymbol{r})\frac{|\phi_i(\boldsymbol{r'})|^2}{|\boldsymbol{r}-\boldsymbol{r}'|}\chi_m(\boldsymbol{r})d\boldsymbol{r'}d\boldsymbol{r'}, \tag{14.44}$$

i.e., plugging $\phi_i = \sum_n C_{ni}\chi_n$, we get

$$(v_H)_{lm} = 2\sum_{n,p=1}^{M} C^*_{ni}C_{pi}\int \chi_l^*(\boldsymbol{r})\frac{\chi^*_n(\boldsymbol{r'})\chi_p(\boldsymbol{r'})}{|\boldsymbol{r}-\boldsymbol{r}'|}\chi_m(\boldsymbol{r})d\boldsymbol{r'}d\boldsymbol{r'}, \tag{14.45}$$

i.e., (in a non-universal notation!!)

$$(v_H)_{lm} = 2\sum_{n,p=1}^{M}(\chi_l\chi_n|\chi_m\chi_p)P_{np}, \tag{14.46}$$

where the **density-matrix** is

$$P_{np} = \sum_{i=1}^{\frac{N}{2}} C^*_{ni}C_{pi}, \qquad n,p=1,..,N \tag{14.47}$$

while (note the order of indices)

$$(\chi_l\chi_n|\chi_m\chi_p) \equiv \int \frac{\chi_l^*(\boldsymbol{r})\chi_n^*(\boldsymbol{r'})\chi_m(\boldsymbol{r})\chi_p(\boldsymbol{r'})}{|\boldsymbol{r}-\boldsymbol{r}'|}d\boldsymbol{r'}d\boldsymbol{r'}, \tag{14.48}$$

is calculable since the corresponding $(g_lg_m|g_mg_p)$ matrix element is calculable analytically, and we can "transform" basis using the "d" matrix to go from the set of $(g_lg_m|g_ng_p)$ to the set of $(\chi_l\chi_m|\chi_n\chi_p)$, as we did earlier

$$(\chi_l\chi_n|\chi_m\chi_p) = \sum_{l'm'n'p'=1}^{M}(g_{l'}g_{m'}|g_{n'}g_{p'})d_{ll'}d_{mm'}d_{nn'}d_{pp'}. \tag{14.49}$$

Finally, we similarly get for the exchange operator (I bold the exchanged indices in the matrix element, "m" and "p")

$$(\hat{k})_{lm} = -\sum_{n,p=1}^{M} (\chi_l \chi_n | \chi_{\boldsymbol{p}} \chi_{\boldsymbol{m}}) P_{np} \tag{14.50}$$

We thus have a well-defined prescription for solving HF in a basis.

Step I: Calculate analytically the 4-index and 2-index matrices involving the Gaussians (g).

Step II: Transform to make 4-index and 2-index involving χ's. Calculate the matrix elements of h.

Step III: Start with a simple guess to the $M \times M$ density-matrix , e.g., with all elements vanishing except

$$P_{ii} = 1 \text{ if } i \leq \frac{N}{2} \tag{14.51}$$

Step IV: Do the SCF loop until convergence:

- From the density-matrix P calculate the matrices of the operators v_H and $\hat{k}$, from which get the Fock matrix.
- Diagonalize the Fock matrix f and from its lowest $\frac{N}{2}$ eigenvectors C calculate the new P.

The density matrix in Hartree-Fock

The density matrix in a basis-set P is fundamental in Hartree-Fock calculations (and in modern DFT), so let's talk about it for a second. It is clear that P is the matrix representation of the density-matrix in coordinate space (**prove this relation!**):

$$P_{kl} = \langle \chi_l | \frac{\rho}{2} | \chi_l \rangle = \int \chi_l^*(\boldsymbol{r'}) \frac{\rho(\boldsymbol{r}, \boldsymbol{r'})}{2} \chi_l(\boldsymbol{r}) d\boldsymbol{r} d\boldsymbol{r'} \tag{14.52}$$

In fact, because of this proximity we often use the same label "density-matrix" to both, even though one is basis-dependent and the other refers to coordinate

The $M \times M$ P matrix has $\frac{N}{2}$ eigenvalues of "1", and the rest $(M - \frac{N}{2})$ will be 0. It is thus a **projection operator,** i.e., $P^2 = P$, i.e., we can rewrite the defining equation $P_{kl} = \sum_{i=1}^{\frac{N}{2}} C_{ki}^* C_{li}$ as

$$P = C \theta C^+ \tag{14.53}$$

where the diagonal occupation matrix fulfills:

$$\theta_{ij} = \begin{cases} \delta_{ij} & i = j \leq \frac{N}{2} \\ 0 & \text{otherwise.} \end{cases} \tag{14.54}$$

Modern QM techniques often aim at getting P directly, without finding all the eigenstates, so that the Hartree-Fock exchange operator can be calculated efficiently, although this is easier said than done (except for specialized cases where P is very much "near-diagonal" in the proper localized basis set, like in simulations of water).

In DFT the density matrix has similar role and properties.

Hartree-Fock efficiency (improvable) and accuracy (not very good)

The Hartree-Fock equations for a basis, as presented above, are very expensive as they scale as $O(N^4)$ with the number of electrons. Solving them on a grid, i.e., getting directly $\phi_i(\boldsymbol{r})$ without using a Gaussian basis, is more efficient for large molecules due to a gentler scaling, $O(N^3)$ or even better. There are also techniques that mix basis sets and grids.

More crucially, however, Hartree-Fock does not get good chemical results. For example, the bond-breaking energy of H_2 is predicted to be $\approx$3.5eV rather than the experimental 4.5eV.

Improvement is possible in several directions:

- Modifying the Fock matrix by "magic" (justifiable) terms – this is the essence of what DFT does, we'll see shortly.
- Applying perturbation theory, where $H_1 \equiv H - H_0$, and H_0 is a solvable Hamiltonian that we get from Hartree-Fock. Later we'll calculate the 2nd order perturbation theory expression (MP2)
- Using more numerically expensive approaches, such as Configuration Interaction, or Coupled Cluster – we'll glance over both.

15 Pre-DFT: Homogeneous Electron Gas (HEG)

Notation:

- Starting this chapter "n" is the electron density $n(\boldsymbol{r})$ so it is a continuous variable; the only integers would be labeled by "m" (and occasionally "i, j").
- Also, my notation to the energy-per-volume as ϵ (and not as $n\varepsilon$) differs from most books, so beware when using the formulae here.

HEG eigenstates

The simplest system in electronic structure is HEG. From HEG we'll derive a lot of the motivational and practical parts of DFT.

Imagine a uniform charge system – uniform positive and uniform negative; the positive and negative charge (using e=1) per volume will be constant as $n(\boldsymbol{r}) = n = fixed$. Assume the charges are in a box of length L, volume $V = L^3$. We'll use periodic boundary conditions (b.c.)

1D

If space was 1D, the 1-electron w.f.'s (see discussion soon) are periodic

$$\psi(x+L) = \psi(x)$$

and would be formally

$$\psi_{k_x}(x) = \frac{e^{ik_x x}}{\sqrt{L}} \tag{15.1}$$

with

$$k_x = \frac{2\pi}{L}m, \qquad m = -\infty, ..., -1, 0, 1, 2, ...\infty. \tag{15.2}$$

Such functions are of course orthogonal. Proof: If $k_x \neq k'_x$ then

$$\begin{aligned}\langle\psi_{k_x}|\psi_{k'_x}\rangle = \int_0^L \psi^*_{k_x}(x)\psi_{k'_x}(x)dx = \frac{1}{L}\int_0^L e^{i(k'_x-k_x)x}dx \\ = \frac{e^{i(k'_x-k_x)L}-1}{i(k'_x-k_x)} = \frac{e^{i\frac{2\pi(m'-m)L}{L}}-1}{i(k'_x-k_x)} = 0 \end{aligned} \tag{15.3}$$

while for $k_x = k'_x$, $\langle\psi_{k_x}|\psi_{k'_x}\rangle = \frac{1}{L}\int_0^L e^{i\cdot 0\cdot x}dx = 1$.

3D

In 3D, the extension is obvious:

$$\psi_{\boldsymbol{k}}(\boldsymbol{r}) = \frac{e^{i\boldsymbol{k}\cdot\boldsymbol{r}}}{\sqrt{V}} = \frac{e^{ik_x x + ik_y y + ik_z z}}{\sqrt{V}} \tag{15.4}$$

and

$$k_x = \frac{2\pi}{L} m_x,\ k_y = \frac{2\pi}{L} m_y,\ k_z = \frac{2\pi}{L} m_z, \tag{15.5}$$

with m_x, m_y, m_z integers. Note the spacing between the momenta: $dk_x = \frac{2\pi}{L}$, so

$$d^3k = dk_x dk_y dk_z = \left(\frac{2\pi}{L}\right)^3. \tag{15.6}$$

HEG eigenstates clarification

Where will these HEG eigenstates be eigenstates of something? In two cases:

- Non-interacting electrons: if we were to ignore the electron electron interaction, replacing it by a constant that negates the effect of the positive background, our Hamiltonian would be kinetic only (with the periodicity constraint)
- Another interesting case is Hartree-Fock for HEG. Here we don't know yet (we'll partially calculate) the Hartree-Fock energies, but we do know that due to the transnational symmetry the eigenstates have to be those $\psi_{\boldsymbol{k}}$.

In either case, we can imagine that the electrons are in a closed-shell Slater determinant, where all eigenstates $\psi_{\boldsymbol{k}}$ with $|\boldsymbol{k}|$ below a certain maximum value k_F (F for "Fermi") are "filled", i.e., are part of the Slater determinant. Let's find this:

Fermi Momentum k_F

Label the (constant) density n. In a big box of volume V, we have

$$N = nV = nL^3 \tag{15.7}$$

electrons. The electrons fill all states till k_F, so (the "2" below for spin)

$$N = 2 \sum_{m_x, m_y, m_z} \Theta\left(k_F^2 - k^2\right), \quad k^2 = k_z^2 + k_y^2 + k_z^2 \tag{15.8}$$

and the k's and the integer m's are related by Eq. (15.5).

Since the k's are closely spaced (they're proportional to 1/L, and L is big), we replace the sum by integral, i.e., multiply by $\frac{d^3k}{d^3k}$, where $d^3k = \left(\frac{2\pi}{L}\right)^3$.

$$N = \frac{2\int \Theta\left(k_F^2 - k^2\right) d^3k}{d^3k} = \frac{2\int_0^{k=k_F} d^3k}{\left(\frac{2\pi}{L}\right)^3} = \frac{L^3}{4\pi^3}\int_0^{k_F} 4\pi k^2 dk = \frac{V}{\pi^2}\frac{k_F^3}{3} \tag{15.9}$$

so

$$n = \frac{N}{V} = \frac{k_F^3}{3\pi^2} \tag{15.10}$$

i.e.,

$$k_F = \left(3\pi^2 n\right)^{\frac{1}{3}} \tag{15.11}$$

i.e., $k_F = \mathrm{O}\left(n^{\frac{1}{3}}\right)$.

Note that the associated **Fermi energy** is

$$E_F \equiv \frac{k_F^2}{2m_e} = \frac{\left(3\pi^2\right)^{\frac{2}{3}}}{2m_e} n^{\frac{2}{3}}. \tag{15.12}$$

Hartree-Fock calculation of an HEG

Kinetic energy per volume

The total kinetic energy is labeled $\langle T\rangle = \langle\Psi|T|\Psi\rangle$, and $|\Psi\rangle$ is the closed-shell determinant made from all occupied eigenstates (with spins up and down), i.e., those with $|\boldsymbol{k}| < k_F$.

We further label the kinetic energy per volume as $t = \frac{\langle T\rangle}{V}$, and calculate it trivially by noting that t/n is the average kinetic energy per electron (since "t" is the average kinetic energy per volume and "n" the average # of electrons per volume).

$$\frac{t}{n} = \text{Average}\left(\frac{k^2}{2m_e}\right) = \frac{\int_0^{k_F} \frac{k^2}{2m_e} d^3k}{\int_0^{k_F} d^3k} = \frac{\frac{1}{2m_e}\int_0^{k_F} k^2 \cdot k^2 dk}{\int_0^{k_F} k^2 dk} = \frac{\frac{1}{2m_e}\frac{k_F^5}{5}}{\frac{k_F^3}{3}} = \frac{3}{10m_e}k_F^2. \tag{15.13}$$

Note that

$$\frac{t}{n} = \frac{3}{5}E_F, \tag{15.14}$$

which makes sense since the average kinetic energy per occupied electron is smaller than the maximum kinetic energy.

Note that by its definition, $\frac{t}{n}$ depends only on the density. Also, for completeness note that

$$t = \frac{3}{10m_e} n k_F^2 = \frac{3}{10m_e} n (3\pi^2 n)^{\frac{2}{3}} = \frac{3}{5} C_T n^{\frac{5}{3}}, \tag{15.15}$$

where

$$C_T \equiv \frac{3(3\pi^2)^{\frac{2}{3}}}{10m_e}, \tag{15.16}$$

i.e., the kinetic energy density scales as the density to the 5/3 power.

Electron-nucleus and Hartree electron-repulsion energies cancel each other.

Each electron sees an overall zero background charge, so that the (classical) Hartree repulsion potential cancels the attraction from the positive background.

Density matrix in HEG (for the exchange energy)

But the exchange energy does not vanish for an HEG!

Calculating it will all constants is quite painstaking, but we can get its scaling with n. For this we first need the density matrix

$$\begin{aligned}\rho(\boldsymbol{r},\boldsymbol{r'}) = 2\sum_{\boldsymbol{k}} \theta(k_F^2 - k^2)\psi_{\boldsymbol{k}}(\boldsymbol{r})\psi_{\boldsymbol{k}}^*(\boldsymbol{r'}) = 2\sum_{\boldsymbol{k}} \theta(k_F^2 - k^2)\frac{e^{i\boldsymbol{k}\cdot\boldsymbol{r}}e^{-i\boldsymbol{k}\cdot\boldsymbol{r'}}}{V} \\ = \frac{2}{V}\sum_{\boldsymbol{k}} \theta(k_F^2 - k^2)e^{i\boldsymbol{k}\cdot\boldsymbol{s}}, \quad \boldsymbol{s} = \boldsymbol{r} - \boldsymbol{r'}\end{aligned} \tag{15.17}$$

(here **"s" is unrelated to spin**). Recalling that $\rho(\boldsymbol{r},\boldsymbol{r}' = \boldsymbol{r}) = n(\boldsymbol{r})$, we avoid recalculating the constants (pi's, etc.) by noting that

$$\frac{\rho(\boldsymbol{r},\boldsymbol{r'})}{n} = \frac{\int_0^{k_F} e^{i\cdot\boldsymbol{k}\cdot\boldsymbol{s}}k^2dk}{\int_0^{k_F} k^2dk} \equiv \frac{\int_0^1 e^{i\cdot\boldsymbol{u}\cdot\boldsymbol{q}}u^2du}{\int_0^1 u^2du} = g(q) = g(k_F s)$$

$$\text{where}: \ \boldsymbol{u} \equiv \frac{\boldsymbol{k}}{k_F}, \quad \boldsymbol{q} \equiv k_F\boldsymbol{s} \tag{15.18}$$

and $g(q)$ is defined by the equation above (from isotropy it clearly depends only on $q = k_F s$, the length of $\boldsymbol{q}$, not on its direction). It can be calculated explicitly (see Baer's 2017 book, Eq. 3.2.4) and is falling off with q (oscillating as it falls off), so it is tiny by the time that $k_F s > 3$.

> Side note : The density matrix exhibits interesting oscillations due to the abrupt separation of occupied from non-occupied states. The oscillations of the density matrix are formally (and distantly) related to **Friedel Oscillations** in the density of metals near the edge of the metal – i.e., the electron wavefunctions near surfaces oscillate as they oscillate in vacuum – see class.

Exchange energy for HEG

The total exchange energy is then

$$K = -\frac{1}{4}\int \frac{|\rho(\boldsymbol{r}.\boldsymbol{r'})|^2}{|\boldsymbol{r}-\boldsymbol{r}'|} d\boldsymbol{r}\, d\boldsymbol{r'} = -\frac{n^2}{4}\int \frac{|g(k_F s)|^2}{s} d\boldsymbol{s} d\boldsymbol{r}, \tag{15.19}$$

where we converted the integral over $\boldsymbol{r}$,$\boldsymbol{r'}$ to an integral over $\boldsymbol{s}, \boldsymbol{r}$. We get rid of the integral over $\boldsymbol{r}$ by defining the exchange energy per volume. (at times we write a 3-d integration volume as $d\boldsymbol{s}$, at times as $d^3\boldsymbol{s}$):

$$\epsilon_X \equiv \frac{K}{V} = -\frac{n^2}{4}\int \frac{|g(k_F s)|^2}{s} d^3\boldsymbol{s} = -\frac{n^2}{4}\int \frac{|g(q)|^2}{\frac{q}{k_F}} \frac{d^3\boldsymbol{q}}{k_F^3} = -\frac{n^2}{4k_F^2} B, \tag{15.20}$$

where $B \equiv \int \frac{|g(q)|^2}{q} d\boldsymbol{q}$ is some purely-numerical constant (independent of n). And of course "X" stands here for exchange.

Therefore, since $k_F = n^{\frac{1}{3}}$, the exchange energy scales as $\mathrm{O}(n^{2-\frac{2}{3}}) = \mathrm{O}(n^{\frac{4}{3}})$, i.e.,

$$\epsilon_X = -\frac{3}{4} C_X n^{\frac{4}{3}} \tag{15.21}$$

where a careful calculation, including that of B from above, shows that

$$C_X = \left(\frac{3}{\pi}\right)^{\frac{1}{3}}. \tag{15.22}$$

Total energy

We saw that the H.F. energy for an HEG is made from only two parts: kinetic and exchange. Let's write the total energy for our volume V as

$$E_{H.F.} = \int t[n] d\boldsymbol{r} + \int \epsilon_X[n] d\boldsymbol{r} \tag{15.23}$$

where we included an [] to denote the formal of the two terms on the density.

It turns out that the total energy for an HEG can be calculated exactly, in a beyond-HF method called Quantum Monte Carlo, and when we do it there is a correction to the energy, labeled "C" for correlation

$$E_C = E - E_{H.F.} = \int \epsilon_C[n] d\boldsymbol{r} \tag{15.24}$$

where $\epsilon_C[n]$ has been numerically tabulated (since the late 1970's).

The reason for this correlation energy is that the Hartree-Fock method, by its nature, ignores most correlation, while in the true many-electron wavefunction we expect that the w.f. will be very small if any two electrons get close to each other. Put differently, the true w.f. would have what we labeled "Coulomb Hole", not just the "Fermi Hole" (see in previous chapters).

Let's now include in the total energy the explicit form of the electron-(positive-background) and electron-electron interaction, even though they cancel for an HEG, and end up with the full expression for the total energy

$$E[n] = \int t(n(\boldsymbol{r}))d\boldsymbol{r} + \int n(\boldsymbol{r})v(\boldsymbol{r})d\boldsymbol{r} + \frac{1}{2}\int \frac{n(\boldsymbol{r})n(\boldsymbol{r'})}{|\boldsymbol{r}-\boldsymbol{r}'|}d\boldsymbol{r}'d\boldsymbol{r'} + E_{XC}[n] \quad (15.25)$$

where the attractive potential due to the positive-charge background is labeled here $v(\boldsymbol{r})$ (i.e., just like the potential for nuclei in Chemical systems – we assume that an HEG has "completely smeared nuclei") and we denoted the dependence on the local density even though at this stage it is uniform. Further, define the <u>exchange-correlation energy</u> as:

$$E_{XC}[n] = \int \epsilon_X(n(\boldsymbol{r}))d\boldsymbol{r} + \int \epsilon_C(n(\boldsymbol{r}))d\boldsymbol{r} \quad (15.26)$$

Note the symbol $E[n]$ – it implies that E is a functional of $n(\mathbf{r})$ i.e., given the density over all space, $\{n(\boldsymbol{r})$ at all $\boldsymbol{r}\}$, a final (single number) for the energy is outputted.

Generalizing to non-uniform backgrounds

The expression for the energy above is very enticing, **once we stipulate that we will apply it to ANY system, not just one with constant background.** This is formally justifiable ONLY in the limit that the electron density varies slowly in space, so around each region we can approximate that it is locally constant, but we can try to apply it for general systems, even heavy atoms where the density varies a lot from core to valence.

> If you're horrified with applying expressions derived from constant-n to a general $n(\mathbf{r})$, remember that in practical thermodynamics we always apply the concept of temperature even to regions where it varies from one place to another, even though temperature is formally defined only in equilibrium, i.e., where it is constant in space.

In fact, within a year or two after the advent of QM, Thomas and Fermi realized this, and tried to apply the energy formula (but without ϵ_X which came in the 1930's and ϵ_C which came in the 1960's-1980's) to atoms and molecules.

Specifically, they looked to extremize the energy subject to the constraint that N=fixed, i.e., define

$$L[n] = E[n] - \mu\left(\int n(\boldsymbol{r})d\boldsymbol{r} - N\right) \quad (15.27)$$

Then, functional differentiation w.r.t. n gives

$$0 = \frac{\delta L}{\delta n(\boldsymbol{r})} = \frac{\delta E}{\delta n(\boldsymbol{r})} - \mu \quad (15.28)$$

i.e.,

$$\frac{\delta E}{\delta n(\boldsymbol{r})} = \mu \tag{15.29}$$

This is the main equation also in DFT. Let's put in the terms in E a little more explicitly (and change at times the dummy $\boldsymbol{r}$ to $\boldsymbol{r''}$ so we don't confuse with $n(\boldsymbol{r})$).

$$\begin{aligned}\mu = \frac{\delta}{\delta n(\boldsymbol{r})}\left(\int t(n(\boldsymbol{r}))d\boldsymbol{r} + \int n(\boldsymbol{r})v(\boldsymbol{r})d\boldsymbol{r} + \frac{1}{2}\int \frac{n(\boldsymbol{r''})n(\boldsymbol{r'})}{|\boldsymbol{r''}-\boldsymbol{r}'|}d\boldsymbol{r}'d\boldsymbol{r} + E_{XC}[n]\right)\\ = \frac{d(t(n(\boldsymbol{r})))}{d(n(\boldsymbol{r}))} + v(\boldsymbol{r}) + v_H(\boldsymbol{r}) + v_{XC}(\boldsymbol{r}) \quad (15.30)\end{aligned}$$

where we defined

$$v_{XC}(\boldsymbol{r}) = \frac{\delta E_{XC}}{\delta n(\boldsymbol{r})}. \tag{15.31}$$

> Question for you: for the local-exchange expression we use, i.e., $\epsilon_X(\boldsymbol{r}) = -\frac{4}{3}C_X n(\boldsymbol{r})^{\frac{4}{3}}$, what will be the exchange contribution to $v_{XC}(\boldsymbol{r})$?

Plugging the previous result of $t = \frac{3}{5}C_T n^{\frac{5}{3}}$, and the definition of the Hartree potential (the classical repulsion by the other electrons) we get

$$\mu = C_T n^{\frac{2}{3}} + v_{tot}(\boldsymbol{r}), \tag{15.32}$$

where we defined the total potential:

$$v_{tot}(\boldsymbol{r}) = v(\boldsymbol{r}) + v_H(\boldsymbol{r}) + v_{XC}(\boldsymbol{r}). \tag{15.33}$$

(and I remind you that in the original Thomas-Fermi approach there was no v_{XC}.) Another way to put it is:

$$n(\boldsymbol{r}) = (\mu - v_{tot}(\boldsymbol{r}))^{\frac{3}{2}}\ . \tag{15.34}$$

> Note: in a way, this equation should remind you of the relation between the potential and kinetic energy for a regular classical particle, where a particle cannot be in a region where the potential is higher than the energy.
>
> And here, in the Thomas-Fermi model, the density is non-zero only in regions where the potential is lower than the chemical potential.
>
> End of Note.

Thomas and Fermi solved this model for an atom, where $v(\boldsymbol{r}) = -\frac{Z}{r}$. The solution is interesting; it shows that the total energy is proportional to $Z^{7/3}$ (see Baer's book, Eq. (2.4.10), and Figure 2.4.1, reproduced below) . Further, for positive ions the density terminates at a finite distance.

Indeed, the total atomic energy is quite close, percentage wise, to the Thomas-Fermi energy (see Baer's figure, below)

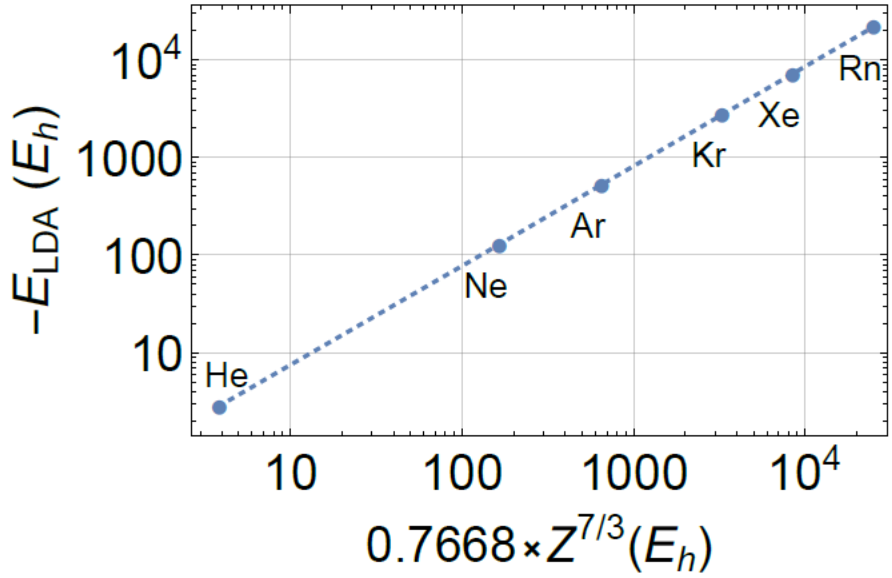


Figure 2.4.1. The density functional (within the local density approximation) energy of rare gas atoms vs the Thomas-Fermi value of Eq. 2.4.10. It should be noted that the HF energies are almost indistinguishable from the LDA values on the scale of this graph.

Unfortunately, the Thomas Fermi by itself predicts no chemical bonding. Adding the exchange part helps, but the results are not of chemical accuracy.

In the last 30 years Thomas-Fermi methods have enjoyed a renaissance, since they can handle giant molecular systems with millions of atoms; the key to make them more numerically useful is to add a new correlation term which, by construction, reproduces the response of a Homogeneous Electron Gas to external fields that oscillate in space. Further, one mixes in other kinetic energy terms that reproduce to a single-electron system (the von-Weizsacker term).

But by and large, DFT took a different route; the crucial point was 1965, when Kohn and Hohenberg first proved that such energy expressions based on n could be taken formally seriously; and then, Kohn and Sham realized that they can formally merge this density-dependent energy approach with Hartree-Fock techniques, taking the best of each. This is in the next chapter.

16 DFT

Note: Some of the discussion below is non-historical. Also, here F denotes a functional of the density.

The Kohn-Hohenberg Lemma

Let's recall he Hamiltonian we know, using "hats" for some operators and explicitly writing the form of the one-body potential (the electron-nucleus attraction) in terms of the density operator:

$$H = \hat{T} + \hat{V} + \hat{U} = \hat{T} + \int v(\boldsymbol{r})\hat{n}(\boldsymbol{r})d\boldsymbol{r} + \hat{U} \tag{16.1}$$

Before proceeding, we need to prove something fairly obvious:

Different many-body potentials lead to different ground-state wavefunctions.

Lemma: the ground-state w.f. is a unique function of the overall potential. I.e., for a given overall potential there's a Hamiltonian, which has a ground state Ψ. We claim that two different potentials, $V_{tot}(\boldsymbol{r}_1, \boldsymbol{r}_2, ..., \boldsymbol{r}_N)$ and $V'_{tot}(\boldsymbol{r}_1, \boldsymbol{r}_2, ..., \boldsymbol{r}_N)$ must produce a different ground-state potential. (Note that in our case $\hat{V}_{tot} = \hat{V} + \hat{U}$, i.e., the total potential is a sum of one-body and two-body terms, but this is true for general potentials).

Proof of Lemma:

The many-electron Schrödinger eq. reads

$$\left(\hat{T} + V_{tot}\right)\Psi = E\Psi \tag{16.2}$$

and therefore, except at points where Ψ vanishes, we get:

$$V_{tot} = -\frac{\hat{T}\Psi}{\Psi} + E \tag{16.3}$$

so the potential is given uniquely by the w.f., except for an overall constant.

Different potentials must lead to different ground-state densities

Hohenberg and Kohn then proved something ostensibly similar, but actually **highly unintuitive.**

> Assuming non-degenerate ground-states, there's a one-to-one match between the density and the one-body potential.

One "direction" is obvious – for a given one-body potential (practically, for a given set of nuclear charges and positions), there's a Hamiltonian, which has a ground state Ψ, from which the density is calculated.

But the other "direction" is very surprising:

> Given a density $n(\boldsymbol{r})$, there's only a single one-body potential $v(\boldsymbol{r})$ which could have produced it ! So in this sense, the potential is a function (or a functional, be as mathematical as you want to be) of the density! (Hohenberg-Kohn)

Note the superficial similarity to the lemma we just proved; however:

- There, different potentials must lead to different overall many-body w.f., which makes sense.
- But here, something much stronger – not only are the w.f. (which are a function of N 3-d position variables) different, the densities (which are a function of only a single 3-d variable) different!

Let's prove by contradiction:

First, say that there were two such different one-body potentials, $v(\boldsymbol{r})$ and $v(\boldsymbol{r'})$, that produced the SAME one-body density $n(\boldsymbol{r})$.

Now note that we proved that the overall potential is different if the one-body potential is different (we don't modify the two-body potential, the e-e repulsion).

So v and v' lead to different many-electron wavefunctions, Ψ and Ψ'; the first fulfills $H\Psi = E\Psi$, and the 2nd $H'\Psi' = E'\Psi'$ (where H' is different from H in the one-body potential).

Then, the variational theorem states that

$$E = \langle\Psi|H|\Psi\rangle, \tag{16.4}$$

and since we assume non-degenerate ground-state, then since $\Psi \neq \Psi'$:

$$E < \langle\Psi'|H|\Psi'\rangle. \tag{16.5}$$

But

$$H = H' + \hat{V} - \hat{V'} = H' + \int \hat{n}(\boldsymbol{r})\,(v(\boldsymbol{r}) - v'(\boldsymbol{r}))\,\hat{n}(\boldsymbol{r}) d\boldsymbol{r}, \tag{16.6}$$

so plugging together the last two equations gives:

$$E < \langle\Psi'|H'|\Psi'\rangle + \langle\Psi'|\hat{V} - \hat{V'}|\Psi'\rangle, \tag{16.7}$$

i.e.,

$$E < E' + \int n(\boldsymbol{r})\,(v(\boldsymbol{r}) - v'(\boldsymbol{r})\,d\boldsymbol{r}, \tag{16.8}$$

where we used $\langle\Psi'|\hat{n}(\boldsymbol{r})|\Psi'\rangle = n(\boldsymbol{r})$, and $\langle\Psi'|H'|\Psi'\rangle = E'$.

The equation above could have also been analogously derived in the opposite direction (starting from $E' < \langle\Psi|H'|\Psi\rangle$), which would give

$$E' < E + \int n(\boldsymbol{r})\,(v'(\boldsymbol{r}) - v(\boldsymbol{r}))\,d\boldsymbol{r} \tag{16.9}$$

But now add the last two eqs., to obtain a contradiction:

$$E + E' < E + E' \tag{16.10}$$

This proves that indeed, only one potential could give rise to the density. So: **if we know the ground state density, then the one–body potential that gives rise to it** (i.e., gives rise to a ground-state wavefunction which has this density) **is set.**

Thus, we can say that $v = v[n]$. (v is a functional of n).

Note however that we don't know what this functional is; further, not all densities $n(\boldsymbol{r})$ are actually ground-state densities associated with a one-body potential! Formally, not all densities are what is called "$v -$ *representable*" (associated with a specific $v(\boldsymbol{r})$.) I.e., all we know is that if they are associated with one $v(\boldsymbol{r})$, they can't be associated with another $v'(\boldsymbol{r})$.

The Levy functional

We circumvent the problem of v-representability by defining a universal functional. Given a density, we search over all wavefunctions that have an associated density $n(\boldsymbol{r})$ (if there are any!) and calculate:

$$F[n] = \min_{\Psi \to n} \langle \Psi | \hat{T} + \hat{U} | \Psi \rangle, \tag{16.11}$$

where the meaning of the subscript "$\Psi \to n$" is that we sample all N-body wavefunctions which have a density $n(\boldsymbol{r})$. Further, we only sample from normalized "Ψ", so therefore the density integrates to give N electrons.

This universal functional is independent of the one-body potential. Of course, we don't know yet what $F[n]$ would be, but we'll soon make approximations to it.

It is easy to see then that, for a given $v(\boldsymbol{r})$, if we know $F[n]$ then our work is over; since the ground-state energy must minimize:

$$E[n] = \min_{n} \left(F[n] + \int v(\boldsymbol{r}) n(\boldsymbol{r}) d\boldsymbol{r} \right) \tag{16.12}$$

where the minimum is taken only over densities that are associated with some normalized Ψ, so $\Psi \to n$.

The proof is simple. The variational principle states

$$E = \min_{\Psi} \langle \Psi | \hat{T} + \hat{U} + \hat{V} | \Psi \rangle. \tag{16.13}$$

(where the minimum is done over normalized many-body w.f.'s) Let's then do the minimization in two stages; for each n search which (if any) many-body wavefunction that has the (normalized) density $n(\boldsymbol{r})$ gives the lowest energy, and then minimize among all these densities, i.e.,

$$E = \min_n \left[\min_{\Psi\to n} \langle \Psi | \hat{T} + \hat{U} + \hat{V} | \Psi \rangle \right] = \min_n \left[\left(\min_{\Psi\to n} \langle \Psi | \hat{T} + \hat{U} | \Psi \rangle + \int v(\boldsymbol{r}) n(\boldsymbol{r}) d\boldsymbol{r} \right) \right]. \tag{16.14}$$

The last term depends only on "n" so we can take it out of the inner-parentheses, i.e.,

$$E = \min_n \left[\left(\min_{\Psi\to n} \langle \Psi | \hat{T} + \hat{U} | \Psi \rangle \right) + \int v(\boldsymbol{r}) n(\boldsymbol{r}) d\boldsymbol{r} \right] = \min_n \left[F[n] + \int v(\boldsymbol{r}) n(\boldsymbol{r}) d\boldsymbol{r} \right]. \quad Q.E.D. \tag{16.15}$$

The Kohn-Sham Approach

Independent-electron Kinetic Energy functional

So far, we advanced theoretically but not practically, since we don't know yet what $F[n]$ is. We could try of course to plug the Thomas-Fermi approach (e.g., use Eq. (15.25)), but as we know that would be very approximate.

The key is then to replace the "big" quantity we don't know $F[n]$ by a known "big" quantity and a "small" unknown part. That was Kohn and Sham's idea.

The original Kohn-Sham approach was: **Let's replace $F[n]$ by a similar term for non-interacting electrons, plus the Hartree repulsion, plus a correction** – and the correction should (hopefully!) be small.

Specifcially, define

$$T_s[n] = \min_{\Psi\to n} \langle \Psi | \hat{T} | \Psi \rangle \tag{16.16}$$

i.e., the minimum-kinetic energy among all w.f. that are associated with the density $n(\boldsymbol{r})$. We'll see right below that it is possible to get this $T_s[n]$ from a Slater determinant (i.e., from non-interacting electrons, i.e., electrons which only interact with a one-body potential) when solving for the minimization of the energy.

Then, we'll write the unknown $F[n]$ as

$$F[n] = T_s[n] + E_H[n] + E_{XC}[n] \tag{16.17}$$

where as we defined earlier $E_H[n] = \frac{1}{2} \int \frac{n(\boldsymbol{r}) n(\boldsymbol{r'})}{|\boldsymbol{r} - \boldsymbol{r'}|} d\boldsymbol{r}\, d\boldsymbol{r'}$, and "XC" stand for exchange-correlation.

In the section after the next, we'll talk about $E_{XC}[n]$; for now just accept that it is a numerical functional of n that we can approximate reasonably.

Minimizing the overall functional

We're looking to minimize the total energy, $F[n] + \int n(\boldsymbol{r}) v(\boldsymbol{r}) d\boldsymbol{r}$, subject to the integrated-density constraint $\int n(\boldsymbol{r}) d\boldsymbol{r} = N$. Based on the eq. above, we would like to extremize therefore the Lagrangian (containing a Lagrange multiplier μ associated with the integrated density constraint)

$$L[n] = E[n] - \mu \left(n(\boldsymbol{r}) d\boldsymbol{r} - N \right), \tag{16.18}$$

i.e., (recall that $v(\boldsymbol{r})$ is the potential due to the nuclei):

$$L[n] = T_s[n] + E_H[n] + E_{XC}[n] + \int n(\boldsymbol{r}) v(\boldsymbol{r}) d\boldsymbol{r} - \mu \left(n(\boldsymbol{r}) d\boldsymbol{r} - N \right) \tag{16.19}$$

Functional differentiation yields

$$0 = \frac{\delta L}{\delta n(\boldsymbol{r})} = \frac{\delta E}{\delta n(\boldsymbol{r})} - \mu = \frac{\delta T_s}{\delta n(\boldsymbol{r})} + v_H(\boldsymbol{r}) + v_{XC}(\boldsymbol{r}) + v(\boldsymbol{r}) - \mu \tag{16.20}$$

where we derived earlier

$$v_H(\boldsymbol{r}) \equiv \frac{\delta E_H}{\delta n(\boldsymbol{r})} = \int \frac{n(\boldsymbol{r'})}{|\boldsymbol{r} - \boldsymbol{r'}|} d\boldsymbol{r'}, \tag{16.21}$$

and we also defined earlier

$$v_{XC}(\boldsymbol{r}) = \frac{\delta E_{XC}}{\delta n(\boldsymbol{r})}. \tag{16.22}$$

The Kohn-Sham equation for orbitals

It seems we have not done much, since we still have $\frac{\delta T_s}{\delta n(\boldsymbol{r})}$ to worry about. But here comes the key: **it could be proved that the lowest kinetic energy is obtained, for a given $n(\boldsymbol{r})$, from a closed-shell Slater-determinant with $\frac{N}{2}$ occupied orbitals** $\phi_i(\boldsymbol{r})$, $i = 1, ..., \frac{N}{2}$, such that

$$n(\boldsymbol{r}) = 2 \sum_{i=1}^{\frac{N}{2}} n_i |\phi_i(\boldsymbol{r})|^2. \tag{16.23}$$

where we introduced the occupation of the states (which for the present think of it as 1 for occupied states and 0 for unoccupied states).

> Note: the reason we introduce the occupation numbers at all is so we can deal later with partially charged system, where the occupation of the HOMO will be less than 1 – or that of the LUMO will be more than 0).

Then, the kinetic energy of this set will be denoted as $T[f, \phi]$ (we should have really wrote it as $T[f, \phi, \{n_i\}]$, but that's too long) and calculated as

$$T = -\frac{1}{2m_e} \cdot 2 \sum_{i=1}^{\frac{N}{2}} n_i \int \phi_i^*(\boldsymbol{r}) \nabla^2 \phi_i(\boldsymbol{r}) \, d\boldsymbol{r}. \tag{16.24}$$

So instead of taking a functional derivative w.r.t $n(\boldsymbol{r})$, take a derivative w.r.t.$\phi_i(\boldsymbol{r})$! When we do it it is useful to remember that a derivative w.r.t.$\phi_i(\boldsymbol{r})$ of a direct functional of $n(\boldsymbol{r})$ is obtained by the chain rule

$$\frac{\delta}{\delta\phi_i^*(\boldsymbol{r})} = \frac{\partial n(\boldsymbol{r})}{\partial\phi_i^*(\boldsymbol{r})}\frac{\delta}{\delta n(\boldsymbol{r})} = 2n_i\phi_i(\boldsymbol{r})\frac{\delta}{\delta n(\boldsymbol{r})} \tag{16.25}$$

where the "2" is due to the spin factor in the closed-shell density.

Further, we need to augment now the Lagrangian; instead of the constraint being now only the total charge=N, it is now that the the eigenstates are orthonormal, or more precisely, that norm of each eigenstate times the occupation of the state is 1 (for occupied states). We multiply by 2 the Lagrange multipliers per convenience (the 2 is due to spin, since we consider closed-shell systems):

$$L[\phi, f] = E(\phi, f) - 2\sum_{ij}\lambda_{ij}\left(n_i\langle\phi_i|\phi_j\rangle - \delta_{ij}\right) \tag{16.26}$$

where differentiating first w.r.t. the wavefunctions gives

$$\begin{gathered}0 = \frac{\delta L}{\delta\phi_i^*(\boldsymbol{r})} = \frac{\delta E}{\delta\phi_i^*(\boldsymbol{r})} - 2\sum_{ij}\lambda_{ij}n_i\phi_j(\boldsymbol{r}) = \\ \frac{\delta T_s}{\delta\phi_i^*(\boldsymbol{r})} + \frac{\delta}{\delta\phi_i^*(\boldsymbol{r})}\left(E_H + E_{XC} + \int n(\boldsymbol{r})v(\boldsymbol{r})d\boldsymbol{r}\right) - 2\sum_{ij}\lambda_{ij}n_i\phi_j(\boldsymbol{r}),\end{gathered} \tag{16.27}$$

i.e.,

$$0 = -2n_i\frac{\nabla}{2m_e}\phi_i(\boldsymbol{r}) + 2n_i\frac{\delta}{\delta n(\boldsymbol{r})}\left(E_H + E_{XC} + \int n(\boldsymbol{r})v(\boldsymbol{r})d\boldsymbol{r}\right)\phi_i(\boldsymbol{r}) - 2n_i\sum_{ij}\lambda_{ij}\phi_j(\boldsymbol{r}), \tag{16.28}$$

So divide by $2n_i$ and use the expressions for the potentials as functional derivatives to get:

$$0 = -\frac{\nabla}{2m_e}\phi_i(\boldsymbol{r}) + \left(v_H(\boldsymbol{r}) + v_{XC}(\boldsymbol{r}) + v(\boldsymbol{r})\right)\phi_i(\boldsymbol{r}) - \sum_{ij}\lambda_{ij}\phi_j(\boldsymbol{r}). \tag{16.29}$$

We now do the same trick as we did in Hartree Fock, rotate the orbitals till the Lagrange multipliers matrix is diagonal (with diagonal value ε_i) so we end up with the so-called **Kohn-Sham eqs.**

$$-\frac{\nabla^2}{2m_e}\phi_i(\boldsymbol{r}) + v_{KS}(\boldsymbol{r})\phi_i(\boldsymbol{r}) = \varepsilon_i\phi_i(\boldsymbol{r}), \tag{16.30}$$

i.e., the phi's are eigenstates of a Kohn-Sham Hamiltonian

$$\hat{f}\phi(\boldsymbol{r}) = \varepsilon_i\phi_i(\boldsymbol{r}) \tag{16.31}$$

which involves a kinetic energy and a local potential

$$\hat{f} = -\frac{\nabla^2}{2m_e} + v_{KS}(\boldsymbol{r}), \tag{16.32}$$

and the Kohn-Sham potential is:

$$v_{KS}(\boldsymbol{r}) = v_H(\boldsymbol{r}) + v_{XC}(\boldsymbol{r}) + v(\boldsymbol{r}). \tag{16.33}$$

Note how we use **the same symbol for the Kohn-Sham 1-electron Hamiltonian here and the Fock-matrix in the Hartree-Fock approach!** This is because we ended up with what looks like Hartree-Fock equations, but with a simple one-body Hamiltonian replacing the Fock-Hamiltonian; basically, the exchange-operator was replaced by a local potential, $v_{XC}(\boldsymbol{r})$!

The quality of the method depends of course then on the quality of the exchange-correlation potential, which is what we'll discuss (with extensions) in the next few sections.

But we need one more step: differentiate w.r.t. the occupation.

The orbital energies as derivatives w.r.t. the occupations

Take the objective we had an differentiate it w.r.t. the occupations (and replace the sum over λ_{ij} by the sum over ε_i)

$$\frac{\partial L}{\partial n_i} = \frac{\partial E}{\partial n_i} - \frac{\partial}{\partial n_i} \sum_i \varepsilon_i \left(n_i \langle \phi_i | \phi_i \rangle - \delta_{ij} \right) = \frac{\partial E}{\partial n_i} - \varepsilon_i \tag{16.34}$$

and if we set the orbitals to be normalized we get:

$$\frac{\partial E}{\partial n_i} = \varepsilon_i \tag{16.35}$$

i.e., the derivative of the energy w.r.t the occupation is the orbital energy!

We can now understand, by analogy with thermodynamics, why the occupations are 0 and 1 (or, if we deal with partially charged system, the HOMO would be partially occupied and all states below would have occupations of 1). Specfcially, if they were not 0 to 1, say one orbital was occupied at 0.7 and another at 0.3, and the second one had a higher ε, then we could lower the energy by moving a little bit of charge from the second orbital to the first (so the occupations would be 0.71 and 0.29, for example). We could continue the process till the occupation of one of the orbitals is 1 (the maximum possible by the Pauli principle) or the other would vanish.

This is the same as classical thermodynamics, where a material would be moving from one phase with higher chemical potential to another, The role of the chemical potentials is played by the orbital energies.

Of course, once a "phase" – in our case a specific level – is fully occupied, $n_i = 1$, nothing can move into it.

> The proof that $n_i \leq 1$ is a little more difficult in practice, since we have to consider the behavior of the one-body density matrix of the full interacting wavefunction, but it is still valid.

The XC potential (and operator)

The simplest exchange-correlation functional: LDA

What is $E_{XC}[n]$ (and from that v_{XC}) then? The simplest functional, labeled LDA (Local Density Approximation) would be based directly on an HEG. Thus, we separate

$$E_{XC}^{LDA}[n] = E_X^{LDA}[n] + E_C^{LDA}[n]. \tag{16.36}$$

We encountered both terms (without putting the LDA superscript) when we discussed the HEG. As a reminder, $E_X^{LDA}[n]$ is simply the local HEG exchange energy,

$$E_X = -\frac{3}{4} C_X \int n^{\frac{4}{3}}(\boldsymbol{r}) d\boldsymbol{r}$$

and the remaining $E_C^{LDA}[n]$ is then

$$E_C = \int \epsilon_C \left(n(\boldsymbol{r})\right) d\boldsymbol{r},$$

where $\epsilon_C(n)$ is obtained by doing a high level ("Monte-Carlo") calculation on a HEG (as HEG is one of the few cases where we can obtain the energy numerically exactly), i.e., for a spatially constant n, and fitting E vs. n to the exact result, as described in the HEG chapter.

LDA gives reasonable but not outstanding results for crystals and fairly weak results for molecules. Thus until the late 1980's LDA was viewed in chemistry circles as just a cheaper theory than HF (since the exact exchange was not calculated), but not as an accurate (or even a more-accurate) alternative.

Beyond LDA: GGA

The first hopes to improve LDA came in the 1970's-1980's with the development of GGA ("Generalized Gradient Approximation") functionals. In GGA, the local exchange-correlation energy per volume depends on the density AND also on on the density-gradient, i.e., on the density shape:

$$E_{XC}^{GGA}[n] = \int \epsilon_{XC}^{GGA} \left(n(\boldsymbol{r}), \nabla n(\boldsymbol{r})\right) d\boldsymbol{r}. \tag{16.37}$$

GGA functionals improve somewhat the description of the LUMO and gaps, and are heavily used in solid-state, but are not good enough for chemistry either.

Hybrid functionals

The revolution that made DFT suitable for chemistry came in the late 1980's-early 1990's, when Becke (based also on some earlier work) took these two so-similar theories, Hartree-Fock and DFT, and mixed them.

Specifically, the optimal mixture was numerically found to be 25% H.F., and 75% DFT; these proportions are purely empirical and give the best fit to a high-level calculation and measurements for a series of tens of molecules (and can be somewhat justified a-posteriori).

Specifically, Becke replaced the Kohn-Sham exchange-correlation functional E_{XC} by:

$$E_{XC} \to 0.75 E_{XC} + 0.25 K \tag{16.38}$$

where K is the Hartree-Fock exchange. Therefore, the only change in the Hamiltonian is the incorporation of Hartree-Fock exchange:

$$\hat{h}^{hybrid} = -\frac{\nabla^2}{2m_e} + v(\boldsymbol{r}) + v_H(\boldsymbol{r}) + 0.75 v_{XC} + 0.25\hat{k} \tag{16.39}$$

where v_{XC} can be an LDA or a GGA potential (i.e., it depends on the local density or also on its derivative).

The most famous such hybrid functional is B3LYP; it is very accurate for reaction energies, especially for organic systems, and most organic chemists use it routinely.

But B3LYP often does very poorly on charging energies (ionization energy and electron affinities) with errors of a few eV, almost as bad as LDA. There is a solution, long-range hybrids, that we'll discuss later (once we discuss some preliminaries). Before that we briefly review for completeness other functionals.

Some other functionals: Meta-GGA, Hybrids with MP2

You should be aware of some new developments in DFT. One class is inspired by hybrid functionals but does not have those expensive summations that the exchange-operator $\hat{k}$ has. One writes instead again $E_{XC} = \int \epsilon_{XC}(\boldsymbol{r}) d\boldsymbol{r}$, but now

$$\epsilon_{XC}(\boldsymbol{r}) = \epsilon_{XC}\left(n(\boldsymbol{r}), \nabla n(\boldsymbol{r}), \tau(\boldsymbol{r})\right), \tag{16.40}$$

so the local exchange-correlation energy per volume, ϵ_{XC}, relies on the density and its gradients, like in GGA; but, in addition, also on the **kinetic-energy density,**

$$\tau(\boldsymbol{r}) \equiv \frac{1}{2m_e} \sum_i n_i \nabla\phi_i^*(\boldsymbol{r}) \cdot \nabla\phi_i(\boldsymbol{r}). \tag{16.41}$$

Such meta-GGA functionals have excellent accuracy at fairly limited computational cost.

Another related set of functionals adds to hybrids other perturbation-energy terms, i.e., the total energy is augmented by (a constant) times the MP2 energy, which we will discuss in later chapters. Basically, this MP2 energy is just the van-der-Waals interaction. The idea is that DFT has a hard time describing the spontaneous polarization which is at the core of vdW interaction.

Still another set adds the vdW interactions empirically, via terms that depend only on the positions of the nuclei (i.e., are modified classical nuclear interactions). Those functionals do surprisingly well for giant biological systems.

There are many other combinations, and a lot of fits of functionals with many terms, which fit either specific sets of molecules (organics, TM's, etc.) or have enough terms to aspire to describe "all elements".

In fact, looking at the literature of DFT, one sympathizes at times with the quote attributed to Sir Rutherford ("Science consists only of physics and stamp-collecting")...

I will describe however in the next section a fundamental property that DFT should fulfill which inspired some excellent functionals, to prove to you that DFT can be on the fundamental side.

Piecewise linearity of the total Energy and implications for DFT

To start, let's discuss:

Energy differences rise with charging state

Our starting point is the realization (we won't prove it, but it makes sense for us as chemists) that in any molecular system, no matter how large or small, the energy differences upon charging rise with the charging state (assuming of course for simplicity, as we always do, that the nuclei are fixed).

For example, for a neutral system with N electrons, the ionization energy is always higher than the electron affinity. In obvious notation

$$E(N) - E(N-1) > E(N+1) - E(N) \tag{16.42}$$

This is true always, even if the system is not neutral.

Let's use this to define:

Energy of a partially charged system

What should be the energy of, for example, benzene with 41.8 electrons, i.e., $C_6H_6^{+0.2}$? (since neutral C_6H_6 has 42 e's.)

> Side note: in the specific case of benzene there would be problems due to degeneracy, so let's assume this is twisted benzene, where the atoms are not in a perfect hexagon, so no state is degenerate,

We define the energy of $C_6H_6^{+0.2}$ as the energy per benzene molecule of an ensemble of, say, 1000 well-separated benzene molecules – of which 800 are neutral and 200 are singly charged.

Also, to avoid writing so many C_6H_6 below, I abbreviate $A \equiv C_6H_6$.

$$E\left(A^{+0.2}\right) \equiv \frac{800 \cdot E\left(A\right) + 200 \cdot E\left(A^{+1}\right)}{1000}, \tag{16.43}$$

i.e.,

$$E(N=41.8) \equiv \frac{800 \cdot E(N=42)+200 \cdot E(N=41)}{1000}. \tag{16.44}$$

You may ask: is this prescription well defined? After all, we could have other combinations of 1000 molecules with the same total number of electrons (41.8*1000=41,800)!

For example, one benzene anion, 798 neutral molecules and 201 singly charged benzenes (all well-separated!). Same total number of electrons. It is easy to see however that this new configuration has a higher energy than the one we defined above. Specfcially, define

$$\begin{aligned}\Delta E \equiv & \left[1 \cdot E\left(A^{-1}\right)+798 \cdot E(A)+201 \cdot E\left(A^{+1}\right)+\right] \\ & -\left[800 \cdot E(A)+200 \cdot E\left(A^{+1}\right)\right]\end{aligned} \tag{16.45}$$

so

$$\Delta E=1 \cdot E\left(A^{-1}\right)-2 \cdot E(A)+1 \cdot E\left(A^{+1}\right)=\left[E\left(A^{+1}\right)-E(A)\right]-\left[E(A)-E\left(A^{-1}\right)\right] \tag{16.46}$$

i.e.,

$$\Delta E=\mathrm{I.E.}(A)-\mathrm{E.A.}(A)>0 \text{ as stipulated.} \tag{16.47}$$

For this reason, the fractional energy is always the weighted average of the energy of the molecules/ions with an integer charge just below and above the fractional charge. So in general

$$E(N+\delta)=(1-\delta) \cdot E(N)+\delta \cdot E(N+1) \quad \delta \in[0,1] \tag{16.48}$$

and of course

$$E(N-\delta)=(1-\delta) \cdot E(N)+\delta \cdot E(N-1) \quad \delta \in[0,1] \tag{16.49}$$

Corollary: $\boldsymbol{I.E.}=-\boldsymbol{\varepsilon_{HOMO}}$

As a corollary of the relation above,

$$-\left.\frac{\partial E(N-\delta)}{\partial \delta}\right|_{\delta \in[0,1]}=E(N-1)-E(N)=I.E.(=\text{ionization energy}). \tag{16.50}$$

i.e., the slope of the energy as function of N on the left side of an integer occupation (i.e., the left-derivative near the neutral point) is the ionization energy.

On the other hand, we explained that as you reduce the total charge slightly, the occupation of the HOMO would be the only one affected (since it is the highest energy one among the occupied states, .i.e., analogous to the highest chemical potential phase with non-zero occupancy). So this, as we saw earlier, implies

$$\frac{\partial E(N-\delta)}{\partial \delta}|_{\delta\in[0,1]} = \varepsilon_{HOMO}, \tag{16.51}$$

and therefore, the "true" DFT must fulfill

$$I.E. = -\varepsilon_{HOMO}. \tag{16.52}$$

Piecewise linear energy curves

The eqs. above indicate that the true molecular energy should be piecewise linear (see plot in class). However, most functionals won't obey this automatically.

Specifically, we can always do an LDA (or other DFT) calculation with non-integer occupations; for example, to calculate benzene with 41.8 electrons, we could do a calculation with 20 closed-shell filled orbitals, and the HOMO orbital with occupation of 0.9, i.e.:

$$n_{HOMO(i=21)} = 0.9 \tag{16.53}$$

$$n_i = 1, \quad i = 1, ..., 20 \tag{16.54}$$

And when we do such calculations, we find that: **LDA (and GGA) is generally convex upward.** This is the same as saying that in LDA and other similar functionals $I.E. \neq \varepsilon_{HOMO}$.

Hartree-Fock has the opposite behavior (see figure in class to make sense of this statement). In fact, one of the reasons for B3LYP's success is that the combination of Hartree-Fock and local-DFT (GGA or LDA) makes the energy more piecewise linear than either method.

Some modern functionals have therefore built in this requirement of piecewise linearity.

We will discuss in a later section one particular such functional where piecewise linearity can be imposed, long-range hybrids. But before, let's see a related problem:

LDA (and most local functionals) do not sufficiently attract a distant electron

A severe fundamental problem in LDA (and other local DFT) is seen when we consider an isolated molecule, large or small, and consider what does the valence electron see when it is fairly far from the molecule.

This question is relevant for charge transfer, where A-B becomes A^+B^-. It is also relevant for Rydberg states, where an excited electron circles an atom or molecule from far away.

First, we know what the exact theory will say: if the valence electron is far from the others (from the nuclei and the core), then it will **see a net charge of $+1$*|e|** (or $+1$ in our e=1 notation), since the core-attraction of $+Z|e|$ is only partially canceled by the $(Z-1)$ other electrons, and the electron does not repel itself.

But in LDA, since almost all the density is near the core and nucleus (including the density of the valence electron), so far away the net electrostatic attraction (due to the combination of the electron-nuclear v with the repulsive v_H) would be close to 0! Further, the local v_{XC} would be too small far away (since it depends monotonically on the density, which is tiny at far-away distances).

Note that this problem DOES NOT AFFECT Hartree-Fock! There, the exact exchange has a term where the electron interaction with itself is removed. Specifically, recall how we removed in Eq. (13.33) the restriction $i \neq j$ since the $i = j$ term in the direct Hartree term (E_H) was canceled by an equal amount in the exchange term (K).

So, to recapture where we are:

- the problem emanates because in DFT an electron is repelled from its own density (which, even for a valence electron, is located mostly close to the molecule, not far away).
- In Hartree-Fock exchange cancels the spurious repulsion.
- This feature of Hartree-Fock is especially important for long-distance repulsion.
- And we don't want to keep the full Hartree-Fock, since Hartree-Fock theory does not have the extra terms that DFT has which allow it to be very accurate (in hybrids).

The solution (Savin, mid 1990's) is then: **keep 100% exchange, but only for long-distances!**

Specfcially, INSTEAD of taking $0.75E_{XC} + 0.25K$, one SEPARATES the electron-electron interaction in the exchange to two terms: long-distance and short-one. For example:

$$\frac{1}{|\boldsymbol{r}-\boldsymbol{r}'|} = \frac{\mathrm{e}^{-\gamma|\boldsymbol{r}-\boldsymbol{r}'|^2}}{|\boldsymbol{r}-\boldsymbol{r}'|} + \frac{\left(1-e^{-\gamma|\boldsymbol{r}-\boldsymbol{r}'|^2}\right)}{|\boldsymbol{r}-\boldsymbol{r}'|}. \tag{16.55}$$

(Note that this separation is only done in the exchange term, not in the Hartree-repulsion part.) In the equation above, the Gaussian in the first term (of the RHS) ensures that this first term will decay off at long distances.

In the exchange term ("K") then:

- The first, short-term part would be treated by DFT;
- The second, long-term part would be treated by Hartree-Fock, where the "1/|**r-r'**|" part is replaced by $\frac{\left(1-e^{-\gamma|\boldsymbol{r}-\boldsymbol{r}'|^2}\right)}{|\boldsymbol{r}-\boldsymbol{r}'|}$. This replacement affects both the total energy and the Kohn-Sham Hamiltonian.

And how to determine the free-parameter γ? The best way would be to impose the piecewise linearity, i.e., one varies γ until Eq. (16.48) is fulfilled, i.e., until the HOMO energy equals the ionization energy, and until the LUMO energy equals the electron affinity.

> Recall that the electron affinity is nothing more than the ionization energy of the negative anion; and the LUMO energy is the HOMO energy for the negative anion. See class for figure.

DFT Scaling

The scaling with number of electrons becomes an issue for hybrid functionals, so that routine simulations cannot be done for systems with hundreds or thousands of atoms, except for special cases (e.g., water, where the density matrix is very localized in space). My research group, colleagues and I are working on methods to reduce this scaling thus enabling large scale simulations for extended systems.

DFT conclusions

To conclude: DFT is the main method for doing electronic structure today. I overviewed the main aspects, there are many more. One very important part that will be covered in 115b is the response to static and especially dynamic perturbations, which affects spectroscopy, etc. It comes under the label: TDDFT (time-dependent DFT).

17 Beyond DFT: perturbation theory (MP2)

In the next three chapters we'll deal with other methods for electronic structure. The first is perturbation theory, and we'll deal with its most basic and important part, 2nd order perturbation theory, also called Moeller-Plessest 2nd order or just MP2.

Notation

- For simplicity, in this and the next chapters we'll consider open-shell, so there are "N" orbitals, and we will not talk much about spins.
- **Subscripts: In this section and henceforth** i, j, k refer to general indices; $b, c, d, ...$refer to occupied states ($b, c, d \leq N$); and $p, q, r, ...$refer to unoccupied states ($N < p, q, r$). (This is **not a universal notation**!)

DFT and Hartree-Fock: Separable Molecular Hamiltonian Approximation

Specifically, in either Hartree-Fock or DFT (regular or hybrid) one solves, as we saw, an eigenvalue equation,

$$\hat{f}\eta_i = \varepsilon_i \eta_i \tag{17.1}$$

and $\hat{f}$ is a one-electron operator.

We can therefore define a trivial many-electron operator from $\hat{f}$:

$$H_0 \equiv c_0 + \hat{f}(1) + ... + \hat{f}(N) = c_0 + \sum_i \hat{f}(i) \tag{17.2}$$

i.e., just a sum of non-interacting-electrons Hamiltonians. Here "c_0" is a constant that we'll tackle shortly.

This (many-electron) operator H_0 is what we call of course one-body operator, since the many electrons don't interact with each other.

The zero-order Slater-determinant

$$|\Psi_0\rangle = \det\{\eta_1, ..., \eta_N\} \tag{17.3}$$

associated with the lowest-energy orbitals in DFT or Hartree-Fock **is an eigenstate of** H_0 :

$$H_0|\Psi_0\rangle = E_0|\Psi\rangle, \tag{17.4}$$

where

$$E_0 = c_0 + \sum_{i=1}^{N} \varepsilon_i \tag{17.5}$$

Proof: a determinant is made from $N!$ different terms. Let's for example consider $N = 3$, a determinant of 3 orbitals. Take arbitrarily one of the N!=6 terms in the determinant, $\eta_2(1)\eta_3(2)\eta_1(3)$. Then

$$\begin{aligned} H_0\left(\eta_2(1)\eta_3(2)\eta_1(3)\right) &= \left(c_0 + \hat{f}\,(1) + \hat{f}\,(2) + \hat{f}\,(3)\right)\left(\eta_2(1)\eta_3(2)\eta_1(3)\right) \\ &= \left(c_0 + \varepsilon_2 + \varepsilon_3 + \varepsilon_1\right)\left(\eta_2(1)\eta_3(2)\eta_1(3)\right) = \\ &\qquad E_0\left(\eta_2(1)\eta_3(2)\eta_1(3)\right) \qquad Q.E.D. \end{aligned} \tag{17.6}$$

Note that when we want to emphasize that we are talking about N electrons, we'll write this ground-state energy as:

$$E_0(N).$$

The constant c_0 in H_0

c_0 is best determined to give the same expectation value of the true Hamiltonian and zero-order Hamiltonian for the $|\Psi_0\rangle$

$$\langle\Psi_0|H|\Psi_0\rangle = \langle\Psi_0|H_0|\Psi_0\rangle. \tag{17.7}$$

But the LHS is just the DFT energy (E_{DFT}) or Hartree Fock energy, $E_{\text{H.F.}}$ which we label as a generic $\langle H\rangle$; the RHS is, from above, $\left(c_0 + \sum_{i=1}^{N}\varepsilon_i\right)$. So together

$$\langle H\rangle = c_0 + \sum_{i=1}^{N}\varepsilon_i \tag{17.8}$$

so, by Eq. (14.31), adopted now for the general (not necessarily closed-shell) case,

$$c_0 = \langle H\rangle - \sum_{i=1}^{N}\varepsilon_i = -\left(E_H + K\right). \tag{17.9}$$

Excited states of H_0 and modified Koopman's theorem

Note that H_0 is nothing more than the regular first-year chem. molecular orbital Hamiltonian; its excited states are determinants where occupied levels are promoted from ground-states, a language you are familiar with.

For example, a singly-excited determinant where the occupied orbital η_b is replaced by the excited orbital η_p (non-occupied in the ground-state, so $N < p$) would be denoted here as $|\Psi_b^p\rangle$

$$|\Psi_b^p\rangle \equiv \det\{\eta_1, ..., \eta_{b-1}\eta_p\eta_{b+1}, ..., \eta_N\}, \quad b \le N, \quad N < p \tag{17.10}$$

and is labeled **either as**

- Singly-excited determinant (a notation some chemists use).
- a particle-hole excitation; here the "particle" is the electron in the excited orbital "p", and the "hole" is the removal of the electron from the occupied orbital "j". Physicists like this notation, as well as some chemists.

We similarly define doubly-excited (or two-particle, two-hole) determinants,

$$|\Psi_{bc}^{pq}\rangle \equiv \det\{\eta_1, ..., \eta_{b-1}\eta_p\eta_{b+1}, ..., \eta_{c-1}\eta_q\eta_{c+1}, ..., \eta_N\}, \quad b < c \leq N, \quad N < p < q \tag{17.11}$$

etc. for triply excited determinants. Note that in this eq., we must order the holes and the particles to avoid double counting.

It is not too hard to prove **(prove yourself)** that the associated excited-state fulfill

$$H_0|\Psi_b^p\rangle = E_b^p|\Psi_b^p\rangle \tag{17.12}$$

where the $b \to p$ excited state energy is just $E_b^p = E_0 + \varepsilon_p - \varepsilon_b$, i.e., the excitation energy is just the orbital-energy difference between the occupied and excited orbital:

$$E_b^p - E_0 = \varepsilon_p - \varepsilon_b \tag{17.13}$$

just like our intuition says.

This eq. above is called at times "Koopman's theorem", although Koopman's theorem really talks about the ionization energy to remove the HOMO being equal to the energy of the HOMO; i.e., if we take the N-1 determinant by removing the HOMO orbital,

$$|^{N-1}\Psi_0\rangle \equiv |\eta_1, ..., \eta_{N-1}\rangle \tag{17.14}$$

then the actual Koopman's theorem states that (in obvious notation) the ionization energy equals the HOMO energy

$$E_0(N-1) - E_0(N) = \varepsilon_N (= \varepsilon_{HOMO}). \tag{17.15}$$

and a similar theorem follows for the LUMO. We won't prove this theorem in Hartree-Fock, it is too long for us.

Note that, as we discussed in the DFT chapter, Koopman's theorem would have been exact if we had the true wavefunctions (or the true functional).

But with Hartree-Fock, Koopman's theorem is approximate for both the one-particle excitation (Eq. (17.13)) and for ionization, Eq. (17.15), for two reasons:

- Hartree-Fock is an approximate theory
- Even with the Hartree-Fock picture, we could expect some orbital change ("relaxation") when an electron is excited from one-orbital to the next. For example, in an A-B molecule, where A is more electronegative than

B, we could excite am electron initially mostly on A to an excited state that's mostly on B, or even ionize it. Then, the other electrons could "rearrange" (orbital relaxation) to respond to the change in the HOMO-electron distribution.

Properly accounting for orbital relaxation takes us outside Hartree-Fock and into what's so-called RPA (also called in this context singles-CI, and also in a sense part of TDDFT) or to Green's functions theories, and will be tackled in 115b/215b.

Finally, obviously, we can repeat the steps to show that

$$E_{bc}^{pq} - E_0 = \varepsilon_p + \varepsilon_q - \varepsilon_b - \varepsilon_c. \tag{17.16}$$

Now that we have H_0 with its eigenstates, we just need the perturbation in order to apply perturbation theory. For this, we'll need some simple clarification:

Connection between the one-body, two-body, Fock operator and the total Hamiltonian

The total Hamiltonian is, I remind you:

$$H = H_{1B} + \hat{U} \tag{17.17}$$

where $H_{1B} = \sum_{i=1}^{N} \hat{h}(i)$, where $\hat{h}$ is, as I remind you, the kinetic plus electron-nuclei interaction, so

$$H = \sum_{i=1}^{N} \hat{h}(i) + \hat{U} \tag{17.18}$$

(We never put a "hat" above H, H_0, H_{1B}, but they are presumed to be operators.)

Similarly, as we saw, $H_0 = c_0 + \sum_{i=1}^{N} \hat{f}(i)$. Therefore, the **perturbation Hamiltonian** is

$$H_1 = H - H_0 = \hat{U} - (H_0 - H_{1B}) = \hat{U} - c_0 - \sum_{i=1}^{N} \left(\hat{f}(i) - \hat{h}(i)\right) \tag{17.19}$$

Matrix elements of the perturbation

Brillouin's theorem: no first-order matrix element with single-electron single-hole determinant

When we apply perturbation theory, a pleasant surprise emerges: there's no matrix element between the ground-state Slater-determinant and singly-excited determinants!

$$\langle \Psi_b^p | H | \Psi_0 \rangle = 0 \tag{17.20}$$

which, since $|\Psi_0\rangle$ and $|\Psi_b^p\rangle$ are different eigenstate of H_0, and therefore $\langle\Psi_b^p|H_0|\Psi_0\rangle = 0$ can also be written as

$$\langle\Psi_b^p|H_1|\Psi_0\rangle = \langle\Psi_b^p|H - H_0|\Psi_0\rangle = 0 \tag{17.21}$$

This is Brillouin's theorem. We won't prove it here, but with the machinery we'll study in the next chapter it could be tackled.

Matrix elements with two-electron excitation

Similarly, I'll quote here (and prove in the next chapter):

$$\langle\Psi_{bc}^{pq}|H|\Psi_0\rangle = (\eta_p\eta_q|\eta_b\eta_c) - (\eta_p\eta_q|\eta_c\eta_b) \tag{17.22}$$

where I defined (reminiscent of Eq. (14.48), but now for the Hartree-Fock (or DFT) molecular orbitals)

$$(\eta_p\eta_q|\eta_b\eta_c) \equiv \int \frac{\eta_p^*(1)\eta_q^*(2)\eta_b(1)\eta_c(2)}{|\boldsymbol{r_1} - \boldsymbol{r}_2|}\, d1\, d2. \tag{17.23}$$

i.e., the matrix element of an excitation is just the antisymmetrized interaction of the excited and occupied electrons!

Note in passing that the last matrix element will vanish unless the spin of "p" and "b" match, and the same for the spins of "q" and "c".

No matrix elements with more than two-electron two-hole excitations

We can prove (hopefully next chapter) that the matrix element of the ground-state with three-electron excited determinants or higher excitations vanishes, $\langle\Psi_{bcd}^{pqr}|H|\Psi_0\rangle = 0$, etc.

MP2 energy-correction

We are thus ready to apply the 2nd order perturbation theory.

As we saw, the 1st order correction vanishes by construction, i.e., $\langle\Psi_0|H - H_0|\Psi_0\rangle = 0$.

So we need to go to 2nd order, i.e., apply Eq. (9.20). We need to sum over all excited-states that couple to the ground-state, i.e., over all two-particle two-hole Slater determinants, so the 2nd-order energy correction to the ground-state reads in this context:

$$\delta E^{MP2} = -\frac{1}{2}\sum_{b<c\le N}^{N<p<q} \frac{|\langle\Psi_{bc}^{pq}|H|\Psi_0\rangle|^2}{E_{bc}^{pq} - E_0}, \tag{17.24}$$

and from the eq. above:

$$\delta E^{MP2} = -\frac{1}{2}\sum_{i<j\le N}^{N<p<q} \frac{\Big|(\eta_p\eta_q|\eta_b\eta_c) - (\eta_p\eta_q|\eta_c\eta_b)\Big|^2}{\varepsilon_p + \varepsilon_q - \varepsilon_b - \varepsilon_c}. \tag{17.25}$$

This MP2-energy eq. is the highlight of this chapter. It is exactly what we learn in first-year chem.; the van-der-Waals interaction has to do with **spontaneous excitations (**electron excited from orbital “b” to unoccupied orbital “p”) **interacting with other spontaneous excitations** (“c” to “q”) – and in this case, antisymmetrization imposes more terms.

Note that the MP2 expression is modified with numerical terms of 4 and 2 when we use closed-shell determinants, but this won’t be presented here.

Beyond MP2

We can continue to 3rd, 4th etc. perturbation theory. Unfortunately, the perturbation theory expression is asymptotic, not convergent, which means that in some order of perturbation theory it will start exploding. Usually, stopping at the level of MP2 gives very good results, and is not too expensive.

At present, as mentioned, a hot trend is to combine MP2 with DFT, as explained in the previous chapter.

Finally, if one wants really accurate results other techniques are needed – primarily full-CI or coupled-clusters; we’ll tackle these in the next chapters.

18 2nd quantization: Creation and destruction operators.

Notation and reminders

- Henceforth, $|\eta_{i_1}\eta_{i_2}\eta_{i_3}...\eta_{i_N}\rangle \equiv \det\{\eta_{i_1}\eta_{i_2}\eta_{i_3}...\eta_{i_N}\}$refers to **a Slater determinant** (with the $1/\sqrt{N!}$ factor), i.e., the **antisymmetrized product of the states.**
- Recall that
$$|...\eta_i...\eta_j...\rangle \equiv -|...\eta_j...\eta_i...\rangle \tag{18.1}$$
and that as a result a determinant with the same two orbitals will vanish
$$|...\eta_i...\eta_i...\rangle \equiv -|...\eta_i...\eta_i...\rangle \rightarrow |...\eta_i...\eta_i...\rangle = 0 \tag{18.2}$$
- We continue to use open-shell case, N separate states.
- We will use a complete basis of molecular orbitals, η_i, where i extends from 1 to infinity; i.e., the first N such functions are the occupied M.O. and the others are the unoccupied ones.
- We define the "vacuum" state, $|\text{vac}\rangle$ or simply $|\rangle$ as a state with no electron. Note that this is not a null (zero) vector; it's norm is $\langle\text{vac}|\text{vac}\rangle = \langle|\rangle = 1$.
- Finally, many of my definitions here are **non-universal and/or non-traditional.**

Creation and destruction operators

These operators have resemblance to (and big differences from) the raising and lowering operators in H.O., and we usually use the same symbol, but this time with a subscript, a_i and a_i^+, as defined below.

Creation operator

The **creation** operator for the i'th state, a_i^+, takes any Slater determinant with, say, N electrons (where N is any integer, including 0) and turns it to a Slater-determinant for N+1 electrons, with the extra electron in the i'th state:

$$a_i^+|\eta_j\eta_k....\rangle = |\eta_i\eta_j\eta_k....\rangle. \tag{18.3}$$

Note that you cannot put two electrons in the same orbitals, so acting with the same creation orbital twice leads to a vanishing result,

$$a_i^+a_i^+|\eta_j\eta_k....\rangle = a_i^+|\eta_i\eta_j\eta_k....\rangle = |\eta_i\eta_i\eta_j\eta_k....\rangle = 0 \tag{18.4}$$

(Mathematically, this is because a determinant with two equal columns will vanish.)

So therefore we can write

$$a_i^+ a_i^+ = 0. \tag{18.5}$$

Similarly, for any determinant

$$a_i^+ a_j^+ |....\rangle = |\eta_i \eta_j ....\rangle = -|\eta_j \eta_i ....\rangle = -a_j^+ a_i^+ |....\rangle, \tag{18.6}$$

so in general

$$a_i^+ a_j^+ = -a_j^+ a_i^+ \tag{18.7}$$

We can therefore write in general

$$\{a_i^+ ,\, a_j^+\} = 0,$$

where we defined the anticommutator

$$\{A, B\} = AB + BA. \tag{18.8}$$

Destruction operator

The destruction operator for a state i , written as a_i, takes a Slater determinant with N electrons in which one is in the i'th states and turns it to a Slater determinant for N-1 electrons where none is in the i'th state, i.e., it removes the i'th orbital from the determinant (and lowers therefore the number of states and electrons).

$$a_i |\eta_i \eta_j \eta_k ...\rangle = |\eta_j \eta_k, ...\rangle. \tag{18.9}$$

If the i'th state is not in the determinant, then the result is 0. As a corollary, if we apply a_i twice, a determinant will vanish.

Example: $a_5 a_5 |\eta_3 \eta_5 \eta_2 \eta_9\rangle = -a_5 a_5 |\eta_5 \eta_3 \eta_2 \eta_9\rangle = -a_5 |\eta_3 \eta_2 \eta_9\rangle = 0$. Thus $a_i a_i = 0$. Similarly,

$$\{a_i, a_j\} = 0.$$

Proof (using colors!) –

$$\begin{aligned}\{a_i, a_j\} |\eta_i \eta_j \eta_k ...\rangle = a_i a_j |\eta_i \eta_j \eta ...\rangle + a_j a_i |\eta_i \eta_j \eta ...\rangle = \\ - a_i a_j |\eta_j \eta_i \eta_k ...\rangle + a_j a_i |\eta_i \eta_j \eta_k ...\rangle = \\ - a_i |\eta_i \eta_k ...\rangle + a_j |\eta_j \eta_k ...\rangle = -|\eta_k ...\rangle + |\eta_k ...\rangle = 0.\end{aligned} \tag{18.10}$$

The last proof was for for $i \neq j$, and for $i = j$ we proved earlier.

Anticommuting a creation and destruction operator:

We can prove along the same lines

$$\{a_i^+,\, a_j\} = 0 \text{ for } i \neq j \text{ (prove it!).} \tag{18.11}$$

A more interesting case is

$$\left\{a_i^+,\, a_i\right\} = 1.$$

Proof: let's apply the anticommutator on a Slater determinant. There are two possibilities: the determinant contains the η_i state, or it does not.

In the first case (presuming η_i is the first function in the determinant, if not then permute the orbitals, perhaps getting an overall minus sign):

$$\left\{a_i^+,\, a_i\right\} |\eta_i\eta_k...\rangle = a_i^+ a_i|\eta_i\eta_k...\rangle + a_i a_i^+|\eta_i\eta_k....\rangle = a_i^+|\eta_k...\rangle + 0 = 1 \cdot |\eta_i\eta_k...\rangle$$

so OK. In the second case:

$$\left\{a_i^+,\, a_i\right\} |\eta_k...\rangle = a_i^+ a_i|\eta_k...\rangle + a_i a_i^+|\eta_k....\rangle = 0 + a_i|\eta_i\eta_k...\rangle = 1 \cdot |\eta_k...\rangle$$

so works again!

So summarizing:

$$\left\{a_i^+,\, a_j^+\right\} = 0 \tag{18.12}$$
$$\{a_i,\, a_j\} = 0 \tag{18.13}$$
$$\left\{a_i^+,\, a_j\right\} = \delta_{ij} \tag{18.14}$$

a_i^+ is indeed the conjugate of a_i

So far we assumed it, let's prove by example:

$$\langle\eta_3\eta_2|a_7|\eta_3\eta_7\eta_2\rangle = -\langle\eta_3\eta_2|a_7|\eta_7\eta_3\eta_2\rangle = -\langle\eta_3\eta_2|\eta\eta_2\rangle = -1$$

$$\langle\eta_3\eta_7\eta_2|a_7^+|\eta_3\eta_2\rangle = \langle\eta_3\eta_7\eta_2|\eta_7\eta_3\eta_2\rangle = -\langle\eta_3\eta_7\eta_2|\eta_3\eta_7\eta_2\rangle = -1, \quad Q.E.D.$$

Slater-Determinants in terms of creation/destruction operators

It is easy to see that a general Slater determinant becomes the result of acting successively with creation operators on the vacuum state. For example, for the Hartree-Fock fully occupied determinant:

$$|\eta_1...\eta_N\rangle = a_1^+ a_2^+ ... a_N^+|\rangle \tag{18.15}$$

The Hamiltonian in terms of creation and destruction operators

The beauty of the 2nd-quantization approach (although it is a misleading name) is that the difficult dealings with Slater determinants become trivial manipulations of creation and destruction operators.

One-body Hamiltonian

Claim: I can replace T, or V, or $H_{1B} \equiv T + V$, or of any-other one-electron operator by a simple related term involving a pair of creation-destruction operator; i.e.,

$$H_{1B} = \sum_{i,j} h_{ij} a_i^{+} a_j. \tag{18.16}$$

To "prove" this (below, actually we'll just exemplify, not prove) we will need to apply this equation to an arbitrary Slater determinant

$$H_{1B}|\Psi\rangle = \sum_{i,j} h_{ij} a_i^{+} a_j |\Psi\rangle. \tag{18.17}$$

Note that the sum extends over all values of i and j from 1 to infinity. Here we define the matrix element of the 1-electron operator in terms of the MO states (both occupied and unoccupied),

$$h_{ij} = \langle \eta_i | \hat{h} | \eta_j \rangle \tag{18.18}$$

and we recall again the relation between H_{1B} and the 1-electron operator; i.e., for an N-electron determinant:

$$H_{1B} = \hat{h}(1) + \hat{h}(2) + ... + \hat{h}(N). \tag{18.19}$$

One-body form with creation/lowering operators: Proof by example

Let's prove by example again, it's more illuminating.

Recall first that for any operator, due to completeness

$$\hat{h} = \sum_{ij} |\eta_i\rangle h_{ij} \langle \eta_j|. \tag{18.20}$$

Now use an N=2 example for simplicity, and be lax with notation.

First, the LHS of Eq. (18.17)

$$H_{1B}|\eta_4\eta_6\rangle = \left(\hat{h}(1) + \hat{h}(2)\right) \frac{(\eta_4(1)\eta_6(2) - (1\leftrightarrow 2 \text{ permutation})}{\sqrt{2}} =$$

$$\frac{\left(\hat{h}(1)\eta_4(1)\right)\eta_6(2) + \eta_4(1)\left(\hat{h}(2)\eta_6(2)\right) - (1\leftrightarrow 2 \text{ permutation})}{\sqrt{2}}, \tag{18.21}$$

and using $\hat{h}|\eta_4\rangle = \sum_i h_{i4}|\eta_i\rangle$, we get

$$H_{1B}|\eta_4\eta_6\rangle = \sum_i \frac{\left(h_{i4}\eta_i(1)\right)\eta_6(2) + \eta_4(1)\left(h_{i6}\eta_i(2)\right) - (1{\leftrightarrow}2 \text{ permutation})}{\sqrt{2}}. \tag{18.22}$$

But the permutation term is:

$$-\,(1{\leftrightarrow}2 \text{ permutation}) = -\left(\eta_6(1)h_{i4}\eta_i(2) + h_{i6}\eta_i(1)\eta_4(2)\right), \tag{18.23}$$

so overall, collecting the last two equations

$$H_{1B}|\eta_4\eta_6\rangle = \sum_i \frac{h_{i4}\left(\eta_i(1)\eta_6(2) - \eta_6(1)\eta_i(2)\right) + h_{i6}\left(\eta_4(1)\eta_i(2) - \eta_6(1)\eta_i(2)\right)}{\sqrt{2}} \tag{18.24}$$

i.e.,

$$H_{1B}|\eta_4\eta_6\rangle = \sum_i h_{i4}|\eta_i\eta_6\rangle + \sum_i h_{i6}|\eta_4\eta_i\rangle. \tag{18.25}$$

Now to the RHS of Eq. (18.17). It is

$$\sum_{i,j} h_{ij}a_i^+ a_j|\Psi\rangle = \sum_{i,j} h_{ij}a_i^+ a_j a_4^+ a_6^+|\rangle, \tag{18.26}$$

but $\{a_j, a_4^+\} = \delta_{j4}$, so $a_j a_4^+ = \delta_{j4} - a_4^+ a_i$ and similarly for $a_j a_6^+$, so

$$\begin{aligned}
\sum_{i,j} h_{ij}a_i^+ a_j|\Psi\rangle &= \sum_{i,j} h_{ij}a_i^+\left(\delta_{j4} - a_4^+ a_j\right)a_6^+|\rangle \\
&= \sum_i h_{i4}a_i^+ a_6^+|\rangle - \sum_{ij} h_{ij}a_i^+ a_4^+ a_j a_6^+|\rangle = \\
&= \sum_i h_{i4}|\eta_i\eta_6\rangle - \sum_{ij} h_{ij}a_i^+ a_4^+\left(\delta_{j6} - a_6^+ a_j\right)|\rangle = \\
&\sum_i h_{i4}|\eta_i\eta_6\rangle - \sum_i h_{i6}|\eta_i\eta_4\rangle + 0
\end{aligned} \tag{18.27}$$

where we used $a_j|\rangle = 0$, i.e., you cannot remove an electron at any level "j" from the vacuum. The RHS of Eq. (18.17) is therefore (since $|\eta_i\eta_4\rangle = -|\eta_4\eta_i\rangle$):

$$\sum_{i,j} h_{ij}a_i^+ a_j|\Psi\rangle = \sum_i h_{i4}|\eta_i\eta_6\rangle + \sum_i h_{i6}|\eta_4\eta_i\rangle, \tag{18.28}$$

i.e., matches the LHS of Eq. (18.17). Therefore Eq. (18.17) is correct for this 2-electron case (N=2), and it turns out to be correct for any N as we asserted.

Two-body Hamiltonian

Avoiding the even more tedious proof (or example) for the two-body term, we'll just quote the results. Note the order of indices (see the color)!

$$\hat{U} = \frac{1}{2}\sum_{ijkl}(\eta_i\eta_j|\eta_l\eta_k)a_i^+a_j^+a_k a_l, \tag{18.29}$$

where we defined earlier the 2-body matrix elements.

Two-body matrix elements

We can now prove Eq. (17.22), $\langle\Psi_{bc}^{pq}|H|\Psi_0\rangle = (\eta_p\eta_q|\eta_b\eta_c) - (\eta_p\eta_q|\eta_c\eta_b)$, i.e., the overlap between a 2-particle 2-hole state and the ground-state is just the antisymmetrized matrix element associated with the specific indices of the holes and particles.

First, note that in our notation

$$|\Psi_{bc}^{pq}\rangle = a_p^+a_q^+a_c a_b|\Psi_0\rangle, \tag{18.30}$$

i.e., we remove two electrons from orbitals "b" and "c" and add two electrons to orbitals "p" and "q". Note that I use a particular order of indices, but we could have exchanged the order of a_c, a_b, and that would just give an overall minus, so it is just a matter of definition.

Now overlap it:

$$\langle\Psi_0|H|\Psi_{bc}^{pq}\rangle = \langle\Psi_0|H_{1B} + U|a_p^+a_q^+a_c a_b\Psi_0\rangle. \tag{18.31}$$

The first term on the RHS, $\langle\Psi_0|H_{1B}|a_p^+a_q^+a_c a_b\Psi_0\rangle$, vanishes since $H_{1B} = \sum_{ih} h_{ij}a_i^+a_j$ can at most do a single hole-particle deexcitaton, so it cannot bring the two-electron two-hole term to overlap back with Ψ_0.

> Side-note, if you prefer to see it in numbers, take N=5, and $p, q, c, b = 6, 8, 5, 3$. Then the associated 2-particle 2-hole excitation would be $a_6^+a_8^+a_5a_3|\Psi_0\rangle = (\pm 1)|\eta_1\eta_2\eta_4\eta_6\eta_8\rangle$, (I am too lazy to work out the sign) and clearly no single $a_i^+a_j$ term can take it back to $|\eta_1\eta_2\eta_3\eta_4\eta_5\rangle = |\Psi_0\rangle$. End of note.

So we only need to take matrix elements of U, i.e.,

$$\langle\Psi_0|H|\Psi_{bc}^{pq}\rangle = \langle\Psi_0|U|a_p^+a_q^+a_c a_b\Psi_0\rangle = \frac{1}{2}\sum_{ijkl}(ij|lk)\langle\Psi_0|a_i^+a_j^+a_k a_l a_p^+a_q^+a_c a_b\Psi_0\rangle. \tag{18.32}$$

This looks horrendous, but is in practice quite easy. Let's use colors:

$$\sum_{ijkl}(ij|lk)\langle\Psi_0|a_i^+a_j^+a_k a_l a_p^+a_q^+a_c a_b\Psi_0\rangle. \tag{18.33}$$

The only non-vanishing terms must have the <u>brown-indices match</u> and the <u>blue-indices match</u>;

(Proof: i.e., the $a_c a_b$ term removes two occupied states, and of the four creation-operators in the term, the $a_p^+ a_q^+$do not create electrons in the occupied states of Ψ_0, so $a_i^+ a_j^+$ must do it., etc.)

Thus, we can restrict the summation to: $i, j = b, c$ or c, b; and similarly to $k, l = p, q$ or q, p.

Let's calculate:

- the 1st term: $i, j = b, c$ and $k, l = p, q$ (and note that all 4-indices (b, c, p, q) must be different from each other). Therefore (I bold exchanged terms

  $$\langle\Psi_0|a_b^+ a_c^+ a_{\boldsymbol{p}} a_{\boldsymbol{q}} a_p^+ a_q^+ a_c a_b|\Psi_0\rangle = (-1)\langle\Psi_0|a_b^+ a_c^+ a_{\boldsymbol{q}} a_p a_p^+ a_q^+ a_c a_b|\Psi_0\rangle = -\langle\Xi|\Xi\rangle = -1 \tag{18.34}$$

  where $|\Xi\rangle = a_p a_q^+ a_c a_b|\Psi_0\rangle$. Therefore, the first contribution to Eq. (18.33) will be $-\frac{1}{2}(pq|cb)$

- The 2nd term $i, j = \boldsymbol{c}, \boldsymbol{b}$ and again $k, l = p, q$ gives $+\frac{1}{2}(pq|bc)$

- The 3rd and 4'th terms duplicate the 2nd and 1st term, respectively.

We therefore end up with

$$\langle\Psi_0|H|\Psi_{bc}^{pq}\rangle = -\frac{1}{2} \cdot 2 \cdot (pq|cb) + \frac{1}{2} \cdot 2 \cdot (pq|bc) = (pq|bc) - (pq|cb). \tag{18.35}$$

as stipulated in the perturbation-theory chapter!

Number-operator

It is interesting and at times useful to define the number operator,

$$\hat{N} = \sum_{i=1}^{\infty} a_i^+ a_i. \tag{18.36}$$

It is easy to see that given a Slater determinant, applying $\hat{N}$ on it amounts to multiplying it by N, the number of orbitals in the determinant (**prove it!**). Also, $\hat{N}$ commutes with H, which makes sense (the Hamiltonian conserves the number of electrons).

The number operator is very useful in statistical mechanics of fermions at finite temperatures in the canonical ensemble, where one does not seek anymore the ground-state energy at a fixed number of particles, but instead is aiming at calculating the partition function,

$$Z = \mathrm{Tr} e^{-\beta(\hat{H}-\mu\hat{N})}. \tag{18.37}$$

19 Higher Order theories – brief overview

There are many other higher order methods that get (or aim at getting) highly accurate results. Most of those techniques aim at establishing **size-consistency,** which means on a practical level that the method does not degrade as the system size increases (or formally that if one takes two separate systems, A and B, then $E(A+B)=E(A)+E(B)$).

A partial list is:

Lower-Accuracy approaches: Tight Binding

Let's actually start by a **lower accuracy** approach, the Tight-binding (or "semi-empirical") family of methods. Here, one uses a minimal basis set (or close to it) and simplifies the Hartree-Fock Hamiltonian (in a basis set) by throwing out altogether the Hartree and exchange terms, or at most keeping a smaller subset of such terms. Instead of these terms, one parametrizes the remaining one-body term so they are not the true matrix elements of $\hat{h}$, but instead are parameterized to fit small-molecules studies.

Tight-binding methods have been very important in the past, but as computer power grows they are reserved now only for very large systems.

The most modern tight-binding approach is the DFT-TB method.

Coupled Cluster:

Here one assumes an ansatz to the true wavefunction $|\Psi\rangle$ as an **exponential** of two or more (usually three) types of terms times a Slater determinant $|\Psi_0\rangle = \det\{\eta_1\eta_2...\eta_N\}$

$$|\Psi\rangle = e^{O_1+O_2+O_3\cdots}|\Psi_0\rangle \tag{19.1}$$

where the operators are one-body, two-body and three-body operators:

$$O_1 = \sum_{b\leq N<p<M} t_b^p a_p^+ a_b, \tag{19.2}$$

$$O_2 = \sum_{b<c\leq N<p<q<M} t_{bc}^{pq} a_q^+ a_p^+ a_c a_b, \tag{19.3}$$

etc. Note that all the terms in O_1 and O_2 involve only destruction operators over the occupied η_b states, and creation operators over the unoccupied states η_p. This makes the math doable, though still very expensive $\mathrm{O}(N^6)-\mathrm{O}(N^8)$ so it can only be applied to small molecules, usually to gauge the quality of DFT functionals.

The presence of the exponential term in the Coupled Clusters ansatz is what makes it extremely accurate, although there are cases it can break too (severe bond-breaking, and other cases where a single zero-order determinant is a bad approximation to the underlying chemistry).

The exponential term also makes Coupled Clusters size-consistent. It also makes it invaluable to use the 2nd quantization language we learned in the previous chapter

GW and other Green's function methods (e.g., GF2).

Those methods are based on the use of the Green's function (hopefully covered in 115b/215b) and the concept of self-energy, an energy-dependent one-electron operator which, if it was exactly known, would give the exact ionization energy and electron affinity and even the optical spectra.

Stochastic GW: With our colleagues, our group was able to invent an extremely efficient (O(N)) GW approach that's faster than the underlying DFT for large systems (the "stochastic GW approach") that does not require a Gaussian basis so has no basis-set convergence issues, and works for thousands of atoms. GW gives excellent charging energies, to about 0.1eV.

GF2: an up and coming approach, which, in lay terms, can be thought of as self-consistent MP2. It also is amenable to **stochastic** treatments, and at present we developed an O(N^2) GF2 technique.

RPA techniques:

Akin to GW, and is complementary to TDDFT, as we'll cover in the next quarter. Gives excellent vdW energies and its scaling is not too severe.

Configuration Interaction methods

Restricted-CI and Full-CI (Configuration Interaction)

The brute-force electronic structure algorithm. In **Restricted-CI** one constructs all Slater-determinant up to a certain order of particle-hole excitations; for example, in CI-doubles one will take the space made from the Hartree-Fock determinant, one-electron excitation, and two-electron (i.e., two-particle, two-holes) excitations

$$|\Psi_0\rangle,\ |\Psi_b^p\rangle,\ |\Psi_{bc}^{pq}\rangle \quad b,c \le N < p,q < M \tag{19.4}$$

The problem with CI-doubles and similar approximations is that it is not size-consistent. (A double excitation on system$A$ and a double excitation on system B, would be a fourth-order excitation on a separated $A + B$).

In **Full-CI** one constructs all possible determinants. The method is self-consistent but scales terribly (combinatorially). For example, if $M = 50$ and $N = 25$ (a fairly tiny system) then there will be $\left(\begin{array}{c} 50 \\ 25 \end{array}\right)$ possible determinants; grouping them by the total spin helps only a little.

At present there are some new algorithm which aim at selecting only those determinants that will be important, by either a random process or by a similarly inspired (but deterministic) process; those can help but practically it is hard to see those methods scaling up to much more than a 10-heavy atom (not hydrogen) systems.

CASSCF and PT2-CASSCF:

CASSCF (Complete Active Space SCF) is similar to full-CI but here one takes a restricted set of valence orbitals. For example, if the basis set has $M = 200$, and there are $N = 50$ electrons, then perhaps one would take all determinants made from the first $N - 5 = 45$ orbitals (the "core") and the last five electrons would be placed in the ten orbitals $46, ..., 55$ just below and above the HOMO.

Then, this set of $\begin{pmatrix} 10 \\ 5 \end{pmatrix}$ possible determinants is used as a basis in which the Hamiltonian is diagonalized; a new density matrix is obtained, and a new set of occupied and virtual (i.e., non-occupied) orbitals is calculated, and the method is then iterated until convergence. This iteration procedure is why the method has SCF in its name, CASSCF.

An even better method is PT2-CASSCF, (PT2 stands for Perturbation Theory of 2nd order) where the effect of the orbitals outside the CASSCF space (in our case the first 45 and orbitals 55-200) is added perturbatively.

Variational techniques:

Here one multiplies the Slater determinant by a function of the electron coordinate. This function is usually an exponential of another term, $|\Psi\rangle = e^{g\cdot}|\Psi_0\rangle$, where g is usually a function of coordinates and is a sum of the electron-electron distance, e.g.

$$g(\boldsymbol{r_1}, \boldsymbol{r_2}, ..., \boldsymbol{r}_N) = \sum_{i \le j \le N} w(|\boldsymbol{r}_i - \boldsymbol{r}_j|) \tag{19.5}$$

and w is a function to be optimized which accounts for the modification of the wavefunction due to correlation when a pair of electrons gets too close to each other. Such a term, and many other possible ones, make the variational wavefunction highly accurate, so $\langle H\rangle = \langle\Psi|H|\Psi\rangle$ is an excellent variational estimate to the true energy; but at the same time they preclude the simple analytic evaluation of $\langle H\rangle$ as in Hartree-Fock and DFT.

Therefore, a Monte-Carlo technique needs to be employed, whereby one chooses randomly (subject to a proper "guiding" function) the positions of the N-electrons in space, $\boldsymbol{r}_1, \boldsymbol{r}_2, ..., \boldsymbol{r}_N$, and then calculate numerically

$$\langle H\rangle \approx \text{average}\left\{\Psi^* \left(\boldsymbol{r}_1, \boldsymbol{r}_2, ..., \boldsymbol{r}_N\right) H\Psi\left(\boldsymbol{r}_1, \boldsymbol{r}_2, ..., \boldsymbol{r}_N\right)\right\} \tag{19.6}$$

Diffusion Monte Carlo

An old and powerful technique which is continuously developed and is becoming chemically relevant. Here, instead of solving the Schrödinger equation, one solves the so-called **imaginary-time** Schrödinger eq.,

$$\frac{\partial|\Psi(\tau)\rangle}{\partial\tau} = -H|\Psi(\tau)\rangle \tag{19.7}$$

This is a diffusion equation which formally results from the Schrödinger equation if we replace τ by an imaginary time, $-it$ (and this is the origin of the name), although in practice we take real-valued τ.

The solution of this diffusion equation is easily shown to be, at long enough τ, just a $\tau-$dependent constant ($e^{-\beta E(\text{ground state})}$) times the ground-state of H.

The diffusion equation is solved by literally representing the wavefunction by thousands of particles; for example, in a problem with $N = 20$ electrons each of these "particles" would have $3N$ coordinates $(x_1, y_{1,}z_{1,}x_{2,}, ..., z_N)$.

At each time-step the potential energy term in the Hamiltonian acts to increase or decrease the weight of each of these particles, and the kinetic energy causes them to randomly diffuse away from their present position.

The one caveat in the technique (in addition to requiring a larger number of particles) is that it treats Ψ itself as a diffused density (not Ψ^2). Therefore, Ψ needs to be positive, or at least its nodes (region where it changes sign) need to be known. But due to the Pauli principle we know that the electronic wavefunction must have negative regions (and positive ones).

In practice, one therefore takes an approximate Slater determinant Ψ_0, and assumes that Ψ has the same nodes as Ψ_0 (so their product, $\Psi\Psi_0$ is never negative). But this **fixed-node** approximation turns out to be quite accurate and does not affect the energies much.

Other techniques

There are many other techniques that are becoming more and more important in chemistry; a partial list includes:

- DMRG and tensor-methods;
- Auxiliary-Fields Monte Carlo (in the Shifted-Contour Nodeless approach)
- Two-body density-matrix methods
- Geminal approaches

and many other techniques.

Embedding and fragment-based techniques

For large systems, we often care more about a specific region and the rest could be treated on a more approximate level. We then say that the inner region is embedded in the larger region. Such embedding methods are prevalent now, e.g.,

- QM-MM (important for biological simulations; an inner region treated quantally and the rest of the atoms classically);

- DFT region within a larger region treated Thomas-Fermi type methods or by Tight-Binding approaches.
- Coupled cluster for the inner region and DFT for the external one;
- Or the related **fragment-based** methods, whereby a large system is fragmented to several smaller systems that are treated quantally each and are "woven" together.

Contents

Source:

Much of the notes is based on Schatz and Ratner's Quantum Mechanics in Chemistry (Dover edition).

General Notation:

We use generally atomic units,$e = m_e = \hbar = 1$. So the basic time unit would be

$$1\text{a.u.(time)} = \hbar/\text{Hartree} = 24.18 \cdot 10^{-18}\text{s} = 0.02418\text{fs} = 24.18\,\text{as(i.e., attosec)}$$

But there will be times where, for clarity, we keep the $\hbar$.

20 Potential curves, the Born-Oppenheimer Approximation, Coupling matrices

Notation:

In this chapter we'll use a different notation from the previous (115a/215a) set: we denote by $\boldsymbol{r}$ the collection of all electron coordinates, both electron and spin (**not just a single-electron coordinate**):

$$\boldsymbol{r} = \{\boldsymbol{r}_1, s_1, \boldsymbol{r}_2, s_2, ...\boldsymbol{r}_N, s_N\} \tag{20.1}$$

Similarly the set of all nuclear coordinates are denoted by $\boldsymbol{R}$**.** Of course, we may switch back (or back-and-forth) the notation in the upcoming chapters.

The Nuclear-Electronic Wavefunction and Hamiltonian

The Schrödinger equation for the full wavefunction (of electrons+nuclei) is of course $H\Psi = E\Psi$,where the full Hamiltonian is

$$H = T_N + H_e \tag{20.2}$$

where the nuclear kinetic energy operator is

$$T_N \equiv \sum_{j=1}^{N_{NUC}} \frac{(-1)}{2M_j} \frac{\partial^2}{\partial \boldsymbol{R}_j^2} \tag{20.3}$$

and we introduced the number of nuclei and M_j is the mass of each nuclei (in atomic units, NOT amu! – i.e. the mass of hydrogen is about 1837 in a.u., etc.). Further, H_e is the electronic Hamiltonian that we dealt with in the previous quarter,

$$H_e = \hat{T} + \hat{V} + \hat{U} = \sum_{i=1}^{N} \frac{(-1)}{2m_e} \frac{\partial^2}{\partial \boldsymbol{r}_j^2} + \hat{V}(\boldsymbol{r}, \boldsymbol{R}) + \hat{U}(\boldsymbol{r}) \tag{20.4}$$

where, to remind you, $\hat{T}, \hat{V}$ and $\hat{U}$ refer, respectively, to the kinetic operator, the electron-nuclei interaction, and the electron-electron repulsion.

Note that H_e depends **parametrically** on $\boldsymbol{R}$. By this we mean that there are no kinetic-energy (derivative) terms w.r.t. the nuclear coordinate $\boldsymbol{R}$ in the electronic Hamiltonian; for a different set of nuclear coordinates the potential will change, but $H_e$will not change the nuclear positions (which is what a derivative w.r.t $\boldsymbol{R}$ does).

We often denote this parametric dependence by writing

$$H_e(\boldsymbol{r};\boldsymbol{R}). \tag{20.5}$$

The full Schrödinger equation

The Schrödinger equation for the electron-nuclear wavefunction reads either as

$$H\Psi_E(\boldsymbol{R},\boldsymbol{r}) = E\Psi_E(\boldsymbol{R},\boldsymbol{r}) \tag{20.6}$$

for time-independent solutions involving the full wavefunctions; or, for time-dependent solutions, replace E by a time-derivative

$$H\Psi(\boldsymbol{R},\boldsymbol{r},t) = i\frac{\partial}{\partial t}\Psi(\boldsymbol{R},\boldsymbol{r},t) \tag{20.7}$$

We'll usually consider the time-indep. version and omit the explicit E subscript, but everything carries to the time-dependent treatment.

Approximate Diabatic Expansion

The "chemically-intuitive" but approximate approach come when we expand the full wavefunction $\Psi(\boldsymbol{r},\boldsymbol{R})$ in terms of electronic states, $\eta_j(\boldsymbol{r})$, that are orthogonal for each electronic position $\int \eta_j^*(\boldsymbol{r};\boldsymbol{R})\eta_l(\boldsymbol{r};\boldsymbol{R})d\boldsymbol{r} = \delta_{jl}$ and are assumed to have a "specific character"

For example, in NaCl, at each nuclear distance (R, distance between the atoms), we can limit ourselves to two such diabatic states; η_a would denote the lowest-energy state where the atoms are neutral, while η_b designate the ionic state. Then[1] the eigenstates of the electronic Hamiltonian would be

$$\Phi_K(\boldsymbol{r};\boldsymbol{R}) = c_{a,K}(\boldsymbol{R})\eta_a(\boldsymbol{r};\boldsymbol{R}) + c_{b,K}(\boldsymbol{R})\eta_b(\boldsymbol{r};\boldsymbol{R}). \quad m = 1,2 \tag{20.8}$$

The Φ_K are **adiabatic nuclear states (K is an index,** has **nothing to do with exchange or kinetic energy!)** that are the eigenstates to the Schrödinger equation,

$$H_e\Phi_K = E_K(\boldsymbol{R})\Phi_K \tag{20.9}$$

---

[1] I use graphics/discussion from http://www.chm.bris.ac.uk/webprojects2002/grant/webcomp/non-crossing.html

while the $c's$ are the "expansion coefficients of the diabatic states in the adiabatic functions." By multiplying with $\eta_a^*(\boldsymbol{r})$ or $\eta_b^*(\boldsymbol{r})$ and integrating w.r.t. $\boldsymbol{r}$ we get a 2-by-2 matrix equation

$$\begin{pmatrix} H_{aa} & H_{ab} \\ H_{ba} & H_{bb} \end{pmatrix} \begin{pmatrix} c_{a,K} \\ c_{b,K} \end{pmatrix} = E_K \begin{pmatrix} c_{a,K} \\ c_{b,K} \end{pmatrix} \tag{20.10}$$

where

$$H_{aa} = \int \eta_a^*(\boldsymbol{r};\boldsymbol{R}) H_e \eta_a(\boldsymbol{r};\boldsymbol{R}) d\boldsymbol{r} \tag{20.11}$$

$$H_{ba} = \int \eta_b^*(\boldsymbol{r};\boldsymbol{R}) H_e \eta_a(\boldsymbol{r};\boldsymbol{R}) d\boldsymbol{r} \tag{20.12}$$

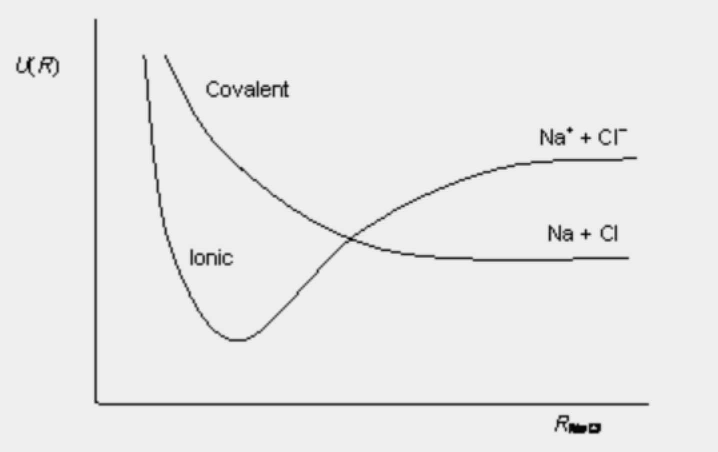
and analogously for $H_{ab} = H_{ba}^*$ and H_{bb}. (In molecular calculations all the states are actually real, but I keep the complex conjugate notation).

The formulae for the eigenstates of 2x2 matrices are well-known (you get them by writing $\det(A - \lambda I) = 0$, where A is any 2x2 matrix and λ an eigenvalue). In this context

$$E_{1,2} = \frac{H_{aa} + H_{bb}}{2} \mp \sqrt{\left(\frac{H_{aa} - H_{bb}}{2}\right)^2 + |H_{ab}|^2}. \tag{20.13}$$

We thus see that to get degeneracy, two conditions have to be fulfilled: the diabatic states need to be degenerate, $H_{aa} = H_{bb}$, and the coupling matrix needs to be vanish $H_{ab} = 0$.

In NaCl, since there's only one important nuclear coordinate (the Na-Cl distance), it will be impossible to fulfill these two conditions. We see in the attached figure (not to scale!) the "diabatic potential surfaces", i.e.:



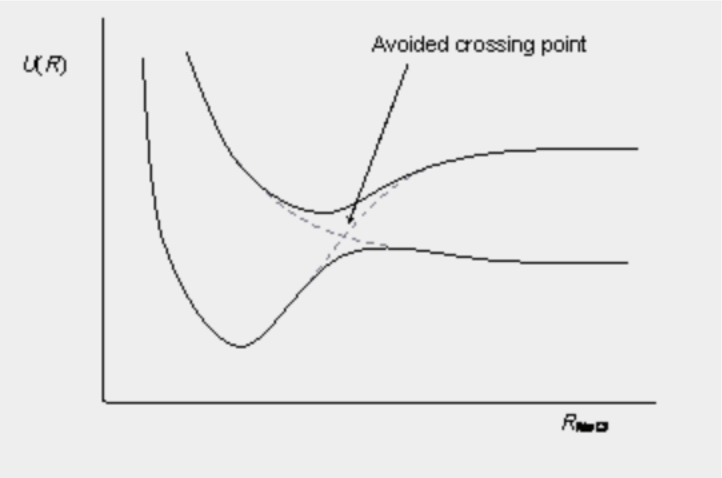
- H_{aa} , the neutrals' ("covalent") potential that rises with decreasing distance due to repulsion,

- and H_{bb}, the ionics' potential, that at large distances is larger than H_{aa} by the difference between the ionization energy of sodium and the electron affinity of chlorine:

$$\begin{aligned} H_{bb}(R\to\infty)-H_{aa}(R\to\infty) &= \left(E\left(\mathrm{Na}^{+}\right)+E\left(\mathrm{Cl}^{-}\right)\right)-\left(E\left(\mathrm{Na}\right)+E\left(\mathrm{Cl}\right)\right) \\ &= \mathrm{I.E}\left(\mathrm{Na}\right)-\mathrm{E.A.}\left(\mathrm{Cl}\right)=5.1\mathrm{eV}\text{-}3.6\mathrm{eV}{=}1.5\mathrm{eV} \end{aligned} \qquad (20.14)$$

  but at short distances H_{bb} will dive below the neutral's I.E. due to the stabilization of the ionic interaction.

The picture below shows the "adiabatic potential surfaces", i.e., $E_1(R)$ and $E_2(R)$ from the equations above:



Note the "avoided crossing" point; here, $H_{aa} = H_{bb}$, so by the formula above the difference between the potential energies is $2 \cdot H_{ab}$, i.e., twice the coupling matrix element.

Conical Intersections

Things get much more interesting when we have more than one distance to vary.

Recall that we will have exact equality of the adiabatic energies if we fulfill two conditions (equal energies and zero coupling matrix element).

Obviously, this is completely analogous to phase transition maps! Recall the map for H_2O, as a function of P and T – the triple point, where the chemical-potential-equality conditions are fulfilled, $\mu_{ice} = \mu_{water}$, and $\mu_{water} = \mu_{vapor}$, is an isolated point.

Further, when we have more degrees of freedom we'll have whole manifolds of distances where the adiabatic eigenvalues are equal!

These points or manifolds where the eigenstates of H_e are degenerate (and often not for symmetry reasons) are called **conical intersections**. They are important for two reasons.

- First, transitions from “excited states” to lower-lying states occur essentially always at the conical intersections or near them.
- Second, and contrary to our intuition, they often have important effects even at much lower energies! This fact, which we’ll explain later, is almost depressing – e.g., if we have a collision where the atoms have thermal energies we may need to model the adiabatic potential at much higher energies (of a few eV) to correctly obtain the transition energies, even though the nuclear wavefunctions of the particles never reach such high energies.

The effects of the conical intersection have to do with phases – when we rotate the nuclei slowly, so we get back to the same state, the electronic eigenstates need not always come back with the same phase!

We saw something similar (and related) when we discussed spin last quarter. Starting with a spin-up wavefunction along the x-axis, where a magnetic field is pointing, we rotated the magnetic field and therefore the axis of the spin slowly in the x-y plane till it pointed back to the x-axis. After rotation we got back the same wavefunction, but it now has an overall minus! I.e., rotation by 360 degrees gave rise to a sign of -1.

For further details on conical intersection, a good resource is Michael Baer’s book, Beyond Born-Oppenheimer: Electronic Nonadiabatic Coupling Terms and Conical Intersections, Wiley (2006).

The Adiabatic Expansion

The need for adiabatic expansion

Now that we understand how important conical intersections are, let’s do a more rigorous job in deriving the appropriate states. You see, the problem is that the states we really know are **only the adiabatic states,** i.e., the eigenstates and eigenvalues of H_e. We don’t actually have rigorously an electronic Hamiltonian in the form of Eq. (20.10)!

Adiabatic expansion: the math

So let’s do the math correctly.

At each $\boldsymbol{R}$, let’s assume that we can calculate the solution of the electronic Schrödinger equation at that $\boldsymbol{R}$. We showed how to do that in the last quarter, at least for the ground state and we’ll talk later about ways to calculate excited states – at least within the extension to DFT called TDDFT, and perhaps with other methods too.

Thus, in the language of the previous part we calculate the solutions to

$$H_e(\boldsymbol{r})\Phi_K(\boldsymbol{r}) = E_K\Phi_K(\boldsymbol{r}) \qquad \text{in our previous notation, with}\boldsymbol{R}\textbf{ assumed implicitly.} \tag{20.15}$$

at a specific set of nuclear coordinates $\mathbf{R}$; we'll now use the more detailed notation:

$$H_e(\boldsymbol{r};\boldsymbol{R})\Phi_K(\boldsymbol{r};\boldsymbol{R}) = E_K(\boldsymbol{R})\Phi_K(\boldsymbol{r},\boldsymbol{R}) \tag{20.16}$$

Here, K is an index to the states – $K = 1$ (or 0, or "G"), is the ground-state, and $K = 2, 3, ...$ are excited states.

These solutions to Eq. (20.16) fulfill orthogonality relations. Because for each $\boldsymbol{R}$ the electronic Hamiltonian H_e is Hermitian, we get of course

$$\int \Phi_K^*(\boldsymbol{r};\boldsymbol{R})\Phi_L(\boldsymbol{r};\boldsymbol{R})d\boldsymbol{r} = \delta_{KL}. \tag{20.17}$$

(We will try to avoid the bra-ket notation so as not to confuse on whether the bra-ket refer to integration over all coordinates or only the electronic ones).

Similarly, the <u>completeness</u> relation for the electronic coordinates is (remember that the nuclear coordinate is just "parametrizing")

$$\sum_K \Phi_K(\boldsymbol{r};\boldsymbol{R})\Phi_K^*(\boldsymbol{r'};\boldsymbol{R}) = \delta(\boldsymbol{r}-\boldsymbol{r'}). \tag{20.18}$$

The full wavefunction for the electrons AND the nuclei can then be expanded in terms of the electronic states :

$$\Psi(\boldsymbol{r},\boldsymbol{R}) = \sum_K \psi_K(\boldsymbol{R})\Phi_K(\boldsymbol{r};\boldsymbol{R}). \tag{20.19}$$

Here, the so called **nuclear wavefunction** $\psi_K(\boldsymbol{R})$ are related of course to the full-wavefunction as

$$\psi_K(\boldsymbol{R}) = \int \Phi_K^*(\boldsymbol{r};\boldsymbol{R})\Psi(\boldsymbol{r};\boldsymbol{R})d\boldsymbol{r} \tag{20.20}$$

(prove this relation based on the eqs. above!).

This is called the adiabatic-states expansion (reason soon to be clarified). There's confusion in the literature – at times the $\Phi_K(\boldsymbol{r};\boldsymbol{R})$ are called adiabatic states; at times the $\psi_K(\boldsymbol{R})$ are called that name. The latter are also called at times **adiabatic nuclear states**, a name we'll use.

Also note that the nuclear states, $\psi_K(\boldsymbol{R})$, depend of course on the total energy E, since the full wavefunction depends on it, but for brevity we do not write them explicitly as $\psi_{K,E}(\boldsymbol{R})$.

The Schrödinger equation for the adiabatic nuclear states

The Schrödinger equation for the full wavefunction, $H\Psi = E\Psi$, becomes

$$\left(\sum_{j=1}^{N_{NUC}} \frac{(-1)}{2M_j}\frac{\partial^2}{\partial \boldsymbol{R}_j^2} + H_e\right)\Psi(\boldsymbol{r},\boldsymbol{R}) = E\Psi(\boldsymbol{r},\boldsymbol{R}) \tag{20.21}$$

i.e.,

$$(T_N + H_e) \sum_K \psi_K(\boldsymbol{R}) \Phi_K(\boldsymbol{r};\boldsymbol{R}) = E \sum_K \psi_K(\boldsymbol{R}) \Phi_K(\boldsymbol{r};\boldsymbol{R}). \tag{20.22}$$

So, since H_e acts only on the electronic coordinates $\boldsymbol{r}$ (we use lower-case k instead of K in the following, sorry for the confusion):

$$\left(\sum_{j=1}^{N_{NUC}} \frac{(-1)}{2M_j} \frac{\partial^2}{\partial \boldsymbol{R}_j^2} \right) \sum_k \psi_k(\boldsymbol{R}) \Phi_k(\boldsymbol{r};\boldsymbol{R}) + \sum_k \psi_k(\boldsymbol{R}) H_e \Phi_k(\boldsymbol{r};\boldsymbol{R}) = E \sum_k \psi_k(\boldsymbol{R}) \Phi_k(\boldsymbol{r};\boldsymbol{R}). \tag{20.23}$$

Now recall that $H_e \Phi_k = E_k \Phi_k$, and the rule for differentiating a product, to get

$$\sum_k \left(\sum_{j=1}^{N_{NUC}} \frac{(-1)}{2M_j} \left(\frac{\partial^2 \psi_k(\boldsymbol{R})}{\partial \boldsymbol{R}_j^2} \Phi_k(\boldsymbol{r};\boldsymbol{R}) + \psi_k(\boldsymbol{R}) \frac{\partial^2 \Phi_k(\boldsymbol{r};\boldsymbol{R})}{\partial \boldsymbol{R}_j^2} + \frac{\partial \psi_k(\boldsymbol{R})}{\partial \boldsymbol{R}_j} \frac{\partial \Phi_k(\boldsymbol{r};\boldsymbol{R})}{\partial \boldsymbol{R}_j} \right) \right)$$
$$= \sum_k (E - E_k)\, \psi_k(\boldsymbol{R}) \Phi_k(\boldsymbol{r};\boldsymbol{R}). \tag{20.24}$$

Finally, dot product with $\Phi_m^*(\boldsymbol{r};\boldsymbol{R})$ and integrate w.r.t. $\boldsymbol{r}$ to get

$$\boxed{T_N \psi_m + \sum_k \left(D_{mk} \psi_k(\boldsymbol{R}) + \sum_{j=1}^{N_{NUC}} \tau_{mk}(j) \frac{\partial \psi_k(\boldsymbol{R})}{\partial \boldsymbol{R}_j} \right) = (E - E_m(\boldsymbol{R}))\, \psi_m(\boldsymbol{R})} \tag{20.25}$$

where we defined defined a matrix D and a set of matrices, $\tau(j)$ such that

$$\tau_{mk}(j) = -\frac{1}{2M_j} \int \Phi_m^*(\boldsymbol{r};\boldsymbol{R}) \frac{\partial \Phi_k(\boldsymbol{r};\boldsymbol{R})}{\partial \boldsymbol{R}_j} d\boldsymbol{r} \tag{20.26}$$

and

$$D_{mk} = \sum_{j=1}^{N_{NUC}} \frac{(-1)}{2M_j} \int \Phi_m^*(\boldsymbol{r};\boldsymbol{R}) \frac{\partial^2 \Phi_k(\boldsymbol{r};\boldsymbol{R})}{\partial \boldsymbol{R}_j^2} d\boldsymbol{r}, \tag{20.27}$$

and from the definition of the kinetic energy,

$$T_N \psi_m(\boldsymbol{R}) \equiv \sum_{j=1}^{N_{NUC}} \frac{(-1)}{2M_j} \frac{\partial^2 \psi_m(\boldsymbol{R})}{\partial \boldsymbol{R}_j^2} \tag{20.28}$$

The boxed equation above is the full Schrödinger equation for the electrons+nuclei. The τ and D matrices are called the **non-adiabatic coupling terms.** In principle they can be deduced from electronic-structure calculations when those yield the adiabatic electronic states, i.e., the Φ_k which are eigenstates of the electronic Hamiltonian H_e

The time-dependent Schrödinger equation

We could do exactly the same steps we did for the time-dependent Schrödinger equation for the nuclei. I.e., replace the energy-dependent Schrödinger equation for the full wavefunction by a time-dependent one

$$i\frac{\partial \Psi}{\partial t} = H\Psi, \tag{20.29}$$

and now label the full-Schroedinger equation as $\Psi(\boldsymbol{r}, \boldsymbol{R},t)$, and expand it in terms of **time-dependent nuclear states**

$$\Psi(\boldsymbol{r},\boldsymbol{R},t) = \sum_k \psi_k(\boldsymbol{R}, t)\Phi_k(\boldsymbol{r}; \boldsymbol{R}), \tag{20.30}$$

and repeating the same steps we did before leads to the boxed-equation above, with E being replaced by $i\frac{\partial}{\partial t}$.

The time-dependent Born-Oppenheimer equation

Since D and τ are proportional to the inverse of the large nuclear masses (large in comparison to the electron's mass), the simplest approximation is to neglect them altogether and end up with a single-level Schrödinger equation

$$T_N\psi_m(\boldsymbol{R}) + E_m(\boldsymbol{R})\psi_m(\boldsymbol{R}) = E\psi_m(\boldsymbol{R}). \tag{20.31}$$

This is the time-independent Born-Oppenheimer equation. It says that the nuclei move on a potential surface , $E_m(\boldsymbol{R})$ which is simply the eigenstate of the electronic Hamiltonian. Similarly we can replace E by a time-derivative and end up with the time-dependent Schrödinger eq.,

$$i\frac{\partial}{\partial t}\psi_m(\boldsymbol{R}, t) = T_N\psi_m(\boldsymbol{R}, t) + E_m(\boldsymbol{R})\psi_m(\boldsymbol{R},t) \tag{20.32}$$

which has the same interpretation, in a time-dependent manner.

Another look at deriving the time-dependent Born-Oppenheimer equation

The Schrödinger equation for the adiabatic states, Eq. (20.32) can be re-derived by writing:

$$i\frac{\partial \Psi(\boldsymbol{r},\boldsymbol{R},t)}{\partial t} = H\Psi(\boldsymbol{r},\boldsymbol{R},t). \tag{20.33}$$

Once we plug in the expansion above we get

$$i\sum_k \frac{\partial \psi_k(\boldsymbol{R}, t)}{\partial t}\Phi_k(\boldsymbol{r}; \boldsymbol{R}) = \sum_k H\psi_k(\boldsymbol{R}, t)\Phi_k(\boldsymbol{r}; \boldsymbol{R}). \tag{20.34}$$

We also recall that, when we neglect non-adiabatic terms, the 0-order Hamiltonian is just the kinetic energy for the nuclei plus the (diagonal) electronic energy for the adiabatic states (the latter contains the kinetic energy for the electrons and their potential); i.e.,

$$H\psi_k(\boldsymbol{R},t)\Phi_k(\boldsymbol{r};\boldsymbol{R}) \simeq \Phi_k(\boldsymbol{r};\boldsymbol{R})\left(T_N + E_k(\boldsymbol{R})\right)\psi_k(\boldsymbol{R},t) \tag{20.35}$$

where the " $\simeq$ " symbol refers to the neglect of the non-adiabatic term, i.e., no gradient w.r.t. the nuclei is taken for the electronic states Φ_k. Defining the nuclear Hamiltonian associated with the k'th potential surface as

$$H_k = T_N + E_k(\boldsymbol{R}) \tag{20.36}$$

gives

$$H\psi_k(\boldsymbol{R},t)\Phi_k(\boldsymbol{r};\boldsymbol{R}) \simeq \Phi_k(\boldsymbol{r};\boldsymbol{R})H_k\psi_k(\boldsymbol{R},t). \tag{20.37}$$

Therefore

$$i\sum_k \frac{\partial\psi_k(\boldsymbol{R},t)}{\partial t}\Phi_k(\boldsymbol{r};\boldsymbol{R}) = \sum_k \Phi_k(\boldsymbol{r};\boldsymbol{R})H_k\psi_k(\boldsymbol{R},t). \tag{20.38}$$

Multiply the equation above times a "bra", $\Phi_m^*(\boldsymbol{r},\boldsymbol{R})$, and integrate w.r.t., $\boldsymbol{r}$, to get

$$i\frac{\partial\psi_m(\boldsymbol{R},t)}{\partial t} = H_m\psi_m(\boldsymbol{R},t), \tag{20.39}$$

where we used $\int \Phi_m^*(\boldsymbol{r};\boldsymbol{R})\Phi_k(\boldsymbol{r};\boldsymbol{R})d\boldsymbol{r} = \delta_{k,m}$, and

Thus, the electronic energy ($E_m(\boldsymbol{R})$, within H_m) plays the role of a potential surface for the nuclei, as we know from first-year chemistry. For example, in a diatomic molecule, often the ground-potential state $E_1(\boldsymbol{R})$ will be bonding, and the first excited surface ($E_2(\boldsymbol{R})$) will be dissociative, see picture.

Time-independent Born-Oppenheimer equation

We wrote the equation above for the time-dependent Schrödinger equation, but it is equally valid for a time-independent picture. In that case, replace $\psi_m(\boldsymbol{R},t)$ by a ro-vibrational state labeled $\chi_{n,m}(\boldsymbol{R})$ where "n" is an index over the rovibrational-state, and "m" is the electronic state. The time-independent Schrödinger equation is then

$$H_m\chi_{n,m}(\boldsymbol{R}) = \varepsilon_{n,m}\chi_{n,m}(\boldsymbol{R}). \tag{20.40}$$

where $\varepsilon_{n,m}$ is the energy of the n'th ro-vibrational state on the m'th potential surface.

Relation between conical intersections the Breakdown of the Born-Oppenheimer equation.

The Born-Oppenheimer eq. is “intuitive” for us chemists. However, it breaks down when the adiabatic nuclear states are degenerate or near degenerate, since then we cannot neglect the “small” non-adiabatic terms. Then the effects of couplings and conical intersections, as explained earlier, kick-in.

Epilogue: Couplings

Here we understood the Born-Oppenheimer (B.O.) approximation and how it is derived from the exact analysis. We’ll use it to:

- First, in the next chapters, understand how light influences motion and dynamics between electronic states
- In the last chapters, if time-permits, we’ll see how the non-B.O. coupling terms we found can be used to study **classical** dynamics of nuclei that will occasionally **hop** between B.O. states due to the coupling terms.

21 Time-dependent perturbation theory: quick derivation of Fermi's Golden Rule

In this chapter we'll do a quick derivation of time-dependent perturbation theory. We'll then re-tackle it in a more elegant fashion in the next chapter, where we'll also talk specifically about electronic absorption, i.e., promotion the electron from one electronic state to another.

A lot of the material in this and the next two chapters is repetitive, which is fine since it is so important that we need to get at it from different angles.

The perturbation

Let's ignore the non-adiabatic couplings, but add to the Hamiltonian the interaction with light, assumed to be at frequency ω.

We assume that **the perturbation is small**. (This is usually valid unless the laser intensity is extremely high, higher than $10^{14}\mathrm{W/cm}^2$.)

That perturbation part can be written, to a good approximation as the dot product of the electric field with the electronic dipole (We'll derive that later too). Thus, we write ($\boldsymbol{r}$ is the electronic coordinate):

$$V(\boldsymbol{r},t) = E_{\text{field}}(t)U(\boldsymbol{r}) \tag{21.1}$$

where we'll explain the terms in sequence. (Note that we don't write the coordinate of the nuclei – that will become relevant in the next chapter.)

$U(\boldsymbol{r})$: Let's assume that the electric field is polarized along the "x" axis, then

$$U(\boldsymbol{r}) = -ex,$$

and therefore the dipole is along x (including the electron's charge). We'll try to keep using the form $U(\boldsymbol{r})$ in this chapter, not just writing "$-ex$". Our results will therefore be general, regardless of the form of U, i.e., for any type of perturbation. But in the next chapter we will be less general and call this perturbation μ, the general symbol for a dipole moment. (And in the chapter after that we'll show that it actually originates more fundamentally from the momentum operator...)

$E_{\text{field}}(t)$: We introduced here the electric field E_{field} (I don't call it E in this chapter so as not to confuse with the symbol for the electronic energies). Then

$$E_{\text{field}}(t) = |E_{\text{field}}|\cos(\omega t - \boldsymbol{k}\cdot\boldsymbol{r} + \phi) = \frac{|E_{\text{field}}|}{2}\left(e^{i(\omega t - \boldsymbol{k}\cdot\boldsymbol{r}+\phi)} + e^{-i(\omega t - \boldsymbol{k}\cdot\boldsymbol{r}+\phi)}\right) \tag{21.2}$$

Here,

- $|E_{\text{field}}|$ is maximum amplitude of the field, and
- ϕ is the phase of the field which in reality it not fixed but varies slowly with time – we'll talk about that phase later. For now we won't put it

into the calculation. The variation of the phase with time is associated with the finite-time coherence of the laser making the electric field

- The phase also has a "spatial part", $\boldsymbol{k} \cdot \boldsymbol{r}$, which does not vary much on the scale of the molecule since the wave-vector $\boldsymbol{k}$**,**is proportional to the inverse of the wavelength of the light – and the latter is thousands of angstrom. So we can ignore it too.

Thus, combining the last equations and the dropping of the $-\boldsymbol{k} \cdot \boldsymbol{r} + \phi$ phase, the interaction is

$$V(\boldsymbol{r},t) = \frac{eE_{\text{field}}}{2} x \left(e^{i\omega t} + e^{-i\omega t}\right) \tag{21.3}$$

It will turn out that **in absorption we can safely retain only the** $e^{-i\omega t}$ **term, which will "kick" from low-energy state to a higher energy one,** and for emission we'll keep the $e^{i\omega t}$. This is called the **rotating wave approximation.** It is valid unless the pulse times are extremely short, i.e., it is valid as long as: $t > \frac{h}{\delta E_{typical}} \sim \frac{h}{2\text{eV}}$ ~2fsec. So any typical laser field with coherence of psec or more will lead to transitions that follow the rotating-wave approximation.

Thus, for absorption, we'll write

$$V(\boldsymbol{r}, t) \simeq bU(\boldsymbol{r})e^{-i\omega t} \tag{21.4}$$

and for emission we'll use $V(\boldsymbol{r},t) \simeq bU(\boldsymbol{r})e^{+i\omega t}$. Here, for simplicity we defined the scaled strength of the electric field:

$$b \equiv \frac{eE_{\text{field}}}{2}. \tag{21.5}$$

Time-dependent perturbation theory: "usual" derivation

Given a set of eigenfunction, $|\chi_j\rangle, j = 1, ..., \infty$, which are eigenfunctions of a "zero-order" static Hamiltonian, H_0, i.e.,

$$H_0|\chi_j\rangle = \varepsilon_j|\chi_j\rangle, \tag{21.6}$$

let's assume then that initially the system is in the n'th state $|\chi_n\rangle$. The Hamiltonian will be the 0th order one plus the perturbation

$$H = H_0 + V(\boldsymbol{r}, t) = H_0 + E_{\text{field}}(t)U(\boldsymbol{r}) \simeq H_0 + bU(\boldsymbol{r})e^{-i\omega t}, \tag{21.7}$$

where in the last equality sign we have put in the rotating wave approximation. We'll use whatever form is more convenient in the equations below.

Let's label as $\psi_n(\boldsymbol{r}, t)$ (or as ket $|\psi_n(t)\rangle$) the **wavefunction that starts at** χ_n **and propagates according to the Schrödinger equation,** i.e.,

$$i\frac{\partial|\psi_n\rangle}{\partial t} = H|\psi_n(t)\rangle, \quad |\psi_n(t = 0)\rangle = |\chi_n\rangle \tag{21.8}$$

and let's expand it in terms of the eigenfunctions of H_0 :

$$|\psi_n(t)\rangle = \sum_j a_{jn}(t)|\chi_j\rangle \tag{21.9}$$

where $a_{jn}(t)$ are time-dependent coefficients which tell us what's the amplitude of the j'th eigenstate of H_0 in true solution of the full Schrödinger equation, $\psi_n(\boldsymbol{r},t)$. Obviously $a_{jn}(t=0)=\delta_{jn}$.

Further, for convenience let's define new coefficients, $c_{jn}(t) = e^{i\varepsilon_j t}a_{jn}(t)$, so therefore $a_{jn}(t) = e^{-i\varepsilon_j t}c_{jn}(t)$, so

$$|\psi_n(t)\rangle = \sum_j e^{-i\varepsilon_j t}c_{jn}(t)|\chi_j\rangle. \tag{21.10}$$

Let's plug it into the Schrödinger equation:

$$\begin{aligned} &i|\frac{\partial\psi_n}{\partial t}\rangle = H|\psi_n(t)\rangle \rightarrow \\ &i\sum_j \frac{\partial\left(e^{-i\varepsilon_j t}c_{jn}(t)\right)}{\partial t}|\chi_j\rangle = \sum_j e^{-i\varepsilon_j t}c_{jn}(t)H|\chi_j\rangle = \sum_j e^{-i\varepsilon_j t}c_{jn}(t)\left(H_0+V(t)\right)|\chi_j\rangle \end{aligned} \tag{21.11}$$

(where $V(t)$ is an abbreviation to $V(\boldsymbol{r},t)$), i.e.,

$$\sum_j e^{-i\varepsilon_j t}\left(\varepsilon_j c_{jn}(t) + i\frac{\partial c_{jn}(t)}{\partial t}\right)|\chi_j\rangle = \sum_j e^{-i\varepsilon_j t}c_{jn}(t)\left(\varepsilon_j + V(t)\right)|\chi_j\rangle. \tag{21.12}$$

Now remove the identical (colored) term from the LHS and RHS, resulting in

$$i\sum_j e^{-i\varepsilon_j t}\frac{\partial c_{jn}(t)}{\partial t}|\chi_j\rangle = \sum_j e^{-i\varepsilon_j t}c_{jn}(t)V(t)|\chi_j\rangle. \tag{21.13}$$

Let's now dot product with the l'th eigenstate of H_0:

$$i\langle\chi_l|\sum_j e^{-i\varepsilon_j t}\frac{\partial c_{jn}(t)}{\partial t}|\chi_j\rangle = \langle\chi_l|\sum_j e^{-i\varepsilon_j t}c_{jn}(t)V(t)|\chi_j\rangle. \tag{21.14}$$

Now in the LHS, due to the overlap $\langle\chi_k|\chi_j\rangle = \delta_{kj}$, only the k'th term contributes so

$$ie^{-i\varepsilon_l t}\frac{\partial c_{ln}(t)}{\partial t} = \sum_j e^{-i\varepsilon_j t}c_{jn}(t)\langle\chi_l|V(t)|\chi_j\rangle \tag{21.15}$$

i.e.,

$$\frac{\partial c_{ln}(t)}{\partial t} = -i\sum_j e^{i(\varepsilon_l-\varepsilon_j)t}c_{jn}(t)\langle\chi_l|V(t)|\chi_j\rangle \tag{21.16}$$

and when we integrate from 0 to t we get

$$c_{ln}(t) - c_{ln}(t=0) = -i\sum_j \int_0^t e^{i(\varepsilon_l - \varepsilon_j)\tau} c_{jn}(\tau)\langle\chi_l|V(\tau)|\chi_j\rangle d\tau. \tag{21.17}$$

i.e.,

$$c_{ln}(t) = \delta_{ln} - i\sum_j \int_0^t e^{i(\varepsilon_l - \varepsilon_j)\tau} c_{jn}(\tau)\langle\chi_l|V(\tau)|\chi_j\rangle d\tau. \tag{21.18}$$

Because of the presence of V in the integral, all the coefficients $c_{jn}(t)$ will be small (1st order) except for the l=n term. I.e., the 0th order approximation will be:

$$c^0_{ln}(t) = \delta_{ln}. \tag{21.19}$$

Then, the 1st order term are obtained on the LHS by plugging into the RHS the zeroth-order term (since the RHS, except for the delta function,has already the 1st order term, “V”):

$$c^1_{ln}(t) = -i\sum_j \int_0^t e^{i(\varepsilon_l - \varepsilon_j)\tau} c^0_{jn}(\tau)\langle\chi_l|V(\tau)|\chi_j\rangle d\tau \tag{21.20}$$

i.e., since $c^0_{jn}(\tau) = \delta_{jn}$, then for any state “$l$” that’s not the initial state,

$$\boxed{c^1_{ln}(t) = -i\int_0^t e^{i(\varepsilon_l - \varepsilon_n)\tau}\langle\chi_l|V(\tau)|\chi_n\rangle d\tau, \quad l \neq n} \tag{21.21}$$

This equation is valid, in perturbation theory, regardless of the form of V (as long as it is small). We’ll omit the “1” superscript henceforth but remember that this equation is only valid to 1st order.

For the type of oscillating perturbation that we usually have, under the rotating wave approximation, we’ll replace V by an oscillating term, $V(\boldsymbol{r},\tau) \simeq bU(\boldsymbol{r})e^{-i\omega\tau}$ resulting in **a separation to the matrix element of the spatial part of the perturbation, times a time-dependent factor.**

$$c_{ln}(t) = -ib\langle\chi_l|U|\chi_n\rangle \int_0^t e^{i(\varepsilon_l - (\varepsilon_n + \omega))\tau} d\tau, \quad l \neq n \tag{21.22}$$

i.e., the **absorption probability after time t,**

$$P_{l\leftarrow n}(t) \equiv |c_{ln}(t)|^2 \tag{21.23}$$

can be written as

$$P_{l\leftarrow n}(t) = b^2|\langle\chi_l|U|\chi_n\rangle|^2 f(t) \quad l \neq n \tag{21.24}$$

where we defined

$$f(t) = \left| \int_0^t e^{i\Delta\tau} d\tau \right|^2, \tag{21.25}$$

where

$$\Delta = \varepsilon_l - (\varepsilon_n + \omega) \tag{21.26}$$

is the deviation from resonance. Obviously, there will be **a significant absorption if the initial energy,** ε_n, **plus the frequency of the perturbing radiation,** ω, **matches or almost matches the final energy,** ε_l, **i.e.** $\Delta \simeq 0$.

Laser Fluctuations

Before proceeding**,** however, we need to incorporate **an important physical factor:** the frequency of the laser undergoes random fluctuations. The simplest way to incorporate these are by multiplying the laser field, $e^{-i\omega\tau}$,by a fluctuating phase, $e^{i\phi(\tau)}$,

$$f(t) = | \int_0^t e^{i\phi(\tau)} e^{i\Delta\tau} d\tau|^2, \qquad l \neq n, \tag{21.27}$$

where the **relative phase will** decay after a while, i.e., **the average of** $e^{i\left(\phi(\tau)-\phi(\tau')\right)}$ will be a **decaying function of the time difference,**

$$\text{average of } e^{i\left(\phi(\tau)-\phi(\tau')\right)} = \exp\left(-\gamma^2(\tau-\tau')^2\right), \tag{21.28}$$

i.e., if $\tau = \tau'$, then the phase cancel perfectly in the exponent, while if $|\tau - \tau'| \gg \gamma^{-1}$ the exponent of the relative phase will cancel. Note that for lasers, the values of γ^{-1} (the coherence time) will typically be quite long compared with the usual dynamics (i.e., more than a picosec).

Transitions with Fluctuations

Let's calculate now the transition probability, i.e., $f(t)$:

$$f(t) = \left(\int_0^t e^{i\phi(\tau)} e^{i\Delta\tau'} d\tau'\right)^* \left(\int_0^t e^{i\phi(\tau)} e^{i\Delta\tau} d\tau\right) = \int_0^t \int_0^t e^{i\Delta(\tau-\tau')} e^{i\left(\phi(\tau)-\phi(\tau')\right)} d\tau\, d\tau'. \tag{21.29}$$

Let's change variables

$$t' = \frac{\tau+\tau'}{2}$$
$$t'' = \tau - \tau'.$$

So:

$$f(t) = \int_0^t \left(\int_{-t'}^{t'} e^{-i\Delta t''} e^{i\left(\phi(\tau)-\phi(\tau')\right)} dt'' \right) dt'. \tag{21.30}$$

Further, let's replace $e^{i\left(\phi(\tau)-\phi(\tau')\right)}$ **by its average,** $e^{-\Gamma^2(\tau-\tau')^2}$, i.e., $e^{-\Gamma^2 t''^2}$!!

> Note: the reason for this replacement is that we do an additional integral, over t'. Physically, what we are saying is that if our total time t, of, say, 1000 psec, and we consider two times, τ and τ' that are separated by, say, 20 psec, then we can get many such time – e.g., $\tau = 110$ and $\tau' = 130$; or $\tau = 250$ and $\tau' = 270$ (all in psec.), etc. .; so the exponential of the phase difference, $e^{i\left(\phi(\tau)-\phi(\tau')\right)}$,will effectively average and can be replaced by its average.

So then:

$$f(t) = \int_0^t \left(\int_{-t'}^{t'} e^{-i\Delta t'} e^{-\gamma^2 t'^2} dt'' \right) dt'. \tag{21.31}$$

Now let's consider very long times, t, so that most of the inner integral (over t') would be associated with long times (much longer than γ^{-1}); therefore, we can **replace the inner limits by** $-\infty$ **to** ∞, getting

$$f(t) = \int_0^t \left(\int_{-\infty}^{\infty} e^{-i\Delta t'} e^{-\gamma^2 t'^2} dt'' \right) dt' = \left(\int_0^t dt' \right) \cdot \left(\int_{-\infty}^{\infty} e^{-i\Delta t'} e^{-\gamma^2 t'^2} dt'' \right) = t \left(\int_{-\infty}^{\infty} e^{-i\Delta t'} e^{-\gamma^2 t'^2} dt'' \right). \tag{21.32}$$

Let's do the Gaussian integral:

$$f(t) = t \cdot \frac{\sqrt{\pi} e^{-\frac{\Delta^2}{4\gamma^2}}}{\gamma} \tag{21.33}$$

The Gaussian factor looks strange; but in fact, **in the limit that** γ **is tiny** (long coherence times – $\underline{\text{but still the coherence time needs to be much smaller than the total time, i.e., } \gamma^{-1} \ll t}$)**, the Gaussian factor is essentially a Dirac-delta function:**

$$\sqrt{\pi} \frac{e^{-\frac{\Delta^2}{4\gamma^2}}}{\gamma} \simeq 2\pi\delta(\Delta) \tag{21.34}$$

Proof: the LHS is obviously very narrow if γ is small, and looks "pointed up"; so we just need to prove that its integral is "2π"

$$\sqrt{\pi} \int_{-\infty}^{\infty} \frac{e^{-\frac{\Delta^2}{4\gamma^2}}}{\gamma} d\Delta = 2\pi \int_{-\infty}^{\infty} \frac{e^{-z^2}}{\sqrt{\pi}} dz = 2\pi \quad \left(z \equiv \frac{\Delta}{2\gamma} \right). \quad Q.E.D.! \tag{21.35}$$

Transition Rate: Fermi's Golden rule (state-to-state)

Thus, collecting from the definition of $f(t)$ we get

$$P_{l \leftarrow n}(t) = b^2 |\langle \chi_l | U | \chi_n \rangle|^2 \cdot t \cdot 2\pi\delta\left(\varepsilon_l - (\omega + \varepsilon_n)\right) \tag{21.36}$$

i.e., **the transition rate, i.e., probability PER TIME to transfer from state "n" to another, "l",**

$$\Gamma_{l\leftarrow n} \equiv \frac{1}{t} P_{l\leftarrow n} \tag{21.37}$$

is

$$\boxed{\Gamma_{l\leftarrow n} = |b^2| \langle \chi_l | U | \chi_n \rangle |^2 \cdot 2\pi \delta \left(\varepsilon_l - (\omega + \varepsilon_n) \right)} \tag{21.38}$$

This is **Fermi's Golden rule** (derived by Heisenberg, not Fermi...). In this form it is valid for state-to-state transitions.

Summing over final states: Fermi's Golden rule (total transfer rate).

Now we want to **move to summing over final states,** for total absorption rates at a given photon frequency ω.

Thus, define a **frequency-dependent absorption rate:**

$$\Gamma_n \equiv \sum_l \Gamma_{l\leftarrow n} = 2\pi \sum_l |b^2| \langle \chi_l | U | \chi_n \rangle |^2 \delta \left(\varepsilon_l - (\omega + \varepsilon_n) \right) \tag{21.39}$$

Interpreting Fermi's Golden Rule with Density of States

What do we mean by "delta" function here?

For a **non-disosociated** small molecule with few states the delta function just means a sum of isolated transitions. (see class for figure)

For a big molecule (or even a small **dissociating molecule**), the "smeared" delta functions overlap. To understand what happens, we'll need to replace the matrix element $|\langle \chi_{\nu',m} | \mu_{mg} | \chi_{\nu,g} \rangle|^2$ by some kind of an averaged transition rate. To do it properly we first need to note that

$$\begin{aligned} \Gamma_n &= 2\pi \sum_l |b^2| \langle \chi_l | U | \chi_n \rangle \delta \left(\varepsilon_l - (\omega + \varepsilon_n) \right) \langle \chi_n | U | \chi_l \rangle \\ &= 2\pi \sum_l |b^2| \langle \chi_l | U | \chi_n \rangle \langle \chi_n | U \delta \left(\varepsilon_l - (\omega + \varepsilon_n) \right) | \chi_l \rangle \\ &= 2\pi \sum_l |b^2| \langle \chi_l | U | \chi_n \rangle \langle \chi_n | U \delta \left(H_0 - (\omega + \varepsilon_n) \right) | \chi_l \rangle \end{aligned} \tag{21.40}$$

i.e., since the sum over a complete number of states is a trace we write

$$\Gamma_n = 2\pi |b|^2 \mathrm{Tr} \left(U | \chi_n \rangle \langle \chi_n | U \delta \left(H_0 - (\omega + \varepsilon_n) \right) \right) \tag{21.41}$$

A side note: Recall

$$\mathrm{Tr} \left(A | \chi_n \rangle \langle \chi_n | \right) = \langle \chi_n | A | \chi_n \rangle$$

(to prove, consider summing over the orthogonal basis we have, where χ_n is one of the basis-members). In our case then:

$$\Gamma_n = 2\pi|b|^2\langle\chi_n|U\delta\left(H_0-(\omega+\varepsilon_n)\right)U|\chi_n\rangle \tag{21.42}$$

This equation will be **very important later when we view it in the time picture.**

End of side-note.

Back to Eq. (21.41). We'll end then at:

$$\boxed{\Gamma_n \equiv 2\pi|b|^2\bar{U}^2\rho(E),} \tag{21.43}$$

where

$$E = \omega + \varepsilon_n$$

is the final energy (=ω+initial energy) , and we defined the **averaged squared-transition matrix**

$$\bar{U}^2 \equiv \bar{U}^2(E) = \frac{\langle\chi_n|U\delta\left(H_0-E\right)U|\chi_n\rangle}{\rho(E)}, \tag{21.44}$$

and the **density of states that the system transfers to:**

$$\rho(E) \equiv \mathrm{Tr}\left(\delta\left(H_0-E\right)\right)$$

A note on the density-of-states:

Let's see why we call it that name. For a given zero-order static Hamiltonian, H_0 define first the **cumulative density of states:**

$$N(E) = \textbf{number of states with energy below } E \quad = \sum_j \theta(E-\varepsilon_j) \tag{21.45}$$

where the θ function is defined as:

$$\theta(x) = \begin{cases} 1 & x<0 \\ 0 & x>0 \end{cases} \tag{21.46}$$

For example, for the electronic Hamiltonian of Hydrogen, $N(E)$ will be 1 between the ground state and the n=2 states i.e.,

$$N(E)=1 \text{ for } \quad \text{-13.6eV} < E < -\frac{13.6\text{eV}}{2^2} = -3.4\text{eV},$$

and similarly $N(E)=5$ between the n=2 and n=3 states,

$$N(E) = 5 \text{ for } \quad \text{-3.4eV} < E < -\frac{13.6\text{eV}}{3^2} = -1.5\text{eV},$$

etc. See plot in class.

We will rewrite

$$N(E) = \sum_j \theta(E-\varepsilon_j) = \sum_j \langle\chi_j|\theta(E-\varepsilon_j)|\chi_j\rangle = \sum_j \langle\chi_j|\theta(E-H_0)|\chi_j\rangle = \mathrm{Tr}\left(\theta(E-H)\right).$$

The density of states is defined as the **derivative of $N(E)$ w.r.t. E:**

$$\rho(E) = \frac{dN(E)}{dE}, \tag{21.47}$$

i.e., the number of states per unit energy.

We need to recall now a mathematical fact,

$$\frac{d\theta(x)}{dx} = \delta(x). \tag{21.48}$$

Proof: both the LHS and RHS are zero when $x \neq 0$. So we just need to prove that their integral is the same, i.e., that the integral of the LHS is 1:

$$\int_{-\infty}^{\infty} \frac{d\theta(x)}{dx} dx = \theta(x=+\infty) - \theta(x=-\infty) = 1 - 0 = 1, \quad Q.E.D. \tag{21.49}$$

so:

$$\rho(E) = \frac{dN(E)}{dE} = \frac{d}{dE}\mathrm{Tr}\left(\theta(E-H)\right) = \mathrm{Tr}\left(\frac{d}{dE}\theta(E-H)\right) = \mathrm{Tr}\left(\delta(E-H)\right). \quad Q.E.D. \tag{21.50}$$

This concludes the main Fermi Golden rule, except that we'll need several addenda. The first is: Partial absorption Cross-Sections. More specifically, so far we considered two extremes:

- Starting at one state and ending up at one other state
- Starting at one state and being absorbed to any other state.

There are important in-between cases. Primarily:

- Starting in one electronic state and asking what's the absorption probability from one state to **a specific group of states.**

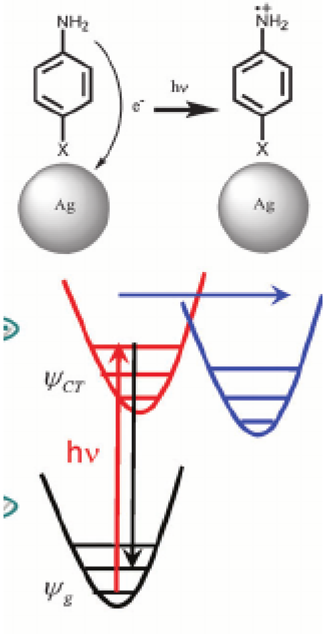
The primary example is electronic spectroscopy. We usually start from a state (or many state) on the ground-electronic state and ask what's the probability to **transfer to another, specific, electronic state** (rather than others).

The formulae are very similar then; basically we now have a formula that looks like $\Gamma \equiv 2\pi \bar{U}^2 \rho(E)$,but now the density of states refers to **only states in the specific desired electronic state** (and all associated ro-vibrational level), and the same thing for the averaging.

We'll see more details in the next chapter.

Transitions without an electric field.

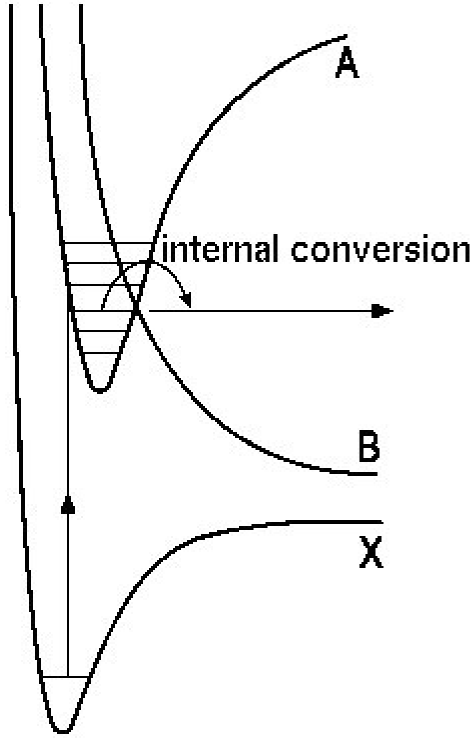
Finally, once we're familiar with all the math, we need to consider the case of **transfer without an electric field.** The primary example is **starting in one electronic state, and then moving to another electronic state.** For example, if due to sun-light, lasers, or collisions, an electron is excited and is on one electronic state, it could move to another state.



Molecule-to-Metal CT

An example is the figure above. If a laser already excited the molecule from ground to an excited state (**in red in the figure**), it will transfer to **a charge transfer state with an electron on the gold and the molecule positively charged (the state in blue).**

Another example is predissociation. Again, a molecule is excited to an electronic state ("A") and then the vibrational state are excited to the dissociating ro-vibrational states on "B"



Luckily, the **same methodology works. The only difference is that** instead of the photon-associated disturbance of $V(\boldsymbol{r},t) \simeq bU(\boldsymbol{r})e^{-i\omega t}$,we will have a **coupling matrix element,** which in the language of the previous chapter we label as, for example per the figure above, as : "$H_{AB}(\boldsymbol{R})$" for transfer from an initial electronic state "A" to a final electronic state. "B"

And the density of states will be **not the total density of states,** but the density of states on the **final electronic state.**

Finally, let's ask a question:

Why is dephasing necessary and shift in absorption type

A qualitatively interesting issue is how the transfer rate changes from very short times to the more usual long-time.

I remind you that we saw

$$c_{ln}(t) = -i\langle\chi_l|U|\chi_n\rangle \int_0^t e^{i(\varepsilon_l - \varepsilon_n - \omega)\tau} d\tau \qquad l \neq n.$$

Therefore, at very short times (shorter than $1/\Delta \equiv 1/\left(\varepsilon_l - \varepsilon_n - \omega\right)$) the integrand will be fairly constant, so

$$c_{ln}(t) \sim -i\langle\chi_l|U|\chi_n\rangle \int_0^t d\tau = -i\langle\chi_l|U|\chi_n\rangle t \quad \text{(for } t \text{ small)},\ l \neq n,. \tag{21.51}$$

Therefore, for short times the **amplitude of transfer will rise as** t^2 :

$$P_{ln}(t) = |c_{ln}(t)|^2 \propto t^2 \tag{21.52}$$

while, as we saw, at long time $P_{ln}(t) \propto t$ (since the rate is fixed).

Thus, **at short time the transfer is coherent and the probability of transfer rises rapidly (like** t^2**),but at long times the transfer is "incoherent" and rises much slower, linearly in time.**

The reason of course is the incoherent dephasing – either due to the laser, and/or due to other degrees of freedom, i.e., the fluctuating environment.

22 Time-dependent perturbation theory: correlation function picture.

We will next use a very instructive **correlation function** approach to understand **electronic** spectroscopy and spectroscopy in general.

Consider the expression we studied in the fairly-dry Fermi's Golden rule description. We can rewrite this expression in a much livelier time-picture. We will do it in two ways in the chapter, and then specialize to electronic spectroscopy.

We'll use a lot of similar techniques to the previous chapter, so when needed we can skip derivations.

Quick derivation of correlation function-based rate from Fermi-Golden rule

The simplest derivation starts with the relation

$$\delta(x) = \frac{1}{2\pi}\int e^{-itx}dt \tag{22.1}$$

(we didn't prove it, but it make sense and is needed for Fourier transforms). Integrals are generally from $-\infty$ to ∞ unless specified differently.

Insert this relation in Eq. (21.42), replacing "x" by $\delta\left(H_0 - (\omega+\varepsilon_n)\right)$, to get

$$\delta\left(H_0 - (\omega+\varepsilon_n)\right) = \frac{1}{2\pi}\int e^{-i(H_0-(\omega+\varepsilon_n))t}dt \tag{22.2}$$

So the expectation value in Eq. (21.42) becomes

$$\Gamma_n = |b|^2\int \langle\chi_n|Ue^{-i(H_0-(\omega+\varepsilon_n))t}U|\chi_n\rangle dt \tag{22.3}$$

i.e.,

$$\boxed{\Gamma_n = |b|^2\int_{-\infty}^{\infty} C(t)e^{i(\omega+\varepsilon_n)t}dt,} \tag{22.4}$$

where we defined the **correlation function**

$$C(t) \equiv \langle\chi_n|Ue^{-iH_0t}U|\chi_n\rangle = \langle g_n|e^{-iH_0t}|g_n\rangle \tag{22.5}$$

where we defined a new state

$$|g_n\rangle \equiv U|\chi_n\rangle. \tag{22.6}$$

Further, define

$$|g_n(t)\rangle = e^{-iH_0t}|g_n\rangle. \tag{22.7}$$

we see that

$$\boxed{C(t) = \langle g_n|g_n(t)\rangle.} \tag{22.8}$$

Thus, we see a physically interesting picture: we start with the initial state, χ_n, then convert it to a new state, $g_n(\boldsymbol{r}) = U(\boldsymbol{r})\chi_n(\boldsymbol{r})$,and then **propagate the new-state with time (using the unperturbed Hamiltonian)**. The **correlation function of this new function at t with itself at the start time**, is then Fourier transformed to give the absorption frequency as a function of frequency.

Two things remain:

- The derivation above is not elegant. The Fermi Golden rule starts with the time-picture and then goes an energy picture. Then we go back to time. There should be a more elegant way that talks only about time. We will do this derivation next
- Then, we'll exemplify (with a few modification) the correlation function picture for **electronic spectroscopy.**

Absorption from time-dependent correlation functions: a time-based derivation.

Recall that we try to solve the Schrödinger equation

$$i\frac{\partial|\psi_n\rangle}{\partial t} = H(t)|\psi_n(t)\rangle, \qquad |\psi_n(t=0)\rangle = |\chi_n\rangle.$$

Write it explicitly:

$$i\frac{\partial|\psi_n\rangle}{\partial t} = (H_0 + V(t))|\psi_n(t)\rangle, \tag{22.9}$$

where recall that $V(t)$ stands for $V(\boldsymbol{r}, t)$.

Claim: we can write the <u>**exact**</u> solution to the Schrödinger equation as

$$\boxed{|\psi_n(t)\rangle = e^{-iH_0 t}|\chi_n\rangle - i\int_0^t e^{-iH_0(t-\tau)}V(\tau)|\psi_n(\tau)\rangle d\tau.} \tag{22.10}$$

Proof: to prove we'll need to differentiate the form on the RHS, which has an integral.

> Recall that 1st-year calculus tells us that when we differentiate an integral that depends on time, we get two terms – the integrand at the upper limit, and an integral where you differentiate the integrand with respect to time, i.e.,
>
> $$\frac{d}{dt}\int_0^t f(t,\tau)d\tau = f(t,\tau = t) + \int_0^t \frac{\partial f(t,\tau)}{\partial t}d\tau.$$

Therefore

$$i\frac{\partial|\psi_n\rangle}{\partial t} = i(-i)H_0 e^{-iH_0 t}|\chi_n\rangle - i^2 e^{-iH_0(t-t)}V(t)|\psi_n(t)\rangle - i^2\int_0^t \left(\frac{\partial}{\partial t}e^{-iH_0(t-\tau)}\right)V(\tau)|\psi_n(\tau)\rangle d\tau$$
$$= H_0 e^{-iH_0 t}|\chi_n\rangle + V(t)|\psi_n(t)\rangle - iH_0\int_0^t e^{-iH_0(t-\tau)}V(\tau)|\psi_n(\tau)\rangle d\tau$$
$$= H_0\left(e^{-iH_0 t}|\chi_n\rangle - i\int_0^t e^{-iH_0(t-\tau)}V(\tau)|\psi_n(\tau)\rangle d\tau\right) + V(t)|\psi_n(t)\rangle. \quad (22.11)$$

But the "magenta" colored part is simply our ansatz for the wavefunction, so we proved that

$$i\frac{\partial|\psi_n(t)\rangle}{\partial t} = H_0|\psi_n(t)\rangle + V(t)\frac{\partial|\psi_n\rangle}{\partial t} = H(t)|\psi_n(t)\rangle, \qquad Q.E.D. \quad (22.12)$$

Perturbation expansion

We'll write

$$|\psi_n(t)\rangle = |\psi_n^0(t)\rangle + |\psi_n^1(t)\rangle \quad (22.13)$$

The boxed-equation above is a good starting point for perturbation. Thus, the 0th order expression is

$$|\psi_n^0(t)\rangle = e^{-iH_0 t}|\chi_n\rangle \quad (22.14)$$

while in the 1st-order expression, we replace the w.f. in $\int_0^t e^{-iH_0(t-\tau)}V(\tau)|\psi_n(\tau)\rangle d\tau$ by $|\psi_n^0(\tau)\rangle$, i.e.,

$$|\psi_n^1(t)\rangle = -i\int_0^t e^{-iH_0(t-\tau)}V(\tau)|\psi_n^0(\tau)\rangle d\tau \quad (22.15)$$

so combining the two equations above gives:

$$|\psi_n^1(t)\rangle = -i\int_0^t e^{-iH_0(t-\tau)}V(\tau)e^{-iH_0\tau}|\chi_n\rangle d\tau. \quad (22.16)$$

Now **let's recall that we assume that the initial state is an eigenstate of the zero-order-Hamiltonian,** i.e., $H_0|\chi_n\rangle = \varepsilon_n|\chi_n\rangle$. Therefore, we can simplify

$$|\psi_n^1(t)\rangle = -i\int_0^t e^{-iH_0(t-\tau)}V(\tau)e^{-i\varepsilon_n\tau}|\chi_n\rangle d\tau = e^{-iH_0 t}\int_0^t e^{i(H_0-\varepsilon_n)\tau}V(\tau)|\chi_n\rangle d\tau \quad (22.17)$$

now recall that we'll use the rotating wave approximation $V(\boldsymbol{r},t) = be^{-i\omega t}U(\boldsymbol{r})$,to give

$$|\psi_n^1(t)\rangle = e^{-iH_0t}|\eta_n\rangle, \tag{22.18}$$

where we introduced a new function

$$|\eta_n\rangle = -ib\int_0^t e^{i(H_0-\omega-\varepsilon_n)\tau}U|\chi_n\rangle d\tau = -ib\int_0^t e^{i(H_0-\omega-\varepsilon_n)\tau}|g_n\rangle d\tau, \tag{22.19}$$

where recall that $g_n \equiv U\chi_n$

Absorption probability

Let's assume that the perturbation has no diagonal matrix element, i.e., $\langle\chi_n|U|\chi_n\rangle = 0$. Then, $|\psi_n^1\rangle$,the first order perturbation, is **orthogonal** to the ground state

Proof:

$$\langle\chi_n|\psi_n^1\rangle = -ib\langle\chi_n|e^{-iH_0t}\int_0^t e^{i(H_0-\varepsilon_n-\omega)\tau}U|\chi_n\rangle d\tau =$$
$$\langle\chi_n|e^{-i\varepsilon_nt}\int_0^t e^{i(\varepsilon_n-\varepsilon_n-\omega)\tau}U|\chi_n\rangle d\tau = e^{-i\varepsilon_nt}\int_0^t e^{-i\omega\tau}\langle\chi_n|U|\chi_n\rangle d\tau = 0 \tag{22.20}$$

Therefore, the **absorption probability** is simply the squared amplitude of the non-ψ^0 part, i.e., of ψ^1 :

$$P_n(t) = \langle\psi_n^1|\psi_n^1\rangle = \langle\eta_n|\eta_n\rangle =$$
$$|b|^2\left(\int_0^t e^{i(H_0-\omega-\varepsilon_n)\tau'}U|\chi_n\rangle d\tau'\right)^\dagger\left(\int_0^t e^{i(H_0-\omega-\varepsilon_n)\tau}U|\chi_n\rangle d\tau\right) =$$
$$|b|^2\int_0^t\int_0^t\langle\chi_n|Ue^{-i(H_0-\omega-\varepsilon_n)\tau'}e^{i(H_0-\omega-\varepsilon_n)\tau}U|\chi_n\rangle d\tau' d\tau \tag{22.21}$$

i.e.,

$$P_n(t) = |b|^2\int_0^t\int_0^t e^{-i(\omega+\varepsilon_n)(\tau-\tau')}C(\tau-\tau')d\tau' d\tau \tag{22.22}$$

This is almost the form we need. Recalling how we broke earlier the double time integral to a factor of "t" times an integral over the time difference, we do the same here , i.e., again write

$$t' = \frac{\tau+\tau'}{2}$$
$$t'' = \tau-\tau',$$

so

$$P_n(t) = |b|^2 \int_0^t \left(\int_{-t'}^{t'} C(t")dt'' \right) dt', \tag{22.23}$$

and again extending the inner limits to ±∞ and dividing by t we get

$$\Gamma_n = \frac{P_n(t)}{t} = |b|^2 \left(\int_{-\infty}^{\infty} e^{-i(\omega+\varepsilon_n)t''} C(t")dt'' \right), \tag{22.24}$$

so changing back the variables $t'' \to t$ gives

$$\Gamma_n = |b|^2 \left(\int_{-\infty}^{\infty} e^{-i(\omega+\varepsilon_n)t} C(t)dt \right). \tag{22.25}$$

This is exactly the equation we had before!

> Actually, there's a cheating in the argument that C(t) decays if the transition is to an electronic state with very few ro-vibrational states, since then there could be infinitely many recurrences in $C(t)$, see figures in Chapter 3.
>
> But even if that's the case, the laser used to excite the system will generally have fluctuations which will cause the correlation function to damp – i.e., if the coherence time of the laser is, say, 1ns, then the "effective" correlation function will decay within that time.
>
> So the more precise statement we should make is that we are **looking at times that are long compared with either the decay time of the correlation function, or longer than the coherence time of the laser or the environment** (if we're looking at spectroscopy in liquids, where there are many collisions)**.**

Example for using time-dependent correlation function: electronic spectroscopy.

The derivation of absorption from correlation function is very similar in the case where nuclei (coordinate $\boldsymbol{R}$) and electronic states are considered.

Specifically, in a previous chapter, we converted the Schrödinger equation for the nuclei and molecules (ignoring non-adiabatic effects!) to an equation that involves electronic-potential surfaces, i.e., Eq. (20.39). Let's follow the same steps, but now add the potential to the Hamiltonian. I.e., modify Eq. (20.38) to read:

$$i \sum_K \frac{\partial \psi_K(\boldsymbol{R}, t)}{\partial t} \Phi_K(\boldsymbol{r}; \boldsymbol{R}) = \sum_K \Phi_K(\boldsymbol{r}; \boldsymbol{R}) \left(H_K + V(\boldsymbol{R}, \boldsymbol{r}, t)\right) \psi_K(\boldsymbol{R}, t). \tag{22.26}$$

and again multiplying the equation above times a "bra", $\Phi_m^*(\boldsymbol{r}, \boldsymbol{R})$, and integrating w.r.t., $\boldsymbol{r}$, to get

$$i\frac{\partial\psi_K(\boldsymbol{R},t)}{\partial t} = H_K\psi_K(\boldsymbol{R},t) + \sum_K V_{K,L}(\boldsymbol{R},t)\psi_L(\boldsymbol{R},t) \tag{22.27}$$

where we defined the matrix element of the perturbation, which generally is

$$V_{K,L}(\boldsymbol{R},t) = \int \Phi_K^*(\boldsymbol{r};\boldsymbol{R})V(\boldsymbol{r},\boldsymbol{R},t)\Phi_L(\boldsymbol{r};\boldsymbol{R})d\boldsymbol{r} \tag{22.28}$$

and in a bra-ket notation

$$\boxed{i\frac{\partial|\psi_K(t)\rangle}{\partial t} = H_K|\psi_K(t)\rangle + \sum_L V_{KL}(t)|\psi_L(t)\rangle} \tag{22.29}$$

In the particular case had above (rotating wave approximation, sinusoidal interaction, field polarized along x) becomes:

$$V_{KL}(\boldsymbol{R},t) = be^{-i\omega t}\mu_{K,L}(\boldsymbol{R}), \tag{22.30}$$

where we defined the dipole-matrix (or more properly the matrix of the dipole along x):

$$\mu_{K,L}(\boldsymbol{R}) = \int \Phi_K^*(\boldsymbol{r};\boldsymbol{R})\, x\, \Phi_L(\boldsymbol{r};\boldsymbol{R})d\boldsymbol{r}. \tag{22.31}$$

Note that in the bra-ket notation we don't write the **R-**dependence of $V_{K,L}(\boldsymbol{R},t)$ but is still there.

Analogies with the earlier derivation of the correlation function

Consider the earlier derivation, which was more general, the Schrödinger equation was then $i\frac{\partial|\psi(t)\rangle}{\partial t} = H_0|\psi(t)\rangle + V(t)|\psi(t)\rangle$, and in the rotating wave approximation we wrote $V(\boldsymbol{r},t) = be^{-i\omega t}U(\boldsymbol{r})$. By inspection, therefore, when we consider electronic states:

- The role of the "zero" order Hamiltonian H_0 is taken over by H_F, the Hamiltonian for the nuclei on the "F"'th electronic state. (F stands for final state, i.e., we are asking how much of the initial wavefunction (on electronic state G, for "ground") ended up on state F.
- The role of the initial state $\chi_n(\boldsymbol{r})$ is taken over by $\chi_{\nu,G}(\boldsymbol{R})$,the ν'th ro-vibrational state on the starting electronic state, G.
- The role of the dipole moment operator $U(\boldsymbol{r})$ is taken over by the dipole-matrix element, $\mu_{F,G}(\boldsymbol{R})$ between the initial "G" and final "F"electronic states.

Then, the correlation function expression for the **absorption of radiation from an initial state $\chi_{\nu,G}(\boldsymbol{R})$ on the G'th electronic state to ANY ro-vibrational state on the final electronic state** F is:

$$C(t) = \langle g|e^{-iH_Ft}|g\rangle = \langle g|g(t)\rangle$$

where now

$$g(\boldsymbol{R}) = \mu_{FG}(\boldsymbol{R})\chi_{\nu,G}(\boldsymbol{R}) \tag{22.32}$$

is the product of the dipole from G to the final state F, times the initial wavefunction. Further, as before,

$$|g(t)\rangle = e^{-iH_F t}|g\rangle \tag{22.33}$$

And finally, from the correlation function we get, as before, the absorption rate to the final (Fth) electronic state,

$$\Gamma_{F\leftarrow(\nu,G)}(t) = \int C(t)e^{i(\omega+\varepsilon_{\nu,G})t} \tag{22.34}$$

Also, below we **usually** (not always!) **abbreviate $\mu_{FG}(\boldsymbol{R})$ as $\mu(\boldsymbol{R})$ to avoid plethora of indices. But recall that it will refer to transitions between two specific states.**

Interpretation

We are finally ready for interpretation.

Electronic absorption can be interpreted as

- We "kick" the original wavefunction from the ground to the excited electronic state; the "kicked" function,g, is, in **R-**space, simply the product of the dipole matrix element and the initial ro-vibrational state on the ground electronic state.
- We let this "kicked" function move under the influence of the electronic Hamiltonian for the state we kicked to, H_F, making $|g(t)\rangle$. We calculate the overlap of the the "moved" function with the initial kicked function, $g(\boldsymbol{R})$
- We then Fourier transform the correlation function at the energy of the initial state+photon. Any peak denotes an absorption at that frequency, i.e., indicates that there's a final ro-vibrational state with an energy that equals $\omega + \varepsilon_{\nu,G}$.

The following figures show an example (a diatom, with one important nuclear coordinate) of a propagating wavepacket ("$g(\boldsymbol{R},t)$") which is the result of excitation from the ground-states to a: bound electronic state; and a dissociating electronic state (photo-dissociation), and then its evolution in time.

See the following figures from Schatz and Ratner:

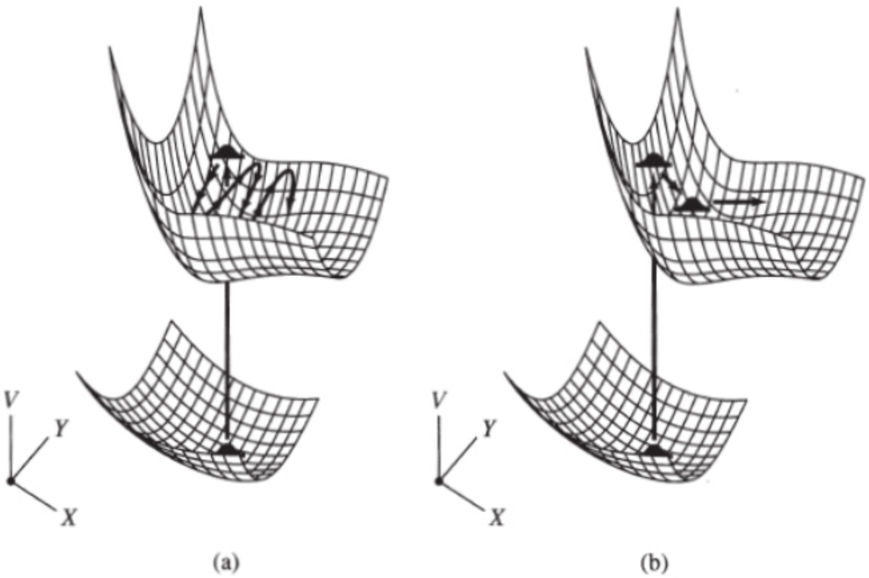


Figure 9.1. Schematic of photoabsorption between two Born–Oppenheimer potential surfaces. On the left, the excited state Franck–Condon wavepacket $\Phi_f(t)$ vibrates many times before dissociating, yielding oscillatory structure in the overlap. On the right, the initial Franck–Condon wavepacket is smoothly accelerated to the right, dissociating without any excited-state vibration. [With permission from E. J. Heller, *Acct. Chem. Res.* 14, 368 (1981), Fig. 1.]

and the associated correlation function:

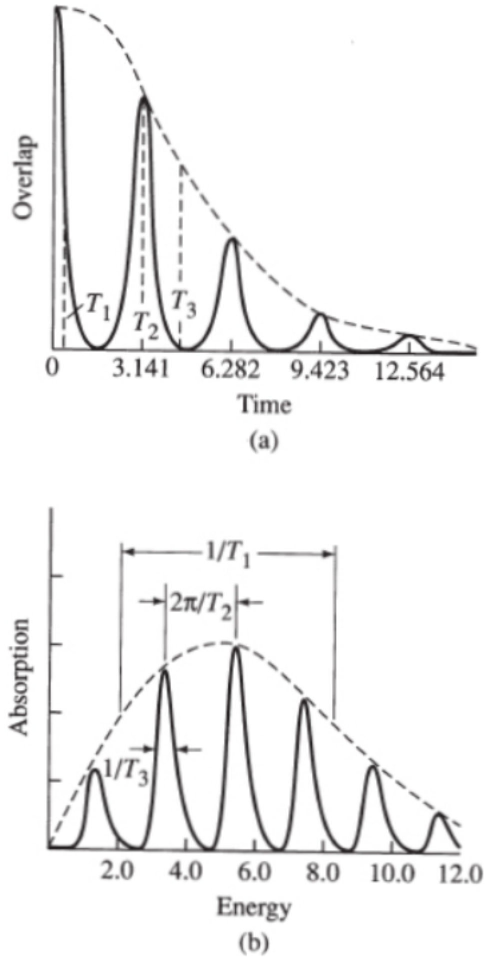


Figure 9.2. Correlation function $\langle \Phi_i(t) | \Phi_f(t + \tau) \rangle$ for the wavepackets shown in Fig. 9.1 (a) and the resulting absorption spectrum (b). The times T_1, T_2, and T_3 arise physically from initial motion of Φ_f on the excited surface, recurrent overlaps of Φ_f with Φ_i, and eventual photodissociation of the molecule. [With permission from E. J. Heller, *Acct. Chem. Res.* 14, 368 (1981), Fig. 3.]

In the bound-excitation case, the wavefunction oscillates, and the absorption lines will then be sharp (widened by the laser fluctuations or by collisions with the environment.

For photo-dissociation, however, the wavepacket immediately "dissipates" so it looks like a short blob in time, which corresponds to a wide-absorption peak.

Emission

There are two types of emission: Stimulated and Spontaneous.

Stimulated Emission

Is exactly like absorption, we just reverse in Eq. (21.26) the sign of $\omega \to -\omega$, since the final-state energy will be lower than the initial one. This can be traced back to Eq. (21.3) where there were two terms, $e^{i\omega t}$ and $e^{-i\omega t}$,where due to the rotating-wave approximation we retain only the $e^{-i\omega t}$; while for emission we need to keep $e^{i\omega t}$. Note that stimulated emission is what it says – the more there's light, the more there will be emission.

Spontaneous Emission:

Exists even if there's no light. We'll describe in more detail in a latter chapter.

Thermal rates

We are now ready for thermal rates – when we can start from several initial vibrational states;

Recall the main equation of our previous section, the initial-rovibrational-state-selected rate of transition from one electronic state to another:

$$\Gamma_{F\leftarrow(\nu,G)} = 2|b|^2 \int_{-\infty}^{\infty} e^{i(\omega+\varepsilon_{\nu,G})t} \langle \chi_{\nu,G}|\mu e^{-iH_F t}\mu|\chi_{\nu,G}\rangle dt \tag{22.35}$$

Now let's assume that there's a **thermal distribution** of the initial ro-vibrational states on the ground electronic state. I.e., the states are distributed by the Boltzmann distribution, so the probability for an initial ro-vibrational state to be "ν" is

$$p_\nu = \frac{e^{-\beta\varepsilon_{\nu,G}}}{Q}, \qquad Q = \sum_\nu e^{-\beta\varepsilon_{\nu,G}} \tag{22.36}$$

(and recall that Q is defined this way so that $\sum_\nu p_\nu = 1$.) Here, $\beta \equiv (k_B T)^{-1}$. Also, we'll assume for simplicity that the higher electronic states are sufficiently high so that thermally they are not excited, i.e., only ro-vibrational states on the ground-electronic state have non-vanishing population.

We then define **the thermal transition rate to an electronic state "m"**

$$\Gamma_{F\leftarrow G} \equiv \sum_\nu p_\nu \Gamma_{F\leftarrow(\nu,G)} \tag{22.37}$$

so

$$\Gamma_{F\leftarrow G} = \frac{|b|^2}{Q}\sum_\nu e^{-\beta\varepsilon_{\nu,G}}\int_{-\infty}^{\infty} e^{i(\omega+\varepsilon_{\nu,G})t}\langle\chi_{\nu,G}|\mu e^{-iH_F t}\mu|\chi_{\nu,G}\rangle dt. \tag{22.38}$$

Note that we colored two terms in blue. This is because we can insert them into the bra-ket, using

$$e^{-\beta\varepsilon_{\nu,G}}e^{i\varepsilon_{\nu,G}t}|\chi_{\nu,G}\rangle = e^{-\beta H_g}e^{itH_G}|\chi_{\nu,G}\rangle \tag{22.39}$$

(since $H_G|\chi_{\nu,G}\rangle = \varepsilon_{\nu,G}|\chi_{\nu,G}\rangle$). Thus,

$$\Gamma_{F\leftarrow G} = \frac{|b|^2}{Q}\int_{-\infty}^{\infty} e^{i\omega t}\sum_\nu\langle\chi_{\nu,G}|\mu e^{-iH_F t}\mu e^{-\beta H_G}e^{itH_G}|\chi_{\nu,G}\rangle dt.$$

Now I colored the summation over terms in brown, because I want to use the **definition of a trace** – the sum over a complete basis of the expectation value.

I.e., since the ground-state wavefunctions are a complete set over the ro-vibrational coordinates, we can write for any operator A

$$\sum_\nu\langle\chi_{\nu,g}|A|\chi_{\nu,g}\rangle = \mathrm{Tr}\,(A)$$

i.e., again

$$\boxed{\Gamma_{F\leftarrow G} = \frac{|b|^2}{Q}\int_{-\infty}^{\infty} e^{i\omega t}C(t)dt} \tag{22.40}$$

where now however

$$\boxed{C(t) \equiv \mathrm{Tr}\left(\mu e^{-iH_F t}\mu e^{-\beta H_G}e^{iH_G t}\right).} \tag{22.41}$$

This is a very elegant expression. It says that the correlation function has to do with **motion back-and-forth,** $e^{iH_G t}$ **and** $e^{-iH_F t}$**, on two different surfaces,** where the motion on the ground state surface **is "weighted"** by the Boltzmann factor.

Importance for Approximations

An advantage of correlation-function expressions is that they are **a convenient starting point for approximations.** Basically, most approximations are good at short times and fail at long times, so by having a time expressions we can limit the time (by putting a convergence factor, $e^{-t^2\Gamma^2}$, multiplying the correlation function, where Γ is an energy-width parameter) thereby only including contributions from early times when the approximations will be valid.

Other expressions with correlation function

So far we talked about linear absorption from one electronic state to another. Similar expressions follow for other absorptions. These include

- Absorption on the same electronic state
- Raman Scattering (a **two-photon** process)
- Scattering of radiation (rather than absorption).
- And, without reference to lasers, Flux-Flux expressions for chemical reactions.
- And related to that, electron-transfer rate

Check, for example, the book of Tannor (David Tannor, Introduction to Quantum Mechanics: A Time-Dependent Perspective) for an expanded view of this correlation-function picture.

Other Spectroscopies: Vibrational Spectroscopy (without electronic transitions).

The correlation function methods we introduced are applicable to other spectroscopies – vibrational, rotational, NMR, and also non-linear spectroscopies (e.g., Raman).

To see, start from the expression we derived earlier, $C(t) \equiv \mathrm{Tr}\left(e^{iH_F t}\mu e^{-iH_G t}e^{-\beta H_G}\mu\right)$. And we recall that a Fourier transformation of the correlation function will give the absorption frequency.

Now imagine that the motion is on the same surface, i.e., **ro-vibrational** spectroscopy (where the **electronic state does not** change).

We'll remove the "F" and "G" subscript therefore.

But we'll also need to be careful about the dipole moment, $\mu(\boldsymbol{R})$ (i.e., $\mu_{gg}(\boldsymbol{R})$). Let's write it as a **constant plus a non-constant part,** with the latter **approximated by its first-derivative part:**

$$\boldsymbol{\mu}(\boldsymbol{R}) \simeq \boldsymbol{\mu}(\boldsymbol{R}_{\mathrm{eq}}) + (\boldsymbol{R} - \boldsymbol{R}_{\mathbf{eq}}) \cdot \boldsymbol{\mu}' \tag{22.42}$$

We should consider only **transitions** from the initial state; if the final state is on the same electronic potential surface (our present case), we'll need to consider only **the non-constant part of μ,** since the constant part will not cause transitions (i.e., if we start on the 3rd ro-vibrational state of the ground electronic state, applying the constant part of mu will keep us on the same ro-vibrational state).

In practice, it is **OK** to keep the constant term, but then we should **subtract from the correlation function its non-zero value at t=0.**

The correlation function becomes then very simply:

$$\boxed{C(t) \equiv \mathrm{Tr}\left(e^{iHt}\boldsymbol{\mu} e^{-iHt} e^{-\beta H}\mu\right)} \tag{22.43}$$

A very interesting expression then emerges when we **go to the Heisenberg picture** (or really to the interaction picture, but let's not dwell on that). Then we can write:

$$\boldsymbol{\mu}(\boldsymbol{t}) = e^{iHt}\boldsymbol{\mu} e^{-iHt} \tag{22.44}$$

so then

$$\boxed{C(t) \equiv \mathrm{Tr}\left(e^{-\beta H}\boldsymbol{\mu}(t)\boldsymbol{\mu}(t=0)\right)}. \tag{22.45}$$

This is a **classical-like expression:** absorption in classical mechanics is the Fourier-transform of the dipole-dipole correlation function, i.e., just this expression!

Also, note that $\boldsymbol{\mu}(t)$ is an **operator** in **R**. (For t=0 it is a local operator, $\mu(t=0) \equiv \mu(\boldsymbol{R})$, but it will not be local for latter times.)

Spontaneous Emission: Prequel

Our discussion so far assumed a **classical** electric field. That discussion explained then stimulated absorption and emission but not spontaneous one.

Physically, spont. emission happens since the true "system" is not just the molecule, but the molecule+radiation modes. Put differently, there are many "modes" the light can have. These (i.e., each allowed momentum) is associated with **modes of the electromagnetic field.** In a given box of lengths $L \cdot L \cdot L$ (and volume $V = L^3$) – and this L has nothing to do with angular momentum – the allowed modes of the periodic electromagnetic field are labeled by the "wavevector".

These modes are exactly the same as the modes of the electron w.f. that we saw when we considered in 115a/215a the electron modes in the Homogeneous Electron Gas, i.e., these **modes**, which are the **momentum of** the modes (i.e., "momentum of the photons"), are similarly labeled as $\boldsymbol{k} = (k_x, k_y, k_z)$, with $k_x = \frac{2\pi}{L} \times integer$. Each electromagnetic mode is also labeled by another parameter – the polarization.

Within each mode, there could be an electric field oscillating at that mode (with that momentum, i.e., wavevector, $\boldsymbol{k}$,and frequency, $\omega = c|\boldsymbol{k}|$). We'll see **that for each mode, the amplitude of the electric field at that mode can be thought of as as an independent Harmonic oscillators.** The **state** of that harmonic oscillator ($n = 0, 1, 2, etc.$) **is the number of photons in that mode.**

Thus, spontaneous emission means that we go from a mode without photons (i.e., with $n = 0$) to a mode with $n = 1$. (It is more general than this, actually, and we'll explain later.)

To understand spontaneous emission, we'll thus need to understand light as a quantum object. This will be the topic of the next chapter.

23 Electromagnetic Fields and interaction with them: proper-classical derivation and quantum derivation (for spontaneous emission).

We have already seen the Fermi Golden rule applied with external fields. Now we need to do a more "proper" derivation, explaining about electric and magnetic fields, and eventually landing in quantum description of the electric fields. This is important for spontaneous emission and also for modern studies of light interacting with matter, at times non-perturbatively.

$\boldsymbol{A}$ and ϕ : vector potential and scalar potential: Classical treatment

Maxwell's equations are for the electric and magnetic fields, $\boldsymbol{E}(\boldsymbol{r}, t)$ and $\boldsymbol{B}(\boldsymbol{r}, t)$. But for quantum mechanics it's easier to work with the electromagnetic vector potential $\boldsymbol{A}(\boldsymbol{r}, t)$ and scalar potential $\varphi(\boldsymbol{r}, t)$; you probably saw them but a reminder – they're defined so that the electromagnetic fields derive from them, i.e.,

$$\boldsymbol{E} = -\boldsymbol{\nabla}\varphi - \frac{1}{c}\frac{\partial A}{\partial t} \tag{23.1}$$

$$\boldsymbol{B} = \boldsymbol{\nabla} \times \boldsymbol{A}. \tag{23.2}$$

Note that the $\boldsymbol{E}$,$\boldsymbol{B}$ have 6 components together that are highly constrained by the Maxwell equations; $\boldsymbol{A}$ and ϕ have only 4, and therefore are easier to work with.

Gauge transformation

Further: we have still some freedom in choosing $\boldsymbol{A}$, φ; the electric and magnetic fields don't change when we do a **gauge transformation**, i.e., pick an arbitrary scalar function $\chi(\boldsymbol{r}, t)$ and modify $\boldsymbol{A}$, φ as follows:

$$\boldsymbol{A} \to \boldsymbol{A} + \boldsymbol{\nabla}\chi$$

$$\varphi \to \varphi - \frac{1}{c}\frac{\partial \chi}{\partial t}. \tag{23.3}$$

(prove that the electric and magnetic fields do not change when we do this transformation, using the fact that the curl of a gradient is zero!).

> Note: Gauge transformations are very important when coupling to quantum mechanics, see later.

Three plausible gauges are:

- One would be $\varphi = 0$, i.e., given some φ, negate it by a gauge transformation $\chi(\boldsymbol{r},t) = c\int_0^t \varphi(\boldsymbol{r},t')dt'$, and end up with only $\boldsymbol{A}$.
- A physically better one would be the Coulomb gauge, $\nabla \cdot \boldsymbol{A} = 0$. (That means that $\boldsymbol{A}$ is "not-longitudinal".) That's the one we will generally use.
- Yet another one is the Lorenz gauge, where $\boldsymbol{A}$ and φ are treated equally.

Choosing the Coulomb gauge, then in empty space (what we usually care about, the effects of atoms on the light for us are small so we for now ignore to 0-th order their effects on the electromagnetic waves which essentially "see" an empty space) then $\varphi = const.$ and the Maxwell eqs. become a single **wave equation for $\boldsymbol{A}$**

$$\nabla^2 \boldsymbol{A} - \frac{1}{c}\frac{\partial^2 \boldsymbol{A}}{\partial t^2} = 0. \tag{23.4}$$

More generally, when we are not in empty space, than in the Coulomb (and to some extent the Lorenz) gauge, the scalar potential has similar properties to the electrostatic potentials we're used to – the Hartree potential and the electron-nuclear potential, while the effects of moving charges cause magnetic fields and the emergence of the vector potential. But that's not for now.

Plane waves:

As presented earlier in the classical discussion of fields, we'll deal in this course with the simplest form of electromagnetic fields, plane waves (the discussion here is repetitive for completeness).

$$\boldsymbol{A}(\boldsymbol{r},t) = \hat{\boldsymbol{\epsilon}} A_0 \cos(\boldsymbol{k}\cdot\boldsymbol{r} - \omega t + \phi_0) = \frac{A_0}{2}\hat{\boldsymbol{\epsilon}} e^{i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t+\phi_0)} + \frac{A_0}{2}\hat{\boldsymbol{\epsilon}} e^{-i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t+\phi_0)}. \tag{23.5}$$

Here, the unit-vector $\hat{\boldsymbol{\epsilon}}$ (not related to the dielectric constant – we don't have enough symbols!) is the **polarization** of the wave. ϕ_0 (omitted in many books) is the phase; A_0 is (up to $\sqrt{2}$ factors) the amplitude; and $e^{i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)}$ corresponds to a wave moving in the $+\boldsymbol{k}$ direction, with a velocity:

$$\frac{\omega}{k} = c. \tag{23.6}$$

It is actually quite more complicated in the general case. The expression above is the phase velocity, $v_{phase} = \frac{\omega}{k}$. A general wave really moves with the group velocity, $v_{group} = \frac{d\omega}{dk}$ (we may prove it later, or you may have seen it). For electromagnetic waves in vacuum the two are identical, due to the linear relation between $\omega = kc$, where $k \equiv |\boldsymbol{k}|$, so we omit the subscript. But for other waves it isn't:

1. For "Schrödinger" waves of free particles with mass in vacuum, the group velocity is twice the phase velocity (since $\omega = \frac{E}{\hbar} = \frac{p^2}{2M\hbar} = \frac{\hbar^2 k^2}{2M\hbar}$, i.e., $\omega \propto k^2$; when we differentiate ω w.r.t k we get a factor of 2, i.e., $\frac{d\omega}{dk} = 2\frac{\omega}{k}$)
2. For electromagnetic waves in a medium, the phase velocity is often higher than speed of light.
3. One can make "crazy" meta-material circuits where the wave goes forward (i.e., the group velocity points forward) but the phase velocity is backward!

Electromagnetic waves are transverse (in the Coulomb gauge).

Since $\nabla(\boldsymbol{k}\cdot\boldsymbol{r}) = \boldsymbol{k}$ **(prove it!)** the Coulomb gauge implies

$$0 = \nabla\cdot\,\boldsymbol{A} = -(\hat{\boldsymbol{\epsilon}}\cdot\boldsymbol{k})\cdot A_0 \sin(\boldsymbol{k}\cdot\boldsymbol{r} - \omega t + \phi_0) \tag{23.7}$$

i.e.,

$$\hat{\boldsymbol{\epsilon}}\cdot\boldsymbol{k} = 0. \tag{23.8}$$

This is what we call a transverse wave – the polarization is **perpendicular** to the direction of motion.

For example, if the wave moves along the $\boldsymbol{k}||z$ direction, then it will oscillate within (i.e., $\hat{\boldsymbol{\epsilon}}$ will be within) the $x-y$ plane. In contrast, of course, to sound-waves that are longitudinal,i.e., oscillate along the direction of the wave motion.

Note that as mentioned you can make crazy electromagnetic waves that have longitudinal components using so called meta-materials (especially "active ones", where you pump energy into them continuously).

Plane wave intensity and # of photons per volume

It is easy to see (see Schatz and Ratner 5.3 for details, we'll just skim over) the relation between the amplitude of the vector potential, A_0, and the # of photons per volume. Specifically, the energy density (per volume) is classically the average (that's what the over-line stands for), over different times, of the squared fields:

$$\text{Energy density}=\frac{\overline{\boldsymbol{E^2}+\boldsymbol{B^2}}}{8\pi} = 2\frac{\overline{E^2}}{8\pi} \tag{23.9}$$

where we used the fact that in average, for a plane wave, the electric and magnetic fields have the same intensity. Further, for a plane wave $\boldsymbol{E^2} = \frac{\omega^2}{c^2}A_0^2\sin^2(\boldsymbol{k}\cdot\boldsymbol{r} - \omega t + \phi_0)$ (prove it!) and since the average of $\sin^2$ is $\frac{1}{2}$, we get

$$\text{Energy density} = 2\frac{\overline{E^2}}{8\pi} = 2\frac{\omega^2}{8\pi c^2}A_0^2 \cdot \frac{1}{2} = \frac{\omega^2}{8\pi c^2}A_0^2 \tag{23.10}$$

Now, the flux of radiation in an electromagnetic wave, i.e., the amount of energy carried per time per area, is shown in E&M classes to be (in the CGS units we use):

$$P = \frac{c}{4\pi}\overline{\boldsymbol{E} \times B} = c\frac{\overline{E^2}}{4\pi} = \frac{\omega^2}{8\pi c}A_0^2 \tag{23.11}$$

(Usually the Poynting vector is denoted by S, but I use P). So the final expression is

$$A_0^2 = \frac{8\pi c}{\omega^2}P. \tag{23.12}$$

Coupling electromagnetism to the particle motion in classical dynamics

Classically, we know that the force for an electron (with charge $-e$, and note that Schatz and Ratner use a charge of $+e$) in electric and magnetic fields is

$$F = -e\left(\boldsymbol{E} + \frac{\boldsymbol{v} \times \boldsymbol{B}}{c}\right) \tag{23.13}$$

This is labeled the Lorenz force. Turns out that we can get this force law from the following Hamiltonian

$$H = \frac{1}{2m}(\boldsymbol{p} + \frac{e}{c}\boldsymbol{A})^2 - e\varphi \tag{23.14}$$

where the vector and scalar potential, $\boldsymbol{A}, \phi$ are what we saw earlier, i.e., such that we can get the electric and magnetic potential from them.

An important point: typically we divide the potentials such that the scalar potential is static – not time-dependent – and is due to the nucleus (and if there are other electrons and we treat them in a "mean-field" manner, than it will be due to the other electrons too). The vector potential would be due to an external electromagnetic field, e.g., a laser pulse. This splitting is feasible with the Coulomb gauge we use.

Hamiltonian Dynamics

I presume you saw Hamiltonian dynamics. But if not, here's a quick review. A Hamiltonian is a function of coordinates and momenta, from which you can derive the force law. Given a Hamiltonian, the equation of motion are

$$\frac{d\boldsymbol{r}}{dt} = \frac{\partial H}{\partial \boldsymbol{p}} \tag{23.15}$$

$$\frac{d\boldsymbol{p}}{dt} = -\frac{\partial H}{\partial \boldsymbol{r}} \tag{23.16}$$

In our case, the Lorentz force is known so what was done was to backtrack, i.e., search for a Hamiltonian that yields, from the previous equations, the Lorenz force law, $m\,\ddot{\boldsymbol{r}} = \boldsymbol{F}$.

It turns out that with some painful algebra we can prove that indeed these last two Eqs., (23.15)-(23.16) yield, when applied to the Hamiltonian, Eq. (23.14), the Lorenz force, Eq. (23.13). See Ratner and Schatz.

Hamiltonian as a tool for quantization

It turns out, and this will be important later, that given the Hamiltonian of a classical system it is "easy" to find the associated quantum Hamiltonian. Specifically, we **replace p** by $-i\nabla$. That's it! This will be useful later.

Side note: p is not m*v here!

The coupling to electromagnetic fields Hamiltonian exemplifies something interesting: p is not necessarily mass*velocity! Specifically, since $H = \frac{1}{2m}(p + \frac{e}{c}\boldsymbol{A})^2 - e\varphi$,then

$$\boldsymbol{velocity} \equiv \frac{d\boldsymbol{r}}{dt} = \frac{\partial H}{\partial \boldsymbol{p}} = \frac{p + \frac{e}{c}\boldsymbol{A}}{m}, \tag{23.17}$$

i.e.,

$$p = m * velocity - \frac{e}{c}\boldsymbol{A}. \tag{23.18}$$

Hamiltonian of QUANTUM particle in presence of CLASSICAL electromagnetic fields.

Knowing the Hamiltonian is a nice feature in classical mechanics, but is crucial in quantum mechanics. Replacing p by $-i\hbar\nabla$ (and setting $\hbar = 1$), gives

$$H = \frac{1}{2m}\left(-i\nabla + \frac{e}{c}\boldsymbol{A}\right)^2 - e\varphi \tag{23.19}$$

Let's expand (reverting back to p):

$$H = \frac{p^2}{2m} + \frac{e}{2mc}\left(\boldsymbol{A}\cdot\boldsymbol{p} + \boldsymbol{p}\cdot\boldsymbol{A}\right) + \frac{e^2}{2m}\boldsymbol{A^2} - e\varphi. \tag{23.20}$$

This expression is easily extended to many-particle systems, see Schatz and Ratner (Eq. 5.26).

There's a simplification in the Coulomb gauge ($\nabla\cdot\,\boldsymbol{A} = 0$). Specifically, the difference between the two terms in the parentheses is just (proportional to) $\nabla\cdot\boldsymbol{A}$, so it is zero in the Coulomb gauge.

$$\begin{aligned}\boldsymbol{p}\cdot\boldsymbol{A} - \boldsymbol{A}\cdot\boldsymbol{p} &= p_x A_x + p_y A_y + p_z A_z - A_x p_x - \dots \\ &= [p_x, A_x] + [p_y, A_y] + \dots = -i\frac{\partial A_x}{\partial x} - i\frac{\partial A_y}{\partial y} - \dots = -i\boldsymbol{\nabla}\cdot\boldsymbol{A} = 0\end{aligned} \tag{23.21}$$

Therefore, $\boldsymbol{A} \cdot \boldsymbol{p} + \boldsymbol{p} \cdot \boldsymbol{A} = 2\boldsymbol{A} \cdot \boldsymbol{p}$.

Now, divide the Hamiltonian to the static part and two disturbances: $H = H_0 + V + \bar{V}$

$$H_0 = \frac{p^2}{2m} - e\phi \tag{23.22}$$

$$V = \frac{e}{mc}\boldsymbol{A} \cdot \boldsymbol{p} \tag{23.23}$$

$$\bar{V} = \frac{e^2}{2mc^2}\boldsymbol{A^2}. \tag{23.24}$$

In the shifted Hamiltonian, Eq. (23.14), the $\frac{e}{c}A$ term is typically much smaller than p so $\bar{V} \ll V \ll H_0$ (see the discussion in Schatz and Ratner, Section 5.3.3). So we typically ignore $\bar{V}$.

An exception is large length scales, over which $\boldsymbol{A}$ changes so much (since it is an integral over a distance) such that $\bar{V}$, which is a potential (i.e., a function of position, $\bar{V} = \bar{V}(\boldsymbol{r})$) can be large; this is the situation in **optical traps**, where one prepares the electromagnetic field so that $\bar{V}(\boldsymbol{r})$ will be small in some region and large in others, so that the particle will be attracted to the small-$\bar{V}(\boldsymbol{r})$ regions.

Dipole approximation

We can now give a derivation of the dipole approximation, i.e., $V = -\mu \cdot \boldsymbol{E}$ (we don't label the field anymore as $E_{\text{field.}}$).

Inserting the equation for $\boldsymbol{A}$ in e.m. waves, Eq. (23.5), and throwing out an unimportant phase, we get that the disturbance is (for an x-polarization)

$$V = \frac{e}{mc}A_x p_x = \frac{e}{2mc}\left(A_0(\omega)e^{i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)} + A_0(\omega)e^{-i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)}\right)p_x \tag{23.25}$$

where $A(\omega)$ is the intensity of the wave.

We note that $kr = \frac{2\pi}{\lambda}r$. So for atomic optical transitions where $r <$ 1nm, $\lambda \sim 100 - 1000$nm, then $kr \ll 1$. We often make therefore the **dipole approximation,** where the $e^{i\boldsymbol{k}\cdot\boldsymbol{r}}$ term is replaced by 1.

Then

$$V = \frac{eA_0(\omega)p_x}{2mc}\left(e^{-i\omega t} + e^{i\omega t}\right). \tag{23.26}$$

Clearly, this looks like the equation we had long ago, Eqs. (21.1)-(21.3) except that instead of the spatial factor of $-ex\frac{|E_{\text{field}}|}{2}$ we now have $\frac{eA_0(\omega)p_x}{2mc}$. I.e., we replaced $-x|E_{\text{field}}\,|$ by $A_0(\omega)\frac{p_x}{mc}$.

We can easily show that **the matrix elements of the two terms are, however, equal,** so they are interchangeable. Specifically, given an initial state χ_i with energy (i.e., eigenvalue of H_0) ε_i and a final state χ_f with energy ε_f,the

matrix element of p_x will be $\langle\phi_f|p_x|\phi_i\rangle$.Here comes a trick: look at the commutator of H with the position operator

$$[x, H_0] = \frac{1}{2m}[x,p_x^2 + p_y^2 + p_z^2] + [x, V(\boldsymbol{r})] = \frac{[x, p_x^2]}{2m} = \frac{[x, p_x]p_x + p_x[x, p_x]}{2m} = i\frac{p_x}{m}. \quad (23.27)$$

and similarly for y and z. Thus:

$$p_x = -im[x, H_0], \quad (23.28)$$

so the matrix element of p_x is

$$\langle\phi_f|p_x|\phi_i\rangle = -im\langle\phi_f|[x, H_0]|\phi_i\rangle = -im\langle\phi_f|xH_0 - H_0x|\phi_i\rangle, \quad (23.29)$$

and using $H_0|\phi_i\rangle = \varepsilon_i|\phi_i\rangle$, $\langle\phi_f|H_0 = \langle\phi_f|\varepsilon_f$, and $\omega = \varepsilon_f - \varepsilon_i$, gives

$$\langle\phi_f|p_x|\phi_i\rangle = -im\langle\phi_f|x\varepsilon_i - \varepsilon_f x|\phi_i\rangle = -im\omega \cdot \langle\phi_f|x|\phi_i\rangle, \quad (23.30)$$

On the other hand, since we are talking about plane wave and we ignore the static potential, φ,then $E = -\frac{1}{c}\frac{\partial \boldsymbol{A}}{\partial t}$; differentiating A w.r.t time gives a factor of ω,so $|E_{\text{field}}| = \frac{\omega}{c}A_0$.

Thus, up to an overall factor of "i" (since we were not careful on phases) we get that, **as far as matrix elements,** $-x|E_{\text{field}}|$ and $A_0\frac{p_x}{mc}$ are equal.

Quantum electromagnetic fields

Next we'll consider quantum treatment of electromagnetic fields, important for stimulated emission and for modern treatments of molecules in cavities. We'll follow here Schatz and Ratner, Sec. 6.2.3-6.2.4, with abbreviations.

Classical Hamiltonian for electromagnetic field

Thus, we recall that classically the electromagnetic field in free space satisfies the wave equation:

$$\nabla^2\boldsymbol{A} = \frac{1}{c}\frac{\partial^2\boldsymbol{A}}{\partial t^2} \quad (23.31)$$

We know what the solutions to these equation are, a plane wave. But we should not use that equation when we derive!

So if we can't use the classical wave equation, what can we use? The answer is that we need to start from the **classical Hamiltonian for the fields themselves** (in addition to the field-particle Hamiltonian that we had earlier). We mentioned above that the "energy density" is $\frac{\overline{\boldsymbol{E^2}+\boldsymbol{B^2}}}{8\pi}$,so the energy, i.e., the

Hamiltonian (we wont distinguish between the two, you need to take analytical mechanics classes to clarify that) is

$$H = \int \frac{\boldsymbol{E}^2 + \boldsymbol{B}^2}{8\pi} d\boldsymbol{r}. \tag{23.32}$$

And relating to the electric and magnetic field gives (again, free space and we ignore φ), we get the **Hamiltonian of an electromagnetic field in free space**

$$H = \frac{1}{8\pi} \int \left(\left(\frac{1}{c^2} \frac{\partial \boldsymbol{A}}{\partial t} \right)^2 + (\boldsymbol{\nabla} \times \boldsymbol{A})^2 \right) d\boldsymbol{r}. \tag{23.33}$$

Now let's write the electromagnetic field in free space as **sum of modes,** i.e., a temporal form times a spatial function (times a convenient factor for later)

$$\boldsymbol{A}(\boldsymbol{r}, t) = \sqrt{\frac{4\pi}{V}} c \sum_i q_{i\sigma}(t) u_{i\sigma}(\boldsymbol{r}) \tag{23.34}$$

without yet "telling" ourselves that $q_{k,\sigma}(t)$ should be $e^{\pm i\omega t}$. The spatial modes should be ones that when inserting to the Hamiltonian expression give a simple expression (not mixing the modes).

It is easy to guess that the spatial form is a plane-wave, i.e., replace the spatial index "i" by the 3-dimensional index of the plane wave $\boldsymbol{k}$, so

$$u_{\boldsymbol{k}\sigma}(\boldsymbol{r}) = e^{i\boldsymbol{k}\cdot\boldsymbol{r}} \tag{23.35}$$

for this homogeneous case.[2] Here, "$\boldsymbol{k}$" stands for the allowed modes (i.e., for a cavity of lengths $L \times L \times L$ and volume $V = L^3$, they will be $\boldsymbol{k} = \frac{2\pi}{L}(m_x, m_y, m_z)$, where the m's are integer – as we saw for the HEG), and σ is the polarization mode (perpendicular to $\mathbf{k}$).

Inserting the last two equation to the Hamiltonian we get indeed a simple form, as promised; the math details are in Schatz and Ratner, eqs. 6.17-6.20, and the final results are:

$$H = \frac{1}{2} \sum_{\boldsymbol{k}} \left(\left(\frac{dq_{\boldsymbol{k}}}{dt} \right)^2 + c^2 \boldsymbol{k}^2 q_{\boldsymbol{k}}^2 \right) = \frac{1}{2} \sum_{\boldsymbol{k}} \left(\left(\frac{dq_{\boldsymbol{k}}}{dt} \right)^2 + \omega_{\boldsymbol{k}}^2 q_{\boldsymbol{k}}^2 \right) \tag{23.36}$$

where we define here of course $\omega_{\boldsymbol{k}} = |\boldsymbol{k}|c$. **We suppress the polarization index,** i.e., for each $\mathbf{k}$ there are really two such oscillators associated with the two transverse directions.

This **still classical** form looks just like a Harmonic Oscillator! (with a "mass" of 1). I.e., a classical electromagnetic wave in a uniform free space has a Hamiltonian (or energy) that is just that of a **sum of Harmonic oscillators.**

It is exactly this form which we can now quantize!

---

[2] It will be different if we were to consider constrained geometries, which is a very hot topic in today's physical chemistry, since we can then induce things like fast spontaneous emission, etc.

Quantization of the electromagnetic field

We first note that, since the "mass" of each mode in the equation above is 1 (this is not the mass of the photon, just an arbitrary constant reflecting the fact that there's no constant dividing $\left(\frac{dq_{\boldsymbol{k}}}{dt}\right)^2$ in the energy), the momentum of the mode is

$$p_{\boldsymbol{k}} = mass * \frac{dq_{\boldsymbol{k}}}{dt} = \frac{dq_{\boldsymbol{k}}}{dt}. \tag{23.37}$$

(Recall that this momentum-velocity relation is **not always valid**, but it is true for Harmonic oscillators). Then quantum mechanically, $[p_{\boldsymbol{k}}, q_{\boldsymbol{k}}] = -i$ (we set henceforth $\hbar = 1$).

All we have to do now for quantization is define **quantum annihilation** and **creation** operators for each mode of the electromagnetic field.

$$b_{\boldsymbol{k}} = \frac{1}{\sqrt{2\hbar\omega_{\boldsymbol{k}}}} \left(\omega_{\boldsymbol{k}} q_{\boldsymbol{k}} + i p_{\boldsymbol{k}}\right) \tag{23.38}$$

$$b^+_{\boldsymbol{k}} = \frac{1}{\sqrt{2\hbar\omega_{\boldsymbol{k}}}} \left(\omega_{\boldsymbol{k}} q_{\boldsymbol{k}} - i p_{\boldsymbol{k}}\right) \tag{23.39}$$

and they fulfill of course $[b^+_{\boldsymbol{k}}, b_{\boldsymbol{k}}] = \delta_{\boldsymbol{k'},\boldsymbol{k}}$, just like what we learned about regular Harmonic oscillators.

For later, let's express the q's and p's in terms of the b's, setting $\hbar = 1$:

$$q_{\boldsymbol{k}} = \frac{b^+_{\boldsymbol{k}} + b_{\boldsymbol{k}}}{\sqrt{2\omega_{\boldsymbol{k}}}} \tag{23.40}$$

$$p_k = i\sqrt{\omega_k}\frac{b^+_k - b_k}{\sqrt{2}} \tag{23.41}$$

Remembering what we learned about regular H.O., we get then

$$H = \sum \omega_{\boldsymbol{k}} \left(b^+_k b_k + \frac{1}{2}\right). \tag{23.42}$$

Reminder: the number operator.

Recall that the **number** operator for a Harmonic oscillator is $N = b^+ b$. Put differently, given the n'th eigenstate of a single Harmonic Oscillator, we proved in 115a/215a (we used "a" rather than "b" then)

$$b|n\rangle = \sqrt{n}|n-1\rangle \tag{23.43}$$

from which follows

$$b^+|n-1\rangle = \sqrt{n}|n\rangle. \tag{23.44}$$

(Note: prove the equivalence of the last two lines!)

Therefore:

$$N|n\rangle = b^+b|n\rangle = b^+\sqrt{n}|n-1\rangle = \sqrt{n}\cdot\left(b^+|n-1\rangle\right) = \sqrt{n}\cdot\left(\sqrt{n}|n\rangle\right) = n|n\rangle. \tag{23.45}$$

This was for a single Harmonic oscillator; for our case that there are many H.O., each labeled by $\boldsymbol{k}$, then the number operator for each mode, that measures how many oscillations (i.e., **photons**) are in that mode, is

$$n_{\boldsymbol{k}} = b^+_{\boldsymbol{k}} b_{\boldsymbol{k}}. \tag{23.46}$$

Therefore,

$$H = \sum \omega_{\boldsymbol{k}}\left(n_{\boldsymbol{k}} + \frac{1}{2}\right). \tag{23.47}$$

Let's # the modes (i.e., the $\boldsymbol{k}$'s) from 1, 2, etc. (so "1" may be associated with a momenta 0, "2" with a slightly higher momenta, etc.)

The overall state of the photons are then denoted as:

$$|\psi\rangle = |n_1, n_2, ...n_{\boldsymbol{k}}, ...\rangle. \tag{23.48}$$

Here, "$n_1, n_2, ...$" means: n_1 photons in the 1st mode; n_2 photons in the 2nd mode, etc.

A better way to write it is

$$|\psi\rangle = (b_1^+)^{n_1}(b_2^+)^{n_2}(b_3^+)^{n_3}|\rangle, \tag{23.49}$$

where $|\rangle$ implies the vacuum state without photons;

I.e., the overall state of the radiation will be associated with creating n_1 photons in the 1st mode; n_2 photons in the 2nd mode, etc. (Note that the b_j^+ commute so it does not matter which order we use, b_1^+ on the left or the right of b_2^+,etc.)

A photon is just **an excitation of a specific mode**; for example, if the 20th mode has 3 photons, it means that its state is $n_{20} = 3$, i.e., the full wavefunction of all the modes has a component $(b_{20}^+)^3$ in it.

The vector potential in terms of the mode-oscillator creation/annihilation(Schatz and Ratner 6.2.4)

Recall that we expanded above $\boldsymbol{A}(\boldsymbol{r}, t) = \sqrt{\frac{4\pi}{V}}c\sum_{\boldsymbol{k}} q_{\boldsymbol{k}}(t)e^{i\boldsymbol{k}\cdot\boldsymbol{r}}$.

We will use henceforth the Schrödinger approach, where $\boldsymbol{A}$ is taken at t=0 (the w.f. change, including the photon modes, but the operators don't change). So

$$\boldsymbol{A}(\boldsymbol{r}) = \sqrt{\frac{4\pi}{V}}c\sum_{\boldsymbol{k}} q_{\boldsymbol{k}}e^{i\boldsymbol{k}\cdot\boldsymbol{r}} \tag{23.50}$$

Now recall, using $q_{\boldsymbol{k}} = \frac{b_{\boldsymbol{k}}^{+} + b_{\boldsymbol{k}}}{\sqrt{2\omega_{\boldsymbol{k}}}}$. So the field becomes (**now we're un-hiding** the polarization part $\boldsymbol{\epsilon}$)

$$\boldsymbol{A}(\boldsymbol{r}) = \sqrt{\frac{4\pi}{V}} c \sum_{\boldsymbol{k}\boldsymbol{\epsilon}} \frac{b_{\boldsymbol{k}\boldsymbol{\epsilon}}^{+} + b_{\boldsymbol{k}\boldsymbol{\epsilon}}}{\sqrt{2\omega_{\boldsymbol{k}}}} e^{i\boldsymbol{k}\cdot\boldsymbol{r}} \boldsymbol{\epsilon} \tag{23.51}$$

Thus the vector potential is made of **photon creation** and **photon annihilation** operators for each mode.

Coupling radiation to matter: the fully quantum expression (Schatz and Ratner 6.2.4)

We explained above that the perturbation potential in the Hamiltonian of a particle due to the interaction with the field is, to first order in the field, just $V = -\frac{e}{mc}\boldsymbol{A} \cdot p$. Therefore, based on the expressions above we see that

$$V = -\frac{e}{m}\sqrt{\frac{2\pi}{V}} \sum_{\boldsymbol{k}} \frac{1}{\sqrt{\omega_{\boldsymbol{k}}}} \left(b_{\boldsymbol{k}\boldsymbol{\epsilon}}^{+} e^{i\boldsymbol{k}\cdot\boldsymbol{r}} + b_{\boldsymbol{k}\boldsymbol{\epsilon}} e^{i\boldsymbol{k}\cdot\boldsymbol{r}} \right) \boldsymbol{\epsilon}\cdot\boldsymbol{p} \tag{23.52}$$

where the c's canceled. For later convenience, we now **change $\boldsymbol{k}$ to $-\boldsymbol{k}$** in the 1st summation (permissible, since it is a dummy index and we have really two different summations here), so

$$V = -\frac{e}{m}\sqrt{\frac{2\pi}{V}} \sum_{\boldsymbol{k}} \frac{1}{\sqrt{\omega_{\boldsymbol{k}}}} \left(b_{-\boldsymbol{k}\boldsymbol{\epsilon}} e^{-i\boldsymbol{k}\cdot\boldsymbol{r}} + b_{\boldsymbol{k}\boldsymbol{\epsilon}}^{+} e^{i\boldsymbol{k}\cdot\boldsymbol{r}} \right) \boldsymbol{\epsilon}\cdot\boldsymbol{p} \tag{23.53}$$

Note that the full Hamiltonian is then:

$$H = H_0 + V \tag{23.54}$$

where the non-coupled matter+interaction 0-order Hamiltonian is now

$$H_0 = H^{\text{molecule}} + H^{\text{field}} \tag{23.55}$$

(Note that we didn't include in the last two equations a 2nd order term in A in the coupling of the field to the molecule, i.e., we ignore the $V' \propto \frac{e^2}{2mc^2}A^2$ second-order term that comes from expanding $\frac{(\boldsymbol{p}+\frac{e}{c}\boldsymbol{A})^2}{2m}$).

The eigenstates of the 0th-order Hamiltonian are products of matter and field states:

$$|\psi_M\rangle = \psi_{m,n_1 n_2,\ldots} = |m\rangle|n_1, n_2, \ldots\rangle. \tag{23.56}$$

where "M" is a label of the combined matter/photon states, $M \equiv m, n_1, n_2, \ldots$

> Note: this "m" has nothing to do with mass, so don't confuse with the $\frac{1}{m}$ below. Also don't confuse V for potential and V for volume, appearing in the denominator too!

We'll talk in the next subsection about transition from this state "M" of the full system (matter + radiation-modes) which is an eigenstate of H_0 to another state "J" that its "bra" will be denoted as

$$\langle\psi_J| = \langle n_1', n_2', ...|\langle j|. \tag{23.57}$$

Fermi's Golden Rule for Stimulated AND spontaneous emission TOGETHER (Schatz and Ratner 6.2.5, almost verbatim)

Fermi's Golden rule states that the transition rate (from the initial state M of the full system to the final state J of the full system) is a product of a density-of-state (which we'll deal with later) and a transition matrix-squared:

$$\Gamma_{JM} = 2\pi\rho(E_J)|V_{JM}|^2 \tag{23.58}$$

where E_J is the energy of the full system, radiation+matter.

So start with he matrix elements of the interaction:

$$\begin{aligned} V_{JM} &= \langle\psi_J|V|\psi_M\rangle = \\ &-\frac{e}{m}\sqrt{\frac{2\pi}{V}}\sum_{\boldsymbol{k\epsilon}}\langle n_1', n_2', ...|\langle j|\left(b_{-\boldsymbol{k\epsilon}}e^{-i\boldsymbol{k\cdot r}} + b^+_{\boldsymbol{k\epsilon}}e^{i\boldsymbol{k\cdot r}}\right)\boldsymbol{\epsilon}\cdot\boldsymbol{p}\,|m\rangle|n_1, n_2, ...\rangle. \end{aligned} \tag{23.59}$$

Further, each term can only have one "**k"** contributing

> e.g.. if state "J" is different from state "M" by having an occupation of the mode $\boldsymbol{k_{specific}}$ being larger or smaller by 1, then only the $b_{\boldsymbol{k}_{\text{specific}}}$ term, or its Hermitian conjugate, will contribute. so that for each matrix element only one "k" can contribute.

So we'll consider then **each transition separately**. So then the squared matrix element becomes, for each transition associated with a photon **k**,

$$|V_{JM}(\boldsymbol{k})|^2 = \left|\langle n_1', n_2', ...|\langle j|\frac{1}{\sqrt{\omega_{\boldsymbol{k}}}}\left(b_{-\boldsymbol{k\epsilon}}e^{-i\boldsymbol{k\cdot r}} + b^+_{\boldsymbol{k\epsilon}}e^{i\boldsymbol{k\cdot r}}\right)\boldsymbol{\epsilon}\cdot\boldsymbol{p}\,|m\rangle|n_1, n_2, ...\rangle\right|^2. \tag{23.60}$$

We'll now assume the dipole approximation, i.e., assume $e^{-i\boldsymbol{k\cdot r}} = 1$. Since the V term is a product of two terms, one associated with the matter (**p**) and a term from the field, the matrix element breaks to two parts; so for absorption, where the number of photons in the final state is less t

$$|V_{JM}^{\text{absorption or emission}}(\boldsymbol{k})|^2 = |V(\text{matter})|^2\cdot\left|\langle V^{\text{absorption or emission}}(\text{radiation})\right|^2 \tag{23.61}$$

where

$$|V(\text{matter})|^2 = \left(\frac{e}{m}\sqrt{\frac{2\pi}{V}}\frac{1}{\sqrt{\omega_{\boldsymbol{k}}}}\right)^2 |\boldsymbol{\epsilon}\cdot\langle j|\boldsymbol{p}|m\rangle|^2 = \tag{23.62}$$

and using, as we saw,

$$\frac{1}{m}\langle j|\boldsymbol{p}|m\rangle = \langle j|[H,\boldsymbol{r}]|m\rangle = (E_j - E_m)\langle j|\boldsymbol{r}|m\rangle = \omega_{\boldsymbol{k}}\langle j|\boldsymbol{r}|m\rangle \tag{23.63}$$

(note that we use "m" here for two different thing, the electron mass and the quantum mechanical state of the system), we get

$$|V(\text{matter})|^2 = \frac{2\pi e^2}{V}\omega_{\boldsymbol{k}}\,|\boldsymbol{\epsilon}\cdot\langle j|\boldsymbol{r}|m\rangle|^2 \tag{23.64}$$

For the radiation, in absorption the k'th mode has <u>one less</u> photon in the final state; no change in the other photons, so the matrix element becomes (we only denote the # of photons associated with the **k**'th mode)"

$$\boxed{|V_{JM}^{\text{absorption}}(\boldsymbol{k})|^2 = |\langle n_{\boldsymbol{k}} - 1\,(b_{-\boldsymbol{k}\boldsymbol{\epsilon}})\,|n_{\boldsymbol{k}}\rangle|^2 = |\sqrt{n_{\boldsymbol{k}}}|^2 = n_{\boldsymbol{k}}} \tag{23.65}$$

But for **emission**, where the final **radiation** state has one **more** electron, then

$$\boxed{|V_{JM}^{\text{emission}}(\boldsymbol{k})|^2 = \left|\langle n_{\boldsymbol{k}} + 1|\left(b^{+}_{\boldsymbol{k}\boldsymbol{\epsilon}}\right)|n_{\boldsymbol{k}}\rangle\right|^2 = \left|\sqrt{n_{\boldsymbol{k}}+1}\right|^2 = n_{\boldsymbol{k}} + 1} \tag{23.66}$$

so the final transition rate is proportional to the occupation (which is proportional to the intensity) for absorption, **but in emission the occupation $n_{\boldsymbol{k}}$ is supplemented by 1 – i.e., is non-zero even if there's no field intensity – and this "1" is spontaneous emission!**

Density-of state factor,

The density of states with $\rho(E_J)$: it is density-of-state of the final state of the FULL system, i.e., it is a product of the density of state of the system (we'll label the system energy as ε_j, and that of the radiation:

$$\rho(E_J) \simeq \rho_{matter}(\varepsilon_j)\rho_{\text{radiation}}(\omega_{\boldsymbol{k}}) \tag{23.67}$$

(Not exactly, but let's not get to details).

The radiation density of states is proportional to $\rho_{\text{radiation}}(\omega_{\boldsymbol{k}})d\omega_{\boldsymbol{k}} \propto V\omega_{\boldsymbol{k}}^2 d\omega_{\boldsymbol{k}}$. (i.e., $\rho_{\text{radiation}}(\omega_{\boldsymbol{k}}) \propto V\omega_{\boldsymbol{k}}^2$.)We saw a similar expression when we considered the density of states for electron in a HEG, where the density of states rose like k^2.

Merging all the terms above us shows that the final expression has no V(=volume), like it shouldn't – for a molecule which is much tinier than the volume, it doesn't matter how big the volume is; but it has a factor of ω^3 (one

ω from the transition matrix element squared, two from the density of states of the radiation), and the dipole moment; i.e.

$$\boxed{\Gamma_{JM} = \frac{\omega_{\boldsymbol{k}}^3}{2\pi\hbar c^3} \left|\boldsymbol{\epsilon} \cdot \langle j|\boldsymbol{r}|m\rangle\right|^2 \rho(\varepsilon_j) d\Omega \cdot \begin{cases} n_{\boldsymbol{k}} & \text{absorption} \\ (n_{\boldsymbol{k}}+1) & \text{emission} \end{cases}} \tag{23.68}$$

where the "$d\Omega$" factor refers to the angle of the beam of photons we consider (absorption or emission)

Relative importance of stimulated and spontaneous emission

For the same laser intensity, if we compare UV vs. IR, then in **UV there will be many more modes (since** $\rho d\omega \propto \omega^2 d\omega = d\omega^3$).

So the intensity **per mode** will be smaller, i.e., $n_{\boldsymbol{k}}$ will be smaller. So the "1" in $n_{\boldsymbol{k}} + 1$ will dominate, i.e., **at high frequencies spontaneous emission dominates;** at low frequencies (IR) stimulated emission will, unless the laser intensity is very weak.

24 Electron Transfer

24.1 Introduction: 2 electronic states and a single nuclear coordinate

Electron transfer is a fascinating subject where we pull a lot of what we learned and get unexpected results.

A generic system is three atoms(or molecules) on a line, A—B—C, for example a positviely charged system; and electron transfer proceeding $A^+B + C \to A + BC^+$, which we label states "1" and "2" (so 1 is a state where C is postively ionized, and 2 where A is). (For example, A and C can be transition metals, and B a ligand.) Let's further assume that A and C are locked in place, but B can move along the line connecting them.

The key is that in states 1 and 2 the equlibrium bond-distances will be different. The equilbirum distances in electronic state 1 are schematically :

$$A^+ \qquad - \qquad B \qquad - \qquad - \qquad - \qquad C$$

and in state 2 the equilibrium configuration is

$$A \qquad - \qquad - \qquad - \qquad B \qquad - \qquad C^+.$$

Define a difference-distance

$$x = distance(BC) - distance(AC). \tag{24.1}$$

In eletctronic state 1 the equilbirum atomic configuratin is at positive x,labeled $x_{\dot{1}}$ –positive since the BC distance is larger than the AB distance. (Similarly, for state 2 the equlirbrium will be at a negative x, labeled x_2.)

In electronic state 1 we'll call the energy of the electrons $V_1(x)$; the label is "V" since, as we learned talking about Born Oppenheimer systems, the electronic energy (including the nucleus-nucleus repulsion) is the potential energy which the nuclear coordinates feel.

Now near a minimum, every function looks like a parabola; so the energy when the system is in electronic state "1" is approximated as a parabola in x

$$V_1(x) = \varepsilon_1 + \frac{\kappa}{2}(x - x_1)^2, \tag{24.2}$$

where κ is the force constant, and V_{1eq} the electronic energy of state 1 in its equilibrium configuration. Similarly, if the electronic configuration is "2" then

$$V_2(x) = \varepsilon_2 + \frac{\kappa}{2}(x - x_2)^2. \tag{24.3}$$

Note that we assume that the force constant is the same in both cases, but the potential minima do not have to match when the system is heterogenous.

The Figure below (taken from Schatz and Ratner, Chap. 10) exemplifies the equations (note that $V_1(x)$ and $V_2(x)$ are labeled in the figure as $V_L(x)$ and $V_R(x)$.

In the figure, note, for example, that when the system is in electronic state "1" (i.e., A is in its cation form, A^+) but the position is x_2 (so B is close to C) then the potential is large, i.e., the electronic energy $V_1(x_2)$ is high, much more positive than $V_1(x_1)$,

$$V_1(x_2) > V_2(x_2). \tag{24.4}$$

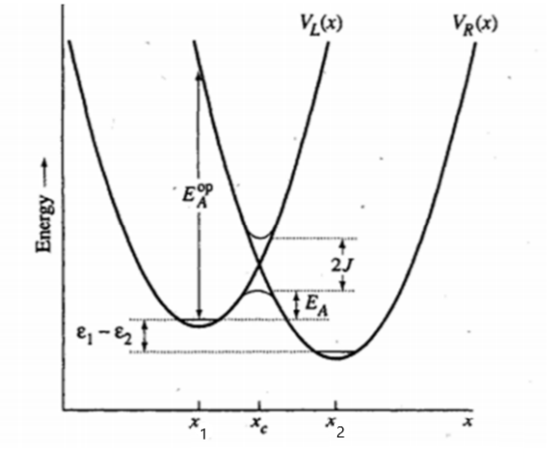
In our previous notation (see the chapter on non-adiabatic dyamics), the $V_1(x)$ and $V_2(x)$ are the **diagonal elements of the diabatic potentials.**

The combined wavefunction has two degrees of freedom associated with it, the electronic state and the nuclear coordinate x, so it is labeled as $\Psi_i(x), \quad i = 1, 2$. The full Hamiltonian contains therefore both the electronic two-state part from and **also a part associated with the** kinetic.

The part of the Hamiltonian associated with the two electronic states (for a fixed nuclear coordinate, x) is just the 2-2 diabatic matrix from Eq. (20.10) which becomes now, **for every given atomic configuration (**i.e., every given x)

$$\begin{pmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{pmatrix} \equiv \begin{pmatrix} V_1(x) & J \\ J & V_2(x) \end{pmatrix} \tag{24.5}$$

(Note that "a" and "b" from before are replaced here by "1" and "2"). Here, J **is the coupling matrix element between the electronic states,** $J \equiv H_{12} = H_{21}$. We'll assume it is distance independent (i.e., independent of x). Further, we'll consider the case the Γ is small, i.e., very little coupling between the electronic states. This is labeled as "non-adiabatic" case.



In addition to this matrix (coupling the electronic states) the Hamiltonian also contains a kinetic energy term, associated with x; this term is $\frac{p^2}{2m}$ where "p" is the mometnum associated with x,and "m" is generally a "reduced-mass". Here it is related to the mas of atom B (there's a factor of 2 we wont discuss). So the total Hamiltoian is then

$$H = \frac{p^2}{2m} + \begin{pmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{pmatrix} = \frac{p^2}{2m} + \begin{pmatrix} V_1(x) & J \\ J & V_2(x) \end{pmatrix}. \qquad (24.6)$$

Let's plug in the expression for $V_1(x), V_2(x)$:

$$H = \frac{p^2}{2m} + \begin{pmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{pmatrix} = \frac{p^2}{2m} + \begin{pmatrix} \frac{1}{2}\kappa(x-x_1)^2 & 0 \\ 0 & \frac{1}{2}\kappa(x-x_2)^2 \end{pmatrix} + \begin{pmatrix} \varepsilon_1 & J \\ J & \varepsilon_2 \end{pmatrix}. \qquad (24.7)$$

Now let's define as usual for harmonic oscillators $\omega \equiv \sqrt{\kappa/m}$, so $\kappa = m\omega^2$;also let's define the equilibrium shift as $x_d = x_2 - x_1$ and shift the coordinate $x \to x - x_1$. Then

$$H = \frac{p^2}{2m} + \frac{m\omega^2}{2} \begin{pmatrix} x^2 & 0 \\ 0 & (x-x_d)^2 \end{pmatrix} + \begin{pmatrix} \varepsilon_1 & J \\ J & \varepsilon_2 \end{pmatrix} \qquad (24.8)$$

This Hamiltonian is almost, but not yet, ready for perturbation theory. That's because in perturbation theory we need to have a **solvable** zero-order Hamiltonian, and a perturbation. Due to the "$x - x_d$" term even the 0-order Hamiltonian (i.e., without Γ, no coupling) is not yet a simple addition of an x-related part and an x-independent electronic state part.

So we'll need an insert to introduce a **transformation** that will do the job:

24.2 Insert: Canonical Transformations

Canonical transformations: general

We explained earlier that in QM only expectation values are measurable. Therefore, we can "rotate" wavefunctions and "rotate" operators such that the expectation value will be the same.

Specfcially, given a wavefunction $|\Psi\rangle$,we can define a "rotated" wavefunction,

$$|\Psi'\rangle = D|\Psi\rangle \qquad (24.9)$$

where D is any unitary operator. The new wavefunction will have the same norm as the old one, since D is unitary (prove!). Note that from the definition of a conjugate operator, the associated bra is

$$\langle\Psi'| = \langle\Psi|D^+. \qquad (24.10)$$

Operators are transformed similarly. Any operator Z will become now

$$Z' = DZD^+ \qquad (24.11)$$

Note that the physical quantity, expectation values, are indeed conserved

$$\langle\Psi'|Z'|\Psi'\rangle = \left(\langle\Psi|D^+\right) DZD^+ \left(D|\Psi\rangle\right) = \langle\Psi|D^+DZDD^+|\Psi\rangle = \langle\Psi|Z|\Psi\rangle, \qquad (24.12)$$

Q.E.D.

Reminiscient of the transformation from Schrödinger to Heisenberg represenations

If the previous discussion looks familiar, there's a good reason for it. We used a particular canonical transformation earlier when we talked about moving from the Schrödinger to the Heisenberg representations. To remind you, first recall that the Schrödinger wavefunction is by definition (assuming a time-independent Hamiltonian): $|\Psi(t)\rangle = e^{-iHt}|\Psi(0)\rangle$. We will define in this case $D = e^{iHt}$. Then

$$|\Psi'(t)\rangle = D|\Psi(t)\rangle = e^{iHt}|\Psi(t)\rangle = e^{iHt}e^{-iHt}|\Psi(0)\rangle = |\Psi(0)\rangle \tag{24.13}$$

i.e., as we recall, the wavefunction does not change in the Heisenberg represetantion; but in the Heisenberg representaiton operatrs do, $Z' = DZD^+ = e^{iHt}Ze^{-iHt}$.

Time-independent D

The subscection above was an insert relating to stuff you learned earlier. Henceforth, we'll only consider canoncial transformation operators (D) that are **time-indepdent.**

If D is time-independent, then the Schrödinger equation for $|\Phi(t)\rangle$ will become a Schrödinger equation for $|\Psi(t)\rangle$,but with a rotated Hamiltonian,

$$H' = DHD^+. \tag{24.14}$$

In other words, conviniently, the original Schrödinger equation is converted to an analogous equation for the rotated funciton:

$$i\frac{\partial|\Psi\rangle}{\partial t} = H|\Psi\rangle \to i\frac{\partial|\Psi'\rangle}{\partial t} = H'|\Psi'\rangle. \tag{24.15}$$

Proof: multiply the original Schrödinger Eq. by D

$$i\frac{\partial|\Psi'\rangle}{\partial t} = i\frac{\partial D|\Psi\rangle}{\partial t} = iD\frac{\partial|\Psi\rangle}{\partial t} = DH|\Psi\rangle = DHD^+D|\Psi\rangle = H'D|\Psi\rangle = H'|\Psi'\rangle \tag{24.16}$$

That's it.

Translation operator

Why would we even want to do a canonical transormation? Since in some cases it can simplify the form of the Hamiltonian, or at least of some parts of it.

To warm-up, let's consider not an electron transfer but a simpler example of a single electronic state (that depends on x). The transformation we will use often is $D = e^{i\alpha p}$ where α is a constant and p is the momentum operator. I claim that this transformation gives a very nice result: it **translates** the wavefunction by a constant:

$$|\Psi\rangle = e^{i\alpha p}|\Phi\rangle \to \Psi(x) = \Phi(x+\alpha) \tag{24.17}$$

Proof (omiting the "ket" symbols):

$$\Psi(x) = e^{i\alpha p}\Phi = e^{i\alpha(-i\frac{d}{dx})}\Phi(x) = e^{\alpha\frac{d}{dx}}\Phi(x) = \\ (1 + \alpha\frac{d}{dx} + \frac{\alpha^2}{2}\frac{d^2}{d^2x} + ...)\Phi(x). \quad (24.18)$$

i.e.,

$$\Psi(x) = \Phi(x) + \alpha\frac{d}{dx}\Phi(x) + \frac{\alpha^2}{2}\frac{d^2\Phi(x)}{d^2x} + ... = \Phi(x+\alpha) \quad (24.19)$$

where the last equality is simply the Taylor expansion theorem. Q.E.D!

So we saw what does $D = e^{i\alpha p}$ does to a wavefunction. Next, we need to do the canonical transformation on operators. Intuitviely, we know the answer: doing a canonical transformation with the translation operator, will modify any operator so that it now refers to a shifted coordinate, i.e., $x \to x + \alpha$,i.e.,

$$x' = x + \alpha \quad \text{(Intuitively)}. \quad (24.20)$$

Let's prove this assertion. First, the definition:

$$x' \equiv DxD^+ = e^{i\alpha p}xe^{-i\alpha p}. \quad (24.21)$$

Next, differentiate the operator x' w.r.t. α (since x' depends on α) :

$$\frac{dx'}{d\alpha} = \frac{de^{i\alpha p}xe^{-i\alpha p}}{d\alpha} = \frac{de^{i\alpha p}}{d\alpha}xe^{-i\alpha p} + e^{i\alpha p}x\frac{de^{-i\alpha p}}{d\alpha} \\ = ie^{i\alpha p}pxe^{-i\alpha p} - ie^{i\alpha p}xpe^{-i\alpha p} = ie^{i\alpha p}[p,x]e^{-i\alpha p} = ie^{i\alpha p}(-i)e^{-i\alpha p} \quad (24.22)$$

i.e.,

$$\frac{dx'}{d\alpha} = 1 \quad (24.23)$$

and since $x'(\alpha = 0) = x$,it follows that

$$x' = x + \alpha. \quad (24.24)$$

Translating a parabolic potential energy (still for 1 electronic state)

Next, let's apply the canonical transformation to a potential energy of the form

$$V = \frac{m\omega^2}{2}(x - x_d)^2. \quad (24.25)$$

After the transformation, the potential will be

$$V' = \frac{m\omega^2}{2}D(x-x_d)^2D^+ = \frac{m\omega^2}{2}D(x-x_d)D^+D(x-x_d)D^+ \\ = \frac{m\omega^2}{2}(x'-x_d)^2 = \frac{m\omega^2}{2}(x+\alpha-x_d)^2 \quad (24.26)$$

where we used

$$D(x-x_d)D^+ = DxD^+ - x_dDD^+ = x' - x_d.$$

So far α was arbitrary, let's choose $\alpha = x_d$. Then

$$V' = \frac{m\omega^2}{2}x^2. \tag{24.27}$$

I.e., the canonical transformation shifted the potential minimum, as expected.

Finally: applying the canonical transformation to the electron-transfer problem.

Recall that for the electron-transfer problem:

$$H = \frac{p^2}{2m}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \frac{m\omega^2}{2}\begin{pmatrix} x^2 & 0 \\ 0 & (x-x_d)^2 \end{pmatrix} + \begin{pmatrix} \varepsilon_1 & 0 \\ 0 & \varepsilon_2 \end{pmatrix} + \begin{pmatrix} 0 & J \\ J & 0 \end{pmatrix} \tag{24.28}$$

where we now split the coupling term, since it is the "perturbation" that couples the two electronic state. Further, the x-kinetic energy term $p^2/2m$, is multiplied by the identity matrix as far as the electronic states, i.e.,

$$\frac{p^2}{2m} \equiv \frac{p^2}{2m}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{24.29}$$

We now need to shift the potential minimum so that it is the same for both states. Therefore, D should be now a 2*2 operator-matrix

$$D = \begin{pmatrix} 1 & 0 \\ 0 & e^{ipx_d} \end{pmatrix} \tag{24.30}$$

so

$$D^+ = \begin{pmatrix} 1 & 0 \\ 0 & e^{-ipx_d} \end{pmatrix}$$

and recall that here p is an operator and x_d is a constant.

Side-note: this transformation matrix (and the associated math) is called in the literature **the polaron transformation,** since it depends on the vibrational shift associated with changing electronic state, which is essentially what a polaron it.

To continude Let's apply:

$$H' = \frac{p^2}{2m}D\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}D^+ + \frac{m\omega^2}{2}D\begin{pmatrix} x^2 & 0 \\ 0 & (x-x_d)^2 \end{pmatrix}D^+ + D\begin{pmatrix} \varepsilon_1 & 0 \\ 0 & \varepsilon_2 \end{pmatrix}D^+ + D\begin{pmatrix} 0 & J \\ J & 0 \end{pmatrix}D^+ \tag{24.31}$$

i.e., since D and D^+ are diagonal, then two of the four terms are unchanged,

$$H' = \frac{p^2}{2m}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \frac{m\omega^2}{2}\begin{pmatrix} 1 * x^2 * 1 & 0 \\ 0 & e^{ipx_d}(x-x_d)^2 e^{-ipx_d} \end{pmatrix} \\ + \begin{pmatrix} \varepsilon_1 & 0 \\ 0 & \varepsilon_2 \end{pmatrix} + D\begin{pmatrix} 0 & J \\ J & 0 \end{pmatrix} D^+ \tag{24.32}$$

Now by construction $e^{ipx_d}(x-x_d)^2 e^{-ipx_d} = x^2$. We're left with the coupling term:

$$D\begin{pmatrix} 0 & \Gamma \\ \Gamma & 0 \end{pmatrix} D^+ = \begin{pmatrix} 1 & 0 \\ 0 & e^{ipx_d} \end{pmatrix}\begin{pmatrix} 0 & J \\ J & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & e^{-ipx_d} \end{pmatrix} \\ = \begin{pmatrix} 0 & J \\ Je^{ipx_d} & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & e^{-ipx_d} \end{pmatrix} = \begin{pmatrix} 0 & Je^{-ipx_d} \\ Je^{ipx_d} & 0 \end{pmatrix}, \tag{24.33}$$

and therefore the shifted Hamiltonian is now:

$$H' = \frac{p^2}{2m}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \frac{m\omega^2}{2}\begin{pmatrix} x^2 & 0 \\ 0 & x^2 \end{pmatrix} + \begin{pmatrix} \varepsilon_1 & 0 \\ 0 & \varepsilon_2 \end{pmatrix} \\ + \begin{pmatrix} 0 & Je^{-ipx_d} \\ Je^{ipx_d} & 0 \end{pmatrix} \tag{24.34}$$

which gives the final form, with the rotated Hamiltonian being a sum of separable and perturbation terms:

$$H' = H_0 + JH_1, \tag{24.35}$$

where the 0-order Hamiltonian is a combination of a vibrational part and two-level constant diagonal part

$$H_0 = H_v + \begin{pmatrix} \varepsilon_1 & 0 \\ 0 & \varepsilon_2 \end{pmatrix} \tag{24.36}$$

$$H_v = \frac{p^2}{2m} + \frac{m\omega^2 x^2}{2}, \tag{24.37}$$

while the perturbation operator (multplied by J, the coupling strength) is

$$H_1 = \begin{pmatrix} 0 & e^{-ipx_d} \\ e^{ipx_d} & 0 \end{pmatrix}. \tag{24.38}$$

In this form, we can finally apply perturbation theory.

24.3 Perturbation theory applied to electron-transfer.

Eqs. (22.40),(22.41), (22.45) were the highlight of perturbation theory as dervied in an earlier chapter. We can adopt them here, with the role of b (the strength of the perturbation) taken by J, the coupling matrix element, and F and G, the initial and final states, replaced by "1" and "2", and "H" replaced by H_0

$$\boxed{\Gamma_{2\leftarrow 1} = \frac{|J|^2}{Q}\int_{-\infty}^{\infty} e^{i\omega t}C(t)dt} \tag{24.39}$$

and μ, the coupling matrix element, replaced by $S \equiv \begin{pmatrix} 0 & 0 \\ e^{ipx_d} & 0 \end{pmatrix}$, the exchage operator coupling the states (and its conjugate):

$$C(t) \equiv \mathrm{Tr}\left(S(t)S^{+}e^{-\beta H_0}\right) = \mathrm{Tr}\left(e^{iH_0 t}Se^{-iH_0 t}S^{+}e^{-\beta H_0}\right) \tag{24.40}$$

i.e.,

$$C(t) = \mathrm{Tr}\left(e^{iH_v t}e^{ipx_d}e^{-iH_v t}e^{-ipx_d}e^{-\beta H_v}\right)e^{i(\varepsilon_1-\varepsilon_2)t}. \tag{24.41}$$

Note that the electronic state part comes out of the expression; this is because in the trace in the previous equation, the $e^{iH_0 t}$ operator acts on the starting state, "1", so the electronic state part gives $e^{i\varepsilon_1 t}$ and due to the exchange operator S acting in front of $e^{-iH_0 t}$,this exponential refers to the ending state, "2" so a factor of $e^{i(\varepsilon_1-\varepsilon_2)t}$ is picked up. This would have been apparent if we did a more careful derivation, but there's so much math already so I skipped a step.

> Note: S has nothing to do with "S-matrix", we dont have enough letters.

The trace in the last equation is written solely in terms of a Harmonic oscillator. I'll just start the derivation it up to you to work out (with guidance) some part of this trace, and just quote the final result.

The math derivation starts from

$$H_v = \omega\left(b^{+}b + \frac{1}{2}\right) \tag{24.42}$$

where

25 Path Integrals and Green's functions

Short time propagator

We'll now do a break from "traditional chemical physics" to a very important topic that grew even more important in chemical physics in the last two decades, as we'll see later.

To understand path-integrals, consider the time-dependent Schrodinger equation, for a general particle (for now a single 3-dimensional of mass "m", but it could be a wavefunction of N particles, i.e., a function of 3N coordinates):

$$\frac{\partial|\psi(t)\rangle}{\partial t} = -iH(t)|\psi(t)\rangle \tag{25.1}$$

where we'll use later $H(t) = \frac{p^2}{2M} + V(\boldsymbol{r},t)$, i.e., a kinetic plus a potential (which can be time-dependent in the general case). Let's **discretize** the Schrödinger equation as

$$\frac{\partial|\psi(t)\rangle}{\partial t} \simeq \frac{|\psi(t+dt)\rangle - |\psi(t)\rangle}{dt} \tag{25.2}$$

(where "dt" is assumed tiny but fixed, i.e., not smaller and smaller). So putting together the last two eqs. gives:

$$|\psi(t+dt)\rangle - |\psi(t)\rangle = -iH(t)\cdot dt\cdot|\psi(t)\rangle \tag{25.3}$$

i.e.,

$$|\psi(t+dt)\rangle = \left(1 - iH(t)\cdot dt\right)|\psi(t)\rangle \tag{25.4}$$

which we rewrite to first order in dt (valid since dt is small) as

$$|\psi(t+dt)\rangle = e^{-iH(t)dt}|\psi(t)\rangle \tag{25.5}$$

Note that we already saw this relation when we talked about time-independent Hamiltonian, when H was not dependent on time, but it is valid in general.

There are two intertwined ways to proceed from this simple yet fundamental relation: **path integrals and Green's functions.** We'll derive first Green's functions.

Recall the completeness relation

$$\int |\boldsymbol{r}_0\rangle\langle\boldsymbol{r}_0|d\boldsymbol{r}_0 = 1 \tag{25.6}$$

where $\boldsymbol{r}_0$ is a 3-dimensional coordinate, or a 1-dimensional one, or even a 3N dimensional one. Insert it to the equation above, to get

$$|\psi(t+dt)\rangle = \int e^{-iH(t)dt}|\boldsymbol{r}_0\rangle\langle\boldsymbol{r}_0|\psi(t)\rangle d\boldsymbol{r}_0 \tag{25.7}$$

Not surprisingly, we'll now dot product with a bra of a general position, that will be labeled $\langle\boldsymbol{r}_1|$. Then:

$$\langle \boldsymbol{r}_1|\psi(t+dt)\rangle = \int \langle \boldsymbol{r}_1|e^{-iH(t)dt}|\boldsymbol{r}_0\rangle\langle \boldsymbol{r}_0|\psi(t)\rangle d\boldsymbol{r}_0. \tag{25.8}$$

By writing $\psi(\boldsymbol{r}_0, t) \equiv \langle \boldsymbol{r}_0|\psi(t)\rangle$ we get the **general relation relating the value of a wavefunction in a general grid point after time dt to its value at the starting time**

$$\boxed{\psi(\boldsymbol{r}_1, t+dt) = \int G(\boldsymbol{r}_1, t+dt; \boldsymbol{r}_0, t_0)\psi(\boldsymbol{r}_0, t_0)d\boldsymbol{r}_0,} \tag{25.9}$$

where we introduced the **time-dependent Green's function,** i.e., the "Kernel" which relates the wavefunction at later times and a different position to all the values of the wavefunction at the earlier time. **You can think of this last equation as the definition of the Green's function**. (Its value of course is: $G\left(\boldsymbol{r}_{\boldsymbol{1}}, t+dt, \boldsymbol{r}_0, t\right) = \langle \boldsymbol{r}_1|e^{-iH(t)dt}|\boldsymbol{r}_0\rangle$ which we'll evaluate later.)

Time-dependent Green's function in general

Let's digress a little to discuss the time-dependent Green's function. **In general, i.e., not only for a short time difference, the Green's function is defined as the Kernel relating the value of the function at one time point to another at an earlier time, i.e., in general**

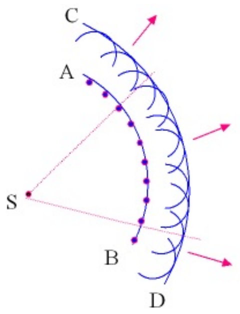
$$\boxed{\psi(\boldsymbol{r}_1, t_1) = \int G(\boldsymbol{r}_1, t_1; \boldsymbol{r}_0, t_0)\psi(\boldsymbol{r}_0, t_0)d\boldsymbol{r}_0.} \tag{25.10}$$

A **(time-dependent) Green's function $G(\boldsymbol{r}_{\boldsymbol{1}}, t_1, \boldsymbol{r}_{\boldsymbol{0}}, t_0)$ can be thought of as: what will be the value of the wavefunction at a time t_1 and a position $\boldsymbol{r}_1$ if at time t_0 it is a delta-function localized at $\boldsymbol{r}_0$.**

Proof: if $\psi(\boldsymbol{r}, t_0) = \delta(\boldsymbol{r} - \boldsymbol{r}_1)$, then (we now use $\boldsymbol{r}$ as dummy index):

$$\psi(\boldsymbol{r}_1, t_1) = \int G(\boldsymbol{r}_1, t_1; r, t_0)\psi(\boldsymbol{r}, t_0)d\boldsymbol{r} = \int G(\boldsymbol{r}_1, t_1; r, t_0)\delta(\boldsymbol{r} - \boldsymbol{r}_0)d\boldsymbol{r} = G(\boldsymbol{r}_1, t_1; \boldsymbol{r}_0, t_0). \tag{25.11}$$

Another interesting property expresses mathematically Huygens principle, which states that any given (intermediate) moment we can imagine that a "sends" amplitude forward from any given point.



Huygens' Principle:
Each wavefront is the envelope of the wavelets. Each point on a wavefront acts as an independent source to generate wavelets for the next wavefront. AB and CD are two wavefronts.

Taken from:
http://www.pstcc.edu/departments/natural_behavioral_sciences
/Web%20Physics/Chapter038.htm

Mathematically:

$$\boxed{G(\boldsymbol{r}_2, t_2, \boldsymbol{r}_0, t_0) = \int G(\boldsymbol{r}_2, t_2, \boldsymbol{r}_1, t_1) G(\boldsymbol{r}_1, t_1; \boldsymbol{r}_0, t_0) d\boldsymbol{r}_1} \tag{25.12}$$

i.e., the Green's function is an accumulation of propagation to an intermediate time step, and propagating from that time-step to a later time. See more pictures in class.

Proof: Let's assume again that $\psi\left(\boldsymbol{r}, t_0\right) = \delta(\boldsymbol{r} - \boldsymbol{r}_0)$,and insert to the defs. (we replace again r0 by r as the dummy variable):

$$\psi\left(\boldsymbol{r}_2, t_2\right) = \int G(\boldsymbol{r}_2, t_2; \boldsymbol{r}, t_0)\psi(\boldsymbol{r}, t_0) d\,\boldsymbol{r} = \int G(\boldsymbol{r}_2, t_2; \boldsymbol{r}, t_0)\delta(\boldsymbol{r} - \boldsymbol{r}_0) d\boldsymbol{r} = G(\boldsymbol{r}_2, t_2; \boldsymbol{r}_0, t_0) \tag{25.13}$$

And similarly,

$$\psi\left(\boldsymbol{r}_1, t_1\right) = G(\boldsymbol{r}_1, t_1; \boldsymbol{r}_0, t_0). \tag{25.14}$$

But we know that, by definition,

$$\psi\left(\boldsymbol{r}_2, t_2\right) = \int G(\boldsymbol{r}_2, t_2; \boldsymbol{r}_1, t_1)\psi(\boldsymbol{r}_1, t_1) d\boldsymbol{r}_1. \tag{25.15}$$

so inserting the last equation gives

$$\psi\left(\boldsymbol{r}_2, t_2\right) = \int G(\boldsymbol{r}_2, t_2; \boldsymbol{r}_1, t_1) G(\boldsymbol{r}_1, t_1; \boldsymbol{r}_0, t_0) d\boldsymbol{r}_1 \tag{25.16}$$

The Magenta-colored terms all match, which proves the Huygens principle.

Back to Short-time propagator: toward path integrals.

As we saw earlier, the value of the Green's function is:

$$G(\boldsymbol{r_1}, t+dt; \boldsymbol{r}_0, t) = \langle \boldsymbol{r}_1 | e^{-iH(t)dt} | \boldsymbol{r}_0 \rangle. \tag{25.17}$$

Let's evaluate it. We know that (we suppress the "t" dependence of V and H):

$$e^{-iHdt} = e^{-i\left(\frac{p^2}{2M}+V\right)dt} \simeq e^{-iV\frac{dt}{2}} e^{-i\frac{p^2}{2M}dt} e^{-iV\frac{dt}{2}} + O(dt^3) \tag{25.18}$$

where the meaning of this last $O(dt^3)$ is that we can break the exponential to separate exponentials, and the error order will be much tinier than linear in dt (i.e., if we make dt smaller by 10, the error in the expression above goes down by 1000). We saw similar relations in 115a/215a, and if you don't believe me expand both sides to 2nd order and match the 0th, 1st and 2nd order term one by one.

> BTW, the **symmetric splitting** which we did, (where we did an exponential of $-iVdt/2$ to the left and the right of the kinetic-energy exponential) leads to an excellent third-order-error expression; if we did a non-symmetric splitting the order would not be so tiny, i.e., $e^{-i\left(\frac{p^2}{2M}+V\right)dt} \simeq e^{-i\frac{p^2}{2M}dt} e^{-iVdt} + O(dt^2)$

Inserting we get

$$\begin{aligned} G(\boldsymbol{r}_1, t+dt; \boldsymbol{r}_0, t) &= \langle \boldsymbol{r}_1 | e^{-iV\frac{dt}{2}} e^{-i\frac{p^2}{2M}dt} e^{-iV\frac{dt}{2}} | \boldsymbol{r}_0 \rangle = \\ &\langle \boldsymbol{r}_1 | e^{-iV(\boldsymbol{r}_1)\frac{dt}{2}} e^{-i\frac{p^2}{2M}dt} e^{-iV(\boldsymbol{r}_0)\frac{dt}{2}} | \boldsymbol{r}_0 \rangle = \\ &e^{-i\frac{V(\boldsymbol{r}_1)+V(\boldsymbol{r}_0)}{2}dt} \langle \boldsymbol{r}_1 | e^{-i\frac{p^2}{2M}dt} | \boldsymbol{r}_0 \rangle. \end{aligned} \tag{25.19}$$

Let's calculate the generic matrix element by introducing a resolution of identity: $\int |\boldsymbol{p}\rangle\langle\boldsymbol{p}| d\boldsymbol{p}$ (note: p here is a dummy index showing the **eigenvalue** of the momentum operator; in the equation above $p^2/2M$ indicated the momentum **operator** squared):

$$\begin{aligned} \langle \boldsymbol{r}_1 | e^{-i\frac{p^2}{2M}dt} | \boldsymbol{r}_0 \rangle &= \int \langle \boldsymbol{r}_1 | \boldsymbol{p} \rangle \langle \boldsymbol{p} | e^{-i\frac{p^2}{2M}dt} | \boldsymbol{r}_0 \rangle d\boldsymbol{p} \\ &= \int e^{-i\frac{p^2}{2M}dt} \langle \boldsymbol{r}_1 | \boldsymbol{p} \rangle \langle \boldsymbol{p} | \boldsymbol{r}_0 \rangle d\boldsymbol{p} = \frac{1}{\left((2\pi)^{3/2}\right)^2} \int e^{i\boldsymbol{p}\cdot\boldsymbol{r}_1} e^{-i\frac{p^2}{2M}dt} e^{-i\boldsymbol{p}\cdot\boldsymbol{r}_0} d\boldsymbol{p} \\ &= \frac{1}{(2\pi)^3} \int e^{i\boldsymbol{p}\cdot(\boldsymbol{r}_1-\boldsymbol{r}_0)} e^{-i\frac{p^2}{2M}dt} d\boldsymbol{p} \end{aligned} \tag{25.20}$$

But we learned before about the Gaussian integral (all integrals go from minus infinity to infinity), which in 3D reads:

$$\int e^{-\frac{a\boldsymbol{y}^2}{2}} e^{i\boldsymbol{b}\cdot\boldsymbol{y}} d\boldsymbol{y} = \left(\frac{2\pi}{a}\right)^{3/2} e^{-\frac{b^2}{2a}}. \tag{25.21}$$

We can then apply it to our formula, setting $\boldsymbol{b} = (\boldsymbol{r}_{t+dt} - \boldsymbol{r}_t)$,$a = \frac{idt}{M}$, and $\boldsymbol{y} = \boldsymbol{p}$, to get (using $-\frac{1}{i} = i$):

$$\langle \boldsymbol{r}_1 | e^{-i\frac{p^2}{2M}dt} | \boldsymbol{r}_0 \rangle = \frac{1}{(2\pi)^3} \left(\frac{2\pi}{a}\right)^{3/2} e^{-\frac{b^2}{2a}} = Ce^{i\frac{M(\boldsymbol{r}_1 - \boldsymbol{r}_0)^2}{2\,dt}} \tag{25.22}$$

where

$$C \equiv \left(\frac{iM}{2\pi dt}\right)^{3/2} \tag{25.23}$$

> You may worry that there's a mathematical problem in applying the Gaussian-integral formula due to the "i" in $e^{-i\frac{p^2}{2M}dt}$, which means that for large p the exponential doesn't fall to 0 but just oscillates faster and faster,; luckily it all works out.

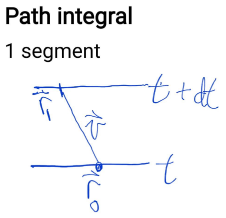
Together we get

$$G(\boldsymbol{r}_1, t+dt; \boldsymbol{r}_0, t) = \langle \boldsymbol{r}_1 | e^{-iH(t)dt} | \boldsymbol{r}_0 \rangle = Ce^{-i\frac{V(\boldsymbol{r}_1)+V(\boldsymbol{r}_0)}{2}dt} e^{i\frac{M(\boldsymbol{r}_1-\boldsymbol{r}_0)^2}{2\,dt}}$$
$$= Ce^{i\left(\frac{V(\boldsymbol{r}_1)+V(\boldsymbol{r}_0)}{2} - \frac{M(\boldsymbol{r}_1-\boldsymbol{r}_0)^2}{2\,dt^2}\right)dt} \tag{25.24}$$

where in the last term we divided and multiplied by dt.

Further, let's think of $(\boldsymbol{r}_1, t+dt), (\boldsymbol{r}_0, t)$ as a single **segment** (see **Figure**); then the **velocity in the segment** will be

$$v(t \to t+dt) \simeq \frac{\boldsymbol{r}_1 - \boldsymbol{r}_0}{dt}$$



Therefore, we write

$$\langle \boldsymbol{r}_1 | e^{-iH(t)dt} | \boldsymbol{r}_0 \rangle = Ce^{iS(t \to t+dt)} \quad (25.25)$$

where we defined the action of the segment as

$$S(t \to t+dt) = \left(\left(\frac{V(\boldsymbol{r}_1) + V(\boldsymbol{r}_0)}{2} \right) - \frac{M}{2} v^2 \right) dt \tag{25.26}$$

i.e., it is essentially the **average potential along the segment minus the kinetic energy**. Those of you who took analytical mechanics will recognize it as the **Lagrangian** of the segment times dt, which is **the action of the segment.** We'll see more soon.

Path Integrals

Let's now use the formula we learned to break a **finite-time** Green's function to a product of many intermediate ones. Specifically, assume we start at a time "t_0" and end up at a time "t'" and break the time-difference into N segments dt such that $t' - t = N \cdot dt$, i.e., define

$$t_0 = t, \; t_1 = t + dt, \; t_2 = t + 2dt, ...t_N = t' = t + N \cdot dt \tag{25.27}$$

Let's further assume for simplicity that the potential is time-independent; the formulae trivially generalize to the the time-dependent case, as we'll present later.

Then, the Green's function between two space-time points, label $\boldsymbol{r}_0, t_0$ and $\boldsymbol{r'}, t'$, becomes

$$\begin{aligned} G(\boldsymbol{r}', t'; \boldsymbol{r}_0, t_0) = \langle \boldsymbol{r'} | e^{-iH(t'-t_0)} | \boldsymbol{r}_0 \rangle &= \langle \boldsymbol{r}_N | e^{-iH(N \cdot dt)} | \boldsymbol{r}_0 \rangle \\ &= \langle \boldsymbol{r'} | e^{-iHdt} e^{-iHdt} ... e^{-iHdt} | \boldsymbol{r}_0 \rangle. \end{aligned} \tag{25.28}$$

Now comes the crucial part. Insert $N-1$ times the resolution of identity, $1 = \int |\boldsymbol{r}\rangle\langle \boldsymbol{r}| d\boldsymbol{r}$, into the expression above (each time labeling it differently, OK since it is a dummy index). Further, label the position $\boldsymbol{r}'$ as $\boldsymbol{r}_N$:

$$\begin{aligned} &\langle \boldsymbol{r}_N | e^{-iH(t'-t)} | \boldsymbol{r}_0 \rangle = \\ &\int \int ... \int \langle \boldsymbol{r}_N | e^{-iHdt} | \boldsymbol{r}_{N-1} \rangle \langle \boldsymbol{r}_{N-1} | e^{-iHdt} | \boldsymbol{r}_{N-2} \rangle \langle \boldsymbol{r}_{N-2} | ... | \boldsymbol{r}_1 \rangle \langle \boldsymbol{r}_1 | e^{-iHdt} | \boldsymbol{r}_0 \rangle d\boldsymbol{r}_1 \, d\boldsymbol{r}_2 ... d\boldsymbol{r}_{N-1} \end{aligned} \tag{25.29}$$

Note: we can think of the relation above as nothing more than a multi-step representation of the Huygens principle, i.e.,

$$\begin{aligned} &G(\boldsymbol{r}_N, t_N; \boldsymbol{r}_0, t_0) = \\ &\int \int ... \int G(\boldsymbol{r}_N, t_N; \boldsymbol{r}_{N-1}, t_{N-1}) G(\boldsymbol{r}_{N-1}, t_{N-1}, \boldsymbol{r}_{N-2}, t_{N-2}) .... G(\boldsymbol{r}_1, t_1; \boldsymbol{r}_0, t_0) d\boldsymbol{r}_1 \, d\boldsymbol{r}_2 ... d\boldsymbol{r}_{N-1} \end{aligned} \tag{25.30}$$

End of note.

The expression above the note can be reworked using the single time-step relation we proved above (the Green-color equation) to be

$$\langle \boldsymbol{r}_N | e^{-iHT} | \boldsymbol{r}_0 \rangle = C^{N-1} \int \int ... \int e^{i\boldsymbol{S}} d\boldsymbol{r}_1\, d\boldsymbol{r}_2 ... d\boldsymbol{r}_{N-1} \tag{25.31}$$

where we defined the **action** of the trajectory as

$$\begin{aligned} S = &\left(\frac{V(\boldsymbol{r}_0)}{2} + V(\boldsymbol{r}_1) + ... + V(\boldsymbol{r}_{N-1}) + \frac{V(\boldsymbol{r}_N)}{2} \right) dt \\ &- \frac{M}{2} \left(\left(\frac{\boldsymbol{r}_N - \boldsymbol{r}_{N-1}}{dt} \right)^2 + \left(\frac{\boldsymbol{r}_{N-1} - \boldsymbol{r}_{N-2}}{dt} \right)^2 + ... + \left(\frac{\boldsymbol{r}_1 - \boldsymbol{r}_0}{dt} \right)^2 \right) dt \end{aligned} \tag{25.32}$$

Now let's think of the whole set $(\boldsymbol{r}_0, t_0), (\boldsymbol{r}_1, t_1), (\boldsymbol{r}_2, t_2) ..., (\boldsymbol{r}_{N-1}, t_{N-1}), (\boldsymbol{r}_N, t_N)$ as a **(multi-segment) trajectory,** see the attached **figure.** Then the **velocity of each segment is simply,** as we saw, the position difference over dt, $\boldsymbol{v}_{t_{j-1}\to t_j} = (\boldsymbol{r}_j - \boldsymbol{r}_{j-1})/dt$, so

$$\begin{aligned} S = &\left(\frac{V(\boldsymbol{r}_0)}{2} + V(\boldsymbol{r}_1) + ... + V(\boldsymbol{r}_{N-1}) + \frac{V(\boldsymbol{r}_N)}{2} \right) dt \\ &- \frac{M}{2} \left(\boldsymbol{v}^2_{t_0\to t_1} + \boldsymbol{v}^2_{t_1\to t_2} + ... + \boldsymbol{v}^2_{t_{N-1}\to t_N} \right) dt \end{aligned} \tag{25.33}$$

i.e., it is made from summation (i.e., integration) over potential terms minus a summation (integration)of the kinetic energy.

Formally,the "**action**" S is then the **integral over time** of this multi-segment expression

$$S \simeq \int_{t_0}^{t_N} (V - K)\, dt = \int_t^{t+T} L dt \tag{25.34}$$

where the Lagrangian, as we mentioned, is defined here as the potential minus the kinetic energy (in other cases it is more complicated, but the relation $S = \int_t^{t+T} L dt$ is always valid).

Note: that the Lagrangian and therefore the action depends on the trajectory; we write it as

$$L = L(\boldsymbol{r}(t), \dot{\boldsymbol{r}}(t)) = V(\boldsymbol{r}(t)) - K(\dot{\boldsymbol{r}}(t)) = V(\boldsymbol{r}(t)) - \frac{M}{2} \dot{\boldsymbol{r}}(t)^2 \tag{25.35}$$

We end up with a very enticing expression:

$$\boxed{G(\boldsymbol{r}_N, t_N; \boldsymbol{r}_0, t_0) = C' \int e^{i\frac{S}{\hbar}} D\boldsymbol{r}} \tag{25.36}$$

where we reinserted $\hbar$, and where $D\boldsymbol{r}$ is an abbreviation for an integral over the **trajectory** that starts at $\boldsymbol{r}_0, t_0$ and ends at $\boldsymbol{r}$;t', i.e.,

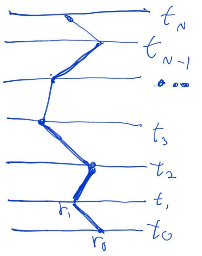
$$D\,\boldsymbol{r} \equiv d\boldsymbol{r}_1\, d\boldsymbol{r}_2 ... d\boldsymbol{r}_{N-1} \tag{25.37}$$

Thus, the **quantum mechanical transition amplitude from one point $\boldsymbol{r}_0$ at a starting time to a final point $\boldsymbol{r}'$ at a later time $\boldsymbol{t}'$** is the **sum of the exponentials of the action of all possible trajectories** between these points.

> Note: you may be unhappy that we considered non-smooth multi-segment trajectories; the theory is as easily written in terms of smooth trajectories, but we wont dwell on it.

The important thing to note is that this is the sum over **all possible trajectories,** that in general will not **not fulfill Newton's relation** (will not conserve energy even if the potential is time-independent, etc.).

Multi-segment
path-integral:



Time-dependent Potentials

Our derivation was for a time-independent potential, but it is valid in general, i.e., the potential appearing in the Lagrangian can depend on time explicitly (e.g., as in lasers). We then write:

$$L = V\left(\boldsymbol{r}(t), t\right) - \frac{M}{2}\dot{\boldsymbol{r}}^2 \tag{25.38}$$

Classical mechanics and path integrals

Unfortunately it is impossible to numerically sum the exponential of the action of each trajectory, due to the "i" in the exponent. In fact, the more refined time step we become, the worse the oscillations become!

Thus these so-called “real-time” path integrals (we’ll explain the name below) have become in practice primarily a tool to build approximations from.

The simplest such approximation is classical mechanics of course, and its “semi-classical” extensions.

Specifically, it is possible to show (see an analytical mechanics course) that **a classical mechanics trajectory** (i.e., fulfilling Newton’s equation) that starts and ends at the required points will **extremize** the action. Therefore, the contribution of **trajectories that are close to a classical trajectory** will **add up in phase.** Therefore, in many cases **classical trajectories dominate the Green’s function**. This is the reason that the limit of quantum mechanics is classical mechanics. The field explaining (and building on) this is called **semi-classical quantum mechanics.**

I don’t want to spend too much time on this, so I just show you some figures to help understanding this concepts. But this is a very rich field (that many a brilliant scientist sank their career in...) so if it interests you I’ll be happy to give some references.

Temperature-Based Path Integrals

Interestingly (and very importantly) the path integral concept carries to thermodynamics. You may have seen it is you took/are taking statistical mechanics, but the more times you see the merrier.

Specifically, the temperature Green’s function is defined as

$$G(\boldsymbol{r}', \boldsymbol{r}_0; \beta) = \langle \boldsymbol{r}' | e^{-\beta H} | \boldsymbol{r}_0 \rangle \tag{25.39}$$

where $\beta = 1/(k_B T)$. Thus, it looks just like the quantum mechanics expression, except that we replaced it by β.

> This is often called **imaginary time expression, since** $e^{-\beta H} = e^{-i\tau H}$**, if we define** $\bar{t} = i\beta$**, i.e., this “time”** $\bar{t}$ **is imaginary.** (And therefore the expressions we had earlier are called real-time expressions.

We can repeat the math from the previous section and we’ll end up with a very similar (but numerically much better!) expression,

$$\boxed{G(\boldsymbol{r}', \boldsymbol{r}_0; \beta) = C' \int e^{-S_E} D\boldsymbol{r}} \tag{25.40}$$

where S_E is the **Euclidean action**,

$$S_E = \int_0^\beta \left(V(\boldsymbol{r}(\tau)) + \frac{1}{2M} \left(\frac{d\boldsymbol{r}}{d\tau} \right)^2 \right) d\tau \tag{25.41}$$

i.e., the Euclidean action is the integral of the local kinetic energy plus local potential energy, i.e., the local total energy, when averaged over the full trajectory $\boldsymbol{r}(\tau)$ where τ is an intermediate “time” that goes from 0 to β.

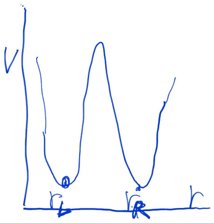
Numerically, since all contributions are positive, the “boxed” equation above is very convenient and can be sampled by Monte-Carlo techniques.

> Note that the reason the kinetic energy has the same sign as the potential energy in the Euclidean action can be traced back to the Gaussian integral, which is the analog of Eq. (25.21), but without the i, i.e., the matrix element of $\langle \boldsymbol{r}_{t+dt}|e^{-\frac{p^2}{2M}d\tau}|\boldsymbol{r}_t\rangle$.

Qualitative example for a matrix element with the Euclidean action: tunneling through a barrier.

Let’s consider a simple exam, tunneling through a barrier in a double-well system (see figure). This could represent the potential for proton transfer, or organic reactions, etc.

Double well



Let’s consider a “transfer” matrix element of a generic form

$$G(\boldsymbol{r}_{\boldsymbol{R}},\boldsymbol{r}_L;\beta)=\langle \boldsymbol{r}_R|e^{-\beta H}|\boldsymbol{r}_L\rangle = C'\int e^{-S_E}D\boldsymbol{r} \tag{25.42}$$

where $\boldsymbol{r}_L, \boldsymbol{r}_R$ are generic positions on the right and left well. This matrix element is a sum over all trajectories that start at the left well position and end up at the right well position (and therefore transfer over the barrier at least once).

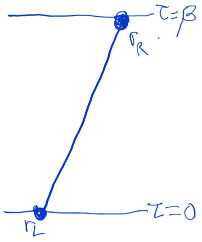
Let’s see qualitatively which trajectories will have the least action, i.e., the least average energy.

- The least “kinetic energy” trajectory is one that over “β” transfers, at the same pace, from left to right (see Figure XXX) i.e., a constant “velocity” trajectory.

  $$\boldsymbol{r}_{\text{least KE}}(\tau)=\boldsymbol{r}_L+\frac{\tau}{\beta}\left(\boldsymbol{r}_{\boldsymbol{R}}-\boldsymbol{r}_L\right) \tag{25.43}$$

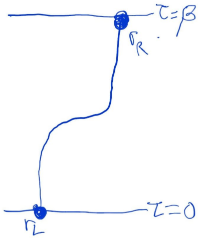
  Note that $\boldsymbol{r}_{\text{least KE}}(\tau=0)=\boldsymbol{r}_L$, and $\boldsymbol{r}_{\text{least KE}}(\tau=\beta)=\boldsymbol{r}_R$, as required. The problem with the least-kinetic-energy trajectory is that the particle spends too-long-a-time in the middle near the barrier peak, where the potential energy is very high

Least-KE
trajectory



- It is possible to get much lower action if we take a trajectory where the particles **spend a lot of time near the left and right well, and quickly transfer from one well to the other in a short time** (see figure below). The particles will have a higher kinetic energy then the "constant-velocity" trajectory, but this is well compensated by the reduction of time that the particle spends in the high-potential region.

Least-action
trajectory
(least time
in barrier)



26 Some Interesting topics:

Here's a list of some interesting topics, which if time permits we'll discuss one or two of them, but you should pursue yourself otherwise:

- Ring-polymer molecular dynamics
- Electron transfer
- Time-dependent Hartree-Fock
- Density matrices, Redfield equations (Lindblad approach)
- Green's functions, GW, Bethe-Salpeter
- Raman Scattering and correlation functions
- Scattering

Conclusions

Overall, research in quantum chemistry is active and there is continuous progress on many of these methods. As new computational tools and techniques are developed, the balance shifts among the different approaches, and I urge you to study and keep up with the developed methodology, and hopefully you can also contribute to it one day!