Quantum Chemistry: Fundamentals and Methods

Chem 115B/215B Lecture Notes

Thomas P Fay

University of California, Los Angeles
thomaspfay@ucla.edu

Contents

Introduction

These lecture notes accompany the Chem 115B/215B Graduate Quantum Chemistry course at UCLA. This course builds on the previous 115A/215A course that introduced the foundations of quantum mechanics and its applications to a range of simple analytically soluble problems. This course focusses on more advanced topics and problems that cannot be solved exactly with analytical techniques, which almost all quantum problems in chemistry are, but instead have to be solved approximately. The first part of this course focusses on time-dependent quantum mechanics and its applications, particularly an introduction to rates of quantum state transitions, density operators and theoretical foundations of spectroscopy. The second part focusses on the electronic structure of atoms and molecules, including the second quantisation for many-electron systems and density functional theory, and concludes with setting out the framework for treating full coupled electron-nuclear systems.

Assumed knowledge

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

- Basics of (finite-dimensional) vectors and matrices. Matrix-vector multiplication, determinants, inverses, eigenvalues and eigenvectors.
- Multivariable calculus, partial differentiation, changes of variables, integration, solving simple differential equations.
- The time independent and time-dependent Schrödinger equation.
- Position-space wave functions, position and momentum operators.
- Wave functions for systems including angular momentum, harmonic oscillators and the hydrogenic atom.
- Time-independent perturbation theory and the variational approach to quantum mechanics.

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

Accompanying Textbooks

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

Acknowledgements

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

1 Time-dependent Perturbation Theory

1.1 Time-Dependent Hamiltonians

Suppose we have a system described by a Hamiltonian which depends on time $\hat{H}(t)$. How do we solve the TDSE for this?

$$i\hbar\frac{\mathrm{d}}{\mathrm{d}t}\left|\Psi(t)\right\rangle = \hat{H}(t)\left|\Psi(t)\right\rangle$$

We formally integrate this (we'll assume $t \geq t_0$ throughout this):

$$\left|\Psi(t)\right\rangle - \left|\Psi(t_0)\right\rangle = -\frac{i}{\hbar}\int_{t_0}^{t} dt_1\,\hat{H}(t_1)\left|\Psi(t_1)\right\rangle$$

$$\Rightarrow \quad \left|\Psi(t)\right\rangle = \left|\Psi(t_0)\right\rangle - \frac{i}{\hbar}\int_{t_0}^{t} dt_1\,\hat{H}(t_1)\left|\Psi(t_1)\right\rangle$$

Now insert this equation for $\left|\Psi(t_1)\right\rangle$:

$$\begin{aligned}\Rightarrow \left|\Psi(t)\right\rangle = \left|\Psi(t_0)\right\rangle &- \frac{i}{\hbar}\int_{t_0}^{t} dt_1\,\hat{H}(t_1)\left|\Psi(t_0)\right\rangle \\ &+ \left(-\frac{i}{\hbar}\right)^2\int_{t_0}^{t}\int_{t_0}^{t_1} dt_1\,dt_2\,\hat{H}(t_1)\hat{H}(t_2)\left|\Psi(t_2)\right\rangle\end{aligned}$$

Notice that $\hat{H}(t_1)\hat{H}(t_2)$ only appears with $t_1 \geq t_2$. We repeat this to find the propagator from t_0 to t:

$$\left|\Psi(t)\right\rangle = \hat{U}(t,t_0)\left|\Psi(t_0)\right\rangle$$

$$\hat{U}(t,t_0) = \hat{1} - \frac{i}{\hbar}\int_{t_0}^{t}\hat{H}(t_1)dt_1 + \left(-\frac{i}{\hbar}\right)^2\int_{t_0}^{t}\int_{t_0}^{t_1}\hat{H}(t_1)\hat{H}(t_2)dt_1dt_2 + \cdots$$

Notice we can write:

$$\int_{t_0}^{t}\int_{t_0}^{t_1}\hat{H}(t_1)\hat{H}(t_2)dt_1dt_2 = \frac{1}{2}\int_{t_0}^{t}\int_{t_0}^{t}\mathcal{T}_+\left(\hat{H}(t_1)\hat{H}(t_2)\right)dt_1dt_2$$

where $\mathcal{T}_+\left(\hat{H}(t_1)\hat{H}(t_2)\right)$ is the time-ordered product:

$$\mathcal{T}_+\left(\hat{H}(t_1)\hat{H}(t_2)\right) = \begin{cases}\hat{H}(t_1)\hat{H}(t_2) & t_1 \geq t_2 \\ \hat{H}(t_2)\hat{H}(t_1) & t_2 > t_1\end{cases}$$

We can repeat this for all higher integrals appearing in $\hat{U}(t,t_0)$ to find:

$$\begin{aligned}\hat{U}(t,t_0) &= \mathcal{T}_+\left(1 - \frac{i}{\hbar}\int_{t_0}^{t}dt_1\,\hat{H}(t_1) + \frac{1}{2}\left(-\frac{i}{\hbar}\right)^2\int_{t_0}^{t}dt_1\int_{t_0}^{t}dt_2\,\hat{H}(t_1)\hat{H}(t_2) + \cdots\right) \\ &= \mathcal{T}_+\exp\left(-\frac{i}{\hbar}\int_{t_0}^{t}dt\,\hat{H}(t)\right)\end{aligned}$$

This is called a time-ordered exponential. It satisfies:

$$\frac{\partial}{\partial t}\hat{U}(t,t_0) = -\frac{i}{\hbar}\hat{H}(t)\hat{U}(t,t_0)$$

$$\frac{\partial}{\partial t_0}\hat{U}(t,t_0) = +\frac{i}{\hbar}\hat{U}(t,t_0)\hat{H}(t_0)$$

It is fairly straight forward to verify that this $\hat{U}$ is unitary:

$$\hat{U}(t,t_0)^\dagger\hat{U}(t,t_0) = \hat{U}(t,t_0)\hat{U}(t,t_0)^\dagger = \hat{1}$$

1.2 The Interaction Picture

We now have the basic tools to tackle time-dependent perturbation theory. Let's start by dividing $\hat{H}$ into a reference part $\hat{H}_0$ and a perturbation $\lambda\hat{V}(t)$ which we now take to be time-dependent:

$$\hat{H}(t) = \hat{H}_0 + \lambda\hat{V}(t)$$

We define the interaction picture state $\left|\Psi^I(t)\right\rangle$ as:

$$\left|\Psi^I(t)\right\rangle = e^{+i\hat{H}_0 t/\hbar}\left|\Psi(t)\right\rangle$$

And we differentiate this to find:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\left|\Psi^I(t)\right\rangle &= \left(+\frac{i\hat{H}_0}{\hbar}\right)e^{+i\hat{H}_0 t/\hbar}\left|\Psi(t)\right\rangle \\ &\quad + e^{+i\hat{H}_0 t/\hbar}\left(-\frac{i}{\hbar}\hat{H}_0 - \frac{i}{\hbar}\lambda\hat{V}(t)\right)\left|\Psi(t)\right\rangle \\ &= e^{+i\hat{H}_0 t/\hbar}\lambda\hat{V}(t)\left|\Psi(t)\right\rangle\end{aligned}$$

Inserting $\hat{1} = e^{-i\hat{H}_0 t/\hbar}e^{+i\hat{H}_0 t/\hbar}$, we find:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\left|\Psi^I(t)\right\rangle &= -\frac{i}{\hbar}e^{+i\hat{H}_0 t/\hbar}\lambda\hat{V}(t)e^{-i\hat{H}_0 t/\hbar}e^{+i\hat{H}_0 t/\hbar}\left|\Psi(t)\right\rangle \\ &= -\frac{i}{\hbar}\lambda\hat{V}^I(t)\left|\Psi^I(t)\right\rangle\end{aligned}$$

where $\hat{V}^I(t) = e^{+i\hat{H}_0 t/\hbar}\hat{V}(t)e^{-i\hat{H}_0 t/\hbar}$. This is called the interaction picture Schrödinger equation. We can always recover $\left|\Psi(t)\right\rangle$ using:

$$\left|\Psi(t)\right\rangle = e^{-i\hat{H}_0 t/\hbar}\left|\Psi^I(t)\right\rangle$$

We can use our general solution from before to solve this as a perturbation series (note $\left|\Psi(0)\right\rangle = \left|\Psi^I(0)\right\rangle$):

$$\begin{aligned}\left|\Psi^I(t)\right\rangle &= \left|\Psi^I(0)\right\rangle - \frac{i}{\hbar}\lambda\int_0^t \hat{V}^I(t_1)\left|\Psi(0)\right\rangle dt_1 \\ &\quad + \left(-\frac{i}{\hbar}\right)^2\lambda^2\int_0^t\int_0^{t_1} dt_1 dt_2\,\hat{V}^I(t_1)\hat{V}^I(t_2)\left|\Psi(0)\right\rangle \\ &\quad + \cdots\end{aligned}$$

We can also write $\left|\Psi^I(t)\right\rangle$ in the eigenbasis of $\hat{H}_0$:

$$\left|\Psi^I(t)\right\rangle = \sum_n \left|n^{(0)}\right\rangle c_n(t) \Rightarrow c_n(t) = \left\langle n^{(0)}\middle|\Psi^I(t)\right\rangle$$

and the coefficients can be expanded as a perturbation series:

$$c_n(t) = c_n^{(0)}(t) + \lambda c_n^{(1)}(t) + \lambda^2 c_n^{(2)}(t) + \cdots$$

Comparing coefficients in powers of λ we find:

$$c_n^{(0)}(t) = \left\langle n^{(0)}\middle|\Psi^I(0)\right\rangle$$

$$c_n^{(1)}(t) = -\frac{i}{\hbar}\int_0^t \langle n^{(0)}|\hat{V}^I(t_1)|\Psi^I(0)\rangle\, dt_1$$

$$c_n^{(2)}(t) = \left(-\frac{i}{\hbar}\right)^2 \int_0^t \int_0^{t_1} dt_1 dt_2 \ \langle n^{(0)}|\hat{V}^I(t_1)\hat{V}^I(t_2)|\Psi^I(0)\rangle$$

If $|\Psi(0)\rangle = \left|i^{(0)}\right\rangle$, then we can simplify the first two equations to:

$$c_n^{(0)}(t) = \left\langle n^{(0)} \middle| i^{(0)} \right\rangle = \delta_{n,i}$$

$$c_n^{(1)}(t) = -\frac{i}{\hbar}\int_0^t dt_1 \ \langle n^{(0)}|\hat{V}^I(t)|i^{(0)}\rangle$$

1.3 Time-Independent Perturbation & Fermi's Golden Rule

Suppose $\hat{V}(t) = \hat{V}$ is independent of time. Let's evaluate $c_n^{(1)}(t)$ in this case:

$$\begin{aligned}
c_n^{(1)}(t) &= -\frac{i}{\hbar}\int_0^t dt_1 \ \langle n^{(0)}|e^{+i\hat{H}_0 t_1/\hbar}\hat{V}e^{-i\hat{H}_0 t_1/\hbar}|i^{(0)}\rangle \\
&= -\frac{i}{\hbar}\int_0^t dt_1 \ e^{+iE_n^{(0)}t_1/\hbar} \langle n^{(0)}|\hat{V}|i^{(0)}\rangle \, e^{-iE_i^{(0)}t_1/\hbar} \\
&= -\frac{i}{\hbar}V_{n,i}\int_0^t dt_1 \ e^{+i(E_n^{(0)}-E_i^{(0)})t_1/\hbar} \\
&= -\frac{i}{\hbar}V_{n,i}\frac{e^{i(E_n^{(0)}-E_i^{(0)})t/\hbar}-1}{i(E_n^{(0)}-E_i^{(0)})/\hbar} \\
&= -V_{n,i}\, e^{i\Delta E_{ni}t/2\hbar}\frac{e^{i\omega_{n,i}t/2}-e^{-i\omega_{n,i}t/2}}{\hbar\omega_{n,i}} \\
&= -\frac{2iV_{n,i}}{\hbar}e^{i\omega_{n,i}t/2}\frac{\sin(\omega_{n,i}t/2)}{\omega_{n,i}}
\end{aligned}$$

where $\omega_{n,i} = (E_n^{(0)} - E_i^{(0)})/\hbar$ and $V_{n,i} = \langle n^{(0)}|\hat{V}|i^{(0)}\rangle$. Now note the population of $\left|n^{(0)}\right\rangle$ is:

$$\begin{aligned}
|\langle n^{(0)}|\Psi(t)\rangle|^2 &= |\langle n^{(0)}|e^{-i\hat{H}_0 t/\hbar}|\Psi^I(t)\rangle|^2 \\
&= |e^{-iE_n^{(0)}t/\hbar}\langle n^{(0)}|\Psi^I(t)\rangle|^2 \\
&= |e^{-iE_n^{(0)}t/\hbar}|\,|c_n(t)|^2 \\
&= |c_n(t)|^2
\end{aligned}$$

So the population of $\left|n^{(0)}\right\rangle$ is (for $n \neq i$) $\lambda^2|c_n^{(1)}(t)|^2$:

$$\lambda^2|c_n^{(1)}(t)|^2 = \frac{\lambda^2|V_{n,i}|^2}{\hbar^2}\frac{\sin^2(\omega_{n,i}t/2)}{(\omega_{n,i}/2)^2}$$

Let's carefully examine the behaviour of this. First note:

$$\int_{-\infty}^{\infty} d\omega \, \frac{\sin^2(\omega a)}{\omega^2} = \pi a \quad \text{for } a > 0$$

So consider $(\omega_{n,i}t/2)$ becoming large:

$$\lim_{t\to\infty}\frac{\sin^2(\omega t/2)}{(\omega_{n,i}/2)^2 t} = \begin{cases} 0 & \text{if } \omega_{n,i} \neq 0 \\ +\infty & \text{if } \omega_{n,i} = 0 \end{cases}$$

but:

$$\int_{-\infty}^{\infty}\frac{\sin^2(\omega_{n,i}t/2)}{(\omega_{n,i}/2)^2 t}d\omega_{n,i} = 2\pi\frac{t}{t} = 2\pi$$

So as $(\omega_{n,i}t/2) \to \infty$, this becomes a δ-function. So:

$$P_n(t) = \lambda^2 |c_n^{(1)}(t)|^2 = \frac{\lambda^2 |V_{n,i}|^2}{\hbar^2}\, t \cdot 2\pi\delta(\omega_{n,i})$$

Noting $\delta(ax) = \frac{1}{|a|}\delta(x)$, this is:

$$P_n(t) = \frac{2\pi\lambda^2 |V_{n,i}|^2}{\hbar}\, t \cdot \delta(E_n^{(0)} - E_i^{(0)})$$

which is valid for small $\lambda^2 |V_{n,i}|^2 t$, i.e. short times. Compare this to a kinetic equation solution:

$$\dot{P}_n(t) = \sum_i k_{i\to n} P_i(t) - \sum_i k_{n\to i} P_n(t)$$

for $P_i(0) = 1$, $P_n(0) = 0$, and $t \to 0$:

$$P_n(t) = k_{i\to n} t \quad \text{or} \quad \dot{P}_n(0) = k_{i\to n}$$

This implies with the above that:

$$k_{i\to n} = \frac{2\pi\lambda^2 |V_{n,i}|^2}{\hbar}\delta(E_n^{(0)} - E_i^{(0)})$$

This is Fermi's Golden Rule. It tells us a perturbation can (to lowest order) only induce transitions between states if:

$$V_{n,i} \neq 0 \quad \text{and} \quad E_n^{(0)} = E_i^{(0)}$$

i.e. the coupling is non-zero and energy must be conserved.

1.4 Time-Dependent Perturbations

Let's consider a time-oscillating perturbation:

$$\begin{aligned}\hat{V}(t) &= \hat{V}\cos(\Omega t + \phi) \\ &= \frac{1}{2}\hat{V}e^{+i\phi}e^{+i\Omega t} + \frac{1}{2}\hat{V}e^{-i\phi}e^{-i\Omega t} \\ &= \hat{V}_+ e^{+i\Omega t} + \hat{V}_- e^{-i\Omega t}\end{aligned}$$

We can now use what we worked out before to write down $c_n^{(1)}(t)$ if $|\Psi(0)\rangle = \left|i^{(0)}\right\rangle$:

$$\begin{aligned}c_n^{(1)}(t) = &-\frac{2i}{\hbar}V_{+,ni}\frac{\sin\left((\omega_{ni}+\Omega)t/2\right)}{\omega_{ni}+\Omega}e^{+i(\omega_{ni}+\Omega)t/2} \\ &-\frac{2i}{\hbar}V_{-,ni}\frac{\sin\left((\omega_{ni}-\Omega)t/2\right)}{\omega_{ni}-\Omega}e^{+i(\omega_{ni}-\Omega)t/2}\end{aligned}$$

We can again find $P_n(t)$ to lowest order in λ:

$$\begin{aligned}P_n(t) = &\frac{\lambda^2|V_{+,ni}|^2}{\hbar^2}\frac{\sin^2\left((\omega_{ni}+\Omega)t/2\right)}{\left((\omega_{ni}+\Omega)/2\right)^2} \\ &+\frac{\lambda^2|V_{-,ni}|^2}{\hbar^2}\frac{\sin^2\left((\omega_{ni}-\Omega)t/2\right)}{\left((\omega_{ni}-\Omega)/2\right)^2} \\ &+\frac{\lambda^2 2\mathrm{Re}\left[V_{+,ni}^* V_{-,ni}e^{-i\Omega t}\right]}{\hbar^2}\frac{\sin\left((\omega_{ni}+\Omega)t/2\right)}{\left((\omega_{ni}+\Omega)/2\right)}\frac{\sin\left((\omega_{ni}-\Omega)t/2\right)}{\left((\omega_{ni}-\Omega)/2\right)}\end{aligned}$$

By the argument we had before, for $|(\omega_{ni} \pm \Omega)t| \gg 1$ and $\lambda \to 0$, the first two terms give:

$$\frac{\lambda^2 |V_{\pm,ni}|^2}{\hbar^2} \cdot t\, 2\pi\delta(\omega_{ni} \pm \Omega)$$

Let's analyse the last term as before: first as $t \to \infty$:

$$\lim_{t\to\infty} \frac{\sin(\omega t/2)}{(\omega/2)t^{1/2}} = \begin{cases} 0 & \omega \neq 0 \\ +\infty & \omega = 0 \end{cases}$$

and the integral satisfies (for $t > 0$):

$$\int_{-\infty}^{+\infty} d\omega \frac{\sin(\omega t/2)}{(\omega/2)t^{1/2}} = \frac{2\pi}{\sqrt{t}}$$

So the last term is involves terms that become:

$$\lim_{t\to\infty} \frac{\sin(\omega t/2)}{(\omega t/2)t^{1/2}} = \frac{2\pi}{t^{1/2}}\delta(\omega)$$

So:

$$\lim_{t\to\infty} \frac{\sin\left((\omega_{ni}+\Omega)t/2\right)}{((\omega_{ni}+\Omega)/2)} \frac{\sin\left((\omega_{ni}-\Omega)t/2\right)}{((\omega_{ni}-\Omega)/2)} = (2\pi)^2\delta(\omega_{ni}+\Omega) \times \delta(\omega_{ni}-\Omega)$$

but for $\Omega \neq 0$, this is zero, so this term vanishes. Overall we find:

$$k_{i\to n} = \frac{\lambda^2|V_{+,ni}|^2}{\hbar^2} \cdot 2\pi\delta(\omega_{ni}+\Omega) + \frac{\lambda^2|V_{-,ni}|^2}{\hbar^2} \cdot 2\pi\delta(\omega_{ni}-\Omega)$$

So a time oscillating perturbation can only cause a transition if the frequency matches the energy gap.

1.4.1 Example: Dipole-transitions

For a system interacting with an oscillating electromagnetic field, $\hat{V}(t)$ is:

$$\hat{V}(t) = -\hat{\mu} \cdot \underline{\mathcal{E}}(t)$$

$$\underline{\mathcal{E}}(t) = \mathcal{E}_0 \underline{n} \cos(\Omega t)$$

where $\underline{n}$ is a unit vector defining the polarization axis, and $\underline{\hat{\mu}}$ is the system dipole moment. This gives:

$$\hat{V}_{\pm} = -\frac{1}{2}\mathcal{E}_0 \underline{n} \cdot \underline{\hat{\mu}}$$

So a transition probability is non-zero if and only if:

$$\langle n^{(0)} | \underline{n} \cdot \underline{\hat{\mu}} | i^{(0)} \rangle \neq 0$$

so at least one component of the transition dipole moment must be non-zero. We also see that the transition probability is proportional to $\mathcal{E}_0^2$ which is directly proportional to the energy intensity of the electromagnetic wave.

2 Adiabatic theorem

Suppose the Hamiltonian of a QM system depends very slowly on time. Previously we saw that time-dependent perturbations can cause transitions between eigenstates, but what happens if the change is very slow?

$\hat{H}(t)$ is hermitian, so it has a time-dependent basis of eigenstates

$$\hat{H}(t)\left|n(t)\right\rangle = E_n(t)\left|n(t)\right\rangle$$

This basis can be used to expand $\left|\psi(t)\right\rangle$

$$\left|\psi(t)\right\rangle = \sum_n c_n(t)\left|n(t)\right\rangle$$

Plugging this into the TDSE gives

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

$$\sum_n\left[\dot{c}_n(t)\left|n(t)\right\rangle + c_n(t)\left|\dot{n}(t)\right\rangle\right] = -\frac{i}{\hbar}\sum_n c_n(t)E_n(t)\left|n(t)\right\rangle$$

Take the inner product with $\left\langle m(t)\right|$ to obtain

$$\dot{c}_m(t) + c_m(t)\left\langle m(t)|\dot{m}(t)\right\rangle + \sum_{n\neq m} c_n(t)\left\langle m(t)|\dot{n}(t)\right\rangle = -\frac{i}{\hbar}c_m(t)E_m(t)$$

$$\Rightarrow \dot{c}_m(t) + c_m(t)\left\langle m(t)|\dot{m}(t)\right\rangle + \frac{i}{\hbar}E_m(t) = \sum_{n\neq m} c_n(t)\left\langle m(t)|\dot{n}(t)\right\rangle$$

Let's consider $\left\langle m(t)|n(t)\right\rangle = \delta_{n,m}$

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\left\langle m(t)|n(t)\right\rangle &= \left\langle \dot{m}(t)|n(t)\right\rangle + \left\langle m(t)|\dot{n}(t)\right\rangle = 0 \\ \Rightarrow \quad \mathrm{Re}\left[\left\langle m(t)|\dot{m}(t)\right\rangle\right] &= 0 \qquad \because \left\langle \dot{m}(t)|m(t)\right\rangle = -\left\langle m(t)|\dot{m}(t)\right\rangle^* \\ \Rightarrow \quad \left\langle m(t)|\dot{m}(t)\right\rangle &= i\gamma_m(t) \end{aligned}$$

where $\gamma_m(t)$ is real. So this term only contributes a phase factor to $c_m(t)$. This is called the Berry phase.

Now let's consider the derivative of $\hat{H}(t)\left|n(t)\right\rangle$

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\hat{H}(t)\left|n(t)\right\rangle\right) = \frac{\mathrm{d}}{\mathrm{d}t}\left(E_n(t)\left|n(t)\right\rangle\right)$$

$$\Rightarrow \left(\frac{\mathrm{d}}{\mathrm{d}t}\hat{H}(t)\right)\left|n(t)\right\rangle + \hat{H}(t)\left|\dot{n}(t)\right\rangle = \dot{E}_n(t)\left|n(t)\right\rangle + E_n(t)\left|\dot{n}(t)\right\rangle$$

Take $\left\langle m(t)\right|$ with this for $m \neq n$

$$\left\langle m(t)\right|\left(\frac{\mathrm{d}}{\mathrm{d}t}\hat{H}(t)\right)\left|n(t)\right\rangle + E_m(t)\left\langle m(t)|\dot{n}(t)\right\rangle = \dot{E}_n(t)\cancelto{0}{\left\langle m(t)|n(t)\right\rangle} + E_n(t)\left\langle m(t)|\dot{n}(t)\right\rangle$$

$$\Rightarrow \left\langle m(t)|\dot{n}(t)\right\rangle = \frac{\left\langle m(t)\right|\left[\frac{\mathrm{d}}{\mathrm{d}t}\hat{H}(t)\right]\left|n(t)\right\rangle}{E_n(t) - E_m(t)}$$

Putting this all together we find

$$\dot{c}_m(t) + i\left(\gamma_m(t) + \frac{E_m(t)}{\hbar}\right)c_m(t) = \sum_{n\neq m} c_n(t)\frac{\langle m(t)|\frac{\mathrm{d}}{\mathrm{d}t}\hat{H}|n(t)\rangle}{E_m(t) - E_n(t)}$$

Assuming $\frac{\mathrm{d}}{\mathrm{d}t}\hat{H}(t)$ is small (i.e. $\hat{H}(t)$ is slowly varying) and $E_m(t) - E_n(t)$ is never zero, the right hand side is approximately zero, so

$$c_m(t) \simeq c_m(0)\exp\left[-i\int_0^t d\tau\left(\gamma_m(\tau) + \frac{1}{\hbar}E_m(\tau)\right)\right]$$

We can include the term on the RHS, the non-adiabatic term $d_{mn}(t) = \langle m(t)|\dot{n}(t)\rangle$, as a perturbation to give

$$\dot{c}_m^{(1)}(t) + i\dot{\theta}_m c_m^{(1)}(t) = \sum_{n\neq m} c_n^{(0)}(t)d_{mn}(t)$$

$$\Rightarrow \frac{\mathrm{d}}{\mathrm{d}t}\left(e^{i\theta_m(t)}c_m^{(1)}(t)\right) = \sum_{n\neq m} e^{i\theta_m(t)}c_n^{(0)}(t)d_{mn}(t)$$

$$\Rightarrow c_m^{(1)}(t) = e^{-i\theta_m(t)}\sum_{n\neq m}\int_0^t d\tau\, e^{i\theta_m(\tau)}c_n^{(0)}(\tau)d_{mn}(\tau)$$

So for very slow changes $|c_m(t)| = |c_m(0)|$ and there are no transitions between eigenstates. If energy states cross however this breaks down and transitions occur.

2.1 Example: Molecular Electronic Transitions

Let's treat a molecule semi-classically. We'll fix the nuclear positions, but let them vary in time, $\boldsymbol{R}(t)$, but treat the electrons quantum-mechanically.

The electronic energy states depend on $\boldsymbol{R}$ through $\hat{H}(\boldsymbol{R})$

$$|E_n\rangle = |E_n(\boldsymbol{R})\rangle$$

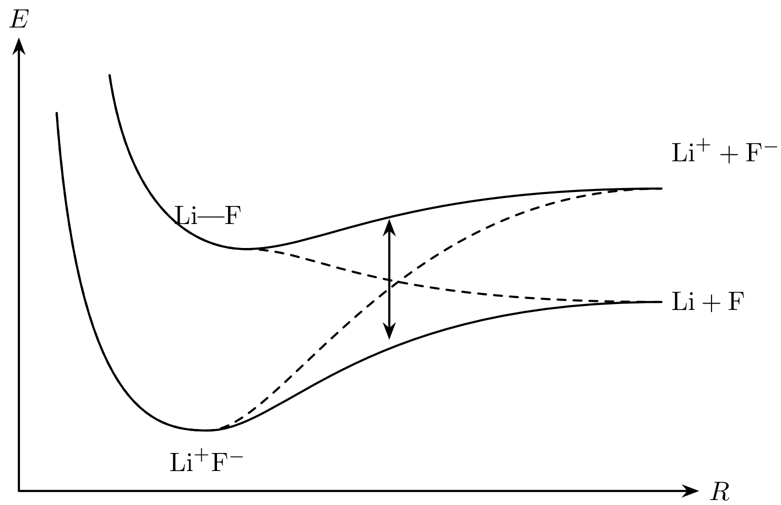
Transitions occur due to nuclear motion

$$E_n(t) = E_n(\boldsymbol{R}(t))$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{H}(\boldsymbol{R}(t)) = \boldsymbol{\nabla}\hat{H}(\boldsymbol{R}(t))\cdot\dot{\boldsymbol{R}}(t)$$

If energy levels cross, coupling becomes significant due to nuclear motion, and a transition can occur.

Points where $E_n(\boldsymbol{R}) = E_m(\boldsymbol{R})$ are called conical intersections, and when $E_n(\boldsymbol{R}) \neq E_m(\boldsymbol{R})$ goes through a minimum is an avoided crossing. An example is LiF



If Li—F is formed in its excited state, it can transition when the energy gap is smallest as Li + F move apart.

2.2 Landau–Zener Transitions

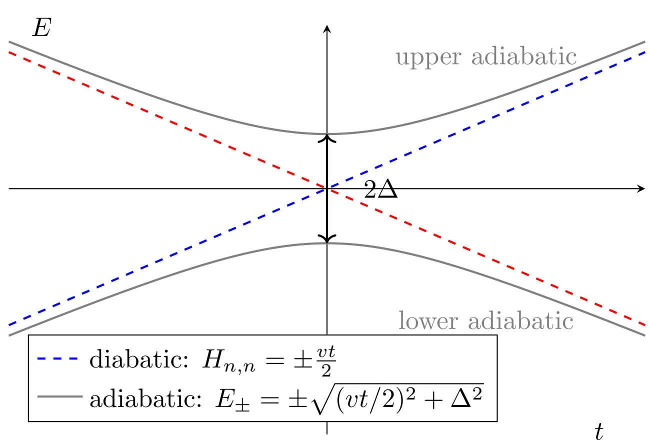
For most systems we cannot analytically calculate the transition probability between states. One special case is the two state Landau–Zener model

$$\begin{aligned}\hat{H}(t) &= \frac{1}{2}vt\,|1\rangle\langle 1| - \frac{1}{2}vt\,|2\rangle\langle 2| + \Delta(|1\rangle\langle 2| + |2\rangle\langle 1|)\\ &= \begin{pmatrix}\frac{1}{2}vt & \Delta\\ \Delta & -\frac{1}{2}vt\end{pmatrix}\end{aligned}$$

$|1\rangle$ and $|2\rangle$ are called diabatic states. The adiabatic (time-dependent) eigenstates are

$$\begin{aligned}|\psi_+\rangle &= \cos(\theta(t))\,|1\rangle + \sin(\theta(t))\,|2\rangle\\ |\psi_-\rangle &= -\sin(\theta(t))\,|1\rangle + \cos(\theta(t))\,|2\rangle\\ E_\pm(t) &= \pm\frac{1}{2}\sqrt{(vt)^2 + 4\Delta^2}\\ \tan(2\theta(t)) &= \frac{2\Delta}{vt}\end{aligned}$$

These are shown here



For $t \to -\infty$

$$|\psi_+\rangle \to |2\rangle \quad , \quad |\psi_-\rangle \to |1\rangle$$

and $t \to +\infty$

$$|\psi_+\rangle \to |1\rangle \quad , \quad |\psi_-\rangle \to |2\rangle$$

We'll assume we start in either $|\psi_+\rangle$ or $|\psi_-\rangle$ at $t \to -\infty$. We want to find the transition probability as $t \to +\infty$. Henceforth we'll assume we start at $t = -\infty$ in $|1\rangle$. We can just work with diabatic states to find

$$|\psi(t)\rangle = c_1(t)\,|1\rangle + c_2(t)\,|2\rangle$$

Applying the TDSE we find

$$\dot{c}_1(t) = -\frac{ivt}{2\hbar}c_1(t) - \frac{i}{\hbar}\Delta c_2(t) \tag{1}$$

$$\dot{c}_2(t) = +\frac{ivt}{2\hbar}c_2(t) - \frac{i}{\hbar}\Delta c_1(t) \tag{2}$$

We can re-write this by introducing

$$a_1(t) = e^{+i\varphi(t)}c_1(t)$$
$$a_2(t) = e^{-i\varphi(t)}c_2(t)$$
$$\varphi(t) = \int_{-T}^{t} \frac{vt'}{2\hbar}\,\mathrm{d}t' = \frac{v(t^2 - T^2)}{4\hbar}$$

so we find

$$\dot{a}_1(t) = -\frac{i\Delta}{\hbar}e^{2i\varphi(t)}a_2(t)$$
$$\dot{a}_2(t) = -\frac{i\Delta}{\hbar}e^{-2i\varphi(t)}a_1(t)$$

We can differentiate the $a_1(t)$ equation again with respect to time to obtain

$$\ddot{a}_1(t) = \frac{ivt}{\hbar}\left(-\frac{i\Delta}{\hbar}e^{2i\varphi(t)}a_2(t)\right) - \frac{i\Delta}{\hbar}e^{2i\varphi(t)}\dot{a}_2(t)$$

The first term on the RHS in brackets is $\dot{a}_1(t)$. The second term can be written in terms of $a_1(t)$ using the equation for $\dot{a}_2(t)$

$$\ddot{a}_1(t) = \frac{ivt}{\hbar}\dot{a}_1(t) - \frac{\Delta^2}{\hbar^2}a_1(t)$$

We note that letting $t \to \infty$, $a_n(t \to \infty)$ becomes a constant, because the phase factor completely captures its evolution where the diagonal term dominates the dynamics of $c_n(t)$. We also know $a_2(t \to -\infty) = 0$.

The last equation we found for $a_1(t)$ can be rearranged to

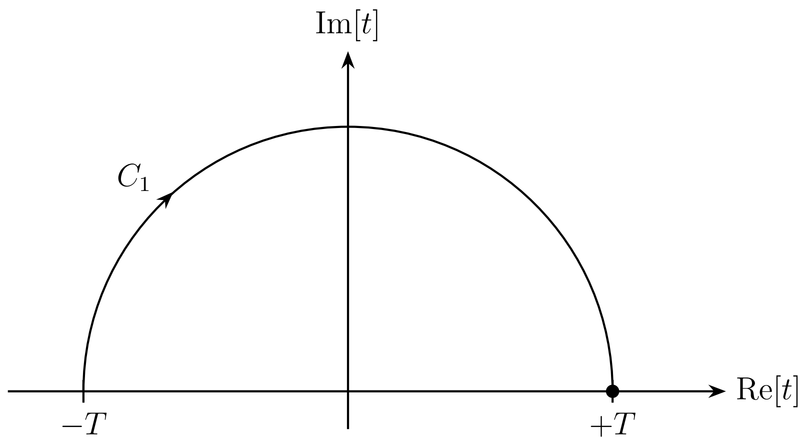
$$\frac{\dot{a}_1(t)}{a_1(t)} = -\frac{i\hbar}{vt}\frac{\ddot{a}_1(t)}{a_1(t)} - i\frac{\Delta^2}{\hbar vt}$$

We now integrate this from $-T$ to T to obtain

$$\ln\!\left(\frac{a_1(T)}{a_1(-T)}\right) = -\int_{-T}^{T} \frac{i\hbar}{vt}\frac{\ddot{a}_1(t)}{a_1(t)}\,\mathrm{d}t - i\int_{-T}^{T}\frac{\Delta^2}{\hbar vt}\,\mathrm{d}t$$

Letting $T \to \infty$, it can be argued using tools from complex analysis that the first integral is 0. This involves taking the integral into the complex plane and integrating over the contour depicted below. Likewise the second integral can be evaluated along the same contour as

$$\int_{-T}^{T}\frac{1}{t}\,\mathrm{d}t = \int_{\pi}^{0}\frac{1}{Te^{i\theta}}iTe^{i\theta}\,\mathrm{d}\theta = -i\pi$$



The end result is

$$\ln\left(\frac{a_1(\infty)}{a_1(-\infty)}\right) = -\frac{\pi\Delta^2}{\hbar v}$$

and noting that $P_1 = |c_1(\infty)|^2 = |a_1(\infty)|^2$, and $|a_1(-\infty)|^2 = 1$, the probability of staying in $|1\rangle$ is

$$P_1 = e^{-2\pi\Delta^2/\hbar v}$$

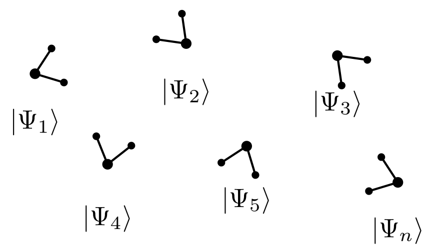
This is the Landau–Zener Formula.

This is consistent with the adiabatic theorem as $v \to 0$, $P_1 \to 0$ so $P_2 \to 1$, so population stays on the adiabatic (eigen-) state. If $v \to +\infty$ the time spent in the crossing region is low. So no population can be transferred from $|1\rangle \to |2\rangle$ and $P_1 \to 1$.

3 Density Operators

3.1 Ensembles and the density operator

So far we have only considered single isolated quantum systems, the state of which is described with a state vector $|\Psi\rangle$. In chemistry we generally have very large numbers $\sim 10^{23}$ molecules in our system of interest. Each molecule has the same Hamiltonian, so in this sense is identical, but it may be in a different state. We label each state for each member of the ensemble $|\Psi_n\rangle$.



If we measure an observable $\hat{O}$ for each system and take the average of the measurements we find

$$\langle O\rangle = \frac{1}{N}\sum_n \langle\Psi_n|\hat{O}|\Psi_n\rangle$$

We re-arrange this to find

$$\begin{aligned}\langle O\rangle &= \frac{1}{N}\sum_n \langle\Psi_n|\hat{O}\hat{1}|\Psi_n\rangle \\ &= \frac{1}{N}\sum_{n,k} \langle\Psi_n|\hat{O}|k\rangle\,\langle k|\Psi_n\rangle \\ &= \sum_k \langle k|\frac{1}{N}\sum_n |\Psi_n\rangle\langle\Psi_n|\,\hat{O}|k\rangle\end{aligned}$$

We define the density operator as

$$\hat{\rho} = \frac{1}{N}\sum_n |\Psi_n\rangle\,\langle\Psi_n|$$

and recall $\mathrm{Tr}[\hat{A}] = \sum_k \langle k|\hat{A}|k\rangle$ so we can write $\langle O\rangle$ as

$$\langle O\rangle = \mathrm{Tr}\Big[\hat{\rho}\hat{O}\Big]$$

The density operator $\hat{\rho}$ contains all information about observables of the ensemble.

3.2 Properties of the density operator

It is straightforward to verify that $\hat{\rho}$ is hermitian

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

Its trace is 1

$$\mathrm{Tr}[\hat{\rho}] = 1$$

and it is positive-semi-definite, i.e. $\forall\,|\phi\rangle \in \mathcal{H}$

$$\langle\phi|\hat{\rho}|\phi\rangle \geq 0$$

From this we can show

$$\frac{1}{d_{\mathcal{H}}} \leq \mathrm{Tr}\left[\hat{\rho}^2\right] \leq 1$$

where $d_{\mathcal{H}}$ is the dimensionality of the Hilbert space.

$\hat{\rho} = \hat{\rho}^\dagger$ is necessary for observables to be real-valued and $\mathrm{Tr}[\hat{\rho}] = 1$ ensures the total probability for finding the ensemble in any set of orthonormal states is 1.

3.3 The purity of the density operator

The $\mathrm{Tr}\left[\hat{\rho}^2\right]$ is called the purity of $\hat{\rho}$. If all members of an ensemble are in the same state $|\Psi_n\rangle$ then

$$\begin{aligned}\hat{\rho} &= |\Psi\rangle\langle\Psi| \\ \hat{\rho}^2 &= |\Psi\rangle\langle\Psi| \\ \Rightarrow \mathrm{Tr}\left[\hat{\rho}^2\right] &= \mathrm{Tr}\left[\hat{\rho}^1\right] = 1\end{aligned}$$

This is called a pure state for the density operator.

If a fraction of the ensemble is in another state $|\Psi'\rangle$ where $|\langle\Psi|\Psi'\rangle| = 0$ then

$$\begin{aligned}\hat{\rho} &= (1-p)|\Psi\rangle\langle\Psi| + p|\Psi'\rangle\langle\Psi'| \\ \hat{\rho}^2 &= (1-p)^2|\Psi\rangle\langle\Psi| + p^2|\Psi'\rangle\langle\Psi'| + p(1-p)(|\Psi\rangle\langle\Psi'| + |\Psi'\rangle\langle\Psi|) \\ \Rightarrow \mathrm{Tr}\left[\hat{\rho}^2\right] &= (1-p)^2 + p^2 = 1 - 2p + 2p^2 < 1\end{aligned}$$

So if any fraction of the ensemble is in a different state $\mathrm{Tr}\left[\hat{\rho}^2\right] < 1$. This type of ensemble (density operator) state is called a mixed state.

If $\hat{\rho} = \frac{1}{d_{\mathcal{H}}}\hat{1}$, this is called the maximally mixed state, which has the minimum purity.

3.4 Thermal Density operators

From Stat. Mech. we know that for a thermal ensemble the expectation value of an observable is

$$\langle O\rangle = \sum_n \frac{1}{Q} e^{-\beta E_n} \langle E_n|\hat{O}|E_n\rangle$$

Note $\beta = 1/k_B T$. Inserting an identity as before we find $\hat{\rho}$ to be

$$\begin{aligned}\hat{\rho} &= \frac{1}{Q}\sum_n e^{-\beta E_n}|E_n\rangle\langle E_n| \\ &= \frac{1}{Q}\sum_n e^{-\beta\hat{H}}|E_n\rangle\langle E_n| \\ &= \frac{1}{Q}e^{-\beta\hat{H}}\end{aligned}$$

and the partition function $Q = \mathrm{Tr}\left[e^{-\beta\hat{H}}\right]$

3.5 Entropy

The entropy of a thermal ensemble is

$$\begin{aligned}
S &= \frac{1}{T}(E - A) = k_B \beta (E - A) \\
A &= -k_B T \ln Q, \quad E = -\frac{\partial}{\partial \beta} \ln Q \\
S &= k_B (\ln Q - \beta \frac{\partial}{\partial \beta} \ln Q) \\
&= k_B (\ln Q \sum_n p_n + \sum_n \beta E_n p_n) \\
&= -k_B \sum_n p_n (\ln Q + \beta E_n) \\
&= -k_B \sum_n p_n \ln p_n
\end{aligned}$$

For a thermal density operator

$$\hat{\rho} \left| E_n \right\rangle = p_n \left| E_n \right\rangle \text{ and } \ln(\hat{\rho}) \left| E_n \right\rangle = \ln(p_n) \left| E_n \right\rangle$$

So

$$\begin{aligned}
S &= -k_B \sum_n \langle E_n | \hat{\rho} \ln \hat{\rho} | E_n \rangle \\
&= -k_B \operatorname{Tr}[\hat{\rho} \ln \hat{\rho}]
\end{aligned}$$

We can use this to define the entropy for any density operator.

3.6 Dynamics of the density operator

How does the density operator evolve in time? Each state in the ensemble obeys the TDSE

$$\frac{d}{dt} \left| \Psi_n(t) \right\rangle = -i \frac{\hat{H}}{\hbar} \left| \Psi_n(t) \right\rangle$$

so this means $\frac{d}{dt}\hat{\rho}(t)$ is

$$\begin{aligned}
\frac{d}{dt}\hat{\rho}(t) &= \frac{1}{N} \sum_n \left((\frac{d}{dt} \left| \Psi_n \right\rangle) \left\langle \Psi_n \right| + \left| \Psi_n \right\rangle \frac{d}{dt} \left\langle \Psi_n \right| \right) \\
&= \frac{1}{N} \sum_n \left(-\frac{i}{\hbar} \hat{H} \left| \Psi_n \right\rangle \left\langle \Psi_n \right| + \frac{i}{\hbar} \left| \Psi_n \right\rangle \left\langle \Psi_n \right| \hat{H} \right) \\
&= -\frac{i}{\hbar} \hat{H} \hat{\rho} + \frac{i}{\hbar} \hat{\rho} \hat{H} = -\frac{i}{\hbar} [\hat{H}, \hat{\rho}]
\end{aligned}$$

This is called the Liouville-von Neumann equation.

Recall that operators on $\mathcal{H}$ themselves form a Hilbert space (in the parlance of linear algebra this is $\mathcal{H} \otimes \mathcal{H}^*$) which we call the **Liouville space**.

It is straightforward to verify that

$$\hat{A}\hat{\rho}\hat{B} \equiv \hat{\hat{\mathcal{C}}}(\hat{\rho})$$

is a linear operator on operators.

$$\begin{aligned}\hat{\hat{\mathcal{C}}}(\hat{\rho}a_1\hat{\beta}_1 + a_2\hat{\beta}_2) &= \hat{A}(a_1\hat{\beta}_1 + a_2\hat{\beta}_2)\hat{B} \\ &= a_1\hat{A}\hat{\rho}_1\hat{B} + a_2\hat{A}\hat{\rho}_2\hat{B} \\ &= a_1\hat{\hat{\mathcal{C}}}(\hat{\rho}) + a_2\hat{\hat{\mathcal{C}}}(\hat{\beta})\end{aligned}$$

So operators on operators of this form are linear and we can unambiguously write $\hat{\hat{\mathcal{C}}}(\hat{\rho}) = \hat{\hat{\mathcal{C}}}\hat{\rho}$. This is called a"super-operator". We indicate them with double hats e.g. $\hat{\hat{\mathcal{A}}}$.

We can also write the Liouville von-Neumann equation as

$$\frac{d}{dt}\hat{\rho}(t) = \hat{\hat{\mathcal{L}}}\hat{\rho}(t) \quad , \quad \hat{\hat{\mathcal{L}}}\hat{\rho} = -\frac{i}{\hbar}[\hat{H}, \hat{\rho}]$$

Because this equation is linear, we can use all the tools of linear algebra to manipulate equations involving density operators.

3.7 Inner products of density operators and observables

As we have shown, density operators form a Hilbert space, much like quantum states of individual isolated systems, but with the dimensionality squared. We will sometimes use Hilbert space "ket" notation for density operators. This will be referred to as a **Liouville space** vector

$$|\rho\rangle\rangle \equiv \hat{\rho}$$

The dual $\langle\langle\rho|$ is defined using the density operator inner product

$$\langle\langle\rho_1|\rho_2\rangle\rangle = \mathrm{Tr}\Big[\hat{\rho}_1^\dagger\hat{\rho}_2\Big].$$

A long time ago in 215A you verified that this does indeed constitute an inner product. Note we should generally try to avoid mixing the Hilbert space notation $\hat{\rho}$ and Liouville space notation $|\rho\rangle\rangle$, in order to avoid confusion.

Using Liouville space notation, we note expectation values of observables are given by

$$\langle O\rangle = \langle\langle O^\dagger|\rho\rangle\rangle$$

and density operators are normalised such that

$$1 = \langle 1\rangle = \langle\langle 1|\rho\rangle\rangle.$$

Note that the inner product

$$\langle\langle\rho|\rho\rangle\rangle$$

is the purity of the density operator.

3.8 Matrix-vector representation of Liouville space

Very similar to how quantum states of composite systems are represented with complex column vectors, we can construct a column vector representation of a density operator. First we pick a basis for the Hilbert space $|n\rangle$ and map each basis vector onto a unit column vector $\boldsymbol{e}_n$.

$$|n\rangle \to \boldsymbol{e}_n$$

The operators $|n\rangle\langle m|$ form a basis for the Liouville space, so we map these onto kronecker products of the Hilbert space basis vectors.

$$|n\rangle\,\langle m| \to \boldsymbol{e}_n \otimes \boldsymbol{e}_m$$

so density operators are mapped onto column vectors as

$$\hat{\rho} \to \sum_{n,m} \langle n|\hat{\rho}|m\rangle\, \boldsymbol{e}_n \otimes \boldsymbol{e}_m$$

From this we can represent super-operators as matrices. Left multiplication by an operator is represented as

$$\begin{aligned}
\hat{A}\hat{\rho} &\to \sum_{n,m,k} \langle n|\hat{A}|k\rangle\ \langle k|\hat{\rho}|m\rangle\, \boldsymbol{e}_n \otimes \boldsymbol{e}_m \\
&= \sum_m (\mathbf{A}\boldsymbol{\rho}^c_m) \otimes \boldsymbol{e}_m \quad \text{[Note } \boldsymbol{\rho}^c_m \text{ denotes column } m \text{ of the matrix of } \hat{\rho}] \\
&= (\mathbf{A} \otimes \mathbf{1})\boldsymbol{\rho}
\end{aligned}$$

and right multiplication is represented as

$$\begin{aligned}
\hat{\rho}\hat{A} &\to \sum_{n,m,k} \langle n|\hat{\rho}|k\rangle\ \langle k|\hat{A}|m\rangle\, \boldsymbol{e}_n \otimes \boldsymbol{e}_m \\
&= \sum_{n,m} \boldsymbol{e}_n \otimes \boldsymbol{e}_m[(\boldsymbol{\rho}^r_n\mathbf{A})]_m \quad \text{[Note } \boldsymbol{\rho}^r_n \text{ denotes row } n \text{ of the matrix of } \hat{\rho}] \\
&= \sum_n \boldsymbol{e}_n \otimes ((\boldsymbol{\rho}^r_n\mathbf{A})^T) \\
&= (\mathbf{1} \otimes \mathbf{A}^T)\boldsymbol{\rho}
\end{aligned}$$

So general super-operators are represented as

$$\begin{aligned}
\hat{A}\hat{\rho}\hat{B} &\to (\mathbf{A} \otimes \mathbf{1})(\mathbf{1} \otimes \mathbf{B}^T)\boldsymbol{\rho} \\
&= (\mathbf{A} \otimes (\mathbf{B}^T))\boldsymbol{\rho}.
\end{aligned}$$

This is a very useful recipe for going from matrix representations of operators and states in Hilbert space to super-operators and density operators in Liouville space for computations.

4 Linear Response Theory

In a typical experiment, e.g. in spectroscopy, we apply a time-dependent perturbation to a QM system and then measure the response in an observable. This is often done to a whole ensemble of systems.

Let's consider the ensemble initially in $\hat{\rho}_0$. It evolves according to

$$\frac{d}{dt}\hat{\rho}(t) = -\frac{i}{\hbar}\left[\hat{H}_0 + f(t)\hat{V}, \hat{\rho}(t)\right]$$

$f(t)\hat{V}$ is the applied perturbation.

As in TDPT we can move to the interaction picture

$$\hat{A}^I(t) = e^{i\hat{H}_0 t/\hbar}\hat{A}(t)e^{-i\hat{H}_0 t/\hbar}.$$

In this frame $\hat{\rho}^I(t)$ obeys

$$\begin{aligned}\frac{d}{dt}\hat{\rho}^I(t) &= +\frac{i}{\hbar}\left[\hat{H}_0, \hat{\rho}^I(t)\right] - \frac{i}{\hbar}e^{i\hat{H}_0 t/\hbar}\left[\hat{H}_0 + f(t)\hat{V}, \hat{\rho}(t)\right]e^{-i\hat{H}_0 t/\hbar} \\ &= -\frac{i}{\hbar}\left[f(t)\hat{V}^I(t), \hat{\rho}^I(t)\right]\end{aligned}$$

We can apply PT to this to find

$$\hat{\rho}^I(t) = \hat{\rho}^I(0) - \frac{i}{\hbar}\int_0^t dt_1\, f(t_1)\left[\hat{V}^I(t_1), \hat{\rho}^I(0)\right] + \mathcal{O}(V^2)$$

This is completely analogous to TDPT with wave-functions.

Suppose we measure an observable $\hat{A}$ in response to the perturbation. Using $\langle A\rangle = \mathrm{Tr}\left[\hat{A}\hat{\rho}\right]$ we find

$$\begin{aligned}\langle A(t)\rangle = &\ \mathrm{Tr}\left[\hat{A}\hat{U}_0\hat{\rho}(0)\hat{U}_0^\dagger\right] \\ &- \frac{i}{\hbar}\int_0^t dt_1\, f(t_1)\,\mathrm{Tr}\left[\hat{A}\hat{U}_0\left[\hat{V}^I(t_1), \hat{\rho}_0\right]\hat{U}_0^\dagger\right]\end{aligned}$$

$\hat{U}_0 = e^{-i\hat{H}_0 t/\hbar}$. Re-arranging gives, together with $\hat{\rho}^I(0) = \hat{\rho}_0$ [Note $\mathrm{Tr}\left[\hat{A}\hat{B}\right] = \mathrm{Tr}\left[\hat{B}\hat{A}\right]$]:

$$\begin{aligned}\langle A(t)\rangle &= \mathrm{Tr}\left[\hat{A}^I(t)\hat{\rho}_0\right] \\ &\quad - \frac{i}{\hbar}\int_0^t dt_1\, f(t_1)\,\mathrm{Tr}\left[\hat{A}^I(t)\left[\hat{V}^I(t_1), \hat{\rho}_0\right]\right] \\ &= \mathrm{Tr}\left[\hat{A}^I(t)\hat{\rho}_0\right] \\ &\quad - \frac{i}{\hbar}\int_0^t dt_1\, f(t_1)\,\mathrm{Tr}\left[\left[\hat{A}^I(t), \hat{V}^I(t_1)\right]\hat{\rho}_0\right]\end{aligned}$$

Write $\langle A\rangle_0 = \mathrm{Tr}\left[\hat{A}\hat{\rho}_0\right]$ to find

$$\langle A(t)\rangle = \langle A^I(t)\rangle_0 - \frac{i}{\hbar}\int_0^t dt_1\, f(t_1)\langle\left[\hat{A}^I(t), \hat{V}^I(t_1)\right]\rangle_0$$

The last term is called the linear response function

$$\chi_{AV}(t, t_1) = \frac{i}{\hbar}\left\langle[\hat{A}^I(t), \hat{V}^I(t_1)]\right\rangle_0$$

This is linear response theory.

4.1 Thermal Initial conditions

The response function can be simplified considerably if $[\hat{\rho}_0, \hat{H}_0] = 0$, e.g. if $\hat{\rho}_0 = \frac{1}{Z_0} e^{-\beta \hat{H}_0}$. First note $\chi_{AB}(t, t_1)$ can be written as

$$\chi_{AB}(t, t_1) = \frac{i}{\hbar}(C_{AB}(t, t_1) - C_{BA}(t_1, t))$$

where $C_{AB}(t, t_1)$ is

$$C_{AB}(t, t_1) = \mathrm{Tr}\left[\hat{A}^I(t)\hat{B}^I(t_1)\hat{\rho}_0\right]$$

Henceforth we drop the $\hat{A}^I(t) = \hat{A}(t)$, but remember the time evolution is generated by $\hat{H}_0$, and not $\hat{H}(t)$. If $[\hat{H}_0, \hat{\rho}_0] = 0$ this can be written as

$$\begin{aligned}
C_{AB}(t, t_1) &= \mathrm{Tr}\left[\hat{U}_0^\dagger(t)\hat{A}\hat{U}_0(t)\hat{U}_0^\dagger(t_1)\hat{B}\hat{U}_0(t_1)\hat{\rho}_0\right] \\
&= \mathrm{Tr}\left[\hat{U}_0(t_1)\hat{U}_0^\dagger(t)\hat{A}\hat{U}_0(t)\hat{U}_0^\dagger(t_1)\hat{B}\hat{\rho}_0\right] \\
&= \mathrm{Tr}\left[\hat{U}_0^\dagger(t - t_1)\hat{A}\hat{U}_0(t - t_1)\hat{B}\hat{\rho}_0\right] \\
&= C_{AB}(t - t_1, 0) \\
&\equiv C_{AB}(t - t_1)
\end{aligned}$$

So $\chi_{AB}(t, t_1) \equiv \chi_{AB}(t - t_1)$. Likewise $\langle A^I(t)\rangle_0 = \langle A\rangle_0$

So linear-response theory simplifies to:

$$\langle A(t)\rangle = \langle A\rangle_0 - \int_0^t f(t_1)\chi_{AV}(t - t_1)\, dt_1$$

4.2 Correlation functions

We see that the linear response is entirely encoded in the correlation function

$$C_{AB}(t) = \mathrm{Tr}\left[\hat{A}(t)\hat{B}\hat{\rho}_0\right]$$

where we will continue to assume $[\hat{H}_0, \hat{\rho}_0] = 0$.

Let's consider some properties of this:

$$\begin{aligned}
C_{AB}(-t) &= \mathrm{Tr}\left[\hat{U}^\dagger(-t)\hat{A}\hat{U}(-t)\hat{B}\hat{\rho}_0\right] \\
&= \mathrm{Tr}\left[\hat{A}\hat{B}(t)\hat{\rho}_0\right] \\
&= \mathrm{Tr}\left[\hat{A}\hat{B}(t)\hat{\rho}_0\right]^{**} \\
&= \mathrm{Tr}\left[\hat{\rho}_0^\dagger\hat{B}^\dagger(t)\hat{A}^\dagger\right] \\
&= C_{B^\dagger A^\dagger}(t)^*
\end{aligned}$$

If $\hat{B} = \hat{B}^\dagger$ and $\hat{A} = \hat{A}^\dagger$,

$$C_{AB}(-t) = C_{BA}(t)^*$$

A short aside on the Kubo-transformed correlation function

For thermal states we can find a more symmetric form of the correlation function as

$$C^K_{AB}(t) = \frac{1}{Z_0\beta}\int_0^\beta d\lambda \ \mathrm{Tr}\left[\hat{A}(t)e^{-(\beta-\lambda)\hat{H}_0}\hat{B}e^{-\lambda\hat{H}_0}\right]$$

This is called the Kubo-transformed correlation function.

This satisfies

$$C^K_{AB}(t) = C^K_{B^\dagger A^\dagger}(-t)^*$$

(the proof is very similar to before.) But also

$$\begin{aligned} C^K_{AB}(t) &= \frac{1}{Z_0\beta}\int_0^\beta d\lambda \ \mathrm{Tr}\left[e^{-\lambda\hat{H}_0}\hat{A}e^{-(\beta-\lambda)\hat{H}_0}\hat{B}(-t)\right] \\ &= \frac{1}{Z_0\beta}\int_0^\beta d\lambda' \ \mathrm{Tr}\left[e^{-(\beta-\lambda')\hat{H}}\hat{A}e^{\lambda'\hat{H}_0}\hat{B}(-t)\right] \\ &= C^K_{BA}(-t) \end{aligned}$$

So for $\hat{A} = \hat{A}^\dagger$ and $\hat{B} = \hat{B}^\dagger$

$$\begin{aligned} C^K_{AB}(t)^* &= C^K_{BA}(-t) \\ &= C^K_{AB}(t) \end{aligned}$$

So $C^K_{AB}(t)$ is real-valued. The Kubo-transformed correlation function contains the same information as the standard thermal $C_{AB}(t)$ but it has a more natural connection to classical dynamics (more on this later).

4.3 Useful Properties of the Response Function

Note that the response function is given by:

$$\chi_{AB}(t) = \frac{i}{\hbar}\left(C_{AB}(t) - C_{BA}(-t)\right)$$

And the conjugate:

$$\begin{aligned} \chi_{AB}(t)^* &= -\frac{i}{\hbar}\left(C_{AB}(t)^* - C_{BA}(-t)^*\right) \\ &= -\frac{i}{\hbar}\left(C_{BA}(-t) - C_{AB}(t)\right) \\ &= \frac{i}{\hbar}\left(C_{AB}(t) - C_{BA}(-t)\right) \\ &= \chi_{AB}(t) \qquad \Longrightarrow \text{ Real} \end{aligned}$$

Checking time reversal:

$$\begin{aligned} \chi_{AA}(-t) &= \frac{i}{\hbar}\left(C_{AA}(-t) - C_{AA}(t)\right) \\ &= -\chi_{AA}(t) \qquad \Longrightarrow \text{ Odd} \end{aligned}$$

So $\chi_{AA}(t)$ is real and odd.

4.4 Linear Response and the Absorption Spectrum

Why is the absorption spectrum related to $\tilde{\chi}_{\mu\mu}(\omega)$?

Consider Maxwell's equations for light propagating along the x direction with polarisation included. Let $\boldsymbol{E}(\boldsymbol{x}, t) = E(x,t)\boldsymbol{e}_z$ (polarised along z).

The wave equation is:

$$\frac{\partial^2 E(x,t)}{\partial x^2} - \frac{1}{c^2}\frac{\partial^2 E(x,t)}{\partial t^2} = \frac{1}{c^2\epsilon_0}\frac{\partial^2 P(x,t)}{\partial t^2}$$

where $P(x,t)$ is the polarisation along the z direction. Since $P(x,t)$ is local, we have:

$$P = \text{average dipole per unit volume}$$
$$P(x,t) = \epsilon_0 \int_{-\infty}^{\infty} dt' \chi(x, t-t') E(x,t')$$

assuming Linear Response, with causality requiring $\chi(x, t<0) = 0$. This falls naturally out of linear response theory. To relate $\chi(t)$ to microscopic quantum linear response theory, note

$$P(x,t) = \rho \langle \mu_z(t) \rangle = \rho \int_{-\infty}^{t} \chi_{\mu_z\mu_z}(t-\tau) E(x,\tau) d\tau$$

Comparing to macroscopic definition:

$$\epsilon_0 \chi(x,t) = \rho\,\Theta(t)\,\chi_{\mu_z\mu_z}(t)$$

Taking the Fourier transform of the above equation (with $e^{i\omega t}$):

$$\frac{\partial^2 \tilde{E}(x,\omega)}{\partial x^2} + \frac{\omega^2}{c^2}\tilde{E}(x,\omega) = -\frac{\omega^2}{c^2}\tilde{\chi}(x,\omega)\tilde{E}(x,\omega)$$
$$\frac{\partial^2 \tilde{E}(x,\omega)}{\partial x^2} + \frac{\omega^2}{c^2}(1+\tilde{\chi}(x,\omega))\tilde{E}(x,\omega) = 0$$

Let's assume we have the material of interest between $x = 0$ and $x = L$. This means $\tilde{\chi}(x,\omega) = 0$ outside of this region.

- **Outside the medium** ($x < 0$ and $x > L$):

$$\frac{\partial^2 \tilde{E}(x,\omega)}{\partial x^2} + \frac{\omega^2}{c^2}\tilde{E}(x,\omega) = 0$$

- **Inside the medium** ($0 \leq x \leq L$):

$$\frac{\partial^2 \tilde{E}(x,\omega)}{\partial x^2} + \frac{\omega^2}{c^2}(1+\tilde{\chi}(\omega))\tilde{E}(x,\omega) = 0$$

Solving these with the source to the left (at $x < 0$):

$$\begin{aligned} x < 0: &\quad \tilde{E}(x,\omega) = \tilde{E}_i(\omega)e^{ik_0x} + \tilde{E}_r(\omega)e^{-ik_0x} \\ 0 \leq x \leq L: &\quad \tilde{E}(x,\omega) = \tilde{E}_f(\omega)e^{ik(\omega)x} + \tilde{E}_b e^{-ik(\omega)x} \\ x > L: &\quad \tilde{E}(x,\omega) = \tilde{E}_t(\omega)e^{ik_0x} \end{aligned}$$

Assuming minimal reflection, $\tilde{E}_r \approx 0, \tilde{E}_b \approx 0$ (if you work through with continuity and derivative continuity this falls out if $|k(\omega)| \approx k_0$). Matching $|\tilde{E}(x,\omega)|^2$ at the boundaries, we find:

$$|\tilde{E}_t(\omega)|^2 = |\tilde{E}_i(\omega)|^2 e^{i(k(\omega)-k(\omega)^*)L}$$

Now, analyzing $k(\omega)$:

$$
\begin{aligned}
k(\omega)^2 = \frac{\omega^2}{c^2}(1+\tilde{\chi}(\omega)) &\implies k(\omega) \approx \frac{\omega}{c}\sqrt{1+\tilde{\chi}(\omega)} \\
\implies k(\omega) &\simeq \frac{\omega}{c} + \frac{\omega}{2c}\tilde{\chi}(\omega) \\
\implies k(\omega) - k(\omega)^* &\simeq \frac{\omega}{2c}(\tilde{\chi}(\omega) - \tilde{\chi}(\omega)^*) \\
&= i\frac{\omega}{c}\,\mathrm{Im}[\tilde{\chi}(\omega)]
\end{aligned}
$$

So the transmitted intensity $I_t \propto |\tilde{E}_t(\omega)|^2$ is:

$$
I_t = I_i e^{-\frac{\omega}{c}\,\mathrm{Im}[\tilde{\chi}(\omega)]L}
$$

Comparing this to the Beer-Lambert law ($I = I_0 e^{-\alpha\rho L}$):

$$
\begin{aligned}
I_t &= I_i e^{-\alpha\rho L} \\
\therefore \quad \alpha\rho &= \frac{\omega}{c}\,\mathrm{Im}[\tilde{\chi}(\omega)]
\end{aligned}
$$

Taking the Fourier transform of the linear response result gives

$$
\tilde{\chi}(\omega) = \frac{\rho}{\epsilon_0}\int_0^\infty e^{i\omega t}\chi_{\mu_z\mu_z}(t)dt
$$

And the spectrum is therefore proportional to $\alpha(\omega) \propto \mathrm{Im}\,\tilde{\chi}_{\mu_z\mu_z}(\omega)$.

Now considering the integral for the imaginary part:

$$
\begin{aligned}
\mathrm{Im}[\tilde{\chi}(\omega)] &= \mathrm{Im}\left[\frac{\rho}{\epsilon_0}\int_0^\infty e^{i\omega t}\chi_{\mu\mu}(t)dt\right] \\
&= \frac{\rho}{\epsilon_0}\frac{1}{2i}\left(\int_0^\infty e^{i\omega t}\chi_{\mu\mu}(t)dt - \int_0^\infty e^{-i\omega t}\chi_{\mu\mu}(t)dt\right) \\
&= \frac{\rho}{2\epsilon_0}\int_{-\infty}^\infty e^{i\omega t}\chi_{\mu\mu}(t)dt
\end{aligned}
$$

(Using the symmetry properties derived above).

4.5 Fermi's Golden Rule from Linear Response

Suppose we have an initial state $|i\rangle$ and a family of final states $|f,\epsilon\rangle$ with density of states $\rho_f(\epsilon)$, such that:

$$
\begin{aligned}
\hat{H}_0\,|i\rangle &= E_i\,|i\rangle \\
\hat{H}_0\,|f,\epsilon\rangle &= \epsilon\,|f,\epsilon\rangle
\end{aligned}
$$

How do we find the rate if a coupling perturbation $\hat{V}$ is turned on at time $t = 0$?

First, note for $t \geq 0$, with the projection operator $\hat{P}_f = \int d\epsilon\,\rho_f(\epsilon)\,|f,\epsilon\rangle\langle f,\epsilon|$:

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\left\langle \hat{P}_f\right\rangle &= \frac{i}{\hbar}\left\langle [\hat{H},\hat{P}_f]\right\rangle \\
&= \frac{i}{\hbar}\left\langle [\hat{V},\hat{P}_f]\right\rangle
\end{aligned}
$$

(Since $[\hat{H}_0,\hat{P}_f] = 0$).

The commutator is:

$$[\hat{V}, \hat{P}_f] = \sum_{n \neq f} \int d\epsilon\, \rho_f(\epsilon)\, (V_{nf}(\epsilon)\, |n\rangle\langle f, \epsilon| - V_{fn}(\epsilon)\, |f, \epsilon\rangle\langle n|)$$

This defines the "flux" operator from $i \to f$ (taking $n = i$):

$$\hat{F}_{if} = \frac{i}{\hbar} \int d\epsilon\, \rho_f(\epsilon)\, (V_{if}(\epsilon)\, |i\rangle\langle f, \epsilon| - V_{fi}(\epsilon)\, |f, \epsilon\rangle\langle i|)$$

The instantaneous rate is $\langle F_{if} \rangle$. Take $\hat{\rho}_0 = |i\rangle\langle i|$ and apply Linear Response theory with $A = F_{if}$ and $\hat{V}(t) = h(t)\hat{V}$:

$$\langle F_{if} \rangle = -\int_0^t d\tau\, \chi_{F_{if}V}(\tau)$$

where the response function is:

$$\chi_{F_{if}V}(t) = \frac{i}{\hbar} \operatorname{Tr}\left[[\hat{F}^I_{if}(t), \hat{V}]\hat{\rho}_0 \right]$$

4.5.1 Evaluation of the Response Function

Evaluating the trace with $\hat{\rho}_0 = |i\rangle\langle i|$:

$$\begin{aligned} \chi_{F_{if}V}(t) &= \frac{i}{\hbar} \langle i|[\hat{F}^I_{if}(t), \hat{V}]|i\rangle \\ &= \frac{i}{\hbar} \Big(\langle i|\hat{F}^I_{if}(t)\hat{V}|i\rangle - \langle i|\hat{V}\hat{F}^I_{if}(t)|i\rangle \Big) \end{aligned}$$

Using the interaction picture definition $\hat{F}^I_{if}(t) = e^{i\hat{H}_0 t/\hbar}\hat{F}_{if}e^{-i\hat{H}_0 t/\hbar}$:

$$\langle i|\hat{F}^I_{if}(t)\hat{V}|i\rangle = \frac{i}{\hbar} \int d\epsilon\, e^{i(E_i - \epsilon)t/\hbar} V_{if}(\epsilon)\, \langle f, \epsilon|\hat{V}|i\rangle$$

Using the property $\int d\epsilon\, \langle f, \epsilon'|f, \epsilon\rangle\, \rho_f(\epsilon') = \delta(\epsilon - \epsilon')$:

$$\langle i|\hat{F}^I_{if}(t)\hat{V}|i\rangle = \frac{i}{\hbar} \int d\epsilon\, e^{i(E_i - \epsilon)t/\hbar} \rho_f(\epsilon)|V_{if}(\epsilon)|^2$$

Substituting this back into the expression for χ:

$$\begin{aligned} \chi_{F_{if}V}(t) &= \left(\frac{i}{\hbar}\right)^2 \int d\epsilon\, \rho_f(\epsilon)|V_{if}(\epsilon)|^2 \left(e^{i(E_i - \epsilon)t/\hbar} + e^{-i(E_i - \epsilon)t/\hbar}\right) \\ &= -\frac{2}{\hbar^2} \int d\epsilon\, \rho_f(\epsilon)|V_{if}(\epsilon)|^2 \cos\left(\frac{E_i - \epsilon}{\hbar} t\right) \end{aligned}$$

4.5.2 The Rate Constant

For short times, the flux has contributions from all $|f, \epsilon\rangle$, but over time the oscillations average out leaving only $E_i = \epsilon$ terms. The long-time limit of the flux is the rate constant:

$$\begin{aligned} k_{i \to f} = \langle F_{if}(t \to \infty) \rangle &= -\int_0^\infty d\tau\, \chi_{F_{if}V}(\tau) \\ &= \frac{2}{\hbar^2} \int d\epsilon\, \rho_f(\epsilon)|V_{if}(\epsilon)|^2 \int_0^\infty d\tau \cos\left(\frac{E_i - \epsilon}{\hbar}\tau\right) \end{aligned}$$

Using the identity $\int_0^\infty \cos(\omega t)dt = \pi\delta(\omega)$:

$$\begin{aligned} k_{i \to f} &= \frac{2}{\hbar^2} \int d\epsilon\, \rho_f(\epsilon)|V_{if}(\epsilon)|^2 \cdot \pi\hbar\delta(E_i - \epsilon) \\ &= \frac{2\pi}{\hbar} |V_{if}(E_i)|^2 \rho_f(E_i) \end{aligned}$$

This is exactly Fermi's Golden Rule!

5 Open Quantum Systems and Quantum Master Equations

5.1 Open Systems and Reduced Density Operators

Suppose a quantum system can be partitioned into two sub systems A and B, so the full Hilbert space is

$$\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$$

and basis states are

$$|n_A, m_B\rangle = |n_A\rangle \otimes |m_B\rangle$$

with $|n_A\rangle \in \mathcal{H}_A$ and $|m_B\rangle \in \mathcal{H}_B$.

The full density operator for this system is $\hat{\rho}$.

We now define the partial trace over A or B by

$$\begin{aligned}\mathrm{Tr}_A[|n_A, m_B\rangle \langle n'_A, m'_B|] &= |m_B\rangle \langle m'_B| \, \delta_{nn'} \\ \mathrm{Tr}_B[|n_A, m_B\rangle \langle n'_A, m'_B|] &= |n_A\rangle \langle n'_A| \, \delta_{mm'}\end{aligned}$$

So this reduces an operator from an operator on $\mathcal{H}$ to $\mathcal{H}_B$ or $\mathcal{H}_A$.

We can define the full trace in terms of partial traces as

$$\mathrm{Tr}\left[\hat{A}\right] = \mathrm{Tr}_A[\mathrm{Tr}_B[\hat{A}]] = \mathrm{Tr}_B[\mathrm{Tr}_A[\hat{A}]]$$

We can now define the partial reduced density operators $\hat{\sigma}_A$ and $\hat{\sigma}_B$ as

$$\hat{\sigma}_A = \mathrm{Tr}_B[\hat{\rho}] \,, \quad \hat{\sigma}_B = \mathrm{Tr}_A[\hat{\rho}]$$

If $\hat{O}_A$ is an operator that only acts on $\mathcal{H}_A$, then it is straightforward to show

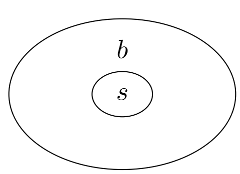
$$\begin{aligned}\langle O_A \rangle &= \mathrm{Tr}\left[\hat{O}_A \hat{\rho}\right] \\ &= \mathrm{Tr}_A[\mathrm{Tr}_B[\hat{O}_A \hat{\rho}]] \\ &= \mathrm{Tr}_A[\hat{O}_A \, \mathrm{Tr}_B[\hat{\rho}]] \\ &= \mathrm{Tr}_A[\hat{O}_A \hat{\sigma}_A]\end{aligned}$$

So $\hat{\sigma}_A$ contains all information about observables of A and likewise $\hat{\sigma}_B$ contains information on observables of B.

Why is this useful?

Often in real systems we have a system of interest, s, and an environment, often called a "bath", b. Our experiments often only probe the system, so which consists of far fewer degrees of freedom than the full system + bath. This is why we often work with $\hat{\sigma}_s$.

This is illustrated here:



$A = s$ the subsystem

$B = b$ the environment or "bath"

The system might be a molecule and the bath might be a solvent or electromagnetic fields or vibrational degrees of freedom of a molecule. The most sensible partitioning depends on the problem at hand.

This is what is generally called an open quantum system. The subsystem of interest can exchange energy (and also potentially particles) with the bath through a system bath interaction, so in this sense it is open, as opposed to a closed quantum system that has a fixed energy and number of particles. What we aim to find is an effective equation of motion for the subsystem recduced density operator, which is analogous to the Liouville von Neumann equation, but which accounts for the effect of the bath. We do this because the full system is generally too complicated to solve in its entirety.

Note we'll play a bit fast-and-loose with use of the term "system" and "subsystem". "System" may refer to the total system or, just the subsystem. I'll try to use the term "total system" when referring to the whole subsystem+bath, and the term "system" will mostly be used to refer just to the subsystem.

5.2 System–bath interactions

The full Hamiltonian of a total system can be written as

$$\hat{H} = \hat{H}_s + \hat{H}_b + \hat{H}_{sb}$$

$\hat{H}_s$ and $\hat{H}_b$ only act on the subsystem and bath, but $\hat{V}_{sb}$ couples the two. $\hat{V}_{sb}$ can always be decomposed as

$$\hat{V}_{sb} = \sum_a \hat{S}_a \hat{B}_a$$

where $\hat{S}_a$ acts on s and $\hat{B}_a$ on b. For simplicity we'll just consider $\hat{V}_{sb} = \hat{S}\hat{B}$, but the generalization to multiple $\hat{S}_a\hat{B}_a$ is relatively simple.

Question: Can we find an equation of motion for $\hat{\sigma}_s$?

Let's try starting from the LVN equation for $\hat{\rho}(t)$

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\hat{\rho}(t) &= -\frac{i}{\hbar}[\hat{H}_s + \hat{H}_b + \hat{H}_{sb}, \hat{\rho}(t)] \\
\frac{\mathrm{d}}{\mathrm{d}t}\hat{\sigma}_s(t) &= \mathrm{Tr}_b\left[\frac{\mathrm{d}}{\mathrm{d}t}\hat{\rho}(t)\right] \\
&= -\frac{i}{\hbar}\,\mathrm{Tr}_b\left[[\hat{H}_s, \hat{\rho}(t)]\right] - \frac{i}{\hbar}\,\mathrm{Tr}_b\left[[\hat{H}_b, \hat{\rho}(t)]\right] \\
&\quad - \frac{i}{\hbar}\,\mathrm{Tr}_b\left[[\hat{H}_{sb}, \hat{\rho}(t)]\right] \\
&= -\frac{i}{\hbar}[\hat{H}_s, \mathrm{Tr}_b[\hat{\rho}(t)]] - \frac{i}{\hbar}\,\mathrm{Tr}_b\left[[\hat{H}_{sb}, \hat{\rho}(t)]\right] \\
&= -\frac{i}{\hbar}[\hat{H}_s, \hat{\sigma}_s(t)] - \frac{i}{\hbar}\,\mathrm{Tr}_b\left[[\hat{H}_{sb}, \hat{\rho}(t)]\right]
\end{aligned}$$

The first term is the LVN equation as if the bath was not present. The second term can't be simplified exactly though.

The bath interaction introduces correlations between the system and bath that cannot be easily simplified.

5.3 Interaction picture

Let's try another approach. First let's move to the interaction picture (as we did in perturbation theory).

$$\begin{aligned}
\hat{H}_0 &= \hat{H}_s + \hat{H}_b\,, \quad \hat{V} = \hat{H}_{sb} \\
\hat{\rho}^I(t) &= e^{+i\hat{H}_0 t/\hbar}\hat{\rho}(t)e^{-i\hat{H}_0 t/\hbar} \quad \text{(Same def. for } \hat{V}^I(t) \text{ etc.)}
\end{aligned}$$

Analogous to the TDSE in the interaction picture we find,

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{\rho}^I(t) = -\frac{i}{\hbar}[\hat{V}^I(t), \hat{\rho}^I(t)] \tag{1}$$

Let's define $\hat{\sigma}_s^I(t) = e^{+i\hat{H}_s t/\hbar}\hat{\sigma}_s(t)e^{-i\hat{H}_s t/\hbar}$. This is equal to $\mathrm{Tr}_b[\hat{\rho}^I(t)]$. Now let's integrate this to give

$$\hat{\rho}^I(t) = \hat{\rho}^I(0) - \frac{i}{\hbar}\int_0^t d\tau\,[\hat{V}^I(\tau), \hat{\rho}^I(\tau)] \tag{2}$$

and substitute this into the equation for $\frac{\mathrm{d}}{\mathrm{d}t}\hat{\rho}^I(t)$

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{\rho}^I(t) = -\frac{i}{\hbar}[\hat{V}^I(t), \hat{\rho}^I(0)] - \frac{i}{\hbar}\int_0^t d\tau\,[\hat{V}^I(t), [\hat{V}^I(\tau), \hat{\rho}^I(\tau)]] \tag{3}$$

We don't seem any closer to an equation just for $\hat{\sigma}_s(t)\ldots$

Let's assume initially the subsystem and bath are uncorrelated.

$$\hat{\rho}(0) = \hat{\sigma}_s(0)\hat{\rho}_b$$

Taking $\mathrm{Tr}_b[\cdot]$ of (1) gives

$$\hat{\sigma}_s^I(t) = \hat{\sigma}_s^I(0) - \frac{i}{\hbar}\int_0^t d\tau\ \mathrm{Tr}_b[[\hat{V}^I(\tau), \hat{\rho}^I(\tau)]]$$

So $\hat{\sigma}_s^I(t) = \hat{\sigma}_s^I(\tau) + \mathcal{O}(V^2)$.

Taking Tr_b of the first term in (3) gives

$$\mathrm{Tr}_b[[\hat{V}^I(t), \hat{\rho}^I(0)]] = \hat{S}^I(t)\,\mathrm{Tr}_b[\hat{B}^I(t)\hat{\rho}_b]\hat{\sigma}_s^I(0) - \hat{\sigma}_s^I(0)\,\mathrm{Tr}_b[\hat{\rho}_b\hat{B}^I(t)]\hat{S}^I(t)$$

Now assuming $\mathrm{Tr}_b[\hat{B}^I(t)\hat{\rho}_b] = 0$, as we generally can, this term vanishes. Often by assuming $\hat{\rho}_b^I(t) = \hat{\rho}_b$ and $\mathrm{Tr}_b[\hat{B}\hat{\rho}_b] = 0$. Now in (3) in the second term we can replace $\hat{\rho}^I(\tau)$ with

$$\begin{aligned}\hat{\rho}^I(\tau) &= \hat{\sigma}^I(\tau)\hat{\rho}_b \quad + \mathcal{O}(V)\\ &= \hat{\sigma}^I(t)\hat{\rho}_b \quad + \mathcal{O}(V)\end{aligned}$$

This is called the Markov approximation. The full equation includes a memory term, so we'd need to know the density operator and previous times to predict its dynamics. The Markovian approximation simplifies this term by removing the memory.

Now taking Tr_b of (3) gives

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{\sigma}_s^I(t) = -\frac{1}{\hbar^2}\int_0^t d\tau\ \mathrm{Tr}_b\left[[\hat{V}^I(t), [\hat{V}^I(\tau), \hat{\sigma}_s^I(t)\hat{\rho}_b]]\right]$$

Now let's transform back to the Schrödinger picture

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\hat{\sigma}_s(t) = &-\frac{i}{\hbar}[\hat{H}_s, \hat{\sigma}_s(t)]\\ &-\frac{i}{\hbar}\int_0^t d\tau\ \mathrm{Tr}_b\left[[\hat{V}, [\hat{V}^I(\tau - t), \hat{\sigma}_s(t)\hat{\rho}_b]]\right]\end{aligned}$$

Let's expand out the last term after $\tau' = t - \tau$ substitution.

$$\begin{aligned}\mathrm{Tr}_b[[\hat{V}, [\hat{V}^I(-\tau'), \hat{\sigma}_s\hat{\rho}_b]]] &= \mathrm{Tr}_b[\hat{S}\hat{B}\hat{S}^I(-\tau')\hat{B}(-\tau')\hat{\sigma}_s\hat{\rho}_b] - \mathrm{Tr}_b[\hat{S}\hat{B}\hat{\sigma}_s\hat{\rho}_b\hat{S}^I(-\tau')\hat{B}(-\tau')]\\ &\quad - \mathrm{Tr}_b[\hat{S}^I(-\tau')\hat{B}(-\tau')\hat{\sigma}_s\hat{\rho}_b\hat{S}\hat{B}] + \mathrm{Tr}_b[\hat{\sigma}_s\hat{\rho}_b\hat{S}^I(-\tau')\hat{B}(-\tau')\hat{S}\hat{B}]\\ &= \mathrm{Tr}_b[\hat{B}\hat{B}(-\tau')\hat{\rho}_b]\hat{S}\hat{S}^I(-\tau')\hat{\sigma}_s - \mathrm{Tr}_b[\hat{B}\hat{\rho}_b\hat{B}(-\tau')]\hat{S}\hat{\sigma}_s\hat{S}^I(-\tau')\\ &\quad - \mathrm{Tr}_b[\hat{B}(-\tau')\hat{\rho}_b\hat{B}]\hat{S}^I(-\tau')\hat{\sigma}_s\hat{S} + \mathrm{Tr}_b[\hat{\rho}_b\hat{B}(-\tau')\hat{B}]\hat{\sigma}_s\hat{S}^I(-\tau')\hat{S}\end{aligned}$$

Define $C(\tau) = \mathrm{Tr}_b[\hat{B}^I(\tau)\hat{B}\hat{\rho}_b]$, and this becomes

$$\begin{aligned}\mathrm{Tr}_b[[\hat{V}, [\hat{V}^I(-\tau'), \hat{\sigma}_s\hat{\rho}_b]]] = &\hat{S}C(\tau')\hat{S}^I(-\tau')\hat{\sigma}_s - C(\tau')^*\hat{S}\hat{\sigma}_s\hat{S}^I(-\tau') \\ &- C(\tau')\hat{S}^I(-\tau')\hat{\sigma}_s\hat{S} + C(\tau')^*\hat{\sigma}_s\hat{S}^I(-\tau')\hat{S}\end{aligned}$$

We see now when we expect the Markov approximation to be valid, which is when the correlation terms decay fast compared to the dynamics of the density operator.

We can expand $\hat{S}^I(-\tau')$ in the eigenstates of $\hat{H}_s$

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

$$\begin{aligned}\hat{S}^I(-\tau') &= \sum_{nm} \hat{U}_s(-\tau') |n\rangle \langle n|\hat{S}|m\rangle \langle m| \hat{U}_s^\dagger(-\tau') \\ &= \sum_{nm} e^{-i(E_n - E_m)t/\hbar} S_{nm} |n\rangle \langle m|\end{aligned}$$

So taking the integral gives terms of the form

$$\int_0^t C(\tau)e^{-i\omega_{nm}t/\hbar} S_{nm} |n\rangle \langle m|$$

Assuming $C(\tau)$ decays rapidly, we can take the integral time limits to ∞

We define

$$\hat{\Lambda} = \int_0^\infty d\tau\, C(\tau)\hat{S}^I(-\tau')$$

and the equation for $\hat{\sigma}_s$ becomes

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\hat{\sigma}_s(t) &= -\frac{i}{\hbar}[\hat{H}_s, \hat{\sigma}_s(t)] - \frac{1}{\hbar^2}[\hat{S}, \hat{\Lambda}\hat{\sigma}_s(t) - \hat{\sigma}_s\hat{\Lambda}^\dagger] \\ &= \hat{\hat{\mathcal{L}}}_s\hat{\sigma}_s(t) + \hat{\hat{\mathcal{R}}}\hat{\sigma}_s(t)\end{aligned}$$

This is the Redfield equation (in its full "non-secular" form). The Redfield superoperator $\hat{R}$ describes the relaxation of the system induced by the bath. Sometimes this is also called the Born-Markov quantum master equation.

5.4 Secular Redfield Theory

In its current form Redfield theory is a little unwieldy. Let's move back to the interaction picture to see what happens.

First let's define

$$\langle a|(\hat{\hat{\mathcal{R}}} |c\rangle \langle l|)|b\rangle = \mathcal{R}_{ab,cl}$$

So

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{\sigma}_s^I(t) = \hat{U}_s(t) \left(\hat{\hat{\mathcal{R}}}(\hat{U}_s(t)\hat{\sigma}_s^I(t)\hat{U}_s^\dagger(t)) \right) \hat{U}_s^\dagger(t)$$

Taking component $\langle a|\hat{\sigma}_s^I(t)|b\rangle = \sigma_{s,ab}^I(t)$ gives

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\sigma_{s,ab}^I(t) &= e^{+i(E_a - E_b)t/\hbar} \sum_{cd} \hat{\hat{\mathcal{R}}} |c\rangle \langle d| \, \sigma_{s,cl}^I(t) e^{-i(E_c - E_d)t/\hbar} \\ &= \sum_{cd} e^{+i(\omega_{ab} - \omega_{cd})t} \mathcal{R}_{ab,cd}\sigma_{s,cd}^I(t)\end{aligned}$$

Integrating this gives

$$\sigma^I_{s,ab}(t) = \int_0^t d\tau \sum_{cd} e^{+i(\omega_{ab}-\omega_{cd})\tau} R_{ab,cd}\sigma^I_{s,cd}(0) + \cdots$$

Doing the integral gives (for $\omega_{ab} \neq \omega_{cd}$)

$$\int_0^t d\tau\, e^{+i(\omega_{ab}-\omega_{cd})\tau} = \frac{e^{i(\omega_{ab}-\omega_{cd})t} - 1}{i(\omega_{ab}-\omega_{cd})}$$

The dominant terms, therefore are $\omega_{ab} = \omega_{cd}$ terms which grow linearly in time. In the secular approximation we retain only these terms, where

$$E_a - E_b - E_c + E_d = 0$$

So we retain terms with

$$\begin{aligned} 1) \quad & a = b\,, \quad c = d \\ \text{or } 2) \quad & a = c\,, \quad b = d \end{aligned}$$

The terms corresponding to case 1) are

$$\mathcal{R}_{aabb}$$

These connect populations in the reduced density operator to populations so they can be interpreted as rate constants

$$\mathcal{R}_{aabb} = k_{a\leftarrow b} = k_{b\rightarrow a}$$

The case 2) terms are

$$\mathcal{R}_{abab} = -\gamma_{ab} - i\Delta\omega_{ab}$$

So these cause decay through the real part and a change in oscillation frequency due to the imaginary part.

5.5 Expressions for Secular RE Redfield Rates

First let's consider R_{aabb} with $a \neq b$

$$\begin{aligned} \mathcal{R}_{aabb} &= -\frac{1}{\hbar^2} \langle a|[\hat{S}, \hat{\Lambda}\,|b\rangle\,\langle b| - |b\rangle\,\langle b|\,\hat{\Lambda}^\dagger]|a\rangle \\ &= +\frac{1}{\hbar^2}\Big(\,\langle a|\hat{S}|b\rangle\,\langle b|\hat{\Lambda}^\dagger|a\rangle + \langle a|\hat{\Lambda}|b\rangle\,\langle b|\hat{S}|a\rangle\Big) \end{aligned}$$

We have

$$\langle a|\hat{\Lambda}|b\rangle = S_{ab}\int_0^\infty d\tau\, C(\tau)e^{-i\omega_{ab}\tau} = S_{ab}\tilde{J}(\omega_{ab})$$

$$\Rightarrow \mathcal{R}_{aabb} = +\frac{1}{\hbar^2}|S_{ab}|^2 \cdot 2\text{Re}[\tilde{J}(\omega_{ab})]$$

And likewise

$$\mathcal{R}_{aa,aa} = -\sum_{b\neq a} R_{bb,aa}$$

In the homework problems you will show that

$$2\text{Re}\int_0^\infty C(\tau)e^{-i\omega\tau} = \int_{-\infty}^{+\infty} C(\tau)e^{-i\omega\tau}d\tau = \tilde{C}(\omega)$$

and also

$$\tilde{C}(-\omega) = \tilde{C}(\omega)e^{+i\hbar\omega\beta}$$

So Redfield theory predicts rates that satisfy

$$\frac{k_{b\to a}}{k_{a\to b}} = \frac{\mathcal{R}_{aabb}}{\mathcal{R}_{bbaa}} = \frac{\tilde{C}(\omega_{ab})}{\tilde{C}(-\omega_{ab})} = e^{-\beta(E_a - E_b)}$$

So the steady state for the system evolving under the Redfield equation will be the thermal distribution.

$$\hat{\sigma}_s(t \to \infty) = \frac{1}{Z_s}\sum_n |n\rangle\langle n| \, e^{-\beta E_n}$$

Coherences $\hat{\sigma}^I_{sab}(t)$ evolve with the rate (frequency shift

$$\begin{aligned}
\mathcal{R}_{abab} &= -\frac{1}{\hbar^2}\langle a|[\hat{S}, \hat{\Lambda}\,|a\rangle\langle b| - |a\rangle\langle b|\,\hat{\Lambda}^\dagger]|b\rangle \\
&= -\frac{1}{\hbar^2}\Big(\langle a|\hat{S}\hat{\Lambda}|b\rangle + \langle a|\hat{\Lambda}^\dagger\hat{S}|b\rangle - \langle a|\hat{S}\,|a\rangle\langle b|\,\hat{\Lambda}^\dagger|b\rangle - \langle a|\hat{\Lambda}\,|a\rangle\langle b|\,\hat{S}|b\rangle\Big) \\
&= -\frac{1}{\hbar^2}\left(\sum_c (S_{ac}S_{ca}\tilde{J}(\omega_{ca}) + S_{bc}S_{cb}\tilde{J}(\omega_{cb})^*)\right) + \frac{1}{\hbar^2}(2S_{aa}S_{bb}\mathrm{Re}[\tilde{J}(0)])
\end{aligned}$$

The sums can be re-written in terms of real and imaginary parts tp give

$$\begin{aligned}
\mathcal{R}_{abab} = &-\frac{1}{\hbar^2}\left(\sum_{c\neq a}|S_{ac}|^2\,\mathrm{Re}\,\tilde{J}(\omega_{ca}) + \sum_{c\neq b}|S_{bc}|^2\,\mathrm{Re}\,\tilde{J}(\omega_{cb})\right) + \frac{1}{\hbar^2}(2S_{aa}S_{bb} - S_{aa}^2 - S_{bb}^2)\mathrm{Re}[\tilde{J}(0)]) \\
&-\frac{i}{\hbar^2}\sum_c\Big(|S_{ac}|^2\,\mathrm{Im}\,\tilde{J}(\omega_{ca}) - |S_{bc}|^2\,\mathrm{Im}\,\tilde{J}(\omega_{ba})\Big) \\
&= +\frac{1}{2}\left(\mathcal{R}_{aaaa} + \mathcal{R}_{bbbb}\right) - \frac{i}{\hbar}(H_{LS,aa} - H_{LS,bb}) - (S_{aa} - S_{bb})^2\,\mathrm{Re}[\tilde{J}(0)]
\end{aligned}$$

This $\mathcal{R}_{abab}$ term has three distinct terms. The first is a population decay term, this ensures coherences decay as populations decay.

The second term $i\omega^{LS}_{ab} = \frac{i}{\hbar}(H_{LS,aa} - H_{LS,bb})$ is imaginary so it contributes to a frequency shift in the dynamics of the system. This is called a Lamb shift term. It's often a small effect so it's frequently just ignored.

The last term is an additional dephasing term that arises due to fluctuations in the diagonal coupling between the subsystem and bath. This is also called an additional decoherence term.

5.6 The Lindblad Equation

The secular Redfield equation can be written as

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\hat{\sigma}_s(t) = &-\frac{i}{\hbar}[\hat{H}_s + \hat{H}_{LS}, \hat{\sigma}_s(t)] \\
&+ \sum_{a,b\neq a} k_{b\to a}\left(\hat{L}_{ab}\hat{\sigma}_s\hat{L}^\dagger_{ab} - \frac{1}{2}\{\hat{L}^\dagger_{ab}\hat{L}_{ab}, \hat{\sigma}_s(t)\}\right) \\
&+ \gamma_d\left(\hat{L}_d\hat{\sigma}_s\hat{L}^\dagger_d - \frac{1}{2}\{\hat{L}^\dagger_d\hat{L}_d, \hat{\sigma}_s(t)\}\right)
\end{aligned}$$

with $\{\hat{A}, \hat{B}\} = \hat{A}\hat{B} + \hat{B}\hat{A}$ as the anticommutator and

$$\begin{aligned}
\hat{H}_{LS} &= \frac{1}{\hbar}\sum_{ac} |S_{ac}|^2 \mathrm{Im}[\tilde{J}(\omega_{ca})] \, |a\rangle \langle a| \\
k_{b\to a} &= \mathcal{R}_{aabb} = \frac{2}{\hbar^2}|S_{ab}|^2 \mathrm{Re}[\tilde{J}(\omega_{ab})] \\
\hat{L}_{ab} &= |a\rangle \langle b| \\
\gamma_a &= \frac{2}{\hbar^2}\mathrm{Re}[\tilde{J}(0)] \\
\hat{L}_a &= \sum_a S_{aa} \, |a\rangle \langle a|
\end{aligned}$$

This is exactly of "Lindblad" form

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{\rho}(t) = -\frac{i}{\hbar}[\hat{H}, \hat{\rho}(t)] + \sum_n \gamma_n \left(\hat{L}_n \hat{\rho} \hat{L}_n^\dagger - \frac{1}{2}\{\hat{L}_n^\dagger \hat{L}_n, \hat{\rho}\} \right)$$

This is the most general form of equation for a density operator that preserves trace, Hermiticity and positivity.

Redfield theory in its non-secular form is not of this form, so it can break the positivity of the density operator (which is unphysical). There are two perspectives on this: 1) we should always have positivity preserving dynamics, so the secular approximation is to be preferred, or 2) violation of positivity gives a clear metric for how accurate we expect the other approximations we have made (perturbation theory and Markovian approximations), so secular Redfield is to be preferred.

6 Introduction to Many-Electron Systems

6.1 One-Electron Systems

Consider a single electron. Fundamental physics tells us this has spin. So the full Hilbert space for a single electron is $\mathcal{H} = \mathcal{H}_{3D} \otimes \mathcal{H}_{spin}$ and a suitable basis is:

$$|\boldsymbol{r}, \sigma\rangle = |\boldsymbol{r}\rangle \otimes |\sigma\rangle$$

where the spin states are $|\alpha\rangle = |s = 1/2, m_s = +1/2\rangle$ and $|\beta\rangle = |s = 1/2, m_s = -1/2\rangle$. The electron spin angular momentum is always fixed at $s = 1/2$ unlike orbital angular momentum.

Single electron states are can be expressed as a "spinor" – a vector of two position wave-functions:

$$|\Psi\rangle = \sum_{\sigma=\alpha,\beta} \psi_\sigma(\boldsymbol{r}) \, |\boldsymbol{r}, \sigma\rangle \implies |\Psi\rangle = \begin{pmatrix} \psi_\alpha(\boldsymbol{r}) \\ \psi_\beta(\boldsymbol{r}) \end{pmatrix}$$

Typical single-electron systems in the absence of magnetic fields or spin-orbit coupling have Hamiltonians that commute with spin operators (e.g., $\hat{S}_a, a = x, y, z$):

$$[\hat{H}, \hat{S}_a] = 0$$

This means eigenstates of $\hat{H}$ can be chosen to be eigenstates of $\hat{S}_z$:

$$|E\rangle \equiv |E, \sigma\rangle$$

6.2 Two-Electron Systems

Two electron systems live in $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2 = \mathcal{H}_{3D,1} \otimes \mathcal{H}_{spin,1} \otimes \mathcal{H}_{3D,2} \otimes \mathcal{H}_{spin,2}$. A suitable basis is $|\boldsymbol{r}_1\sigma_1, \boldsymbol{r}_2\sigma_2\rangle$. The wave-functions are in general:

$$|\Psi\rangle = \sum_{\sigma_1\sigma_2} \int d\boldsymbol{r}_1 \int d\boldsymbol{r}_2 \, \Psi_{\sigma_1\sigma_2}(\boldsymbol{r}_1, \boldsymbol{r}_2) \, |\boldsymbol{r}_1\sigma_1, \boldsymbol{r}_2\sigma_2\rangle$$

The typical Hamiltonian for simple atomic systems like He, H_2, Li^+, etc., is (in atomic units):

$$\hat{H} = \hat{T}_1 + \hat{T}_2 + V(\hat{\boldsymbol{r}}_1) + V(\hat{\boldsymbol{r}}_2) + \frac{1}{|\hat{\boldsymbol{r}}_1 - \hat{\boldsymbol{r}}_2|}$$

The electrons are identical so they feel the same external (nuclear) potential $V(\boldsymbol{r})$. The Hamiltonian commutes with total spin $\hat{S}_a = \hat{S}_{1a} + \hat{S}_{2a}$.

$$[\hat{H}, \hat{S}_a] = 0 \implies [\hat{S}^2, \hat{H}] = 0$$

So eigenstates can be chosen to be eigenstates of $\hat{S}^2$ and $\hat{S}_z$ as well.

For two coupled spins ($s_1 = 1/2, s_2 = 1/2$), the allowed total S is $S = 0$ or 1.

- $S = 0$: Singlets
- $S = 1$: Triplets

The eigenstates $|E, S, M_s\rangle$ can be separated into Space and Spin parts:

$$|\Psi_{E,S,M_s}\rangle = |\Psi_{space}\rangle \otimes |S, M_s\rangle$$

The spin functions are (see homework on coupled angular momentum from 215A):

$$\text{Singlet } |S\rangle = \frac{1}{\sqrt{2}}(|\alpha_1\beta_2\rangle - |\beta_1\alpha_2\rangle) \quad (S=0, M_s=0)$$
$$\text{Triplet } |T_+\rangle = |\alpha_1\alpha_2\rangle \quad (S=1, M_s=1)$$
$$\text{Triplet } |T_0\rangle = \frac{1}{\sqrt{2}}(|\alpha_1\beta_2\rangle + |\beta_1\alpha_2\rangle) \quad (S=1, M_s=0)$$
$$\text{Triplet } |T_-\rangle = |\beta_1\beta_2\rangle \quad (S=1, M_s=-1)$$

6.3 Symmetry of 2-Electron Systems

Let's define the hermitian (you'll prove this in the homework) exchange (permutation) operator $\hat{P}_{12}$ by:

$$\hat{P}_{12}\,|\boldsymbol{r}_1\sigma_1, \boldsymbol{r}_2\sigma_2\rangle = |\boldsymbol{r}_2\sigma_2, \boldsymbol{r}_1\sigma_1\rangle$$

Physically realistic Hamiltonians for electrons commute with $\hat{P}_{12}$, i.e., $[\hat{H}, \hat{P}_{12}] = 0$. Because $\hat{P}_{12}^2 = \hat{1}$, the eigenvalues are $p = \pm 1$.

Quantum field theory tells us that for electrons, the only physically allowed states are antisymmetric ($p = -1$):

$$\hat{P}_{12}\,|\Psi\rangle = -\,|\Psi\rangle$$

Note that $\hat{P}_{12} = \hat{P}_{12}^{space}\hat{P}_{12}^{spin}$.

- For Singlet: $\hat{P}_{12}^{spin}\,|S\rangle = -\,|S\rangle \implies$ Space must be Symmetric (+).
- For Triplet: $\hat{P}_{12}^{spin}\,|T_m\rangle = +\,|T_m\rangle \implies$ Space must be Antisymmetric (−).

So:

$$|\Psi, S\rangle = \left|\Psi_{space}^{(+)}\right\rangle \otimes |S\rangle$$
$$|\Psi, T_m\rangle = \left|\Psi_{space}^{(-)}\right\rangle \otimes |T_m\rangle$$

where $\left|\Psi_{space}^{(\pm)}\right\rangle$ is a spatial wave function with $\pm$ permutation symmetry.

Note that this means electrons cannot be in the same state at the same time. So if two electrons are in the same spin-state, they cannot be at the same point in space, but if they are in opposite spin states, then they can be at the same point in space.

6.4 A Simple First System: Helium (He)

He has nuclear charge $Z = 2$ and two electrons. The Hamiltonian is:

$$\hat{H} = \hat{T}_1 + \hat{T}_2 - \frac{2}{r_1} - \frac{2}{r_2} + \frac{1}{|\boldsymbol{r}_1 - \boldsymbol{r}_2|}$$

Neglecting repulsion ($\hat{H} \approx \hat{H}_0$), the ground state is $1s^2$. The spatial wave-function is:

$$\left|1s^2\right\rangle = |\Psi_{space}\rangle = \int d\boldsymbol{r}_1 d\boldsymbol{r}_2\, \psi_{1s}(\boldsymbol{r}_1)\psi_{1s}(\boldsymbol{r}_2)\,|\boldsymbol{r}_1, \boldsymbol{r}_2\rangle$$

This is symmetric, so it must pair with the singlet spin function $|S\rangle$. This has a ground state energy (0th order): $E_0 = 2E_{1s}(Z=2) = 2(-2) = -4$ Hartrees.

The first order repulsion energy is:

$$E_0^{(1)} = \langle V_{12}\rangle = \frac{5}{8}Z = \frac{5}{4}$$

Total energy $\approx -4 + 1.25 = -2.75$.

Note of effective nuclear charge

In order to improve the estimate of the ground state energy we can also treat this with orbitals with an effective nuclear charge which is treated as a variational parameter using $Z_{eff} < Z$. The orbitals expand to reduce the electron-electron repulsion. You do exactly this in the homework problems.

6.5 Excited States of Helium

The lowest energy excited states from $\hat{H}_0$ will have one electron in the $n = 2$ shell, so the configuration will either be 1s2s or 1s2p.

Consider the $1s2s$ configuration. The overall zero-order wavefunctions for the singlet (S) and triplet (T_m) states are:

$$|1s2s, S\rangle = \frac{1}{\sqrt{2}} \left(|1s_1 2s_2\rangle + |2s_1 1s_2\rangle\right) \otimes |S\rangle$$
$$|1s2s, T_m\rangle = \frac{1}{\sqrt{2}} \left(|1s_1 2s_2\rangle - |2s_1 1s_2\rangle\right) \otimes |T_m\rangle$$

6.6 First-Order Repulsion Energy

Evaluating the repulsion energy to first order for the singlet state:

$$\begin{aligned}
E^{(1)}_{1s2s,S} &= \langle 1s2s, S|\hat{V}_{12}|1s2s, S\rangle \\
&= \frac{1}{2} \left(\langle 1s_1 2s_2| + \langle 2s_1 1s_2|\right) \hat{V}_{12} \left(|1s_1 2s_2\rangle + |2s_1 1s_2\rangle\right) \\
&= \frac{1}{2} \langle 1s_1 2s_2|\hat{V}_{12}|1s_1 2s_2\rangle + \frac{1}{2} \langle 1s_1 2s_2|\hat{V}_{12}|2s_1 1s_2\rangle + \frac{1}{2} \langle 2s_1 1s_2|\hat{V}_{12}|1s_1 2s_2\rangle + \frac{1}{2} \langle 2s_1 1s_2|\hat{V}_{12}|2s_1 1s_2\rangle \\
&= J + K
\end{aligned}$$

The simplification to J and K uses symmetry of the integrals.

- **Coulomb Term (J):** Represents classical Coulomb repulsion between the two electron densities:

$$J = \int d\boldsymbol{r}_1 d\boldsymbol{r}_2 |\phi_{1s}(\boldsymbol{r}_1)|^2 \frac{1}{r_{12}} |\phi_{2s}(\boldsymbol{r}_2)|^2$$

- **Exchange Term (K):** Arises from exchange symmetry with no classical analogue:

$$K = \int d\boldsymbol{r}_1 d\boldsymbol{r}_2 \phi^*_{1s}(\boldsymbol{r}_1)\phi^*_{2s}(\boldsymbol{r}_2) \frac{1}{r_{12}} \phi_{2s}(\boldsymbol{r}_1)\phi_{1s}(\boldsymbol{r}_2)$$

For the triplet states, the calculation yields $E^{(1)}_{1s2s,T_m} = J - K$. Thus, the triplet is stabilized relative to the singlet by $2K$.

6.6.1 Probability Densities and the Fermi Hole

The two-electron density $\rho_2(\boldsymbol{r}_1, \boldsymbol{r}_2) = \Psi_{space}(\boldsymbol{r}_1, \boldsymbol{r}_2)^* \Psi_{space}(\boldsymbol{r}_1, \boldsymbol{r}_2)$ explains this stabilization of the triplet:

$$\rho^S_2(\boldsymbol{r}_1, \boldsymbol{r}_2) = \rho_{1s}(\boldsymbol{r}_1)\rho_{2s}(\boldsymbol{r}_2) + \phi_{1s}(\boldsymbol{r}_1)\phi_{2s}(\boldsymbol{r}_2)\phi_{1s}(\boldsymbol{r}_2)\phi_{2s}(\boldsymbol{r}_1)$$
$$\rho^T_2(\boldsymbol{r}_1, \boldsymbol{r}_2) = \rho_{1s}(\boldsymbol{r}_1)\rho_{2s}(\boldsymbol{r}_2) - \phi_{1s}(\boldsymbol{r}_1)\phi_{2s}(\boldsymbol{r}_2)\phi_{1s}(\boldsymbol{r}_2)\phi_{2s}(\boldsymbol{r}_1)$$

As $|\boldsymbol{r}_1 - \boldsymbol{r}_2| \to 0$, $\rho^T_2 \to 0$ (the Fermi hole), while ρ^S_2 remains finite. Like-spin electrons in the triplet state avoid each other, reducing their mutual repulsion.

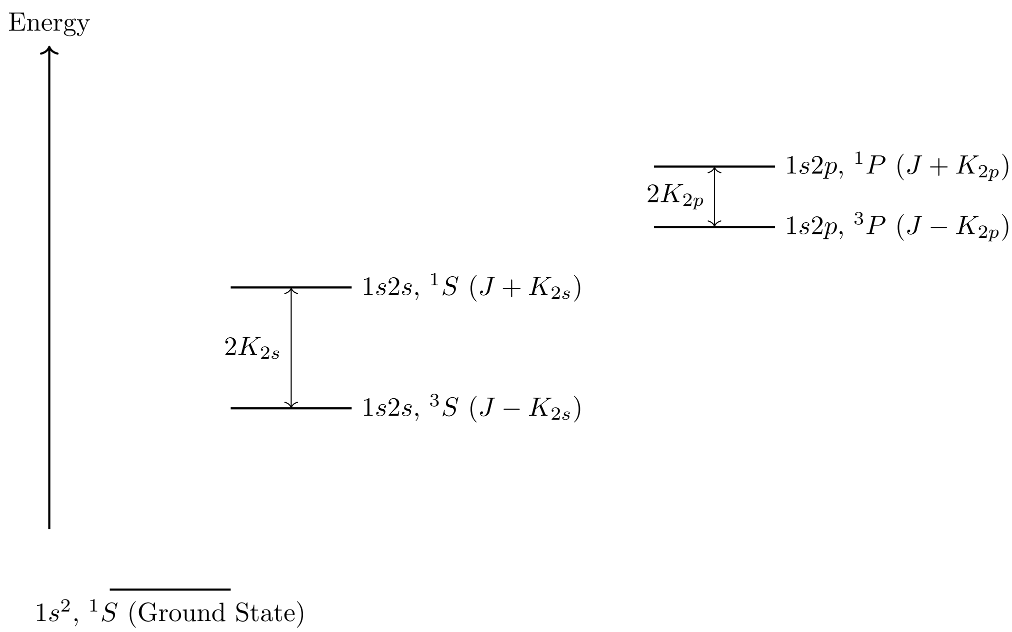
Comparison of $1s2s$ and $1s2p$ Configurations

For hydrogenic orbitals with charge Z, the integrals for the 1s2s and 1s2p are:

Configuration	J	K
$1s2s$	$\frac{17}{81}Z$	$\frac{16}{729}Z$
$1s2p$	$\frac{59}{243}Z$	$\frac{32}{6561}Z$

The $1s2s$ configuration has lower repulsion (J) because the $2s$ orbital is more penetrating and electrons are further apart from each other on average while feeling a stronger nuclear attraction.

Energy Level Diagram



6.7 Dynamic Electron Correlation

Let's consider the 2-electron Schrödinger equation where $r_{12} \to 0$. Locally the e-e repulsion dominates.

$$E\Psi \approx \hat{T}_1\Psi + \hat{T}_2\Psi + \frac{1}{r_{12}}\Psi$$

In terms of relative coordinates $\boldsymbol{r} = \boldsymbol{r}_1 - \boldsymbol{r}_2$ and centre of mass $\boldsymbol{R}$:

$$\hat{T}_1 + \hat{T}_2 = -\frac{1}{4}\nabla_R^2 - \nabla_r^2$$

Separating variables $\Psi = \Psi(\boldsymbol{R})\Psi_r(\boldsymbol{r})$. The relative aprt can be split into different relative angular momentum terms (just like the hydrogenic atom Schrödinger equation) $\Psi_r(\boldsymbol{r}) = \sum_{L,M} \Phi_L(r) Y_{L,M}(\theta, \varphi)$ and following a very similar separation of variables we find

$$\left(-\frac{d^2}{dr^2} - \frac{2}{r}\frac{d}{dr} + \frac{L(L+1)}{r^2} + \frac{1}{r}\right)\Phi_L(r) = E\Phi_L(r)$$

We found before that $\Phi_L(r) = r^L u_L(r)$, so plugging this in above gives

$$\begin{aligned} Er^L u_L &= (-L(L-1)r^{L-2}u_L - 2Lr^{L-1}u_L' - r^L u_L'') + (-2Lr^{L-2}u_L - 2r^{L-1}u_L') + L(L+1)r^{L-2}u_L + r^{L-1}u_L \\ &= -2(L+1)r^{L-1}u_L' - r^L u_L'' + r^{L-1}u_L \end{aligned}$$

So dividing by $r^{L-1}u_L$ and considering small r we find

$$\lim_{r\to 0} \frac{1}{u_L}\frac{\mathrm{d}u_L}{\mathrm{d}r} = \frac{1}{2(L+1)}$$

This is the **cusp condition**. It implies that for like-spin electrons (which must be antisymmetric spatial so $L = 1, 3, ...$), the wavefunction goes as r^L (linear) or higher powers, reducing probability density at coalescence. On the other hand for opposite spin electrons $L = 0$ components can contribute at the wavefunction can be non-zero at small r_{12} because the spatial function has to have even L contributions. (Recall that the parity i.e. exchange symmetry of the angular part is $(-1)^L$, c.f. s orbitals are symmetric but p orbitals are anti-symmetric.)

Zero-order mean-field wavefunctions (like our symmetrised product ansatz here) do not satisfy this cusp condition exactly, because they are only approximations to the true wave function. This missing energy is called **dynamic correlation**. It can be added back using perturbation theory:

$$E_n^{(2)} = \sum_{m\neq n} \frac{|\langle m^{(0)}|\hat{V}_{12}|n^{(0)}\rangle|^2}{E_n - E_m}$$

Wave functions of multi-electron systems built from (symmetrised) products of orbitals never include dynamic correlation. They only include the simple direct Coulomb and Exchange energies.

6.8 Molecules

For molecules we have an electron Hamiltonian

$$\hat{H} = \hat{T} + \frac{\hbar^2}{m}\sum_{i,A}\frac{Z_A}{|\boldsymbol{r}_i - \boldsymbol{R}_A|} + \sum_{i\neq j}\frac{1}{|\boldsymbol{r}_i - \boldsymbol{r}_j|} + \sum_{A,B>A}\frac{Z_A Z_B}{|\boldsymbol{R}_A - \boldsymbol{R}_B|}$$

To gain a qualitative understanding, let's first ignore e-e repulsion ($\hat{V}_{12}$) and set $\hat{H}_0 = \hat{T} - \hat{V}_{12}$. This is separable, so solutions are appropriately symmetrized products of 1e wavefunctions i.e. molecular orbitals.

We also know $\hat{H} \to \hat{H}_A$ for $\boldsymbol{r}_i \to \boldsymbol{R}_A$, so locally solutions should resemble atomic orbitals. This motivates using the linear-combination of atomic orbitals ansatz for the spatial 1e functions.

$$|\psi\rangle = \sum_\nu c_\nu |\chi_\nu\rangle$$

Applying the variational theorem we find

$$\sum_\mu \epsilon_j S_{\mu\nu} c_\nu = \sum_\mu H_{\mu\nu} c_\nu$$

where $H_{\mu\nu}$ is the 1e Hamiltonian matrix element.

Let's see how this works for H_2 ...

For H_2 let's use a linear combination of 1s orbitals on atoms A and B

$H_{AA} = \alpha = H_{BB}$ and $H_{AB} = H_{BA} = \beta$. Let's approximate $S_{AB} = 0$ and we know $S_{AA} = S_{BB} = 1$. This gives two molecular orbitals

$$|\sigma\rangle = \frac{1}{\sqrt{2}}\left(|1s_A\rangle + |1s_B\rangle\right), \quad \epsilon_\sigma = \alpha + \beta$$
$$|\sigma^*\rangle = \frac{1}{\sqrt{2}}\left(|1s_A\rangle - |1s_B\rangle\right), \quad \epsilon_{\sigma^*} = \alpha - \beta$$

Note $H_{AA} \approx E_{1s} < 0$ and also $\beta < 0$, so the σ orbital is lowest in energy. Now the allowed many electron wavefunctions are

$$\begin{aligned}
\left|\sigma^2, S\right\rangle &= \left|\sigma_1\sigma_2\right\rangle \otimes \left|S\right\rangle \\
\left|\sigma^{*2}, S\right\rangle &= \left|\sigma^*\sigma_2^*\right\rangle \otimes \left|S\right\rangle \\
\left|\sigma\sigma^*, S\right\rangle &= \frac{1}{\sqrt{2}}\left(\left|\sigma_1\sigma_2^*\right\rangle + \left|\sigma_1^*\sigma_2\right\rangle\right) \otimes \left|S\right\rangle \\
\left|\sigma\sigma^*, T_m\right\rangle &= \frac{1}{\sqrt{2}}\left(\left|\sigma_1\sigma_2^*\right\rangle - \left|\sigma_1^*\sigma_2\right\rangle\right) \otimes \left|T_m\right\rangle
\end{aligned}$$

The lowest energy state is $\left|\sigma^2, S\right\rangle$. Let's evaluate its first-order energy including repulsion. We'll use the shorthand $\left|A_1B_2\right\rangle = \left|1s_{A,1}1s_{B,2}\right\rangle$ etc.

$$\begin{aligned}
E^{(0)}_{\sigma^2 S} &= 2(\alpha+\beta) + \frac{1}{R_{AB}} \\
E^{(1)}_{\sigma^2 S} &= \langle\sigma_1\sigma_2|\hat{V}_{12}|\sigma_1\sigma_2\rangle \\
&= \int d\boldsymbol{r}_1 \int d\boldsymbol{r}_2 \frac{\rho_\sigma(\boldsymbol{r}_1)\rho_\sigma(\boldsymbol{r}_2)}{r_{12}}
\end{aligned}$$

$\rho_\sigma(\boldsymbol{r})$ is given by

$$\rho_\sigma(\boldsymbol{r}) = \frac{1}{2}\left(\rho_{AA}(\boldsymbol{r}) + \rho_B(\boldsymbol{r}) + 2\rho_{AB}(\boldsymbol{r})\right)$$

Let's approximate with $\rho_{AB}(\boldsymbol{r}) = \chi_{1s_A}(\boldsymbol{r})\chi_{1s_B}(\boldsymbol{r})$ and further let's approximate

$$\begin{aligned}
\int d\boldsymbol{r}_1 \int d\boldsymbol{r}_2 \frac{\rho_{AB}(\boldsymbol{r}_1)\rho_{AB}(\boldsymbol{r}_2)}{r_{12}} &\approx 0 \\
\int d\boldsymbol{r}_1 \int d\boldsymbol{r}_2 \frac{\rho_{AA}(\boldsymbol{r}_1)\rho_{BB}(\boldsymbol{r}_2)}{r_{12}} &\approx \frac{1}{R_{AB}} \\
\int d\boldsymbol{r}_1 \int d\boldsymbol{r}_2 \frac{\rho_{AA}(\boldsymbol{r}_1)\rho_{AB}(\boldsymbol{r}_2)}{r_{12}} &\approx 0
\end{aligned}$$

So the electron repulsion energy becomes

$$E^{(1)}_{\sigma^2 S} = \frac{1}{4}\left(J_{AA} + J_{BB} + 2J_{AB}\right) = \frac{1}{2}J + \frac{1}{2}\frac{1}{R_{AB}}$$

Likewise $\alpha \approx E_{1s} \approx \frac{1}{R_{AB}}$ so the total energy is

$$\begin{aligned}
E_{\sigma^2 S} &= 2E_{1s} - \frac{2}{R_{AB}} + \frac{1}{R_{AB}} + \frac{1}{2}\frac{1}{R_{AB}} + \frac{1}{2}J + 2\beta \\
&= 2E_{1s} - \frac{1}{2}\frac{1}{R_{AB}} + \frac{1}{2}J + 2\beta
\end{aligned}$$

This is valid for medium to large R_{AB}, as long as the overlap is small.

Is this physically correct? Let's consider $R_{AB} \to \infty$

$$E_{\sigma^2 S} = 2E_{1s} + \frac{1}{2}J$$

and the spatial wavefunction becomes

$$\left|\sigma^2\right\rangle_{\text{space}} = \frac{1}{2}\left(\left|A_1A_2\right\rangle + \left|A_1B_2\right\rangle + \left|B_1A_2\right\rangle + \left|B_1B_2\right\rangle\right)$$

so the wavefunction corresponds to the state $\frac{1}{2}\left(\left|H^-H^+\right\rangle + \left|H^+H^-\right\rangle + 2\left|H\ H\right\rangle\right)$ and it has high energy ionic contributions. This is wrong for large R_{AB} because H_2 should dissociate into two neutral H atoms (this is the ground state for infinite R_{AB})! The spurious ionic contribution explains the spurious $-\frac{1}{2}\frac{1}{R_{AB}}$ in the energy.

In order to fix this we need to account for static correlation, which arises from the electron repulsion coupling different molecular orbital configurations.

6.9 Static Correlation in H_2

Let's evaluate what we need for the 2nd order perturbation energy for H_2 in the $\left|\sigma^2, S\right\rangle$ state. First we find

$$\left\langle\sigma\sigma^*, S\middle|\hat{V}_{12}\middle|\sigma^2, S\right\rangle = 0$$
$$\left\langle\sigma\sigma^*, T_M\middle|\hat{V}_{12}\middle|\sigma^2, S\right\rangle = 0$$

and the only non-zero term is

$$\left\langle\sigma^{*2}, S\middle|\hat{V}_{12}\middle|\sigma^2, S\right\rangle = \frac{1}{4}\left(J_{AA} + J_{BB} - J_{AB} - J_{BA}\right) \approx \frac{1}{2}J - \frac{1}{2}\frac{1}{R_{AB}}$$

So the electron-repulsion couples σ^2 and σ^{*2} configurations. The $\left|\sigma^{*2}, S\right\rangle$ energy is (up to 1st order)

$$E_{\sigma^{*2}S} = 2E_{1s} - \frac{1}{2}\frac{1}{R_{AB}} + \frac{1}{2}J - 2\beta$$

This is all we need to do a variational calculation using a multi-configurational ansatz

$$|\Psi\rangle = C_1\left|\sigma^2, S\right\rangle + C_2\left|\sigma^{*2}, S\right\rangle = C_1|1\rangle + C_2|2\rangle$$

The variational (secular) equations are

$$\begin{pmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{pmatrix}\begin{pmatrix} C_1 \\ C_2 \end{pmatrix} = E\begin{pmatrix} C_1 \\ C_2 \end{pmatrix}$$

with

$$H_{11} = 2E_{1s} - \frac{1}{2}\frac{1}{R_{AB}} + \frac{1}{2}J + 2\beta$$
$$H_{22} = 2E_{1s} - \frac{1}{2}\frac{1}{R_{AB}} + \frac{1}{2}J - 2\beta$$
$$H_{12} = H_{21} = \frac{1}{2}J - \frac{1}{2}\frac{1}{R_{AB}}$$

For large separations $\beta \to 0$ and the lowest energy solution is

$$|\Psi_-\rangle = \frac{1}{\sqrt{2}}\left(\left|\sigma^2, S\right\rangle - \left|\sigma^{*2}, S\right\rangle\right)$$

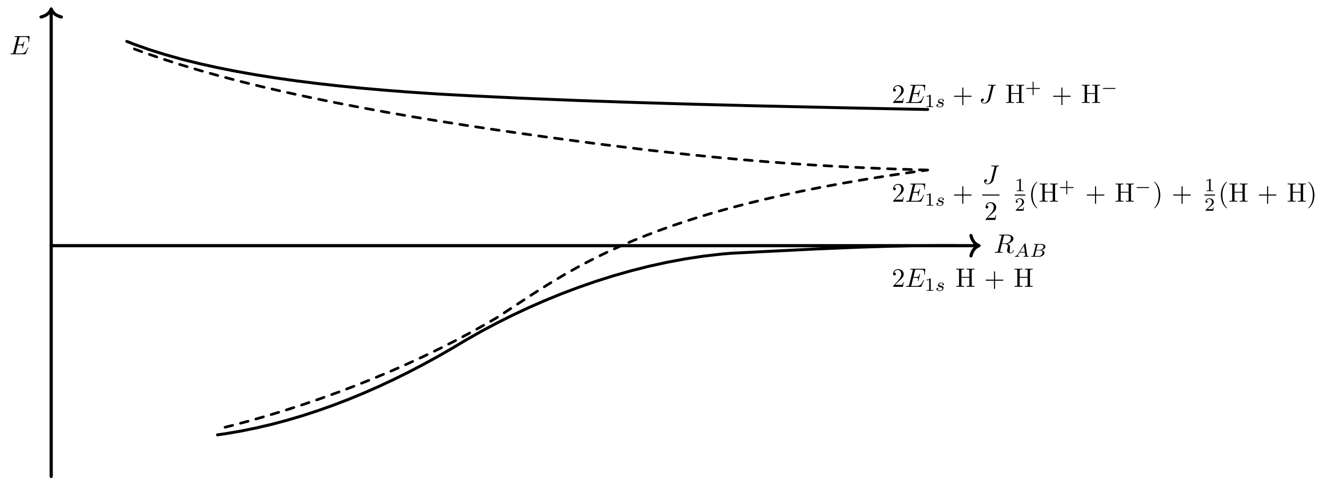
with

$$E_- = 2E_{1s} - \frac{1}{2}\frac{1}{R_{AB}} + \frac{1}{2}J - \left(\frac{1}{2}J - \frac{1}{2}\frac{1}{R_{AB}}\right) = 2E_{1s}$$

and the wavefunction is

$$|\Psi_-\rangle = \frac{1}{\sqrt{2}}\left(|A_1B_2\rangle + |B_1A_2\rangle\right) \otimes |S\rangle$$

which is purely atomic with no ions. Notice that the erroneous $-\frac{1}{2R_{AB}}$ dependence in the energy is removed.



6.10 Many electron wave-functions

We have been able to put together a treatment of two electron systems, but how to many-electron systems work? Let's consider Li using the same analysis as above. Ignoring the two-electron repulsion $\hat{V}_{ee}$, the zero order Hamiltonian has hydrogenic solutions, so we fill these states. Because electrons can only be in one of two spin states, we can only have two electrons in 1s spatial orbitals and respect the anti-symmetry under exchange. The third electron therefore has to occupy the 2s orbital.

We can't easily factorise spatial and spin functions anymore, but we can write down the anti-symmetrised wave function by considering $|1s_1 1s_2 2s_3\rangle\, |S_{12}, \alpha_3\rangle$, which is anti-symmetric for 1 and 2, but not all three. We therefore symmetrised by adding in the states generated by applying $-\hat{P}_{13}$ and $-\hat{P}_{23}$ to this reference partially symmetrised state. This gives

$$|\Psi\rangle = \frac{1}{\sqrt{3}}\Big(|1s_1 1s_2 2s_3\rangle\, |S_{12}, \alpha_3\rangle - |2s_1 1s_2 1s_3\rangle\, |S_{32}, \alpha_1\rangle - |1s_1 2s_2 1s_3\rangle\, |S_{13}, \alpha_2\rangle \Big)$$

If we want to generalise this to 4, 5, 6, or more electrons, this rapidly gets very difficult to deal with. The general approach to generate an anti-symmetrised state from a set of occupied, orthonormal orbitals is to use a **Slater determinant**. First let's introduce the notation for the space-spin basis states $|\mathbf{x}_k\rangle = |\boldsymbol{r}_k \sigma_k\rangle$, so $\phi(\mathbf{x}_k) = \phi_{\sigma_k}(\boldsymbol{r}_k)$. This is a common short-hand for dealing with the electron spin in compact notation (but it can take a bit of time to get your head around it). States can be written as

$$|\phi\rangle = \int \mathrm{d}\mathbf{x}\, \phi(\mathbf{x})\, |\mathbf{x}\rangle = \sum_\sigma \int \mathrm{d}\boldsymbol{r}\, \phi_\sigma(\boldsymbol{r})\, |\boldsymbol{r}\sigma\rangle = \sum_\sigma \int \mathrm{d}\boldsymbol{r}\, \phi(\boldsymbol{r}\sigma)\, |\boldsymbol{r}\sigma\rangle$$

The function $\phi(\mathbf{x})$ is called a **spin-orbital** because it is a function of both space and spin coordinates.

The Slater determinant for N electrons in N spin-orbitals is defined as follows

$$\Phi(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \phi_1(\mathbf{x}_1) & \phi_1(\mathbf{x}_2) & \cdots & \phi_1(\mathbf{x}_N) \\ \phi_2(\mathbf{x}_1) & \phi_2(\mathbf{x}_2) & \cdots & \phi_2(\mathbf{x}_N) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_N(\mathbf{x}_1) & \phi_N(\mathbf{x}_2) & \cdots & \phi_N(\mathbf{x}_N) \end{vmatrix}$$

This is a determinant so by construction this is anti-symmetric with respect to exchange (which corresponds to swapping columns).

Example: He $1s^2$

As a sanity check for a $1s^2$ configuration of He, $\phi_1 = 1s_\alpha$ and $\phi_2 = 1s_\beta$. For these $\phi_1(\boldsymbol{r}\alpha) = \chi_{1s}(\boldsymbol{r})$ and $\phi_1(\boldsymbol{r}\beta) = 0$ and $\phi_2(\boldsymbol{r}\beta) = \chi_{1s}(\boldsymbol{r})$ and $\phi_1(\boldsymbol{r}\alpha) = 0$

$$
\begin{aligned}
\Phi(\mathbf{r}_1\alpha, \mathbf{r}_2\alpha) &= \Phi(\mathbf{r}_1\beta, \mathbf{r}_2\beta) = 0 \\
\Phi(\mathbf{r}_1\alpha, \mathbf{r}_2\beta) &= -\Phi(\mathbf{r}_1\beta, \mathbf{r}_2\alpha) = \frac{1}{\sqrt{2}}\chi_{1s}(\boldsymbol{r}_1)\chi_{1s}(\boldsymbol{r}_2)
\end{aligned}
$$

so this is exactly the $\left|1s^2, S\right\rangle$ state we had before. We denote this Slater determinant $\Phi_{1s_\alpha 1s_\beta}$ (we'll use analogous notation below for other states of He).

We can much more compactly represent the Li wave function above using a Slater determinant with $\phi_1 = 1s_\alpha$, $\phi_2 = 1s_\beta$ and $\phi_3 = 2s_\alpha$.

Question: Can we always represent states with single Slater Determinants?

We can represent the He singlet $|1s2s, T_+\rangle$ states as single Slater determinants with $\phi_1 = 1s_\alpha$, $\phi_2 = 2s_\alpha$, but try as you may you cannot write the $|1s2s, S\rangle$ or $|1s2s, T_0\rangle$ states as single Slater determinants. These are given by

$$
\begin{aligned}
\Psi_{1s2s,S}(\mathbf{x}) &= \frac{1}{\sqrt{2}}\Phi_{1s_\alpha 2s_\beta}(\mathbf{x}) - \frac{1}{\sqrt{2}}\Phi_{1s_\beta 2s_\alpha}(\mathbf{x}) \\
\Psi_{1s2s,T_0}(\mathbf{x}) &= \frac{1}{\sqrt{2}}\Phi_{1s_\alpha 2s_\beta}(\mathbf{x}) + \frac{1}{\sqrt{2}}\Phi_{1s_\beta 2s_\alpha}(\mathbf{x})
\end{aligned}
$$

The single slater determinants $\Phi_{1s_\alpha 2s_\beta}(\mathbf{x})$ are not eigenfunctions of $\hat{S}^2$, and in general Slater determinants are not guaranteed to be eigenfunctions of $\hat{S}^2$. This is why some states cannot be represented with single Slater determinants.

In general electronic wave functions that satisfy anti-symmetry are linear combinations of Slater determinants

$$
\Psi(\mathbf{x}) = \sum_J C_J \Phi_J(\mathbf{x})
$$

where each Slater determinant Φ_J can be associated with a different set of orbitals $\phi_k^{(J)}$.

Configuration state functions (CSFs) are linear combinations of Slater determinants with spins flipped appropriately, and using Clebsch-Gordon angular momentum coupling coefficients that can are by construction eigenstates of $\hat{S}^2$ and $\hat{S}_z$, so they are always "spin pure". What we constructed above for two electron systems are strictly speaking configuration state functions. Slater determinants find wider use in computational chemistry because it's much easier to evaluate matrix elements between Slater determinants (as we'll see when we examine the second quantisation) compared to CSFs, although CSFs find their use in some electronic structure methods and codes (for example they are used in the "complete active space self consistent field" (CASSCF) method).

We'll see more how Slater determinants are useful when we learn about the second quantisation.

7 Second Quantisation

In the previous course we saw the consequences of symmetry in quantum mechanics, and how it plays a role in several aspects of quantum mechanics. When considering systems consisting of multiple identical and physically indistinguishable particles, there is one more important symmetry: particle exchange. If we swap the labels between two indistinguishable particles, any physical observable, must be unchanged. This implies the two particle permutation operator, $\hat{P}_{ij}$ is a symmetry of a system of indistingushable particles. In this chapter we will explore the consequences of this and set-up the so-called “second quantisation” formalism for describing systems consisting of many indistinguishable particles, for example the electrons in atoms and molecules, or photons interacting with matter.

7.1 Setup and symmetrised states

For systems with identical and indistinguishable particles, the Hamiltonian governing them, and any physically relevant observables, must be invariant to any permutation of two particle indices

$$\hat{H} = \hat{P}_{ij}\hat{H}\hat{P}_{ij}$$

and therefore the Hamiltonian is invariant to any permutation. Because $\hat{P}_{ij}^2 = \hat{1}$, and because it commutes with the Hamiltonian, we know the states of any physical system of indistinguishable particles must obey

$$\hat{P}_{ij} |\Psi\rangle = \pm |\Psi\rangle$$

and because any permutation can be decomposed into a product of other permutations, they must all have the same sign. The symmetry is determined by the type of particle + for *bosons* and − for *fermions.* This turns out to be determined by the intrinsic spin of the particle, but showing this connection goes beyond this course. We will simply accept this for now. One immediate consequence for this is that an acceptable basis for N bosons is

$$|\mathbf{k}\rangle = |k_1, k_2, \ldots, k_N\rangle = \frac{1}{\sqrt{N! \prod_j n_j(\mathbf{k})!}} \sum_P \hat{P} |k_1(1), k_2(2), \ldots, k_3(N)\rangle$$

where $n_j(\boldsymbol{k})$ is the number of times basis function j appears in the list $\mathbf{k}$. For fermions the antisymmetrised basis functions are

$$|\mathbf{k}\rangle = |k_1, k_2, \ldots, k_N\rangle = \frac{1}{\sqrt{N!}} \sum_P \sigma_P \hat{P} |k_1(1), k_2(2), \ldots, k_3(N)\rangle$$

where σ_P is -1 if the permutation P consists of an odd number of two-particle permutations, and $+1$ otherwise. In general we can write these in terms of the symmetrisation and anti-symmetrisation operators as

$$\hat{S}_\pm = \sum_P \hat{P}\sigma_P^\pm$$

where $\sigma_P^+ = 1$ and $\sigma_P^- = \sigma_P$. In general

$$\hat{P}_{ij}\hat{S}_\pm = \pm\hat{S}_\pm$$

We can now define an enlarged Hilbert space, that includes not only the N particle states, states with any number of particles N from 0 upwards. This is called the Fock space, and its basis is defined by the set of all possible $|\mathbf{k}\rangle$ states for all N. We note that this set of states is over-complete because $|\mathbf{k}\rangle = \sigma_P^\pm |P\mathbf{k}\rangle$, but this set at least contains the basis.

The one state we should carefully define is the zero particle state or “vacuum” state

$$|\rangle \equiv |\text{vac}\rangle\,.$$

This is a state that contains no particles, but it is *not* the zero-vector. Sometimes the notation $|\rangle$, $|\Omega\rangle$, $|0\rangle$ or $|\mathbf{0}\rangle$ is also used for this state. We define this state to be normalised to 1, so $\langle\text{vac}|\text{vac}\rangle = 1$.

7.2 Creation/annihilation operators

We can now define the creation operator on the Fock space by its action on an arbitrary

$$\hat{a}_k^\dagger |k_1, k_2, \ldots, k_N\rangle = \sqrt{n_k(\mathbf{k}) + 1}\, |k_1, k_2, \ldots, k_N, k\rangle$$

and we note that if $k = k_i$ for any i when this is a fermionic state that this is zero. The creation operator just appends k to the end of the list of states occupied in the many-body basis state. We also note that we can construct the one-particle states as $|k\rangle = \hat{a}_k^\dagger |\text{vac}\rangle$.

We will now distinguish between fermions and bosons by using $\hat{c}_k^\dagger$ for fermions and $\hat{b}_k^\dagger$ for bosons.

For bosons we have

$$\hat{b}_k^\dagger \hat{b}_j^\dagger |\mathbf{k}\rangle = \hat{b}_j^\dagger \hat{b}_k^\dagger |\mathbf{k}\rangle$$

for any j and k. For fermions we have (due to the anti-symmetry)

$$\hat{c}_k^\dagger \hat{c}_j^\dagger |\mathbf{k}\rangle = -\hat{c}_j^\dagger \hat{c}_k^\dagger |\mathbf{k}\rangle$$

and therefore specifically

$$\hat{c}_k^\dagger \hat{c}_k^\dagger |\mathbf{k}\rangle = 0$$

For fermions this follows because $|\mathbf{k}, j, k\rangle = -|\mathbf{k}, k, j\rangle = \hat{P}_{N+1,N+2} |\mathbf{k}, k, j\rangle$ because the "parent" non-symmetrised state for the two anti-symemtrised states are related by $\hat{P}_{N+1,N+2}$, and so therefore are the symmetrised states.

So from this we deduce that

$$[\hat{b}_j^\dagger, \hat{b}_k^\dagger] = 0 \quad \{\hat{c}_j^\dagger, \hat{c}_k^\dagger\} = 0$$

where $\{\hat{A}, \hat{B}\} = \hat{A}\hat{B} + \hat{B}\hat{A}$ is the anti-commutator.

What about the adjoints of these operators? Remember that they act on the Fock space that consists of all correctly symmetrised states with all possible number of particles. We know that

$$\langle \mathbf{k}' | \mathbf{k} \rangle = 0$$

if the states have different numbers of particles or if the occupied basis states are different, and $\sigma_P^{\pm}$ if $\mathbf{k}$ and $\mathbf{k}'$ are related by a permutation P. This means the only non-zero matrix elements of $\hat{a}_k^\dagger$ are

$$\langle \mathbf{k}, k | \hat{a}_k^\dagger | P\mathbf{k} \rangle$$

where P is an arbitrary permutation of the sequence $\mathbf{k}$ considering the conjugate of this we find that

$$\langle k, \mathbf{k} | \hat{a}_k^\dagger | P\mathbf{k} \rangle^* = \langle P\mathbf{k} | \hat{a}_k | k, \mathbf{k} \rangle$$

so the non-zero matrix elements of $\hat{a}_k$ are between states with $N-1$ and N particles. As such we say $\hat{a}_k$ is an annihilation operator, because it removes a particle from the state k to produce a correctly symmetrised state with one fewer particles occupying this basis state. Overall we deduce that

$$\hat{a}_k |\mathbf{k}\rangle = \sigma_P^{\pm} \sqrt{n_k(\mathbf{k})}\, |\mathbf{k}_k^-\rangle$$

where $\mathbf{k}_k^-$ denotes the sequence $\mathbf{k}$ with k removed and P is the permutation such that $P\mathbf{k} = (\mathbf{k}_k^-, k)$.

From this we can conclude that the adjoint is well-defined, and it has the same (anti-)commutation relations as $\hat{a}_k^\dagger$

$$[\hat{b}_j, \hat{b}_k] = 0 \quad \{\hat{c}_j, \hat{c}_k\} = 0.$$

In order to fully define the algebra of the creation and annihilation operators in Fock space, we need to deduce how $\hat{a}_k^\dagger$ and $\hat{a}_j$ act together. Firstly for bosons for

$$\hat{b}_j \hat{b}_k^\dagger |\mathbf{k}\rangle = \sqrt{n_k(\mathbf{k}) + 1}\hat{b}_j |\mathbf{k}, k\rangle$$

This is zero if j is not in $\mathbf{k}$ and because $k \neq j$ by assumption, and denoting $\left|\mathbf{k}_j^-, k\right\rangle$ with $\mathbf{k}_j^-$ as the state with j removed once from the sequence we have

$$\hat{b}_j \hat{b}_k^\dagger \left|\mathbf{k}\right\rangle = \sqrt{n_k(\mathbf{k})+1}\sqrt{n_j(\mathbf{k})}\left|\mathbf{k}_j^-, k\right\rangle$$

Repeating this for $\hat{b}_k^\dagger \hat{b}_j$ we get the same result so $[\hat{b}_j, \hat{b}_k^\dagger] = 0$ for $j \neq k$. Now considering the case where $j = k$, we know $\hat{b}_k^\dagger$ adds a particle in state k, so

$$\hat{b}_k \hat{b}_k^\dagger \left|\mathbf{k}\right\rangle = (\sqrt{n_k(\mathbf{k})+1})^2 \left|\mathbf{k}\right\rangle = (n_k(\mathbf{k})+1)\left|\mathbf{k}\right\rangle$$

now because $\hat{b}_k$ removes a particle from k, for $\hat{b}_k^\dagger \hat{b}_k$ we have

$$\hat{b}_k^\dagger \hat{b}_k \left|\mathbf{k}\right\rangle = \sqrt{n_k(\mathbf{k})}\hat{b}_k^\dagger \left|\mathbf{k}_k^-\right\rangle = n_k(\mathbf{k})\left|\mathbf{k}\right\rangle$$

so overall we have

$$[\hat{b}_j, \hat{b}_k^\dagger] = \delta_{j,k}$$

For fermions, we have

$$\hat{c}_j \hat{c}_k^\dagger \left|\mathbf{k}\right\rangle = -\sigma_P^- \left|\mathbf{k}_j^-, k\right\rangle$$

where P is the permutation such that $P(\mathbf{k}_j^-, j) = (\mathbf{k})$ if k is not in $\mathbf{k}$ and j is in $\mathbf{k}$, and it is zero otherwise. Likewise for $\hat{c}_k^\dagger \hat{c}_j$ this is zero unless k is not in $\mathbf{k}$ and j is in $\mathbf{k}$ in which case it reduces to

$$\hat{c}_k^\dagger \hat{c}_j \left|\mathbf{k}\right\rangle = \sigma_P^- \hat{c}_k^\dagger \left|\mathbf{k}_j^-\right\rangle = \sigma_P^- \left|\mathbf{k}_j^-, k\right\rangle$$

so for $j \neq k$ we have $\{\hat{c}_j, \hat{c}_k^\dagger\} = 0$. For $j = k$ the $\hat{c}_k \hat{c}_k^\dagger \left|\mathbf{k}\right\rangle$ is zero unless k does not appear in $\mathbf{k}$, in which case it simply returns $\left|\mathbf{k}\right\rangle$, so overall

$$\hat{c}_k \hat{c}_k^\dagger \left|\mathbf{k}\right\rangle = (1 - n_k(\mathbf{k}))\left|\mathbf{k}\right\rangle$$

Likewise $\hat{c}_k^\dagger \hat{c}_k \left|\mathbf{k}\right\rangle$ is zero unless k does appear in $\mathbf{k}$, and it is zero otherwise, so

$$\hat{c}_k^\dagger \hat{c}_k \left|\mathbf{k}\right\rangle = n_k(\mathbf{k})\left|\mathbf{k}\right\rangle$$

so overall we have

$$\{\hat{c}_j, \hat{c}_k^\dagger\} = \delta_{j,k}.$$

This fully defines the algebra of the creation and annihilation operators. We now consider how to construct second quantized representations of symemtric one and two-body operators.

Summary

To briefly summarise, the creation and annihilation operators act on (anti-)symmetrised states as

$$\hat{a}_k^\dagger \left|\mathbf{k}\right\rangle = \sqrt{n_k(\mathbf{k})+1}\left|\mathbf{k}, k\right\rangle \quad \hat{a}_k \left|\mathbf{k}\right\rangle = \sigma_P^\pm \sqrt{n_k(\mathbf{k})}\left|\mathbf{k}_k^-\right\rangle$$

We note that for Fermions the state $\left|\mathbf{k}, k\right\rangle = 0$ if $k \in \mathbf{k}$ and $\left|\mathbf{k}_k^-\right\rangle = 0$ if $k \notin \mathbf{k}$.

The creation/annihilation operators also obey the following (anti-)commutation relations for bosons ($\hat{b}_k$) and fermions ($\hat{c}_k$)

$$[\hat{b}_j^\dagger, \hat{b}_k^\dagger] = 0 \quad \{\hat{c}_j^\dagger, \hat{c}_k^\dagger\} = 0$$

$$[\hat{b}_j, \hat{b}_k] = 0 \quad \{\hat{c}_j, \hat{c}_k\} = 0$$

$$[\hat{b}_j, \hat{b}_k^\dagger] = \delta_{j,k} \quad \{\hat{c}_j, \hat{c}_k^\dagger\} = \delta_{j,k}$$

Note that this means we can write down the many-body basis states as

$$\left|\mathbf{k}\right\rangle = a_{k_N}^\dagger a_{k_{N-1}}^\dagger \cdots a_{k_2}^\dagger a_{k_1}^\dagger \left|\text{vac}\right\rangle = \left(\prod_{i=1}^{N} \hat{a}_{k_i}^\dagger\right)\left|\text{vac}\right\rangle$$

7.3 Number operators and Fock states

The number operators $\hat{n}_k$ are defined as

$$\hat{n}_k = \hat{a}_k^\dagger \hat{a}_k$$

This operator clearly preserves the total number of particles in the system (or reduces it to the zero vector). It is fairly straightforward to verify from the (anti-)commutation properties of the creations/annihilation operators that

$$[\hat{a}_j, \hat{n}_k] = \delta_{j,k} \quad [\hat{a}_j^\dagger, \hat{n}_k] = -\delta_{j,k}$$

both both fermions and bosons. This is identical in structure to the harmonic oscillator raising/lowering operators. This means the operator $\hat{n}_k$ acts like $\hat{a}^\dagger \hat{a} = (\hat{H} - \hbar\omega/2)/\hbar\omega$ for the harmonic oscillator to measure how many particles are in state k.

Likewise the operator $\hat{n}_k$ acts on $|\mathbf{k}\rangle$ to give

$$\hat{n}_k |\mathbf{k}\rangle = n_k(\mathbf{k}) |\mathbf{k}\rangle$$

so these states are eigenstates of $\hat{n}_k$. The number operators will always return 0 when acting on the vacuum state

$$\hat{n}_k |\text{vac}\rangle = 0$$

Just as for the harmonic oscillator, we can construct states that are simultaneous eigenstates of $\hat{n}_k$ as

$$|\mathbf{n}\rangle \equiv |n_1, n_2, \ldots, n_k \ldots\rangle = \prod_k \frac{1}{\sqrt{n_k!}} (\hat{a}_k^\dagger)^{n_k} |\text{vac}\rangle$$

where the product is understood such that the largest $\hat{a}_k^\dagger$ appears to the left. This set of states is complete, so forms a basis for the Fock space.

We note that the total number operator can be written as

$$\hat{N} = \sum_k \hat{n}_k$$

which measures the total number of particles in the state.

7.4 Basis changes

Suppose we change basis from $|k\rangle$, to $|\tilde{k}\rangle$ where

$$|\tilde{k}\rangle = \sum_j |j\rangle S_{jk}$$

where $\langle j|\tilde{k}\rangle = S_{jk}$. The transformation $\hat{S} = \sum_{jk} |j\rangle\langle k| S_{jk}$ is unitary, and transforms $|k\rangle$ to $|\tilde{k}\rangle$.

The non-symmetrised many-body state $\left|\tilde{k}_1(1), \tilde{k}_2(2), \ldots, \tilde{k}_3(N)\right\rangle$ can be written as

$$\left|\tilde{k}_1(1), \tilde{k}_2(2), \ldots, \tilde{k}_3(N)\right\rangle = \sum_{j_1 \cdots j_N} \prod_{i=1}^{N} S_{j_i,k_i} |j_1(1), j_2(2), \ldots, j_N(N)\rangle$$

so this extends to the symmetrised states

$$|\tilde{\mathbf{k}}\rangle = \sum_{j_1 \cdots j_N} \prod_{i=1}^{N} S_{j_i,k_i} \frac{\mathcal{N}_{\tilde{\mathbf{k}}}}{\mathcal{N}_{\mathbf{j}}} |\mathbf{j}\rangle$$

Noting that

$$|\mathbf{j}\rangle = \mathcal{N}_{\mathbf{j}} \prod_{i=1}^{N} \hat{a}^{\dagger}_{j_i} |\text{vac}\rangle$$

we find

$$\left|\tilde{k}_1(1), \tilde{k}_2(2), \ldots, \tilde{k}_3(N)\right\rangle = \mathcal{N}_{\tilde{\mathbf{k}}} \prod_{i=1}^{N} \left(\sum_{j_i} \hat{a}^{\dagger}_{j_i} S_{j_i,k_i} \right) |\text{vac}\rangle$$

So this implies

$$\hat{a}^{\dagger}_{\tilde{k}} = \sum_j \hat{a}^{\dagger}_j \langle j|\tilde{k}\rangle .$$

and likewise for the annihilation operators

$$\hat{a}_{\tilde{k}} = \sum_j \hat{a}_j \langle j|\tilde{k}\rangle^* = \sum_j \hat{a}_j \langle \tilde{k}|j\rangle .$$

7.5 Second quantised one-body operators

A general symmetrised one-body operator acting on an N particle state can be written in the form

$$\hat{A} = \sum_{i=1}^{N} \sum_{pq} A_{pq} |p(i)\rangle\langle q(i)|$$

where $|p(i)\rangle$ denotes the state where particle i is in basis state p. Let's consider how to construct

$$\hat{B}_{pq} = \sum_i |p(i)\rangle\langle q(i)|$$

in a second quantized form.

Firstly we note that

$$\hat{B}_{pp} = \sum_i |p(i)\rangle\langle p(i)|$$

simply measures the total number of particles in state p. This means $\hat{B}_{pp}$ maps onto the number operator $\hat{n}_p$.

Now we consider $p \neq q$. We can construct the operator $\hat{B}^x_{pq}$ as

$$\begin{aligned} \hat{B}^x_{pq} &= \hat{B}_{pq} + \hat{B}^{\dagger}_{pq} \\ &= \sum_i (|+_x(i)\rangle\langle +_x(i)| - |-_x(i)\rangle\langle -_x(i)|) \end{aligned}$$

where $\langle p|\pm_x\rangle = 1/\sqrt{2}$ and $\langle p|\pm_x\rangle = \pm 1/\sqrt{2}$. Using our change of basis formula (left as an exercise), we can express $\hat{B}^x_{pq}$ as

$$\hat{B}^x_{pq} \to \left(\hat{a}^{\dagger}_p \hat{a}_q + \hat{a}^{\dagger}_q \hat{a}_p\right)$$

We can repeat this with $\hat{B}^y_{pq} = i(\hat{B}_{pq} - \hat{B}^{\dagger}_{pq})$ to find

$$\hat{B}^y_{pq} \to i\left(\hat{a}^{\dagger}_p \hat{a}_q - \hat{a}^{\dagger}_q \hat{a}_p\right)$$

and noting that $\hat{B}_{pq} = (\hat{B}^x_{pq} - i\hat{B}^y_{pq})/2$ gives

$$\hat{B}_{pq} \to \hat{a}^{\dagger}_p \hat{a}_q$$

for all p and q.

Alternative proof: A direct derivation goes as follows. Firstly for bosons let's consider

$$\sum_i |p(i)\rangle\langle q(i)| \, |\mathbf{k}\rangle = \frac{1}{\sqrt{N!\prod_j n_j(\mathbf{k})!}} \sum_P \sum_i |p(i)\rangle\langle q(i)| \, \hat{P} \, |k_1(1), k_2(2), \ldots, k_3(N)\rangle$$

If q does not appear in $\mathbf{k}$ then this is zero. If q appears exactly once in the sequence $\mathbf{k}$ at the index i_q then we can use the permutational invariance of $\hat{B}_{pq}$ to deduce that

$$\begin{aligned}
\sum_i |p(i)\rangle\langle q(i)| \, |\mathbf{k}\rangle &= \frac{1}{\sqrt{N!\prod_j n_j(\mathbf{k})!}} \sum_P \hat{P} \sum_i |p(i)\rangle\langle q(i)| \, |k_1(1), k_2(2), \ldots, k_3(N)\rangle \\
&= \frac{1}{\sqrt{N!\prod_j n_j(\mathbf{k})!}} \sum_P \hat{P} \, |k_1(1), k_2(2), \ldots, p(i_q), \ldots, k_3(N)\rangle \\
&= \left|p, \mathbf{k}_q^-\right\rangle
\end{aligned}$$

In general if q appears $n_q(\mathbf{k})$ times then the expression above is multiplied by $n_q(\mathbf{k})$ (and i_q is now the first index for which $k_n = q$) so

$$\begin{aligned}
\sum_i |p(i)\rangle\langle q(i)| \, |\mathbf{k}\rangle &= \frac{n_q(\mathbf{k})}{\sqrt{N!\prod_j n_j(\mathbf{k})!}} \sum_P \hat{P} \, |k_1(1), k_2(2), \ldots, p(i_q), \ldots, k_3(N)\rangle \\
&= \sqrt{n_q(\mathbf{k})}\sqrt{n_p(\mathbf{k}_q^-)+1} \left|p, \mathbf{k}_q^-\right\rangle
\end{aligned}$$

The additional factors are introduced to produce the correct normalisation for $\left|p, \mathbf{k}_q^-\right\rangle$ We note that this is exactly equivalent to

$$\hat{b}_p^\dagger \hat{b}_q \, |\mathbf{k}\rangle \, .$$

For Fermions the argument is simpler because q can only appear either 0 or 1 times in the sequence $\mathbf{k}$ so

$$\sum_i |p(i)\rangle\langle q(i)| \, |\mathbf{k}\rangle = n_q(\mathbf{k})(1 - n_p(\mathbf{k}_q^-))\sigma_P^- \left|p, \mathbf{k}_q^-\right\rangle$$

This is equivalent to

$$\hat{c}_p^\dagger \hat{c}_q \, |\mathbf{k}\rangle \, .$$

So overall we can conclude that in order to go from one-body operators in the first quantisation to the second quantisation we use the replacement

$$\sum_i |p(i)\rangle\langle q(i)| \to \hat{a}_p^\dagger \hat{a}_q$$

7.6 Two body operators

A general two-body operator in the first quantisation can be written as

$$\hat{A} = \frac{1}{2} \sum_{pqrs} \langle p(1)q(2)|\hat{A}_{12}|r(1)s(2)\rangle \sum_{i,j\neq i} |p(i)q(j)\rangle\langle r(i)s(j)|$$

This motivates finding a second quantised form of

$$\hat{B}_{pqrs} = \sum_{i,j\neq i} |p(i)q(j)\rangle\langle r(i)s(j)|$$

We start by noting this can we written as

$$\hat{B}_{pqrs} = \hat{B}_{pr}\hat{B}_{qs} - \sum_i |p(i)\rangle\langle r(i)| \, |q(i)\rangle\langle s(i)| = \hat{B}_{pr}\hat{B}_{qs} - \delta_{r,q}\hat{B}_{ps}$$

In second quantised form this is

$$\begin{aligned}
\hat{B}_{pqrs} &= \hat{a}_p^\dagger \hat{a}_r \hat{a}_q^\dagger \hat{a}_s - \delta_{r,s} \hat{a}_p^\dagger \hat{a}_s \\
&= \hat{a}_p^\dagger (\pm \hat{a}_q^\dagger \hat{a}_r + \delta_{r,q}) \hat{a}_s - \delta_{r,q} \hat{a}_p^\dagger \hat{a}_s \\
&= \pm \hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_r \hat{a}_s \\
&= (\pm 1)^2 \hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_s \hat{a}_r \\
&= \hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_s \hat{a}_r
\end{aligned}$$

Note the change in order of r and s. Some authors use a slightly different definition of the two-particle matrix elements $A_{pqrs} = \langle p(1)q(2)|\hat{A}_{12}|r(1)s(2)\rangle$

7.7 Second quantised Hamiltonians

Overall we find that for Hamiltonians containing up to two-body interactions, if the first quantised Hamiltonian is

$$\hat{H} = \sum_i \sum_{pq} h_{pq} |p(i)\rangle\langle q(i)| + \frac{1}{2} \sum_{i,j\neq i} \sum_{pqrs} V_{pqrs} |p(i)q(j)\rangle\langle r(i)s(j)|$$

then the corresponding second quantised Hamiltonian is

$$\hat{H} = \sum_{pq} h_{pq} \hat{a}_p^\dagger \hat{a}_q + \frac{1}{2} \sum_{pqrs} V_{pqrs} \hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_s \hat{a}_r$$

In chemistry the matrix elements are sometimes written is so-called chemist's notation as

$$(pr|\hat{V}|qs) = V_{pqrs}$$

For an interaction which is just a function of $\hat{\boldsymbol{r}}_1 - \hat{\boldsymbol{r}}_1$ the chemist's integrals are

$$(pr|\hat{V}|qs) = \int \mathrm{d}\mathbf{x}_1 \int \mathrm{d}\mathbf{x}_2 \, \phi_p^*(\mathbf{x}_1)\phi_r(\mathbf{x}_1) V(\hat{\boldsymbol{r}}_{12}) \phi_q^*(\mathbf{x}_2)\phi_s(\mathbf{x}_2)$$

where we have introduced the space-spin coordinate $\phi(\mathbf{x}) \equiv \phi(\boldsymbol{r}, \sigma)$ and the integral over this coordinate is denotes an integral over space and a sum over spin coordinates. If $V(\boldsymbol{r}_{12})$ is the Coulomb interaction this is often just denoted $(pr|qs)$ for short.

7.8 Hartree-Fock theory

Hartree-Fock theory is the simplest theory for the many-electron wave-function fully accounting for the Coulomb interaction. It formalises the idea of molecular orbitals to include electron-electron intractions Our ansatz for the Hartree-Fock method is

$$|\Phi\rangle = \prod_{i=1}^{N} \hat{a}_i^\dagger |\mathrm{vac}\rangle$$

each orbital i is expressed as a sum of some basis (orthogonal) orbitals

$$|i\rangle = \sum_\mu c_{\mu i} |\mu\rangle$$

and we assume we can evaluated the one and two-body Hamiltonians in this basis. We enforce that

$$\langle i|j\rangle = \delta_{i,j} = \sum_\mu c_{\mu i}^* c_{\mu j}$$

With this ansatz the energy of the Hartree-Fock state can be evaluated, and then we can apply the **variational principle** to optimise the HF molecular orbitals to minimise the energy of the approximate wave function and give us a best estimate for the true ground state. Starting with the one-body term we have

$$\langle H_1 \rangle = \sum_{pq} h_{pq} \langle \Phi | \hat{a}_p^\dagger \hat{a}_q | \Phi \rangle = \sum_{pq} h_{pq} \delta_{p,q} \langle \Phi | \hat{n}_p | \Phi \rangle = \sum_{i=1}^{N} h_{ii}$$

For the interaction term we have

$$\langle V \rangle = \frac{1}{2} \sum_{pqrs} V_{pqrs} \langle \Phi | \hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_s \hat{a}_r | \Phi \rangle$$

because of the action of the creation annihilation operators, only $q = r$, $p = s$ or $q = s$, $p = r$ terms are non zero, and only for the occupied orbitals (denoted i, j, k etc.) so

$$\langle V \rangle = \frac{1}{2} \sum_{pq} (V_{pqpq} \langle \Phi | \hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_q \hat{a}_p | \Phi \rangle + V_{pqqp} \langle \Phi | \hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_p \hat{a}_q | \Phi \rangle)$$

using the anti-commutation relations we get

$$\langle V \rangle = \frac{1}{2} \sum_{pq} (V_{pqpq} - V_{pqqp}) \langle \Phi | \hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_q \hat{a}_p | \Phi \rangle$$

which again simplifies to

$$\langle V \rangle = \frac{1}{2} \sum_{i,j \neq i} (V_{ijij} - V_{ijji}) = \frac{1}{2} \sum_{i,j} (V_{ijij} - V_{ijji})$$

The V_{ijij} term is the classical coulomb repulsion, as can be seen by writing it in chemist's notation

$$V_{ijij} = (ii|jj)$$

whereas the $V_{ijji} = (ij|ji)$ term is the exchange integral that arises from the anti-symmetry.

We can write this energy out in terms of the wave-function coefficients as

$$\langle E \rangle = \sum_{\mu\nu} \sum_{i} c_{\mu i}^* c_{\nu i} h_{\mu\nu} + \frac{1}{2} \sum_{\mu\nu\mu'\nu'} \sum_{i,j} (\mu\nu|\mu'\nu')(c_{\mu i}^* c_{\nu i} c_{\mu' j}^* c_{\nu' j} - c_{\mu i}^* c_{\nu j} c_{\mu' j}^* c_{\nu' i})$$

This can be written in terms of the density matrix

$$D_{\mu\nu} = \sum_{i} c_{\mu i}^* c_{\nu i}$$

after some manipulation of indices in sums as

$$\langle E \rangle = \sum_{\mu\nu} D_{\mu\nu} h_{\mu\nu} + \frac{1}{2} \sum_{\mu\nu\mu'\nu'} [(\mu\nu|\mu'\nu') - (\mu\nu'|\mu'\nu)] D_{\mu\nu} D_{\mu'\nu'}$$

The basis functions need not always be orthogonal, but in general

$$\delta_{i,j} = \sum_{\mu\nu} c_{\mu i}^* S_{\mu\nu} c_{\nu j} \equiv g_{ij}[c]$$

We can enforce orthogonality of the Hartree-Fock orbitals using lagrange multipliers, so we minimise the functional

$$\mathcal{L} = \langle E \rangle - \sum_{ij} \lambda_{ij} (g_{ij} - \delta_{i,j})$$

Differentiating this gives

$$\frac{\partial \mathcal{L}}{\partial c^*_{\mu i}} = \sum_{\nu} F_{\mu\nu}[D] c_{\nu i} - \sum_{\nu}\sum_{ij} \lambda_{ij} S_{\mu\nu} c_{\nu j} = 0$$

$$\frac{\partial \mathcal{L}}{\partial c_{\mu i}} = \sum_{\nu} F_{\mu\nu}[D]^* c^*_{\nu i} - \sum_{\nu}\sum_{ij} \lambda_{ij} S^*_{\mu\nu} c^*_{\nu j} = 0$$

The matrix $F_{\mu\nu}[D]$ is called the Fock matrix, and it is given by

$$F_{\mu\nu}[D] = h_{\mu\nu} + \sum_{\mu'\nu'} (\mu\nu|\mu'\nu') D_{\mu'\nu'} - \sum_{\mu'\nu'} (\mu\nu|\mu'\nu) D_{\mu'\nu'}$$

The second term is denoted as the Coulomb matrix $J_{\mu\nu}$ and it represents the classical electron repulsion, and the third term is denoted the exchange matrix $K_{\mu\nu}[D]$ which represents stablisation due to anti-symmetry of the wave-function.

If we choose $c_{\mu i}$ to solve the generalised eigenvalue problem $F_{\mu\nu}[D]$, then they will automatically be orthogonal and $\lambda_{ij} = 0$ for $i \neq j$. We find that $\lambda_{ii} = \epsilon_i$ where ϵ_i is the eigenvalue for the generalised eigenvalue problem solves the above equations when $c_{\mu i}$ solves the generalised eigenvalue problem.

With this the total energy can be re-written as

$$\langle E \rangle = \sum_i \epsilon_i - \frac{1}{2} \sum_{\mu\nu\mu'\nu'} [(\mu\nu|\mu'\nu') - (\mu\nu'|\mu'\nu)] D_{\mu\nu} D_{\mu'\nu'}$$

Note that the because $F_{\mu\nu}[D]$ is itself a function of $c_{\mu i}$, so we have to **self-consistently** solve the generalised eigenvalue problem until the equations are satisfied. So in the Hartree-Fock method we solve the generalised eigenvalue problem

$$\mathbf{F}[D]\mathbf{c}_i = \epsilon_i \mathbf{S}\mathbf{c}_i$$

but the density matrix $D_{\mu\nu}$ depends on the coefficients of the occupied orbitals $\mathbf{c}_i$, so this equation needs to be solved self-consistently (this is actually a set of cubic equations for $c_{\mu,i}$). In practice we first take some guess of the occupied orbitals $\mathbf{c}_i^{(0)}$, build the density matrix $D^{(0)}_{\mu\nu}$ and Fock matrix, then diagonalise $\mathbf{F}[D^{(0)}]$ and take the lowest energy states to be occupied, and then update the density matrix $D^{(1)}_{\mu\nu}$ and use this to build a new Fock matrix. We then repeat the diagonalisation and the other steps until the orbitals (i.e. density matrix) stop changing. [In practice we often need some other tricks like damping, or Pulay mixing to get efficient convergence.]

7.8.1 Restricted and unrestricted HF

In the above the labels i for the HF orbitals are spin-space functions and $|\mu\rangle$ are spin-space one electron states. We can enforce singlet spin symmetry by associating both an α and β spin state with an identical spatial orbital

$$|i\rangle = |\varphi_i^{\text{space}}\rangle \otimes |\sigma_i\rangle$$

with $\sigma_i = \alpha, \beta$. This is less general than the form we had above. This is called the unrestricted HF (UHF) theory. Note that this ansatz will not necessarily be a true eigenstates of $\hat{S}^2$, although it will be an eigenstates of $\hat{S}_z$.

For a system with an even number of electrons we can restrict the ansatz further such that each there are always two electrons with different spins in a given spatial orbital, e.g.

$$|2i\rangle = |\varphi_i^{\text{space}}\rangle \otimes |\alpha\rangle$$

$$|2i+1\rangle = |\varphi_i^{\text{space}}\rangle \otimes |\beta\rangle$$

This defines RHF theory. This state will always be a singlet $\hat{S}^2 |\Phi_{\text{RHF}}\rangle = 0$.

When accounting for spin (see homework problems) we use a basis which is a product of a spatial function basis and α/β spin functions: $|\mu\sigma\rangle = |\mu\rangle \otimes |\sigma\rangle$. The sum over the basis μ above has to be replaced with a sum over both space and spin states $\sum_\mu \to \sum_{\mu,\sigma}$, and using the fact that the one and two electron terms are spin-indpendent, things simplify a lot. We define the spin-free density matrix as $\bar{D}_{\mu\nu} = \sum_\sigma D_{\mu\sigma,\nu\sigma}$ which for the RHD state is $\bar{D}_{\mu\nu} = 2\sum_i c^*_{\mu,i} c_{\nu,i}$ where i are the occupied spatial orbitals (so for N electrons only $N/2$ orbitals are occupied and included in this sum). The resulting RHF energy equation is

$$\langle E\rangle = \sum_{\mu\nu} \bar{D}_{\mu\nu} h_{\mu\nu} + \frac{1}{2}\sum_{\mu\nu\mu'\nu'}[(\mu\nu|\mu'\nu') - \frac{1}{2}(\mu\nu'|\mu'\nu)]\bar{D}_{\mu\nu}\bar{D}_{\mu'\nu'}$$

where we have to interpret $(\mu\nu|\mu'\nu')$ as an integral just over spatial degrees of freedom for the spatial part of the basis functions, so these electron-repulsion integrals are spin-free as well. The exchange term is halved because the exchange interaction is only present if electrons have the same spin. The Hartree-Fock equations can be derived in the same way as above for this system, and the resulting Fock matrix is

$$F_{\mu\nu}[\bar{D}] = h_{\mu\nu} + \sum_{\mu'\nu'}(\mu\nu|\mu'\nu')\bar{D}_{\mu'\nu'} - \frac{1}{2}\sum_{\mu'\nu'}(\mu\nu|\mu'\nu)\bar{D}_{\mu'\nu'} = h_{\mu\nu} + J_{\mu\nu}[\bar{D}] - \frac{1}{2}K_{\mu\nu}[\bar{D}]$$

In RHF we still need to solve the HF equation self-consistently, but with the RHF form of the Fock matrix.

In UHF we can always choose orbitals to be real-valued, which improves computational efficiency, and the wave function has the correct $\hat{S}_z$ symmetry by design. Likewise in RHF the HF equations simplify because there are only half as many independent wave-function coefficients, so this is more computationally efficient, plus total spin symmetry is enforced. Note also there is restricted-open HF (ROHF) which can treat systems with multiple unpaired electrons and yields eigenstates of $\hat{S}^2$ with $S > 0$.

7.9 Electron correlation

Hartree-Fock theory captures the exact electronic energy fairly well, however it misses "correlation" due the electron repulsion causing electrons to avoid each other. We know for example that the electronic wave-function should be "cusped" as two electrons approach each other, but the Hartree-Fock ansatz does not capture this. Hartree-Fock theory does provide the zeroth order approximation for more complicated ansätze that account for the missing electron correlation.

The simplest approximations in corporate excited electronic configurations into the wave-function

$$|\Psi\rangle = |\Phi\rangle + \sum_{ai} t_i^a \hat{a}_a^\dagger \hat{a}_i |\Phi\rangle + \sum_{abij} t_{ij}^{ab} \hat{a}_b^\dagger \hat{a}_a^\dagger \hat{a}_j \hat{a}_i |\Phi\rangle + \cdots$$

where the amplitudes t_i^a etc. are either determined perturbatively, variationally, or some other way. If we incorporate all possible excitations and determine the amplitudes variationally, then we would recover the exact energy. Because perturbation theory tells us that high energy excited states contribute less to the wave-function, we generally hope that truncating this expansion at low excitation levels will give a good approximation to the exact electronic energy.

7.10 Brillouin's theorem

In order to more concretely analyse electron correlation, let's re-write the electronic Hamiltonian as

$$\hat{H} = \hat{F} + \hat{v}$$

where $\hat{F}$ is the Fock operator,

$$\hat{F} = \sum_{pq} F_{pq} \hat{a}_p^\dagger \hat{a}_q$$

where we note that $F_{pq} = \delta_{p,q}\epsilon_p$, and $\hat{v}$ is the correlation potential , given by the remainder

$$\hat{v} = \frac{1}{2}\sum_{pqrs} \hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_s \hat{a}_r (pr|qs) - \sum_{pqi}((pq|ii) - (pi|iq))\hat{a}_p^\dagger \hat{a}_q$$

We can evaluate the matrix element between $|\Phi\rangle$ and

$$|\Phi_i^a\rangle = \hat{a}_a^\dagger \hat{a}_i |\Phi\rangle$$

For the two-body part we have

$$\langle\Phi|\hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_s \hat{a}_r \hat{a}_a^\dagger \hat{a}_i|\Phi\rangle$$

This will be non-zero only if $r = a$, $p = i$ and $q = s = j \in \text{occ}$, and $j \neq i$, or some permutation of these conditions. So the two-body part is

$$\langle\Phi|\hat{V}|\Phi_i^a\rangle = \sum_j (V_{ijaj} - V_{ijja}) = \sum_j [(ia|jj) - (ij|ja)]$$

The sum can be taken over all occupied j because the $j = i$ terms vanish. The one-body term in the fluctuation matrix element is non-zero only if $q = a$ and $p = i$. So clearly these two terms cancel.

Overall the only excited determinants that couple to the HF state through the fluctuation potential are doubly excited states. A similar argument to the above reveals that

$$\left|\Phi_{ij}^{ab}\right\rangle = \hat{a}_b^\dagger \hat{a}_j \hat{a}_a^\dagger \hat{a}_i |\Phi\rangle = \hat{a}_a^\dagger \hat{a}_b^\dagger \hat{a}_j \hat{a}_i |\Phi\rangle$$

is coupled by the matrix element

$$\left\langle\Phi\middle|\hat{v}\middle|\Phi_{ij}^{ab}\right\rangle = (ia|jb) - (ib|ja)$$

These results are enough to immediately apply second-order perturbation theory to obtain a correction to the Hartree-Fock energy that accounts for electron correlation

$$E^{(2)} = -\sum_{n\neq 0} \frac{|\langle 0|\hat{V}|n\rangle|^2}{E_n^{(0)} - E_0^{(0)}}$$

Identifying $|\Phi\rangle = |0\rangle$ and noting the only coupled states excited states are $|n\rangle = \left|\Phi_{ij}^{ab}\right\rangle$, this can be evaluated to give the "MP2" energy.

Aside: The state $|\Phi\rangle$ is sometimes called a Slater determinant because the wave-function can be written as

$$\Phi(\mathbf{x}_1, \ldots, \mathbf{x}_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \phi_1(\mathbf{x}_1) & \phi_1(\mathbf{x}_2) & \cdots & \phi_1(\mathbf{x}_N) \\ \phi_2(\mathbf{x}_1) & \phi_2(\mathbf{x}_2) & \cdots & \phi_2(\mathbf{x}_N) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_N(\mathbf{x}_1) & \phi_N(\mathbf{x}_2) & \cdots & \phi_N(\mathbf{x}_N) \end{vmatrix}$$

where $|\cdots|$ denotes the determinant. We saw this before in the our fundamentals of electronic structure.

Note that MP2 is only the beginning for how we can treat electron correlation. Other methods include CASSCF, CISD, CCSD, and many others. We won't aim to cover these here, but you should now have the basic tools to start to understand these theories.

8 Density functional theory

We saw in the previous section that it is quite a formidable task to evaluate the energy eigenvalues for general owing to the very large number of excited configurations $\hat{a}_a^\dagger \hat{a}_b^\dagger \cdots \hat{a}_i \hat{a}_j \cdots |\Phi_0\rangle$ that we need to account for in order to obtain accurate energies. In this section we will explore a very powerful formal theory to circumvent this problem. **Density functional theory** or **DFT**.

8.1 Background

8.1.1 Field operators and the electron density

It is useful to re-write the second quantised many-electron Hamiltonian in terms of **field operators** $\hat{\psi}(\mathbf{x})^\dagger \equiv \hat{\psi}_\sigma(\boldsymbol{r})^\dagger$. $\hat{\psi}_\sigma(\boldsymbol{r})^\dagger$ creates an electron in spin state $\sigma = \alpha$ or β at the point $\boldsymbol{r}$

$$\hat{\psi}_\sigma(\boldsymbol{r})^\dagger |\text{vac}\rangle = |\boldsymbol{r}\sigma\rangle$$

These can be written in terms of the standard creation/annihilation operators (using the basis transformation we found in setting up the second quantisation) as

$$\hat{\psi}_\sigma(\boldsymbol{r})^\dagger = \sum_p \hat{a}_p^\dagger \langle p|\boldsymbol{r}\sigma\rangle$$

and they obey the expected anti-commutation relations

$$\{\hat{\psi}_\sigma(\boldsymbol{r}), \hat{\psi}_{\sigma'}(\boldsymbol{r}')\} = \{\hat{\psi}_\sigma(\boldsymbol{r})^\dagger, \hat{\psi}_{\sigma'}(\boldsymbol{r}')^\dagger\} = 0$$
$$\{\hat{\psi}_\sigma(\boldsymbol{r}), \hat{\psi}_{\sigma'}(\boldsymbol{r}')^\dagger\} = \delta(\boldsymbol{r} - \boldsymbol{r}')\delta_{\sigma,\sigma'}$$

The second quantised Hamiltonian can now be written as

$$\hat{H} = \hat{T} + \int \mathrm{d}\boldsymbol{r}\, V_{\text{ext}}(\boldsymbol{r}) \sum_\sigma \hat{\psi}_\sigma(\boldsymbol{r})^\dagger \hat{\psi}_\sigma(\boldsymbol{r}) + \hat{U}$$

$$\hat{U} = \frac{1}{2} \int \mathrm{d}\boldsymbol{r} \int \mathrm{d}\boldsymbol{r}'\, U(\boldsymbol{r} - \boldsymbol{r}') \sum_{\sigma,\sigma'} \hat{\psi}_\sigma(\boldsymbol{r})^\dagger \hat{\psi}_{\sigma'}(\boldsymbol{r}')^\dagger \hat{\psi}_{\sigma'}(\boldsymbol{r}') \hat{\psi}_\sigma(\boldsymbol{r})$$

here $U(\boldsymbol{r}) = 1/|\boldsymbol{r}|$ is the standard Coulomb repulsion between electrons. $V_{\text{ext}}(\boldsymbol{r})$ is the external potential generated (typically) by the nuclear charges, but in DFT this is generalised to an arbitrary potential.

In DFT the central quantity is not the many-electron wave function $|\Psi\rangle$ but instead the electron density $n(\boldsymbol{r})$

$$n(\boldsymbol{r}) = \sum_\sigma \langle\Psi|\hat{\psi}_\sigma^\dagger(\boldsymbol{r})\hat{\psi}_\sigma(\boldsymbol{r})|\Psi\rangle$$

This is just the number of electrons per unit volume at a point in space. This quantity is positive $n(\boldsymbol{r}) \geq 0$ and integrates to the total number of electrons N_e

$$N_e = \int \mathrm{d}\boldsymbol{r}\, n(\boldsymbol{r})$$

The density is measured with the density operator $\hat{n}(\boldsymbol{r})$ (the operator that measures the density) which can be defined using the field operators as (using a direct product basis for the one electron Hilbert space $|p\sigma\rangle = |\varphi_p\rangle_{\text{space}} \otimes |\sigma\rangle_{\text{spin}}$)

$$\hat{n}(\boldsymbol{r}) = \sum_\sigma \hat{\psi}_\sigma(\boldsymbol{r})^\dagger \hat{\psi}_\sigma(\boldsymbol{r}) = \sum_{p,q,\sigma} \varphi_p^*(\boldsymbol{r})\varphi_q(\boldsymbol{r})\hat{a}_{p\sigma}^\dagger \hat{a}_{q\sigma}$$

Note that this is not the same as the density operator $\hat{\rho}$ describe the state of a specific quantum system, althoguh confusingly it is called the same thing.

8.1.2 Functionals

In what follows we will be dealing a lot with **functionals**. These are a special class of functions $F[n]$ that take functions as their arguments and return a $\boldsymbol{r}$ independent real or complex number. A simple example is

$$N[n] = \int \mathrm{d}\boldsymbol{r}\, n(\boldsymbol{r})$$

but they can also be more complicated, for example they may involve multiple integrals or gradients of the input function. In fact they can be very abstract indeed, and we may not always be able to write down an analytic closed form expression for a particular functional.

8.2 N-Representability

One key result needed to formulate DFT is the fact that any physical electron density $n(\boldsymbol{r})$ which is non-negative everywhere $n(\boldsymbol{r}) \geq 0$ and has a positive integer integral $\int_{\mathbb{R}^3} \mathrm{d}\boldsymbol{r}\, n(\boldsymbol{r}) = N$ can be obtained the density of an antisymmetrised N electron wave function.

The proof of this goes as follows. For a given density let us define

$$\varphi_k(\boldsymbol{r}) = \sqrt{n(\boldsymbol{r})/N}\, \exp(2\pi k f_k(x) i)$$

with $k = 0, \ldots, N-1$ with

$$f(x) = \frac{1}{N} \int_{-\infty}^{x} \mathrm{d}x \int_{-\infty}^{\infty} \mathrm{d}y \int_{-\infty}^{\infty} \mathrm{d}z\, n(\boldsymbol{r})$$

It's fairly straightforward (do this as an exercise) to show that these orbitals are orthonormal, and if there are all occupied an a Hartree-Fock state $|\Phi\rangle$ (with arbitrary spin functions) then the density is $n(\boldsymbol{r})$. So there are a whole family of quantum states that can represent the density $n(\boldsymbol{r})$.

8.3 First Hohenberg-Kohn Theorem: Densities determine external potentials

The first important theorem of DFT states that the external potential $V_{\text{ext}}(\boldsymbol{r})$ is (up to an additive constant) uniquely determined by the ground state electron density $n_0(\boldsymbol{r})$. Clearly the external potential, through the Hamiltonian also determines the ground state density, so this overall means that there is a **one-to-one** correspondence between the ground state electron density and the external potential.

Proof: The proof of this follows a proof by contradiction set-up. Suppose we have two external potentials $V(\boldsymbol{r})$ and $V'(\boldsymbol{r})$ and

$$V(\boldsymbol{r}) \neq V'(\boldsymbol{r}) + C$$

Each potential defines a many electron Hamiltonian $\hat{H}$ and $\hat{H}'$. These have non-degenerate ground state wave functions $|\Psi_0\rangle$ and $|\Psi_0'\rangle$. Now suppose these give rise to the <u>same</u> ground state density $n_0(\boldsymbol{r})$.

We know by the variational theorem that if $|\Psi_0\rangle \neq e^{i\phi} |\Psi_0'\rangle$

$$E_0 = \langle \Psi_0 | \hat{H} | \Psi_0 \rangle < \langle \Psi_0' | \hat{H} | \Psi_0' \rangle$$

But the right hand side can also be written as

$$\begin{aligned}
\langle \Psi_0' | \hat{H} | \Psi_0' \rangle &= \langle \Psi_0' | (\hat{H} - \hat{H}') | \Psi_0' \rangle + \langle \Psi_0' | \hat{H}' | \Psi_0' \rangle \\
&= \langle \Psi_0' | (\hat{V} - \hat{V}') | \Psi_0' \rangle + E_0' \\
&= \int \mathrm{d}\boldsymbol{r}\, (V(\boldsymbol{r}) - V'(\boldsymbol{r})) n_0(\boldsymbol{r}) + E_0'
\end{aligned}$$

So overall we have

$$E_0 < \int \mathrm{d}\boldsymbol{r}\, (V(\boldsymbol{r}) - V'(\boldsymbol{r}))n_0(\boldsymbol{r}) + E'_0$$

We can also repeat this argument swapping the two Hamiltonians and wave functions, so we also find

$$E'_0 < \int \mathrm{d}\boldsymbol{r}\, (V'(\boldsymbol{r}) - V(\boldsymbol{r}))n_0(\boldsymbol{r}) + E_0$$

Adding these together gives

$$E_0 + E'_0 < E_0 + E'_0 \implies 0 < 0$$

This is a contradiction, so our original premise that $V(\boldsymbol{r}) \neq V'(\boldsymbol{r}) + C$ must be false. We conclude that the ground state density also determines the external potential.

Consequences: Formally through the full many body Hamiltonian, the external potential determines the ground state wave function, which also determines the ground state density

$$V_{\mathrm{ext}}(\boldsymbol{r}) \to \hat{H} \to |\Psi_0\rangle \to n_0(\boldsymbol{r})$$

The first HK theorem shows also that the density determines the potential $n_0(\boldsymbol{r}) \to V_{\mathrm{ext}}(\boldsymbol{r})$, so this also means

$$n_0(\boldsymbol{r}) \to V_{\mathrm{ext}}(\boldsymbol{r}) \to \hat{H} \to |\Psi_0\rangle$$

So this means that the groundstate wavefunction and Hamiltonian are formally (operator and wave-function valued) functionals of the ground state density

$$|\Psi_0\rangle \equiv |\Psi_0[n]\rangle \quad \hat{H} \equiv \hat{H}[n]$$

Actually writing down a closed form expression for these functionals is not possible, but formally this relationship exists.

Note: Above we assumed the ground states to be non-degenerate. It is possible to relax this assumption, but these get more fiddly so we'll only consider non-degenerate ground states here.

8.4 Second Hohenberg-Kohn Theorem: A Universal Variational Energy Functional

The first HK theorem showed us that the ground state density uniquely determines the external potential, and therefore the many-body Hamiltonian and ground state wave function. The second theorem states that the ground state energy is the minimum of a universal functional $E[n]$

$$E_0 = \min_n E[n]$$

where $n(\boldsymbol{r})$ integrates to a particular fixed total number of electrons N_e.

Furthermore this functional can be written as

$$E[n] = F[n] + \int \mathrm{d}\boldsymbol{r}\, n(\boldsymbol{r})V_{\mathrm{ext}}(\boldsymbol{r})$$

Where $F[n]$ encodes the electron kinetic and interaction energy as a functional of the density. The external potential term we will denote as $V[n]$ for short. This is independent of the external potential, so in this sense it is universal.

Proof: The proof of the second HK theorem uses the variational theorem energy, together with the first HK theorem.

Let us define $F[n]$ as

$$F[n] = \min_{\Psi \to n} \langle \Psi | (\hat{T} + \hat{U}) | \Psi \rangle$$

The $\min_{\Psi \to n}$ here means minimising over wave functions that yield a particular fixed density $n(\boldsymbol{r})$.

Using this we know that for any wave function $|\Psi\rangle$ yielding a density $n(\boldsymbol{r})$

$$\langle \Psi | \hat{H} | \Psi \rangle \geq F[n] + \int \mathrm{d}\boldsymbol{r}\, V_{\mathrm{ext}}(\boldsymbol{r}) n(\boldsymbol{r})$$

This means the ground state energy can be found by minimising the energy functional above with respect to the density because

$$\min_{\Psi} \langle \Psi | (\hat{T} + \hat{U}) | \Psi \rangle = \min_{n} \min_{\Psi \to n} \langle \Psi | (\hat{T} + \hat{U}) | \Psi \rangle$$

so the ground state energy is given by

$$E_0 = \min_{\Psi} \langle \Psi | \hat{H} | \Psi \rangle = \min_{n} (F[n] + V[n])$$

This strengthens the first theorem to a variational theorem.

Consequences: Formally if we know the external potential felt by electrons, we can variationally determine the ground state energy by minimising $E[n]$. Furthermore the kinetic/interaction functional is independent of the external potential, so we only formally have to determine it for an arbitrary external potential, and we can then apply it to any system.

Unfrotunately the formal definition of $F[n]$ still involves constructing wave functions, so it's completely impractical. In what follows we'll outline the steps to constructing more practically useful DFT.

8.5 Kohn-Sham DFT

8.5.1 Splitting the universal functional

We have established that formally there exists a functional $E[n]$ that when minimised with respect to the density yields the ground state energy of the electronic system. Unfortunately the expression we found still involves full-dimensional wavefunctions, so it's not particularly useful. To make progress, it is useful to split $F[n]$ into a kinetic energy and potential energy part

$$F[n] = T[n] + U[n]$$

We can also examine the interaction term

$$U[n] = \frac{1}{2} \int \mathrm{d}\boldsymbol{r} \int \mathrm{d}\boldsymbol{r'}\, U(\boldsymbol{r} - \boldsymbol{r'}) \sum_{\sigma,\sigma'} \langle \Psi | \hat{\psi}_{\sigma}(\boldsymbol{r})^{\dagger} \hat{\psi}_{\sigma'}(\boldsymbol{r'})^{\dagger} \hat{\psi}_{\sigma'}(\boldsymbol{r'}) \hat{\psi}_{\sigma}(\boldsymbol{r}) | \Psi \rangle$$

We can write this in terms of the two-particle density

$$n^{(2)}_{\sigma,\sigma'}(\boldsymbol{r}, \boldsymbol{r'}) = \langle \Psi | \hat{\psi}_{\sigma}(\boldsymbol{r})^{\dagger} \hat{\psi}_{\sigma'}(\boldsymbol{r'})^{\dagger} \hat{\psi}_{\sigma'}(\boldsymbol{r'}) \hat{\psi}_{\sigma}(\boldsymbol{r}) | \Psi \rangle$$

This can be split into a classical uncorrelated density term, and an "exchange-correlation" hole term

$$n^{(2)}_{\sigma,\sigma'}(\boldsymbol{r}, \boldsymbol{r'}) = n_{\sigma}(\boldsymbol{r}) n_{\sigma'}(\boldsymbol{r'}) + n_{\sigma}(\boldsymbol{r}) n_{\sigma'}(\boldsymbol{r'}) h_{\sigma,\sigma'}(\boldsymbol{r}, \boldsymbol{r'})$$

where the exchange correlation hole is defined as

$$h_{\sigma,\sigma'}(\boldsymbol{r}, \boldsymbol{r'}) = \frac{n^{(2)}_{\sigma,\sigma'}(\boldsymbol{r}, \boldsymbol{r'})}{n_{\sigma}(\boldsymbol{r}) n_{\sigma'}(\boldsymbol{r'})} - 1$$

This function should decay to zero for large separations $|\boldsymbol{r}-\boldsymbol{r}'|$ because electrons should become uncorrelated at this distances (at least for molecules). This exchange correlation hole contains all of the exchange antisymmetry effects and correlation effects of the electron interaction.

This division of the two-particle density in this way allows us to write the interaction term as

$$U[n] = \frac{1}{2}\int \mathrm{d}\boldsymbol{r} \int \mathrm{d}\boldsymbol{r'}\, n(\boldsymbol{r})U(\boldsymbol{r}-\boldsymbol{r}')n(\boldsymbol{r}') + E_{xc}[n] \equiv E_H[n] + E_{xc}[n]$$

The first term is just the classical coulomb energy of a charge density, also called the Hartree energy. The second term $E_{xc}[n]$ is the exchange-correlation functional that formally contains all the non-classical and correlation effects. The first term normally makes up a large part of the overall energy of a system, but exchange-correlation effects are still very important.

The simple Coulomb term includes a self-interaction. Just using the Coulomb term on its own for a one-electron system would not give the correct total energy, because the electron interacts with itself. An exchange contribution is needed to exactly cancel this for one-electron systems. All approximate density functionals contain this so-called self-interaction error.

8.5.2 The kinetic energy functional

The kinetic energy functional $T[n]$ formally only depends on the density. However in practice it is challenging to obtain an exact form of the kinetic energy functional. We note however that in the entire derivation of the Hohenberg-Kohn theorems, we never assumed anything about the form of the electron-electron interaction. This means it also holds true for systems where electrons are non-interacting. This this case we know the exact wave function is a simple slater determinant

$$|\Phi\rangle = \prod_i \hat{a}_i^\dagger \,|\mathrm{vac}\rangle$$

and the kinetic energy is

$$\langle T\rangle = \sum_i \langle\varphi_i|\hat{T}|\varphi_i\rangle$$

so in this limit we know the exact exchange correlation functional, which we denote $F_0[n]$

$$F_0[n] = T_0[n] = \min_{\Phi\to n} \sum_i \langle\varphi_i|\hat{T}|\varphi_i\rangle$$

where $\min_{\Phi\to n}$ means taking the minimum over single Slater determinant wave functions, which is a much smaller space of wave functions than the full anti-symmetrised Hilbert space. $|\varphi_i\rangle$ is a one electron state, and these states are orthonormal.

8.5.3 Kohn-Sham DFT

We have seen that we can write-down the exact universal kinetic/interaction functional in the case of non-interacting electrons. We can also represent any density $n(\boldsymbol{r})$ with a single Slater determinant $|\Phi\rangle$ with

$$n(\boldsymbol{r}) = \sum_{i,\sigma} |\varphi_{i,\sigma}(\boldsymbol{r})|^2$$

where $\varphi_{i,\sigma}(\boldsymbol{r})$ are the orthonormal orbitals occupied by the non-interacting electrons. Using this we can represent ground state energy as

$$\begin{aligned}
E_0 &= \min_n (F[n] + V[n]) \\
&= \min_n \left(T_0[n] + (F[n] - T_0[n]) + V[n]\right) \\
&= \min_n \min_{\Phi \to n} \Big(\langle \Phi | \hat{T} | \Phi \rangle + (F[n[\Phi]] - T_0[n[\Phi]]) + V[n[\Phi]] \Big) \\
&= \min_{\Phi} \Big(\langle \Phi | \hat{T} | \Phi \rangle + (F[n[\Phi]] - T_0[n[\Phi]]) + V[n[\Phi]] \Big)
\end{aligned}$$

In the last two lines we introduced an explicit dependence of $n(\boldsymbol{r})$ on Φ, but we will drop this below. The new functional

$$F_{\text{KS}}[n] = F[n] - T_0[n] = E_H[n] + E_{xc}[n] + \Delta T[n]$$

contains electron Hartree energy, exchange-correlation energy and corrections to the kinetic energy. Evaluating the energy of the effective non-interacting electronic system is more involved than just working with the density, but the non-interacting kinetic energy captures most of the interacting kinetic energy. This proves invaluable in developing an accurate and useful approximate density functional theory. Often the (unknown) correction to the kinetic energy functional is lumped in with the (also unknown) exchange correlation functional, although it is important to remember that in Kohn-Sham DFT this term is formally present. Almost all DFT calculations use the Kohn-Sham formalism, where the density and energy is constructed using the fictitious Kohn-Sham non-interacting electron system given here.

8.5.4 Exchange and correlation hole

Above we defined the full spin-dependent exchange-correlation hole. Let us define the spin-free exchange correlation hole as

$$h_{xc}(\boldsymbol{r}, \boldsymbol{r}') = \frac{n^{(2)}(\boldsymbol{r}, \boldsymbol{r}')}{n(\boldsymbol{r})n(\boldsymbol{r}')} - 1 \quad n^{(2)}(\boldsymbol{r}, \boldsymbol{r}') = \sum_{\sigma,\sigma'} n^{(2)}(\boldsymbol{r}, \boldsymbol{r}')$$

The exchange hole $h_x(\boldsymbol{r}, \boldsymbol{r}')$ captures the portion of the exchange-correlation hole that arises purely from Pauli exchange effects, i.e. the fact that no two electrons can occupie the same quantum state (point in space with same spin) at the same time. The Kohn-Sham reference non-interacting electron system provides a definition for the exchange hole $h_x(\boldsymbol{r}, \boldsymbol{r}')$. This is because the the Kohn-Sham system has exchange symmetry built in (because it is a non-interacting system of indistinguishable electrons), and it formally exactly reproduces the ground state energy and density, but the Kohn-Sham wave function has no correlation effects (because the electrons do not interact). This means its exchange-correlation hole cannot contain any correlation effects, just exchange, so we can define the exchange hole as

$$h_x(\boldsymbol{r}, \boldsymbol{r}') = \frac{n^{(2)}_{\text{KS}}(\boldsymbol{r}, \boldsymbol{r}')}{n(\boldsymbol{r})n(\boldsymbol{r}')} - 1$$

The correlation hole arises from the electrons avoiding each other in space due to repulsion. The full exchange-correlation hole is the sum of correlation and exchange parts, so the correlation hole is defined as

$$h_c(\boldsymbol{r}, \boldsymbol{r}') = h_{xc}(\boldsymbol{r}, \boldsymbol{r}') - h_x(\boldsymbol{r}, \boldsymbol{r}')$$

These also allow us to formally define the exchange and correlation functionals in terms of these holes

$$E_x[n] = \int \mathrm{d}\boldsymbol{r}\, n(\boldsymbol{r}) \frac{1}{|\boldsymbol{r} - \boldsymbol{r}'|} n(\boldsymbol{r}') h_x(\boldsymbol{r}, \boldsymbol{r}')$$

$$E_c[n] = \int \mathrm{d}\boldsymbol{r}\, n(\boldsymbol{r}) \frac{1}{|\boldsymbol{r} - \boldsymbol{r}'|} n(\boldsymbol{r}') h_c(\boldsymbol{r}, \boldsymbol{r}')$$

This means the exchange correlation hole is also a functional of the density. Using this the problem of finding approximate functionals can be tackled by finding an approximate model for the exchange-correlation holes in terms of the density. Unfortunately this remains an exceptionally hard problem.

8.5.5 Functional derivatives

We have an expression for the Kohn-Sham DFT energy in. Assuming we have the KS exchange correlation functional, or some reasonable approximation to it, in order to obtain this energy we need to minimise the functional with respect to the occupied single electron orbitals φ_i.

Formally we can do this using **functional differentiation**. Let us assume that the functional can be written as

$$F[f] = \int \mathrm{d}\boldsymbol{r}\, L(f(\boldsymbol{r}))$$

where $L(f)$ is some function. This is a common functional form that we encounter, and we will generalise this below. Let us consider how this changes on addition of a small change to the function f, $f(\boldsymbol{r}) \to f(\boldsymbol{r}) + \delta g(\boldsymbol{r})$.

$$\delta F = \int \mathrm{d}\boldsymbol{r}\, (L(f(\boldsymbol{r}) + \epsilon g(\boldsymbol{r})) - L(f(\boldsymbol{r}))) = \int \mathrm{d}\boldsymbol{r}\, \frac{\partial L(f)}{\partial f} \delta g(\boldsymbol{r})$$

So in order for δF to be zero for any small function $\delta g(\boldsymbol{r})$, we need

$$\frac{\partial L(f)}{\partial f} = 0$$

The functional derivative $\frac{\delta F}{\delta f}$ is defined such that

$$\delta F = \int \mathrm{d}\boldsymbol{r}\, \frac{\delta F}{\delta f} \delta g(\boldsymbol{r})$$

We can repeat the above analysis for different forms of L, e.g. $L = L(f, \nabla f)$

$$\delta F = \int \mathrm{d}\boldsymbol{r} \left(\frac{\partial L(f, \nabla f)}{\partial f} \delta g(\boldsymbol{r}) + \sum_\alpha \frac{\partial L(f, \nabla f)}{\partial(\partial_\alpha f)} \frac{\partial}{\partial r_\alpha} \delta g(\boldsymbol{r}) \right)$$

Using integration by parts (and assuming $\delta g(\boldsymbol{r})$ vanishes at the integral limits) we find

$$\delta F = \int \mathrm{d}\boldsymbol{r} \left(\frac{\partial L(f, \nabla f)}{\partial f} - \sum_\alpha \frac{\partial}{\partial r_\alpha} \frac{\partial L(f, \nabla f)}{\partial(\partial_\alpha f)} \right) \delta g(\boldsymbol{r})$$

So the functional derivative in this case is

$$\frac{\delta F}{\delta f} = \frac{\partial L(f, \nabla f)}{\partial f} - \sum_\alpha \frac{\partial}{\partial r_\alpha} \frac{\partial L(f, \nabla f)}{\partial(\partial_\alpha f)}$$

We can also generalise this to integrands involving higher derivatives of f and more complex functionals, for example those involving multiple integrals like the Hartree energy.

8.5.6 The Kohn-Sham equations

Equipped with the basic tools functional differentiation, we can find equations for orthonormal orbitals as

$$\frac{\delta \mathcal{L}[\Phi]}{\delta \varphi^*_{i,\sigma}} = 0$$

where the Lagrangian $\mathcal{L}[\Phi]$ is

$$\mathcal{L}[\Phi] = E[\Phi] - \sum_{i \geq j} \epsilon_{i,j} \left(\sum_\sigma \int \mathrm{d}\boldsymbol{r}\, \varphi^*_{i,\sigma}(\boldsymbol{r}) \varphi_{j,\sigma}(\boldsymbol{r}) - \delta_{i,j} \right)$$

The Lagrange multipliers $\epsilon_{i,j}$ enforce the orthonormality. Performing the functional differentiation we obtain

$$\frac{\delta E[\Phi]}{\delta \varphi_{i,\sigma}^*} = \sum_{i \geq j} \epsilon_j \varphi_{j,\sigma}(\boldsymbol{r})$$

Now due to the symmetry of the energy functional in φ_i, because it is defined in terms of the real-valued density, if we choose

$$\frac{\delta E[\Phi]}{\delta \varphi_{i,\sigma}^*} = \epsilon_i \varphi_{i,\sigma}(\boldsymbol{r})$$

orthogonality of the orbitals is automatically enforces, so the Lagrange multipliers are $\epsilon_{i,j} = \epsilon_i \delta_{i,j}$. The ϵ_i are the Kohn-Sham orbital energies. They don't have any direct physical meaning apart from the highest occupied moelcular orbital (HOMO) energy. For the HOMO obtained with the **exact** density functional the HOMO energy is the negative of the first ionisation energy for the system

$$\epsilon_{\text{HOMO}} = -E_{\text{IE}}$$

For approximate density functionals this will not be true, so it can serve as a useful test for the applicability of a particular approximate density functional.

Now let us state the result for the final Kohn-Sham equations

$$\epsilon_i \varphi_{i,\sigma}(\boldsymbol{r}) = \hat{T} \varphi_{i,\sigma}(\boldsymbol{r}) + V_{\text{ext}}(\boldsymbol{r}) \varphi_{i,\sigma}(\boldsymbol{r}) + V_H(\boldsymbol{r}) \varphi_{i,\sigma}(\boldsymbol{r}) + V_{xc}(\boldsymbol{r}) \varphi_{i,\sigma}(\boldsymbol{r})$$

the Hartree potential V_H is the potential due to all other electrons in the system

$$V_H(\boldsymbol{r}) = \int \mathrm{d}\boldsymbol{r}\, U(\boldsymbol{r} - \boldsymbol{r}') n(\boldsymbol{r}')$$

and the exchange-correlation potential is given by

$$V_{xc}(\boldsymbol{r}) = \frac{\delta E_{xc}}{\delta n}(\boldsymbol{r})$$

When we expand the orbitals in a particular basis $\varphi_i = \sum_\mu c_{\mu i} \chi_\mu$ we can use the KS equations to obtain a set of self-consistent equations for coefficients

$$\mathbf{F}_{\text{KS}}[\mathbf{D}] \boldsymbol{c}_i = \epsilon_i \mathbf{S} \boldsymbol{c}_i$$

where $\boldsymbol{D}$ is the density matrix we encountered before in Hartree-Fock theory. The main difference between HF and KS DFT is in the exchange term appearing in the Fock matrix $\mathbf{F}[\mathbf{D}]$. In HF this only includes exchange effects, given by a non-local potential, whereas in KS-DFT it includes both exchange and correlation effects in the V_{xc} term.

8.6 Adiabatic Connection

One final thing we will discuss is the adiabatic connection. This is a useful theorem for connecting Hartree-Fock and other wave function theories to DFT, and it's also a useful starting point for building approximations to the exact density functional (DFAs).

We know that for non-interacting electrons the exact ground state wave function is given by a single Hartree-Fock state. Likewise our analysis of Hartree-Fock theory and MP2 (perturbationt theory) showed that if the electron-electron interaction is scaled by λ then the energy is given by

$$E(\lambda) = E_{\text{HF}}(\lambda) + \lambda^2 E_{\text{MP2}}(\lambda) + \mathcal{O}(\lambda^3)$$

so the exact energy is given by the Hartree Fock energy, with an error $\mathcal{O}(\lambda^2)$. This means any DFA should reduce to HF as $\lambda \to 0$. In terms of the DFA we require that

$$E^\lambda[n] = \min_{\Phi \to n} E_{\text{HF}}^\lambda[\Phi] \quad \text{as } \lambda \to 0$$

This motivates adding some “exact” exchange into a DFA (these are called hybrid functionals – see below).

A more formal framework for doing this is the **adiabatic connection** where it is possible to show that

$$E_{xc}[n] = \int_0^1 \mathrm{d}\lambda\, W_\lambda[n] - E_H[n]$$

with

$$W_\lambda[n] = \langle \Psi_\lambda[n] | \hat{U} | \Psi_\lambda[n] \rangle$$

where $|\Psi_\lambda[n]\rangle$ is the many-body quantum state which minimises the energy with the density constrained to $n(\boldsymbol{r})$ for the electron-electron interaction strength scaled by λ. We know that for non-interacting electrons, when $\lambda = 0$, $|\Psi_\lambda[n]\rangle = |\Phi[n]\rangle$ where $|\Phi[n]\rangle$ is the Kohn-Sham non-interacting state that produces the density $n(\boldsymbol{r})$. This means

$$W_{\lambda=0}[n] = \langle \Phi[n] | \hat{U} | \Phi[n] \rangle$$

We can approximate $W_\lambda[n]$ using known limits or other approximations, for example

$$W_\lambda[n] \approx \langle \Phi[n] | \hat{U} | \Phi[n] \rangle + f(\lambda)(W_{\mathrm{DFA}}[n] - \langle \Phi[n] | \hat{U} | \Phi[n] \rangle)$$

with $f(\lambda \to 0) = 0$. Integrating this yields an XC energy functional that contains a portion of “exact” Hartree-Fock type exchange. This justifies mixing in HF exchange into approximate DFAs. This can also be used to justify a range of other functionals, for example range-separated and double-hybrid type functionals (see below).

8.7 Density Functionals

There are many recipes for concocting density functionals. They all have their advantages and disadvantages.

Density functionals generally are written in terms of a **local exchange-correlation energy density** $\varepsilon_{xc}[n](\boldsymbol{r})$

$$E_{xc}[n] = \int \mathrm{d}\boldsymbol{r}\, n(\boldsymbol{r}) \varepsilon_{xc}[n](\boldsymbol{r})$$

There are many way to cosntruct DFAs. Some are based on using known limits (such as the uniform electron gas), and/or using the adiabatic conenction. Others are more empirical. In either case the DFAs vary in complexity based on what math the complexity of the mathematical function $\varepsilon_{xc}[n](\boldsymbol{r})$.

8.7.1 Types of Exchange–Correlation Functionals

The exchange–correlation functional $E_{xc}[n]$ formally contains all of the complex quantum effects of exchange and correlation, but its exact form is unknown. In practice, various approximate forms are used. These are often organised according to how much information about the electron density they use, forming what is sometimes called **Jacob’s ladder** of density functional approximations. Each rung of this ladder increases in both accuracy and computational cost.

(1) Local Density Approximation (LDA) The simplest approximation is the **Local Density Approximation** (LDA). Here the exchange–correlation energy density at each point is assumed to be the same as that of a uniform electron gas with the same local density,

$$E_{xc}^{\mathrm{LDA}}[n] = \int \mathrm{d}\boldsymbol{r}\ n(\boldsymbol{r})\, \varepsilon_{xc}^{\mathrm{UEG}}(n(\boldsymbol{r})). \tag{1}$$

LDA works surprisingly well for systems with slowly varying electron densities, such as simple metals, but tends to overestimate binding energies and underestimate band gaps in molecules and semiconductors.

(2) Generalized Gradient Approximation (GGA) The **Generalized Gradient Approximation** (GGA) improves upon the LDA by including the local gradient of the electron density, allowing the functional to respond to spatial inhomogeneities:

$$E_{xc}^{\mathrm{GGA}}[n] = \int \mathrm{d}\boldsymbol{r}\; f(n(\boldsymbol{r}), \nabla n(\boldsymbol{r})). \tag{2}$$

This additional information makes GGAs much more accurate for molecular geometries, atomization energies, and surface properties. Popular examples include PBE (Perdew–Burke–Ernzerhof) and BLYP (Becke–Lee–Yang–Parr).

(3) Meta-GGA (MGGA) A further refinement is the **Meta-Generalized Gradient Approximation** (Meta-GGA or MGGA), which depends not only on the density and its gradient but also on higher-level quantities such as the kinetic energy density $\tau(\boldsymbol{r})$ or the Laplacian $\nabla^2 n(\boldsymbol{r})$. This provides additional information about the local orbital structure and type of bonding:

$$E_{xc}^{\mathrm{MGGA}}[n] = \int \mathrm{d}\boldsymbol{r}\; f(n(\boldsymbol{r}), \nabla n(\boldsymbol{r}), \tau(\boldsymbol{r})). \tag{3}$$

The **kinetic energy density**, denoted $\tau(\boldsymbol{r})$, represents the local contribution of the Kohn–Sham orbitals to the total noninteracting kinetic energy. It is defined as

$$\tau(\boldsymbol{r}) = \frac{1}{2} \sum_i^{\mathrm{occ}} |\nabla \psi_i(\boldsymbol{r})|^2, \tag{4}$$

where the sum runs over all occupied Kohn–Sham orbitals $\psi_i(\boldsymbol{r})$. The integral of $\tau(\boldsymbol{r})$ over all space gives the total Kohn–Sham kinetic energy,

$$T_s = \int \tau(\boldsymbol{r})\; \mathrm{d}\boldsymbol{r}\,. \tag{5}$$

In meta-GGA functionals, $\tau(\boldsymbol{r})$ provides information about the local orbital character of the density, helping to distinguish between regions of different bonding types (e.g. covalent, metallic, or weakly bound).

Examples include the SCAN and TPSS functionals, which can achieve near-hybrid accuracy with moderate computational cost.

(4) Hybrid Functionals Finally, **Hybrid Functionals** mix a fraction of exact Hartree–Fock exchange with the DFT exchange and correlation components,

$$E_{xc}^{\mathrm{hybrid}} = a\, E_x^{\mathrm{HF}} + (1-a)\, E_x^{\mathrm{DFT}} + E_c^{\mathrm{DFT}}, \tag{6}$$

where a is an empirical mixing parameter, typically around 0.2–0.25. Well-known hybrid functionals include B3LYP, PBE0, and HSE06. Hybrids generally improve thermochemical and electronic property predictions, at the cost of greater computational effort.

Each rung of this hierarchy improves the accuracy of the exchange–correlation energy at the expense of additional computational cost. The appropriate choice of functional depends on the type of system and the required balance between accuracy and efficiency.

The exact exchange-correlation potential should asymptotically become $-1/r$. One way to enforce this is through range-separated hybrids, which use Hartree fock exchange for the long-range part of the Coulomb potential. Examples of this include ωB97X-D3 and ωPBE.

(5) Non-local functionals Dispersion interactions are relatively long range, $\sim 1/R^6$, and the long range part cannot be captured with the local density descriptors above. Sometimes this is corrected for empirically using simple functions of the atom identity and distances (e.g. D2, D3, D3(BJ), D4), but these don't alter the density, this is just an empirical correction. Non-local functionals can more rigorously account for this by adding functional terms like

$$E_{nl} = \frac{1}{2}\int \mathrm{d}\boldsymbol{r}\, n(\boldsymbol{r}) \int \mathrm{d}\boldsymbol{r'}\, \phi_{nl}(|\boldsymbol{r} - \boldsymbol{r'}|) n(\boldsymbol{r'})$$

where $\phi_{nl}(r) \sim 1/r^6$ at large distances. The price to pay for using these is that they are much more expensive to self-consistently evaluate, so often they're added in non-self-consistently after an SCF KS calculation with the rest of the functional.

(6) Double-hybrid In double hybrid functionals a proportion of MP2 correlation energy is added in. This also helps capture dispersion, but it adds significant cost and basis set-dependence because now the accuracy depends not only on how well the occupied orbitals are represented in the chosen basis set, but also the virtual orbitals need to be well represented.

9 Molecular Hamiltonians

In this section of the course we will explore how to put together a framework for describing molecular properties and dynamics with QM.

9.1 The Hamiltonian

The basic Hamiltonian for N_e electrons interacting with N_n nuclei, with charges Z_A is

$$\hat{H} = \hat{T}_n + \hat{T}_e + \hat{V}_{nn} + \hat{V}_{en} + \hat{V}_{ee}$$

$$\hat{T}_n = -\frac{1}{2}\sum_{A}^{N_n} \frac{1}{M_A}\hat{P}_A^2$$

$$\hat{T}_e = -\frac{1}{2}\sum_{i=1}^{N_e} \hat{p}_i^2$$

$$\hat{V}_{nn} = +\frac{1}{2}\sum_{A,B\neq A} \frac{Z_A Z_B}{|\boldsymbol{R}_A - \boldsymbol{R}_B|}$$

$$\hat{V}_{en} = -\sum_{A,i} \frac{Z_A}{|\boldsymbol{r}_i - \boldsymbol{R}_A|}$$

$$\hat{V}_{ee} = +\frac{1}{2}\sum_{i\neq j} \frac{1}{|\boldsymbol{r}_i - \boldsymbol{r}_j|}$$

We have used atomic units so $m_e = 1$, $M_A = \frac{M_A}{m_e}$

In general this is very difficult to solve. But we can make some simplifying approximations to gain physical insight.

9.2 For The electronic Hamiltonian

First we note that

$$\frac{m_e}{M_A} < \frac{1}{2000}$$

So generally we expect, from perturbation theory, the effect of $\hat{T}_n$ to generally be small compared to other terms. This motivates of splitting $\hat{H}$ into

$$\hat{H} = \hat{H}_e + \hat{T}_n$$

Now $\hat{H}_e$ is diagonal in the nuclear position basis:

$$|\boldsymbol{R}\rangle = |\boldsymbol{R}_1\rangle \otimes |\boldsymbol{R}_2\rangle \otimes \cdots \otimes |\boldsymbol{R}_{N_n}\rangle$$

$$\langle \boldsymbol{R}'|\hat{H}_e|\boldsymbol{R}\rangle = \delta(\boldsymbol{R} - \boldsymbol{R}')\hat{H}_e(\boldsymbol{R})$$

where $\hat{H}_e(\boldsymbol{R})$ is just a function of $\boldsymbol{R}_1$, $\boldsymbol{R}_2$, ... but still an operator on the electrons.

$$\hat{H}_e(\boldsymbol{R}) = \hat{T}_e + \hat{V}_{ee} + \sum_{A,i} \frac{Z_A}{|\boldsymbol{r}_i - \boldsymbol{R}_A|} + \frac{1}{2} \sum_{A,B\neq A} \frac{Z_A Z_B}{|\boldsymbol{R}_A - \boldsymbol{R}_B|}$$

In principle this can be solved to give a set of eigenstates & eigenvalues that depend parametrically on $\boldsymbol{R}$.

The states $|\Phi_k(\boldsymbol{R})\rangle$ are the electronic eigenstates for a given $\boldsymbol{R}$, and we denote the eigenvalues as $V_j(\boldsymbol{R})$

$$\hat{H}_e(\boldsymbol{R})|\Phi_k(\boldsymbol{R})\rangle = V_k(\boldsymbol{R})|\Phi_k(\boldsymbol{R})\rangle$$

and we note that these states are orthonormal, $\langle\Phi_j(\boldsymbol{R})|\Phi_k(\boldsymbol{R})\rangle = \delta_{jk}$, or equivalently

$$\int \mathrm{d}\boldsymbol{x}\, \Phi_j(\boldsymbol{x}|\boldsymbol{R})^* \Phi_k(\boldsymbol{x}|\boldsymbol{R}) = \delta_{jk}$$

and they form a basis for the electronic Hilbert space for any choice of $\boldsymbol{R}$.

The states $|\Phi_k(\boldsymbol{R})\rangle$ are called the **adiabatic** electronic states, because they are the electronic energy eigenstates that adiabatically follow the nuclear coordinates $\boldsymbol{R}$ if these coordinates are very slowly changed.

9.3 Born-Huang expansion

We now return to the full molecular Hamiltonian. The states $|\Phi_k(\boldsymbol{R})\rangle$ form an electronic basis, so we can expand any state of the nuclei + electrons as with this for any $\boldsymbol{R}$. The states

$$|\boldsymbol{R}, \Phi_k(\boldsymbol{R})\rangle \equiv |\boldsymbol{R}\rangle \otimes |\Phi_k(\boldsymbol{R})\rangle$$

form a basis, so we can write any full quantum state of the nuclear-electronic system as

$$|\Psi\rangle = \sum_k \int d\boldsymbol{R} |\boldsymbol{R}\rangle \otimes |\Phi_k(\boldsymbol{R})\rangle\langle \boldsymbol{R}, \Phi_k(\boldsymbol{R})|\Psi\rangle$$

$$= \sum_k \int d\boldsymbol{R} \chi_k(\boldsymbol{R}) |\boldsymbol{R}\rangle \otimes |\Phi_k(\boldsymbol{R})\rangle$$

or written in terms of wave-functions

$$\Psi(\boldsymbol{x}, \boldsymbol{R}) = \sum_k \chi_k(\boldsymbol{R}) \Phi_k(\boldsymbol{x}|\boldsymbol{R})$$

where $\boldsymbol{x} = (r_1, s_1, r_2, s_2, \ldots)$ is the vector of electron space-spin coordinates, and $\langle \boldsymbol{x}|\Phi_k(\boldsymbol{R})\rangle = \Phi_k(\boldsymbol{x}|\boldsymbol{R})$.

This is called the **Born-Huang expansion**. Sometimes this is also called the wave function in the **adiabatic representation**.

We can extract the nuclear wave-function associated with each electronic state , using the orthonormality of the electronic states, as

$$\chi_j(\boldsymbol{R}) = \int \mathrm{d}\boldsymbol{x}\, \Phi_j(\boldsymbol{x}|\boldsymbol{R})^* \Psi(\boldsymbol{x}, \boldsymbol{R}) = \langle \boldsymbol{R}, \Phi_j(\boldsymbol{R})|\Psi\rangle$$

and we can define the nuclear states as

$$|\chi_j\rangle = \int \mathrm{d}\boldsymbol{R}\, |\boldsymbol{R}\rangle\, \chi_j(\boldsymbol{R}) = \int \mathrm{d}\boldsymbol{R}\, |\boldsymbol{R}\rangle\, \langle \boldsymbol{R}, \Phi_j(\boldsymbol{R})|\Psi\rangle\, .$$

9.4 The Schrödinger equation in the Born-Huang expansion

In order to find a more practical eigenvalue equation, or dynamical equation, for $|\Psi\rangle$ we need to evaluate $\hat{H}|\Psi\rangle$ in the Born-Huang expansion.

We recall that

$$\hat{H}_e(\boldsymbol{R})|\Phi_k(\boldsymbol{R})\rangle = V_k(\boldsymbol{R})|\Phi_k(\boldsymbol{R})\rangle$$

and $\langle\Phi_j(\boldsymbol{R})|\Phi_k(\boldsymbol{R})\rangle = \delta_{jk}$

Electronic energy term: With this we can evaluate $\langle\boldsymbol{R},\Phi_k(\boldsymbol{R})|\hat{H}|\Psi\rangle$ from the $\hat{H}_e$ part:

$$\langle\boldsymbol{R},\Phi_k(\boldsymbol{R})|\hat{H}_e|\Psi\rangle = V_k(\boldsymbol{R})\langle\boldsymbol{R},\Phi_k(\boldsymbol{R})|\Psi\rangle = V_k(\boldsymbol{R})\chi_k(\boldsymbol{R})$$

Nuclear kinetic energy term: The $\hat{T}_n$ part is more complex. Let's consider

$$\hat{\boldsymbol{P}}_A|\boldsymbol{R},\Phi_k(\boldsymbol{R})\rangle\chi_k(\boldsymbol{R})$$

or in the wave function representation

$$\hat{\boldsymbol{P}}_A\left(\chi_k(\boldsymbol{R})\Phi_k(\boldsymbol{x}|\boldsymbol{R})\right) = \left(-i\nabla_A\chi_k(\boldsymbol{R})\right)\Phi_k(\boldsymbol{x}|\boldsymbol{R}) + \chi_k(\boldsymbol{R})\left(-i\nabla_A\Phi_k(\boldsymbol{x}|\boldsymbol{R})\right)$$

Importantly $\frac{\partial}{\partial R_{A\alpha}}\Phi_k(\boldsymbol{x}|\boldsymbol{R}) \neq 0$ due to the parametric dependence of $\Phi_k(\boldsymbol{x}|\boldsymbol{R})$ on $\boldsymbol{R}$.
Let us note that

$$\hat{\boldsymbol{P}}_A\left(\chi_k(\boldsymbol{R})\Phi_k(\boldsymbol{x}|\boldsymbol{R})\right) = \left(-i\nabla_A\chi_k(\boldsymbol{R})\right)\Phi_k(\boldsymbol{x}|\boldsymbol{R}) + \chi_k(\boldsymbol{R})\left(-i\nabla_A\Phi_k(\boldsymbol{x}|\boldsymbol{R})\right)$$

and

$$\hat{\boldsymbol{P}}_A\cdot\hat{\boldsymbol{P}}_A\left(\chi_k(\boldsymbol{R})\Phi_k(\boldsymbol{x}|\boldsymbol{R})\right) = \left(-\nabla_A^2\chi_k(\boldsymbol{R})\right)\Phi_k(\boldsymbol{x}|\boldsymbol{R}) + 2\left(-i\nabla_A\chi_k(\boldsymbol{R})\right)\left(-i\nabla_A\Phi_k(\boldsymbol{x}|\boldsymbol{R})\right) + \chi_k(\boldsymbol{R})\left(-\nabla_A^2\Phi_k(\boldsymbol{x}|\boldsymbol{R})\right)$$

Taking the sum over A & k and projecting with $\langle\Phi_j(\boldsymbol{R})|$ gives $\langle\boldsymbol{R},\Phi_j(\boldsymbol{R})|\hat{T}_n|\Psi\rangle$

$$\langle\boldsymbol{R},\Phi_j(\boldsymbol{R})|\hat{T}_n|\Psi\rangle = -\frac{1}{2}\sum_A\frac{1}{M_A}\nabla_A^2\chi_j(\boldsymbol{R}) - \frac{i}{2}\sum_A\sum_k\frac{1}{M_A}\boldsymbol{A}_{jk}^{A*}\cdot\left(-i\nabla_A\chi_k(\boldsymbol{R})\right) - \frac{1}{2}\sum_A\frac{1}{M_A}G_{jk}^A\chi_k(\boldsymbol{R})$$

with

$$\boldsymbol{A}_{jk}^A = i\langle\Phi_j(\boldsymbol{R})|\nabla_A\Phi_k(\boldsymbol{R})\rangle$$

$$G_{jk}^A = \langle\Phi_j(\boldsymbol{R})|\nabla_A^2\Phi_k(\boldsymbol{R})\rangle$$

We note that

$$-i\nabla_A\cdot\boldsymbol{A}_{jk}^A = -i\nabla_A\cdot\langle\Phi_j(\boldsymbol{R})|i\nabla_A\Phi_k(\boldsymbol{R})\rangle = \langle i\nabla_A\Phi_j(\boldsymbol{R})|i\nabla_A\Phi_k(\boldsymbol{R})\rangle + G_{jk}^A$$

Inserting $\hat{1} = \sum_\ell|\Phi_\ell(\boldsymbol{R})\rangle\langle\Phi_\ell(\boldsymbol{R})|$ we find

$$-i\nabla_A\cdot\boldsymbol{A}_{jk}^A = G_{jk}^A + \sum_\ell\boldsymbol{A}_{\ell j}^{A*}\cdot\boldsymbol{A}_{\ell k}^A$$

We note also that

$$i\nabla_A\langle\Phi_k|\Phi_j\rangle = -\langle i\nabla_A\Phi_k|\Phi_j\rangle + \langle\Phi_k|i\nabla_A\Phi_k\rangle = -\boldsymbol{A}^{A*}_{kj} + \boldsymbol{A}^A_{kj} = i\nabla_A\delta_{jk} = 0$$

$\Rightarrow \boldsymbol{A}^{A*}_{\ell j} = \boldsymbol{A}^A_{j\ell}$

Combining this we arrive at

$$\langle\boldsymbol{R},\Phi_j(\boldsymbol{R})|\hat{T}_n|\Psi\rangle = +\frac{1}{2}\sum_{\ell,k}\sum_A\frac{1}{M_A}\left(-i\nabla_A\delta_{j\ell} + \boldsymbol{A}^A_{j\ell}\right)\cdot\left(-i\nabla_A\delta_{\ell k} + \boldsymbol{A}^A_{\ell k}\right)\chi_k(\boldsymbol{R})$$

The Schrödinger equation in the Born-Huang expansion: We can put the above terms together concisely to find the action of the full Hamiltonian in the Born-Huang expansion (a.k.a. the adiabatic representation)

$$\langle\boldsymbol{R},\Phi_j(\boldsymbol{R})|\hat{H}|\Psi\rangle = \sum_k \hat{H}_{jk}\chi_k(\boldsymbol{R})$$

$$\hat{H}_{jk} = \frac{1}{2}\sum_\ell\sum_A\frac{1}{M_A}\left(-i\nabla_A\delta_{j\ell} - \boldsymbol{A}^A_{j\ell}\right)\left(-i\nabla_A\delta_{\ell k} - \boldsymbol{A}_{\ell k}\right) + \delta_{jk}V_j(\boldsymbol{R})$$

Putting this together we can write the time-dependent Schrödinger equation by noting that with the Born Huang expansion to electronic states are time-independent so

$$|\Psi(t)\rangle = \sum_j\int \mathrm{d}\boldsymbol{R}\,\chi_j(\boldsymbol{R},t)\,|\boldsymbol{R},\Phi_j(\boldsymbol{R})\rangle$$

so the nuclear wave functions satisfy

$$i\frac{\partial}{\partial t}\chi_j(\boldsymbol{R},t) = \sum_k \hat{H}_{jk}\chi_k(\boldsymbol{R},t)$$

or for the energy eigenvalue equation $\hat{H}\,|\Psi\rangle = E\,|\Psi\rangle$

$$E\chi_j(\boldsymbol{R}) = \sum_k \hat{H}_{jk}\chi_k(\boldsymbol{R}).$$

Importantly $\hat{H}_{jk}$ is non-zero for $j \neq k$, so nuclear wave functions associated with different adiabatic electronic states are coupled together by terms dependent on the coupling vectors $\boldsymbol{A}^A_{jk}$.

9.5 Hellmann-Feynman Theorem for couplings

The coupling term $\boldsymbol{A}^A_{jk}$ can be written as

$$\langle\Phi_j|i\nabla_A\Phi_k\rangle = \boldsymbol{A}^A_{jk}$$

Let's consider $j \neq k$

$$0 = \langle\Phi_j|\hat{H}_e|\Phi_k\rangle$$

$$\Rightarrow 0 = i\nabla_A\langle\Phi_j|\hat{H}_e|\Phi_k\rangle = -\langle i\nabla_A\Phi_j|\hat{H}_e|\Phi_k\rangle + \langle\Phi_k|\hat{H}_e|i\nabla_A\Phi_k\rangle + \langle\Phi_j|(i\nabla_A\hat{H}_e)|\Phi_k\rangle$$

$$= -V_k\boldsymbol{A}^{A*}_{kj} + V_j\boldsymbol{A}^A_{jk} + \langle\Phi_j|(i\nabla_A\hat{H}_e)|\Phi_k\rangle$$

$$\Rightarrow \boldsymbol{A}^A_{jk} = -\frac{i\langle\Phi_j|\nabla_A \hat{H}_e|\Phi_k\rangle}{V_j - V_k}$$

So as long as $|V_j - V_k|$ is large, $\boldsymbol{A}^A_{jk}$ will be small.

9.6 The Born-Oppenheimer Approximation

For positions where adiabatic surfaces are well-separated in energy from each other, the couplings will be small. This means if initially $\Psi = \chi_j(\boldsymbol{R})\Phi_j(\boldsymbol{x}|\boldsymbol{R})$, other adiabatic states will never be populated so we just have

$$-\frac{1}{2}\sum_A \frac{1}{M_A}\nabla^2_A \chi_j(\boldsymbol{R}) + V_j(\boldsymbol{R})\chi_j(\boldsymbol{R}) = -i\partial_t \chi_j(\boldsymbol{R})$$

So the system evolves on a single adiabatic surface, as if the other electronic states weren't even there. This is precisely the **Born-Oppenheimer** approximation. We can use the same approximation in the time-independent Schrödinger equation to define the Born-Oppenheimer energy eigenstates. We can always formally correct for the Born-Oppenheimer approximation because we are essentially doign perturbation theory. Our zeroth order Hamiltonian is the diagonal Born-Oppenheimer Hamiltonian, where any term involing $\nabla_A \Phi_k(\boldsymbol{x}|\boldsymbol{R})$ is ignored. The effects of these terms can be added back in with perturbation theory.

This is essential for understanding lots of chemical reactivity, because we can treat nuclear motion on a single potential energy surface with wells and barriers. It's also essential for spectroscopy like vibrational, rotational and electronic spectra (see below).

9.7 Electronic Spectroscopy and the Franck-Condon approximation

If we make the Born-Oppenheimer approximation we can always write eigenstates of a molecular Hamiltonian as (sometimes called vibronic states)

$$\Psi_n(\boldsymbol{x}, \boldsymbol{R}) = \langle \boldsymbol{R}, \boldsymbol{x}|E_n\rangle = \chi_n(\boldsymbol{R})\Phi_{j_n}(\boldsymbol{x}|\boldsymbol{R})$$

where j_n is the adiabatic state index associated with the total molecular eigenstate n, and χ_n is the associated nuclear wave function. It's straightforward to show that within this approximation the energy is

$$E_n = E^{\text{nuc}}_n + V_{j_n}(\boldsymbol{R}_{j_n,0})$$

where the first term is a nuclear energy and the second term is the energy at the minimum of adiabat j_n.

The transition probability between these states in spectroscopy is given by the the transition dipole moment

$$\langle E_n|\hat{\boldsymbol{\mu}}|E_m\rangle = \boldsymbol{\mu}_{nm}$$

$$\hat{\boldsymbol{\mu}} = \sum_A Z_A \hat{\boldsymbol{R}}_A - \sum_i \hat{\boldsymbol{r}}_i$$

Using the Born-Oppenheimer approximation we can simplify this to

$$\boldsymbol{\mu}_{nm} = \langle\chi_n|\ \langle\Phi_{j_n}|\hat{\boldsymbol{\mu}}|\Phi_{j_m}\rangle_{\text{e}}\,|\chi_m\rangle$$

the $\langle\Phi_{j_n}|\hat{\boldsymbol{\mu}}|\Phi_{j_m}\rangle_{\text{e}} = \boldsymbol{\mu}_{j_n j_m}(\hat{\boldsymbol{R}})$ term is an operator on the nuclear degrees of freedom, the subscript means we just integrate over electronic coordinates. For electronic transitions we can further make the **Franck-Condon** approximation where we treat this as a constant at the minimum energy geometry of the lower state $\boldsymbol{R}_0$ (this is called the "Franck Condon point")

$$\boldsymbol{\mu}_{j_n j_m}(\hat{\boldsymbol{R}}) \approx \boldsymbol{\mu}_{j_n j_m}(\boldsymbol{R}_0)$$

and with this the transition dipole moment is

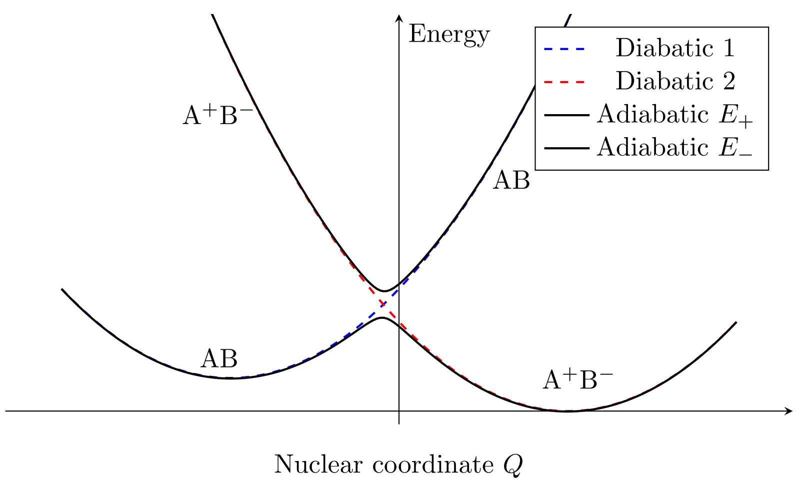
$$\boldsymbol{\mu}_{nm} \approx \boldsymbol{\mu}_{j_n j_m}(\boldsymbol{R}_0) \langle \chi_n | \chi_m \rangle$$

and because the nuclear wave functions are associated with different adiabatic potential energy surfaces the overlap $\langle \chi_n | \chi_m \rangle$ is not simply 0 or 1. This factor is called the Franck-Condon factor. Together the electronic transition dipole moment $\boldsymbol{\mu}_{nm}(\boldsymbol{R}_0)$ at the Franck-Condon point, and the Frank Condon factor $\langle \chi_n | \chi_m \rangle$ determine the probability of a given electronic-vibrational (vibronic) energy state transition.

9.8 Quasi-diabatic states

Adiabatic electronic states form the basis of many phenomena in chemistry, such as reactivity at typical temperatures. But there are situations where they break down. One example is in electron transfer reactions (and excitation energy transfer as in FRET). In these situations two adiabatic states come very close in energy, so the couplings diverge. This makes it impossible to apply perturbation theory to this type of problem.

Looking more closely at the two adiabatic potential energy curves we have. A typical situation for electron transfer is illustrated below. As a function of some nuclear coordinate Q on the left hand side the lower adiabat corresponds to molecular fragments A and B both being neutral, but the upper surface an electron has transferred from A to be to give A^+B^-. On the right hand side however the situation is reversed and the lower adiabat corresponds to the charge transfer state and the upper corresponds to the neutral state. In between where the two curves come very close lower and upper adiabatcs correspond to partial charge transfer state $A^{+0.5}B^{-0.5}$.



If the gap is large, then the electron transfer is adiabatic, so if we start on the lower surface, we stay there and the electron continuously transfers. However in the limit where the adiabatic gap is small we can instead define new states that are called (quasi-)diabatic states,

$$|\Xi_1(\boldsymbol{R})\rangle = c_{1,+}(\boldsymbol{R}) |\Phi_+(\boldsymbol{R})\rangle + c_{1,-}(\boldsymbol{R}) |\Phi_-(\boldsymbol{R})\rangle$$
$$|\Xi_2(\boldsymbol{R})\rangle = c_{2,+}(\boldsymbol{R}) |\Phi_+(\boldsymbol{R})\rangle + c_{2,-}(\boldsymbol{R}) |\Phi_-(\boldsymbol{R})\rangle$$

These states are linear combinations of the adibatic states, where the position dependent coefficients are chosen to remove the divergent terms in the non-adiabatic coupling. We previously encountered this idea looking at the adiabatic theorem and LiF, but now we have a more formal framework to understand this.

The molecular hamiltonian then approximately reduces to

$$\hat{H} \approx \hat{T}_n + \sum_{k=1,2} U_k(\boldsymbol{R}) |\Xi_k(\boldsymbol{R})\rangle\langle\Xi_k(\boldsymbol{R})| + \Delta_{12}(\boldsymbol{R}) |\Xi_1(\boldsymbol{R})\rangle\langle\Xi_2(\boldsymbol{R})| + \Delta_{12}^*(\boldsymbol{R}) |\Xi_2(\boldsymbol{R})\rangle\langle\Xi_1(\boldsymbol{R})|$$

The diagonal terms $U_k(\boldsymbol{R})$ are called the diabatic energies and the off-diagonal terms $\Delta_{12}(\boldsymbol{R})$ are called the diabatic couplings. The diabatic coupling will typically be small, which allows perturbation theory to now be applied (more on this later).

Formally quasi-diabatic states have a nuclear coordinate dependence, so the derivative couplings $\langle \Xi_j | \nabla_A \Xi_k \rangle$ are non-zero, which means the nuclear kinetic energy operator in the quasi-diabatic basis is not exactly diagonal. However quasi-diabatic state derivative couplings **do not** follow the Hellman-Feynman theorem, so they do not diverge at state intersections. In general we can more safely assume the derivative couplings with quasi-diabats are small, unlike with adiabatic states.

Diabatic states are **not unique**. This means they are not universally useful or even well-defined. There are many ways to define them, one such way is to define an electronic operator $\hat{Q}_A$ and $\hat{Q}_B$ that measures the electronic charge on fragments A and B, and to diagonalise the difference operator in a basis of adiabatic states

$$\Delta Q_{ij}(\boldsymbol{R}) = \langle \Phi_i(\boldsymbol{R}) | (\hat{Q}_A(\boldsymbol{R}) - \hat{Q}_B(\boldsymbol{R})) | \Phi_j(\boldsymbol{R}) \rangle$$

The resulting eigenvectors define the diabatic states.

This type of situation is also naturally encountered when spin-orbit coupling is included. The diabats are then called the **spin-diabats**.