

Supplementary subject: Quantum Chemistry

Perturbation theory

6 lectures, (Tuesday and Friday, weeks 4-6 of Hilary term)

Chris-Kriton Skylaris

(*chris-kriton.skylaris @ chem.ox.ac.uk*)

Physical & Theoretical Chemistry Laboratory
South Parks Road, Oxford

February 24, 2006

Bibliography

All the material required is covered in “Molecular Quantum Mechanics” fourth edition by Peter Atkins and Ronald Friedman (OUP 2005). Specifically, Chapter 6, first half of Chapter 12 and Section 9.11.

Further reading:

“Quantum Chemistry” fourth edition by Ira N. Levine (Prentice Hall 1991).

“Quantum Mechanics” by F. Mandl (Wiley 1992).

“Quantum Physics” third edition by Stephen Gasiorowicz (Wiley 2003).

“Modern Quantum Mechanics” revised edition by J. J. Sakurai (Addison Wesley Longman 1994).

“Modern Quantum Chemistry” by A. Szabo and N. S. Ostlund (Dover 1996).

Contents

1	Introduction	2
2	Time-independent perturbation theory	2
2.1	Non-degenerate systems	2
2.1.1	The first order correction to the energy	4
2.1.2	The first order correction to the wavefunction	5
2.1.3	The second order correction to the energy	6
2.1.4	The closure approximation	8
2.2	Perturbation theory for degenerate states	8
3	Time-dependent perturbation theory	12
3.1	Revision: The time-dependent Schrödinger equation with a time-independent Hamiltonian	12
3.2	Time-independent Hamiltonian with a time-dependent perturbation . . .	13
3.3	Two level time-dependent system - Rabi oscillations	17
3.4	Perturbation varying “slowly” with time	19
3.5	Perturbation oscillating with time	20
3.5.1	Transition to a single level	20
3.5.2	Transition to a continuum of levels	21
3.6	Emission and absorption of radiation by atoms	23
4	Applications of perturbation theory	28
4.1	Perturbation caused by uniform electric field	28
4.2	Dipole moment in uniform electric field	28
4.3	Calculation of the static polarizability	29
4.4	Polarizability and electronic molecular spectroscopy	30
4.5	Dispersion forces	32
4.6	Revision: Antisymmetry, Slater determinants and the Hartree-Fock method	35
4.7	Møller-Plesset many-body perturbation theory	36

1 Introduction

In these lectures we will study perturbation theory, which along with the variation theory presented in previous lectures, are the main techniques of approximation in quantum mechanics. Perturbation theory is often more complicated than variation theory but also its scope is broader as it applies to any excited state of a system while variation theory is usually restricted to the ground state.

We will begin by developing perturbation theory for stationary states resulting from Hamiltonians with potentials that are independent of time and then we will expand the theory to Hamiltonians with time-dependent potentials to describe processes such as the interaction of matter with light. Finally, we will apply perturbation theory to the study of electric properties of molecules and to develop Møller-Plesset many-body perturbation theory which is often a reliable computational procedure for obtaining most of the correlation energy that is missing from Hartree-Fock calculations.

2 Time-independent perturbation theory

2.1 Non-degenerate systems

The approach that we describe in this section is also known as “Rayleigh-Schrödinger perturbation theory”. We wish to find approximate solutions of the time-independent Schrödinger equation (TISE) for a system with Hamiltonian \hat{H} for which it is difficult to find exact solutions.

$$\hat{H}\psi_n = E_n\psi_n \quad (1)$$

We assume however that we know the exact solutions $\psi_n^{(0)}$ of a “simpler” system with Hamiltonian $\hat{H}^{(0)}$, i.e.

$$\hat{H}^{(0)}\psi_n^{(0)} = E_n^{(0)}\psi_n^{(0)} \quad (2)$$

which is not too different from \hat{H} . We further assume that the states $\psi_n^{(0)}$ are *non-degenerate* or in other words $E_n^{(0)} \neq E_k^{(0)}$ if $n \neq k$.

The small difference between \hat{H} and $\hat{H}^{(0)}$ is seen as merely a “perturbation” on $\hat{H}^{(0)}$ and all quantities of the system described by \hat{H} (the perturbed system) can be expanded as a Taylor series starting from the unperturbed quantities (those of $\hat{H}^{(0)}$). The expansion is done in terms of a parameter λ .

We have:

$$\hat{H} = \hat{H}^{(0)} + \lambda\hat{H}^{(1)} + \lambda^2\hat{H}^{(2)} + \dots \quad (3)$$

$$\psi_n = \psi_n^{(0)} + \lambda\psi_n^{(1)} + \lambda^2\psi_n^{(2)} + \dots \quad (4)$$

$$E_n = E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \dots \quad (5)$$

The terms $\psi_n^{(1)}$ and $E_n^{(1)}$ are called the first order corrections to the wavefunction and energy respectively, the $\psi_n^{(2)}$ and $E_n^{(2)}$ are the second order corrections and so on. The task of perturbation theory is to approximate the energies and wavefunctions of the perturbed system by calculating corrections up to a given order.

Note 2.1 *In perturbation theory we are assuming that all perturbed quantities are functions of the parameter λ , i.e. $\hat{H}(\lambda)$, $E_n(\lambda)$ and $\psi_n(\mathbf{r}; \lambda)$ and that when $\lambda \rightarrow 0$ we have $\hat{H}(0) = \hat{H}^{(0)}$, $E_n(0) = E_n^{(0)}$ and $\psi_n(\mathbf{r}; 0) = \psi_n^{(0)}(\mathbf{r})$. You will remember from your maths course that the Taylor series expansion of say $E_n(\lambda)$ around $\lambda = 0$ is*

$$E_n = E_n(0) + \left. \frac{dE_n}{d\lambda} \right|_{\lambda=0} \lambda + \frac{1}{2!} \left. \frac{d^2 E_n}{d\lambda^2} \right|_{\lambda=0} \lambda^2 + \frac{1}{3!} \left. \frac{d^3 E_n}{d\lambda^3} \right|_{\lambda=0} \lambda^3 + \dots \quad (6)$$

By comparing this expression with (5) we see that the perturbation theory “corrections” to the energy level E_n are related to the terms of Taylor series expansion by: $E_n^{(0)} = E_n(0)$, $E_n^{(1)} = \left. \frac{dE_n}{d\lambda} \right|_{\lambda=0}$, $E_n^{(2)} = \frac{1}{2!} \left. \frac{d^2 E_n}{d\lambda^2} \right|_{\lambda=0}$, $E_n^{(3)} = \frac{1}{3!} \left. \frac{d^3 E_n}{d\lambda^3} \right|_{\lambda=0}$, etc. Similar relations hold for the expressions (3) and (4) for the Hamiltonian and wavefunction respectively.

Note 2.2 *In many textbooks the expansion of the Hamiltonian is terminated after the first order term, i.e. $\hat{H} = \hat{H}^{(0)} + \lambda \hat{H}^{(1)}$ as this is sufficient for many physical problems.*

Note 2.3 *What is the significance of the parameter λ ?*

In some cases λ is a physical quantity: For example, if we have a single electron placed in a uniform electric field along the z -axis the total perturbed Hamiltonian is just $\hat{H} = \hat{H}^{(0)} + \mathcal{E}_z(e\hat{z})$ where $\hat{H}^{(0)}$ is the Hamiltonian in the absence of the field. The effect of the field is described by the term $e\hat{z} \equiv \hat{H}^{(1)}$ and the strength of the field \mathcal{E}_z plays the role of the parameter λ .

In other cases λ is just a fictitious parameter which we introduce in order to solve a problem using the formalism of perturbation theory: For example, to describe the two electrons of a helium atom we may construct the zeroth order Hamiltonian as that of two non-interacting electrons 1 and 2, $\hat{H}^{(0)} = -1/2\nabla_1^2 - 1/2\nabla_2^2 - 2/\mathbf{r}_1 - 2/\mathbf{r}_2$ which is trivial to solve as it is the sum of two single-particle Hamiltonians, one for each electron. The entire Hamiltonian for this system however is $\hat{H} = \hat{H}^{(0)} + 1/|\mathbf{r}_1 - \mathbf{r}_2|$ which is no longer separable, so we may use perturbation theory to find an approximate solution for $\hat{H}(\lambda) = \hat{H}^{(0)} + \lambda/|\mathbf{r}_1 - \mathbf{r}_2| = \hat{H}^{(0)} + \lambda \hat{H}^{(1)}$ using the fictitious parameter λ as a “dial” which is varied continuously from 0 to its final value 1 and takes us from the model problem to the real problem.

To calculate the perturbation corrections we substitute the series expansions of equations (3), (4) and (5) into the TISE (1) for the perturbed system, and rearrange and

group terms according to powers of λ in order to get

$$\begin{aligned}
& \{\hat{H}^{(0)}\psi_n^{(0)} - E_n^{(0)}\psi_n^{(0)}\} \\
+ & \lambda\{\hat{H}^{(0)}\psi_n^{(1)} + \hat{H}^{(1)}\psi_n^{(0)} - E_n^{(0)}\psi_n^{(1)} - E_n^{(1)}\psi_n^{(0)}\} \\
+ & \lambda^2\{\hat{H}^{(0)}\psi_n^{(2)} + \hat{H}^{(1)}\psi_n^{(1)} + \hat{H}^{(2)}\psi_n^{(0)} - E_n^{(0)}\psi_n^{(2)} - E_n^{(1)}\psi_n^{(1)} - E_n^{(2)}\psi_n^{(0)}\} \\
+ & \dots = 0
\end{aligned} \tag{7}$$

Notice how in each bracket terms of the same order are grouped (for example $\hat{H}^{(1)}\psi_n^{(1)}$ is a second order term because the sum of the orders of $\hat{H}^{(1)}$ and $\psi_n^{(1)}$ is 2). The powers of λ are linearly independent functions, so the only way that the above equation can be satisfied for all (arbitrary) values of λ is if the coefficient of each power of λ is zero. By setting each such term to zero we obtain the following sets of equations

$$\hat{H}^{(0)}\psi_n^{(0)} = E_n^{(0)}\psi_n^{(0)} \tag{8}$$

$$(\hat{H}^{(0)} - E_n^{(0)})\psi_n^{(1)} = (E_n^{(1)} - \hat{H}^{(1)})\psi_n^{(0)} \tag{9}$$

$$(\hat{H}^{(0)} - E_n^{(0)})\psi_n^{(2)} = (E_n^{(2)} - \hat{H}^{(2)})\psi_n^{(0)} + (E_n^{(1)} - \hat{H}^{(1)})\psi_n^{(1)} \tag{10}$$

...

To simplify the expressions from now on we will use bra-ket notation, representing wavefunction corrections by their state number, so $\psi_n^{(0)} \equiv |n^{(0)}\rangle$, $\psi_n^{(1)} \equiv |n^{(1)}\rangle$, etc.

2.1.1 The first order correction to the energy

To derive an expression for calculating the first order correction to the energy $E^{(1)}$, take equation (9) in ket notation

$$(\hat{H}^{(0)} - E_n^{(0)})|n^{(1)}\rangle = (E_n^{(1)} - \hat{H}^{(1)})|n^{(0)}\rangle \tag{11}$$

and multiply from the left by $\langle n^{(0)}|$ to obtain

$$\langle n^{(0)}|(\hat{H}^{(0)} - E_n^{(0)})|n^{(1)}\rangle = \langle n^{(0)}|(E_n^{(1)} - \hat{H}^{(1)})|n^{(0)}\rangle \tag{12}$$

$$\langle n^{(0)}|\hat{H}^{(0)}|n^{(1)}\rangle - E_n^{(0)}\langle n^{(0)}|n^{(1)}\rangle = E_n^{(1)}\langle n^{(0)}|n^{(0)}\rangle - \langle n^{(0)}|\hat{H}^{(1)}|n^{(0)}\rangle \tag{13}$$

$$E_n^{(0)}\langle n^{(0)}|n^{(1)}\rangle - E_n^{(0)}\langle n^{(0)}|n^{(1)}\rangle = E_n^{(1)} - \langle n^{(0)}|\hat{H}^{(1)}|n^{(0)}\rangle \tag{14}$$

$$0 = E_n^{(1)} - \langle n^{(0)}|\hat{H}^{(1)}|n^{(0)}\rangle \tag{15}$$

where in order to go from (13) to (14) we have used the fact that the eigenfunctions of the unperturbed Hamiltonian $\hat{H}^{(0)}$ are normalised and the Hermiticity property of $\hat{H}^{(0)}$ which allows it to operate to its eigenket on its left

$$\langle n^{(0)}|\hat{H}^{(0)}|n^{(1)}\rangle = \langle(\hat{H}^{(0)}n^{(0)})|n^{(1)}\rangle = \langle(E_n^{(0)}n^{(0)})|n^{(1)}\rangle = E_n^{(0)}\langle n^{(0)}|n^{(1)}\rangle \tag{16}$$

So, according to our result (15), the first order correction to the energy is

$$E_n^{(1)} = \langle n^{(0)} | \hat{H}^{(1)} | n^{(0)} \rangle \quad (17)$$

which is simply the expectation value of the first order Hamiltonian in the state $|n^{(0)}\rangle \equiv \psi_n^{(0)}$ of the unperturbed system.

Example 1 Calculate the first order correction to the energy of the n th state of a harmonic oscillator whose centre of potential has been displaced from 0 to a distance l .

The Hamiltonian of the unperturbed system harmonic oscillator is

$$\hat{H}^{(0)} = -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + \frac{1}{2}k\hat{x}^2 \quad (18)$$

while the Hamiltonian of the perturbed system is

$$\hat{H} = -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + \frac{1}{2}k(\hat{x} - l)^2 \quad (19)$$

$$= -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + \frac{1}{2}k\hat{x}^2 - lk\hat{x} + l^2 \frac{1}{2}k \quad (20)$$

$$= \hat{H}^{(0)} + l\hat{H}^{(1)} + l^2\hat{H}^{(2)} \quad (21)$$

where we have defined $\hat{H}^{(1)} \equiv -k\hat{x}$ and $\hat{H}^{(2)} \equiv \frac{1}{2}k$ and l plays the role of the perturbation parameter λ . According to equation 17,

$$E_n^{(1)} = \langle n^{(0)} | \hat{H}^{(1)} | n^{(0)} \rangle = -k \langle n^{(0)} | \hat{x} | n^{(0)} \rangle. \quad (22)$$

From the theory of the harmonic oscillator (see earlier lectures in this course) we know that the diagonal matrix elements of the position operator within any state $|n^{(0)}\rangle$ of the harmonic oscillator are zero ($\langle n^{(0)} | \hat{x} | n^{(0)} \rangle = 0$) from which we conclude that the first order correction to the energy in this example is zero.

2.1.2 The first order correction to the wavefunction

We will now derive an expression for the calculation of the first order correction to the wavefunction. Multiply (9) from the left by $\langle k^{(0)} |$, where $k \neq n$, to obtain

$$\langle k^{(0)} | \hat{H}^{(0)} - E_n^{(0)} | n^{(1)} \rangle = \langle k^{(0)} | E_n^{(1)} - \hat{H}^{(1)} | n^{(0)} \rangle \quad (23)$$

$$(E_k^{(0)} - E_n^{(0)}) \langle k^{(0)} | n^{(1)} \rangle = -\langle k^{(0)} | \hat{H}^{(1)} | n^{(0)} \rangle \quad (24)$$

$$\langle k^{(0)} | n^{(1)} \rangle = \frac{\langle k^{(0)} | \hat{H}^{(1)} | n^{(0)} \rangle}{E_n^{(0)} - E_k^{(0)}} \quad (25)$$

where in going from (23) to (24) we have made use of the orthogonality of the zeroth order wavefunctions ($\langle k^{(0)} | n^{(0)} \rangle = 0$). Also, in (25) we are allowed to divide with $E_n^{(0)} - E_k^{(0)}$

because we have assumed non-degeneracy of the zeroth-order problem (i.e. $E_n^{(0)} - E_k^{(0)} \neq 0$).

To proceed in our derivation for an expression for $|n^{(1)}\rangle$ we will employ the identity operator expressed in the eigenfunctions of the unperturbed system (zeroth order eigenfunctions):

$$|n^{(1)}\rangle = \hat{1}|n^{(1)}\rangle = \sum_k |k^{(0)}\rangle \langle k^{(0)}|n^{(1)}\rangle \quad (26)$$

Before substituting (25) into the above equation we must resolve a conflict: k must be different from n in (25) but not necessarily so in (26). This restriction implies that the first order correction to $|n\rangle$ will contain no contribution from $|n^{(0)}\rangle$. To impose this restriction we require that $\langle n^{(0)}|n\rangle = 1$ (this leads to $\langle n^{(0)}|n^{(j)}\rangle = 0$ for $j \geq 1$. Prove it!) instead of $\langle n|n\rangle = 1$. This choice of normalisation for $|n\rangle$ is called *intermediate normalisation* and of course it does not affect any physical property calculated with $|n\rangle$ since observables are independent of the normalisation of wavefunctions. So now we can substitute (25) into (26) and get

$$|n^{(1)}\rangle = \sum_{k \neq n} |k^{(0)}\rangle \frac{\langle k^{(0)}|\hat{H}^{(1)}|n^{(0)}\rangle}{E_n^{(0)} - E_k^{(0)}} = \sum_{k \neq n} |k^{(0)}\rangle \frac{H_{kn}^{(1)}}{E_n^{(0)} - E_k^{(0)}} \quad (27)$$

where the matrix element $H_{kn}^{(1)}$ is defined by the above equation.

2.1.3 The second order correction to the energy

To derive an expression for the second order correction to the energy multiply (10) from the left with $\langle n^{(0)}|$ to obtain

$$\begin{aligned} \langle n^{(0)}|\hat{H}^{(0)} - E_n^{(0)}|n^{(2)}\rangle &= \langle n^{(0)}|E_n^{(2)} - \hat{H}^{(2)}|n^{(0)}\rangle + \langle n^{(0)}|E_n^{(1)} - \hat{H}^{(1)}|n^{(1)}\rangle \\ 0 &= E_n^{(2)} - \langle n^{(0)}|\hat{H}^{(2)}|n^{(0)}\rangle - \langle n^{(0)}|\hat{H}^{(1)}|n^{(1)}\rangle \end{aligned} \quad (28)$$

where we have used the fact that $\langle n^{(0)}|n^{(1)}\rangle = 0$ (section 2.1.2). We now solve (28) for $E_n^{(2)}$

$$E_n^{(2)} = \langle n^{(0)}|\hat{H}^{(2)}|n^{(0)}\rangle + \langle n^{(0)}|\hat{H}^{(1)}|n^{(1)}\rangle = H_{nn}^{(2)} + \langle n^{(0)}|\hat{H}^{(1)}|n^{(1)}\rangle \quad (29)$$

which upon substitution of $|n^{(1)}\rangle$ by the expression (27) becomes

$$E_n^{(2)} = H_{nn}^{(2)} + \sum_{k \neq n} \frac{H_{nk}^{(1)} H_{kn}^{(1)}}{E_n^{(0)} - E_k^{(0)}}. \quad (30)$$

Example 2 Let us apply what we have learned so far to the “toy” model of a system which has only two (non-degenerate) levels (states) $|1^{(0)}\rangle$ and $|2^{(0)}\rangle$. Let $E_1^{(0)} < E_2^{(0)}$ and assume that there is only a first order term in the perturbed Hamiltonian and that the

diagonal matrix elements of the perturbation are zero, i.e. $\langle m^{(0)} | \hat{H}^{(1)} | m^{(0)} \rangle = H_{mm}^{(1)} = 0$. For this simple system we can solve exactly for its perturbed energies up to infinite order (see Atkins):

$$E_1 = \frac{1}{2}(E_1^{(0)} + E_2^{(0)}) - \frac{1}{2}[(E_1^{(0)} - E_2^{(0)})^2 + 4|H_{12}^{(1)}|^2]^{\frac{1}{2}} \quad (31)$$

$$E_2 = \frac{1}{2}(E_1^{(0)} + E_2^{(0)}) + \frac{1}{2}[(E_1^{(0)} - E_2^{(0)})^2 + 4|H_{12}^{(1)}|^2]^{\frac{1}{2}} \quad (32)$$

According to equation 30 the total perturbed energies up to second order are

$$E_1 \simeq E_1^{(0)} - \frac{|H_{12}^{(1)}|^2}{E_2^{(0)} - E_1^{(0)}} \quad (33)$$

$$E_2 \simeq E_2^{(0)} + \frac{|H_{12}^{(1)}|^2}{E_2^{(0)} - E_1^{(0)}} \quad (34)$$

These sets of equations show that the effect of the perturbation is to lower the energy of the lower level and raise the energy of the upper level. The effect increases with the strength of the perturbation (size of $|H_{12}^{(1)}|^2$ term) and decreasing separation between the unperturbed energies ($E_2^{(0)} - E_1^{(0)}$ term).

2.1.4 The closure approximation

We will now derive a very crude approximation to the second order correction to the energy. This approximation is computationally much simpler than the full second order expression and although it is not very accurate it can often be used to obtain qualitative insights. We begin by approximating the denominator of (30) by some kind of “average” energy difference $\Delta E \simeq E_k^{(0)} - E_n^{(0)}$ which is independent of the summation index k , and thus can be taken out of the summation. Using ΔE , (30) becomes:

$$E_n^{(2)} \simeq H_{nn}^{(2)} - \sum_{k \neq n} \frac{H_{nk}^{(1)} H_{kn}^{(1)}}{\Delta E} = H_{nn}^{(2)} - \frac{1}{\Delta E} \sum_{k \neq n} H_{nk}^{(1)} H_{kn}^{(1)} \quad (35)$$

We can now see that the above expression could be simplified significantly if the sum over k could be made to include n as this would allow us to eliminate it by using the completeness (or closure) property ($\hat{1} = \sum_k |k^{(0)}\rangle \langle k^{(0)}|$) of the zeroth order wavefunctions. We achieve just that by adding and subtracting the $H_{nn}^{(1)} H_{nn}^{(1)} / \Delta E$ term:

$$E_n^{(2)} \simeq H_{nn}^{(2)} - \frac{1}{\Delta E} \sum_k H_{nk}^{(1)} H_{kn}^{(1)} + \frac{1}{\Delta E} H_{nn}^{(1)} H_{nn}^{(1)} \quad (36)$$

$$\simeq H_{nn}^{(2)} - \frac{1}{\Delta E} \langle n^{(0)} | \hat{H}^{(1)} \hat{H}^{(1)} | n^{(0)} \rangle + \frac{1}{\Delta E} H_{nn}^{(1)} H_{nn}^{(1)} \quad (37)$$

This approximation can only be accurate if $n = 0$ (the ground state) and all excited states are much higher in energy from the ground state than their maximum energy separation. This assumption is usually not valid. Nevertheless, from a mathematical viewpoint, it is always possible to find a value for ΔE that makes the closure approximation exact. To find this value we just need to equate (37) to the righthand side of (30) and solve for ΔE to obtain

$$\Delta E = \frac{H_{nn}^{(1)} H_{nn}^{(1)} - \langle n^{(0)} | \hat{H}^{(1)} \hat{H}^{(1)} | n^{(0)} \rangle}{\sum_{k \neq n} \frac{H_{nk}^{(1)} H_{kn}^{(1)}}{E_n^{(0)} - E_k^{(0)}}} . \quad (38)$$

This expression is of course of limited practical interest as the computational complexity it involves is actually higher than the exact second order formula (30).

2.2 Perturbation theory for degenerate states

The perturbation theory we have developed so far applies to non-degenerate states. For systems with degenerate states we need to modify our approach. Let us assume that our zeroth order Hamiltonian has d states with energy $E_n^{(0)}$. We will represent these zeroth order states as

$$\psi_{n,i}^{(0)} = |(n, i)^{(0)}\rangle, \quad i = 1, \dots, d \quad (39)$$

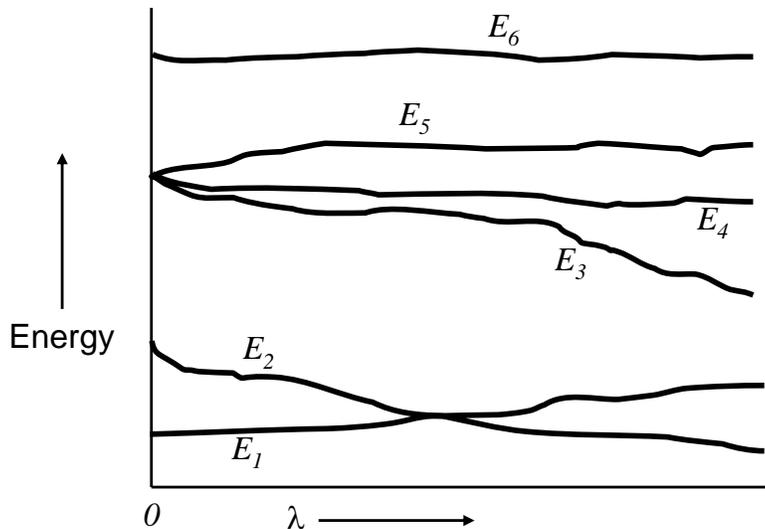


Figure 1: Effect of a perturbation on energy levels. In this example the perturbation removes all the degeneracies of the unperturbed levels.

where now we use two indices to represent each state: the first index n runs over the different energy eigenvalues while the second index i runs over the d degenerate states for a particular energy eigenvalue. Since we have d degenerate states of energy $E_n^{(0)}$, any linear combination of these states is also a valid state of energy $E_n^{(0)}$. However, as the perturbation parameter λ is varied continuously from 0 to some finite value, it is likely that the degeneracy of the states will be lifted (either completely or partially). The question that arises then is whether the states $\psi_{n,i}^{(0)}$ of equation (39) are the “correct” ones, i.e. whether they can be continuously transformed to the (in general) non-degenerate perturbed states. It turns out that this is usually not the case and one has to first find the “correct” zeroth order states

$$\phi_{n,j}^{(0)} = \sum_{i=1}^d |(n,i)^{(0)}\rangle c_{ij} \quad j = 1, \dots, d \quad (40)$$

where the coefficients c_{ij} that mix the $\psi_{n,i}^{(0)}$ are specific to the perturbation $\hat{H}^{(1)}$ and are determined by its symmetry.

Here we will find a way to determine the “correct” zeroth order states $\phi_{n,j}^{(0)}$ and the first order correction to the energy. To do this we start from equation 9 with $\phi_{n,i}^{(0)}$ in place of $\psi_{n,i}^{(0)}$

$$(\hat{H}^{(0)} - E_n^{(0)})\psi_{n,i}^{(1)} = (E_{n,i}^{(1)} - \hat{H}^{(1)})\phi_{n,i}^{(0)} \quad (41)$$

Notice that we include in the notation for the first order energy $E_{n,i}^{(1)}$ the index i since the

perturbation may split the degenerate energy level $E_n^{(0)}$. Figure 1 shows an example for a hypothetical system with six states and a three-fold degenerate unperturbed level. Note that the perturbation splits the degenerate energy level. In some cases the perturbation may have no effect on the degeneracy or may only partly remove the degeneracy.

The next step involves multiplication from the left by $\langle(n, j)^{(0)}|$

$$\langle(n, j)^{(0)}|\hat{H}^{(0)} - E_n^{(0)}|(n, i)^{(1)}\rangle = \langle(n, j)^{(0)}|E_{n,i}^{(1)} - \hat{H}^{(1)}|\phi_{n,i}^{(0)}\rangle \quad (42)$$

$$0 = \langle(n, j)^{(0)}|E_{n,i}^{(1)} - \hat{H}^{(1)}|\phi_{n,i}^{(0)}\rangle \quad (43)$$

$$0 = \sum_k \langle(n, j)^{(0)}|E_{n,i}^{(1)} - \hat{H}^{(1)}|(n, k)^{(0)}\rangle c_{ki} \quad (44)$$

where we have made use of the Hermiticity of $\hat{H}^{(0)}$ to set the left side to zero and we have substituted the expansion (40) for $\phi_{n,i}^{(0)}$. Some further manipulation of (44) gives:

$$\sum_k [\langle(n, j)^{(0)}|\hat{H}^{(1)}|(n, k)^{(0)}\rangle - E_{n,i}^{(1)}\langle(n, j)^{(0)}|(n, k)^{(0)}\rangle] c_{ki} = 0 \quad (45)$$

$$\sum_k (H_{jk}^{(1)} - E_{n,i}^{(1)} S_{jk}) c_{ki} = 0 \quad (46)$$

We thus arrive to equation 46 which describes a system of d simultaneous linear equations for the d unknowns c_{ki} , ($k = 1, \dots, d$) for the “correct” zeroth order state $\phi_{n,i}^{(0)}$. Actually, this is a *homogeneous* system of linear equations as all constant coefficients (i.e. the righthand side here) are zero. The trivial solution is obviously $c_{ki} = 0$ but we reject it because it has no physical meaning. As you know from your maths course, in order to obtain a non-trivial solution for such a system we must demand that the determinant of the matrix of the coefficients is zero:

$$|H_{jk}^{(1)} - E_{n,i}^{(1)} S_{jk}| = 0 \quad (47)$$

We now observe that as $E_{n,i}$ occurs in every row, this determinant is actually a d th degree polynomial in $E_{n,i}$ and the solution of the above equation for its d roots will give us *all* the $E_{n,i}^{(1)}$ ($i = 1, \dots, d$) first order corrections to the energies of the d degenerate levels with energy $E_n^{(0)}$. We can then substitute each $E_{n,i}^{(1)}$ value into (46) to find the corresponding non-trivial solution of c_{ki} ($k = 1, \dots, d$) coefficients, or in other words the function $\phi_{n,i}^{(0)}$. Finally, you should be able to verify that $E_{n,i}^{(1)} = \langle\phi_{n,i}^{(0)}|\hat{H}^{(1)}|\phi_{n,i}^{(0)}\rangle$, i.e. that the expression (17) we have derived which gives the first order energy as the expectation value of the first order Hamiltonian in the zeroth order wavefunctions still holds, provided the “correct” degenerate zeroth order wavefunctions are used.

Example 3 *A typical example of degenerate perturbation theory is provided by the study of the $n = 2$ states of a hydrogen atom inside an electric field. In a hydrogen atom all*

four $n = 2$ states (one $2s$ orbital and three $2p$ orbitals) have the same energy. The lifting of this degeneracy when the atom is placed in an electric field is called the Stark effect and here we will study it using first order perturbation theory for degenerate systems.

Assuming that the electric field \mathcal{E} is applied along the z -direction, the form of the perturbation is

$$\lambda \hat{H}^{(1)} = e\mathcal{E}_z z \quad (48)$$

where the strength of the field \mathcal{E}_z plays the role of the parameter λ . Even though we have four states, based on parity and symmetry considerations we can show that only elements between the $2s$ and $2p_z$ orbitals will have non-zero off-diagonal $\hat{H}^{(1)}$ matrix elements and as a result the 4×4 system of equations (46) is reduced to the following 2×2 system (note that here all states are already orthogonal so the overlap matrix is equal to the unit matrix):

$$e\mathcal{E}_z \begin{pmatrix} \langle 2s|z|2s\rangle & \langle 2s|z|2p_z\rangle \\ \langle 2p_z|z|2s\rangle & \langle 2p_z|z|2p_z\rangle \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = E^{(1)} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \quad (49)$$

which after evaluating the matrix elements becomes

$$\begin{pmatrix} 0 & -3e\mathcal{E}_z a_0 \\ -3e\mathcal{E}_z a_0 & 0 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = E^{(1)} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} . \quad (50)$$

The solution of the above system results in the following first order energies and “correct” zeroth order wavefunctions

$$E^{(1)} = \pm 3e\mathcal{E}_z a_0 \quad (51)$$

$$\phi_{n,1}^{(0)} = \frac{1}{\sqrt{2}}(|2s\rangle - |2p_z\rangle) , \quad \phi_{n,2}^{(0)} = \frac{1}{\sqrt{2}}(|2s\rangle + |2p_z\rangle) \quad (52)$$

Therefore, the effect of the perturbation (to first order) on the energy levels can be summarised in the diagram of Figure 2.

Finally, we should mention that the energy levels of the hydrogen atom are also affected in the presence of a uniform magnetic field \mathbf{B} . This is called the Zeeman effect, and the form of the perturbation in that case is $\hat{H}^{(1)} = \frac{e}{2m}(\mathbf{L} + 2\mathbf{S}) \cdot \mathbf{B}$ where \mathbf{L} is the orbital angular momentum of the electron and \mathbf{S} is its spin angular momentum.

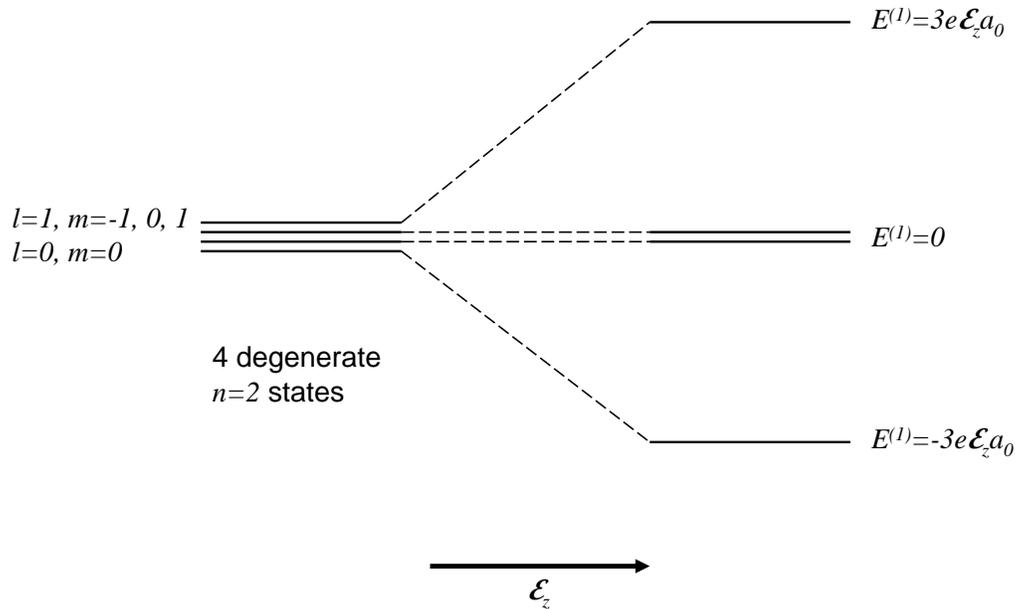


Figure 2: Pattern of Stark splitting of hydrogen atom in $n = 2$ state. The fourfold degeneracy is partly lifted by the perturbation. The $m = \pm 1$ states remain degenerate and are not shifted in the Stark effect.

3 Time-dependent perturbation theory

3.1 Revision: The time-dependent Schrödinger equation with a time-independent Hamiltonian

We want to find the (time-dependent) solutions of the time-dependent Schrödinger equation (TDSE)

$$i\hbar \frac{\partial \Psi^{(0)}}{\partial t} = \hat{H}^{(0)} \Psi^{(0)} \quad (53)$$

where we assume that $\hat{H}^{(0)}$ *does not* depend on time. Even though in this section we are not involved with perturbation theory, we will still follow the notation $\hat{H}^{(0)}$, $\Psi^{(0)}$ of representing the exactly soluble problem as the zeroth order problem as this will prove useful in the derivation of time-dependent perturbation theory that follows in the next section. According to the mathematics you have learned, the solution to the above equation can be written as a product

$$\Psi^{(0)}(\mathbf{r}, t) = \psi^{(0)}(\mathbf{r})T^{(0)}(t) \quad (54)$$

where the $\psi^{(0)}(\mathbf{r})$, which depends only on position coordinates \mathbf{r} , is the solution of the energy eigenvalue equation (TISE)

$$\hat{H}^{(0)}\psi_n^{(0)}(\mathbf{r}) = E_n^{(0)}\psi_n^{(0)}(\mathbf{r}) \quad (55)$$

and the expression for $T(t)$ is derived by substituting the right hand side of the above to the time-dependent equation 53. Finally we obtain

$$\Psi^{(0)}(\mathbf{r}, t) = \psi_n^{(0)}(\mathbf{r})e^{-iE_n^{(0)}t/\hbar} . \quad (56)$$

Now let us consider the following linear combination of $\psi_n^{(0)}$

$$\Psi^{(0)}(\mathbf{r}, t) = \sum_k a_k \psi_k^{(0)}(\mathbf{r})e^{-iE_k^{(0)}t/\hbar} \quad (57)$$

where the a_k are constants. This is also a solution of the TDSE (prove it!) because the TDSE consists of linear operators. This more general “superposition of states” solution of course contains (56) (by setting $a_k = \delta_{nk}$) but unlike (56) it is not, in general, an eigenfunction of $\hat{H}^{(0)}$. Assuming that the $\psi_n^{(0)}$ have been chosen to be orthonormal, which is always possible, we find that the expectation value of the Hamiltonian is

$$\langle \Psi^{(0)} | \hat{H}^{(0)} | \Psi^{(0)} \rangle = \sum_k |a_k|^2 E_k^{(0)} \quad (58)$$

We see thus that in the case of equation 56 the system is in a state with definite energy $E_n^{(0)}$ while in the general case (57) the system can be in any of the states with an average energy given by (58) where the probability $P_k = |a_k|^2$ of being in the state k is equal to the square modulus of the coefficient a_k . Both (56) and (57) are time-dependent because of the “phase factors” $e^{-iE_k^{(0)}t/\hbar}$ but the probabilities P_k and also the expectation values for operators that do not contain time (such as the $\hat{H}^{(0)}$ above) are time-independent.

3.2 Time-independent Hamiltonian with a time-dependent perturbation

We will now develop a perturbation theory for the case where the zeroth order Hamiltonian is time-independent but the perturbation terms are time-dependent. Thus our perturbed Hamiltonian has the following form

$$\hat{H}(t) = \hat{H}^{(0)} + \lambda \hat{H}^{(1)}(t) + \lambda^2 \hat{H}^{(2)}(t) + \dots \quad (59)$$

To simplify our discussion, in what follows we will only consider up to first order perturbations in the Hamiltonian

$$\hat{H}(t) = \hat{H}^{(0)} + \lambda \hat{H}^{(1)}(t) . \quad (60)$$

We will use perturbation theory to approximate the solution $\Psi(\mathbf{r}, t)$ to the time-dependent Schrödinger equation of the perturbed system.

$$i\hbar \frac{\partial \Psi}{\partial t} = \hat{H}(t)\Psi \quad (61)$$

At any instant t , we can expand the $\Psi(\mathbf{r}, t)$, in the complete set of eigenfunctions $\psi_k^{(0)}(\mathbf{r})$ of the zeroth order Hamiltonian $\hat{H}^{(0)}$,

$$\Psi(\mathbf{r}, t) = \sum_k b_k(t) \psi_k^{(0)}(\mathbf{r}) \quad (62)$$

but of course the expansion coefficients $b_k(t)$ vary with time as $\Psi(\mathbf{r}, t)$ does. In fact let us define $b_k(t) = a_k(t)e^{-iE_k^{(0)}t/\hbar}$ in the above equation to get

$$\Psi(\mathbf{r}, t) = \sum_k a_k(t) \psi_k^{(0)}(\mathbf{r}) e^{-iE_k^{(0)}t/\hbar} \quad (63)$$

Even though this expression looks more messy than (62), we prefer it because it will simplify the derivation that follows and also it directly demonstrates that when the $a_k(t)$ lose their time dependence, i.e. when $\lambda \rightarrow 0$ and $a_k(t) \rightarrow a_k$, (63) reduces to (57).

We substitute the expansion (63) into the time-dependent Schrödinger equation 53 and after taking into account the fact that the $\psi_n^{(0)} = |n^{(0)}\rangle$ are eigenfunctions of $\hat{H}^{(0)}$ we obtain

$$\sum_n a_n(t) \lambda \hat{H}^{(1)}(t) |n^{(0)}\rangle e^{-iE_n^{(0)}t/\hbar} = i\hbar \sum_n \frac{da_n(t)}{dt} |n^{(0)}\rangle e^{-iE_n^{(0)}t/\hbar} \quad (64)$$

The next step is to multiply with $\langle k^{(0)}|$ from the left and use the orthogonality of the zeroth order functions to get

$$\sum_n a_n(t) \lambda \langle k^{(0)} | \hat{H}^{(1)}(t) | n^{(0)} \rangle e^{-iE_n^{(0)}t/\hbar} = i\hbar \frac{da_k(t)}{dt} e^{-iE_k^{(0)}t/\hbar} \quad (65)$$

Solving this for $da_k(t)/dt$ results in the following differential equation

$$\frac{da_k(t)}{dt} = \frac{\lambda}{i\hbar} \sum_n a_n(t) H_{kn}^{(1)}(t) e^{i(E_k^{(0)} - E_n^{(0)})t/\hbar} = \frac{\lambda}{i\hbar} \sum_n a_n(t) H_{kn}^{(1)}(t) e^{i\omega_{kn}t} \quad (66)$$

where we have defined $\omega_{kn} = (E_k^{(0)} - E_n^{(0)})/\hbar$ and $H_{kn}^{(1)}(t) = \langle k^{(0)} | \hat{H}^{(1)}(t) | n^{(0)} \rangle$. We now integrate the above differential equation from 0 to t to obtain

$$a_k(t) - a_k(0) = \frac{\lambda}{i\hbar} \sum_n \int_0^t a_n(t') H_{kn}^{(1)}(t') e^{i\omega_{kn}t'} dt' \quad (67)$$

The purpose now of the perturbation theory we will develop is to determine the time-dependent coefficients $a_k(t)$. We begin by writing a perturbation expansion for the coefficient $a_k(t)$ in terms of the parameter λ

$$a_k(t) = a_k^{(0)}(t) + \lambda a_k^{(1)}(t) + \lambda^2 a_k^{(2)}(t) + \dots \quad (68)$$

where you should keep in mind that while λ and t are not related in any way, we take $t = 0$ as the “beginning of time” for which we know exactly the composition of the system so that

$$a_k(0) = a_k^{(0)}(0) \quad (69)$$

which means that $a_k^{(l)}(0) = 0$ for $l > 0$. Furthermore we will assume that

$$a_g^{(0)}(0) = \delta_{gj} \quad (70)$$

which means that at $t = 0$ the system is exclusively in a particular state $|j^{(0)}\rangle$ and all other states $|g^{(0)}\rangle$ with $g \neq j$ are unoccupied. Now substitute expansion (68) into (67) and collect equal powers of λ to obtain the following expressions

$$a_k^{(0)}(t) - a_k^{(0)}(0) = 0 \quad (71)$$

$$a_k^{(1)}(t) - a_k^{(1)}(0) = \frac{1}{i\hbar} \sum_n \int_0^t a_n^{(0)}(t') H_{kn}^{(1)}(t') e^{i\omega_{kn}t'} dt' \quad (72)$$

$$a_k^{(2)}(t) - a_k^{(2)}(0) = \frac{1}{i\hbar} \sum_n \int_0^t a_n^{(1)}(t') H_{kn}^{(1)}(t') e^{i\omega_{kn}t'} dt' \quad (73)$$

$$\dots \quad (74)$$

We can observe that these equations are *recursive*: each of them provides an expression for $a_f^{(m)}(t)$ in terms of $a_f^{(m-1)}(t)$. Let us now obtain an explicit expression for $a_f^{(1)}(t)$ by first substituting (71) into (72), and then making use of (70):

$$a_f^{(1)}(t) = \frac{1}{i\hbar} \sum_n \int_0^t a_n^{(0)}(0) H_{fn}^{(1)}(t') e^{i\omega_{fn}t'} dt' = \frac{1}{i\hbar} \int_0^t H_{fj}^{(1)}(t') e^{i\omega_{fj}t'} dt' \quad (75)$$

The probability that the system is in state $|f^{(0)}\rangle$ is obtained in a similar manner to equation 58 and is given by the squared modulus of the $a_f(t)$ coefficient

$$P_f(t) = |a_f(t)|^2 \quad (76)$$

but of course a significant difference from (58) is that $P_f = P_f(t)$ now changes with time. Using the perturbation expansion (68) for $a_f(t)$ we have

$$P_f(t) = |a_f^{(0)}(t) + \lambda a_f^{(1)}(t) + \lambda^2 a_f^{(2)}(t) + \dots|^2 \quad (77)$$

Note that in most of the examples that we will study in these lectures we will confine ourselves to the first order approximation which means that we will also approximate the above expression for $P_f(t)$ by neglecting from it the second and higher order terms.

Note 3.1 The previous derivation of time-dependent perturbation theory is rather rigorous and is also very much in line with the approach we used to derive time-independent perturbation theory. However, if we are only interested in obtaining only up to first order corrections, we can follow a less strict but more physically motivated approach (see also Atkins).

We begin with (67) and set λ equal to 1 to obtain

$$a_k(t) - a_k(0) = \frac{1}{i\hbar} \sum_n \int_0^t a_n(t') H_{kn}^{(1)}(t') e^{i\omega_{kn}t'} dt' \quad (78)$$

This equation is exact but it is not useful in practice because the unknown coefficient $a_k(t)$ is given in terms of all other unknown coefficients $a_n(t)$ including itself! To proceed we make the following approximations:

1. Assume that at $t = 0$ the system is entirely in an initial state j , so $a_j(0) = 1$ and $a_n(0) = 0$ if $n \neq j$.
2. Assume that the time t for which the perturbation is applied is so small that the change in the values of the coefficients is negligible, or in other words that $a_j(t) \simeq 1$ and $a_n(t) \simeq 0$ if $n \neq j$.

Using these assumptions we can reduce the sum on the righthand side of equation 78 to a single term (the one with $n = j$ for which $a_j(t) \simeq 1$). We will also rename the lefthand side index from k to f to denote some "final" state with $f \neq j$ to obtain

$$a_f(t) = \frac{1}{i\hbar} \int_0^t H_{fj}^{(1)}(t') e^{i\omega_{fj}t'} dt' \quad (79)$$

This approximate expression for the coefficients $a_f(t)$ is correct to first order as we can see by comparing it with equation 75.

Example 4 Show that with a time-dependent Hamiltonian $\hat{H}(t)$ the energy is not conserved.

We obviously need to use the time-dependent Schrödinger equation

$$i\hbar \frac{\partial \Psi}{\partial t} = \hat{H}(t)\Psi \Leftrightarrow \frac{\partial \Psi}{\partial t} = -\frac{i}{\hbar} \hat{H}(t)\Psi \quad (80)$$

where the system is described by a time-dependent state Ψ . We now look for an expression for the derivative of the energy $\langle H \rangle = \langle \Psi | \hat{H}(t) | \Psi \rangle$ (expectation value of the Hamiltonian) with respect to time. We have

$$\frac{\partial \langle H \rangle}{\partial t} = \left\langle \frac{\partial \Psi}{\partial t} \middle| \hat{H}(t) \middle| \Psi \right\rangle + \left\langle \Psi \middle| \frac{\partial \hat{H}(t)}{\partial t} \middle| \Psi \right\rangle + \left\langle \Psi \middle| \hat{H}(t) \middle| \frac{\partial \Psi}{\partial t} \right\rangle \quad (81)$$

Now using equation 80 to eliminate the $\frac{\partial \Psi}{\partial t}$ terms we obtain

$$\frac{\partial \langle H \rangle}{\partial t} = \langle \Psi | \frac{\partial \hat{H}(t)}{\partial t} | \Psi \rangle \neq 0 \quad (82)$$

which shows (in contrast with the case of a time-independent Hamiltonian!) that the time derivative of the energy is no longer zero and therefore the energy is no longer a constant of the motion. So, in the time-dependent perturbation theory we develop here it is pointless to look for corrections to the energy levels. Nevertheless, we will continue to denote the energy levels of the unperturbed system as zeroth order, $E_n^{(0)}$, for consistency with our previously derived formulas of time-independent perturbation theory.

3.3 Two level time-dependent system - Rabi oscillations

Let us look at the simplified example of a quantum system with only two stationary states (levels), $\psi_1^{(0)}$ and $\psi_2^{(0)}$ with energies $E_1^{(0)}$ and $E_2^{(0)}$ respectively. Since we only have two levels, equation 66 becomes

$$\frac{da_1(t)}{dt} = \frac{\lambda}{i\hbar} \left[a_1(t)H_{11}^{(1)}(t) + a_2(t)H_{12}^{(1)}(t)e^{i\omega_{12}t} \right] \quad (83)$$

for $\frac{da_1(t)}{dt}$ and a similar equation holds for $\frac{da_2(t)}{dt}$. We will now impose two conditions:

- We assume that the diagonal elements of the time-dependent perturbation are zero, i.e. $H_{11}^{(1)}(t) = H_{22}^{(1)}(t) = 0$.
- We will only consider a particular type of perturbation where the off-diagonal element is equal to a constant $H_{12}^{(1)}(t) = \hbar V$ for t in the interval $[0, T]$ and equal to zero for all other times. Of course we must also have $H_{21}^{(1)}(t) = \hbar V^*$ since $\hat{H}^{(1)}(t)$ must be a Hermitian operator.

Under these conditions we obtain the following system of two differential equations for the two coefficients $a_1(t)$ and $a_2(t)$

$$\frac{da_1(t)}{dt} = \frac{1}{i\hbar} a_2(t)H_{12}^{(1)}(t)e^{-i\omega_{12}t} \quad \text{and} \quad \frac{1}{i\hbar} a_2(t)H_{21}^{(1)}(t)e^{-i\omega_{21}t} \quad (84)$$

We can now solve this system of differential equations by substitution, using the initial condition that at $t = 0$ the system is definitely in state 1, or in other words $a_1(0) = 1$ and $a_2(0) = 0$. The solution obtained under these conditions is

$$a_1(t) = \left[\cos \Omega t + \frac{i\omega_{21}}{2\Omega} \sin \Omega t \right] e^{-i\omega_{21}t/2} \quad , \quad a_2(t) = -\frac{i|V|}{\Omega} \sin \Omega t e^{i\omega_{21}t/2} \quad (85)$$

where

$$\Omega = \frac{1}{2}(\omega_{21}^2 + 4|V|^2)^{\frac{1}{2}} \quad . \quad (86)$$

Note 3.2 *In this section we are not really applying perturbation theory: The two level system allows us to obtain the exact solutions for the coefficients $a_1(t)$ and $a_2(t)$ (up to infinite order in the language of perturbation theory).*

The probability of the system being in the state $\psi_2^{(0)}$ is

$$P_2(t) = |a_2(t)|^2 = \left(\frac{4|V|^2}{\omega_{21}^2 + 4|V|^2} \right) \sin^2 \frac{1}{2}(\omega_{21}^2 + 4|V|^2)^{1/2} t \quad (87)$$

and of course since we only have two states here we will also have $P_1(t) = 1 - P_2(t)$.

Let us examine these probabilities in some detail. First consider the case where the two states are degenerate ($\omega_{21} = 0$). We then have

$$P_1(t) = \cos^2 |V|t \quad , \quad P_2(t) = \sin^2 |V|t \quad (88)$$

which means that the system oscillates freely between the two states $|1^{(0)}\rangle$ and $|2^{(0)}\rangle$ and the only role of the perturbation is to determine the frequency $|V|$ of the oscillation. The other extreme is the case where the levels are widely separated in comparison with the strength of the perturbation in the sense that $\omega_{21}^2 \gg |V|^2$. In this case we obtain

$$P_2(t) \simeq \left(\frac{2|V|}{\omega_{21}} \right)^2 \sin^2 \frac{1}{2}\omega_{21}t \quad (89)$$

which shows that the probability of the system occupying state $|2^{(0)}\rangle$ can not get any larger than $(2|V|/\omega_{21})^2$ which is a number much smaller than 1. Thus the system remains almost exclusively in state $|1^{(0)}\rangle$. We should also observe here that the frequency of oscillation is independent of the strength of the perturbation and is determined only by the separation of the states ω_{21} .

3.4 Perturbation varying “slowly” with time

Here we will study the example of a very slow time-dependent perturbation in order to see how time-dependent theory reduces to the time-independent theory in the limit of very slow change. We define the perturbation as follows

$$\hat{\mathcal{H}}^{(1)}(t) = \begin{cases} 0, & t < 0 \\ \hat{H}^{(1)}(1 - e^{-kt}), & t \geq 0 \end{cases} . \quad (90)$$

where $\hat{H}^{(1)}$ is a time-independent operator, which however may not be a constant as for example it may depend on \hat{x} , and so on. The entire perturbation $\hat{\mathcal{H}}^{(1)}(t)$ is time-dependent as $\hat{H}^{(1)}$ is multiplied by the term $(1 - e^{-kt})$ which varies from 0 to 1 as t increases from 0 to infinity. Substituting the perturbation into equation (75) we obtain

$$a_f^{(1)}(t) = \frac{1}{i\hbar} H_{fj}^{(1)} \int_0^t (1 - e^{-kt'}) e^{i\omega_{fj}t'} dt' = \frac{1}{i\hbar} H_{fj}^{(1)} \left[\frac{e^{i\omega_{fj}t} - 1}{i\omega_{fj}} + \frac{e^{-(k-i\omega_{fj})t} - 1}{k - i\omega_{fj}} \right] \quad (91)$$

If we assume that we will only examine times very long after the perturbation has reached its final value, or in other words $kt \gg 1$, we obtain

$$a_f^{(1)}(t) = \frac{1}{i\hbar} H_{fj}^{(1)} \left[\frac{e^{i\omega_{fj}t} - 1}{i\omega_{fj}} + \frac{-1}{k - i\omega_{fj}} \right] \quad (92)$$

and finally that the rate in which the perturbation is switched is slow in the sense that $k^2 \ll \omega_{fj}^2$, we are left with

$$a_f^{(1)}(t) = -\frac{H_{fj}^{(1)}}{\hbar\omega_{fj}} e^{i\omega_{fj}t} \quad (93)$$

The square of this, which is the probability of being in state $|f^{(0)}\rangle$ to first order is

$$P_f(t) = |a_f^{(1)}(t)|^2 = \frac{|H_{fj}^{(1)}|^2}{\hbar^2\omega_{fj}^2} . \quad (94)$$

We observe that the resulting expression for $P_f(t)$ is no longer time-dependent. In fact, it is equal to the square modulus $|\langle f^{(0)}|j^{(1)}\rangle|^2$ of the expansion coefficient in $|f^{(0)}\rangle$ of the first order state $|j^{(1)}\rangle$ as given in equation 25 of time-independent perturbation theory. Thus in the framework of time-independent theory (94) is interpreted as being the fraction of the state $|f^{(0)}\rangle$ in the expansion of $|j^{(1)}\rangle$ while in the time-dependent theory it represents the probability of the system being in state $|f^{(0)}\rangle$ at a given time.

3.5 Perturbation oscillating with time

3.5.1 Transition to a single level

We will examine here a harmonic time-dependent potential, oscillating in time with angular frequency $\omega = 2\pi\nu$. The form of such a perturbation is

$$\hat{H}^{(1)}(t) = 2V \cos \omega t = V(e^{i\omega t} + e^{-i\omega t}) \quad (95)$$

where V does not depend on time (but of course it could be a function of coordinates, e.g. $V = V(x)$). This in a sense is the most general type of time-dependent perturbation as any other time-dependent perturbation can be expanded as a sum (Fourier series) of harmonic terms like those of (95). Inserting this expression for the perturbation $\hat{H}^{(1)}(t)$ into equation 75 we obtain

$$a_f^{(1)}(t) = \frac{1}{i\hbar} V_{fj} \int_0^t (e^{i\omega t'} + e^{-i\omega t'}) e^{i\omega_{fj} t'} dt' = \frac{1}{i\hbar} V_{fj} \left[\frac{e^{i(\omega_{fj} + \omega)t} - 1}{i(\omega_{fj} + \omega)} + \frac{e^{i(\omega_{fj} - \omega)t} - 1}{i(\omega_{fj} - \omega)} \right] \quad (96)$$

where $V_{fj} = \langle f^{(0)} | V | j^{(0)} \rangle$. If we assume that $\omega_{fj} - \omega \simeq 0$, or in other words that $E_f^{(0)} \simeq E_j^{(0)} + \hbar\omega$, only the second term in the above expression survives. We then have

$$a_f^{(1)}(t) = \frac{i}{\hbar} V_{fj} \frac{1 - e^{i(\omega_{fj} - \omega)t}}{(\omega_{fj} - \omega)} \quad (97)$$

from which we obtain

$$P_f(t) = |a_f^{(1)}(t)|^2 = \frac{4|V_{fj}|^2}{\hbar^2(\omega_{fj} - \omega)^2} \sin^2 \frac{1}{2}(\omega_{fj} - \omega)t \quad (98)$$

This equation shows that due to the time-dependent perturbation, the system can make transitions from the state $|j^{(0)}\rangle$ to the state $|f^{(0)}\rangle$ by absorbing a quantum of energy $\hbar\omega$. Now in the case where $\omega_{fj} = \omega$ exactly, the above expression reduces to

$$\lim_{\omega \rightarrow \omega_{fj}} P_f(t) = \frac{|V_{fj}|^2}{\hbar^2} t^2 \quad (99)$$

which shows that the probability increases quadratically with time. We see that this expression allows the probability to increase without bounds and even exceed the (maximum) value of 1. This is of course not correct, so this expression should be considered valid only when $P_f(t) \ll 1$, according to the assumption behind first order perturbation theory through which it was obtained.

Our discussion so far for the harmonic perturbation has been based on the assumption that $E_f^{(0)} > E_j^{(0)}$ so that the external oscillating field causes *stimulated absorption* of energy in the form of quanta of energy $\hbar\omega$. However, the original equation 96 for

$a_f^{(1)}(t)$ also allows us to have $E_f^{(0)} < E_j^{(0)}$. In this case we can have $E_f^{(0)} \simeq E_j^{(0)} - \hbar\omega$ and then the first term in equation 96 dominates from which we can derive an expression analogous to (98):

$$P_f(t) = |a_f^{(1)}(t)|^2 = \frac{4|V_{fj}|^2}{\hbar^2(\omega_{fj} + \omega)^2} \sin^2 \frac{1}{2}(\omega_{fj} + \omega)t \quad (100)$$

This now describes *stimulated emission* of quanta of frequency $\omega/2\pi$ that is caused by the time-dependent perturbation and causes transitions from the higher energy state $E_j^{(0)}$ to the lower energy state $E_f^{(0)}$. One can regard the time-dependent perturbation here as an inexhaustible source or sink of energy.

3.5.2 Transition to a continuum of levels

In many situations instead of a single final state $|f^{(0)}\rangle$ of energy $E_f^{(0)}$ there is usually a group of closely-spaced final states with energy close to $E_f^{(0)}$. In that case we should calculate the probability of being in any of those final states which is equal to the sum of the probabilities for each state, so we have

$$P(t) = \sum_{n, E_n^{(0)} \simeq E_f^{(0)}} |a_n^{(1)}(t)|^2 \quad (101)$$

As the final states form a continuum, it is customary to count the number of states $dN(E)$ with energy in the interval $(E, E + dE)$ in terms of the density of states $\rho(E)$ at energy E as

$$dN(E) = \rho(E) dE \quad (102)$$

Using this formalism, we can change the sum of equation 101 into an integral

$$P(t) = \int_{E_f^{(0)} - \Delta E}^{E_f^{(0)} + \Delta E} \rho(E) |a_E^{(1)}(t)|^2 dE \quad (103)$$

where the summation index n has been substituted by the continuous variable E . According to our assumption $E \simeq E_f^{(0)}$ so the above expression after substitution of (98) becomes

$$P(t) = \int_{E_f^{(0)} - \Delta E}^{E_f^{(0)} + \Delta E} 4 \frac{|V_{fj}|^2}{\hbar^2} \frac{\sin^2 \frac{1}{2}(E/\hbar - E_j^{(0)}/\hbar - \omega)t}{(E/\hbar - E_j^{(0)}/\hbar - \omega)^2} \rho(E) dE \quad (104)$$

where the integral is evaluated in a narrow region of energies around $E_f^{(0)}$. The integrand above contains a term that, as t grows larger it becomes sharply peaked at $E = E_j^{(0)} + \hbar\omega$ and sharply decaying to zero away from this value (see Figure 3). This then allows us to approximate it by treating $|V_{fj}|$ as a constant and also the density of states as a

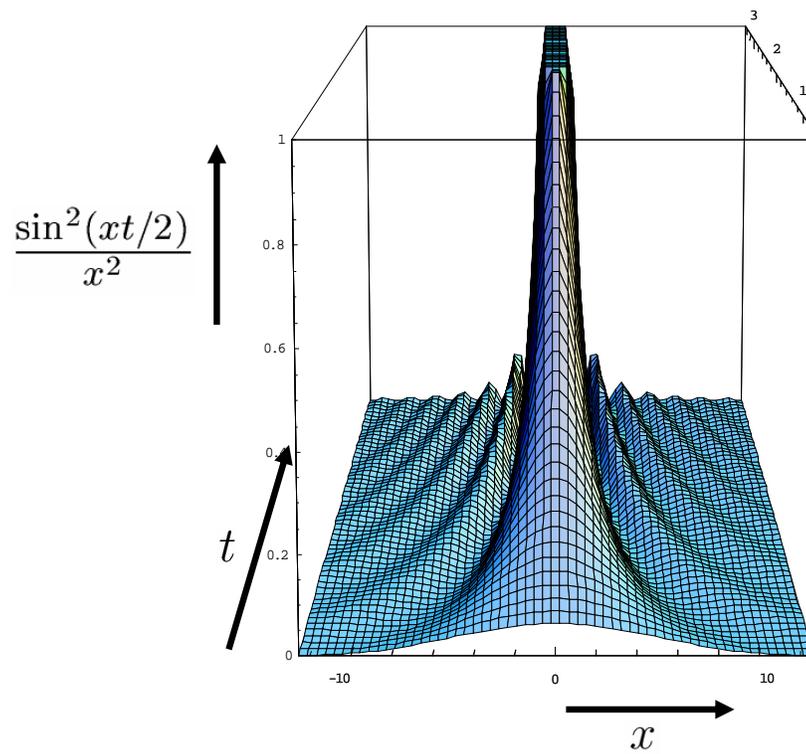


Figure 3: A plot of $\frac{\sin^2(xt/2)}{x^2}$ as a function of x and t . Notice that as t increases the function turns into a sharp peak centred at $x = 0$.

constant in terms of its value $\rho(E_f^{(0)})$ at $E_f^{(0)}$. These constants can then be taken out of the integral. What remains inside the integral is the trigonometric function. We now extend the range of integration from $[E_f^{(0)} - \Delta E, E_f^{(0)} + \Delta E]$ to $(-\infty, \infty)$ as this allows us to evaluate it but it barely affects its value due to the peaked shape of the trigonometric function. Evaluation of the integral then results in the following expression

$$P(t) = \frac{2\pi}{\hbar} t |V_{fj}|^2 \rho(E_f^{(0)}) . \quad (105)$$

Its derivative with respect to time is the *transition rate* which is the rate at which the initially empty levels become populated.

$$W(t) = \frac{dP}{dt} = \frac{2\pi}{\hbar} |V_{fj}|^2 \rho(E_f^{(0)}) \quad (106)$$

This succinct expression, which is independent of time, is sometimes called *Fermi's golden rule*.

3.6 Emission and absorption of radiation by atoms

We will now use the theory for a perturbation oscillating with time to study the interaction of an atom with an electromagnetic wave. The electromagnetic wave is approximated by an electric field ¹ oscillating in time ²

$$\mathcal{E}(t) = 2\mathcal{E}_z \mathbf{n}_z \cos \omega t \quad (107)$$

where \mathbf{n}_z is a unit vector along the direction of the wave, which for convenience here we have chosen it to lie along the direction of the z axis. The factor of 2 is again included for computational convenience as in the previous section. The interaction of the atom with the radiation field is given by the electric dipole interaction

$$\hat{H}^{(1)}(t) = -\hat{\boldsymbol{\mu}} \cdot \mathcal{E}(t) = -2\mu_z \mathcal{E}_z \cos \omega t . \quad (108)$$

The $\hat{\boldsymbol{\mu}}$ is the dipole moment operator for the atom

$$\hat{\boldsymbol{\mu}} = -e \sum_{k=1}^Z \mathbf{r}_k \quad (109)$$

¹We will neglect the magnetic interaction of the radiation with atoms as it is usually small compared to the interaction with the electric field

²Actually the electric field oscillates both in space and in time and has the following form

$$\mathcal{E}(t) = 2\mathcal{E}_z \mathbf{n}_z \cos(\mathbf{k} \cdot \mathbf{r} - \omega t)$$

where the wavelength of the radiation is $\lambda = 2\pi/|\mathbf{k}|$ and its angular frequency is $\omega = c|\mathbf{k}|$. However, here we work under the assumption that λ is very large compared to the size of the atom and thus we neglect the spatial variation of the field. This approach is called the electric dipole approximation.

where the sum over k runs over all the electrons, and the position vector of the k th electron is \mathbf{r}_k . The nucleus of the atom is assumed to be fixed at the origin of coordinates ($\mathbf{r} = 0$).

We can immediately see that the work of section 3.5 for a perturbation oscillating with time according to a harmonic time-dependent potential applies here if we set $V = \mu_z \mathcal{E}_z$ in equation 95 and all the expressions derived from it. In particular, equation 98 for the probability of absorption or radiation for transition from state $|j^0\rangle$ to the higher energy state $|f^{(0)}\rangle$ takes the form

$$P_{fj}(t) = \frac{4|\mu_{z,fj}|^2 \mathcal{E}_z^2(\omega)}{\hbar^2(\omega_{fj} - \omega)^2} \sin^2 \frac{1}{2}(\omega_{fj} - \omega)t . \quad (110)$$

You will notice in the above expression that we have written \mathcal{E}_z as $\mathcal{E}_z(\omega)$ in order to remind ourselves that it *does depend* on the angular frequency ω of the radiation. In fact the above expression is valid only for monochromatic radiation. Most radiation sources produce a continuum of frequencies, so in order to take this fact into account we need to integrate the above expression over all angular frequencies

$$P_{fj}(t) = \int_{-\infty}^{\infty} \frac{4|\mu_{z,fj}|^2 \mathcal{E}_z^2(\omega)}{\hbar^2(\omega_{fj} - \omega)^2} \sin^2 \frac{1}{2}(\omega_{fj} - \omega)t d\omega \quad (111)$$

$$= \frac{4|\mu_{fj}|^2 \mathcal{E}_z^2(\omega_{fj})}{\hbar^2} \int_{-\infty}^{\infty} \frac{\sin^2[\frac{1}{2}(\omega - \omega_{fj})t]}{(\omega - \omega_{fj})^2} d\omega \quad (112)$$

$$= \frac{2\pi t |\mu_{fj}|^2 \mathcal{E}_z^2(\omega_{fj})}{\hbar^2} \quad (113)$$

where we have evaluated the above integral using the same technique we used for the derivation of Fermi's golden rule in the previous section. The rate of absorption of radiation is equal to the time derivative of the above expression and as we are interested in atoms in the gas phase, we average the above expression over all directions in space. It turns out that this is equivalent to replacing $|\mu_{z,fj}|^2$ by the mean value of x , y and z components, $\frac{1}{3}|\mu_{fj}|^2$, which leads to

$$W_{f \leftarrow j}(t) = \frac{2\pi |\mu_{fj}|^2 \mathcal{E}_z^2(\omega_{fj})}{3\hbar^2} \quad (114)$$

A standard result from the classical theory of electromagnetism is that the energy density $\rho_{rad}(\omega_{fj})$ (i.e. energy contained per unit volume of space for radiation of angular frequency ω_{fj}) of the electromagnetic field is

$$\rho_{rad}(\omega_{fj}) = 2\varepsilon_0 \mathcal{E}_z^2(\omega_{fj}) \quad (115)$$

which upon substitution into (114) gives

$$W_{f \leftarrow j}(t) = \frac{2\pi |\mu_{fj}|^2}{6\varepsilon_0 \hbar^2} \rho_{rad}(\omega_{fj}) \quad (116)$$

We can also write this equation as

$$W_{f \leftarrow j} = B_{jf} \rho_{rad}(\omega_{fj}) \quad (117)$$

where the coefficient

$$B_{jf} = \frac{2\pi |\mu_{fj}|^2}{6\varepsilon_0^2 \hbar^2} \quad (118)$$

is the *Einstein coefficient of stimulated absorption*. As we know from the theory of section 3.5, it is also possible to write a similar equation for stimulated emission in which case the *Einstein coefficient of stimulated emission* B_{fj} will be equal to the B_{jf} as a result of the Hermiticity of the dipole moment operator. If the system of atoms and radiation is in thermal equilibrium, at a temperature T , the number of atoms N_f in state $|f^{(0)}\rangle$ and the the number of atoms N_j in state $|j^{(0)}\rangle$ should not change with time, which means that there should be no net transfer of energy between the atoms and the radiation field:

$$N_j W_{f \leftarrow j} = N_f W_{f \rightarrow j} \quad (119)$$

Given that $B_{fj} = B_{jf}$ this equation leads to the result $N_j = N_f$ which shows that the populations of the two states are equal. This can not be correct: we know from the generally applicable principles of statistical thermodynamics that the populations of the two states should obey the Boltzmann distribution

$$\frac{N_f}{N_j} = e^{-E_{fj}/kT} \quad (120)$$

To overcome this discrepancy, Einstein postulated that there must also be a process of *spontaneous emission* in which the upper state $|f^{(0)}\rangle$ decays to the lower state $|j^{(0)}\rangle$ independently of the presence of the radiation of frequency ω_{fj} . According to this the rate of emission should be written as

$$W_{f \rightarrow j} = A_{fj} + B_{fj} \rho_{rad}(\omega_{fj}) \quad (121)$$

where A_{fj} is the *Einstein coefficient of spontaneous emission* which does not need to be multiplied by $\rho_{rad}(\omega_{fj})$ as spontaneous emission is independent of the presence of the radiation ω_{fj} . The expression for this coefficient is (see Atkins for a derivation):

$$A_{fj} = \frac{\hbar \omega_{fj}^3}{\pi^2 c^3} B_{fj} \quad (122)$$

As we saw, spontaneous emission was postulated by Einstein as it is not predicted by combining a quantum mechanical description of the atoms with a classical description of the electric field. It is predicted though by the theory of quantum electrodynamics where the field is also quantized. The types of interaction of radiation with atoms that we have studied here are summarized in Figure 4.

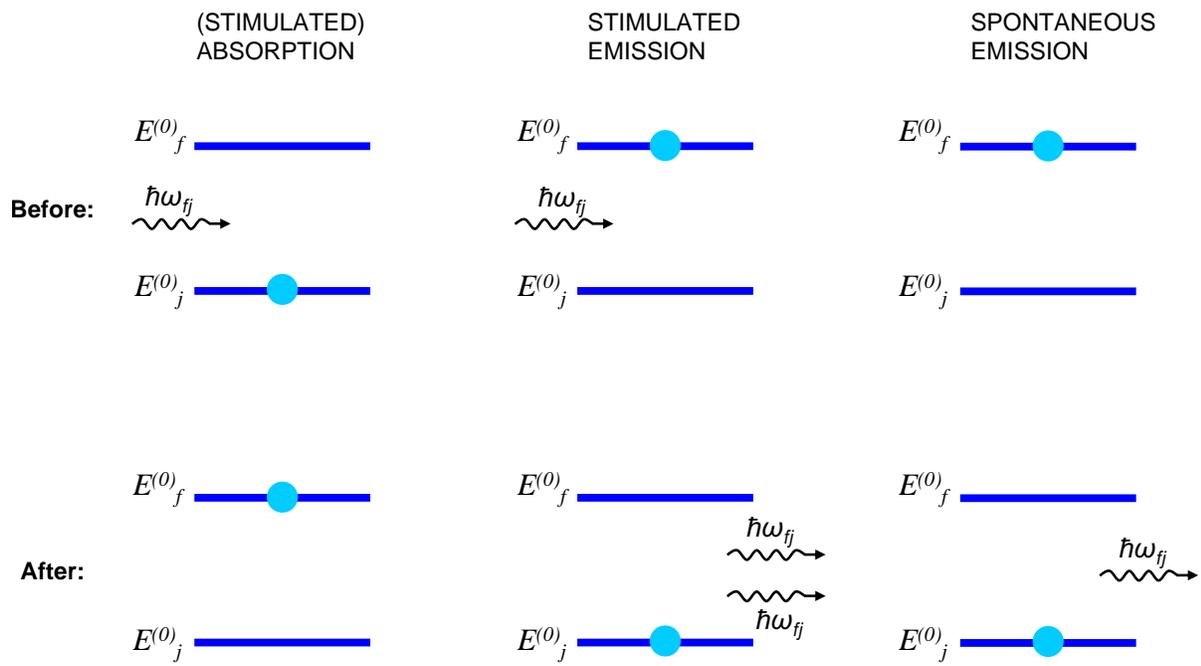


Figure 4: Schematic representation of stimulated absorption, stimulated emission and spontaneous emission.

We should note that the Einstein coefficients, while derived for thermal equilibrium, are completely general and hold also for non-equilibrium conditions. The operation of the laser (Light Amplification by Stimulated Emission of Radiation) is based on this principle. The idea behind this is to have some means of creating a non-equilibrium population of states (population inversion) where $N_f > N_j$. Then, from (117) and (121), and under the assumption of negligible spontaneous emission ($A_{fj} \ll B_{fj}\rho_{rad}(\omega_{fj})$) we will have

$$\frac{N_f W_{f \rightarrow j}}{N_j W_{f \leftarrow j}} = \frac{\text{rate of emission}}{\text{rate of absorption}} \simeq \frac{N_f}{N_j} > 1 \quad (123)$$

which shows that the applied frequency ω_{fj} will be amplified in intensity by the interaction process, resulting in more radiation emerging than entering the system. This process will reduce the population of the upper state until equilibrium is re-established, so the operation of a laser also depends on having a different process which maintains the population inversion of the states. As A_{fj} grows with the third power (see equation 122) of the angular frequency ω_{fj} we can expect that spontaneous emission will dominate at high frequencies leading to significant uncontrolled loss of energy and thus making population inversion difficult to maintain. A practical consequence of this is that X-ray lasers are difficult to make.

4 Applications of perturbation theory

In this section we will see how perturbation theory can be used to derive formulas for the calculation of properties of molecules.

4.1 Perturbation caused by uniform electric field

To study a molecule inside a uniform electric field \mathcal{E} we need to add the following term to the Hamiltonian

$$\lambda \hat{H}^{(1)} = -\hat{\mu} \cdot \mathcal{E} \quad (124)$$

which describes the interaction of the molecule with the electric field using the dipole moment operator $\hat{\mu}$ which is defined as

$$\hat{\mu} = \sum_i q_i \mathbf{r}_i \quad (125)$$

To simplify the notation in what follows we will always assume that the electric field is applied along the z-axis in which case the dot product of (124) becomes

$$\lambda \hat{H}^{(1)} = -\hat{\mu}_z \mathcal{E}_z = -\mathcal{E}_z \sum_i q_i \hat{z}_i \quad (126)$$

You will notice that we are already using the notation of perturbation theory as we are representing the above term as a first order Hamiltonian. The role of the perturbation parameter λ is played here by the z-component of the electric field \mathcal{E}_z .

4.2 Dipole moment in uniform electric field

So far we have been using perturbation theory to find corrections to the energy and wavefunctions. Here we will use perturbation theory to find corrections to the dipole moment as a function of the electric field \mathcal{E}_z , which plays the role of the perturbation parameter λ . We begin by applying the Hellmann-Feynman theorem to the energy with \mathcal{E}_z as the parameter. Since $\hat{H} = \hat{H}^{(0)} + \hat{H}^{(1)}(\mathcal{E}_z)$, or in other words only $\hat{H}^{(1)}$ depends on the parameter \mathcal{E}_z , we obtain

$$\frac{dE}{d\mathcal{E}_z} = \left\langle \frac{dH}{d\mathcal{E}_z} \right\rangle = \left\langle \frac{dH^{(1)}}{d\mathcal{E}_z} \right\rangle = \left\langle \frac{d(-\mu_z \mathcal{E}_z)}{d\mathcal{E}_z} \right\rangle = -\langle \mu_z \rangle \quad (127)$$

Let us now write the energy as a Taylor series expansion (perturbation expansion) with respect to the \mathcal{E}_z parameter at the point $\mathcal{E}_z = 0$:

$$E = E^{(0)} + \left(\frac{dE}{d\mathcal{E}_z} \right)_{\mathcal{E}_z=0} \mathcal{E}_z + \frac{1}{2!} \left(\frac{d^2E}{d\mathcal{E}_z^2} \right)_{\mathcal{E}_z=0} \mathcal{E}_z^2 + \frac{1}{3!} \left(\frac{d^3E}{d\mathcal{E}_z^3} \right)_{\mathcal{E}_z=0} \mathcal{E}_z^3 + \dots \quad (128)$$

where $\left(\frac{dE}{d\mathcal{E}_z}\right)_{\mathcal{E}_z=0}$ is the first derivative of the energy with respect to \mathcal{E}_z evaluated at $\mathcal{E}_z = 0$, etc. Of course, the zeroth order term $E^{(0)}$ is the value of the energy at $\mathcal{E}_z = 0$. If we now differentiate the above Taylor expansion with respect to \mathcal{E}_z , and substitute for the left hand side what we found in (127), we obtain an expression for the dipole moment in non-zero electric field

$$\langle\mu_z\rangle = -\left(\frac{dE}{d\mathcal{E}_z}\right)_{\mathcal{E}_z=0} - \left(\frac{d^2E}{d\mathcal{E}_z^2}\right)_{\mathcal{E}_z=0} \mathcal{E}_z - \frac{1}{2} \left(\frac{d^3E}{d\mathcal{E}_z^3}\right)_{\mathcal{E}_z=0} \mathcal{E}_z^2 + \dots \quad (129)$$

We usually write the above expression as

$$\langle\mu_z\rangle = \mu_{0z} + \alpha_{zz}\mathcal{E}_z + \frac{1}{2}\beta_{zzz}\mathcal{E}_z^2 + \dots \quad (130)$$

where by comparison with (129) we define following quantities as derivatives of the energy with respect to the electric field at zero electric field ($\mathcal{E}_z = 0$).

The *permanent dipole moment*

$$\mu_{0z} = -\left(\frac{dE}{d\mathcal{E}}\right)_{\mathcal{E}_z=0} = -\langle 0^{(0)}|\hat{\mu}_z|0^{(0)}\rangle \quad (131)$$

which is the first order energy correction to the ground state wavefunction.

The *polarizability*

$$\alpha_{zz} = -\left(\frac{d^2E}{d\mathcal{E}^2}\right)_{\mathcal{E}_z=0} \quad (132)$$

and the first *hyperpolarizability*

$$\beta_{zzz} = -\left(\frac{d^3E}{d\mathcal{E}^3}\right)_{\mathcal{E}_z=0} . \quad (133)$$

4.3 Calculation of the static polarizability

We can readily derive a formula for the calculation of the polarizability from the expression for the second order correction to the energy, equation 30. Here we apply it to the calculation of the polarizability of the ground state

$$\alpha_{zz} = -2E_0^{(2)} = -2 \sum_{n \neq 0} \frac{\langle 0^{(0)}|\hat{\mu}_z|n^{(0)}\rangle \langle n^{(0)}|\hat{\mu}_z|0^{(0)}\rangle}{E_0^{(0)} - E_n^{(0)}} . \quad (134)$$

The above is an explicit expression for the polarizability of a molecule in terms of integrals over its wavefunctions. We can write it in the following more compact form

$$\alpha_{zz} = 2 \sum_{n \neq 0} \frac{\mu_{z,0n}\mu_{z,n0}}{\Delta E_{n0}} \quad (135)$$

where we have defined the dipole moment matrix elements $\mu_{z,mn} = \langle m^{(0)} | \hat{\mu}_z | n^{(0)} \rangle$ and the denominator $\Delta E_{n0} = E_n^{(0)} - E_0^{(0)}$. This compact form can be used to express the *mean polarizability* which is the property that is actually observed when a molecule is rotating freely in the gas phase or in solution and one measures the average of all its orientations to the applied field:

$$\alpha = \frac{1}{3}(\alpha_{xx} + \alpha_{yy} + \alpha_{zz}) = \frac{2}{3} \sum_{n \neq 0} \frac{\mu_{x,0n} \mu_{x,n0} + \mu_{y,0n} \mu_{y,n0} + \mu_{z,0n} \mu_{z,n0}}{\Delta E_{n0}} \quad (136)$$

$$= \frac{2}{3} \sum_{n \neq 0} \frac{\boldsymbol{\mu}_{0n} \cdot \boldsymbol{\mu}_{n0}}{\Delta E_{n0}} \quad (137)$$

$$= \frac{2}{3} \sum_{n \neq 0} \frac{|\mu_{0n}|^2}{\Delta E_{n0}} \quad (138)$$

At this point we can also use the closure approximation (37) to eliminate the sum over states and derive a computationally much simpler but also more approximate expression for the polarizability.

$$\alpha \simeq \frac{2}{3\Delta E} \sum_{n \neq 0} \boldsymbol{\mu}_{0n} \cdot \boldsymbol{\mu}_{n0} = \frac{2}{3\Delta E} \left(\sum_n \boldsymbol{\mu}_{0n} \cdot \boldsymbol{\mu}_{n0} - \boldsymbol{\mu}_{00} \cdot \boldsymbol{\mu}_{00} \right) = \frac{2(\langle \mu^2 \rangle - \langle \mu \rangle^2)}{3\Delta E} \quad (139)$$

4.4 Polarizability and electronic molecular spectroscopy

As we saw in the previous section the polarizability depends on the square of *transition dipole moments* μ_{n0} between states $|n^{(0)}\rangle$ and $|0^{(0)}\rangle$. If we now re-write expression 138 as

$$\alpha = \frac{\hbar^2 e^2}{m_e} \sum_{n \neq 0} \frac{f_{n0}}{\Delta E_{n0}^2} \quad (140)$$

where we have used the *oscillator strengths* f_{n0} defined as

$$f_{n0} = \left(\frac{4\pi m_e}{3e^2 \hbar} \right) \nu_{n0} |\mu_{n0}|^2 \quad (141)$$

The oscillator strengths can be determined from the intensities of the electronic transitions of a molecule and the energies ΔE_{n0} from the frequencies where these transitions occur. From expression 140 we can observe that a molecule will have a large polarizability the higher the intensity and the lower the frequency of its electronic transitions. We can now further approximate (140) by replacing ΔE_{n0} by its average ΔE to obtain

$$\alpha \simeq \frac{\hbar^2 e^2}{m_e \Delta E^2} \sum_{n \neq 0} f_{n0} \quad (142)$$

This allows us to make use of the following standard result which is known as the *Kuhn-Thomas sum rule*

$$\sum_n f_{n0} = N_e \quad (143)$$

where N_e is the total number of electrons in the molecule. Notice that the sum rule involves a summation over all states, including $n = 0$, but this is compatible with (142) as $f_{00} = 0$ by definition. We therefore obtain

$$\alpha \simeq \frac{\hbar^2 e^2 N_e}{m_e \Delta E^2} \quad (144)$$

which again shows that the polarisability increases with increasing number of electrons and decreasing mean excitation energy. We therefore expect molecules composed of heavy atoms to be highly polarizable.

Example 5 *Prove the Kuhn-Thomas sum rule (143).*

Let us first prove the following relation in one dimension

$$\sum_n (E_n^{(0)} - E_a^{(0)}) |\langle n^{(0)} | \hat{x} | a^{(0)} \rangle|^2 = \frac{\hbar^2}{2m} . \quad (145)$$

Start with the following commutation relation,

$$[\hat{x}, \hat{H}^{(0)}] = \frac{i\hbar}{m} \hat{p}_x \quad (146)$$

that you can prove quite trivially if you take into account that the Hamiltonian is a sum of a kinetic energy operator and a potential energy operator. We next sandwich this commutator between $\langle n^{(0)} |$ and $| a^{(0)} \rangle$ to obtain

$$\langle n^{(0)} | \hat{x} \hat{H}^{(0)} | a^{(0)} \rangle - \langle n^{(0)} | \hat{H}^{(0)} \hat{x} | a^{(0)} \rangle = \frac{i\hbar}{m} \langle n^{(0)} | \hat{p}_x | a^{(0)} \rangle \quad (147)$$

$$(E_a^{(0)} - E_n^{(0)}) \langle n^{(0)} | \hat{x} | a^{(0)} \rangle = \frac{i\hbar}{m} \langle n^{(0)} | \hat{p}_x | a^{(0)} \rangle \quad (148)$$

$$\langle n^{(0)} | \hat{x} | a^{(0)} \rangle = \frac{i\hbar \langle n^{(0)} | \hat{p}_x | a^{(0)} \rangle}{m(E_a^{(0)} - E_n^{(0)})} \quad (149)$$

where we have made use of the Hermiticity of $\hat{H}^{(0)}$.

Now substitute this relation into the left hand side of (145) as follows

$$\begin{aligned}
& \sum_n (E_n^{(0)} - E_a^{(0)}) |\langle n^{(0)} | \hat{x} | a^{(0)} \rangle|^2 \\
&= \frac{1}{2} \sum_n (E_n^{(0)} - E_a^{(0)}) (\langle a^{(0)} | \hat{x} | n^{(0)} \rangle \langle n^{(0)} | \hat{x} | a^{(0)} \rangle + \langle a^{(0)} | \hat{x} | n^{(0)} \rangle \langle n^{(0)} | \hat{x} | a^{(0)} \rangle) \\
&= \frac{i\hbar}{2m} \sum_n (\langle a^{(0)} | \hat{p}_x | n^{(0)} \rangle \langle n^{(0)} | \hat{x} | a^{(0)} \rangle - \langle a^{(0)} | \hat{x} | n^{(0)} \rangle \langle n^{(0)} | \hat{p}_x | a^{(0)} \rangle) \\
&= \frac{i\hbar}{2m} \langle a^{(0)} | \hat{p}_x \hat{x} - \hat{x} \hat{p}_x | a^{(0)} \rangle \\
&= \frac{i\hbar}{2m} \langle a^{(0)} | [\hat{p}_x, \hat{x}] | a^{(0)} \rangle = \frac{i\hbar}{2m} \langle a^{(0)} | -i\hbar | a^{(0)} \rangle = \frac{\hbar^2}{2m}
\end{aligned}$$

where in the last line we have made use of the well known commutator between momentum and position, $[\hat{p}_x, \hat{x}] = -i\hbar$.

Having proved (145) it is straightforward to show that by rearranging it, multiplying with appropriate coefficients, generalising it to three dimensions and to N_e electrons results in the Kuhn-Thomas sum rule (143).

4.5 Dispersion forces

Here we will see how we can use perturbation theory to study the *dispersion force* (also called London or Van der Waals force) which is the weakest form of interaction between uncharged molecules. Dispersion forces are too weak to be considered as chemical bonds but they are nevertheless of great chemical relevance as for example these are the attractive forces which keep the carbon sheets of materials like graphite in place, cause attraction between DNA base pairs on adjacent planes, make possible the existence of liquid phases for noble gases, etc.

Dispersion forces are caused by the interaction between electric dipoles on different species. These dipoles are not permanent but are brought about by instantaneous fluctuations in the charge distribution of the species. Here we will use perturbation theory to calculate the interaction energy due to dispersion between two species A and B which are not charged and do not have permanent dipole moments. Our zeroth order Hamiltonian is the sum of the Hamiltonians for A and B

$$\hat{H}^{(0)} = \hat{H}_A^{(0)} + \hat{H}_B^{(0)}, \quad \hat{H}_A^{(0)} |n_A^{(0)}\rangle = E_{n_A}^{(0)} |n_A^{(0)}\rangle \quad \hat{H}_B^{(0)} |n_B^{(0)}\rangle = E_{n_B}^{(0)} |n_B^{(0)}\rangle \quad (150)$$

which of course means that its zeroth order energies are the sum of the energies for the isolated A and B species and its eigenfunctions are the products of the eigenfunctions of the A and B species

$$\hat{H}^{(0)} |n_A^{(0)} n_B^{(0)}\rangle = (E_{n_A}^{(0)} + E_{n_B}^{(0)}) |n_A^{(0)} n_B^{(0)}\rangle, \quad |n_A^{(0)} n_B^{(0)}\rangle = |n_A^{(0)}\rangle |n_B^{(0)}\rangle. \quad (151)$$

This Hamiltonian completely ignores all interactions between A and B. We will now add to it the following first order Hamiltonian

$$\hat{H}^{(1)} = \frac{1}{4\pi\epsilon_0 R^3} (\hat{\mu}_{Ax}\hat{\mu}_{Bx} + \hat{\mu}_{Ay}\hat{\mu}_{By} - 2\hat{\mu}_{Az}\hat{\mu}_{Bz}) \quad (152)$$

which (can be proved using classical electrostatics) describes the interaction between a dipole moment on A and a dipole moment on B, the two dipoles being a distance of R apart. Note that we have implicitly assumed here the Born-Oppenheimer approximation which means that we are only working with the electronic wavefunctions and the distance R is not a variable in our wavefunctions but it is just a parameter on which our calculations depend. Here we will study only the ground state.

As we have assumed that A and B have no permanent dipole moments it is easy to show that the first order correction to the energy $\langle 0_A^{(0)} 0_B^{(0)} | \hat{H}^{(1)} | 0_A^{(0)} 0_B^{(0)} \rangle$ is zero (show this!). We therefore turn our attention to the second order energy as defined by equation 30:

$$E^{(2)} = \sum_{n_A, n_B \neq (0_A, 0_B)} \frac{\langle 0_A^{(0)} 0_B^{(0)} | \hat{H}^{(1)} | n_A^{(0)} n_B^{(0)} \rangle \langle n_A^{(0)} n_B^{(0)} | \hat{H}^{(1)} | 0_A^{(0)} 0_B^{(0)} \rangle}{E_{0_A 0_B}^{(0)} - E_{n_A n_B}^{(0)}} \quad (153)$$

$$= - \sum_{n_A, n_B \neq (0_A, 0_B)} \frac{\langle 0_A^{(0)} 0_B^{(0)} | \hat{H}^{(1)} | n_A^{(0)} n_B^{(0)} \rangle \langle n_A^{(0)} n_B^{(0)} | \hat{H}^{(1)} | 0_A^{(0)} 0_B^{(0)} \rangle}{\Delta E_{n_A 0_A}^{(0)} + \Delta E_{n_B 0_B}^{(0)}} \quad (154)$$

where we have defined $\Delta E_{n_A 0_A}^{(0)} = E_{n_A}^{(0)} - E_{0_A}^{(0)}$ which is a positive quantity. We now substitute the expression for $\hat{H}^{(1)}$ which consists of 3 terms and therefore results in 9 terms. However out of the nine terms only the 3 diagonal terms are non-zero (see Atkins for a justification) and each of the non-zero terms has the following form:

$$\langle 0_A^{(0)} | \hat{\mu}_{Ax} | n_A^{(0)} \rangle \langle n_A^{(0)} | \hat{\mu}_{Ax} | 0_A^{(0)} \rangle \langle 0_B^{(0)} | \hat{\mu}_{Bx} | n_B^{(0)} \rangle \langle n_B^{(0)} | \hat{\mu}_{Bx} | 0_B^{(0)} \rangle \quad (155)$$

$$= \frac{1}{9} \langle 0_A^{(0)} | \hat{\boldsymbol{\mu}}_A | n_A^{(0)} \rangle \cdot \langle n_A^{(0)} | \hat{\boldsymbol{\mu}}_A | 0_A^{(0)} \rangle \langle 0_B^{(0)} | \hat{\boldsymbol{\mu}}_B | n_B^{(0)} \rangle \cdot \langle n_B^{(0)} | \hat{\boldsymbol{\mu}}_B | 0_B^{(0)} \rangle \quad (156)$$

Upon substitution of this expression into (154) we obtain the following expression

$$E^{(2)} = -\frac{2}{3} \left(\frac{1}{4\pi\epsilon_0 R^3} \right)^2 \sum_{n_A, n_B \neq (0_A, 0_B)} \frac{(\boldsymbol{\mu}_{A, 0_A n_A} \cdot \boldsymbol{\mu}_{A, n_A 0_A})(\boldsymbol{\mu}_{B, 0_B n_B} \cdot \boldsymbol{\mu}_{B, n_B 0_B})}{\Delta E_{n_A 0_A}^{(0)} + \Delta E_{n_B 0_B}^{(0)}} \quad (157)$$

from which we can deduce that the interaction is attractive ($E^{(2)} < 0$) and that the interaction energy is proportional to $1/R^6$.

We can do some further manipulations to obtain a more approximate yet physically meaningful expression for $E^{(2)}$. To proceed we apply the closure approximation by

replacing $\Delta_{n_A 0_A}^{(0)} + \Delta_{n_B 0_B}^{(0)}$ with the an average value $\Delta E_A + \Delta E_B$ and apply equation 37:

$$\begin{aligned} E^{(2)} &\simeq -\frac{2}{3} \left(\frac{1}{4\pi\epsilon_0 R^3} \right)^2 \left(\frac{1}{\Delta E_A + \Delta E_B} \right) \sum_{n_A, n_B \neq (0_A, 0_B)} (\boldsymbol{\mu}_{A, 0_A n_A} \cdot \boldsymbol{\mu}_{A, n_A 0_A}) (\boldsymbol{\mu}_{B, 0_B n_B} \cdot \boldsymbol{\mu}_{B, n_B 0_B}) \\ &\simeq - \left(\frac{1}{24\pi^2 \epsilon_0^2 R^6} \right) \left(\frac{1}{\Delta E_A + \Delta E_B} \right) \langle \mu_A^2 \rangle \langle \mu_B^2 \rangle \end{aligned}$$

where $\langle \mu_A^2 \rangle = \langle 0_A^{(0)} | \hat{\mu}_A^2 | 0_A^{(0)} \rangle$ and there is no $\langle \mu_A \rangle^2$ term since we assumed that the permanent dipole moments of A and B are zero. Having reached this stage, we can re-express the dispersion energy by using relation (139) between the mean square dipole moment and the polarizability (in the absence of a permanent dipole moment, $\langle \mu_A^2 \rangle \simeq \frac{3}{2} \alpha_A \Delta E_A$) to obtain

$$E^{(2)} \simeq - \left(\frac{3}{32\pi^2 \epsilon_0^2} \right) \left(\frac{\Delta E_A \Delta E_B}{\Delta E_A + \Delta E_B} \right) \frac{\alpha_A \alpha_B}{R^6} . \quad (158)$$

Finally we approximate the mean excitation energy with the ionization energy of each species $\Delta E_A \simeq I_A$ to arrive at the London formula for the dispersion energy between two non-polar species

$$E^{(2)} \simeq - \left(\frac{3}{32\pi^2 \epsilon_0^2} \right) \left(\frac{I_A I_B}{I_A + I_B} \right) \frac{\alpha_A \alpha_B}{R^6} . \quad (159)$$

This very approximate expression can provide chemical insight from “back of the envelope calculations” of the dispersion energy between atoms based on readily available quantities such as the polarizabilities and the ionization energies. Based on this formula we expect large, highly polarisable atoms to have strong dispersion interactions.

4.6 Revision: Antisymmetry, Slater determinants and the Hartree-Fock method

The Pauli exclusion principle follows from the postulate of (non-relativistic) quantum mechanics that a many-electron wavefunction must be *antisymmetric* with respect to interchange of the coordinates of any two electrons ³

$$\Phi(\mathbf{x}_1, \dots, \mathbf{x}_i, \dots, \mathbf{x}_j, \dots, \mathbf{x}_{N_e}) = -\Phi(\mathbf{x}_1, \dots, \mathbf{x}_j, \dots, \mathbf{x}_i, \dots, \mathbf{x}_{N_e}) \quad (160)$$

where $\mathbf{x}_j = \{\mathbf{r}_j, \sigma_j\}$ collectively denotes the space (\mathbf{r}_j) and spin (σ_j) coordinates of electron j .

We often choose to approximate the many-electron wavefunction as a product of single-electron wavefunctions (spinorbitals). Such a simple product of spin orbitals (also known as a Hartree product) is not antisymmetric. To overcome this limitation we define the wavefunction as a Slater determinant, which is antisymmetric as the interchange of any of its rows, which correspond to its electron coordinates, will change its sign.

In Hartree-Fock theory, we assume that the many-electron wavefunction has the form of a Slater determinant and we seek to find the best possible such wavefunction (for the ground state). To achieve this goal we use the variational principle which states that the total energy for the optimum determinant which we seek is going to be lower than the energy calculated from any other determinant

$$E_0^{HF} = \langle \Psi_0^{(0)} | \hat{H} | \Psi_0^{(0)} \rangle \leq \langle \Psi | \hat{H} | \Psi \rangle \quad (161)$$

where we have assumed that the Hartree-Fock solution $\Psi_0^{(0)}$ and all trial Slater determinants Ψ are normalized. E_0^{HF} is the Hartree-Fock energy for the ground state which we are seeking. The full Hamiltonian for the electrons in a material (e.g. a molecule or a portion of solid) has the following form

$$\hat{H} = \frac{\hbar^2}{2m_e} \sum_{i=1}^{N_e} \nabla_i^2 - \sum_{i=1}^{N_e} \sum_{I=1}^{N_N} \frac{Z_I e^2}{4\pi\epsilon_0 |\mathbf{r}_I - \mathbf{r}_i|} + \frac{1}{2} \sum_{i,j} \sum_{i \neq j} \frac{e^2}{4\pi\epsilon_0 |\mathbf{r}_i - \mathbf{r}_j|} \quad (162)$$

where we have assumed that the material consists of N_e electrons and N_N nuclei. The first term is the sum of the kinetic energy of each electron and the second term is the sum of the electrostatic attraction of each electron from the N_N nuclei, each of which is fixed (Born Oppenheimer approximation) at position \mathbf{r}_I . The final term is the repulsive electrostatic (Coulomb) interaction between the electrons and consists of a sum over all distinct pairs of electrons.

³More generally, the postulate states that a wavefunction must be antisymmetric with respect to interchange of any pair of identical fermions (=particles with half-integer spin quantum number such as electrons and protons) and symmetric with respect to interchange of any pair of identical bosons (=particles with integer spin quantum number, such as photons and α -particles)

The variational principle (161) results into single-electron Schrödinger equations of the form

$$\hat{f}_i \chi_i(\mathbf{x}) = \varepsilon_i \chi_i(\mathbf{x}) \quad (163)$$

for the spinorbitals χ_i that make up $\Psi_0^{(0)}$. However, the difficulty is that the *Fock operator* \hat{f}_i above is constructed from the (unknown!) solutions χ_i . In practice the way we solve these equations is by guessing a form for the χ_i , using it to build an approximate \hat{f}_i from which we solve the eigenvalue problem (163) to obtain a fresh (better) set of χ_i s. We then repeat this procedure until the χ_i s we obtain do not change any more - this condition is often referred to as *Self-Consistency*. In the literature Hartree-Fock (HF) calculations are also called Self-Consistent-Field (SCF) calculations.

4.7 Møller-Plesset many-body perturbation theory

In this section we will see how time-independent perturbation theory can be used as an improvement on the Hartree-Fock approximation. Let us rewrite the Hamiltonian (162) in the following form:

$$\hat{H} = \sum_{i=1}^{N_e} \left[\frac{\hbar^2}{2m_e} \nabla_i^2 - \sum_{I=1}^{N_N} \frac{Z_I e^2}{4\pi\epsilon_0 |\mathbf{r}_I - \mathbf{r}_i|} \right] + \frac{1}{2} \sum_{i,j}^{N_e} \frac{e^2}{4\pi\epsilon_0 |\mathbf{r}_i - \mathbf{r}_j|} \quad (164)$$

$$= \sum_{i=1}^{N_e} \hat{h}_i + \frac{1}{2} \sum_{i,j}^{N_e} \frac{e^2}{4\pi\epsilon_0 |\mathbf{r}_i - \mathbf{r}_j|} \quad (165)$$

which demonstrates the fact that the first two terms are “separable” into sums of one-electron Hamiltonians \hat{h}_i while this is obviously not possible for the last term as each $1/|\mathbf{r}_i - \mathbf{r}_j|$ can not be “broken” into a sum of a term for electron i and a term for electron j . The problem of the sum of one-electron Hamiltonians $\sum_{i=1}^{N_e} \hat{h}_i$ is computationally trivial as its solutions are antisymmetrised products (Slater determinants) of *one-electron wavefunctions* (=molecular spinorbitals). In contrast, because of the non-separability of the third term, such a simple solution is not possible for \hat{H} . Its solution is extremely complicated and computationally tractable only for very small systems (e.g. the hydrogen molecule). Thus this is a case where perturbation theory can be very useful for approximating the solution to \hat{H} .

As a first attempt to apply perturbation theory we may treat the $\sum_{i=1}^{N_e} \hat{h}_i$ part of (165) as the zeroth order Hamiltonian and the remaining part as the perturbation. This is not a very good choice though as the perturbation is of similar magnitude to the zeroth order Hamiltonian. Instead, we will define the zeroth order Hamiltonian as follows

$$\hat{H}^{(0)} = \sum_{i=1}^{N_e} \left(\hat{h}_i + \hat{v}_i^{HF} \right) = \sum_{i=1}^{N_e} \hat{f}_i \quad (166)$$

as a sum of Fock operators \hat{f}_i for each electron i . The *Hartree-Fock potential* for electron i is defined as

$$\hat{v}_i^{HF} = \sum_{a=1}^{N_e} \left(\hat{\mathcal{J}}_a(i) - \hat{\mathcal{K}}_a(i) \right) \quad (167)$$

where N_e is the total number of occupied spinorbitals and $\hat{\mathcal{J}}_a(i)$ is the *Coulomb* operator made of spinorbital χ_a acting on electron i , and the *Exchange* operator $\hat{\mathcal{K}}_a(i)$ is defined in a similar manner. With this choice of $\hat{H}^{(0)}$, the $\hat{H}^{(1)}$ is given by

$$\hat{H}^{(1)} = \hat{H} - \hat{H}^{(0)} = \frac{1}{2} \sum_{i,j}^{N_e} \frac{e^2}{4\pi\epsilon_0|\mathbf{r}_i - \mathbf{r}_j|} - \sum_{i=1}^{N_e} \hat{v}_i^{HF} \quad (168)$$

where now we can see, at least in a qualitative manner, that $\hat{H}^{(1)}$ defined in this way is much smaller than $\hat{H}^{(0)}$ and it is therefore plausible to treat it as a perturbation. Many-body perturbation theory using this choice of $\hat{H}^{(0)}$ is called Møller-Plesset (MP) perturbation theory.

Here we will develop MP theory for the ground state. First of all we observe that any Slater determinant made of N_e spinorbitals, each of which is a solution of the Hartree-Fock eigenvalue equation 163, is an eigenfunction of $\hat{H}^{(0)}$ according to (prove this!)

$$\hat{H}^{(0)}\Phi_0^{(0)} = \hat{H}^{(0)}|\chi_a(1)\chi_b(2)\dots\chi_z(N_e)| = (\varepsilon_a + \varepsilon_b + \dots + \varepsilon_z)|\chi_a(1)\chi_b(2)\dots\chi_z(N_e)| \quad (169)$$

where here we have used the ground state determinant $\Phi_0^{(0)} = |\chi_a(1)\chi_b(2)\dots\chi_z(N_e)|$ which consists of the N_e lowest energy spinorbitals and as we can see from this equation, the zeroth order energy for this state is the sum of the energies of these spinorbitals

$$E_0^{(0)} = \varepsilon_a + \varepsilon_b + \dots + \varepsilon_z \quad (170)$$

The first order energy is given by equation 17:

$$E_0^{(1)} = \langle \Phi_0^{(0)} | \hat{H}^{(1)} | \Psi_0^{(0)} \rangle \quad (171)$$

We can now observe that the sum $E_0^{(0)} + E_0^{(1)}$ is equal to the Hartree-Fock energy E_0^{HF} for the ground state

$$E_0^{(0)} + E_0^{(1)} = \langle \Psi_0^{(0)} | \hat{H}^{(0)} + \hat{H}^{(1)} | \Psi_0^{(0)} \rangle = \langle \Psi_0^{(0)} | \hat{H} | \Psi_0^{(0)} \rangle = E_0^{HF} \quad (172)$$

This means that we need to go beyond the first order energy correction to obtain an improvement to the Hartree-Fock energy, and more specifically to recover (at least some part of) the correlation energy. Here we will confine ourselves to second order MP theory

which is often referred to as “MP2”. According to (30) the second order correction to the energy is

$$E_0^{(2)} = \sum_{J \neq 0} \frac{\langle \Phi_0^{(0)} | \hat{H}^{(1)} | \Phi_J^{(0)} \rangle \langle \Phi_J^{(0)} | \hat{H}^{(1)} | \Phi_0^{(0)} \rangle}{E_0^{(0)} - E_J^{(0)}} \quad (173)$$

We need to evaluate the matrix elements $\langle \Phi_0^{(0)} | \hat{H}^{(1)} | \Phi_J^{(0)} \rangle$. Using the orthogonality property of different Slater determinants, we see that for $J \neq 0$

$$\langle \Phi_0^{(0)} | \hat{H}^{(0)} | \Phi_J^{(0)} \rangle = \langle \Phi_0^{(0)} | E_J^{(0)} | \Phi_J^{(0)} \rangle = E_J^{(0)} \langle \Phi_0^{(0)} | \Phi_J^{(0)} \rangle = 0 \quad (174)$$

which leads to

$$0 = \langle \Phi_0^{(0)} | \hat{H}^{(0)} | \Phi_J^{(0)} \rangle \Leftrightarrow 0 = \langle \Phi_0^{(0)} | \hat{H} - \hat{H}^{(1)} | \Phi_J^{(0)} \rangle \Leftrightarrow \langle \Phi_0^{(0)} | \hat{H} | \Phi_J^{(0)} \rangle = \langle \Phi_0^{(0)} | \hat{H}^{(1)} | \Phi_J^{(0)} \rangle \quad (175)$$

which shows that the matrix elements of the \hat{H} and $\hat{H}^{(1)}$ operators are equal. Using this result, and the knowledge that $\langle \Phi_0^{(0)} | \hat{H} | \Phi_J^{(0)} \rangle$ is nonzero only when $\Phi_J^{(0)}$ differs from $\Phi_0^{(0)}$ by two excitations (spinorbitals) we arrive at the following result

$$E_0^{(2)} = \frac{1}{4} \sum_{x,y} \sum_{r,s} \frac{\langle xy || rs \rangle \langle rs || xy \rangle}{\varepsilon_x + \varepsilon_y - \varepsilon_r - \varepsilon_s} \quad (176)$$

where the x, y indices run over all occupied spinorbitals (from 1 to N_e) while the indices r, s run over all virtual (unoccupied) spinorbitals (from $N_e + 1$ upwards). The two-electron integrals are defined as follows

$$\langle xy || rs \rangle = \int \int \frac{\chi_x^*(\mathbf{x}_1) \chi_y^*(\mathbf{x}_2) \chi_r(\mathbf{x}_1) \chi_s(\mathbf{x}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} d\mathbf{x}_1 d\mathbf{x}_2 - \int \int \frac{\chi_x^*(\mathbf{x}_1) \chi_y^*(\mathbf{x}_2) \chi_s(\mathbf{x}_1) \chi_r(\mathbf{x}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} d\mathbf{x}_1 d\mathbf{x}_2 \quad (177)$$

in terms of spinorbitals $\chi_i(\mathbf{r})$. The second order energy correction as given by equation 176 is widely used in *ab initio* calculations which include electron correlation and is available in many Quantum Chemistry software packages (e.g. GAUSSIAN, NWCHEM, GAMESS etc.). Of course, by definition, before doing an MP2 calculation one needs to have the Hartree-Fock solutions (spinorbitals) and their energies, so in practice MP2 calculations are performed as a post-processing step after a Hartree-Fock calculation.

Example 6 Equation 175 shows that the matrix elements of $\hat{H}^{(1)}$ between Slater determinants are the same with the matrix elements of the full Hamiltonian \hat{H} . Given that $\langle \Phi_0^{(0)} | \hat{H} | \Phi_J^{(0)} \rangle$ is non-zero only when $\Phi_J^{(0)} = \Phi_{xy}^{rs(0)}$, or in other words when $\Phi_J^{(0)}$ is constructed by replacing no more and no less than two of any χ_x, χ_y ground state spinorbitals by any two excited state spinorbitals χ_r, χ_s respectively, derive (176) from (173). Also given is the value of the non-zero matrix element: $\langle \Phi_0^{(0)} | \hat{H} | \Phi_{xy}^{rs(0)} \rangle = \langle xy || rs \rangle$.

According to the above we have

$$\langle \Phi_0^{(0)} | \hat{H}^{(1)} | \Phi_{xy}^{rs(0)} \rangle = \langle \Phi_0^{(0)} | \hat{H} | \Phi_{xy}^{rs(0)} \rangle = \langle xy || rs \rangle \quad (178)$$

We now re-write (173) confining its summations to only doubly-excited determinants

$$E_0^{(2)} = \frac{1}{2} \sum_{x,y=1}^{N_e} \frac{1}{2} \sum_{r,s=N_e+1}^{\infty} \frac{\langle \Phi_0^{(0)} | \hat{H}^{(1)} | \Phi_{xy}^{rs(0)} \rangle \langle \Phi_{xy}^{rs(0)} | \hat{H}^{(1)} | \Phi_0^{(0)} \rangle}{E_0^{(0)} - E_{xy}^{rs(0)}} \quad (179)$$

where the factors of 1/2 are introduced in order to make sure that each distinct pair of indices is used only once (e.g. if we have the pair $x=1$ and $y=5$, we will also have the same pair when $y=1$ and $x=5$, so we multiply with 1/2 to make sure we count this distinct pair only once) while the cases where $x = y$ and/or $r = s$ lead to zero matrix elements so it does not matter that they are included in the sum. We now substitute (178) into the above expression to obtain

$$E_0^{(2)} = \frac{1}{4} \sum_{x,y=1}^{N_e} \sum_{r,s=N_e+1}^{\infty} \frac{\langle xy || rs \rangle \langle rs || xy \rangle}{E_0^{(0)} - E_{xy}^{rs(0)}} \quad (180)$$

Finally, we need to express the denominator in terms of spinorbital energies. According to (170) we have:

$$\begin{aligned} E_0^{(0)} - E_{xy}^{rs(0)} &= \varepsilon_a + \dots + \varepsilon_x + \dots + \varepsilon_y + \dots + \varepsilon_z - (\varepsilon_a + \dots + \varepsilon_r + \dots + \varepsilon_s + \dots + \varepsilon_z) \\ &= \varepsilon_x + \varepsilon_y - \varepsilon_r - \varepsilon_s \end{aligned}$$

Using this result for $E_0^{(0)} - E_{xy}^{rs(0)}$ we get the expression for the MP2 energy in terms of spinorbitals and their energies

$$E_0^{(2)} = \frac{1}{4} \sum_{x,y=1}^{N_e} \sum_{r,s=N_e+1}^{\infty} \frac{\langle xy || rs \rangle \langle rs || xy \rangle}{\varepsilon_x + \varepsilon_y - \varepsilon_r - \varepsilon_s} \quad (181)$$

MP2 calculations with their ability to include at least some of the correlation energy, are a definite improvement over HF calculations. Figure 5 demonstrates this with some examples of bond lengths of small molecules calculated with the two methods and compared with experiment.

We should observe however that MP theory is also qualitatively different from HF theory. The total (perturbed) Hamiltonian in MP theory (165) is the exact one, involving the true electron-electron interactions (the $1/|\mathbf{r}_i - \mathbf{r}_j|$ terms). In contrast the HF Hamiltonian (zeroth order, $\hat{H}^{(0)}$) corresponds to a system of non-interacting particles that move in an effective (averaged) potential. Thus, MP theory includes electron

Molecule Method	 CH ₄	 NH ₃	 H ₂ O	 FH
HF	2.048	1.897	1.782	1.703
MP2	2.048	1.912	1.816	1.740
Experiment	2.050	1.913	1.809	1.733

Figure 5: Comparison of HF and MP2 calculations of equilibrium bond lengths (in atomic units) of some hydrides of first row elements.

correlation and the perturbed wavefunction does take into account the instant interactions between electrons: the modulus of the wavefunction (and hence the probability distribution) decreases as a function of the positions of any pair of electrons when they are approaching each other in space. This *dynamical correlation* is absent from a HF wavefunction (Slater determinant). In section 4.5 we saw that dispersion interactions between molecules are due to instantaneous fluctuations of their electronic distributions. We expect that HF calculations would be incapable of predicting dispersion interactions while MP calculations should be able to. This is indeed the case. For example, a HF calculation predicts zero binding between two Ne atoms, while an MP2 calculation predicts binding with an equilibrium distance of 6.06 a.u. and a binding energy of 2.3 meV. The “exact” values for these quantities are 5.84 a.u. and 3.6 meV respectively. There are numerous cases where dispersion interactions play a key role. In computational simulations involving such cases methods like MP theory need to be used. Figure 6 shows some examples of materials with dispersion interactions.

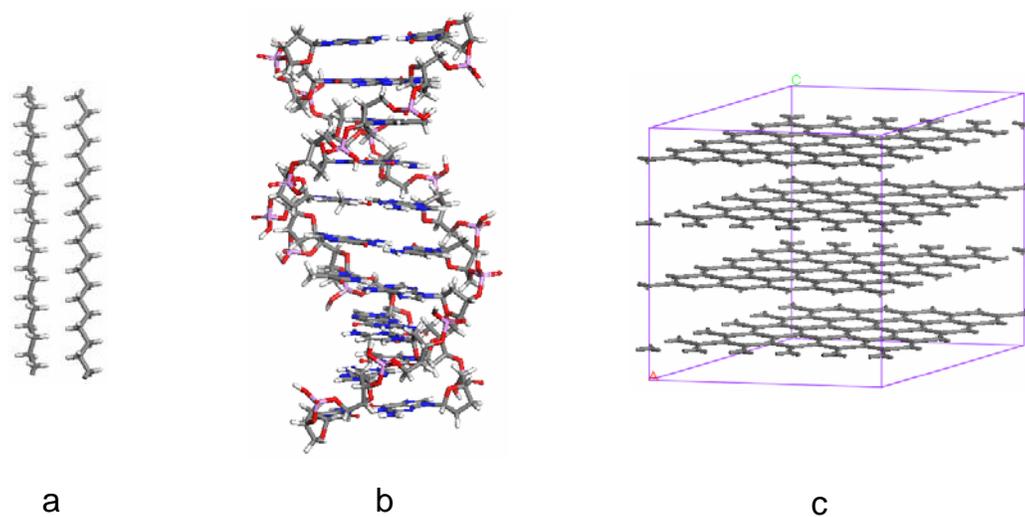


Figure 6: Examples of dispersion interactions in molecular structure. (a) Fragments of polyethylene, held together by dispersion forces; (b) The two strands in DNA are held together by hydrogen bonds but they are also stabilized by dispersion forces between the bases (planes) of each strand; (c) The structure of graphite consists of sheets of carbon held together by dispersion forces.