Perturbation Theory

Although quantum mechanics is beautiful stuff, it suffers from the fact that there are relatively few, analytically solveable examples. The classical solvable examples are basically piecewise constant potentials, the harmonic oscillator and the hydrogen atom. One can always find particular solutions to particular problems by numerical methods on the computer. An alternative is to use analytical approximation techniques. I will introduce you to two such techniques which are in common usage: perturbation theory and variational techniques. There are many such techniques developed over years, but these two are among the simplest, most fundamental, and most widely applied. In this chapter we will discuss time independent perturbation theory.

1st order Perturbation Theory

The perturbation technique was initially applied to classical orbit theory by Isaac Newton to compute the effects of other planets on the orbit of a given planet. The use of perturbative techniques in celestial mechanics led directly to the discovery of Neptune in 1846. Physicists used the pattern of irregularities in the orbit of Uranus, to predict the general location of the perturbing planet now known as Neptune.

Perturbation theory is used to estimate the energies and wave functions for a quantum system described by a potential which is only slightly different than a potential with a known solution. The approach is to develop a Taylor series in the perturbation which we will typically write as $\Delta V(x)$. We will show shortly that the lowest order correction to the energy or ΔE is $\langle \psi | \Delta V(x) | \psi \rangle$ which is just the expectation value $\langle \Delta V(x) \rangle$. For atomic systems change in energies due to practical perturbations using externally applied electric or magnetic fields is often quite small compared to typical atomic energy scales of several eV.[†] For example consider building an electric field by putting 10,000 volts between two

---

[†] However the perturbations on an electron due to the electrical fields present in adjacent atoms can be very large in molecules or multielectron atoms. We will discuss these in our chapter on Variational Methods.

electrodes separated by 1 cm. In this case $E_0 = |\vec{E}| = 10,000 \ V/0.01 \ m = 1 \times 10^6 \ V/m = 1 \times 10^{-3} \ V/nm$. This field (if directed along the $\hat{z}$ axis) will create a perturbation of $\Delta V = eE_0 z$ and one naively expects $\langle \Delta V \rangle \approx eE \ a_0$ since typical atomic dimension are on the order of the Bohr radius. Hence $\Delta E = e(10^{-3} \ V/nm)(0.05 \ nm) = 5 \times 10^{-5} \ eV$ which is very, very small on the scale of typical atomic binding energies of several eV although ΔE's of this size can be measured experimentally. The second order correction will be negligibly small so the Taylor series should be an excellent estimate when done properly such as in our section on the Stark Effect in this chapter.

So how does one build a Taylor expansion out of a function such as $\Delta V(x)$? The trick is to introduce an expansion parameter "g" and write: $\check{H}(g) = H_o + g\Delta V(x)$. Hence the Hamiltonian becomes a function of g as well as positions and derivatives. The wave function and energy also become functions of g: $\psi(g)$ and $E(g)$. The Schrödinger equation then becomes

$$\check{H}(g) \ \psi(g) = E(g) \ \psi(g) \tag{1}$$

Our approach will be to expand $H(g)$, $E(g)$, and $\psi(g)$ in powers of g and demand consistency order by order. If you want, feel free to think of setting $g \ll 1$ as a way of forcing the perturbation to be vanishingly small so that the Taylor series converges rapidly. If it turns out ΔV is small enough we can then restore $g = 1$ and be confident that the true answer is very close to the sum of just a few terms in the expansion. Although we are expanding in g, each factor of g will bring a power of ΔV along with it so essentially we will produce a Taylor expansion in powers of ΔV.

In keeping with this philosophy we expand both the wave function and energy for the perturbed system order by order as follows:

$$E_n = E_n^{(o)} + g \ E_n^{(1)} + g^2 \ E_n^{(2)} + ...$$

$$\psi_n = \psi_n^{(o)} + g \ \psi_n^{(1)} + g^2 \ \psi_n^{(2)} + ... \tag{2}$$

A few words on the notation are in order. The subscript n refers to the principal quantum number of the state ($n = 1$ might be ground , $n = 2$ might be first excited etc.). The $^{(o)}$ superscript denotes the zero order or unperturbed quantity. Hence $E_n^{(o)}$ is the ground state energy of the unperturbed system and $\psi_n^{(o)}(x)$ is the unperturbed ground state wavefunction. We will generally be content to compute $E_n^{(1)}$ and $\psi_n^{(1)}$ which are the lowest order (first order) corrections to the unperturbed system. We next insert the wave function and energy expansion given in Eq. (2) into a time independent Schrödinger Equation based on the Hamiltonian given in Eq. (1) or $H(g)\ \psi_n(g) = E_n(g)\ \psi_n(g)$:

$$(H_o + g\ \Delta V(x))\ \left(\psi_n^{(o)} + g\ \psi_n^{(1)} + g^2\ \psi_n^{(2)} + ...\right)$$

$$= \left(E_n^{(o)} + g\ E_n^{(1)} + g^2\ E_n^{(2)} + ...\right)\ \left(\psi_n^{(o)} + g\ \psi_n^{(1)} + g^2\ \psi_n^{(2)} + ...\right) \qquad (3)$$

Since Eq. (3) must be true independent of the strength of g as it is varied from $0 < g < 1$, it must be true that the Eq. (3) is true order by order. Grouping terms of the same order (or power of g) together we have:

$$H_o\ \psi_n^{(o)} = E_n^{(o)}\ \psi_n^{(o)} \quad \text{(zero order)}$$

$$g\ \left(H_o\ \psi_n^{(1)} + \Delta V(x)\ \psi_n^{(o)}\right) = g\ \left(E_n^{(o)}\ \psi_n^{(1)} + E_n^{(1)}\ \psi_n^{(o)}\right) \quad \text{(1st order)} \qquad (4)$$

There is nothing new in the zero order term of Eq. (4) since we already know that $\psi_n^{(o)}$ are eigenfunctions of the unperturbed Hamiltonian H_o. We can solve

$$H_o\ \psi_n^{(1)} + \Delta V(x)\ \psi_n^{(o)} = E_n^{(o)}\ \psi_n^{(1)} + E_n^{(1)}\ \psi_n^{(o)} \qquad (5)$$

to solve for the 1st order correction to the energy $E_n^{(1)}$ by simply forming brackets

by "multiplying" both sides by $< \psi_n^{(o)} |$:

$$< \psi_n^{(o)} \mid H_o \mid \psi_n^{(1)} > + < \psi_n^{(o)} \mid \Delta V(x) \mid \psi_n^{(o)} >$$

$$= E_n^{(o)} < \psi_n^{(o)} |\psi_n^{(1)} > + E_n^{(1)} < \psi_n^{(o)} \mid \psi_n^{(o)} > \qquad (6)$$

We note that

$$< \psi_n^{(o)} \mid H_o \mid \psi_n^{(1)} > = E_n^{(o)} < \psi_n^{(o)} \mid \psi_n^{(1)} > \qquad (7)$$

since we can operate H_o on the left because it is a self-adjoint operator. Note that $< \psi_n^{(o)} \mid \psi_n^{(o)} > = 1$ from orthonormality of the unperturbed wave functions. We thus have

$$E_n^{(o)} < \psi_n^{(o)} \mid \psi_n^{(1)} > + < \psi_n^{(o)} \mid \Delta V(x) \mid \psi_n^{(o)} > = E_n^{(o)} < \psi_n^{(o)} |\psi_n^{(1)} > + E_n^{(1)}$$
$$(8)$$

Cancelling terms we have much desired first order correction:

$$E_n \approx E_n^{(o)} + E_n^{(1)}$$

$$\text{where } E_n^{(1)} = \langle \psi_n^{(o)} \mid \Delta V(x) \mid \psi_n^{(o)} \rangle = \int\limits_{-\infty}^{+\infty} dx \ \left(\psi_n^{(o)}(x) \right)^* \Delta V(x) \ \psi_n^{(o)}(x) \quad (9)$$

The first order contribution is as memorable as it is reasonable. Eq. (9) says the shift in energy away from an unperturbed solution due to a perturbation $\Delta V(x)$ is just the expectation of the perturbing potential or

$$\Delta E \ \approx < \psi_n^{(o)} \mid \Delta V(x) \mid \psi_n^{(o)} >$$

<u>Perturbed Wave Functions</u>

Of course in quantum mechanics, the wave functions are often as important as the energy levels themselves. Is there a way of approximating the wave function for the perturbed system? You bet! Since we assume that the original unperturbed wave functions are complete, we can expand any function, including $\psi_n^{(1)}(x)$. Lets write this expansion as:

$$\psi_n^{(1)}(x) = \sum_k a_{nk} \ \psi_k^{(o)}(x) \tag{10}$$

where the amplitudes are for finding the k'th unperturbed solution in the first order wave function is a_{nk}. We now insert this expansion into our first order consistency equation given by Eq. (5):

$$\left(\sum_k a_{nk} \ H_o \mid \psi_k^{(o)} >\right) + \Delta V \ \psi_n^{(o)} = E_n^{(o)} \left(\sum_m a_{nk} \ \psi_k^{(o)}\right) + E_n^{(1)} \ \psi_n^{(o)} \tag{11}$$

We can extract the desired a_{nm} amplitudes by forming brackets by "multiplying" both sides of Eq. (11) by $< \ \psi_m^{(o)} \mid$

$$\left(\sum_k a_{nk} \ < \ \psi_m^{(o)} \mid H_o \mid \psi_k^{(o)} >\right) + < \ \psi_m^{(o)} \mid \Delta V \mid \psi_n^{(o)} >$$

$$= E_n^{(o)} \left(\sum_k a_{nk} \ < \psi_m^{(o)} \mid \psi_k^{(o)} >\right) + E_n^{(1)} \ < \ \psi_m^{(o)} \mid \psi_n^{(o)} > \tag{12}$$

We can simplify by exploiting the orthonormality condition $< \ \psi_m^{(o)} \mid \psi_k^{(o)} > = \delta_{mk}$ where δ_{mk}, known as the Kronecker delta, is defined as $\delta_{mk} = 1$ if $m = k$ and $\delta_{mk} = 0$ if $m \neq k$. We further note that $< \ \psi_m^{(o)} \mid H_o \mid \psi_k^{(o)} > = E_m^{(o)} \ \delta_{mk}$. Simplifying Eq. (12) we have:

$$\left(\sum_k a_{nk} \ E_m^{(o)} \ \delta_{mk}\right) + < \ \psi_m^{(o)} \mid \Delta V \mid \psi_n^{(o)} >$$

$$= E_n^{(o)} \left(\sum_k a_{nk} \ \delta_{mk}\right) + E_n^{(1)} \ \delta_{mn} \tag{13}$$

We now terms like $(\sum_k a_{nk} \ \delta_{mk})$ are only non-zero when $k = n$ and simplify to:

$(\sum_k a_{nk} \delta_{mk}) = a_{nm}$. We thus have:

$$E_m^{(o)} a_{nm} + < \psi_m^{(o)} |\Delta V | \psi_n^{(o)} > = E_n^{(o)} a_{nm} + E_n^{(1)} \delta_{nm} \qquad (14)$$

Solving for a_{nm} and inserting $E_n^{(1)} = < \psi_n^{(o)} | \Delta V(x) | \psi_n^{(o)} >$ we have:

$$a_{nm} = \frac{< \psi_m^{(o)} |\Delta V | \psi_n^{(o)} > (1 - \delta_{mn})}{E_n^{(o)} - E_m^{(o)}} \qquad (15)$$

We note that the Kronecker delta in the numerator of Eq. (15) means that $a_{nm} = 0$ whenever $n = m$. We can insert our expression for a_{nm} back into the expansion Eq. (10) to obtain:

$$\psi_n^{(1)} = \sum_{m \neq n} \frac{< \psi_m^{(o)} |\Delta V | \psi_n^{(o)} >}{E_n^{(o)} - E_m^{(o)}} \psi_m^{(o)} \quad \text{or}$$

$$\psi_n \approx \psi_n^{(o)} + \sum_{m \neq n} \frac{< \psi_m^{(o)} |\Delta V | \psi_n^{(o)} >}{E_n^{(o)} - E_m^{(o)}} \psi_m^{(o)} \qquad (16)$$

There is a subtlety in Eq. (16). In fact the δ_{mn} in Eq. (15) means that we can only find a_{mn} if $m \neq n$. Eq. (16) makes the much stronger claim that $a_{mm} = 0$, i.e. there is no correction to the leading $\psi_n^{(o)}$ term to first order in ΔV. The reason that such a a_{mm} term can be excluded is that it would break the normalization of the corrected wave function to first order in ΔV. All other terms break normalization to *second order* in ΔV which is ok since we are only computing things to first order.

In the limit that $\Delta V \to 0$, Equation (16) tells us that the wave function of principal quantum number n for the perturbed system approaches the wave function for the unperturbed system. Again this is a first order correction since it is first order in the perturbing potential ΔV. As the perturbation is turned on, the wave function for principal quantum number becomes a linear combination of

unperturbed wave functions with other principal quantum numbers. For example, the perturbed ground state will include contributions from the unperturbed excited state wave functions. The amplitude for "mixing" with a given excited state depends on how effective ΔV is in connecting the ground to a given excited state as well as the inverse of the energy difference between the ground and the excited state. The energy difference denominator means, for example, that strongest contributions to a perturbed ground state will from the first, second, and third excited state. Contributions to a perturbed ground state from a highly excited wave function tend to be small. This is often summarized by saying that "mixing" is strongest among "close" quantum states.

Often the symmetry of the wave function and the potential will cause $< \psi_m^{(o)} | \Delta V | \psi_n^{(o)} >$ to vanish. For example, if $\Delta V(x)$ is an odd function, and the unperturbed system is a symmetric well, the product $\psi_3^{(o)}(x) \times \Delta V(x) \times \psi_1^{(o)}(x)$ for the ground and second excited state is an odd function since both $\psi_1^{(o)}(x)$ and $\psi_3^{(o)}(x)$ are even functions. This means that the corresponding contribution of the second excited state unperturbed wave function will vanish after integration of $< \psi_3^{(o)} | \Delta V | \psi_1^{(o)} >$. Under this set of circumstances the ground state will only be able to "mix" with the first, third, fifth excited state, and so on.

A perturbation calculation is very reliable to the extent that the mixing of the wave functions is fairly small. According to Equation (16), first order perturbation theory should be fairly reliable so long as:

$$< \psi_m^{(o)} | \Delta V | \psi_n^{(o)} > \quad \ll \quad E_n^{(o)} - E_m^{(o)} \tag{17}$$

Other than going to higher orders, this is perhaps the simplest check to make to insure one is in a perturbative regime. Eq. (16) also tells us that we are in big trouble as $E_n^{(o)} \to E_m^{(o)}$. A new technique, which we discuss shortly , called *degenerate perturbation theory* is used when more than one state has the same energy in the unperturbed system.

Perturbation Theory Example

I've assigned the below system as a homework problem in the Physics 485 chapter **Bound States in One Dimension**, and discussed the limiting case where $E \ll V_o$ as model for an ammonia atomic clock.
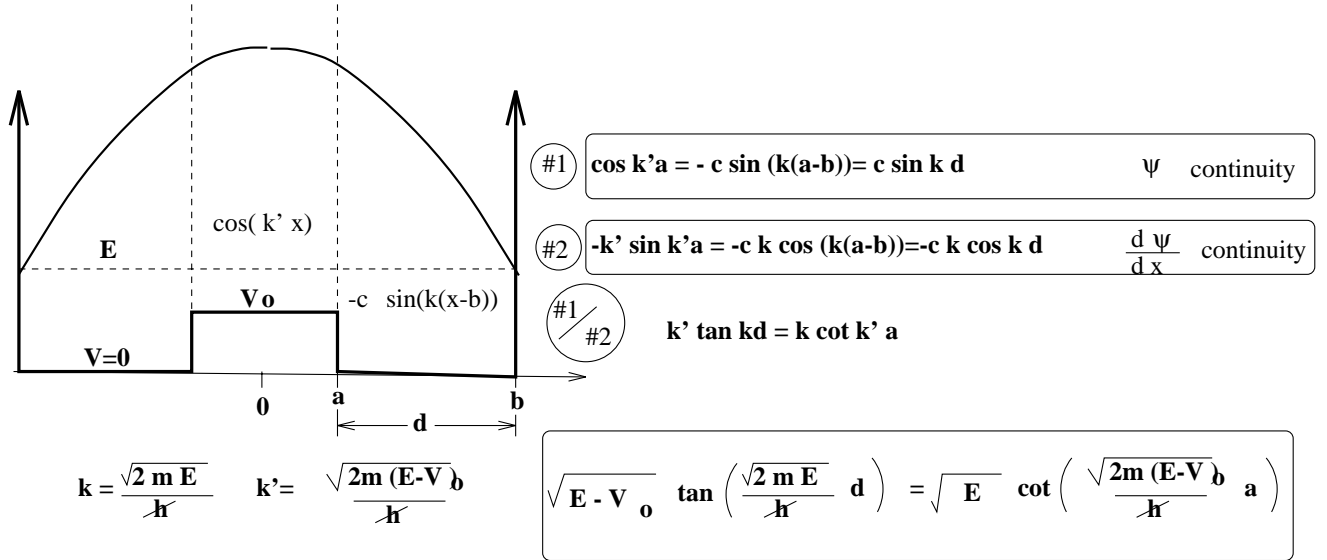
In this section we discuss the opposite limit where $V_o \ll E$ and we have a classically allowed wave function in all three regions of the well. As shown in the figure, it is straightforward to obtain a transcental equation which will allow one to solve for all even bound states. For sufficiently small V_o, we can also estimate the energy using first order perturbation theory. This example allows us the luxury of comparing the approximate perturbation energy estimate with the exact solution of the transcendental equation.

We wish to compute the shift in the ground state energy ΔE due to the potential hump, V_o, which extends over the range $-a < x < a$. The approximate energy shift is given by $\Delta E = <\psi_o|\Delta V|\psi_o>$ where $\psi_o(x)$ is the ground state wave function of the *unperturbed system*. Here the unperturbed ground state (*ie* the ground state for the system where $V_o = 0$) is the ground state of the infinite square well of width $2b$.

$$\psi_o(x) = \sqrt{\frac{1}{b}} \cos\left(\frac{\pi x}{2b}\right) \quad , \quad E_o = \frac{\hbar^2 k^2}{2m} = \frac{\hbar^2 \pi^2}{8mb^2} \tag{18}$$

Inserting this wave function into $< \psi_o | \Delta V | \psi_o >$ and realizing that the perturbation only exists over the range $-a < x < a$, we have:

$$\Delta E = < \psi_o | \Delta V | \psi_o >= \int_{-a}^{a} dx \sqrt{\frac{1}{b}} \cos\left(\frac{\pi x}{2b}\right) V_o \sqrt{\frac{1}{b}} \cos\left(\frac{\pi x}{2b}\right)$$

$$= \frac{V_o}{b} \int_{-a}^{a} dx \cos^2\left(\frac{\pi x}{2b}\right) = \frac{V_o}{b} \int_{-a}^{a} dx \frac{1 + \cos\left(\frac{\pi x}{b}\right)}{2} \tag{19}$$

In obtaining Eq. (19) we make use of the identity $\cos^2 \theta = (1 + \cos(2\theta))/2$ since it is easy to integrate.

$$\Delta E = \frac{V_o}{2b} \int_{-a}^{a} dx \, 1 + \cos\left(\frac{\pi x}{b}\right) = \frac{V_o}{b} \int_{0}^{a} dx \, 1 + \cos\left(\frac{\pi x}{b}\right) = V_o \left(\frac{a}{b} + \frac{1}{\pi} \sin\left(\frac{\pi a}{b}\right)\right) \tag{20}$$

Combining ΔE with the ground state energy of the unperturbed system , E_o , would give us approximate new ground state energy of:

$$E_{\text{ground}} = \frac{\hbar^2 \pi^2}{8mb^2} + V_o \left(\frac{a}{b} + \frac{1}{\pi} \sin\left(\frac{\pi a}{b}\right)\right) \tag{21}$$

A simple check of this result is provided by going to the limit where $a \rightarrow b$. We begin with the "exact" result. In this limit, the entire floor of the well has been elevated to the potential $V = V_o$ and the problem has degenerated into a one dimensional infinite square well.
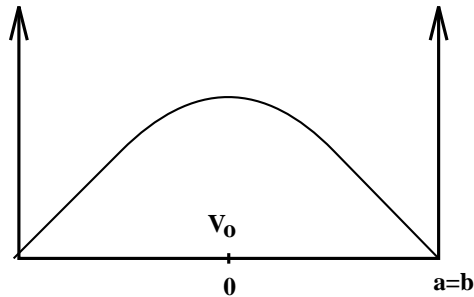
The correct wave function for this well has a wavelength of $\lambda = 4b$, and a kinetic energy of $\mathrm{KE} = \hbar^2\pi^2/(8\ m\ b^2)$. This kinetic energy must be added to the potential energy $V = V_o$ to get an exact total energy of:

$$E^{(\mathrm{exact})} = \frac{\hbar^2\pi^2}{8mb^2} + V_o \tag{22}$$

The answer from perturbation theory is:

$$E^{(\mathrm{perturb})} = \frac{\hbar^2\pi^2}{8mb^2} + V_o\ \left(\frac{a}{b} + \frac{1}{\pi}\ \sin\left(\frac{\pi a}{b}\right)\right) \rightarrow \frac{\hbar^2\pi^2}{8mb^2} + V_o\ \left(1 + \frac{1}{\pi}\ \sin\left(\pi\right)\right)$$

$$= \frac{\hbar^2\pi^2}{8mb^2} + V_o \tag{23}$$

Both the exact and perturbation answers must agree in the limit $V_o << E_o$ where the perturbation theory is valid. As you see they do agree.

We next compare the exact and perturbation solutions for an electron placed in the well illustrated below. The hump perturbation we are considering is a 5 eV on a system whose unperturbed ground state-first excited state difference is only 7.05 eV and we might justifiably be concerned that that the first order approximation might not be that good.
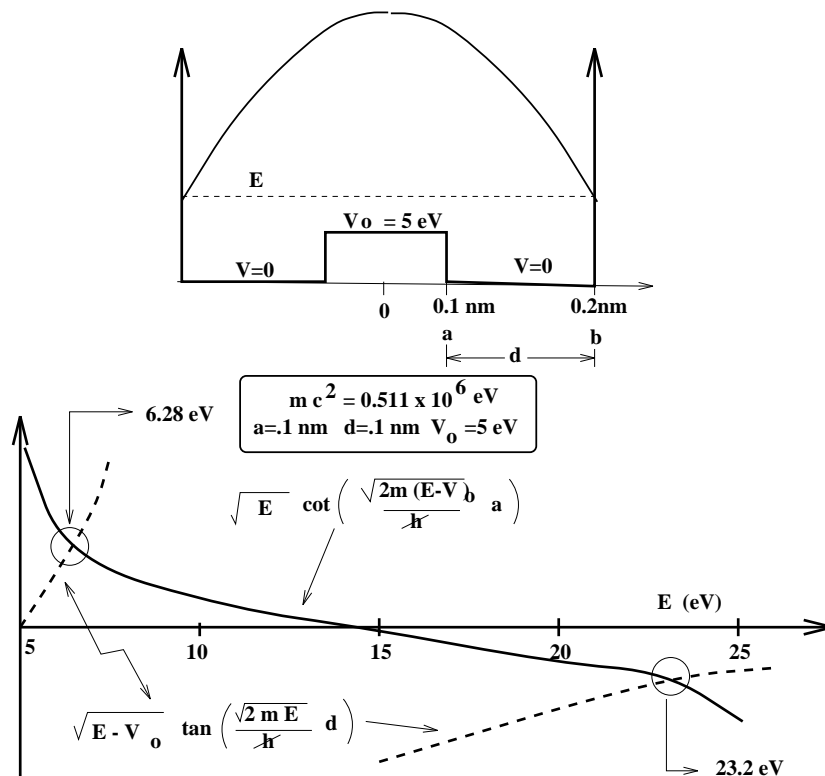
This is the graphic solution of one of the **Bound State** Physics 485 exercises. As indicated in the figure, the exact solution of the electron's ground state is about $E_1 = 6.28\ eV$ which implies an energy shift above the unperturbed ground state of $\Delta E = 3.93\ eV$. According to Eq. (21) approximate energy shift in perturbation theory is:

$$\Delta E = <\psi_o|\Delta V|\psi_o> = V_o\ \left(\frac{a}{b} + \frac{1}{\pi}\ \sin\left(\frac{\pi a}{b}\right)\right)$$

$$\Delta E = 5\left(\frac{0.1}{0.2} + \frac{1}{\pi}\ \sin\left(\frac{\pi 0.1}{0.2}\right)\right) = 5\left(0.5 + \frac{1}{\pi}\right) = 4.09\ eV \qquad (24)$$

Our ground state energy estimation from first order perturbation theory is then: $E_{\text{perturb}} = E_o + \Delta E = 2.35 + 4.09 = 6.44\ eV$. The perturbation estimate is about 2.5 % higher than the true solution which I find remarkably good considering the magnitude of the perturbation.

Degenerate Perturbation Theory

Some interesting things happen to 1st order perturbation theory in multi-dimensional problems where there is usually degeneracy or several states with exactly the same energy owing to some symmetry. One immediate problem follows from Eq. (16) which gives the first order correction to the wave function. This formula says that the perturbed wave function for a given is composed of all other states that connect to the given state through the perturbation ΔV. The problem is the amplitude contains a denominator that depends on the difference of energies. This denominator will blow up if there is another state with the same energy as the given state. The only escape clause is we might get a well defined result if the matrix element in the numerator vanishes as well.

But perhaps the more telling point is illustrated below. Recall the strength parameter g that turns the perturbation on as it is increased from $0 \rightarrow 1$. We illustrate the shift in energy as g is increased for the case of a set of non-degenerate states to the left. In this case each of three states a, b and c responds to the perturbation differently and there is a clear path from the unperturbed system to the perturbed system.
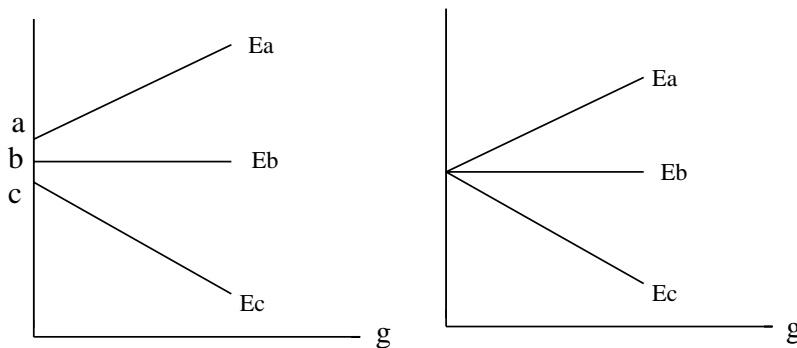
We now contrast this situation to the the triply degenerate system illustrated on the right. We see in this case the perturbation has *lifted* or broken the degeneracy creating three energy levels. The perturbation breaks the degeneracy by breaking the underlying symmetry responsible for the degeneracy. Evidently the three degenerate eigenfunctions re-arrange themselves to match the symn-

metry of the perturbation which then pulls the energy levels apart But what unperturbed wave function should we insert in our 1st order expression?

Before describing the (very elegant) technique used to find these re-arranged wave functions and compute these energies, let us give a simple example. Consider the two dimensional, isotropic harmonic oscillator in the Cartesian (x , y) basis. As you recall, the 1st excited state is doubly degenerate with an energy of $2\hbar\omega$

$$\psi_1(x, y) = \mathcal{N} \ x \ \exp(-\gamma(x^2 + y^2)/2) \ \ , \psi_2(x, y) = \mathcal{N} \ y \ \exp(-\gamma(x^2 + y^2)/2)$$

where $\gamma = m\omega/\hbar$. We can normalize the wave functions using our Gaussian integral results

$$I_n = \int\limits_{-\infty}^{\infty} dx \ x^n \ exp(-\gamma x) \ , \ I_{n+2} = -\frac{\partial I_n}{\partial \gamma} \ , \ I_0 = \sqrt{\pi} \ \gamma^{-1/2} \ , \ I_2 = \sqrt{\pi} \ \frac{1}{2} \ \gamma^{-3/2} \ldots$$

$$1 = \int\limits_{-\infty}^{\infty} dx \int\limits_{-\infty}^{\infty} dy \ (\psi_1(x, y))^2 = \mathcal{N}^2 \int\limits_{-\infty}^{\infty} x^2 dx \ e^{-\gamma x} \int\limits_{-\infty}^{\infty} dy \ e^{-\gamma y} = \mathcal{N}^2 \ I_2 \ I_0 =$$

$$\mathcal{N}^2 \left(\sqrt{\pi} \ \frac{1}{2} \ \gamma^{-3/2}\right) \left(\sqrt{\pi} \ \gamma^{-1/2}\right) = \frac{\pi}{2\gamma^2} \rightarrow \mathcal{N} = \gamma\sqrt{2/\pi}$$

In this case the two wave functions are both real and their corresponding PDF's have independent reflection symmetry about the x and y axis since the PDF is both even in x and even in y. Of course any linear combination of $\psi = \alpha\psi_1 + \beta\psi_2$ will also be a valid stationary state with an energy of $2\hbar\omega$ which you can easily confirm by applying $\check{H}_o$ to ψ, but as we will show, a linear combination of ψ_1 and ψ_2 can produce a PDF that breaks this double reflection symmetry.

Now imagine applying a perturbation of the form $\Delta V = \beta\ x\ y$.[1] Although the perturbation has some interesting symmetries of its own, it breaks independent x or y reflection symmetry. For example for a given positive y value, it is positive for $x > 0$ and negative for $x < 0$ (assuming $\beta > 0$). If we were to throw caution to the winds and try to compute the shift in energy of ψ_1 or ψ_2 state by $\Delta E_1 =< 1|\Delta V|1 >$ or $\Delta E_2 =< 2|\Delta V|2 >$ we would (erroneously) conclude the first-order energy shifts are zero owing to the reflection symmetry of the PDF's. For a potential that only depends on position since $PDF(x,y) = \psi^*(x,y)\psi(x,y)$ we can write:

$$\Delta E = \int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty} dx\ dy\ PDF(x,y)\ \beta xy$$

and for every given y value the x integral integrand is odd and vice versa, thus canceling all ΔE contributions.

We can easily construct wave functions that produce PDF's with pieces with the same underlying symmetry as ΔV such that the naive 1st order perturbation estimate of ΔE won't vanish. The trick is to switch bases to $\psi_\pm = (\psi_1 \pm \psi_2)/\sqrt{2}$ which produce PDF's of

$$PDF_\pm = \frac{\mathcal{N}^2}{2}\left(\underline{(x^2 + y^2)} \pm 2xy\right)\exp(-\gamma(x^2+y^2)) = \frac{\gamma^2}{\pi}\left(\underline{(x^2 + y^2)} \pm 2xy\right)\exp(-\gamma(x^2+y^2))$$

Assuming we can, if we try to construct $\Delta E_\pm =< \psi_\pm|\beta xy|\psi_\pm >$ by integrating the PDF $\times \beta xy$, it is apparent that the underlined part will integrate to zero leaving:

$$\Delta E_\pm = \pm\beta\frac{2\gamma^2}{\pi} \times \int\limits_{-\infty}^{+\infty} dx\ x^2\ e^{-\gamma\ x^2} \times \int\limits_{-\infty}^{+\infty} dy\ y^2\ e^{-\gamma\ y^2} = \pm\frac{2\beta\gamma^2}{\pi} \times I_2 \times I_2 = \pm\frac{\beta}{2\gamma}$$

$$(25)$$

If all of this is true, it suggests that the βxy perturbation, breaks the 2 fold

---

1 This perturbation could be created by a set of hyperbolic electrostatic plates creating an electric quadrupole field if the the harmonic oscillator is charged.

degeneracy of the two $2\hbar\omega$ states by raising the energy of the state ψ_+ and lowering the energy of the state ψ_-.

This exercise raises an interesting question which we will answer shortly. Although the PDF for $\psi_\pm$ contain pieces with the same symmetry as the perturbation, other linear combinations of ψ_1 and ψ_2 such as $\psi' = (4\psi_1 + \psi_2)/\sqrt{17}$ will contain a piece like this as well. If we apply the perturbation and observe the splitting of single state into two states – how do we know that the upper state is ψ_+ rather than ψ' ?

Eigenvector Equation Approach

Lets examine the first order expression again:

$$H_o \; \psi^{(1)} + \Delta V(x) \; \psi^{(o)} = E^{(o)} \; \psi^{(1)} + E^{(1)} \; \psi^{(o)}$$

This is a version of Eq. (5) except we have suppressed the n index. The main thing that we don't know for the case of degenerate states is which combination of states do we want to call $\psi^{(0)}$. We do know that it is a linear combination of the unperturbed degenerate eigenfunctions which we will write as $\psi^{(0)} = \sum_i a_i \psi_i = \sum_i a_i |i>$. The key to entangling the degenerate situation is to find the amplitudes a_i. It is remarkably easy to. First insert our form for $\psi^{(0)}$

$$H_o \; \psi^{(1)} + \Delta V(x) \; \sum_i a_i |i >= E^{(o)} \; \psi^{(1)} + E^{(1)} \; \sum_i a_i |i >$$

Next multiply through by $< j|$ where ψ_j is another degenerate solution to the unperturbed system.

$$< j|H_o \; \psi^{(1)} > + < j|\Delta V(x) \; \sum_i a_i |i >= E^{(o)} \; < j|\psi^{(1)} > + E^{(1)} \; < j|\sum_i a_i |i >$$

Because $\check{H}_o$ is self-adjoint we have $< j|H_o \; \psi^{(1)} >= E^{(o)} < j|\psi^{(1)} >$ thus canceling all dependence on the unknown $\psi^{(1)}$ - thank goodness! We also have

$< j| \sum_i a_i|i >= \sum_i a_i < j|i >= a_j$ owing to the assumed orthonormality of the unperturbed wave functions. We thus have the remarkably simple result:

$$\sum_i < j|\Delta V|i > a_i = E^{(1)}a_j \text{ or in matrix notation}$$

$$\begin{pmatrix} < 1|\Delta V|1 > & < 1|\Delta V|2 > & < 1|\Delta V|3 > & \ldots \\ < 2|\Delta V|1 > & < 2|\Delta V|2 > & < 2|\Delta V|3 > & \ldots \\ < 3|\Delta V|1 > & < 3|\Delta V|2 > & < 3|\Delta V|3 > & \ldots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \end{pmatrix} = E^{(1)} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \end{pmatrix} \quad (26)$$

We thus have the rather wonderful result, that the initially degenerate states that a perturbation separates are the eigenfunctions of a finite matrix constructed from the perturbation matrix elements. The 1st order corrections to the energy levels for a given state are just the eigenvalues of the matrix.

Lets try it out for the $\Delta V = \beta xy$ problem starting from the basis

$$\psi_1(x,y) = \gamma\sqrt{\frac{2}{\pi}} \ x \ \exp(-\gamma(x^2 + y^2)/2) \ , \psi_2(x,y) = \gamma\sqrt{\frac{2}{\pi}} \ y \ \exp(-\gamma(x^2 + y^2)/2)$$

Here we have a 2×2 matrix. We argued from symmetry before that $< 1|\Delta V|1 >=< 2|\Delta V|2 >= 0$. Because the perturbation and wave functions are real we know $< 2|\Delta V|1 >=< 1|\Delta V|2 >$ which leaves us with one inner product to evaluate:

$$< 1|\Delta V|2 > \frac{2\beta\gamma^2}{\pi} \times \int\limits_{-\infty}^{+\infty} dx \ x^2 \exp(-\gamma \ x^2) \times \int\limits_{-\infty}^{+\infty} dy \ y^2 \exp(-\gamma \ y^2) = \frac{2\beta\gamma^2}{\pi} I_2 I_2 = \frac{\beta}{2\gamma}$$

Looks familiar? Lets write out the matrix expression and subtract the right-side from the left-side of the equation:

$$\begin{pmatrix} 0 & \beta/(2\gamma) \\ \beta/(2\gamma) & 0 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = E^{(1)} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \rightarrow \begin{pmatrix} -E^{(1)} & \beta/(2\gamma) \\ \beta/(2\gamma) & -E^{(1)} \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = 0$$

Recalling the (hopefully) familiar argument, either both amplitudes are zero, or

the matrix must be singular and thus have a zero determinent:

$$\begin{vmatrix} -E^{(1)} & \beta/(2\gamma) \\ \beta/(2\gamma) & -E^{(1)} \end{vmatrix} = 0 \rightarrow E^{(1)} = \pm\frac{\beta}{2\gamma}$$

We insert the eigenvalues to find the eigenvectors to within a factor whose modulus can set by normalization and whose phase is arbitrary.

$$\begin{pmatrix} -\beta/(2\gamma) & \beta/(2\gamma) \\ \beta/(2\gamma) & -\beta/(2\gamma) \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = 0 \rightarrow a_1 = a_2 = 1/\sqrt{2} \rightarrow \psi_+ \text{ with } E^{(1)} = +\frac{\beta}{2\gamma}$$

$$\begin{pmatrix} +\beta/(2\gamma) & \beta/(2\gamma) \\ \beta/(2\gamma) & +\beta/(2\gamma) \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = 0 \rightarrow \frac{\psi_1 - \psi_2}{\sqrt{2}} = \psi_- \text{ with } E^{(1)} = -\frac{\beta}{2\gamma}$$

Exactly the same answer that we got before in Eq. (25) but without the need for obtaining $\psi_\pm$ through an inspired guess.

But was our guess that inspired?? In retrospect we were looking for an eigenfunction of the ΔV operator. Recall our old theorem on simultaneous eigenfunctions. If an operator $\check{Q}$ commutes with $[\Delta V, \check{Q}] = 0$, it should be possible to find simultaneous eigenfunctions of $\check{Q}$ and ΔV. As we tried to point out earlier, symmetry appears to be the key to understanding degenerate perturbation theory – hence an obvious candidate for $\check{Q}$ would be a symmetry operator. If we can arrange the degenerate eigenfunctions such that are eigenfunctions of a symmetry operator of the perturbation, we have a good chance at guessing the eigenfunctions of the ΔV matrix and if we fail we can always solve the secular equation. The symmetry of $\Delta V = \beta xy$ that we exploited in our guess was $x \leftrightarrow y$ exchange. Lets call this operator $\mathcal{S}$ where $\mathcal{S}\psi(x,y) = \psi(y,x)$. Clearly $\check{\mathcal{S}}\Delta V f(x,y) = \Delta V \check{\mathcal{S}} f(x,y)$ where $f(x,y)$ is a test function. This implies $[\check{\mathcal{S}}, \Delta V] = 0$. Can we construct wave functions out of the degenerate wave functions such that $\mathcal{S}\psi(x,y) = \pm\psi(x,y)$[1] or $\psi(x,y) = \pm\psi(y,x)$? You can see that

---

1 We know the eigenvalues of $\check{\mathcal{S}}$ must be ± 1 since $\mathcal{S}^2 = 1$

since $\check{\mathcal{S}}\psi_1 = \psi_2$, it must be true that $(\psi_1 \pm \psi_2)/\sqrt{2}$ fits the bill with symmetry eigenvalues of ± 1. We now illustrate degenerate perturbation theory with one of the most degenerate of all physical systems – the hydrogen atom. We work the a classic problems of hydrogen solved by countless generations of physics undergraduates: the Stark Effect.

The Stark Effect

Hydrogen is a great testing ground for quantum mechanical ideas since it is generally solveable and its energy levels can be measured spectroscopically with remarkable precision. Lets consider applying a strong electric field $\mathcal{E}$ along the $\hat{z}$ direction to the 4 degenerate $n = 2$ spacial orbitals of hydrogen. The perturbation is of the form $\Delta V = +e\mathcal{E}z$ where $e = +1.6 \times 10^{-19}$ C. Essentially this means that work is required to place the electron on top of the proton (*i.e.* with $z > 0$) in a uniform electric field $\vec{\mathcal{E}} = \mathcal{E}\hat{z}$ since the electrostatic forces will try to flip the atom over.

If we are clueless as to what combination of wave functions will be affected by the perturbation, we would begin by constructing the 4×4 $< j|\Delta V|i >$ matrix. Each matrix element will involve a three dimensional integral over the $n = 2$ hydrogen wave functions. Sounds daunting but in fact we only will need to do one integral and that one is actually pretty easy!

Our chief tool in deciding which $< j|\Delta V|i >$ elements survive is parity symmetry or what happens to functions when $\vec{r} \to -\vec{r}$. Recall from our discussion of hydrogen in Physics 485 $\check{\mathcal{P}} Y_\ell^m = (-1^\ell)Y_\ell^m$. In other words if ℓ is odd , the parity is odd. Clearly the parity of our perturbation is odd since $\check{\mathcal{P}}z = -z$. One will only get a non-zero integral if the integrand has even parity. We need to compute integrals of the form $< j|z|i >$. Since the parity of z is odd, we can only get a non-zero matrix element if the bra and cket state have opposite parity. In one stroke we have eliminated any diagonal contributions since in this case the bra and cket states are the same and thus have the same parity. We have also eliminated matrix elements connecting any two of the $\ell = 1$ states since they

have the same parity. Hence only $\ell = 1$ and $\ell = 0$ states can connect through a z perturbation.

Hence of the 10 original, upper triangle of integrals we would have to consider, we are down to just three:

$$< \psi_{211}|z|\psi_{200} > \quad , \quad < \psi_{210}|z|\psi_{200} > \quad , \quad < \psi_{21-1}|z|\psi_{200} >$$

But actually only the $< \psi_{210}|z|\psi_{200} >$ matrix element will survive integration over ϕ, remember why? If $z = r\cos\theta$ our integral will be.

$$< 21m|z|200 >= \int\limits_{0}^{\infty} r^2\, dr \int\limits_{-1}^{+1} d\cos\theta \int\limits_{0}^{2\pi} d\phi\, \psi_{21m}^*(r,\theta,\phi) \times r\cos\theta \times \psi_{200}(r,\theta,\phi)$$

$$\propto \int\limits_{0}^{2\pi} d\phi\, e^{-im\phi} = 0 \quad \text{unless } m = 0$$

where we explicitly put in the $\exp(im\phi)$ part of $< 21m|$ and showed the ϕ integral only survives if $m = 0$.

Hence only $< 210|z|200 >=< 200|z|210 >^*$ are non-zero. The Stark perturbation will therefore split two linear combinations of $a_1\,|200 > \; + \; a_2\,|210 >$. We thus only need to "diagonalize" or solve the eigenvector equation for a 2×2 matrix of the form:

$$+e\mathcal{E} \begin{pmatrix} 0 & < 200|z|210 > \\ < 210|z|200 > & 0 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = E^{(1)} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$$

This is essentially a 2×2 version of Eq. (26). We now need the relevant hydrogen wave function to constuct our sole surviving integral.

$$\psi_{200} = \frac{1}{4\sqrt{2\pi}} \left(\frac{Z}{a_o}\right)^{3/2} \left(2 - \frac{Zr}{a_o}\right) e^{-Zr/2a_o}$$

19

$$\psi_{210} = \frac{1}{4\sqrt{2\pi}} \left(\frac{Z}{a_o}\right)^{3/2} \frac{Zr}{a_o} e^{-Zr/2a_o} \cos\theta$$

Inserting these forms in our integral expression we get:

$$< 210|z|200 > = \frac{1}{32\pi a_o^3} \left[\int_0^\infty r^3 \, dr \, e^{-r/a_o} \left(2 - \frac{r}{a_o}\right)\frac{r}{a_o}\right] \left[\int_{-1}^{+1} d\cos\theta \, \cos^2\theta\right] \int_0^{2\pi} d\phi$$

The ϕ integral is just 2π. The $\cos\theta$ integral is just $2/3$. We can simplify the radial integral through the substitution $v = r/a_o$ which brings in 4 powers of a_0 to get:

$$< 210|z|200 > = \frac{a_0}{24} \int_0^\infty dv \, v^4 \, (2 - v) \, \exp(-v)$$

We are getting close. Radial integrals over the hydrogen wave function are always really simple in light of the remarkably cool definite integral:

$$\int_0^\infty dy \, y^n \, \exp(-ay) = \frac{n!}{a^{n+1}} \text{ if } n \in \{0, 1, 2, \ldots\}$$

Hence we get a very simple result after all of this work:

$$< 210|z|200 > = \frac{a_0}{24} (2 \times 4! - 5!) = -3a_o$$

Thus we need to diagonalize

$$+3e\mathcal{E} \, a_o \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = E^{(1)} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \rightarrow +3e\mathcal{E}a_o \begin{pmatrix} -\lambda & -1 \\ -1 & -\lambda \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = 0$$

where $E^{(1)} = 3e\mathcal{E}a_o\lambda$. As usual we set the determinant to zero to obtain the characteristic equation

$$\lambda^2 - 1 = 0 \rightarrow \lambda = \pm 1 \rightarrow E^{(1)} = \pm 3e\mathcal{E}a_o$$

Inserting $\lambda = 1$ into our matrix equation gives $a_1 = -a_2 = 1/\sqrt{2}$ where we threw in the $1/\sqrt{2}$ to preserve normalization. Inserting $\lambda = -1$ into our matrix

equation gives $a_1 = a_2 = 1/\sqrt{2}$. Thus of the four initially degenerate $n = 2$ states, the electric field will cause the combination $(\psi_{200} - \psi_{210})/\sqrt{2}$ to raise in energy by an amount $\Delta E = +3e\mathcal{E}a_o$ and the combination $(\psi_{200} + \psi_{210})/\sqrt{2}$ to lower its energy by an amount $\Delta E = -3e\mathcal{E}a_o$ relative to its unperturbed energy given by the Bohr formula. The two combinations $\psi_{21\pm1}$ will still be degenerate at their unperturbed energy given by the Bohr equation as illustrated below:
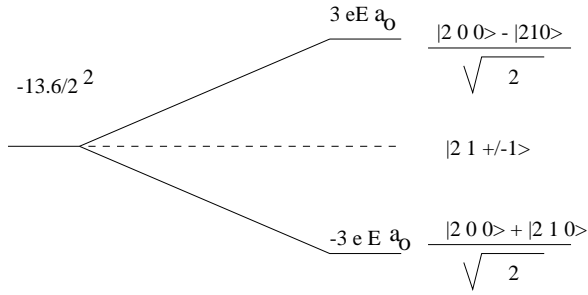
$$\boxed{\text{Important Points}}$$

1. If we have a perturbed quantum system where a small perturbation $\Delta V(x)$ is added to an unperturbed Hamiltonian H_o, to first order the energies of the perturbed system are given by:

$$E_n \approx E_n^{(o)} + <\ \psi_n^{(o)}\ |\ \Delta V\ |\ \psi_n^{(o)}\ >$$

where $E_n^{(o)}$ are the energies and $\psi_n^{(o)}$ are the wave functions of the unperturbed system.

2. The wave functions for the perturbed system to first order are given by:

$$\psi_n \approx \psi_n^{(o)} + \sum_{m \neq n} \frac{<\ \psi_m^{(o)}\ |\Delta V\ |\ \psi_n^{(o)}\ >}{E_n^{(o)} - E_m^{(o)}}\ \psi_m^{(o)} \tag{12}$$

This means that states closest in energy tend to have the largest mixing as long as symmetry considerations do not cause $<\ \psi_m^{(o)}\ |\Delta V\ |\ \psi_n^{(o)}\ > = 0$.

3. The perturbation estimates should be reasonably accurate as long as

$$< \psi_m^{(o)} |\Delta V | \psi_n^{(o)} > \ll E_n^{(o)} - E_m^{(o)}$$

Very roughly, this means we should be accurate as long as the perturbation is small compared to the typical splitting between quantum levels.

4. For the case of n degenerate states, one must use degenerate perturbation theory. To first order, one solves the eigenvalue problem for an $n \times n$ marix.

$$\begin{pmatrix} < 1|\Delta V|1 > & < 1|\Delta V|2 > & < 1|\Delta V|3 > & \ldots \\ < 2|\Delta V|1 > & < 2|\Delta V|2 > & < 2|\Delta V|3 > & \ldots \\ < 3|\Delta V|1 > & < 3|\Delta V|2 > & < 3|\Delta V|3 > & \ldots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \end{pmatrix} = E^{(1)} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \end{pmatrix}$$

The 1st order correction to the energy is the eigenvalue $E^{(1)}$ and the eignvectors are the state that have their degeneracy broken by the perturbation.

5. We discussed a classic example of degenerate perturbation theory applied to hydrogen the Stark effect (or splitting of the hydrogen levels due to an intense electric field into combinations with a dipole moment).