Many important problems in quantum mechanics cannot be solved exactly. In these cases, perturbation theory is an important component of our quantum mechanical toolkit. To capture light-matter interactions and spectroscopy, we will ultimately need to use time-dependent perturbation theory, but we will first build up our mathematical formalism using time-independent perturbation theory. Time-independent perturbation theory comes in two flavors: non-degenerate (when all the eigenvalues are distinct) and degenerate (when some eigenvalues may be identical).

1 Non-degenerate perturbation theory

1.1 Laying out the problem

We start by splitting the Hamiltonian into two parts:

\[ \hat{H} = \hat{H}^{(0)} + \lambda \hat{H}^{(1)} \]  

where \( \hat{H}^{(0)} \) is a “zeroth-order” reference Hamiltonian whose eigenvectors and eigenvalues we know, and which is not too different from our present unknown problem. \( \hat{H}^{(1)} \) is the perturbation, a disturbance with coefficient \( \lambda \) that should be “small.” This \( \lambda \) coefficient is really there for bookkeeping purposes, as we will see shortly. The scale for what is considered small is typically the energy level separations of \( \hat{H}^{(0)} \). Some examples of problems that can be treated within a small perturbation picture include \( \mathbf{L} \cdot \mathbf{S} \) coupling in light atoms, anharmonic oscillators, and the static Stark and Zeeman effects.

Since \( \hat{H} \) has a nice power series expansion in \( \lambda \) as written, let’s develop analogous power series expansions for the \( j^{th} \) eigenvector and eigenvalue solutions:

\[ \hat{H}|\psi_j(\lambda)\rangle = E_j(\lambda)|\psi_j(\lambda)\rangle \quad (2) \]
\[ E_j(\lambda) = E_j^{(0)} + \lambda E_j^{(1)} + \frac{1}{2} \lambda^2 E_j^{(2)} + \cdots \quad (3) \]
\[ |\psi_j(\lambda)\rangle = |\psi_j^{(0)}\rangle + \lambda |\psi_j^{(1)}\rangle + \frac{1}{2} \lambda^2 |\psi_j^{(2)}\rangle + \cdots \quad (4) \]

The task before us is to develop expressions for \( E_j^{(n)} \) and \( |\psi_j^{(n)}\rangle \) \((n > 0)\) in terms of known quantities: our reference Hamiltonian \( \hat{H}^{(0)} \), the perturbation \( \hat{H}^{(1)} \), and our reference eigenvectors and eigenvalues \( E_j^{(0)} \) and \( |\psi_j^{(0)}\rangle \).
Assuming no time-dependence for the moment, we will substitute the expansions above into the time-independent Schrödinger equation:

\[ \hat{H}\psi_j(\lambda) = E_j(\lambda)\psi_j(\lambda) \]  \hspace{1cm} (5)

RHS: \[ (\hat{H}^{(0)} + \lambda\hat{H}^{(1)}) (|\psi_j^{(0)}\rangle + \lambda|\psi_j^{(1)}\rangle + \frac{1}{2}\lambda^2|\psi_j^{(2)}\rangle + \cdots) \]  \hspace{1cm} (6)

LHS: \[ (E_j^{(0)} + \lambda E_j^{(1)} + \frac{1}{2}\lambda^2E_j^{(2)} + \cdots) (|\psi_j^{(0)}\rangle + \lambda|\psi_j^{(1)}\rangle + \frac{1}{2}\lambda^2|\psi_j^{(2)}\rangle + \cdots) \]  \hspace{1cm} (7)

The above expression can only hold true for arbitrary values of \( \lambda \) if the coefficients of each power of \( \lambda \) are equivalent on both right and left hand sides. So, grouping in powers of \( \lambda \):

\[ \lambda^0 : \quad \hat{H}^{(0)}|\psi_j^{(0)}\rangle = E_j^{(0)}|\psi_j^{(0)}\rangle \quad \checkmark \quad \text{our solved reference problem} \]  \hspace{1cm} (8)

\[ \lambda^1 : \quad \hat{H}^{(0)}|\psi_j^{(1)}\rangle + \hat{H}^{(1)}|\psi_j^{(0)}\rangle = E_j^{(0)}|\psi_j^{(1)}\rangle + E_j^{(1)}|\psi_j^{(0)}\rangle \]  \hspace{1cm} (9)

\[ \lambda^2 : \quad \hat{H}^{(1)}|\psi_j^{(1)}\rangle + \frac{1}{2}\hat{H}^{(0)}|\psi_j^{(2)}\rangle = \frac{1}{2}E_j^{(2)}|\psi_j^{(0)}\rangle + E_j^{(1)}|\psi_j^{(1)}\rangle + \frac{1}{2}E_j^{(0)}|\psi_j^{(2)}\rangle \]  \hspace{1cm} (10)

\[ \ldots \]

It’s clear that this is a recursive problem. We can find the first-order quantities \( E_j^{(1)} \) and \( |\psi_j^{(1)}\rangle \) in terms of the known zeroth-order eigenvalues and eigenfunctions. The second-order quantities can be expressed in terms of known zeroth- and first-order quantities, etc.

### 1.2 First-order quantities

Let’s start by considering the first-order quantities by projecting our first-order expression in Eqn. 9 with \( \langle \psi^{(0)} | \) from the left:

\[ \langle \psi_j^{(0)} | \hat{H}^{(0)} |\psi_j^{(1)}\rangle + \langle \psi_j^{(0)} | \hat{H}^{(1)} |\psi_j^{(0)}\rangle = \langle \psi_j^{(0)} | E_j^{(0)} |\psi_j^{(1)}\rangle + \langle \psi_j^{(0)} | E_j^{(1)} |\psi_j^{(0)}\rangle \]  \hspace{1cm} (11)

\[ = E_j^{(0)}\langle \psi_j^{(0)} | \psi_j^{(1)}\rangle + \langle \psi_j^{(0)} | \hat{H}^{(1)} |\psi_j^{(0)}\rangle = E_j^{(0)}\langle \psi_j^{(0)} | \psi_j^{(1)}\rangle + E_j^{(1)}\langle \psi_j^{(0)} | \psi_j^{(0)}\rangle \]  \hspace{1cm} (12)

\[ \rightarrow E_j^{(1)} = \langle \psi_j^{(0)} | \hat{H}^{(1)} |\psi_j^{(0)}\rangle \]  \hspace{1cm} (13)

Note the identical terms which we cancel on the left and right-hand sides.

The diagonal elements of \( \hat{H}^{(1)} \) therefore give us the first-order corrections to the energy \( E_j^{(1)} \).

To solve this problem to first order, there is no need to diagonalize \( \hat{H}^{(1)} \)!

**Aside:** To give a concrete example, if we have a 1D harmonic oscillator (HO) with \( \hat{H}^{(0)} = KE + \frac{1}{2}kx^2 \), then we might consider the quartic oscillator as a perturbation, with \( \hat{H}^{(1)} = \lambda cx^4 \). Provided \( \lambda \) is small, we expect that the \( j^{th} \) new energy level will diverge from the original \( j^{th} \) HO energy level by \( \sim E_j^{(1)} = \lambda c \langle \psi_j^{(0)} | x^4 |\psi_j^{(0)}\rangle \).
What is now left to us is the first-order wavefunctions $|\psi^{(1)}_j\rangle$. Let’s return to our $\lambda^1$ expression from Eqn. 9. Since we now know $E_j^{(1)}$, we only have one remaining unknown:

$$\hat{H}^{(0)}|\psi^{(1)}_j\rangle + \hat{H}^{(1)}|\psi^{(0)}_j\rangle = E_j^{(0)}|\psi^{(1)}_j\rangle + E_j^{(1)}|\psi^{(0)}_j\rangle \quad (14)$$

The first-order correction to the $j^{th}$ wavefunction $|\psi^{(1)}_j\rangle$ can be expressed as an expansion of our zeroth-order wavefunctions (which form a complete, orthonormal basis set):

$$|\psi^{(1)}_j\rangle = \sum_k c^{(1)}_{kj}|\psi^{(0)}_k\rangle \quad (15)$$

**Aside:** Before we go any further, it’s useful to require that our total wavefunction solutions be normalized. In other words:

$$\langle \psi_j(\lambda)|\psi_j(\lambda)\rangle = 1 \quad (16)$$

where $|\psi_j(\lambda)\rangle = |\psi^{(0)}_j\rangle + \lambda|\psi^{(1)}_j\rangle + \frac{1}{2}\lambda^2|\psi^{(2)}_j\rangle + \cdots \quad (17)$

This implies

$$\left[ |\psi^{(0)}_j\rangle + \lambda|\psi^{(1)}_j\rangle + \frac{1}{2}\lambda^2|\psi^{(2)}_j\rangle + \cdots \right] \left[ |\psi^{(0)}_j\rangle + \lambda|\psi^{(1)}_j\rangle + \frac{1}{2}\lambda^2|\psi^{(2)}_j\rangle + \cdots \right] = 1 \quad (18)$$

Let’s use our trusty trick of grouping by powers of $\lambda$. We know that all factors multiplying non-zero powers of $\lambda$ must go to zero if we want the expression above to evaluate to 1 for arbitrary $\lambda$:

$$\lambda^0 : \quad \langle \psi^{(0)}_j|\psi^{(0)}_j\rangle = 1 \quad (19)$$

$$\lambda^1 : \quad \langle \psi^{(0)}_j|\psi^{(1)}_j\rangle + \langle \psi^{(1)}_j|\psi^{(0)}_j\rangle = 0 \quad (20)$$

$$\lambda^2 : \quad \frac{1}{2}\langle \psi^{(0)}_j|\psi^{(2)}_j\rangle + \frac{1}{2}\langle \psi^{(2)}_j|\psi^{(0)}_j\rangle + \langle \psi^{(1)}_j|\psi^{(1)}_j\rangle = 0 \quad (21)$$

Note that in the $\lambda^1$ expression, if the $\langle \psi^{(0)}_0|\psi^{(1)}_0\rangle$ product is real, $|\psi^{(0)}_0\rangle$ and $|\psi^{(1)}_0\rangle$ must be orthogonal. Without loss of generality, we can choose $\langle \psi^{(0)}_0|\psi^{(1)}_0\rangle$ to be real, and therefore take $c^{(1)}_{kk} = 0$. In other words, $|\psi^{(1)}_j\rangle$ has no contribution from the zeroth-order eigenfunction with the same index $j$:

$$|\psi^{(1)}_j\rangle = \sum_{k\neq j} c^{(1)}_{kj}|\psi^{(0)}_k\rangle \quad (22)$$
The other $c^{(1)}_{kj}$ coefficients can be found by projecting Eqn. 14 with $|\psi^{(0)}_k\rangle$:

$$
\langle \psi^{(0)}_k | \hat{H}^{(0)} | \psi^{(1)}_j \rangle + \frac{1}{2} \langle \psi^{(0)}_k | \hat{H}^{(1)} | \psi^{(2)}_j \rangle = \frac{E^{(2)}_j}{2} \langle \psi^{(0)}_j | \psi^{(0)}_j \rangle + \frac{E^{(1)}_j}{2} \langle \psi^{(1)}_j | \psi^{(0)}_j \rangle + \frac{E^{(0)}_j}{2} \langle \psi^{(0)}_j | \psi^{(2)}_j \rangle
\tag{23}
$$

$$
= \langle \psi^{(0)}_k | \hat{H}^{(0)} | \psi^{(0)}_j \rangle + \frac{1}{2} \langle \psi^{(0)}_k | \hat{H}^{(1)} | \psi^{(1)}_j \rangle + \frac{1}{2} \langle \psi^{(0)}_k | \hat{H}^{(1)} | \psi^{(2)}_j \rangle + E^{(1)}_j \langle \psi^{(1)}_j | \psi^{(0)}_j \rangle + \frac{E^{(0)}_j}{2} \langle \psi^{(0)}_j | \psi^{(2)}_j \rangle
\tag{24}
$$

$$
= \sum_{m \neq j} c^{(1)}_{mj} \langle \psi^{(0)}_k | \psi^{(0)}_m \rangle + \langle \psi^{(0)}_k | \hat{H}^{(1)} | \psi^{(1)}_j \rangle + \frac{1}{2} \langle \psi^{(0)}_k | \hat{H}^{(1)} | \psi^{(2)}_j \rangle
\tag{25}
$$

$$
= \langle \psi^{(0)}_k | \hat{H}^{(0)} | \psi^{(0)}_j \rangle + 2 \frac{1}{2} \langle \psi^{(0)}_k | \hat{H}^{(1)} | \psi^{(1)}_j \rangle + \frac{1}{2} \langle \psi^{(0)}_k | \hat{H}^{(1)} | \psi^{(2)}_j \rangle
\tag{26}
$$

$$
\rightarrow c^{(1)}_{kj} \frac{\hat{H}^{(1)}_{kj}}{E^{(0)}_j - E^{(0)}_k}
\tag{27}
$$

So we have found the $k^{th}$ expansion coefficient of $|\psi^{(1)}_j\rangle$.

### 1.3 Second-order quantities

It’s not much more work to find the second-order correction to the energy eigenvalues, $E^{(2)}_j$. This time we will take our $\lambda^2$ expression in Eqn. 10 and project with $|\psi^{(0)}_j\rangle$:

$$
\langle \psi^{(0)}_j | \hat{H}^{(1)} | \psi^{(1)}_j \rangle + \frac{1}{2} \langle \psi^{(0)}_j | \hat{H}^{(0)} | \psi^{(2)}_j \rangle = \frac{E^{(2)}_j}{2} \langle \psi^{(0)}_j | \psi^{(0)}_j \rangle + \frac{E^{(1)}_j}{2} \langle \psi^{(1)}_j | \psi^{(0)}_j \rangle + \frac{E^{(0)}_j}{2} \langle \psi^{(0)}_j | \psi^{(2)}_j \rangle
\tag{28}
$$

$$
= \langle \psi^{(0)}_j | \hat{H}^{(1)} | \psi^{(0)}_j \rangle + \frac{1}{2} \langle \psi^{(0)}_j | \hat{H}^{(0)} | \psi^{(2)}_j \rangle
\tag{29}
$$

$$
\rightarrow E^{(2)}_j = \langle \psi^{(0)}_j | \hat{H}^{(1)} | \psi^{(1)}_j \rangle
\tag{30}
$$

Since the second-order correction to the energy involves the first-order correction to the ground state wavefunction, it also involves all of the zeroth-order excited states:

$$
\frac{1}{2} \lambda^2 E^{(2)}_j = \lambda^2 \langle \psi^{(0)}_j | \hat{H}^{(1)} | \psi^{(1)}_j \rangle
\tag{31}
$$

$$
= \lambda^2 \sum_{k \neq j} c^{(1)}_{kj} \langle \psi^{(0)}_j | \hat{H}^{(1)} | \psi^{(0)}_j \rangle
\tag{32}
$$

$$
= \lambda^2 \sum_{k \neq j} c^{(1)}_{kj} \hat{H}^{(1)}_{jk}
\tag{33}
$$

$$
= \lambda^2 \sum_{k \neq j} \frac{\hat{H}^{(1)}_{jk} \cdot \hat{H}^{(1)}_{kj}}{E^{(0)}_j - E^{(0)}_k}
\tag{34}
$$
1.4 Example: A molecule in a static electric field

Let’s apply what we’ve learned to the example of a molecule in a static electric field. The perturbed Hamiltonian is given by:

\[ \hat{H} = \hat{H}^{(0)} + \xi \cdot \hat{\mu} \]  

where \( \hat{H}^{(0)} \) is the bare molecular Hamiltonian, \( \xi \) is the applied static electric field and \( \hat{\mu} \) is the molecular dipole operator given by

\[ \hat{\mu} = \sum_i q_i \cdot r_i \]  

which describes the spatial distribution of charges in the molecule. Note that here we use hats to represent quantum operators and underlined quantities to represent vectors.

The first order response of the ground state energy to the field is simply:

\[ E_0^{(1)} = \xi \cdot \langle \psi_0^{(0)} | \hat{\mu} | \psi_0^{(0)} \rangle \equiv \xi \cdot \langle \hat{\mu} \rangle_0 \]  

where \( \langle \hat{\mu} \rangle_0 \) is the permanent dipole moment of the molecule in its ground state. This is the Stark shift!

The second order response to the field is

\[ E_0^{(2)} = \xi^\dagger \cdot \left( \sum_{k>0} \frac{E_{0k} \cdot E_{k0}^\dagger}{E_0^{(0)} - E_k^{(0)}} \right) \cdot \xi \equiv \xi^\dagger \alpha \xi \]  

where \( \alpha \) is the polarizability tensor and \( \alpha \xi \) represents an induced dipole moment of the molecule in the field. The magnitude of the induced dipole depends on (i) how low-lying the excited states of the molecule are, which provide small denominators in Eqn. 38, and (ii) the strength of the dipole matrix elements. This is known as the quadratic Stark effect.

2 Degenerate perturbation theory

Above, we examined how perturbation theory may be used to derive first and second-order corrections to the solutions of the zeroth-order Hamiltonian in the case where the eigenfunctions of \( \hat{H}^{(0)} \) are unique. However, for many systems, degenerate eigenstates abound. Consider the angular momentum eigenstates of the rigid rotor for instance, which feature degeneracies of \( 2J + 1 \), where \( J \) is the angular momentum quantum number. Non-degenerate perturbation theory cannot be applied in this case – just consider the denominators of the second-order energy eigenstate and first-order eigenfunction corrections above. We’ll need to do things slightly differently in the degenerate case.

Consider a system in which the ground state of the \( \hat{H}^{(0)} \) Hamiltonian is \( q \)-fold degenerate. If the perturbing term \( \lambda \hat{H}^{(1)} \) completely lifts this degeneracy then there will be \( q \) distinct energy levels. The goal of perturbation theory is to calculate these new energies and the altered wavefunctions associated with the now distinct energy eigenvalues. As before, we can expand the \( j \)th first-order wavefunction in terms of the zeroth-order eigenstates:

\[ |\psi_j^{(1)}\rangle = \sum_k c_{kj}^{(1)} |\psi_k^{(0)}\rangle \]  

5
where the coefficients are given by

\[ c_{kj}^{(1)} = \frac{\hat{H}_{kj}^{(1)}}{E_j^{(0)} - E_k^{(0)}} \]  

(40)

if the ground state is \( q \)-fold degenerate, then

\[ E_1^{(0)} = E_2^{(0)} = \cdots = E_q^{(0)} \]  

(41)

and the \( \{c_{kj}^{(1)}\} \) are will explode for \( j, k \leq q \) as the denominator goes to 0. How do we get out of this predicament? The idea is to construct a new set of basis functions based on particular linear combinations of our reference \( \{|\psi_k^{(0)}\rangle\} \) which diagonalize the \( q \times q \) submatrix block of \( \hat{H}^{(1)} \) with \( j, k \leq q \). Once diagonalized, the \( \hat{H}^{(1)}_{jk} \) matrix elements with \( j \neq k \) are, by definition, zero. The corresponding \( \{c_{kj}^{(1)}\} \) coefficients also vanish as the numerator goes to zero, and our problem is solved. The remainder of the problem may be solved exactly as we treated it in the first-order non-degenerate case.

Let’s call \( |\bar{\psi}_j\rangle \) the first-order wavefunctions which diagonalize \( \hat{H}^{(1)} \) with \( j, k \leq q \). We can generate \( |\bar{\psi}_j\rangle \) as linear combinations of the zeroth-order wavefunctions:

\[ |\bar{\psi}_j\rangle = \sum_i a_{ji} |\psi_i^{(0)}\rangle \]  

(42)

Assuming that all other \( \hat{H}^{(0)} \) eigenstates are non-degenerate, then the reference basis is now given by:

\[ \{ |\bar{\psi}_1\rangle, |\bar{\psi}_2\rangle, \cdots, |\bar{\psi}_q\rangle, |\psi_{q+1}^{(0)}\rangle, |\psi_{q+2}^{(0)}\rangle, \cdots \} \]  

(43)

Conveniently, the diagonal elements of the \( q \times q \) submatrix of \( \hat{H}^{(1)} \) are the first-order energy corrections \( E_j^{(1)} \) for \( j \leq q \):

\[ E_j^{(1)} = \langle \bar{\psi}_j | \hat{H}^{(1)} | \bar{\psi}_j \rangle = \hat{H}_{jj}^{(1)} \quad (j \leq q) \]  

(44)

We just emphasize this conceptual approach and won’t go into further detail here regarding how one actually goes about diagonalizing \( \hat{H}^{(1)} \) here.