Lecture 15

Degenerate perturbation theory

Normally if \( H = H_0 + V \) and

\( H_0 \) is easy to solve it will often be because there are some simplifying symmetries.

If \( SH = HS \)

and \( H |\psi\rangle = E |\psi\rangle \)

then \( HS |\psi\rangle = SH |\psi\rangle = E S |\psi\rangle \)

which means \( S |\psi\rangle \) is also an eigenstate of \( H \) with the same eigenvalue.

In this case non-degenerate perturbation theory is not applicable.
Assume that $H_0$ has $N$ eigenstates with the same eigenvalue. Relabel states for $|\psi_1\rangle \ldots |\psi_n\rangle$ all have eigenvalue $E^0$ and $E^0 \neq E^n$ for all $n > N$.

Without loss of generality, we can choose $|\psi_k\rangle \ 1 \leq k \leq N$ satisfying

$$
\langle \psi_n | \psi_m \rangle = \delta_{nm} \quad n, m \leq N
$$

Let $P_n$ be the orthogonal projection operator on the subspace of the Hilbert space spanned by the $N$ eigenstates with eigenvalue $E^0$. 
\[ P_N = \sum_{n=1}^{N} \langle \psi_n^{\circ} | \psi_n^{\circ} \rangle \]

Note that

\[ P_N = \rho_N = \rho_N^2 \quad P_N H_0 = E^0 P_N \]

Define

\[ Q_0 = I - P_N \]

It follows that

\[ Q_N = Q_N^+ = Q_0 \]

\[ Q_N P_N = P_N Q_0 = 0 \]

\[ P_N + Q_N = I \]

\[ Q_N = \sum_{n=N+1}^{\infty} \langle \psi_n^{\circ} | \psi_n^{\circ} \rangle \]

Using these relations

\[ H = H_0 + V = \]

\[ (P_N + Q_N)(H_0 + V)(P_N + Q_N) \]

\[ E^0 P_N + P_N V P_N + Q_N H_0 Q_N + \]

\[ Q_N V P_N + P_N V Q_N + Q_0 V Q_0 \]
Note that

\[ P_n H_n Q_n = Q_n H_n P_n = E^o P_n Q_n = 0 \]

The strategy is to replace \( H_0 \) and \( V \) by

\[ H_0' = P_n H_0 P_n + Q_n H_0 Q_n \]

\[ V' = P_n V Q_n + Q_n V P_n + Q_n V Q_n \]

Then from the decomposition on the last page

\[ H = H_0 + V = H_0' + V' \]

To use perturbation theory, we need to diagonalize

\[ H_0' \]

The \( Q_n H_0 Q_n \) part is already diagonal

\[ H_0' |\psi_n\rangle = E_n |\psi_n\rangle \quad n > N \]
for \( n \in \mathbb{N} \)

\[
\langle \psi^0_n | H_0 | \psi^0_n \rangle = E^0_n \psi^0_n + \langle \psi^0_n | V | \psi^0_n \rangle
\]

This is an \( N \times N \) Hermitian matrix \( H_{mn} \) that can be diagonalized.

Eigenvalue:

\[
\det \left( E^\prime S_{mn} - H_{mn} \right) = 0
\]

has \( N \) roots \( E_n^\prime \), \( n = 1 \ldots N \)

\[
\Pi_n = \frac{\prod_{m \neq n} (H_{mn}^\prime - E_m^\prime)}{\prod_{m \neq n} (H_{nn}^\prime - E_n^\prime)}
\]

Project in eigenvector \( \psi_n^\prime \) of \( H^\prime \) with eigenvalue \( E_n^\prime \), \( 1 \leq n \leq N \)

Since each \( \langle \psi^0_n | \) is a linear combination of the \( \langle \psi^0_n | \) for \( n \leq N \)
\[ \langle \psi_n | \psi_m \rangle = 0 \quad n \leq N \quad m > N \]

Normally it is enough to stop after \( H_0 \) has been diagonalized. Because it is not diagonal in the original basis, the new eigenstates will not generally have the same eigenvalue - even when two of them have the same eigenvalue

\[ \langle \psi_n' | \psi_m' \rangle = 0 \quad n,m \leq N \]

because \( \psi' \) has a \( Q_0 \) on the right on left and

\[ Q_0 | \psi_n \rangle = 0 \quad n \leq N \]

so there are never any terms where one is dividing by \( \sqrt{(E_n - E_m)} \).
diagonalizing \( H_0 \) is the same problem and diagonalizing \( \Lambda \)

When the eigenvalues are all the same there are many possible choices of \( |\Psi_n> \)

A lot of work can be saved if they are chosen so

\[
\forall |\Psi_n> = \zeta_n |\Psi_n> 
\]

* The problem is to diagonalize \( \rho_{11} \).

Put \( P_n \). The resulting solution has terms of \( n-1 \) powers of the potential
Example: Hydrogen atom with spin

\[ |S_n \rangle = |n \rangle |+ \rangle \]  \hspace{1cm} (1)

or

\[ |S_n \rangle = |n \rangle |\pm \rangle |m_s \rangle \]  \hspace{1cm} (2)

These are 2 possible bases

for an electron interacting with a proton with a
coulomb interaction. There are small corrections

for example the electron sees a proton moving in
a circle – This creates a magnetic field that couples
to the electron’s magnet moment. This results in

a V of the \[ \frac{e^2}{2m_e} \frac{1}{r^3} \hat{\mathbf{L}} \cdot \hat{\mathbf{S}} = V_{Ls}(r) \hat{\mathbf{L}} \cdot \hat{\mathbf{S}} \]
In the absence of this spin-orbit interaction, eigenstates with different \( \mu, m_\epsilon, m_s \) all have the same eigenvalues provided \( n, \ell \) are the same.

Since \( J^2 = (\ell + \hat{s})^2 = \ell^2 + \hat{s}^2 + 2 \ell \cdot \hat{s} \),
we have
\[
\ell \cdot \hat{s} = \frac{1}{2} (J^2 - \ell^2 - \hat{s}^2).
\]

This means that
\[
\langle \text{nu}\mu\epsilon s | \mathbf{V}_0 | \ell \cdot \hat{s} | \text{nu}\mu\epsilon s \rangle = \langle \text{nu}\mu\epsilon s | \mathbf{V}_0 | \text{nu}\mu\epsilon s \rangle \delta_{\ell \mu} \frac{1}{2} (\ell (\ell + 1) - \epsilon (\epsilon + 1) - s (s + 1)).
\]

Notes:

* This choice of basis diagonalizes \( \mathbf{V} \).

* There is still a \( 2\ell + 1 \)fold degeneracy.
state with different $|z_i\rangle$ will have their eigenvalues shifted.

If we use the basis $|n, m, m\rangle$ we would get a non-diagonal matrix in $m, m$.

The problem could still be done but it would be necessary to diagonalize a matrix. Clearly the matrix that does this is

$$ \langle m, m | \sum c_{m, m'} | m, m' \rangle $$

If the atom is put in a magnetic field then
the energy is shifted by

\[ V = -\mu \cdot \vec{B} = -\frac{e\hbar}{mc} \cdot \vec{B} \]

If we choose coordinates so \( \vec{B} \) is in the \( z \) direction then

\[ V = -\frac{e\hbar}{mc} S_z \cdot \vec{B} \]

Unlike the spin orbit interaction, this is diagonal in the basis \( |m_s \rangle \rightarrow |m_s \rangle \) but not in the basis \( |m_s \rangle \rightarrow |m_s \rangle \).

When both interactions appear, then the choice of basis normally depends on which \( V \) is larger, however in both cases the full interaction matrix is not diagonal.
Next we consider the case when $H = H_0 + V(t)$, where here the perturbing interaction may depend explicitly on time.

This can be an explicitly time dependent interaction or a time independent interaction that is turned on at a given time.

To treat this situation we use the interaction picture which is "between" the Schrödinger and Heisenberg pictures.
Recall

Schrodinger Picture

\[ |\psi(t)\rangle = e^{-iHt} |\psi(0)\rangle \]

\[ \langle 0 | (t) \rangle = \langle \psi(t) | 0 | \psi(t) \rangle \]

\[ = \langle \psi(0) | e^{iHt} 0 e^{-iHt} | \psi(0) \rangle \]

Heisenberg picture

\[ 0(t) = e^{iHt} \]

\[ |\psi(0)\rangle \]

\[ \langle 0(t) | = \langle \psi(t) | 0(0) | \psi(t) \rangle \]

\[ = \langle \psi(0) | e^{iHt} 0 e^{-iHt} | \psi(0) \rangle \]

Both pictures give identical observables. For the interaction picture
\[ H = H_0 + V(t) \]

\[ \psi_+(t) = e^{-iH_0 t} \psi_+(t) \]

\[ O_+ \psi_+(t) = e^{-iH_0 t} O(t) e^{iH_0 t} \psi_+(t) \]

In this case

\[ \langle O(t) \psi_+(t) \rangle = \langle \psi_+(t) | O(t) \psi_+(t) \rangle = \]

\[ \langle \psi_+(t) | e^{-iH_0 t} O(t) e^{iH_0 t} \psi_+(t) \rangle \]

\[ = \langle \psi_+(t) | O_+ \psi_+(t) \rangle = \langle \psi_+(t) | O(t) \psi_+(t) \rangle \]

all three pictures give the same probabilities, expectation values and ensemble averages.

\[ \frac{d}{dt} \psi_+(t) = iH_0 \psi_+(t) + iV(t) \psi_+(t) \]

\[ = e^{-iH_0 t} (-iH_0 \psi_+(t)) \]

\[ = e^{(-iH_0 + iH_0) t} \psi_+(t) \]

\[ = e^{-iV(t) \psi_+(t)} \]
This gives
\[ \frac{d\psi_{\pm}(t)}{dt} = -i V_x(t) \psi_{\pm}(t) \]

with initial condition
\[ \psi_{\pm}(0) = \psi_{\pm}(0) = \psi_{\pm}(0) \]

The differential equation and initial condition can be expressed as an integral equation
\[ \psi_{\pm}(t) = \psi_{\pm}(0) - i \int_0^t V_x(t') \psi_{\pm}(t') \, dt' \]

This can be formally solved by iteration:
\[ \psi_{\pm}(t) = \lim_{n \to \infty} \psi_{\pm}^{(n)}(t) \]
\[ \psi_{\pm}^{(n)}(t) = \psi_{\pm}(0) \]
\[ \psi_{\pm}(t) = \psi_{\pm}(0) \]
\[ |\Psi_{I}^{(n)}(t)\rangle = |\Psi_{I}^{0}(t)\rangle - i \int_{0}^{t} V_{I}(t') |\Psi_{I}^{0}(t')\rangle \, dt' \]

In order to discuss the convergence first consider the second order term

\[ |\Psi_{I}^{(2)}(t)\rangle = |\Psi_{I}^{0}(t)\rangle - \int_{0}^{t} V_{I}(t') \int_{0}^{t'} V_{I}(t'') |\Psi_{I}^{0}(t'')\rangle \, dt' \, dt'' \]

Remarks

(1) Since \([H_{0}, V_{I}] \neq 0\) in general

\([V_{I}(t'), V_{I}(t'')]=0\)

i.e. the interactions at different times do not generally commute

(2) \(t \geq t'' \geq t'\) has

\[ V_{I}(t'') V_{I}(t') \]

with \(V_{I}(t')\) to the right of \(V_{I}(t'')\)
If we consider 0 ≤ t', t'' < t there are 2 possible orderings, 

\[ t' > t'' \quad \text{and} \quad t'' > t' \]

\[
\int_0^t dt' \int_0^{t''} V_{1}(t') V_{1}(t'') = (t' > t'')
\]

\[
\int_0^t dt'' \int_0^{t'} dt' V_{1}(t') V_{1}(t'') = (t'' > t')
\]

\[
= \frac{1}{2} \int_0^t dt' \int_0^{t''} \left( \Theta(t'' - t') V_{1}(t'') \left| V_{1}(t') \right| - \Theta(t'' - t') V_{1}(t') \left| V_{1}(t'') \right| \right)
\]

In general if there are

N times 0 ≤ t_i ≤ t there are N! possible orderings

\[
T \left( V_{1}(t) \left| V_{1}(t) \right| \right) =
\]

\[
\Theta(t_1 - t) \Theta(t_2 - t) \cdots \Theta(t_N - t) + V_{1}(t_1) V_{2}(t_2) \cdots V_{1}(t_N)
\]

similar contributions from each ordering
This gives the following expression for $|\psi_1(t)\rangle$

$$|\psi_1(t)\rangle = |\psi_5(0)\rangle + \sum_{n=1}^{\infty} \frac{1}{n!} \int_0^t dt_1 \int_0^{t_1} dt_2 \ T(V_2(t_2), V_2(t_2)) \ T(V_3(t_1), V_3(t_1)) |\psi_5(0)\rangle$$

This is called the Dyson expansion. The advantage of writing things this way is that there are not variables in the limit of integration. If $V_5$ is bounded

$$\| T(V_2(t_1), V_3(t_1)) \| = e^{\| H \| t_1} \leq e^{\| V \| t_1}{\epsilon} |V| - \frac{iHt_n}{n!}$$
It follows that

\[ \| \mathbf{v} \| \leq \frac{1}{n!} \int_0^t dt_1 \cdots dt_n \left( V_{\frac{1}{2}}(t_1) \cdots V_{\frac{1}{2}}(t_n) \right) \left| Y_0(t) \right| \]

\[ \leq \frac{i}{n!} t^n \| \mathbf{V} \| \| Y_0 \| = e^{t \| \mathbf{V} \|} < \infty \]

This means that if \( \| \mathbf{V} \| \| Y_0 \| < \infty \) that the series converges strongly.