Chapter 1

Perturbation Theory

In section 1 we will give some descriptions of how and why approximation methods are used in solving problems in quantum mechanics. In section 2, the fundamental formulas of perturbation theory are derived.

1.1 Stationary Perturbation Theory and Its Underlying Assumptions

The description of any physical system is only approximate. It is true, of course, that it is possible to solve some models exactly, but in turn these models only approximately describe real physical systems. For example, the harmonic oscillator describes a vibrating diatomic molecule only when the vibrations are not too violent and anharmonic forces are neglible. In a similar vein, it is possible to find the energy levels of the hydrogen atom after the nonrelativistic Hamiltonian is known, but the nonrelativistic Hamiltonian or Schrodinger equation is only approximate. To take into account the spin of the electron, the Pauli equation or Pauli Hamiltonian must be used to describe the electron. The Dirac equation is a relativistic equation that describes a spin 1/2 particle. To take into account spin, relativistic corrections, and the fact that electromagnetic signals travel between the proton and electron at the speed of light, the Bethe-Salpeter equation can be employed. Even when the electron is traveling at speeds far less than that of light, the electron interacts electromagnetically with itself, causing small but experimentally detectable deviations from the energy level calculated from the Pauli equation or Pauli Hamiltonian. Therefore if a physicist wishes to calculate numbers that can be compared with experimental values, approximations are inevitable.

The most frequently used approxiamation method in quantum mechanics is based upon splitting the Hamiltonian into two parts.

$$H = H_0 + H_1 \tag{1.1}$$

This splitting is done such that H_0 is the Hamiltonian of a known system whose eigenvalue equation can be solved exactly, such as the oscillator or rotator. It is often called the free or interaction free Hamiltonian. H_1 describes some small interaction which perturbs the interaction free system slightly. This means the physical system with the exact Hamiltonian H is *similar* to the sytem described by the free Hamiltonian H_0. To explain what it means that the quantum physical systems with H and H_0 are *similar*, we first consider the rotator of section III.4 in the presence of a magnetic field. This is a situation for which one does not need perturbation theory.

The Hamiltonian of the interaction free system is the rotator Hamiltonian, $H_0 = \frac{1}{2I}J^2$ of (III.2.11). To conjecture the interaction Hamiltonian, H_1, for the interaction with an external magnetic field $\vec{B}$, we follow the usual procedure and start with the energy of the corresponding classical system in a magnetic field and make the transition to the quantum system by replacing the classical observables by their corresponding quantum operators. The magnetic field acts on a magnetic dipole moment $\vec{\mu}$. The energy of a classical system with a dipole magnetic moment $\vec{\mu}$ in a magnetic field $\vec{B}$ is given by

$$E = -\vec{\mu} \cdot \vec{B}. \tag{1.2}$$

The magnetic moment of a rotating or spinning object with angular momentum $\vec{j}$ is in general given by

$$\vec{\mu} = g\frac{q}{2m}\vec{j} \tag{1.3}$$

where q is the electric charge and m the mass of the object. The factor g, called the Landé factor or gyromagnetic ratio, is a parameter which is specific for the physical object. For example, if the physical system is a mass-point of mass m with charge q moving in a circular orbit with orbital angular momentum $\vec{l}$, then the magnetic moment is given by (1.3) with $g = 1$:

$$\vec{\mu} = \frac{q}{2m}\vec{l}. \tag{1.4}$$

If the physical system is a classical relativistic spinning particle with angular momentum $\vec{s}$ about its center of mass and with no intrinsic structure then one can show that $g = 2$ (sect. IX.3a). The electron with $q = -e = -1.602 * 10^{-19}C$ is a nearly structurless relativistic particle and its Landé factor is nearly $g_{electron} \approx 2$ (precisely $g_{electron} = 2(1+\frac{\alpha}{2\pi} - 0.328(\frac{\alpha}{2\pi})^2)$ due to QED corrections), thus

$$\vec{\mu}_{electron} \cong 2(\frac{-e}{2m_e})\vec{s} \tag{1.5}$$

where m_e and $\vec{s}$ are the mass and spin of the electron, respectively. A proton has intrinsic structure and its magnetic moment is given by

$$\vec{\mu}_{proton} \cong (2 + 3.59)\frac{e}{2m_p}\vec{s} \tag{1.6}$$

where m_p is the mass of the proton. Thus the Landé factor of the proton is $g_{proton} = 2+3.59$ where $3.59\frac{e}{2m_p}\vec{s}$, called the *anomalous* magnetic moment. It is due to the intrinsic structure of the proton, which unlike the electron, is not a nearly structureless particle. A neutron has structure but no net charge. Its magnetic moment is therefore only anomalous and is empirically determined to be

$$\mu_{neutron} = (0 - 1.92)\frac{e}{2m_n}\vec{s}. \tag{1.7}$$

In quantum mechanics the magnetic dipole moment operator $\hat{\mu}$ is obtained by replacing the classical angular momenta by the angular momentum operators:

$$\vec{\hat{\mu}}^{op} = g\frac{q}{2m}\vec{\hat{J}} \tag{1.8}$$

where J_i fulfill the algebra of angular momentum (III.2.16). The J_i can be the orbital angular momentum L_i, the spin angular momentum S_i, or the sum of spin and orbital angular momentum. Take for instance the case of a circulating and spinning electron where the magnetic dipole operator is:

$$\vec{\mu}^{op} = \vec{\mu}_{orbital} + \vec{\mu}_{spin} = -\frac{e}{2m_e}(\vec{L} + 2\vec{S}). \tag{1.9}$$

When the physical system is a rotating molecule a magnetic dipole moment is produced by the rotational motion of the whole molecule. These magnetic moments result from the rotation of the positively charged nuclei about the center of mass and the rotation of the negatively charged electron clouds which move with the nuclei. This rotation of the charge distribution results in a magnetic moment which again is proportional to the angular momentum of the molecule and inversely proportional to the mass of the nucleus,

$$\vec{\mu}^{op}_{molecule} = g\frac{q}{2M_{nucleus}}\vec{J} \tag{1.10}$$

The factor g is a measure of the current distribution of the rotating molecule and is determined empirically from the Zeeman splitting of the energy levels. (e.g. for the diatomic molecule H_2, $g = 0.883$)

In the above formulas, we have masses of different orders of magnitude, m_e, $m_p \approx m_n \approx 1836 m_e$, and $M_{nucleus} \approx m_p$. Therefore in quantum mechanics two quantities with units of a magnetic dipole moment are often used:

$$\text{Bohr magneton } \mu_0 \equiv \frac{e\hbar}{2m_e} = 9.272 * 10^{-24} J/Tesla \tag{1.11}$$

$$\text{Nuclear Bohr magneton } \mu_N \equiv \frac{e\hbar}{2m_p} = 5.049 * 10^{-27} J/Tesla \tag{1.12}$$

The magnetic moment of a proton or a nucleus is thus due to the mass ratio three orders of magnitude smaller than the magnetic moments of electrons. Only when the latter are absent, do these small nuclear moments give noticeable effects.

The magnetic dipole moment produced by the rotatational motion of the whole molecule is of the same order of magnitude as that of the magnetic moments of the nuclei. Most molecules in the ground state have no electronic angular momentum. For these molecules the magnetic dipole moment is produced by the nuclei of the molecule or by the rotational motion of the entire molecule and is given by

$$\vec{\mu} = g\mu_N \frac{1}{\hbar}\vec{J} \tag{1.13}$$

which is of the order of magnitude given by the nuclear Bohr magneton (1.12).

If a quantum system with a magnetic moment $\vec{\mu}^{op}$ and a Hamiltonian operator H_0, e.g. $H_0 = \frac{1}{2I}J^2$ for the simple rotator, is placed into a magnetic field $\vec{B}$, the Hamiltonian operator of the quantum rotator in the magnetic field is obtained by correspondence from the classical interaction energy (1.3) as

$$H = H_0 + H_1 = H_0 - \vec{\mu}^{op} \cdot \vec{B}. \tag{1.14}$$

This Hamiltonian is no longer rotationally invariant, which means that

$$[H, J_i] = [\vec{\mu}^{op}, J_i]\vec{B} = g\frac{q}{2m}[J_k, J_i]\vec{B}_k \neq 0. \tag{1.15}$$

The magnetic field $\vec{B}$ distinguishes a direction and destroys the isotropy of space. The interaction Hamiltonian H, due to the magnetic field $\vec{B}$ has broken the rotational symmetry.

In summary, the Hamiltonian that we need to consider for the rotator in an external magnetic field is

$$H = \frac{1}{2I}J^2 - g\frac{q}{2m}\vec{B}\cdot\vec{J}, \tag{1.16}$$

where the free or "unperturbed" Hamiltonian is given by (see also III.2.11)

$$H_0 = \frac{1}{2I}J^2 \tag{1.17}$$

and the perturbation or ineraction Hamiltonian is according to (1.10) and (1.2),

$$H_1 = -g\frac{q}{2m}\vec{B}\cdot\vec{J}. \tag{1.18}$$

For the Hamiltonian (1.16) one does not need perturbation theory if one chooses as the 3-direction (i.e. the arbitrary direction of the J_3 component of angular momentum whose eigenvectors $|j, j_3\rangle$ one uses) the direction of the magnetic field, $\vec{B} = (0, 0, B_3)$. Then (1.16) becomes

$$H = \frac{1}{2I}J^2 - g\frac{q}{2m}B_3 J_3. \tag{1.19}$$

The eigenvectors of this Hamiltonian are already known to us. They are the eigenvectors $|j, j_3\rangle$ of the c.s.c.o. J^2, J_3 found in section III.3. The eigenvalues of the free Hamiltonian H_0 are

$$E_j^0 = \frac{1}{2I}j(j+1) \tag{1.20}$$

The eigenvalues of the exact Hamiltonian $H|j,j_3\rangle = E|j,j_3\rangle$, are immediately obtained by applying H of (1.19) to $|j,j_3\rangle$. The eigenvalues of H depend not only on j like the E^0_j, but also j_3:

$$E = E_{j,j_3} = \frac{1}{2I}j(j+1) + g\frac{q}{2m}B_3 j_3. \tag{1.21}$$

Each eigenvalue E^0_j has a $2j+1$-fold degeneracy. This means to each E^0_j correspond $2j+1$ eigenvectors $|j,j_3\rangle$ or a $2j+1$ dimensional space R^j. The eigenvalues E_{j,j_3} of H are not degenerate; to each E_{j,j_3} corresponds one eigenvector $|j,j_3\rangle$ or the one-dimensional space $R^j_{j_3}$. Therefore to each degenerate energy level of H_0 correspond $(2j+1)$ energy levels of H which merge into E^0_j when the external magnetic field is switched off, i.e. when $B_3 \to 0$.

$$\begin{array}{l} \underline{\qquad} E_{j,j_3=-j} \\ \underline{\qquad} E_{j,j_3=-j+1} \quad \xrightarrow{B_3\to 0} \; E^0_j \underline{\qquad} \\ \vdots \\ \underline{\qquad} E_{j,j_3=j} \end{array} \tag{1.22}$$

The splitting of the energy levels in a magnetic field is called Zeeman splitting (or the Zeeman effect) and it occurs for atoms, molecules and any other rotationally invariant system that is placed into a magnetic field which breaks the symmetry.

If we had used the Hamiltonian (1.16) instead of (1.19), then the eigenvectors of H_0 and J_3 could not have been eigenvectors of H, because for H of (1.16), $[H, J_3] \neq 0$. In order to obtain eigenvectors of (1.16), we would have to use instead of $|j,j_3\rangle$ a basis system $|j,j_n\rangle$ of eigenvectors of the c.s.c.o. $J_k, \vec{n}\cdot\vec{J} \equiv J_n$, where $\vec{n}$ is the unit vector in the direction of the magnetic field and proceed in the same way as above using $|j,j_n\rangle$ in place of $|j,j_3\rangle$. However H and H_0 of (1.19) and (1.17) (or (1.16) and (1.17)) do not represent the general case of the exact and free Hamiltonian, because they have another special relationship with each other, which in general is not fulfilled, namely $[H, H_0] = 0$. For a general quantum system(not the rotator) in a magnetic field with the interaction Hamiltonian (1.18) this is not fulfilled since the free Hamiltonian H_0 is in general not the rotator Hamiltonian (1.16) (or any $H_0 = f(J^2)$).

In the general case the exact and free Hamiltonian do not commute, $[H, H_0] \neq 0$, as is the case for atoms and molecules in a magnetic field, for which H_0 would not be given by a function $f(J^2)$. But H_0 could still be a spherically symmetric Hamiltonian, i.e. fulfill $[H_0, J_i] = 0$. These systems will also have Zeeman splitting, however the eigenvalues E_{j,j_3} cannot be calculated as in (1.21), though the physical situation is very similar. If H and H_0 do not commute one has to use a procedure called perturbation theory.

We now consider this general case. We have two Hamiltonians H_0 and $H = H_0 + H_1$. Except for very special cases (such as the one dimensional harmonic oscillator) one has in addition to the energy quantum number other quantum numbers(such as j_3 in the case of the rotator). Then we have to select other operators, which together with H_0, form a c.s.c.o. and whose eigenvalues are known. We denote this set of operators $B_1, B_2, \ldots, B_{N-1}$ collectively by B and their eigenvalues(quantum numbers) $b_1, b_2, \ldots, b_{N-1}$ we denote by b.

It is of great advantage if the same set of operators B also form a c.s.c.o. with the exact Hamiltonian H, i.e. if $[H, B] = 0$. Then we have two c.s.c.o.'s consisting of N operators

$$H_0, B_1, B_2, \ldots, B_{N-1} \text{ which we abbreviate } H_0, B \tag{1.23}$$
$$H, B_1, B_2, \ldots, B_{N-1} \text{ which we abbreviate } H, B \tag{1.24}$$

If this is not the case, i.e. if one or more of the B_i do not commute with H then one tries to find another system of operators $A_1, A_2, \ldots, A_{N-1} = A$ which commute with H_0 as well as H. This we will discuss below. The two c.s.c.o.'s (1.23) and (1.24) can be treated as if they were c.s.c.o.'s that consists of N=2 operators.

We have the following situation. The physical system described by (1.23) has solutions (i.e. eigenvectors $|E_n^0, b\rangle = |E_n^0, b_1, b_2, ..., b_{N-1}\rangle$ and eigenvalues $E_n^0, b = E_n^0, b_1, b_2, ..., b_{N-1}$ that are completely known. We want to find the eigenvectors $|E, b \succ = |E, b_1, b_2, ..., b_{N-1} \succ$ and eigenvalues $E, b = E, b_1, b_2, ..., b_{N-1}$ of the physical system described by (1.24) and express them in terms of the solutions for the free systems. For the special case that $[H, H_0] = 0$ as in (1.19), one can choose $|E_n, b \succ = |E_n^0, b\rangle$ since the eigenvectors of H are those of H_0. For the case that $[H, H_0] \neq 0$, one must use perturbation theory to calculate the unknown eigenvalues $E = E_n(b)$ and eigenvectors $|E, b \succ$ of H from the known eigenvalues E_n^0 and eigenvectors $|E_n^0, b\rangle$ of H_0.

The situation to which one applies perturbation theory is otherwise very similar to the situation depicted in (1.22). One has energy levels E_n^0 of the interaction free system which are degenerate (except for the very special case that B has only one eigenvalue or the observable B is not part of the system). To each of the energy levels E_n^0 corresponds a finite or infinite number of energy levels $E_{n,b} = E_n(b)$ of the system with interaction H_1, such that for zero interaction (i.e. $H_1 \to 0$ in some manner) all these energy levels, $E_{n,b}$ merge into the same energy value E_n^0.

$$\begin{array}{l} \underline{\qquad} E_{n,1} \\ \underline{\qquad} E_{n,2} \\ \quad\vdots \\ \underline{\qquad} E_{n,b} \end{array} \quad \overset{H_1 \to 0}{\longrightarrow} \; E_n^0 \underline{\qquad} \tag{1.25}$$

However, in distinction to case (1.22), to each $E_{n,b}$ corresponds a vector $|E_n, b \succ$ different from the known eigenvector $|E_n^0, b\rangle$ of the c.s.c.o. (1.23). These vectors have to be calculated under the condition that

$$|E_n, b \succ \to |E_n^0, b\rangle \text{ for } H_1 \to 0 \quad \forall \text{ b=1, 2, \ldots, (finite or infinite)} \tag{1.26}$$

Thus in the typical situation to which perturbation theory applies, there are several(finite or infinite) known eigenvectors $|E_n^0, b\rangle$ belonging to the one known eigenvalue E_n^0 of H_0. They span the finite or infinite dimensional space of eigenvectors H_0 with eigenvalue E_n^0. The perturbation H_1 then splits the energy level into the sublevel $E_{n,b}$ (the degeneracy of the eigenvalue E_n^0 is removed). Often- as is the case of the rotator in the magnetic field- the degeneracy stems from the symmetries in the unperturbed Hamiltonian, H_0. The perturbation Hamiltonian, H_1, then breaks the symmetry, thereby removing the degeneracy[1].

In order for the scheme (1.25) and (1.26) to make physical sense the splitting inside a multiplet (i.e. between the $E_{n,b}$ for a given value of n and different values of b) should be small compared to the splitting between different multiplets (i.e. between $E_{n,b}$ and $E_{n',b}$ with $n' \neq n$). This means the interaction operator must be "small" in a certain mathematical sense (e.g. $|\langle E_n^0, b|H_1|E_{n'}^0, b'\rangle| << |E_{n'}^0 - E_n^0|$). How to precisely formulate the mathematical conditions on the operators H_0 and H such that (1.25) and (1.26) result and such that the perturbation procedure converges is a difficult mathematical problem. Physicists use the procedure to calculate numbers and hope that H_1 has been chosen right so that the perturbation series somehow converges.

Before we describe in the next section how the unknown eigenvalues E and the eigenvectors $|E, b \succ$ of H are determined in terms of the known eigenvalues E_n^0 and known eigenvectors $|E_n^0, b\rangle$ of H_0 by an approximated perturbation series, we want to discuss what to do in the case that $[H_0, B] = 0$ but $[H, B] \neq 0$. In this case we could still use b as labels for the eigenvectors of H, but these eigenvectors of H are not eigenvectors of B and the label b in $|E(b) \succ$ would only be an "approximate" quantum number defined by the limit of (1.26). (This is what is often done in scattering theory for the continuous eigenvalues E of H where (b) are the asymptotic momenta $\vec{p}$ with $\frac{\vec{p}^2}{2m} = E^0 =$ eigenvalue of H_0). For H and H_0 with discrete spectra it is better if one does not use $|E(b) \succ$, but tries to find an operator A, or a system of operators $A_1, A_2, \ldots, A_{N-1}$, which together with H forms a c.s.c.o.

$$[H, A] = 0 \tag{1.27}$$

and which also has the property that

$$[H_0, A] = 0. \tag{1.28}$$

One then obtains a new system of basis vectors of the interaction free system $|E_n^0, a\}$, which are eigenvectors of the c.s.c.o. H_0, A, by a linear transformation from the known free basis vectors $|E_n^0, b\rangle$:

$$|E_n^0, a\} = \sum_b |E_n^0, b\rangle\langle E_n^0, b|E_n^0, a\} \tag{1.29}$$

These basis vectors $|E_n^0, a\}$ are like the $|E_n^0, b\rangle$ eigenvectors of H_0 with eigenvalue E_n^0. The exact eigenvectors $|E, a\}$, which are eigenvectors of the c.s.c.o. H, A are then obtained from the known $|E_n^0, a\}$ by perturbation theory. This means for the case that $[H, B] \neq 0$, perturbation for a degenerate energy spectrum has been reduced by the linear transformation

[1]There is a very special case, e.g. the one-dimensional harmonic oscillator, which is of little practical value, in which there is no degeneracy and there is only one eigenvector $|E_n^0\rangle$ of H_0 for each value of E_n^0. Then no splitting can occur and there is also only one eigenvalue E_n of H, which is shifted slightly from E_n^0. This case is usually discussed in most textbooks under the heading of "nondegenerate perturbation theory."

in (1.29) to the previous case (1.23) and (1.24), which can be treated like a non-degenerate perturbation since the degeneracy labels b are not affected. The transformation matrix, $\langle E_n^0, b|E_n^0, a\}$ depends upon the properties of the operators B and A and not upon H and H_0.

An example of this will be discussed in detail in chapter IX. The c.s.c.o. H_0, B is

$$H_0, L^2, L_3, S^2, S_3 \tag{1.30}$$

where $\vec{L}$ is the orbital angular momentum and $\vec{S}$ is the spin, and H_0 is a function of the orbital observables only and not of the spin. The eigenvectors $|E_n^0, b\rangle$ of (1.30) are the direct product basis vectors

$$|E_n^0, b\rangle = |E_n^0, l, l_3, s, s_3\rangle = |E_n^0, l, l_3\rangle \otimes |s, s_3\rangle \tag{1.31}$$

where $s = \frac{1}{2}$ and $l = 0, 1, 2, \ldots$.

Since the exact Hamiltonian H, of section IX.3, contains an $\vec{L}$-$\vec{S}$ coupling

$$H = H_0 + f\vec{L}\cdot\vec{S} \tag{1.32}$$

(where f may or may not be a function of operators that commute with L_i, and S_i, e.g. $f(\vec{Q}^2)$ where Q_i are the components of the position operator Q) the exact Hamiltonian H will not commute with L_i and S_i:

$$[H, L_i] \neq 0 \tag{1.33}$$

$$[H, S_i] \neq 0 \tag{1.34}$$

However, the exact Hamiltonian will commute with the total angular momentum operator $\vec{J} = \vec{L} + \vec{S}$.

$$[H, J_i] = 0 \tag{1.35}$$

This one would expect on the basis of rotational invariance of the problem but one can also see it explicitly from (1.32) since

$$\vec{L}\cdot\vec{S} = \frac{1}{2}(J^2 - L^2 - S^2) \tag{1.36}$$

and H_0 commutes with J_i. Therefore

$$H_0, J^2, J_3, L^2, S^2 \;\Leftrightarrow\; H_0, A \tag{1.37}$$

as well as

$$H, J^2, J_3, L^2, S^2 \;\Leftrightarrow\; H, A \tag{1.38}$$

are c.s.c.o.'s which differ only in the operators H and H_0. The eigenvectors

$$|E_n^0, a\} = |E_n^0, j, j_3, l, s\} \tag{1.39}$$

of (1.37) are obtained from the direct product vectors (1.31) by (cf. V.2.12),

$$|E_n^0, j, j_3, l, s\} = \sum_{l_3 s_3} |E_n^0, l, l_3, s, s_3\rangle\langle l, l_3, s, s_3|j, j_3\rangle, \tag{1.40}$$

where $\langle l, l_3, s, s_3|j, j_3\rangle$ are the Clebsch Gordon coefficients obtained in section V.2.

There is a special case, not covered above, which requires special consideration. This is the case that $[H, B] \neq 0$ and one cannot find an operator A (or it is not practical to find an operator A) which has the property (1.27) and (1.28) and together with H and H_0 forms a c.s.c.o. In this case one has to apply another procedure usually called "degenerate perturbation theory."

1.2 Stationary Perturbation Series

In this section we derive approximate formulas for the energy eigenvalues and eigenvectors of H in terms of the energy eigenvectors and eigenvalues of H_0. Without loss of generality, we restrict our attention to systems for which only two operators are required to form a complete set of commuting observables. The generalization to systems for which more operators are required is not difficult and will be considered in the examples. The formulas we derive can be used for the two types of problems listed below:

Case (a): A hermitian operator B is known which satisfies

$$[B, H_0] = [B, H] = 0. \tag{1.41}$$

Case (b): The eigenvalues of H_0 are nondegenerate.

In either case we assume that the free Hamiltonian (unperturbed Hamiltonian) H_0 has been solved in a basis $|E_n^0, b\rangle$ satisfying:

$$H_0|E_n^0, b\rangle = E_n^0|E_n^0, b\rangle \tag{1.42}$$

$$B|E_n^0, b\rangle = b|E_n^0, b\rangle. \tag{1.43}$$

Usually the eigenvalues of H_0 depend only on the quantum number n but not on b although there are exceptions. For later convenience, the eigenvectors $|E_n^0, b\rangle$ are normalized as follows:

$$\langle E_{n'}^0, b'|E_n^0, b\rangle = \delta_{n'n}\delta_{b'b}. \tag{1.44}$$

For case (a) we seek eigenvectors $|E_{n,b}, b \succ$ of the "exact" Hamiltonian H

$$H|E_{n,b}, b \succ = E_{n,b}|E_{n,b}, b \succ, \tag{1.45}$$

that satisfy

$$|E_{n,b}, b \succ \to |E_n^0, b\rangle \text{ for } H_1 \to 0 \tag{1.46}$$

The energy eigenvalues E_n^0 are degenerate, i.e. they do not depend upon b, but the eigenvalues of H are functions of b, $E_n(b) = E_{n,b}$.

For case (b) we seek eigenvectors $|E_{n,b}, b \succ$ of the "exact" Hamiltonian H

$$H|E_{n,b}, b \succ = E|E_{n,b}, b \succ, \tag{1.47}$$

that satisfy

$$|E_{n,b}, b \succ \to |E_n^0, b\rangle \text{ for } H_1 \to 0 \tag{1.48}$$

(If H commuted with some other operator, the eigenvectors $|E_{n,b}, b \succ$ would be labeled by additional quantum numbers which are omitted here because they are not of consequence for our discussion.) Because, for this case, the energy eigenvalues of H_0 are nondegenerate, the eigenvector of H_0 that is the best approximation to $|E_{n,b}, b \succ$ is the eigenvector of H_0, $|E_n^0, b\rangle$, with the same value of the quantum number n.

The perturbation calculation is begun by considering the following matrix element:

$$\langle E_{n'}^0, b|H|E_{n,b}, b \succ = \langle E_{n'}^0, b|H_0 + H_1|E_{n,b}, b \succ . \tag{1.49}$$

where $E_n = E_n(b) \equiv E_{n,b}$. The matrix element on the left-hand side of (1.49) is simplified using the fact that $|E_{n,b}, b \succ$ is, by definition, an eigenvector of H with (unknown) eigenvalue $E_{n,b}$. The first matrix element on the right-hand side of (1.49) is evaluated using the hermiticity of H_0 and (1.42). We thus obtain for (1.49)

$$E_{n,b}\langle E_{n'}^0, b|E_{n,b}, b \succ = E_{n'}^0\langle E_{n'}^0, b|E_{n,b}, b \succ + \langle E_{n'}^0, b|H_1|E_{n,b}, b \succ . \tag{1.50}$$

Since the perturbation is assumed to be small, the energy eigenvalues $E_{n,b}$ do not differ by much from the value E_n^0. For the same reason, the eigenvectors $|E_{n,b}, b \succ$ and $|E_n^0, b\rangle$ are in a certain way also close to each other. This in particular means that for $n' = n$ the matrix element $\langle E_{n'=n}^0, b|E_{n,b}, b \succ \approx \langle E_n^0, b|E_n^0, b\rangle = 1$. Therefore we can divide (1.50) by this matrix element and obtain one of the two fundamental equations in stationary perturbation theory:

$$E_{n,b} = E_n^0 + \frac{\langle E_n^0, b|H_1|E_{n,b}, b \succ}{\langle E_n^0, b|E_{n,b}, b \succ} = E_n^0 + \Delta_{n,b}. \tag{1.51}$$

This means that the exact energy values $E_{n,b}$ are equal to the energy value of the interaction free system E_n^0 plus a small perturbation which depends upon the quantum numbers b. $\Delta_{n,b}$ is the energy shift of the unperturbed energy value E_n^0 of the interaction free system and the energy values $E_{n,b}$ of the system with interaction. It is given by the matrix element of the "small" interaction Hamiltonian H_1.

If $n' \neq n$ in (1.50), then the energy difference $(E_{n,b} - E_{n'}^0)$ is not a small shift like $\Delta_{n,b}$ but will for every b be substantially different from zero, $(E_{n,b} - E_{n'}^0) \approx (E_n^0 - E_{n'}^0) \neq 0$.Therefore one can divide (1.50) by $(E_{n,b} - E_{n'}^0)$ obtaining

$$\langle E_{n'}^0, b|E_{n,b}, b \succ = \frac{\langle E_{n'}^0, b|H_1|E_{n,b}, b \succ}{E_{n,b} - E_{n'}^0} \text{ for } n \neq n'. \tag{1.52}$$

To derive the second fundamental equation of stationary perturbation theory, we expand the "exact" eigenvector $|E_{n,b}, b \succ$ of H in terms of the eigenvectors $|E_{n'}^0, b\rangle$ of H_0. The eigenvectors $|E^0_n, b\rangle$ of the c.s.c.o. $H_0, B_1, \ldots, B_{N-1} = H_0, B$ form a complete basis system of the space $\mathcal{H}$ of state vectors. Thus every vector $\phi \,\epsilon\, \mathcal{H}$ can be expanded in terms of this basis system.

$$\phi = \sum_{n',b'} |E^0_{n'}, b'\rangle\langle E^0_{n'}, b'|\phi\rangle \tag{1.53}$$

In particular each eigenvector $\phi = |E_{n,b}, b \succ$ (with given value n, b) of the c.s.c.o. $H, B_1, \ldots, B_{N-1} = H, B$ can be expanded as in (1.53)

$$|E_{n,b}, b \succ = \sum_{n',b'} |E^0_{n'}, b'\rangle\langle E^0_{n'}, b'|E_{n,b}, b \succ = \sum_{n'} |E^0_{n'}, b\rangle\langle E^0_{n'}, b|E_{n,b}, b \succ \tag{1.54}$$

The second equality holds because

$$\langle E^0_{n'}, b|E_{n,b}, b' \succ = \langle E^0_{n'}, b|E_{n,b}, b \succ \text{ for } b' \neq b \tag{1.55}$$

since $|E_{n'}, b'\rangle$ and $|E_{n,b}, b\rangle$ are eigenvectors of the same system of hermitian operators B with different eigenvalues. Therefore in the expansion on the right hand side of (1.53), only vectors $|E^0_n, b\rangle$ with the same value as on the left hand side appear and we can write (1.53) as

$$|E_{n,b}, b \succ = \sum_{n'} |E^0_{n'}, b\rangle\langle E^0_{n'}|E_{n,b}, b \succ = |E^0_n, b\rangle\langle E^0_n, b|E_{n,b}, b \succ + \sum_{n'\neq n} |E^0_{n'}, b\rangle\langle E^0_{n'}, b|E_{n,b}, b\rangle. \tag{1.56}$$

For a given value of n the terms with $n' = n$ and $n' \neq n$ in (1.56) are of quite different magnitude, since the matrix elements $\langle E^0_n, b|E_{n,b}, b \succ \approx \langle E^0_n, b|E^0_n, b\rangle = 1$ for $n' = n$ and $\langle E^0_n, b|E_{n,b}, b \succ \approx \langle E^0_n, b|E^0_n, b\rangle = 0$ for $n' \neq n$. Therefore one writes the eigenvector $|E_{n,b}, b\rangle$ of H as the corresponding eigenvector $|E^0_n, b\rangle$ of H_0 plus a (infinite) sum of terms of much smaller magnitude. To obtain the traditional form we add and subtract $|E^0_n, b\rangle$ to the first term in (1.56) and insert (1.52) in the second term (the finite sum over n') of (1.56). Then (1.56) becomes

$$|E_{n,b}, b \succ = |E^0_n, b\rangle + (\langle E^0_n, b|E_{n,b}, b \succ -1)|E^0_n, b\rangle + \sum_{n'\neq n} |E^0_{n'}, b\rangle \frac{\langle E^0_{n'}, b|H_1|E_{n,b}, b \succ}{E_{n,b} - E^0_{n'}}. \tag{1.57}$$

Except for the first term of the right hand side, the coefficients of the eigenvectors $|E^0_n, b\rangle$ are all of small magnitude. Equations (1.51) and (1.57) are the two fundamental equations in stationary perturbation theory.

Equations (1.51) and (1.57) are exact equations for the energy eigenvalues $E_{n,b}$ and eigenvectors $|E_{n,b}, b \succ$ of H. But since $E_{n,b}$ and $|E_{n,b}, b \succ$ appear on the right-hand sides of these equations, the equations must again be solved for these quantities. We do this by successively inserting the left-hand sides into the right-hand sides. In carrying out this iterative procedure, it is important to keep in mind that H_1 is a small perturbation. Thus, the second term on the right-hand side of (1.51) is small compared with the first term.

Similarly, the second term and the sum over n' on the right-hand side of (1.57) have small coefficients compared with the first term $|E_n^0, b\rangle$.

Thus the zeroth-order approximation for the eigenvalues $E_{n,b}$ and eigenvector $|E_{n,b}, b \succ$, denoted by $E_{n,b}^{(0)}$ and $|E_{n,b}, b\rangle^{(0)}$ respectively, is obtained by setting $H_1 = 0$ in (1.57):

$$E_{n,b} \to E_{n,b}^{(0)} = E_n^0 \tag{1.58}$$

$$|E_{n,b}, b \succ \to |E_{n,b}, b\rangle^{(0)} = |E_n^0, b\rangle. \tag{1.59}$$

Since the terms on the right-hand side of (1.51) and (1.57) that involve E_n or $|E_{n,b}, b \succ$ are small, we obtain the first-order approximations to E_n and $|E_{n,b}, b \succ$ by substituting their zeroth-order approximations (1.58), (1.59) on the right-hand sides of (1.51) and (1.57), respectively. Denoting the first-order approximation for energy by $E_{n,b}^{(1)}$, (1.51) becomes

$$E_{n,b}^{(1)} = E_n^0 + \frac{\langle E_n^0, b|H_1|E_{n,b}, b\rangle^{(0)}}{\langle E_n^0, b|E_{n,b}, b\rangle^{(0)}}. \tag{1.60}$$

Substituting (1.58) and using the orthogonality relation (1.44) for the eigenvectors $|E_n^0, b\rangle$, this becomes

$$E_{n,b}^{(1)} = E_n^0 + \langle E_n^0, b|H_1|E_n^0, b\rangle = E_n^0 + \Delta_{n,b}^{(1)} \tag{1.61}$$

To first order the energy shift, $\Delta_{n,b}^{(1)}$ of the degenerate energy level E_n^0 of the free or unperturbed Hamiltonian equals the matrix element of the perturbation hamiltonian H_1 between the unperturbed, known eigenstates $|E_n^0, b\rangle$.

Denoting the first-order approximation for energy eigenvectors by $|E_{n,b}\rangle^{(1)}$ and using (1.58 and (1.59, (1.57) becomes

$$|E_{n,b}\rangle^{(1)} = |E_n^0, b\rangle + (\langle E_n^0, b|E_{n,b}, b\rangle^{(0)} - 1)|E_{n'}^0, b\rangle + \sum_{n'\neq n} |E_{n'}^0, b\rangle \frac{\langle E_{n'}^0, b|H_1|E_{n,b}, b\rangle^{(0)}}{E_n^{(0)} - E_{n'}^0}. \tag{1.62}$$

Again using equations (1.58),(1.59) for the zeroth-order approximations and the normalization (1.44), the above equation takes the form

$$|E_{n,b}\rangle^{(1)} = |E_n^0, b\rangle + \sum_{n'\neq n} |E_{n'}^0, b\rangle \frac{\langle E_{n'}^0, b|H_1|E_n^0, b\rangle}{E_n^0 - E_{n'}^0}. \tag{1.63}$$

One obtains the second order approximation by inserting first-order approximations for $E_{n,b}$ and $|E_{n,b}, b \succ$ into the right-hand sides of (1.51) and (1.57). The results (See problem 14 and 15) are as follows:

$$E_{n,b}^{(2)} = E_n^0 + \langle E_n^0, b|H_1|E_n^0, b\rangle + \sum_{n'\neq n} \frac{|\langle E_n^0, b|H_1|E_{n'}^0, b\rangle|^2}{E_n^0 - E_{n'}^0} \tag{1.64}$$

$$|E_n, b\rangle^{(2)} = |E_n^0, b\rangle + \sum_{n'\neq n} |E_{n'}^0, b\rangle \frac{\langle E_{n'}^0, b|H_1|E_n^0, b\rangle}{E_n^{(1)} - E_{n'}^0} + \sum_{n'\neq n} \sum_{n''\neq n} |E_{n'}^0, b\rangle \frac{\langle E_{n'}^0, b|H_1|E_{n''}^0, b\rangle}{E_n^{(1)} - E_{n'}^0} \frac{\langle E_{n''}^0, b|H_1|E_n^0\rangle, b}{E_n^0 - E_{n''}^0}. \tag{1.65}$$

Continuing in this fashion, it is possible to obtain higher order approximations for the exact energy eigenvalues and eigenvectors.

Note that the eigenvectors $|E_n, b\rangle^{(1)}$ and $|E_n, b\rangle^{(2)}$ are neither normalized nor orthogonal. However, since the matrix element of H_1 is small, from (1.63) and (1.65) and we see that the deviations from orthonormality are small.

The exact eigenvectors $|E_{n,b}, b \succ$ in (1.57) could have been normalized to unity. Instead it is conventional to choose the normalization

$$\langle E_n^0, b|E_{n,b}, b \succ = 1 \tag{1.66}$$

This also fixes the relative phase between the exact eigenvector $|E_{n,b}, b \succ$ and the free eigenvector $|E_n^0, b\rangle$. With the above choice for normalization, the two fundamental (exact) equations (1.51) and (1.57) become, respectively,

$$E_n = E_{n,b}^0 + \langle E_n^0, b|H_1|E_{n,b}, b \succ \tag{1.67}$$

$$|E_{n,b}, b \succ = |E_n^0, b\rangle + \sum_{n' \neq n} |E_{n'}^0, b\rangle \frac{\langle E_{n'}^0|H_1|E_{n,b}, b \succ}{E_{n,b} - E_{n'}^0} \tag{1.68}$$

If the iteration procedure for solving (1.51) and (1.57) is to make sense, the exact eigenvector $|E_{n,b}, b \succ$ must not differ much from $|E_n^0, b\rangle$. From (1.57) we then conclude that the following condition will have to be fulfilled:

$$|\langle E_{n',b}^0|H_1|E_{n,b}, b \succ | << |E_{n,b} - E_{n'}^0|. \tag{1.69}$$

But even if the above condition is satisfied, it does not always ensure that the lowest order solution is actually an approximation to the exact solution. However, there is a large class of problems for which the approximation does work, including the spectra of atoms and molecules where H_0 is given by the Coulomb interaction. The interaction Hamiltonian H_1 can describe the interaction with a weak electromagnetic field or the interaction between the spin of the electrons and the magnetic field due to the nuclear motion.