13 Stationary Perturbation Theory and the Variational Method

In most cases of interest, there is no exact analytical solution to the Schrödinger equation for the given problem. In other cases, the exact solution can be found, but its structure is so complicated that it lacks transparency or becomes impractical. Nowadays, it is possible to resort to computational methods, which can be very powerful but have the disadvantage of hiding the qualitative behavior of the solution. In such cases, it is useful or even at times necessary to use a method that allows an approximate solution to be found. There is no universal recipe: in each instance the method to be used must be chosen carefully, because although there may be more than one method that can serve similar purposes, one of them may have advantages over the others for some specific reason.

In this chapter we will review two of the most useful approximation methods for time-independent problems: perturbation theory and the variational method, with applications of specific interest.

13.1 Introduction to Time-Independent Perturbation Theory

It is often the case that a problem can be successfully analyzed using the methods of perturbation theory if it is a minor modification of a problem whose analytical solution is known. For example, in the previous chapter we saw that for an atom subject to an external magnetic field, the Hamiltonian is $\widehat{H} = \widehat{H}_0 + \widehat{H}_{\text{mag}}$, where $\widehat{H}_0$ is the Hamiltonian of the atom in the absence of the field and $\widehat{H}_{\text{mag}} = (\mu_0/\hbar)\boldsymbol{B} \cdot \widehat{\boldsymbol{L}}$ represents the interaction between the orbiting electron and the field. In this case, as long as the magnetic field is not excessively high, it will probably only slightly alter the behavior due to $\widehat{H}_0$; therefore we can consider the initial problem as one due to $\widehat{H}_0$, but perturbed by the magnetic interaction. The energy levels and eigenfunctions of $\widehat{H}$ are then expected to be very close to the original ones, with small corrections due to the perturbation. This is precisely the idea behind the perturbation theory, which will be developed below.[1]

We start by writing the total Hamiltonian of the problem as $\widehat{H} = \widehat{H}_0 + \widehat{V}$, where $\widehat{V}$ is the perturbative potential, or the Hamiltonian of the perturbation. As the magnetic example illustrates, $\widehat{V}$ can depend on position as well as on other operators. In any

[1] Quantum perturbation theory has its origins in classical perturbation theory. The prototypical example of the latter is the solar system, in which the orbit of each planet can be studied separately, as if the other planets and other minor elements of the system did not exist. The combined effects of all these neglected bodies can normally be considered as small disturbances, which are to be calculated separately, one by one. It is on this idea that the astronomical perturbation theory was proposed and successfully developed over the last three centuries.

https://doi.org/10.1017/9781009679633.014 Published online by Cambridge University Press

case, in this chapter we will focus on the study of stationary problems, so the equation to be solved is

$$\widehat{H}\Psi_n = (\widehat{H}_0 + \widehat{V})\Psi_n = E_n\Psi_n, \tag{13.1}$$

where both H_0 and V are time-independent. We write the unperturbed equation, whose solution is assumed to be completely known, as

$$\widehat{H}_0\psi_n = E_n^{(0)}\psi_n, \tag{13.2}$$

to distinguish the eigenfunctions ψ_n and the eigenvalues $E_n^{(0)}$ of the original Hamiltonian $\widehat{H}_0$ from the corresponding ones of the total Hamiltonian $\widehat{H}$.

In this section we restrict ourselves to the case of nondegenerate states, for which it is assumed that the wave function of the perturbed state n differs little from that of the corresponding unperturbed state n. Later we will see that this is not necessarily true for degenerate states (i.e., when more than one state corresponds to the same energy level), if the perturbation breaks the degeneracy. So for now, we propose that $\Psi_n = \psi_n + \delta\psi_n$, where $\delta\psi_n$ is the modification of ψ_n due to the perturbation, which we will consider to be "small."

The basic idea is to use the unperturbed eigenfunctions ψ_n as a basis for expanding the unknown function $\delta\psi_n$,

$$\Psi_n = \psi_n + \sum_k C_{nk}\psi_k. \tag{13.3}$$

We substitute this expression in Eq. (13.1), subtract Eq. (13.2) from the result, and write the total (perturbed) energy E_n in the form

$$E_n = E_n^{(0)} + \delta E_n, \tag{13.4}$$

to get

$$\begin{aligned}\sum_k C_{nk}E_k^{(0)}\psi_k + \widehat{V}\psi_n + \sum_k C_{nk}\widehat{V}\psi_k \\ = \sum_k C_{nk}E_n^{(0)}\psi_k + \delta E_n\psi_n + \delta E_n\sum_k C_{nk}\psi_k.\end{aligned} \tag{13.5}$$

To transform this differential problem into an algebraic one we multiply the equation by ψ_l^* and integrate. Writing the matrix elements of V in the usual form $V_{nl} = \langle n|\widehat{V}|l\rangle$, we get

$$\sum_k V_{lk}C_{nk} + V_{ln} - (E_n^{(0)} - E_l^{(0)} + \delta E_n)C_{nl} = \delta E_n\delta_{nl}. \tag{13.6}$$

The index n denotes the state we are studying while l can take any possible value; thus, Eq. (13.6) represents a system of algebraic equations for the coefficients C_{nl} of the expansion (13.3) of the wave function and for the corrections δE_n to the energy. Although in general it is not possible to solve this system exactly (the equations are nonlinear and there may be a very large or even infinite number of them), we can draw some general conclusions as follows.

Taking $l = n$ in the preceding equation, we get

$$\sum_k V_{nk}C_{nk} + V_{nn} - \delta E_n C_{nn} = \delta E_n. \tag{13.7}$$

https://doi.org/10.1017/9781009679633.014 Published online by Cambridge University Press

We now extract from the sum the term with $k = n$ and solve for δE_n, to obtain

$$\delta E_n = V_{nn} + \frac{\sum' V_{nk} C_{nk}}{1 + C_{nn}}, \tag{13.8}$$

where the prime means that the term with $k = n$ is to be excluded from the sum, $\sum' \equiv \sum_{k \neq n}$.

This expression shows that the effects of the perturbation on the energy are manifested both directly through the average potential term V_{nn}, and indirectly through the modification of the wave function represented by the coefficients C_{nk}. As we will see later (e.g., in Eqs. (13.18) and (13.19)), the direct effect on the energy is of first order, while the indirect effect is of at least second order in the perturbation, precisely because of its indirect nature (the coefficients C_{nk} generally depend on the perturbation at all orders).

We now take $l \neq n$ in (13.6) and eliminate δE_n using (13.8), to obtain, after minor rearrangements,

$$\sum_{k \neq n,l} V_{lk} C_{nk} + (1 + C_{nn}) V_{ln} - \left(E_n^{(0)} - E_l^{(0)} + V_{nn} - V_{ll} + \frac{\sum' V_{nk} C_{nk}}{1 + C_{nn}} \right) C_{nl} = 0 \quad (n \neq l). \tag{13.9}$$

This (obviously nonlinear) system determines, in principle, the coefficients C_{nk} for $n \neq k$. To determine C_{nn} we impose the normalization condition on Ψ_n (assuming, of course, that the ψ_n are normalized),

$$\begin{aligned} \langle \Psi_n | \Psi_n \rangle &= \langle \psi_n + \sum_k C_{nk}^* \psi_k | \psi_n + \sum_{k'} C_{nk'} \psi_{k'} \rangle \\ &= 1 + C_{nn} + C_{nn}^* + \sum_k |C_{nk}|^2 = 1, \end{aligned} \tag{13.10}$$

which leads to

$$|C_{nn}|^2 + 2\mathrm{Re} C_{nn} + \sum{}' |C_{nk}|^2 = 0. \tag{13.11}$$

Knowing C_{nk} for $n \neq k$, this equation determines the function $(\mathrm{Re} C_{nn})^2 + 2\mathrm{Re} C_{nn} + (\mathrm{Im} C_{nn})^2$. Since there are no more conditions to satisfy, we can conveniently take $\mathrm{Im} C_{nn} = 0$, that is, C_{nn} real. Thus, the preceding equation gives

$$1 + C_{nn} = \sqrt{1 - \sum{}' |C_{nk}|^2} \tag{13.12}$$

(the negative solution is unacceptable, since by hypothesis $|C_{nn}|$ is small compared to unity). Taken together, Eqs. (13.9) and (13.12) determine all the coefficients C_{kl} of the expansion, whereby the wave functions and (with the help of (13.8)) the energy eigenvalues are determined, in principle. However, the complexity of the previous system of equations, which is exact, makes it necessary in practically all cases to look for an approximate solution. In the following sections we present a systematic way of proceeding in perturbation theory.

A note of caution is in order. There are situations in which perturbation theory is not appropriate. The problem at hand is normally nonlinear, and nonlinear systems

https://doi.org/10.1017/9781009679633.014 Published online by Cambridge University Press

can exhibit unpredictable and unexpected behavior. In such cases perturbation theory may lead to unreliable results. At the very least, it is necessary to check the convergence of your result. Problems arise when even small perturbations of the system lead to large changes in the response, which is typical behavior for nonlinear systems.

13.2 Perturbation Theory for Nondegenerate States

As just mentioned, our aim is to develop a method that allows us to find approximate solutions to Eqs. (13.9) and (13.12). For the sake of clarity, it is convenient to rewrite the perturbing potential in the form $\lambda\widehat{V}$, where the auxiliary parameter λ (which will be made equal to one in all final results) helps us to determine the order of smallness of a given term by simple inspection of the factor λ^r, which indicates that the respective term is of order r. We express both the coefficient C_{nl} and the correction to the energy δE_n as the sum of the contributions of first, second, third, … order, as follows,[2]

$$C_{nl} = \lambda C_{nl}^{(1)} + \lambda^2 C_{nl}^{(2)} + \lambda^3 C_{nl}^{(3)} + \dots \tag{13.13}$$

$$\delta E_n = \lambda \delta E_n^{(1)} + \lambda^2 \delta E_n^{(2)} + \lambda^3 \delta E_n^{(3)} + \dots \tag{13.14}$$

where the factor λ^r $(r = 1, 2, 3, \dots)$ indicates the order of smallness of the term. Substituting in (13.12) we get (summing over k is understood)

$$\begin{aligned} 1 + \lambda C_{nn}^{(1)} + \lambda^2 C_{nn}^{(2)} + \dots &= \left[1 - {\sum}' |\lambda C_{nk}^{(1)} + \lambda^2 C_{nk}^{(2)} + \dots|^2\right]^{1/2} \\ = 1 - \tfrac{1}{2}\lambda^2 {\sum}' |C_{nk}^{(1)}|^2 - \tfrac{1}{2}\lambda^3 {\sum}' (C_{nk}^{(1)*} C_{nk}^{(2)} + C_{nk}^{(1)} C_{nk}^{(2)*}) + \dots \end{aligned} \tag{13.15}$$

Equating the coefficients of each power of λ that appear on either side of this expression, we get the system

$$C_{nn}^{(1)} = 0; \quad C_{nn}^{(2)} = -\tfrac{1}{2} {\sum}' |C_{nk}^{(1)}|^2; \dots \tag{13.16}$$

We see that the first-order correction to C_{nn} is zero, the second-order correction is $-\frac{1}{2}\sum_{n\neq k} |C_{nk}^{(1)}|^2$, and so on.

To obtain the correction to the energy for different orders of approximation, we substitute Eqs. (13.13, 13.14, 13.16) into (13.8),

$$\begin{aligned} \delta E_n &= \lambda \delta E_n^{(1)} + \lambda^2 \delta E_n^{(2)} + \dots \\ &= \lambda V_{nn} + \lambda {\sum_k}' V_{nk} (\lambda C_{nk}^{(1)} + \lambda^2 C_{nk}^{(2)} + \dots)(1 - \lambda^2 C_{nn}^{(2)} + \dots) \\ &= \lambda V_{nn} + \lambda^2 {\sum_k}' C_{nk}^{(1)} V_{nk} + \lambda^3 {\sum}' C_{nk}^{(2)} V_{nk} \\ &\quad + \lambda^4 {\sum_k}' (C_{nk}^{(3)} - C_{nk}^{(1)} C_{nk}^{(2)}) V_{nk} + \dots \end{aligned} \tag{13.17}$$

[2] This is equivalent to assuming that the solutions are analytic functions of the parameter λ, which is not necessarily true. In the latter case it usually means that the perturbative expansion is not convergent.

https://doi.org/10.1017/9781009679633.014 Published online by Cambridge University Press

Equating the coefficients of each power of λ we obtain

$$\delta E_n^{(1)} = V_{nn} = \langle n|\widehat{V}|n\rangle, \tag{13.18}$$

$$\delta E_n^{(2)} = \sum{}' C_{nk}^{(1)} \langle n|\widehat{V}|k\rangle, \tag{13.19}$$

and so on for the higher orders. Formula (13.18) is without doubt the most important result of perturbation theory; it shows that, up to the first order, the correction to the energy of the state n is always given by the expectation value of the perturbing Hamiltonian, calculated with the *unperturbed* wave function of the corresponding state n.

We will analyze formula (13.19) after determining the corrections to the wave functions. To do so, we proceed as before, substituting Eqs. (13.13, 13.14, 13.16) into (13.9),

$$\begin{aligned}\lambda\left[V_{ln} - (E_n^{(0)} - E_l^{(0)})C_{nl}^{(1)}\right] + \lambda^2\Bigg[\sum_{k\neq l,n} C_{nk}^{(1)} V_{lk} - (E_n^{(0)} - E_l^{(0)})C_{nl}^{(2)} \\ -(V_{nn} - V_{ll})C_{nl}^{(1)}\Bigg] + \lambda^3\Bigg[V_{ln}C_{nn}^{(2)} + \sum_{k\neq l,n} C_{nk}^{(2)} V_{lk} - (E_n^{(0)} - E_l)C_{nl}^{(3)} \\ -(V_{nn} - V_{ll})C_{nl}^{(2)} - \sum{}' V_{nk}^{(0)} C_{nk}^{(1)} C_{nl}\Bigg] + \cdots = 0.\end{aligned} \tag{13.20}$$

For this system to be satisfied up to all orders of perturbation theory, the coefficient of each power of λ must again vanish separately. Up to the second order, we get

$$\begin{aligned} V_{ln} - (E_n^{(0)} - E_l^{(0)})C_{nl}^{(1)} = 0, \\ \sum_{k\neq n,l} V_{lk} C_{nk}^{(1)} - (E_n^{(0)} - E_l^{(0)})C_{nl}^{(2)} - (V_{nn} - V_{ll})C_{nl}^{(1)} = 0.\end{aligned} \tag{13.21}$$

If $E_n^{(0)} - E_l^{(0)} \neq 0$ for all l and n, that is, if the level n is not degenerate (or more generally, if there are no pairs of very close energy levels), the previous system can be solved satisfactorily, giving

$$C_{nl}^{(1)} = \frac{V_{ln}}{E_n^{(0)} - E_l^{(0)}}, \tag{13.22}$$

$$\begin{aligned} C_{nl}^{(2)} &= \frac{\sum_{k\neq n,l} V_{lk} C_{nk}^{(1)} - (V_{nn} - V_{ll})C_{nl}^{(1)}}{E_n^{(0)} - E_l^{(0)}} \\ &= \frac{1}{E_n^{(0)} - E_l^{(0)}}\left[\sum_{k\neq n,l} \frac{V_{lk}V_{kn}}{E_n^{(0)} - E_k^{(0)}} + \frac{(V_{ll} - V_{nn})V_{ln}}{E_n^{(0)} - E_l^{(0)}}\right], \end{aligned} \tag{13.23}$$

and so on for higher-order corrections. Perturbation theory is usually only carried up to first or second order, so we will not write final results beyond what is required.

It is important to underline that Eq. (13.22) is inconsistent if there is degeneracy, since in such a case it would be necessary that $\langle l|\widehat{V}|n\rangle = 0$ for arbitrary $\widehat{V}$, which is obviously not true in general. We will develop the theory for the degenerate case later, but this observation already anticipates that it will be necessary to move to a basis that diagonalizes the perturbing Hamiltonian.

https://doi.org/10.1017/9781009679633.014 Published online by Cambridge University Press

The formulas (13.22) and (13.23) allow us to calculate correctly normalized wave functions up to first and second order, respectively. We can now introduce (13.22) in (13.19) to obtain an explicit formula for the second-order energy correction

$$\delta E_n^{(2)} = \sum_{k \neq n} V_{nk} C_{nk}^{(1)} = \sum_{k \neq n} V_{nk} \frac{V_{kn}}{E_n^{(0)} - E_k^{(0)}}, \tag{13.24}$$

which can be rewritten as follows, assuming that $\widehat{V}$ is Hermitian,

$$\delta E_n^{(2)} = \sum_{k \neq n} \frac{|V_{nk}|^2}{E_n^{(0)} - E_k^{(0)}}. \tag{13.25}$$

This is also an important formula because it is often the case that the first-order correction is zero and the main effect on the energy levels is then of second order. When n refers to the ground state, each term of (13.25) will be negative, so this correction will be negative; that is, the ground-state energy up to the first-order is always in excess.

To show that this conclusion can be generalized to any ψ real, we write the perturbed wave function in the form $\Psi = \psi G$ and substitute into Eq. (13.1), subtracting the unperturbed equation (13.2) from the result. After dividing by G, we get

$$\psi \delta E = \widehat{V}\psi - \frac{\hbar^2}{2m}\left[\left(\frac{\nabla^2 G}{G}\right)\psi + 2\left(\frac{\nabla G}{G}\right)\cdot\nabla\psi\right]. \tag{13.26}$$

An integration by parts shows that

$$\int \psi^* \frac{\nabla G}{G}\cdot\nabla\psi\, dx = -\int\left[\psi\frac{\nabla G}{G}\cdot\nabla\psi^* + \psi^*\frac{\nabla^2 G}{G}\psi - \psi^*\left(\frac{\nabla G}{G}\right)^2\psi\right]dx. \tag{13.27}$$

If ψ is real (e.g., for s states), the first term on the right-hand side cancels the one on the left-hand side, whence

$$2\int \psi^* \frac{\nabla G}{G}\cdot\nabla\psi\, dx = -\int \psi^*\left[\frac{\nabla^2 G}{G} - \left(\frac{\nabla G}{G}\right)^2\right]\psi\, dx. \tag{13.28}$$

With this result, the expectation value of (13.26) reduces to

$$\delta E = \left\langle V - \frac{\hbar^2}{2m}\left(\frac{\nabla G}{G}\right)^2\right\rangle. \tag{13.29}$$

From Eq. (13.26) it is easy to see that G can be taken real if V is real. Therefore, for ψ real, the first-order correction $\langle V\rangle$ to the energy is always in excess, since the correction up to all orders of perturbation theory,

$$(-\hbar^2/2m)\langle(\nabla G/G)^2\rangle, \tag{13.30}$$

is negative or zero.

Returning to the formula (13.25) for the second-order energy correction, we see that for $E_k^{(0)} > E_n^{(0)}$, the energy of state n decreases, while for $E_k^{(0)} < E_n^{(0)}$, E_n increases with respect to $E_n^{(0)}$; in both cases the levels E_n and E_k tend to move further apart. This phenomenon is known as *level repulsion*.

https://doi.org/10.1017/9781009679633.014 Published online by Cambridge University Press

13.2.1 The harmonic oscillator in a uniform electric field

As an example of application of the previous results, we study the problem of a 1D HO immersed in a uniform and constant electric field applied in the direction of oscillation. The problem is interesting because its perturbative solution is simple and illustrative, and because it can be solved exactly, which allows the results to be compared. In addition, it makes us aware of the inherent advantages and limitations of the perturbative method, which can be summarized schematically as follows.

Suppose that perturbation theory has led us, up to a certain approximation, to a solution of the form $f(x) = 1 - x + \frac{1}{2}x^2$; this is clearly useful information, but from that alone it is impossible to know whether the exact solution is, say, $f_{\text{exac}}(x) = e^{-x}$, or some other function of x with the same first three terms. In other words, the behavior of the system for values of x approaching 1 remains hidden. It is clear that the closed (analytical) form gives much more information (both qualitatively and quantitatively) than its perturbative approximation, although the two may coincide numerically within the range of validity of the approximation, and the latter may be correct and sufficient for many purposes. Such limitations (and possibilities) are inherent in the perturbative method and must be taken into account in its applications.

13.2.1.1 Approximate perturbative solution

The Hamiltonian of the problem is

$$\widehat{H} = \widehat{H}_0 + \widehat{V} = \widehat{H}_0 - e\mathcal{E}x, \tag{13.31}$$

where $\mathcal{E}$ represents the intensity of the applied electric field and $\widehat{H}_0$ is the Hamiltonian of the oscillator of frequency ω. The matrix elements of the perturbing potential are $\langle n|\widehat{V}|m\rangle = -e\mathcal{E}\langle n|x|m\rangle$, where $|n\rangle$ is an eigenvector of $\widehat{H}_0$. We know from Chapter 9 that the only nonzero matrix elements are

$$\langle n|\widehat{V}|n-1\rangle = -\tfrac{1}{\sqrt{2}}e\mathcal{E}\alpha_0\sqrt{n}, \tag{13.32}$$

$$\langle n|\widehat{V}|n+1\rangle = -\tfrac{1}{\sqrt{2}}e\mathcal{E}\alpha_0\sqrt{n+1}, \tag{13.33}$$

where $\alpha_0 = \sqrt{\hbar/m\omega}$. Therefore, up to first order in the perturbation there is no change in the energy,

$$\delta E_n^{(1)} = \langle n|\widehat{V}|n\rangle = 0. \tag{13.34}$$

For the second-order energy correction, we obtain from Eq. (13.25), using (9.33) for $E_n^{(0)}$,

$$\begin{aligned}\delta E_n^{(2)} &= {\sum}' \frac{|V_{nm}|^2}{E_n^{(0)} - E_m^{(0)}} = \frac{1}{h\omega}{\sum}' \frac{|V_{nm}|^2}{n-m} \\ &= \frac{1}{\hbar\omega}\left[|V_{n,n-1}|^2 - |V_{n,n+1}|^2\right] = \frac{e^2\mathcal{E}^2\alpha_0^2}{2\hbar\omega}(n-n-1),\end{aligned} \tag{13.35}$$

therefore,

$$\delta E^{(2)} \equiv \delta E_n^{(2)} = -\frac{e^2\mathcal{E}^2}{2m\omega^2}. \tag{13.36}$$

https://doi.org/10.1017/9781009679633.014 Published online by Cambridge University Press

This result shows that the energy of all states decreases by the same amount $\frac{1}{2}m\omega^2 x_0^2$, where $x_0 = e\mathcal{E}/m\omega^2$; hence up to second order the spectral lines are not modified by the perturbation.

To obtain the first-order corrections to the wave functions we use (13.22),

$$C_{nl}^{(1)} = \frac{V_{l\,n}}{E_n^{(0)} - E_l^{(0)}} = \frac{V_{l\,n}}{\hbar\omega(n-l)}. \tag{13.37}$$

Thus, $C_{n,n-1}^{(1)} = -q\sqrt{n}$, $C_{n,n+1}^{(1)} = q\sqrt{n+1}$, where $q = e\mathcal{E}\alpha_0/\sqrt{2}\hbar\omega = x_0/\sqrt{2}\alpha_0$. Using (13.3), we obtain for the first-order perturbed wave function

$$\Psi_n = \psi_n - q\sqrt{n}\psi_{n-1} + q\sqrt{n+1}\psi_{n+1}, \quad q = \frac{x_0}{\sqrt{2}\alpha_0}. \tag{13.38}$$

Exercise E13.1. Comparison with the exact solution for the harmonic oscillator in a constant electric field

Find the exact solution for the oscillator perturbed by a constant electric field and compare with the approximate solution obtained earlier using perturbation theory.

Solution. To solve the problem exactly, we write the Hamiltonian in the form

$$\begin{aligned}\widehat{H} &= \frac{\widehat{p}^2}{2m} + \tfrac{1}{2}m\omega^2 x^2 - e\mathcal{E}x \\ &= \frac{\widehat{p}^2}{2m} + \tfrac{1}{2}m\omega^2\left(x - \frac{e\mathcal{E}}{m\omega^2}\right)^2 - \frac{e^2\mathcal{E}^2}{2m\omega^2},\end{aligned} \tag{13.39}$$

which shows that the effect of the field is to shift the equilibrium position of the oscillator by a distance $x_0 = e\mathcal{E}/m\omega^2$ and to add to all levels a constant negative term to the energy. The exact Schrödinger equation is (we take into account that $e^2\mathcal{E}^2/2m\omega^2 = m\omega^2 x_0^2/2$)

$$E_n\Psi_n = (E_n^{(0)} + \delta E_n)\Psi_n = \widehat{H}_0\Psi_n - \tfrac{1}{2}m\omega^2 x_0^2\Psi_n, \tag{13.40}$$

where the relation between the (exact) perturbed wave function and the unperturbed one is given by

$$\Psi_n(x) = \psi_n(x - x_0). \tag{13.41}$$

Since the change to the variable $y = x - x_0$ does not modify the energy levels, we have $E_n^{(0)}\Psi_n = \widehat{H}_0\Psi_n$; subtracting this equation from (13.40) we get Eq. (13.36), but here it is obtained as an exact result. Expanding $\psi_n(x - x_0)$ in Taylor series around x and approximating to first order, we get

$$\Psi_n(x) = \psi_n(x - x_0) = \psi_n(x) - x_0\frac{d\psi_n}{dx}. \tag{13.42}$$

Equations (9.61,9.62), based on the property of the Hermite polynomials $H_n' = 2nH_{n-1}$, give for the derivative:

$$\frac{d\psi_n}{dx} = \frac{1}{\sqrt{2}\alpha_0}\left(\sqrt{n}\psi_{n-1} - \sqrt{n+1}\psi_{n+1}\right). \tag{13.43}$$

https://doi.org/10.1017/9781009679633.014 Published online by Cambridge University Press

Substituting this expression into the previous one, we obtain precisely (13.38). The comparison between Eqs. (13.38) and (13.41) clearly illustrates the power and simplicity of the perturbative method, as well as one of its intrinsic limitations.

A general lesson to be drawn from this exercise is that it is useful to pay attention to certain features of the problem that may suggest a simple way of solving it – such as the symmetries, or its relationship to a problem whose solution is known – before embarking on a complicated mathematical procedure that may not give the best result.

13.3 Perturbation Theory for Degenerate States

The exact equations of the theory developed in Section 13.1 and, in particular, Eqs. (13.8), (13.9), and (13.12) are obviously also applicable to the study of degenerate states. It is only in Section 13.2 when making the approximations in the first of the equations (13.14) that we introduced the hypothesis that limits the perturbative analysis to the nondegenerate case. To see this, we rewrite Eq. (13.6) for $n \neq l$, introducing the parameter λ and assuming that the states n and l are degenerate, that is, that $E_l^{(0)} = E_n^{(0)}$ for some l; to first order in the energy and the C_{nk}, we have

$$\lambda(V_{ln} + \sum_k \lambda V_{lk} C_{nk}^{(1)} - \lambda \delta E_n^{(1)} C_{nl}^{(1)}) = 0. \tag{13.44}$$

This condition mixes first- and second-order terms and can only be satisfied if $V_{ln} = 0$, which is not generally true. However, if we allow the coefficients C_{nk} to have a zero-order component, the conflict disappears, as the resulting equation

$$V_{ln} + \sum_k V_{lk} C_{nk}^{(0)} - \delta E_n^{(1)} C_{nl}^{(0)} + \cdots = 0 \tag{13.45}$$

has a solution (all terms explicitly written here are of first order). The presence of zero-order components in the C_{nk} coefficients means that in the case of degenerate states, the different (unperturbed) wave functions belonging to the same energy *become mixed.*

We can look at this behavior from a slightly different angle. If we construct new zero-order wave functions that diagonalize V, then the $V_{n'l'}$ (the prime denoting the new basis) are zero for $n' \neq l'$ and the conflict disappears. Physically, this means that the wave functions modified by the perturbation according to the formulas derived in Section 13.2 are not the original ψ_n (which were determined completely independently of $\widehat{V}$), but certain linear combinations of them, $\varphi_n = \sum A_{nl} \psi_l$, which are eigenfunctions of the *complete* Hamiltonian, that is, they are determined taking into account the perturbing potential. These are the *correct wave functions* for the perturbed system.

In order to avoid confusion, we introduce a slight change in notation. Suppose that the level n is degenerate and that there are g different states with the same energy $E_n^{(0)}$; we will denote these states by n_i, $i = 1, 2, \ldots g$, and focus our attention on them, since the nondegenerate levels will be treated as before, with the method presented in Section 13.2. The exact equation (13.6) is now rewritten for $n = n_i$ and $l = n_j$,

https://doi.org/10.1017/9781009679633.014 Published online by Cambridge University Press

$$\sum_k V_{n_j k} C_{n_i k} + V_{n_j n_i} - \delta E_{n_i} C_{n_i n_j} = \delta E_{n_i} \delta_{n_i n_j}. \tag{13.46}$$

Further, we rewrite the coefficient $C_{n_i k}$ as the sum of a zero-order term (a mixture of the degenerate states) denoted with $a_{n_i k}$, and higher-order terms (containing all powers of λ) denoted now collectively with $B_{n_i k} = B_{n_i k}(\lambda)$,

$$C_{n_i k} = a_{n_i k} + B_{n_i k}. \tag{13.47}$$

Since we wish to mix only degenerate states, we take

$$a_{n_i k} = \begin{cases} a_{n_i k} & \text{if } k \in \{n_i\} \\ 0 & \text{otherwise.} \end{cases} \tag{13.48}$$

So, for example,

$$\sum_k V_{n_j k} C_{n_i k} = \sum_{n_s} V_{n_j n_s} a_{n_i n_s} + \sum_k V_{n_j k} B_{n_i k}. \tag{13.49}$$

We use these expressions to rewrite Eq. (13.46); at the same time, we add and subtract some terms to explicitly write all first-order contributions,

$$\begin{aligned} &\lambda \sum_{n_s} V_{n_j n_s} a_{n_i n_s} + \lambda \sum_k V_{n_j k} B_{n_i k} + \lambda V_{n_j n_i} - \lambda \delta E^{(1)}_{n_i} a_{n_i n_j} \\ &-(\delta E_{n_i} - \lambda \delta E^{(1)}_{n_i}) a_{n_i n_j} - \delta E_{n_i} B_{n_i n_j} = \lambda \delta E^{(1)}_{n_i} \delta_{n_i n_j} + (\delta E_{n_i} - \lambda \delta E^{(1)}_{n_i}) \delta_{n_i n_j}. \end{aligned} \tag{13.50}$$

We can now split this equation into two parts, one of first order and another containing all the remaining terms, and cancel each of them separately. This gives

$$\sum_{n_s} V_{n_j n_s} a_{n_i n_s} + V_{n_j n_i} - \delta E^{(1)}_{n_i} a_{n_i n_j} = \delta E^{(1)}_{n_i} \delta_{n_i n_j}, \tag{13.51}$$

$$\begin{aligned} \lambda \sum_k V_{n_j k} B_{n_i k} - (\delta E_{n_i} - \lambda \delta E^{(1)}_{n_i}) a_{n_i n_j} - \delta E_{n_i} B_{n_i n_j} \\ = (\delta E_{n_i} - \lambda \delta E^{(1)}_{n_i}) \delta_{n_i n_j}. \end{aligned} \tag{13.52}$$

Equation (13.52) determines the coefficients B_{nk} once the a_{nk} are known, that is, it serves to calculate both the corrections to the correct wave functions (i.e., to the appropriate mixtures of degenerate states), as well as the second- or higher-order corrections to the energy. This system is used *after* solving Eq. (13.51) for the correct (zero-order) wave functions and the first-order energy corrections. Since the solution of (13.52) is equivalent to the nondegenerate problem already discussed, we will not analyze it further here.

To solve Eq. (13.51), we rewrite it in terms of new coefficients A defined as

$$A_{n_i n_j} = \delta_{n_i n_j} + a_{n_i n_j}, \tag{13.53}$$

which gives

$$\delta E^{(1)}_{n_i} A_{n_i n_j} - \sum_{n_s} V_{n_j n_s} A_{n_i n_s} = 0. \tag{13.54}$$

This system of equations is a generalization to the degenerate case of Eq. (13.18), which gives the first-order correction to the energy in the nondegenerate case. In fact, if in (13.53) we take $A_{ij} = \delta_{ij}$, that is, if we cancel all the a's (which can only be done if the

https://doi.org/10.1017/9781009679633.014 Published online by Cambridge University Press

state is nondegenerate), Eq. (13.54) reduces for $i = j$ to $\delta E_{n_i}^{(1)} - V_{n_i n_i} = 0$, which is precisely (13.18). An expression for the energy correction in the degenerate case can be obtained by setting $i = j$ in (13.54),

$$\delta E_{n_i}^{(1)} = V_{n_i n_i} + \frac{\sum_{n_s \neq n_i} V_{n_i n_s} A_{n_i n_s}}{A_{n_i n_i}}. \tag{13.55}$$

However, this expression is not useful in practice, since in general the energy corrections are determined first, and these in turn are used to obtain the A coefficients. To do this in a systematic way, we rewrite (13.54) in the equivalent form

$$\sum_{n_s} \left[V_{n_j n_s} - \delta E_{n_i}^{(1)} \delta_{n_j n_s} \right] A_{n_i n_s} = 0. \tag{13.56}$$

For this homogeneous system of coupled linear equations to have a nontrivial solution, its determinant must be zero. Applying this condition, we obtain a *secular equation* of degree g (the notation has been simplified by writing $V_{n_i n_j} = V_{ij}$),

$$\begin{vmatrix} V_{11} - \delta E^{(1)} & V_{12} & \cdots & V_{1g} \\ V_{21} & V_{22} - \delta E^{(1)} & \cdots & V_{2g} \\ \vdots & \vdots & \ddots & \vdots \\ V_{g1} & V_{g2} & \cdots & V_{gg} - \delta E^{(1)} \end{vmatrix} = 0. \tag{13.57}$$

The solutions of this equation are the g values (equal or different) for $\delta E_{n_i}^{(1)}$, which are the first-order corrections to each of the g degenerate states corresponding to the energy $E_n^{(0)}$ (the order of the numbering is entirely arbitrary). By successively substituting each of these values of $\delta E_{n_i}^{(1)}$ into the system of equations (13.56), we obtain the corresponding coefficients of the A's. We then choose, say, $A_{n_i n_j}$, so that the zero-order functions are correctly normalized (the explicit condition is written later). In this way, the zero-order normalized wave functions and the corresponding first-order energy are fully determined. To continue the calculation to higher order in the perturbation, we use the system of equations (13.52), as already mentioned.

The solution of the secular equation is simplified in the frequent case where the perturbation preserves certain symmetries of the original Hamiltonian. In such a case there are operators, which we will call $\widehat{Q}$, that commute with both $\widehat{H}$ and $\widehat{H}_0$, and therefore with $\widehat{V} = \widehat{H} - \widehat{H}_0$. Consequently, the eigenfunctions of the perturbed Hamiltonian $\widehat{H}$ can be chosen as simultaneous eigenfunctions of $\widehat{Q}$, and therefore only wave functions belonging to the *same* eigenvalue q of $\widehat{Q}$ will be mixed. This means that in this basis the secular equation is factorized into as many independent factors as there are different values of q.

We conclude that an effective use of symmetry properties can help to reduce considerably the algebraic work involved in the perturbative solution of a degenerate problem. For example, suppose that a system with spherical symmetry is immersed in a uniform weak electric field oriented along the z-axis; since the symmetry around the z-axis is preserved, $\widehat{L}_z$ commutes with both Hamiltonians $\widehat{H}_0$ and $\widehat{H}$. In this case the perturbation will only mix wave functions corresponding to the same quantum number m, as will be illustrated in Section 13.4.

https://doi.org/10.1017/9781009679633.014 Published online by Cambridge University Press

13.3.1 Meaning of the transformed basis; diagonalization of V

We will now study some properties of the correct wave functions φ_{n_i}, which are, by definition, the new zero-order wave functions. To do this we write

$$\Psi_{n_i} = \psi_{n_i} + \sum_k (a_{n_i k} + B_{n_i k})\psi_k = \psi_{n_i} + \sum_{n_s} a_{n_i n_s}\psi_{n_s} + \sum_k B_{n_i k}\psi_k$$
$$= \sum_{n_s} A_{n_i n_s}\psi_{n_s} + \sum_k B_{n_i k}\psi_k = \varphi_{n_i} + \chi_{n_i}, \tag{13.58}$$

where

$$\varphi_{n_i} = \sum_{n_s} A_{n_i n_s}\psi_{n_s} \tag{13.59}$$

are the correct (zero-order) wave functions and

$$\chi_{n_i} = \sum_k B_{n_i k}\psi_k \tag{13.60}$$

are the corrections to them. Considering ψ_{n_s} and φ_{n_i} as column vectors of g elements, we rewrite Eq. (13.59) in matrix form

$$\varphi = \widehat{A}\psi, \tag{13.61}$$

where $\widehat{A}$ is a square matrix of dimension $g \times g$ with elements $A_{n_i n_j}$. The condition that the φ are properly normalized gives $1 = \langle\varphi|\varphi\rangle = \langle\widehat{A}\psi|\widehat{A}\psi\rangle = \langle\psi|\widehat{A}^\dagger A|\psi\rangle$, therefore

$$\widehat{A}^\dagger\widehat{A} = 1, \tag{13.62}$$

that is, the matrix $\widehat{A}$ that transforms the initial basis ψ into the correct basis φ is unitary (see Problem P13.5). Explicitly, this means that

$$\sum_{n_s} A_{n_i n_s} A^*_{n_j n_s} = \delta_{ij} \quad (\widehat{A}\widehat{A}^\dagger = 1), \tag{13.63}$$

$$\sum_{n_s} A^*_{n_s n_i} A_{n_s n_j} = \delta_{ij} \quad (\widehat{A}^\dagger\widehat{A} = 1). \tag{13.64}$$

The matrix elements of $\widehat{V}$ in the correct basis are

$$\int dx \varphi^*_{n_j}\widehat{V}\varphi_{n_i} = \sum_{n_s, n_{s'}} \int dx A^*_{n_j n_{s'}}\psi^*_{n_{s'}}\widehat{V}A_{n_i n_s}\psi_{n_s}$$
$$= \sum_{n_s, n'_s} A^*_{n_j n_{s'}} A_{n_i n_s} V_{n_{s'} n_s}, \tag{13.65}$$

or, using Eqs. (13.56) and (13.63),

$$\int dx \varphi^*_{n_j}\widehat{V}\varphi_{n_i} = \delta E^{(1)}_{n_i} \sum_{n_{s'}} A^*_{n_j n_{s'}} A_{n_i n_{s'}} = \delta E^{(1)}_{n_i}\delta_{ij}. \tag{13.66}$$

We see that the transformed vector basis (13.61) is exactly the one that diagonalizes the perturbation Hamiltonian, as anticipated in the introductory discussion. This means

https://doi.org/10.1017/9781009679633.014 Published online by Cambridge University Press

that the g wave functions φ_{n_i}, which are eigenfunctions of the original Hamitonian H_0, are also eigenfunctions of $\widehat{H} = \widehat{H}_0 + \widehat{V}$, with eigenvalues $E_n^{(0)} + \delta E_{n_i}^{(1)}$,

$$\widehat{H}\varphi_{n_i} = (E_n^{(0)} + \delta E_{n_i}^{(1)})\varphi_{n_i}. \tag{13.67}$$

Mathematically, the problem we have solved is the diagonalization of the matrix $\widehat{V}$, that is, the construction of the eigenvectors φ_n of $\widehat{V}$. This is a very common problem in QM, and the procedure presented here is useful for dealing with it in general.

13.3.2 Two coupled harmonic oscillators

As an example of application of the theory just presented, we consider the problem of two harmonic oscillators oscillating on perpendicular axes and coupled by a bilinear potential. The problem has the double interest of providing a simple illustration of the perturbative method in the presence of degeneracy, and, at the same time, of accepting an exact solution, which allows us to compare results and get a better feeling of the merits and limitations of the perturbative method for the present more elaborate case.

13.3.2.1 Perturbative solution

We consider two oscillators of equal mass m and natural frequency ω, oscillating along the x and y axes, respectively, and coupled by a potential of the form

$$V = m\omega^2\beta xy, \tag{13.68}$$

where β is a dimensionless coupling parameter that can be positive or negative. The complete Hamiltonian is

$$\widehat{H} = \frac{1}{2m}\left(\widehat{p}_x^2 + \widehat{p}_y^2\right) + \frac{1}{2}m\omega^2(x^2 + y^2) + m\omega^2\beta xy. \tag{13.69}$$

For the unperturbed problem of the two independent oscillators we write the wave function as the product $\psi_{n_1}(x)\psi_{n_2}(y)$; the sum of their energies,

$$E_n^{(0)} = \hbar\omega(n_1 + n_2 + 1), \tag{13.70}$$

depends only on the total quantum number $n = n_1 + n_2$.

Since n_1 can take the values $0, 1, 2, \ldots, n$, while n_2 takes the values $n, n-1, n-2, \ldots, 0$, the degeneracy of the energy level E_n is of order $g = n + 1$. The $n + 1$ unperturbed degenerate states, denoted by $|n_i\rangle$, $i = 1, 2, 3, \ldots, n + 1$, are

$$|n_i\rangle = |i - 1\rangle|n - i + 1\rangle. \tag{13.71}$$

The matrix element $V_{ij} = \langle n_i|V|n_j\rangle$ is therefore proportional to the product $\langle i - 1|x|j - 1\rangle\langle n - i + 1|y|n - j + 1\rangle$. Using Eqs. (9.58) and (9.59) we get

$$V_{i,i+1} = \tfrac{1}{2}\beta\hbar\omega\sqrt{i(n - i + 1)}, \tag{13.72}$$

$$V_{i,i-1} = \tfrac{1}{2}\beta\hbar\omega\sqrt{(i - 1)(n - i + 2)}, \tag{13.73}$$

and all other elements V_{ij} are zero. With this preliminary material, we can apply perturbation theory.

https://doi.org/10.1017/9781009679633.014 Published online by Cambridge University Press

To illustrate, we take the simplest case, $n = 1$. The degenerate states are $(n_1, n_2) = (0, 1)$ and $(1, 0)$, so i can take the values $i = 1, 2$. The corresponding matrix elements (13.72) and (13.73) are

$$V_{12} = V_{21} = \tfrac{1}{2}\beta\hbar\omega \equiv V_0, \tag{13.74}$$

and the secular equation (13.57) reduces to

$$\begin{vmatrix} -\delta E^{(1)} & V_0 \\ V_0 & -\delta E^{(1)} \end{vmatrix} = 0, \tag{13.75}$$

which gives $(\delta E^{(1)})^2 - V_0^2 = 0$ for the first-order corrections to the energy: thus,

$$\delta E^{(1)}_{1,2} = \pm V_0 = \pm\tfrac{1}{2}\beta\hbar\omega. \tag{13.76}$$

To obtain the correct wave functions we first take the case $i = 1$, with $\delta E^{(1)}_1 = V_0$. Substituting into the system of equations (13.56) we get

$$\begin{aligned} -\delta E^{(1)}_1 A_{11} + V_0 A_{12} &= 0, \\ V_0 A_{11} - \delta E^{(1)}_1 A_{12} &= 0. \end{aligned} \tag{13.77}$$

With the value of $\delta E^{(1)}_1$ selected and the wave functions normalized to 1, it follows that $A_{11} = A_{12} = \frac{1}{\sqrt{2}}$, whence, using (13.71), we get

$$|1\rangle = A_{11}|1_1\rangle + A_{12}|1_2\rangle = \tfrac{1}{\sqrt{2}}\left[|0\rangle|1\rangle + |1\rangle|0\rangle\right], \quad \delta E^{(1)}_1 = V_0. \tag{13.78}$$

Analogously for the case $i = 2$, with $\delta E^{(1)}_2 = -V_0$, we obtain $A_{21} = -A_{22} = \frac{1}{\sqrt{2}}$, so that

$$|2\rangle = A_{21}|1_1\rangle + A_{22}|1_2\rangle = \tfrac{1}{\sqrt{2}}\left[|0\rangle|1\rangle - |1\rangle|0\rangle\right], \quad \delta E^{(1)}_2 = -V_0. \tag{13.79}$$

The interaction between the oscillators has broken the energy degeneracy expressed in (13.70), resulting in two energy levels separated by $2V_0$.

The next more complex case is $n = 2$, which is triply degenerate, with $i = 1, 2, 3$ corresponding to the unperturbed states $(n_1, n_2) = (0, 2), (1, 1), (2, 0)$, respectively. With the nonzero matrix elements given by

$$V_{12} = V_{21} = V_{23} = V_{32} = \tfrac{1}{\sqrt{2}}\beta\hbar\omega = \sqrt{2}V_0, \tag{13.80}$$

the secular equation is

$$\begin{vmatrix} -\delta E^{(1)} & \sqrt{2}V_0 & 0 \\ \sqrt{2}V_0 & -\delta E^{(1)} & \sqrt{2}V_0 \\ 0 & \sqrt{2}V_0 & -\delta E^{(1)} \end{vmatrix} = 0, \tag{13.81}$$

or explicitly, $\delta E^{(1)}[(\delta E^{(1)})^2 - 4V_0^2] = 0$. From the solutions of this equation it follows that the first-order energy corrections are

$$i = 1, \quad \delta E^{(1)}_1 = -2V_0 = -\beta\hbar\omega, \tag{13.82}$$

$$i = 2, \quad \delta E^{(1)}_2 = 0, \tag{13.83}$$

$$i = 3, \quad \delta E^{(1)}_3 = 2V_0 = \beta\hbar\omega. \tag{13.84}$$

https://doi.org/10.1017/9781009679633.014 Published online by Cambridge University Press

Substituting successively each of these values of $\delta E^{(1)}$ into (13.56) and solving for the coefficients A_{ij} we obtain, after normalizing each state vector to unity, that

$$|1\rangle = \tfrac{1}{2}\Big(|0\rangle|2\rangle - \sqrt{2}|1\rangle|1\rangle + |2\rangle|0\rangle\Big), \quad \delta E_1^{(1)} = -2V_0, \tag{13.85}$$

$$|2\rangle = \tfrac{1}{\sqrt{2}}\Big(|0\rangle|2\rangle - |2\rangle|0\rangle\Big), \quad \delta E_2^{(1)} = 0, \tag{13.86}$$

$$|3\rangle = \tfrac{1}{2}\Big(|0\rangle|2\rangle + \sqrt{2}|1\rangle|1\rangle + |2\rangle|0\rangle\Big), \quad \delta E_3^{(1)} = 2V_0. \tag{13.87}$$

In this case the original energy level has been split into three.

Note that in both cases, the correct wave functions – in addition to being orthonormal – turn out to be either symmetric or antisymmetric with respect to the exchange of oscillators 1 and 2, i.e., they have well-defined parity. This is an important feature, which we might have expected, since the two oscillators are identical and therefore their exchange should have no physical consequences.

Exercise E13.2. Exact solution of the two coupled harmonic oscillators
Solve the problem of two harmonic oscillators of equal mass and frequency, linearly coupled by the potential $V = m\omega^2\beta xy$.

Solution. As shown in Section 9.5, the transformation from x, y to the coordinates of the (independent or decoupled) normal modes of oscillation is the usual way to proceed in the case of two similar classical, coupled harmonic oscillators. It is easy to see that if we take

$$x = \tfrac{1}{\sqrt{2}}(x_1 + x_2), \qquad y = \tfrac{1}{\sqrt{2}}(x_1 - x_2), \tag{13.88}$$

the Hamiltonian (13.69) reduces to

$$\widehat{H} = \frac{\widehat{p}_1^2}{2m} + \frac{\widehat{p}_2^2}{2m} + \tfrac{1}{2}m\omega^2[(1+\beta)x_1^2 + (1-\beta)x_2^2]. \tag{13.89}$$

This is the Hamiltonian of two quasi-particles of mass m, oscillating independently with frequencies

$$\omega_1 = \omega\sqrt{1+\beta}, \qquad \omega_2 = \omega\sqrt{1-\beta}. \tag{13.90}$$

Since we are now dealing with two independent oscillators, the solutions of the Schrödinger equation are of the form

$$\Psi_{ni}(x_1, x_2) = \psi_{i-1}(x_1; \omega_1)\psi_{n-i+1}(x_2; \omega_2), \tag{13.91}$$

and the energy is the sum of the respective energies,

$$\begin{aligned} E_{ni} &= \hbar\Big[\omega_1(i - 1 + \tfrac{1}{2}) + \omega_2(n - i + 1 + \tfrac{1}{2})\Big] \\ &= \hbar\omega\Big[(i - \tfrac{1}{2})\sqrt{1+\beta} + \Big(n - i + \tfrac{3}{2}\Big)\sqrt{1-\beta}\Big] \\ &= E_n^{(0)}\sqrt{1-\beta} + \Big(i - \tfrac{1}{2}\Big)\Big(\sqrt{1+\beta} - \sqrt{1-\beta}\Big)\hbar\omega. \end{aligned} \tag{13.92}$$

https://doi.org/10.1017/9781009679633.014 Published online by Cambridge University Press

We have thus obtained the exact solution of the problem. To compare with the perturbative solution, we expand the above expression in a Taylor series around $\beta = 0$; for small values of β we obtain

$$E_n = \hbar\omega\left[n + 1 - \tfrac{1}{2}\beta(n + 2 - 2i) - \tfrac{1}{2}\beta^2(n + 1) + \dots\right] \tag{13.93}$$

or rearranging,

$$E_{ni} = E_n^{(0)}\left(1 - \tfrac{1}{2}\beta^2 + \dots\right) - \tfrac{1}{2}\beta\hbar\omega(n + 2 - 2i) + \dots \quad (i = 1, 2, \dots, n + 1). \tag{13.94}$$

We see that the first-order corrections (linear in β) go from $-\frac{1}{2}\beta\hbar\omega n$ (for $i = 1$) to $\frac{1}{2}\beta\hbar\omega n$ (for $i = n + 1$), in equal steps of $\beta\hbar\omega$; they completely break the initial degeneracy. On the other hand, the second-order corrections are independent of i, that is, they are the same for all degenerate states. For $n = 1$ ($i = 1, 2$), (13.94) gives

$$\delta E_{1,2}^{(1)} = \pm\tfrac{1}{2}\beta\hbar\omega = \pm V_0, \tag{13.95}$$

while for $n = 2$ ($i = 1, 2, 3$),

$$\delta E_{1,2,3}^{(1)} = -2V_0, 0, 2V_0. \tag{13.96}$$

These results coincide with those provided by perturbation theory, Eqs. (13.76) and (13.82–13.84).

From this analysis it is clear that perturbation theory can only give a satisfactory answer for values of β small enough that the Taylor expansion of the radicals $\sqrt{1-\beta}$ and $\sqrt{1+\beta}$ to first or second order can be considered a good approximation. This limitation is due to the fact that only for $\beta \ll 1$ can the perturbation energy be considered small compared to the other energy terms of the Hamiltonian (in terms of their expectation values). However, it often happens that even when this is not entirely true, the information provided by perturbation theory is qualitatively acceptable, even up to higher orders. We will have a specific example of this situation later when we study the He atom, for which perturbation theory gives good qualitative information, despite the fact that the interaction energy between the electrons is comparable to that of each electron with the atomic nucleus. The basic criterion is that as long as the perturbation does not cause the energy levels to cross or come too close together (i.e., as long as they remain sufficiently differentiated), perturbation theory can give reliable results, at least qualitatively.

13.4 The Stark Effect

The preceding perturbation theory was developed by Schrödinger (and published in the third of his series of four foundational papers on QM) based on the classical perturbative techniques of Lord Rayleigh. For this reason it is known as Rayleigh–Schrödinger perturbation theory. Schrödinger applied it to the study of the Stark

https://doi.org/10.1017/9781009679633.014 Published online by Cambridge University Press

effect, making it the first quantum problem to be successfully solved using perturbative methods with his theory.[3]

In 1913, after a long search, J. Stark[4] observed that the lines of the H atom's Balmer series split into several components in the presence of a uniform electric field. This phenomenon, characteristic of all atoms and known since then as the *Stark effect*, could not be properly explained by Bohr's theory, although it was possible to understand some aspects of it. The explanation given by QM is simple and direct, and the results agree in detail with the observations, so that this phenomenon provided important support for the further development of Schrödinger's theory.

The perturbing potential is written as

$$V = -e\boldsymbol{E}\cdot\boldsymbol{r}, \tag{13.97}$$

where $\boldsymbol{E}$ is the electric field, which for simplicity is taken to be uniform and constant. The energy levels of the atom up to second order in the perturbation are, according to Eqs. (13.18), (13.19), and (13.22),

$$E_n = E_n^{(0)} - e\boldsymbol{E}\cdot\boldsymbol{r}_{nn} + e^2\sum_{k\neq n}\frac{(\boldsymbol{E}\cdot\boldsymbol{r}_{nk})(\boldsymbol{E}\cdot\boldsymbol{r}_{kn})}{E_n^{(0)} - E_k^{(0)}}. \tag{13.98}$$

If n refers to a *nondegenerate* level, the unperturbed wave function ψ_n has well-defined parity because the Hamiltonian of a central problem is invariant to an inversion of the coordinates with respect to the origin, that is, $\widehat{H}_{0P} \equiv \widehat{P}\widehat{H}_0\widehat{P}^{-1} = \widehat{H}_0$ where $\widehat{P}$ is the parity operator (see Section 8.7.3). Hence it follows that

$$\boldsymbol{r}_{nn} = \langle n|\boldsymbol{r}|n\rangle = \langle n|(-\widehat{P}^{\dagger}\boldsymbol{r}\widehat{P})|n\rangle = -\langle n\widehat{P}|\boldsymbol{r}|\widehat{P}n\rangle = -\langle n|\boldsymbol{r}|n\rangle, \tag{13.99}$$

which implies that $\boldsymbol{r}_{nn} = 0$. Therefore, nondegenerate atomic (or molecular) states, including the ground state, do not show a linear Stark effect (proportional to the electric field) but only a quadratic Stark effect.

The H atom is a special case. As we have seen, for the Coulomb potential (and only for it) the excited energy levels are degenerate with respect to the angular-momentum quantum number l ($l = 0, 1, \ldots, n-1$). Since the n degenerate states have different parity given by $(-1)^l$, the matrix element of $\boldsymbol{r}$ need not vanish. The conclusion is that the excited H atom has a permanent electric dipole and therefore exhibits a first-order (linear) Stark effect.

To calculate the effect of the perturbation on the dipole moment of the atom, we write the perturbed dipole as

$$\boldsymbol{d} = e\int \Psi_n^*\boldsymbol{r}\Psi_n d^3r \tag{13.100}$$

and introduce the expression (13.3) for the perturbed wave function

$$\Psi_n = (1 + C_{nn})\psi_n + \sum_k{}' C_{nk}\psi_k, \tag{13.101}$$

[3] A little earlier, Heisenberg had solved the problem of a nonlinear HO as a first step in matrix mechanics.

[4] The German physicist Johannes Stark (1874–1957) was awarded the Nobel Prize in 1919 for his work on the splitting of atomic spectral lines in the presence of an electric field. The Stark effect is used today in applications as diverse as imaging the firing activity of neurons and measuring the pressure broadening of spectral lines in plasmas.

https://doi.org/10.1017/9781009679633.014 Published online by Cambridge University Press

thus obtaining

$$\boldsymbol{d} = e(1+C_{nn})^2\boldsymbol{r}_{nn} + e(1+C_{nn})\sum_{k\neq n}[C_{nk}\boldsymbol{r}_{nk} + C_{nk}^*\boldsymbol{r}_{kn}] + e\sum_{k\neq n}\sum_{k'\neq n} C_{nk}^* C_{nk'}\boldsymbol{r}_{kk'}. \tag{13.102}$$

Using the perturbative expansions (13.14) and (13.16) and formula (13.22) for the first-order coefficients, we obtain

$$\boldsymbol{d} = e\boldsymbol{r}_{nn} - e^2\sum_{k\neq n}\frac{(\boldsymbol{r}_{kn}\cdot\boldsymbol{E})\,\boldsymbol{r}_{nk} + (\boldsymbol{r}_{nk}\cdot\boldsymbol{E})\,\boldsymbol{r}_{kn}}{E_n^{(0)} - E_k^{(0)}} + \ldots \tag{13.103}$$

The first term, which does not depend on the field, represents the permanent (or intrinsic) dipole moment,

$$\boldsymbol{d}_0 = \langle n|e\boldsymbol{r}|n\rangle. \tag{13.104}$$

The remaining terms in (13.103) represent an induced dipole moment that is created by the presence of the field and disappears with it. This suggests introducing the polarizability tensor $\overleftrightarrow{\alpha}$ with components

$$\alpha_{ij} = \frac{\partial d_i}{\partial E_j}, \tag{13.105}$$

which measures the capacity of the system to acquire a dipole moment (i.e., to polarize) in different directions as a result of the action of the external field oriented in a given direction. From Eq. (13.103) we see that to lowest order, the polarizability is

$$\alpha_{ij}^{(0)} \equiv \alpha_{ij}(\boldsymbol{E}=0) = e^2\sum_{k\neq n}\frac{r_{i\,kn}r_{j\,nk} + r_{i\,nk}r_{j\,kn}}{E_k^{(0)} - E_n^{(0)}}. \tag{13.106}$$

The total dipole moment of the system is then

$$\boldsymbol{d} = \boldsymbol{d}_0 + \int_0^{\mathcal{E}} \overleftrightarrow{\alpha}\cdot d\boldsymbol{E} = \boldsymbol{d}_0 + \overleftrightarrow{\alpha}^{(0)}\cdot\boldsymbol{E} + \ldots \tag{13.107}$$

and the interaction energy is

$$V = -\int_0^{\mathcal{E}} \boldsymbol{d}\cdot d\boldsymbol{E} = -\boldsymbol{d}_0\cdot\boldsymbol{E} - \tfrac{1}{2}\boldsymbol{E}\cdot\overleftrightarrow{\alpha}^{(0)}\cdot\boldsymbol{E} + \ldots \tag{13.108}$$

These last two expressions agree with the first-order perturbative expansions (13.103) and (13.98), respectively, and help to better understand the physical meaning of the different terms.

13.5 The Dielectric Constant

In a dielectric medium the electrons do not flow through the material because they are bound to their atoms. Therefore, in response to an applied electric field, they can only produce a local polarization, quantified as the density of (induced) electric dipoles. An average macroscopic measure of the polarizability of a medium is the dielectric constant ϵ, defined by the equation

$$\boldsymbol{P} = \frac{\epsilon - 1}{4\pi}\boldsymbol{E}, \tag{13.109}$$

https://doi.org/10.1017/9781009679633.014 Published online by Cambridge University Press

where $\boldsymbol{P}$ is the polarization density induced by the applied electric field $\boldsymbol{E}$. Therefore, if we know the polarization produced by the field, using Eq. (13.109), we can determine its dielectric constant (or permittivity).

The problem is interesting for a number of reasons. From a physical point of view, it is a question of determining from the microscopic structure of the medium how the polarization, and hence the dielectric constant, varies with frequency, that is, of solving one of the most important problems in the theory of light scattering by a dielectric medium. From a methodological point of view, it is a problem in which it is more important to determine the modification of the wave function due to the perturbation than the modification of the atomic levels. Finally, from the point of view of perturbation theory, this is a case where the perturbation potential is time-dependent, so it is necessary to first eliminate the time dependence in order to apply the theory developed so far. (Chapter 17 develops the theory of time-dependent perturbations more generally.)

Exercise E13.3. Dielectric constant of a transparent material

Use first-order perturbation theory to determine the dielectric constant of a material that has no permanent electric dipole and is transparent to incident radiation.

Solution. We restrict the problem to one dimension, and consider incident radiation of sufficiently long wavelength so that we can assume that its electric vector does not vary appreciably in the dimensions of an atom; in this case we can write the incident electric field in the form $E = \mathcal{E}\cos\omega t$, where $\omega = 2\pi c/\lambda$, and the potential perturbing the atomic electrons becomes

$$V = e_0 \mathcal{E} x \cos\omega t \tag{13.110}$$

(we neglect the effects of the magnetic field). To remove the time dependence, we consider the atomic electrons initially in an unperturbed state, $\psi_n = \varphi_n e^{-i\omega_n t}$ and write the perturbed state as $\psi = \psi_n + \delta\psi$. Substituting these expressions into the time-dependent Schrödinger equation and expanding, we obtain to first order

$$\left(i\hbar\frac{\partial}{\partial t} - \widehat{H}_0\right)\delta\psi = \tfrac{1}{2}e_0\mathcal{E}x\varphi_n(e^{-i(\omega_n-\omega)t} + e^{-i(\omega_n+\omega)t}). \tag{13.111}$$

This result suggests that we write the induced change in the wave function in the form

$$\delta\psi = u_- e^{-i(\omega_n-\omega)t} + u_+ e^{-i(\omega_n+\omega)t}. \tag{13.112}$$

By substituting, we can separate the preceding equation into two time-independent equations,

$$[\hbar(\omega_n \pm \omega) - \widehat{H}_0]u_\pm = \tfrac{1}{2}e_0\mathcal{E}x\varphi_n. \tag{13.113}$$

First we write $u_- = \sum_k a_k\varphi_k$ and substitute, taking into account that $\widehat{H}_0\varphi_k = E_k\varphi_k$; after multiplying the result by φ_l^* and integrating over all space we obtain for the coefficient a_k

$$a_k = -\frac{e_0\mathcal{E}}{2\hbar}\frac{\langle k|x|n\rangle}{\omega_{kn}+\omega}, \quad \omega_{kn} \equiv \omega_k - \omega_n = \frac{E_k - E_n}{\hbar}. \tag{13.114}$$

https://doi.org/10.1017/9781009679633.014 Published online by Cambridge University Press

The second solution is obtained by replacing ω by $-\omega$. Collecting the results and simplifying, we get

$$\psi = e^{-i\omega_n t}\left[\varphi_n - \frac{e_0\mathcal{E}}{\hbar}\sum_k \frac{\langle k|x|n\rangle}{\omega_{kn}^2-\omega^2}\varphi_k(\omega_{kn}\cos\omega t - i\omega\sin\omega t)\right]. \tag{13.115}$$

Note that the perturbation contains the coefficients x_{kn} for $k \neq n$, which represent the dipolar response of the atomic electrons to the field modes of frequency ω_{kn}; these are the modes that induce a polarization of the atoms and determine the refraction index of the medium. The average induced electric dipole moment of the perturbed atoms is $\langle d\rangle = -e_0\int \psi^* x\psi d^3x$. If N is the number of atoms of the dielectric per unit volume, the polarization vector is given by

$$P = N\langle d\rangle = \frac{2Ne_0^2\mathcal{E}}{\hbar}\sum_k \frac{\omega_{kn}|\langle n|x|k\rangle|^2}{\omega_{kn}^2-\omega^2}\cos\omega t. \tag{13.116}$$

Using Eq. (13.109), we obtain for the dielectric constant as a function of the frequency, in the dipole approximation,

$$\epsilon = 1 + \frac{8\pi Ne_0^2}{\hbar}\sum_k \frac{\omega_{kn}}{\omega_{kn}^2-\omega^2}|\langle n|x|k\rangle|^2. \tag{13.117}$$

Note that according to this result, at resonance ($\omega = \omega_{kn}$) the value of the dielectric constant is infinite, which is obviously incorrect. By taking into account the dissipative effect due to the radiation reaction of the electrons (proportional to $\dddot{x}$) – as done, for example, in Section 10.2 – the denominator of Eq. (13.117) acquires a complex contribution $i\gamma(\omega)$, which leads to the expression

$$\sum_k \frac{f_{kn}}{\omega_{kn}^2-\omega^2+i\gamma} = \sum_k f_{kn}\frac{\omega_{kn}^2-\omega^2-i\gamma\,(\omega)}{\left(\omega_{kn}^2-\omega^2\right)^2+\gamma^2\,(\omega)}. \tag{13.118}$$

The real part of this expression now becomes finite for $\omega^2 = \omega_{kn}^2$ and small $\gamma\,(\omega)$ (instead of being infinite); the imaginary term, proportional to $\gamma(\omega)$, describes the losses in the dielectric.

Second, we note that it is customary to write (13.117) in the form

$$\epsilon = 1 + \frac{4\pi Ne_0^2}{m}\sum_k \frac{f_{kn}}{\omega_{kn}^2-\omega^2}, \tag{13.119}$$

where f_{kn} are the so-called oscillator intensities, given by

$$f_{kn} = \frac{2m}{\hbar}\omega_{kn}|\langle n|x|k\rangle|^2. \tag{13.120}$$

In classical theory the oscillator intensities are integers, corresponding to the number of electrons per atom in the eigenfrequency state $\omega_\alpha = \omega_{kn}$ but this is not the case in quantum theory or experiment. In fact, the observation that the oscillator intensities were not integer led to the first doubts about the validity of the classical description.

https://doi.org/10.1017/9781009679633.014 Published online by Cambridge University Press

13.6 *The Canonical Transformation Method

In the previous exposition we developed a specific procedure to carry out the perturbative analysis, which corresponds to the Rayleigh–Schrödinger method. An alternative way of doing a perturbative expansion is to go from the unperturbed to the perturbed problem by means of a unitary transformation. Although here we will illustrate this method as applied to perturbation theory, its use in solving complex problems is much broader, including cases where it is not possible to perform perturbative expansions. A classic example of the latter situation is Bogolyubov's[5] successful use of the method of canonical transformations to solve the superconductivity problem (for an illustration, see Problem P13.13).

We assume that there is a unitary operator $\widehat{S}$ that transforms the unperturbed basis $|n\rangle \equiv |\psi_n\rangle$ into the perturbed one $|\Psi_n\rangle$,

$$|\Psi\rangle = \widehat{S}|\psi\rangle. \tag{13.121}$$

From the Schrödinger equation it follows immediately that $\langle m|\widehat{H}_0 + \lambda\widehat{V}|\Psi_n\rangle = E_n\langle m|\Psi_n\rangle$, that is,

$$\lambda\langle m|\widehat{V}|\Psi_n\rangle = \left(E_n - E_m^{(0)}\right)\langle m|\Psi_n\rangle. \tag{13.122}$$

For $|m\rangle = |n\rangle$ it follows that

$$E_n - E_n^{(0)} = \lambda\frac{\langle n|\widehat{V}|\Psi_n\rangle}{\langle n|\Psi_n\rangle} = \lambda\frac{\langle n|\widehat{V}\widehat{S}|n\rangle}{\langle n|\widehat{S}|n\rangle} = \lambda\sum_k \frac{\langle n|\widehat{V}|k\rangle\langle k|\widehat{S}|n\rangle}{\langle n|\widehat{S}|n\rangle}, \tag{13.123}$$

from where

$$E_n = E_n^{(0)} + \lambda\langle n|\widehat{V}|n\rangle + \lambda{\sum_k}'\langle n|\widehat{V}|k\rangle\frac{\langle k|\widehat{S}|n\rangle}{\langle n|\widehat{S}|n\rangle}. \tag{13.124}$$

For $|m\rangle \neq |n\rangle$, we see from (13.122) that

$$\lambda\sum_{k\neq m}\langle m|\widehat{V}|k\rangle\langle k|\widehat{S}|n\rangle = \left(E_n - E_m^{(0)} - \lambda\langle m|\widehat{V}|m\rangle\right)\langle m|\widehat{S}|n\rangle. \tag{13.125}$$

By iteratively using (13.124) and (13.125) it is possible to determine the corrections to the energy and the matrix elements of $\widehat{S}$. For example, to obtain E_n to first order in λ it is sufficient to write $\widehat{S} = 1$ in (13.124), which gives

$$E_n = E_n^{(0)} + \lambda\langle n|\widehat{V}|n\rangle. \tag{13.126}$$

To go further we write $\widehat{S} = 1 + \lambda\widehat{S}_1 + \ldots$ in (13.125) and get

$$\lambda\langle m|\widehat{V}|n\rangle\langle n|1 + \lambda\widehat{S}_1|n\rangle = (E_n - E_m)\langle m|\lambda\widehat{S}_1|n\rangle, \tag{13.127}$$

that is,

$$\langle m|\widehat{S}_1|n\rangle = \lambda\frac{\langle m|\widehat{V}|n\rangle}{E_n - E_m}, \tag{13.128}$$

[5] Nicolai Nikolayevich Bogolyubov (1909–1992) was an outstanding Russian mathematician and theoretical physicist who made many first-rate contributions to statistical mechanics, quantum field theory, and the theory of dynamical systems.

https://doi.org/10.1017/9781009679633.014 Published online by Cambridge University Press

which gives a second-order correction to the energy, and so on. Note that in the denominator we write $E_n - E_m^{(0)} - \lambda\langle m|\widehat{V}|m\rangle$ as $E_n - E_m$, using (13.126). To first order in λ it would have been sufficient to write $E_n^{(0)} - E_m^{(0)}$, as in the case of the Rayleigh–Schrödinger method, to obtain an expansion in powers of λ.

13.7 *The Feynman and Hellman Method

An alternative method can be obtained from a generalization of the *Feynman–Hellman formula*, which applies to a perturbative Hamiltonian of the form $\widehat{H} = \widehat{H}_0 + \lambda\widehat{V}$. This formula states that if $\widehat{F}(\lambda)$ is a Hermitian operator depending on a parameter λ, and $|\psi(\lambda)\rangle$, $f(\lambda)$ are its eigenvectors and eigenvalues, then

$$\langle\psi(\lambda)|\frac{\partial\widehat{F}}{\partial\lambda}|\psi(\lambda)\rangle = \frac{\partial f}{\partial\lambda}. \tag{13.129}$$

We first generalize this expression, to later apply it to the perturbative problem. Let $|\psi_n(\lambda)\rangle, |\psi_m(\lambda)\rangle$ be two eigenvectors of $\widehat{F}(\lambda)$, with eigenvalues f_n, f_m, respectively; the $|\psi\rangle$ are normalized to unity, so the product

$$\langle\psi_n(\lambda)|\psi_m(\lambda)\rangle = \delta_{nm} \tag{13.130}$$

does not depend on λ. Multiplying the eigenvalue equation

$$\widehat{F}(\lambda)|\psi_n(\lambda)\rangle = f_n(\lambda)|\psi_n(\lambda)\rangle \tag{13.131}$$

by $\langle\psi_m(\lambda)|$ and deriving with respect to λ, one gets

$$\frac{\partial\langle\psi_m|}{\partial\lambda}\widehat{F}|\psi_n\rangle + \langle\psi_m|\widehat{F}\frac{\partial|\psi_n\rangle}{\partial\lambda} + \langle\psi_m|\frac{\partial\widehat{F}}{\partial\lambda}|\psi_n\rangle = \frac{\partial f_n}{\partial\lambda}\delta_{nm}. \tag{13.132}$$

The sum of the first two terms can be written as follows,

$$\begin{aligned}\frac{\partial\langle\psi_m|}{\partial\lambda}\widehat{F}|\psi_n\rangle + \langle\psi_m|\widehat{F}\frac{\partial|\psi_n\rangle}{\partial\lambda} &= f_n\frac{\partial\langle\psi_m|}{\partial\lambda}|\psi_n\rangle + f_m\langle\psi_m|\frac{\partial|\psi_n\rangle}{\partial\lambda}\\ = \tfrac{1}{2}(f_n+f_m)\frac{\partial}{\partial\lambda}\langle\psi_m|\psi_n\rangle &+ \tfrac{1}{2}(f_n-f_m)\left[\frac{\partial\langle\psi_m|}{\partial\lambda}|\psi_n\rangle - \langle\psi_m|\frac{\partial|\psi_n\rangle}{\partial\lambda}\right].\end{aligned} \tag{13.133}$$

Since the first term vanishes due to (13.130), we obtain the desired generalization after simplifying,

$$\langle\psi_m|\frac{\partial\widehat{F}}{\partial\lambda}|\psi_n\rangle + \tfrac{1}{2}(f_n - f_m)\left(\frac{\partial\langle\psi_m|}{\partial\lambda}|\psi_n\rangle - \langle\psi_m|\frac{\partial|\psi_n\rangle}{\partial\lambda}\right) = \frac{\partial f_n}{\partial\lambda}\delta_{nm}. \tag{13.134}$$

With $m = n$ we recover the Feynman–Hellman formula.

To apply this result to perturbation theory we take $\widehat{F}(\lambda) = \widehat{H}_0 + \lambda\widehat{V}$, so that $\partial\widehat{F}/\partial\lambda = \widehat{V}$; (13.134) gives then

$$\langle\psi_m|\widehat{V}|\psi_n\rangle + \tfrac{1}{2}(E_n - E_m)\left(\frac{\partial\langle\psi_m|}{\partial\lambda}|\psi_n\rangle - \langle\psi_m|\frac{\partial|\psi_n\rangle}{\partial\lambda}\right) = \frac{\partial E_n}{\partial\lambda}\delta_{nm}. \tag{13.135}$$

In particular, for $n = m$ we get

$$\frac{\partial E_n}{\partial\lambda} = \langle\psi_n|\widehat{V}|\psi_n\rangle, \tag{13.136}$$

https://doi.org/10.1017/9781009679633.014 Published online by Cambridge University Press

while for $m \neq n$,

$$2\langle\psi_m|\widehat{V}|\psi_n\rangle + (E_n - E_m)\left(\frac{\partial\langle\psi_m|}{\partial\lambda}|\psi_n\rangle - \langle\psi_m|\frac{\partial|\psi_n\rangle}{\partial\lambda}\right) = 0. \tag{13.137}$$

These exact expressions are widely used in quantum chemistry, in particular for the calculation of intermolecular forces. The perturbative expansions can be obtained from them in a systematic way, following a procedure similar to that used earlier (compare Eqs. (13.124) and (13.126)).

13.8 The Variational Method: Minimum Energy of the Quantum States

In 1870, while trying to solve the vibration problem of organ pipes, Lord Rayleigh observed that it was possible to determine the frequencies of the normal modes of oscillation of a system by means of a variational principle. In particular, he showed that the fundamental frequency of oscillation of a classical system is a minimum, in the sense that small variations in the amplitudes of the oscillations give rise to second-order changes in frequency. This Rayleigh principle, which is widely used in technical applications, was generalized by W. Ritz[6] in 1900, when he proposed a variational method for the approximate solution of an eigenvalue equation. To recognize the contributions of both scientists, the theory was later named the *Rayleigh–Ritz method*.

The Rayleigh–Ritz (or Ritz) method has been successfully applied in QM to a wide variety of systems, including atoms, molecules, and the scattering of particles and radiation. It is particularly useful in cases where perturbation theory fails because the full Hamiltonian is not close to one with known eigenvalues and eigenstates, such as in the case of polyatomic molecules.

To illustrate the possibilities of this method, we will first use it to determine an upper bound for the energy of the ground state of a quantum system, and then apply it to more complicated cases.

Let $\widehat{H}$ be the Hamiltonian of the system of interest and ψ_k the eigenfunction corresponding to the eigenvalue E_k (with ψ_k and E_k unknown),

$$\widehat{H}\psi_k = E_k\psi_k. \tag{13.138}$$

The index k is chosen such that the energy increases with it, $E_k > E_{k'}$ for $k > k'$; the ground state is denoted by $k = 0$ and its energy is E_0. Let Φ be an arbitrary function normalized to unity, expressed in terms of the basis ψ_k,

$$\Phi = \sum_k a_k\psi_k, \qquad \sum_k |a_k|^2 = 1. \tag{13.139}$$

Using Eq. (13.139), we obtain for the expectation value of H

$$\langle\Phi|\widehat{H}|\Phi\rangle = \sum_{m,n} a_m^* a_n \int \psi_m^* \widehat{H}\psi_n dx = \sum_{m,n} a_m^* a_n E_n \delta_{mn}$$

[6] Walther Ritz (1878–1909) was a Swiss painter, engineer, and theoretical physicist, known for several methods that bear his name. He died of tuberculosis at the age of 31.

https://doi.org/10.1017/9781009679633.014 Published online by Cambridge University Press

$$= \sum_n |a_n|^2 E_n \geq E_0 \sum_n |a_n|^2, \tag{13.140}$$

or, since $E_n \geq E_0$,

$$\langle \Phi | \widehat{H} | \Phi \rangle \geq E_0. \tag{13.141}$$

This shows that the expectation value of $\widehat{H}$ calculated with an arbitrary function Φ (normalized to unity) is greater than, or at best equal to, the ground state energy. In other words, *the eigenfunction corresponding to the ground state is the (normalized) function that minimizes the expectation value.* We can use this result to obtain an upper bound for E_0; it suffices to make Φ depend on one or more parameters and minimize $\langle \Phi | \widehat{H} | \Phi \rangle$ with respect to each one of them. This procedure will give us the optimal value of the parameters and the best possible estimate of the eigenvalue sought, for a wave function of the proposed type.

The quality of the result depends to a large extent on the quality of the test function Φ, so its choice must be made judiciously. In general, one starts with functions that are mathematically easy to handle (analytically or numerically, as the case may be), and that have suitable properties for the given problem. In particular, for the ground state the function should have no nodes (with possible exceptions such as the one noted in Section 15.3, possess the correct parity, decrease exponentially at infinity in the case of bound states, and reflect the symmetries of the problem. If chosen properly, one or two parameters are usually sufficient to obtain reasonable approximations in simple cases.

The method can be extended to the excited states, taking in each case the test function orthogonal to those of the lower states. Thus, if the solutions up to state $n-1$ are known, we write for state n

$$\Phi_n = \sum_{k \geq n} a_k \psi_k. \tag{13.142}$$

With this selection, $\langle \psi_l | \Phi_n \rangle = 0$ for $l = 0, 1, \ldots n-1$, and we readily obtain

$$\langle \Phi_n | \widehat{H} | \Phi_n \rangle = \sum_{k \geq n} |a_k|^2 E_k \geq E_n. \tag{13.143}$$

If only approximate wave functions are known for $k \leq n-1$, the accumulated errors mean that this method can only be used successfully for the first excited states. The treatment of highly excited states can then be done by resorting to the semiclassical (WKB) approximation.

The Ritz variational method can be extended to find approximate solutions to any eigenvalue equation, allowing an infinite number of degrees of freedom of a system to be reduced to a finite number, thus making analysis easier and more practical (see Exercise 13.44).

13.8.1 The electron distribution minimizes the mean energy

Equation $\widehat{H}\psi_k = E_k \psi_k$ implies that the spatial distribution of particles is just that which minimizes the mean energy of the system. This is an extremely interesting property of Schrödinger's theory, which holds in general because it is an eigenvalue problem. To see this, we write the expectation value of the energy

$$E = \langle \psi | \widehat{H} | \psi \rangle, \quad \text{with} \quad \langle \psi | \psi \rangle = 1, \tag{13.144}$$

https://doi.org/10.1017/9781009679633.014 Published online by Cambridge University Press

and introduce the modification $|\psi\rangle \to |\psi\rangle + |\delta\psi\rangle$, $\langle\psi| \to \langle\psi| + \langle\delta\psi|$, where $|\delta\psi\rangle$ and $\langle\delta\psi|$ are arbitrary, but such that the modified function is also normalized; to first order, the change in E is

$$\delta E = \langle\psi|\widehat{H}|\delta\psi\rangle + \langle\delta\psi|\widehat{H}|\psi\rangle, \quad \langle\delta\psi|\psi\rangle + \langle\psi|\delta\psi\rangle = 0. \tag{13.145}$$

The condition of stationarity of E means that $\delta E = 0$. Introducing the normalization condition with the method of undetermined Lagrange multipliers, we obtain

$$\langle\psi|(\widehat{H} - \lambda)|\delta\psi\rangle + \langle\delta\psi|(\widehat{H} - \lambda)|\psi\rangle = 0. \tag{13.146}$$

This condition must be satisfied for any pair $|\delta\psi\rangle$, $\langle\delta\psi|$ and it is met whenever both the Schrödinger equation and its adjoint are satisfied,

$$\widehat{H}|\psi\rangle = \lambda|\psi\rangle, \quad \langle\psi|\widehat{H} = \lambda\langle\psi|, \quad \lambda = \langle\widehat{H}\rangle = E. \tag{13.147}$$

This can be easily seen by choosing, in particular, $|\delta\psi\rangle = \varepsilon(\widehat{H} - \lambda)|\psi\rangle$ with ε infinitesimal and real, which gives $2\varepsilon\langle\psi|(\widehat{H} - \lambda)(\widehat{H} - \lambda)|\psi\rangle = 0$. So the fact that $|\psi\rangle$ is a solution of the Schrödinger eigenvalue equation guarantees that the mean energy at each discrete level has an extreme value (we have already seen that it is actually a minimum) against small arbitrary changes of the corresponding stationary distribution. Alternatively, we can say that the preceding is equivalent to a variational derivation of the Schrödinger equation.

There are two interesting observations here. The first is that nothing in the preceding derivation attaches a specific interpretation to the operator $\widehat{H}$. This means that *all quantum integrals of motion* (which are solutions of an eigenvalue equation) *are extremal*. The second is that the previous result explains why in first-order perturbation theory, the energy correction is simply $\delta E^{(1)} = \langle V\rangle$, where V is the perturbation Hamiltonian. The reason is that, being $E^{(0)}$ extremal with respect to first-order changes in $|\psi\rangle$, the modification of the wave function generated by the perturbation cannot affect $\langle\widehat{H}\rangle$ to first order.

13.8.2 Variational method applied to the H atom

As an illustrative exercise, we will apply the variational method to the H atom. The ground state of the H atom is an s-state ($l = 0$), so its wave function does not depend on either θ or ϕ, that is, $\psi_0 = \psi_0(r)$. Since it must have no nodes and should decay exponentially at infinity, we propose as a test function (with $\alpha \geq 0$)

$$\psi_0 = Ae^{-\alpha r}, \quad A = \left(\frac{\alpha^3}{\pi}\right)^{1/2}, \tag{13.148}$$

where the value of A was determined by normalizing ψ_0. This gives

$$\langle\widehat{H}\rangle = 4\pi\frac{\alpha^3}{\pi}\int_0^\infty \left[-\frac{\hbar^2}{2m}e^{-\alpha r}\nabla^2 e^{-\alpha r} - \frac{e^2}{r}e^{-2\alpha r}\right] r^2 dr = \frac{\hbar^2}{2m}\alpha^2 - e^2\alpha. \tag{13.149}$$

As a function of α this expression has its minimum nonzero value for

$$\alpha = \frac{me^2}{\hbar^2} = \frac{1}{a_0}, \tag{13.150}$$

https://doi.org/10.1017/9781009679633.014 Published online by Cambridge University Press

where a_0 is the Bohr radius, which gives the correct value of the ground-state energy,

$$E_0 \equiv E_{1s} \leq \min\langle \widehat{H} \rangle = \langle \widehat{H}(1/a_0) \rangle = -\frac{e^2}{2a_0}. \tag{13.151}$$

The best approximation for the corresponding wave function is therefore

$$\psi_0 \equiv \psi_{1s} = \frac{1}{\sqrt{\pi a_0^3}} e^{-r/a_0}. \tag{13.152}$$

Due to the simplicity of the problem we were able to find the exact solution in this case; things are usually not that simple.

To gain further experience, we will determine the solution for the first excited s-state (for states with $l \neq 0$ the test function must be of the form $R(r)Y_l^m(\theta, \phi)$, which makes the problem more difficult, but still manageable). We choose the test function such that: (a) it is normalized to unity; (b) it is orthogonal to ψ_{1s}; (c) it has a node on the r-axis for $r > 0$; and (d) it decays exponentially. With these restrictions the simplest possibility is

$$\psi_{2s} = A\left(1 - \beta \frac{r}{a_0}\right) e^{-\alpha r/a_0}. \tag{13.153}$$

Condition (b) leads to $\beta = (1+\alpha)/3$, and from the normalization condition it follows that $A^2 = 3\alpha^5/\pi a_0^3(\alpha^2 - \alpha + 1)$; thus the test function (13.153) depends on a single variational parameter, α. The expectation value of $\widehat{H}$ is

$$\langle \widehat{H} \rangle = -\frac{e^2}{a_0}\left[\frac{\alpha^2}{2(\alpha^2 - \alpha + 1)} + \tfrac{7}{6}\alpha^2 + \tfrac{1}{2}\alpha\right], \tag{13.154}$$

and the condition that it takes a stationary value, $\partial\langle \widehat{H} \rangle/\partial\alpha = 0$, is fulfilled for $\alpha = 1/2$. With this choice we obtain

$$E_{2s} \leq -\frac{e^2}{8a_0}, \tag{13.155}$$

and the corresponding eigenfunction,

$$\psi_{2s} = \frac{1}{\sqrt{8\pi a_0^3}}\left(1 - \frac{r}{2a_0}\right) e^{-r/2a_0}. \tag{13.156}$$

Both results are again exact thanks to the relative simplicity of the problem, which leads to an excellent test function.

So far, the method has allowed us to obtain an upper bound on the eigenvalue, but it is also possible to obtain a lower bound, as follows. Let ψ_n and E_n be the eigenfunction and eigenvalue of the original problem, $\widehat{H}\psi_n = E_n\psi_n$, and ψ' and $E' = \langle \psi' | \widehat{H} | \psi' \rangle$ the respective approximate expressions obtained by applying the variational procedure. We introduce an error vector $|\chi\rangle$ and an estimated bound Δ, defined as

$$|\chi\rangle = \left(\widehat{H} - E'\right)|\psi'\rangle, \quad \Delta = \langle \chi \, | \chi \rangle^{1/2}. \tag{13.157}$$

We have, on the one hand,

$$\begin{aligned} \Delta^2 &= \langle \chi \, | \chi \rangle = \langle \psi' | \left(\widehat{H} - E'\right)^2 | \psi' \rangle, \\ \Delta^2 &= \langle \psi' | \, \widehat{H}^2 \, | \psi' \rangle - E'^2, \end{aligned} \tag{13.158}$$

https://doi.org/10.1017/9781009679633.014 Published online by Cambridge University Press

and on the other hand, inserting a spectral expansion,

$$\begin{aligned}\Delta^2 &= \sum_{k,k'} \langle \psi' | \left(E_k - E'\right)\left(E_{k'} - E'\right) |\psi_k\rangle \langle \psi_k | \psi_{k'}\rangle \langle \psi_{k'} | \psi'\rangle \\ &= \sum_k \left(E_k - E'\right)^2 \langle \psi' | \psi_k\rangle \langle \psi_k | \psi'\rangle \\ &= \sum_k \left(E_k - E'\right)^2 \left|\langle \psi' | \psi_k\rangle\right|^2 . \end{aligned} \tag{13.159}$$

Since E' is a best estimate of E_n, the quantity $\left(E_k - E'\right)^2$ should take a minimum value for $k = n$, so we can write

$$\begin{aligned}\Delta^2 &\geq \left(E_n - E'\right)^2 \sum_k \left|\langle \psi' | \psi_k\rangle\right|^2 \\ &= \left(E_n - E'\right)^2 \sum_k \langle \psi' | \psi_k\rangle \langle \psi_k | \psi'\rangle \\ &= \left(E_n - E'\right)^2 \langle \psi' | \psi'\rangle = \left(E_n - E'\right)^2 . \end{aligned} \tag{13.160}$$

Since $E' \geq E_n$, it follows that

$$\Delta \geq E' - E_n \quad \text{and} \quad \Delta \geq -\left(E' - E_n\right). \tag{13.161}$$

The second inequality is equivalent to $-\Delta \leq E' - E_n$, hence

$$E' - \Delta \leq E_n \leq E' + \Delta, \tag{13.162}$$

to which it must be added that $E_n \leq E'$. In deriving this result, the hypothesis that E_n is the eigenvalue closest to E' was essential; if this is not true, it is clear that the previous estimate is no longer reliable. More advanced textbooks on QM contain other methods that give more precise results for the eigenvalue bounds, although they are, of course, more complicated.

13.8.3 Other applications: Partial shielding of the Coulomb potential

Before concluding this section, we apply the variational method to a problem whose exact solution we do not know, so that it is no longer a trivial example. The aim is to determine the ground-state energy for a particle subject to a partially shielded Coulomb potential produced by a neutral H atom. The effective potential, obtained by calculating the average (over the atomic ground state) of the potential produced by the nucleus and that produced by the orbiting electron, has the form

$$V(r) = \frac{k}{r} e^{-\alpha r}, \tag{13.163}$$

where α depends on the structure of the atom, A reasonable choice of trial function is somewhat similar to that for the ground state of the H atom, since this guarantees a good approximation for small values of α. Once normalized, this function can be written as

$$\psi(r) = \left(\frac{\beta^3}{\pi}\right)^{1/2} e^{-\beta r}, \tag{13.164}$$

and the expectation value of the Hamiltonian becomes

https://doi.org/10.1017/9781009679633.014 Published online by Cambridge University Press

$$E(\beta) = \langle\psi|\widehat{H}|\psi\rangle = \frac{\hbar^2\beta^2}{2m} - \frac{4k\beta^3}{(\alpha+2\beta)^2}. \tag{13.165}$$

The complicated dependence of this expression on β forces us to carry out a graphical or numerical analysis to determine its minimum. However, some conclusions can easily be drawn from Eq. (13.165). For example, $E(\beta) \leq 0$ for all values of β between 0 and a maximum (which can be easily determined), when $0 \leq \alpha \leq 1$ (in atomic units, with $k = 1$). For all these values of α and $\beta > 0$ there is a minimum of $E(\beta)$, which is the upper limit of the energy of the corresponding ground state. For example, for $\alpha = 0.5$ we get $\beta \simeq 0.9$ and $E_{\min} = E' \simeq -0.15$ a.u. (remember that for the H atom, $E_0 = -0.5$ a.u.). According to these calculations, the potential (13.163) ceases to have bound states at $\alpha = 1$, but a more precise analysis shows that this actually occurs around the value $\alpha \simeq 1.2$.

Once the value of β for which the minimum occurs has been obtained, that is, an upper bound for the ground-state energy has been found, and a lower bound can be obtained by combining Eqs. (13.162) and (13.158). To apply these, it is only necessary to calculate $\left\langle\psi'\right|\widehat{H}^2\left|\psi'\right\rangle$; the calculation is a bit tedious, but straightforward. For example, for $\alpha = 0.5$ we get $\left\langle\psi'\right|\widehat{H}^2\left|\psi'\right\rangle = 0.026$ a.u. and $\Delta = 0.059$ a.u., which means that $E_0 > -0.15 - 0.059$ a.u. in this case.

As mentioned earlier, the variational method is successfully applied to solve many important problems. When the number of parameters used is small and the problem is simple, analytical methods of solution are often possible. For systems with a larger number of components, computational algorithms have been developed that successfully use variational methods, with increasingly smaller margins of error. In Section 16.1 we will have the opportunity to introduce the variational method for the theoretical study of systems of general interest that do not have an exact solution, and in Section 16.3.3 we apply it in particular to the He atom.

Exercise E13.4. General use of the Rayleigh–Ritz method

The Ritz method for the approximate solution of an eigenvalue equation consists of expressing the eigenfunctions sought in terms of a conveniently truncated basis and requiring that the approximate eigenvalues computed with these functions are stationary with respect to small arbitrary changes in the development coefficients. Show that this procedure leads to a secular equation that solves the problem and that the smallest eigenvalue computed is an upper bound on the corresponding exact eigenvalue.

Solution. We write the equation to be solved in the form

$$\widehat{L}\Phi_k = \Lambda_k\Phi_k. \tag{13.166}$$

Instead of the exact problem (13.166) we propose the approximate problem $\widehat{L}\varphi_k = \lambda_k\varphi_k$, where the approximate eigenfunctions φ_k are constructed in terms of a proper basis $\{u_n\}$, but using only a finite number of terms,

$$\varphi_k = \sum_{n=0}^{N} c_{kn}u_n. \tag{13.167}$$

If L_{mn} denotes the matrix elements of the operator $\widehat{L}$ in the u representation, it follows from the previous two expressions that

https://doi.org/10.1017/9781009679633.014 Published online by Cambridge University Press

$$\int \varphi_k^* \widehat{L} \varphi_k dx = \sum_{m,n}^{N} c_{km}^* c_{kn} L_{mn} = \lambda_k \sum_{m,n}^{N} c_{km}^* c_{kn} \delta_{mn}. \tag{13.168}$$

λ_k is now a function of the coefficients c_{kl}, c_{kl}^*; requiring it to be stationary with respect to them (i.e., imposing the conditions $\partial\lambda_k/\partial c_{kl} = 0$, $\partial\lambda_k/\partial c_{kl}^* = 0$), we obtain by derivation of the previous equation that

$$\sum_{n}^{N} c_{kn} L_{mn} = \lambda_k \sum_{n}^{N} c_{kn} \delta_{mn}, \tag{13.169}$$

which can be expressed as a secular equation

$$\sum_{n}^{N} (L_{mn} - \lambda_k \delta_{mn})\, c_{kn} = 0, \tag{13.170}$$

as requested. Solving this equation by the usual algebraic methods gives the approximate eigenvalues λ_k and the coefficients c_{kl} of the expansion (13.167). In the limit $N \to \infty$ the solution becomes exact.

To show that this method gives an upper bound on the lowest eigenvalue, we proceed as follows. First, we order the eigenvalues, both exact and approximate, from smallest to largest:

$$\Lambda_0 < \Lambda_1 < \Lambda_2 < \ldots, \quad \lambda_0 < \lambda_1 < \lambda_2 < \ldots \tag{13.171}$$

Then we express the approximate functions in terms of the exact ones,

$$\varphi_k = \sum_{m} a_{km} \Phi_m. \tag{13.172}$$

From the preceding equations, it follows that

$$\lambda_k = \frac{\int \varphi_k^* \widehat{L} \varphi_k dx}{\int \varphi_k^* \varphi_k dx} = \frac{\sum_{m,n} a_{km}^* a_{kn} \Lambda_n \delta_{mn}}{\sum_n |a_{kn}|^2} = \frac{\sum_n \Lambda_n |a_{kn}|^2}{\sum_n |a_{kn}|^2}. \tag{13.173}$$

Since $\Lambda_n \geq \Lambda_0$, substituting in the previous expression each Λ_k by Λ_0 we see that $\lambda_k \geq \Lambda_0$. In particular, taking $k = 0$ gives $\lambda_0 \geq \Lambda_0$, which is precisely the requested result. In a similar way it is shown that

$$\lambda_1 \geq \Lambda_1 - (\Lambda_1 - \Lambda_0) \frac{|a_{10}|^2}{\sum_n |a_{1n}|^2}. \tag{13.174}$$

Problems

P13.1 Show that if $\widehat{f}$ is a dynamical variable of a perturbed system, its first-order perturbed matrix elements are

$$f_{nm} = f_{nm}^{(0)} + \sum_{k \neq n} \frac{V_{nk} f_{km}^{(0)}}{E_n^{(0)} - E_k^{(0)}} + \sum_{k \neq m} \frac{V_{km} f_{nk}^{(0)}}{E_m^{(0)} - E_k^{(0)}}. \tag{13.175}$$

https://doi.org/10.1017/9781009679633.014 Published online by Cambridge University Press

P13.2 The potential of a one-dimensional anharmonic oscillator can be approximated by the expression

$$V = \tfrac{1}{2}m\omega^2 x^2 \left[1 + \alpha\frac{x}{x_0} + \beta\left(\frac{x}{x_0}\right)^2\right], \tag{13.176}$$

where $x_0 = \sqrt{\hbar/m\omega}$. Use perturbation theory to determine the corrections to the energy up to second-order and the first-order wave functions. Use your results to calculate $\langle x\rangle$, $\langle\widehat{p}\rangle$, $\langle(\Delta x)^2\rangle$, $\langle(\Delta\widehat{p})^2\rangle$.

P13.3 A particle moves on a vertical circle of radius R; neglecting the effects of friction, but taking into account the force of gravity, the Hamiltonian of the system is

$$\widehat{H} = \frac{\widehat{L}_z^2}{2mR^2} + mgR\sin\varphi. \tag{13.177}$$

Determine the unperturbed solutions, as well as the first- and second-order corrections to the energy, treating the gravitational term as a perturbation. Under what conditions is this solution valid?

P13.4 Study the normal Zeeman effect of spinless particles using the methods of perturbation theory.

P13.5 Show that the matrix $\widehat{A}$ defined by Eq. (13.61) also satisfies the condition $\widehat{A}\widehat{A}^\dagger = 1$, which, taken together with $\widehat{A}^\dagger\widehat{A} = 1$, guarantees that $\widehat{A}$ is unitary.

P13.6 Explain why the linear Stark effect increases with the principal quantum number n.

P13.7 Calculate the Stark effect on the H atom for the levels with $n = 3$. Make maximum use of the symmetry properties of the system, using spherical coordinates.

P13.8 Calculate the intensities of the components of the H_α line of hydrogen, split by the linear Stark effect.

P13.9 Two coupled oscillators have the Hamiltonian

$$\widehat{H} = \frac{1}{2m}\left(\widehat{p}_1^2 + \widehat{p}_2^2\right) + \tfrac{1}{2}m\omega_1^2 x_1^2 + \tfrac{1}{2}m\omega_2^2 x_2^2 + m\omega^2\beta x_1 x_2, \tag{13.178}$$

where $\omega^2 = \omega_1\omega_2$ and β is arbitrary. Study this problem with perturbative methods when:

(a) ω_1/ω_2 is an irrational number;
(b) ω_1/ω_2 is a rational number.

Solve the problem exactly and compare results.

P13.10 Determine the emission spectrum of the coupled oscillators of the previous problem, and compare it with the corresponding spectrum of two independent oscillators.

P13.11 Consider a particle inside a 3D square well of side L. Find the corrections to the correct energy levels and wave functions of the problem perturbed by gravity.

P13.12 Two identical flat rotors are coupled so that their Hamiltonian is

$$\widehat{H} = A\left(p_{\theta_1}^2 + p_{\theta_2}^2\right) - B\cos(\theta_1 - \theta_2), \quad A, B > 0.$$

https://doi.org/10.1017/9781009679633.014 Published online by Cambridge University Press

Determine the energy spectrum of the system and the corresponding zero-order eigenfunctions:

a) up to the linear terms in B, assuming that $B \ll A\hbar^2$;
b) reducing the problem to that of a harmonic oscillator if $B \gg A\hbar^2$.

P13.13 Consider the Hamiltonian

$$\widehat{H} = \hat{a}^\dagger \hat{a} + \tfrac{1}{2} + \eta(\hat{a}^\dagger \hat{a}^\dagger + \hat{a}\hat{a}), \quad |\eta| < \tfrac{1}{2}, \tag{13.179}$$

representing an oscillator perturbed by an interaction with coupling constant η. Consider the linear transformation

$$\hat{a} = \alpha\hat{\lambda} + \beta\hat{\lambda}^\dagger, \quad \hat{a}^\dagger = \alpha\hat{\lambda}^\dagger + \beta\hat{\lambda}, \tag{13.180}$$

with real α, β , and $\alpha^2 - \beta^2 = 1$, where $\hat{\lambda}, \hat{\lambda}^\dagger$ are two new operators that comply with the usual commutation rule (which allows us to equate them with creation and annihilation operators), $\left[\hat{\lambda}, \hat{\lambda}^\dagger\right] = \mathbb{I}$. This leads to a Hamiltonian without interaction

$$\widehat{H} = \sqrt{1 - 4\eta^2}\left(\hat{\lambda}^\dagger\hat{\lambda} + \tfrac{1}{2}\right), \tag{13.181}$$

if

$$\begin{aligned} \alpha^2 + \beta^2 + 4\eta\alpha\beta &= \sqrt{1 - 4\eta^2}, \\ \tfrac{1}{2} + 2\eta\alpha\beta + \beta^2 &= \tfrac{1}{2}\sqrt{1 - 4\eta^2}. \end{aligned} \tag{13.182}$$

This is the so-called *canonical Bogolyubov transformation*. Determine the operator that performs this transformation.

P13.14 To solve the harmonic oscillator problem with the variational method, the following test functions are proposed:

(i) $A_0 e^{-\alpha x^2}$ for the ground state;
(ii) $A_1 x e^{-\alpha x^2}$ for the first excited state;
(iii) $A_2(1 + bx^2)e^{-\alpha x^2}$ for the second excited state.

Justify this selection and determine the wave functions and the energy eigenvalues for the first three states.

P13.15 Study the problem of a particle bound by the Yukawa[7] potential,

$$V = -g^2 \frac{e^{-\alpha r}}{r}, \tag{13.183}$$

using the variational method. Find the maximum value of α for which there is at least one bound state. Go to the limit of the Coulomb potential and show that for this potential there is always a discrete spectrum. Find the optimal ground-state energy and wave function and in the Coulomb limit. *Hint:* For Coulomb solutions to be exact, it is convenient to use $\psi = Ae^{-\beta r}$ as a test function.

[7] Hideki Yukawa (1907–1981) was a Japanese physicist who won the Nobel Prize in Physics in 1949 for successfully predicting the existence of the meson. He is also known for introducing a potential of the form (13.183) as a model for the nuclear force.

https://doi.org/10.1017/9781009679633.014 Published online by Cambridge University Press

P13.16 In nuclear physics, a truncated harmonic-oscillator potential $V = \frac{1}{2}m\omega^2(x^2 - a^2)$ for $|x| \leq a$ and $V = 0$ for $|x| > a$ is often used as a model for an attractive well.

a) Use the variational method to estimate the energy of the ground state and the first excited state.

b) Estimate the preceding eigenvalues using the WKB method. Under what conditions can the well contain only one bound state?

Compare the results and determine which are more reliable.

P13.17 Use the variational method to show that every purely attractive 1D potential has at least one bound state, regardless of its depth. *Hint*: A Gaussian test function works well for this purpose.

Bibliographical Notes

An introduction to the perturbation theory of classical systems and its application to the H atom can be found in José and Saletan (1998) and in Corben and Stehle (1960), chapter 13. The latter treats the classical Zeeman effect.

The excellent book on the H and He atoms by Bethe and Salpeter (1957) contains a detailed discussion of the Stark effect in H, including a comparison with the experiment. For very strong fields, the problem is treated using the WKB method.

A detailed discussion of perturbation theory can be found in Merzbacher (1998), chapter 18.

Fernández (2020) takes a practical approach to perturbation theory that addresses its actual implementation.

For an advanced textbook on variational principles, see Basdevant (2023).

https://doi.org/10.1017/9781009679633.014 Published online by Cambridge University Press