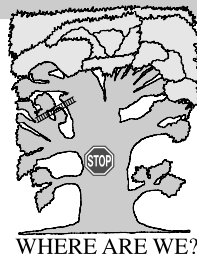


Chapter 5

TWO FUNDAMENTAL APPROXIMATE METHODS



Where are we?

We are moving upwards in the central parts of the TREE trunk.

An example

We are interested in properties of the ammonia molecule in its ground and excited states, e.g., we would like to know the mean value of the nitrogen–hydrogen distance. Only quantum mechanics gives a method for calculation this value (p. 24): we have to calculate the mean value of an operator with the ground-state wave function. But where could this function be taken from? As a solution of the Schrödinger equation? Impossible. This equation is too difficult to solve (14 particles, cf. problems with exact solutions, Chapter 4).

The only possibility is somehow to obtain an *approximate* solution to this equation.

What is it all about

We need mathematical methods which *will allow us to obtain approximate solutions* of the Schrödinger equation. These methods are: the variational method and the perturbational approach.

Variational method (▲)

p. 196

- Variational principle
- Variational parameters
- Ritz method

Perturbational method (△◆)

p. 203

- Rayleigh–Schrödinger approach (△)
- Hylleraas variational principle (◆)
- Hylleraas equation (◆)
- Convergence of the perturbational series (◆)

Why is this important?

We have to know how to calculate wave functions. The exact wave function is definitely out of our reach, therefore in this chapter we will learn how to calculate the approximations.

What is needed?

- Postulates of quantum mechanics (Chapter 1, needed).

- Hilbert space (Appendix B, p. 895, necessary).
- Matrix algebra (Appendix A, p. 889, needed).
- Lagrange multipliers (Appendix N, on p. 997, needed).
- Orthogonalization (Appendix J, p. 977, occasionally used).
- Matrix diagonalization (Appendix K, p. 982, needed).
- Group theory (Appendix C, p. 903, occasionally used in this chapter).

Classical works

The variational method of linear combinations of functions was formulated by Walther Ritz in a paper published in *Zeitschrift für Reine und Angewandte Mathematik*, 135 (1909) 1. ★ The method was applied by Erwin Schrödinger in his first works “*Quantisierung als Eigenwertproblem*” in *Annalen der Physik*, 79 (1926) 361, *ibid.* 79 (1926) 489, *ibid.* 80 (1926) 437, *ibid.* 81 (1926) 109. Schrödinger also used the perturbational approach when developing the theoretical results of Lord Rayleigh for vibrating systems (hence the often used term Rayleigh–Schrödinger perturbation theory). ★ Egil Andersen Hylleraas, in *Zeitschrift der Physik*, 65 (1930) 209 showed for the first time that the variational principle may be used also for separate terms of the perturbational series.

5.1 VARIATIONAL METHOD

5.1.1 VARIATIONAL PRINCIPLE

variational
function

Let us write the Hamiltonian \hat{H} of the system under consideration¹ and take an *arbitrary* (variational) function Φ , which satisfies the following conditions:

- it depends on the same coordinates as the solution to the Schrödinger equation;
- it is of class Q, p. 73 (which enables it to be normalized).

We calculate the number ε that depends on Φ (i.e. ε is a functional of Φ)

$$\varepsilon[\Phi] = \frac{\langle \Phi | \hat{H} | \Phi \rangle}{\langle \Phi | \Phi \rangle}. \quad (5.1)$$

The variational principle states:

- $\varepsilon \geq E_0$, where E_0 is the *ground-state* energy of the system
- in the above inequality $\varepsilon = E_0$ happens, if and only if, Φ equals the exact ground-state wave function ψ_0 of the system, $\hat{H}\psi_0 = E_0\psi_0$.

¹We focus here on the non-relativistic case (eq. (2.1)), where the lowest eigenvalue of \hat{H} is bound from below ($> -\infty$). As we remember from Chapter 3, this is not fulfilled in the relativistic case (Dirac’s electronic sea), and may lead to serious difficulties in applying the variational method.

Proof (expansion into eigenfunctions):

The unknown eigenfunctions $\{\psi_i\}$ of the Hamiltonian \hat{H} represent a complete set (we may be assured of its orthonormality, see Appendix B on p. 895) in the Hilbert space of our system. This means that any function belonging to this space can be represented as a linear combination of the functions of this set

$$\Phi = \sum_{i=0}^{\infty} c_i \psi_i, \quad (5.2)$$

where c_i assure the normalization of Φ , i.e. $\sum_{i=0}^{\infty} |c_i|^2 = 1$, because

$$\langle \Phi | \Phi \rangle = \sum_{i,j} c_j^* c_i \langle \psi_j | \psi_i \rangle = \sum_{i,j} c_j^* c_i \delta_{ij} = \sum_i c_i^* c_i = 1.$$

Let us insert this into the expression for the mean value of the energy $\varepsilon = \langle \Phi | \hat{H} \Phi \rangle$

$$\begin{aligned} \varepsilon - E_0 &= \langle \Phi | \hat{H} \Phi \rangle - E_0 = \left\langle \sum_{j=0}^{\infty} c_j \psi_j \left| \hat{H} \sum_{i=0}^{\infty} c_i \psi_i \right. \right\rangle - E_0 \\ &= \sum_{i,j=0}^{\infty} c_j^* c_i E_i \langle \psi_j | \psi_i \rangle - E_0 = \sum_{i,j=0}^{\infty} c_j^* c_i E_i \delta_{ij} - E_0 = \sum_{i=0}^{\infty} |c_i|^2 E_i - E_0 \cdot 1 \\ &= \sum_{i=0}^{\infty} |c_i|^2 E_i - E_0 \sum_{i=0}^{\infty} |c_i|^2 = \sum_{i=0}^{\infty} |c_i|^2 (E_i - E_0) \geq 0. \end{aligned}$$

Note that the *equality* (in the last step) is satisfied only if $\Phi = \psi_0$. This therefore proves the variational principle (5.1): $\varepsilon \geq E_0$.

In several places in this book we will need similar proofs using Lagrange multipliers. This is why we will demonstrate how to prove the same theorem using this technique (Appendix N on p. 997).

Proof using Lagrange multipliers:

Take the functional

$$\varepsilon[\Phi] = \langle \Phi | \hat{H} \Phi \rangle. \quad (5.3)$$

We want to find a function that assures a minimum of the functional and satisfies the normalization condition

$$\langle \Phi | \Phi \rangle - 1 = 0. \quad (5.4)$$

We will change the function Φ a little (the change will be called “variation”) and see, how this will change the value of the functional $\varepsilon[\Phi]$. In the functional we

have, however, Φ and Φ^* . It seems that we have, therefore, to take into account in Φ^* the variation made in Φ . *However, in reality there is no need to do that:* it is sufficient to make the variation either in Φ or in Φ^* (the result does not depend on the choice²). *This makes the formulae simpler.* We decide to choose the variation of Φ^* , i.e. $\delta\Phi^*$.

Now we apply the machinery of the Lagrange multipliers (Appendix N on p. 997). Let us multiply eq. (5.4) by (for the time being) unknown Lagrange multiplier E and subtract afterwards from the functional ε , resulting in an auxiliary functional $G[\Phi]$

$$G[\Phi] = \varepsilon[\Phi] - E(\langle\Phi|\Phi\rangle - 1).$$

The variation of G (which is analogous to the differential of a function) represents a linear term in $\delta\Phi^*$. For an extremum the variation has to be equal to zero:

$$\delta G = \langle\delta\Phi|\hat{H}\Phi\rangle - E\langle\delta\Phi|\Phi\rangle = \langle\delta\Phi|(\hat{H} - E)\Phi\rangle = 0.$$

Since this has to be satisfied for *any* variation $\delta\Phi^*$, then it can follow only if

$$(\hat{H} - E)\Phi_{\text{opt}} = 0, \quad (5.6)$$

which means that the optimal $\Phi \equiv \Phi_{\text{opt}}$ is a solution of the Schrödinger equation³ with E as the energy of the stationary state.

²Let us show this, because we will use it several times in this book. In all our cases the functional (which depends here on a single function $\phi(x)$, but later we will also deal with several functions in a similar procedure) might be rewritten as

$$\varepsilon[\phi] = \langle\phi|\hat{A}\phi\rangle, \quad (5.5)$$

where \hat{A} is a Hermitian operator. Let us write $\phi(x) = a(x) + ib(x)$, where $a(x)$ and $b(x)$ are *real functions*. The change of ε is equal to

$$\begin{aligned} \varepsilon[\phi + \delta\phi] - \varepsilon[\phi] &= \langle a + \delta a + ib + i\delta b|\hat{A}(a + \delta a + ib + i\delta b)\rangle - \langle a + ib|\hat{A}(a + ib)\rangle \\ &= \langle \delta a + i\delta b|\hat{A}\phi\rangle + \langle\phi|\hat{A}(\delta a + i\delta b)\rangle + \text{quadratic terms} \\ &= \langle \delta a|\hat{A}\phi + (\hat{A}\phi)^*\rangle + i\langle \delta b|(\hat{A}\phi)^* - \hat{A}\phi\rangle + \text{quadratic terms.} \end{aligned}$$

The variation of a function only represents a linear part of the change, and therefore $\delta\varepsilon = \langle \delta a|\hat{A}\phi + (\hat{A}\phi)^*\rangle + i\langle \delta b|(\hat{A}\phi)^* - \hat{A}\phi\rangle$. At the extremum the variation has to equal zero *at any* variations of δa and δb . This may happen only if $\hat{A}\phi + (\hat{A}\phi)^* = 0$ and $(\hat{A}\phi)^* - \hat{A}\phi = 0$. This means $\hat{A}\phi = 0$ or, *equivalently*, $(\hat{A}\phi)^* = 0$.

The first of the conditions would be obtained if in ε we made the variation in ϕ^* *only* (the variation in the extremum would then be $\delta\varepsilon = \langle\delta\phi|\hat{A}\phi\rangle = 0$), hence, from the arbitrariness of $\delta\phi^*$ we would get $\hat{A}\phi = 0$, the second, if we made the variation in ϕ *only* (then, $\delta\varepsilon = \langle\phi|\hat{A}\delta\phi\rangle = \langle\hat{A}\phi|\delta\phi\rangle = 0$ and $(\hat{A}\phi)^* = 0$) and the result is exactly the same. This is what we wanted to show: *we may vary either ϕ or ϕ^* and the result is the same.*

³In the variational calculus the equation for the optimum Φ , or the conditional minimum of a functional ε , is called the Euler equation. As one can see in this case the Euler equation is identical with the Schrödinger one.

Now let us multiply eq. (5.6) by Φ_{opt}^* and integrate. We obtain

$$\langle \Phi_{\text{opt}} | \hat{H} \Phi_{\text{opt}} \rangle - E \langle \Phi_{\text{opt}} | \Phi_{\text{opt}} \rangle = 0, \quad (5.7)$$

which means that the conditional minimum of $\varepsilon[\Phi]$ is $E = \min(E_0, E_1, E_2, \dots) = E_0$ (the ground state). Indeed, eq. (5.7) may be written as the mean value of the Hamiltonian

$$\left\langle \frac{1}{\sqrt{\langle \Phi_{\text{opt}} | \Phi_{\text{opt}} \rangle}} \Phi_{\text{opt}} \left| \hat{H} \right| \frac{1}{\sqrt{\langle \Phi_{\text{opt}} | \Phi_{\text{opt}} \rangle}} \Phi_{\text{opt}} \right\rangle = E = \varepsilon \left[\frac{1}{\sqrt{\langle \Phi_{\text{opt}} | \Phi_{\text{opt}} \rangle}} \Phi_{\text{opt}} \right], \quad (5.8)$$

and the lowest possible eigenvalue E is E_0 . Hence, for any other Φ we obtain a *higher* energy value, therefore $\varepsilon \geq E_0$.

The same was obtained when we expanded Φ into the eigenfunction series.

Variational principle for excited states

The variational principle (5.1) has been proved for an approximation to the ground-state wave function. What about excited states? If the variational function Φ is *orthogonal to exact solutions to the Schrödinger equation that correspond to all the states of lower energy than the state we are interested in*, the variational principle is still valid.⁴ If the wave function k being sought represents the *lowest* state among those belonging to a given irreducible representation of the symmetry group of the Hamiltonian, then the orthogonality mentioned above is automatically guaranteed (see Appendix C on p. 903). For other excited states, the variational principle cannot be satisfied, except that function Φ does not contain lower-energy wave functions, i.e. is orthogonal to them, e.g., because the wave functions have been cut out of it earlier.

variational
principle for
excited states

Beware of mathematical states

We mentioned in Chapter 1 that not all solutions of the Schrödinger equation are acceptable. Only those are acceptable which satisfy the symmetry requirements with respect to the exchange of labels corresponding to identical particles (Postulate V). The other solutions are called *mathematical*. If, therefore, an incautious scientist takes a variational function Φ with a non-physical symmetry, the variational principle, following our derivation exactly (p. 197), will still be valid, but with respect to the *mathematical ground state*. The mathematical states may correspond to energy eigenvalues *lower* than the physical ground state (they are called the *underground* states, cf. p. 76). All this would end up as a catastrophe, because the mean value of the Hamiltonian would tend towards the non-physical underground mathematical state.

underground
states

⁴The corresponding proof will only be slightly modified. Simply in the expansion eq. (5.2) of the variational function Φ , the wave functions ψ_i that correspond to lower energy states (than the state in which we are interested), will be absent. We will therefore obtain $\sum_{i=1} |c_i|^2 (E_i - E_k) \geq 0$, because state k is the lowest among all the states i .

5.1.2 VARIATIONAL PARAMETERS

The variational principle (5.1) may seem a little puzzling. We insert an *arbitrary* function Φ into the integral and obtain a result related to the ground state of the *system under consideration*. And yet the arbitrary function Φ may have absolutely nothing to do with the molecule we consider. The problem is that the integral contains the most important information about our system. The information resides in \hat{H} . Indeed, if someone wrote down the expression for \hat{H} , we would know right away that the system contains N electrons and M nuclei, we would also know the charges on the nuclei, i.e. the chemical elements of which the system is composed.⁵ This is important information.

trial function

The variational method represents an application of the variational principle. The *trial wave function* Φ is taken in an analytical form (with the variables denoted by the vector \mathbf{x} and automatically satisfying Postulate V). In the “key positions” in the formula for Φ we introduce the parameters $\mathbf{c} \equiv (c_0, c_1, c_2, \dots, c_P)$, which we may change smoothly. The parameters play the role of tuning, their particular values listed in vector \mathbf{c} result in a certain shape of $\Phi(\mathbf{x}; \mathbf{c})$. The integration in the formula for ε pertains to the variables \mathbf{x} , therefore the result depends uniquely on \mathbf{c} . Our function $\varepsilon(\mathbf{c})$ has the form

$$\varepsilon(c_0, c_1, c_2, \dots, c_P) \equiv \varepsilon(\mathbf{c}) = \frac{\langle \Phi(\mathbf{x}; \mathbf{c}) | \hat{H} \Phi(\mathbf{x}; \mathbf{c}) \rangle}{\langle \Phi(\mathbf{x}; \mathbf{c}) | \Phi(\mathbf{x}; \mathbf{c}) \rangle}.$$

Now the problem is to find the minimum of the function $\varepsilon(c_0, c_1, c_2, \dots, c_P)$.

In a general case the task is not simple, because what we are searching for is the *global* minimum. The relation

$$\frac{\partial \varepsilon(c_0, c_1, c_2, \dots, c_P)}{\partial c_i} = 0 \quad \text{for } i = 0, 1, 2, \dots, P,$$

therefore represents only a *necessary* condition for the global minimum.⁶ This problem may be disregarded, when:

- the number of minima is small,
- in particular, when we use Φ with the linear parameters \mathbf{c} (in this case we have a single minimum, see below).

The above equations enable us to find the optimum set of parameters $\mathbf{c} = \mathbf{c}_{\text{opt}}$. Then,

in a given class of the trial functions Φ the best possible approximation to ψ_0 is $\Phi(\mathbf{x}; \mathbf{c}_{\text{opt}})$, and the best approximation to E_0 is $\varepsilon(\mathbf{c}_{\text{opt}})$.

⁵And yet we would be unable to decide whether we have to do with matter or antimatter, or whether we have to perform calculations for the benzene molecule or for six CH radicals (cf. Chapter 2).

⁶More about global minimization may be found in Chapter 6.

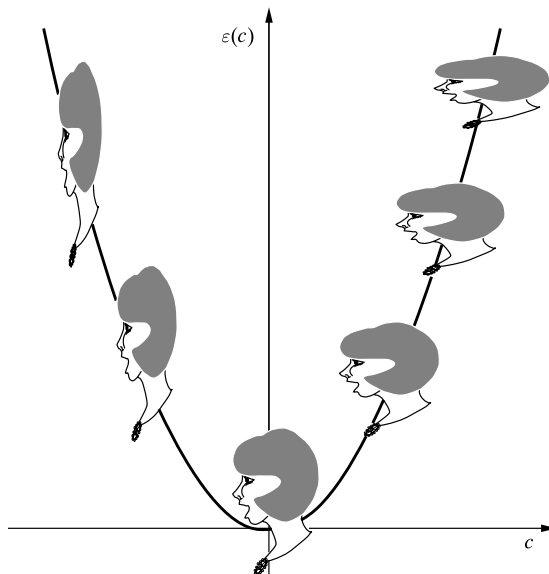


Fig. 5.1. The philosophy behind the variational method. A parameter c is changed in order to obtain the best solution possible. Any commentary would obscure the very essence of the method.

Fig. 5.1. shows the essence of the variational method.⁷

Let us assume that someone does not know that the hydrogen-like atom (the nucleus has a charge Z) problem has an exact solution.⁸ Let us apply the simplest version of the variational method to see what kind of problem we will be confronted with.

An important first step will be to decide which class of trial functions to choose. We decide to take the following class⁹ (for $c > 0$) $\exp(-cr)$ and after normalization of the function: $\Phi(r, \theta, \phi; c) = \sqrt{\frac{c^3}{\pi}} \exp(-cr)$. The calculation $\varepsilon[\Phi] = \langle \Phi | \hat{H} | \Phi \rangle$ is shown in Appendix H on p. 969. We obtain $\varepsilon(c) = \frac{1}{2}c^2 - Zc$. We very easily find the minimum of $\varepsilon(c)$ and the optimum c is equal to $c_{\text{opt}} = Z$, which, as we know from Chapter 4, represents the exact result. In practise (for atoms or molecules), we would never know the exact result. The optimal ε might then be obtained after many days of number crunching.¹⁰

⁷The variational method is used in everyday life. Usually we determine the target (say, cleaning the car), and then by trial, errors and corrections we approach the ideal, but never fully achieve it.

⁸For a larger system we will not know the exact solution either.

⁹A particular choice is usually made through scientific discussion. The discussion might proceed as follows.

The electron and the nucleus attract themselves, therefore they will be close in space. This assures many classes of trial functions, e.g., $\exp(-cr)$, $\exp(-cr^2)$, $\exp(-cr^3)$, etc., where $c > 0$ is a single variational parameter. In the present example we pretend not to know, which class of functions is most promising (i.e. which will give lower ε). Let us begin with class $\exp(-cr)$, and other classes will be investigated in the years to come. The decision made, we rush to do the calculations.

¹⁰For example, for $Z = 1$ we had to decide a starting value of c , say, $c = 2$; $\varepsilon(2) = 0$. Let us try $c = 1.5$, we obtain a lower (i.e. better) value $\varepsilon(1.5) = -0.375$ a.u., the energy goes down. Good direction, let us

5.1.3 RITZ METHOD¹¹

basis functions

The Ritz method represents a special kind of variational method. The trial function Φ is represented as a *linear* combination of the *known* basis functions $\{\Psi_i\}$ with the (for the moment) *unknown* variational coefficients c_i

$$\Phi = \sum_{i=0}^P c_i \Psi_i. \quad (5.9)$$

Then

$$\varepsilon = \frac{\langle \sum_{i=0}^P c_i \Psi_i | \hat{H} | \sum_{i=0}^P c_i \Psi_i \rangle}{\langle \sum_{i=0}^P c_i \Psi_i | \sum_{i=0}^P c_i \Psi_i \rangle} = \frac{\sum_{i=0}^P \sum_{j=0}^P c_i^* c_j H_{ij}}{\sum_{i=0}^P \sum_{j=0}^P c_i^* c_j S_{ij}} = \frac{A}{B}. \quad (5.10)$$

complete basis set

In the formula above $\{\Psi_i\}$ represents the chosen *complete basis set*.¹² The basis set functions are usually non-orthogonal, and therefore

$$\langle \Psi_i | \Psi_j \rangle = S_{ij}, \quad (5.11)$$

overlap matrix

where S stands for the *overlap matrix*, and the integrals

$$H_{ij} = \langle \Psi_i | \hat{H} | \Psi_j \rangle \quad (5.12)$$

are the matrix elements of the Hamiltonian. Both matrices (S and H) are calculated once and for all. The energy ε becomes a function of the linear coefficients $\{c_i\}$. The coefficients $\{c_i\}$ and the coefficients $\{c_i^*\}$ are not independent (c_i can be obtained from c_i^*). Therefore, as the linearly independent coefficients, we may treat either $\{c_i\}$ or $\{c_i^*\}$. When used for the minimization of ε , both choices would give the same. We decide to treat $\{c_i^*\}$ as variables. For each $k = 0, 1, \dots, P$ we

try, therefore $c = 1.2$; $\varepsilon(1.2) = -0.48$ a.u. Indeed, a good direction. However, when we continue and take $c = 0.7$, we obtain $\varepsilon = -0.455$, i.e. a higher energy. We would continue this way and finally obtain something like $c_{\text{opt}} = 1.0000000$. We might be satisfied by 8 significant figures and decide to stop the calculations. We would never be sure, however, whether other classes of trial functions would provide still better (i.e. lower) energies. In our particular case this, of course, would never happen, because “accidentally” we have taken a class which contains the exact wave function.

¹¹Walther Ritz was the Swiss physicist and a former student of Poincaré. His contributions, beside the variational approach, include perturbation theory, the theory of vibrations etc. Ritz is also known for his controversy with Einstein on the time flow problem (“time flash”), concluded by their joint article “An agreement to disagree” (W. Ritz, A. Einstein, *Phys. Zeit.* 10 (1909) 323).

¹²Such basis sets are available in the literature. A practical problem arises as to how many such functions should be used. In principle we should have used $P = \infty$. This, however, is unfeasible. We are restricted to a finite, *preferably small number*. And this is the moment when it turns out that some basis sets are more effective than others, that this depends on the problem considered, etc. This is how a new science emerges, which might facetiously be called basology.

have to have in the minimum

$$\begin{aligned}
 0 = \frac{\partial \varepsilon}{\partial c_k^*} &= \frac{(\sum_{j=0}^P c_j H_{kj})B - A(\sum_{j=0}^P c_j S_{kj})}{B^2} \\
 &= \frac{(\sum_{j=0}^P c_j H_{kj})}{B} - \frac{A}{B} \frac{(\sum_{j=0}^P c_j S_{kj})}{B} = \frac{(\sum_{j=0}^P c_j (H_{kj} - \varepsilon S_{kj}))}{B},
 \end{aligned}$$

what leads to the *secular equations*

secular
equations

$$\left(\sum_{j=0}^P c_j (H_{kj} - \varepsilon S_{kj}) \right) = 0 \quad \text{for } k = 0, 1, \dots, P. \quad (5.13)$$

The unknowns in the above equation are the coefficients c_j and the energy ε . With respect to the coefficients c_j , eqs. (5.13) represent a homogeneous set of linear equations. Such a set has a non-trivial solution if the *secular determinant* is equal to zero (see Appendix A)

secular
determinant

$$\det(H_{kj} - \varepsilon S_{kj}) = 0. \quad (5.14)$$

This happens however only for some particular values of ε satisfying the above equation. Since the rank of the determinant is equal $P + 1$, we therefore obtain $P + 1$ solutions ε_i , $i = 0, 1, 2, \dots, P$. Due to the Hermitian character of operator \hat{H} , the matrix \mathbf{H} will be also Hermitian. In Appendices J on p. 977 and L on p. 984, we show that the problem reduces to the diagonalization of some transformed \mathbf{H} matrix (also Hermitian). This guarantees that all ε_i will be real. Let us denote the lowest ε_i as ε_0 , to represent an approximation to the ground state energy.¹³ The other ε_i , $i = 1, 2, \dots, P$, will approximate the excited states of the system. We obtain an approximation to the i -th wave function by inserting the calculated ε_i into eq. (5.13), and then, after including the normalization condition, we find the corresponding set of c_i . The problem is solved.

5.2 PERTURBATIONAL METHOD

5.2.1 RAYLEIGH-SCHRÖDINGER APPROACH

The idea of the perturbational approach is very simple. We know everything about a non-perturbed problem. Then we slightly perturb the system and everything changes. If the perturbation is small, it seems there is a good chance that there

¹³ Assuming we used the basis functions that satisfy Postulate V (on symmetry).

will be no drama: the wave function and the corresponding energy will change only a little (if the changes were large, the perturbational approach would simply be inapplicable). The whole perturbational procedure aims at finding these tiny changes with satisfactory precision.

Perturbational theory is notorious for quite clumsy equations. Unfortunately, there is no way round if we want to explain how to calculate things. *However, in practise only a few of these equations will be used – they will be highlighted in frames.*

unperturbed
operator

Let us begin our story. We would like to solve the Schrödinger equation

$$\hat{H}\psi_k = E\psi_k, \quad (5.15)$$

and as a rule we will be interested in a single particular state k , most often the ground state ($k = 0$).

We apply a perturbational approach, when¹⁴

$$\hat{H} = \hat{H}^{(0)} + \hat{H}^{(1)},$$

perturbation

where the so called *unperturbed operator* $\hat{H}^{(0)}$ is “large”, while the *perturbation operator* $\hat{H}^{(1)}$ is “small”.¹⁵ We assume that there is no problem whatsoever with solving the unperturbed Schrödinger equation

$$\hat{H}^{(0)}\psi_k^{(0)} = E_k^{(0)}\psi_k^{(0)}. \quad (5.16)$$

intermediate
normalization

We assume that $\psi_k^{(0)}$ form an orthonormal set, which is natural.¹⁶ We are interested in the fate of the wave function $\psi_k^{(0)}$ after the perturbation is switched on (when it changes to ψ_k). We choose the *intermediate normalization*, i.e.

$$\langle \psi_k^{(0)} | \psi_k \rangle = 1. \quad (5.17)$$

perturbational
parameter

The intermediate normalization means that ψ_k , as a vector of the Hilbert space (see Appendix B on p. 895), has the normalized $\psi_k^{(0)}$ as the component along the unit basis vector $\psi_k^{(0)}$. In other words,¹⁷ $\psi_k = \psi_k^{(0)} +$ terms orthogonal to $\psi_k^{(0)}$.

We are all set to proceed. First, we introduce the *perturbational parameter* $0 \leq \lambda \leq 1$ in Hamiltonian \hat{H} , making it, therefore, λ -dependent¹⁸

$$\hat{H}(\lambda) = \hat{H}^{(0)} + \lambda\hat{H}^{(1)}.$$

¹⁴We assume all operators are Hermitian.

¹⁵In the sense that the energy spectrum of $\hat{H}^{(0)}$ is only slightly changed after the perturbation $\hat{H}^{(1)}$ is switched on.

¹⁶We can always do that, because $\hat{H}^{(0)}$ is Hermitian (see Appendix B).

¹⁷The intermediate normalization is convenient, but not necessary. Although convenient for the derivation of perturbational equations, it leads to some troubles when the mean values of operators are to be calculated.

¹⁸Its role is technical. It will enable us to transform the Schrödinger equation into a sequence of perturbational equations, which must be solved one by one. Then the parameter λ *disappears from the theory*, because we put $\lambda = 1$.

When $\lambda = 0$, $\hat{H}(\lambda) = \hat{H}^{(0)}$, while $\lambda = 1$ gives $\hat{H}(\lambda) = \hat{H}^{(0)} + \hat{H}^{(1)}$. In other words, we tune the perturbation at will from 0 to $\hat{H}^{(1)}$. It is worth noting that $\hat{H}(\lambda)$ for $\lambda \neq 0, 1$ may not correspond to any physical system. It does not need to. We are interested here in a mathematical trick, we will come back to reality by putting $\lambda = 1$ at the end.

We are interested in the Schrödinger equation being satisfied for all values $\lambda \in [0, 1]$

$$\hat{H}(\lambda)\psi_k(\lambda) = E_k(\lambda)\psi_k(\lambda).$$

Now this is a key step in the derivation. We expect that both the energy and the wave function can be developed in a power series¹⁹ of λ

$$E_k(\lambda) = E_k^{(0)} + \lambda E_k^{(1)} + \lambda^2 E_k^{(2)} + \dots, \quad (5.18)$$

$$\psi_k(\lambda) = \psi_k^{(0)} + \lambda \psi_k^{(1)} + \lambda^2 \psi_k^{(2)} + \dots, \quad (5.19)$$

where $E_k^{(i)}$ stand for some (unknown for the moment) coefficients, and $\psi_k^{(i)}$ represents the functions to be found. We expect the two series to converge (Fig. 5.2).

In practise we calculate only $E_k^{(1)}, E_k^{(2)}$ and quite rarely $E_k^{(3)}$, and for the wave function, we usually limit the correction to $\psi_k^{(1)}$.

How are these corrections calculated?

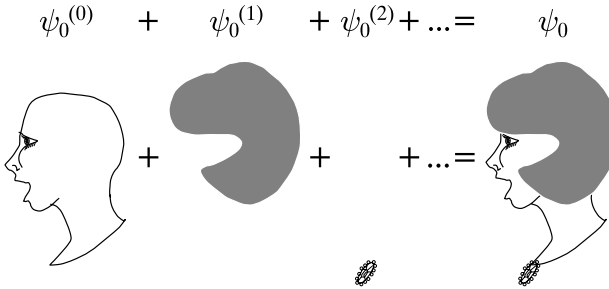


Fig. 5.2. Philosophy of the perturbational approach (optimistic version). The ideal ground-state wave function ψ_0 is constructed as a sum of a good zero-order approximation ($\psi_0^{(0)}$) and consecutive small corrections ($\psi_0^{(n)}$). The first-order correction ($\psi_0^{(1)}$) is still quite substantial, but fortunately the next corrections amount to only small cosmetic changes.

¹⁹It is in fact a Taylor series with respect to λ . The physical meaning of these expansions is the following: $E_k^{(0)}$ and $\psi_k^{(0)}$ are good approximations of $E_k(\lambda)$ and $\psi_k(\lambda)$. The rest will be calculated as a sum of small correction terms.

We insert the two perturbational series for $E_k(\lambda)$ and $\psi_k(\lambda)$ into the Schrödinger equation

$$\begin{aligned} &(\hat{H}^{(0)} + \lambda \hat{H}^{(1)})(\psi_k^{(0)} + \lambda \psi_k^{(1)} + \lambda^2 \psi_k^{(2)} + \dots) \\ &= (E_k^{(0)} + \lambda E_k^{(1)} + \lambda^2 E_k^{(2)} + \dots)(\psi_k^{(0)} + \lambda \psi_k^{(1)} + \lambda^2 \psi_k^{(2)} + \dots) \end{aligned}$$

and, since the equation has to be satisfied for *any* λ belonging to $0 \leq \lambda \leq 1$, this may happen only if

the coefficients at the same powers of λ on the left- and right-hand sides are equal.

This gives a sequence of an infinite number of perturbational equations to be satisfied by the unknown $E_k^{(n)}$ and $\psi_k^{(n)}$. *These equations may be solved consecutively allowing us to calculate $E_k^{(n)}$ and $\psi_k^{(n)}$ with larger and larger n .* We have, for example:

perturbational
equations

$$\begin{aligned} \text{for } \lambda^0: & \hat{H}^{(0)} \psi_k^{(0)} = E_k^{(0)} \psi_k^{(0)} \\ \text{for } \lambda^1: & \hat{H}^{(0)} \psi_k^{(1)} + \hat{H}^{(1)} \psi_k^{(0)} = E_k^{(0)} \psi_k^{(1)} + E_k^{(1)} \psi_k^{(0)} \\ \text{for } \lambda^2: & \hat{H}^{(0)} \psi_k^{(2)} + \hat{H}^{(1)} \psi_k^{(1)} = E_k^{(0)} \psi_k^{(2)} + E_k^{(1)} \psi_k^{(1)} + E_k^{(2)} \psi_k^{(0)} \\ & \dots \\ & \text{etc.}^{20} \end{aligned} \tag{5.20}$$

Doing the same with the intermediate normalization (eq. (5.17)), we obtain

$$\langle \psi_k^{(0)} | \psi_k^{(n)} \rangle = \delta_{0n}. \tag{5.21}$$

The first of eqs. (5.20) is evident (the unperturbed Schrödinger equation does not contain any unknown). The second equation involves two unknowns, $\psi_k^{(1)}$ and $E_k^{(1)}$. To eliminate $\psi_k^{(1)}$ we will use the Hermitian character of the operators. Indeed, by making the scalar product of the equation with $\psi_k^{(0)}$ we obtain:

$$\begin{aligned} &\langle \psi_k^{(0)} | (\hat{H}^{(0)} - E_k^{(0)}) \psi_k^{(1)} + (\hat{H}^{(1)} - E_k^{(1)}) \psi_k^{(0)} \rangle \\ &= \langle \psi_k^{(0)} | (\hat{H}^{(0)} - E_k^{(0)}) \psi_k^{(1)} \rangle + \langle \psi_k^{(0)} | (\hat{H}^{(1)} - E_k^{(1)}) \psi_k^{(0)} \rangle \\ &= 0 + \langle \psi_k^{(0)} | (\hat{H}^{(1)} - E_k^{(1)}) \psi_k^{(0)} \rangle = 0, \end{aligned}$$

i.e.

²⁰We see the construction principle of these equations: we write down all the terms which give a given value of the sum of the upper indices.

the formula for the *first-order correction to the energy*

$$E_k^{(1)} = H_{kk}^{(1)}, \quad (5.22)$$

where we defined

$$H_{kn}^{(1)} = \langle \psi_k^{(0)} | \hat{H}^{(1)} | \psi_n^{(0)} \rangle. \quad (5.23)$$

Conclusion: the first order correction to the energy, $E_k^{(1)}$, represents the mean value of the perturbation with the unperturbed wave function of the state in which we are interested (usually the ground state).²¹

first-order
correction

Now, from the perturbation equation (5.20) corresponding to $n = 2$ we have²²

$$\begin{aligned} & \langle \psi_k^{(0)} | (\hat{H}^{(0)} - E_k^{(0)}) \psi_k^{(2)} \rangle + \langle \psi_k^{(0)} | (\hat{H}^{(1)} - E_k^{(1)}) \psi_k^{(1)} \rangle - E_k^{(2)} \\ & = \langle \psi_k^{(0)} | \hat{H}^{(1)} \psi_k^{(1)} \rangle - E_k^{(2)} = 0, \end{aligned}$$

and hence

$$E_k^{(2)} = \langle \psi_k^{(0)} | \hat{H}^{(1)} \psi_k^{(1)} \rangle. \quad (5.24)$$

For the time being we cannot compute $E_k^{(2)}$, because we do not know $\psi_k^{(1)}$, but soon we will. In the perturbational equation (5.20) for λ^1 let us expand $\psi_k^{(1)}$ into the complete set of the basis functions $\{\psi_n^{(0)}\}$ with as yet unknown coefficients c_n :

$$\psi_k^{(1)} = \sum_{n(\neq k)} c_n \psi_n^{(0)}.$$

Note that because of the intermediate normalization (5.17) and (5.21), we did not take the term with $n = k$. We get

$$(\hat{H}^{(0)} - E_k^{(0)}) \sum_{n(\neq k)} c_n \psi_n^{(0)} + \hat{H}^{(1)} \psi_k^{(0)} = E_k^{(1)} \psi_k^{(0)},$$

²¹This is quite natural and we use such a perturbative estimation all the time. What it really says is: we do not know what the perturbation exactly does, but let us estimate the result by assuming *that all things are going on as they were before the perturbation was applied*. In the first-order approach, insurance estimates your loss by *averaging* over similar losses of others. A student score in quantum chemistry is often close to its *a posteriori* estimation from his/her other scores, etc.

²²Also through a scalar product with $\psi_k^{(0)}$.

and then transform

$$\sum_{n(\neq k)} c_n (E_n^{(0)} - E_k^{(0)}) \psi_n^{(0)} + \hat{H}^{(1)} \psi_k^{(0)} = E_k^{(1)} \psi_k^{(0)}.$$

We find c_m by making the scalar product with $\psi_m^{(0)}$. Due to the orthonormality of functions $\{\psi_n^{(0)}\}$ we obtain

$$c_m = \frac{H_{mk}^{(1)}}{E_k^{(0)} - E_m^{(0)}},$$

first-order
correction to
wave function

which gives the following formula for the *first-order correction to the wave function*

$$\psi_k^{(1)} = \sum_{n(\neq k)} \frac{H_{nk}^{(1)}}{E_k^{(0)} - E_n^{(0)}} \psi_n^{(0)}, \quad (5.25)$$

second-order
energy

and then the formula for the *second-order correction to the energy*

$$E_k^{(2)} = \sum_{n(\neq k)} \frac{|H_{kn}^{(1)}|^2}{E_k^{(0)} - E_n^{(0)}}. \quad (5.26)$$

From (5.25) we see that the contribution of function $\psi_n^{(0)}$ to the wave function deformation is large if the coupling between states k and n (i.e. $H_{nk}^{(1)}$) is large, and the closer in the energy scale these two states are.

The formulae for higher-order corrections become more and more complex. We will limit ourselves to the low-order corrections in the hope that the perturbational method converges fast (we will see in a moment how surprising the perturbational series behaviour can be) and further corrections are much less important.²³

5.2.2 HYLLERAAS VARIATIONAL PRINCIPLE²⁴

The derived formulae are rarely employed in practise, because we only very rarely have at our disposal all the necessary solutions of eq. (5.16). The eigenfunctions of the $\hat{H}^{(0)}$ operator appeared as a consequence of using them as the complete set of functions (e.g., in expanding $\psi_k^{(1)}$). There are, however, some numerical methods

²³Some scientists have been bitterly disappointed by this assumption.

²⁴See his biographic note in Chapter 10.

that enable us to compute $\psi_k^{(1)}$ using the complete set of functions $\{\phi_i\}$, which *are not* the eigenfunctions of $\hat{H}^{(0)}$.

Hylleraas noted²⁵ that the functional

$$\mathcal{E}[\tilde{\chi}] = \langle \tilde{\chi} | (\hat{H}^{(0)} - E_0^{(0)}) \tilde{\chi} \rangle \quad (5.27)$$

$$+ \langle \tilde{\chi} | (\hat{H}^{(1)} - E_0^{(1)}) \psi_0^{(0)} \rangle + \langle \psi_0^{(0)} | (\hat{H}^{(1)} - E_0^{(1)}) \tilde{\chi} \rangle \quad (5.28)$$

exhibits its minimum at $\tilde{\chi} = \psi_0^{(1)}$ and for this function the value of the functional is equal to $E_0^{(2)}$. Indeed, inserting $\tilde{\chi} = \psi_0^{(1)} + \delta\chi$ into eq. (5.28) and using the Hermitian character of the operators we have

$$\begin{aligned} [\psi_0^{(1)} + \delta\chi] - [\psi_0^{(1)}] &= \langle \psi_0^{(1)} + \delta\chi | (\hat{H}^{(0)} - E_0^{(0)}) (\psi_0^{(1)} + \delta\chi) \rangle \\ &\quad + \langle \psi_0^{(1)} + \delta\chi | (\hat{H}^{(1)} - E_0^{(1)}) \psi_0^{(0)} \rangle \\ &\quad + \langle \psi_0^{(0)} | (\hat{H}^{(1)} - E_0^{(1)}) (\psi_0^{(1)} + \delta\chi) \rangle \\ &= \langle \delta\chi | (\hat{H}^{(0)} - E_0^{(0)}) \psi_0^{(1)} + (\hat{H}^{(1)} - E_0^{(1)}) \psi_0^{(0)} \rangle \\ &\quad + \langle (\hat{H}^{(0)} - E_0^{(0)}) \psi_0^{(1)} + (\hat{H}^{(1)} - E_0^{(1)}) \psi_0^{(0)} | \delta\chi \rangle \\ &\quad + \langle \delta\chi | (\hat{H}^{(0)} - E_0^{(0)}) \delta\chi \rangle = \langle \delta\chi | (\hat{H}^{(0)} - E_0^{(0)}) \delta\chi \rangle \geq 0. \end{aligned}$$

This proves the Hylleraas variational principle. The last equality follows from the first-order perturbational equation, and the last inequality from the fact that $E_0^{(0)}$ is assumed to be the lowest eigenvalue of $\hat{H}^{(0)}$ (see the variational principle).

What is the minimal value of the functional under consideration? Let us insert $\tilde{\chi} = \psi_0^{(1)}$. We obtain

$$\begin{aligned} \mathcal{E}[\psi_0^{(1)}] &= \langle \psi_0^{(1)} | (\hat{H}^{(0)} - E_0^{(0)}) \psi_0^{(1)} \rangle \\ &\quad + \langle \psi_0^{(1)} | (\hat{H}^{(1)} - E_0^{(1)}) \psi_0^{(0)} \rangle + \langle \psi_0^{(0)} | (\hat{H}^{(1)} - E_0^{(1)}) \psi_0^{(1)} \rangle \\ &= \langle \psi_0^{(1)} | (\hat{H}^{(0)} - E_0^{(0)}) \psi_0^{(1)} + (\hat{H}^{(1)} - E_0^{(1)}) \psi_0^{(0)} \rangle + \langle \psi_0^{(0)} | \hat{H}^{(1)} \psi_0^{(1)} \rangle \\ &= \langle \psi_0^{(1)} | 0 \rangle + \langle \psi_0^{(0)} | \hat{H}^{(1)} \psi_0^{(1)} \rangle = \langle \psi_0^{(0)} | \hat{H}^{(1)} \psi_0^{(1)} \rangle = E_0^{(2)}. \end{aligned}$$

5.2.3 HYLLERAAS EQUATION

The first-order perturbation equation (p. 206, eq. (5.20)) after inserting

$$\psi_0^{(1)} = \sum_{j=1}^N d_j \phi_j \quad (5.29)$$

²⁵E.A. Hylleraas, *Zeit. Phys.* 65 (1930) 209.

takes the form

$$\sum_{j=1}^N d_j (\hat{H}^{(0)} - E_0^{(0)}) \phi_j + (\hat{H}^{(1)} - E_0^{(1)}) \psi_0^{(0)} = 0.$$

Making the scalar products of the left- and right-hand side of the equation with functions ϕ_i , $i = 1, 2, \dots$, we obtain

$$\sum_{j=1}^N d_j (\hat{H}_{ij}^{(0)} - E_0^{(0)} S_{ij}) = -(\hat{H}_{i0}^{(1)} - E_0^{(1)} S_{i0}) \quad \text{for } i = 1, 2, \dots, N,$$

where $\hat{H}_{ij}^{(0)} \equiv \langle \phi_i | \hat{H}^{(0)} | \phi_j \rangle$, and the overlap integrals $S_{ij} \equiv \langle \phi_i | \phi_j \rangle$. Using the matrix notation we may write the *Hylleraas equation*

$$(\mathbf{H}^{(0)} - E_0^{(0)} \mathbf{S}) \mathbf{d} = -\mathbf{v}, \quad (5.30)$$

where the components of the vector \mathbf{v} are $v_i = \hat{H}_{i0}^{(1)} - E_0^{(1)} S_{i0}$. All the quantities can be calculated and the set of N linear equations with unknown coefficients d_i remains to be solved.²⁶

5.2.4 CONVERGENCE OF THE PERTURBATIONAL SERIES

The perturbational approach is applicable when the perturbation only slightly changes the energy levels, therefore not changing their order. This means that the unperturbed energy level separations have to be much larger than a measure of perturbation such as $\hat{H}_{kk}^{(1)} = \langle \psi_k^{(0)} | \hat{H}^{(1)} | \psi_k^{(0)} \rangle$. However, even in this case we may expect complications.

The subsequent perturbational corrections need not be monotonically decreasing. However, if the perturbational series eq. (5.19) converges, for any $\varepsilon > 0$ we may choose such N_0 that for $N > N_0$ we have $\langle \psi_k^{(N)} | \psi_k^{(N)} \rangle < \varepsilon$, i.e. the vectors $\psi_k^{(N)}$ have smaller and smaller length in the Hilbert space.

Unfortunately, perturbational series are often divergent in a sense known as *asymptotic convergence*. A divergent series $\sum_{n=0}^{\infty} \frac{A_n}{z^n}$ is called an *asymptotic series* of a function $f(z)$, if the function $R_n(z) = z^n [f(z) - S_n(z)]$, where $S_n(z) = \sum_{k=0}^n \frac{A_k}{z^k}$, satisfies the following condition: $\lim_{z \rightarrow \infty} R_n(z) = 0$ for any fixed n . In other words, the error of the summation, i.e. $[f(z) - S_n(z)]$ tends to 0 as $z^{-(n+1)}$ or faster.

Despite the fact that the series used in physics and chemistry are often asymptotic, i.e. divergent, we are able to obtain results of high accuracy with them provided we limit ourselves to appropriate number of terms. The asymptotic character

²⁶We obtain the *same equation*, if in the Hylleraas functional eq. (5.28), the variational function χ is expanded as a linear combination (5.29), and then vary d_i in a similar way to that of the Ritz variational method described on p. 202.

of such series manifests itself in practise in such a way that the partial sums $S_n(z)$ stabilize and we obtain numerically a situation typical for convergence. For example, we sum up the consecutive perturbational corrections and obtain the partial sums changing on the eighth, then ninth, then tenth significant figures. This is a very good time to stop the calculations, publish the results, finish the scientific career and move on to other business. The addition of further perturbational corrections ends up in catastrophe, cf. Appendix X on p. 1038. It begins by an innocent, very small, increase in the partial sums, they just begin to change the ninth, then the eighth, then the seventh significant figure. Then, it only gets worse and worse and ends up by an explosion of the partial sums to ∞ and a very bad state of mind for the researcher (I did not dare to depict it in Fig. 5.2).

In perturbation theory we assume that $E_k(\lambda)$ and $\psi_k(\lambda)$ are analytical functions of λ (p. 205). In this *mathematical* aspect of the physical problem we may treat λ as a complex number. *Then the radius of convergence ρ of the perturbational series on the complex plane is equal to the smallest $|\lambda|$, for which one has a pole of $E_k(\lambda)$ or $\psi_k(\lambda)$.* The convergence radius ρ_k for the energy perturbational series may be computed as (if the limit exists²⁷)

$$\rho_k = \lim_{N \rightarrow \infty} \frac{|E_k^{(N)}|}{|E_k^{(N+1)}|}.$$

For physical reasons $\lambda = 1$ is most important. It is, therefore, desirable to have $\rho_k \geq 1$. Note (Fig. 5.3), that if $\rho_k \geq 1$, then the series with $\lambda = 1$ is convergent together with the series with $\lambda = -1$.

Let us take as the unperturbed system the harmonic oscillator (the potential energy equal to $\frac{1}{2}x^2$) in its ground state, and the operator $\hat{H}^{(1)} = -0.000001 \cdot x^4$ as its perturbation. In such a case the perturbation seems to be small²⁸ in comparison with the separation of the eigenvalues of $\hat{H}^{(0)}$. And yet the perturbational series carries the seed of catastrophe. It is quite easy to see why a catastrophe has to happen. After the perturbation is added, the potential becomes *qualitatively* different from $\frac{1}{2}x^2$. For large x , instead of going to ∞ , it will tend to $-\infty$. The perturbation is not small at all, it is a monster. This will cause the perturbational series to diverge. How will it happen in practise? Well, in higher orders we have to calculate the integrals $\langle \psi_n^{(0)} | \hat{H}^{(1)} | \psi_m^{(0)} \rangle$, where n, m stand for the vibrational quantum numbers. As we recall from Chapter 4 high-energy wave functions have large values for large x , where the perturbation changes as x^4 and gets larger and larger as x increases. This is why the integrals will be large. Therefore, the better we do our job (higher orders, higher-energy states) the faster we approach catastrophe.

Let us consider the opposite perturbation $\hat{H}^{(1)} = +0.000001 \cdot x^4$. Despite the fact that everything looks good (the perturbation does not qualitatively change the potential), the series will diverge sooner or later. It is bound to happen, because the

²⁷If the limit does not exist, then nothing can be said about ρ_k .

²⁸As a measure of the perturbation we may use $\langle \psi_0^{(0)} | \hat{H}^{(1)} | \psi_0^{(0)} \rangle$, which means an integral of x^4 multiplied by a Gaussian function (cf. Chapter 4). Such an integral is easy to calculate and, in view of the fact that it will be multiplied by the (small) factor 0.000001, the perturbation will turn out to be small.

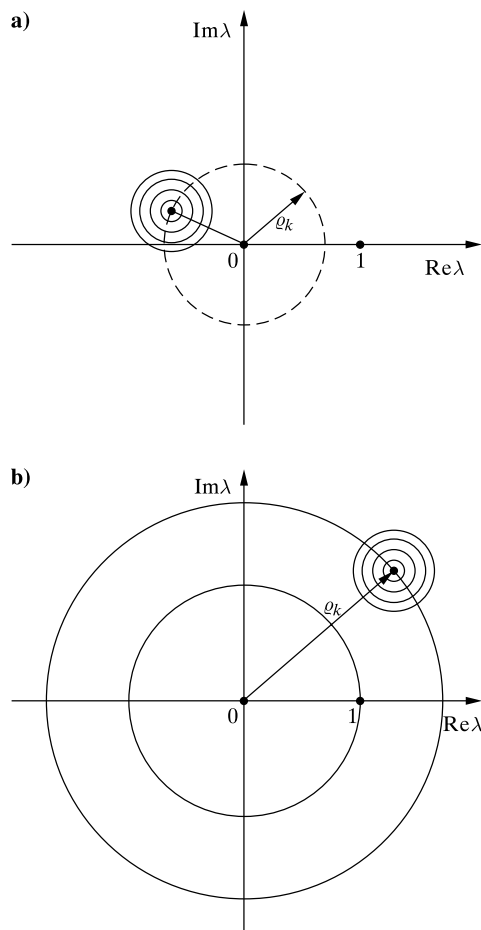


Fig. 5.3. The complex plane of the λ parameter. The physically interesting points are at $\lambda = 0, 1$. In perturbation theory we finally put $\lambda = 1$. Because of this the convergence radius ρ_k of the perturbational series has to be $\rho_k \geq 1$. However, if any complex λ with $|\lambda| < 1$ corresponds to a pole of the energy, the perturbational series will diverge in the physical situation ($\lambda = 1$). The figure shows the position of a pole by concentric circles. (a) The pole is too close ($\rho_k < 1$) and the perturbational series diverges; (b) the perturbational series converges, because $\rho_k > 1$.

convergence radius does not depend on the sign of the perturbation. A researcher might be astonished when the corrections begin to explode.

Quantum chemistry experiences with perturbational theories look quite consistent:

- low orders may give excellent results,
- higher orders often make the results worse.²⁹

Summary

There are basically two numerical approaches to obtain approximate solutions to the Schrödinger equation, variational and perturbational. In calculations we usually apply variational methods, while perturbational is often applied to estimate some small physical effects.

²⁹Even orders as high as 2000 have been investigated in the hope that the series will improve the results...

The result is that most concepts (practically all we know) characterizing the reaction of a molecule to an external field come from the perturbational approach. This leads to such quantities (see Chapter 12) as dipole moment, polarizability, hyperpolarizability, etc. The computational role of perturbational theories may, in this context, be seen as being of the second order.

• Variational method

- The method is based on the variational principle, which says that, if for a system with Hamiltonian \hat{H} we calculate the number $\varepsilon = \frac{\langle \Phi | \hat{H} | \Phi \rangle}{\langle \Phi | \Phi \rangle}$, where Φ stands for an arbitrary function, then the number $\varepsilon \geq E_0$, with E_0 being the ground-state energy of the system. If it happens that $\varepsilon[\Phi] = E_0$, then there is only one possibility: Φ represents the exact ground-state wave function ψ_0 .
- The variational principle means that to find an approximate ground-state wave function we can use the *variational method*: minimize $\varepsilon[\Phi]$ by changing (varying) Φ . The minimum value of $\varepsilon[\Phi]$ is equal to $\varepsilon[\Phi_{\text{opt}}]$ which approximates the ground-state energy E_0 and corresponds to Φ_{opt} , i.e. an approximation to the ground-state wave function ψ_0 .
- In practise the variational method consists from the following steps:
 - * make a decision as to the *trial function class*, among which the $\Phi_{\text{opt}}(\mathbf{x})$ will be sought³⁰
 - * introduce into the function the *variational parameters* $\mathbf{c} \equiv (c_0, c_1, \dots, c_P)$: $\Phi(\mathbf{x}; \mathbf{c})$. In this way ε becomes a function of these parameters: $\varepsilon(\mathbf{c})$
 - * minimize $\varepsilon(\mathbf{c})$ with respect to $\mathbf{c} \equiv (c_0, c_1, \dots, c_P)$ and find the optimal set of parameters $\mathbf{c} = \mathbf{c}_{\text{opt}}$
 - * the value $\varepsilon(\mathbf{c}_{\text{opt}})$ represents an approximation to E_0
 - * the function $\Phi(\mathbf{x}; \mathbf{c}_{\text{opt}})$ is an approximation to the ground-state wave function $\psi_0(\mathbf{x})$
- The Ritz procedure is a special case of the variational method, in which the parameters \mathbf{c} enter Φ *linearly*: $\Phi(\mathbf{x}; \mathbf{c}) = \sum_{i=0}^P c_i \Psi_i$, where $\{\Psi_i\}$ are some *known* basis functions that form (or more exactly, in principle form) the complete set of functions in the Hilbert space. This formalism leads to a set of homogeneous linear equations to solve (“secular equations”), from which we find approximations to the ground- and excited states energies and wave functions.

• Perturbational method

We assume that the solution to the Schrödinger equation for the *unperturbed system* is *known* ($E_k^{(0)}$ for the energy and $\psi_k^{(0)}$ for the wave function, usually $k = 0$, i.e. the ground state), but when a *small perturbation* $\hat{H}^{(1)}$ is added to the Hamiltonian, then the solution changes (to E_k and ψ_k , respectively) and is to be sought using the perturbational approach. Then the key assumption is: $E_k(\lambda) = E_k^{(0)} + \lambda E_k^{(1)} + \lambda^2 E_k^{(2)} + \dots$ and $\psi_k(\lambda) = \psi_k^{(0)} + \lambda \psi_k^{(1)} + \lambda^2 \psi_k^{(2)} + \dots$, where λ is a parameter that tunes the perturbation. The goal of the perturbational approach is to compute corrections to the energy: $E_k^{(1)}, E_k^{(2)}, \dots$ and to the wave function: $\psi_k^{(1)}, \psi_k^{(2)}, \dots$. We assume that because the perturbation is small, only a few such corrections are to be computed, in particular,

$$E_k^{(1)} = \langle \psi_k^{(0)} | \hat{H}^{(1)} | \psi_k^{(0)} \rangle, \quad E_k^{(2)} = \sum_{n (\neq k)} \frac{|H_{kn}^{(1)}|^2}{E_k^{(0)} - E_n^{(0)}}, \quad \text{where } H_{kn}^{(1)} = \langle \psi_k^{(0)} | \hat{H}^{(1)} | \psi_n^{(0)} \rangle.$$

³⁰ \mathbf{x} symbolizes the set of coordinates (space and spin, cf. Chapter 1).

Main concepts, new terms

variational principle (p. 196)	secular determinant (p. 203)
variational method (p. 196)	perturbational method (p. 203)
variational function (p. 196)	unperturbed system (p. 204)
variational principle for excited states (p. 199)	perturbed system (p. 204)
underground states (p. 199)	perturbation (p. 204)
variational parameters (p. 200)	corrections to energy (p. 205)
trial function (p. 200)	corrections to wave function (p. 205)
Ritz method (p. 202)	Hylleraas functional (p. 209)
complete basis set (p. 202)	Hylleraas variational principle (p. 209)
secular equation (p. 203)	Hylleraas equation (p. 210)
	asymptotic convergence (p. 210)

From the research front

In practise, the Ritz variational method is used most often. One of the technical problems to be solved is the size of the basis set. Enormous progress in computation and software development now facilitate investigations which 20 years ago were absolutely beyond the imagination. The world record in quantum chemistry means a few billion expansion functions. To accomplish this quantum chemists have had to invent some powerful methods of applied mathematics.

Ad futurum...

The computational technique impetus we witness nowadays will continue in the future (maybe in a modified form). It will be no problem to find some reliable approximations to the ground-state energy and wave function for a molecule composed of thousands of atoms. We will be effective. We may, however, ask whether such effectiveness is at the heart of science. Would it not be interesting to know what these ten billion terms in our wave function are telling us about and what we could learn from this?

Additional literature

E. Steiner, "The Chemistry Maths Book", Oxford University Press, Oxford, 1996.

A very good textbook. We may find some useful information there about the secular equation.

W.H. Press, B.P. Flannery, S.A. Teukolsky, W.T. Vetterling, "Numerical Recipes. The Art of Scientific Computing", Cambridge University Press, 1986. p. 19–77, 274–326, 335–381.

Probably the best textbook in computational mathematics, some chapters are very closely related to the topics of this chapter (diagonalization, linear equations).

H. Margenau and G.M. Murphy, "The Mathematics of Physics and Chemistry", D. van Nostrand Co., 1956.

An excellent old book dealing with most mathematical problems which we may encounter in chemistry and physics, including the variational and perturbational methods.

J.O. Hirschfelder, W. Byers Brown, S.T. Epstein, "Recent Developments in Perturbation Theory", *Adv. Quantum Chem.* 1 (1964) 255.

A long article on perturbation theory. For many years obligatory for those working in the domain.

Questions

- Variational method (Φ stands for the trial function, \hat{H} the Hamiltonian, E_0 the exact ground-state energy, and ψ_0 the exact ground-state wave function, $\varepsilon = \frac{\langle \Phi | \hat{H} | \Phi \rangle}{\langle \Phi | \Phi \rangle}$). If $\varepsilon = E_0$, this means that:
a) $\psi_0 = \Phi$; b) $|\Phi|^2 = 1$; c) $\psi_0 \leq \Phi$; d) $\psi_0 = E_0 \Phi$.
- In the Ritz method (Φ stands for the trial function, \hat{H} the Hamiltonian, E_0 the exact ground-state energy, ψ_0 the exact ground-state wave function, $\varepsilon = \frac{\langle \Phi | \hat{H} | \Phi \rangle}{\langle \Phi | \Phi \rangle}$) the trial function Φ is always a linear combination of:
a) orthonormal functions; b) unknown functions to be found in the procedure; c) eigenfunctions of \hat{H} ; d) known functions.
- A trial function used in the variational method for the hydrogen atom had the form: $\psi = \exp(-c_1 r) + c_2 \exp(-r/2)$. From a variational procedure we obtained:
a) $c_1 = c_2 = 0$; b) $c_1 = 1, c_2 = 0$; c) $c_1 = 0, c_2 = 1$; d) $c_1 = 1, c_2 = 1$.
- In the variational method applied to a molecule:
a) we search an approximate wave function in the form of a secular determinant;
b) we minimize the mean value of the Hamiltonian computed with a trial function;
c) we minimize the trial function with respect to its parameters;
d) we minimize the secular determinant with respect to the variational parameters.
- In a variational method, four classes of trial functions have been applied and the total energy computed. The exact value of the energy is equal to -50.2 eV. Choose the best approximation to this value obtained in correct calculations:
a) -48.2 eV; b) -50.5 eV; c) -45.3 eV; d) -43.0 eV.
- In the Ritz method (M terms) we obtain approximate wave functions only for:
a) the ground state; b) the ground state and M excited states; c) M states; d) one-electron systems.
- In the perturbational method for the ground state ($k = 0$):
a) the first-order correction to the energy is always negative;
b) the second-order correction to the energy is always negative;
c) the first-order correction to the energy is the largest among all the perturbational corrections;
d) the first-order correction to the energy is $E_k^{(1)} = \langle \psi_k^{(0)} | \hat{H}^{(1)} | \psi_k^{(0)} \rangle$, where $\psi_k^{(0)}$ stands for the unperturbed wave function, $\hat{H}^{(1)}$ is the perturbation operator and $\psi_k^{(1)}$ is the first-order correction to the wave function.
- Perturbation theory [$\hat{H}, \hat{H}^{(0)}, \hat{H}^{(1)}$ stand for the total (perturbed), unperturbed and perturbation Hamiltonian operators, $\psi_k^{(0)}$ the normalized unperturbed wave function of state k corresponding to the energy $E_k^{(0)}$]. The first-order correction to energy $E_k^{(1)}$ satisfies the following relation:
a) $E_k^{(1)} = \langle \psi_k^{(0)} | \hat{H}^{(1)} | \psi_k^{(0)} \rangle$; b) $E_k^{(1)} = \langle \psi_k^{(0)} | \hat{H}^{(0)} | \psi_k^{(0)} \rangle$; c) $E_k^{(1)} = \langle \psi_k^{(0)} | \hat{H}^{(0)} | \psi_k^{(1)} \rangle$;
d) $E_k^{(1)} = \langle \psi_k^{(1)} | \hat{H}^{(0)} | \psi_k^{(0)} \rangle$.
- In perturbation theory:
a) we can obtain accurate results despite the fact that the perturbation series diverges (converges asymptotically);

- b) the first-order correction to the energy has to be larger than the second-order correction;
- c) the wave function $\psi_k^{(0)}$ should not have any nodes;
- d) $E_k^{(1)} > 0$.
10. Perturbation theory [$\hat{H}, \hat{H}^{(0)}, \hat{H}^{(1)}$ stand for the total (perturbed), unperturbed and perturbation Hamiltonian operators, $\psi_k^{(0)}$ the normalized unperturbed wave function of state k corresponding to energy $E_k^{(0)}$]. The following equation is satisfied:
- a) $\hat{H}^{(0)}\psi_k^{(0)} = E_k^{(1)}\psi_k^{(1)} + E_k^{(0)}\psi_k^{(1)}$; b) $\hat{H}^{(1)}\psi_k^{(0)} = E_k^{(1)}\psi_k^{(0)} + E_k^{(0)}\psi_k^{(1)}$; c) $\hat{H}^{(0)}\psi_k^{(1)} + \hat{H}^{(1)}\psi_k^{(0)} = E_k^{(1)}\psi_k^{(0)} + E_k^{(0)}\psi_k^{(1)}$; d) $\hat{H}^{(0)}\psi_k^{(1)} + \hat{H}^{(1)}\psi_k^{(0)} = E_k^{(1)}\psi_k^{(0)}$.

Answers

1a, 2d, 3b, 4b, 5a, 6c, 7b, 8a, 9a, 10c