

Chapter 2

GENERAL THEOREMS AND PRINCIPLES

1. The variation principle

1.1. The Rayleigh quotient

For the solutions of the Schrödinger equations the expectation value (quantum mechanical average value) of the energy is equal to the eigenvalue of the Hamiltonian \hat{H} . In fact, the exact solutions are “energy eigenstates”:

$$\hat{H}\Psi = E\Psi . \quad (2.1)$$

We multiply (2.1) with Ψ^* and integrate:

$$\langle \Psi | \hat{H} | \Psi \rangle = E \langle \Psi | \Psi \rangle . \quad (2.2)$$

Dividing both sides with $\langle \Psi | \Psi \rangle$ we obtain the energy as an *expectation value* (“Rayleigh quotient”):

$$E = \frac{\langle \Psi | \hat{H} | \Psi \rangle}{\langle \Psi | \Psi \rangle} . \quad (2.3)$$

For approximate wave functions only the expectation value can be used because $\frac{\hat{H}\Psi}{\Psi} \neq \text{const.}$

1.2. The variation principle for the ground state

In the class of the wave functions satisfying the boundary conditions of the problem considered, the lowest possible expectation value E_0 of the energy E belongs to the exact ground state-wave function. Therefore, for any Ψ

$$E = \frac{\langle \Psi | \hat{H} | \Psi \rangle}{\langle \Psi | \Psi \rangle} \geq E_0 . \quad (2.4)$$

(The equality holds only for the exact ground-state solution.) Therefore an energy computed by using an approximate wave function gives an “upper limit” (upper bound) for the exact ground-state energy.

Proof.

The Hamiltonian \hat{H} is Hermitian; consequently its eigenvectors Ψ_i form a complete orthonormalized set of functions and any wave function Ψ can be expanded as

$$\Psi = \sum_i c_i \Psi_i . \quad (2.5)$$

(If \hat{H} has a continuous spectrum, then the sum should be replaced—or augmented—by an integration, and other complications arise; we shall not consider this possibility as it does not influence the essence.)

We calculate

$$\langle \Psi | \Psi \rangle = \left\langle \sum_i c_i \Psi_i \left| \sum_j c_j \Psi_j \right. \right\rangle = \sum_{i,j} c_i^* c_j \langle \Psi_i | \Psi_j \rangle = \sum_{i,j} c_i^* c_j \delta_{ij} = \sum_i |c_i|^2 \quad (2.6)$$

and

$$\begin{aligned} \langle \Psi | \hat{H} | \Psi \rangle &= \left\langle \sum_i c_i \Psi_i \left| \hat{H} \left| \sum_j c_j \Psi_j \right. \right. \right\rangle = \sum_{i,j} c_i^* c_j \langle \Psi_i | \hat{H} | \Psi_j \rangle \\ &= \sum_{i,j} c_i^* c_j E_j \langle \Psi_i | \Psi_j \rangle = \sum_{i,j} c_i^* c_j E_j \delta_{ij} = \sum_i |c_i|^2 E_i . \end{aligned} \quad (2.7)$$

Therefore:

$$E = \frac{\langle \Psi | \hat{H} | \Psi \rangle}{\langle \Psi | \Psi \rangle} = \frac{\sum_i |c_i|^2 E_i}{\sum_i |c_i|^2} . \quad (2.8)$$

Because $E_i \geq E_0$, we get

$$E \geq \frac{\sum_i |c_i|^2 E_0}{\sum_i |c_i|^2} = E_0 . \quad (2.9)$$

Q.E.D.

The fact that the expectation value of the energy represents a minimum indicates that the energy calculated in that manner is usually much more accurate than the wave function from which it is determined. In fact, in the minimum the energy as the functional of the wave function must be stationary, i.e., it must have a vanishing first variation. Therefore, the error in the energy is quadratic with respect to the error in the wave function: if the deviation of the wave function from the exact one is proportional to some (small) parameter ε , then the expectation value of the energy will differ from the exact one by a quantity

proportional to ε^2 . In practice this means that if we have determined the wave function expansion parameters with the accuracy of k decimal digits, then the energy is expected to be accurate up to about $2k$ digits, or even more.

1.3. The variation principle as an equivalent of the Schrödinger equation: A useful formulation of the variation principle

The Schrödinger equation can also be formulated as a variational problem: one has to find the wave function for which the energy (energy functional, expectation value of the Hamiltonian) is stationary:

$$\delta E = \delta \left[\frac{\langle \Psi | \hat{H} | \Psi \rangle}{\langle \Psi | \Psi \rangle} \right] = 0 \quad (2.10)$$

The vanishing first variation of the energy and the Schrödinger equation are equivalent to each other—not only for the ground state but for excited states, as well. (Not sufficiently recognized.) If one uses a minimization procedure then the orthogonality to the states with lower energy should also be required. Exceptions are the lowest states of each symmetry species for which the orthogonality is satisfied automatically—they are essentially the ground state of their own species.

In (2.10) one may perform the variation of the fraction in a manner quite analogous to differentiation (for the analogy between variations and differentials we refer to Appendix III):

$$\delta E = \frac{\langle \delta \Psi | \hat{H} | \Psi \rangle \langle \Psi | \Psi \rangle - \langle \delta \Psi | \Psi \rangle \langle \Psi | \hat{H} | \Psi \rangle}{\langle \Psi | \Psi \rangle^2} + c.c. = 0 \quad (2.11)$$

(*c.c.* = complex conjugate).

As $\delta \Psi$ is arbitrary, its phase is arbitrary as well (i.e., $\delta \Psi$ contains an arbitrary phase factor), therefore the expression written out explicitly and its complex conjugate should separately be equal to 0. After multiplying by $\langle \Psi | \Psi \rangle \neq 0$ (we consider only normalizable wave functions):

$$\underbrace{\langle \delta \Psi | \hat{H} | \Psi \rangle}_{1} \underbrace{\frac{\langle \Psi | \Psi \rangle}{\langle \Psi | \Psi \rangle}}_1 - \underbrace{\langle \delta \Psi | \Psi \rangle}_{1} \underbrace{\frac{\langle \Psi | \hat{H} | \Psi \rangle}{\langle \Psi | \Psi \rangle}}_E = 0. \quad (2.12)$$

That means

$$\langle \delta \Psi | \hat{H} - E | \Psi \rangle = 0, \quad (2.13)$$

which is a very useful form of the variation principle that we shall use throughout. As $\delta \Psi$ is arbitrary if no restrictions are imposed on Ψ , condition (2.13) is

also fully equivalent to the Schrödinger equation

$$(\hat{H} - E)\Psi = 0 . \quad (2.14)$$

The form in (2.13) we have derived for the variation principle is applicable also in the cases when we do not use an arbitrary trial wave function Ψ in the variation procedure but approximate the exact wave function by a given type (“Ansatz”) of variational wave function. In this case the variation $\delta\Psi$ is not fully arbitrary any more, but it must correspond to a variation of Ψ performed *within* the given class of wave functions. That means that the wave function $\Psi + \delta\Psi$ obtained after the variation must also belong to the given class, which puts restrictions on the possible variations $\delta\Psi$. In such cases (2.13) is not equivalent to the Schrödinger equation but permits to derive the respective variational equations corresponding to the given class of trial wave functions.

The use of Lagrangian multipliers

The equivalence of the variation principle with the Schrödinger equation can also be established by using the method of Lagrangian multipliers. The problem is then formulated in the following manner: let us request the functional (integral) $F = \langle \Psi | \hat{H} | \Psi \rangle$ to be stationary under the condition $\langle \Psi | \Psi \rangle = 1$. We form the auxiliary functional F' by adding to F the left-hand side of the auxiliary condition rewritten as

$$\langle \Psi | \Psi \rangle - 1 = 0 \quad (2.15)$$

multiplied by the undetermined Lagrangian multiplier λ :

$$F' = \langle \Psi | \hat{H} | \Psi \rangle + \lambda(\langle \Psi | \Psi \rangle - 1) \quad (2.16)$$

and request $\delta F'$ to vanish. One should join to the equation resulting from this variation also condition (2.15), or—which is the same—request F' to be stationary as a function of the scalar λ , as well. [From the requirement $\frac{\partial F'}{\partial \lambda} = 0$ one immediately recovers condition (2.15).] We get

$$\langle \delta\Psi | \hat{H} | \Psi \rangle + \lambda \langle \delta\Psi | \Psi \rangle + c. c. = 0 . \quad (2.17)$$

As $\delta\Psi$ is arbitrary, and also contains an arbitrary phase factor, it follows from (2.17) that

$$\hat{H}\Psi + \lambda\Psi = 0 . \quad (2.18)$$

We now multiply (2.18) by Ψ^* and integrate by taking into account (2.15), and obtain

$$\langle \Psi | \hat{H} | \Psi \rangle + \lambda = 0 , \quad (2.19)$$

i.e., $\lambda = -E$. Substituting this into (2.18) we have

$$\hat{H}\Psi - E\Psi = 0 . \quad (2.20)$$

Q.E.D.

1.4. Eckart's inequality

The variation principle tells us that the exact ground state has the absolute minimum of the energy, and Eckart's inequality shows that—once we are close enough to the solution—decreasing the energy of the trial function may be expected to improve the wave function, as well; in the limiting case, when the energy is approaching the exact value, then the wave function should also approach the exact solution.

For sake of simplicity, we shall assume all wave functions considered in this section to be normalized to 1.

Let us consider a trial wave function Ψ ; it is a linear combination of the exact eigenfunctions of the Hamiltonian with the coefficients c_i :

$$\Psi = \sum_{i=0} c_i \Psi_i \quad (2.21)$$

and we have requested

$$\langle \Psi | \Psi \rangle = \sum_{i=0} |c_i|^2 = 1. \quad (2.22)$$

We consider the overlap (scalar product, inner product) of the wave function Ψ and the exact ground-state solution Ψ_0

$$S = \langle \Psi_0 | \Psi \rangle = c_0 \quad (2.23)$$

where the last equality follows from the orthonormalization of the exact solutions Ψ_i . One may write

$$\begin{aligned} E &= \sum_{i=0} |c_i|^2 E_i = |S|^2 E_0 + \sum_{i=1} |c_i|^2 E_i \quad (\text{as } E_i \geq E_1 \text{ if } i \geq 1) \\ &\geq |S|^2 E_0 + \sum_{i=1} |c_i|^2 E_1 = |S|^2 E_0 + E_1 \sum_{i=1} |c_i|^2. \end{aligned} \quad (2.24)$$

As $\sum_{i=0} |c_i|^2 = 1$ and $|c_0|^2 = |S|^2$, we have obviously $\sum_{i=1} |c_i|^2 = 1 - |S|^2$, so

$$E \geq |S|^2 E_0 + (1 - |S|^2) E_1 \quad (2.25)$$

from which

$$|S|^2 \geq \frac{E_1 - E}{E_1 - E_0}. \quad (2.26)$$

Therefore, if $E \rightarrow E_0$, then $|S| \rightarrow 1$. ($|S| \leq 1$ according to its definition.)

Q.E.D.

1.5. Excited states

The Hamiltonian is a Hermitian operator; its eigenvectors belonging to different eigenvalues are automatically orthogonal; those corresponding to the same (degenerate) eigenvalue can also be chosen orthogonal. However, as already noted in Section 1.3, when one searches for the eigenvector corresponding to an excited state, then one may meet the problem of the “variational collapse”: the calculation is “trapped” by the ground-state solution, i.e. it converges back to the latter instead of yielding the required excited state. In such situations it may be necessary to introduce the explicit requirement of the orthogonality to the lower lying state(s) already determined. (We shall consider explicitly only one such state; the generalization to the case of several such ones is trivial.) The following considerations are applicable also in the case when the ground-state wave function considered is not the exact one but only some approximation to the latter.

Accordingly, we should solve the following problem. One has to find the wave functions $|\Psi\rangle$ having a stationary energy within the subspace of functions which are orthogonal to some (exact or approximate) ground-state wave function $|\Psi_0\rangle$. (It is convenient to use the “bra-ket” notations —see Appendix VII—in this section.) That means that we should request

$$(1 - \hat{P})|\Psi\rangle = |\Psi\rangle \quad (2.27)$$

where

$$\hat{P} = |\Psi_0\rangle\langle\Psi_0| \quad (2.28)$$

is the projector on the normalized ground state wave function $|\Psi_0\rangle$.

As we are searching a wave function within the subspace orthogonal to $|\Psi_0\rangle$, the variation should not move $|\Psi\rangle$ out of this subspace, either. Consequently, one must consider only variations $|\delta\Psi\rangle$ that are also within the same subspace; this can be achieved by writing the most general permitted variation in the form $|\delta\Psi\rangle = (1 - \hat{P})|\delta\Phi\rangle$ where $|\delta\Phi\rangle$ is fully arbitrary. Substituting this variation in the general form (2.13) of the variation principle, we get

$$\langle\delta\Phi|(1 - \hat{P})(\hat{H} - E)|\Psi\rangle = 0 . \quad (2.29)$$

Utilizing that $|\delta\Phi\rangle$ is fully arbitrary, this gives

$$(1 - \hat{P})\hat{H}|\Psi\rangle = E(1 - \hat{P})|\Psi\rangle . \quad (2.30)$$

This equation remains valid if one substitutes the restriction $(1 - \hat{P})|\Psi\rangle = |\Psi\rangle$ on the right-hand side:

$$(1 - \hat{P})\hat{H}|\Psi\rangle = E|\Psi\rangle . \quad (2.31)$$

In fact, by multiplying (2.31) with $(1 - \hat{P})$ and utilizing the idempotency of this operator, we see that all solutions of (2.31) belonging to nonzero eigenvalues E are automatically orthogonal to $|\Psi_0\rangle$: $|\Psi\rangle = (1 - \hat{P})|\Psi\rangle$.

Furthermore, one can utilize this equality once again, and insert a projector $(1 - \hat{P})$ on the left-hand side of (2.31) and get the eigenvalue equation of a Hermitian operator:

$$(1 - \hat{P})\hat{H}(1 - \hat{P})|\Psi\rangle = E|\Psi\rangle, \quad (2.32)$$

which is essentially the eigenvalue equation of the Hamiltonian projected on the subspace of functions orthogonal to the ground state $|\Psi_0\rangle$. It has the property that the ground-state wave function $|\Psi_0\rangle$ is also its eigenvector with a zero eigenvalue. (Therefore, the required orthogonality is also guaranteed by the orthogonality of the eigenvectors of the Hermitian operators.)

The use of Lagrangian multipliers

The previous task can also be solved by using the method of Lagrangian multipliers. The problem is then formulated in the following manner: let us request the functional $F = \langle\Psi|\hat{H}|\Psi\rangle$ to be stationary under the conditions $\langle\Psi|\Psi\rangle = 1$ and $\langle\Psi|\Psi_0\rangle = 0$.

$\langle\Psi|\Psi\rangle$ and $\langle\Psi|\hat{H}|\Psi\rangle$ are necessarily real, but the condition $\langle\Psi|\Psi_0\rangle = 0$ in the general complex case represents two real conditions $Re\langle\Psi|\Psi_0\rangle = 0$ and $Im\langle\Psi|\Psi_0\rangle = 0$. Accordingly, we form the auxiliary functional F' by introducing three Lagrangian multipliers, λ , μ , and ν :

$$F' = \langle\Psi|\hat{H}|\Psi\rangle + \lambda(\langle\Psi|\Psi\rangle - 1) + \mu Re\langle\Psi|\Psi_0\rangle + \nu Im\langle\Psi|\Psi_0\rangle. \quad (2.33)$$

This can appropriately rewritten as follows. One has

$$Re\langle\Psi|\Psi_0\rangle = \frac{1}{2}(\langle\Psi|\Psi_0\rangle + \langle\Psi_0|\Psi\rangle) \quad (2.34)$$

and

$$Im\langle\Psi|\Psi_0\rangle = \frac{1}{2i}(\langle\Psi|\Psi_0\rangle - \langle\Psi_0|\Psi\rangle), \quad (2.35)$$

We substitute these expressions into (2.33), collect the terms containing $\langle\Psi|\Psi_0\rangle$ and $\langle\Psi_0|\Psi\rangle$, introduce the notations $\tau = \frac{1}{2}(\mu - i\nu)$ and $\tau^* = \frac{1}{2}(\mu + i\nu)$ and get

$$F' = \langle\Psi|\hat{H}|\Psi\rangle + \lambda(\langle\Psi|\Psi\rangle - 1) + \tau\langle\Psi|\Psi_0\rangle + \tau^*\langle\Psi_0|\Psi\rangle. \quad (2.36)$$

We request $\delta F' = 0$, i.e.,

$$\delta F' = \langle\delta\Psi|\hat{H}|\Psi\rangle + \lambda\langle\delta\Psi|\Psi\rangle + \tau\langle\delta\Psi|\Psi_0\rangle + c.c. = 0. \quad (2.37)$$

From this ($\delta\Psi$ is arbitrary, contains also an arbitrary phase factor) we get the system of equations

$$\begin{aligned}\hat{H}|\Psi\rangle + \lambda|\Psi\rangle + \tau|\Psi_0\rangle &= 0 \\ \langle\Psi|\Psi\rangle &= 1 \\ \langle\Psi|\Psi_0\rangle &= 0\end{aligned}\quad (2.38)$$

By multiplying the first equation of (2.38) with $\langle\Psi|$ and $\langle\Psi_0|$, respectively, and utilizing the auxiliary conditions—the second and third equations of the system—we get

$$\lambda = -E \quad (2.39)$$

and

$$\tau = -\langle\Psi_0|\hat{H}|\Psi\rangle. \quad (2.40)$$

($|\Psi_0\rangle$ has been assumed normalized.) By substituting these values into the first equation of (2.38) we arrive at

$$\hat{H}|\Psi\rangle - |\Psi_0\rangle\langle\Psi_0|\hat{H}|\Psi\rangle = E|\Psi\rangle \quad (2.41)$$

that is

$$(1 - \hat{P})\hat{H}|\Psi\rangle = E|\Psi\rangle \quad (2.42)$$

We have obtained (2.31) that has already been discussed; it may be Hermitized in the same manner as described previously.

2. The Hellmann–Feynman theorem

When deriving the variation principle, we have considered arbitrary “mathematical” variations of the wave function for a given system (given Hamiltonian) in order to investigate the behavior of the energy functional under the slight changes of the wave function in the vicinity of the exact solution. Contrary to this, the Hellmann–Feynman theorem deals with the changes induced by the variations in the physical parameters of the system (parameters defining its Hamiltonian). Nevertheless, they have many aspects in common.

2.1. The differential Hellmann–Feynman theorem

Let us consider the case when some parameter α describing the system undergoes a change

$$\alpha \rightarrow \alpha + d\alpha. \quad (2.43)$$

Therefore, there will be a change in the Hamiltonian \hat{H} and, as a consequence, in Ψ and E , as well:

$$\hat{H} \rightarrow \hat{H} + \frac{\partial\hat{H}}{\partial\alpha}d\alpha$$