# Chapter 5
# Approximating Molecular Schrödinger Equation

**Abstract** Theoretical basis of the approximate *perturbational* and *variational* approaches in quantum chemistry is outlined and the *adiabatic* separation of the fast (electronic) motions from slow (nuclear) movements in molecular systems is established. The rudiments of the *Ritz* method, a linear variant of the variational treatment, are summarized and the criteria for an effective mixing of quantum states are formulated. The illustrative applications of the perturbative and variational methods to helium atom are discussed and compared. The elements of the *orbital approximation* of the *many*-electron wave functions are introduced and selected properties of the *Slater determinant*, defined by the antisymmetrized product of the occupied *spin orbitals*, are examined in the context of the *Pauli exclusion principle*. The relevant expression for the expectation value of the electronic energy in orbital theories is derived and the *Slater–Condon* rules for matrix elements of the electronic Hamiltonian between determinantal wave functions are given. The additional possibilities of reducing the complexity of the molecular electronic Schrödinger equation by using the *pseudopotentials* are briefly outlined. These *core* potentials reflect a resultant influence of the “frozen” (chemically inactive) *inner*-shell electrons and the system nuclei in the effective Schrödinger equation for the (chemically active) *valence* shell electrons of constituent atoms, coordinates of which are treated explicitly in the approximate wave functions.

## 5.1 Rudiments of Perturbational and Variational Approaches

The stationary (time-independent) Schrödinger equation, i.e., the eigenvalue problem of the system Hamiltonian, can be solved analytically only for simple model systems. The quantum mechanical determination of the electronic structure of molecules, and particularly the complicated systems of interest in contemporary chemistry, requires adequate approximate methods of sufficient accuracy. In recent decades a remarkable progress of applying quantum mechanics to diverse problems in physics, chemistry, and molecular biology was possible due to spectacular

R.F. Nalewajski, *Perspectives in Electronic Structure Theory*,
DOI 10.1007/978-3-642-20180-6_5, © Springer-Verlag Berlin Heidelberg 2012

developments in the approximate theories of molecular electronic structure, covering original and sometimes ingenious new concepts and efficient algorithms, as well as a steadily increasing capability of modern computers and new software techniques of the advanced computational tools of modern quantum chemistry and solid state physics.

It is the main purpose of this chapter to summarize the main strategies used in reducing the complexity of the molecular Schrödinger equation and approximating its electronic wave function. It is intended to provide an overview of the successive levels of reducing the immense computational complexity of treating the coupled $N$-electron and $m$-nuclei problem of the molecular quantum mechanics. These perturbational and variational methods use the adiabatic, *Born–Oppenheimer* (BO) separation of the electronic and nuclear motions in molecules, as well as the orbital (Slater determinant) approximation of the trial $N$-electron wave functions, which automatically satisfy the requirements of the Pauli exclusion principle.

### 5.1.1 Perturbation Theory

It is the often encountered scenario in quantum theory that we have to estimate solutions of the Schrödinger equation for a more complicated (perturbed) *real* system from the known solutions of a simpler (unperturbed) *model* system, e.g., the stationary states and the associated energy levels of an anharmonic oscillator from the known (analytical) results for the harmonic oscillator. This goal summarizes the basic purpose of the *perturbation theory* (PT), which has also been used in classical mechanics. Its simplest variant within the Rayleigh–Schrödinger theory, for the nondegenerate energy levels and time-independent perturbations, will be summarized below.

Let us assume that the Hamiltonian $\hat{\mathrm{H}}$ of the real (perturbed) system can be expressed as the sum of the simpler, model Hamiltonian $\hat{\mathrm{H}}^0$, representing the associated unperturbed system the eigensolutions of which are assumed to be available, and the *perturbation* $\hat{\mathrm{h}} \equiv \lambda \hat{\mathrm{h}}$' including weak interactions compared with those already comprised in $\hat{\mathrm{H}}^0$. The perturbation approach can be then used to generate corrections to eigensolutions of $\hat{\mathrm{H}}^0$, due to a presence of the perturbation, to approximate the exact eigensolutions of $\hat{\mathrm{H}}$. Formally, this assumption of a relative "smallness" of $\hat{\mathrm{h}}$ can be expressed by the condition involving the perturbation parameter $\lambda$, $|\lambda| \ll 1$,

$$\hat{\mathrm{H}} = \hat{\mathrm{H}}^0 + \lambda \hat{\mathrm{h}}' = \hat{\mathrm{H}}^0 + \hat{\mathrm{h}}(\lambda) \equiv \hat{\mathrm{H}}(\lambda). \tag{5.1}$$

It controls the *order* of corrections to the known unperturbed solutions,

$$\hat{\mathrm{H}}^0 |n^{(0)}\rangle = E_n^{(0)} |n^{(0)}\rangle, \quad n = 0, 1, 2, \ldots, \quad \langle n^{(0)} \mid m^{(0)} \rangle = \delta_{n,m}, \tag{5.2}$$

for the nondegenerate energy levels $E_0^{(0)} < E_1^{(0)} < E_2^{(0)} < \ldots$, with $n = 0$ corresponding to the ground state of the model system, introduced to approximate the unknown stationary states of the perturbed system:

$$\hat{\mathrm{H}}(\lambda)|n(\lambda)\rangle = E_n(\lambda)|n(\lambda)\rangle, \quad n = 0, 1, 2, \ldots, \quad \langle n(\lambda) \mid m(\lambda)\rangle = \delta_{n,m}. \tag{5.3}$$

These corrections appear as coefficients in the corresponding power series expansions of the perturbed eigenstates and the associated eigenvalues,

$$\begin{aligned} |n(\lambda)\rangle &= \sum_{i=0}^{\infty} \left|n^{(i)}\right\rangle \lambda^i \equiv \left|n^{(0)}\right\rangle + \sum_{i=1}^{\infty} \left|\Delta n^{(i)}\right\rangle, \\ E_n(\lambda) &= \sum_{j=0}^{\infty} E_n^{(j)} \lambda^j \equiv E_n^{(0)} + \sum_{j=1}^{\infty} \Delta E_n^{(j)}, \end{aligned} \tag{5.4}$$

which define the *k*th-order corrections to the *n*th unperturbed state,

$$\left|\Delta n^{(k)}\right\rangle = \lambda^k \left|n^{(k)}\right\rangle \quad \text{and} \quad \Delta E_n^{(k)} = \lambda^k E_n^{(k)}, \quad k = 1, 2, \ldots$$

They can be determined by substituting these expansions into (5.3):

$$\sum_{i=0}^{\infty} \lambda^i (\hat{\mathrm{H}}^0 + \lambda \hat{\mathrm{h}}') \left|n^{(i)}\right\rangle = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \lambda^{i+j} E_n^{(j)} \left|n^{(i)}\right\rangle. \tag{5.5}$$

Indeed, by comparing the coefficients at the given power *k* of the enhancement parameter $\lambda$ in both sides of the preceding equation, one arrives at the following system of equations determining the corrections to the *n*th unperturbed state and its energy:

$$\begin{aligned} &\lambda^0: && \hat{\mathrm{H}}^0 \left|n^{(0)}\right\rangle = E_n^{(0)} \left|n^{(0)}\right\rangle; \\ &\lambda^1: && \hat{\mathrm{H}}^0 \left|n^{(1)}\right\rangle + \hat{\mathrm{h}}' \left|n^{(0)}\right\rangle = E_n^{(0)} \left|n^{(1)}\right\rangle + E_n^{(1)} \left|n^{(0)}\right\rangle; \\ &\lambda^2: && \hat{\mathrm{H}}^0 \left|n^{(2)}\right\rangle + \hat{\mathrm{h}}' \left|n^{(1)}\right\rangle = E_n^{(0)} \left|n^{(2)}\right\rangle + E_n^{(1)} \left|n^{(1)}\right\rangle + E_n^{(2)} \left|n^{(0)}\right\rangle; \\ &\lambda^p: && \hat{\mathrm{H}}^0 \left|n^{(p)}\right\rangle + \hat{\mathrm{h}}' \left|n^{(p-1)}\right\rangle = \sum_{j=0}^{p} E_n^{(j)} \left|n^{(p-j)}\right\rangle. \end{aligned} \tag{5.6}$$

As expected, the $\lambda^0$-equation repeats the eigenvalue problem (5.2) of the unperturbed Hamiltonian. The subsequent elimination of corrections from these equations recognizes the completeness of the unperturbed solutions $\left\{\left|n^{(0)}\right\rangle\right\}$, which allows one to expand any state of the system, including all unknown corrections $\left\{\left|n^{(p)}\right\rangle\right\}$ or $\left\{\left|\Delta n^{(p)}\right\rangle\right\}$, in this basis set.

For example, one can expand the *first*-order correction $|n^{(1)}\rangle$,

$$\left|n^{(1)}\right\rangle = \sum_{j=0}^{\infty} \left|j^{(0)}\right\rangle c_{j,n}^{(1)} = \sum_{j\neq n} \left|j^{(0)}\right\rangle c_{j,n}^{(1)}, \quad c_{j,n}^{(1)} = \left\langle j^{(0)}\middle|n^{(1)}\right\rangle, \tag{5.7}$$

or the resultant state $\hat{\mathrm{h}}'\left|n^{(0)}\right\rangle$:

$$\hat{\mathrm{h}}'\left|n^{(0)}\right\rangle = \sum_{j=0}^{\infty} \left|j^{(0)}\right\rangle h_{j,n}, \quad h_{j,n} = \left\langle j^{(0)}\middle|\hat{\mathrm{h}}'\middle|n^{(0)}\right\rangle. \tag{5.8}$$

In (5.7) we have recognized that $c_{n,n}^{(1)} = \left\langle n^{(0)}\middle|n^{(1)}\right\rangle = 0$, since the *direction* of the unperturbed state vector $|n^{(0)}\rangle$ can be modified only by combining this state with the remaining states $\{|j^{(0)}\rangle, j \neq n\}$, which are orthogonal to $|n^{(0)}\rangle$.

Projecting $\lambda^1$-equation (5.6) onto $|n^{(0)}\rangle$ and $|k^{(0)}\rangle$, $k \neq n$, respectively, gives the associated equations for determining the *first*-order corrections we seek:

$$\begin{aligned} \left\langle n^{(0)}\middle|\hat{\mathrm{H}}^0\middle|n^{(1)}\right\rangle + \left\langle n^{(0)}\middle|\hat{\mathrm{h}}'\middle|n^{(0)}\right\rangle &= E_n^{(0)}\left\langle n^{(0)}\middle|n^{(1)}\right\rangle + h_{n,n} = h_{n,n} \\ &= E_n^{(0)}\left\langle n^{(0)}\middle|n^{(1)}\right\rangle + E_n^{(1)}\left\langle n^{(0)}\middle|n^{(0)}\right\rangle = E_n^{(1)}, \end{aligned} \tag{5.9}$$

$$\begin{aligned} \left\langle k^{(0)}\middle|\hat{\mathrm{H}}^0\middle|n^{(1)}\right\rangle + \left\langle k^{(0)}\middle|\hat{\mathrm{h}}'\middle|n^{(0)}\right\rangle &= E_k^{(0)}\left\langle k^{(0)}\middle|n^{(1)}\right\rangle + h_{k.n} = E_k^{(0)}c_{k,n}^{(1)} + h_{k,n} \\ &= E_n^{(0)}\left\langle k^{(0)}\middle|n^{(1)}\right\rangle + E_n^{(1)}\left\langle k^{(0)}\middle|n^{(0)}\right\rangle = E_n^{(0)}c_{k,n}^{(1)}. \end{aligned} \tag{5.10}$$

A straightforward rearrangements of these equations then give the following explicit expressions for the *first*-order corrections to $E_n^{(0)}$,

$$E_n^{(1)} = \left\langle n^{(0)}\middle|\hat{\mathrm{h}}'\middle|n^{(0)}\right\rangle = h_{n,n} \quad \text{or} \quad \Delta E_n^{(1)} = \left\langle n^{(0)}\middle|\hat{\mathrm{h}}\middle|n^{(0)}\right\rangle, \tag{5.11}$$

and to $|n^{(0)}\rangle$:

$$c_{k,n}^{(1)} = h_{k,n}/[E_n^{(0)} - E_k^{(0)}] \quad \text{or} \quad \left|\Delta n^{(1)}\right\rangle = \sum_{k\neq n} \frac{\left\langle k^{(0)}\middle|\hat{\mathrm{h}}\middle|k^{(0)}\right\rangle}{E_n^{(0)} - E_k^{(0)}} \left|k^{(0)}\right\rangle. \tag{5.12}$$

When determining the *second*-order corrections one similarly expands

$$\left|n^{(2)}\right\rangle = \sum_{j=0}^{\infty} \left|j^{(0)}\right\rangle c_{j,n}^{(2)} = \sum_{j\neq n} \left|j^{(0)}\right\rangle c_{j,n}^{(2)}, \quad c_{j,n}^{(2)} = \left\langle j^{(0)}\middle|n^{(2)}\right\rangle, \tag{5.13}$$

again realizing that $c_{n,n}^{(2)} = \left\langle n^{(0)}\middle|n^{(2)}\right\rangle = 0$. The corresponding projections of the $\lambda^2$-equation (5.6) onto $|n^{(0)}\rangle$ and $|k^{(0)}\rangle$, $k \neq n$, respectively, gives the relevant equations determining the expansion coefficients $\{c_{j,n}^{(2)}\}$ and the *second*-order energy:

$$\begin{aligned}\langle n^{(0)}|\hat{\mathrm{H}}^0|n^{(2)}\rangle + \langle n^{(0)}|\hat{\mathrm{h}}'|n^{(1)}\rangle &= E_n^{(0)}\langle n^{(0)}|n^{(2)}\rangle + \sum_{l\neq n} c_{l,n}^{(1)}\langle n^{(0)}|\hat{\mathrm{h}}'|l^{(0)}\rangle \\ = \sum_{l\neq n} c_{l,n}^{(1)} h_{n,l} = E_n^{(0)}\langle n^{(0)}|n^{(2)}\rangle &+ E_n^{(1)}\langle n^{(0)}|n^{(1)}\rangle + E_n^{(2)}\langle n^{(0)}|n^{(0)}\rangle = E_n^{(2)},\end{aligned} \tag{5.14}$$

$$\begin{aligned}\langle k^{(0)}|\hat{\mathrm{H}}^0|n^{(2)}\rangle + \langle k^{(0)}|\hat{\mathrm{h}}'|n^{(1)}\rangle &= E_k^{(0)}\langle k^{(0)}|n^{(2)}\rangle + \sum_{l\neq n} c_{l,n}^{(1)}\langle k^{(0)}|\hat{\mathrm{h}}'|l^{(0)}\rangle \\ &= E_k^{(0)} c_{k,n}^{(2)} + \sum_{l\neq n} c_{l,n}^{(1)} h_{k,l} \\ &= E_n^{(0)}\langle k^{(0)}|n^{(2)}\rangle + E_n^{(1)}\langle k^{(0)}|n^{(1)}\rangle + E_n^{(2)}\langle k^{(0)}|n^{(0)}\rangle \\ &= E_n^{(0)} c_{k,n}^{(2)} + E_n^{(1)} c_{k,n}^{(1)}.\end{aligned} \tag{5.15}$$

Subsequent substitution to (5.14) of the known *first*-order solutions gives the following expression for the *second*-order correction to $E_n^{(0)}$:

$$\begin{aligned}E_n^{(2)} &= \langle n^{(0)}|\hat{\mathrm{h}}'|n^{(1)}\rangle = \sum_{l\neq n} h_{n,l} c_{l,n}^{(1)} = \sum_{l\neq n} h_{n,l} h_{l,n}/[E_n^{(0)} - E_l^{(0)}] \quad \text{or} \\ \Delta E_n^{(2)} &= \langle n^{(0)}|\hat{\mathrm{h}}|\Delta n^{(1)}\rangle = \sum_{l\neq n} \left|\langle n^{(0)}|\hat{\mathrm{h}}|l^{(0)}\rangle\right|^2/[E_n^{(0)} - E_l^{(0)}].\end{aligned} \tag{5.16}$$

A similar rearrangement of (5.15) gives the expansion coefficients

$$c_{k,n}^{(2)} = \frac{1}{E_n^{(0)} - E_k^{(0)}}\left[\sum_{l\neq n} \frac{h_{k,l} h_{l,n}}{E_n^{(0)} - E_l^{(0)}} - \frac{h_{k,n} h_{n,n}}{E_n^{(0)} - E_k^{(0)}}\right]$$

determining the associated correction to $|n^{(0)}\rangle$:

$$|\Delta n^{(2)}\rangle = \sum_{k\neq n}\left(\sum_{l\neq n} \frac{\langle k^{(0)}|\hat{\mathrm{h}}|l^{(0)}\rangle\langle l^{(0)}|\hat{\mathrm{h}}|n^{(0)}\rangle}{(E_n^{(0)} - E_l^{(0)})(E_n^{(0)} - E_k^{(0)})} - \frac{\langle k^{(0)}|\hat{\mathrm{h}}|n^{(0)}\rangle\langle n^{(0)}|\hat{\mathrm{h}}|n^{(0)}\rangle}{(E_n^{(0)} - E_k^{(0)})^2}\right)|k^{(0)}\rangle. \tag{5.17}$$

Obviously, one could similarly extract the *higher* order corrections, but the above explicit expressions for the *first*- and *second*-order corrections are sufficient for most applications included in this book.

### 5.1.2 Variational Method

The alternative *variational* method of determining the approximate solutions of the time-independent Schrödinger equation guarantees that the successive approximations of increasing accuracy approach from above the exact energy level $E_0$ of the molecular ground state $|\psi_0\rangle$. In other words, this exact eigenvalue represents the lower bound of all approximate estimates of the system average energy: $\langle E\rangle \geq E_0$.

Indeed, the eigenstates $\{|\psi_n\rangle\}$ of the system Hamiltonian (the quantum mechanical observable),

$$\hat{\mathrm{H}}|\psi_n\rangle = E_n|\psi_n\rangle, \quad n = 0, 1, 2, \ldots, \quad E_0 \leq E_1 \leq E_2 \leq \ldots, \tag{5.18}$$

form the basis of the energy representation (see Sect. 2.7) in the molecular Hilbert space, so that any approximate state $|\phi\rangle$ can be expanded in this set:

$$|\phi\rangle = \sum_n |\psi_n\rangle\langle\psi_n|\phi\rangle = \sum_n |\psi_n\rangle C_n, \tag{5.19}$$

with $|C_n|^2 = P(\psi_n|\phi)$ measuring the conditional probability of observing $E_n$ in state $|\phi\rangle$ (see Postulate II of Sect. 3.2). Hence, any approximate estimate of the system average energy can be expressed as the mean value of the exact energy levels (see also Postulate IV.3 of Sect. 3.3.3):

$$\langle E\rangle_\phi = \langle\phi|\hat{\mathrm{H}}|\phi\rangle = \sum_n P(\psi_n|\phi)E_n, \quad \sum_n P(\psi_n|\phi) = 1.$$

We thus conclude that $\langle E\rangle_\phi = E_0$ can be reached only for $P(\psi_0|\phi) = 1$ and $\{P(\psi_{n>0}|\phi) = 0\}$, and hence $|\phi\rangle = |\psi_0\rangle$. Any deviation from this exact solution implies a finite probability of observing one of the higher (excited) energy levels, and hence $\langle E\rangle_\phi > E_0$. These deductions constitute the essence of the Rayleigh–Ritz variational principle of quantum mechanics: for any approximate state $|\phi\rangle$ the average energy

$$\langle E\rangle_\phi \geq E_0. \tag{5.20}$$

Thus, the more accurately $|\phi\rangle$ approximates $|\psi_0\rangle$, the lower $\langle E\rangle_\phi$ level, and hence the smaller $\langle E\rangle_\phi - E_0$ error gap.

This general statement gives rise to the efficient computational technique, the *variational method*, which dominates the modern quantum mechanical calculations of molecular electronic structure. The main idea behind this computational tool is to use the parametrically defined trial state including several *variational parameters* $\boldsymbol{\lambda} = \{\lambda_t, t = 1, 2, \ldots, s\}$, $|\phi\rangle = |\phi(\lambda_1, \lambda_2, \ldots, \lambda_s)\rangle$. The domain of their admissible values then determines the whole range of the approximate (variational) states. In accordance with the variational principle, the best approximation of the molecular ground state in this family of trial states is then obtained for the optimum values of variational parameters $\boldsymbol{\lambda}^{opt.}$ which correspond to the minimum of the system average energy $\langle E(\boldsymbol{\lambda})\rangle_{\phi} = \langle\phi|(\boldsymbol{\lambda})|\hat{\mathrm{H}}|\phi(\boldsymbol{\lambda})\rangle$:

$$\min_{\boldsymbol{\lambda}}\langle E(\boldsymbol{\lambda})\rangle_{\phi} = \langle E(\boldsymbol{\lambda}^{opt})\rangle_{\phi} \geq E_0 \quad \text{or}$$
$$\left.\frac{\partial\langle E(\boldsymbol{\lambda})\rangle_{\phi}}{\partial\lambda_t}\right|_{\boldsymbol{\lambda}^{opt.}} = 0, \quad t = 1, 2, \ldots, s. \tag{5.21}$$

Both *linear* and *nonlinear* parameters $\boldsymbol{\lambda}$ are used to provide the trial state vectors or the associated wave functions exhibiting a sufficient variational flexibility, so that they are capable to adjust to the interactions embodied in the system Hamiltonian, in order to lower the energy, and thus to resemble the most the true ground state of the molecular system in question. The former, e.g., the coefficients multiplying the adopted set of the ("frozen") *basis functions*, are more easily handled, giving rise to a system of linear *secular equations* for determining the optimum values of the expansion coefficients. The latter, e.g., the exponents of the *Slater*-type orbitals (STO) or *Gaussian*-type orbitals (GTO), the popular analytical functions used to approximate the atomic or molecular orbitals, although relatively more efficient in modifying the trial wave functions, are more difficult to handle, requiring more advanced, nonlinear optimization techniques.

Consider the illustrative application of this procedure to the hydrogen-like atom of Chap. 4, for simplicity adopting a.u. of Sect. 4.6. Suppose that we take the trial wave function in the general form of a parametric family of the spherically symmetric, exponentially decaying functions defined by a single nonlinear variational parameter $\lambda$, $\phi(r, \vartheta, \varphi; \lambda) = N(\lambda) \exp(-\lambda r)$, with $N(\lambda)$ standing for the appropriate normalization factor [see (4.62)]: $N(\lambda) = (\lambda^3/\pi)^{1/2}$. It gives rise to the average electronic energy, the expectation value of the Hamiltonian (4.40),

$$E(\lambda) = \frac{1}{2}\lambda^2 - \lambda Z. \tag{5.22}$$

The optimum value of $\lambda$, which identifies the best approximation to the ground state, is then obtained for the minimum of $E(\lambda)$, $dE(\lambda)/d\lambda|_{opt.} = \lambda^{opt.} - Z = 0$, or $\lambda^{opt.} = Z$, thus correctly predicting the true ground state of (4.62).

Not knowing the true asymptotic behavior of the ground state at large distances from the nucleus, one could alternatively try the spherical Gaussian function $\varphi(r, \vartheta, \varphi; \xi) = N(\xi) \exp(-\xi r^2)$ as an approximate representation of the ground state wave function in this *one*-electron atom, which gives:

$$E(\xi) = \frac{3}{2}\xi - \sqrt{\frac{8\xi}{\pi}}Z, \quad \xi^{opt.} = \frac{8}{9\pi}Z^2, \quad E(\xi^{opt.}) = \frac{4}{3\pi}Z^2 \cong -0.424Z^2. \tag{5.23}$$

Therefore, on the basis of the variational criterion one concludes that the exponential form of the variational wave function provides a better representation of the electronic wave function in the *one*-electron atom, since it generates lower energy compared with that resulting from the optimum Gaussian state.

The *linear* variant of the variational approach is known as the *Ritz method*. The trial state $|\varphi\rangle$ is then given as the linear combination of the adopted basis states $|\boldsymbol{\chi}\rangle = \{|\chi_p\rangle, p = 1, 2, \ldots, w\}$ (*row* vector) defined by the expansion coefficients $\boldsymbol{C} = \langle\boldsymbol{\chi}|\varphi\rangle = \{C_p\}$ (*column* vector):

$$|\varphi\rangle = \sum_{p=1}^{w} |\chi_p\rangle C_p \equiv |\boldsymbol{\chi}\rangle \boldsymbol{C}. \tag{5.24}$$

However, since $w$ basis functions define $w$ linearly independent combinations $|\boldsymbol{\varphi}\rangle = \{|\varphi^{(s)}\rangle, s = 1, 2, \ldots, w\}$ (*row* vector), we can generalize the above expression:

$$\left|\varphi^{(s)}\right\rangle = \sum_{p=1}^{w} \left|\chi_p\right\rangle C_{p,s} = |\boldsymbol{\chi}\rangle \boldsymbol{C}^{(s)} \equiv |\varphi_{s-1}\rangle, \quad s = 1, 2, \ldots, w, \tag{5.25}$$

or in the joint, matrix notation:

$$|\boldsymbol{\varphi}\rangle = |\boldsymbol{\chi}\rangle \mathbf{C}, \;\; \mathbf{C} = (\boldsymbol{C}^{(1)}|\boldsymbol{C}^{(2)}|\ldots|\boldsymbol{C}^{(s)}|\ldots|\boldsymbol{C}^{(w)}) \equiv (\boldsymbol{C}_0|\boldsymbol{C}_1|\ldots|\boldsymbol{C}_{s-1}|\ldots|\boldsymbol{C}_{w-1}). \tag{5.26}$$

In general, the basis vectors give rise to a nonunit metric tensor defined by the *overlap matrix* $\mathbf{S} = \langle\boldsymbol{\chi}|\boldsymbol{\chi}\rangle = \{S_{p,q} = \langle\chi_p|\chi_q\rangle\}$, while the Hamiltonian is representated by the *energy matrix* $\mathbf{H} = \langle\boldsymbol{\chi}|\hat{\mathrm{H}}|\boldsymbol{\chi}\rangle = \left\{H_{p,q} = \left\langle\chi_p\middle|\hat{\mathrm{H}}\middle|\chi_q\right\rangle\right\}$.

In what follows we shall assume that the optimum combinations are ordered in accordance with their increasing energies $\left\{\left\langle E^{(s)}\right\rangle = \left\langle\varphi^{(s)}\middle|\hat{\mathrm{H}}\middle|\varphi^{(s)}\right\rangle \equiv \langle E_{s-1}\rangle\right\}$:

$$[\langle E^{(1)}\rangle \equiv \langle E_0\rangle] \leq [\langle E^{(2)}\rangle \equiv \langle E_1\rangle] \leq \ldots \leq [\langle E^{(w)}\rangle \equiv \langle E_{w-1}\rangle]. \tag{5.27}$$

The optimum combination $|\varphi^{(1)}\rangle = |\varphi_0\rangle$ corresponding to the lowest energy $\langle E^{(1)}\rangle \equiv \langle E_0\rangle$ will then approximate the system ground state $|\psi_0\rangle$, while the remaining orthonormal combinations will approach the corresponding excited states.

In the last three equations, we have relabeled the *upper* indices of the eigenvectors, the associated columns in the $(w \times w)$ square matrix $\mathbf{C}$ grouping the combination coefficients, the linear variational parameters of the Ritz method, and the associated energy estimates into the corresponding subscripts conforming to the customary labeling of the molecular energy levels of (5.18), with $|\varphi_0\rangle$ and $\langle E_0\rangle$ denoting the *ground* state approximations and the remaining states corresponding to successive *excited* states:

$$|\varphi^{(s)}\rangle \equiv |\varphi_{s-1}\rangle, \quad \mathbf{C} = \{\, \boldsymbol{C}^{(s)} \equiv \boldsymbol{C}_{s-1}\}, \quad \{\langle E^{(s)}\rangle \equiv \langle E_{s-1}\rangle\}.$$

Let us first consider a single combination of (5.24). The expectation value of the system energy in state $|\varphi\rangle$ reads: $\langle E\rangle = \left\langle \tilde{\varphi}\middle|\hat{\mathrm{H}}\middle|\tilde{\varphi}\right\rangle = \langle\varphi|\hat{\mathrm{H}}|\varphi\rangle/\langle\varphi|\varphi\rangle$, where the denominator $\langle\varphi \mid \varphi\rangle = \sum_{p=1}^{w}\sum_{q=1}^{w} C_p^* S_{p,q} C_q = \boldsymbol{C}^\dagger \mathbf{S}\boldsymbol{C}$ is due to the normalization constant of $|\varphi\rangle$ in the *normalized* trial state $|\tilde{\varphi}\rangle = \langle\varphi\,|\,\varphi\rangle^{-1/2}|\varphi\rangle$. Hence,

$$\langle E\rangle\langle\varphi|\varphi\rangle = \langle E\rangle \sum_{p=1}^{w}\sum_{q=1}^{w} C_p^* S_{p,q} C_q = \langle\varphi|\hat{\mathrm{H}}|\varphi\rangle. \tag{5.28}$$

One further observes that the expansion coefficients $\boldsymbol{C}$ are in general complex numbers. Therefore, the unknowns in this linear variational problem consist of their real and imaginary parts, $\boldsymbol{C} = \mathrm{Re}(\boldsymbol{C}) + i\mathrm{Im}(\boldsymbol{C})$, where: $\mathrm{Re}(\boldsymbol{C}) = (\boldsymbol{C} + \boldsymbol{C}^*)/2$ and $\mathrm{Im}(\boldsymbol{C}) = (\boldsymbol{C} - \boldsymbol{C}^*)/2i$. Thus, one can alternatively designate the coefficients $\boldsymbol{C}$ and their complex conjugates $\boldsymbol{C}^*$ as independent variational parameters, since they uniquely identify both parts of $\boldsymbol{C}$. In fact, due to the Hermitian character of $\mathbf{H}$ and the symmetrical character of the metric $\mathbf{S}$, the secular equations for the optimum values of the linear variational parameters derived from the independent variations of $\boldsymbol{C}^*$ and $\boldsymbol{C}$, respectively, are identical.

The optimum solutions must minimize the system energy function $\langle E(\boldsymbol{C}^*, \boldsymbol{C})\rangle$ [see (5.21)]:

$$\left.\frac{\partial\langle E(\boldsymbol{C}^*,\boldsymbol{C})\rangle}{\partial C_p^*}\right|_{\min} = 0 \quad \text{and} \quad \left.\frac{\partial\langle E(\boldsymbol{C}^*,\boldsymbol{C})\rangle}{\partial C_p}\right|_{\min} = 0, \quad p = 1, 2, \ldots, w. \tag{5.29}$$

Differentiating (5.28) with respect to $C_p{}^*$ and taking into account the condition of the energy minimum of (5.29) then gives:

$$\frac{\partial\langle E(C^*,C)\rangle}{\partial C_p^*}\sum_{p=1}^{w}\sum_{q=1}^{w} C_p^* S_{p,q} C_q + \langle E\rangle \sum_{q=1}^{w} S_{p,q}C_q = \langle E\rangle \sum_{q=1}^{w} S_{p,q}C_q = \sum_{q=1}^{w} H_{p,q}C_q \text{ or}$$

$$\sum_{q=1}^{w}\left(H_{p,q} - \langle E\rangle S_{p,q}\right) C_q = 0, \quad p = 1, 2, \ldots, w. \tag{5.30}$$

This system of the *secular* (linear, homogeneous) *equations* has in fact only $w - 1$ independent unknowns. The additional, nonhomogeneous equation required to specify $\boldsymbol{C}$ uniquely is provided by the normalization condition for the combination in question:

$$\langle\varphi|\varphi\rangle = \sum_{p=1}^{w}\sum_{q=1}^{w} C_p^* S_{p,q} C_q = \boldsymbol{C}^\dagger \boldsymbol{S}\boldsymbol{C} = 1. \tag{5.31}$$

It then directly follows from the Cramer rules of Algebra that the necessary condition for the physically meaningful, nontrivial solutions $\boldsymbol{C} \neq \boldsymbol{0}$ of these secular equations is the vanishing determinant of coefficients before the unknowns in these homogeneous equations, called the *secular determinant*:

$$\begin{vmatrix} H_{1,1}-\langle E\rangle S_{1,1} & H_{1,2}-\langle E\rangle S_{1,2} & \ldots & H_{1,w}-\langle E\rangle S_{1,w} \\ H_{2,1}-\langle E\rangle S_{2,1} & H_{2,2}-\langle E\rangle S_{2,2} & \ldots & H_{2,w}-\langle E\rangle S_{2,w} \\ \ldots\ldots\ldots\ldots & \ldots\ldots\ldots\ldots & \ldots & \ldots\ldots\ldots\ldots \\ H_{w,1}-\langle E\rangle S_{w,1} & H_{w,2}-\langle E\rangle S_{w,2} & \ldots & H_{w,w}-\langle E\rangle S_{w,w} \end{vmatrix} \equiv \left|H_{p,q}-\langle E\rangle S_{p,q}\right| = 0. \tag{5.32}$$

Hence, by expanding the determinant one arrives at the equation of degree $w$ for the unknown $\langle E\rangle$. Its ordered solutions $\left\{\left\langle E^{(s)}\right\rangle \equiv \langle E_{s-1}\rangle\right\}$ (5.27) approximate the exact energy levels of the system ground and the first $(w-1)$ excited states (5.18).

To summarize, one first solves (5.32) for the approximate energy levels $\{\langle E^{(s)}\rangle\}$, the knowledge of which is required to uniquely specify the coefficients of the secular equations (5.30) supplemented by (5.31). Selecting $\langle E\rangle = \langle E^{(s)}\rangle$ in these equations gives the coefficients $\boldsymbol{C}^{(s)}$ determining $|\varphi^{(s)}\rangle$, etc.

Fortunately, this rather cumbersome procedure in terms of determinants can be recast in the form of the standard *matrix diagonalization* problem, which is easily handled in computer calculations. For this purpose, we arrange the energy estimates $\{\langle E^{(s)}\rangle\}$ as diagonal elements of the *eigenvalue matrix* $\mathbf{E} = \{E_{s,s'} = \langle E^{(s)}\rangle\delta_{s,s'}\}$ and rewrite the secular equations (5.30) for $s$th combination of (5.25):

$$\sum_{q=1}^{w}\left(H_{p,q}-\left\langle E^{(s)}\right\rangle S_{p,q}\right)C_{q,s} = \sum_{q=1}^{w}H_{p,q}C_{q,s} - \sum_{q=1}^{w}\sum_{s'=1}^{w}S_{p,q}C_{q,s'}E_{s',s} = 0 \text{ or}$$
$$\mathbf{HC}{=}\mathbf{SCE}. \tag{5.33}$$

This equation must be supplemented by the matrix equation combining the relevant orthonormality requirements for the optimum combinations, which are summarized by the requirement of the unit metric tensor defined by $|\boldsymbol{\varphi}\rangle = |\boldsymbol{\chi}\rangle\mathbf{C}$,

$$\langle\boldsymbol{\varphi}|\boldsymbol{\varphi}\rangle = \mathbf{C}^{\dagger}\langle\boldsymbol{\chi}\mid\boldsymbol{\chi}\rangle\mathbf{C} = \mathbf{C}^{\dagger}\mathbf{SC} = \mathbf{I}. \tag{5.34}$$

As already shown in Sect. 3.3.2, the nonorthogonal basis vectors $|\boldsymbol{\chi}\rangle$ can be transformed into the symmetrically orthogonalized analogs $|\tilde{\boldsymbol{\chi}}\rangle = |\boldsymbol{\chi}\rangle\,\mathbf{S}^{-1/2}$ of Löwdin, strongly resembling the original basis vectors $|\boldsymbol{\chi}\rangle$, which can be subsequently "rotated" in the unitary transformation $\mathbf{U}$ to the final optimum combinations we seek:

$$|\boldsymbol{\varphi}\rangle = |\boldsymbol{\chi}\rangle\mathbf{C} \equiv \left(|\boldsymbol{\chi}\rangle\mathbf{S}^{-1/2}\right)\mathbf{U} = |\tilde{\boldsymbol{\chi}}\rangle\mathbf{U}, \quad \mathbf{UU}^{\dagger}{=}\mathbf{U}^{\dagger}\mathbf{U} = \mathbf{I}. \tag{5.35}$$

This way of arriving at orthonormal combinations thus automatically satisfies (5.34). In this Löwdin orthogonalized representation the only unknown part of

$\mathbf{C} = \mathbf{S}^{-1/2}\mathbf{U}$ is its $\mathbf{U} = \mathbf{S}^{1/2}\mathbf{C}$ factor, where we have used the relation $\mathbf{S}^{-1/2}\mathbf{S}^{1/2} = \mathbf{S}^0 = \mathbf{I}$ (3.47). A straightforward transcription of (5.33) multiplied from the left by $\mathbf{S}^{-1/2}$ then gives:

$$(\mathbf{S}^{-1/2}\mathbf{H}\mathbf{S}^{-1/2})(\mathbf{S}^{1/2}\mathbf{C}) \equiv \tilde{\mathbf{H}}\mathbf{U} = (\mathbf{S}^{-1/2}\mathbf{S})\mathbf{C}\mathbf{E} = (\mathbf{S}^{1/2}\mathbf{C})\mathbf{E} = \mathbf{U}\mathbf{E}. \qquad (5.36)$$

Hence, by multiplying the preceding equation from the left by $\mathbf{U}^{\dagger}$ finally gives:

$$\mathbf{U}^{\dagger}\tilde{\mathbf{H}}\mathbf{U} = \mathbf{E}. \qquad (5.37)$$

The determination of the optimum coefficients $\mathbf{C}$, the linear variational parameters of the Ritz method, and of the associated average energy estimates $\mathbf{E}$ is thus simultaneously accomplished by the diagonalization in the unitary transformation $\mathbf{U}$ of the Hermitian matrix $\tilde{\mathbf{H}} = \mathbf{S}^{-1/2}\mathbf{H}\mathbf{S}^{-1/2}$. The latter constitutes the matrix representation of the Hamiltonian in the symmetrically orthogonalized basis set $|\tilde{\boldsymbol{\chi}}\rangle$,

$$\tilde{\mathbf{H}} = \langle\tilde{\boldsymbol{\chi}}|\hat{\mathrm{H}}|\tilde{\boldsymbol{\chi}}\rangle = \mathbf{S}^{-1/2}\langle\boldsymbol{\chi}|\hat{\mathrm{H}}|\boldsymbol{\chi}\rangle\,\mathbf{S}^{-1/2} = \mathbf{S}^{-1/2}\mathbf{H}\mathbf{S}^{-1/2}, \qquad (5.38)$$

where we have observed that $\langle\tilde{\boldsymbol{\chi}}| = \left(|\boldsymbol{\chi}\rangle\mathbf{S}^{-1/2}\right)^{\dagger} = \mathbf{S}^{-1/2}\langle\boldsymbol{\chi}|$, since $\mathbf{S}^{-1/2}$ is the real, symmetric matrix. This linear variational procedure thus amounts to the standard algorithmic problem in the matrix algebra.

We conclude this section by examining general criteria for an effective mixing of quantum states in the linear combination of (5.24). In textbooks on quantum chemistry such an analysis is carried out in the context of mixing AO into *Molecular Orbitals* (MO), when the prototype chemical bond is being formed, say between atoms A and B. To simplify these qualitative considerations, we reduce the problem to two AO states $|\boldsymbol{\chi}\rangle = (|A\rangle, |B\rangle)$, originating from atoms A and B, respectively, which are assumed to be normalized but nonorthogonal (overlapping):

$$\mathbf{S} = \begin{bmatrix} 1 & S \\ S & 1 \end{bmatrix}, \quad \mathbf{H} = \begin{bmatrix} \alpha_A & \beta \\ \beta & \alpha_B \end{bmatrix}, \qquad (5.39)$$

where for definiteness we put $S = \langle A \mid B\rangle > 0$ and $\alpha_A = \langle A|\hat{\mathrm{H}}|A\rangle \le \alpha_B = \langle B|\hat{\mathrm{H}}|B\rangle < 0$ (Fig. 5.1). The (negative) *Coulomb* integrals $\{\alpha_p\}$ reflect the energy levels associated with the individual AO and hence the corresponding negative ionization potentials (see the Koopmans theorem of Sect. 6.1.2 and the Janak theorem of Sect. 7.3.6), $\alpha_p \cong -I_p$, $p = A, B$, while the *resonance* integral $\beta = \langle A|\hat{\mathrm{H}}|B\rangle = \langle B|\hat{\mathrm{H}}|A\rangle$ measures their mutual interaction (coupling) in the bond formation process. In the semiempirical theories of the molecular electronic structure, it was adequately approximated as being proportional to the AO overlap integral $S$ and an average value (Av), arithmetic, geometric, or harmonic, of the corresponding diagonal elements of the Hamiltonian: $\beta \propto S\mathrm{Av}(\alpha_A, \alpha_B) \equiv S\langle\alpha\rangle \approx -S\mathrm{Av}(I_A, I_B) < 0$.

It then directly follows from the eigenvalue equation (5.32)

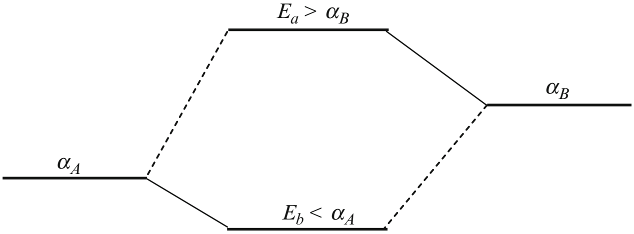

**Fig. 5.1** A qualitative diagram of the chemical interaction between two AO

$$\begin{vmatrix} \alpha_A - \langle E \rangle & \beta - \langle E \rangle S \\ \beta - \langle E \rangle S & \alpha_B - \langle E \rangle \end{vmatrix} = 0 \text{ or } (\alpha_A - \langle E \rangle)(\alpha_B - \langle E \rangle) = (\beta - \langle E \rangle S)^2 > 0, \quad (5.40)$$

that the two AO energy levels "repel" each other as a result of their quantum mechanical coupling into MO. More specifically, the preceding equation allows the two optimum MO energies, which are simultaneously either above or below both AO energies:

$$\begin{aligned} &[\alpha_A - \langle E \rangle > 0 \text{ and } \alpha_B - \langle E \rangle > 0] \Rightarrow \langle E \rangle = E_b < \alpha_A \quad \text{or} \\ &[\alpha_A - \langle E \rangle < 0 \text{ and } \alpha_B - \langle E \rangle < 0] \Rightarrow \langle E \rangle = E_a > \alpha_B. \end{aligned}$$

As a result the two MO energy estimates are obtained: the bonding level $\langle E^{(b)} \rangle = E_b < \alpha_A$ and the antibonding level $\langle E^{(a)} \rangle = E_a > \alpha_B$, which are also shown in the schematic diagram of Fig. 5.1.

For a general case of nonequal AO energy levels, the secular equation (5.40) gives the following expression for the *bonding energy*:

$$\alpha_A - E_b = (\beta - E_b S)^2 / (\alpha_B - E_b), \quad (5.41)$$

which satisfies the following inequalities:

$$0 < \alpha_A - E_b < (\beta - E_b S)^2 / (\alpha_B - \alpha_A), \quad (5.42)$$

since $\alpha_B - E_b > \alpha_B - \alpha_A > 0$ (see Fig. 5.1).

We thus conclude from the preceding equation that the larger the difference between the energy levels of the mixed states the smaller the bonding effect of their interaction. Indeed the strongest bonding results for $\alpha_B = \alpha_A = \alpha$ when

$$\begin{aligned} &E_b = (\alpha + \beta)/(1 + S) < \alpha, \; |\varphi_b\rangle = (|A\rangle + |B\rangle)/(2 + 2S)^{-1/2} \\ &E_a = (\alpha - \beta)/(1 - S) > \alpha, \; |\varphi_a\rangle = (|A\rangle - |B\rangle)/(2 - 2S)^{-1/2}. \end{aligned} \quad (5.43)$$

It follows from these equations that for the overlapping AO the antibonding effect $E_a - \alpha$ always exceeds its bonding companion $\alpha - E_b$. This explains why no net chemical bonding results in this simple orbital description from the interaction between the two fully occupied AO, e.g., in $He_2$ and $Be_2$. Moreover, due to an approximate proportionality relation $\beta \propto S\langle\alpha\rangle$, the r.h.s. of (5.42), which marks the upper limit of the bonding effect,

$$\alpha_A - E_b < (\beta - E_b S)^2/(\alpha_B - \alpha_A) \approx S^2(\langle\alpha\rangle - E_b)^2/(\alpha_B - \alpha_A),$$

identically vanishes, when there is no overlap between AO. For example, at finite separations between atoms, this can be due to the symmetry restrictions in the valence shell or the "narrowness" of electron distributions in the *inner* shells of both atoms. Together with the nucleus these chemically inactive electrons of the *inner* shells define the atomic "*core*," which remains largely unaffected by the chemical bonds formed in the valence shell. At very large internuclear distances, in the *separated atoms* limit, the AO overlap also vanishes, so that no chemical interaction is predicted. We thus conclude that a large AO overlap is conducive for a strong chemical bonding originating from the orbital interaction in a molecule.

Let us next consider the *squared* secular equation for the unknown coefficients of the combination:

$$(\alpha_A - \langle E\rangle)C_A + (\beta - \langle E\rangle S C_B = 0 \Rightarrow (\alpha_A - \langle E\rangle)^2 C_A^2 = (\beta - \langle E\rangle S)^2 C_B^2. \quad (5.44)$$

Using the expression for $(\beta - \langle E\rangle S)^2$ from (5.40) gives the following ratio of the squares of coefficients, reflecting a relative participation (conditional probability) of AO in the combination,

$$C_A^2/C_B^2 = (\alpha_B - \langle E\rangle)/(\alpha_A - \langle E\rangle) = |\alpha_B - \langle E\rangle|/|\alpha_A - \langle E\rangle|. \quad (5.45)$$

Indeed, for $\alpha_B = \alpha_A$ both AO participate equally in MO and $C_A = \pm C_B$, in accordance with (5.43). In a general case of Fig. 5.1, one predicts for $\langle E\rangle = E_b$: $\alpha_B - E_b > \alpha_A - E_b$, so that orbital $|A\rangle$ dominates the bonding combination $|\varphi_b\rangle$: $C_A^2 > C_B^2$. One similarly predicts a stronger similarity of $|\varphi_a\rangle$ to $|B\rangle$ for $\langle E\rangle = E_a$, since then $|\alpha_B - E_a| < |\alpha_A - E_a|$. Therefore, with increasing gap $\alpha_B - \alpha_A$ of the AO energies the bonding combination $|\varphi_b\rangle$ more strongly resembles $|A\rangle$ and the antibonding combination $|\varphi_a\rangle$ becomes more like $|B\rangle$.

## 5.2 Adiabatic Separation of Electronic and Nuclear Motions

To a good approximation, when describing the state of (light) electrons in a molecule, one can treat the system (heavy) nuclei as being at rest, in view of the drastic difference in masses of these two *micro*-objects. Indeed, the motions of the former are very fast compared with the slow movements of the latter. This physical intuition suggests that for the nuclear dynamics the instantaneous positions of electrons are unimportant, with only the average effect of their fast motions influencing the forces acting on nuclei in the molecular system under consideration.

The formal basis of this separation of the electronic and nuclear degrees of freedom in the molecular (stationary) quantum mechanics is the familiar *adiabatic* approximation of Born and Oppenheimer (1927).

Consider the molecular wave function $\Psi(\mathbf{q}, \mathbf{Q})$ of $N$ electrons at positions $\mathbf{r} = \{\boldsymbol{r}_i\}$ exhibiting the spin orientations $\boldsymbol{\sigma} = \{\sigma_i\}$, or in the combined notation $\mathbf{q} \equiv (\mathbf{r}, \boldsymbol{\sigma}) = \{\boldsymbol{r}_i, \sigma_i\} \equiv \{\boldsymbol{q}_i\}$, and $m$ nuclei of masses $\{M_\alpha\}$ and charges $\{Z_\alpha\}$ in positions $\mathbf{R} = \{\boldsymbol{R}_\alpha\}$ with spins $\boldsymbol{\Sigma} = \{\Sigma_\alpha\}$, which determine the corresponding position-spin variables $\mathbf{Q} \equiv (\mathbf{R}, \boldsymbol{\Sigma}) = \{\boldsymbol{R}_\alpha, \Sigma_\alpha\} \equiv \{\boldsymbol{Q}_\alpha\}$. It generates the associated probability distribution of the *joint*, electronic-nuclear events: $P(\mathbf{q}, \mathbf{Q}) = |\Psi(\mathbf{q}, \mathbf{Q})|^2$, which satisfies the relevant overall and partial normalizations:

$$\iint P(\mathbf{q}, \mathbf{Q})\, d\mathbf{q}\, d\mathbf{Q} = \int \pi(\mathbf{Q})\, d\mathbf{Q} = \int \rho(\mathbf{q})\, d\mathbf{q} = 1, \tag{5.46}$$

where $\pi(\mathbf{Q})$ and $\rho(\mathbf{q})$ denote the partially integrated nuclear and electronic probability distributions, respectively.

The essence of the adiabatic approximation lies in extracting from this joint distribution the probability density of the heavy (slow) nuclei as the reference (parameter) distribution:

$$P(\mathbf{q}, \mathbf{Q}) = \pi(\mathbf{Q}) \frac{P(\mathbf{q}, \mathbf{Q})}{\pi(\mathbf{Q})} \equiv \pi(\mathbf{Q}) p(\mathbf{q}|\mathbf{Q}), \quad \int p(\mathbf{q}|\mathbf{Q})\, d\mathbf{q} = 1. \tag{5.47}$$

In the *conditional* probability density of electrons, $p(\mathbf{q}|\mathbf{Q})$, the nuclear variables thus play the role of parameters, as indeed reflected by the above normalization condition. This further implies the associated factorization of the system wave function in terms of the nuclear, $\chi(\mathbf{Q})$, and electronic, $\varphi(\mathbf{q}|\mathbf{Q})$, functions,

$$\Psi(\mathbf{q}, \mathbf{Q}) \cong \chi(\mathbf{Q}) \varphi(\mathbf{q}|\mathbf{Q}). \tag{5.48}$$

They accordingly represent the nuclear and (*conditional*) electronic amplitudes of the associated probability distributions:

$$\pi(\mathbf{Q}) = |\chi(\mathbf{Q})|^2 \quad \text{and} \quad p(\mathbf{q}|\mathbf{Q}) = |\varphi(\mathbf{q}|\mathbf{Q})|^2. \tag{5.49}$$

Therefore, in the *Born–Oppenheimer* (BO) approximation, the nuclear wave function is not *explicitly* dependent upon the electronic positions, while the electronic state $\varphi(\mathbf{q}|\mathbf{Q})$ is defined for the fixed geometry of the molecular system, defined by specified, parametric positions of the nuclei. The relevant orthonormality relations satisfied by different adiabatic states $\{\chi_i\}$ and $\{\varphi_r\}$ thus read:

$$\int \chi_i^*(\mathbf{Q}) \chi_j(\mathbf{Q})\, d\mathbf{Q} \equiv \langle \chi_i \mid \chi_j \rangle_{\mathbf{Q}} = \delta_{i,j} \quad \text{and}$$

$$\int \varphi_r^*(\mathbf{q}|\mathbf{Q}) \varphi_t(\mathbf{q}|\mathbf{Q})\, d\mathbf{q} \equiv \langle \varphi_r(\mathbf{Q}) \mid \varphi_t(\mathbf{Q}) \rangle_{\mathbf{q}} = \delta_{r,t}. \tag{5.50}$$

The molecular (Coulombic) Hamiltonian in the position representation,

$$\begin{aligned}\hat{\mathrm{H}}(\mathbf{q},\mathbf{Q}) &= \hat{\mathrm{T}}_n(\mathbf{Q}) + [\hat{\mathrm{T}}_e(\mathbf{q}) + V_{ne}(\mathbf{q},\mathbf{Q}) + V_{ee}(\mathbf{q}) + V_{nn}(\mathbf{Q})] \\ &\equiv \hat{\mathrm{T}}_n(\mathbf{Q}) + \hat{\mathrm{H}}_e(\mathbf{q},\mathbf{Q}) \equiv \hat{\mathrm{T}}_n(\mathbf{Q}) + \hat{\mathrm{H}}^e(\mathbf{q},\mathbf{Q}) + V_{nn}(\mathbf{Q}),\end{aligned} \tag{5.51}$$

groups operators of the following (dominating) contributions to the molecular energy (a.u.):

kinetic energy of nuclei: $\hat{\mathrm{T}}_n(\mathbf{Q}) = -\sum_{\alpha=1}^{m} \frac{1}{2M_\alpha}\Delta_\alpha, \quad \Delta_\alpha = \nabla^2_{\boldsymbol{R}_\alpha};$

kinetic energy of electrons: $\hat{\mathrm{T}}_e(\mathbf{q}) = -\frac{1}{2}\sum_{j=1}^{N}\Delta_j, \quad \Delta_j = \nabla^2_{\boldsymbol{r}_j};$

nuclear-electron attraction energy: $V_{ne}(\mathbf{q},\mathbf{Q}) = -\sum_{j=1}^{N}\left(\sum_{\alpha=1}^{m}\frac{Z_\alpha}{|\boldsymbol{R}_\alpha - \boldsymbol{r}_j|}\right) \equiv \sum_{j=1}^{N} v(\boldsymbol{r}_j,\mathbf{Q});$

electron repulsion energy: $V_{ee}(\mathbf{q}) = \sum_{i=1}^{N-1}\sum_{j=i+1}^{N}\frac{1}{|\boldsymbol{r}_i - \boldsymbol{r}_j|} \equiv \sum_{i=1}^{N-1}\sum_{j=i+1}^{N} g(i,j), \; g(i,j) = 1/r_{i,j};$

nuclear repulsion energy: $V_{nn}(\mathbf{Q}) = \sum_{\alpha=1}^{m-1}\sum_{\beta=\alpha+1}^{m}\frac{Z_\alpha Z_\beta}{|\boldsymbol{R}_\alpha - \boldsymbol{R}_\beta|} = \sum_{\alpha=1}^{m-1}\sum_{\beta=\alpha+1}^{m}\frac{Z_\alpha Z_\beta}{R_{\alpha,\beta}}.$

Above, $v(\boldsymbol{r}, \mathbf{Q})$ denotes the external potential for an electron in position $\boldsymbol{r}$ due to the nuclei in their "frozen" positions $\{\boldsymbol{R}_\alpha\}$.

The *electronic Hamiltonian* $\hat{\mathrm{H}}_e(\mathbf{q},\mathbf{Q})$ defined in (5.51) groups all these terms except the nuclear kinetic energy operator. Since the nuclear repulsion energy does not affect the electronic states, representing just the irrelevant additive constant in $\hat{\mathrm{H}}_e(\mathbf{q},\mathbf{Q})$, it is sometimes neglected in the Schrödinger equation for electrons, defined by the eigenvalue problem of the redefined electronic Hamiltonian $\hat{\mathrm{H}}^e(\mathbf{q},\mathbf{Q}) \equiv \hat{\mathrm{H}}_e(\mathbf{q},\mathbf{Q}) - V_{nn}(\mathbf{Q})$.

Therefore, in the BO approximation of (5.48), the molecular states $\Psi_{r,k}(\mathbf{q},\mathbf{Q}) = \varphi_r(\mathbf{q}|\mathbf{Q})\chi_k(\mathbf{Q})$ must satisfy the stationary Schrödinger equation:

$$[\hat{\mathrm{T}}_n(\mathbf{Q}) + \hat{\mathrm{H}}_e(\mathbf{q},\mathbf{Q})]\varphi_r(\mathbf{q}|\mathbf{Q})\chi_k(\mathbf{Q}) = E_{r,k}\varphi_r(\mathbf{q}|\mathbf{Q})\chi_k(\mathbf{Q}), \tag{5.52}$$

where $E_{r,k}$ stands for the molecular energy in the adiabatic state combining $r$th electronic and $k$th nuclear states. Since both factors depend, at least parametrically, on the nuclear positions the action of the nuclear kinetic energy operator on adiabatic wave function gives:

$$\begin{aligned}\hat{\mathrm{T}}_n(\varphi_r\chi_k) &= -\sum_{\alpha=1}^{m}\frac{1}{2M_\alpha}[\chi_k(\Delta_\alpha\varphi_r) + 2(\nabla_\alpha\varphi_r)\cdot(\nabla_\alpha\chi_k) + \varphi_r(\Delta_\alpha\chi_k)] \\ &= (\hat{\mathrm{T}}_n\varphi_r)\chi_k - \sum_{\alpha=1}^{m}\frac{1}{M_\alpha}(\nabla_\alpha\varphi_r)\cdot(\nabla_\alpha\chi_k) + \varphi_r(\hat{\mathrm{T}}_n\chi_k).\end{aligned} \tag{5.53}$$

The adiabatic approximation assumes that the kinetic energy operator $\hat{\mathrm{T}}_n(\mathbf{Q})$ constitutes a minor perturbation compared with the electronic Hamiltonian $\hat{\mathrm{H}}_e(\mathbf{q},\mathbf{Q})$. One can therefore envisage the perturbation approach constructed on the basis of the unperturbed Hamiltonian $\hat{\mathrm{H}}^0 = \hat{\mathrm{H}}_e(\mathbf{q},\mathbf{Q})$, in which there are no gradient operations with respect to nuclear positions. Therefore, in *zeroth* order approximation the nuclear positions are treated as parameters and one neglects the second term in r.h.s. of (5.53) as negligible, eventually to be taken into account in *higher* orders of PT. In other words, one assumes that nuclear gradient of the electronic wave function is generally small compared with the associated action of the electronic Hamiltonian. One could also neglect the first (small) Laplacian term, as in the original BO approach, but this contribution can be easily accounted for without any serious complication of the emerging formalism.

Therefore, neglecting only the second term in (5.53), which involves the scalar product of the nuclear gradients of both factors in the adiabatic form of the molecular wave function, multiplying from the left (5.52) by $\varphi_r{}^*$, and "integrating" the result over the electronic position-spin variables $\mathbf{q}$, denoted by $\langle\rangle_\mathbf{q}$, then give the following effective Schrödinger equation for the nuclear function $\chi_k(\mathbf{Q})$:

$$\begin{aligned}
&[\hat{\mathrm{T}}_n(\mathbf{Q}) + \langle\varphi_r(\mathbf{q}|\mathbf{Q})|\hat{\mathrm{H}}_e(\mathbf{q},\mathbf{Q})|\varphi_r(\mathbf{q}|\mathbf{Q})\rangle_\mathbf{q} + \langle\varphi_r(\mathbf{q}|\mathbf{Q})|\,\hat{\mathrm{T}}_n(\mathbf{Q})\,|\varphi_r(\mathbf{q}|\mathbf{Q})\rangle_\mathbf{q}]\chi_k(\mathbf{Q})\\
&\qquad\equiv \{\hat{\mathrm{T}}_n(\mathbf{Q}) + [E_r^e(\mathbf{Q}) + T_r^n(\mathbf{Q})]\}\chi_k(\mathbf{Q}) \equiv [\hat{\mathrm{T}}_n(\mathbf{Q}) + U_r^{adiab.}(\mathbf{Q})]\chi_k(\mathbf{Q})\\
&\qquad= E_{r,k}\chi_k(\mathbf{Q}).
\end{aligned}\tag{5.54}$$

This equation contains the effective adiabatic potential in the electronic state $\varphi_r$, $U_r^{adiab.}(\mathbf{Q})$, the dominant component of which is the average electronic energy, the associated expectation value of the electronic Hamiltonian:

$$E_r^e(\mathbf{Q}) = \langle\varphi_r(\mathbf{q}|\mathbf{Q})|\hat{\mathrm{H}}_e(\mathbf{q},\mathbf{Q})|\varphi_r(\mathbf{q}|\mathbf{Q})\rangle_\mathbf{q}, \tag{5.55}$$

called the *Potential Energy Surface* (PES). It parametrically depends on nuclear positions (molecular geometry) and carries the influence of the average electronic distribution on the system nuclei. It follows from (5.54) that adiabatic potential also includes a (small) diagonal correction due to $\hat{\mathrm{T}}_n(\mathbf{Q})$ in state $\varphi_r$,

$$T_r^n(\mathbf{Q}) = \langle\varphi_r(\mathbf{q}|\mathbf{Q})|\,\hat{\mathrm{T}}_n(\mathbf{Q})\,|\varphi_r(\mathbf{q}|\mathbf{Q})\rangle_\mathbf{q}, \tag{5.56}$$

which has been neglected in the original, *crude*-adiabatic BO approximation.

It thus follows from the nuclear Schrödinger equation (5.54) that it requires the knowledge of the whole electronic PES $E_r^e(\mathbf{Q})$ of electrons in the specified adiabatic state $\varphi_r(\mathbf{q}|\mathbf{Q})$, the eigenfunction of the electronic Schrödinger equation:

$$\hat{\mathrm{H}}_e(\mathbf{q},\mathbf{Q})\varphi_r(\mathbf{q}|\mathbf{Q}) = E_r^e(\mathbf{Q})\varphi_r(\mathbf{q}|\mathbf{Q}). \tag{5.57}$$

Clearly, the parametric dependence of the electronic energy on nuclear coordinates can be extracted only from a very large number of solutions of the preceding equation, for a sufficient multitude of the fixed nuclear configurations $\{\mathbf{Q}^{(i)}\}$, by an analytical interpolation of the known energies $\{E_r^e(\mathbf{Q}^{(i)})\}$, points on the resulting PES.

To summarize, solving the molecular Schrödinger equation in the adiabatic (BO) approximation first involves solving the *fixed*-nuclei, electronic eigenvalue problem for a large number of molecular geometries, in order to extract the effective potential of forces acting on the system nuclei, averaged over the instantaneous positions of the *fast*-moving electrons. In the second, nuclear stage one uses this effective adiabatic potential to solve the nuclear Schrödinger equation (5.54), which generates the amplitude functions of the nuclear probability distributions and the molecular energy levels containing the kinetic energy of the *slowly* moving nuclei.

As we have already mentioned earlier in this section, the nonadiabatic effects can be accounted for in the *higher* order of the perturbation theory in which the kinetic energy of nuclei represents the perturbation to the unperturbed, electronic Hamiltonian. Therefore, the electronic states $\{\varphi_r(\mathbf{q}|\mathbf{Q})\}$ span the complete basis of the *zeroth* order solutions [see (5.57)], in terms of which the nonadiabatic states can be expanded. Consider the dominating, *first*-order corrections to the adiabatic electronic state $\varphi_r(\mathbf{q}|\mathbf{Q})$ (see Sect. 5.1):

$$\Delta\varphi_r^{(1)}(\mathbf{q}\,|\mathbf{Q}) = \sum_{t\neq r} c_{t,r}^{(1)}(\mathbf{Q})\;\varphi_t(\mathbf{q}\,|\mathbf{Q}). \tag{5.58}$$

It follows from (5.12) that this expansion coefficient is given by the following ratio:

$$c_{t,r}^{(1)}(\mathbf{Q}) = \frac{\langle\varphi_t(\mathbf{q}|\mathbf{Q})|\hat{\mathrm{T}}_n(\mathbf{Q})|\varphi_r(\mathbf{q}|\mathbf{Q})\rangle_{\mathbf{q}}}{E_r^e(\mathbf{Q}) - E_t^e(\mathbf{Q})}. \tag{5.59}$$

The adiabatic approximation is thus adequate only, when the numerator in this expression is small compared with the denominator. Indeed, the degeneracy or *near*-degeneracy of electronic states (small value of the denominator) would generate a large nonadiabatic correction thus contradicting the basic assumption of the adiabatic approximation. The same would be true for a large value of the numerator, signifying a strong nuclear-motion coupling between electronic states.

Therefore, the adiabatic approximation breaks down when for some molecular geometries several electronic states exhibit very close values of the electronic energy. This is the case in the familiar Jahn–Teller effect (removal of the electronic degeneracy by spontaneous distortion of the molecule) and the related Renner effect, due to the vibronic coupling between electronic and nuclear motions, which have profound structural and spectroscopic implications. Let us recall that the Jahn–Teller theorem states that in any nonlinear system there exists some vibrational mode that removes the degeneracy of an electronically (orbitally) degenerate state by lowering the system symmetry. The vibronic coupling between the degenerate electronic states of linear molecules and the deformation (bending) vibrations of the nuclei are responsible for splitting these energy levels in the Renner effect. A proper quantum mechanical description of these processes calls for an explicit dependence of electronic states on nuclear coordinates in the nonadiabatic molecular wave functions, which are customarily represented as linear combinations of several adiabatic states [see (5.58)]. The nuclear motions, of the paramount importance for molecular dynamics (e.g., Murrell et al. 1984; Murrell and Bosanac 1989) and spectroscopy (e.g., Longuet-Higgins 1961), are not covered by this book.

## 5.3 Orbital Approximation of Electronic Wave Functions

The quantum theory of electronic structure of molecules is based upon the *one*-electron approach to electronic functions of *many*-electron systems, known as the *orbital approximation*. It has greatly influenced the existing terminology of quantum chemistry and the chemical concepts used in interpretations of diverse chemical processes. It ascribes to each electron in the system the *one*-electron function called the *spin orbital* (SO, see Sect. 4.5).

Let us recall that the internal stationary state of the hydrogen-like atom discussed in Chap. 4 has been described by a single SO, $\psi(\boldsymbol{q}) = \varphi(\boldsymbol{r})\zeta(\sigma)$, given by the product of the spatial function, the *orbital* $\varphi(\boldsymbol{r})$, and one of the two admissible *spin functions* $\zeta(\sigma) = \{\alpha(\sigma) = \langle\sigma|\alpha\rangle,\ \beta(\sigma) = \langle\sigma|\beta\rangle\}$ of an electron [see (3.76)]. When the same orbital is used to generate two SO, thus describing a pair of electrons with the opposite spin orientations, as in (3.76), one adopts the so-called *spin-restricted* version of the orbital approximation. Accordingly, in the *spin-unrestricted* description of such two spin-paired electrons, one uses different orbitals for different spins:

$$\{\psi^{+}(\boldsymbol{q}) \equiv \varphi_{\alpha}(\boldsymbol{r})\alpha(\sigma), \quad \psi^{-}(\boldsymbol{q}) \equiv \varphi_{\beta}(\boldsymbol{r})\beta(\sigma)\}. \tag{5.60}$$

Let us now examine the Slater (1929, 1931, 1960) method of constructing in the orbital approximation the $N$-electron wave functions $\Psi(\mathbf{q}|\mathbf{Q}) \equiv \Psi(N)$, which automatically satisfy the basic requirement of the Fermi–Dirac statistics, the Pauli postulate of their antisymmetry with respect to an exchange of any two indistinguishable fermions. Should the electronic states be exactly independent, the $N$-electron wave function would then be exactly given by the product of $N$ orthonormal SO attributed to each particle,

$$\Psi(N) = \prod_{i=1}^{N} \psi_i(\boldsymbol{q}_i) \equiv \prod_{i=1}^{N} \psi_i(i) = \psi_1(1)\psi_2(2)\ldots\psi_i(i)\ldots\psi_j(j)\ldots\psi_N(N). \tag{5.61}$$

Indeed, the $N$-electron probability distribution would then be given by the product of distributions of independent *one*-electron events:

$$p(N) = |\Psi(N)|^2 = \prod_{i=1}^{N} |\psi_i(i)|^2 = \prod_{i=1}^{N} p_i(i). \tag{5.62}$$

Obviously, due to a finite electric charge, electrons repel each other, so that this independent particle approximation can at best be considered only as a first step in a more adequate treatment, which recognizes the dependence (correlation) between their instantaneous positions. Besides this *Coulomb* correlation the electron probabilities must also reflect the constraints imposed by the antisymmetry principle of Pauli, thus additionally exhibiting the *Fermi* (exchange) correlation, which severely conditions the simultaneous probability distributions of the spin-like electrons in the physical space.

The product trial function of (5.61), which has been used as the variational wave function in the Hartree (1928) method, clearly violates this antisymmetry requirement, since each electron is distinguished by the identity of the SO to which it has been individually ascribed. Thus, the permutation $\hat{\mathrm{P}}(i,j)$ of electrons $i$ and $j$, of exchanging the wave function arguments $\boldsymbol{q}_i$ and $\boldsymbol{q}_j$, instead of changing only the sign of $\Psi(N)$ transforms it into an entirely different function, in which electrons are attributed to different SO:

$$\hat{\mathrm{P}}(i,j)\Psi(N) = \psi_1(1)\psi_2(2)\ldots\psi_j(i)\ldots\psi_i(j)\ldots\psi_N(N) \neq -\Psi(N). \tag{5.63}$$

This shortcoming can be remedied by the appropriate antisymmetrization operation $\hat{\mathrm{A}}$ performed on the product function of (5.61). It is effected by combining all product functions obtained by permuting all $N$ electrons between all $N$ occupied SO. Each permutation $P$ is now identified by the *list of electrons*

$$\boldsymbol{l}(P) = \{l_i(P)\} = [l_1(P), l_2(P), \ldots, l_N(P)], \quad l_i(P) \in (1, 2, \ldots, N),$$

attributed to orbitals $\psi_i$ in the ordered list $\{\psi_i\} = (\psi_1, \psi_2, \ldots, \psi_N)$. Thus, the permutation $\boldsymbol{l}(P) = (4, 2, \ldots, 1)$ symbolizes the product function $\psi_1(4)\ \psi_2(2) \ldots \psi_N(1)$, etc. One could alternatively identify the current permutation $P$ by the *list of orbitals* $\{\psi_j(P)\}$, identified by their labels

$$\boldsymbol{k}(P) = \{k_j(P)\} = [k_1(P), k_2(P), \ldots, k_N(P)], \quad k_j(P) \in (1, 2, \ldots, N),$$

which are attributed to the ordered list of electrons $\{j\} = (1, 2, \ldots, N)$. Thus, the permutation $\boldsymbol{k}(P) = (4, 2, \ldots, 1)$ stands for the product function $\psi_4(1)\ \psi_2(2) \ldots \psi_1(N)$.

For the chosen type of permuting the products of SO, one then introduces the appropriate sign convention for each of $N!$ permutations in the antisymmetrized combination. In order to enforce the change of sign of the wave function, when the current permutation is obtained by an *odd* number of elementary replacements of pairs of electrons in the ordered permutation $P_0 = (1, 2, \ldots, N)$ of (5.61), one introduces the permutation parity $p(P)$, which counts the number of such pair exchanges required to bring the current permutation $P$ to the initial permutation $P_0$, with $p(P_0) = 0$, and puts the sign $(-1)^{p(P)}$ before the associated product function.

The antisymmetric combination of such $N!$ product functions corresponding to either all permutations of electrons among the ordered list of SO, or all permutations of SO among the ordered list of electrons, thus determines the Slater determinant:

$$\begin{aligned}
\Psi_A(N) = \hat{\mathrm{A}}\Psi(N) &= \frac{1}{\sqrt{N!}}\sum_P (-1)^{p(P)}\psi_1(l_1(P))\,\psi_2(l_2(P))\ldots\psi_N(l_N(P)) \\
&= \frac{1}{\sqrt{N!}}\sum_P (-1)^{p(P)}\psi_{k_1(P)}(1)\,\psi_{k_2(P)}(2)\ldots\psi_{k_N(P)}(N) \\
&= \frac{1}{\sqrt{N!}}\begin{vmatrix} \psi_1(1) & \psi_1(2) & \ldots\ldots. & \psi_1(N) \\ \psi_2(1) & \psi_2(2) & \ldots\ldots. & \psi_2(N) \\ \ldots\ldots. & \ldots\ldots. & \ldots\ldots. & \ldots\ldots. \\ \psi_N(1) & \psi_N(2) & \ldots\ldots. & \psi_N(N) \end{vmatrix} \\
&\equiv |\psi_1\psi_2\ldots\psi_N| \equiv \det(\psi_1\psi_2\ldots\psi_N).
\end{aligned} \tag{5.64}$$

Here, the constant before the determinant assures the normalization for the orthonormal set of SO: $\int \psi_i^*(\boldsymbol{q})\,\psi_j(\boldsymbol{q})\,d\boldsymbol{q} = \langle i\mid j\rangle = \delta_{i,j}$. Since exchanging two electrons amounts to the permutation of two columns in this determinantal wave function, the correct result of a change of sign of $\Psi_A(N)$ is obtained, $\hat{\mathrm{P}}(i,j)\Psi_A(N) = -\Psi_A(N)$, as indeed required by the Pauli antisymmetry postulate for fermions. One also observes that this form of wave function automatically satisfies the Pauli exclusion principle that two electrons cannot be described by identical SO. More specifically, should this be the case, the two *rows* in the Slater determinant would then be identical, thus automatically implying $\Psi_A(N) = 0$.

A more subtle implication also follows, when two spin-like electrons near-coalesce in the same position, when $\boldsymbol{q}_i \cong \boldsymbol{q}_j$. This limiting proximity of two electrons exhibiting the same spin orientation gives rise to two identical *columns* in $\Psi_A(N)$, thus again predicting $\Psi_A(N) \cong 0$. In other words, the probability of such an event becomes very small indeed. This implies that spin-like electrons are statistically correlated, avoiding nearby positions in space. This effect is called the *Fermi* or *exchange* correlation between electrons. It should be emphasized that no such restrictions on the instantaneous positions of electrons intervene for the electrons with opposite spins, since then spatial coalescence of two electrons does not imply equality of their position-spin variables: $\boldsymbol{q}_i = (\boldsymbol{r}, \uparrow) \neq \boldsymbol{q}_j = (\boldsymbol{r}, \downarrow)$. Therefore, electrons with different orientations of their spins exhibit only the Coulomb correlation, resulting from their electric charge, while the movements of the spin-like electrons are influenced by both the Fermi and Coulomb correlations.

Thus, in the orbital approximation of the Slater determinant (5.64), the spin-like electrons are not independent, being already Fermi correlated by the exchange symmetry of the electronic wave function. It should be emphasized, however, that this variational wave function ignores completely the Coulomb correlation of all electrons. Therefore, the latter effect should be relatively more important in interactions between electrons exhibiting different spin states, since the spin-like electrons have already been Fermi correlated. It could be also expected that accounting for this missing effect within the spin-restricted approach should most influence the simultaneous probabilities of two electrons occupying the same orbital, the movements of which are confined to the same part of space, probed by the square of their common spatial function (orbital).

It should be observed that the correct symmetry of the analogous orbital wave function for the set of $N$ identical *bosons* would call for the related symmetrization operation $\hat{\mathrm{S}}$ performed on the product wave function:

$$\begin{aligned}\Psi_S(N) = \hat{\mathrm{S}}\Psi(N) &= \frac{1}{\sqrt{N!}}\sum_P \psi_1[l_1(P)]\psi_2[l_2(P)]\ldots\psi_N[l_N(P)] \\ &= \frac{1}{\sqrt{N!}}\sum_P \psi_{k_1(P)}(1)\psi_{k_2(P)}(2)\ldots\psi_{k_N(P)}(N). \qquad (5.65)\end{aligned}$$

Indeed, this symmetrical combination of the permuted product functions satisfies the symmetry postulate for bosons, $\hat{\mathrm{P}}(i,j)\Psi_S(N) = \Psi_S(N)$, since such an operation only exchanges two product functions in the sum of all $N!$ terms of $\Psi_S(N)$.

It should be realized that the set $\boldsymbol{\psi}$ of singly occupied SO defining the Slater determinant is not unique. Indeed, any unitary transformation $\boldsymbol{\psi}' = \boldsymbol{\psi}\mathbf{T}$, $\mathbf{T}\mathbf{T}^{\dagger} = \mathbf{I}$, which preserves the orbital orthonormality, replaces the rows $\{\psi_i\}$ of the original determinant (5.64), $\Psi_A = \det(\boldsymbol{\psi})$, with their combinations $\{\psi_k{}' = \sum_j \psi_j T_{j,k}\}$ in the transformed determinant $\Psi_A{}' = \det(\boldsymbol{\psi}')$. It thus follows from the elementary properties of determinants that these two functions are identical: $\Psi_A{}' = \Psi_A$. The two sets of SO which define them are called the *equivalent orbitals*.

One thus encounters various types of *molecular orbitals* (MO) in the theory of electronic structure, selected for their numerical or interpretative convenience. For example, in the two most popular computational methods, formulated within the *Hartree–Fock* (HF) (Fock 1930) and *Kohn–Sham* (KS) (Kohn and Sham 1965) theories, the two canonical sets of orbitals are introduced, which are delocalized throughout the whole molecule and reflect the system spatial symmetry. They provide a useful orbital picture of the spectroscopic and electron ionization phenomena, satisfying important theorems linking their energies and decay behavior with the molecular ionization potentials. The *Natural Orbitals* (NO) of the *Configuration Interaction* (CI) theory similarly generate a compact representation of the Coulomb correlation effects. Finally, the *Localized Orbitals* (LO), describing the diatomic chemical bonds and lone electronic pairs, are useful in providing the orbital interpretations of the near-additivity of several molecular properties and in explaining the remarkable invariance of the given type of $\sigma$ bonds in different molecular environments. It should be also noticed that the very criteria for the orbital localization are not unique either, so that a variety of alternative sets of the physically equivalent LO have been reported in scientific literature (e.g., Boys and Foster 1960; Edmiston and Ruedenberg 1963), which generate the same determinantal wave function of the molecular system as a whole.

The orbital approximation provides a firm basis for the classification and understanding of electronic states and configurations in atomic and molecular systems. Since to a good approximation the length of the resultant spin $\boldsymbol{S}$ of all electrons and its $z$-component $S_z$ are sharply defined simultaneously with the system electronic energy $E_e$, the electronic wave functions are required to be eigenfunctions of the associated quantum mechanical operators $\hat{\mathrm{S}}^2$ and $\hat{\mathrm{S}}_z$, which commute with the electronic Hamiltonian. In the spin-unrestricted form the Slater determinant does not generally satisfy this requirement, while the spin-restricted functions

$$\Psi_A(N = 2p) = \left|\varphi_1^+\varphi_1^-\varphi_2^+\varphi_2^- \ldots \varphi_p^+\varphi_p^-\right|, \tag{5.66}$$

$$\Psi_A(N = 2p + q) = \left|\varphi_1^+\varphi_1^-\varphi_2^+\varphi_2^-{\ldots}\varphi_p^+\varphi_p^-\varphi_{p+1}^+\varphi_{p+2}^+{\ldots}\varphi_{p+q}^+\right|, \tag{5.67}$$

are eigenfunctions of these two resultant-spin operators, corresponding to the quantum numbers $S$ and $M_S$ determining the associated eigenvalues: $|\boldsymbol{S}|^2 = S(S+1)\hbar^2$ and $S_z = M_S\hbar$, $M_S = -S, -S + 1, \ldots, S - 1, S$. Hence, the state spin-multiplicity $2S + 1$ determines the overall degeneracy of the electronic state with respect to alternative orientations of the resultant spin. For example, the multiplicities of the representative wave functions of (5.66) and (5.67), which correspond to $S = 0$ and $S = q/2$, respectively, are 1 and $q + 1$.

The occupation numbers of shells and orbitals define the system *electron configuration*. When the (doubly occupied) spin-restricted orbitals of (5.66) involve all symmetry-related (degenerate) orbitals of each electronic subshell, this wave function is said to describe the *closed-shell* state of the molecule. Accordingly the *open-shell* state is either characterized by the singly occupied MO, as in (5.67), or it involves doubly occupied subset of the symmetry-related (degenerate) orbitals of the occupied electronic subshell(s).

## 5.4 Matrix Elements of Electronic Hamiltonian in Orbital Approximation

In order to apply the Slater determinants in the *variational* determination of the approximate electronic states, we have to derive the associated expression for the expectation value of the system electronic energy in the orbital approximation. Moreover, when mixing different determinantal wave functions in a more accurate CI variant, capable of accounting for the Coulomb correlation between electrons, one needs the related expressions for the matrix elements of the electronic Hamiltonian in such a basis set of $N$-electron functions. A short derivation of these missing elements in the independent electron approximation is the main goal of this section.

Let us first verify the normalization constant of the Slater determinant. Expanding the $\Psi_A$ and $\Psi_A{}^*$ determinants in the normalization integral gives:

$$\begin{aligned}\langle\Psi_A|\Psi_A\rangle &= \int\ldots\int \Psi_A^*(\mathbf{q})\Psi_A(\mathbf{q})\,d\mathbf{q}\\ &= \frac{1}{N!}\sum_P\sum_{P'}(-1)^{p(P)+p(P')}\prod_{j=1}^{N}\left\langle k_j(P)\mid k_j(P')\right\rangle.\end{aligned} \tag{5.68}$$

Therefore, for the orthonormal orbitals $\{\psi_k(\boldsymbol{q}) = \varphi_k(\boldsymbol{r})\zeta_k(\sigma) \equiv \langle\boldsymbol{q}|k\rangle\}$, when $\langle k|l\rangle = \delta_{k,l}$, one obtains a nonvanishing contribution in this sum only when for all electrons $k_j(P) = k_j(P')$, i.e., when the lists of orbitals $\{k_j(P)\}$ and $\{k_j(P')\}$ in permutations $P$ and $P'$ are identical, i.e., when $P = P'$, and hence

$$\langle\Psi_A|\Psi_A\rangle = \frac{1}{N!}\sum_P(-1)^{2p(P)} = \frac{N!}{N!} = 1,$$

where we have recognized that there are $N!$ distinct permutations involved in the Slater determinant of (5.64).

Let us now separately combine all *one*- and *two*-electron contributions in the electronic Hamiltonian of (5.51):

$$\begin{aligned}\hat{\mathrm{H}}^e(\mathbf{q},\mathbf{Q}) &= \sum_{i=1}^{N}\hat{\mathrm{h}}(i) + \sum_{i=1}^{N-1}\sum_{j=i+1}^{N} g(i,j) \equiv \hat{\mathcal{F}}(\mathbf{q},\mathbf{Q}) + \hat{\mathcal{G}}(\mathbf{q})\\ &\equiv \hat{\mathcal{F}}(N) + \hat{\mathcal{G}}(N);\end{aligned} \tag{5.69}$$

here, the *one*-electron Hamiltonian $\hat{\mathrm{h}}(i)$ groups the operators of the kinetic energy of $i$th electron and its attraction energy to all nuclei in their specified, fixed positions, which generate the external potential $v(i)$ (5.51),

$$\hat{\mathrm{h}}(i) = -\frac{1}{2}\Delta_i + v(i), \tag{5.70}$$

while the multiplicative operator $g(i,j)$ corresponds to the Coulomb repulsion between the indicated pair of electrons. Thus, the expectation value of the electronic energy in the state (5.64) is determined by the trivial nuclear-repulsion contribution and the sums of *one*- and *two*-electron contributions in the $N$-electron system:

$$\begin{aligned}\langle E_e\rangle_{\Psi_A} &= \langle \Psi_A \mid \hat{\mathrm{H}}_e|\Psi_A\rangle = \left\langle \Psi_A \mid \hat{\mathrm{H}}^e|\Psi_A\right\rangle + V_{nn} \equiv \langle E^e\rangle_{\Psi_A} + V_{nn},\\ \langle E^e\rangle_{\Psi_A} &= \left\langle \Psi_A \,\middle|\, \hat{\mathcal{F}}\middle|\Psi_A\right\rangle + \left\langle \Psi_A \mid \hat{\mathcal{G}}|\Psi_A\right\rangle \equiv \left\langle\hat{\mathcal{F}}\right\rangle_{\Psi_A} + \left\langle\hat{\mathcal{G}}\right\rangle_{\Psi_A}.\end{aligned} \tag{5.71}$$

Consider first the *one*-electron energy $\langle\mathcal{F}\rangle_{\Psi_A}$. Expanding the two determinants as in (5.68) gives:

$$\langle\mathcal{F}\rangle_{\Psi_A} = \frac{1}{N!}\sum_P\sum_{P'}(-1)^{p(P)+p(P')}\sum_{i=1}^{N}\langle k_i(P)\,|\,\hat{\mathrm{h}}(i)|k_i(P')\rangle\prod_{j\neq i}\langle k_j(P)\,|\,k_j(P')\rangle. \tag{5.72}$$

Again, a nonvanishing product of the overlap integrals in this expression can appear only when the two permutations are identical: $P = P'$. One also realizes that due to indistinguishability of $N$ electrons in the Slater determinant, each of them gives the same contribution as the representative electron “1” so that the above expression can be further simplified:

$$\langle\mathcal{F}\rangle_{\Psi_A} = \frac{N}{N!}\sum_P(-1)^{2p(P)}\langle k_1(P)\mid\hat{\mathrm{h}}(1)|k_1(P)\rangle = \frac{1}{(N-1)!}\sum_P \bar{h}_{k_1(P),k_1(P)}. \tag{5.73}$$

The above summation over permutations can be replaced by the equivalent summation over $N$ different choices of spin orbital $\psi_{k_1}$ describing electron 1, which defines the matrix elements $\{\bar{h}_{k_1,k_1}\}$ in the SO basis. These *one*-electron integrals should be then multiplied by their multiplicity in all permutations $P$, equal to the number $(N-1)!$ of all permutations of the remaining $(N-1)$ occupied SO $\{k_j \neq k_1\}$ among $(N-1)$ electrons $(2, 3, \ldots, N)$. Hence,

$$\langle\mathcal{F}\rangle_{\Psi_A} = \frac{(N-1)!}{(N-1)!}\sum_{k=1}^{N}\bar{h}_{k,k} = \sum_{k=1}^{N}\bar{h}_{k,k}. \tag{5.74}$$

One similarly arrives at the corresponding expression for the *two*-electron energy $\langle\mathcal{G}\rangle_{\Psi_A}$. Expanding the determinantal wave functions and taking into account the indistinguishability of electrons give:

$$\begin{aligned}\langle \mathcal{G} \rangle_{\Psi_A} &= \frac{1}{N!}\sum_{P}\sum_{P'}(-1)^{p(P)+p(P')}\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}\langle k_i(P)k_j(P) \mid g(i,j)|k_i(P')k_j(P')\rangle \\ &\quad\times \prod_{l\neq(i,j)}\langle k_l(P) \mid k_l(P')\rangle \\ &= \frac{1}{2(N-2)!}\sum_{P}\sum_{P'}(-1)^{p(P)+p(P')}\langle k_1(P)k_2(P) \mid g(1,\, 2)|k_1(P')k_2(P')\rangle \\ &\quad\times \prod_{l\neq(1,2)}\langle k_l(P) \mid k_l(P')\rangle,\end{aligned}$$

where we have recognized that each of the $N(N-1)/2$ electronic pairs gives the same contribution as the representative pair (1, 2). A subsequent examination of the overlap integrals in the product indicates that all SO for remaining electrons $l \neq (1,\, 2)$ in permutations $P$ and $P'$ must be identical for electrons $(3,\, 4,\, \ldots,\, N)$. Therefore, the nonvanishing contributions arise only when the two permutations are identical, $P = P'$, or when they differ only in orbitals describing electrons 1 and 2: $P = \hat{\mathrm{P}}(1,\, 2)P'$. In the former case, the parities of both permutations are equal, giving rise to factor $(-1)^{p(P)+p(P')} = (-1)^{2p(P)} = 1$, while in the latter case they differ by one exchange of two electrons, so that $(-1)^{p(P)+p(P')} = -1$. Moreover, for each choice of the two SO describing electrons 1 and 2, we thus have $(N-2)!$ permutations of the remaining $(N-2)$ orbitals $\{k_l \neq (k_1,\, k_2)\}$ among $(N-2)$ electrons $(3,\, 4,\, \ldots,\, N)$. Therefore, the preceding expression can be expressed in terms of contributions from *two*-electron integrals:

$$\begin{aligned}\langle \mathcal{G} \rangle_{\Psi_A} &= \frac{(N-2)!}{2(N-2)!}\sum_{k=1}^{N}\sum_{l=1}^{N}\left[\langle k(1)l(2)\,|\,g(1,2)|k(1)l(2)\rangle - \langle k(1)l(2)\,|\,g(1,2)|l(1)k(2)\rangle\right] \\ &\equiv \frac{1}{2}\sum_{k=1}^{N}\sum_{l=1}^{N}\left[\langle kl|g|kl\rangle - \langle kl|g|lk\rangle\right] \equiv \frac{1}{2}\sum_{k=1}^{N}\sum_{l=1}^{N}\left[(kk|ll) - (kl|lk)\right] \\ &\equiv \frac{1}{2}\sum_{k=1}^{N}\sum_{l=1}^{N}\left[\bar{J}_{k,l} - \bar{K}_{k,l}\right] = \sum_{k=1}^{N-1}\sum_{l=k+1}^{N}\left[\bar{J}_{k,l} - \bar{K}_{k,l}\right].\end{aligned} \tag{5.75}$$

The *two*-electron integrals for the specified pair of SO describing the states of the representative electrons "1" and "2,"

$$\begin{aligned}\bar{J}_{k,l}[\psi_k, \psi_l] &= \langle k(1)l(2) \mid g(1,\, 2)|k(1)l(2)\rangle = \iint |\psi_k(1)|^2 g(1,2)\left|\psi_l^2(1)\right|^2 d\boldsymbol{q}_1 d\boldsymbol{q}_2 \\ &\equiv \bar{J}_{k,l},\end{aligned}$$

$$\begin{aligned}\bar{K}_{k,l}[\psi_k, \psi_l] &= \langle k(1)l(2) \mid g(1,\, 2)|l(1)k(2)\rangle \\ &= \iint \psi_k^*(1)\psi_l(1)g(1,\, 2)\,\psi_l^*(2)\psi_k(2)d\boldsymbol{q}_1 d\boldsymbol{q}_2 \equiv \bar{K}_{k,l},\end{aligned} \tag{5.76}$$

are called the *Coulomb* and *exchange* integrals, respectively. The former is indeed seen to measure the Coulomb interaction between the charge distributions of electrons occupying SO $\psi_k$ and $\psi_l$, respectively.

Since *two*-electron "integrations" involve summation over spin variables, the exchange integrals identically vanish for the two electrons with opposite spins, due to orthogonality of their spin functions (3.76),

$$\begin{aligned}\bar{K}_{k,l}[\psi_k,\psi_l] =& \left(\iint \varphi_k^*(\boldsymbol{r}_1)\varphi_l(\boldsymbol{r}_1)g(\boldsymbol{r}_1,\boldsymbol{r}_2)\,\varphi_l^*(\boldsymbol{r}_2)\varphi_k(\boldsymbol{r}_2)d\boldsymbol{r}_1 d\boldsymbol{r}_2\right)\left[\sum_{\sigma_1}\zeta_k^*(\sigma_1)\zeta_l(\sigma_1)\right]\\ &\times\left[\sum_{\sigma_2}\zeta_l^*(\sigma_2)\zeta_k(\sigma_2)\right] = K_{k,l}[\varphi_k,\varphi_l]\delta_{\zeta_k,\zeta_l}\equiv K_{k,l}\delta_{\zeta_k,\zeta_l},\end{aligned} \tag{5.77}$$

where $K_{k,l}$ stands for the exchange integral defined by the specified pair of the spatial functions (*orbitals*).

It should be observed that no such restriction intervenes in calculating the Coulomb integrals:

$$\begin{aligned}\bar{J}_{k,l}[\psi_k,\psi_l] =& \left(\iint |\varphi_k(\boldsymbol{r}_1)|^2 g(\boldsymbol{r}_1,\boldsymbol{r}_2)\left|\varphi_l^2(\boldsymbol{r}_2)\right|^2 d\boldsymbol{r}_1 d\boldsymbol{r}_2\right)\left[\sum_{\sigma_1}|\zeta_k(\sigma_1)|^2\right]\left[\sum_{\sigma_2}|\zeta_l(\sigma_2)|^2\right]\\ &= J_{k,l}[\varphi_k,\varphi_l]\equiv J_{k,l},\end{aligned} \tag{5.78}$$

where the sums in the square brackets are both equal to 1 by the normalization condition of the spin states [see (3.76)]:

$$\begin{aligned}\langle\alpha|\alpha\rangle &= \sum_\sigma\langle\alpha|\sigma\rangle\langle\sigma|\alpha\rangle = \sum_\sigma|\alpha(\sigma)|^2 = \langle\beta|\beta\rangle = \sum_\sigma\langle\beta|\sigma\rangle\langle\sigma|\beta\rangle\\ &= \sum_\sigma|\beta(\sigma)|^2 = 1.\end{aligned} \tag{5.79}$$

It also follows from (5.77) and (5.78) that $J_{k,k} = K_{k,k}$, $J_{k,l} = J_{l,k}$, and $K_{k,l} = K_{l,k}$, since the value of the electron repulsion energy must be independent of the subjectively assigned labels of electrons. This justifies the final expression in (5.75), involving only the *off*-diagonal terms in the double summation $\sum_{k<l}$.

The same is true for the *one*-electron integrals $\bar{h}_{k,k}\equiv\bar{h}_{k,k}[\psi_k]$:

$$\begin{aligned}\bar{h}_{k,k}[\psi_k] &= \int\psi_k^*(\boldsymbol{q})\,\hat{\mathrm{h}}(\boldsymbol{r})\,\psi_k(\boldsymbol{q})\,d\boldsymbol{q} = \left(\int\varphi_k^*(\boldsymbol{r})\,\hat{\mathrm{h}}(\boldsymbol{r})\,\varphi_k(\boldsymbol{r})\,d\boldsymbol{r}\right)\left[\sum_\sigma|\zeta_k(\sigma)|^2\right]\\ &= h_{k,k}[\varphi_k]\equiv h_{k,k}.\end{aligned} \tag{5.80}$$

For example, for the wave function (5.66), one obtains the following functional for the expectation value of the electronic energy:

$$\langle E^e \rangle_{\Psi_A} = 2\sum_{i=1}^{p} h_{i,i} + \sum_{i=1}^{p}\sum_{j=1}^{p}(2J_{i,j} - K_{i,j}). \tag{5.81}$$

Clearly, its numerical value depends on the shapes of the $p$ doubly occupied orbitals defining the associated *one*- and *two*-electron integrals. In the closed-shell ground state configuration $[1s^2]$ of the helium atom, one thus finds

$$\langle E^e(\mathrm{He}) \rangle_{\Psi_0} = 2h_{1s,1s} + J_{1s,1s}, \tag{5.82}$$

while for the ground state of beryllium, defined by configuration $[1s^2 2s^2]$, one predicts

$$\langle E^e(\mathrm{Be}) \rangle_{\Psi_0} = 2h_{1s,1s} + 2h_{2s,2s} + J_{1s,1s} + J_{2s,2s} + 4J_{1s,2s} - 2K_{1s,2s}. \tag{5.83}$$

The same result can be heuristically derived by summing the elementary *one*-electron energies of all $N$ electrons, the expectation values of the Hamiltonian (5.70), and the repulsion energies in each of the $N(N-1)/2$ different electronic pairs. Indeed, the average interaction energy between two (indistinguishable) electrons (1, 2) occupying spin orbitals $\psi_i$ and $\psi_j$ is given by the expectation value of the $g(1, 2)$ in the Slater determinant $\Psi_A(2) = |\psi_i\psi_j| = 2^{-1/2}[\psi_i(1)\psi_j(2) - \psi_i(2)\psi_j(1)]$:

$$\langle \Psi_A(2)|g(1,2)|\Psi_A(2)\rangle = \bar{J}_{i,j} - \bar{K}_{i,j} = \begin{cases} J_{i,j}, & \text{for different spin states} \\ J_{i,j} - K_{i,j}, & \text{for identical spin states} \end{cases}.$$

Thus, for the two spin-paired electrons in He atom, when $\bar{K}_{i,j} = 0$, one reproduces the result of (5.82). It can be also easily verified that for beryllium atom in the ground state $\Psi_A(4) = |1s^+1s^-2s^+2s^-|$, one recovers (5.83).

The expectation value $\left\langle \Psi_A \mid \hat{\mathrm{H}}^e \middle| \Psi_A \right\rangle = \left\langle \Psi_0 \mid \hat{\mathrm{H}}^e \middle| \Psi_0 \right\rangle$ (5.71) represents a particular, ground state case of a general *diagonal* matrix element of the electronic Hamiltonian, $\left\langle \Psi_n \mid \hat{\mathrm{H}}^e \middle| \Psi_n \right\rangle$, for any antisymmetric electronic state (Slater determinant) $\Psi_n$ specified by alternative choices of $N$ occupied, orthonormal SO. The same energy formulas also apply to excited electron configurations $\Psi_{n>0}$ obtained by replacing some of the SO occupied in $\Psi_0$, by the *virtual* orbitals, unoccupied in $\Psi_0$. Thus, given the modified list of SO occupied in $\Psi_n$, *occd.*$[n] \equiv [n]$, a general formula for the expectation value of the electronic energy remains unaffected:

$$\begin{aligned}\langle E^e \rangle_{\Psi_n} &= \left\langle \Psi_n \mid \hat{\mathrm{H}}^e \middle| \Psi_n \right\rangle = \sum_{k}^{[n]} \bar{h}_{k,k} + \frac{1}{2}\sum_{k}^{[n]}\sum_{l}^{[n]}\left[\bar{J}_{k,l} - \bar{K}_{k,l}\right] \\ &\equiv \sum_{k}^{[n]}\left\{\langle k|\hat{\mathrm{h}}|k\rangle + \sum_{l}^{[n]}\langle kl|\,g\,|kl - lk\rangle\right\}, \quad \langle \Psi_n|\Psi_n\rangle = \prod_{i}^{[n]}\langle i \mid i\rangle = 1.\end{aligned} \tag{5.84}$$

Finally, let us examine the *off-diagonal* matrix elements $\left\langle \Psi_n \mid \hat{\mathrm{H}}^e \middle| \Psi_{n'} \right\rangle$ between two electron configurations, differing in the list of the occupied SO. Due to the orthogonality of SO such configurations can be shown to be also automatically orthogonal: $\langle \Psi_n \mid \Psi_{n'} \rangle = 0$ [see (5.68)]. The sets of the occupied SO in these two states may still exhibit some common SO, be it in different positions (rows) of two determinantal functions. Therefore, for definiteness, we assume that by appropriate exchange of rows in one of these two Slater determinants, the two configurations have been brought to the *maximum coincidence* form, in which the rows of the common SO of both configurations appear in the same positions in both determinants. We already know from the properties of the Slater determinant that such exchanges can at best change the sign of the wave function, which can be diagnosed from the known number of the *row* permutations in the original Slater determinant required for reaching this maximum coincidence. This sign can then be used to multiply the matrix element obtained from the maximum coincidence rules.

There are three general types of differences between such prearranged lists of SO in both Slater determinants, giving rise to the associated expressions for the matrix element of the electronic Hamiltonian. They can be derived in a way analogous to that used to derive the diagonal element, by expanding both determinants in terms of the permuted products of SO, applying the SO orthonormality relations, and recognizing the indistinguishability of electrons. The relevant cases are summarized by the following *Slater–Condon rules* (Slater 1929; Condon 1930):

1. Configurations $\Psi_n$ and $\Psi_{n'}$ differ only in a single SO, with $\psi_p$ of the former being replaced by $\psi_r$ in the latter, as a result of the electron excitation $\psi_p \to \psi_r$,

$$\left\langle \Psi_n \middle| \hat{\mathrm{H}}^e \middle| \Psi_{n'} \right\rangle = \langle p|\hat{\mathrm{h}}|r\rangle + \sum_j^{[n]} \langle pj|g|rj - jr\rangle,$$
$$\langle \Psi_n|\Psi_{n'}\rangle = \langle p \mid r\rangle \prod_{i\neq p}^{[n]} \langle i \mid i\rangle = 0; \tag{5.85}$$

2. Configurations $\Psi_n$ and $\Psi_{n'}$ differ only in two SO, as a result of the double excitation $(\psi_p \to \psi_{r,}\ \psi_q \to \psi_s)$ or $(\psi_{p,}\ \psi_q) \to (\psi_r,\ \psi_s)$,

$$\left\langle \Psi_n \middle| \hat{\mathrm{H}}^e \middle| \Psi_{n'} \right\rangle = \langle pq|g|rs - sr\rangle,$$
$$\langle \Psi_n|\Psi_{n'}\rangle = \langle p|r\rangle\langle q|s\rangle \prod_{i\neq(p,q)}^{[n]} \langle i \mid i\rangle = 0. \tag{5.86}$$

3. Configurations $\Psi_n$ and $\Psi_{n'}$ differ in more than two SO, thus reflecting the triple or higher excitations,

$$\left\langle \Psi_n \middle| \hat{\mathrm{H}}^e \middle| \Psi_{n'} \right\rangle = 0, \quad \langle \Psi_n|\Psi_{n'}\rangle = 0. \tag{5.87}$$

To summarize, the determinantal functions corresponding to the system ground and excited configurations, in which electrons have been excited from some $\Psi_0$-occupied to the corresponding $\Psi_0$-virtual SO of the molecule, form the orthonormal basis of $N$-electron functions. The Slater–Condon rules allow one to express their average energies and the coupling matrix elements between them in terms of the elementary *one*- and *two*-electron integrals involving SO, the elementary *one*-electron functions of the orbital approximation.

## 5.5 Example: Helium Atom

As an illustration we shall now apply the approximate methods to estimate the ground state energy of the helium atom, when its two electrons occupy the lowest orbital 1$s$, from the known solutions for the *one*-electron atom [(4.58) and (4.62)]. The a.u. are used throughout.

### 5.5.1 Perturbation Approximation

The internal (Coulomb) Hamiltonian of two electrons moving in the field of the helium nucleus ($Z = 2$) determines the *perturbed* Hamiltonian (a.u.),

$$\begin{aligned} &\hat{\mathrm{H}}(\boldsymbol{r}_1,\boldsymbol{r}_2) = \hat{\mathrm{H}}^0(\boldsymbol{r}_1,\boldsymbol{r}_2) + \hat{\mathrm{h}}(\boldsymbol{r}_1,\boldsymbol{r}_2), \quad &&\hat{\mathrm{h}}(\boldsymbol{r}_1,\boldsymbol{r}_2) = g(1,2) = |\boldsymbol{r}_1-\boldsymbol{r}_2|^{-1} = {r_{1,2}}^{-1},\\ &\hat{\mathrm{H}}^0(\boldsymbol{r}_1,\boldsymbol{r}_2) = \hat{\mathrm{H}}_1(\boldsymbol{r}_1) + \hat{\mathrm{H}}_2(\boldsymbol{r}_2), &&\hat{\mathrm{H}}_i(\boldsymbol{r}_i) = -\frac{1}{2}\Delta_i - \frac{Z}{r_i}, \quad i = 1,2, \end{aligned} \tag{5.88}$$

with the electron repulsion operator representing the *perturbation* and the separable Hamiltonian $\hat{\mathrm{H}}^0(\boldsymbol{r}_1,\boldsymbol{r}_2)$ given by the sum of the hydrogen-like operators $\{\hat{\mathrm{H}}_i(\boldsymbol{r}_i)\}$ of two electrons determining the *unperturbed* Hamiltonian. This assumption thus defines the unperturbed solutions:

$$E_0^{(0)} = -Z^2 = -4, \quad \Psi_0^{(0)}(\boldsymbol{r}_1,\boldsymbol{r}_2) = \psi_{1s}(\boldsymbol{r}_1)\psi_{1s}(\boldsymbol{r}_2) = \frac{Z^3}{\pi}\exp[-Z(r_1+r_2)], \tag{5.89}$$

since the hydrogen-like solutions for each electron read:

$$\psi_{1s}(\boldsymbol{r}_i) = \sqrt{\frac{Z^3}{\pi}}\exp(-Zr_i), \quad \hat{\mathrm{H}}_i(\boldsymbol{r}_i)\,\psi_{1s}(\boldsymbol{r}_i) = -\frac{Z^2}{2}\psi_{1s}(\boldsymbol{r}_i), \quad i = 1,2. \tag{5.90}$$

This energy estimate should be compared with the experimental value $E_0 = -2.9037$.

The first correction to this crude estimate [see (5.11)] determines the expectation value in state $\Psi_0^{(0)}$ of the repulsion energy operator (perturbation):

$$\Delta E_0^{(1)} = \left\langle \Psi_0^{(0)} \middle| \hat{\mathrm{h}} \middle| \Psi_0^{(0)} \right\rangle = \iint \psi_{1s}^*(\boldsymbol{r}_1)\psi_{1s}^*(\boldsymbol{r}_2)\frac{1}{r_{1,2}}\psi_{1s}(\boldsymbol{r}_1)\psi_{1s}(\boldsymbol{r}_2)\, d\boldsymbol{r}_1\, d\boldsymbol{r}_2. \tag{5.91}$$

To calculate this integral we first assume the spherical coordinates of electron "1" relative to the nucleus, $\boldsymbol{r}_1 = (r_1, \vartheta_1, \varphi_1)$, adopt the relative spherical coordinates of electron "2" with respect to electron "1," $\boldsymbol{r}_{1,2} = \boldsymbol{r}_2 - \boldsymbol{r}_1 = (r_{1,2}, \theta, \phi)$, and use Carnot's cosine formula to express the *inter*-electron separation:

$$r_{1,2} = (r_1^2 + r_2^2 - 2r_1r_2\cos\theta)^{1/2}. \tag{5.92}$$

Hence, the *first*-order correction of (5.91) reads:

$$\begin{aligned}\Delta E_0^{(1)} = &\frac{Z^6}{\pi^2}\left(\int_0^\pi \sin\vartheta_1\, d\vartheta_1 \int_0^{2\pi} d\varphi_1\right)\left(\int_0^{2\pi} d\phi\right)\\ &\times\left(\int_0^\infty r_1^2 e^{-2Zr_1}\int_0^\infty r_2^2 e^{-2Zr_2}\int_0^\pi (r_1^2 + r_2^2 - 2r_1r_2\cos\theta)^{-1/2}\sin\theta\, d\theta\, dr_1 dr_2\right).\end{aligned} \tag{5.93}$$

We then substitute $x = \cos\theta$:

$$\begin{aligned}\int_0^\pi (r_1^2 + r_2^2 - 2r_1r_2\cos\theta)^{-1/2}\sin\theta\, d\theta &= \int_{-1}^1 (r_1^2 + r_2^2 - 2r_1r_2x)^{-1/2} dx\\ &= -\frac{1}{r_1r_2}\int_{r_{1,2}(x=-1)}^{r_{1,2}(x=1)} dr_{1,2} = \begin{cases}\frac{2}{r_1}, & r_1 > r_2\\ \frac{2}{r_2}, & r_1 < r_2\end{cases}.\end{aligned} \tag{5.94}$$

Using this result and typical integral $\int_0^\infty y^n \exp(-ay)dy = n!/a^{n+1}$in (5.93) gives:

$$\begin{aligned}\Delta E_0^{(1)} &= 16Z^6\int_0^\infty r_1^2 e^{-2Zr_1}\left(\frac{1}{r_1}\int_0^{r_1} r_2^2 e^{-2Zr_2} dr_2 + \int_{r_1}^\infty r_2 e^{-2Zr_2} dr_2\right) dr_1\\ &= 4Z^3\int_0^\infty r_1^2 e^{-2Zr_1}\left[\frac{1}{r_1} - \left(Z + \frac{1}{r_1}\right)e^{-2Zr_1}\right] dr_1 = \frac{5Z}{8}.\end{aligned} \tag{5.95}$$

Thus, in the *first* order of the perturbation theory, one estimates the ground state energy of the helium atom as being much closer to the experimental value,

$$E_0^{(1)} = -Z^2 + 5Z/8 = -2.75. \tag{5.96}$$

### 5.5.2 Variational Estimates

The simplest trial wave function of this *two*-electron system in the spin-restricted ($R$) approximation is given by the product function of (5.89), when one replaces $Z$ with an effective nuclear charge $\zeta$, a nonlinear variational parameter,

$$\Phi(\boldsymbol{r}_1,\boldsymbol{r}_2;\zeta) = \frac{\zeta^3}{\pi}\exp[-\zeta(r_1+r_2)] = \psi_{1s}(\boldsymbol{r}_1;\zeta)\psi_{1s}(\boldsymbol{r}_2;\zeta) \equiv \Phi^R(\boldsymbol{r}_1,\boldsymbol{r}_2;\zeta), \quad (5.97)$$

where the normalized trial orbital $\psi_{1s}(\boldsymbol{r}_i;\zeta)$ satisfies the energy eigenvalue equation for the *one*-electron atom described by Hamiltonian $\hat{\mathrm{H}}(\boldsymbol{r};\zeta) = -\frac{1}{2}\Delta - \frac{\zeta}{r}$,

$$\psi_{1s}(\boldsymbol{r}_i;\zeta) = \sqrt{\frac{\zeta^3}{\pi}}\exp(-\zeta\, r_i), \quad \hat{\mathrm{H}}(\boldsymbol{r}_i;\zeta)\,\psi_{1s}(\boldsymbol{r}_i;\zeta) = -\frac{\zeta^2}{2}\psi_{1s}(\boldsymbol{r}_i;\zeta), \quad i=1,2. \quad (5.98)$$

Since each electron in the helium atom experiences a diminished attraction compared with that in He$^+$, due to a partial screening of the nucleus by the other electron, one expects the optimum value of this variational parameter to be in the range $1 < \zeta < 2$.

In order to express the average electronic energy, the expectation value of the Hamiltonian $\hat{\mathrm{H}}(\boldsymbol{r}_1,\boldsymbol{r}_2)$ (5.88), we first express the latter in terms of the effective *one*-electron Hamiltonians $\{\hat{\mathrm{H}}(\boldsymbol{r}_i;\zeta)\}$:

$$\hat{\mathrm{H}}(\boldsymbol{r}_1,\boldsymbol{r}_2) = \hat{\mathrm{H}}(\boldsymbol{r}_1;\zeta) + \hat{\mathrm{H}}(\boldsymbol{r}_2;\zeta) + (\zeta - Z)(r_1^{-1} + r_2^{-1}) + r_{12}^{-1}. \quad (5.99)$$

Hence, using the eigenvalues of (5.98) and the electron repulsion energy of (5.95) for $Z = \zeta$, gives

$$\begin{aligned} E(\zeta) &= \iint \Phi^*(\boldsymbol{r}_1,\boldsymbol{r}_2;\zeta)\hat{\mathrm{H}}(\boldsymbol{r}_1,\boldsymbol{r}_2)\,\Phi(\boldsymbol{r}_1,\boldsymbol{r}_2;\zeta)\,d\boldsymbol{r}_1 d\boldsymbol{r}_2 \\ &= \frac{\zeta^6}{\pi^2}\iint e^{-2\zeta(r_1+r_2)}[-\zeta^2 + (\zeta - Z)(r_1^{-1} + r_2^{-1}) + r_{12}^{-1}]\,d\boldsymbol{r}_1 d\boldsymbol{r}_2 \\ &= -\zeta^2 + 2(\zeta - Z)\zeta + 5\zeta/8 = \zeta^2 - 27\zeta/8. \end{aligned} \quad (5.100)$$

The optimum value of the effective nuclear charge is then determined by the variational condition:

$$\left.\frac{dE(\zeta)}{d\zeta}\right|_{\zeta^{opt.}} = 2\zeta^{opt.} - 27/8 = 0 \quad \text{or} \quad \zeta^{opt.} = 27/16 \cong 1.69 < Z. \quad (5.101)$$

It gives the following estimate of the helium ground state energy (Hylleraas 1828):

$$E(\zeta^{opt.}) = -(27/16)^2 \cong -2.848 > E_0 = -2.9037. \tag{5.102}$$

It should be realized that the full electronic wave function for the helium atom in the spin-*restricted* variant also includes the spin singlet function of two electrons,

$$\begin{aligned}\Psi[(\boldsymbol{r}_1, \sigma_1), (\boldsymbol{r}_2, \sigma_2)] &= \Psi^R(\boldsymbol{q}_1, \boldsymbol{q}_2)\\ &= \Phi^R(\boldsymbol{r}_1, \boldsymbol{r}_2)\frac{1}{\sqrt{2}}[\alpha(\sigma_1)\beta(\sigma_2) - \beta(\sigma_1)\alpha(\sigma_2)].\end{aligned} \tag{5.103}$$

In the spin-*unrestricted* (*U*) approximation, using different orbitals for different spins, the spatial function depends on two nonlinear variational parameters:

$$\Phi^U(\boldsymbol{r}_1, \boldsymbol{r}_2; \zeta_1, \zeta_2) = \frac{1}{\sqrt{2}}[\psi_{1s}(\boldsymbol{r}_1; \zeta_1)\psi_{1s'}(\boldsymbol{r}_2; \zeta_2) + \psi_{1s'}(\boldsymbol{r}_1; \zeta_2)\psi_{1s}(\boldsymbol{r}_2; \zeta_1)]. \tag{5.104}$$

Its symmetrization with respect to the exchange of the position vectors of two electrons is required by the Pauli exclusion principle. The product of the spatial (symmetric) wave function and the singlet (antisymmetric) spin function is then antisymmetric with respect to the permutation of two electrons. The optimum values of these exponents of the two 1*s* orbitals, $\zeta_1^{opt.} = 1.19$ and ${\zeta_2}^{opt.} = 2.18$, give $E({\zeta_1}^{opt.}, {\zeta_2}^{opt.}) = -2.876 < E(\zeta^{opt.})$ thus offering a better approximation (variational flexibility) of the ground state wave function compared with the spin-restricted analog.

Indeed, it follows from (5.104) that two electrons in the spin-unrestricted state correlate (radially) their movements around the nucleus: when one electron occupies a more compact orbital, thus being on average closer to the nucleus, the other electron occupies the more diffuse orbital, thus exhibiting larger average separation from the nucleus. Therefore, the average *inter*-electron distance in the spin unrestricted state is expected to be larger compared with that in the spin-restricted case, in which both electrons are kept within confines of the same orbital.

## 5.6 Idea of a Pseudopotential

As we have already observed at the end of Sect. 5.1.2, the atomic cores are predicted to remain largely inactive (invariant) in the chemical processes of the bond forming and/or bond breaking. Indeed, it directly follows from the AO-mixing criteria that the system chemical bonds must be shaped by the *valence* shells of constituent atoms. This observation is particularly important for heavy *many*-electron atoms, in which the number of the *inner*-shell electrons ($n$) is much larger than the complementary number of the chemically active, *valence* shell electrons: $N - n \ll n$. It is thus tempting to formulate the chemical bond theory focused solely on the quantum states of the valence electrons, since such a reduction of the problem gives rise to a significant conceptual and mathematical simplification while still retaining all significant sources of the chemical binding in molecular systems.

It is evident from the orthogonality relations between orbitals, which generate the familiar nodal structure of valence orbitals in the regions of atomic cores, that any change of the orbitals in one set affects the shapes of orbitals in the other set. Therefore, such constraints appear to prevent any formulation of the "valence-only" theory, in which the optimized valence orbitals satisfy the orthogonality constraints to the "frozen" core orbitals. However, the Pauli antisymmetrization postulate for fermions (Sect. 5.3) is automatically satisfied by the determinantal wave function, no matter whether the orbitals defining the Slater determinant are mutually orthogonal or not. Thus, the requirement that the valence orbitals be orthogonal to the core orbitals is not actually needed to satisfy the exclusion principle. The pseudopotential theory makes use of this very property in designing the "valence-only" theory of molecular electronic structure.

Achieving this goal calls for a nontrivial replacement of the valence core orthogonality requirement by an equivalent theoretical concept, which turns out to represent an additional (nonclassical) operator or the associated local potential term in the effective Hamiltonian of the valence electron, called the *pseudopotential* (Hellmann 1935; Gombas 1967; Slater 1960, 1974; Szasz 1985). It explicitly depends on the shapes and energies of the core orbitals. The *pseudoorbitals*, the eigenfunctions of this effective Hamiltonian, are determined as if the core did not exist at all, its presence being felt exclusively through the pseudopotential. The lowest pseudoorbital exhibits no nodes in the atomic core regions and generates the maximum probability density in the system valence shell.

The first attempt in this direction was made by Hellmann (Hn) (Hellmann 1935), who introduced the very idea of replacing the orthogonality requirement by the pseudopotential within the statistical Thomas–Fermi (TF) model (Thomas 1927; Fermi 1928; Gombas 1949; March 1975). He also proposed the use in atomic and molecular calculations of the model atomic potential for the valence electron, say in Na atom, including the Coulomb attraction $V_q(r)$ due to the core net charge $q = Z - n$ and the model pseudopotential $V_p(r)$:

$$V_{\mathrm{Hn}}(r) = V_q(r) + V_p(r) = -\frac{q}{r} + A\frac{\exp(-\kappa r)}{r}. \tag{5.105}$$

Here, the adjustable parameters $A$ and $\kappa$ are determined by fitting the predicted eigenvalues of the associated effective Hamiltonian for the valence electron,

$$\hat{\mathrm{H}}_{\mathrm{Hn}}(r) = -(1/2)\Delta + V_{\mathrm{Hn}}(r) \equiv \hat{\mathrm{H}}_{\mathrm{Na}}^{val.}(1), \tag{5.106}$$

to the atom experimental energy spectrum in the valence shell regime.

Notice that the pseudopotential part of this effective potential generates at small distances from the nucleus the repulsive wall, which prevents the valence orbital (nodeless in the core range) to assume large values in the region occupied by the inner-shell electrons. In other words, this repulsive barrier prevents the valence electron from falling into the core, despite the lack of the core–valence orthogonality. Thus, in the pseudopotential approach, the usual *geometrical* constraints of the orbital orthogonality are replaced by equivalent *physical* barriers preventing the valence orbitals to collapse into the inner-shell region.

Hellmann has applied the same idea to atoms with many valence electrons. For example, by determining the model potential of (5.105) to fit as accurately as possible the valence energy spectrum of $Mg^+$, containing a single valence electron, one writes the valence-only Hamiltonian for the neutral atom, containing two valence electrons, in the following form:

$$\hat{\mathrm{H}}_{\mathrm{Mg}}^{val.}(1,\,2) = \hat{\mathrm{H}}_{\mathrm{Hn}}(1) + \hat{\mathrm{H}}_{\mathrm{Hn}}(2) + 1/r_{12}. \tag{5.107}$$

The molecular applications, e.g., to $Na_2 = Na_A - Na_B$, is also valid. Since $Na^+$ cores do not participate in this $\sigma$ bond, the effective Hamiltonian for the two valence electrons in the molecular sodium can be simplified as follows:

$$\hat{\mathrm{H}}_{\mathrm{Na}_2}^{val.}(1,\,2) = \sum_{i=1}^{2}\left[-\frac{1}{2}\Delta_i + V_{\mathrm{Hn}}(r_{iA}) + V_{\mathrm{Hn}}(r_{iB})\right] + \frac{1}{r_{12}}, \quad r_{iX} = |\boldsymbol{r}_i - \boldsymbol{R}_X|. \tag{5.108}$$

The pseudopotential method generates the exact “valence only” formalism for atoms and molecules. It can be formulated either as a model (semiempirical) procedure or as the ab initio theory in the spirit of the Phillips and Kleinman (PK) (Phillips and Kleinman 1959, 1960) treatment, which defines the *local* pseudopotential corresponding to the *pseudopotential operator* representing the Pauli exclusion principle (core–valence orthogonality). The latter aspect will be addressed in Sect. 6.1.5.

Having outlined the basic idea of the local pseudopotential, without attempting at this stage to present the ab initio theory in a comprehensive way, let us complete this section with just a short comment on some of the method’s most attractive aspects. The effects of the exclusion principle and orbital orthogonality have been shown to be *exactly* representable in the form of the associated effective local pseudopotential. This localization of operators in the PK theory is very much in spirit of that later used in the semiempirical methods of quantum chemistry and in the modern DFT, to establish the effective Hamiltonian determining the KS orbitals. The local pseudopotentials are more suitable for an analytical representation, in the form of model potentials, both semiempirical and those having ab initio justifications as their background. It brings a deeper, physical understanding of the core–valence separation problem. Indeed, the local pseudopotential has permitted a plausible physical interpretation and opened the way to wide range of applications and modeling. It has provided the causal picture of the quantum states of valence electrons in atomic or molecular systems, as moving in the resultant potential generated by atomic cores, including both their electrostatic potentials and the “Pauli” term mimicking the valence–core orthogonality constraint.

The pseudopotential theory was first applied in 1930 and 1940 to problems of the solid state physics, to build up the quantum theory of metals, and then – starting from 1960 – it has been developed as an alternative theory of atoms and molecules. This conceptually appealing approach is mathematically coherent and elegant, particularly in its ab initio formulation. Its conceptual simplicity, still combined with remarkable accuracy, facilitates a subsequent modeling of atomic and molecular phenomena. These models also include those derived from the ab initio theory. A good exposition of the method origins and capabilities is given in the monograph by Szasz (1985), the chief proponent of the molecular applications of the pseudopotential theory.

## References

Born M, Oppenheimer JR (1927) Ann Phys 64:457
Boys SF, Foster J (1960) Rev Mod Phys 32:305
Condon EU (1930) Phys Rev 36:1121
Edmiston C, Ruedenberg K (1963) Rev Mod Phys 32:457
Fermi E (1928) Z Phys 48:73
Fock VA (1930) Z Phys 61:126
Gombas P (1949) The statistical theory of atoms and its applications. Springer, Vienna
Gombas P (1967) Pseudopotentials. Springer, Vienna
Hartree DR (1928) Proc Camb Philos Soc 24:89
Hellmann H (1935) J Chem Phys 3:61
Hylleraas EA (1828) Z Physik 48:469
Kohn W, Sham LJ (1965) Phys Rev 140A:1133
Longuet-Higgins HC (1961) Adv Spectrosc 2:429
March NH (1975) Self-consistent fields in atoms. Pergamon, Oxford
Murrell JN, Bosanac SD (1989) Introduction to the theory of atomic and molecular collisions. Wiley, New York
Murrell JN, Carter S, Farantos SC, Huxley P, Varandas AJC (1984) Molecular potential energy functions. Wiley, New York
Phillips JC, Kleinman L (1959) Phys Rev 116:287
Phillips JC, Kleinman L (1960) Phys Rev 118:1153
Slater JC (1929) Phys Rev 34:1293
Slater JC (1931) Phys Rev 38:1109
Slater JC (1960) Quantum theory of atomic structure, vols 1 and 2. McGraw-Hill, New York
Slater JC (1974) Quantum theory of molecules and solids, vol 4: The self-consitent field for molecules and solids. McGraw-Hill, New York
Szasz L (1985) Pseudopotential theory of atoms and molecules. Wiley, New York
Thomas LH (1927) Proc Camb Philos Soc 23:542