Chapter 13
If *Truncated* Wave Functions of Excited State Energy Saddle Points Are Computed as Energy *Minima*, Where Is the Saddle Point?

N. C. Bacalis

Abstract Theoretical computations tend to compute electronic properties of increasingly larger systems. To understand the properties, we should rather need small truncated but concise and comprehensible wave functions. For electronic processes, in particular charge transfer, which occur in excited states, we need both the energy and the wave function in order to draw and predict correct conclusions. But the excited states are saddle points in the Hilbert space, and, as shown here, the standard methods for excited states, based on the Hylleraas-Undheim and MacDonald (HUM) theorem, compute indeed the correct energy but may give misleadingly incorrect truncated wave functions, because they search for *an energy minimum, not a saddle point* (many functions can have the correct energy). Then, **where is the saddle point?** We shall see the use of a functional F_n of the wave function that has a local minimum **at** the excited state saddle point, without using orthogonality to approximants of lower-lying states, provided these approximants are reasonable, even if they are crude. Therefore F_n finds a correct, albeit small and concise, thus comprehensible truncated wave function, approximant of the desired excited state saddle point, allowing correct predictions for the electronic process. This could also lead to computational developments of more appropriate (to excited state) truncated basis sets. It is further shown that, via a correct approximant of the 1st excited state, we can improve the ground state. Finally it is shown that, in iterative computations, in cases of "root flipping" (which would deflect the computation), we can use F_n to identify the flipped root. For all the above, demonstrations are given for excited states of He and Li. The grand apophthegm is that HUM finds an energy **minimum** which, only if the expansion is increased, can approach the excited state saddle point, whereas F_n has local minimum **at** the saddle point, so it finds it independently of the size of the expansion.

N. C. Bacalis (✉)
Theoretical and Physical Chemistry Institute, National Hellenic Research Foundation, Athens, Greece
e-mail: nbacalis@eie.gr

13.1   Introductory Remark

Suppose that we want to know the behavior of some material when we bring atoms close together, for example, we want to know the behavior of a catalyst or we want to invent a new battery. After the contact of the atoms, where will the electrons go? Will they be excited? How will they drive the surrounded nuclei? Will they go to the lowest unoccupied orbital of the ground state or to the highest occupied orbital of the excited state? When the charge transfer occurs in appropriate excited electronic states, the electrons are transferred to appropriate orbitals of the surrounding species. Therefore, we must know not only the excitation energy **but also the wave function of the excited state**. But, the excited states are saddle points in the Hilbert space, and the standard computational methods (except for some that aim to the saddle nature of the answer), based on the Hylleraas-Undheim and MacDonald (HUM) theorem [1], minimize the energy of the desired root of the secular equation, since in optimizing the "desired root," the other roots, all mutually orthogonal, get deteriorated. **However**, **a local minimum** (i.e., HUM) **cannot be simultaneously saddle point** *in the same* **restricted subspace**; therefore, the standard methods, finding a local energy minimum, do not find the saddle point sought, with a danger, if they use small truncated, supposedly easily comprehensible expansions, to get probably a deceiving answer. For this reason the standard methods resort to approaching the saddle point using huge expansions (rather inappropriate for large systems). We can use a proposed functional [2] that, even within small expansions, has minimum **at** the saddle point, successfully tested, up to now, by direct multidimensional minimization, which is hard to use and time-consuming. It will be much more beneficial to the scientific community, both theoretical and practical, if we try to transform it, so that it is solved, faster and safer, self-consistently, as we shall see below, but this has not been done yet. Presently we shall see the presentation of the functional along with some computational applications due to it, e.g., (i) immediately improving a ground state approximant if we know a better excited state saddle point approximant and (ii) avoiding "root flipping." Then, we shall see demonstrations, for excited He, using both Hylleraas coordinates and configuration interaction (CI) in standard coordinates, and for Li using CI.

13.2   Overview

In order to study electronic processes occurring via excited states, we must know not only the excitation energy but also the wave function of the excited state, which, as is well known, is a saddle point in the Hilbert space. The correct wave function

Fig. 13.1 E_n is a saddle point
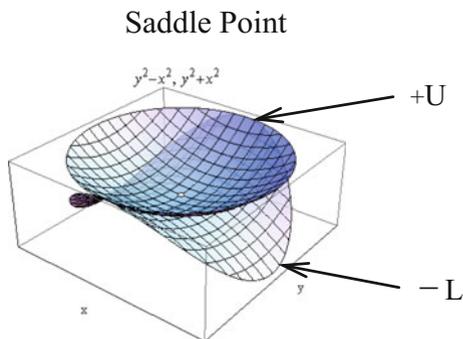
Saddle Point

is needed because if the computed wave function is away from the saddle point, incorrect conclusions may be deduced, concerning the physical process of charge transfer, i.e., of the main information needed both for the understanding of the electronic process and for the prediction of the path of the process, desired by any interested social-minded organization, private or public. This danger exists because we expand the wave functions in truncated bases, mimicking an ideal expansion in the complete orthonormal basis of the unknown exact Hamiltonian eigenfunctions, $\psi_0, \psi_1, \dots$ ($|0\rangle, |1\rangle, \dots$) with energies $E_0 < E_1 < \dots$ (assumed in the present analysis normalized real and non-degenerate). That is, if these were known, we could expand in terms of them any normalized wave function, in particular an approximant $|\phi_n\rangle$ of the nth excited state, as

$$|\phi_n\rangle = \sum_{i \neq n} |i\rangle \, \langle i|\phi_n\rangle + |n\rangle \sqrt{1 - \sum_{i \neq n} \langle i|\phi_n\rangle^2},$$

where the expansion coefficients $\langle i| \phi_n\rangle$ would be small. Then the energy would be

$$\langle \phi_n | H \, |\phi_n\rangle = E\,[\phi_n] = E = E_n - L + U \qquad (13.1)$$

where the lower term, L, and the higher term, U, would be

$$L = \sum_{i<n} (E_n - E_i) \, \langle i|\phi_n\rangle^2 > 0, \quad U = \sum_{i>n} (E_i - E_n) \, \langle i|\phi_n\rangle^2 > 0, \qquad (13.2)$$

which, in the Hilbert space (of the wave functions), are parabolas, the L downward and the U upward. In other words, the unknown sought exact eigenfunction, ψ_n, i.e., the stationary point of the parabolas, is a saddle point with n downward parabolas as indicated pictorially in Fig. 13.1.

Thus, the ground state, ψ_0 ($n = 0$), where the term L (with $n = 0$) does not exist, can be computed by minimizing the energy, $E = E_0 + U$, of a trial normalized approximant ϕ_0. And also, if the term L (with $n \neq 0$) were artificially absent, the excited states, ψ_n, could be computed similarly by minimizing the energy of

a normalized trial approximant ϕ_n, by orthogonalizing the trial function ϕ_n to all lower **exact** ψ_i, $i < n$ (if they were known), that would nullify all coefficients in L, $\langle i| \phi_n \rangle = 0$, and would leave only the remaining term U to be minimized.

However, when, in the absence of the **exact** eigenfunctions, ψ_n, we have only normalized **approximants** ϕ_n (expanded in truncated bases, e.g., Gaussians), we cannot use orthogonalization to a **known** ϕ_0, approximant of ψ_0, not even to approach the 1st excited state ψ_1: Because, if (most probably) the – known – normalized approximant ϕ_0 is not exactly orthogonal to the – unknown – ψ_1, the orthogonal to ϕ_0 subspace (call this subspace $\{\backslash \Phi_0\}$) *does not contain* ψ_1 (see Fig. 13.2). Therefore, ψ_1 cannot be found orthogonally to the normalized **approximant** ϕ_0. So, the best achievement of a normalized trial ϕ_1 varied orthogonally to ϕ_0 (i.e. belonging to the subspace $\{\backslash \Phi_0\}$) is to be *closest* to ψ_1 in the subspace $\{\backslash \Phi_0\}$, i.e. with the largest overlap $\langle 1| \phi_1 \rangle$. At the *closest,* here called ϕ_1^+, it has no other components out of the 2D space of $\{\phi_0, \psi_1\}$ (because any other components would diminish the largest overlap $\langle 1| \phi_1 \rangle$), i.e. ϕ_1^+ is the Gram-Schmidt orthogonal to ϕ_0 in the 2D space of $\{\phi_0, \psi_1\}$:

$$|\phi_1^+\rangle = \frac{|1\rangle - |\phi_0\rangle \langle \phi_0|1\rangle}{\sqrt{1 - \langle 1|\phi_0\rangle^2}}; \quad \text{with} \;\; E\left[\phi_1^+\right] = E_1 - \frac{E_1 - E\left[\phi_0\right]}{1 - \langle 1|\phi_0\rangle^2} \langle 1|\phi_0\rangle^2. \quad (13.3)$$

Observe, however, that (if ϕ_0 is a reasonable approximant of ψ_0 with $E[\phi_0] < E_1$) the last (subtracted) energy term in Eq. (13.3), is positive, and ϕ_1^+, the *closest to the saddle point* ψ_1 in the orthogonal to ϕ_0 subspace, has **lower** energy than the exact E_1, without being any stationary point. Therefore, the minimization of the energy of ϕ_1 orthogonally to the known normalized **approximant** ϕ_0 will end up even lower, departing even more from the saddle point (of the sought excited ψ_1), farther than ϕ_1^+. The final (converged in minimization) function, being orthogonal to a reasonable ground state approximant, will not be collapsed to the ground state, but it will simply be veered away from the saddle point ψ_1. This problem has been known since the early application tries of quantum mechanics.

However, fortunately (for the energy – but "unfortunately" for the wave function) the theorem of Hylleraas-Undheim and MacDonald (HUM) [1] ensures that the $(n + 1)^{\text{th}}$ root (eigenvalue) of the secular equation of the expansion coefficients in *truncated* basis, when varied, cannot take values below the **exact** (unknown) Hamiltonian eigenvalue, i.e., it has a **minimum** at (or just above) the exact value E_n. Thus, almost all computational methods of excited states (except some that aim to the saddle nature of the excited states, cf. [3–9]) are based on the HUM theorem, cf. [3–5, 10–65], achieving correct energies by **minimizing** the energy of the desired root. But **the finally converged point, being a *minimum, cannot be simultaneously a saddle point* in the same, restricted, subspace**; thus, only when the expansions are huge can the wave functions approach the exact excited state saddle point eigenfunctions (the larger, the closer to the exact). If they are truncated, not huge, "unfortunately," the HUM wave functions (themselves – not their energies) *avoid the saddle points* of the exact and unknown eigenfunctions, ψ_n, many times deceivingly.
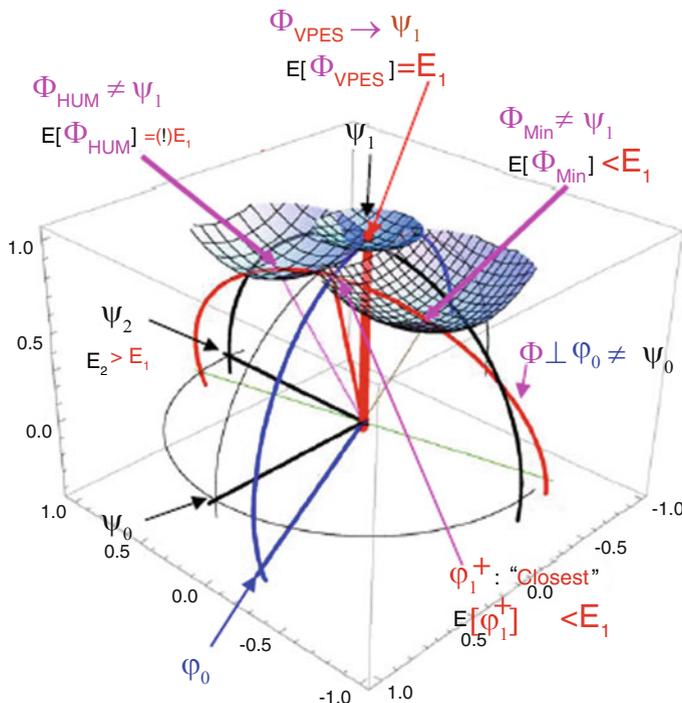
Fig. 13.2 Schematic representation of states. All states are assumed normalized: ψ_0, ψ_1, ψ_2, ... with energies $E_0 < E_1 < E_2 < ...$ are the unknown exact eigenstates; φ_0 is a known approximant of ψ_0. The subspace $S = \{\Phi, \varphi_1^+\}$ (red circle orthogonal to the blue circle of $\{\varphi_0, \psi_1\}$) is orthogonal to φ_0, and if φ_0 is not (accidentally) orthogonal to the unknown ψ_1, the subspace $S = \{\Phi, \varphi_1^+\}$ does not contain ψ_1. In the subspace $S = \{\Phi, \varphi_1^+\}$, the closest approximant to ψ_1 is φ_1^+ (the trace of the red on the blue cycle), and, as explained in the text, φ_1^+ lies below ψ_1: $E\left[\varphi_1^+\right] < E_1$. In going, while in S, orthogonally to φ_0, from φ_1^+ toward a state near ψ_2 (the green vector near ψ_2), i.e., in going, in S, from $E\left[\varphi_1^+\right] < E_1$ toward $E \approx E_2 > E_1$, one passes from E_1, i.e., from states, φ_1, of S, orthogonal to φ_0, but having energy $E[\varphi_1] = E_1$. If, in optimizing φ_1 by HUM theorem, φ_0 is the lowest (deteriorated, as explained in the text) root of the secular equation, then the 2nd "root," $\varphi_1 = \Phi_{HUM}$, is one of these states, "φ_1," with lowest possible energy $E[\varphi_1] = E[\Phi_{HUM}] = E_1$. But it is not ψ_1. It might be desirable to continue optimization in S toward, at least, φ_1^+, the closest, in S, to ψ_1. But HUM theorem forbids such a continuation, since the 2nd root must always be higher than E_1. In an attempt to approach, as much as possible, ψ_1, one might try, by other means, i.e., by direct minimization, to minimize the energy, in S, orthogonally to φ_0, toward φ_1^+. But φ_1^+ is not a critical point, and the minimum in S, orthogonal to φ_0, lies even lower: $E[\Phi_{Min}] < E_1$. Φ_{Min} does not suffer from variational collapse, since it is orthogonal to φ_0. Φ_{Min} is not a "bad" approximant of ψ_0: it is an approximant of ψ_1, probably as good (or as bad) as Φ_{HUM}. Both Φ_{HUM} and Φ_{Min} are veered away from φ_1^+, in S, and, therefore, from ψ_1. On the other hand, the here reported "variational principle for excited states" F_n (VPES) approaches ψ_1 (in general the exact excited states ψ_n), $\Phi_{VPES} \to \psi_1$, independently of the orthogonality to φ_0 (to lower lying approximants), and regardless of the accuracy of the latter, i.e., of their closeness to the exact saddle point, provided that the lower approximants, used in VPES F_n, are reasonable approximants, as explained in the text

Why "unfortunately"? Because by optimizing one root of the secular equation, all others get deteriorated: For example, the expansion of ψ_0 in the *truncated* basis of the roots – all mutually orthogonal to each other – $\Phi_0, \Phi_1, ..., \Phi_N$, leaves an unknown remainder Y; therefore, even if we optimize *at will* one of them, say Φ_1, $\langle \psi_1 | \Phi_1 \rangle^2 \to 1$, we shall have

$$\langle \psi_0 | \Phi_0 \rangle^2 + \langle \psi_0 | \Phi_1 \rangle^2 + \langle \psi_0 | \Phi_2 \rangle^2 + \cdots + \langle \psi_0 | \Phi_N \rangle^2 \leq 1 - Y^2$$
$$\Rightarrow \langle \psi_0 | \Phi_0 \rangle^2 < 1 - \langle \psi_0 | \Phi_2 \rangle^2 - \cdots - \langle \psi_0 | \Phi_N \rangle^2 < 1$$

i.e., Φ_0 will be deteriorated, *without being able to approach* ψ_0 *at will*, and the optimized Φ_1 will be orthogonal to a *deteriorated* Φ_0. So, if the expansion is large, then it tends to the exact eigenfunction, but if the expansion is small (necessarily for large systems, like molecules, catalysts, etc.), then, in a *truncated* basis, it is worse than just always being "$\Phi_1 \neq \psi_1$." For example, in a subspace of mainly two configurations, if the *deteriorated* normalized function is $\Phi_0 = \alpha \psi_0 + \beta \psi_1$ (+corrections), where β is not negligible, $\alpha < \sqrt{1 - \beta^2}$, then the HUM-*optimized* (i.e., with $E[\Phi_1] > E_1$) $\Phi_1 = \alpha \psi_1 - \beta \psi_0$ (+corrections) *is equally deteriorated*, and the worse, it is not even close to Φ_1^+, which is *the closest to the saddle point* ψ_1, among all normalized functions orthogonal to Φ_0: $\Phi_1^+ = \left((1 - \beta^2) \psi_1 - \alpha \beta \psi_0 \right) / \sqrt{1 - \beta^2}$ (+corrections) with $E[\Phi_1^+] < E_1$, because the HUM theorem necessitates that the *optimized* root Φ_1 shall *not go below* E_1 (in order to try to approach Φ_1^+).

Therefore, the HUM-optimized root Φ_1 is much more away from the saddle point ψ_1, although its energy approaches E_1. (Since there are normalized functions Φ orthogonal to ϕ_0 with $E[\Phi] < E_1$, it will eventually lead to one of the infinitely many normalized functions having $E[\Phi] = E_1$. Indeed, doing the analysis from any function Ψ if we consider a normalized function $\phi_1^{\perp +}$ orthogonal to both ϕ_0 and ϕ_1^+ and diagonalize the Hamiltonian operator between $\phi_1^{\perp +}$ and ϕ_1^+, this will open their energy gap, giving eigenfunctions Ψ^- and Ψ^+ (both orthogonal to ϕ_0). Then the function

$$\Phi = \Psi^- \sqrt{\frac{E[\Psi^+] - E_1}{E[\Psi^+] - E[\Psi^-]}} \pm \Psi^+ \sqrt{\frac{E_1 - E[\Psi^-]}{E[\Psi^+] - E[\Psi^-]}}$$

is orthogonal to ϕ_0 and has *equal to E_1* energy $E[\Phi] = E_1$, **without being ψ_1**. Even worse, from any Ψ, consider a normalized function $\Psi_\perp$ orthogonal to both ψ_0 and ψ_1. Then the function

$$\Phi = \sqrt{\frac{E[\Psi_\perp] - E_1}{E[\Psi_\perp] - E_0}} \psi_0 + 0\psi_1 - \sqrt{\frac{E_1 - E_0}{E[\Psi_\perp] - E_0}} \Psi_\perp$$

has *equal to E_1* **energy:** $E[\Phi] = E_1$, but is **orthogonal, to ψ_1**.) Thus, the only rescue of HUM optimization of a higher root, in order to approach the excited state

saddle point, is to use huge expansions ($\alpha \to 1$) (impracticable for large systems). The above are explained pictorially in Fig. 13.2.

However, in order to correctly comprehend charge transfer that will enable us to guide the applied research toward the correct and not (as seen below) toward a deceiving direction, we must find the *saddle point*, ψ_n, and, as explained above, with *truncated bases, this cannot be approached by neither orthogonal optimization (OO) nor by HUM optimization.* On the contrary, by both of these methods (OO, HUM), with *truncated* bases, *the saddle point ψ_n, is avoided!* (At least, by HUM the correct energy is approached). **For this reason, a safe approach of the saddle point, ψ_n, is needed, using *truncated* bases, thus, allowing both an understanding and a correct prediction, of practical interest, about the behavior of the electrons concerning charge transfer.** Specifically:

Because, by HUM, the space spanned by a *small* basis is not completely exploitable, since, **necessarily**, normalized functions that are closer to the *saddle point ψ_1* – and in general to ψ_n (i.e., with lower energy than the HUM answer, toward Φ_n^+) – are **excluded** (**forbidden**) by HUM, we shall see a developed functional, F_n [2], which has *local minimum at the saddle points, ψ_n*, which uses lower-lying normalized approximants Φ_i, $i < n$, but, practically, **does not depend on Φ_i's accuracy** (provided only that the Φ_is are reasonable):

$$F_n [\Phi_0, \Phi_1, \ldots; \Phi_n] \equiv E [\Phi_n] + 2 \sum_{i<n} \frac{\langle \Phi_i | H - E [\Phi_n] | \Phi_n \rangle^2}{E [\Phi_n] - E [\Phi_i]} \left[1 - \sum_{i<n} \langle \Phi_i | \Phi_n \rangle^2 \right]^{-1}.$$
(13.4)

As shown below, F_n has a *minimum* at each excited energy saddle point, even in pathological cases where at the excited eigenvalue there is no definite Hessian (to count its negative eigenvalues). As a "difficult" check, F_n easily passes the pathological Rellich test [5] of the Hermitian matrix:

$$\begin{pmatrix} -\sin x & \sin y \\ \sin y & \sin x \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix} = E \begin{pmatrix} X \\ Y \end{pmatrix} \Rightarrow E_{0,1} = \pm \sqrt{\sin^2 x + \sin^2 y}.$$

The "excited" E_1 has minimum at $x = 0$, $y = 0$ (as parameters), with indefinite Hessian there. The eigenvectors ψ_0, ψ_1 are easily computed:

$$\frac{Y_{0,1}}{X_{0,1}} = \frac{\pm \sqrt{\sin^2 x + \sin^2 y} + \sin x}{\sin y}, \quad \psi_{0,1}(x, y) = \begin{pmatrix} X_{0,1} \\ Y_{0,1} \end{pmatrix} / \sqrt{X_{0,1}^2 + Y_{0,1}^2}.$$

If we choose in F_1 (i.e., for $n = 1$), as fixed function, a **crude but reasonable** "ground state" expansion, in terms of $\{\psi_0, \psi_1\}$, namely, $\Phi_0 = 0.1 \ \psi_1 + \psi_0 \sqrt{1 - 0.1^2}$ (where 0.1 is "close" to 0, but, obviously, not too close) and if, by using this fixed function Φ_0, we minimize F_1 by optimizing $\Phi_1 = a \ \psi_1 +$
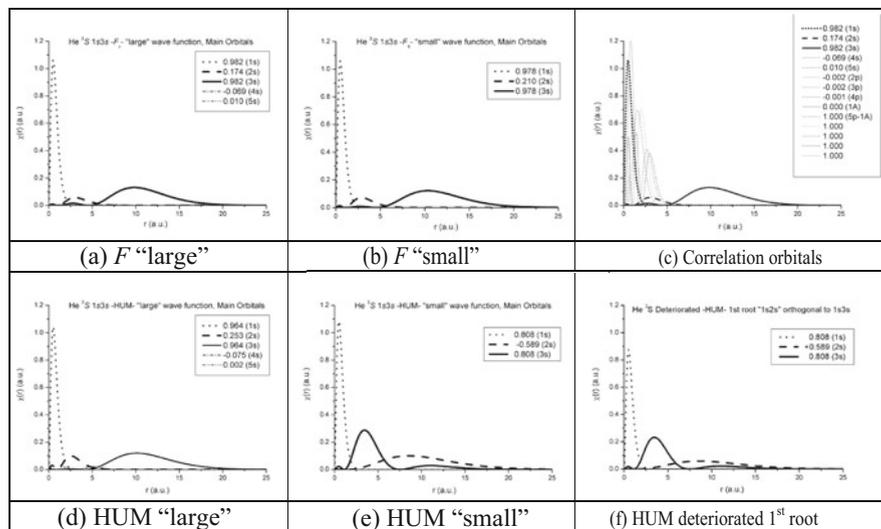
Fig. 13.3 CI wave functions (main orbitals): for $\mathbf{He^3S1s3s}$. Dotted: $1s$. Dashed: $2s$. Solid: $3s$. Clearly, the "large" expansions, either "F" (**a**) or HUM (**d**), are equivalent and have the same main orbitals as the F "small" (**b**), i.e., $1s3s$, and the $2s$ just serves for dynamic correlation. However, the HUM "small" (**e**) is, incorrectly, mainly $1s2s$, and the $3s$ "tries to improve" the wave function (and the energy) as static correlation, but by no means can $3s$ be taken as "LUMO" (and $2s$ as "HOMO") orbital. Note its orthogonal deteriorated HUM 1st "root" (**f**): they have the same main orbitals but with different signs. Note also that the "large" expansion needs static correlation corrections (from other CI terms) (**c**), to "split" electrons in areas where they are compacted together in the core. The diffuse orbital 3s essentially is not affected

$\psi_0\sqrt{1-a^2}$ with respect to the parameters $\{a, x, y\}$ around $(a, x, y) \sim (1, 0, 0)$, the resulting minimum is easily obtained correctly at $x = 0, y = 0$ (for any a).

The above F_1 has been successfully tested for excited states of the atoms He and C [2, 66] using large basis sets, where it agrees with bibliography and with our HUM – *large*-bases – computations, as well as *using small bases* for He, where, for example, it gives correct excited wave functions $1s3s\,(+\text{corrections})^1S\,or^3S$, whereas, on the contrary, small-bases-HUM gives, incorrectly, $1s\mathbf{2}s\,(+\text{corrections})^1S\,or^3S$ [2]. See also Fig. 13.3, where it is shown that, in optimizing the 2nd HUM root, the 1st root, Φ_0, is indeed deteriorated: It is *neither* $1s^2$ *nor* $1s2s$, the nodes are at unexpected "unphysical" distances from the nucleus, and the energy deviates significantly from the ground state energy. Its orthogonal **optimized** 2nd HUM root consists of the same "main" orbitals but with different signs, i.e., mainly $(\beta, -\alpha)$, although its energy **approaches** E_1, but remains above it, $E[\Phi] > E_1$, in accordance with the HUM theorem. It is not allowed to approach even Φ_1^+, *the closest to the saddle point ψ_1 in the space orthogonal to the *deteriorated* Φ_0 (because Φ_1^+ has energy $E\left[\Phi_1^+\right] < E_1$). Thus, the optimized 2nd HUM root **avoids** *the exact saddle point eigenfunction* $\psi_1 = 1s3s$. This is clearly shown in Fig. 13.4, where the functional F_n [2], with a small basis, finds the
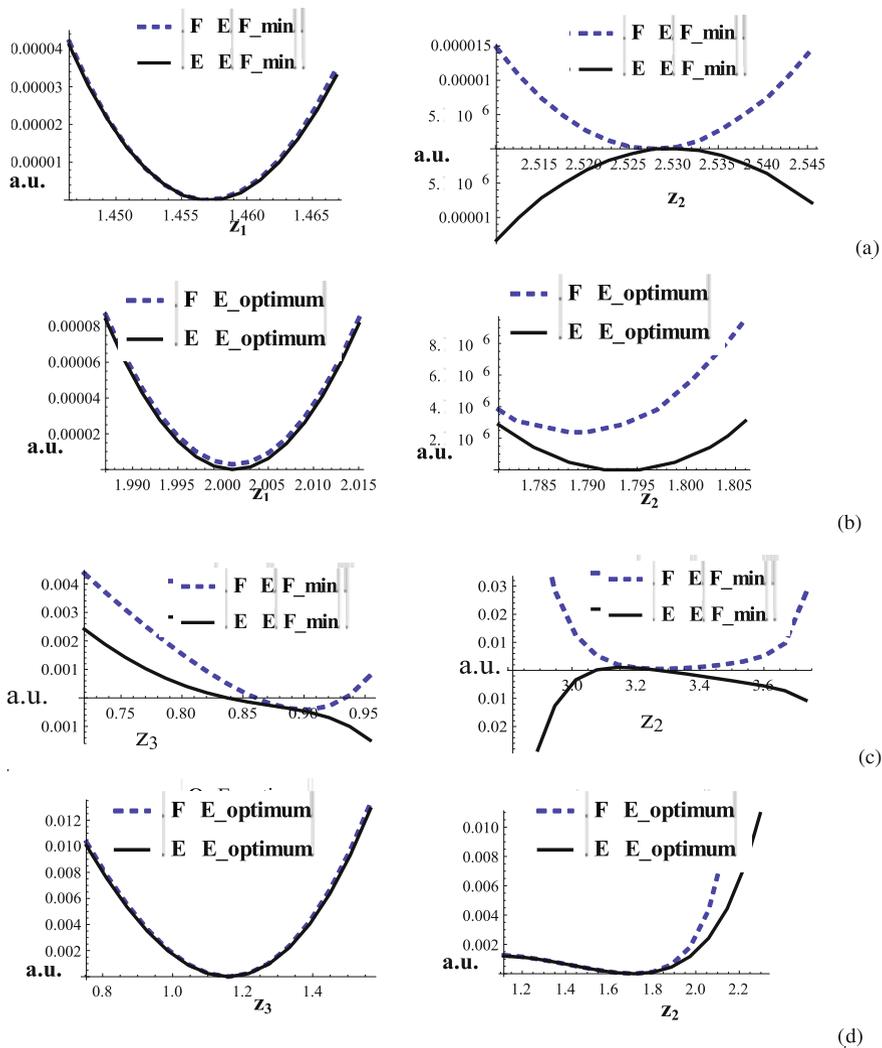
Fig. 13.4 He^{3}S 1s3s. (**a**) F_1 "small" expansion: Saddle point of energy at the F minimum. (**b**) HUM "small" expansion: Although the optimized 2nd root has "perfect" energy, it is not a saddle point and, of course, does not correspond to an F minimum. (**c**) F_1 "large" expansion: Saddle point of energy at the F minimum. (**d**) HUM "large" expansion: Although the optimized 2nd root has more "perfect" energy, it is not a saddle point

excited state saddle point, whereas the HUM optimization, with a small basis, does not find it, while, with a large basis, although HUM finds a minimum – as expected, not the saddle point – it approaches it more reasonably (the larger the basis, the better the approximation).

From Fig. 13.3 it is shown that the HUM optimization tries to correct the 2nd root as much as possible, via other orbitals (in the same small basis) as static (just descriptive) correlation corrections. Thus, although the **energy** may approach the correct eigenvalue E_1, the normalized **wave function** remains orthogonal to the *deteriorated* Φ_0, and far from Φ_1^+, and much farther from the saddle point ψ_1.

Similar results are found for Li ^{2}S and ^{4}S low-lying excited states (cf. Table 13.1).

Since the optimization of higher HUM roots fails for the *smallest atoms* He and Li, it does not provide any guarantee for larger systems, where, most probably, small basis sets are needed.

Presently, the direct minimization of the functional F_n has been tested by varying the parameters of the trial normalized expansion Φ_n (while the lower normalized approximants $\Phi_{i<n}$ must be kept unvaried), but this procedure is hard to use and is time-consuming. It is desirable, by nullifying (setting equal to 0) the derivatives of F_n with respect to the parameters, to reduce the minimization to an eigenvalue problem, so that the lowest eigenvalue would be used. This would be immediately accomplishable if the differentiation yielded a set of linear equations.

However, because of the form of F_n in the denominator, the Rayleigh-Ritz quotient yields a nonlinear system of equations of higher degree. Thus, there should be a plan to solve the problem self-consistently, i.e., to consider Φ_n in certain locations in the formula of the Rayleigh-Ritz quotient, except one location, as constant (initial guess), so as, by nullifying the (varied) derivatives, to use the lowest root as a new guess Φ_n in the other (artificially un-varied) locations, hoping to converge. A choice that has not been yet exhaustively tested, but seems to work efficiently, is to consider as initial guess (*ig*) all locations of the 2nd term of the F_n formula, except inside the quadratic term (because this term always converges quadratically), i.e., specifically,

$$F_n\left[\Phi_{ig}, \Phi_n\right] \equiv E\left[\Phi_n\right] + 2\sum_{i<n} \frac{\langle\Phi_i|\,H - E\left[\Phi_{ig}\right]|\Phi_n\rangle^2}{E\left[\Phi_{ig}\right] - E\left[\Phi_i\right]}\left[1 - \sum_{i<n}\left\langle\Phi_i|\Phi_{ig}\right\rangle^2\right]^{-1}$$

which is of the form

$$F_n\left[\Phi_{ig}, \Phi_n\right] \equiv \frac{\langle\Phi_n|\,H\,|\Phi_n\rangle}{\langle\Phi_n|\Phi_n\rangle} + 2Z\left[\Phi_{ig}\right]\sum_{i<n}\frac{\langle\Phi_i|\,H - E\left[\Phi_{ig}\right]|\Phi_n/\sqrt{\langle\Phi_n|\Phi_n\rangle}\rangle^2}{E\left[\Phi_{ig}\right] - E\left[\Phi_i\right]},$$

where Φ_n is yet unnormalized, and which reduces to a standard generalized eigenvalue problem:

$$F_n\left[\Phi_{ig}, \Phi_n\right]\langle\Phi_n|\Phi_n\rangle \equiv \langle\Phi_n|\,H\,|\Phi_n\rangle + 2Z\left[\Phi_{ig}\right]\sum_{i<n}\frac{\langle\Phi_i|\,H - E\left[\Phi_{ig}\right]|\Phi_n\rangle^2}{E\left[\Phi_{ig}\right] - E\left[\Phi_i\right]}.$$

Table 13.1 The CI expansion of various excited states of Li ^{2}S and ^{4}S. The nth optimized HUM root is indicated by n**r**; the nth excited state saddle point, local minimum of F_n, is indicated by f**n**; <r> is the RMS extent of the indicated orbital; <f|r>/NCI is the overlap of the corresponding wave functions consisting of NCI configurations. A 2nd derivative $E[\Phi]_{z_i z_i}$ (<0 for the saddle point, >0 for the HUM root **minimum**) is shown with respect to the indicated variation parameter in the eighth column

Φ	Main configurations	$F_n[\Phi]$ (a.u.)	$E[\Phi]$ (a.u.)	<r> (a.u.)	<r> (a.u.)	<r> (a.u.)	$E[\Phi]_{z_i z_i}$ [a]	|<f|r>|/NCI
Li^{2}S								
f0 1s^{2}2s(= **1r**)	0.996 1s^{2}2s − 0.053 1s^{2}3s	−7.45	−7.45	1s: 0.69	2s: 4.13	3s: 1.29		1/18
2r 1s^{2}3s	0.838 1s^2**2s** − 0.545 1s^{2}3s	−7.295	−7.296	1s: 0.64	3s: 4.93	2s: 11.3		b
f1 1s^{2}3s	0.935 1s^{2}3s − 0.354 1s3s2s	−7.31	−7.33	1s: 0.49	3s: 14.4	2s: 1.12	**z₃: −0.003**	0.89/8
2r 1s^{2}3s	0.964 1s^{2}3s − 0.264 1s^{2}2s	−7.298	−7.299	1s: 0.64	3s: 11.6	2s: 2.94	z₃: +0.002	c
3r 1s^{2}4s	0.820 1s^2**3s** − 0.44 1s^{2}4s -...	−7.25	−7.260	1s: 0.64	4s: 7.71	3s: 18.5		
f2 1s^{2}4s	0.974 1s^{2}4s − 0.222 1s4s2s	−7.285	−7.290	1s: 0.55	4s: 22.9	2s: 1.26	**z₄: −0.016**	0.94/20
3r 1s^{2}4s	0.979 1s^{2}4s + 0.188 1s^{2}2s	−7.264	−7.265	1s: 0.64	4s: 21.5	2s: 2.68	z₄: +0.020	
f3 1s^{2}5s	0.863 1s^{2}5s − 0.490 1s^{2}4s	−7.250	−7.250	1s: 0.64	5s: 32.6	4s: 28.5	**z₃: −0.0002**	0.99^d/40
4r 1s^{2}5s	0.981 1s^{2}5s + 0.165 1s^{2}2s	−7.248	−7.250	1s: 0.64	5s: 35.1	2s: 2.54	z₃: +0.0001	
Li^{4}S								
f0 1s2s3s(= **1r**)	0.993 1s2s3s + 0.112 1s2s4s	−5.204	−5.204	1s: 0,58	2s: 2.56	3s: 8.95		1
2r 1s2s4s	0.836 1s2s**3s**−0.547 1s2s4s	−5.153	−5.154	1s: 0,57	2s: 2.70	3s: 18.6		b
f1 1s2s4s	0.999 1s2s4s−0.030 1s4s5s	−5.147	−5.156	1s: 0,58	2s: 2.67	4s: 21.8	**z₂: −0.624**	0.90/10
2r 1s2s4s	0.961 1s2s4s + 0.273 1s2s3s	−5.154	−5.154	1s: 0,58	2s: 2.58	4s: 18.8	z₂: +0.193	c
3r 1s2s5s	0.580 1s2s3s−0.547 1s3s**4s**	−5.120	−5.123	1s: 0,57	2s: 5.17	3s: 39.5		
f2 1s2s5s	0.999 1s2s5s + 0.035 1s4s5s	−5.130	−5.136	1s: 0.58	2s: 2.66	5s: 39.3	**z₃: −0.009**	0.79/10
3r 1s2s5s	0.980 1s2s5s + 0.188 1s2s3s	−5.134	−5.134	1s: 0.58	2s: 2.55	5s: 29.9	z₃: +0.156	

[a] 2nd derivative $E[\Phi]_{z_i z_i}$ at the **optimal point** with respect to the parameter z_i, indicated in each case, to check whether it is **saddle** for f**n**, or **minimum** for n**r** (indicating that the HUM root, n**r**, is **not the saddle point**)

[b] Initially we found this "optimized" 2nd root (for ^{2}S: 1s^2**2s** instead of the expected 1s^{2}3s − and similarly for the ^{4}S). These parameters are more appropriate to the ground state basis functions. Starting from this we minimized F_1 that changed the basis functions more appropriate to the excited state, getting f1; then, starting from f1 we found a better 2nd root, 2r, more appropriate to the excited state basis functions, which we used for the comparisons

[c] Similarly for the 3rd root (for ^{2}S:1s^2**3s** instead of the expected 1s^{2}4s − and similarly for the ^{4}S)

[d] The CI expansion is already large

After normalization, the normalized Φ_n is used as a new $\Phi_{ig} \leftarrow \Phi_n$, until convergence. Note that, this way, we need only the extra computation of $\langle \Phi_i | \Phi_n \rangle$ and $\langle \Phi_i | H | \Phi_n \rangle$, which is standard feature in most standard electronic structure codes, with Φ_n expanded in the chosen (presumably small) basis.

In other words, this would extend the known MCSCF (multi-configuration self-consistent field) methodology with a Rayleigh-Ritz quotient of not only the energy (of the 1st term of F_n) under the normalizing condition of the trial function but of the whole functional F_n. Note that now, having used the $\Phi_{i < n}$ lower, rather crude, approximants, the desired root of the resulting secular equation is the *lowest* root above the (n^{th} HUM root $- \delta E$) where δE should be a reasonably small quantity. (This precaution should be taken especially for the initial guess.)

Since the above self-consistency is presently still being tested, in the following demonstrations, for the linear part of the process, the optimization of the expansion coefficients minimizing F_n is performed on the 1st term of F_n while checking whether the 2nd term vanishes (being essentially Schrodinger's equation "dotted" on the unvaried $\Phi_{i < n}$). Thus, the lowest (not the nth) HUM root, if it is above all lower-lying (crude) energies and if it does not nullify the 2nd term and is not a saddle point, is rejected. But, in this way (by checking every time), the correct local minimum may be missed, that is why a targeted but self-consistent nullification of the derivatives of the Rayleigh-Ritz quotient from F_n itself is rather needed.

If the method succeeds for larger atoms, then it could be extended to molecules and larger systems using CI expansions in standard Gaussian bases, so that it may be proven useful for large systems demanding small bases, appropriate for both time saving and easy understanding of the (small and comprehensible) wave functions. These, however, are future plans; here we shall see a presentation of F_n and some demonstrations of its applications.

13.3 The Construction of F_n

The central idea is to invert the sign of L in Eq. 13.1, in order to have all parabolas upward and then to make a continuation from the unknown exact set of $\{\psi_i\}$ to known approximants $\phi_{i < n}$.

Consider the energy (Eq. 13.1) of a normalized approximant ϕ_n of the nth excited state, expanded in the exact eigenfunctions $\{\psi_i\}$

$$E[\phi_n] = E_n - L + U,$$

where L and U are the downward and upward paraboloids defined in Eq. 13.2.

Invert the sign of L to introduce the functional

$$F[\phi_n] = E_n + L + U;$$

now all parabolas are upward and $F[\phi_n]$ has a local minimum at ψ_n: $F[\phi_n] = E_n$.

Substitute from Eq. 13.1 the unknown quantities $E_n + U = E[\phi_n] + L$:

$$F[\phi_n] = E[\phi_n] + 2L; \tag{13.5}$$

now $F[\phi_n]$ is expressed in terms of only the lower exact eigenfunctions $\psi_{i<n}$ in paraboloid L, inverted upward. But all $\psi_{i<n}$ are unknown; suppose, however, that we already have computed normalized approximants $\phi_{i<n}$ for all $\psi_{i<n}$. Then make a continuation of $F[\phi_n]$ from $\psi_{i<n}$ to $\phi_{i<n}$.

Expand each $\phi_{i<n}$ in the $\{\psi_i\}$ basis, and obtain to leading order in coefficients

$$\begin{aligned}
\langle\phi_i|\phi_n\rangle &= \langle i|\phi_n\rangle + \langle n|\phi_i\rangle + \cdots \\
\langle\phi_i|H|\phi_n\rangle &= E_i \langle i|\phi_n\rangle + E_n \langle n|\phi_i\rangle + \cdots .
\end{aligned} \tag{13.6}$$

Solve Eq. 13.6 for $\langle i|\phi_n\rangle$, and substitute to each term of $L = \sum_{i<n}(E_n - E_i)\langle i|\phi_n\rangle^2$ in Eq. 13.5 to get, to leading order, $L = \sum_{i<n}(E_n\langle\phi_i|\phi_n\rangle - \langle\phi_i|H|\phi_n\rangle)^2/(E_n - E_i)$, which suggests an examination, in terms the known *approximants* $\phi_{i<n}$, of the expression

$$S \equiv \sum_{i<n} \frac{(E[\phi_n]\langle\phi_i|\phi_n\rangle - \langle\phi_i|H|\phi_n\rangle)^2}{E[\phi_n] - E[\phi_i]} . \tag{13.7}$$

When both $\phi_i = \psi_i$ and [in Eq. (13.1)] $U = \sum_{i>n}(E_i - E_n)\langle i|\phi_n\rangle^2 \to 0$, the quantity S of Eq. 13.7 reduces, as directly verified, to

$$S \equiv L\left(1 - \sum_{i<n}\langle\phi_i|\phi_n\rangle^2\right) \Rightarrow L = \frac{S}{1 - \sum_{i<n}\langle\phi_i|\phi_n\rangle^2}.$$

Then, for $U \neq 0$ and for ϕ_n close to ψ_n, the functional $F[\phi_n]$ in Eq. 13.5 adequately behaves as

$$F_n[\phi_0, \phi_1, \ldots; \phi_n] \equiv E[\phi_n] + 2\sum_{i<n}\frac{\langle\phi_i|H - E[\phi_n]|\phi_n\rangle^2}{E[\phi_n] - E[\phi_i]}\left[1 - \sum_{i<n}\langle\phi_i|\phi_n\rangle^2\right]^{-1},$$

where the adequacy depends on the Hessian determinant, A_n^n, and its principal minors along the main diagonal, $A_n^{k<n}$, at $\phi_n = \psi_n$. According to the standard theorems of calculus (cf. Sylvester theorem), if ψ_n is still a critical point and $A_n^{k<n} > 0$, $A_n^n > 0$, then $F[\phi_n]$ has a local minimum at $\phi_n = \psi_n$.

Indeed, at $\phi_n = \psi_n$, $E[\phi_n] = E_n$, and the 2nd term vanishes because the operator becomes $(H - E[\phi_n])\phi_n = (H - E_n)\psi_n = 0$, (i.e., Schrödinger's equation) so that

$$F_n\left[\phi_0, \phi_1, \ldots; \psi_n\right] = E_n,$$

i.e., ψ_n *is* a critical point, the saddle point, ***regardless* of the quality of** $\phi_{i<n}$.

Incidentally, notice that, since E_n is a saddle point, the energy $E[\phi_n]$ of an approximant ϕ_n **is not necessarily an upper bound of E_n**; it must only be **an upper bound of $(E_n - L)$**: Since $E[\phi_n] + L - E_n = U > 0$, then $E_n - L < E[\phi_n]$, or

$$E_n - \sum_{i<n} (E_n - E_i)\, \langle i|\phi_n\rangle^2 < E\left[\phi_n\right], \tag{13.8}$$

i.e., it may be *below E_n* (that is why the aforementioned "orthogonal" minimization (OO) does not "collapse"). This means that since, in practice, with a truncated expansion, the minimization procedure of F_n, within a tolerance, stops around (but not exactly at) ψ_n, if it stops at a side of upward parabola, we will have $E_n < E[\phi_n] < F_n$, but if it stops at a side of downward parabola, we will have $E_n - L < E[\phi_n] < E_n < F_n$. In this case, the final (converged) ϕ_n, although slightly below $E[\mathrm{HUM}]$ (by $\leq L$), will be certainly acceptable. The value of L could be estimated using accurate approximants $\phi_{i<n}$ and checked to be of the order of the tolerance.

Further, at $\phi_n = \psi_n$, if we denote by $\left|\phi_i^{\perp\{n\}}\right\rangle$ the projection of $|\phi_i\rangle$ on the subspace of the higher than-n eigenfunctions, the Hessian determinant is

$$A_n^n = 2^{n+1} \prod_{i=0}^{n-1} (E_n - E_i)\left(E\left[\phi_n^{\perp\{n\}}\right] - E_n\right) \times$$

$$\times \left\{ 1 + 2 \left[\begin{array}{l} n \sum_{i=0}^{n-1} \langle n|\phi_i\rangle^2 + \\[2mm] + \sum_{i=0}^{n-1} \frac{\left(E\left[\phi_n^{\perp\{n\}}\right] - E_n\right)\left(E\left[\phi_i^{\perp\{n\}}\right] - E_n\right) - \left(\left\langle \phi_i^{\perp\{n\}}\middle| H - E_n \middle|\phi_n^{\perp\{n\}}\right\rangle\right)^2}{(E_n - E_i)\left(E\left[\phi_i^{\perp\{n\}}\right] - E_n\right)}\left\langle \phi_i^{\perp\{n\}}|\phi_i\right\rangle^2 \\[3mm] - 2\sum_{i=0}^{n-1}\sum_{j=i+1}^{n-1}\frac{\left((E_n - E_j)\langle j|\phi_i\rangle + (E_n - E_i)\langle i|\phi_j\rangle\right)^2}{(E_n - E_i)(E_n - E_j)} + O[\text{coefficients}]^3 \end{array} \right] \right\}, \tag{13.9}$$

and its principal minors are

$$A_n^{k<n} = 2^{k+1} \prod_{i=0}^{k} (E_n - E_i) \times$$

$$\times \left\{ 1 + 2 \left[\begin{array}{l} (k+1)\sum_{i=0}^{n-1}\langle n|\phi_i\rangle^2 + \sum_{j=k+1}^{n-1}\sum_{i=0}^{k}\frac{E_n - E_i}{E_n - E_j}\langle i|\phi_j\rangle^2 + \sum_{i=0}^{k}\frac{E\left[\phi_i^{\perp\{n\}}\right] - E_n}{E_n - E_i}\left\langle \phi_i^{\perp\{n\}}|\phi_i\right\rangle^2 \\[3mm] - \sum_{j=k+1}^{n-1}\sum_{i=0}^{k}\frac{E_n - E_j}{E_n - E_i}\langle j|\phi_i\rangle^2 - 2\sum_{i=0}^{k}\sum_{j=i+1}^{k}\frac{\left((E_n - E_j)\langle j|\phi_i\rangle + (E_n - E_i)\langle i|\phi_j\rangle\right)^2}{(E_n - E_i)(E_n - E_j)} \end{array} \right] \right\}$$
$$+ O[\text{coefficients}]^3. \tag{13.10}$$

(For clarity in the above expressions, care has been taken to be expressed in terms of positive quantities.) In these expressions, if the lower approximants $\phi_{i<n}$ are accurate enough, the coefficients inside the "2[square brackets]" are normally small (squared: even smaller), and the factor of "1," before them, normally dominates over them, so that $A_n^{k<n} > 0$, $A_n^n > 0$. Therefore, if the lower approximants $\phi_{i<n}$, **without being very accurate**, fulfill the conditions $A_n^{k<n} > 0$, $A_n^n > 0$, the functional $F[\phi_n]$ has a local minimum at $\phi_n = \psi_n$. (In practice, if these conditions are not fulfilled, $F[\phi_n]$ drops down to $-\infty$).

It is important to note that **no orthogonality to the lower approximants $\phi_{i<n}$ has been assumed**, whatsoever; it is not needed, the functional $F[\phi_n]$ has a local minimum at $\phi_n = \psi_n$ for any **reasonable** $\phi_{i<n}$ (that would satisfy Eqs. 13.9 and 13.10). Also, as mentioned before, it is not necessary to compute and diagonalize the Hessian and count its negative eigenvalues to identify the saddleness at $\phi_n = \psi_n$.

However, a simple criterion of being at the saddle point is to compute the 2nd derivative of the energy $E[\phi_n]$ with respect to the parameters: There should be n down-parabolas. Of course, the corresponding 2nd derivatives of F_n would be positive.

13.4 Improving a Ground State Approximant ϕ_0 Via an Accurate ϕ_1

The central idea is that: if our approximant ϕ_1 is more accurate than an approximant ϕ_0, then diagonalizing the Hamiltonian between the two, opens their energy gap and improves ϕ_0 more than ϕ_1 (since ϕ_1 is more accurate than ϕ_0). Then repeat the process by working in the space orthogonal to ϕ_1, by diagonalizing between the new ϕ_0 and any function orthogonal to both ϕ_1 and the new ϕ_0, and so on, until no further improvement.

13.4.1 *Improving ϕ_0 Orthogonally to the Exact ψ_1 (Analysis)*

First, if we had the exact ψ_1, then we could immediately improve an approximant ϕ_0 orthogonally to ψ_1 [2]: In the subspace of $\{\phi_0, \psi_1,\}$, the highest Hamiltonian eigenvector, Ψ^+, is ψ_1 itself:

$$\Psi^+ = \psi_1. \tag{13.11}$$

The lowest, Ψ^-, is orthogonal to ψ_1,

$$\Psi^- = \phi_0^+ \equiv \frac{\phi_0 - \psi_1 \langle \psi_1 | \phi_0 \rangle}{\sqrt{1 - \langle \psi_1 | \phi_0 \rangle^2}}, \tag{13.12}$$

with energy

$$E\left[\phi_0^+\right] = E\left[\phi_0\right] - \frac{\left(E\left[\psi_1\right] - E\left[\phi_0\right]\right)\langle\psi_1|\phi_0\rangle^2}{1 - \langle\psi_1|\phi_0\rangle^2} \leq E\left[\phi_0\right] \tag{13.13}$$

(i.e., Ψ^- is same or better than ϕ_0). Further, rotating ϕ_0^+ around ψ_1 would improve ϕ_0^+ as follows: Introduce (e.g., by one more configuration) a function $\phi_0^{(2+)}$ orthogonal to both $\{\phi_0^+, \psi_1\}$. Then, in the subspace of $\{\phi_0^+, \phi_0^{(2+)}\}$, (both orthogonal to ψ_1), the Hamiltonian opens their energy gap, so the lowest eigenvector $\Psi^- \equiv \phi_0^-$ has energy $E\left[\phi_0^-\right] \leq E\left[\phi_0^+\right]$, closer to $E[\psi_0]$ (in a three-dimensional function space $\{\psi_0, \psi_1, \psi_k\}$, this would be exactly $E[\psi_0]$ as directly verified). $E\left[\phi_0^-\right]$ could be further improved by further rotating around ψ_1 similarly: i.e., after introducing another function $\phi_0^{(3+)}$ orthogonal to both $\{\phi_0^-, \psi_1\}$ in the subspace of $\{\phi_0^-, \phi_0^{(3+)}\}$ (both orthogonal to ψ_1), the lowest Hamiltonian eigenvector is $\Psi^- \equiv \phi_0^{(2-)}$ with energy $E\left[\phi_0^{(2-)}\right] \leq E\left[\phi_0^-\right]$ (even closer to $E[\psi_0]$) and so on.

Now let us consider the approximants.

13.4.2   Improving ϕ_0 Orthogonally to ϕ_1

Since ψ_1 is never exactly known, we can still improve ϕ_0 orthogonally to ϕ_1, our **best** approximant of ψ_1 (obtained either via F_1 or via a "large" HUM expansion), by first computing ϕ_0^+ orthogonal to our ϕ_1,

$$\phi_0^+ \equiv \frac{\phi_0 - \phi_1\langle\phi_1|\phi_0\rangle}{\sqrt{1 - \langle\phi_1|\phi_0\rangle^2}}, \tag{13.14}$$

if the condition $E\left[\phi_0^+\right] \leq E\left[\phi_0\right]$ is attainable: Indeed, by expanding about ϕ_1, as directly verified, the condition $E\left[\phi_0^+\right] \leq E\left[\phi_0\right]$, to leading order, reads $(E_1 - E_0)\left(1 - \langle\psi_1|\phi_0\rangle^2\right) \geq \left(E\left[\phi_0^{\perp\{1\}}\right] - E_0\right)\left\langle\phi_0^{\perp\{1\}}|\phi_0\right\rangle^2$, which is not impossible. Here, as defined in Eqs. (13.9) and (13.10), $\phi_0^{\perp\{1\}}$ is the normalized function, orthogonal to both $\{\psi_0, \psi_1\}$, collecting all higher than ψ_1 terms (the last terms forming U) of Eqs. (13.1) and (13.2) for $\phi_n = \phi_0$. For ϕ_0, ϕ_1 very close to ψ_0, ψ_1, as directly verified by expanding about ϕ_0 as well, the condition is satisfied when $\langle\psi_0|\phi_1\rangle^2 \leq \langle\psi_1|\phi_0\rangle^2$, i.e., when ϕ_1 attains better orthogonality to ψ_0 than ϕ_0 to ψ_1, which is essentially the original assumption that our ϕ_1 is more accurate than our ϕ_0.

In fact, if our ϕ_1 is close to ψ_1, then a 2×2 diagonalization of the Hamiltonian between ϕ_1 and ϕ_0 (or ϕ_0^+) should give as a higher eigenvector $\phi_1 \approx \psi_1$ unaffected

and as a lower eigenvector ϕ_0^+ itself as improved ϕ_0. This, by just a 2×2 diagonalization, is another simple criterion of being at the saddle point.

Incidentally, all other (small) components are less relevant, so that the opposite procedure of optimizing ϕ_1 orthogonally to ϕ_0 can lead to ϕ_1^{MIN}, unpredictably *far from* ψ_1, even orthogonal to ψ_1, with *still* $E\left[\phi_1^{\text{MIN}}\right] \lesssim E\left[\psi_1\right]$, as shown (as analysis) in the following counterexample: Even in the subspace $\{\psi_0, \psi_1, \psi_2\}$ (supposedly known for the present counterexample), the orthonormal trial functions $\phi_0 = a\psi_0 + 0\psi_1 + b\psi_2$ and $\phi_1 = b\psi_0 - a\psi_2$ with $a = \sqrt{\frac{(E_1 - \varepsilon) - E_0}{E_2 - E_0}}$, $b = \sqrt{\frac{E_2 - (E_1 - \varepsilon)}{E_2 - E_0}}$, (small ε), where, by construction, ϕ_0 is orthogonal to ψ_1 and ϕ_1 is orthogonal to both ψ_1 and ϕ_0, have energies $\langle\phi_0|H|\phi_0\rangle = a^2 E_0 + b^2 E_2 =$, $E[\phi_0] = E_0 + E_2 - (E_1 - \varepsilon) > E_0$, and $E[\phi_1] = E_1 - \varepsilon$, while ϕ_0 may approximate ψ_0. (To make the counterexample numerical, consider He ^{1}S, for which it is known that, in a.u., $E_0 = -2.903$, $E_1 = -2.146$, $E_2 = -2.06$. In terms of $\{\psi_0 (=1s^2),$ $\psi_1(=1s2s), \psi_2(=1s3s)\}$ (supposedly known for the present counterexample), the function $\phi_0 = 0.9476 \ \psi_0 + 0.3194 \ \psi_2$ would have $E[\phi_0] = -2.832$, and the function $\phi_1 = 0.3194 \ \psi_0 - 0.9476 \ \psi_2$ *chosen orthogonal to both ψ_1 and ϕ_0* would have $E[\phi_1] = -2.146 = E_1$ (i.e., it would be a function with the **energy** of ψ_1, although **orthogonal** to ψ_1), so that, any function orthogonal to the same ϕ_0, between this ϕ_1 and $\phi_1^+ = \psi_1$ (if ϕ_0 is orthogonal to ψ_1, the "closest" to ψ_1 is ψ_1 itself), could be a minimization result, ϕ_1^{MIN}, with arbitrary $\langle\psi_1|\phi_1^{\text{MIN}}\rangle$ and with $E_1 - \varepsilon \leq E\left[\phi_1^{\text{MIN}}\right] \leq E_1$. In this counterexample, *using F_1* this danger would not exist: Minimization of F_1, for the same, not particularly accurate, unvaried ϕ_0 (0.3194 is not very small), by varying $\phi_1 = c \ \psi_0 + d \ \psi_2 + \psi_1\sqrt{1 - c^2 - d^2}$, yields: $c < \text{tol}(=10^{-8})$ and $d < \text{tol}$, i.e., $\phi_1 = \psi_1$, with $E[\phi_1] = -2.146$ [so that, from Eq. 13.14, $\phi_0^+ = \phi_0$]. Below we shall see an actual demonstration.

13.4.3 Further Improvement of ϕ_0

If $E\left[\phi_0^+\right] \leq E\left[\phi_0\right]$, then, by rotating around ϕ_1, as described above [after Eq. 13.13], since the Hamiltonian always opens the energy gap between mutually orthogonal functions (all orthogonal to ϕ_1), ϕ_0^+ can be further improved (until $\langle\psi_0|\phi_1\rangle^2 > \langle\psi_1|\phi_0\rangle^2$), by always computing the 2×2 eigenfunctions (both orthogonal to ϕ_1) and taking the lowest current eigenfunction Ψ^- as: $\phi_0^{(m-)} = \Psi^-$. (At any step, $\phi_0^{(m-)}$ could be used as a new ϕ_0 to improve ϕ_1 via F_1 of Eq. 13.4.) In the above example (using the above He values), rotating ϕ_0 around ϕ_1 immediately gives $\phi_0^{(1-)} = \Psi^- = \psi_0$ (and $\Psi^+ = \psi_2$). Below we shall see an actual demonstration, performing a few "rotating" cycles.

13.5   Identifying a Flipped Root Around an Avoided Crossing

When atoms approach each other, then at a certain nuclear separation for each pair of molecular electronic states (of the same symmetry), the Hamiltonian opens their energy gap, and an avoided crossing may occur, changing their order: In passing the crossing, the continuation of the ground state becomes excited, and the continuation of the excited state becomes lowest. At a **fixed** nuclear separation, **near** the crossing, the (avoided) exact and sought eigenstates ψ_0 and ψ_1 are **distinct**. However, their approximants ϕ_0 and ϕ_1 are computed by varying their (truncated) expansion parameters, and in the parameter space, also an avoided crossing occurs, because, for any set of their parameters, the Hamiltonian opens their energy gap. However, there is a significant difference between avoided crossing in the parameter space and avoided crossing in the nuclear separation space. In the nuclear separation space, the wave function in the united-atom limit is completely different than in the separated-atom limit. However, in the parameter space, if $\phi_1 \approx \psi_1$ is, say 1s2s, then, just beyond the crossing and away from the sought ψ_1, ϕ_1 is still 1s2s, but with slightly different values of the same parameters (i.e., of the exponents, of the expansion coefficients, etc.). Thus, during optimization, the new (last suggested by the optimization method) trial parameters do not provide any information as to whether they are beyond or before the crossing. And if, toward convergence, we consider as ϕ_1 the ϕ_0 (instead of ϕ_1, erroneously), just because, with the new (last suggested) trial parameters, ϕ_0 has flipped to be higher, i.e., beyond the crossing, then the computation gets lost. We should feed ϕ_1 **with the orbitals of** $\boldsymbol{\phi_1}$ at the new (last suggested) trial parameters beyond the crossing, despite their flipping [9, 61, 64, 65, 67–78], i.e., **in spite of the fact that** at the new (last suggested) trial parameters, $\boldsymbol{\phi_1}$ **is lower than** $\boldsymbol{\phi_0}$. Root flipping may be avoided by an appropriate representation of the excited state, [65] but this may not be known in advance. If we do not know where the crossing is, we cannot identify the flipped functions by looking at the values of their parameters, but we can, by checking their (individually computed) F_1 using an **unvaried** (at least crude but reasonable) lower approximant of ψ_0, say Φ_0 (presumably unambiguously known from nearby nuclear separations): We shall see that: independently of whether ϕ_1 is, the 2nd (correct) or the 1st (flipped) root, by computing the two functionals $F_1[\Phi_0;\phi_1]$ and $F_1[\Phi_0;\phi_0]$, we will have

$$F_1\left[\Phi_0; \phi_1\right] < F_1\left[\Phi_0; \phi_0\right]. \tag{13.15}$$

Indeed: Consider the two lowest states [see also Ref. 75]. The eigenfunctions of the secular equation, "roots," depend on variational parameters **p** (both linear and nonlinear) to be optimized, and say the **optimal** functions are $\phi_0(\mathbf{p}_0)$, $\phi_1(\mathbf{p}_1)$. For example, in 1-electron hydrogen-like S-states $\phi_0(r) \sim e^{-\alpha Zr}$, $\phi_1(r) \sim (1 - \gamma Zr/2)e^{-\beta Zr/2}$, $\mathbf{p} = (\alpha, \beta, \gamma)$, so that at their F_1 minimum $\mathbf{p}_0 = (1, 0, 0)$, $\mathbf{p}_1 = (0, 1, 1)$. At the specified nuclear positions, under examination, ϕ_0, ϕ_1 are continuous functions of their parameters; their characteristics are maintained at

the region of any value of $\mathbf{p}$. Of course, we are seeking the **optimal** positions, in the parameter space, i.e., $\mathbf{p}_0$ for ϕ_0 and $\mathbf{p}_1$ for ϕ_1, ***not*** the position of the crossing, which is not the optimal. Now, it is important to realize that if there is no other crossing closer to $\mathbf{p}_0$ and $\mathbf{p}_1$ (otherwise we would consider the closest, judged by the overlap $\langle \Phi_0 | \phi_0 \rangle$), the **optimal** functions ϕ_0 (lower), ϕ_1 (higher) are unflipped *both* at $\mathbf{p}_0$ (where $\phi_0(\mathbf{p}_0)$ is the lowest and all other roots (deteriorated) including $\phi_1(\mathbf{p}_0)$ are above it) **and** at $\mathbf{p}_1$ (where $\phi_1(\mathbf{p}_1)$ is excited, having the lower root(s) (deteriorated) below it, including $\phi_0(\mathbf{p}_1)$): Since $E_0 < E_1$, and at $\mathbf{p}_0$, $\phi_0(\mathbf{p}_0) \simeq \psi_0$ is the lowest lying, then every other root, including $\phi_1(\mathbf{p}_0)$, lies higher, i.e., at $\mathbf{p}_0$, $E_0 \approx E[\phi_0(\mathbf{p}_0)] < E[\phi_1(\mathbf{p}_0)]$; also, at $\mathbf{p}_1$, $\phi_1(\mathbf{p}_1) \simeq \psi_1$, $E_1 \approx E[\phi_1(\mathbf{p}_1)] > E_0$, and any lower-lying root must describe (well or badly) the only lower-lying exact ψ_0 (or, in general, one of the lower lying $\psi_{i<n}$), i.e., it is the deteriorated approximant $\phi_0(\mathbf{p}_1)$, so, also, $E_0 < E[\phi_0(\mathbf{p}_1)] < E[\phi_1(\mathbf{p}_1)] \approx E_1$; otherwise, if we had $E[\phi_1(\mathbf{p}_1)] \simeq \psi_1] < E[\phi_0(\mathbf{p}_1)]$, then $\phi_1(\mathbf{p}_1) \simeq \psi_1$ would not be excited, contrary to our assumption. Thus, the order $E[\phi_0(\mathbf{p})] < E[\phi_1(\mathbf{p})]$ holds at *both* ***optimal*** positions $\mathbf{p}_0$ (of ϕ_0) and $\mathbf{p}_1$ (of ϕ_1). Let us call the whole union of the $\mathbf{p}$-regions where $E[\phi_0(\mathbf{p})] < E[\phi_1(\mathbf{p})]$: "*in front of*" the crossing or "*before*" root flipping. Thus, both optimal positions $\mathbf{p}_0$ and $\mathbf{p}_1$ are located "*before*" root flipping. At the (closest) crossing (say at some value $\mathbf{p_c}$), since there is no other crossing in between, the characteristics of the roots $\phi_0(\mathbf{p}_0)$, $\phi_1(\mathbf{p}_1)$ in the region of $\mathbf{p_c}$ are maintained, i.e., they are continuous at $\mathbf{p_c}$. In the variational parameter $\mathbf{p}$-space, "root flipping" $E[\phi_0(\mathbf{p})] > E[\phi_1(\mathbf{p})]$ means that the parameters $\mathbf{p}$ are such that the continuation *behind* the crossing of the sought excited state ϕ_1 lies lower than the continuation of ϕ_0, *behind* the crossing, i.e., the **lower** "root" *behind* the crossing consists mainly, i.e., has, *behind* the crossing, mainly the characteristics, of the **sought excited** state wave function which we must detect and recognize, in order to extrapolate ***it*** (i.e., the "root" that is lower *behind* the crossing) by some method, e.g., by quasi-Newton, to regions *in front of* the crossing, close to $\mathbf{p}_1$ (where the unknown optimal $E[\phi_1(\mathbf{p}_1)]$ must occur – higher than $E[\phi_0(\mathbf{p}_1)]$), while the other root, that is higher beyond the crossing, must, beyond the crossing, mostly resemble the crude approximant $\Phi_0 \sim \phi_0(\mathbf{p}_0)$.

Near the crossing, let the indices "$-$"/"$+$" indicate "just before"/"just behind" the crossing, so that, while ϕ_0 and ϕ_1 are **continuous** near the crossing, i.e., $\phi_{0-} = \phi_{0+} = \phi_0$ and $\phi_{1-} = \phi_{1+} = \phi_1$, the higher (blindly taken) "2nd root," $\Phi_1^{"2r"}$, "just before" (unflipped) is $\Phi_{1-}^{"2r"} = \phi_{1-}$ and "just behind" (flipped) is $\Phi_{1+}^{"2r"} = \phi_{0+}$, whereas the lower (blindly taken) "1st root," $\Phi_0^{"1r"}$, "just before" (unflipped) is $\Phi_{0-}^{"1r"} = \phi_{0-}$ and "just behind" (flipped) is $\Phi_{0+}^{"1r"} = \phi_{1+}$. Now, near the crossing, by computing the two functionals $F_1[\Phi_0; \phi_1]$ and $F_1[\Phi_0; \phi_0]$, we have for the higher (blindly taken) "2nd root": $F_1\left[\Phi_0; \Phi_{1-}^{"2r"}\right] = F_1[\Phi_0; \phi_{1-}]$ and $F_1\left[\Phi_0; \Phi_{1+}^{"2r"}\right] = F_1[\Phi_0; \phi_{0+}]$, similarly for the lower (blindly taken) 1st root: $F_1\left[\Phi_0; \Phi_{0-}^{"1r"}\right] = F_1[\Phi_0; \phi_{0-}]$ and $F_1\left[\Phi_0; \Phi_{0+}^{"1r"}\right] = F_1[\Phi_0; \phi_{1+}]$. (These could be anyway recognizable from known $\mathbf{p}$-points a little away from the crossing, i.e., where unambiguously $F_1\left[\Phi_0; \Phi_1^{"2r"}\right] = F_1[\Phi_0; \phi_1]$ and $F_1\left[\Phi_0; \Phi_0^{"1r"}\right] = F_1[\Phi_0; \phi_0]$.) Now, the fixed Φ_0, independent of the presently

varied parameters, already optimized at its own energy minimum, is close to ϕ_0, i.e., $\langle\phi_0|\,\Phi_0\rangle^2 \sim 1 > \langle\phi_1|\,\Phi_0\rangle^2 \sim 0$, so, near the crossing, the denominators in Eq. 13.4 are $[1 - \langle\phi_1|\,\Phi_0\rangle^2]^{-1} < [1 - \langle\phi_0|\,\Phi_0\rangle^2]^{-1}$ (for the two lowest states $\ll$ holds), while the numerators in Eq. 13.4, $\langle\Phi_0|H - E[\phi_0]|\phi_0\rangle^2$, $\langle\Phi_0|H - E[\phi_1]|\phi_1\rangle^2$, remain finite. But, in optimizing the 2nd "root," the 1st "root" deteriorates. So, if the (supposedly good) lower approximant Φ_0 is better than the deteriorated $\Phi_0^{"1r"}$ near the crossing, i.e., if $E[\Phi_0] < E\left[\Phi_0^{"1r"}\right] \sim E\left[\Phi_1^{"2r"}\right]$, then the 2nd terms (i.e., the sums $2\sum\ldots$ – the "annexations," so to speak, to the energy) of both individually computed $F_1\left[\Phi_0; \Phi_1^{"2r"}\right]$ and $F_1\left[\Phi_0; \Phi_0^{"1r"}\right]$ are positive (cf. Eq. 13.4), and, due to the smaller denominator, the F_1 of (the continuous) ϕ_0 is normally **larger** than the F_1 of (the continuous) ϕ_1. Therefore, in passing the crossing, for the blindly taken 2nd root, we have that $F_1\left[\Phi_0; \Phi_1^{"2r"}\right]$ jumps up from $F_1\left[\Phi_0; \Phi_{1-}^{"2r"}\right] = F_1[\Phi_0; \phi_{1-}] = F_1[\Phi_0; \phi_1]$ to $F_1\left[\Phi_0; \Phi_{1+}^{"2r"}\right] = F_1[\Phi_0; \phi_{0+}] = F_1[\Phi_0; \phi_0]$:

$$F_1\left[\Phi_0; \Phi_{1-}^{"2r"}\right] = F_1[\Phi_0; \phi_1] < F_1[\Phi_0; \phi_0] = F_1\left[\Phi_0; \Phi_{1+}^{"2r"}\right], \qquad (13.16)$$

while for the blindly taken 1st root, we have that $F_1\left[\Phi_0; \Phi_0^{"1r"}\right]$ jumps down from $F_1\left[\Phi_0; \Phi_{0-}^{"1r"}\right] = F_1[\Phi_0; \phi_{0-}] = F_1[\Phi_0; \phi_0]$ to $F_1\left[\Phi_0; \Phi_{0+}^{"1r"}\right] = F_1[\Phi_0; \phi_{1+}] = F_1[\Phi_0; \phi_1]$:

$$F_1\left[\Phi_0; \Phi_{0-}^{"1r"}\right] = F_1[\Phi_0; \phi_0] > F_1[\Phi_0; \phi_{1+}] = F_1\left[\Phi_0; \Phi_{0+}^{"1r"}\right]. \qquad (13.17)$$

In both cases

$$F_1[\Phi_0; \phi_1] < F_1[\Phi_0; \phi_0].$$

Hence, near the crossing, if $E[\Phi_0] < E\left[\Phi_0^{"1r"}\right] \sim E\left[\Phi_1^{"2r"}\right]$, *we can recognize ϕ_1: It is the (blindly taken) "root" that has the lowest (individually computed) F_1.* Of course, it is a continuation of an unambiguous value of $F_1[\Phi_0; \phi_1]$ well before the crossing – if it is known. It might be possible to just check the closeness to Φ_0, or, for the nth state, to check $\prod_{i<n}(1 - \langle\Phi_i|\phi_n\rangle^2)^{-1}$ vs. $\prod_{i<n}(1 - \langle\Phi_i|\phi_{n-1}\rangle^2)^{-1}$ because all terms are close to 1, except the one resembling some lower crude approximant $\Phi_{i<n}$, that would make it close to 0^{-1}, but this possibility needs future investigation. (Generally, the recognition of ϕ_1 could be used in MCSCF to feed the next iteration.)

Yet, by just recognizing ϕ_1 **at** the crossing, we have not found the sought **optimal** position of $\mathbf{p_1}$, which is "**before**" the crossing. Therefore, at every quasi-Newton cycle, we must "suggest" new trial $\mathbf{p}$ values that are "**before**" the crossing, without being stuck at the crossing. To ensure that we are not stuck at the crossing, we can use the procedure shown in Table 13.2, which extrapolates the quasi-Newton $\mathbf{p}$ values every three steps by twice the final step – or more if convergence is slow, the question being to decide, via F_1, whether a given value of $\mathbf{p}$, near the crossing, is "before" or "beyond" the crossing.

Table 13.2 Three-step
quasi-Newton (qN)
extrapolation

```
if F₁[Φ₀;Φ¹ˢᵗ ʳᵒᵒᵗ(p)]< F₁[Φ₀;Φ²ⁿᵈ ʳᵒᵒᵗ(p)] then
    n(p) ← 1
else
    n(p) ← 2.

if |p3-p2|>|p2-p1| then
    h ← p3
else
    h ← p3+(p3-p2).

if n(p1)≡1 then
    { p2 ← qN(p1) and
    if n(p2)≡1 then
        [ p3 ← qN(p2) and
        if n(p3)≡1 then
            p ← h
        else if n(p3)(≡2)≠1 then
            p ← qN(p3)
        ]
    else if n(p2)(≡2)≠1 then
        p ← qN(p2)
    }
else if n(p1)(≡2)≠1 then
    p ← qN(p1).
```

(Incidentally, note that **"state averaging"** *at* the crossing does not provide any information about the correct, and sought, optimal points p_0 and p_1 that must be located "***before***" the crossing.) The recognition is demonstrated below for 1-electron atomic ion S-states.

13.6 Demonstrations

Below demonstrations of the above are presented for He ^1S ($1s^2$ and $1s2s$), by using Hylleraas variables $s = r_1 + r_2$, $t = r_1 - r_2$ and $u = |\vec{r}_1 - \vec{r}_2|$ [1], so as to establish rather accurate basis-functions out of variationally optimized state-specific Laguerre-type orbitals [79], where the polynomial coefficients and exponents are optimized, allowing few term (small-size) series expansions in terms of $s^i(t^2)^j u^k$: For example, by selecting 24 terms up to 0.001 $s^2(t^2)^2 u^3$, the obtained energy is $E_0 \simeq -2.90372$ a.u. However, for demonstration reasons, all 27 terms up to $s^2(t^2)^2 u^2$ are used ($E_0 \simeq -2.90371$ a.u., $E_1 \simeq -2.14584$ a.u.), along with all 8 terms up to $s^1(t^2)^1 u^1$ whose ϕ_0 will be immediately improved via the 27-term ϕ_1. (Pekeris' 9-term ϕ_1 is still unbound; he reports larger than 95 terms wave functions [80] – indicating that the present optimized Laguerre-type orbitals reduce the size efficiently.) Evidently, if E_1 were minimized, instead of F_1, the $2s$ Laguerre-type orbital $(1 - ar)\,e^{-\zeta_1 r}$ would collapse to $1s$ ($a \to 0$, $\zeta_1 \to \zeta_0$), transparently showing E_1's "saddleness" at $|\phi_1\rangle = |1\rangle$.

In the next section, the formalism will be described (see [81] for *matrix elements of* $\langle \phi_i | \phi_n \rangle$, $\langle \phi_i | H | \phi_n \rangle$ *in Hylleraas coordinates*). Then, the results will be presented, including the immediate improvement of ϕ_0 via a more accurate ϕ_1. And then, a demonstration of recognizing a "flipped root" near the crossing will follow. After the demonstrations using the *Hylleraas coordinates* demonstrations for He and Li by conventional CI computation, appropriate for any number of electrons, will be presented.

13.6.1   Formalism in Hylleraas Coordinates

For two-electron atomic ions of nuclear charge Z, the wave function will consist of a single Slater determinant multiplied by a truncated power series of $s = r_1 + r_2$, $t = r_1 - r_2$, and $u = |\vec{r}_1 - \vec{r}_2|$: $\sum_{i_s, i_t, i_u = 0}^{n_s, n_t, n_u} c_{i_s, i_t, i_u} s^{i_s} t^{2i_t} u^{i_u}$. Due to the spin antisymmetry, the Slater determinant is reduced to a symmetric sum of products and the power series to a symmetric function of t, i.e., of t^2. And since the Hamiltonian is also symmetric, the t-integrals could be evaluated only for $t > 0$ (eventually multiplied by 2), so that, with volume element $2\pi^2(s^2 - t^2)u\, dt\, du\, ds$, the limits of integration be $0 < t \leq u \leq s < \infty$ [82].

The spin-orbitals are composed of Laguerre-type radial orbitals, where their polynomial coefficients are treated as variational parameters [79]. For low-lying *singlet* states, only s- such orbitals will be considered. Thus, the spatial orbitals will be

$$\chi\left(n, r; z_n, \{a_{n,k}\}\right) = \frac{4\sqrt{\pi}\sqrt{(n-1)!n!}}{n^2} z_n^{3/2} \sum_{k=0}^{n-1} \frac{a_{n,k}(-2rz_n/n)^k e^{-rz_n/n}}{k!\,(k+1)!\,(n-k-1)!}$$

where the variational parameters are z_n and $a_{n,k}$ factors [the latter are expected to have values near 1 (for $k = 0$, $a_{n,0} \equiv 1$)] for state-specific functions, allowing also the possibility for non-state-specific description: $a_{n,0} \equiv 1$, $a_{n,k>0} = 0$. (In wider parameter space, the free non-state-specific functions turn out to be slightly more accurate than the state-specific.) The prefactors assert orbital orthonormality for one-electron ions (all $z_n = Z$ and all $a_{n,k} = 1$). The power series will be truncated at most up to the 2nd power, $s^2(t^2)^2u^2$, sufficient for the purposes of the present demonstrations.

The two-electron Hamiltonian in terms of the s, t, u variables is given, e.g., in Refs. [82] or [83], but:

(i) by selecting r_1 or r_2 via the sum $(1 - q)r_1 + qr_2$, $(q = 0, 1)$, (so that the symmetric sum of products be expressed as a monomial sum, e.g.,

$$\sum_{q=0}^{1} \psi\left((1-q)\,r_1 + qr_2\right) \chi\left(qr_1 + (1-q)\,r_2\right)),$$

(ii) by binomial expanding $r_{1,2}^k \sim (s \pm t)^k$,

(iii) by using $\left(s^2 - t^2\right)\varphi = \sum\limits_{p=0}^{1} (-1)^p s^{2(1-p)} t^{2p}\varphi$, and,

(iv) by writing the required derivatives as $\frac{\mathrm{d}}{\mathrm{d}x}\left(x^n e^{-zx}\right) = \sum\limits_{b=0}^{1} [n(1-b) - zb]$
$x^{n-1+b}e^{-zx}$,

all terms can be reduced to monomial sums, so that the total symmetric two-electron function be written as

$$\phi\left(\overrightarrow{r}_1, \overrightarrow{r}_2\right) = \Phi^{\left(n_{1\Phi},n_{2\Phi},n_{s\Phi},n_{t\Phi},n_{u\Phi}\right)}(s,t,u) = \frac{1}{\pi}\sqrt{(n_{1\Phi}-1)!n_{1\Phi}!}$$
$$\sqrt{(n_{2\Phi}-1)!n_{2\Phi}!}\left(1 - \tfrac{1}{2}\delta_{n_{1\Phi},n_{2\Phi}}\right)$$

$$\sum_{i_{s\Phi}=0}^{n_{s\Phi}}\sum_{i_{t\Phi}=0}^{n_{t\Phi}}\sum_{i_{u\Phi}=0}^{n_{u\Phi}}\sum_{k_{2\Phi}=0}^{n_{2\Phi}-1}\sum_{k_{1\Phi}=0}^{n_{1\Phi}-1}\sum_{q\Phi=0}^{1}\sum_{j_{2\Phi}=0}^{k_{2\Phi}}\sum_{j_{1\Phi}=0}^{k_{1\Phi}}\left[(-1)^{j_{1\Phi}+k_{1\Phi}+k_{2\Phi}}\right.$$
$$(2q_\Phi - 1)^{j_{1\Phi}+j_{2\Phi}}n_{1\Phi}^{-2-k_{1\Phi}}n_{2\Phi}^{-2-k_{2\Phi}}z_{\Phi,n_{1\Phi}}^{\frac{3}{2}+k_{1\Phi}}z_{\Phi,n_{2\Phi}}^{\frac{3}{2}+k_{2\Phi}}e^{-s\left(\frac{z_{\Phi,n_{1\Phi}}}{2n_{1\Phi}}+\frac{z_{\Phi,n_{2\Phi}}}{2n_{2\Phi}}\right)}$$
$$-t\left(\tfrac{1}{2}-q_\Phi\right)\left(\frac{z_{\Phi,n_{1\Phi}}}{n_{1\Phi}}-\frac{z_{\Phi,n_{2\Phi}}}{n_{2\Phi}}\right)s^{i_{s\Phi}-j_{1\Phi}-j_{2\Phi}+k_{1\Phi}+k_{2\Phi}}t^{2i_{t\Phi}+j_{1\Phi}}$$
$$\left. + j_{2\Phi}u^{i_{u\Phi}}a_{\Phi,n_{1\Phi},k_{1\Phi}}a_{\Phi,n_{2\Phi},k_{2\Phi}}c_{\Phi,i_{s\Phi},2i_{t\Phi},i_{u\Phi}}\right]/$$
$$\left[j_{1\Phi}!j_{2\Phi}!\left(k_{1\Phi}+1\right)!\left(k_{1\Phi}-j_{1\Phi}\right)!\left(k_{2\Phi}+1\right)!\left(k_{2\Phi}-j_{2\Phi}\right)!\left(n_{1\Phi}-1-k_{1\Phi}\right)!\left(n_{2\Phi}-1-k_{2\Phi}\right)!\right]$$

where $n_{i\Phi}$ is the degree of the associated Laguerre polynomial of the ith spin-orbital in the Φ-Slater determinant (which is multiplied by the power series expansion). Hence, the total wave function is characterized by the set of integers $(n_1, n_2, n_s, n_t, n_u)$ (where $n_i - 1$ is the number of nodes of the ith polynomial). Thus, the ground state's, $1s^2$, approximant is characterized by $(1, 1, n_s, n_t, n_u)$ and the 1st excited state's, $1s2s$, approximant, by $(1, 2, n_s, n_t, n_u)$ or equivalently by $(n_1 = 2, n_2 = 1, n_s, n_t, n_u)$, in state-specific description, or also by the 2nd root of "$(1, 1, n_s, n_t, n_u)a = 0$" in non-state-specific description (where "$a = 0$" means that all parameters $a_{n,k \neq 0} = 0$), and these can be calculated either by using (minimizing) the F functional – indicated by an index F, or not.

By applying Green's theorem in the integrals, in order to deal with first derivatives, Φ_s, Φ_t, Φ_u, instead of Laplacians, the Hamiltonian and overlap matrix elements are:

$$\langle\Phi|H|\Psi\rangle = 2\pi^2\int_0^\infty\int_0^s\int_0^u\left[-4Z\,s\,u\,\Phi\,\Psi + \sum_{p=0}^{1}(-1)^p(s^{2(1-p)}t^{2p}u\,(\Phi_s\Psi_s\right.$$
$$\left.+\Phi_t\Psi_t + \Phi_u\Psi_u) + u^{2(1-p)}t^{2p}s\,(\Phi_u\Psi_s + \Phi_s\Psi_u) + s^{2(1-p)}u^{2p}t\,(\Phi_u\Psi_t + \Phi_t\Psi_u)\right.$$
$$\left.+s^{2(1-p)}t^{2p}\Phi\,\Psi)\right]\mathrm{d}t\,\mathrm{d}u\,\mathrm{d}s$$

and

$$\langle \Phi | \Psi \rangle = 2\pi^2 \int_0^\infty \int_0^s \int_0^u \sum_{p=0}^1 (-1)^p s^{2(1-p)} t^{2p} u \; \Phi \; \Psi \; \mathrm{d}t \; \mathrm{d}u \; \mathrm{d}s$$

The integrals, of the form $J(z_s, z_t; n_s, n_t, n_u) = \int_0^\infty \int_0^s \int_0^u e^{-z_s s - z_t t} s^{n_s} t^{n_t} u^{n_u} \mathrm{d}t \; \mathrm{d}u \; \mathrm{d}s$, are reported in [81].

13.6.1.1 F_n Minimization Procedure

To minimize F_n, the fixed lower approximants $\{\phi_i, i < n\}$ (that can be crude, but reasonable, needed to compute F_n) have been computed (independently of each other, and not necessarily orthogonal to each other, since, in F_n, Schrödinger's equation is "dotted" into them). For each (varied) value of the parameters z_n and $a_{n,k}$, the critical points of F_n (saddle points of E_n) are found, and F_n is minimized (*without* demanding orthogonality to the lower approximants .$\{\phi_i, i < n\}$ Orthogonality, among the functions that minimize each F_i, $i < n$, should be an outcome). In finding the saddle points of E_n, for certain trial values of the parameters z_n and $a_{n,k}$, the linear part, $\partial \langle \phi_n | H | \phi_n \rangle / \partial c_{i_s, 2i_t, i_u} = 0$, can be solved either by direct variation of the c-coefficients or, preferably, by reducing to a generalized eigenvalue problem (requiring $\langle \phi_n | \phi_n \rangle$ for normalization). Thus, the *lowest* eigenvalue (1st root) *above the known* (highest) $E[\phi_{n-1}]$ and the c-coefficients of its eigenvector are substituted in Eq. 13.4 of F_n, and improved values of z_n and $a_{n,k}$ are sought until minimization of F_n.

13.6.1.2 Establishing "Exact" Wave Functions ψ_0, ψ_1 and Truncated Approximants ϕ_0, ϕ_1

First a reliable wave function basis $|\psi_0\rangle \cong |0\rangle$, $|\psi_1\rangle \cong |1\rangle$ is established by taking 27 terms, i.e., up to $s^2(t^2)^2 u^2$, $(n_s, n_t, n_u) = (2, 2, 2)$, adequate to achieve coincidence of the HUM and F_1 minima: ($E[\phi_0 \sim \psi_0] = -2.90371$ a.u., $z_0 = 1.9549$), ($E\left[\psi_1^{\mathrm{HUM}}\right] = -2.14584$, $z_{1,1} = 1.8348$, $z_{1,2} = 1.9745$, $a_{1,2,1} = 0$; or $E\left[\psi_1^{F_1}\right] = -2.14577$, $z_{1,1} = 1.930501$, $z_{1,2} = 1.827298$, $a_{1,2,1} = 0.799760$; $\left\langle \psi_1^{F_1} | \psi_1^{\mathrm{HUM}} \right\rangle = 0.99996$) – thus, going up to $(n_s, n_t, n_u) = (2, 2, 2)$ is sufficient. Then, in order to exhibit the aforesaid demonstrations, truncations up to $s^1(t^2)^1 u^1$, $(n_s, n_t, n_u) = (1, 1, 1)$, i.e., 8 terms, will be used.

In F_1, a fixed 1-term normalized ϕ_0, i.e., up to $s^0(t^2)^0 u^0$, $(n_s, n_t, n_u) = (0, 0, 0)$ ($z_0 = 1.6875$, $E[\phi_0] = -2.84766$ a.u.), is used.

13.6.1.3 Results

For He ($Z = 2$) ^1S, the exact eigenvalues are [80] $E_0 = -2.90372$ a.u., $E_1 = -2.14597$ a.u., $E_2 = -2.06127$ a.u.

The Lowest Orthogonal to ϕ_0 (OO)

Before proceeding to HUM and F_n results and their comparisons, it will be demonstrated that minimizing the excited state energy while keeping orthogonality to a normalized *approximant*, ϕ_0, of the ground state (e.g., the 1-term "fixed," used in F_1) leads to a function, ϕ_1^{LE}, lying lower than the exact, therefore veered away from the exact.

From a trial function ϕ (unnormalized), subtract its projection to (normalized) ϕ_0, to obtain the required normalized orthogonal, and its energy:

$$\phi_\perp \equiv \Phi_1 = \frac{\phi - \phi_0 \langle \phi_0 | \phi \rangle}{\sqrt{\langle \phi | \phi \rangle - |\langle \phi_0 | \phi \rangle|^2}} \; ;$$

$$\langle \Phi_1 | H | \Phi_1 \rangle = \frac{\langle \phi | H | \phi \rangle - 2 \langle \phi_0 | H | \phi \rangle \langle \phi_0 | \phi \rangle + \langle \phi_0 | H | \phi_0 \rangle |\langle \phi_0 | \phi \rangle|^2}{\langle \phi | \phi \rangle - |\langle \phi_0 | \phi \rangle|^2},$$

(13.18)

which will be used as a Rayleigh-Ritz quotient.

To obtain the energy minimum, either all parameters ($\{z\},\{a\},\{c\text{'s}\}$) of the trial function ϕ can be varied, or only the nonlinear ($\{z\},\{a\}$) can be varied, while (for every trial $\{z\},\{a\}$) the linear part, $\{c\}$, can be reduced to a generalized eigenvalue problem $(\mathbf{A} - \lambda \mathbf{B}) \cdot \mathbf{c} = 0$, where $\mathbf{A}_{ij}$ are the coefficients of $c_i c_j$ of the numerator and $\mathbf{B}_{ij}$ of the denominator of the Rayleigh-Ritz quotient and where λ is a Lagrange multiplier, whose lowest value (1st root) is minimized by varying ($\{z\},\{a\}$). The corresponding lowest-lying eigenvector $\{c\}$ provides the coefficients of ϕ, which, along with ($\{z\},\{a\}$), minimize $\langle \Phi_1 | H | \Phi_1 \rangle$, to obtain the final ϕ_1^{LE}, by normalizing the function $\phi - \phi_0 \langle \phi_0 | \phi \rangle$.

After minimization of the above Rayleigh-Ritz quotient, (without using ψ_1 or ϕ_1^+) ϕ_1^{LE} has energy -2.14762 a.u., certainly below the exact or of the above 27-term ψ_1, -2.14584. The 27-term minimizing function ϕ forms $\phi - \phi_0 \langle \phi_0 | \phi \rangle$, whose main orbital part has the form $(\upsilon(r_1)\chi(r_2) + \upsilon(r_2)\chi(r_1) - 1)\omega(r_1)\omega(r_2)$, where $\omega(r)$ is the 1s orbital of ϕ_0 (after normalization) and $\upsilon(r)$ and $\chi(r)$ are shown in Fig. 13.5. They remind 1s and 2s (where the 2s is more remote), which means that ϕ_1^{LE}, despite its lower energy, is not collapsed but, of course, is veered away from the exact.

$(n_s, n_t, n_u) = (2, 2, 2)$: 27 Terms

First establish a reliable basis $|\psi_0\rangle \cong |0\rangle$, $|\psi_1\rangle \cong |1\rangle$, (27 terms) in order to compare (project on it) the truncated $(n_s, n_t, n_u) = (1, 1, 1)$ 8-term approximants:

Fig. 13.5 Main orbitals of
the 27-term (solid line) and
8-term (dashed line) optimal
wave functions. (**a**) Ground
state $1s^2$. (**b**) HUM excited
state $1s1s'$ (either $a = 0$, or
$a \approx 0$ as explained in the text,
compared to $1s^2$. (**c**) F_1
excited state $1s2s$ compared
to $1s^2$. (**d**) Lowest orthogonal
to ϕ_0 (OO)
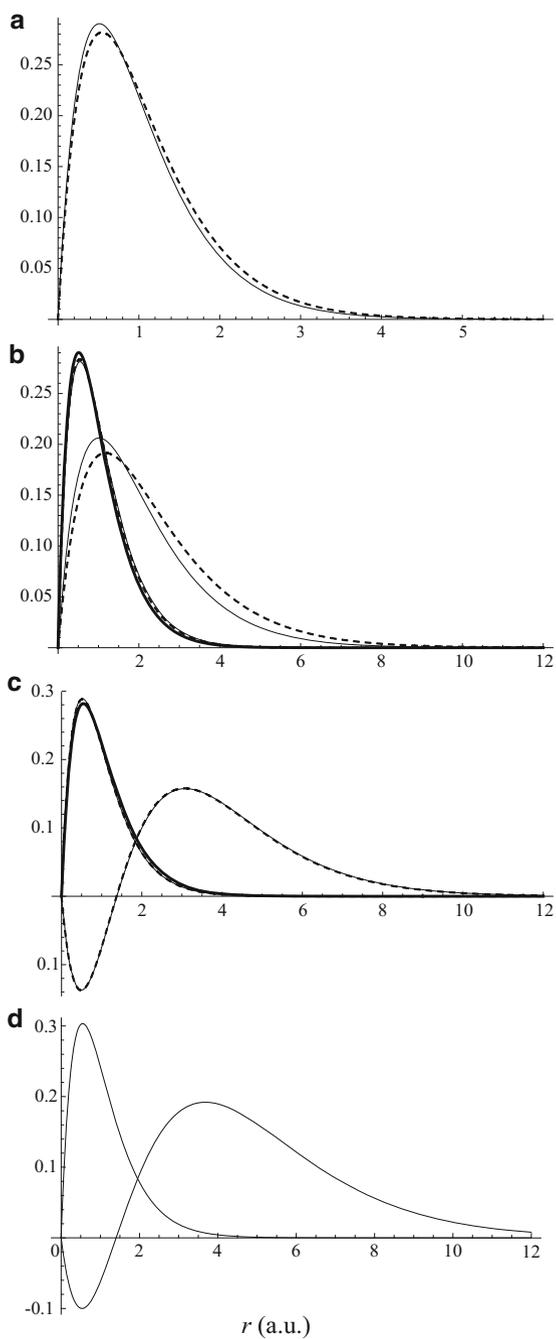
Table 13.3 Energies and overlaps of the computed states. Notation "tSzaR": $t = \{\psi$: 27 terms, Φ: 8 terms$\}$, $S = \{$via optimization of: 0: ground state, 1: excited state$\}$, $z = \{0: z_1$ only, 1: $z_1, z_2\}$, $a = \{0: a = 0, 1: a \neq 0\}$, $R = \{0:$ 1st root, 1: 2nd root, H: HUM, F: $F_1\}$

| Φ = tSzaR	E (a.u.)	$(\psi_0\|\Phi)$	$(\psi_1\|\Phi)$	$(\Phi010H\|\Phi)$
$\psi0000 \rightarrow \psi_0$	**−2.90371**			
($\psi0001$)	(−2.1391)			
$\psi0100$	−2.90371	1		
($\psi0101$)	(−2.1306)			
$\psi110H \rightarrow \psi_1$	**−2.14584**	0.000172		
($\psi1100$)	(−2.90327)			
$\psi111F$	−2.14577	0.000457	0.99996	
ϕ_0 (1-term)	−2.84766	0.993		
$\Phi0000$	−2.903121	0.999958	1.3 10^{-5}	
($\Phi0001$)	(−2.01016)			
$\Phi100H$	−2.07215	1.97 10^{-5}	0.875	
($\Phi1000$)	(−2.89748)			
$\Phi100F$	−2.07215			
$\Phi010H$	**−2.903123**			
($\Phi0101$)	(−2.0196)			
$\Phi110H$	−2.14449			0.0026
($\Phi1100$)	(−2.8886)			
$\Phi110F$	−2.14449			
$\Phi111H \rightarrow 1s1s'$	**−2.14449**	**0.0033**	**0.9986**	
$\Phi111F \rightarrow 1s2s$	**−2.145152**	**0.00490**	**0.999807**	0.0186

Ground State 27 Terms

The optimized ground state approximant $0^{(1,1,2,2,2)}$ [cf. monosyllabic "tSzaR" = $\psi0000$ in Table 13.3] has lowest root energy $E[0^{(1,1,2,2,2)}] = -2.90371$ a.u. and parameters $z_{0,1} = 1.954881$, and the 27 c-expansion coefficients are given in Ref. [81]. (The optimized 2nd root with the same $z_{0,1}$ is rather high: (tSzaR = $\psi0001$), $E = -2.1391$ a.u.)

Also, by including the $z_{1,2}$ parameter, but $a_{1,2,1} = 0$, the optimized 1st root $0^{(1,2,2,2,2)a = 0}$ (tSzaR = $\psi0100$) has $E[0^{(1,2,2,2,2)a = 0}] = -2.90371$ a.u., $z_{1,1} = 1.93945$, $z_{1,2} = 3.8926$ and corresponding c-expansion coefficients as given in Ref. [81]. (The optimized 2nd root with the same $z_{1,1}, z_{1,2}$ is rather high again: (tSzaR = $\psi0101$) $E = -2.1306$ a.u.)

As seen in Fig. 13.6, the minimum is very flat, but indeed, within the accuracy of the computation (~ 6 digits[1]), the 2nd function has, as expected, $z_{1,2} = 2z_{1,1}$, i.e., both represent the same $1s$ orbital, with $\{c\}$ coefficients very similar to the 1st, and $\langle 0^{(1,1,2,2,2)}|0^{(1,2,2,2,2)a = 0}\rangle = 1. = (\psi0000|\psi0100)$. Among the above

---

[1]The computation is performed in 15 digit "double precision," but the coexistence in the secular matrix of very large numbers with small ones reduces the accuracy to ~ 6 digits
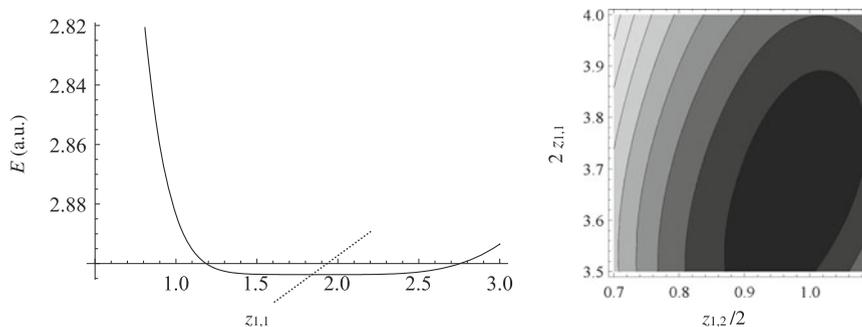
Fig. 13.6 1st root's energy (in a.u.) v.s. $z_{1,1}$ (left) and 2nd root's energy v.s. $z_{1,1}, z_{1,2}$ (right). Dotted line: Some other property might not have 0 slope at the (flat) energy minimum

two approximants representing the same unique ground state the simplest is used, $0^{(1,1,2,2,2)}$, (tSzaR $= \psi 0000$), as a quite reliable approximant of ψ_0.

Incidentally, because of the flatness of the minimum, any traditional convergence criterion (e.g., $\delta E < \varepsilon = 10^{-6}$) can be easily satisfied anywhere between $1.5 < z_{1,1}$, < 2.3, so that any other property not having extremum at ψ_0 but rather would pass from it with a finite slope would be computed arbitrarily incorrectly; some reference could give it positive, while others would give it negative. The same comment holds also for the 1st excited state – [cf. below and Fig. 13.6] – which is also quite flat in varying the $z_{1,1}, z_{1,2}$ parameters. Thus the exact F minimum is required.

2nd HUM Root 27 Terms

The optimized 2nd HUM root $1^{(1,2,2,2,2)a=0}$ has (tSzaR $= \psi 110H$) $E[1^{(1,2,2,2,2)a=0}]$ $= -2.14584$ a.u., $z_{1,1} = 1.8348$, $z_{1,2} = 1.9745$ and corresponding c-expansion coefficients as given in Ref. [81]. (Of course, this is orthogonal to its deteriorated 1st root Φ_0^{1r}, (tSzaR $= \psi 1100$) $E = -2.90327$ compared to $E[0^{(1,1,2,2,2)}] = -2.90371$).

Its overlap is $\langle 0^{(1,1,2,2,2)} | 1^{(1,2,2,2,2)a=0} \rangle = 0.000172 = (\psi 0000 | \psi 110H)$, and its main two orbitals are as in Fig. 13.5. However, these resemble $1s, 1s'$, rather than $1s, 2s$, but since its F-value is the same ($F - E = 3 \times 10^{-9}$), because its 27-term series is large, it can be considered close to the excited state saddle point.

Also, by allowing $a_{1,2,1} \neq 0$, the optimized 2nd HUM root has $a_{1,2,1} \approx 0$, introducing a node very far from the nucleus (rendering it *literally* "2"s!, but *essentially* $1s'$), and its energy is negligibly better than the above with $a_{1,2,1} = 0$. Actually there is also another minimum at $a_{1,2,1} = 0.539$, with $E[1^{(1,2,2,2,2)}] = -2.126$ a.u. Although $z_{1,1} = 2.106$, $z_{1,2} = 2.478$, i.e., the *main* (without the series expansion) orbital $2s$ is orthogonal to the *main* $1s$, this high lying function is rejected. This suggests that using a (perhaps habitual) criterion *of orthogonality of the **main** orbitals,* in computing excited states, ***is not sufficient***, because many random functions may have mutually orthogonal *main* orbitals.
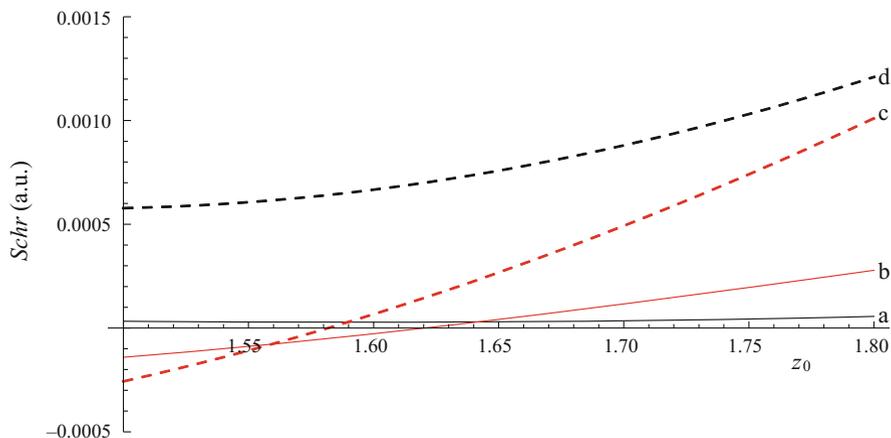
Fig. 13.7 The difference Schr $\equiv \langle 0^{(1,1,0,0,0)}|(H - E[\Phi])|\Phi\rangle$ vs. the exponent z_0 of normalized $0^{(1,1,0,0,0)}$, for optimized $\Phi = (a)$: $1^{(1,2,2,2,2)a=0}$ HUM 2nd root, (b) $1^{(1,2,2,2,2)F,a}$, (c) $1^{(1,2,1,1,1)F,a}$, (d) $1^{(1,2,1,1,1)a=0}$ HUM 2nd root

"F_1" Root, 27 Terms

The optimized "F_1" root $\Phi_1^F \Phi_1^F$ (the 1st above the fixed $E[\phi_0] = -2.84766$ a.u.) has (tSzaR = ψ111F) $E[1^{(1,2,2,2,2)F}] = -2.14577$ a.u., $(F - E = 3 \times 10^{-8})$ a.u., $z_{1,1} = 1.930501$, $z_{1,2} = 1.827298$, $a_{1,2,1} = 0.799760$, and corresponding c-expansion coefficients as given in Ref. [81]. Its main orbitals are as in Fig. 13.5, resembling $1s$, $2s$. It has overlap $\langle 0^{(1,1,2,2,2)}| 1^{(1,2,2,2,2)F}\rangle = 0.000457 = (\psi 0000|\psi 111F)$ and $\langle 1^{(1,2,2,2,2)a=0}| 1^{(1,2,2,2,2)F}\rangle = 0.99996 = (\psi 110H|\psi 111F)$. Although their main orbitals differ, their series expansions, up to $(n_s, n_t, n_u) = (2, 2, 2)$, are large, and the two wave functions essentially coincide.

Checking the Satisfaction of Schrödinger's Equation

If these two functions Φ_1^{2r} and Φ_1^F are close to the exact, they should satisfy Schrödinger's equation for any normalized "ϕ_0" not orthogonal to Φ_1:

$$\text{Schr} \equiv \langle \phi_0| (H |\Phi_1\rangle - E[\Phi_1]|\Phi_1\rangle) = \langle \phi_0| H |\Phi_1\rangle - E[\Phi_1] \langle \phi_0|\Phi_1\rangle \sim 0.$$

(In principle it might be possible for some $\langle \phi_0|$ to be accidentally almost orthogonal to *both* $|\Phi_1\rangle$ *and* $H|\Phi_1\rangle$, making Schr artificially small without Φ_1 being close to the exact, that could pull F_1 below E_1: $E < F_1 < E_1$. To avoid this, just use some other ϕ_0 for the same Φ_1.)

Using the above fixed ϕ_0, the Schr difference equals $\sim 3 \times 10^{-5}$ for Φ_1^{2r} and $\sim 8 \times 10^{-5}$ for Φ_1^F. Figure 13.7 shows the Schr difference for both Φ_1^{2r} and Φ_1^F, along with their more truncated versions (cf. below) $1^{(1,2,1,1,1)a=0}$ (Schr $\sim 8 \times 10^{-4}$) and $1^{(1,2,1,1,1)F}$ (Schr $\sim 4 \times 10^{-4}$), for various ϕ_0s, i.e., various values

of the exponent z_0. Indeed, Φ_1^{2r} has slightly better Schr values than Φ_1^F, contrary to the more truncated versions.

Thus, a quite reliable basis (27 terms) has been established, adequate for the attempted demonstration, by adopting $\{\psi_0 = 0^{(1,1,2,2,2)}, \psi_1 = 1^{(1,2,2,2,2)a = 0}\}$, since the state-specific expansion, despite its more reasonable main orbitals, is slightly inferior in Schr than the expansion of the (large) HUM 2nd root.

$(n_s, n_t, n_u) = (1, 1, 1)$: 8 Terms

Minimal Basis

Ground State 8 Terms

The optimized ground state 8-term approximant $0^{(1,1,1,1,1)}$ has lowest root energy (cf. Table 13.3: tSzaR = Φ0000) $E[0^{(1,1,1,1,1)}] = -2.903121$ a.u. and parameters $z_{0,1} = 1.84250$, $c_{0,0,0,0} = 1$, $c_{0,0,0,1} = 0.290798$, $c_{0,0,1,0} = 0.190212$, $c_{0,0,1,1} = -0.0765151$, $c_{0,1,0,0} = 0.00688611$, $c_{0,1,0,1} = 0.0139948$, $c_{0,1,1,0} = 0.016833$, $c_{0,1,1,1} = 0.0120625$, normalization constant $N = 1/1.47861$, while its orthogonal 2nd root $1^{[0^{(1,1,1,1,1)}]}$ (same $z_{0,1}$) lies too high: (tSzaR = Φ0001) $E[1]^{[0^{(1,1,1,1,1)}]} = -2.01016$ a.u. Since $0^{(1,1,1,1,1)}$ has $\langle \psi_0 | 0^{(1,1,1,1,1)} \rangle = 0.999958$ and $\langle \psi_1 | 0^{(1,1,1,1,1)} \rangle = 1.3 \times 10^{-5}$, it is nearly "perfect." However, for the 1st excited state, as seen below, this truncation (1,1,1,1,1) is not adequate: A richer function (1,2,1,1,1) is needed.

2nd HUM Root and "F_1" Root 8 Terms

The optimized 8-term 2nd HUM root $1^{(1,1,1,1,1)}$ (see also Fig. 13.5) has (tSzaR = Φ100H) $E[1^{(1,1,1,1,1)}] = -2.07215$ a.u., $z_{1,1} = 1.44234$, and corresponding c-expansion coefficients $\{c\} = \{1, 0.199163, 0.0540545, 0.0273041, -0.418026, -0.0402484, -0.0720681, -0.040584\}$, $N = 1/1.3566$, while its orthogonal 1st root $0^{[1^{(1,1,1,1,1)}]}$ (same $z_{1,1}$) is, of course, deteriorated: (tSzaR = Φ1000) $E[0]^{[1^{(1,1,1,1,1)}]} = -2.89748$ a.u., $\{c\} = \{1, 0.145873, 0.160241, -0.0416745, -0.245673, -0.00154526, -0.0197971, 0.00659241\}$. The overlap between the normalized independently optimized approximants is $\langle 0^{(1,1,1,1,1)} | 1^{(1,1,1,1,1)} \rangle = -1.97349 \times 10^{-5} = (\psi_0 | \Phi 100H)$, meaning that the **HUM root** $1^{(1,1,1,1,1)}$ has contributions from other higher states orthogonal to the 1st root (its energy is close to the 3rd). The overlap with the above established $\psi_1 = 1^{(1,2,2,2,2)}$ is $\langle \psi_1 | 1^{(1,1,1,1,1)} \rangle = 0.875$. By using F_1 the optimized "F" root $1^{(1,1,1,1,1)F}$ (tSzaR = Φ100F), in lack of other parameters, nearly coincides with the 2nd HUM root near the minimum.

Richer Basis

In a richer parameter space, including $z_{1,2}$ and, optionally, $a_{1,2,1} \neq 0$ (state-specific description), the optimized 8-term functions are as follows:

$a_{1,2,1} = 0$

HUM 1st Root 8 Terms, $a_{1,2,1} = 0$

The optimized 8-term HUM 1st root $0^{(1,2,1,1,1)}$ has (tSzaR = Φ010H) $E = -2.903123$ a.u, $z_{0,1} = 1.93393$, $z_{0,2} = 3.50164$, $\{c\} = \{1, 0.290154, 0.184912, -0.077204, 0.006390, 0.014435, 0.017008, 0.011521\}$, $N = 1/2.94444$, with 2nd root severely deteriorated (tSzaR = Φ0101), E = -2.0196 a.u. Thus, the 8-term $0^{(1,2,1,1,1)}$ is not essentially improved over the above simpler 8-term $0^{(1,1,1,1,1)}$.

HUM 2nd Root 8 Terms, $a_{1,2,1} = 0$

However, the optimized 8-term HUM 2nd root $1^{(1,2,1,1,1)}$ has (tSzaR = Φ110H) $E = -2.14449$ a.u., $z_{1,1} = 0.851359$, $z_{1,2} = 3.73405$ (or $z_{1,1} = 1.86703$, $z_{1,2} = 1.70272$), $\{c\} = \{1, 0.180637, -0.127452, 0.019460, -0.367640, -0.065135, -0.046299, -0.003814\}$, $N = 1/3.19489$, with 1st root deteriorated (tSzaR = Φ1100), E = -2.8886 a.u. The overlap $\langle 0^{(1,2,1,1,1)} | 1^{(1,2,1,1,1)} \rangle = 0.0026$ =(Φ010H$|\Phi$110H).

"F_1" Root 8 Terms, $a_{1,2,1} = 0$

By minimizing F_1 (using the above fixed 1-term ϕ_0) first with $a_{1,2,1} = 0$, the same function is obtained, because the 1st root lies below ϕ_0 – and the lowest root above $E[\phi_0]$ is the 2nd root. [With $a_{1,2,1} = 0$ there are no other parameters to vary, and, as seen in the next subsection, with $a_{1,2,1} \neq 0$, a minimum of F_1 is obtained with the same ϕ_0 above; so, it is not necessary to use any lower (fixed) ϕ_0.]

$a_{1,2,1} \neq 0$

HUM 2nd Root, 8 Terms, $a_{1,2,1} \neq 0$

Using $a_{1,2,1} \neq 0$, the optimized 8-term HUM 2nd root $1^{(1,2,1,1,1)a}$, having $a_{1,2,1} = 1.5773 \times 10^{-7}$, remains essentially the same: (tSzaR = Φ111H) $E = -2.14449$, $z_{1,1} = 1.86703$, $z_{1,2} = 1.70272$. The main orbitals resemble **1s1s′** (cf. Fig. 13.5). The overlaps with the above established $\psi_0 = 0^{(1,2,2,2,2)}$ and $\psi_1 = 1^{(1,2,2,2,2)}$ are $\langle \psi_0 | 1^{(1,2,1,1,1)} \rangle = 0.0033$, and $\langle \psi_1 | 1^{(1,2,1,1,1)} \rangle = 0.9986$.

Generally, in a complete (**infinite**) space, both HUM and F should yield the same wave function, but in a truncated space, the 2nd HUM root may be worse than the lowest root above the known (highest) $E[\phi_{n-1}]$. Indeed, in 8-term He$_2$ ^{1}S, *at* the F_1 minimum, the lowest three roots above the known 1-term $E[\phi_0] = -2.824$ are $\{-2.145, -2.028, -1.898\}$, whereas in optimizing the 2nd HUM root, the lowest three HUM roots are $\{-2.889, -2.144, -1.993\}$ (As seen later in detail, F_1 yields $E[\phi_1] = -2.1452$, much closer to the exact -2.1459 than the 2nd HUM root $E\left[\phi_1^{2r}\right] = -2.144$, while its 1st HUM root, $E\left[\phi_1^{1r}\right] = -2.889$, is much deteriorated).

Obviously, fulfilling the inherent restrictions of the HUM functions [84], the 8-term 2nd HUM root $1^{(1,1,1,1,1)}$ has lower quality than $0^{(1,1,1,1,1)}$ ($\approx 0^{(1,2,1,1,1)}$) since $0.875^2 < (1 - (1.3 \times 10^{-5})^2)$, but also the optimized HUM 2nd root $1^{(1,2,1,1,1)a}$ has lower quality as well, since $0.9986^2 < 1 - (1.3 \times 10^{-5})^2$ – assuming near orthogonality (0.0026) to the optimized HUM 1st root.

"F_1"Root 8 Terms, $a_{1,2,1} \neq 0$

In contrast, by minimizing F_1, with the same set of parameters, $(z_{1,1}, z_{1,2}, a_{1,2,1})$, a much better 8-term approximant $1^{(1,2,1,1,1)F,a}$ is obtained: (tSzaR = $\Phi111F$) $E = -2.145152$, $F = -2.145151$, $F - E = 5.3 \times 10^{-7}$, $z_{1,1} = 1.938718$, $z_{1,2} = 1.817736$, $a_{1,2,1} = 0.799286$, $\{c\} = \{1., 0.196506, 0.086316, -0.026204, 0.043226, 0.016366, 0.015810, 0.002479\}$, $N = 1/3.637910$, **where the main orbitals resemble 1s2s** [cf. Figure 13.5]. The overlap with the optimized 8-term HUM 1st root is $\langle 0^{(1,2,1,1,1)} | 1^{(1,2,1,1,1)F,a} \rangle = 0.0186 = (\Phi010H | \Phi111F)$. The corresponding overlaps with the above-established basis are $\langle \psi_0 | 1^{(1,2,1,1,1)F,a} \rangle = 0.00490$, and $\langle \psi_1 | 1^{(1,2,1,1,1)F,a} \rangle = 0.999807$. Clearly, the F_1-minimizing function, although truncated (to 8 terms), is much closer to the exact saddle point. This is also depicted in Fig. 13.7, where *it satisfies the Schrödinger equation better* ($-0.0003 < \text{Schr} < 0.0009$) *than the 8-term HUM 2nd root* ($0.0006 < \text{Schr} < 0.0012$).

13.6.1.4   The Main Orbitals: F_1 (and OO) Give 1s2s; HUM Gives 1s1s′

Observe that the HUM 2nd root's main orbitals are **1s1s′** and it is "corrected" by the c-series, the (OO) function, although veered away and below the saddle point (thus, abandoned) has correct main orbitals 1s2s (cf. Eq. 13.18 ff), whereas F_1 directly finds physically correct main orbitals **1s2s** (and better energy in truncated space). If this HUM phenomenon emerges already in a system of two-electrons, or three-electrons as foresaid and will be seen for Li (cf. Table 13.1), there is no guarantee (and might be more important) in **large systems**, where **small truncated space is unavoidable** and where the nature of the main (HOMO/LUMO) orbitals is decisive in order to **predict** or **cause** desired reactions; **then, the F functional will be needed**. The above energies and overlaps are summarized in Table 13.3.

13.6.1.5   Fulfillment of the Saddle Point Criteria by F_1 and Immediate Improvement of Φ_0

The functions obtained by F_1, fulfill the saddle point **criteria** mentioned in the introduction:

Firstly, to check ***the 27-term approximant ψ_1***, a 2×2 $(H_{ij} - ES_{ij})$ generalized diagonalization between ψ_1 and the 1-term Φ_0 indeed leaves ψ_1 practically unaffected, as "high 2×2 root" $\phi_1 = (\psi_1 - 4.7 \times 10^{-5} \Phi_0)$, with energy changed by only 1.57×10^{-9} a.u, while, by opening their energy "gap," improves Φ_0

to the "low 2×2 root" $\phi_0 = (0.999\ \Phi_0 - 0.048\ \psi_1)$, with (indeed lower than $E[\Phi_0] = -2.84766$) energy: $E[\phi_0] = -2.84926$ a.u. differing from the energy of *the exact orthogonal to* ψ_1 (in the $\{\Phi_0, \psi_1\}$ subspace [cf. Eq. 13.13]) by -1.57×10^{-9} a.u. (which means that the "low 2×2 root" ϕ_0 is actually orthogonal to ψ_1). This verifies that ψ_1 is indeed very close to the exact excited saddle point eigenfunction and that ϕ_0, in order to further approach ψ_0, needs be rotated only orthogonally to ψ_1.

Similarly, a 2×2 diagonalization between *the F_1 27-term Φ_1* and the 1-term Φ_0 yields: $\phi_1 = (\psi_1 - 1.4 \times 10^{-4}\ \Phi_0)$, $E[\phi_1] - E[\Phi_1] = -1.34 \times 10^{-8}$ a.u., $\phi_0 = (0.999\ \Phi_0 - 0.048\ \psi_1)$, with lower energy: $E[\phi_0] = -2.84934$ a.u. differing from the energy of *the exact orthogonal to* Φ_1 (in the $\{\Phi_0, \Phi_1\}$ subspace [cf. Eq. 13.13]) by -1.34×10^{-8} a.u., which verifies that the F_1 27-term Φ_1 is also very close to the exact saddle point, as expected.

Now, to check F_1, a 2×2 diagonalization between **the 8-term Φ_1** and the 1-term Φ_0 indeed leaves Φ_1 also almost unaffected, as "high root" $\phi_1 = (\psi_1 - 6.2 \times 10^{-4}\ \Phi_0)$, $E[\phi_1] - E[\Phi_1] = -2.69 \times 10^{-7}$ a.u., and, by opening their "gap," improves Φ_0 to the "low root" $\phi_0 = (0.999\ \Phi_0 - 0.052\ \psi_1)$, with lower energy: $E[\phi_0] = -2.84957$ a.u. differing from the energy of *the exact orthogonal to* Φ_1 (in the $\{\Phi_0, \Phi_1\}$ subspace [cf. Eq. 13.13]) by -2.69×10^{-7} a.u., which means that the "low 2×2 root" ϕ_0 is practically orthogonal to Φ_1 and verifies that Φ_1 $1s2s$ is indeed very close to the exact excited saddle point eigenfunction; ϕ_0 could still be further improved by rotating it orthogonally to Φ_1.

Repeating the check with *the 8-term HUM 2nd root* $1s1s'$ and the 1-term Φ_0, yields: $\phi_1 = (0.999\ \psi_1 - 1.2 \times 10^{-3}\ \Phi_0)$, $E[\phi_1] - E[\Phi_1] = -1.02 \times 10^{-6}$ a.u., $\phi_0 = (0.999\ \Phi_0 - 0.049\ \psi_1)$, with lower energy: $E[\phi_0] = -2.84937$ a.u. differing from the energy of *the exact orthogonal to* Φ_1 (in the $\{\Phi_0, \Phi_1\}$ subspace [cf. Eq. 13.13]) by -1.02×10^{-6} a.u., which verifies that 8-term HUM 2nd root $1s1s'$ is indeed **inferior** than the F_1 8-term $1s2s$ function.

Secondly, as shown in Fig. 13.8, F_1 **has local minimum**, always above the energy, while (*thirdly*) the energy $E[\Phi_1]$ is a **saddle point at the minimum of F_1**, showing that it has approached the exact excited state $1s2s$. For some parameters (z, c, etc.), the energy has local minimum, while for others ($a_{1,2}$) it has local maximum.

13.6.1.6 Quick Check of Reasonableness via the Main Orbitals and via 2nd Derivatives

Figure 13.5 shows the main orbitals of both the 27-term and the 8-term optimized wave functions of the HUM 2nd root and also of ϕ_1^{LE} (the lowest orthogonal to ϕ_0) and of F_1 excited state solutions, compared to the ground state. Observe that the HUM answer is $1s1s'$, instead of $1s2s$ (because $a_{1,2} \approx 0$, as mentioned above), and the HUM 8-term function deviates from the 27-term ψ_1, as anticipated in the introduction (also the HUM $1s$ orbital is, unexpectedly, slightly more *diffuse* than $1s^2$); therefore the HUM function is veered away from the exact, whereas the F_1 answer is indeed $1s2s$, the F_1 8-term function already almost coincides with the

Fig. 13.8 The 27-term
energy and the F_1 values
(a.u.) near the minimum of
F_1, versus: Top: a z-parameter
($z_{1,2}$). Middle: an
$a \neq 1$-parameter ($a_{1,2}$).
Bottom: a c-parameter
($c_{0,0,1}$). Always: $F_1 \geq E_1$
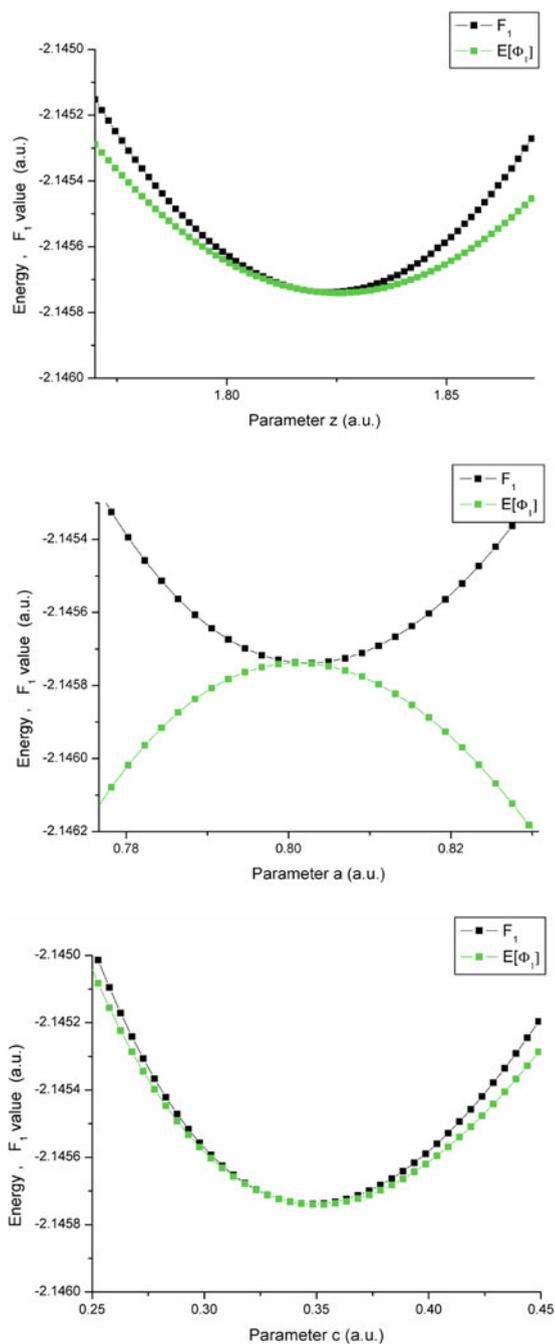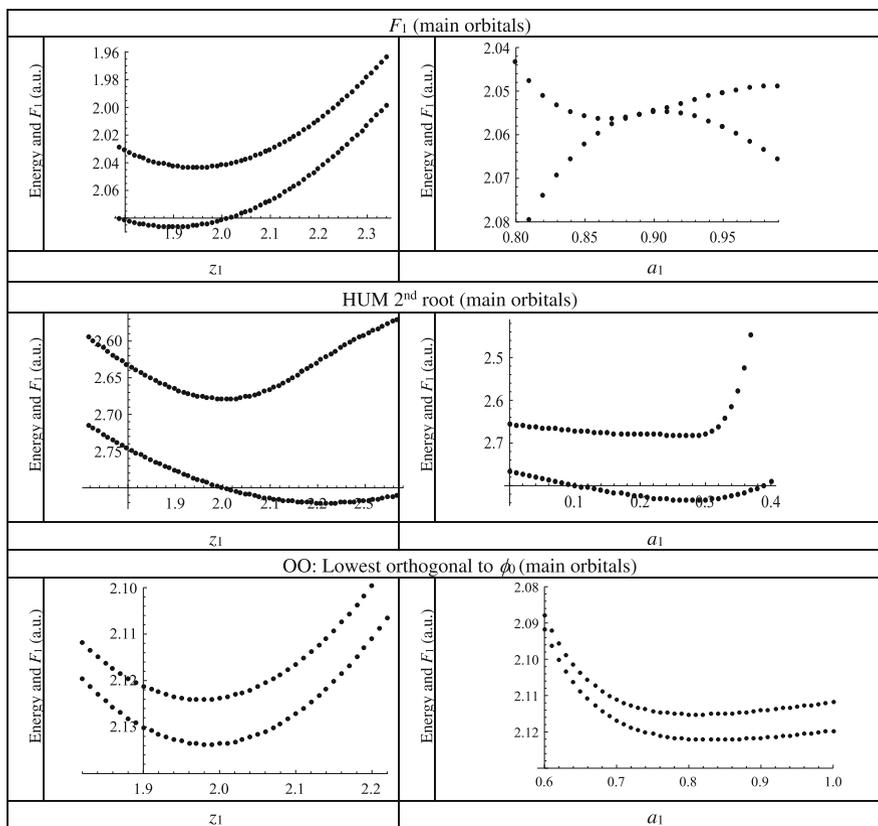
Fig. 13.9 The behavior of F_1 (upper curves in each frame) above the energy due to the *main-orbitals* (lower curves in each frame) (in a.u.) in varying z_1 (left frames) and a_1 (right frames). Upper frames: F_1 $1s2s$ approximant. Middle frames: HUM $1s1s'$ approximant. Lower frames: OO: the lowest orthogonal to ϕ_0, whose behavior (at the minimum) was expected, i.e., both F_1 and energy have minimum. But note that the HUM energies (and F_1) also have minimum (not saddle)

exact, and the F_1 $1s$ orbital is indeed more *compact* than $1s^2$, pushed toward the nucleus by the $2s$ electron, as intuitively expected.

Using F_1 we can quickly decide about the quality of the main orbitals. If the function of the main orbitals is close to the saddle point, then F_1 will have a minimum above it. Fig. 13.9 shows that F_1 behaves "reasonably" around the F_1 $1s2s$ 8-term main orbitals (i.e., without the series expansion), but neither HUM $1s1s'$ 8-term nor the 27-term (OO) lowest orthogonal to ϕ_0 behave "reasonably."

We shall see that similar results hold also for Li (cf. Table 13.1) where the 2nd derivative is also displayed, (positive for HUM, negative with respect to one parameter for F_n).

### 13.6.1.7   The Energy of ϕ_1^+ (the Closest to ψ_1 Orthogonal to ϕ_0)

Now, by using the above-established Φ_1^{2r} as ψ_1 and a truncated approximant of ψ_0 as ϕ_0, it is demonstrated that the corresponding $|\phi_1^+\rangle$ [cf. Eq. 13.3], i.e., the closest to ψ_1 while orthogonal to ϕ_0, lies lower than $E[\psi_1]$ (actually, than $E[\Phi_1^{2r}]$). Since the 8-term HUM 1st root $0^{(1,1,1,1,1)}$ is, as seen above, "almost perfect," $\langle\phi_0|1\rangle^2 = 10^{-10}$, $E[\phi_1^+]$ is lower than $E_1 = -2.14584$ by only 10^{-10} a.u., well beyond the accuracy of the present demonstration, so, a less accurate ϕ_0 will be used in the demonstration, e.g., the 1-term "fixed," used in F_1, with $\langle\phi_0|1\rangle = 0.04786$; then by replacing ψ_1 by Φ_1^{2r} in $|\phi_1^+\rangle$: $\langle\phi_1^+|H|\phi_1^+\rangle = -2.15470$ a.u., whereas by replacing it in the exact formula [cf. Eq. 13.3]: $E[\phi_1^+] = -2.14745$ a.u. The 0.3% discrepancy is due to the inexactness of Φ_1^{2r}. We expect that if the energy were further minimized while keeping orthogonality to this 1-term $\phi_0 = 0^{(0,1,0,0,0)}$, we should end up with a function orthogonal to this ϕ_0, lying lower than -2.15470 a.u. (if this were the correct value based on the exact ψ_1), veered well away from ψ_1. This is demonstrated below. If the deteriorated HUM 1st root, orthogonal to the optimized HUM 2nd root $1^{(12111)a=0}$, is used as ϕ_0, with $\langle\phi_0|1\rangle = -0.00405$, then the corresponding $\langle\phi_1^+|H|\phi_1^+\rangle = -2.14591$ a.u. (and by the exact Eq. 13.3: $E[\phi_1^+] = -2.14585$ a.u.), which is slightly below $E[\Phi_1^{2r}]$, while the exact eigenvalue is slightly even lower, which confirms that it is not necessary for the 1st root to approach ψ_0, $\phi_0 \to \psi_0$, although deteriorated, **only orthogonality to ψ_1, $\langle\phi_0|\psi_1\rangle \to 0$, is adequate**.

13.6.1.8   Comparison Between F, HUM, and the Lowest Orthogonal to ϕ_0 (OO)

Up to now it has been demonstrated that the optimized HUM 2nd root Φ_1^{2r} is veered away from the exact saddle point because it must be orthogonal to a deteriorated approximant of the ground state while lying higher than the exact (in accordance with the HUM theorem). Also, previously it was demonstrated that (OO) minimizing the excited state energy orthogonally to a normalized *approximant* ϕ_0, of ψ_0, leads to a function, ϕ_1^{LE}, lying lower than the exact (not collapsed), therefore, also veered away from the exact. Thus the F_1-minimizing function, $\phi_1^{F_1}$, is the most reliable truncated approximant among the three $\{\phi_1^{F_1}, \phi_1^{LE}, \text{and } \Phi_1^{2r}\}$.

13.6.1.9   Remarks

1. If $\langle 1|\phi_0\rangle = 0$, then the three functions $\{\phi_1^{F_1}, \phi_1^{LE}, \text{and } \Phi_1^{2r}\}$ coincide with ψ_1 and ϕ_1^+, because the orthogonal to ϕ_0 subspace will contain ψ_1; hence ϕ_1^+ will coincide with ψ_1: not lying lower. Thus the minimizing function ϕ_1^{LE} will coincide with ϕ_1^+ and, therefore, with ψ_1. Also, Φ_1^{2r}, having ϕ_0 as 1st root, will
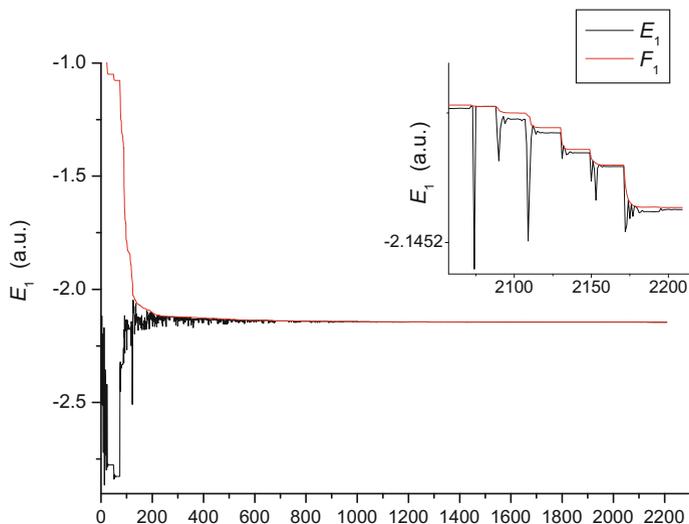
Fig. 13.10 The various (E,F) values visited in F_1-minimization, sorted by descending F_1. Observe that many E values are deceptively low, having large F_1 (the first points). Those having good F_1, e.g., the 2074th or the 2010th zoomed in the inset, may have acceptable wave function (see text). The correct is the last one

belong to the orthogonal subspace that contains ψ_1, and, since there is no other minimum than that of $\phi_1^{LE} = \psi_1$, it will coincide with ψ_1. Therefore, for the three functions to coincide the condition $\langle 1|\phi_0\rangle \to 0$ suffices; it is not necessary to meet the condition $\phi_0 \to \psi_0$.

2. Finding the minimum of F_n is necessary: Fig. 13.10 shows all computed 8-term energy values, $E[\phi_1]$, in this work along their F_1 values, sorted by F. Of course, the first points have large F and low E, and the last points tend to the minimum of F_1. Nevertheless, some of the last points, as shown in the inset, have reasonable F_1, but rather low $E[\phi_1]$. The correct point is the last one at the minimum of F_1, where the E and F values coincide.

3. Since E_n is a saddle point, some wave functions ϕ_n near the F_n minimum are still acceptable even if they lie below E_n, as long as they lie above $E_n - L$ [cf. Eq. 13.8], or, in general, above the convex combination of all lower eigenvalues up to E_n, which is of the form $(1 - \sum_i a_i)E_n + \sum_i a_i E_i$; if the expansion coefficients (weights $a_i = \langle i|\phi_n\rangle^2$) are small, or $L < \varepsilon$ (an acceptable tolerance), the point is near E_n and ϕ_n is acceptably near ψ_n [79]. The value of L, already at such accuracies near the F_n minimum, can be easily checked. This criterion of validity holds for any ϕ_n, not necessarily obtained via F_n. For example, ϕ_1^{LE}, above, has $L = (E_1 - E_0)\langle 0|\phi_1^{LE}\rangle^2 = 0.0018$, which is a quite large tolerance; therefore, ϕ_1^{LE} is rather unacceptable.

13.6.2 Demonstration of Identifying a "Flipped Root"

Let the ground and 1st excited state wave functions of a hydrogen-like ion be parametrized as

$$\psi_0(z_0; r) = a_0(z_0)\, e^{-Z\, r\, z_0}, \quad \psi_1(z_1, g; r) = a_1(z_1, g)\, e^{-z_1 Z\, r/2}\, (1 - g\, Z\, r/2)$$

where Z is the nuclear charge, z_0, z_1, g are variational parameters, and $a_0(z_0)$, $a_1(z_1, g)$ are normalization constants. These functions are not orthonormal, unless $z_0 = 1, z_1 = 1, g = 1$, when they form eigenfunctions of the Hamiltonian $H\psi_j(r) = -\psi_j''(r)/2 - \psi_j'(r)/r - Z\psi_j(r)/r$.

In their 2×2 subspace create an orthonormal basis

$$\Psi_0(r) = \psi_0(r), \quad \Psi_1(r) = (\psi_1(r) - \psi_0(r)\,\langle\psi_0|\psi_1\rangle)\,/\sqrt{1 - \langle\psi_0|\psi_1\rangle^2},$$

whose overlap matrix is $\delta_{i,j}$. In diagonalizing their Hamiltonian matrix $H_{i,j} = \int_0^\infty 4\pi r^2\, \Psi_i(r) H \Psi_j(r)\mathrm{d}r$, let the two normalized eigenfunctions be $\Phi^{1r}(r)$, $\Phi^{2r}(r)$, with their eigenvalues ("roots") depending on z_0, z_1, g.

Now, around $z_0 \approx 2$ a root crossing occurs for a wide range around $z_1 \approx 1$ and $g \approx 1$. Near and "before" the crossing the continuation of $\psi_0(z_0; r)$ is $\Phi^{1r}(r)$ and the continuation of $\psi_1(z_1, g; r)$ is $\Phi^{2r}(r)$, whereas "beyond" the crossing the continuation of $\psi_1(z_1, g; r)$ is $\Phi^{1r}(r)$ and the continuation of $\psi_0(z_0; r)$ is $\Phi^{2r}(r)$. The question is to decide, via F_1, whether a given value ("point") of (z_0, z_1, g), near the crossing, is "before" or "beyond" the crossing, in order to use the continuation, ε, of (always) $E[\psi_1(z_1, g; r)]$ in an optimization algorithm. (In the present demonstration Newton-Raphson (NR) is used: $\boldsymbol{\varepsilon}' := (\partial\varepsilon/\partial z_0, \partial\varepsilon/\partial z_1, \partial\varepsilon/\partial g) = 0$ is solved by proceeding iteratively to a new point $\mathbf{p} + \delta\mathbf{p} = \mathbf{p} - \mathbf{J}^{-1} \cdot \boldsymbol{\varepsilon}'$ – or less if the method diverges – having started at some point $\mathbf{p} = (z_0, z_1, g)$, where $\mathbf{J}$ is the Jacobian matrix, Hessian of ε [85].)

Thus, consider

$$F_1\left[\phi_0; \Phi^{nr}\right] = E\left[\Phi^{nr}\right] + 2\frac{(\langle\phi_0|\, H\, |\Phi^{nr}\rangle - E\left[\Phi^{nr}\right]\langle\phi_0|\Phi^{nr}\rangle)^2}{(E\left[\Phi^{nr}\right] - E\left[\phi_0\right])\left(1 - \langle\phi_0|\Phi^{nr}\rangle^2\right)}, \quad n = 1, 2$$

where the fixed predetermined (deliberately not very accurate) ground state approximant is $\phi_0\left(z_0^0; r\right) = a_0\left(z_0^0\right)e^{-Z\, r\, z_0^0}$ with $z_0^0 = 1-0.05$. (By direct minimization, $F_1[\phi_0; \Phi^{2r}]$ is minimized to $F_1 \to E[\Phi^{2r}(r)] = E[\psi_1(1, 1; r)] = -0.125$ at $z_1 = 1$, $g = 1$, ($z_0 \approx 1$).) As explained above in the theory (cf. Eq. 13.15), among the two "roots," the continuation of the excited state, near the crossing, is the one with the lowest F_1. Indeed, for $Z = 1$, using (first traditionally) blindly the "2nd root," i.e., keeping ε to $E[\Phi^{2r}(r)]$, i.e., regardless of which n (n^{th} root) the lowest F_1 suggests, "root flipping" occurs: The sequence of Table 13.4 is obtained, even by using half NR-step.

Table 13.4 Root flipping: no convergence of the 2nd root

			n	E^{2r} (a.u.)
2.7	0.95	0.8	1	1.09915
1.89975	0.657935	0.618574	1	0.0312309
1.79561	0.457167	0.5843	2	−0.05816
1.81109	0.303372	0.471712	2	−0.0491109
1.88811	0.0253053	0.0834495	2	−0.0056792
−4.64705	1.13597	3.82639	1	14.4125
−4.64705	1.13597	3.82639	1	14.4125
−1.79846	0.463128	1.53227	1	3.19065
0.416586	−0.117902	−0.379791	2	0.0138916
0.647726	−0.168901	−0.60433	2	0.0232773
0.913718	−0.21854	−0.88652	2	0.0336386
1.15883	−0.254786	−1.19969	2	0.0428655
...	...	...	...	...

Table 13.5 Same as Table 13.4, but by consulting F_1

z_0	z_1	g	n	ε (a.u.)
2.7	0.95	0.8	**1**	−0.142331
2.62754	0.95588	0.907362	1	−0.127151
2.37234	0.971824	0.98181	1	−0.124935
1.92938	0.992178	1.00683	1	−0.125094
1.3233	1.02512	0.953813	**2**	−0.124875
1.18358	1.01644	0.958059	2	−0.12496
1.07193	1.01041	0.964645	2	−0.124988
0.980565	1.00646	0.97107	2	−0.124997
0.905827	1.00399	0.976612	2	−0.124999
0.844753	1.00248	0.981153	2	−0.125

On the contrary, by consulting F_1 the continuation of the excited state is recognized near the crossing and is used (until finally, at convergence, only $n = 2$, the 2nd root, is suggested by the lowest F_1) [cf. Table 13.5].

Observe that at the beginning, "beyond" the crossing, the lowest F_1 dictates to use, for the next step, the (**lower than E_1**) value of $\varepsilon = E[\Phi^{1r}(r)]$ ($n = 1$, the lowest function at that point).

Similarly, using blindly only the 2nd root (regardless of which n is dictated by the lowest F_1) and starting again "beyond" the crossing ($n = 1$), Table 13.6 is obtained. In this case, despite the original irregularities due to root flipping, the 2nd root finally happened to remain "before" the crossing ($n = 2$) and converged (0.3 of NR-step was used.)

By starting from the same point but by consulting F_1 no irregularities occurred [cf. Table 13.7].

Note that finally, near the minimum of the 2nd root where $E[\Psi_1(\mathbf{p}_1)] > E[\Psi_0(\mathbf{p}_1)]$, "before" the crossing, the convergence *should* use the 2nd root. If (and when), while in $n = 1$, it were stuck approaching a point of the 1st root, i.e., "beyond" the crossing, the NR-step should be increased somewhat in the direction of the last

Table 13.6 Accidental convergence of the 2nd root

z_0	z_1	g	n	E^{2r} (a.u.)
2.7	1.2	1.1	1	0.99477
2.20242	0.940034	0.917342	1	0.256091
1.89495	0.731999	0.818135	1	−0.0551044
1.58749	0.523964	0.718929	2	−0.0866029
−3.05737	−10.9139	−12.3444	1	69.042
−2.07399	−7.38161	−8.30417	1	33.6488
−1.38465	−4.90769	−5.47902	1	16.3024
−0.695314	−2.43378	−2.65387	1	5.07481
−0.695314	−2.43378	−2.65387	1	5.07481
−0.414382	−1.43604	−1.53652	1	2.2856
−0.214196	−0.731385	−0.76036	1	0.910672
−0.0140099	−0.0267279	0.0157957	2	0.0214185
0.164968	0.31467	−0.18661	2	−0.0598432
0.326748	0.626253	−0.334901	2	−0.103163
0.401263	0.771636	−0.381191	2	−0.116354
...	...	...	...	...
0.518132	1.00162	−0.429903	2	−0.124998
0.516411	0.998281	−0.42876	2	−0.125

Table 13.7 Same as in Table 13.5, but by consulting F_1

z_0	z_1	g	n	ε(a.u.)
2.7	1.2	1.1	**1**	−0.119716
1.93961	1.03415	0.976157	1	0.126188
1.76764	0.886719	1.08156	**2**	−0.121279
1.61807	0.892447	1.16453	2	−0.121751
1.41702	0.918759	1.21697	2	−0.123372
1.20892	0.9443	1.21744	2	−0.124476
1.04199	0.963778	1.18615	2	−0.12487
0.924215	0.977699	1.15139	2	0.124969
0.841179	0.986541	1.12222	2	0.124992
0.780975	0.991767	1.09888	2	0.124998
0.735948	0.994853	1.0804	2	−0.124999
0.701294	0.996712	1.06577	2	−0.125

Table 13.8 Overpassing a false convergence

z_0	z_1	g	n	E (a.u.)
2.5	1.2	0.9	1	−0.138096
2.62454	1.03587	0.935687	1	−0.127460
2.48591	0.989487	0.976356	1	−0.125154
2.34727	0.943106	1.01703	**1**	−0.124859
1.41343	1.02160	0.963025	**2**	−0.124868
1.25337	1.01623	0.964009	2	−0.124954
1.12615	1.01049	0.968734	2	−0.124986
1.02330	1.00655	0.974106	2	−0.124996
0.939738	1.00405	0.978974	2	−0.124999
0.871820	1.00250	0.983044	2	−0.125

"false"-converged $\delta\mathbf{p}$ in order to send it farther to $n = 2$, i.e., "before" the crossing where the real minimum should be [cf. Table 13.8]. (Such a "false" convergence should always be checked). While being in $n = 1$, it approached p = (2.34, 0.94, 1.01); in order to send it farther to $n = 2$ "before" the crossing, p was extrapolated every three steps by twice the final step – or more if convergence is slow, according to the logical code exposed in Table 13.2. Such an extrapolation was also used in the first example above; in the second it was not needed.

Note that the graphs of all converged functions, above,

$$\Phi^{2r}(r) = -0.0074983\, e^{-0.844753\, r} + (0.206437 - 0.101273\, r)\, e^{-0.50124\, r},$$
$$\Phi^{2r}(r) = 0.615983\, e^{-0.516411\, r} - (0.416497 + 0.0892886\, r)\, e^{-0.499141\, r},$$
$$\Phi^{2r}(r) = 0.0178076\, e^{-0.701294\, r} + (0.182078 - 0.0970265\, r)\, e^{-0.498356\, r},$$
$$\Phi^{2r}(r) = -0.00706569\, e^{-0.87182\, r} + (0.205952 - 0.10123\, r)\, e^{-0.50125\, r}$$

are practically identical to the exact

$$\psi_1(1, 1; r) = (1 - r/2)\, e^{-r/2}/\sqrt{8\pi} = (0.199471 - 0.0997356\, r)\, e^{-r/2};$$

they differ by at most 0.001 at $r = 0$ (the second differs at most by $1.5\ 10^{-5}$).

13.6.3 *Application to Conventional Configuration Interaction Treatment*

The functional F_n has been preliminarily [2, 66] used for the computation of atomic excited states in standard coordinates (r, θ, φ) by configuration interaction (CI), where the configurations are symmetry-adapted linear combinations of Slater determinants (SD) composed of analytic Laguerre-type orbitals, whose polynomial prefactors are variationally optimized (AVOLTOs), thus providing small, concise, and comprehensible analytic wave functions, with accuracy comparable to numerical MCSCF [79]:

$$\langle r|n, \ell, m \rangle = A^{n,\ell,m} L_{n,\ell}(r) Y^{\ell,m}(\theta, \phi),$$

where $A^{n,\ell,m}$ are normalization factors, $Y^{\ell,m}(\theta, \phi)$ are spherical harmonics, and $L_{n,\ell}(r)$ are the AVOLTOs,

$$L_{n,\ell}(r) = \sum_{k=0}^{n-\ell-1} c_k^{n,\ell} g_k^{n,\ell} r^{(\ell+k)} e^{-z_{n,\ell}\frac{r}{n}} + b^{n,\ell} e^{-q^{n,\ell} z_{n,\ell}\frac{r}{n}} \delta_{\ell,0}. \qquad (13.19)$$

Here $c_k^{n,\ell}$ are the usual associated Laguerre polynomial coefficients, $z_{n,\ell}$, $b^{n,\ell}$, $q^{n,\ell}$ are variational parameters, and $g_k^{n,\ell}$ are factors determined by orthogonalization to desired, only, orbitals, i.e., by solving (for $g_k^{n,\ell}$) the relations $\langle n_i, \ell, m | n_j, \ell, m\rangle = \delta_{i,j}$, so that not all orbitals are mutually orthogonal. Therefore the general formalism is non-orthogonal, which allows spin unrestricted as well as open shell computation.

Thus, for an atom with nuclear charge Z_{nuc} and N electrons, with space and spin coordinates $\mathbf{r}_1 s_1, ..., \mathbf{r}_N s_N$, and for a given symmetry type and electron occupancy, the desired $N -$ electron normalized wave function, of the nth excited state, consisting of N_{conf} (predetermined) configurations, out of N_{det} SDs, is written as

$$\phi_n\,(\mathbf{r}_1 s_1, \ldots, \mathbf{r}_N s_N) = \sum_{p=1}^{N_{\text{conf}}} d_p \sum_{a=1}^{N_{\text{det}}} f_{p,a} D_a; \quad |\phi_n|^2 = 1 \qquad (13.20)$$

and is obtained by minimization of F_n, with Hamiltonian

$$H = -\frac{1}{2} \sum_{i=1}^{N} \left(\nabla_i^2 + \frac{Z_{\text{nuc}}}{|\,r_i\,|} \right) + \sum_{i>j}^{N} \frac{1}{|\,r_i - r_j\,|} \equiv \sum_{i=1}^{N} h_i + \sum_{i>j}^{N} g_{i,j}.$$

In Eq. 13.20, if N_{det} is the number of SDs, D_α, consistent with the desired electronic state, N_{orb} is the predetermined number of spinorbitals, a_i, under optimization, and $f_{p,a}$ are all ($N_{\text{conf}} \times N_{\text{det}}$) consistent corresponding coefficients, then, the linear parameters d_p are determined from a desired root of the secular equation ($N_{\text{conf}} \times N_{\text{conf}}$) with ($p,q$) matrix elements

$$\sum_{a,b=1}^{N_{\text{det}}} f_{p,a} f_{q,b} \langle D_a\,|H - E|\,D_b\rangle,$$

and ϕ_n is a critical point of both the energy and F_n, while all $\phi_i\ i < n$, remain unvaried, rather crude, approximants of the lower states.

The one- and two-electron terms between SDs are computed by

$$\langle D_a| \sum_{i=1}^{N} h_i\,|D_b\rangle = \frac{1}{\sqrt{D_{aa} D_{bb}}} \sum_{i,j=1}^{N} \langle a_i\,|h|\,b_j\rangle D_{ab}\,(a_i b_j)$$

where $D_{ab} = \det\,|\langle a_1|b_1\rangle\langle a_2|b_2\rangle\cdots\langle a_n|b_n\rangle|$, a_i (and b_j, etc.) are the spin-orbitals, and where $D_{ab}(a_i b_j)$ is the cofactor of the element $\langle a_i|b_j\rangle$ in the determinant D_{ab}, and D_{aa}, D_{bb} are similar normalization factors. Also,

$$\langle D_a| \sum_{i>j}^{N} g_{i,j} |D_b\rangle = \frac{1}{\sqrt{D_{aa} D_{bb}}} \sum_{j>l}^{N} \sum_{i>t}^{N} \langle a_i a_t |g| \, b_j b_l\rangle \, D_{ab} \left(a_i a_t b_j b_l\right)$$

where $D_{ab}(a_i a_t b_j b_l)$ is the cofactor of D_{ab} defined by deleting the rows and columns containing $\langle a_i | b_j\rangle$ and $\langle a_t | b_l\rangle$ and attaching a factor $(-1)^{i+j+t+l}$ to the resultant minor. The determinant cofactors are most efficiently computed via the inverse matrix, after adding to all matrix elements small random numbers of the order of machine accuracy, in order to avoid occasional (but harmless) vanishing of the determinant.

13.6.3.1   Results

He

To check the reliability of the CI computation, before demonstrating comparisons between F_n and HUM – small and large – expansions, a comparison with the literature is shown:

For the 1st excited state of He ^1S $1s2s$, using, in ϕ_1, 11 AVOLTOs forming 76 SDs and 22 configurations and a fixed crude ϕ_0 approximant of $1s^2$ of only two AVOLTOs in a 2×2 CI ($E[\phi_0] = -2.88$ a.u.), then $E[\phi_1] = -2.1458140$ a.u. ($F_1 = -2.1458139$ a.u.). The wave function is primarily $\phi_1 = 0.9993$ ($1sB\,2sA -$ $1sA\,2sB$) $+ 0.0190$ ($2p_1A\,3p_{-1}B - 2p_1B\,3p_{-1}A - 2p_0A\,3p_0B + 2p_0B\,3p_0A + 2p_{-1}A$ $3p_1B - 2p_{-1}B\,3p_1A$) $+ 0.0178\,3s^2$, where A,B denote the spin, the {r_{rms} distance from the nucleus (a.u.); and $z_{n,\ell}$, $b^{n,\ell}$, $q^{n,\ell}$ values} are, respectively, for $1s$: {0.8699666; 1.9690983, 0.0881525, 1.2060984}, for $2s$: {5.6429592; 1.1248561, 0, 1}, for $2p$: {1.1438142; 4.7885604, 0, 1}, and for $3s$: {1.0879814; 5.6155136, 0, 1}, the g_k-factors are for $2s$: (0.8186816, 1) and for $3s$: (4.0287651, 2.0322568, 1), making them both orthogonal to another $1s$, a little more diffuse: {1.0374887; 3.0931600, -0.9341100, 0.9519696}. This energy value can be compared to (i) -2.1457316 a.u., with 10 AVOLTOs forming 68 SDs and 18 configurations, exactly orthogonal to a quite accurate ϕ_0 of $E = -2.9031501$ a.u. with 15 AVOLTOs, 157 SDs, and 40 configurations and to (ii) -2.145873 a.u. using 10 numerical MCHF orbitals, with a comparably highly accurate ground state of $E = -2.9031173$ E_h [86].

For the 2nd excited state of He ^1S $1s3s$, using, in ϕ_2, 11 AVOLTOs, 77 SDs, 23 configurations, and again 2×2 ϕ_0 and ϕ_1, then $E[\phi_2] = -2.0612263$ a.u. ($F_1 = -2.0611758$ a.u.). The wave function is primarily $\phi_2 = 0.9788$ ($1sA\,3sB -$ $1sB\,3sA$) $+ 0.2035$ ($3sA\,2sB - 3sB\,2sA$) $+ 0.0170$ ($1sA\,2sB - 1sB\,2sA$); the {r_{rms}; z, b, q} values are for $1s$: {1.1051964; 1.61507, 2.0564, 0.95729}, for $3s$: {12.9875477; 1.09458, 0, 1}, and for $2s$: {1.6490337; 3.49456, 0, 1}; the g_k-factors are for $3s$: (0.7660135, 0.9277702, 1), making it orthogonal to $1s$ and to the previous $2s$, and for $2s$: (1.5813926, 1), making it orthogonal to $1s$. This energy value can be compared to -2.0612681, obtained by *B-splines* [87]. By increasing to 19 AVOLTOs, 263
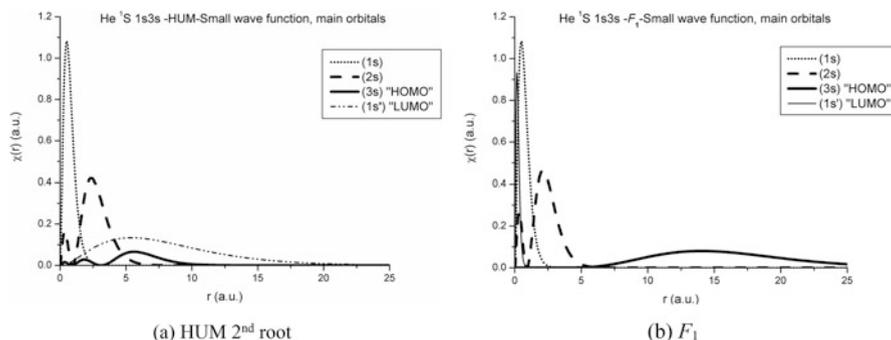
Fig. 13.11 CI wave functions (main orbitals): for **He^{1}S1s3s**. Dotted: 1s. Dashed: 2s. Solid: 3s. The HUM main orbitals (**a**) are, incorrectly, **1s,1$\underline{s}'$**, whereas the F main orbitals (**b**) are, correctly, 1s3s. Notice that the "LUMO" is just a static correlation orbital very close to the nucleus, not an excitation orbital

SDs, 53 configurations, the improved energy is $E[\phi_2] = -2.0612522$ a.u. ($= F_1$) [88].

Therefore, the CI computation is reliable. To compare with HUM, as "exact" ψ_n a "large" expansion in **1s, 2s, 3s, 4s, 5s, 2p, 3p, 4p, 5p, 3d, 4d, 5d, 4f, 5f** was used. For ^1S: $E_0 \approx -2.90324$ a.u., $E_1 \approx -2.14594$ a.u., $E_2 \approx -2.06125$ a.u. (exact: -2.06127 a.u [80].), for ^3S: $E_0 \approx -2.17521$ a.u. (compared to -2.17516 of ref. [89]), $E_1 \approx -2.06869$ a.u. (exact: $-2.17536, -2.06881$ a.u [80]). As "truncated" trial functions ϕ_n a "small" expansion in **1s, 2s,** and **3s** was used.

He ^1S

He ^1S 1s2s (1st Excited State ϕ_1)

As found above with Hylleraas coordinates, for He ^1S 1s2s the "small" HUM and F functions, in lack of other parameters, are essentially the same. There is nothing interesting to compare.

He ^1S 1s3s (2nd Excited State ϕ_2)

For the singlet **He ^{1}S 1s3s**, the comparison is shown in Fig. 13.11. F "small" is, correctly, mainly 1s3s, whereas HUM "small" is, **misleadingly**, mainly 1$s$$\underline{1s}'$, instead of 1$s3s$. Notice that the lowest unoccupied orbital, "LUMO," acts as a static correlation orbital, improving the *total* wave function *near the nucleus*. By no means should it be considered as the first candidate orbital to be occupied by an excited electron. For details see Ref. [90].

He ^3S

He ^3S 1s3s (1st Excited State ϕ_1)

The lowest state of He ^3S is 1s2s, obtainable by simply minimizing the energy. For the 1st excited state ϕ_1, as seen in Fig. 13.3 (a), again it is demonstrated that "large" functions, HUM and *F*, are practically equivalent. The *F* "small" function (b) has the same main orbitals, 1*s* and 3*s*, as both the "large" functions *F* (a) and HUM (d), where the 2*s* just adds some static correlation correction near the nucleus (as well as all higher orbitals of the "large" expansion (c)). But the HUM "small" expansion (e), orthogonal to a deteriorated 1st root ^3S "1*s*2*s*" (f), **has main orbitals 1*s*2*s*** (with opposite sign), while the 3*s* orbital acts as a static correlation correction of the total wave function, to approach the correct energy. Thus, the HUM solution **incorrectly** predicts, as "HOMO" orbital, the 2*s* instead of the 3*s*; therefore, it **is misleading**. Of course, blindly considering as "LUMO" the 1st higher "unoccupied" orbital, i.e., the HUM 3*s*, is completely out of question.

The above conclusions are clearly confirmed by Fig. 13.4, where all HUM functions have minimum, whereas all *F* functions, either "large" or "small," have energy saddle point. (Of course, the larger the expansion, the better is the approximation, but for larger systems this is not feasible; rather "small" but reliable functions are needed.)

Li

Similar results are found for Li: Several low-lying excited states of Li ^2S and ^4S are shown in Table 13.1, where the main configuration, the energy *E* and the F_n value are displayed, along with the RMS extent of the main orbitals. The low-lying HUM "small" functions were misleading. However, by appropriately changing the basis, guided by the F_n functional, much better HUM "small" functions were found, used for final comparisons with the F_n functions. This, as well as the corresponding aforementioned results for He, suggests that small expansions for excited states would be more correctly described in terms of more appropriate basis functions. The last (9th) column shows the number of configurations along with the overlap $\langle HUM|F_n \rangle$ of the corresponding functions. Clearly, the larger the expansion, the better the approximation; nevertheless, the 2nd derivatives in the eighth column assert that the HUM functions always have energy minimum and therefore are **not** the excited state saddle points, whereas the F_n functions are saddle points.

13.6.4 *Immediate Improvement of a Lowest State Approximant*

A very accurate ϕ_1 approximant of ψ_1 was used for He ^1S 1*s*2*s*, in terms of 19 AVOLTOs up to 5 *g* with unrestricted *s* orbitals up to 4*s*, 53 CI terms, 263 SDs,

with $E[\phi_1] = -2.14593739906107$ a.u., $F_1 = -2.14593739906075$, and, as ϕ_0, a crude approximant $1s^2$, restricted s, with $E[\phi_0] = -2.847656$ a.u. In their 2D space their 2×2 overlap matrix, s_{ij}, and Hamiltonian matrix, h_{ij}, are $s_{11} = 1.0$; $s_{12} = s_{21} = -0.047857$; $s_{22} = 1.0$ and $h_{11} = -2.847656$; $h_{12} = h_{21} = 0.102699$; $h_{22} = -2.145937$, as expected, since ϕ_1 is very close to ψ_1 and $(s_{12} h_{22} - h_{12}) = 5 \times 10^{-8}$ (i.e., in other words: $\langle \phi_0|H|\phi_1 \rangle = E_1 \langle \phi_0|\phi_1 \rangle$).

The (orthonormal) eigenvectors are indeed $\Psi^+ = 1.2 \, 10^{-6} \phi_0 + 1.00 \phi_1 = \psi_1$, $E^+ = E_1 = -2.145937$ a.u. (unchanged, since it was already essentially a Hamiltonian eigenfunction) and the improved $\Psi^- = 1.0011471 \phi_0 + 0.047910626 \phi_1$, $E^- = -2.8492667$ a.u. This is equal to [also computed independently via Eq. (13.14)] ϕ_0^+ exactly orthogonal to $\phi_1 = \psi_1$. By using a partially optimized wave function $\mathbf{f_2 = 2s3s}$, a function $\phi_0^{(2+)}$ orthogonal to both $\{\phi_0^+, \psi_1\}$ was obtained, then, in the 2×2 subspace of $\{\phi_0^+, \phi_0^{(2+)}\}$ (both orthogonal to ψ_1), the lowest Hamiltonian eigenvector $\Psi^- \equiv \phi_0^- = 0.09305114 \phi_1 + 0.29149129 \phi_0 + 0.715803 \, f_2$ has energy $E[\phi_0^-] = -2.8673305$ a.u. By further introducing an individually optimized wave function $\mathbf{f_3 = 4s^2}$, then $\Psi^- \equiv \phi_0^{2-} = 0.093057191 \phi_1 + 0.29149218 \phi_0 + 0.71580519 \, f_2 - 0.0025808019 \, f_3$ was obtained with $E_0^{-2} = -2.8673481$ a.u. By repeating with an individually optimized wave function $\mathbf{f_4 = 2p3p}$, then $\Psi^- \equiv \phi_0^{3-} = 0.09284577 \phi_1 + 0.29145852 \phi_0 + 0.71572254 \, f_2 - 0.0025805039 \, f_3 - 0.013914576 \, f_4$ was obtained with $E_0^{3-} = -2.8677801$ a.u., and so on. All these improvements were done immediately within negligible time.

13.7   Final Remark

The excited states are energy saddle points in the Hilbert space of wave functions. To compute excited states with **truncated** wave functions, most standard methods are based on HUM theorem, which says that, optimizing the desired higher root of the secular equation approaches the correct energy from above; it does not say anything about the quality of the wave function. However, the wave function is needed if we want to understand the behavior of the electrons and to compute other properties besides the energy. For large systems we need rather small and comprehensible (truncated) wave functions. We have seen that using small wave function expansions, HUM theorem, allows us to approach the energy (from above, i.e., by minimization), **but it prevents us from approaching the exact excited state saddle point wave function**. Thus, based on HUM theorem, we must use large expansions (the larger, the better), which is impracticable for large systems. We have also seen that **the functional F_n (Eq. 13.4) has local minimum at the excited state saddle point, regardless of the size of the expansion**. Therefore, it could be used for large systems with small wave functions. We saw that it has been tested for small atoms and demonstrated the above statements. So, it is worth trying it for larger systems. The implementation needs the computation, beyond the energy $\langle \phi_n|H|\phi_n \rangle$, of simple matrix elements $\langle \phi_{i<n}|\phi_n \rangle$ and $\langle \phi_{i<n}|H|\phi_n \rangle$, that

are standard subroutines of most of the common computational codes. The lower lying approximants $\phi_{i < n}$ need not be very accurate; it is sufficient to be **crude** but reasonable. F_n can be also used to identify a "flipped root," a usual problem near an avoided energy crossing in the variational parameter space. It can also be used to immediately improve a ground state approximant if the 1st excited state is more accurately known.

References

1. E.A. Hylleraas, B. Undheim, Z. Phys. **65**, 759 (1930); J.K.L. McDonald, Phys. Rev. **43**, 830 (1933)
2. N.C. Bacalis, Z. Xiong, J. Zang, D. Karaoulanis, AIP Conference Proceedings, **1790** UNSP 020007 (2016) https://doi.org/10.1063/1.4968633; N.C. Bacalis, Z. Xiong, Z.X. Wang, Int. J. Quantum Chem (2017) (submitted)
3. J. Golab, D. Yeager, P. Jørgensen, Chem. Phys. **93**, 83 (1985)
4. P. Jørgensen, J. Olsen, D. Yeager, J. Chem. Phys. **75**, 5802 (1981)
5. M. Reed, B. Simon, *Methods of Modern Mathematical Physics. Analysis of Operators*, vol IV (Academic, New York, 1978)
6. P. Jørgensen, P. Swanstrøm, D. Yeager, J. Chem. Phys. **78**, 347 (1983)
7. H. Jensen, P. Jørgensen, J. Chem. Phys. **80**, 1204 (1984)
8. H. Jensen, Electron correlation in molecules using direct second order MCSCF, in *Relativistic and Electron Correlation Effects in Molecules and Solids* (Plenum, New York, 1994), pp. 179–206
9. E. Cancès, H. Galicher, M. Lewin, J. Comput. Phys. **212**, 73 (2006)
10. H. Nakatsuji, K. Hirao, J. Chem. Phys. **68**, 2053 (1978)
11. D. Hegarthy, M.A. Robb, Mol. Phys. **38**, 1795 (1979)
12. R.H.E. Eade, M.A. Robb, Chem. Phys. Lett. **83**, 362 (1981)
13. O. Christiansen, H. Koch, P. Jørgensen, J. Chem. Phys. **103**, 7429 (1995)
14. H. Koch, O. Christiansen, P. Jørgensen, T. Helgaker, A.S. de Meras, J. Chem. Phys. **106**, 1808 (1997)
15. C. Haettig, F. Weigend, J. Chem. Phys. **113**, 5154 (2000)
16. C. Haettig, Adv. Quantum Chem. **50**, 37 (2005)
17. E.S. Nielsen, P. Jørgensen, J. Oddershede, J. Chem. Phys. **73**, 6238 (1980)
18. S.P.A. Sauer, J. Phys. B **30**, 3773 (1997)
19. J.J. Eriksen, S. Sauer, K.V. Mikkelsen, H.J.A. Jensen, J. Kongsted, J. Comput. Chem. **33**, 2012 (2013)
20. J. Schirmer, Phys. Rev. A **26**, 2395 (1982)
21. A.B. Trofimov, J. Schirmer, J. Phys. B **28**, 2299 (1995)
22. J.H. Starcke, M. Wormit, A. Dreuw, J. Chem. Phys. **130**, 024104 (2009)
23. M. Tassi, I. Theophilou, S. Thanos, Int. J. Quantum Chem. **113**, 690 (2013)
24. D.H. Weinstein, Proc. Natl. Acad. Sci. U. S. A. **20**, 529 (1934)
25. J.E. Subotnik, J. Chem. Phys. **135**, 071104 (2011)
26. A. Dreuw, M. Head-Gordon, J. Am. Chem. Soc. **126**, 4007 (2004)

27. J. Autschbach, ChemPhysChem **10**, 1757 (2009)
28. S. Grimme, F. Neese, J. Chem. Phys. **127**, 154116 (2007)
29. H.J. Monkhorst, Int. J. Quantum Chem. Symp. **11**, 421 (1977)
30. E. Dalgaard, H.J. Monkhorst, Phys. Rev. A **28**, 1217 (1983)
31. H. Koch, P. Jørgensen, J. Chem. Phys. **93**, 3333 (1990)
32. H. Koch, H.J.A. Jensen, P. Jørgensen, T. Helgaker, J. Chem. Phys. **93**, 3345 (1990)
33. K. Emrich, Nucl. Phys. A **351**, 397 (1981)
34. H. Sekino, R.J. Bartlett, Int. J. Quantum Chem. Symp. **18**, 255 (1984)
35. J. Geertsen, M. Rittby, R.J. Bartlett, Chem. Phys. Lett. **164**, 57 (1989)
36. J.F. Stanton, R.J. Bartlett, J. Chem. Phys. **98**, 7029 (1993)
37. D.C. Comeau, R.J. Bartlett, Chem. Phys. Lett. **207**, 414 (1993)
38. R.J. Rico, M. Head-Gordon, Chem. Phys. Lett. **213**, 224 (1993)
39. J.D. Watts, An introduction to equation-of-motion and linear-response coupled-cluster methods for electronically excited states of molecules, in *Radiation Induced Molecular Phenomena in Nucleic Acids*, ed. by M. K. Shukla, J. Leszczynski (Springer, Dordrecht, 2008), pp. 65–92
40. M. Nooijen, R.J. Bartlett, J. Chem. Phys. **107**, 6812 (1997)
41. J.F. Stanton, J. Chem. Phys. **101**, 8928 (1994)
42. M. Musial, R.J. Bartlett, J. Chem. Phys. **134**, 034106 (2011)
43. C. Hättig, A. Hellweg, A. Köhn, J. Am. Chem. Soc. **128**, 15672 (2006)
44. B.M. Wong, J.G. Cordaro, J. Chem. Phys. **129**, 214703 (2008)
45. A.I. Krylov, Chem. Phys. Lett. **350**, 522 (2001)
46. Y. Shao, M. Head-Gordon, A.I. Krylov, J. Chem. Phys. **118**, 4807 (2003)
47. S.V. Levchenko, A.I. Krylov, J. Chem. Phys. **120**, 175 (2004)
48. E.J. Sundstrom, M. Head-Gordon, J. Chem. Phys. **140**, 114103 (2014)
49. B.O. Roos, Ab initio methods, in *Quantum Chemistry, Part 2, Advances in Chemical Physics*, ed. by K. P. Lawley, vol. 69 (Wiley, Hoboken, 1987), pp. 399–442
50. K. Andersson, P.-A. Malmqvist, B.O. Roos, J. Chem. Phys. **96**, 1218 (1992)
51. P.M. Zimmerman, F. Bell, M. Goldey, A.T. Bell, M. Head-Gordon, J. Chem. Phys. **137**, 164110 (2012)
52. D. Casanova, J. Chem. Phys. **137**, 084105 (2012)
53. M. Wormit, D.R. Rehn, P.H.P. Harbach, J. Wenzel, C.M. Krauter, E. Epifanovsky, A. Dreuw, Mol. Phys. **112**, 774–784 (2014)
54. P.-O. Löwdin, Phys. Rev. **97**, 1509 (1955)
55. P.-O. Löwdin, Adv. Chem. Phys. **2**, 207 (1959)
56. R. Shepard, Adv. Chem. Phys. **69**, 63 (1987)
57. M. Born, R. Oppenheimer, Ann. Phys. **84**, 457 (1927)
58. H.-J. Werner, Adv. Chem. Phys. **69**, 1 (1987)
59. H.-J. Werner, W. Meyer, J. Chem. Phys. **73**, 342 (1980)
60. M. Frisch, I. Ragazos, M. Robb, H. Schlegel, Chem. Phys. Lett. **189**, 524 (1992)
61. K. Ruedenberg, L.M. Cheung, S.T. Elbert, Int. J. Quantum Chem. **16**, 1069 (1979)
62. K. Docken, J. Hinze, J. Chem. Phys. **57**, 4928 (1972)
63. H.-J. Werner, W. Meyer, J. Chem. Phys. **74**, 5794 (1981)
64. D. Yeager, D. Lynch, J. Nichols, P. Jørgensen, J. Olsen, J. Phys. Chem. **86**, 2140 (1982)
65. J. Olsen, P. Jørgensen, D. Yeager, J. Chem. Phys. **76**, 527 (1982)
66. Z. Xiong, N.C. Bacalis, Chin. Phys. B **19**, 023601 (2010). https://doi.org/10.1088/1674-1056/19/2/023601
67. T.C. Chang, W.H.E. Schwarz, Theor. Chim. Acta (Berl) **44**, 45 (1977)
68. S. Matsumoto, M. Toyama, Y. Yasuda, T. Uchide, R. Ueno, Chem. Phys. Lett. **157**, 142 (1989)
69. M.R. Hoffmann, C.D. Sherrill, M.L. Leininger, H.F. Schaefer III, Chem. Phys. Lett. **355**, 183 (2002)
70. J.J. Dorando, J. Hachmann, G.K.-L. Chan, J. Chem. Phys. **127**, 084109 (2007)
71. M. McCourt, J.W. McIver Jr., J. Comput. Chem. **8**, 454 (1987)
72. E.R. Davidson, L.Z. Stenkamp, Int. J. Quantum Chem. Symp. **10**, 21 (1976)
73. C.S. Sharma, S. Srirankanathan, Mol. Phys. **40**, 1021 (1980)

74. H.G. Miller, R.M. Dreizler, Nucl. Phys. A **316**, 32 (1979)
75. G.J. Atchity, S.S. Xantheas, K. Ruedenberg, J. Chem. Phys. **95**, 1862 (1991)
76. P.J. Knowles, H.-J. Werner, Theor. Chim. Acta **84**, 95 (1992)
77. M. Lewin, J. Math. Chem. **44**, 967 (2008)
78. J. Liu, W. Liang, J. Chem. Phys. **135**, 014113 (2011)
79. Z. Xiong, N.C. Bacalis, Commun. Math. Comput. Chem. 53, 283 (2005); Z. Xiong, M. Velgakis, N.C. Bacalis, Int. J. Quantum Chem. **104**, 418 (2005); Chin Phys. **15**, 992 (2006)
80. C.L. Pekeris, Phys. Rev. **126**, 1470 (1962)
81. N.C. Bacalis, J. Comput. Methods Sci. Eng. **16**, 253 (2016)
82. H.A. Bethe, E.E. Salpeter, *Quantum Mechanics of One- and Two-Electron Atoms* (Plenum, New York, 1977), pp. 146–162
83. M.B. Ruiz, Int. J. Quantum Chem. **101**, 246 (2005)
84. N.C. Bacalis, Z. Xiong, D. Karaoulanis, J. Comput. Methods Sci. Eng. **8**, 277 (2008). http://iospress.metapress.com/content/9270636750564km0/
85. W.H. Press, S.A. Teukolsky, W.T. Vetterling, B.P. Flannery, *Numerical Recipes in FORTRAN 77: The Art of Scientific Computing* (Cambridge University Press, Cambridge, 1992), p. 374
86. C. Froese Fischer, T. Brage, P. Jönsson, *Computational Atomic Structure: An MCHF Approach* (Institute of Physics Publishing, Bristol, 1997), pp. 92–96
87. M.-K. Chen, J. Phys. B Atomic Mol. Phys. **27**, 865 (1994)
88. Z. Xiong, Z.-X. Wang, N.C. Bacalis, Acta Phys. Sin. **63**, 053104 (2014)
89. A. Sarsa, E. Buendía, F.J. Gálvez, J. Phys. BQ At. Mol. Phys. **49**, 145003 (2016)
90. Z. Xiong, J. Zang, H.J. Liu, D. Karaoulanis, Q. Zhou, N.C. Bacalis, J. Comput. Methods Sci. Eng. **17**(3), 347–361 (2017)