Variation-Perturbation Expansions and Padé Approximants to the Energy*

E. Brändas and O. Goscinski

Quantum Theory Project, University of Florida, Gainesville, Florida 32601
and Quantum Chemistry Group, Uppsala University, Uppsala, Sweden

(Received 18 September 1969)

The variational solutions to the Rayleigh-Schrödinger inhomogeneous equations are derived by simple matrix algebra. It is shown how the approximate perturbation energies and wave functions obtained in this way obey the same type of equations as the exact ones. This permits the formulation of a variation-iteration theory convenient for numerical applications. Results for the He-like ions show that the conventional linear series is less suitable than the present method in obtaining accurate energies. The use of inner projection techniques leads to an efficient calculation of Padé approximants to the energy series, which show remarkable convergence properties.

I. INTRODUCTION

The combined use of perturbation and variation theory has several advantageous features due to a variety of factors. Perturbation energies can, in most cases, only be determined by variational approximations.[1] Low-order perturbation theory may be inadequate and the convergence of the perturbation series has to be studied in detail. These problems have been studied by a number of authors,[2-14] and our emphasis here is to obtain expressions involving the exact perturbation energies or the approximate ones in a systematic way. An operator approach, with approximations defined by inner projections[12] of the resolvent, is employed.

A convenient test case for studying perturbation problems is given by the Z expansions of two-electron systems.[6,15,16] We have examined the rich material available for He-like ions. One of the results emerging from our analysis is that Padé approximants[17] to the perturbation series follow immediately from inner projections of the reaction operator in the Brillouin-Wigner (BW) case.[18,19] In Rayleigh-Schrödinger (RS) theory, the same formulas may be adopted. We have carried out a Padé analysis of the He series, finding remarkable stability and convergence properties, superior in our opinion to any of the extrapolation techniques previously employed in this connection.

In order to introduce notations, and for the sake of comparison, we first review the conventional approach and then proceed to derive and illustrate our approximations.

II. CONVENTIONAL RAYLEIGH-SCHRÖDINGER PERTURBATION THEORY

We seek an eigenfunction Ψ and its eigenvalue E to the Hamiltonian H,

$$H\Psi = E\Psi \ . \tag{1}$$

Given the splitting

$$H = H_0 + V \ , \tag{2}$$

the eigenfunction ϕ_0 and the eigenvalue E_0 for the corresponding state of the unperturbed Hamiltonian H_0,

$$H_0\phi_0 = E_0\phi_0 \ , \tag{3}$$

RS perturbation theory in its conventional formulation assumes expansions of the type:

$$\Psi = \sum_{n=0}^{\infty} \lambda^n \phi_n, \quad E = \sum_{n=0}^{\infty} \lambda^n \epsilon_n, \quad (\epsilon_0 = E_0) \ . \tag{4}$$

The nth-order perturbation equation is

$$(E_0 - H_0)\phi_n = V\phi_{n-1} - \sum_{l=0}^{n-1} \epsilon_{n-l}\phi_l \ . \tag{5}$$

A detailed treatment of RS theory, with emphasis on a comparison of the resolvent and the inhomogeneous equation approach, was given by Löwdin.[11]

Variational principles for the perturbation energies of even order, ϵ_{2n}, have been obtained by Hylleraas[1] ($n=1$), Scherr and Knight,[6] and Stewart.[7] A short derivation of these principles illustrates some general features of approximate perturbation theory. We use a matrix formulation previously used to obtain variation solutions in the BW case.[10] We thus expand an approximate eigenfunction $\tilde{\Psi}$ and an approximate eigenvalue $\tilde{E}$:

$$\tilde{\Psi} = \sum_{n=0}^{\infty} \lambda^n \tilde{\phi}_n, \quad (\tilde{\phi}_0 = \phi_0), \quad \tilde{E} = \sum_{n=0}^{\infty} \lambda^n \tilde{\epsilon}_n \ . \tag{6}$$

Insertion of (6) into the variational expression

$$\langle\Psi|H|\Psi\rangle/\langle\Psi|\Psi\rangle \geq E \ , \tag{7}$$

and collecting terms after even powers of λ gives, assuming ϕ_k, $k<n$, to be known exactly, i.e.,

$$\tilde{\phi}_k = \phi_k, \quad k<n; \tag{8}$$

$$\begin{aligned}\epsilon_{2n}\langle\phi_0|\phi_0\rangle \leq 2\Big(&\sum_{l=0}^{n-1}\langle\phi_{2n-l}|H_0-E_0|\phi_l\rangle \\ &+\sum_{l=1}^{n-1}\langle\phi_{2n-l}|V-\epsilon_1|\phi_{l-1}\rangle - \sum_{l=2}^{n-1} \\ &\times\sum_{p=2}^{l}\epsilon_p\langle\tilde{\phi}_{2n-l}|\phi_{l-p}\rangle\Big)+\langle\tilde{\phi}_n|H_0-E_0|\tilde{\phi}_n\rangle \\ &-2\langle\tilde{\phi}_n|V-\epsilon_1|\phi_{n-1}\rangle-2\sum_{l=2}^{n}\epsilon_l\langle\tilde{\phi}_n|\phi_{n-l}\rangle \\ &-\sum_{l=1}^{n-1}\sum_{p=n-l+1}^{2n-l}\epsilon_p\langle\phi_l|\tilde{\phi}_{2n-l-p}\rangle \ .\end{aligned} \tag{9}$$

It follows from (3) and (5) that the three terms inside the brackets vanish. The following inequality, relating ϵ_{2n} and its approximation $\tilde{\epsilon}_{2n}$, is then obtained:

$$\begin{aligned}\epsilon_{2n} \leq \tilde{\epsilon}_{2n} = \langle\phi_0|\phi_0\rangle^{-1}\big(&\langle\tilde{\phi}_n|H_0-E_0|\tilde{\phi}_n\rangle \\ &+2\langle\tilde{\phi}_n|V-\epsilon_1|\phi_{n-1}\rangle-2\sum_{l=2}^{n}\epsilon_l\langle\tilde{\phi}_n|\phi_{n-l}\rangle \\ &-\sum_{l=1}^{n-1}\sum_{p=n-l+1}^{2n-l}\epsilon_p\langle\phi_l|\tilde{\phi}_{2n-l-p}\rangle\big),\end{aligned} \tag{10}$$

with $\tilde{\phi}_n$ determined from

$$\begin{aligned}\delta\tilde{\epsilon}_{2n} = \langle\delta\tilde{\phi}_n|\big(&(H_0-E_0)\tilde{\phi}_n \\ &+(V-\epsilon_1)\phi_{n-1}-\sum_{l=2}^{n}\epsilon_l\phi_{n-l}\rangle = 0 \ .\end{aligned} \tag{11}$$

In most practical cases, we only know ϕ_0 exactly, which leads to $\tilde{\phi}_1$, $\tilde{\epsilon}_2$ (and $\tilde{\epsilon}_3$).

Note that the variational principle for $\tilde{\epsilon}_{2n}$ and $\tilde{\phi}_n$ involves the knowledge of the *exact* ϕ_0, ϕ_1, ..., ϕ_{n-1} and ϵ_0, ϵ_1, ϵ_2, ..., ϵ_{2n-1}. For $n=1$, Eq. (11) becomes the well-known Hylleraas bound[1]:

$$\epsilon_2 \leq \tilde{\epsilon}_2 = \langle\phi_0|\phi_0\rangle^{-1}\big(\langle\tilde{\phi}_1|H_0-E_0|\tilde{\phi}_1\rangle + 2\langle\tilde{\phi}_1|V-\epsilon_1|\phi_0\rangle\big) \ , \tag{12}$$

and

$$\delta\tilde{\epsilon}_2 = \langle\delta\tilde{\phi}_1|(H_0-E_0)\tilde{\phi}_1+(V-\epsilon_1)\phi_0\rangle = 0 \ . \tag{13}$$

Let us expand the approximate functions in the linearly independent basis $\{|h_1\rangle|h_2\rangle\cdots|h_n\rangle\} \equiv |\underline{h}\rangle$ and introduce the following definitions:

$$\begin{aligned}&\underline{\Delta} = \langle\underline{h}|\underline{h}\rangle, \quad \underline{R} = \langle\underline{h}|E_0-H_0|\underline{h}\rangle^{-1}, \\ &\underline{V} = \langle\underline{h}|V-\epsilon_1|\underline{h}\rangle, \quad \underline{a} = \langle\underline{h}|V-\epsilon_1|\phi_0\rangle \ .\end{aligned} \tag{14}$$

We then seek $\tilde{\phi}_1 = |\underline{h}\rangle\underline{c}_1$, which we obtain from $\delta\tilde{\epsilon}_2 = 0$. With the choice $\delta\tilde{\phi}_1 = |\underline{h}\rangle\underline{c}$, where $\underline{c}$ is arbitrary, we get

$$\underline{c}^\dagger(\langle\underline{h}|H_0-E_0|\underline{h}\rangle\underline{c}_1+\langle\underline{h}|V-\epsilon_1|\phi_0\rangle) = 0 \ , \tag{15}$$

and hence

$$\langle\underline{h}|H_0-E_0|\underline{h}\rangle\underline{c}_1+\langle\underline{h}|V-\epsilon_1|\phi_0\rangle = 0 \ , \tag{16}$$

with the solution ($\langle\phi_0|\phi_0\rangle = 1$ from now on)

$$\underline{c}_1 = \underline{R}\,\underline{a}; \quad \tilde{\epsilon}_2 = \underline{a}^\dagger\underline{R}\,\underline{a} \ . \tag{17}$$

The explicit solution to the Hylleraas problem was obtained by Weiss and Martin[20] and its invariance properties were discussed by Miller.[21] A derivation by means of an inner projection of the resolvent was given by Lindner and Löwdin.[22] It bypasses the formulation of an *ad hoc* variational principle.

Formally, it is now possible to get $\tilde{\phi}_n$ from $\delta\tilde{\epsilon}_{2n} = 0$, provided we assume all corrections $\tilde{\phi}_k$ to the total wave function to be expanded in the same basis $|\underline{h}\rangle$, i.e., $\tilde{\phi}_k = |\underline{h}\rangle\underline{c}_k$. From (11), we obtain

$$\begin{aligned}&\underline{c}_n = \underline{R}\,\underline{V}\,\underline{c}_{n-1} - \sum_{l=2}^{n-1}\tilde{\epsilon}_l\underline{R}\,\underline{\Delta}\,\underline{c}_{n-l}+\tilde{\epsilon}_n\underline{R}\langle\underline{h}|\phi_0\rangle, \\ &n = 2,3,\cdots, \qquad \underline{c}_1 = \underline{R}\,\underline{a} \ ,\end{aligned} \tag{18}$$

$$\tilde{\epsilon}_l = \underline{a}^\dagger\underline{c}_{l-1} - \sum_{k=1}^{l-2}\epsilon_{l-k}\underline{c}_k^\dagger\langle\underline{h}|\phi_0\rangle \ .$$

However, (18) was obtained under the formal assumption that ϕ_k, $k<n$ was known exactly in order to remove the three first terms in (9). Since in practice all the perturbed wave functions are constructed as linear combinations of the same basis,

it is obvious, see also Midtdal,[23] that the condition $\delta\tilde{\epsilon}_{2n}=0$ with the choice $\delta\tilde{\phi}_n=\tilde{\phi}_{2n-l}$, $l=0,1,\ldots,n-1$, have the same effect on (9). Bounds to ϵ_{2n}[1,6,7,9,22,24] are, however, no longer obtainable. The previous matrix formulation can be applied to the case when ϕ_0 is only approximately known. The variation principle for E_0 with $\phi_0=|\underline{h}\rangle\underline{c}_0$ gives

$$\langle\underline{h}|\tilde{E}_0-H_0|\underline{h}\rangle\underline{c}_0=0, \tag{19}$$

with the normalization condition $\underline{c}_0^{+}\langle h|h\rangle\underline{c}_0=1$. The previous formulas then hold with the following modifications (h orthogonal to $\tilde{\phi}_0$):

$$\underline{R}=\langle\underline{h}|\tilde{E}_0-H_0|\underline{h}\rangle^{-1}, \tag{20}$$

$$V=\langle\underline{h}|V-\tilde{\epsilon}_1|\underline{h}\rangle,\quad \underline{a}=\underline{V}\underline{c}_0,$$

with $\tilde{\epsilon}_1$ determined from

$$\underline{c}_0^{\dagger}\underline{V}\underline{c}_0=0. \tag{21}$$

Since the inclusion of approximate quantities eliminates the possibility to give upper and lower bounds, one might question the value of higher-order perturbation theory, unless the basis used is practically complete. This fact leads us to an approach which involves bounds for the total energy and not for the individual ϵ_n.

In Sec. III, we will derive a modified expansion obtained from the variation principle. The method has been presented elsewhere[25] in the case of exact quantities. Here, we will give a more general formulation, which also applies to approximate quantities based on the previous analysis.

III. VARIATION-ITERATION THEORY

The Schrödinger equation in an arbitrary basis $\{\phi_i\}$ can be written in the form

$$\underline{M}\underline{c}=\underline{0}, \tag{22}$$

with the notation

$$\underline{M}=\underline{M}(\mathcal{E})=\langle\phi|H-\epsilon|\phi\rangle\ \phi=\{\phi_0,\phi_1,\ldots\}. \tag{23}$$

Given the splitting, $H=H_0+V$, we choose ϕ_0 as the zeroth-order ground-state solution of $H_0\Psi_n^0=E_n^0\Psi_n^0$, i.e., $\phi_0\equiv\Psi_0^0$, $E_0\equiv E_0^0$. Introducing $\epsilon_n=\langle\phi_n|V|\phi_{n-1}\rangle$, we get

$$M_{00}=E_0+\epsilon_1-\mathcal{E},$$
$$M_{0n}=\epsilon_{n+1};\quad n\geq 1. \tag{24}$$

We have here assumed also the normalization $\langle\phi_0|\phi_n\rangle=\delta_{n0}$. The partitioning technique[5] applied to (22) leads to the following set of equations:

$$M_{00}c_0+\underline{M}_{0b}\underline{c}_b=0,$$
$$\underline{M}_{b0}c_0+\underline{M}_{bb}\underline{c}_b=\underline{0}, \tag{25}$$

which corresponds to the partition of the matrix $\underline{M}$:

$$\underline{M}\underline{c}=\begin{pmatrix}M_{00} & \underline{M}_{0b}\\ \underline{M}_{b0} & \underline{M}_{bb}\end{pmatrix}\begin{pmatrix}c_0\\ \underline{c}_b\end{pmatrix}=\underline{0}. \tag{26}$$

Irrespective of the detailed form of $\underline{M}_{bb}$, provided its inverse exists, we get the solution

$$\underline{c}_b+\underline{M}_{bb}^{-1}\underline{M}_{b0}c_0=\underline{0},$$
$$M_{00}-\underline{M}_{0b}\underline{M}_{bb}^{-1}\underline{M}_{b0}=0,\qquad c_0\neq 0. \tag{27}$$

With the notation $\lambda_n=c_{n-1}/c_0$, we obtain

$$\mathcal{E}=E_0+\epsilon_1+\sum_{n=2}\lambda_n\epsilon_n. \tag{28}$$

If $\{\phi_n\}$ is complete it is clear that $\mathcal{E}=E$; otherwise $E\leq\mathcal{E}$. Note that $\lambda_n=\lambda_n(\mathcal{E})$ and that (28) is the bracketing function.[5]

In this formulation of the problem, we have not yet specified ϕ_n, where $n\geq 1$. In the following we will choose ϕ_n as the nth order correction to ϕ_0 in a certain perturbation scheme, yet to be specified. The usual statement about a convergent expansion

$$\Psi=\sum_{n=0}^{\infty}\phi_n,\quad E=E_0+\epsilon_1+\sum_{n=2}^{\infty}\epsilon_n \tag{29}$$

implies in our formalism

$$\underline{c}_b=-\underline{M}_{bb}^{-1}\underline{M}_{b0}c_0=\underline{1}c_0. \tag{30}$$

A truncated basis of finite order N yields

$$E\leq E_0+\epsilon_1+\sum_{n=2}^{N+1}\lambda_n\epsilon_n, \tag{31}$$

which, contrary to linear expansions, gives a rigorous upper bound to the total energy. In order to compute λ_n, which for instance can be obtained by the Nesbet[26] method, we need the explicit form of $\underline{M}_{bb}$. In the BW case, a simple expression is obtained.[18] In the RS scheme it can be expressed in terms of the corrections ϵ_n and the overlaps $\langle\phi_i|\phi_j\rangle$.[11]

In practice the ϵ_n's and the ϕ_n's are not exactly known. As we will see, *the theory also holds in this approximate case.*

From now on we will restrict our attention to RS theory. We substitute the perturbation expansion into the Schrödinger equation and collect terms after powers of V. The set of inhomogeneous Eqs. (5) is, of course, the starting point:

$$\begin{aligned}
&(E_0 - H_0)\,\phi_0 = 0\,,\\
&(E_0 - H_0)\phi_1 = (V - \epsilon_1)\phi\,,\\
&\quad\vdots\\
&(E_0 - H_0)\phi_n = V\phi_{n-1} - \sum_{k=0}^{n-1} \epsilon_{n-k}\phi_k\,.
\end{aligned} \tag{32}$$

We now construct approximate $\tilde{\phi}_i$'s, which are subject to the conditions

$$\begin{aligned}
&\langle \tilde{\phi}_n | \phi_0 \rangle = 0; n \geq 1\,,\\
&\langle \tilde{\phi}_1 | E_0 - H_0 | \tilde{\phi}_1 \rangle = \langle \tilde{\phi}_1 | V - \epsilon_1 | \phi_0 \rangle\,,\\
&\quad\vdots\\
&\langle \tilde{\phi}_n | E_0 - H_0 | \tilde{\phi}_n \rangle\\
&\qquad = \langle \tilde{\phi}_n | V - \epsilon_1 | \tilde{\phi}_{n-1} \rangle - \sum_{k=1}^{n-2} \epsilon_{n-k} \langle \phi_n | \phi_k \rangle\,.
\end{aligned} \tag{33}$$

These conditions can be simply implemented: Assume a linearly independent basis $|\underline{h}\rangle = |h_1, h_2, \ldots, h_n\rangle$ and expand

$$\tilde{\phi}_k = |\underline{h}\rangle \underline{c}_k \tag{34}$$

in the same way as in (15)–(18). From (33) we get

$$\underline{c}_1^{\dagger}[\langle \underline{h} | E_0 - H_0 | \underline{h} \rangle \underline{c}_1 - \langle \underline{h} | V - \epsilon_1 | \phi_0 \rangle] = 0\,, \tag{35}$$

or in the more compact notation

$$\underline{c}_1^{\dagger}[\underline{R}^{-1}\underline{c}_1 - \underline{a}] = 0. \tag{36}$$

It is immediately clear that the choice $\underline{c}_1 = \underline{R}\underline{a}$ leads to the desired result. In addition, it is the optimal result within the linear space spanned by $\underline{h}$, since $\underline{c}_1 = \underline{R}\,\underline{a}$ is the solution of the Hylleraas problem.[1] Proceeding similarly to $\tilde{\phi}_n$, the obvious result is

$$\begin{aligned}
&\underline{c}_n = \underline{R}\underline{V}\underline{c}_{n-1} - \sum_{k=1}^{n-2} \tilde{\epsilon}_{n-k} \underline{R}\,\underline{\Delta}\,\underline{c}_k + \tilde{\epsilon}_n \underline{R}\langle \underline{h} | \phi_0 \rangle\,,\\
&\qquad n = 2, 3, \ldots,
\end{aligned} \tag{37}$$

which is identical to the variational result of Sec. II.[6,7] The relevant formulas in the Schrödinger theory are now easily rederived in an entirely similar manner, but with the important distinction that they apply to approximate quantities obtained from the Hylleraas, Scherr, Knight, and Stewart method. Since the following formulas hold in the approximate, as well as in the exact case, we will from now on drop the tildes. From (33) is obtained (see Löwdin[11])

$$\langle \phi_m | V | \phi_n \rangle = \langle \phi_{m+1} | V | \phi_{n-1} \rangle + \sum_{k=0}^{m} \epsilon_{m+1-k} \langle \phi_n | \phi_k \rangle - \sum_{l=0}^{n-1} \epsilon_{n-l} \langle \phi_{m+1} | \phi_l \rangle\,, \tag{38}$$

from which follows

$$V_{mn} = \langle \phi_m | V | \phi_n \rangle = \sum_{k=0}^{m} \sum_{l=0}^{n} \epsilon_{m+n+1-k-l} \langle \phi_k | \phi_l \rangle\,. \tag{39}$$

If $\langle \phi_i | \phi_0 \rangle = \delta_{i0}$, (39) reduces to[11]

$$V_{mn} = \epsilon_{m+n+1} + \sum_{k=1}^{m} \sum_{l=1}^{n} \epsilon_{m+n+1-k-l} \langle \phi_k | \phi_l \rangle\,. \tag{40}$$

The matrix elements $M_{mn} = \langle \phi_m | H - \mathcal{E} | \phi_n \rangle$ are then given by

$$M_{mn} = V_{mn} - V_{m,n-1} + (E_0 - \mathcal{E})\langle \phi_m | \phi_n \rangle + \sum_{l=0}^{n-1} \epsilon_{n-l} \langle \phi_m | \phi_l \rangle\,. \tag{41}$$

To see that M_{mn} is Hermitian, we insert V_{mn} from (39) into (41), thus obtaining

$$M_{mn} = \sum_{k=0}^{m}\sum_{l=0}^{n}(\epsilon_{m+n+1-k-l} - \epsilon_{m+n-k-l})\langle\phi_k|\phi_l\rangle + \sum_{k=0}^{m}\epsilon_{m-k}\langle\phi_k|\phi_n\rangle$$

$$+ \sum_{k=0}^{n}\epsilon_{n-k}\langle\phi_k|\phi_m\rangle - \mathcal{E}\langle\phi_m|\phi_n\rangle . \tag{42}$$

For the perturbation energies, we get

$$\epsilon_{2n} = \langle\phi_0|V|\phi_{2n-1}\rangle - \sum_{l=0}^{2n-2}\epsilon_{2n-l}\langle\phi_0|\phi_l\rangle, \quad \epsilon_{2n+1} = \langle\phi_0|V|\phi_{2n}\rangle - \sum_{l=0}^{2n-1}\epsilon_{2n+1-l}\langle\phi_0|\phi_l\rangle, \tag{43}$$

and with $\langle\phi_i|\phi_0\rangle = \delta_{i0}$, (43) gives[4]

$$\epsilon_{2n} = \langle\phi_0|V|\phi_{2n-1}\rangle = \langle\phi_n|V|\phi_{n-1}\rangle - \sum_{k=1}^{n}\sum_{l=1}^{n-1}\epsilon_{2n-k-l}\langle\phi_k|\phi_l\rangle,$$

$$\epsilon_{2n+1} = \langle\phi_0|V|\phi_{2n}\rangle = \langle\phi_n|V|\phi_n\rangle - \sum_{k=1}^{n}\sum_{l=1}^{n}\epsilon_{2n+1-k-l}\langle\phi_k|\phi_l\rangle . \tag{44}$$

In order to assure orthogonality to ϕ_0, we simply replace $\underline{V}$ by

$$\underline{V} - \underline{a}\langle\phi_0|\underline{h}\rangle - \langle\underline{h}|\phi_0\rangle\underline{a}^{\dagger}, \quad \text{and} \quad \underline{\Delta} \text{ by } \underline{\Delta} - \langle\underline{h}|\phi_0\rangle\langle\phi_0|\underline{h}\rangle ,$$

which means that we have changed our manifold $|\underline{h}\rangle$ to $|\underline{h}\rangle - |\phi_0\rangle\langle\phi_0|\underline{h}\rangle$. The matrices $\underline{R}$ and $\underline{a}$ remain unchanged.

Note that the results given here also hold for an approximate $\tilde{\phi}_0 = |\underline{h}\rangle\underline{c}_0$ and $\tilde{E}_0 = \langle\tilde{\phi}_0|H_0|\tilde{\phi}_0\rangle$, subject to the variational condition

$$\underline{R}^{-1}\underline{c}_0 = 0, \quad \underline{c}_0^{\dagger}\langle\underline{h}|\underline{h}\rangle c_0 = 1 \tag{45}$$

and with $\underline{R}$, $\underline{V}$, and $\underline{a}$ modified accordingly (see Sec. II). As a consistency check we write

$$M_{00} = \langle\tilde{\phi}_0|H_0 + V - \mathcal{E}|\tilde{\phi}_0\rangle = \tilde{E}_0 + \tilde{\epsilon}_1 - \mathcal{E}, \quad M_{0n} = \langle\phi_0|H_0 + V - \mathcal{E}|\tilde{\phi}_n\rangle$$

$$= -\underline{c}_0^{\dagger}\underline{R}^{-1}\underline{c}_n + \sum_{l=0}^{n}\tilde{\epsilon}_{n+1-l}\langle\tilde{\phi}_0|\tilde{\phi}_l\rangle + (\tilde{E}_0 - \epsilon)\langle\tilde{\phi}_0|\tilde{\phi}_n\rangle ; \quad n > 0 . \tag{46}$$

Since $\langle\tilde{\phi}_n|\tilde{\phi}_0\rangle = 0$, we get, from (45) and (46),

$$M_{0n} = \tilde{\epsilon}_{n+1} , \tag{47}$$

which is the desired result.

IV. PADÉ APPROXIMANTS

Given a function of a complex variable, its power-series representations are the simplest available, simplicity being not always compatible with convergence or existence. Just as the high-temperature expansion of a thermodynamic-property (e.g., a susceptibility) perturbation theory offers an example where the coefficients of a power series are known in principle and which sometimes can be computed in practice quite accurately. The convergence properties of such expansion are determined among other things by singularities of the function represented. The power series break down at such points and are useless as devices for locating phase transitions. The Padé[27] technique, successfully applied to a number of problems of this type,[17] consists in constructing rational approximants to the function represented based upon the coefficients of the power-series expansion. There is an excellent review by Baker[17] of the Padé technique.

The upper and lower bounds that we have previously discussed in connection with continued fraction energy expansions of BW type[18] and of polarizabilities[28] are recognized to be Padé approximants

of type $[N,N]$ and $[N,N-1]$ (in the notation by Baker[17]), respectively.[19] The approximants $[N,N]$ and $[N,N-1]$ furnish upper and lower bounds to the function represented when the corresponding power series is of Stieltjes type, this being the case in the BW series and in the moment expansion of polarizabilities.[19] One of the characteristics of the inner projection approach is that the form of the approximants is conveniently expressed as the expectation value of an inverse matrix, and not as a ratio of two determinants. The computations are hence very simple.

Given a series like (4) and (6), it is immediately obvious that Padé approximants will yield results of interest about $E(\lambda)$ or $\Psi(\lambda)$. Unfortunately not all series are of Stieltjes type and some of the convergence properties of the Padé approximants themselves are still unknown.[17,29] We discuss elsewhere some of the matters,[19] and it will suffice here to give the explicit form of the approximants. Given a series for the energy like (4), one obtains

$$E_1(\lambda)=E_0+\lambda\epsilon_1+\lambda^2(\epsilon_2+\lambda\underline{\epsilon}^{\dagger}\underline{\Delta}^{-1}\underline{\epsilon}) , \tag{48}$$

where $\underline{\epsilon}^{\dagger}$ is a row vector

$$\underline{\epsilon}^{\dagger}=(\epsilon_3,\ldots,\epsilon_{N+2}) , \tag{49}$$

and $\underline{\Delta}$ is a square matrix with elements

$$\Delta_{kl}=\epsilon_{k+l+1}-\lambda\epsilon_{k+l+2} \quad k,l=1,\ldots N . \tag{50}$$

Note that (48) for a BW series determines a lower bound.[18] Another set of approximants which are easily obtained is given by

$$E_1^{1}(\lambda)=E_0+\lambda\epsilon_1+\lambda^2\underline{e}^{\dagger}\underline{D}^{-1}\underline{e} . \tag{51}$$

The row vector $\underline{e}^{\dagger}$ has the form

$$\underline{e}^{\dagger}=(\epsilon_2,\epsilon_3,\ldots,\epsilon_{N+1}) , \tag{52}$$

and $\underline{D}$ has matrix elements

$$D_{kl}=\epsilon_{k+l}-\lambda\epsilon_{k+l+1} \quad k,l=1,\ldots,N . \tag{53}$$

For a BW series, (51) determines upper bounds.[18] It should be emphasized that both (48) and (51) obtained by inner projections in the case of BW series correspond in all cases to a Padé analysis of the power series

$$E(\lambda)-E_0-\lambda\epsilon_1=\sum_{n=2}^{\infty}\lambda^n\epsilon_n . \tag{54}$$

This means that of all conceivable Padé approximants a special class is automatically selected. In fact, given the Padé approximants $[N,N]$ and $[N,N-1]$ to the series (54) one can write (48) and (51) in the form

$$E_1(\lambda)=E_0+\lambda\epsilon_1-\lambda^2[N,N] , \tag{55}$$

and

$$E_1^{1}(\lambda)=E_0+\lambda\epsilon_1-\lambda^2[N,N-1] , \tag{56}$$

respectively, though for computational purposes the forms (48) and (51) are convenient. They are also suitable for locating the singularities of the function $E(\lambda)$. This is achieved by searching for the zeroes of $\underline{\Delta}$ and $\underline{D}$.[30]

In conclusion, given the linear series (4) and the variation-perturbation, expressions like (31), (48), and (51) provide a good alternate way to compute the energy and to examine the convergence and properties of $E(\lambda)$.

V. NUMERICAL RESULTS

The isoelectronic sequence of He was used as a test example. The Hamiltonian for a two-electron system with nuclear charge Z is given (in atomic units) by

$$\begin{aligned} &H=H_0+V, \quad V=r_{12}^{-1} ,\\ &H_0=-\tfrac{1}{2}(\nabla_1^2+\nabla_2^2)-Z/r_1-Z/r_2 , \end{aligned} \tag{57}$$

which, when transformed to the scaled variables $(\vec{r}_i'=Z\vec{r}_i,\ i=1,2)$, leads to

$$H=-\tfrac{1}{2}(\nabla_1^2+\nabla_2^2)-r_1^{-1}-r_2^{-1}+Z^{-1}\cdot r_{12}^{-1}, \tag{58}$$

where the energy is now expressed in Z^{-2} a.u. The form (58) is particularly suitable for considering a perturbation expansion in the parameter Z^{-1}. The conventional linear expression is given by

$$E=Z^{-1}(E_0+Z^{-1}\epsilon_1+\sum_{n=2}^{\infty}Z^{-n}\epsilon_n) , \tag{59a}$$

$$\Psi=Z^3\sum_{n=0}^{\infty}Z^{-n}\phi_n . \tag{59b}$$

If (59b) is truncated to N terms, Eq. (31) gives

$$E\leq Z^2\{E_0+Z^{-1}\epsilon_1+\sum_{n=2}^{N+1}[\lambda_n(Z)Z^{-n}]\epsilon_n\} , \tag{60}$$

where $\lambda_n(Z)$ are rational functions of Z determined by Eq. (27). Our test is based on extremely accurate perturbation-theory calculations by Midtdal, Lyslo, and Aashamar.[15] The basis employed was constructed with functions of the type

TABLE I. Various expansions of Rayleigh-Schrödinger quantities for H^-. The number in parentheses indicates the highest perturbation order which is included in the calculation. ϵ_n in tables redefined as $\epsilon_n Z^{-n+2}$.

N	$\epsilon_0+\epsilon_1-[N,N]$ $(2N+2)$	$\epsilon_0+\epsilon_1-[N,N-1]$ $(2N+1)$	$\epsilon_0+\epsilon_1+\sum_{n=2}^{N+1}\lambda_n\epsilon_n$ $(2N+1)$	$\epsilon_0+\epsilon_1+\sum_{n=2}^{2N+1}\epsilon_n$ $(2N+1)$
1	−0.5247737	−0.5244223	−0.5010702	−0.5239675
2	−0.5272299	−0.5242083	−0.5207961	−0.5258925
3	−0.5276480	−0.5274622	−0.5255182	−0.5268776
4	−0.5278070	−0.5276770	−0.5269640	−0.5272861
5	−0.5277385	−0.5277468	−0.5274631	−0.5274856
6	−0.5277449	−0.5277424	−0.5276435	−0.5275917
7	−0.5277499	−0.5277507	−0.5277105	−0.5276518
8	−0.5277512	−0.5277501	−0.5277357	−0.5276874
9	−0.5277508	−0.5277508	−0.5277452	−0.5277093
10	−0.5277507	−0.5277508	−0.5277488	−0.5277231
11	−0.5277508	−0.5277510	−0.5277501	−0.5277320
12	−0.5277509	−0.5277509	−0.5277508	−0.5277379
20	−0.5277508	−0.5277508	−0.5277508	−0.5277500

$$v(s,u,t)_{nml}=e^{-\frac{1}{2}s}s^n u^m t^l,$$

$$s=r_1+r_2,\quad 0\le|t|\le u\le s<\infty;$$

$$t=r_1-r_2,\quad n,m,l \text{ positive integers,} \quad (61)$$

$$u=r_{12};\qquad l \text{ even;}$$

and 204 terms were selected according to their maximum contribution to ϵ_2. For details we refer to the work of Midtdal *et al.*[15] Numerical results are given in Tables I–V. We report values for H^-, He, Li^+, and Be^{++} because they illustrate the increasing convergence with larger Z. In the first four tables we compare the linear series (6), the variation-perturbation formula (31), and the Padé approximants obtained by inner projection (55) and (56). Clearly, the variation-perturbation formula is always an upper bound, whereas the other three do not yield, in this case, rigorous bounds.

The comparison of convergence rates is of interest. The linear series does not differ too much in convergence properties with respect to the variation-perturbation result, and in the case of H^-, the latter converges faster in the long run. The expressions involving Padé approximants converge appreciably faster, though a bit erratically at the beginning, when $E_1(1/Z)$ and $E_1^1(1/Z)$ agree with each other, one seems to have reached a very good result. For H^- this requires, as input information, about ϵ_0 up to ϵ_{20}, and to recover a similar result from the linear series, one needs twice as many perturbation energies, and something similar is required to reach the same accuracy by the variation-perturbation scheme.

It should be observed that the variation-perturbation result for $N=1$ is not very good and that $E_1^1(\lambda)$; $\lambda=1/Z$, for $N=1$, is a "geometric" approximation

TABLE II. Various expansions for He.

N	$E=\epsilon_0+\epsilon_1-[N,N]$ $(2N+2)$	$E=\epsilon_0+\epsilon_1-[N,N-1]$ $(2N+1)$	$E=\epsilon_0+\epsilon_1+\sum_{N=2}^{N+1}\lambda_n\epsilon_n$ $(2N+1)$	$E=\epsilon_0+\epsilon_1+\sum_{N=2}^{2N+1}\epsilon_n$ $(2N+1)$
1	−2.9035283	−2.9034334	−2.8954538	−2.9033169
2	−2.9037212	−2.9031624	−2.9032771	−2.9036686
3	−2.9037243	−2.9037237	−2.9037013	−2.9037186
4	−2.9037244	−2.9037243	−2.9037322	−2.9037236
5	−2.9037244	−2.9037244	−2.9037243	−2.9037243
6	−2.9037244	−2.9037244	−2.9037244	−2.9037244

TABLE III. Various expansions for Li^{+}.

N	$E=\epsilon_0+\epsilon_1-[N,N]$ $(2N+2)$	$E=\epsilon_0+\epsilon_1-[N,N-1]$ $(2N+1)$	$E=\epsilon_0+\epsilon_1+\sum_{n=2}^{N+1}\lambda_n\epsilon_n$ $(2N+1)$	$E=\epsilon_0+\epsilon_1+\sum_{n=2}^{2N+1}\epsilon_n$ $(2N+1)$
1	−7.2798622	−7.2798191	−7.2759642	−7.2797667
2	−7.2799131	−7.2802763	−7.2798262	−7.2799039
3	−7.2799134	−7.2799133	−7.2799116	−7.2799130
4	−7.2799134	−7.2799134	−7.2799134	−7.2799134

TABLE IV. Various expansions for Be^{++}.

N	$E=\epsilon_0+\epsilon_1-[N,N]$ $(2N+2)$	$E=\epsilon_0+\epsilon_1-[N,N-1]$ $(2N+1)$	$E=\epsilon_0+\epsilon_1+\sum_{n=2}^{N+1}\lambda_n\epsilon_n$ $(2N+1)$	$E=\epsilon_0+\epsilon_1+\sum_{n=2}^{2N+1}\epsilon_n$ $(2N=1)$
1	−13.655546	−13.655521	−13.653276	−13.655492
2	−13.655566	−13.655598	−13.655539	−13.655563
3	−13.655566	−13.655566	−13.655566	−13.655566

to the series (6)

$$E_1^{1}(\lambda)=\epsilon_0+\lambda\epsilon_1-\lambda^2[N,N-1]$$

$$=\epsilon_0+\lambda\epsilon_1+\frac{\lambda^2\epsilon_2}{1-(\epsilon_3/\epsilon_2)\lambda}\ . \tag{62}$$

The last formula can be obtained by assuming that the higher orders form a geometric series with ratio $\epsilon_n/\epsilon_{n-1}=k$. Of course, this does not hold, but nevertheless (62) is a very good approximation since it is a lower bound to an upper bound.[31]

In Table V we present results obtained with *extrapolated* values of the ϵ_n according to the prescriptions given by Midtdal.[23]

VI. DISCUSSION

The variational solutions to the energy corrections of even order in the Rayleigh-Schrödinger perturbation theory were shown to be identical to a matrix formulation of the RS inhomogeneous equations. This includes the solution to the Hylleraas problem as well as those of higher orders.

The exact RS theory and the approximate one obtained here turned out to have the same structure. We showed how this allows to construct matrix elements of the Hamiltonian

$$H=H_0+V \tag{63}$$

with respect to a basis consisting of the zeroth-order wave function and its approximate higher-order corrections.

A previously discussed variation-iteration method was shown to give upper bounds to the total energy in the form

$$E\leq E_0+\epsilon_1+\sum_{n=2}^{N+1}\lambda_n\epsilon_n\ , \tag{64}$$

where the ϵ_n could be chosen as approximate perturbation energies obtained by conventional variational procedures.

Calculations for the ground state of the isoelectronic sequence of He were carried out. The results show that the combined use of the variational principle and perturbation theory is always better than linear perturbation expansions. The methods outlined here, compared to simple variation theory, retain the formal properties of perturbation meth-

TABLE V. Energies for He in various treatments.

Upper bound and result of Padé analysis	−2.903724362
Pekeris's extrapolation[a]	−2.903724376
Midtdal extrapolation[b]	−2.903724376
Padé analysis of Midtdal's[b] extrapolation	−2.903724377

[a] See C. L. Pekeris, Phys. Rev. 126, 1470 (1962).
[b] See Ref. 23.

ods, show numerical stability, and have good convergence properties. Furthermore the particular approach used leads to the construction of certain approximants of Padé type, which yield alternate ways of computing the final results and thus provide useful checks. There is an additional feature about the Padé approximants considered, namely, their use in locating singularities of the function $E(\lambda)$ and in finding the radius of convergence of the expansion (4).

There has been much work about convergence of the Z expansions of two-electron ions.[32] Midtdal *et al.*[33] used a very interesting procedure, namely, investigating the zeroes of the norm of the wave functions

$$\langle \tilde{\Psi}_p | \tilde{\Psi}_p \rangle = \sum_{n=0}^{2p} Z^{-n} \sum_{k=0}^{n} \langle \tilde{\phi}_{n-k} | \tilde{\phi}_k \rangle , \qquad (65)$$

and increased the number p of terms employed. They applied a theorem by Jentzch[34] stating that for every power series, every point of the circle of convergence is a limit-point of zeroes of partial sums, and could extrapolate a power series not having zeroes outside the circle of convergence. Stillinger[35] approached the problem by using *ad hoc* numerical considerations. We would like to emphasize the use of (48) and (51) in order to locate singularities of $E_1(\lambda)$. We are carrying out such calculations now and they will be reported elsewhere.[30]

ACKNOWLEDGMENTS

We thank Professor Per-Olov Löwdin for stimulating discussions concerning the problem. We also thank Professor J. Midtdal for very kindly sending us energy coefficients and overlap matrices prior to publication.

---

*Sponsored in part by the Swedish National Research Council and in part by the Air Force Office of Scientific Research (OSR) through the European Office of Aerospace Research (OAR) U. S. Air Force under Grant No. AF EOAR 67-50.

[1]E. Hylleraas, Z. Physik 65, 209 (1930).
[2]P. Goldhammer and E. Feenberg, Phys. Rev. 101, 1233 (1956).
[3]R. C. Young, L. C. Biedenharn, and E. Feenberg, Phys. Rev. 106, 1151 (1957).
[4]J. G. Kirkwood, Physik Z. 33, 57 (1932).
[5]P. O. Löwdin, J. Mol. Spectry. 10, 12 (1963).
[6]C. W. Scherr and R. E. Knight, Rev. Mod. Phys. 35, 436 (1963).
[7]A. L. Stewart, Advan. Phys. 12, 299 (1963).
[8]J. O. Hirschfelder, J. Chem. Phys. 39, 2009 (1963).
[9]K. M. Sando and J. O. Hirschfelder, Proc. Natl. Acad. Sci. (U. S.) 52, 434 (1964).
[10]W. J. Meath and J. O. Hirschfelder, J. Chem. Phys. 41, 1628 (1964).
[11]P. O. Löwdin, J. Math. Phys. 6, 1341 (1965).
[12]P. O. Löwdin, Phys. Rev. 139, A357 (1965).
[13]J. Midtdal, Phys. Rev. 138, 1010 (1965).
[14]For a review, see J. O. Hirschfelder, W. B. Brown and S. T. Epstein, Advan. Quant. Chem. 1, 255 (1964).
[15]J. Midtdal, G. Lyslo, and K. Aashamar, Physica Norvegica 3, 163 (1969).
[16]R. L. Matcha and W. B. Brown, J. Chem. Phys. 48, 74 (1968).
[17]For a review, see G. A. Baker, Jr., Advan. Theoret. Phys. 1, 1 (1965).
[18]O. Goscinski, Intern. J. Quant. Chem. 1, 769 (1967).
[19]O. Goscinski and E. Brändas, Proceedings of the Vilnius International Symposium on the Shell Structure of Atoms and Molecules, 1969 (to be published).
[20]A. W. Weiss and J. B. Martin, Phys. Rev. 132, 2118 (1963).
[21]W. H. Miller, J. Chem. Phys. 44, 2198 (1966).
[22]P. Lindner and P. O. Löwdin, Intern. J. Quant. Chem. 2S, 161 (1968).
[23]J. Midtdal, Phys. Math. Univ. Osloensis, No. 32, pg. 1 (1966).
[24]S. Prager and J. O. Hirschfelder, J. Chem. Phys. 39, 3289 (1963).
[25]E. Brändas and O. Goscinski, Intern. J. Quant. Chem. 38, 383 (1969).
[26]R. K. Nesbet, J. Chem. Phys. 43, 311 (1965).
[27]H. Padé, Thèse Ann. Ecole Nor. 9, Suppl. 1 (1892).
[28]O. Goscinski, Intern. J. Quant. Chem. 2, 761 (1968).
[29]For some recent results, see G. A. Baker, Jr., J. Math. Phys. 10, 814 (1969).
[30]O. Goscinski and E. Brändas (to be published).
[31]O. Goscinski and E. Brändas, Chem. Phys. Letters 2, 299 (1968).
[32]T. Kato, Y. Fac. Sci., Univ. Tokyo, Sec. I6, 145 (1951).
[33]J. Midtdal, K. Aashamar, and G. Lyslo, Physica Norvegica 3, 11 (1968).
[34]R. Jentzch, Acta Math. 41, 219 (1917).
[35]F. H. Stillinger, J. Chem. Phys. 45, 3623 (1966).