Information-theoretic approach to the convergence of perturbation expansions

Jeremiah N. Silverman*
Department of Chemistry, University of Ottawa, Ottawa, Ontario, Canada K1N 9B4

Danail Bonchev
Higher School of Chemical Technology, BU-8010 Burgas, Bulgaria

Oskar E. Polansky
Max-Planck-Institut für Strahlenchemie, D-4330 Mülheim (Ruhr), Federal Republic of Germany
(Received 6 February 1986)

Two information-theoretic relationships have been derived for assessing the convergence behavior of arbitrary Rayleigh-Schrödinger (RS) perturbation expansions; these provide sensitive global information indices for quantitatively estimating the *rapidity* and *regularity* of convergence. The procedure has been applied to nine different high-order RS eigenvalue series for the ground states of H^- and He, arising from different partitionings of the Hamiltonian operator for the same system; our findings for the order of merit of these series are in complete accord with previous numerical and theoretical studies. Furthermore, the theory has been applied to an idealized RS series modeled on a geometric progression, thereby gaining insight enabling one to correlate the rapidity with the regularity of convergence for actual RS series; in this context, the so-called doctrine of "small perturbations" has also been investigated, and, in corroboration of earlier work, it is concluded that the doctrine is an unreliable guide for selecting favorable partitionings of the Hamiltonian. To our knowledge, this is the first time that information theory has been applied to the study of perturbational convergence.

I. INTRODUCTION

The central problem of Rayleigh-Schrödinger (RS) perturbation theory is the convergence behavior of the RS perturbation series for the systems under consideration. Conventional RS perturbation theory,[1] however, presents such formidable computational difficulties in proceeding to higher order that, until recently, only low- (first- or second-) order RS series were normally available; hence, the issue of convergence could seldom be investigated quantitatively. The situation has now radically changed with the advent of the new discipline of large-order perturbation theory (LOPT) involving techniques whereby very-high-order (e.g., 100th-order) RS perturbation series can be readily computed for a variety of systems. As a result, RS convergence is currently being intensively studied in the context of LOPT; several recent reviews[2,3] of these developments are available.

Via the methods of LOPT, RS series to almost arbitrarily high order have been generated *exactly* for the eigenvalues of a number of one-particle systems,[2,3] (e.g., the anharmonic oscillator, and the Stark and Zeeman effects for hydrogenic ions). For larger systems, exact RS series cannot be obtained, but excellent approximations to them can be computed by other LOPT methods which combine the variational principle with perturbation theory, such as the Hylleraas[4]-Scherr-Knight[5] (HSK) procedure[6] and the perturbational-variational Rayleigh-Ritz (PV-RR) matrix formalism.[7–9] The HSK procedure has primarily been applied to calculate high-order $1/Z$ expansions[10] for atomic isoelectronic sequences, while the flexible PV-RR formalism has also been recently applied to atomic $1/Z$ expansions,[11] as well as to compute high-order RS series for molecular-orbital eigenvalues in molecular topological studies[12–14] and for the hydrogenic Stark effect.[15]

As these LOPT series become increasingly available, the following three key questions assume ever greater relevance: (1) In general, is a given series convergent, i.e., does it possess a nonvanishing radius of convergence r_*; (2) more practically, how rapidly, if at all, does a series for which it is known that $r_* \neq 0$, actually converge for selected physically significant values of the perturbing parameter λ; and (3) given two or more RS series for the same system, resulting from two or more different perturbational partitionings of the Hamiltonian operator, which of these series has superior convergence properties?

Of course, one can always attempt to answer such questions of convergence by straightforward numerical summation of the series for various values of λ, forming and comparing partial sums of increasingly higher order. Frequently, however, subtle effects arise, particularly in connection with the third question, which cannot be readily investigated with a purely numerical approach: For example, for a given system, one series may converge more rapidly in the lower orders, and then be overtaken by another series in the higher orders.[16] In such cases, one would like to have a procedure for objectively computing a weighted index of convergence which characterizes the *global* behavior of the series under comparison.

The traditional analytical method of investigating series convergence involves singularity studies and radius-of-

 ©1986 The American Physical Society

convergence considerations. Thus, great studies towards answering the first of the above three questions in an *a priori* manner have been made possible by the masterly work of Kato.[17,18] At first glance, it might appear that the Kato general theory should also be a suitable tool for addressing the second and third questions concerning rapidity of convergence.[19] Unfortunately, this is not so because, as has been shown recently,[10] the Kato procedure yields disappointingly weak lower bounds for the r_* of representative RS series, i.e., these bounds are typically an order of magnitude too small. We may illustrate the nature of the difficulty by considering $1/Z$ expansions which Kato proved,[17] in general, to be convergent RS series: For the ground state of the helium isoelectronic sequence, the initial[17(a)] Kato estimate[20] of r_* was $1/7.64=0.131$, thus implying *divergence* of the eigenvalue series for all members of the sequence with $Z\leq 7$ (N^{5+}), very slow convergence for $Z=8$ (O^{6+}), and only slowly increasing rapidity of convergence[19] for $Z=9,10,\ldots$ (F^{7+},Ne^{8+}, . . .). In fact, as has been demonstrated both numerically[21] and theoretically,[10,11] r_* is slightly greater than unity in harmony with the observation[5,22] that the eigenvalue series converges very slowly for $Z=1$ (H^-), moderately rapidly for $Z=2$ (He), and with greatly increasing rapidity of convergence for $Z=3,4,\ldots$ (Li^+,Be^{2+}, . . .). Recently, another method has been put forth[23] for estimating radii of convergence, which yields considerably more accurate results than the Kato procedure. Nevertheless, in view of the complexity of the subject and the many associated difficulties, there remains a real need for the development of additional, completely independent methods for assaying RS convergence.

Qualitatively, one might anticipate some linkage between perturbation theory and information theory. In this context, the use of perturbation theory formally resembles a basic problem in any experimental science, namely, selection of the minimum number of experiments to gain the maximum amount of information about a given system. In perturbation theory, the successive perturbation corrections (orders) may be regarded as the "experiments," and achieving a high rate of convergence to the series limit corresponds to minimizing the number of perturbation orders required.

In this paper we show how these ideas may be quantified: We present and illustrate here a novel method for investigating the *rapidity* of convergence of RS series, based upon combining information theory with perturbation considerations. The new procedure compliments existing ones[2,3] in that it is completely free of the necessity of the complicated study of singularity structure and/or of the asymptotic behavior of large-order series coefficients. In addition, our information-theoretic approach also sheds some light on a related problem, namely, the *regularity* of convergence. To our knowledge, this is the first time that information theory has been applied to the study of perturbation convergence.

This paper is organized as follows. In Sec. II, the theory is developed; in Sec. III, the procedure is tested and illustrated by application to a comparative study of nine different high-order RS perturbation series for the ground states of H^- and He. In Sec. III we also summarize the so-called doctrine of "small perturbations" and contrast it with our procedure; in this context, we study the relationship between rapidity and regularity of convergence by considering a model RS series based on a geometric series. Finally, in Sec. IV our results are discussed and our conclusions are presented.

II. THEORY

A. General considerations

Two fundamental relationships of information theory[24–26] provide our starting point. The first of these, the Shannon equation,

$$\bar{I}=-\sum_{i=1}^{\mu} p_i \log_2 p_i \text{ bits }, \tag{1}$$

specifies the entropy of a probability distribution $P_\mu(p_1,p_2,\ldots,p_\mu)$, where p_i is the probability of the ith outcome and the logarithm is taken to the base 2 in order to obtain the entropy in bits. Here, $\bar{I}$ is a measure of the uncertainty of an experimental outcome or, equivalently, a measure of the information obtained in the experiment which reduces the uncertainty. Therefore, $\bar{I}$ is often called the information index or the information content. Now consider the special case, relevant to this study, where all μ outcomes are equiprobable so that $p_1=p_2=\cdots=p_\mu=\mu^{-1}$. Then, Eq. (1) yields the maximum entropy $\bar{I}_\mu=\log_2\mu$. Furthermore, let two distributions P_μ and P_ν, having, respectively, μ and ν equiprobable outcomes, be compared. One obtains, for the difference in the information content, $\Delta\bar{I}=\bar{I}_\mu-\bar{I}_\nu$,

$$\Delta\bar{I}=\log_2\frac{\mu}{\nu} \text{ bits }, \tag{2}$$

which is the basic equation we shall employ in this study. Note that when $\nu<\mu$, $\Delta\bar{I}>0$, i.e., information is gained due to the reduction in the outcome uncertainty; conversely, when $\nu>\mu$, then $\Delta\bar{I}<0$, corresponding to the loss of information with the increase in the number of possible outcomes.

We now show in a general way how RS perturbation series are susceptible to analysis in information theory. The RS series for an arbitrary quantum-mechanical expectation value E (e.g., the energy) of a given system in any state has the form[1]

$$E=E(\lambda)=\sum_{i=0}^{\infty}\epsilon_i\lambda^i=\sum_{i=0}^{\infty}E_i , \tag{3}$$

where the ϵ_i are the RS expansion coefficients, λ is a real natural (variable) or dummy (fixed at unity) coupling parameter, $E_i=\epsilon_i\lambda^i$, and for convergent series, $E(\lambda)$ is the exact series limit for a given λ; in what follows, we shall fix λ at the various physically significant values of interest, and consider the E_i on the right-hand side of (3). It is convenient to partition the RS series (3) into the partial sums of the leading j orders and of the remainder of the series by writing

$$E=E^{(j)}+\Delta E^{(j+1)},\quad j=0,1,2,\ldots \tag{4a}$$

$$E^{(j)}=\sum_{i=0}^{j} E_i \ , \tag{4b}$$

$$\Delta E^{(j+1)}=\sum_{i=j+1}^{\infty} E_i \ , \tag{4c}$$

with the convention that $E^{(-1)}=0$ and $E=\Delta E^{(0)}$. Note that, in general, the jth-order perturbation correction E_j is given by

$$E_j=E^{(j)}-E^{(j-1)}=\Delta E^{(j)}-\Delta E^{(j+1)} \ . \tag{4d}$$

We may relate Eq. (2) to the RS perturbation series as follows. Consider the RS series of any expectation value summed through an arbitrary order, say the kth. It is crucial to our procedure to recognize that *the different numerical values which can be expected for $E^{(k)}$ are the outcomes of the Shannon scheme.* Since these computed expectation values behave as continuous variables, one cannot work with individual discrete values but rather with a range or interval of values which is proportional to the number of possible outcomes;[24,25] full details of how these intervals are selected according to different criteria are presented in Sec. II B. Since any of these outcomes of $E^{(k)}$ can be expected in summing the RS series through kth order, their probabilities are equal in the corresponding probability distribution $P_\nu^{(k)}(p_1,p_2,\ldots,p_\nu)$. Therefore, in principle, Eq. (2) is applicable for the calculation of the amount of information gained or lost in $E^{(k)}$ as compared to the partial sum through one order lower, $E^{(k-1)}$, with the corresponding probability distribution $P_\mu^{(k-1)}(p_1,p_2,\ldots,p_\mu)$: The uncertainty in the perturbation series summation is reduced when the convergence limit E is approach more closely on passing from $E^{(k-1)}$ to $E^{(k)}$, which implies that $\nu<\mu$; conversely, the uncertainty is increased and some information is lost when a "local divergence" occurs so that $E^{(k)}$ is a poorer approximation than $E^{(k-1)}$, i.e., $\nu>\mu$. Furthermore, the ratio μ/ν in Eq. (2) will be larger for a more rapid approach to the convergence limit of the series, and vice versa. Thus, Eq. (2) provides a quantitative means of calculating the information gained or lost with each added perturbation order, as well as yielding numerical indices characterizing the *rapidity* (rate) of convergence of the series. In addition, Eq. (2) also furnishes numerical estimates of the *regularity* of convergence of the RS series: When the series converges regularly, the standard deviation σ of the information index $\Delta\bar{I}$ will be smaller than that occurring in series having local divergencies as well as other irregular behavior; evidently $\sigma(\Delta\bar{I})$ provides a statistical estimate of regularity.

The above discussion provides a general background for the applicability of the basic Eq. (2) to RS perturbation series. The actual implementation of Eq. (2) can be achieved in a variety of ways. In Sec. II B, we present two such methods.

B. Rapidity and regularity analysis

As previously indicated, in order to apply Eq. (2) to the problem of perturbational convergence, one must specify the number of equiprobable outcomes (i.e., the number of expected variable values) of the perturbation series summed through a given order. For example, based on the accuracy of the perturbation series analyzed in Sec. III, one could treat two outcomes of a series summed through kth order, $E^{(k)}$, as different if they differed by not less than a unit in the eighth digit after the decimal point. As is customary, however, with continuous variables, we work with intervals of values which are proportional to the number of outcomes, rather than directly with the latter. Our information-theoretic approach is thus based upon judicious selection of these intervals.

In the great majority of perturbation series encountered in quantum chemistry, all partial sums $E^{(k)}$ have the *same sign* and their absolute values fall somewhere in the interval $0<|E^{(k)}|<2|E|$, where E is the series limit. Thus, for convenience of presentation, but without loss of generality, we assume that these conditions are fulfilled, and work with absolute magnitudes measured on a scale $2|E|$ in length; at the end of this section, we shall show how this scheme can be readily modified to deal with the anomalous situation where some $E^{(k)}$ differ in sign and/or some $|E^{(k)}|>2|E|$.

To be concrete, consider initially the simple case when the series limit E is negative (e.g., atomic and molecular eigenvalues) and is approached monotonically from, say, above so that for all k, $0>E^{(k-1)}>E^{(k)}>E$. This situation is depicted in Fig. 1 in terms of absolute magnitudes (i.e., $0<|E^{(k-1)}|<|E^{(k)}|<|E|$), where all relevant quantities are shown schematically. It is evident from Eq. (4a), as well as from Fig. 1, that the number of outcomes of two consecutive perturbational summations, $E^{(k-1)}$ and $E^{(k)}$, may be taken as respectively equal to the number of outcomes of the corresponding tails of the series $\Delta E^{(k)}$ and $\Delta E^{(k+1)}$. If we let μ and ν denote, respectively, the number of outcomes of $\Delta E^{(k)}$ and $\Delta E^{(k+1)}$, we can write

$$\mu=K\,|\Delta E^{(k)}| \ , \tag{5a}$$

$$\nu=K\,|\Delta E^{(k+1)}| \ ; \tag{5b}$$

here, K is a proportionality constant giving the number of outcomes per unit width of the tail (in our case, $K\approx10^8$),

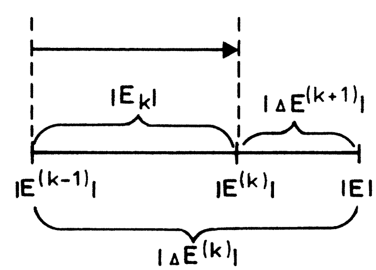

FIG. 1. Monotonic convergence from the left at the $(k-1)$th and kth order of an RS perturbation series. The kth-order perturbation correction, the kth-order summation, the corresponding series tail, and the series limit are, respectively, E_k, $E^{(k)}$, $\Delta E^{(k+1)}$, and E; cf. text for the use of absolute quantities.

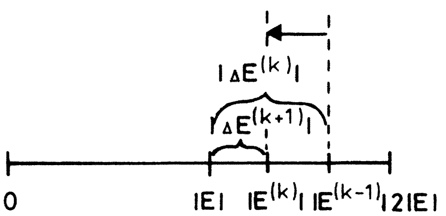

FIG. 2. Monotonic convergence from the right at the $(k-1)$th and kth order of an RS perturbation series. The scale has been enlarged to $2|E|$ on the right.

and one must, of course, take absolute values. Substitution of Eqs. (5) into Eq. (2) yields the first equation of our rapidity and regularity analysis:

$$\Delta \bar{I}_1^{(k)} = \log_2 \left| \frac{\Delta E^{(k)}}{\Delta E^{(k+1)}} \right| \text{ bits}, \quad k=0,1,2,\ldots \qquad (6)$$

At first glance, it might appear that Eq. (6) is limited to the special type of convergent series described above. In the following discussion, however, we remove these restrictions and show that Eq. (6) is, in fact, generally applicable to all types of perturbation series.

(1) Note that since Eq. (6) only involves the absolute magnitudes of the series tails, the sign of the series limit E is irrelevant. Thus, Fig. 1 and Eq. (6) apply equally well to negative E monotonically approached from above or positive E approached from below, both of which may be described as convergence from the left.

(2) Furthermore, the case of monotonic convergence from the right, $|E| < |E^{(k)}| < |E^{(k-1)}| < 2|E|$ (cf. Fig. 2), is also, evidently, entirely equivalent to the case of monotonic convergence from the left, for the same reason as in remark (1) above.

(3) Similarly, the more complex case of alternating convergence, where two consecutive series summations bracket the series limit, i.e., $|E^{(k-1)}| < |E| < |E^{(k)}|$ or vice versa, can be formally regarded as monotonic convergence. This is evident from Fig. 3 where it is seen that the use of the absolute value of, say, $|\Delta E^{(k+1)}|$ is equivalent to the inversion of $|E^k|$ into $|\tilde{E}^{(k)}| = |E| - |\Delta E^{(k+1)}|$ so that $|E^{(k)}|$ and $|\tilde{E}^{(k)}|$ are equidistant from $|E|$, and $|E^{(k-1)}| < |\tilde{E}^{(k)}| < |E|$. Subsequently, we shall present another method of dealing with alternating convergence which is more sensitive to the alternating character of the signs of consecutive perturbation corrections, E_k and E_{k+1}.

(4) Monotonic or alternating convergence implies that, for each k, $|\Delta E^{(k+1)}| < |\Delta E^{(k)}|$. In these situations, Eq. (6) yields $\Delta \bar{I}_1^{(k)} > 0$, thus correctly reflecting that information is gained at each consecutive perturbational order in approaching the series limit more closely.

(5) If for some k, the monotonic or alternating approach to the limit is interrupted and $|\Delta E^{(k+1)}| > |\Delta E^{(k)}|$, which we have previously termed the case of a local divergence, Eq. (6) yields $\Delta \bar{I}_1^{(k)} < 0$, correctly indicating that information is lost in this step in moving away from the series limit; this behavior is illustrated in Fig. 4 for convergence from the left.

(6) Finally, consider asymptotically convergent series which occur frequently.[2,3,27] Here, for some range of k, $0 \le k \le k_{\text{crit}}$, $|\Delta E^{(k+1)}| < |\Delta E^{(k)}|$, and the series converges; subsequently, for $k > k_{\text{crit}}$, $|\Delta E^{(k+1)}| > |\Delta E^{(k)}|$, and the series diverges. Thus, Eq. (6) will correctly exhibit initial gain of information for $0 \le k \le k_{\text{crit}}$, and then loss of information for $k > k_{\text{crit}}$.

(7) The initial condition, i.e., the value to be assigned to $|\Delta E^{(k)}|$ for $k=0$, remains to be specified. For all series considered above, including the case of asymptotically convergent series for k not too much larger than k_{crit}, all partial sums $E^{(k)}$ lie within a range of $|E|$ above or below E. Thus, the number of equiprobable outcomes of $E_0 = E^{(0)}$, before performing the experiment of evaluating E_0, is proportional to $|E|$, and we may take $|\Delta E^{(0)}| = |E|$; formally, this result may be derived from Eq. (4a) by the convention that $E^{(-1)}=0$.

(8) We conclude from the above discussion that Eq. (6) should be applicable to all types of series. In practice, one computes $\Delta \bar{I}_1^{(k)}$ for $k=0,1,2,\ldots$ to the highest order available.[28] These results are then averaged to find the mean value $\langle \Delta \bar{I}_1 \rangle$ and the standard deviation $\sigma(\Delta \bar{I}_1)$, which furnish global indices for, respectively, the rapidity and regularity of convergence; the larger $\langle \Delta \bar{I}_1 \rangle$, the more rapid the convergence, and the smaller $\sigma(\Delta \bar{I}_1)$, the more regular. If two or more RS perturbation series are available for the same system (cf. Secs. III and IV), comparison of their respective information indices provides a

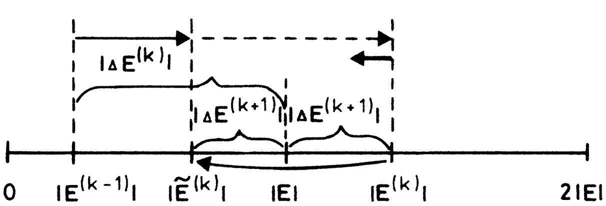

FIG. 3. Alternating convergence at the $(k-1)$th and kth order of an RS perturbation series formally treated as monotonic convergence from the left.

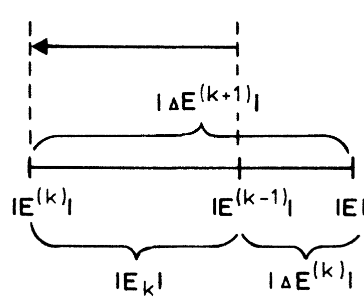

FIG. 4. The case of a local divergence at the $(k-1)$th and kth order of an RS perturbation series.

simple method for determining which series has superior convergence properties, thus answering the key questions raised in Sec. I.

Although offering a convenient procedure for the *a posteriori* analysis of RS perturbation series, Eq. (6) suffers from the following two disadvantages. Firstly, evaluation of the quantities $|\Delta E^{(k)}|$ and $|\Delta E^{(k+1)}|$ requires a rather precise knowledge of the series limit E which, in some cases, may not be available; therefore, Eq. (6) is not well suited for *a priori* analysis of series. Secondly, as previously mentioned, Eq. (6) cannot distinguish between monotonic and alternating convergence, thus possibly losing some of the information content of the series. We now show how both of these difficulties can be overcome simultaneously by a simple reformulation of Eq. (6).

Consider three consecutive perturbational summations which, for later convenience, we reindex as $E^{(k-2)}$, $E^{(k-1)}$, and $E^{(k)}$. If we now regard $E^{(k-1)}$ and $E^{(k)}$, respectively, as the limits approached by $E^{(k-2)}$ and $E^{(k-1)}$, then by precisely the same arguments used in deriving Eq. (6), we obtain via Eq. (4d) our second equation for rate analysis,

$$\Delta \bar{I}_2^{(k)} = \log_2 \left| \frac{E_{k-1}}{E_k} \right| \text{ bits}, \quad k=0,1,2,\ldots . \tag{7}$$

The following comments will be useful in interpreting and applying Eq. (7).

(1) Since the *perturbation orders themselves,* rather than the series tails, appear in Eq. (7), there is no longer any requirement for an accurate value of E. Thus, Eq. (7) lends itself to almost *a priori* analysis, and in any case, at an earlier stage than Eq. (6).

(2) Like Eq. (6), Eq. (7) is applicable to all types of series. Unlike (6), however, (7) exhibits sensitivity to alternating convergence. This is highlighted in Fig. 5, which illustrates typical alternating behavior. Here, $|E^{(k-1)}|$ and $|E^{(k)}|$ lie approximately equidistant from $|E|$, so Eq. (6) would yield $\Delta \bar{I}_1^{(k)} \approx 0$; on the other hand, $|E_k| > |E_{k+1}|$, so Eq. (7) would yield $\Delta \bar{I}_2^{(k+1)} > 0$, thus correctly reflecting the increase in information.

(3) Now consider the initial condition for $k=0$ to be used in applying Eq. (7). For monotonic convergence, i.e., when the low order E_k all display the same sign, we take $|E_{-1}| = |E|$ for the same reasons given previously [cf. remark (7) above] in deriving the initial condition for Eq. (6). If, however, the low-order E_k display alternating behavior by changing signs, it is appropriate to take $|E_{-1}| = 2|E|$ since the zeroth-order term could lie anywhere in the interval $0 \le 2|E|$. In this context, it should be noted that it suffices completely to use a very rough, approximate value of E in the initial condition since it only affects the value of $\Delta \bar{I}_2^{(0)}$.

(4) Aside from the possible distinction in the initial condition, Eq. (7) is applied in the same manner as Eq. (6) [cf. remark (8) above] to compute comparative global indices characterizing the rate and regularity of convergence.

(5) Finally, consider the relationship between Eqs. (6) and (7). Evidently, Eq. (7) can be regarded as an approximation to Eq. (6) obtained by truncating the infinite-series tails to their lead terms. Further, for convergent series, as one proceeds to higher order, Eq. (7) should agree more and more closely with Eq. (6) since the widths of the tails will diminish. Despite this similarity, however, Eq. (7) differs fundamentally from Eq. (6) because of the asymmetry introduced by the former in dealing with alternating series.

Note that in some applications, it is advantageous to apply Eqs. (6) and (7) separately to different portions of the *same* RS series. In general, this will occur when the series under consideration displays different convergence behavior in different regions and/or when the use of the one or the other information index introduces apparent singularities. For example, consider the important special case when the series has primarily an alternating character but one or more perturbation orders E_k vanish identically (cf. Sec. III C): Evidently, Eq. (7) is not defined for vanishing E_k, and it is mandatory to use Eq. (6) to deal with these orders; for the remainder of such a series, however, it is appropriate to use Eq. (7).

In conclusion, for those anomalous cases when some $E^{(k)}$ differ in sign and/or some $|E^{(k)}| > 2|E|$, which can occur due to strong local divergences, we proceed formally as follows. The scale from 0 to $2|E|$ is extended sufficiently to include the anomalous values by adding equal multiples of $|E|$ at the left and right, where the series limit plays the role of a symmetry center to the scale; this procedure is illustrated in Fig. 6. Again a rough value of E suffices. Denote the right and left ends of the extended scale by R and L, respectively, where $R=(n+1)|E|>0$ and $L=-(n-1)|E|<0$, $n=1,2,\ldots$; on subtracting L from R, one then obtains $R-L=n2|E|$. This step is equivalent to a transformation to a new scale ranging from 0 to $(n2|E|+|L|)$ which is shifted on the right by adding $|L|$. Then, in a

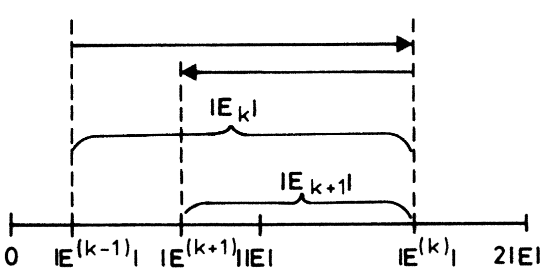

FIG. 5. Direct treatment of alternating convergence at the $(k-1)$th and kth order using the perturbation corrections E_k and E_{k+1}; cf. text and Eq. (7).

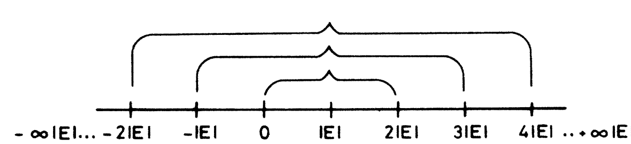

FIG. 6. An extended convergence scale for use in cases of strong local divergences.

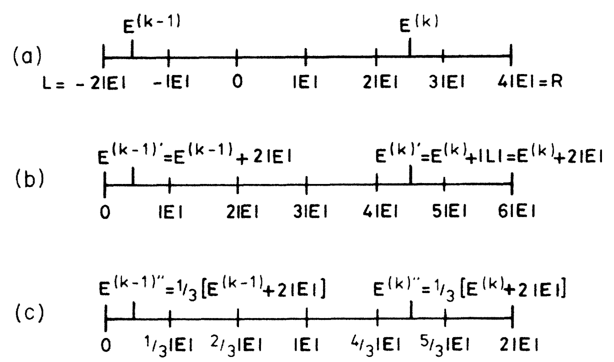

FIG. 7. Transformation of the convergence scale of Fig. 6 to the range from 0 to $2|E|$: (a) the scale before transformation; (b) after the first transformation; (c) after the second transformation.

second transformation, we reduce the scale to the original range of 0 to $2|E|$ by dividing with n. An example of these successive steps is shown in Fig. 7. The information indices $\Delta\bar{I}_1^{(k)}$ and $\Delta\bar{I}_2^{(k)}$ are evidently invariant under these two transformations since Eqs. (6) and (7) only involve the ratios of interval widths which do not change. Thus, after reduction of the data to the new scale, Eqs. (6) and (7) may be applied as before.

III. APPLICATIONS AND RESULTS

A. Types of series considered

In order to subject our information indices, Eqs. (6) and (7), to a rigorous and unbiased test, we have applied them to a comparative study of a number of high-order RS eigenvalue perturbation series computed for different partitionings of the Hamiltonians for the same systems; specifically, we have considered nine such series for the ground states of He and H^-, three for the former presented in Sec. III B, and six for the latter in Sec. III C. The convergence properties of all of these series have previously been thoroughly investigated[16(a),29] by other means, and in this exploratory study, our goal was to determine if our information-theoretic assessments of relative rates of convergence were in accord with these previous findings. Furthermore, in Sec. III D, we have applied our procedure to an idealized RS series modeled on a geometric progression.[30]

The motivation for different partitionings of a given Hamiltonian is, of course, to find that partitioning scheme which leads to the most rapid convergence of the RS series.[31] In this context, there is a well-known, widely used intuitive doctrine which is cited in many contemporary texts[32] on quantum mechanics: This is the doctrine of small perturbations which states that if the perturbation is "small" in some imprecisely defined sense, the perturbational convergence will be rapid, and, as a corollary, if several different partitioning schemes for the same Hamiltonian are compared, that scheme with the smallest perturbation will lead to the most rapid convergence. Since this doctrine is closely related to the subject matter of our study, we now discuss it briefly. Consider two different partitionings of a given Hamiltonian operator H,

$$H = H_0 + H_1 \,, \tag{8a}$$

$$H = \hat{H}_0 + \hat{H}_1 \,, \tag{8b}$$

with corresponding RS eigenvalue perturbation series

$$E = \sum_{i=0}^{\infty} E_i \,, \tag{9a}$$

$$E = \sum_{i=0}^{\infty} \hat{E}_i \,. \tag{9b}$$

A criterion which is often used to define the smallness of a perturbation is the width of the tail of the eigenvalue series from second-order onward (since the zeroth- and first-order terms can be computed from the zeroth-order eigenfunction), i.e., in our notation, $|\Delta E^{(2)}| = |E-(E_0+E_1)|$. If, for example,

$$|\Delta\hat{E}^{(2)}| < |\Delta E^{(2)}| \,, \tag{10}$$

the doctrine would predict that series (9b) converges more rapidly than series (9a). Equation (10) offers the frequently cited justification of partitioning the Hamiltonian via the Dalgarno-Stewart screening transformation[32,1,10] or the choice of the Hartree or Hartree-Fock Hamiltonian[33,16] as H_0. A detailed analysis of the doctrine of small perturbations and its shortcomings is presented elsewhere.[29] Here, we need only note that the series we investigate in this section were all computed to test this doctrine, and that despite a superficial similarity, Eq. (10) differs widely from our information indices, Eqs. (6) and (7). We may anticipate our findings by stating that, in agreement with previous conclusions,[16,29] the doctrine of small perturbations, taken alone, is far from being a reliable guide to rapidity of convergence.

B. RS eigenvalue series for $1s^2\,{}^1S$ He

Riley and Dalgarno[16(a)] (RD) investigated the doctrine of small perturbations by computing three high-order RS eigenvalue series for the ground state of He, resulting, respectively, from a hydrogenic, Hartree, and a modified Hartree partitioning of the Hamiltonian; full details of their partitioning schemes are given in their paper. Both $\Delta\bar{I}_1^{(k)}$, Eq. (6), and $\Delta\bar{I}_2^{(k)}$, Eq. (7), were computed for all three series, and these results are displayed, order by order, in Table I, as well as $\langle\Delta\bar{I}_1\rangle$, $\sigma(\Delta\bar{I}_1)$, $\langle\Delta\bar{I}_2\rangle$, and $\sigma(\Delta\bar{I}_2)$. Table II presents a summary of these results, comparing $\langle\Delta\bar{I}_1\rangle$, $\langle\Delta\bar{I}_2\rangle$, $\sigma(\Delta\bar{I}_1)$, and $\sigma(\Delta\bar{I}_2)$ with the rapidity and regularity of convergence behavior found by RD.

C. RS eigenvalue series for $1s^2\,{}^1S$ H^-

Recently, Silverman and Olbrich[29] (SO) have systematically studied the doctrine of small perturbations by generalizing the Dalgarno-Stewart screening transformation[32] to higher order. Their procedure may be summarized as follows. For atomic isoelectronic sequences,[10] the conventional unscreened $1/Z$ expansions are obtained by parti-

TABLE I. Information indices for the hydrogenic, Hartree, and modified Hartree RS eigenvalue series for $1s^2\,{}^1S$ He. See Tables II, III, and IV (entries for $\alpha=2.3$) of Ref. 16(a); all three are 19th-order series.

Series	Hydrogenic		Hartree		Modified Hartree	
k	$\Delta\bar{I}_1^{(k)a}$	$\Delta\bar{I}_2^{(k)b}$	$\Delta\bar{I}_1^{(k)a}$	$\Delta\bar{I}_2^{(k)b}$	$\Delta\bar{I}_1^{(k)a}$	$\Delta\bar{I}_2^{(k)b}$
0	1.405	0.538	1.443	1.661	8.116	1.005
1	2.834	1.678	4.667	0.840	−2.879	5.442
2	5.286	2.987	2.749	4.409	3.138	−0.366
3	3.280	5.181	2.509	2.716	4.012	3.207
4	1.126	4.306	2.254	2.468	2.665	4.346
5	1.747	0.752	1.846	2.175	3.424	2.289
6	1.678	1.777	1.472	1.756	3.509	3.685
7	1.609	1.710	1.294	1.423	0.720	2.692
8	1.571	1.628	1.238	1.278	3.239	1.566
9	1.577	1.568	1.252	1.242	2.585	2.856
10	1.608	1.562	1.208	1.239	−1.000	2.732
11	1.692	1.568	1.322	1.242		
12	2.115	1.536	1.103	1.256		
13		1.786	1.585	1.239		
14		1.858	0.585	1.263		
15				1.322		
16				1.263		
17				1.322		
18				1.322		
19				1.000		
$\langle\Delta\bar{I}\rangle$	2.118	2.029	1.769	1.584	2.503	2.678
$\sigma(\Delta\bar{I})$	1.075	1.192	0.949	0.841	2.714	1.514

[a]In bits; computed with Eq. (6).
[b]In bits; computed with Eq. (7).

tioning the Hamiltonian hydrogenically so that the interelectronic repulsive potential alone is taken as the perturbation, whereas the screened expansions in $1/\hat{Z}$, $\hat{Z}=Z-s$, result from also including a portion of the nuclear attractive potential, weighted by the factor s, in the perturbation; here s is a disposable screening parameter. The eigenvalue E in ordinary a.u. is related to the unscreened and screened eigenvalues in scaled a.u., ϵ and $\hat{\epsilon}$, respectively, by

$$E=Z^2\epsilon=\hat{Z}^2\hat{\epsilon}\,, \tag{11}$$

where

$$\epsilon=\sum_{i=0}^{\infty}\epsilon_i\lambda^i\,, \tag{12a}$$

$$\lambda=Z^{-1}\,, \tag{12b}$$

$$\hat{\epsilon}=\sum_{i=0}^{\infty}\hat{\epsilon}_i\hat{\lambda}^i\,, \tag{12c}$$

$$\hat{\lambda}=\hat{Z}^{-1}\,. \tag{12d}$$

TABLE II. Comparison of global information indices for the hydrogenic, Hartree, and modified Hartree RS eigenvalue series [see Tables II, III, and IV (entries for $\alpha=2.3$) of Ref. 16(a)] for $1s^2\,{}^1S$ He with observed convergence behavior.

Series	$\langle\Delta\bar{I}_1\rangle^a$	$\langle\Delta\bar{I}_2\rangle^b$	$\sigma(\Delta\bar{I}_1)$	$\sigma(\Delta\bar{I}_2)$	Observed behavior[c,d] Rapidity	Regularity
Modified Hartree	2.503	2.678	2.714	1.514	1	3
Hydrogenic	2.118	2.029	1.075	1.192	2	2
Hartree	1.769	1.584	0.949	0.841	3	1

[a]In bits; computed with Eq. (6).
[b]In bits; computed with Eq. (7).
[c]See discussion in Ref. 16(a).
[d]On a scale where 1 denotes most rapid and most regular, etc.

If s is treated as a constant, Eqs. (11) and (12) yield the standard screening (Euler) transformations relating the screened $\hat{\epsilon}_i$ to the unscreened ϵ_j and s, where s is still arbitrary; in particular, the choice of $\bar{s}=-\epsilon_1/2\epsilon_0$ causes $\hat{\epsilon}_1$ to vanish, and the resultant zeroth-order energy $\hat{Z}^2\hat{\epsilon}_0$ contains the ϵ series correct through first order and a portion of the second-order term. The SO generalization consists of treating s as a function of λ, where

$$s=s(\lambda)=\sum_{i=0}^{\infty} s_i\lambda^i . \tag{13}$$

This yields more complicated transformation equations relating the screened $\hat{\epsilon}_i$ to the unscreened ϵ_j and s_k, which include the standard screening transformations as a special case. The assignment of the values

$$\bar{s}_0=-\epsilon_1/2\epsilon_0 , \tag{14a}$$

$$\bar{s}_1=\frac{1}{2}\left[\bar{s}_0^{\,2}-\frac{\epsilon_2}{\epsilon_0}\right] , \tag{14b}$$

and, in general, the recursively computed

$$\bar{s}_{n-1}=\frac{1}{2}\left[\sum_{j=0}^{n-2}\bar{s}_j\bar{s}_{n-j-2}-\frac{\epsilon_n}{\epsilon_0}\right], \quad n=1,2,\ldots , \tag{14c}$$

to the successive orders s_i, $i=0,1,\ldots,n-1$, Eq. (13), causes the screened $\hat{\epsilon}_1,\hat{\epsilon}_2,\ldots,\hat{\epsilon}_n$ to vanish simultaneously, and the resultant zeroth-order energy $\hat{Z}^2\hat{\epsilon}_0$ contains the ϵ series correct through nth order. Effectively, this generalized screening repartitions the Hamiltonian in a step-by-step manner so as to include more and more of the correlation effects in the zeroth-order problem, thus reducing the width of the series tail $|\Delta\hat{E}^{(2)}|$, Eq. (10), to any desired extent. SO applied their procedure to accurate high-order unscreened RS eigenvalue series[22(a)] for the helium isoelectronic sequence to generate a number of screened series with varying degrees of screening. In particular, we have applied our information-theoretic approach to analyze six of their series, designated as nos. 0,1,2, . . . , 5, for the slowly convergent case of the ground state of H^-; here, in obvious notation, no. 0 refers to the unscreened series, no. 1 to the series where $\hat{\epsilon}_1$ vanishes, no. 2 to the series where $\hat{\epsilon}_1$ and $\hat{\epsilon}_2$ vanish, etc. Equation (7) was used to compute information indices for those portions of these series which exhibit alternating behavior while Eq. (6) was used to deal with vanishing orders and regions of monotonic convergence. Since both Eqs. (6) and (7) were employed in these calculations, it is convenient to suppress the subindex in this case, and to denote the information index as simply $\Delta\bar{I}^{(k)}$, etc. Table

TABLE III. Information indices for six 25th-order RS eigenvalue series (Ref. 29) (nos. 0,1, . . . , 5) for $1s^2\,{}^1S$ H^- with increasing degrees of screening. See Sec. III C.

	$\Delta\bar{I}^{(k)\,a}$					
k	No. 0	No. 1	No. 2	No. 3	No. 4	No. 5
0	0.078	3.260	5.350	6.533	7.264	7.911
1	0.678	0	0	0	0	0
2	1.987	3.486	0	0	0	0
3	4.180	2.246	3.589	0	0	0
4	3.291	0.729	1.194	4.259	0	0
5	0.639	2.751	−0.040	0.889	1.672	0
6	0.577	0.601	0.810	−0.185	0.157	0.517
7	0.512	0.495	0.188	0.685	−0.156	−0.395
8	0.470	0.557	0.618	0.184	0.341	−0.313
9	0.440	0.577	0.211	0.719	0.160	0.052
10	0.415	0.509	0.497	0.363	0.370	0.064
11	0.394	0.494	0.211	0.864	0.217	0.215
12	0.376	0.461	0.418	0.521	0.324	0.231
13	0.360	0.469	0.208	1.242	0.196	0.355
14	0.347	0.417	0.364	0.818	0.260	0.413
15	0.336	0.444	0.204	0.097	0.167	0.606
16	0.325	0.393	0.326	−0.662	0.215	0.848
17	0.316	0.419	0.201	−1.480	0.152	1.894
18	0.308	0.377	0.296	0.258	0.188	−0.182
19	0.300	0.397	0.198	−0.168	0.145	−1.333
20	0.294	0.363	0.274	0.511	0.171	−0.492
21	0.288	0.379	0.194	0.143	0.139	−0.243
22	0.282	0.353	0.256	0.736	0.158	−0.080
23	0.277	0.355	0.191	0.332	0.133	0.021
24	0.272	0.349	0.241	1.142	0.146	0.122
25	0.267	0.372	0.188	0.547	0.127	0.209
$\langle\Delta\bar{I}\rangle$	0.693	0.818	0.626	0.706	0.483	0.401
$\sigma(\Delta\bar{I})$	0.948	0.937	1.163	1.498	1.392	1.595

[a]In bits; computed with both Eqs. (6) and (7); see text.

TABLE IV. Comparison of global information indices for six RS eigenvalue series (Ref. 29) for $1s^2\,{}^1S$ H^- with previous assessment.

Series no.	$\langle \Delta\bar{I} \rangle$ [a]	$\sigma(\Delta\bar{I})$	Order of merit[b,c]
1	0.818	0.937	1
3	0.706	1.498	2
0	0.693	0.948	3
2	0.626	1.163	4
4	0.483	1.392	5
5	0.401	1.595	6

[a]In bits; computed with both Eqs. (6) and (7); see text.
[b]Reference 29.
[c]Rapidity of convergence on a scale where 1 denotes most rapid, etc.

TABLE V. Information indices for an RS eigenvalue series modeled on a geometric progression.

r [a]	$E^{(1)}/E$	n [b]	$\langle \Delta\bar{I}_1 \rangle$ [c,d,e]	$\langle \bar{I}_2 \rangle$ [c,e]	$\sigma(\Delta\bar{I}_2)$
0.1	0.99	7	3.332	2.926	1.048
−0.1	0.99	7	3.332	3.014	0.813
0.3	0.91	15	1.737	1.661	0.296
−0.3	0.91	15	1.737	1.667	0.270
0.5	0.75	26	1	1	0
−0.5	0.75	26	1	0.978	0.110
0.7	0.51	51	0.515	0.538	0.168
−0.7	0.51	51	0.515	0.509	0.038
0.9	0.19	174	0.152	0.170	0.239
−0.9	0.19	174	0.152	0.152	0.006

[a]See Eqs. (15) and (17).
[b]The highest order to which the series must be summed to obtain a relative accuracy of 1×10^{-8}.
[c]In bits; computed with Eq. (17) and appropriate zeroth-order conditions.
[d]The standard deviation $\sigma(\Delta\bar{I}_1)$ vanishes for all r and, hence, is not tabulated.
[e]Averaged over $n+1$ terms.

III displays, order by order, the indices $\Delta\bar{I}^{(k)}$ for these six series, as well as the corresponding $\langle \Delta\bar{I} \rangle$ and $\sigma(\Delta\bar{I})$. Finally, Table IV collects in summary form a comparison of $\langle \Delta\bar{I} \rangle$ and $\sigma(\Delta\bar{I})$ for all six series together with the convergence assessments of SO.

D. RS eigenvalue series modeled on a geometric progression

The geometric progression has frequently been used[30] in atomic calculations to approximate various RS series and to estimate their sums. In order to shed additional light on the relationship between the rapidity and regularity of convergence, we have also computed information indices for an idealized RS eigenvalue series modeled on a geometric progression. Thus, the series has the form

$$E=\sum_{i=0}^{\infty} E_i = E_0 \sum_{i=0}^{\infty} r^i \,, \tag{15a}$$

$$E_i = E_0 r^i \,, \tag{15b}$$

and the convergence limit for $|r|<1$

$$E = E_0/(1-r) \,; \tag{16}$$

the convergence is, respectively, monotonic and alternating for $0<r<1$ and $-1<r<0$. It is elementary[35] to show that in this special case, *both* Eqs. (6) and (7) reduce to the same expression,

$$\Delta\bar{I}_1^{(k)} = \Delta\bar{I}_2^{(k)} = \log_2 \left| \frac{1}{r} \right| \text{ bits, } k=1,2,\ldots \,; \tag{17}$$

for $k=0$, $\Delta\bar{I}_1^{(0)}$ is also given by Eq. (17) but $\Delta\bar{I}_2^{(0)}$, in accordance with our conventions, assumes either the value of $\log_2|1/(1-r)|$ or $\log_2|2/(1-r)|$, depending on whether the convergence is monotonic or alternating.

Our results for $\langle \Delta\bar{I}_1 \rangle$, $\sigma(\Delta\bar{I}_1)$, $\langle \Delta\bar{I}_2 \rangle$, and $\sigma(\Delta\bar{I}_2)$ are presented in Table V for various positive and negative choices of r. The mean values and standard deviations are computed over $n+1$ terms, where n (also collected in Table V) is the highest order for the given value of r to which the series must be summed to achieve a relative accuracy of 1×10^{-8}; Table V also displays the ratio $E^{(1)}/E$ in order to demonstrate the rapidity and regularity of convergence as a function of the smallness of the perturbation.

IV. DISCUSSION AND CONCLUSIONS

First, consider the three RS eigenvalue series[16(a)] for the ground state of He. In accordance with the doctrine of small perturbations, Eq. (10), the Hartree series should converge most rapidly since the Hartree energy $E_{\rm Har}$ is the best first-order approximation[36] to E. In fact, RD found, perhaps to their surprise, that the simple hydrogenic series actually converged more rapidly than the Hartree; to overcome this difficulty, they then introduced the modified Hartree series. By numerical summation, RD showed that the relative order of the rapidity of convergence of the series, in decreasing order, was modified Hartree > hydrogenic > Hartree, requiring, respectively, a 10th-, 14th-, and 19th-order series to converge to within an accuracy of one unit in the eighth decimal place. It is seen from Tables I and II that both of our global information indices, $\langle \Delta\bar{I}_1 \rangle$ and $\langle \Delta\bar{I}_2 \rangle$, reflect precisely this order of rapidity of convergence. Further, both of the standard deviations, $\sigma(\Delta\bar{I}_1)$ and $\sigma(\Delta\bar{I}_2)$, indicate the reverse order of decreasing regularity of convergence, i.e., Hartree > hydrogenic > modified Hartree. Although the Hartree and hydrogenic series differ only slightly in regularity, the modified Hartree series is markedly irregular (as noted by RD); in our approach, this is evidenced by the negative information indices shown in Table I which lead to the large standard deviations for the modified Hartree series. We shall shortly return to the problem of correlating rapidity and regularity of convergence.

It is also of interest to examine the predictive power of the indices $\Delta\bar{I}_1^{(k)}$ and $\Delta\bar{I}_2^{(k)}$, when averaged over only the first two terms, $k=0,1$, in Table I. When this is done, the indices computed with Eq. (6) predict the order of decreasing rapidity of convergence to be Hartree > modified Hartree > hydrogenic, in agreement with the doctrine of small perturbations; one could have anticipated these results since Eq. (6), when restricted to the zeroth- and first-order terms, closely resembles Eq. (10). On the other hand, the indices computed with Eq. (7), when averaged in the same manner, predict the order to be modified Hartree > Hartree > hydrogenic, thus correctly identifying the most rapidly convergent series even at this early stage; this may be attributed to the greater sensitivity of Eq. (7) to the alternating character of the modified Hartree series.

Now consider the six RS eigenvalue series[29,37] for the ground state of H^-. On the basis of radius-of-convergence considerations and numerical analysis, SO concluded that the order of merit for these series was no. 1 > no. 3 > no. 0 > no. 2 > no. 4 > no. 5, where rapidity of convergence played the dominant role in their assessment. It is seen from Tables III and IV that, again, our global information index $\langle \Delta\bar{I} \rangle$ reflects this order of rapidity of convergence perfectly. The order of decreasing regularity is not specified by SO, but we find from $\sigma(\Delta\bar{I})$ that one has no. 1 > no. 0 > no. 2 > no. 4 > no. 3 > no. 5. Note that for the H^- series, unlike the He series, the order of rapidity and of regularity differ but slightly, the sole difference being the displacement of no. 3 from the second to the fifth position due to its large number, i.e., 4, of negative information indices (cf. Table III).

As noted by SO, these six series represent strong counterexamples to the doctrine of small perturbations since the values of $E^{(1)}$ in a.u. for nos. 0,1, . . . ,5, are, respectively, $-0.375\,000\,000$, $-0.472\,656\,250$, $-0.514\,813\,509$, $-0.522\,052\,785$, $-0.524\,316\,896$, and $-0.525\,588\,007$, which monotonically approach the limit $E=-0.527\,751\,016$. Thus, Eq. (10) unequivocally predicts the order of decreasing rapidity of convergence to be no. 5 > no. 4 > no. 3 > no. 2 > no. 1 > no. 0; one also recovers this result by averaging $\Delta\bar{I}^{(k)}$ over the entries for $k=0,1$ in Table III. Despite the steady improvement in the informational start as one proceeds to the higher numbered series, their global convergence behavior is increasingly dominated by oscillatory local divergences leading to the negative information indices shown in Table III.

The question naturally arises as to when, if ever, the intuitive doctrine of small perturbations coincides with reality. Further, to what degree can one correlate rapidity and regularity of convergence? Partial answers to these questions can be obtained by studying the information indices and other data presented in Table V for a model RS series based on a geometric progression. The extreme regularity of a geometric progression is reflected by the fact that both Eqs. (6) and (7) reduce to the same form, Eq. (17); indeed, it is not hard to show that simultaneous satisfaction of Eq. (17) by $\Delta\bar{I}_1^{(k)}$ and $\Delta\bar{I}_2^{(k)}$ is a necessary and sufficient condition for the series to be geometric. It is evident from the entries in Table V that a geometric series fully satisfies the doctrine of small perturbations: As $|r|\to 0$, $E^{(1)}/E\to 1$, $\Delta E^{(2)}\to 0$, and $n\to 0$; as $|r|\to 1$, $E^{(1)}/E\to 0$, $\Delta E^{(2)}\to E$, and $n\to\infty$. This is reflected by the growth of $\langle \Delta\bar{I}_1 \rangle$ and $\langle \Delta\bar{I}_2 \rangle$ as $|r|\to 0$. Further, since the $\Delta\bar{I}_1^{(k)}$ and $\Delta\bar{I}_2^{(k)}$ assume the same constant value for a given r, $\sigma(\Delta\bar{I}_1)$ vanishes, as would $\sigma(\Delta\bar{I}_2)$ were it not for the asymmetric zeroth-order conditions for $\Delta\bar{I}_2^{(0)}$; here, however, as $|r|\to 1$, n increases rapidly, and the averaging effect over many terms drastically reduces $\sigma(\Delta\bar{I}_2)$ as shown in Table V.

We now show how the above analysis provides us with useful insight for interpreting the convergence properties of real RS series from the standpoint of information theory. In general, the actual RS series encountered have a far more complex structure than the simple geometric progression, although it is known[30] that in many cases the tails of these series exhibit approximate geometric behavior. Consider initially a series with a good informational start which then converges rapidly in a nongeometric manner: Since the convergence is taken to be rapid (n small) and nongeometric, the global information indices $\langle \Delta\bar{I}_1 \rangle$ and $\langle \Delta\bar{I}_2 \rangle$ will be relatively large, and the individual $\Delta\bar{I}_1^{(k)}$ and $\Delta\bar{I}_2^{(k)}$ will display considerable fluctuations (i.e., they are not approximately constant as they would be for geometric behavior); it follows immediately that the standard deviations $\sigma(\Delta\bar{I}_1)$ and $\sigma(\Delta\bar{I}_2)$ will also be relatively large. An excellent example of such a series is furnished by the modified Hartree series for He (cf. Table II); another less pronounced example is provided by the no. 3 series for the more difficult case of H^- (cf. Table IV). In this context, the no. 5 series for H^- is a slowly convergent series with nongeometric behavior (seven negative information indices); hence, the excellent informational start, the best for all H^- series, is soon downgraded, $\langle \Delta\bar{I} \rangle$ is relatively small and $\sigma(\Delta\bar{I})$ relatively large; much the same applies to no. 4. Finally, consider series with fair to good informational starts, and approximately geometric tails; examples are provided by the remaining series, hydrogenic and Hartree for He and nos. 0,1,2 for H^-. Here, the primary factor governing the global indices $\langle \Delta\bar{I} \rangle$ and $\sigma(\Delta\bar{I})$ is the approximate value of r in the geometric tail (cf. Table V): For small $|r|$, the convergence will be rapid and $\langle \Delta\bar{I} \rangle$ and $\sigma(\Delta\bar{I})$ will be relatively large, but $\sigma(\Delta\bar{I})$ will not be as large as for nongeometric behavior; for large $|r|$, the convergence will be slow, and both $\langle \Delta\bar{I} \rangle$ and $\sigma(\Delta\bar{I})$ will be relatively small. These predictions are completely borne out by the convergence behavior of the hydrogenic ($r\approx 0.30$) and the Hartree ($r\approx 0.40$) series. Thus, the hydrogenic series has only a fair informational start but its smaller value of r enables it to overtake and surpass the Hartree series, despite the superior informational start of the latter. Series nos. 1, 0, and 2 for H^- may be discussed in a similar manner; the r values of these series are relatively large (respectively, $r\approx 0.75$, 0.80, 0.85) which accounts for their slow convergence[37] and relatively small $\langle \Delta\bar{I} \rangle$ in comparison to the hydrogenic and Hartree states.

It will not have escaped attention that our information-theoretic approach, particularly Eq. (7), resembles the classic ratio test[35] for testing the convergence of a series. There is, however, a fundamental distinction: In the ratio test, one seeks to evaluate the limiting ratio $\lim_{k\to\infty}|E_{k+1}/E_k|$ to determine whether the series converges or not; in our approach, however, we do not discard the values of these ratios (actually, their reciprocals) before the limit is reached, but use them all to obtain a statistical estimate of the global behavior of the series.

We may summarize our results and conclusions as follows.

(1) We have derived two information-theoretic relationships for assessing the convergence behavior of arbitrary RS perturbation series. These provide sensitive global indices for estimating the rapidity and regularity of convergence. Our procedure is simple to implement since it does not depend upon the difficult analysis of singularity structure.

(2) We have applied our procedure to nine different high-order RS eigenvalue series and obtained results in complete agreement with previous numerical and theoretical studies.

(3) We have investigated the doctrine of small perturbations quantitatively from the standpoint of information theory by studying a model RS series based on a geometric series. It is shown that in the case of pure geometric behavior, the doctrine is strictly obeyed. For *real* RS series, however, with nongeometric or mixed behavior, we conclude, in agreement with previous findings, that the doctrine is unreliable.

(4) Further, on the basis of our study of the geometric series, we have been able to correlate, in a semiquantitative manner, the rapidity and regularity of convergence of RS perturbation series. It is shown that, in general, the more rapid the convergence, the more irregular the series.

(5) Finally, our results suggest that it may also be possible to correlate the relative magnitudes of our global information indices with the relative magnitudes of the corresponding radii of convergence of the RS series under consideration.

ACKNOWLEDGMENTS

This work was conceived and largely implemented as a result of a collaboration of the authors at the Max-Planck-Institut für Strahlenchemie, Mülheim/Ruhr, Federal Republic of Germany.

---

*Present address: Department of Chemistry, University of Stirling, Stirling, FK9 4LA, Scotland.

[1]See, for example, J. O. Hirschfelder, W. B. Brown, and S. T. Epstein, Adv. Quantum Chem. **1**, 256 (1964).

[2]J. Zinn-Justin, Phys. Rep. **70**, 109 (1981).

[3]B. Simon, Int. J. Quantum Chem. **21**, 3 (1982); J. Čížek and E. R. Vrscay, *ibid.* **21**, 27 (1982); and other papers in this review issue which is devoted to an international workshop on LOPT.

[4]E. A. Hylleraas, Z. Phys. **65**, 209 (1930).

[5]C. W. Scherr and R. E. Knight, Rev. Mod. Phys. **35**, 436 (1963).

[6]For a modified form of the HSK procedure, see A. Dalgarno and G. W. F. Drake, Chem. Phys. Lett. **3**, 349 (1969); for a discussion of the relationship between the original and modified HSK procedures, see F. C. Sanders, *ibid.* **17**, 291 (1972).

[7]J. N. Silverman and Y. Sobouti, Astron. Astrophys. **62**, 355 (1978).

[8]For the extension of the PV-RR formalism to LOPT, see J. N. Silverman, J. Phys. A **16**, 3471 (1983).

[9]J. N. Silverman and A. H. Pakiari (unpublished).

[10]For a recent survey of atomic $1/Z$ expansions with many references, see J. N. Silverman, Phys. Rev. A **23**, 441 (1981); for these RS series, in Z-scaled units, $\lambda = Z^{-1}$ where Z is the nuclear charge.

[11]J. N. Silverman, B. S. Sudhindra, and G. Olbrich, Phys. Rev. A **30**, 1554 (1984).

[12]I. Motoc, J. N. Silverman, and O. E. Polansky, Phys. Rev. A **28**, 3673 (1983).

[13]I. Motoc, J. N. Silverman, and O. E. Polansky, Chem. Phys. Lett. **103**, 285 (1984).

[14]I. Motoc, J. N. Silverman, O. E. Polansky, and G. Olbrich, Theor. Chim. Acta **67**, 63 (1985).

[15]J. N. Silverman and J. Hinze, Chem. Phys. Lett. (unpublished).

[16](a) For an interesting example of this effect which is considered in detail in Secs. III and IV of this paper, see M. E. Riley and A. Dalgarno, Chem. Phys. Lett. **9**, 382 (1971); (b) see also the earlier calculations of A. W. Weiss and J. B. Martin, Phys. Rev. **132**, 2118 (1963); F. W. Byron and C. J. Joachain, *ibid.* **157**, 1 (1967).

[17](a) T. Kato, J. Fac. Sci. Univ. Tokyo. Sec. I **6**, 145 (1951); (b) Trans. Am. Math. Soc. **70**, 195 (1951).

[18]For an exhaustive review of RS convergence theory, see T. Kato, *Perturbation Theory for Linear Operators,* 2nd ed. (Springer, Berlin, 1976).

[19]In general, for a given RS series, convergence is assured for all λ which satisfy $|\lambda| < r_*$, where the more strongly the inequality is satisfied, the more rapid will be the convergence. Thus, in those cases where r_* can be estimated or determined *accurately,* a useful criterion for gauging rapidity of convergence of RS series is furnished by the normalized index $\tau_\lambda = |\lambda| / r_* < 1$; see J. N. Silverman, B. S. Sudhindra, and G. Olbrich, Ref. 11, for the introduction and quantitative application of the convergence index τ_λ.

[20]For later, more accurate, estimates using the Kato procedure, see Ref. 18, pp. 410–413 and R. Ahlrichs, Phys. Rev. A **5**, 605 (1972). Thus, via a screening modification, Ahlrichs has been able to improve the Kato-type estimate of r_* for the $1s^2\,{}^1S$ state of the helium isoelectronic sequence to the extent that convergence is predicted for $Z \geq 1.98$, implying very slow convergence for the important case of $Z=2$ (He) but still failing for $Z=1$ (H^-); furthermore, as shown in Ref. 10, the estimates of r_* obtained in this manner for other atomic isoelectronic sequences remain unrealistically small.

[21]F. H. Stillinger, Jr., J. Chem. Phys. **45**, 3623 (1966); J. Midtdal, K. Aashamar, and G. Lyslo, Phys. Norwegica **3**, 11 (1968); E. Brändas and O. Goscinski, Int. J. Quantum Chem. **4**, 571 (1970); Int. J. Quantum Chem. Symp. **6**, 59 (1972).

[22](a) J. Midtdal, G. Lyslo, and K. Aashamar, Phys. Norwegica **3**, 163 (1969); (b) J. Midtdal, Phys. Rev. **138**, A1010 (1965).

[23]This method is derived in Ref. 10.

[24]C. Shannon and W. Weaver, *Mathematical Theory of Communication* (University of Illinois Press, Urbana, 1949).

[25] L. Brillouin, *Science and Information Theory* (Academic, New York, 1956).
[26] A. Moshovitz, Bull. Math. Biophys. **30**, 225 (1968).
[27] J. N. Silverman, Phys. Rev. A **28**, 498 (1983).
[28] Due to the loss of significant digits in the calculations, the indices from the highest perturbation orders may not be sufficiently accurate to use.
[29] J. N. Silverman and G. Olbrich, (unpublished).
[30] For a study of the geometric progression in atomic calculations with references to previous work, see J. N. Silverman and J. C. van Leuven, J. Chem. Phys. **67**, 2955 (1977).
[31] In conventional RS perturbation theory (cf. Ref. 1), partitioning the Hamiltonian operator was limited by the necessity of finding a tractable zeroth-order problem; in the modified HSK and PV-RR approaches, however, all partitioning schemes can be treated with equal facility; cf. Refs. 8 and 11–14.
[32] See, for example, A. Messiah, *Quantum Mechanics* (North Holland, Amsterdam, 1969), Vol. II, p. 687.
[33] A. Dalgarno and A. L. Stewart, Proc. R. Soc. London, Ser. A **257**, 534 (1960).
[34] C. Møller and M. S. Plesset, Phys. Rev. **46**, 618 (1934).
[35] See, for example, W. Kaplan, *Advanced Calculus*, 3rd ed. (Addison-Wesley, Reading, Mass, 1984), pp. 316, 317, 340, and 341.
[36] In terms of RS perturbation theory, $E_{\text{Har}}=E_{\text{Har}}^{(1)}=E_{0,\text{Har}}+E_{1,\text{Har}}=-2.861\,683\,55$ a.u., cf. Ref. 16(b); the corresponding first-order values for the hydrogenic and modified Hartree series are $E_{\text{hyd}}^{(1)}=-2.750\,000\,00$ a.u. and $E_{\text{MH}}^{(1)}=-2.826\,701\,49$ a.u.; these three He eigenvalue series converge to the limiting value of $E=-2.903\,723\,725$ a.u.
[37] None of these six H^- eigenvalue series is fully converged at 25th order, although nos. 1, 3, and 0 have, respectively, converged to within 1.7 digits in the sixth, 2.2 digits in the sixth, and 1 digit in the fifth decimal place.