METHODS IN COMPUTATIONAL PHYSICS

Advances in Research and Applications

Edited by

BERNI ALDER
Lawrence Radiation Laboratory
Livermore, California

SIDNEY FERNBACH
Lawrence Radiation Laboratory
Livermore, California

MANUEL ROTENBERG
University of California
La Jolla, California

Volume 2

Quantum Mechanics

1963
ACADEMIC PRESS
NEW YORK AND LONDON

Estimating Convergence Rates of Variational Calculations[1]

CHARLES SCHWARTZ

DEPARTMENT OF PHYSICS, UNIVERSITY OF CALIFORNIA, BERKELEY, CALIFORNIA

I. Introduction

IT IS GENERALLY ACCEPTED that a calculation using a variational principle will give the "best" answer—for a fixed degree of complexity of the approximate solution—for some problem which cannot be solved exactly. But how can one judge *how* good the result from such a calculation really is? This question has certainly been asked many times, but to the best of our knowledge has never been answered in any constructive manner. Complete and precise answers to this question are, of course, not to be expected, since this would involve the complete solution of the (unsolvable) problem; nevertheless, we shall present in this paper

[1] This work was supported in part by the U. S. Atomic Energy Commission and by the U. S. Air Force under grant AF-AFOSR 62-121 monitored by the AF Office of Scientific Research of the Air Research and Development Command; and in part by the Advanced Research Projects Administration through the U. S. Office of Naval Research.

an analytical procedure which, at least semiquantitatively, seems to be both correct and useful.

Now it is well known that for many problems the variational calculation gives rigorously a bound (say an upper bound) on the exact answer. This is an advantage of only limited value and does not really aid one much in deciding just how close the answer is.[2] Thus the early history of the labors involved in some particular problem—consider the ground state of helium as the best known example—consists of a series of results, each of which was either "better" than the previous because of one or a few added parameters, or else more efficient than the previous because a "wiser" choice of terms gave an equally good answer with less effort. Although most people placed a great deal of faith in these results, no reliable estimate of the accuracy was to be had. ("With fourteen parameters one can certainly fit any function extremely well.") However, some people who were more critical tried to point out how very bad, even impossible, these variational wave functions were (Bartlett *et al.*, 1935).

The moderate approach, which is certainly recognized by many people today (see Coolidge and James, 1937), is to steer between these two extremes and admit that the variational calculation is an imperfect fitting of some trial function to the exact, but unknown solution. Then, with a sufficiently (infinitely) flexible trial function it is assumed that one will converge, in some sense which we do not bother to define rigorously (see Kato, 1951), to the exact answer; and the appropriate question is then *how rapidly* does the answer converge?

To be somewhat specific, assume we are considering the Ritz variational principle for the energy of a bound state,

$$E = \frac{(\psi H \psi)}{(\psi \psi)}, \tag{1}$$

and suppose the trial function ψ is represented by a finite number of convenient functions u_n in the time-honored linear manner,

$$\psi = \sum_{n=1}^{N} C_n u_n . \tag{2}$$

Then there will be obtained, as a result of varying the coefficients C_n so as to make (1) stationary, a sequence of energy values E_N which

[2] There are methods which will, in some cases, lead to lower bounds, thus allowing a rigorous bracketing of the exact answer. This generally requires much more labor; and the results, where this labor has been expended (see Kinoshita, 1959), do not seem very rewarding.

will (by assumption) converge to the correct energy value E; and thus what we really want to study is the problem: How fast do the increments $E_N - E_{N-1}$ go to zero as N goes to ∞?

We shall not be able to answer this question exactly; to do so is probably tantamount to completely solving the original unsolvable problem. What we shall attempt will be the construction of approximate formulas for the rate of decrease of these increments in the asymptotic region $N \to \infty$. In practice one never does reach this region, but the use of high-speed computing machines, today and in the future, does allow us to reach regions in which these asymptotic formulas are probably good approximations. Such a program as we are undertaking, to allow one to understand the asymptotic rate of convergence of a large-scale calculation, should be an essential part of current and future high-accuracy calculations. The virtue of this analysis will be twofold. First, it will allow one to estimate beforehand how accurate an answer may be expected from some planned amount of work; thus the wasting of effort in hopeless ventures would be avoided. Secondly, it will add greatly to the confidence in a final result obtained if the observed rate of convergence is shown to be in agreement with some theoretical expectation.

The beautiful work of Pekeris (1958) marks, for us, the dawning of the new era of high-accuracy calculations. What Pekeris did was to carry out the original Hylleraas (1929) program for the calculation of the helium ground state to an accuracy several orders of magnitude beyond any previous attempt—in fact, beyond 1000 terms. To many individuals this work represents the "finishing" of one problem; but to us it means much more, for it allows the introduction of a whole new attitude toward elaborate computations, in which "convergence" is the watchword.

The first essential in talking of convergence rates is to have an orderly plan of procedure.[3] That is, one must choose a set of basis functions to be used and then gradually add more and more of these terms to the variational calculation in some systematic manner. The old habit of picking the "best" (chosen by art) choice of a fixed number of terms is to be discarded if one wants to see how the problem converges. Once one embarks on any very large-scale program, such an orderly plan of attack would be natural just from the bookkeeping considerations; this

[3] We do not wish to imply that Pekeris was the first to plan an orderly calculation, but his work is the most striking, both for its scope and for its success. The work reported by King *et al.* (1958) on the helium ground state certainly looks like an orderly program; however, this work represents an example of false (misleading) convergence as we have recently discussed (Schwartz, 1962).

is, however, essential for any mathematical analysis of the convergence rates.

Pekeris' results on helium do show an extremely smooth behavior, indicating that convergence most probably exists and apparently is quite rapid. His fitting of the convergence rate as an "almost-geometrical series" is, as we shall show, not theoretically correct. This experimental (numerical) approach to the analysis of a completed calculation is certainly to be encouraged, however, since when no theoretical predictions are available, it is the best means of assessing the value of the answers obtained. Yet such experimentally drawn conclusions cannot be proven experimentally (and may sometimes be incorrect[3]) and there will always be skeptics who can doubt the accuracy of any approximate calculation. Thus it behooves us to seek some analytical approach that will lead to at least an estimate of the asymptotic rate of convergence to be expected in any problem.

In the next section we lay the mathematical background for the real problem in the form of several simple model problems; following this, we show how we propose to adapt these simple results to the real problem of variational approximations to solutions of Schrödinger's equation. A number of examples from the study of two-electron atoms shows that our approach, although still rather coarse in some ways, does appear to be both valid and helpful.

II. Model Problems: Integrated Least-Square Fitting of Functions

We shall now study the simple and familiar mathematical problem of the "best" fitting of some given function F with a fixed number N of basis functions

$$u_n\,, \qquad n = 0, 1, 2, \cdots, N-1\,.$$

We require that the mean square error

$$I = \int \left| F - \sum_{n=0}^{N-1} C_n u_n \right|^2 dv \tag{3}$$

should be a minimum under variation of the coefficients C_n. The solution is easily found if one starts with, or arrives at by means of a suitable linear transformation, a set of basis functions which are orthonormal,

$$\int u_n u_{n'}\, dv = \delta_{nn'}\,. \tag{4}$$

Then the appropriate choice for the coefficients C_n is

$$C_n = \int F u_n \, dv \,, \tag{5}$$

and the irreducible value of I (the error at the Nth stage) is

$$I(N) = \sum_{n=N}^{\infty} C_n{}^2 \,. \tag{6}$$

Thus the rate of convergence is the rate at which the squared coefficients C_n go to zero as n approaches infinity. Typical forms for this rate will be

$C_n{}^2 \sim \alpha^n$ (for some parameter $\alpha < 1$), which is the fast "exponential" rate of convergence; and

$C_n{}^2 \sim 1/n^p$, which is the slower "power" rate of convergence. Combinations of these two forms will represent all our present findings.

For our model problems we shall consider functions of one variable

$$0 \leqq r \leqq \infty \,,$$

representative of the radial variable in the three-dimensional Schrödinger equation. The volume element (or metric) will be of the general form

$$dv = r^a \, dr \,; \tag{7}$$

and the trial functions will be chosen to be of the form of powers of r with an exponential envelope, again representative of the common forms used for atomic wave functions.

The appropriate linear combinations of the powers of r which will give the orthogonality relations (4) are some sort of Laguerre polynomials. It will be most convenient for us to represent these basis functions in terms of a generating function; thus we take

$$u_n(r) = A_n(s^n: \frac{\exp\left[-\frac{1}{2} kr(1+s)/(1-s)\right]}{(1-s)^{a+1}} \,, \tag{8}$$

where the symbol $(s^n:$ means "the coefficient of s^n in the following expression." The orthonormality integral is

$$\begin{aligned} \int_0^\infty u_n(r)\, u_m(r)\, r^a \, dr &= A_n A_m (s^n: (t^m: \frac{1}{(1-s)^{a+1}(1-t)^{a+1}} \\ &\quad \times \int_0^\infty r^a \, dr \exp\left[\frac{-kr(1-st)}{(1-s)(1-t)}\right] \\ &= A_n A_m (s^n: (t^m: \frac{a!}{k^{a+1}(1-st)^{a+1}} \,, \end{aligned} \tag{9}$$

which is obviously zero unless $m = n$ and is made equal to unity with the normalization

$$A_n = \left[\frac{(n+a)!}{(n)!\, k^{a+1}}\right]^{-1/2}. \tag{10}$$

The scale parameter k will be left free for use as needed.

With this basis, we shall now take particular functions $F(r)$ and calculate the expansion coefficients C_n according to (5). The manner of grouping and of adding terms will be the natural one in which at the Nth step, we take all polynomials of order $n < N$.

A. Smooth Shape Fitting

First we consider a function F which is infinitely differentiable over the entire region 0-∞, has a power series expansion about the origin (as do the trial functions u_n), and has also an exponential decay at ∞. Here the fitting functions have the easiest job of simply fitting the particular shape and will show the fastest rate of convergence. For example, consider the simple choice

$$F = r^b\, e^{-hr/2}. \tag{11}$$

The integral (5) is easy and the result is

$$C_n = A_n(s^n: \frac{(a+b)!\,(1-s)^b\, 2^{a+b+1}}{[(h+k)-(h-k)\,s]^{a+b+1}}. \tag{12}$$

Since the choice $h = k$ makes the problem too easy and uninteresting, we avoid it. The foregoing result is easily expanded and we get the exact solution in the form of a finite series,

$$C_n = \left[\frac{n!\, k^{a+1}}{(n+a)!}\right]^{1/2} \left(\frac{2}{h+k}\right)^{a+b+1} \sum_{i=0}^{b} (-1)^i \frac{b!\,(a+b+n-i)!}{i!\,(b-i)!\,(n-i)!} \times \left(\frac{h-k}{h+k}\right)^{n-i}. \tag{13}$$

Now because we are really only interested in the form of C_n for very large n, we make the approximation

$$\frac{(n+x)!}{n!} \cong n^x \qquad (n \gg x), \tag{14}$$

and thus have

$$C_n \sim \left[\frac{2k^{1/2}}{h+k}\right]^{a+1} \left[\frac{4k}{k^2-h^2}\right]^b n^{\frac{1}{2}a+b} \left[\frac{h-k}{h+k}\right]^n. \tag{15}$$

The final rate of convergence for this model problem is then

$$\boxed{C_n^2 \sim n^{a+2b}\left(\frac{h-k}{h+k}\right)^{2n}.} \tag{16}$$

This is the exponential convergence which we believe is typical of all problems wherein the trial functions are similar in kind to the function to be fitted, but are different only in details of the shape. Whether this is considered fast or slow convergence in a practical sense depends on how close k is to h, and to a lesser degree on the values of a and b. The convergence will be slowest here when one tries to fit a sharply varying function with slowly varying trial functions ($h \gg k$), or vice versa.

B. Behavior at the Origin

Suppose the function F has a power-series development about $r = 0$ which starts off with some leading power less than that of the trial functions u_n. It will then take a great many of the higher terms in the expansion to build up the proper initial behavior and the convergence will not be very fast. This situation may be typified by repeating the problem of Section II, A, but with the positive exponent b replaced by a negative value, say $-d$. For convergence of the integrals, it is required that $a - 2d$ is nonnegative. The integral for C_n is then of the same form as Eq. (12), but the expression to be expanded is

$$\frac{1}{(1-s)^d\,[1-s(h-k)/(h+k)]^{a-d+1}}. \tag{17}$$

It will now be legitimate to make the simplifying assumption that $h = k$. This may be justified by noting that the coefficient of s^n, for very large n, is governed by the behavior of the expression to be expanded near the limit of convergence (i.e., $s = 1$). Clearly, in the foregoing expression the term $(1 - s)^{-d}$ varies most rapidly in this region. With this choice, we have

$$\begin{aligned} C_n &= \left[\frac{n!\,k^{a+1}}{(n+a)!}\right]^{1/2} \frac{(a-d)!\,(n+d-1)!}{k^{a-d+1} n!\,(d-1)!} \\ &\underset{n\to\infty}{\sim} \frac{(a-d)!}{(d-1)!\,k^{\frac{1}{2}a+\frac{1}{2}-d}} \frac{1}{n^{\frac{1}{2}a-d+1}}, \end{aligned} \tag{18}$$

and the convergence rate is

$$\boxed{C_n^2 \sim 1/n^{a-2d+2}.} \tag{19}$$

It is interesting to note that this result is dependent only on the nature of the integral $\int F^2\,dv$, and is independent of just how many initial powers the fitting functions u_n are lacking. The worst case, $a = 2d$, converges so slowly as to be probably useless for most situations; the remaining error after N terms is $1/N$ and it is generally wholly unrewarding to pursue such an ill-chosen enterprise that shows this behavior.

A special case is that function which has a logarithmic singularity

$$F = r^b \ln re^{-kr/2}\,. \tag{20}$$

Here we get, after a scale change in the integral,

$$C_n = A_n(s^n: \frac{(1-s)^b}{k^{a+b+1}} \int_0^\infty r^{a+b}\,dr\,e^{-r}[\ln(r/k) + \ln(1-s)]\,; \tag{21}$$

and again, since we are only interested in very large n, we can ignore the first term in the integral and get

$$C_n \sim \frac{(a+b)!\,n^{-\frac{1}{2}a}}{k^{\frac{1}{2}a+\frac{1}{2}+b}}(s^n: (1-s)^b \ln(1-s)\,. \tag{22}$$

Now we can write

$$(1-s)^b \ln(1-s) = \frac{d}{db}(1-s)^b\,. \tag{23}$$

The second half of expression (22) is then

$$(s^n:, \text{ etc.} = \frac{d}{db}(-1)^n \frac{b!}{n!\,(b-n)!} = \frac{b!\,(n-b-1)!}{n!}(-1)^{b+1}\,, \tag{24}$$

if we assume b is an integer. Finally, in the large n limit,

$$C_n \sim (-1)^{b+1} \frac{b!\,(a+b)!}{k^{\frac{1}{2}a+\frac{1}{2}+b}} \frac{1}{n^{\frac{1}{2}a+b+1}}\,, \tag{25}$$

and the convergence rate is

$$\boxed{C_n^2 \sim 1/n^{a+2b+2}\,.} \tag{26}$$

This is essentially the same as the result (19).

C. Behavior at Infinity

Suppose the function F does not die off exponentially as r goes to infinity but, say, as some negative power of r; then it will again take many terms of the u_n series to build up the proper tail of the function and we will find again the slower power rate of convergence. Consider the choice

$$F = \frac{1}{(r_0 + r)^d}, \tag{27}$$

wherein we have had to avoid any singularity at the origin in order not to confuse these two separate problems. We will need the restriction $d - a - 1 \geqslant 0$. After a scale change we have

$$C_n = A_n(2/k)^{a+1-d}\,(s^n{:}\int_0^\infty \frac{dx\, x^a\, e^{-x}(1+s)^{d-a-1}}{[(1+s)\,x_0 + (1-s)\,x]^d} \tag{28}$$

$$(x_0 = \tfrac{1}{2}kr_0),$$

which can be written as

$$C_n = A_n(2/k)^{a+1-d}\,(s^n{:}\frac{(d-1)!}{a!}\Big(-\frac{d}{dx_0}\Big)^{d-a-1} \times \int_0^\infty \frac{dx\, x^a\, e^{-x}}{[(1+s)x_0 + (1-s)x]^{a+1}}$$

$$= A_n(2/k)^{a+1-d}\,\frac{(d-1)!}{a!}\Big(-\frac{d}{dx_0}\Big)^{d-a-1}\frac{(n+a)!}{n!a!}\int_0^\infty \frac{dx\, x^a\, e^{-x}}{(x+x_0)^{a+1}}\Big(\frac{x-x_0}{x+x_0}\Big)^n. \tag{29}$$

Another scale change yields

$$C_n = A_n(2/k)^{a+1-d}\,\frac{(d-1)!\,(n+a)!}{(a!)^2\,n!}\int_0^\infty \frac{dy\, y^{d-1}\, e^{-x_0 y}}{(y+1)^{a+1}}\Big(\frac{y-1}{y+1}\Big)^n. \tag{30}$$

This integral over y will get significant contributions from only two regions as n gets very large. The first region is near $y = 0$; here we can approximate

$$\frac{e^{-x_0 y}}{(y+1)^{a+1}} \sim 1,$$

$$\Big(\frac{y-1}{y+1}\Big)^n \sim (-1)^n\, e^{-2ny}. \tag{31}$$

The first portion of the integral is

$$\sim (-1)^n \frac{(d-1)!}{(2n)^d}. \tag{32}$$

The second region is at some $y \gg 1$ and we approximate

$$\frac{y^a}{(y+1)^{a+1}} \sim \frac{1}{y}, \qquad \left(\frac{y-1}{y+1}\right)^n \sim e^{-2n/y}. \tag{33}$$

The remaining integral is

$$\int_0^\infty \frac{dy}{y} e^{-x_0 y} e^{-2n/y}, \tag{34}$$

which is peaked about a point $y = (2n/x_0)^{1/2}$ and has approximately the value

$$\pi^{1/2}(x_0/2n)^{1/4} \exp\left[-2(2nx_0)^{1/2}\right]. \tag{35}$$

The first term (32) will dominate at very large n and is the result of the long tail of the function F; the second part (35) is probably due to fitting the shape in the region of $r = r_0$, and may be the dominant term for some intermediate region of the n series if kr_0 is small. This is an example of the compound structure of the convergence that may often occur.

Putting back all the factors, for this problem we have for $n \to \infty$

$$C_n \sim \left(\frac{(d-1)!}{a!}\right)^2 (-1)^n \left(\frac{2}{k^{1/2}}\right)^{a+1-2d} \frac{1}{n^{d-a/2}}, \tag{36}$$

and the final convergence rate is

$$\boxed{C_n^2 \sim 1/n^{2d-a}.} \tag{37}$$

Thus an expansion of a function F which decays as $1/r$ (with volume element dr) will show the minimal convergence rate $\sim 1/n^2$.

D. Other Discontinuities

We have seen in the foregoing examples that the slower power rate of convergence results when the trial functions do not have the same analytical behavior as the function being represented. This conclusion is generally correct, not only for the end points of the region, but as well for any interior point where the function is not infinitely differentiable (as it is assumed the trial functions are).

The well-known expansion of the delta function (with unit metric, $a = 0$) gives the expansion coefficients

$$C_n = \int u_n(r_0)\, \delta(r - r_0)\, dr = u_n(r_0)\,, \tag{38}$$

where r_0 is the position of the singularity. If this expression is treated to some number, say $b + 1$, of partial integrations, we have

$$\int dr \left[\left(\int\right)^{b+1} \delta(r - r_0)\right] \left(-\frac{d}{dr}\right)^{b+1} u_n(r) = u_n(r_0)\,. \tag{39}$$

The mean value of $u_n(r_0)$ is of the order of magnitude unity, and the mean value of its $(b + 1)$st derivative is of order n^{b+1}. Thus the expansion of a function F which has a finite jump in its bth derivative will have expansion coefficients of the order of magnitude

$$C_n = \int dr \left[\left(\int\right)^{b+1} \delta(r - r_0)\right] u_n(r) \sim 1/n^{b+1}\,, \tag{40}$$

and therefore the convergence rate is

$$\boxed{C_n^2 \sim 1/n^{2b+2}\,.} \tag{41}$$

The boxed results of this section on model problems are probably very well known to a number of people, and can probably be found, in one form or another, in some mathematics texts. The analytical techniques used are really quite simple; and it is our chief purpose to encourage more people, working at computational problems in quantum mechanics (and other fields), to think along these semiformal lines.

III. The Real Problems: Variational Approximations to Solutions of Schrödinger's Equation

Given the Hamiltonian operator H for some particular system, we are really interested in evaluating the quantity

$$J = \int \phi^*(E - H)\phi\, dv\,, \tag{42}$$

and making it stationary with respect to variations of the trial function ϕ. This is the sense in which the variational principle tells us how to get the "best" approximation to the exact eigenfunction ψ. The form (42)

describes, with appropriate specification of the asymptotic behavior of the wave function, both the Rayleigh-Ritz variational principle for the discrete energy eigenvalues and the Kohn-Hulthen variational principle for the scattering problem. The question we wish to investigate is the following. If the trial function ϕ is expressed in terms of N basis functions in the standard form

$$\phi = \sum_{n=0}^{N-1} C_n u_n , \tag{43}$$

how rapidly will J converge to its correct value as N goes to ∞? No complete answers to this question are forthcoming, but we shall now show how the analysis of the preceding section can be used to obtain at least qualitative answers.

Since, by definition,

$$(E - H)\psi = 0 , \tag{44}$$

we can rewrite J in the form

$$J = \int (\psi - \phi)^* (E - H)(\psi - \phi)\, dv . \tag{45}$$

This is now in the form of a generalized least-square fitting problem (3). The operator $(E - H)$ will be looked on as a generalization of the metric of the fitting problem. There are two obvious difficulties in the job of evaluating the stationary value of J. First, the function ψ, unlike the F of our model problems, is not known. Secondly, the appropriate basis functions u_n which are orthonormal with respect to the operator $E - H$ are not known; these are in principle, of course, just the unknown eigenfunctions of H, one of which is the particular ψ under study.

What we have learned in the foregoing is that the asymptotic rate of convergence is controlled by the singularities (broadly speaking) of the function being fitted. All this information—behavior at the origin, behavior at infinity, existence of any discontinuities—is contained in the Hamiltonian, and we simply have to pick these clues out by inspection. The procedure we advocate is as follows. The complete expression (45) is to be butchered into several sections, each containing selected parts of the operators and selected regions of integration. The convergence rate of each section will be separately analyzed as the simple problems of Section II; and then the most slowly converging part will be taken as describing the asymptotic convergence rate of the entire problem. This is certainly not any sort of a rigorous solution to the

general problem; but, if some care is exercised, we expect that this procedure can give the correct rate of the convergence, or, at the very least, it will give some "feeling" for the convergence rate. The numerical coefficients of the results of our model problems should certainly not be used for more than an order of magnitude indication.

The most obvious partitioning of the terms in (45) would be a breakup into the unit, kinetic energy, and potential energy operators,

$$\begin{aligned} J(E) &= \int (\psi - \phi)^* E(\psi - \phi)\, dv\,, \\ J(T) &= -\frac{\hbar^2}{2m} \int (\psi - \phi)^* \nabla^2(\psi - \phi)\, dv\,, \\ J(V) &= \int (\psi - \phi)^* V(\psi - \phi)\, dv\,. \end{aligned} \tag{46}$$

It may then be relatively easy to find the appropriate orthonormal basis for each of these least-square fitting problems.

One danger must be pointed out. Since the foregoing breakup is arbitrarily chosen, one could imagine another in which two terms, each of a very singular character, were made to appear in the separated problems when they might even exactly cancel each other in the complete problem. (This leads one to suspect that the convergence rates calculated may represent upper bounds, but this point is not at all clear.) The rule one must follow to avoid this danger is that the separation is made in order to allow us to focus only on those singularities of the complete problem that are not *precisely* represented by the trial functions.

As an illustration, consider the radial wave function for a state of angular momentum l. The radial operator from the kinetic energy is

$$\frac{1}{r}\frac{d^2}{dr^2} r - \frac{l(l+1)}{r^2}\,. \tag{47}$$

If one were to study these two terms separately, one would get a spurious result, as long as the trial function was designed to have precisely the correct behavior in the region $r \to 0$, namely, $\phi \sim r^l$. Thus we may say that only "uncompensated" singular properties are to be studied. If, in this illustration, the trial function had an initial behavior starting at some power higher than l, then it would be correct to study either of the above terms separately to arrive at the convergence rate characteristic of this particular error.

We readily admit that the procedure we have just outlined gives something rather far removed from a rigorous answer to the

question of convergence rates; there appears to be at least as much art as science involved in the program. But the problem, in its entirety, is a very hard one; and we believe there is much to be gained from even so crude an approach as ours. In support of the merit of these prescriptions we now present several interesting examples drawn from large-scale computations on the two-electron atomic system. The (subsection) listing, A, B, C, and D, of these examples is the same as that of the corresponding model problems of Section II, with which they are represented.

IV. Examples

Pekeris (1958, 1959) has carried out very extensive computations using Hylleraas-type trial functions for the ground states of helium. The form of the trial function is

$$\phi = \exp\left[-\epsilon(r_1 + r_2)\right] \sum C_{lmn} r_{12}{}^l r_1{}^m r_2{}^n \tag{48}$$

where the exponential parameter ϵ is not varied but fixed at the square root of the negative of the energy eigenvalue E (in atomic units). The order of approximation is designated by the value of the parameter ω which gives the maximum power, $1 + m + n$, used at each stage of the work; this replaces the index N previously used.

A. Pekeris' Calculation of the 2^3S State of Helium

The ground state of orthohelium is most simply described as a hydrogenic $(1s, 2s)$ configuration and the simplest wave function would have the form

$$\phi \sim \exp(-Z_1 r_1) \exp(-\tfrac{1}{2} Z_2 r_2)(1 - \tfrac{1}{2} Z_2 r_2) - (\text{exchanged}) \tag{49}$$

where Z_1 and Z_2 are the screened nuclear charges, approximately equal to 2 and 1, respectively. The form (48) used by Pekeris forces the two radial coordinates to have the same exponential coefficient, and this is a very poor starting approximation, particularly for the $2s$ part. The major job of the power series will then be to correct this exponential behavior and we shall interpret this as an example of what we have called smooth shape fitting.

Foley and Traub (1958) have carried out a good variational calculation for this state, varying separately the exponential coefficients of r_1 and r_2, and they find the values $Z_1 = 2.06$ and $\frac{1}{2} Z_2 = 0.60$, while Pekeris' value is $\epsilon = 1.475$. The value for our exponential convergence factor

is $[(h-k)/(h+k)]^2 = 0.18$ for the $2s$ part, and 0.027 for the $1s$ part. The complete calculation, of course, mixes up both these parts, but since the $2s$ part is more slowly converging, we will ignore the other. The other parameters that go into our formula (16) are $a = 2$ and $b = 1$. Since the pertinent region is at large r, we will look only at the E term (unit operator) or the T term (second derivative on r). A small correction should be made because of the fact that Pekeris' specific choice $\epsilon^2 = -E$ causes these two operators to cancel, in their leading terms, for $r \to \infty$; thus the value of a should be reduced by one: $a = 1$. Our prediction for the convergence rate is thus

$$C_\omega^2 \sim \omega^3(0.18)^\omega . \tag{50}$$

TABLE I

COMPARISON OF PREDICTED AND OBSERVED CONVERGENCE RATES IN PEKERIS' CALCULATION OF THE 2^3S STATE OF HELIUM

ω	Pekeris' $[E(\omega) - E(\omega-1)] \times 10^{11}$	Exptl. ratios	$\omega^3(0.18)^\omega \times 10^{11}$	Theoret. ratios
11				
12	827927		855000	
13			200000 } 1101000	
		0.0142	45700	
14				0.0119
15	11792		10300	
16			2275 } 13070	
		0.0120	497	
17				0.0103
18	142		107.4	
19			23.0 } 135.3	
			4.87	

This formula is compared with the differences of Pekeris' results $E(\omega)$ in Table I. Unfortunately, Pekeris gives results for only every third value of ω, and so the comparison cannot be made as finely as we would wish. The things to look at are the ratios of succeeding differences; these are seen to be fairly well represented by our formula(50). The small error is in the sense that the numerical results converge slightly more slowly than (50) predicts. For comparison, the formula

$$C_\omega^2 \sim \omega^4(0.18)^\omega \tag{51}$$

gives the ratios 0.0150 and 0.0126 which are actually closer to the observed ratios, and is in the direction of predicting slightly too slow a convergence rate.

We do not consider our "error" of approximately one unit in the power of the predicted decay law (50) to be a large error, considering the crudeness with which we have reduced the complicated two-electron problem to a simple model problem. An estimate of the numerical coefficient of the theoretical prediction (50) may be obtained from (15), with a factor to normalize the wavefunction (49); and we find a result of approximately $\frac{1}{2}$, which is quite consistent (we would be satisfied with an order-of-magnitude agreement here) with the relative magnitudes of the increments shown in Table I.

It is typical of many problems that at the early, or even at intermediate, stages the convergence is governed by a general shape-fitting problem and appears exponentially convergent. But it may often occur that eventually this part of the problem is well satisfied and the convergence settles down to a slower power rate in response to some weak singularity that contributes only a very small part of the entire problem. This circumstance probably does apply to the 2^3S state, as we shall see in the next section.

B. Pekeris' Calculation of the Ground State of Helium

The ground state of parahelium, described as the $1s^2$ configuration of hydrogen, is not expected to have the severe shape-fitting problem just described; thus we must look farther to find a singularity in the problem which will dictate the asymptotic convergence rate. Fock (1954), and also Bartlett (1937) have pointed out that a formal expansion of the helium wave function will show a weak logarithmic singularity at the point where both electrons approach the nucleus simultaneously; we would like to explain the convergence rate of Pekeris' calculation of the gound state in terms of this behavior.

Fock showed that a power series expansion of the helium wave function about the origin would have the form

$$\psi \sim 1 - Z(r_1 + r_2) + \tfrac{1}{2} r_{12} + (C \sin\alpha \cos\theta + D \cos\alpha) R \ln R + \cdots, \quad (52)$$

where the linear terms are very familiar, but the logarithmic term is something new to be worried about. R is equal to $(r_1^2 + r_2^2)$ and we have evaluated the constants to be $D = 0$, $C = (Z/3)[(1/\pi) - \frac{1}{2}] = -0.121$ for helium. The angle α is such that $R \sin\alpha = 2r_1 r_2$ and θ is the usual angle between the radial position vectors of the two electrons.

Fock does his analysis in a four-dimensional space where R is the radial coordinate and the volume element is $R^3 dR/R$. For our fitting problem we would include the potential energy term $\sim R^{-1/2}$ and then

from (26) with the parameters $a = \frac{3}{2}$, $b = 1$, we get the convergence rate

$$C_n^2 \sim 1/n^{5.5}. \tag{53}$$

This convergence rate is what we would expect if the trial function was a power series in the variable R, but is unrealistic since odd as well as even powers of the coordinates r_1 and r_2 are actually used in the calculation. Thus the convergence rate should be faster than indicated by (53).

Another extreme would be to consider working in the six-dimensional space spanned by the two vectors $\mathbf{r}_1$ and $\mathbf{r}_2$. Here the radial coordinate will be called ρ,

$$\rho = (\mathbf{r}_1^2 + \mathbf{r}_2^2)^{1/2} = R^{1/2}, \tag{54}$$

and the parameters in the fitting problem are rather different;

$$a = 4, \; b = 2,$$

and we get from (26) the convergence rate

$$C_n^2 \sim 1/n^{10}. \tag{55}$$

But this may be expected to be too fast a convergence rate since the trial function used does not have all the pieces of the odd power of ρ.

TABLE II

APPROXIMATE FITTING OF THE OBSERVED CONVERGENCE RATES IN PEKERIS' CALCULATION OF THE 1^1S STATE OF HELIUM

ω	Pekeris' $[E(\omega) - E(\omega-1)] \times 10^9$	Exptl. ratios	$(10/\omega)^8$	Ratios	$(10/\omega)^7$	Ratios
9	489		2.32		2.09	
		0.467		0.431		0.478
10	233		1.000		1.000	
		0.502		0.467		0.514
11	117		0.467		0.514	
		0.53		0.498		0.543
12	62		0.233		0.279	
		1.06		0.984		1.122
13			0.122		0.159	
14	66		0.068 } 0.229		0.095 } 0.313	
15		0.21	0.039	0.204	0.059	0.249
16			0.0233		0.0372	
17	14		0.0143 } 0.0467		0.0244 } 0.0779	
18		0.3	0.0091	0.267	0.0163	0.316
19			0.0059		0.0112	
20	5		0.0039 } 0.0124		0.0078 } 0.0246	
21			0.0026		0.0056	

A correct analysis of this problem must be done working with the actual three coordinates used in the Hylleraas expansion and this is quite complicated. We shall have to content ourselves here with noting that the numerical results of Pekeris' calculation, shown in Table II, converge at a rate well fitted in between our two very rough estimates (53) and (55).

Further evidence for the presence of the logarithmic term in (52) can be found from Pekeris' observations of the numerical coefficients of the linear terms in this expansion. Again we overlook all angular problems and just consider the one-variable fitting of $r^2 \ln re^{-1/2kr}$ by Laguerre functions with a volume element $r^a\,dr$. Let $\tilde{F}(r)$ stand for the function as represented by the best fitting with N terms; we want to look at $\tilde{F}(0)$ and $\tilde{F}'(0)$. As $N \to \infty$, these both go to zero, and we easily find the rates

$$\tilde{F}(0) \sim \frac{(a+1)(a+2)}{k^2} N^{-2}, \tag{56}$$

and

$$\tilde{F}'(0) \sim \frac{(a+2)}{k} N^{-1}. \tag{57}$$

It should not be surprising to find that $F'(0)$ converges so slowly, since $F''(0)$ must diverge. The errors, noted by Pekeris, in the coefficients T and U of the linear terms of (52) do indeed decrease at the rate $1/\omega$ and the numerical value of these errors is of the order of magnitude of what we predict here.

These logarithmic terms will appear in the triplet states only with a higher power of R. Thus only at extremely large ω values, after the shape-fitting problem discussed earlier has been satisfied, will one expect to see a power convergence rate for those states. Then it should be a more rapid rate than that for the singlet states.

C. Zero-Energy Electron-Hydrogen Scattering

For s-wave scattering at zero energy it is well known that the asymptotic form of the wave function for large distances from the scattering center is

$$\psi \sim 1 - A/r + \cdots, \tag{58}$$

where A is the scattering length and the succeeding terms depend on the rate at which the scattering potential dies off. O'Malley *et al.* (1960) have pointed out that for the case of a potential which falls off as $1/r^4$

the next term in (58) should go as $1/r^2$. In the scattering of a charged particle from a neutral system the same conclusion holds, the $1/r^4$ potential being a representation (in the adiabatic approximation) of the effects of polarizing the target. In recently reported calculations (Schwartz, 1961), the Kohn variational principle was applied to calculations of electron scattering from the hydrogen atom. The trial function used was of the form (at zero energy)

$$\phi = 1 - A/r + \sum C_n u_n , \tag{59}$$

where the u_n were exponentially damped polynomials in the coordinates r. The convergence rate of the results was very slow; this may be explained as the problem of fitting the $1/r^2$ tail of the correct ψ. According to our formula (37), with the parameters $a = 2$, $d = 2$, we expect the convergence rate

$$C_n^2 \sim 1/n^2 . \tag{60}$$

Formula (60) does roughly describe the rather limited data on the observed rate of decrease of the increments in the variational calculation of the scattering length A (see Fig. 1). On the basis of this fit an extra-

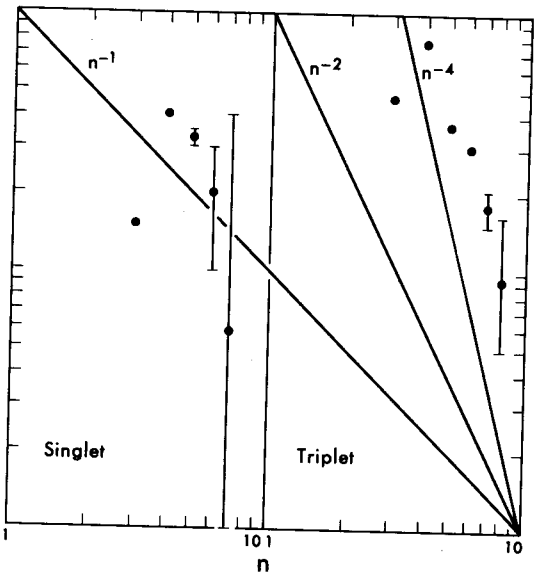

FIG. 1. First calculation of e-H scattering lengths. Vertical bars indicate numerical uncertainties. Straight lines whose slope represents the simple power rate of convergence $1/n^p$ are drawn for comparison.

polation was made according to the general formula

$$\sum_{n=N}^{\infty} T_n \approx \frac{N}{p-1} T_N, \qquad \text{if } T_n \approx 1/n^p. \tag{61}$$

An improved calculation was carried out in which the $1/r^2$ part of the wave function was explicitly represented. These results converged much faster and proved the approximate validity of the above extrapolation. The convergence rate of these improved results should be governed by the next terms in (58) after those accounted for exactly; this would be a $1/r^3$ term from the quadrupolar polarizability and would lead to an expected rate of convergence ($a = 2$, $d = 3$)

$$C_n^2 \sim 1/n^4, \tag{62}$$

which is again in reasonably good accord with results of the computations (see Fig. 2) and helps reassure us of the accuracy of the final result obtained.

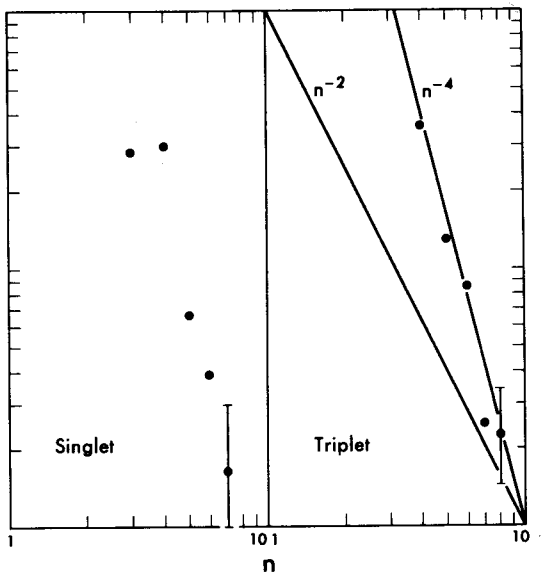

FIG. 2. Second, improved, calculation of e-H scattering lengths. Vertical bars indicate numerical uncertainties. Straight lines whose slope represents the simple power rate of convergence $1/n^p$ are drawn for comparison.

D. Angle-Independent Approximation in the Helium Ground State

If a trial wave function for the two-electron atom is constructed using only the two radial coordinates r_1 and r_2, the effective Hamiltonian is the angular average of the original one and the interaction potential-energy term is $1/r_>$, where the subscripts $<$, $>$ designate, respectively, the lesser and the greater of the two distances r_1 and r_2. This function has a finite discontinuity in its first derivative at the point $r_1 = r_2$, and is the remnant of the singularity of the Coulomb potential after the angular integration. It will then be required that the correct wave-function for this reduced problem have a finite jump in its third derivative.

If one uses as trial functions the completely continuous basis functions

$$r_1^m r_2^n \exp\left[-k(r_1 + r_2)\right], \tag{63}$$

then a convergence rate according to some power law is to be expected.

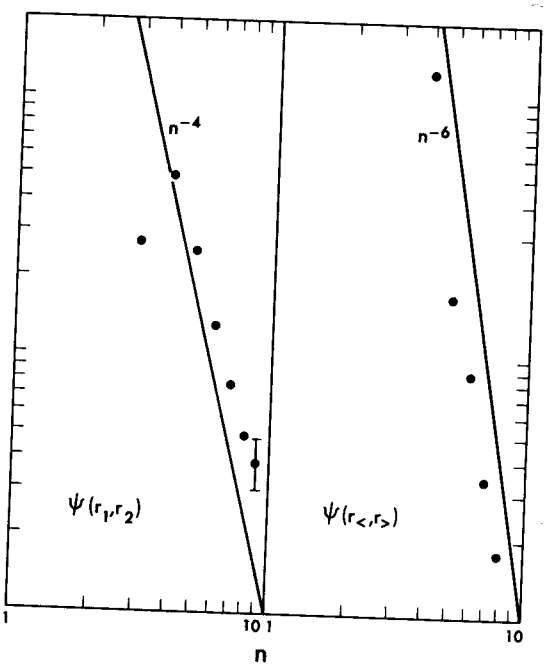

FIG. 3. Two calculations of the angle-independent approximation to the helium ground state. Vertical bars indicate numerical uncertainties. Straight lines whose slope represents the simple power rate of convergence $1/n^p$ are drawn for comparison.

Allowing for the fact that we are expanding through the Laplace operator, we predict from (41) the convergence rate

$$C_n^2 \sim 1/n^6 . \tag{64}$$

The results from such a variational calculation are shown in Fig. 3 along with the results of a second calculation using the more flexible basis functions

$$r_<^m r_>^n \exp[-k(r_1 + r_2)] . \tag{65}$$

(These results have been previously reported by Schwartz, 1962.) It appears that the first problem converges experimentally as $1/n^4$ and the second as $1/n^6$. We are unable to account for this apparent error in our analysis; but again we will be satisfied at seeing that we have at least some qualitative feeling for the convergence rate in this problem.

V. Further Calculations on the Two-Electron Atom: The l-Expansion in the Angle

We now consider another problem which, although not essentially related to variational problems, is concerned with a convergence rate. Consider the ground state of a two-electron atom described in terms of an expansion in inverse powers of the nuclear charge Z. The zeroth-order wave function, in units of a_0/Z for length, is simply

$$\psi_0 = \frac{4}{4\pi} \exp[-(r_1 + r_2)] , \tag{66}$$

and we are interested in studying the first-order correction ψ_1 defined by the inhomogeneous Schrödinger equation

$$(E_0 - H_0)\psi_1 = (H_1 - E_1)\psi_0 , \tag{67}$$

where $E_0 = -1$, $E_1 = 5/8$, $H_1 = 1/r_{12}$, and we are using units of e^2Z^2/a_0 for energy. Of interest is the second-order energy E_2, which is readily obtained once ψ_1 is known,

$$E_2 = \int \psi_0(H_1 - E_1)\psi_1 \, dv . \tag{68}$$

The problem may be reduced by expanding ψ_1 in a series of Legendre polynomials of the angle θ between the two electrons.

$$\begin{aligned} \psi_1 &= \sum_{l=0}^{\infty} \psi_1^{(l)}(r_1, r_2) P_l(\cos\theta) , \\ E_2 &= \sum_{l=0}^{\infty} E_2(l); \end{aligned} \tag{69}$$

$$\left[\tfrac{1}{2}\nabla_1^2 + \tfrac{1}{2}\nabla_2^2 + \frac{1}{r_1} + \frac{1}{r_2} - 1\right]\psi_1^{(l)}(r_1, r_2) = \left[\frac{r_<^{\,l}}{r_>^{\,l+1}} - \tfrac{5}{8}\delta_{l,0}\right]\psi_0,$$

$$\nabla^2 \to \frac{1}{r^2}\frac{\partial}{\partial r}r^2\frac{\partial}{\partial r} - \frac{l(l+1)}{r^2}, \tag{70}$$

$$E_2(l) = \int \psi_0 \frac{r_<^{\,l}}{r_>^{\,l+1}}\psi_1^{(l)}\frac{dv_1\,dv_2}{(2l+1)} \qquad (l > 0).$$

This two-dimensional problem has never been solved; what we shall attempt is to get an approximate solution valid in the limit of large l in order to see how rapidly the series (69) for E_2 converges. First, look at the inhomogeneous term in the differential equation (70) for $\psi_1(r_1, r_2)$. For large l values the quantity $r_<^{\,l}/r_>^{\,l+1}$ is strongly peaked about the point $r_1 = r_2$ and falls rapidly to zero as one goes away from this point. This is simply an expression of the singularity of the Coulomb potential, and it tells us that the important contributions to the function $\psi_1^{(l)}$ come from the immediate neighborhood of this point.

In order to look more closely at this region we introduce the Hylleraas coordinates

$$s = r_1 + r_2, \qquad t = -r_1 + r_2. \tag{71}$$

Setting

$$\psi_1^{(l)} = \psi_0 f_l(s, t), \tag{72}$$

we have the differential equation

$$\left[\frac{d^2}{ds^2} + \frac{d^2}{dt^2} + \frac{4s}{s^2 - t^2}\frac{d}{ds} - \frac{4t}{s^2 - t^2}\frac{d}{dt} - \frac{4l(l+1)(s^2 + t^2)}{(s^2 - t^2)^2} - 2\frac{d}{ds}\right] f_l$$

$$= 2\frac{(s - |t|)^l}{(s + |t|)^{l+1}}. \tag{73}$$

We now make a further change of variables introducing $x = t/s$ instead of t. The term on the right-hand side is now

$$\frac{2}{s}\frac{(1 - |x|)^l}{(1 + |x|)^{l+1}}, \tag{74}$$

and we want to use the expansion

$$\left(\frac{1 - x}{1 + x}\right)^P = \exp\left\{-2Px\left(1 + \frac{x^2}{3} + \frac{x^4}{5} + \cdots\right)\right\}. \tag{75}$$

Now, making the final scale change

$$y = |x|\lambda^{1/2}, \qquad \lambda = (l + \tfrac{1}{2})^2, \tag{76}$$

we have the equation

$$\left[s^2\frac{d^2}{ds^2} - 2y\frac{d}{dy}s\frac{d}{ds} + y^2\frac{d^2}{dy^2} + 2y\frac{d}{dy} - 2s^2\frac{d}{ds} + 2sy\frac{d}{dy}\right.$$
$$\left. - \frac{4}{(1-y^2/\lambda)}\left(2y\frac{d}{dy} - s\frac{d}{ds}\right) + \lambda\frac{d^2}{dy^2} - 4\left(\lambda - \frac{1}{4}\right)\frac{(1+y^2/\lambda)}{(1-y^2/\lambda)^2}\right] f_l$$
$$= \frac{2se^{-2y}\exp\left[-(2y^3/3\lambda) - (2y^5/5\lambda^2) - \cdots\right]}{(1-y^2/\lambda)^{1/2}}. \tag{77}$$

In removing the absolute value sign we are left with the boundary condition

$$\frac{d}{dy}f_l\Big|_{y=0} = 0. \tag{78}$$

The variable y has the range $0 - \lambda^{1/2}$, but because of the term e^{-2y} on the right the effective range of f_l is only over values of y of order unity. It is then completely straightforward to make an expansion in inverse powers of λ:

$$f_l = \lambda^{-1}f^{(-1)} + \lambda^{-2}f^{(-2)} + \cdots. \tag{79}$$

The equation of leading order, determining $f^{(-1)}$, is

$$\left[\frac{d^2}{dy^2} - 4\right] f^{(-1)} = 2se^{-2y}, \tag{80}$$

which we easily solve.

$$f^{(-1)} = -\tfrac{1}{4}se^{-2y}(1+2y). \tag{81}$$

With only slightly more labor, we find the next term

$$f^{(-2)} = -\tfrac{1}{4}se^{-2y}[(-\tfrac{1}{3}y^4 - \tfrac{4}{3}y^3 - y^2 + \tfrac{3}{4}y + \tfrac{3}{8}) - s(\tfrac{2}{3}y^3 + y^2 + y + \tfrac{1}{2})]. \tag{82}$$

The desired result is $E_2(l)$, Eq. (70), which in the new variables is

$$E_2(l) = \frac{1}{\lambda}\int_0^\infty ds\, s^4e^{-2s}\int_0^{\lambda^{1/2}} dy\, e^{-2y}\frac{(1-y^2/\lambda)^2}{(1-y^2/\lambda)^{1/2}}f_l\exp\left[-\frac{2y^3}{3\lambda} - \frac{2y^5}{5\lambda^2}\cdots\right]. \tag{83}$$

This also allows of a power series expansion in λ^{-1} and the first two terms yield our desired result

$$E_2(l) = -\frac{45}{256}\frac{1}{(l+\frac{1}{2})^4}\left[1 - \frac{19/8}{(l+\frac{1}{2})^2} + O\left(\frac{1}{l^4}\right)\right]. \tag{84}$$

This shows the rather slow l^{-4} rate of convergence which we have discussed in a recent paper (Schwartz, 1962). The second term indicates that one does not have to go to very large l values before this asymptotic formula is usable.

VI. Concluding Remarks

We have seen how a store of simple model expansion problems can serve as a rough guide for the analysis of the complicated question of the convergence rate in several problems of real interest. The connections we have made between the models and the real problems are certainly far from being at all rigorous, but it is our belief that the type of crude analysis we have applied does have some considerable validity. We would be very happy to see this procedure put in a more definite and proper form; however, for the present we think it will be already a large step forward if many people, working at such large-scale computations, would learn to think, even in these rough terms, of the convergence-rate question as an important problem to be faced.

Even if no thought whatsoever is given to an attempt at predicting the expected convergence of some problem, the "experimental" procedure of finding the convergence rate shown by a set of results is very valuable. When the results of an attempted computation appear poorly convergent, one should, rather than simply quitting, start looking back to find what element of the problem was responsible for the slow convergence. In this author's experience several poorly initiated computations have, upon further study of the theoretical aspects of convergence rates, led to a clearer understanding of the problems and, eventually, to a satisfying conclusion of the original calculation. In any case, an attempt to extrapolate one's results is always worthwhile. This will generally lead not only to one's best guess as to the exact answer, but can also yield an estimate of the residual uncertainty in the final answer quoted.

One other piece of purely empirical advice may be of use to others: it seems that quite generally, working with computing machines of finite numerical accuracy, a poor choice of basis functions (implying a slow convergence rate) will be signaled by large and rapidly growing roundoff errors. The best cure for this disease will lie not in any improvements in the numerical capacity of the machine, but requires the selection of a basis more appropriate for the problem at hand. Thus it appears that the two problems of getting an accurate (numerical) answer at each step, and having these successive answers converge rapidly, are strongly interwoven.

ACKNOWLEDGMENT

The author wishes to thank Dr. C. Zemach for many very helpful conversations.

REFERENCES

BARTLETT, J. H. (1937). *Phys. Rev.* **51**, 661.
BARTLETT, J. H., GIBBONS, J. J., and DUNN, C. G. (1935). *Phys. Rev.* **47**, 679.
COOLIDGE, A. S., and JAMES, H. M. (1937). *Phys. Rev.* **51**, 855.
FOCK, V. A. (1954). *Izvest. Akad. Nauk S.S.S.R. Ser. Fiz.* **18**, 161.
FOLEY, H. M., and TRAUB, J. (1958). *Phys. Rev.* **111**, 1098.
HYLLERAAS, E. A. (1929). *Z. Physik* **54**, 347.
KATO, T. (1951). *Trans. Am. Math. Soc.* **70**, 212.
KING, K. M., TYCKO, D. H., and THOMAS, L. H. (1958). *Phys. Rev.* **109**, 369.
KINOSHITA, T. (1959). *Phys. Rev.* **115**, 366.
O'MALLEY, T. F., ROSENBERG, L., and SPRUCH, L. (1960). *Phys. Rev. Letters* **5**, 375.
PEKERIS, C. L. (1958). *Phys. Rev.* **112**, 1649.
PEKERIS, C. L. (1959). *Phys. Rev.* **115**, 1216.
SCHWARTZ, C. (1961). *Phys. Rev.* **124**, 1468.
SCHWARTZ, C. (1962). *Phys. Rev.* **126**, 1015.