THE
PHYSICAL REVIEW

A journal of experimental and theoretical physics established by E. L. Nichols in 1893

SECOND SERIES, VOL. 166, NO. 5 25 FEBRUARY 1968

Variational Principles for Expectations

SANFORD ARANOFF
Physics Department, Rutgers University, Newark, New Jersey

AND

JEROME K. PERCUS*
Courant Institute of Mathematical Sciences, New York University, New York, New York

(Received 23 August 1967)

Variational principles for expectation values of physical quantites other than the energy are derived. The expressions implicitly require Green's-function estimations, and higher-order corrections are available. One principle involves a subsidiary minimization, and the other involves the difference between two quantities which are minimum at the stationary point; consequently, numerical computations can be made with both of these principles. As a simple example, the mean-square radius of the hydrogen atom for an incorrect wave function is corrected, with excellent results. Application is also made to the mean-square radius of a model triton.

I. INTRODUCTION

ONE can obtain an excellent approximation to the ground-state energy of a quantum system, even if the wave function cannot be determined, by constructing a variational principle for the energy. This principle, the Rayleigh-Ritz variational principle, consists of selecting from a given class of functions that which minimizes a certain expression, namely, the energy expectation itself. The ground state at fixed values of constants of the motion is treated in precisely the same way, so that many excited state energies are available as well. A wave function which is good for the energy may be poor for other quantities,[1] but unfortunately there is no analogous known method for approximating the expectation of an arbitrary physical quantity by seeking a wave function which minimizes an appropriate expression for this expectation. Methods do exist, however, which permit the calculation of quantities which are probably good approximations to the desired expectation.[2–4] These methods require finding the stationary point of an expression rather than its minimum and are called variational principles, in contrast to the Rayleigh-Ritz principle, which is a minimum principle. In both cases, if one makes a change of order ϵ in the wave function, the approximation will change by order ϵ^2. Although the variational approach lacks rigor, in that the approximation may actually be quite far from the correct value and in an unknown direction, one can say that it is very probable that an improved value results. It is necessary to use variational principles, for there may be no other means of finding an accurate value of the desired expectation or of estimating the error obtained by using a wave function determined by the Raleigh-Ritz principle. There are methods[5] of finding upper bounds to errors in expectations, but these are likely to be very conservative.

Variational principles for expectations have been discussed by many people. Biedenharn and Blatt[6] have derived a variational principle which, however, requires a complete set of trial wave functions. Frost *et al.*[7] have given a procedure for finding an approximate wave function by minimizing the energy dispersion at a finite set of points. Others who have derived variational

* Supported in part by the U. S. Atomic Energy Commission, Contract No. AT(30-1)-1480.

[1] C. Ekart, Phys. Rev. **36**, 878 (1930); W. Kauzman, *Quantum Chemistry* (Academic Press Inc., New York, 1957); P. O. Löwdin, Ann. Rev. Phys. Chem. **11**, 107 (1960); H. Preuss, Z. Naturforsch. **16a**, 598 (1961).

[2] M. Cohen and A. Dalgarno, Proc. Roy. Soc. (London) **A280**, 258 (1964). See this paper for references to previous work.

[3] Y. Rasiel and D. R. Whitman, J. Chem. Phys. **42**, 2124 (1965); D. D. Cheng and Y. Rasiel, *ibid.* **44**, 1819 (1966).

[4] J. Musher, J. Chem. Phys. **46**, 1 (1967). Here it is shown that various perturbation methods for atoms and molecules correspond to keeping leading terms of various Taylor-series expansions of an inverse operator.

[5] S. Aranoff and J. K. Percus, Nucl. Phys. **A98**, 263 (1967).

[6] L. C. Biedenharn and J. M. Blatt, Phys. Rev. **93**, 230 (1954).

[7] A. A. Frost, R. E. Kellogg, B. M. Gimarc, and J. D. Scargle, J. Chem. Phys. **35**, 827 (1961).

Copyright © 1968 by The American Physical Society.

principles are Roberts,[8] Vitchinkin,[9] and Sandro and Hirschfelder.[10] Delves[11] has obtained a variational principle for the expectation values of arbitrary operators. His principle has the drawback that it involves a saddle point, and so in general a computer cannot be used to obtain numerical results. A computer can reasonably be used only to find the minimum of an expression, not a saddle point. Bazley and Fox[12] have derived a variational principle which involves the integral of a Green's function. Such an integral can be performed by a computer for a one-dimensional problem; it is not practical for a problem in three dimensions.

In this paper, a variational principle for arbitrary physical quantities is developed which assumes the form of quantities to be minimized and allows one to use a computer to find the stationary point. In Sec. II, Delves's variational principle is rederived. Possible means of forming corrections to the wave function are discussed in Sec. III. These corrections involve an appropriate Green's function. In Sec. IV, a suitable variational expression for this Green's function is established. This is then used in Sec. V to find a variational expression for expectations, involving a subsidiary minimization. The hydrogen atom is taken as a simple test case in Sec. VI. In Sec. VII, the variational principle is reduced, instead, to a pair of minimization problems—rather than having a single subsidiary minimum as in Sec. V—by exploiting an obvious property of the saddle point. In Sec. VIII, the method is tested with the hydrogen atom and compared with the method of Sec. VI. Application to a model triton is given in Sec. IX, and methods and conclusions are summarized in Sec. X.

II. PROTOTYPE VARIATIONAL EXPECTATION

The most straightforward approach to determining a variational expression for an arbitrary expectation is to relate it to an appropriate energy expectation. Indeed, one knows from first-order perturbation theory that as $\lambda \to 0$, the ground-state energy of $H+\lambda Z$ [assuming nondegenerate $E_0(H)$] is given by

$$E_0(H+\lambda Z)=E_0(H)+\lambda\langle Z\rangle_\psi, \tag{2.1}$$

where

$$\langle Z\rangle_\psi \equiv \langle\psi|Z|\psi\rangle \tag{2.2}$$

denotes the expectation of Z in the true ground state of H. Thus

$$\langle Z\rangle_\psi = \frac{\partial}{\partial\lambda} E_0(H+\lambda Z)\Big|_{\lambda=0}. \tag{2.3}$$

Suppose that ϕ is an approximate ground-state wave function for the system Hamiltonian H; thus we must of course use the ansatz $\phi+\lambda\chi$ to represent the state of $H+\lambda Z$ and vary χ for a minimum. This clearly leads to the approximation

$$\bar{Z}=\frac{\partial}{\partial\lambda}\frac{\langle\phi+\lambda\chi|H+\lambda Z|\phi+\lambda\chi\rangle}{\langle\phi+\lambda\chi|\phi+\lambda\chi\rangle}\Big|_{\lambda=0}, \tag{2.4}$$

or, since we may choose ϕ as normalized and χ orthogonal to ϕ,

$$\langle\phi|\phi\rangle=1, \quad \langle\phi|\chi\rangle=0, \tag{2.5}$$

we have the result[11]

$$\bar{Z}=\langle\phi|Z|\phi\rangle+\langle\phi|H|\chi\rangle+\langle\chi|H|\phi\rangle. \tag{2.6}$$

The expression (2.6) is in fact variational with respect to both ϕ and χ. Inserting (2.5) via Lagrange parameters and varying with respect to χ^* and ϕ^*, the stationary functions, say, ξ and ψ, are given by

$$\begin{aligned} H\psi&=E\psi,\\ (Z-\gamma)\psi&=(E-H)\xi,\\ \bar{Z}&=\langle\psi|Z|\psi\rangle. \end{aligned} \tag{2.7}$$

For solutions to exist, the Lagrange parameters E and γ become energy and $\langle Z\rangle$, respectively. More to the point, if we expand about the stationary solutions

$$\begin{aligned} \phi&=\psi+\delta\phi,\\ \chi&=(E-H)^{-1}(Z-\langle Z\rangle)\psi+\delta\chi, \end{aligned} \tag{2.8}$$

we readily find

$$\begin{aligned} \bar{Z}-\langle\psi|Z|\psi\rangle=&\langle\delta\phi|Z-\langle Z\rangle|\delta\phi\rangle\\ &+\langle\delta\phi|H-E|\delta\chi\rangle+\langle\delta\chi|H-E|\delta\phi\rangle, \end{aligned} \tag{2.9}$$

second order in the pair $\delta\phi$, $\delta\chi$.

The procedure then is to find functions ϕ and χ which make (2.6) stationary. If (2.6) were a minimum principle, this procedure would be straightforward. Unfortunately, it is a saddle point, and so although it is certainly effective in analytic calculations, its form is not at all appropriate to numerical work. There are a number of powerful techniques using large-scale digital computers for minimizing functions of many variables, but the determination of a saddle point is beset by obvious difficulties. We shall return to the saddle-point problem in Sec. VII, explicitly exhibiting the saddle-point characteristic and giving methods for reducing it to stationary forms more suitable for computer solution. First, let us examine the general problem of the construction of variational principles.

III. CORRECTION TO THE WAVE FUNCTION

By a variational principle for a quantity Q, we generally mean an expression $Q[x]$, in some independent quantity x, which takes on the value Q at a special value x_0 and converts a first-order error in x_0 to a second-order

[8] E. M. Roberts, Phys. Rev. **128**, 1381 (1962).
[9] S. I. Vetchinkin, Dokl. Akad. Nauk SSSR **147**, 1328 (1962) [English transl.: Soviet Phys.—Doklady **7**, 1132 (1963)].
[10] K. M. Sandro and J. O. Hirschfelder, Proc. Natl. Acad. Sci. U. S. **52**, 434 (1964).
[11] L. M. Delves, Nucl. Phys. **41**, 497 (1963).
[12] N. W. Bazley and D. W. Fox, J. Math. Phys. **7**, 415 (1966).

error in Q:

$$Q[x_0+\delta x]=Q+O((\delta x)^2). \tag{3.1}$$

An important consequence of this definition is that it can be pyramided to an F depending on Q, since clearly

$$F[Q[x_0+\delta x]]=F[Q]+O(\delta x)^2), \tag{3.2}$$

as well. This means that the obvious way of obtaining a variational principle for $\langle\psi|Z|\psi\rangle$, where $H\psi=E\psi$, is to construct a principle for $\psi[\phi]$ which corrects any input ϕ to a solution to $H\psi=E\psi$, and then substitute into $\langle\psi|Z|\psi\rangle$. We shall proceed in this direction.

Generalizing again, there are innumerable iterative techniques for transforming a first-order error in the solution to

$$G[\psi]=0 \tag{3.3}$$

into a second-order error. Two approaches may be distinguished. First, there is the Newton-Raphson[13] iteration, more familiar in its application to algebraic equations. To solve $G[\psi]=0$ when our guess is ϕ, we simply set $\psi=\phi+\delta\phi$ and expand:

$$G[\psi]=G[\phi]+\delta\phi\cdot\frac{\partial G[\phi]}{\partial\phi}+\tfrac{1}{2}\delta\phi\delta\phi:\frac{\partial^2 G[\phi]}{\partial\phi\partial\phi}+\cdots, \tag{3.4}$$

where the notation has been tailored to G and ϕ being vectors; for functions, each derivative is a variational derivative. Though second order, this yields at once for the solution to (3.3) $\delta\phi=-[\partial G[\phi]/\partial\phi]^{-1}G[\phi]$, or

$$\psi=\phi-[\partial G[\phi]/\partial\phi]^{-1}G[\phi], \tag{3.5}$$

and in fact the error can be made $O((\delta\phi)^s)$ for any s by including further terms from (3.4) before solving for $\delta\phi$.

Returning to the problem at hand, the Schrödinger equation can be written in nonlinear form without specifying the energy:

$$0=G[\psi]=H\psi-\frac{\langle\psi|H|\psi\rangle}{\langle\psi|\psi\rangle}\psi, \tag{3.6}$$

and we find

$$\frac{\delta G[\phi]}{\delta\phi}=H-\frac{\langle\phi|H|\phi\rangle}{\langle\phi|\phi\rangle}-\frac{|\phi\rangle\langle\phi|H}{\langle\phi|\phi\rangle}+\frac{\langle\phi|H|\phi\rangle}{\langle\phi|\phi\rangle^2}|\phi\rangle\langle\phi|. \tag{3.7}$$

Thus, if ϕ is normalized, $\langle\phi|H|\phi\rangle\equiv\xi$, and we define the projection onto ϕ

$$P=|\phi\rangle\langle\phi|, \tag{3.8}$$

(3.5) becomes $\psi=\phi-[H-PH-(1-P)\mathcal{E}]^{-1}(H-\mathcal{E})\phi$, or just

$$\psi=\phi-(1-P)[H/(H-\mathcal{E})]\phi. \tag{3.9}$$

A second approach is less rigid and maintains closer contact with common physical concepts. Our wave-function guess ϕ, instead of being regarded as approximately satisfying the exact Schrödinger equation, can trivially be taken as exactly satisfying the modified Schrödinger equation

$$H_0\phi=\mathcal{E}\phi, \quad \text{where} \quad H_0=H-W \tag{3.10}$$

for any perturbation W for which

$$W\phi=(H-\mathcal{E})\phi. \tag{3.11}$$

In this form, the determination of ψ to any order follows from any perturbation method which adds W to the "unperturbed" $H-W$. For example, Rayleigh-Schrödinger perturbation theory[14] then yields at once

$$\psi=\phi-[1/(H_0-\mathcal{E})](1-P)W\phi, \tag{3.12}$$

with a second-order error in $\delta\phi$. An obvious form for H_0 is that part of H which does not connect ϕ to its orthogonal complement, i.e., if

$$H=\begin{pmatrix}H_{00} & H_{01}\\ H_{10} & H_{11}\end{pmatrix},$$

decomposed according to ϕ plus the complementary space to ϕ, then

$$H_0=\begin{pmatrix}H_{00} & 0\\ 0 & H_{11}\end{pmatrix}.$$

Hence

$$\begin{aligned} H_0&=PHP+(1-P)H(1-P),\\ W&=PH(1-P)+(1-P)HP,\\ \mathcal{E}&=\langle\phi|H|\phi\rangle, \end{aligned} \tag{3.13}$$

and (3.12) reduces a bit to

$$\psi=\phi-(H_0-\mathcal{E})^{-1}W\phi. \tag{3.14}$$

We will employ the perturbational form (3.14), although many others exist, among them an expression based upon the resolvent-operator technique, which gives rise precisely to (3.9).

IV. EVALUATION OF THE CORRECTION

Under a number of circumstances, (3.14) can be evaluated in relatively explicit form. For example, if ϕ is an independent-particle state of many-fermion system,

$$\phi=\sum_p(-1)^p\prod_{i=1}^{N}u_{p\cdot i}(i), \tag{4.1}$$

where p refers to an index permutation, then a solution can be found in terms of a system of separable two-body equations. In fact, for determining expectations of one-

[13] H. Margenau and G. M. Murphy, *Mathematics of Physics and Chemistry* (D. Van Nostrand Co., Inc., New York, 1956), p. 492.

[14] Cf. E. Beckenbach, *Modern Mathematics* (McGraw-Hill Book Co., New York, 1956), p. 396.

body sums, a system of one-body equations suffices. For further details, see Appendix A.

In general, however, for initial wave functions which are not mixed configurations of type (4.1) but are of more involved intrinsic structure, a solution to (3.14) of the above-mentioned form is unavailable. In such a case, one may carry the variational process one step further back and introduce a variational expression for the Green's function

$$G_0=(H_0-\mathcal{E})^{-1}. \tag{4.2}$$

This is most easily obtained by guessing a result G, noting then that $H_0-\mathcal{E}$ is close to G^{-1}, and hence writing $G_0=[G^{-1}+(H_0-\mathcal{E}-G^{-1})]^{-1}$, which becomes an expansion

$$G_0=G-G(H_0-\mathcal{E}-G^{-1})G +G(H_0-\mathcal{E}-G^{-1})G(H_0-\mathcal{E}-G^{-1})G+\cdots. \tag{4.3}$$

If $H_0-\mathcal{E}$ is non-negative (its lowest eigenvalue will usually be precisely zero), (4.3) in fact supplies an upper bound to the operator G_0 when truncated at any even order.

In particular, a second-order variational principle results from truncation at second order, which yields

$$\bar{G}=2G-G(H_0-\mathcal{E})G. \tag{4.4}$$

The minimum character of this principle is directly evident from the relation

$$(H_0-\mathcal{E})^{-1}-2G+G(H_0-\mathcal{E})G =[G(H_0-\mathcal{E})-1](H_0-\mathcal{E})^{-1}[G(H_0-\mathcal{E})-1]^{\dagger}. \tag{4.5}$$

More generally, the upper-bound assertion for even-order truncation is a consequence of the identity

$$(A+B)^{-1}=\sum_0^{2s-1}(-1)^j(A^{-1}B)^jA^{-1} +(A^{-1}B)^s(A+B)^{-1}(BA^{-1})^s. \tag{4.6}$$

V. SUBSIDIARY MINIMUM PRINCIPLE

Let us now proceed to variational expressions for expectations. We seek the correction term

$$\Delta Z=\langle\psi|Z|\psi\rangle-\langle\phi|Z|\phi\rangle. \tag{5.1}$$

Inserting (3.14) and retaining only first order, we have

$$\Delta Z=-2\,\mathrm{Re}\langle Z\phi|\langle H_0-\mathcal{E})^{-1}|W\phi\rangle, \tag{5.2}$$

and to avoid the possibility of confusion we can subtract the ϕ component of $Z\phi$ without changing the value of the expression. We thus have

$$\Delta Z=-2\,\mathrm{Re}\langle(Z-\langle Z\rangle)\phi|\,(H_0-\mathcal{E})^{-1}|\,(H-\langle H\rangle)\phi\rangle, \tag{5.3}$$

where unspecified expectations hereafter refer to the initial state ϕ.

Now we are interested, not in the full Green's-function matrix $(H_0-\mathcal{E})^{-1}$ itself, but rather in specific matrix elements. Hence (4.4) is to be applied in the form

$$\langle\chi_1|\bar{G}|\chi_2\rangle=2\langle\chi_1|G|\chi_2\rangle-\langle\chi_1|G(H_0-\mathcal{E})G|\chi_2\rangle. \tag{5.4}$$

This equation must be used with discretion, because of the fact that one is really only varying the vectors $\xi_1=G\chi_1$ and $\xi_2=G\chi_2$ in it. The most direct use of this equation consists of writing

$$\langle\chi_1|\bar{G}|\chi_2\rangle=\langle\xi_1|\chi_2\rangle+\langle\chi_1|\xi_2\rangle-\langle\xi_1|H_0-\mathcal{E}|\xi_2\rangle, \tag{5.5}$$

where

$$\langle\xi_1|\chi_2\rangle=\langle\chi_1|\xi_2\rangle,\quad \chi_1\equiv(Z-\langle Z\rangle)\phi,\quad \chi_2\equiv(H-\langle H\rangle)\phi,$$

and varying with respect to ξ_1 and ξ_2 [the subsidiary condition in (5.5) need not be specifically imposed; it is a consequence of the variation]. However, (5.5) has the unsatisfactory saddle-point character of (2.6); we shall return to it in Sec. VII.

As an effective alternative approach, we may determine, say, ξ_2 by a separate variational—even minimum—principle, and then substitute into the desired matrix element. In other words, we observe that $\langle\xi_1|H_0-\mathcal{E}|\xi_2\rangle-\langle\xi_1|\chi_2\rangle-\langle\chi_1|\xi_2\rangle$ is stationary with respect to variation of ξ_1 at $\xi_2=(H_0-\mathcal{E})^{-1}\chi_2$, and that $\langle\chi_1|\bar{G}|\chi_2\rangle$ of (5.5) has the value $\langle\chi_1|\xi_2\rangle$ at this point. But the stationary point ξ_2 can also be found by locating the minimum with respect to ξ_2 of

$$M=\langle\xi_2|H_0-\mathcal{E}|\xi_2\rangle-\langle\chi_2|\xi_2\rangle-\langle\xi_2|\chi_2\rangle, \tag{5.6}$$

after which we set

$$\langle\chi_1|\bar{G}|\chi_2\rangle=\langle\chi_1|\xi_2\rangle. \tag{5.6'}$$

Equations (5.6) and (5.6') constitute an apparently practical computational technique.

The subsidiary variational principle (5.6) of course has its dual,

$$\langle\chi_1|\bar{G}|\chi_2\rangle=\langle\xi_1|\chi_2\rangle, \tag{5.7}$$

at minimum of

$$M'=\langle\xi_1|H_0-\mathcal{E}|\xi_1\rangle-\langle\chi_1|\xi_1\rangle-\langle\xi_1|\chi_1\rangle,$$

corresponding to $\xi_1=(H_0-\mathcal{E})^{-1}\chi_1$, but the usefulness of (5.6) and (5.7) is surprisingly nonsymmetric. This is a consequence of the situation that often in practice the wave functions available for variation are in some prescribed linear class. Suppose that (5.3) is to be evaluated by use of (5.6). We are thus called upon to minimize

$$M=\langle\xi|H_0-\mathcal{E}|\xi\rangle-\langle\xi|H-\mathcal{E}|\phi\rangle-\langle\phi|H-\mathcal{E}|\xi\rangle, \tag{5.8}$$

or, using (3.13) to express H_0 in terms of H and the projection $P=|\phi\rangle\langle\phi|$, we find

$$M=\langle\phi-\xi|H-\mathcal{E}|\phi-\xi\rangle-\langle\phi|H-\mathcal{E}|\phi\rangle -\langle\xi|\phi\rangle\langle\phi|H(1-P)|\xi\rangle-\langle\xi|(1-P)H\phi\rangle\langle\phi|\xi\rangle. \tag{5.9}$$

Now the minimization of (5.9) can be carried out under the condition $\langle\phi|\xi\rangle=0$, but if $\phi-\xi$ varies over the same linear class as does ϕ, the absolute minimum of

$\langle\phi-\xi|H-\mathcal{E}|\phi-\xi\rangle$ is zero at $\xi=0$. Thus no improvement is obtained.

On the other hand, the dual principle (5.7) involves $\chi_1=(Z-\langle Z\rangle)\phi$, and so we minimize

$$M'=\langle\xi|H_0-\mathcal{E}|\xi\rangle-\langle\phi|Z-\langle Z\rangle|\xi\rangle -\langle\xi|Z-\langle Z\rangle|\phi\rangle. \quad (5.10)$$

The optimal form for $\xi=(H_0-\mathcal{E})^{-1}(Z-\langle Z\rangle)\phi$ then gives a nontrivial result. Coupling (5.3), (5.10), and (5.11) and inserting the explicit form (3.13) for H_0, we conclude that

$$\Delta Z=-2\,\mathrm{Re}\langle\phi|H-\mathcal{E}|\xi\rangle \quad (5.11a)$$

at the minimum of

$$M'=\langle\xi|H-\mathcal{E}|\xi\rangle-\langle\phi|Z-\langle Z\rangle|\xi\rangle -\langle\xi|Z-\langle Z\rangle|\phi\rangle+N, \quad (5.11b)$$

where either

$$N=-\langle\xi|H-\mathcal{E}|\phi\rangle\langle\phi|\xi\rangle-\langle\xi|\phi\rangle\langle\phi|H-\mathcal{E}|\xi\rangle, \quad (5.11c)$$

or $N=0$ and M' is evaluated subject to

$$\langle\xi|\phi\rangle=0, \quad (5.11c')$$

with (5.11c) or (5.11c′) used as convenient. Equations (5.11) are the major results of this section.

VI. ANALYSIS OF A TEST CASE

To get a feeling as to how effective the variational technique (5.11) will be, we shall set up a simple, extremely solvable model—the hydrogen atom. Thus, in suitable notation,

$$H=-\frac{\hbar^2}{2m}\nabla^2-\frac{\hbar^2}{2m}\frac{\alpha}{r}, \quad (6.1)$$

$$\psi_0=\mathrm{K}e^{-\alpha r/2}, \quad E_0=-(\hbar^2/8m)\alpha^2.$$

As our approximate wave function, we choose, adopting units such that $\alpha\approx1$ and eliminating the angular coordinates,

$$\phi=e^{-r/2}/\sqrt{2}. \quad (6.2)$$

Finally, for the observable Z, we select the square radius r^2, so that

$$\langle\psi|r^2|\psi\rangle=12/\alpha^2, \quad \langle r^2\rangle=12, \quad (6.3)$$

or

$$\Delta Z=24(1-\alpha)+\cdots.$$

The task of the variational principle is essentially to compute $\xi=(H_0-\mathcal{E})^{-1}(r^2-12)\phi$, and by the method of Appendix A, Eq. (A19), this indeed has the exact value

$$\xi=(2m/\hbar^2)(\tfrac{1}{3}r^3+2r^2-44)\phi. \quad (6.4)$$

Now for ξ we choose instead only a linear function of r multiplying ϕ, and orthogonality to ϕ hence requires

$$\xi=A(3-r)\phi. \quad (6.5)$$

A trivial calculation thus establishes that

$$M'=\langle\xi|H-\mathcal{E}|\xi\rangle-\langle\phi|r^2|\xi\rangle-\langle\xi|r^2|\phi\rangle =A^2(\hbar^2/2m)(\tfrac{5}{2}-\tfrac{3}{2}\alpha)+48A, \quad (6.6)$$

with a minimum at

$$A=\frac{-24(2m/\hbar^2)}{(\frac{5}{2}-\frac{3}{2}\alpha)}. \quad (6.7)$$

We conclude that

$$\langle\xi|H|\phi\rangle=\tfrac{1}{2}A(1-\alpha)(\hbar^2/2m) =-12(1-\alpha)/(\tfrac{5}{2}-\tfrac{3}{2}\alpha), \quad (6.8)$$

so that

$$\Delta Z=-2\,\mathrm{Re}\langle\xi|H|\phi\rangle =24(1-\alpha)/(\tfrac{5}{2}-\tfrac{3}{2}\alpha), \quad (6.9)$$

which is in fact the exact result through first order in $(1-\alpha)$. The very crude form (6.5), coupled with the variational principle, is in this case entirely effective.

The above comparison is only through first-order perturbation. As a further check, we may consider a full numerical comparison. Suppose that we choose the approximation Ce^{-ar} to the hydrogen atom, where $a=1.8A^{-1}$. The fractional error in mean-square radius is found to be

$$\Delta r^2/\langle r^2\rangle=-0.086. \quad (6.10)$$

Now with the ansatz (6.5), here reading

$$\xi=A(3-2ar)\phi, \quad (6.11)$$

we find

$$\Delta r^2/\langle r^2\rangle\approx-0.0942, \quad (6.12)$$

which is indeed quite close. Since the effectiveness of the simple (6.11) may appear to be accidental, we can generalize (6.11) to

$$\xi=A[3-2(1+b)ar+ba^2r^2]\phi, \quad (6.13)$$

and we now find

$$\Delta r^2/\langle r^2\rangle\approx-0.092, \quad (6.14)$$

which is even closer. The optimal parameter is now $b=-1.825$, so that (6.11) and (6.13) are quite different but nonetheless nearly equivalent insofar as Δr^2 is concerned.

VII. SADDLE-POINT REDUCTION

A saddle-point variational method is amenable to standard numerical procedures if it can be reduced to a pair of minimization problems. In principle, this can always be done. In practice, let us examine the principle (5.5), dropping the unnecessary subsidiary restriction:

$\Delta Z=$ stationary value of

$$2\,\mathrm{Re}[\langle\xi_1|H_0-\mathcal{E}|\xi_2\rangle-\langle\xi_1|\chi_2\rangle-\langle\chi_1|\xi_2\rangle], \quad (7.1)$$

where

$$\chi_1=(Z-\langle Z\rangle)\phi, \quad \chi_2=(H-\langle H\rangle)\phi.$$

To convert (7.1) to the desired form, we set

$$\xi_1=(1/c)(A+B), \quad \xi_2=c(A-B), \quad A=\tfrac{1}{2}[c\xi_1+(1/c)\xi_2], \quad B=\tfrac{1}{2}[c\xi_1-(1/c)\xi_2], \tag{7.2}$$

for some constant c, so that we must consider (real part understood)

$$[\langle A|H_0-\mathcal{E}|A\rangle-(1/c)\langle A|\chi_2\rangle-c\langle\chi_1|A\rangle] -[\langle B|H_0-\mathcal{E}|B\rangle+(1/c)\langle B|\chi_2\rangle-c\langle\chi_1|B\rangle]. \tag{7.3}$$

Now (7.3) will clearly be stationary where each member is minimum. Hence we ask instead for the minimum of

$$[\langle A|H_0-\mathcal{E}|A\rangle-(1/c)\langle A|\chi_2\rangle-c\langle\chi_1|A\rangle] +[\langle B|H_0-\mathcal{E}|B\rangle+(1/c)\langle B|\chi_2\rangle-c\langle\chi_1|B\rangle]. \tag{7.4}$$

Substituting back from (7.2), we now have the subsidiary minimum principle

$$\Delta Z=2\,\mathrm{Re}[\langle\xi_1|H_0-\mathcal{E}|\xi_2\rangle-\langle\xi_1|\chi_2\rangle-\langle\chi_1|\xi_2\rangle], \tag{7.5}$$

when

$$\mathfrak{Z}\equiv c^2[\langle\xi_1|H_0-\mathcal{E}|\xi_1\rangle-2\,\mathrm{Re}\langle\chi_1|\xi_1\rangle] +(1/c^2)[\langle\xi_2|H_0-\mathcal{E}|\xi_2\rangle-2\,\mathrm{Re}\langle\xi_2|\chi_2\rangle]$$

is minimum. Again, if we restrict ξ by $\langle\xi|\phi\rangle=0$, then by virtue of (3.13) we can replace H_0 by H. Now we have two separate minimizations, each one a replica of one of our previous subsidiary minimizations (5.6) and (5.7). But (7.5) and (5.7) are not independent, for if we restrict ξ_2 by $\langle\xi_2|\phi\rangle=0$ and ξ_2 is expanded in the same linear set as ϕ, then $\xi_2=0$ necessarily results, in fact converting (7.5) to the demonstrably useful (5.7) and (5.11). On the other hand, if ξ_2 covers a larger set, e.g., if nonlinear parameters are involved, then (7.5) may be more likely to strike a balance between optimization of ξ_1 and ξ_2. We shall consider two examples of such a situation.

VIII. FURTHER ANALYSIS OF A TEST CASE

If the conditions $\langle\xi_1|\phi\rangle=\langle\xi_2|\phi\rangle=0$ are imposed, H_0 may be replaced by H in (7.5). Let us return to the hydrogen atom of Sec. VI and observe that, by restricting ϕ to (6.2) but allowing (6.5) for ξ_1, we are indeed allowing a larger class for the latter. In the present case, we need ξ_2 as well, and accordingly set [following the notation of (6.1)]

$$\xi_2=B(3-r)\phi. \tag{8.1}$$

Consequently,

$$\langle\xi_2|H-\mathcal{E}|\xi_2\rangle-\langle\xi_2|H-\mathcal{E}|\phi\rangle-\langle\phi|H-\mathcal{E}|\xi_2\rangle =B^2(\hbar^2/2m)(\tfrac{5}{2}-\tfrac{3}{2}\alpha)-B(\hbar^2/2m)(1-\alpha), \tag{8.2}$$

yielding

$$B=(1-\alpha)/(5-3\alpha). \tag{8.3}$$

Hence

$$\Delta Z=2(\langle\xi_1|H-\mathcal{E}|\xi_2\rangle-\langle\xi_1|H|\phi\rangle-\langle\phi|r^2|\xi_2\rangle) =AB(\hbar^2/2m)(5-3\alpha)-A(\hbar^2/2m)(1-\alpha)+48B =24(1-\alpha)/(\tfrac{5}{2}-\tfrac{3}{2}\alpha), \tag{8.4}$$

precisely the approximate result of (6.9), correct through first order as $(1-\alpha)$, which is all that can be expected.

Again, it is informative to test (7.5) for the full numerical error. We choose once more the hydrogen atom, with approximation and error given by (6.10). If the form (6.13) is assumed for each of ξ_1 and ξ_2, but with arbitrary linear terms as well, and if the subsidiary condition of (5.5) is imposed via a Lagrange parameter, we find

$$\Delta r^2/\langle r^2\rangle\approx-0.053, \tag{8.5}$$

an undercorrection rather than the overcorection of (6.14). If the sibsidiary condition is not imposed, i.e., (7.5) is used unchanged, then the computation is simpler, and

$$\Delta r^2/\langle r^2\rangle\sim-0.050 \tag{8.6}$$

is only slightly worse.

IX. TEST CASE WITH NONLINEAR PARAMETERS

When the variational functions available contain nonlinear parameters, the technique (7.5) has more to recommend it. Indeed, the first-order wave-function correction ξ_2 is a byproduct of the computation. The method has been applied to a model triton, a nucleus of three particles bound by the potential

$$V=-V_0\sum e^{-\mu r_{ij}^2}.$$

For this system, a polarized Gaussian wave function

$$\phi=\mathrm{Sym}\sum_{i=1}^{2} c_i\exp[-\tfrac{1}{2}(\gamma/f_i) \times(r_{12}^2+a_i(\mathbf{r}_3-\tfrac{1}{2}\mathbf{r}_1-\tfrac{1}{2}\mathbf{r}_2)^2)] \tag{9.1}$$

has been found[5,15] to be a very good approximation. For Laskar's potential,[16] $V_0=37.40$ MeV, $\mu=0.2699$ F^{-2}, the results obtained were $\langle\phi|H|\phi\rangle=-8.50$ MeV, $\langle\phi|r^2|\phi\rangle^{1/2}=1.65$ F, where r is the radius from the center of mass; the optimizing polarization parameters were $a_1=0.337$, $a_2=2.096$. Since the asymptotic form of (9.1) cannot be correct, it is important to determine the correction to $\langle r^2\rangle$.

To apply (7.5) in this case, the correction functions ξ_i were chosen to be of the same form as ϕ:

$$\xi_i=\mathrm{Sym}\,d_i\sum_{j=1}^{2} c_j\exp[-\tfrac{1}{2}(\gamma/f_{ji}) \times(r_{12}^2+a_{ji}(\mathbf{r}_3-\tfrac{1}{2}\mathbf{r}_1-\tfrac{1}{2}\mathbf{r}_2)^2)]. \tag{9.2}$$

On numerical minimization, the stationary values of ξ_i were in fact found to represent highly polarized com-

[15] R. C. Herndon and Y. C. Yang, in *Methods of Computational Physics*, edited by B. Alder, S. Fernbach, and M. Rotenberg (Academic Press Inc., New York, 1966), Vol. VI, pp. 153–234.
[16] W. Laskar, Ann. Phys. (N. Y.) 17, 436 (1962).

pressed states, because of the large values of the a_{ij}:

$$a_{11}=7.245,\ a_{12}=9.087,\ a_{21}=3.086,\ a_{22}=2.255. \quad (9.3)$$

Using the corresponding ξ_i, the mean-square error correction became the very small $\Delta r^2/\langle r^2\rangle=0.0157$, which is a further indication of the excellence of the polarized Gaussian wave function ϕ.

X. CONCLUSIONS

A class of variational principles for expectation values of arbitrary physical operators has been derived which permits the stationary point to be found with a computer. Each involves the subsidiary minimization of one or more expressions. The simplest principles require only a single minimization, but their utility depends very much on the precise nature of the class of variational functions. One principle which does not appear restricted in this fashion has been applied to the simple case of the mean-square radius of the hydrogen atom with excellent results. A second principle was determined by a general method of converting stationary saddle to pairs of minima. This is especially suitable for nonlinear parameters in variational functions. We have tested it on the above hydrogen-atom problem and applied it to a model triton. In spite of the fact that a variational principle gives only a probable improvement, not a rigorous improvement, results suggest the use of the minimum-type principles to find improved values of expectations of physical quantities.

ACKNOWLEDGMENTS

We would like to thank Dr. M. Sobol and Dr. J. Leavitt for their minimization routines which were used in this work. We would also like to thank the staff at the A.E.C. Computing and Applied Mathematics Center, Courant Institute of Mathematical Sciences, New York University, for their help during the preparation of this paper.

APPENDIX A: ANALYTIC EVALUATION OF THE CORRECTIONS

For the purpose of developing corrections to expectations, it is sufficient to find the first-order wave-function correction

$$\begin{aligned}\xi_2&=\phi-\psi\\&=(H_0-\mathcal{E})^{-1}W\phi,\end{aligned} \quad (A1)$$

where H_0 is the "unperturbed" Hamiltonian of which ϕ is an eigenfunction, $H_0\phi=\mathcal{E}\phi$, and W the perturbation

$$W=H-H_0-\langle H-H_0\rangle. \quad (A2)$$

Alternatively, for first-order corrections, one can instead determine

$$\xi_1=(H_0-\mathcal{E})^{-1}(Z-\langle Z\rangle)\phi, \quad (A3)$$

which may be simpler. In any event, if ϕ is an independent-particle state of a many-fermion system

$$\phi=\sum_p(-1)^p\prod_{i=1}^{N}u_{p\cdot i}(i), \quad (A4)$$

where p refers to an index permutation, then we may choose the one-body sum

$$H_0=\sum_1^N h(i), \quad (A5)$$

where

$$h(i)u_j(i)=\epsilon_j u_j(i),$$

and (A1) and (A3) can be evaluated in relatively explicit form.

Since $(Z-\langle Z\rangle)\phi$ has no ϕ component, there can be no trouble with resonance denominators in (A3) or (A1). In fact, this statement holds for each particle: For any s-body operator $U(1,\cdots,s)$ present in $Z-\langle Z\rangle$, we can assume that particles 1, $\cdots$, s have no components in the Fermi sea S of u_1, $\cdots$, u_N. Let us show this explicitly. Consider first a one-body sum

$$U=\sum U(i), \quad (A6)$$

and define the projections on and off the Fermi sea

$$\begin{aligned}\mathcal{P}(i)&=\sum_{h\in S}|u_{h(i)}\rangle\langle u_h(i)|,\\ \mathcal{Q}(i)&=1-\mathcal{P}(i).\end{aligned} \quad (A7)$$

Then

$$\sum U(i)\phi=\sum\mathcal{Q}(i)U(i)\phi+\sum_{l,m\in S}u_l(i)\langle u_l|U|u_m\rangle\phi_{im},$$

where

$$\phi_{im}=\partial\phi/\partial u_m(i) \quad (A8)$$

s the mi cofactor of ϕ. But

$$\textstyle\sum_i u_l(i)\phi_{im}=\delta_{lm}\phi \quad (A9)$$

and $\sum\langle u_l|U|u_l\rangle=\langle\sum U(i)\rangle$, and so we conclude that

$$\sum U(i)\phi=\sum\mathcal{Q}(i)U(i)\phi+\langle\sum U(i)\rangle\phi. \quad (A10)$$

In particular, if $U=Z-\langle Z\rangle$, so that $\langle U\rangle=0$, then only the desired $\mathcal{Q}(i)U(i)$ is present.

For a two-body sum

$$V=\sum_{i\neq j}'V(i,j), \quad (A11)$$

one proceeds similarly. Now

$$\begin{aligned}\sum'V(i,j)\phi=&\sum'\mathcal{Q}(i)\mathcal{Q}(j)V(i,j)\phi\\&+\sum'\mathcal{Q}(i)\langle u_l(y)|V(i,y)|u_m(y)\rangle u_l(j)\phi_{jm}\\&+\sum'\mathcal{Q}(j)\langle u_l(x)|V(x,j)|u_m(x)\rangle u_l(i)\phi\\&+\sum'\langle u_k(x)u_l(y)|V(x,y)|u_m(x)u_m(y)\rangle\\&\qquad\times u_l(j)u_k(i)\phi_{jm,im},\end{aligned}$$

where $\phi_{jm,im}$ is the $(N-2)\times(N-2)$ cofactor. Hence,

using (A9) and its $(N-2)\times(N-2)$ analog, we obtain

$$\sum{}' V(i,j)\phi=\sum{}' \mathcal{Q}(i)\mathcal{Q}(j)V(i,j)\phi +2\sum \mathcal{Q}(i)\langle u_l(y)|V(i,y)|u_l(y)\rangle\phi -2\sum \mathcal{Q}(i)\langle u_l(y)|V(i,y)|u_m(y)\rangle u_l(i)\phi_{im} +2\langle\sum{}' V(i,j)\rangle\phi. \quad \text{(A12)}$$

Again, if $\langle V\rangle=0$, the Fermi-sea components of each operator are missing. One can readily extend (A10) and (A12) to s-body operators.

Now, if we fix a permutation p and a term $A(1,\cdots,s)$ in Z, our problem is to solve

$$(H_0-\mathcal{E})\xi=Z(1,\cdots,s)\prod_1^N u_{p\cdot i}(i), \quad \text{(A13)}$$

where $Z(1,\cdots,s)\prod_1^s u_{p\cdot i}(i)$ has no components in the Fermi sea. Since $\mathcal{E}=\sum \epsilon_{p\cdot i}$, we have at once[17]

$$\xi=\xi_s\prod_{s+1}^N u_{p\cdot i}(i), \quad \text{(A14)}$$

where

$$\sum_1^s [h(i)-\epsilon_{p\cdot i}]\xi_s=Z(1,\cdots,s)\prod_1^s u_{p\cdot i}(i).$$

The problem thus reduces to that of the s-body Green's function on the space above the Fermi sea for the separable operator $\sum_1^s [h(i)-\epsilon_{p\cdot i}]$. In general, one can solve numerically or by a complete expansion in orbitals:

$$\xi_s=\sum\langle u_{k_1}(1)\cdots u_{k_s}(s)|Z(1,\cdots,s)|u_{p\cdot 1}(1)\cdots u_{p\cdot s}(s)\rangle \times\prod_1^s u_{k_i}(i)/(\epsilon_{k_i}-\epsilon_{p\cdot i}). \quad \text{(A15)}$$

In special cases, however, closed-form results are available. Thus, if the one-body Hamiltonian h has the standard form

$$h=p^2/2m+v(\mathbf{r}), \quad \text{(A16)}$$

then the one-body singular Green's function corresponding to an s state can be written down at once, i.e., if

$$\left(-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial r^2}+v(r)\right)\phi(r)=\epsilon\phi(r), \quad \text{(A17)}$$

then it is easy to see that

$$-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial r^2}+v(r)-\epsilon=-\frac{\hbar^2}{2m}\frac{1}{\phi(r)}\frac{\partial}{\partial r}\phi(r)^2\frac{\partial}{\partial r}\frac{1}{\phi(r)}, \quad \text{(A18)}$$

so that, if $\psi(r)$ is orthogonal to $\phi(r)$,

$$\left(-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial r^2}+v(r)-\epsilon\right)^{-1}\psi(r)=\left(-\frac{\hbar^2}{2m}\right)^{-1}\phi(r) \int_0^r\frac{1}{\phi^2(r'')}\int_0^{r''}\phi(r')\,\psi\,(r')dr'dr''. \quad \text{(A19)}$$

[(A19) may have an irrelevant ϕ component.]

Another way of expressing (A19) is to observe[18] that if (A16) holds and we define

$$\xi_s(1,\cdots,s)=F_s(1,\cdots,s)\prod_1^s u_{p\cdot i}(i), \quad \text{(A20)}$$

then on commuting $\sum[h(i)-\epsilon_{p\cdot i}]$ through, and using $[h(i)-\epsilon_{pi}]\,u_{p\cdot i}\,(i)=0$, (A14) becomes simply

$$-\frac{\hbar^2}{2m}\sum_1^s(\nabla_i^2F+\nabla_i\ln u_{p\cdot i}\cdot\nabla_iF) =[\prod_1^s u_{p\cdot i}(i)]^{-1}Z(1,\cdots,s)\prod_1^s u_{p\cdot i}(i), \quad \text{(A21)}$$

which, e.g., for $s=1$ and radial $u_{p\cdot 1}$ is immediately solvable.

[17] O. Sinanoglu, Proc. Roy. Soc. (London) **A260**, 379 (1961).

[18] S. Borowitz and M. O. Vassell, J. Quant. Spectr. Radiative Transfer 4, 663 (1964).