Revista Mexicana de Física 19 (1970) 329-338

ON THE MINIMIZATION OF THE VARIANCE

O. Novaro

Instituto Mexicano del Petróleo

and

A. Calles

Instituto de Física, Universidad Nacional de México

(Recibido: julio 10, 1970)

ABSTRACT: A discussion is given here of the implications and relative advantages of the use of the minimization of the energy variance, as an alternative to the well known Ritz principle. To supplement the arguments, some variational calculations of a couple of simple and completely soluble problems are done, using the different minimization criteria. It is shown that when the trial functions are chosen on a physical basis, convergence is quite as good in the variance as in the Ritz procedure. However the variance technique is impractical if the sole criterium for the election of the trial function is a mathematical one as, for instance, analytical convenience only. The convergence of the Temple lower bound (in Conroy's modified version) is also studied.

1. INTRODUCTION

The minimal variance method stands out among the various minimization criteria regularly employed in variational calculations of atomic and molecular systems, in the sense that it is designed to optimize the wave function itself and not just a particular physical property of the system. For instance, the optimization of the ground state energy as is done in the well-known Ritz procedure does not guarantee that the adjusted function will be reliable when used for the calculation of the other properties of the system.

The optimization of the function is obtained by adjusting the variational parameters, so that it becomes an approximate eigen-function of the physical hamiltonian with the exact energy eigenvalue. We do this by minimizing the expectation value of the operator

$$(H - E_{ex})^2 \tag{1}$$

with respect to our trial functions. We then define the variance as the expectation value

$$U^2 = \langle \Psi | (H - E_{ex})^2 | \Psi \rangle \tag{2}$$

which, by construction is positive definite. It should be emphasized here that we are assuming that in (2), H is the exact hamiltonian and also that E_{ex} is the exact energy eigenvalue. As in general this energy is not known beforehand, we can use either the experimental value or any extrapolation of a calculated one.

To avoid this we could use, as is often done[2], the Ritz energy instead of E_{ex} in eq. (2). However, this alternative introduces certain disadvantages, as we shall see. We will now analize two completely soluble problems, where E_{ex} is known; we shall then be able to draw some conclusions about the accuracy and flexibility of (2).

The present work is intended as a test of the variance method, as used by Conroy in his study of small molecules[1]. He uses carefully built trial functions, which satisfy all the characteristics of the problem and that include many-body correlations; he then uses them in eq. (2), proceeding by numerical integration, because of the complexity of the functions. The use of the Monte Carlo method to perform the integration, implies iterations that would normally tend to magnify the error introduced by the use of the Ritz, or any other approximate form of the energy in eq. (2). As Conroy's results give the best ab initio potential energy surfaces for H_3^+ and H_3 yet available, a test of his variance method seems of interest.

Let us first establish how the variance method works. The quantity U^2 is positive definite, so its smallness is a measure of the adequacy of the trial function Ψ. In fact if $U^2 = 0$, then Ψ is an eigenfunction of H with eigenvalue E_{ex} and viceversa. The equivalent situation in the Ritz principle is when $<\Psi|H|\Psi> = E_{ex}$. However if $U^2 > 0$ and $\mathcal{E}_R > E_{ex}$, the Ritz energy $\mathcal{E}_R$ gives only a ***relative*** measure of the deviation as only the ratio

$$\frac{\mathcal{E}_R - E_{ex}}{E_{ex}}$$

has an absolute meaning; using the variance, on the other hand, the value of U^2 is an ***absolute*** measure of the deviation from the exact solution. In fact the variance minimization consists in the approximate annihilation of the function $|\varphi> = (H - E_{ex})|\Psi>$ and the adequacy of the chosen trial function $|\Psi>$ is given by the smallness of the norm

$$<\varphi|\varphi> \approx 0. \tag{3}$$

It has been noted[3] that if the overlap between a trial function and the eigenfunction Ψ_{ex} is optimized, one can in some instances get nearly a 100% overlap but still the function Ψ cannot be used to give the atomic or molecular properties with high accuracy. Or, in other words, the maximum overlap criterium:

$$<\Psi_{ex}|\Psi_{trial}> \approx 1 \tag{4}$$

does not give an absolute measure. This is so, because the two functions can differ in a reduced, yet physically significant region of space. This is avoided in (3) since the operator $(H - E_{ex})^2$ contains all the physical information of the system and in a sense "weighes" the importance of the different regions of space.

The upper and lower bounds we are going to obtain from the variance are the following: for the upper bound, the expectation value of H with respect to functions that have been adjusted before to optimize U^2. This is not only an upper bound to the exact energy, but also by construction to the Ritz value. As a lower bound, we will use the well known Temple criterium[4] using a modified form introduced by Conroy[1]:

$$E_{ex}^{(0)} \geqslant \lambda = \mathcal{E}_R - \frac{U^2}{E_{ex}^{(1)} - \mathcal{E}_R} \quad , \tag{5}$$

where $\mathcal{E}_R$ is the Ritz value, U^2 is defined in eq. (2) and $E_{ex}^{(0)}$ and $E_{ex}^{(1)}$ are the exact eigenvalues of the ground and the first excited state of the same symmetry. Again, the criticism that in eq. (5) we are using two exact values ($E_{ex}^{(0)}$ is implicit in U^2) to obtain an approximate energy, is not relevant, because in Conroy's work eq. (5) is only used to extrapolate a value for the iteration process and, in our problems, we **do** know $E_{ex}^{(0)}$ and $E_{ex}^{(1)}$ exactly; we are only testing the accuracy and convergence properties.

2. VARIANCE MINIMIZATION IN PSEUDO-LITHIUM

A pseudo-atom is a mathematical model that consists of a system of particles of spin 1/2 that interact among themselves through repulsive harmonic oscillator forces of unit frequency and that move in a common oscillator potential of frequency $\sqrt{Z}$. This is a soluble many body problem introduced by Moshinsky and Calles[8] with the idea of using it as a trial ground for approximation techniques developed in atomic physics.

The $2e^-$ pseudo atom has been solved by numerous approximation techniques in a recent paper[9]. The Ritz variance minimizations were carried out using the harmonic oscillator (h.o.) basis both in the standard way and taking the Hartree-Fock solution as a starting point. We use here these results to calculate the Temple-Conroy lower bound defined in eq. (5).

The results for the pseudo-lithium ion ($Z = 3$) are shown in figures 1, 2 and 3, where we have used different energy units for the standard and HF calculations, for practical reasons, as discussed in ref. (9). As the pseudo atom is a fictitious system, what really matters are the convergence properties. The curve marked "variance upper bound" corresponds to the energies derived from the variance minimization. We see that they differ very little from the Ritz values. Also the convergence of the variance is quite satis-

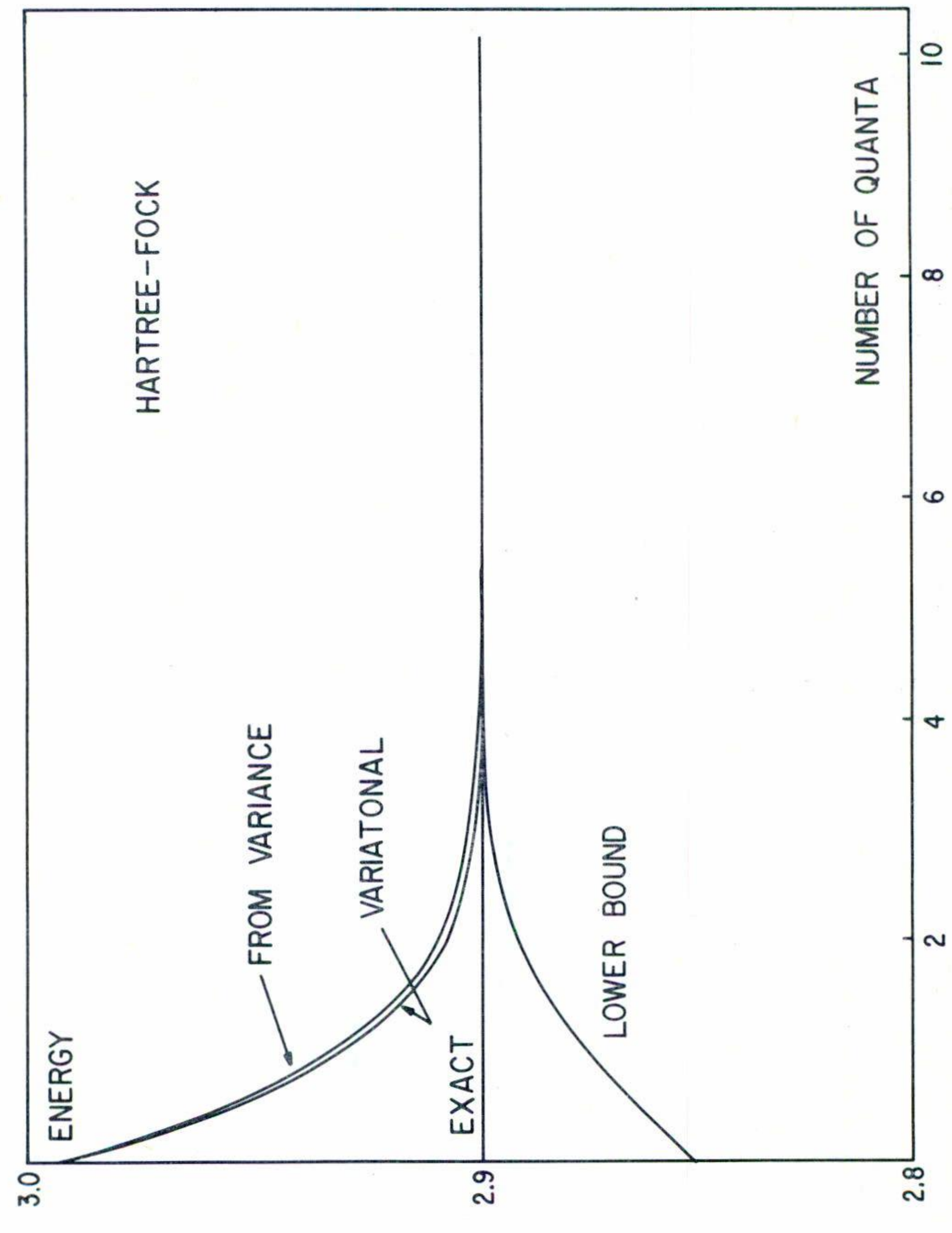

Fig. 1 Upper and lower bounds as a function of the number of quanta, for the standard case.

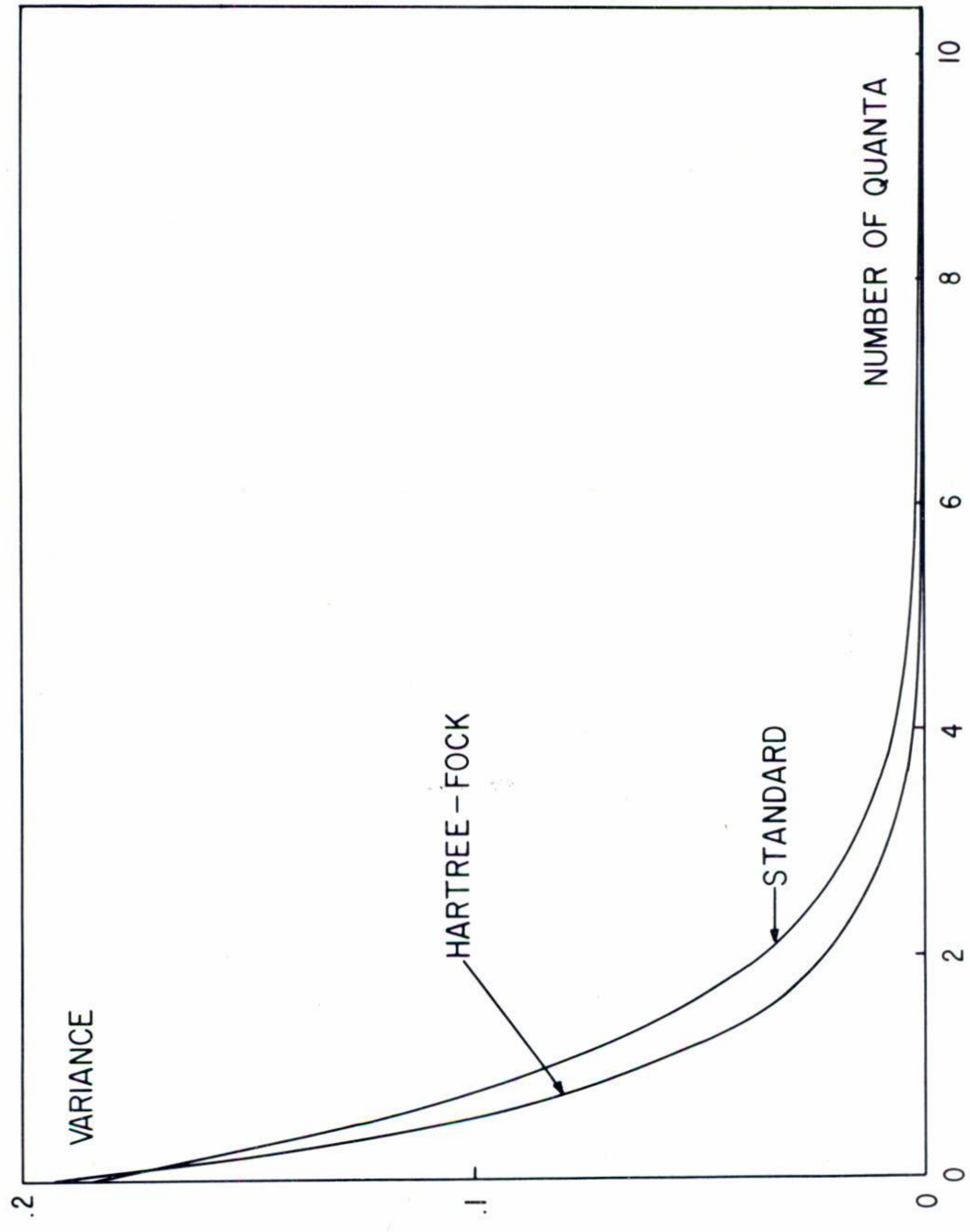

Fig. 2 Upper and lower bounds as a function of the number of quanta, using Hartree-Fock functions.

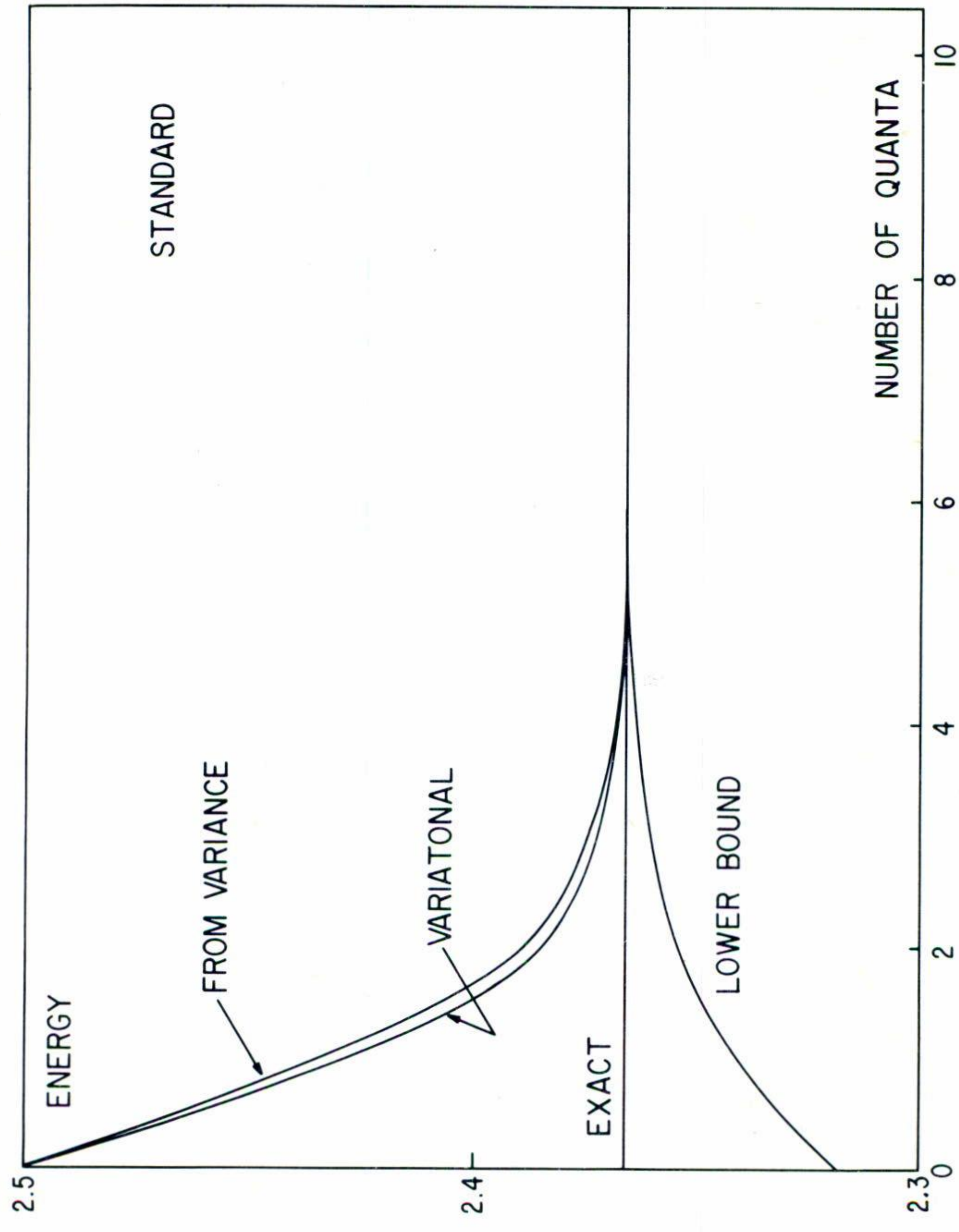

Fig. 3 The variance value as function of number of quanta, for the standard and Hartree-Fock procedures.

factory; the lower bounds, on the other hand, start off very well but are slower in their convergence.

3. VARIANCE MINIMIZATION IN HYDROGEN

Though the H atom is of course a soluble problem, it is still commonly used as trial ground for approximation methods. In a previous paper by Moshinsky and Novaro[6] a Ritz variational calculation for the ground state energy of several molecules and atoms including hydrogen were carried out using h.o. states, with the intention both of testing the practical possibilities of the h.o. basis in atomic and molecular physics, and also of showing the analytical simplicity provided by the group-theorical techniques that have developed in nuclear physics. We shall here develop a parallel calculation, minimizing the variance instead of the hamiltonian operator, and compare our result with those of reference (6). As we are using a real orthogonal basis we have to minimize the following expression:

$$U^2 = \int (H\Psi - E_{\text{ex}}\Psi)^2 \, dV =$$

$$= \sum_{n,m} A_n A_m \left[\int (H\varphi_n H\varphi_m - 2E_{\text{ex}} \varphi_n H\varphi_m) \, dV + E_{\text{ex}}^2 \delta_{nm} \right] \quad (6)$$

where Ψ is the trial function, given as a linear combination of h.o. states φ_m, with coefficients A_m. We shall follow the notation of (6) using the variational parameter

$$\epsilon = \sqrt{\hbar\omega} \sqrt{\frac{2\hbar^2}{me^4}}$$

associated to the frequency ω of the h.o. functions and use the dimensionless form for the hamiltonian:

$$H = \frac{1}{2} \epsilon^2 p^2 - \frac{\sqrt{2}\,\epsilon Z}{r} \quad . \qquad (7)$$

In this way the ground state energy has the value $E_{ex} = -1$.

The integrals involved in eq. (6) can be simply calculated, by an expansion in terms of Talmi integrals using tabulated coefficients[7]. We simply list here the results. For zero quanta in the h.o. expansion we get the value $U^2 = 0.42245$, which is certainly not a small one; if we use the value of the parameter $\epsilon = 0.416$ that minimizes U^2 to calculate the energy, we get $\mathcal{E}_{\mu^2} = -0.53400$ which is a very poor value, especially if it is compared to the value $\mathcal{E}_R = -0.849$ we get using the Ritz procedure[6]. To test the convergence we go to a higher number of quanta, remembering that with the Ritz procedure we get a value of -0.962 for 10 quanta. If we minimize U^2 for 10 quanta in the h.o. expansion we see that convergence is extremely slow, since we get $U^2 = 0.410881$; of course, the corresponding energy $\mathcal{E}_{\mu^2} = -0.53434$, does not fit the correct value.

The lower bound, as defined in eq. (5), was also calculated; although its convergence is of course as slow as that of U^2, the value $\lambda = -1.539$ is not much worse than the upper bound obtained from U^2.

Therefore, we have a situation in which the Ritz principle gives reasonable values for the energy[6] and in which we can optimize the overlap of the 0 quanta function to reach 96% with the exact function[3]; yet, as we have shown here, this function does not satisfy the H atom Schrödinger equation (1).

In conclusion the mere selection of a trial function, only because of analytical simplicity to calculate the integrals, gives a very bad value for the variance; this is because in the physically relevant regions of space (in this case inside Borh's radius) the function does not nearly resemble the exact function[3]. On the other hand the Ritz method, and also the maximum overlap criteria, mask this fact, because they concentrate on a single property of the system.

This indicates that the variance method is actually useful only when we choose the functions with great care and when we want them to give a complete information about the system (as is the actual case nowadays, in most atomic and molecular calculations). In nuclear physics where H itself, let alone E_{ex}, is in doubt and where we can profit on the obtention of a single property, the Ritz principle may be preferred.

In conclusion we see that when the h.o. trial functions are used in a problem for which they are well suited, the variance method is quite superior to the Ritz procedure as it gives excellent convergence, lower and upper bounds quite as satisfactory and, of course, one can rely more on the functions when using them to calculate other physical properties; one example is the calculation of excited states where the Ritz minimization only gives stationary values and U^2 gives absolute minima.

ACKNOWLEDGEMENTS

We thank Drs. M. Moshinsky and J. Flores for their interest in this work.

REFERENCES

1. H. Conroy, J. Chem. Phys. 41 (1964) 1336.
2. A.A. Frost, J. Chem. Phys. 10 (1942) 240.
3. J. Flores, M. Dubovoy, Rev. Mex. Fís. 17 (1968) 289.
4. G. Temple, Proc. Roy. Soc. (London) A119 (1928) 276.
5. S. Larsen, Preprint 173 Uppsala University (June 1966).
6. M. Moshinsky, O. Novaro, J. Chem. Phys. 48 (1968) 4162.
7. T. Brody, M. Moshinsky, G. Jacobs, Nucl. Phys. 17 (1960) 16, M. Moshinsky, T. Brody, Tables of Transformation Brackets, (Gordon & Breach Publ. Inc., New York, 1967).
8. M. Moshinsky, Am. J. Phys. 36 (1968) 52, A. Calles, M. Moshinsky, Am. J. Phys. 38 (1970) 456.
9. M. Moshinsky, O. Novaro, A. Calles. To appear in Journal de Physique.

RESUMEN

En este trabajo se discuten las implicaciones y ventajas relativas del uso de la minimización de la variancia como alternativa al bien conocido principio de Ritz. Para complementar los argumentos, se resuelven variacionalmente dos sistemas simples, con solución exacta conocida, usando los diferentes criterios de minimización.

Se muestra que cuando las funciones de prueba se escogen bajo un criterio físico, la convergencia es tan buena en el método de la variancia como en el de Ritz. Sin embargo, la técnica de la variancia no es práctica si el único criterio de elección de la función de ensayo es matemático como, por ejemplo, por mera conveniencia analítica. La convergencia de la cota inferior de Temple (en la versión modificada por Conroy) también se analiza.