

STUDIES WITH THE LOCAL ENERGY METHOD

Mihaly Mezei

Department of General and Inorganic Chemistry

Doctoral dissertation
1971

Supervisor: Dr. Ferenc Török

Eötvös Loránd University, Budapest, HUNGARY

INTRODUCTION

Quantum chemistry, that seeks to determine the quantum-mechanical state function of chemical systems, assumes that it is determined by the Schrodinger equation.

Exact solutions of Schrodinger's equation (SE) for chemical systems can only be obtained for a few simple cases (e.g., the hydrogen atom). For more complex systems, where the exact solution is not known, approximate solutions are the best we can do.

These approximate solutions either rely on some mathematical approximation (ab initio calculations) or use some experimental data in conjunction with a mathematical approximation (semiempirical methods).

The subject of this study, the Local Energy Method (LEM) of A.A. Frost, is essentially an ab initio method although some experimental data is used to facilitate the calculations.

The LEM was introduced by A.A. Frost in 1942⁷ and (after the wide availability of computers) several calculations were published from 1960 to 1964, some using a large basis set (e.g., the Li atom); application to larger systems, however, has remained so far unsuccessful. The reasons for this will be summarized in Sec. 3.

Modifications of the LEM are trying to make it applicable to larger systems. This is the aim of the modifications discussed in the present work.

Calculations were performed on the hydrogen molecular ion (H_2^+) and on the helium atom, using the Gier computer at MAVEMI. The computer program was written in Gier Algol 4.

1. *The foundations of the LEM.*

The LEM considers the solution of the SE as a function approximation problem, i.e., for each function $\varphi(\mathbf{x})$ in a given function space defines a scalar measure F characterizing the deviation of this function from the solution sought; and looks for the function $\varphi_0(\mathbf{x})$ for which F is minimal.

A natural consequence of the method is the fact that if this search is done in a function space T that contains the exact solution(s) of the SE then — with the proper choice of F — the exact solution(s) will be found.

Such a proper measure could be:

$$F1 = \frac{\int_V \rho(\mathbf{x})(\hat{H}\varphi(\mathbf{x}) - E\varphi(\mathbf{x}))^2 dv}{\int_V \rho(\mathbf{x})\varphi^2(\mathbf{x})dv} \quad (1.1)$$

where

$$\rho(\mathbf{x}) \geq 0 \quad (1.2)$$

and

$$E = \frac{\int_V \rho(\mathbf{x})\varphi(\mathbf{x})\hat{H}\varphi(\mathbf{x})dv}{\int_V \rho(\mathbf{x})\varphi^2(\mathbf{x})dv} \quad (1.3)$$

Here x is a point of the configuration space V , $\varphi(\mathbf{x})$ is a wave function, \hat{H} is the Hamiltonian, $\rho(\mathbf{x})$ is a weight function that does not destroy the convergence of the integrals in (1.1) and (1.3) used to influence the precision of the approximation at different regions of the configuration space. In the simplest case, $\rho(\mathbf{x})$ is constant in a bounded region of the configuration space and zero outside. In this case the integration can be limited to this region and $\rho(\mathbf{x})$ can be dropped from the formalism.

Here and in the followings bold face denotes a vector and the caret denotes an operator. It can be seen that the minimum of $F1$ is zero, and it is a necessary and sufficient condition for $\varphi(\mathbf{x})$ to be the solution of the SE that it be zero everywhere (apart from a set of zero Lebesgue measure).

In this notation, the SE is written as

$$\hat{H}\varphi(\mathbf{x}) = E\varphi(\mathbf{x}) \quad (1.4)$$

The integral in the numerator of (1.1) is more difficult than the integral encountered with variational type methods (we need not only $\int \varphi(\mathbf{x})\hat{H}\varphi(\mathbf{x})dv$ but $\int \hat{H}\varphi(\mathbf{x})\hat{H}\varphi(\mathbf{x})dv$, it is thus appropriate to apply a numerical quadrature. The LEM obtains the solution of the SE by minimizing integrals in such a way that the integrals are approximated with numerical quadratures.

Based on the above, the variants of the LEM can be categorized from several point of views:

1. the choice of the measure F .
2. The choice of the function space T .
3. The choice of the numerical quadrature used in the approximation.

The classical LEM of A.A. Frost minimized the functional $F1$ described by (1.1) to obtain the solution $\varphi(\mathbf{x})$ of the SE.

$\varphi(\mathbf{x})$ is a polynomial of the coordinates multiplied by an exponential. This choice has the advantage over the Slater determinant used in the variational type methods that it is easy to include the electron correlation, but its drawback is that for larger systems a large number of terms are needed to ensure the antisymmetry of the wavefunction.

2. *Description of the method.*

The three points of views used for categorization in Sec. 1 can also form the basis of the presentation of the method.

2.1. *The choice of the measure F .*

As we saw in Sec 1, the measure $F1$ defined by eq. (1.1) is capable to characterize the deviation a function $\varphi(\mathbf{x})$ from the solution the the SE. This measure can also be obtained as follows.

Let e_p be the local energy at the point \mathbf{x}_p of the configuration space, calculated with eq. (2.1) below^{7,10}:

$$e_p = \frac{\hat{H}\varphi(\mathbf{x}_p)}{\varphi(\mathbf{x}_p)} \quad (2.1)$$

If $\varphi(\mathbf{x})$ is the solution of the SE, eq. (1.4), then it is constant independent of the choice of \mathbf{x}_p and, conversely, if for a given $\varphi(\mathbf{x})$ it is constant everywhere, it follows that that $\varphi(\mathbf{x})$ is the solution of the SE.

If, on the other hand, $\varphi(\mathbf{x})$ is not a solution of the SE then the fluctuation of the e_p values will be a measure of its deviation from the solution of the SE. Accordingly, we can define a measure $F2$:

$$F2 = \frac{\int_V \rho'(\mathbf{x}) \left(\frac{\hat{H}\varphi(\mathbf{x})}{\varphi(\mathbf{x})-E} \right)^2 dv}{\int_V \rho'(\mathbf{x}) dv} \quad (2.2)$$

where

$$\rho'(\mathbf{x}) \geq 0, \quad \int_V \rho'(\mathbf{x}) dv < +\infty \quad (2.3)$$

Clearly, the minimum of $F2$ is 0 and it is reached only when $\varphi(\mathbf{x})$ is the solution of the SE (apart from a set of measure zero).

The application of $F2$ encounters serious difficulties with linear trial functions (see Sec. 2.2.1.1.). These difficulties can be eliminated, however, if $\rho'(\mathbf{x})$ is of the form

$$\rho'(\mathbf{x}) = \rho(\mathbf{x})\varphi^2(\mathbf{x}) \quad (2.4)$$

where $\rho(\mathbf{x}) \geq 0$. Substituting (2.4) into (2.2) we find that $F2$ is identical to $F1$. In our opinion, however, the $\rho'(\mathbf{x})$ defined by (2.4) is not defined unequivocally (since it depends on the solutions sought), so it could only be considered a ‘pseudo’ weight function.

This explains why Sec. 1 features the measure (1.1) and not (2.2), used in the literature. It should also be mentioned that measure of type $F2$ was used for estimating errors, but in that case the approximating function was obtained with a different method^{1,20}.

The following comments are in order concerning the measure (1.1):

S.F. Boys proved² that if the elements of the matrices \mathbf{S} , \mathbf{H} , \mathbf{G} are obtained with the same quadrature then the quadrature error in the energy will be only of second order. If we consider the quadrature error as getting the integral of $\rho(\mathbf{x})f(\mathbf{x})$ instead of the integral of $f(\mathbf{x})$ where $\rho(\mathbf{x})$ is constant if the integration is exact and ‘nearly constant’ if the integration is approximate then the acceptance of the measure $F1$ as a ‘good’ measure can be considered a heuristic proof of Boys’s theorem.

2.2. The choice of the function space.

The calculations in the literature used linear trial functions in most of the cases, i.e., $\varphi(\mathbf{x})$ was of the form^{3,7,10–13,19,22,23}.

$$\varphi(\mathbf{x}) = \sum_{i=1}^{nb} c_i \varphi_i(\mathbf{x}) \quad (2.5)$$

Here $\varphi_i(\mathbf{x})$ is a given function ($i = 1, \dots, nb$). The set of $\varphi_i(\mathbf{x})$ s is called the basis and nb is the size of the basis. In this case the procedure to find the c_i is independent of the choice of $\varphi_i(\mathbf{x})$ s while the evaluation of $\hat{H}\varphi(\mathbf{x})$, $\hat{H}^2\varphi(\mathbf{x})$ depend only on the choice of $\varphi_i(\mathbf{x})$ s since \hat{H} and \hat{H}^2 are linear operators, therefore

$$\hat{H} \left(\sum_{i=1}^{nb} c_i \varphi_i(\mathbf{x}) \right) = \sum_{i=1}^{nb} c_i \hat{H}\varphi_i(\mathbf{x}) \quad (2.6)$$

2.2.1. Searching for the c_i s minimizing $F1$.

Expanding (1.1) in detail we get

$$\begin{aligned} F1 &= \frac{\int_V \rho(\mathbf{x})(\hat{H}\varphi(\mathbf{x}) - E\varphi(\mathbf{x}))^2 dv}{\int_V \rho(\mathbf{x})\varphi^2(\mathbf{x})dv} = \\ &= \frac{\int_V \rho(\mathbf{x})(\hat{H}\varphi(\mathbf{x}))^2 - 2E \int_V \rho(\mathbf{x})\hat{H}\varphi(\mathbf{x})dv + E^2 \int_V \rho\varphi^2(\mathbf{x})dv}{\int_V \rho(\mathbf{x})\varphi^2(\mathbf{x})dv} \end{aligned} \quad (2.7)$$

We define the following quantities:

$$s = \int_V \rho(\mathbf{x})\varphi^2(\mathbf{x})dv \quad (2.8)$$

$$h = \int_V \rho(\mathbf{x})\varphi(\mathbf{x})\hat{H}\varphi(\mathbf{x})dv \quad (2.9)$$

$$g = \int_V \rho(\mathbf{x})(\hat{H}\varphi(\mathbf{x}))^2 dv \quad (2.10)$$

$$M(E) = h/s = E \quad (c.f.(1.2)) \quad (2.11)$$

$$M(E^2) = g/s \quad (2.12)$$

Eqs. (2.11) and (2.12) express the expectation value of the energy and its square.

Substituting (2.8) - (2.10) into (2.7) we get

$$F1 = \frac{(g - 2Eh + E^2s)}{s} = \frac{(gs - h^2)}{s^2} \quad (2.13)$$

Our task is thus to minimize (2.13). This will be done based on the fact that for at the minimum the partial derivatives of $F1$ with respect to the c_i s are all zero.

Let's denote the partial derivative with respect to c_i with prime and let

$$S_{ij} = \int_V \rho(\mathbf{x}) \varphi_i(\mathbf{x}) \varphi_j(\mathbf{x}) dv \quad (2.14)$$

$$H_{ij} = \frac{1}{2} \left(\int_V \rho(\mathbf{x}) \varphi_i(\mathbf{x}) \hat{H} \varphi_j(\mathbf{x}) dv + \int_V \rho(\mathbf{x}) \varphi_j(\mathbf{x}) \hat{H} \varphi_i(\mathbf{x}) dv \right) \quad (2.15)$$

$$G_{ij} = \int_V \rho(\mathbf{x}) \hat{H} \varphi_i(\mathbf{x}) \hat{H} \varphi_j(\mathbf{x}) dv \quad (2.16)$$

For the minimum, we have the equations

$$\frac{\partial F1}{\partial c_i} = F1' = \frac{(gs - h^2)'s^2 - 2(gs - h^2)ss'}{s^4} = 0 \quad (2.17)$$

for all i . As $s > 0$, whenever $\varphi(\mathbf{x})$ is not identically zero, it is enough to require

$$(gs - h^2)'s^2 - 2(gs - h^2)ss' = 0. \quad (2.18)$$

This can be expressed as

$$(g's + gs' - 2hh')'s^2 - 2(gs - h^2)ss' = g's^3 + gs^2s' - 2hs^2h' - 2gs^2s' + 2h^2ss' = 0 \quad (2.19)$$

Simplification and division by s gives

$$g's^2 - gss' - 2hsh' + 2h^2s' = 0 \quad (2.20)$$

Further simplifications and use of (2.11) and (2.12) gives

$$g' - 2M(E)h' + (2M(E)^2 - M(E^2))s' = 0 \quad (2.21)$$

Taking the partial derivative of (2.8)-(2.10) gives

$$s' = \sum_{j=1}^{nb} S_{ij}c_j \quad (2.22)$$

$$h' = \sum_{j=1}^{nb} H_{ij}c_j \quad (2.23)$$

$$g' = \sum_{j=1}^{nb} G_{ij}c_j \quad (2.24)$$

Substituting these into (2.21) we get

$$\sum_{j=1}^{nb} c_j (G_{ij} - 2M(E)H_{ij} + (2M(E)^2 - M(E^2))S_{ij}) = 0 \quad (2.25)$$

Let us define the matrix \mathbf{A} as

$$\mathbf{A} = \mathbf{G} - 2M(E)\mathbf{H} + (2M(E)^2 - M(E^2))\mathbf{S} \quad (2.26)$$

where the elements of \mathbf{G} , \mathbf{H} and \mathbf{S} are $G_{i,j}$, $H_{i,j}$ and $S_{i,j}$, respectively. Using this, (2.25) can be written in the form

$$\mathbf{A}\mathbf{c} = \mathbf{0} \quad (2.27)$$

where $\mathbf{0}$ denotes a matrix whose elements are all zero. In the followings we will refer to eqs. (2.25) or (2.27) as the Frost equation^{3,7,10}(FE).

The main difficulty in solving the FE stems from the fact that the equations only appear to be linear because $M(E)$ and $M(E^2)$ both depend on the \mathbf{c}_i s. However, the fact that the system of equations can be brought into this ‘pseudo’ linear form will be of tremendous help.

2.2.1.1. *The role of the weight factor $\varphi^2(x)$ in F2.*

We mentioned in Sec. 2.1 that serious difficulties arise if we want to use F2 without the $\varphi^2(\mathbf{x})$ factor. It was, however, possible, to find an equation determining the \mathbf{c}_i s that contain only the integrals of the basis functions that are (unlike $\varphi(\mathbf{x})$) independent of the c_i s. This was possible because there is no division in the integrands and in general it is true that for a given $\psi(\mathbf{x})$, $\varphi_i(\mathbf{x})(i = 1, 2, \dots, n)$

$$\int \left(\sum_{i=1}^n c_i \varphi_i(\mathbf{x}) \right) \psi(\mathbf{x}) dv = \sum_{i=1}^n c_i \int \varphi_i(\mathbf{x}) \psi(\mathbf{x}) dv \quad (2.28)$$

If, however, $\rho(\mathbf{x})$ does not contain $\varphi^2(\mathbf{x})$ then we get integrals of the type

$$\int \frac{\hat{H}\varphi(\mathbf{x})}{\varphi(\mathbf{x})} dv \quad (2.30)$$

for which this partition is possible only for very special choice of \hat{H} and $\varphi_i(\mathbf{x})$ (e.g., $\varphi_i(\mathbf{x})$ is a power and \hat{H} is the identity operator) and, furthermore, the integral can diverge.

2.2.1.2. *Simplifications of the Frost equation³.*

If we keep $M(E)$ constant (like the experimentally obtained energy of the system) then we can proceed the following way:

It can be seen from (2.27) that this changes the FE into a homogenous linear system of equations. It follows then that it can have nontrivial solution only if the determinant D of \mathbf{A} is nonzero¹⁹. As $M(E)$ is now fixed, $M(E^2)$ should be obtained from this condition. To this end, we calculate both $M(E^2)$ and D from randomly chosen \mathbf{c} s. This would give points of the function $D = D(M(E^2))$ and thus the zero of this function can be obtained by interpolation. Fixing the $M(E^2)$ value so found as well, the FE became a ‘true’ homogeneous system of linear equations. We did not use this method in our calculations.

2.2.1.3. *Solution of the Frost Equation with iteration*¹⁰.

An obvious procedure is to start with a trial \mathbf{c} , calculate with it both $M(E)$ and $M(E^2)$ and solve the resulting homogeneous system of linear equations. Using the solution, we calculate again $M(E)$ and $M(E^2)$. If the new values differ (with more than a threshold value) from the ones obtained in the previous step, we repeat the procedure, otherwise we take the \mathbf{c} obtained in the last step as the solution of the FE.

The procedure does not necessarily converge, we found cases during our calculations when it diverged.

For the applicability of the method we have to prove that the determinant of the matrix prepared with any \mathbf{c} is zero, i.e., (2.27) has nontrivial solution.

This would also prove that the FE has a nontrivial solution. The proof can be found in Appendix 1.

2.2.1.4. *An other possible iteration.*

Based on the procedure described in Sec. 2.2.1.2, we developed the following iteration:

Let's find the \mathbf{c} vector that is the solution of the simplified FE as described in Sec. 2.2.1.2, but consider it a solution only if the calculated $M(E)$ agrees with the value $M(E)$ was fixed to. If it does not, fix $M(E)$ to this new value and repeat the procedure.

Calculations were not done with this method either.

2.2.1.5. *Solution of the Frost equation with relaxation*¹⁰.

In this case the equations in (2.24) are solved one-by-one, in such a way that except for the c_i belonging to the equation being solved, all the other c_i s are considered fixed. The solutions are repeated until the minimum of $F1$ is reached (i.e., the c_i s found satisfy the FE). According to data in the literature, this method is about 10 times slower than the iteration described in Sec. 2.2.1.3.

2.2.1.6. *Determination of the minimum of $F1$ with a gradient method.*

In this case the minimum of the function $F1 = F1(\mathbf{c})$ is found directly. If the search for the minimum is done by a general minimizer, the the FE is not needed at all.

The algorithm described below takes advantage of the fact that the partial derivatives (i.e., its gradient) of $F1 = F1(\mathbf{c})$ can be obtained analytically - they are just the components of the FE.

The calculation proceeds as follows:

We start with a minimum stepsize, r_0 and set the actual stepsize, r , to this value. With an initial \mathbf{c}^0 vector we calculate the gradient of $F1$ and make a step, i.e., select a new vector \mathbf{c}^1 :

$$\mathbf{c}^1 = \mathbf{c}^0 + r\Delta(F1) \tag{2.31}$$

If $F1(\mathbf{c}^1) < F1(\mathbf{c}^0)$ then we perform an other step, but first we double the stepsize s . When for some i $F1(\mathbf{c}^i) < F1(\mathbf{c}^{i+1})$ then the gradient at \mathbf{c}^i is calculated and the minimization continues in the new direction.

We could possibly find lower values of $F1$ in the old direction also by repeating the last step with a reduced stepsize, but this appears unnecessary since the calculation of $F1$ for a given \mathbf{c} is of the same order of difficulty than the calculation of the gradient (i.e., requires about the same computer time). Also, it was found for the gradient methods in general²¹ that it is not efficient to find the minimum in a given direction with too much precision, instead it is preferable to calculate a new gradient earlier.

As the calculation of the gradient is independent of the stepsize, it makes sense to set the initial stepsize in a new direction large enough so that in each direction $F1$ should be evaluated only a few times. If we assume that the sum of the steps in any given direction is approximately constant then from this sum for the i -th direction we can estimate this sum for the $(i + 1)$ th direction. Therefore, if we could perform n steps (with doubling each time) in the i -th direction then it is reasonable to start with a step of size $r * 2^{n-1}$ because in this case the $(i + 1)$ -th direction can be ‘covered’ in one step. To avoid large fluctuations, it is advisable to increase the stepsize based on the average of two steps. For example, if the sum of steps in two subsequent directions exceeded 7 and in the last direction it exceeded 3 the in the next direction the initial stepsize will be four times the previous stepsize (this way we saved two steps).

This choice of stepsize will have as a consequence that the calculation will reach a stage where the initial stepsize is already too large and in this case it has to be reduced. In our procedure we reduced the stepsize in such case by 4, i.e., restored it to its previous value. When the stepsize is reduced to the initial minimal stepsize, we stop the calculation and consider the current \mathbf{c} vector the solution of the FE.

In our calculations we used $r_0 = 10^{-4}$. The gradient vector was normalized so that the largest absolute value of its components is 1 and the \mathbf{c} vector was normalized so that its first element was 1. These normalizations were necessary because the FE only determines \mathbf{c} up to a constant factor. Furthermore, r is only unequivocal if the gradient vector has the same norm all the time.

It should be pointed out that the method described above is always applicable when there is an analytical formula for the partial derivatives with respect to the parameters since we did not take advantage of the special form of the derivative. Thus the method is applicable to the determination of parameters of arbitrary type (see Sec. 2.2.3). If the parameters are not linear, the normalization of \mathbf{c} should, of course, be omitted.

2.2.2. Calculation of $\hat{H}\varphi_i(\mathbf{x})$.

As \hat{H} is a differential operator, there is no difficulty in principle in the calculation of $\hat{H}\varphi_i(\mathbf{x})$. For the computational efficiency, however, the way it is calculated is of importance. As the differentiation of the $\varphi_i(\mathbf{x})$ functions generally used follows strict rules, it is enough, in principle, to specify the function and calculate $\hat{H}\varphi_i(\mathbf{x})$ with a general algorithm. For

larger systems, it is likely that this route will be followed because the required derivations will become more complex and cumbersome.

For a given choice of functional form, however, the computation time can be reduced significantly if the formula for $\hat{H}\varphi_i(\mathbf{x})$ is derived at the outset, and with the choice of appropriate coordinates. Internal coordinates appear to be the best choice.

2.2.2.1. The form of the operator \hat{H} in internal coordinates.

The Hamiltonian can be written as

$$\hat{H} = \hat{T} + \hat{V} \quad (2.32)$$

Here the operator \hat{V} denotes multiplication by the potential energy V and \hat{T} is the kinetic energy operator:

$$\hat{T} = -\frac{1}{2} \sum_{i=1}^n \frac{1}{m_i} \Delta_i \quad (2.33)$$

where n is the number of particles, m_i is the mass of the i -th particle and

$$\Delta_i = \frac{\partial^2}{\partial x_i^2} + \frac{\partial^2}{\partial y_i^2} + \frac{\partial^2}{\partial z_i^2} \quad (2.34)$$

where x_i , y_i , and z_i are the coordinates of the i -th particle.

Let r_{ij} denote the distance between particles i and j (particle here can be an electron or a nucleus), $\vartheta_{ij,ik}$ denote the angle between the directions from particle i to particles j and k and z_i the charge of particle i . Assuming that only Coulomb forces are present, we obviously have

$$V = \sum_{i < j} z_i z_j / r_{ij} \quad (2.35)$$

Using the chain rule gives for \hat{T} (without showing the details)

$$\begin{aligned} \hat{T} = \sum_{i < j} \left[-\frac{1}{2} \left(\frac{1}{m_i} + \frac{1}{m_j} \right) \left(\frac{\partial^2}{r_{ij}^2} + \frac{2}{r_{ij}} \frac{\partial}{r_{ij}} + \frac{z_i z_j}{r_{ij}} \right) \right] \\ - \sum_i \frac{1}{m_i} \sum_j \sum_{k \neq i, j < k} \cos \vartheta_{ij,ik} \frac{\partial^2}{\partial r_{ij} \partial r_{ik}} \end{aligned} \quad (2.36)$$

Substituting (2.35) and (2.36) into (2.32) we obtain \hat{H} in terms of internal coordinates.

2.2.2.2. Numerical calculation of $\hat{H}\varphi_i(\mathbf{x})$ ²³.

We saw in Sec. 2.2.2.1 that the difficulties arising in evaluating the Hamiltonian stem from the kinetic energy operator, $\frac{1}{2} \sum_i \frac{1}{m_i} \Delta_i$. If we perform the differentiations numerically, we obtain a procedure that is practically independent of the choice of $\varphi_i(\mathbf{x})$. According

to published results, using a stepsize of 0.0001 the difference between the numerically and analytically calculated derivatives appears only in the fourth decimal digit.

2.2.3. *Optimizing the non-linear parameters.*

In general the $\varphi_i(\mathbf{x})$ contain numerical parameters and the degree to which our solution approximates the solution of the SE depends on them. It is obvious that by increasing the number of $\varphi_i(\mathbf{x})$ s i.e., extending the basis, the approximation can be improved. We will see, however (Sec. 3.5.7) that this would drastically increase the computation time. It is thus self-evident that it is worth trying to improve the selection of these numerical parameters without extending the basis.

This means that the function space T where we are looking for the solution of the SE is not linear, therefore the solution is not obtained from the FE. Studies in the literature show that optimizing the nonlinear parameters is especially significant for calculations on excited states¹³.

2.2.3.1. *Optimizing the nonlinear parameters with the help of the linear parameters*³⁰.

If we have some idea about the range in which the nonlinear parameters fall, the original basis functions can be replaced by a set of functions where the nonlinear parameters are set to different values. The linear parameters of this extended basis will give different weights to these parameter values. Based on this, the calculation might be repeated with a finer sampling and finally a good approximation could be obtained for the optimal value of the nonlinear parameters.

2.2.3.3. *A special form of $\hat{H}\varphi_i(\mathbf{x})$ and $\varphi_i(\mathbf{x})$.*

If we plan to optimize also the nonlinear parameters in our calculations, it appears to be advantageous¹⁸ to use a special form of $\hat{H}\varphi_i(\mathbf{x})$ and $\varphi_i(\mathbf{x})$:

It can be shown that if the basis functions can be written as a polynomial multiplied by an exponential (and calculations to date used such functions) then a function $q(\mathbf{x})$ can be found such that

$$q(\mathbf{x})\hat{H}\varphi_i(\mathbf{x}) = \sum_{j=1}^{nb} \varphi_j h_{ji} \quad (2.37)$$

$$q(\mathbf{x})\varphi_i(\mathbf{x}) = \sum_{j=1}^{nb} \varphi_j s_{ji} \quad (2.38)$$

where the h_{ji} and s_{ji} can be determined knowing $q(\mathbf{x})$. We did not use this form in our calculations because we kept the nonlinear parameters constant (see Sec. 3.1).

2.3. *The choice of the quadrature points.*

Calculations in the literature used three different types of quadratures:

1. Equidistant^{7,10,11}.

2. Gauss-Laguerre quadrature^{11–13,16}.
3. Random quadrature^{3,22,23}.

The first description of the method^{7,10} used points in the configuration space with uniform distribution, thus essentially used an equidistant quadrature, although reference is made to the advantages of the Gaussian quadrature. In subsequent calculations they successfully used the Gauss-Laguerre quadrature.

Random quadrature was first used by H. Conroy³, but his method differs from the ‘conventional’ LEM. In his calculations he used the simplification mentioned in Sec. 2.2.1.2.

T.A. Rourke and E.T. Steward also used a random quadrature^{22,23}. In our calculations we also employed a random quadrature but without the simplification mentioned in Sec. 2.2.1.2. The random quadrature has the following advantages:

1. The number of quadrature points can be arbitrary, there is no special number to use. With the Gauss-Laguerre quadrature, the number of quadrature points have to take special values and these values increase exponentially (if we want to use n point in one direction, the total number of points will be k^n).

2. The contributions of each point to the matrices \mathbf{S} , \mathbf{H} and \mathbf{G} can be very simply calculated, there are no weight factors or volume elements to worry about, thus the method is rather general and simple. Its drawback is that it is difficult to reach high precision with it.

2.3.1. *Random quadrature (Monte Carlo method)*²⁵.

The random quadrature is based on the following equality:

If $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N$ are vectors uniformly distributed in a bounded closed region of volume V of the configuration space then

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N f(\mathbf{x}_i) = \frac{1}{V} \int_B f(\mathbf{x}) dv \quad (2.39)$$

A proof of (2.39) is given in Appendix 2.

The proof considers the $f(\mathbf{x}_i)$ as probability variables and show that both sides of (2.39) is equal to the expectation value of $f(\mathbf{x})$.

If the values $f(\mathbf{x}_i)$ are the values of a probability variable $f(\mathbf{x})$ then by definition

$$\sum_{i=1}^N \frac{1}{N} f(\mathbf{x}_i) = M(f(\mathbf{x})). \quad (2.40)$$

It is customary to select points for a Monte Carlo method with nonuniform distribution (importance sampling). If the probability density function of the points \mathbf{x}_i is $p(\mathbf{x}) > 0$ then

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N \frac{f(\mathbf{x}_i)}{p(\mathbf{x}_i)} = M(f(\mathbf{x})) = \int_B f(\mathbf{x}) dv \quad (2.42)$$

This can be seen as follows:

Starting from the obvious identity

$$\int_B f(\mathbf{x})dv = \int_B \frac{f(\mathbf{x})}{p(\mathbf{x})}p(\mathbf{x})dv \quad (2.43)$$

we have, based on the definition of the expectation value

$$\int_B \frac{f(\mathbf{x})}{p(\mathbf{x})}p(\mathbf{x})dv = M\left(\frac{f(\mathbf{x})}{p(\mathbf{x})}\right) \quad (2.44)$$

since $p(\mathbf{x})$ is the probability density of $f(\mathbf{x})$. The law of large numbers³¹ gives

$$M\left(\frac{f(\mathbf{x})}{p(\mathbf{x})}\right) = \lim_{N \rightarrow \infty} \sum_{i=1}^N \frac{f(\mathbf{x}_i)}{p(\mathbf{x}_i)}, \quad (2.45)$$

proving (2.42).

Based on (2.42) we can prove that the $\rho(\mathbf{x})$ in (1.1) can be included into our calculations in such a way that it only appears in the formalism for the selection of the points: since the region is bounded, for bounded $\rho(\mathbf{x})$, $\int_V \rho(\mathbf{x})$ is bounded thus it can be normalized to 1 and $\rho(\mathbf{x})$ can be considered a probability density. If we select the points according to the probability distribution $\rho(\mathbf{x})$ we have

$$M(f(\mathbf{x})) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N \frac{\rho(\mathbf{x}_i)f(\mathbf{x}_i)}{\rho(\mathbf{x}_i)} = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N f(\mathbf{x}_i) \quad (2.46)$$

thus $\rho(\mathbf{x})$ dropped out from the quantity to be calculated at each point. Besides simplifying the calculations, this result also proves that the calculation of various local energies indeed are equivalent to minimizing (1.1).

It is to be noted that the $\rho(\mathbf{x})$ weight factor can also be used in the ‘conventional’ way, i.e., it does not have to be the probability distribution of the points.

2.3.2. Estimation of the error of a random quadrature.

If we have to minimize an integral with respect to some parameters then we need an estimate of the error in minimization that is due to the quadrature error. This estimate will provide also a criterion for the decision on increasing the number of points.

For a random quadrature, using the distribution function $F(X) = P(f(\mathbf{x}) \leq X)$ introduced with the proof of (2.39), we succeeded in constructing such an estimate:

Let $f_o(\mathbf{x})$ be an element of the function space T (T is defined by the parameters of the function to be integrated) such that its integral is minimal and let

$$\text{MIN} = \int_B f_o(\mathbf{x})dv \quad (2.47)$$

Furthermore, let $f_1(\mathbf{x})$ be a function in T that minimizes

$$\frac{1}{N_A} \sum_{\mathbf{x} \in A} f(\mathbf{x}_i) \quad (f(\mathbf{x}) \in T)$$

on a set A of randomly (e.g., with uniform distribution) selected but then fixed points from B , i.e., the approximation of $\int_B f(\mathbf{x})dv$ with a random quadrature and K be an other set of points from B that follow the same distribution as the points in A but they are chosen independent of the points in A .

In this case it can be shown that

$$M \left(\frac{1}{N_A} \sum_{\mathbf{x}_i \in A} f_1(\mathbf{x}_i) \right) \leq \text{MIN} \leq M \left(\frac{1}{N_K} \sum_{\mathbf{x}_i \in K} f_1(\mathbf{x}_i) \right) \quad (2.48)$$

where N_A and N_K are the number of points in A and K , respectively.

We can prove (2.48) as follows:

Using (2.42) we have

$$M \left(\frac{1}{N_A} \sum_{\mathbf{x}_i \in A} f_1(\mathbf{x}_i) \right) = \int_B f_1(\mathbf{x})dv \quad (2.49)$$

The righ-hand inequality follows form the definition of MIN and from (2.42). By the definition of $f_1(\mathbf{x})$ it follows that

$$\frac{1}{N_A} \left(\sum_{\mathbf{x}_i \in A} f_1(\mathbf{x}_i) \right) \leq \frac{1}{N_A} \left(\sum_{\mathbf{x}_i \in A} f_o(\mathbf{x}_i) \right) \quad (2.50)$$

giving

$$M \left(\frac{1}{N_A} \sum_{\mathbf{x}_i \in A} f_1(\mathbf{x}_i) \right) \leq M \left(\frac{1}{N_A} \sum_{\mathbf{x}_i \in A} f_o(\mathbf{x}_i) \right) \quad (2.51)$$

Since the righ-hand side of (2.51) is just MIN, we proved the lefthand inequality of (2.48) as well.

Based on the inequalites (2.48) we can thus give a lower and upper bound for the minimum value of the integral. Our approximation can be considered refined enough when the two expectation values, i.e., $M \left(\frac{1}{N_A} \sum_{\mathbf{x}_i \in A} f_1(\mathbf{x}_i) \right)$ and $M \left(\frac{1}{N_A} \sum_{\mathbf{x}_i \in A} f_o(\mathbf{x}_i) \right)$ in (2.48) agree with each other and thus with MIN.

Equality of expectation values that were obtained from statistical methods can also be decided by statistical methods. The Student t test appears to be appropriate for this, but the equality of the variances first have to be demonstrated, e.g., using the F test.

We should mention in closing that if the sum $\frac{1}{N_A} \left(\sum_{\mathbf{x}_i \in A} f_1(\mathbf{x}_i) \right)$ is not minimized by $f_1(\mathbf{x})$ in T then the agreement between the two expectation values is only necessary but not sufficient condition for $f_1(\mathbf{x})$ to minimize the integral.

For the traditional LEM $F1$, as defined by (1.1) is to be minimized. It should be stressed that our estimate is valid for any measure written as an integral.

2.3.2.1. *The estimates of T.A. Rourke and E.T. Stewart*²².

The authors calculated the variance (σ_n) from different systems using 2-200 points on the wave function obtained by the variational principle (atoms and molecules with 1-4 electrons). From this data they set up the regression

$$\log \sigma_n = \log \sigma - A \log n + B \quad (2.54)$$

where A and B are regression parameters, n is the number of points and σ is the variance calculated on the 200 points. Thus, knowing σ_n gives us an estimate for σ that characterizes the quality of the actual approximation and also indicates the precision obtainable with more points.

In our opinion this estimate becomes erroneous if for a given calculation the basis set is significantly better or worse than the basis set used to establish the regression parameters (e.g., if the basis is complete, i.e., contains the solution of the SE, then we have to have $A = B = 0$).

3. *Calculations.*

As mentioned in the introduction, calculations were performed on the hydrogen molecular ion and on the helium atom. We give the details of the calculations in this section.

3.1. *The choice of $\varphi_i(x)$.*

Nothing specific was said yet about the $\varphi_i(\mathbf{x})$ of (2.5). Clearly, the precision of the calculation depends on their choice.

If we choose the $\varphi_i(\mathbf{x})$ in such a way that the solution of the SE can be written in the form (2.5) with this choice then we can find the exact solution of the SE, and it is enough to take as many points as the number of $\varphi_i(\mathbf{x})$ s. If the solution of the SE can not be so written then still we should look for a ‘best choice’ for the $\varphi_i(\mathbf{x})$. As discussed in Sec. 2.2.3., the precision can be increased by increasing appropriately the number of the $\varphi_i(\mathbf{x})$ s, but this results in a drastic increase (see Sec 3.5.7.) in the computation time.

By appropriate increase we mean that the precision resulting from a calculation using n $\varphi_i(\mathbf{x})$ functions can only be improved if the $\varphi_{n+1}(\mathbf{x})$ is linearly independent of the first n ,

i.e.,

$$\sum_{i=1}^n c_i \varphi_i(\mathbf{x}) = \varphi_{n+1}(\mathbf{x}) \quad (3.1)$$

can not hold with any set of c_i s. In general, if the $\varphi_i(\mathbf{x})$ are not linearly independent, at least one of them can be dropped without reducing the precision of the calculation.

The linear dependence among the $\varphi_i(\mathbf{x})$ is to be avoided not only because it causes unnecessary complications but because the expression $\sum_{i,j} c_i c_j S_{ij}$ can become zero (see Appendix 1) that would cause problems when trying to solve the FE.

If we can avoid this difficulty, we are still faced with the problem that even the ground-state solution of the FE is not necessarily unique.

For the H_2^+ we used the following $\varphi_i(\mathbf{x})$:

$$\varphi_i(\mathbf{x}) = e^{-\xi(r_A+r_B)}(r_A^{n_i} r_B^{m_i} + r_A^{m_i} r_B^{n_i}) \quad (3.2)$$

Here r_A and r_B are the distances of the electron from the two nuclei, ξ a constant and the values of n_i and m_i are given in Table 1.

Table 1. The values of n_i and m_i

i	1	2	3	4	5	6	7	8
n_i	0	1	2	1	3	2	2	3
m_i	0	0	0	1	0	1	2	1

Remark: The values up to $i = 6$ were used in the original work on the LEM⁷.

For the He atom the $\varphi_i(\mathbf{x})$ were of the following form:

$$\varphi_i(\mathbf{x}) = e^{-\xi(r_1+r_2)}(r_1^{n_1} r_2^{n_2} r_{12}^{n_3} + r_1^{n_2} r_2^{n_1} r_{12}^{n_3}) \quad (3.3)$$

Here r_1 and r_2 are the distances of the electron from the nucleus, r_{12} is the distance between the two electrons ξ a constant and the values of n_1 , n_2 and n_3 are given in Table 2.

Table 2. The values of n_1 , n_2 and n_3

i	1	2	3	4	5	6
n_1	0	0	1	0	0	1
n_2	0	1	1	0	1	1
n_3	0	0	0	1	1	1

The values of the constants were selected in such a way that at infinity the energy expression should give the experimental value. This gave $\xi = 1.41$ for H_2^+ and $\xi = 1.816$ for He.

It should be pointed out that at this point the method ceases to be purely ab initio and becomes semiempirical. For systems where the energy is not known, this procedure is

not applicable (according to some authors, this is the most serious obstacle in extending the method to larger systems¹³). The optimization of the nonlinear parameters becomes important just in such cases.

The basis functions of the form (3.2) and (3.3) have the property that if the exponent values are different for each $\varphi_i(\mathbf{x})$, i.e. there is no pair of $\varphi_i(\mathbf{x})$ with the same set of exponents, then they are linearly independent. This will be proven in Appendix 3.

3.2. The expressions for $\hat{H}\varphi_i(x)$.

The special cases of the expressions (2.32), (2.35), (2.36) for the Hamiltonian can be written once the form of $\varphi_i(\mathbf{x})$ is known. In our case a further restriction is that we consider the nuclei fixed and we assume that $\varphi_i(\mathbf{x})$ is a product of some functions (this latter can also be considered an extension). The two special expressions that we need are special cases of these.

For the H_2^+ molecular ion⁷

$$\begin{aligned} \hat{H}\varphi_i(\mathbf{x}) &= e^{-\xi(r_A+r_B)} * \\ & [r_A^{n_i} r_B^{m_i} (\xi^2 (1 + \cos \vartheta_{AB}) - (\xi - 1) \left(\frac{1}{r_A} + \frac{1}{r_B} \right) - n_i \frac{\xi}{r_A} (1 + \cos \vartheta_{AB}) \\ & - m_i \frac{\xi}{r_B} (1 + \cos \vartheta_{AB}) + n_i (n_i + 1) \frac{1}{2r_A^2} + m_i (m_i + 1) \frac{1}{2r_B^2} + n_i m_i \frac{\cos \vartheta_{AB}}{r_A r_B}] \\ & + r_A m_i r_B n_i (\xi^2 (1 + \cos \vartheta_{AB}) - (\xi - 1) \left(\frac{1}{r_A} + \frac{1}{r_B} \right) - m_i \frac{\xi}{r_A} (1 + \cos \vartheta_{AB}) \\ & - n_i \frac{\xi}{r_B} (1 + \cos \vartheta_{AB}) + m_i (m_i + 1) \frac{1}{2r_A^2} + n_i (n_i + 1) \frac{1}{2r_B^2} + n_i m_i \frac{\cos \vartheta_{AB}}{r_A r_B}] \end{aligned} \quad (3.4)$$

where ϑ_{AB} is the angle between the directions from the electron to the two nuclei.

For the He atom¹⁴ we have

$$\begin{aligned} \frac{\hat{H}\varphi(\mathbf{x})}{\varphi_i(\mathbf{x})} &= -2n_3(n_3 + 1)/r_3^2 + 2V - \\ & 2\xi^2 + \frac{1}{2} \left((n_1 + 1) \left(\frac{n_1}{r_1^2} - \frac{2\xi}{r_1} \right) (n_2 + 1) \left(\frac{n_2}{r_2^2} - \frac{2\xi}{r_2} \right) \right) \\ & + \left(\xi - \frac{n_1}{r_1} \right) \frac{n_3}{r_3} \cos \vartheta_{13} + \left(\xi - \frac{n_2}{r_2} \right) \frac{n_3}{r_3} \cos \vartheta_{23} \\ & + \frac{1}{2} \left((n_2 + 1) \left(\frac{n_2}{r_1^2} - \frac{2\xi}{r_1} \right) + (n_1 + 1) \left(\frac{n_1}{r_2^2} - \frac{2\xi}{r_2} \right) \right) \\ & + \left(\xi - \frac{n_2}{r_2} \right) \frac{n_3}{r_3} \cos \vartheta_{13} + \left(\xi - \frac{n_1}{r_1} \right) \frac{n_3}{r_3} \cos \vartheta_{23} \end{aligned} \quad (3.5)$$

Here ϑ_{23} is the angle between the directions from the nucleus to the two electrons.

3.3. Solution of the Frost equation, round-off errors.

In our calculations we used the method described in 2.2.1.6 for the solution of the FE. As mentioned there, the threshold stepsize for the decision on the improbability of a solution was set to 10^{-4} . If we tried a smaller value, then the solutions so obtained turned out to be worse than the solution obtained with the 10^{-4} threshold (i.e., the value of the new ‘minimum’ became larger), due to round-off errors. (Clearly, the optimal value of the threshold is not unique and depends on the problem, and within the problem on the number of points used.)

The resulting coefficients were substituted back into the equation to be solved. As our experiences also showed, the error of the solution increased with the number of points. Using a small number of points (about 50) gave 10^{-5} - 10^{-6} instead of 0 on the right hand side, with a large number of points (about 1000) this value grew to 10^{-2} - 10^{-3} . This is a consequence of the increase in the magnitude of the elements of \mathbf{S} , \mathbf{H} and \mathbf{G} . It follows from their definition that they are obtained as a sum of terms of the same sign and of similar order of magnitude. Thus the values of the matrix elements increase proportionally with n , while the order of magnitude of the contributions are unchanged, thus during the summation we are losing as many significant digits as the difference between the order of magnitudes of the matrix element and the individual contribution.

We also tried the iteration described in Sec. 2.2.1.3. but in the cases tried it did not converge.

Round off errors showed up here, too. The determinant of the resulting system of linear equations was around 10^{-10} - 10^{-12} instead of zero, and the right hand side of the equation dropped we obtained, after substitution, 10^{-3} - 10^{-6} instead of zero.

3.4. The choice of the quadrature points.

As mentioned in Sec. 2.3., we used a random quadrature in our calculations. Three kinds of weight functions were used:

$$r \leq K \quad \rho(\mathbf{x}) = (4/3)\pi K^3 \quad (3.6)$$

$$r > K \quad \rho(\mathbf{x}) = 0$$

$$r \leq K \quad \rho(\mathbf{x}) = r/K \quad (3.7)$$

$$r > K \quad \rho(\mathbf{x}) = 0$$

$$r \leq K \quad \rho(\mathbf{x}) = Ne^{-\xi r} \quad (3.8)$$

$$r > K \quad \rho(\mathbf{x}) = 0$$

These weight functions refer to the coordinate of one electron. For the helium atom, the weight function is the product of the weight functions of the two electrons.

In (3.6)-(3.8) r is the distance of the electron from the origin, N is a normalization factor and K determines the region within which the wave function is considered non zero, i.e. the ‘radius’ of the atom or the molecule. We used different K values in our calculations. The points with the required distribution were obtained from the random number generator

of the Gier Algol-4 compiler that generates pseudo random numbers on the $[-1,1]$ closed interval. (A sequence of numbers is called pseudo random numbers when, in spite of being determined by its first element, the elements possess given statistical properties.)

Points from the distributions (3.6)-(3.8) can be obtained from the uniform distribution — with an appropriate transformation. If the coordinates of a sequence of points are uniformly distributed in $[-1,1]$ then the points are uniformly distributed in the cube of edge two centered at the origin. Dropping the points that fall outside the unit sphere, we obtain a sequence uniformly distributed points in the unit sphere.

The distribution corresponding to (3.6) is uniform in a sphere of radius K . Such points can be obtained if the coordinates of the points uniformly distributed in the unit sphere are multiplied with K .

The distance of the points corresponding to the distribution (3.6) from the origin is uniform. (This distribution differs from (3.6)!.) We can obtain such points if we first generate a number r_1 from the uniform distribution in the $[0, K]$ interval then generate a point from the uniform distribution in the unit sphere and multiply its x , y and z coordinates with $r_1/\sqrt{(x^2 + y^2 + z^2)}$.

To procedure to generate points according to the distribution (3.8) follows the procedure for (3.7). These are points whose distance from the origin follows exponential distribution. It is known²³ that a point r_2 from the $\rho(x)$ distribution can be obtained from a point r_1 in the uniform distribution in the $[0,1]$ interval from the condition

$$\int_0^{r_2} \rho(\mathbf{x}) dr = r_1 \quad (3.9)$$

Substituting (3.8) into (3.9) we get

$$N \int_0^{r_2} e^{-\xi r} dr = r_1 \quad (3.10)$$

$$N \left(\frac{1}{\xi} - \frac{1}{\xi} e^{-\xi r_2} \right) = r_1 \quad (3.11)$$

$$r_2 = - \left(\ln \left(1 - \frac{\xi r_1}{N} \right) \right) / \xi \quad (3.12)$$

N is obtained from the normalization condition

$$N \int_0^\infty e^{-\xi r} dr = 1 = \int_0^K e^{-\xi r} dr = N \left(\frac{1}{\xi} - \frac{1}{\xi} e^{-\xi K} \right) \quad (3.13)$$

giving

$$N = \frac{\xi}{1 - e^{-\xi K}} \quad (3.14)$$

From (3.12) and (3.14) we get for the transformed r_2

$$r_2 = -(\ln[1 - r_1(1 - e^{-\xi K})])/\xi \quad (3.15)$$

Once r_2 is obtained, the transformation used for the distribution (3.2) can be used to obtain the required point from a uniform distribution.

There is one more thing to consider when selecting the quadrature points. It can be seen from the form of $\hat{H}\varphi_i(\mathbf{x})$ that the expressions are not valid when the coordinates of an electron and a nucleus coincide or if the coordinates of two electrons coincide. (since the expressions contain $1/r_1, 1/r_2, 1/r_A, 1/B, 1/r_{12}$). This can occur with very small probability, but the precision of the calculation degrades significantly when any of the expressions above becomes too large ($\hat{H}\varphi_i(\mathbf{x})$ will be obtained as a small difference between large values and this necessarily brings a loss of precision). Thus, when selecting the quadrature points, it is advisable to eliminate such situation and exclude points in the vicinity of the nucleus and reject points where the two electrons are too close to each other. In our calculations the cutoff distance from the nucleus was 10^{-2} and for the electrons 10^{-3} .

3.4.1. *The domain of integration.*

The weight functions (3.6)-(3.8) are all limit the integration from the full configuration space to a closed bounded region. The use of such weight functions was necessary not only to avoid calculations where the solution is known ($\equiv 0$) but also because the random quadrature is only valid in a finite domain since we have to divide with the volume of the domain (The Gauss-Laguerre quadrature does not have this restriction).

We can, however, ask whether it makes sense to limit the integration to a domain outside which the wavefunction is not even approximately zero. Clearly, if the exact eigenfunction is in the function space T defined by the basis functions then it can be obtained from an arbitrarily small region of the configuration space.

If the function space T does not contain the exact eigenfunction, and we minimize the integral (1.1) in a domain outside which the eigenfunction is not zero then the $\varphi(\mathbf{x})$ obtained after minimization can be considered an approximation of the exact eigenfunction only in this domain. In this case, however, we can not obtain its normalization factor since we don't know the function in the full configuration space, or we don't consider it to be an approximation to the exact eigenfunction outside this domain.

It also follows that in this case the approximation (1.3) to the expectation value of the energy will be less appropriate and the average of the local energies seems to be a better approximation.

Our calculations on the helium atom (Table 5) appears to support this.

3.5. *The computer program.*

The calculations described in this paper were performed with a program written in the Gier Algol 4 language. The deviations from Algol 60 are documented in a MAVEMI publication to this effect.

3.5.1. *Major subroutines.*

The program has been written to be as general as possible, i.e., it is applicable to arbitrary systems with only small modifications. Information referring the system in question (name, number of electrons, expressions for $\hat{H}\varphi_i(\mathbf{x})$, $\varphi_i(\mathbf{x})$, the positions of the nuclei (the region to be avoided during the generation of the points)) is contained in three subroutines (**procedure**): INIT, FHF and KIEG. The expectation and variance of the energy is computed from the matrices **S**, **H**, and **G** and from the vector **c** by the subroutine VV. The FE is solved by the subroutines SOLVE and SOLVEI, the former with the gradient method and the latter with the iteration described in Sec. 2.2.1.3. SOLVEI uses the Gier library routine for solving systems of linear equations.

There are several simple subroutines in the program for printing error messages, for inputting numbers, logical values, vectors and matrices where the program prompts the user for the quantity to be inputted.

3.5.2. *The running of the program.*

The program starts with reading interactively the specifics of the run (to be echoed at the top of each output page), i.e., the size of the basis, the radius of the system (**K**) the type of the weight factor, the random number generator seed, the precision required (**eps**) than from paper tape it reads the data characterizing the basis set (in our case the n_i , m_i or the n_1 , n_2 , n_3 values as well as an initial trial solution (a vector **c**). After that it again interactively reads the number of points in the quadrature set and un the control set. This is followed by the generation of the quadrature points, the calculation of $\hat{H}\varphi_i(\mathbf{x})$, $\varphi_i(\mathbf{x})$, and the matrices **S**, **H**, and **G** based on the quadrature points generated and solves the FE so obtained. This is followed by the preparation of the control points. Next, the program calculates the energy and the variance on them, performs the statistical tests needed and asks for a new number of control points. If this is given as zero, it combines the matrices **S**, **H**, and **G** calculated from the quadrature points and the control points, merges the control point set into the quadrature point set, calculates $\hat{H}\varphi_i(\mathbf{x})$, $\varphi_i(\mathbf{x})$ at the new points and solves the latest FE one last time. Before stopping, it also prints on a separate page the individual local energy, the final **c** vector, the energy and its variance as well as the mean local energy and its variance. (clearly, if our calculation is exact then the mean local energy should agree with the energy thus the difference between the two kinds of energy value is an additional indicator of the precision of the calculation.) The final results and the intermediate results (for each step the energies, variances, mean local energies and their variances, the F and t values of the statistical tests) are also written on paper tape to serve as an input to additional calculations or analysis programs. The intermediate results are also tabulated on the last output page.

3.5.3. *Options to change the calculation.*

During the calculations, it is possible to change the course by setting the KB register on the console to 1. Three changes are possible:

1. Stop the solving of the FE.

2. Enter a new eps.

3. Switch from the currently used method for solving the FE to the other one (from the gradient method to iteration or back). This is the only way to use the iteration method because the program always starts with the gradient method.

3.5.4. *Printing of the partial results.*

The program can print the following partial results:

1. The **S**, **H**, and **G** matrices.

2. The values of $\hat{H}\varphi_i(\mathbf{x})$, $\varphi_i(\mathbf{x})$.

3. At each step during the solution of the FE the vector **c**, the gradient vector, the energy and its variance.

4. The local energies.

5. The mean and variance of $(\hat{H}\varphi(\mathbf{x}) - E\varphi(\mathbf{x}))^2$ (the former is the variance of the energy).

Printing of these items can be requested at the start of the run. Also, by setting the KA register on the console to 1, the program will print these, independent of the way the run was started.

3.5.6. *Debugging the program.*

Experience showed that a computer program can have bugs showing up in a variety of contexts. The verification of a program is not just a practical matter but it is an important problem in and of itself.

The algorithms for the solution of the FE — as any algorithm solving a system of equations — can be reliably checked by substituting the result into the equations - as discussed in Sec 3.3.

By printing intermediate results, as well as examining special cases (one or two basis functions, 1 – 2 points) the program part calculating the matrices **S**, **H**, and **G** can be checked by hand calculations. For larger number of points, a good test was the preparation of the matrices **S**, **H**, and **G** in several steps in different ways, and comparing the printed values.

The most difficult to checking is the part of the program calculating $\hat{H}\varphi_i(\mathbf{x})$. The printed values can, in principle, can be checked by hand valcation but it is rather lengthy and thus not too reliable. Instead, we obtained the expression of $\hat{H}\varphi_i(\mathbf{x})$ in cartesian system also, and using a different program based on these formulae, calculated $\hat{H}\varphi_i(\mathbf{x})$ in an alternative way and compared the values from the different programs. Incidentally, this also provided a check on the formulae used.

Finally, we would like to point out that the comparison of the calculated and experimental energies can also be considered a check on the correctness of the program.

3.5.7. *The dependence of the computation time on the size of the problem.*

To a good approximation, the computation time is spent on three different tasks. The computing time spent on them can be expressed in terms of the number of quadrature points, P , and the size of the basis, nb , as follows:

$c_1 * P * nb$: preparation of $\varphi_i(\mathbf{x})$, $\hat{H}\varphi_i(\mathbf{x})$

$c_2 * P * nb^2$: preparation of the matrices \mathbf{S} , \mathbf{H} , and \mathbf{G}

$c_3 * L * nb^2 * (S + 1)$: solution of the FE (gradient method)

Here c_1 , c_2 , c_3 are proportionality constants, of which c_2 and c_3 depend on the system studied, L is the number of steps needed to solve the FE and S is the number of steps made in one direction.

We can see that for larger NB the time required for the preparation of $\varphi_i(\mathbf{x})$, $\hat{H}\varphi_i(\mathbf{x})$ becomes negligible relative to the rest of the work, thus for larger basis sets no significant gain can be expected from using a special formula for $\hat{H}\varphi_i(\mathbf{x})$. In our case c_1 was of the order of $3c_3$, so for us the gain was still significant.

c_3 can be considered proportional to the number of gradient calculations needed during the solution of the FE, but this clearly depends on the trial solution. If we perform the calculations in several steps, i.e., in each step we only take a small number of quadrature points then the solution of the last FE will differ little from the solution of the previous FE, thus we will need fewer gradients, so c_3 will become smaller.

The factor $nb^2 * (S + 1)$ results from the necessity of calculating one gradient and S function values in each direction, but each requires the elements of the matrices \mathbf{S} , \mathbf{H} , and \mathbf{G} . From our choice of the trial solution S averages around 2-3.

The computing time required for the solution of the FE with iteration is

$$c'_3 * L * nb^3$$

(because each iteration requires the solution of a system of linear equations), where c'_3 is a constant that, similarly to the gradient method, is proportional to the number of iterations required.

We could not find neither theoretical or experimental guide for the relative number of iterations with the two methods that could decide which method is more efficient. If they are assumed to be of the same order of magnitude (independent of the size of the basis) then the gradient method used by us is the more efficient.

4. *Results.*

4.1. *Energies.*

The results of the calculations on H_2^+ and He are summarized in Tables 3-5.

Due to the small size of the basis it is difficult to compare the results with data in the literature but they do demonstrate the correctness of the method.

4.2. *The dependence of the result on the choice of region.*

For both systems it can be seen that the results deteriorate for too large a region ($K = 4.5$).

The results also seem to support (Table 5) the statement in Sec. 3.4.1 that if the wavefunction is significantly nonzero outside the region of the calculations then the mean local energy is a better approximation to the experimental energy than the expectation value of the energy.

4.3. *The stability of the results.*

Table 6 compares the results of calculations with different quadrature point sets. It can be seen that the variance of the energy is rather stable showing that the precision can not be significantly increased by taking more points since the variance would remain about the same. The energy values fluctuate more, however.

4.4. *The criterion for stopping the calculation.*

Figure 1 shows the variance on the quadrature and control points together with the Student t value. The variance curves follow the behaviour predicted by (2.43).

Finally, we should point out that there is nothing unique about the choice of the t test that is designed establish that two samples are different. We could just as well use a test that is designed for demonstrating the equality of two samples.

5. *Evaluation of the method.*

We would like to summarize the main advantages and drawbacks of the LEM.

1. It is a fundamentally positive feature that by increasing the size of the basis and of the quadrature points it can provide the exact solution of the SE. This, of course, is mainly a question of principle.

2. An additional attractive feature is the fact that if the studies are performed in a function space that contains the solution of the SE then this solution can be obtained much faster than with variational methods.

3. The fact that we don't have to evaluate complicated integrals during the calculations is not an advantage per se, because, as we saw, the method is essentially based on the approximation of an integral; if we need high precision then this integral has to be evaluated

very precisely and this may turn out to be more difficult than the evaluation of the integrals needed with other methods.

4. It is true, however, that round-off errors resulting from the subtraction of large numbers of similar magnitude will not arise during the calculation of this integral. Such round-off errors, however, may occur during the calculation of $\hat{H}\varphi_i(\mathbf{x})$. (If we used numerical differentiation for the calculation of $\hat{H}\varphi_i(\mathbf{x})$, such round off errors would surely occur.)

5. The method is only partly sensitive to the complexity of the functions forming the function space where we search for the solution of the SE. By ‘partly’ we mean that if these functions are simple functions of the parameters but complicated functions of the variables then the statement holds. If, however they depend on the parameters (also) in a complicated way, then the determination of these parameters may be more difficult.

6. Perhaps the biggest problem — a problem for all methods that don’t use a basis set built from Slater determinants — is that for larger systems the antisymmetry of the wave function can only be ensured by the use a very many terms.

7. It follows from item 3 in this list that the method is not economical if the results have to be very ‘precise’. If, however we only require approximate results then we can get the result with less work (computation time).

8. Finally, in view of the approximate nature of the method, it is a non-negligible advantage that without any extra work it provides an error estimate (by estimating the norm of $\hat{H}\varphi(\mathbf{x}) - E\varphi(\mathbf{x})$) and other error related quantities are also easy to estimate since the estimates of the matrix elements of \hat{H}^2 are available.

The version of the method used in this work has the following additional characteristics:

1. As any Monte Carlo method, it is less precise than methods designed for the exact evaluation of integrals.

2. The simple nature of the quadrature used ensures that by increasing the number of quadrature points the previous work is not lost.

3. The large number of quadrature points make especially important our statement in Sec. 3.3 regarding round-off errors of the matrices \mathbf{S} , \mathbf{H} , \mathbf{G} .

Table 3. Results of the calculations on H_2^+ : the expectation value and variance of the energy ($E_{\text{exp}}=-1.10263$)

K	$\rho(\mathbf{x})$	$nb = 4$	$nb = 6$	$nb = 8$
3.5	(3.6)	-1.1011 0.0008	-1.0916 0.0006	-1.1003 0.0003
3.5	(3.7)	-1.0927 0.0004	-1.0943 0.0003	-1.0975 0.0003
3.5	(3.8)	-1.1016 0.0007	-1.0947 0.0005	-1.0951 0.0001
4.5	(3.6)	-1.0981 0.0007	-1.0954 0.0004	-1.0992 0.0004
4.5	(3.7)	-1.0928 0.0004	-1.0913 0.0003	-1.0999 0.00007
4.5	(3.8)	-1.0915 0.0006	-1.0935 0.0006	-1.0972 0.00002

Table 4. Results of the calculations on H_2^+ : the expectation value and variance of the local energy

K	$\rho(\mathbf{x})$	$nb = 4$	$nb = 6$	$nb = 8$
3.5	(3.6)	-1.1066 0.0003	-1.0916 0.004	-1.0986 0.0004
3.5	(3.7)	-1.0878 0.0004	-1.0796 0.0014	-1.1072 0.0011
3.5	(3.8)	-1.0512 0.0019	-1.0913 0.0010	-1.1015 0.0015
4.5	(3.6)	-1.1202 0.0005	-1.0677 0.0010	-1.1044 0.0017
4.5	(3.7)	-1.0857 0.0003	-1.0486 0.0054	-1.1513 0.0016
4.5	(3.8)	-1.0912 0.0004	-1.0851 0.0019	-1.1305 0.0482

Note: Tables 3 and 4 describe the same calculations. The random number seed was 0.1 (so the quadrature points for a given K and $\rho(\mathbf{x})$ were the same). The number of quadrature points was 1900.

Table 5. The results of the calculations on He: the expectation value and variance of the energy, the expectation value and variance of the local energy for nb=6 ($E_{\text{exp}}=-2.90372$).

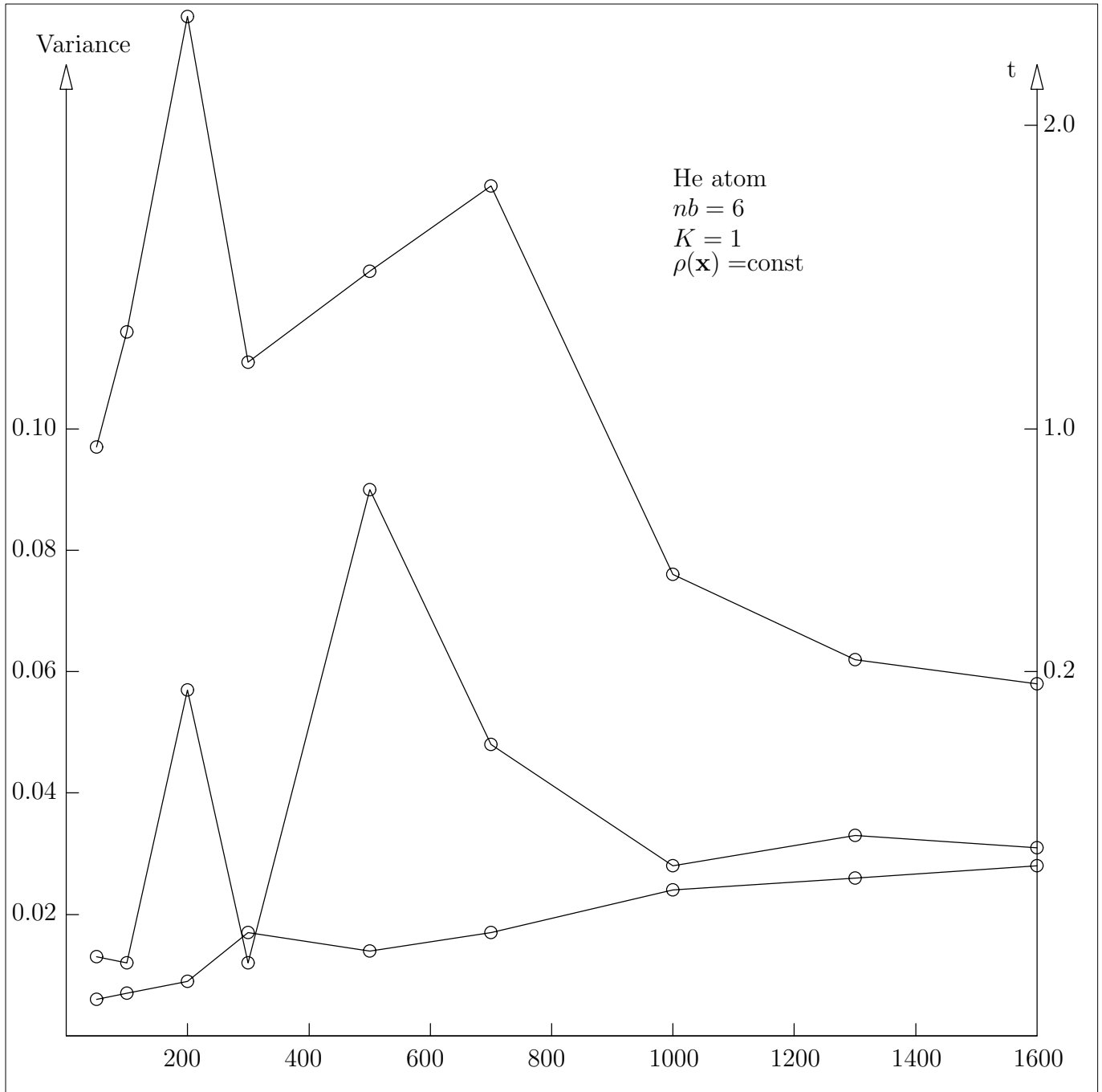
K	$\rho(\mathbf{x}): (3.6)$		$\rho(\mathbf{x}): (3.8)$	
	Energy	Local energy	Energy	Local energy
0.5	-2.6537	-2.6990	-2.8739	-2.9657
	0.0273	0.0244	0.0981	0.1874
1.0	-2.7454	-2.8347	-2.9066	-3.0324
	0.0227	0.0266	0.0981	0.5931
2.5	-2.7248	-3.0801	-2.8894	-3.0214
	0.0143	1.988	0.1074	0.4202
3.5	-2.8734	-3.3376	-2.8884	-3.0278
	0.0088	9.217	0.1071	0.4603
4.5	-2.9972	-3.4446	-2.8882	-3.0299
	0.0041		0.1070	0.4829

Note: For all calculations the random number seed was 0.1 and 1900 quadrature points were used.

Table 6: Comparison of the energies calculated on different quadrature point sets.

K	Seed=0.1	Seed=0.2	Seed=0.4
Energy 3.5	3.5	-1.1001	-1.1008
		0.0008	0.0007
Local energy 3.5	3.5	-1.1066	-1.0911
		0.0003	0.0005
Energy 4.5	4.5	-1.0981	-1.0931
		0.0007	0.0005
Local energy 4.5	4.5	-1.1202	-1.0600
		0.0005	0.0014

Note: The calculations were on H_2^+ , with $nb = 6$, using the weight function (3.6) on 1900 points.



The variance on the quadrature points (bottom curve), variance on the control points (middle curve) and the Student t value (top curve) for calculations on the He atom with $nb = 6$, $K = 1$ and $\rho(\mathbf{x}) = \text{const}$.

Appendix 1

Proof of the existence of a non-trivial solution of the homogeneous system of linear equations obtained from the Frost equation.

Eq. (2.27) is written as

$$\mathbf{A}\mathbf{c} = \mathbf{0} \quad (\text{A1.1})$$

where, according to (2.26)

$$\mathbf{A} = \mathbf{G} - 2M(E)\mathbf{H} + (2M(E)^2 - M(E^2))\mathbf{S} \quad (\text{A1.2})$$

A necessary and sufficient condition for the existence of a nontrivial solution for a homogeneous system of linear equations is that the determinant of the matrix of the system of equations be zero²⁷. The determinant of a matrix is nonzero if a nonzero vector \mathbf{v} can be found such that the expression

$$\mathbf{v}'\mathbf{A}\mathbf{v} \quad (\text{A1.3})$$

is zero (prime denotes the transpose). Let us select \mathbf{v} as

$$\mathbf{v} = \mathbf{c} \quad (\text{A1.4})$$

Then (A1.3) will become

$$\mathbf{c}'\mathbf{A}\mathbf{c} = \mathbf{c}'\mathbf{G}\mathbf{c} - 2M(E)\mathbf{c}'\mathbf{H}\mathbf{c} + (2M(E)^2 - 2M(E))\mathbf{c}'\mathbf{S}\mathbf{c} \quad (\text{A1.5})$$

Substituting (2.11) and (2.12) for $M(E)$ and $M(E^2)$ we get

$$\mathbf{c}'\mathbf{A}\mathbf{c} = \mathbf{c}'\mathbf{G}\mathbf{c} - 2(h/s)\mathbf{c}'\mathbf{H}\mathbf{c} + (2(h/s)^2 - (g/s))\mathbf{c}'\mathbf{S}\mathbf{c} \quad (\text{A1.6})$$

From the definitions of s , h , and g ((2.8)-(2.10)) it follows that

$$s = \sum_{i,j=1}^{nb} c_i S_{ij} c_j = \mathbf{c}'\mathbf{S}\mathbf{c} \quad (\text{A1.7})$$

$$h/s = \sum_{i,j=1}^{nb} c_i H_{ij} c_j / s = \mathbf{c}'\mathbf{H}\mathbf{c} / \mathbf{c}'\mathbf{S}\mathbf{c} \quad (\text{A1.8})$$

$$g/s = \sum_{i,j=1}^{nb} c_i G_{ij} c_j / s = \mathbf{c}'\mathbf{G}\mathbf{c} / \mathbf{c}'\mathbf{S}\mathbf{c} \quad (\text{A1.9})$$

Substituting (A1.7)-(A1.9) into (A1.6) we get

$$\mathbf{c}'\mathbf{A}\mathbf{c} = \mathbf{c}'\mathbf{G}\mathbf{c} - \frac{2\mathbf{c}'\mathbf{H}\mathbf{c}}{\mathbf{c}'\mathbf{S}\mathbf{c}}\mathbf{c}'\mathbf{H}\mathbf{c} + 2\frac{(\mathbf{c}'\mathbf{H}\mathbf{c})^2}{(\mathbf{c}'\mathbf{S}\mathbf{c})^2}\mathbf{c}'\mathbf{S}\mathbf{c} - \frac{\mathbf{c}'\mathbf{G}\mathbf{c}}{\mathbf{c}'\mathbf{S}\mathbf{c}}\mathbf{c}'\mathbf{S}\mathbf{c} \quad (\text{A1.10})$$

Simplifying (A1.10) we get

$$\mathbf{c}'\mathbf{A}\mathbf{c} = \frac{0}{\mathbf{c}'\mathbf{S}\mathbf{c}} \quad (\text{A1.11})$$

The denominator of (A1.11) can only be zero if the determinant of \mathbf{S} is zero, and this can only occur if the $\varphi_i(\mathbf{x})$ s are linearly dependent, but this we excluded at the outset.

Appendix 2

Proof of the fundamental equation of the Monte Carlo method.

We have to prove the equality (2.39):

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N f(\mathbf{x}_i) = \frac{1}{V} \int_B f(\mathbf{x}) dv \quad (\text{A2.1})$$

where $f(\mathbf{x})$ is a continuous bounded function in a domain B of volume V and \mathbf{x}_i is a point in B . The lim here denotes stochastic convergence. For the proof we define the probability distribution

$$F(X) = P(f(\mathbf{x}) \leq X) \quad (\text{A2.2})$$

that is obviously continuous whenever $f(\mathbf{x})$ is continuous, therefore it has a probability density function, denoted here with $r(X)$.

From the law of large numbers³¹, that states that the mean of a large number of probability variables stochastically to the expectation value of the probability variables, it follows that the left hand side of (A2.1) is the expectation value of the distribution $F(X)$. We thus have to prove that this is true for the right hand side as well.

Expressing the expectation value with the probability density the statement to be proven becomes

$$\int_u^v r(X) X dX = \frac{1}{V} \int_B f(\mathbf{x}) dv \quad (\text{A2.3})$$

where u and v are numbers such that

$$u \leq f(\mathbf{x}) \quad (\mathbf{x} \in B) \quad (\text{A2.4})$$

(Such u , v exist because $f(\mathbf{x})$ is bounded in B .) The integrals on the two sides of (A2.3) can be written as the limits of sums (over points chosen at uniform intervals):

$$\lim_{\Delta_1 \rightarrow 0} \sum_{j=1}^{N_1} \Delta_1 r(X_j) X_j = \frac{1}{V} \lim_{\Delta_2 \rightarrow 0} \sum_{i=1}^{N_2} \Delta_2 f(\mathbf{x}_i) \quad (\text{A2.5})$$

where

$$\Delta_1 = (v - u)/N_1 \quad (\text{A2.6})$$

$$\Delta_2 = V/N_2 \quad (\text{A2.7})$$

If m_j of the $f(\mathbf{x}_i)$ s fall into the j -th subinterval of $[u, v]$ then

$$\frac{1}{V} \lim_{\Delta_2 \rightarrow 0} \sum_{i=1}^{N_2} \Delta_2 f(\mathbf{x}_i) = \frac{1}{V} \lim_{\Delta_2 \rightarrow 0} \sum_{j=1}^{N_1} m_j f(\mathbf{x}_i) \Delta_2 = \lim_{\Delta_2 \rightarrow 0} \sum_{j=1}^{N_1} \frac{m_j}{N_2} f(\mathbf{x}_i) \quad (\text{A2.8})$$

Here we used a common \mathbf{x}_j for each subinterval j containing m_j points. The error introduced by this transformation goes to zero in the limit since $f(\mathbf{x})$ is continuous.

Also, by the definition of $r(X)$ we have

$$r(X_j)\Delta_1 = P(X_j \leq f(\mathbf{x}_j) \leq X_j + \Delta_1) \quad (\text{A2.9})$$

By Bernoulli's law or large numbers, according to which the limit of the frequency of occurrence is the expectation value, the probability on the right hand side of (A2.9) is

$$P(X_j \leq f(\mathbf{x}_j) \leq X_j + \Delta_1) = \frac{m_j}{N_2} \quad (\text{A2.10})$$

thus, comparing (A2.10) and (A2.9) we get

$$r(X_j)\Delta_1 = \frac{m_j}{N_2} \quad (\text{A2.11})$$

Substituting this into the right hand side of (A2.8) and using (A2.7) we get the equality (A2.3) we were after and this in turn proves (A2.1)

The most important result of this proof is the introduction of the probability distribution $F(X)$ because we will use it for error estimate.

Appendix 3

The linear independence of the basis functions used for He and H_2^+ .

Linear independence of nb functions means that that none of them can be expressed as the linear combination of the others. This is equivalent to saying that

$$\sum_{i=1}^{nb} c_i \psi_i(\mathbf{x}) = 0 \quad (\text{A3.1})$$

can only hold for all \mathbf{x} when $c_1 = c_2 = \dots = c_{nb}$. For the He atom (A3.1) is of the form

$$\sum_{i=1}^{nb} c_i r_1^{ni1} r_2^{ni2} r_{12}^{ni3} e^{-\xi(r_1+r_2)} = 0 \quad (\text{A3.2})$$

Since for finite r_1 and r_2 $e^{-\xi(r_1+r_2)}$ can not be zero, we can divide (A3.2) with it, giving the equivalent condition

$$\sum_{i=1}^{nb} c_i r_1^{ni1} r_2^{ni2} r_{12}^{ni3} = 0 \quad (\text{A3.2})$$

If we disregard the fact that r_1 , r_2 and r_{12} are not independent of each other and can not be negative then (A3.3) can be considered a polynomial in three variables. In general, a polynomial of n variables can only be identically zero if all its coefficients are zero. To see this, we can consider the polynomial of n variables as a polynomial of the n -th variable whose coefficients themselves are polynomials of $n - 1$ variables. The fundamental theorem

of algebra²⁷ states, however that a polynomial of order k in one variable can only have more than k zeros if all the coefficients are zeros. Thus a polynomial of N variables can only be identically zero if there is a polynomial of $n - 1$ variables that also can be identically zero. Continuing the argument we reach the conclusion that a polynomial of N variables can only be identically zero if there is a polynomial of one variable that also can be identically zero. This, however, contradicts the fundamental theorem of algebra referred to above.

In our case the variables r_1 , r_2 and r_{12} have to satisfy the following constraints:

$$\begin{aligned} r_1 &> 0 \\ r_2 &> 0 \\ r_{12} &> 0 \\ r_1 + r_2 &> r_{12} \end{aligned} \tag{A3.4}$$

The argument above is still applicable because even considering the constraints they can take infinite many different values (thus it is not possible that the constraints would limit r_1 , r_2 and r_{12} to exactly the roots of the polynomial).

For H_2^+ the condition

$$\sum_{i=1}^{nb} c_i r_A^{n_i} r_B^{m_i} e^{-\xi(r_A+r_B)} = 0 \tag{A3.5}$$

is to be considered where the r_A and r_B are constrained by

$$\begin{aligned} r_A &> 0 \\ r_B &> 0 \\ (r_A - r_B)^2 &< R^2 \end{aligned} \tag{A3.6}$$

where R is the distance between the two H atoms.

The argument used for the He atom can be applied for this case as well since the the conditions (A3.6) also allow infinite many r_A and r_B values and we can again divide by $e^{-\xi(r_A+r_B)}$.

Appendix 4

Historical overview.

While it is possible to follow the development of the method through the references in the text, a small historical overview still appears to be in order.

A4.1. *Local energy method in the strict sense.*

The first description of the method is in the 1942 paper of A.A. Frost⁷. Calculations illustrating the method were done on H_2^+ and the energy was obtained within 1% using only 9-15 points. The calculations described in this work used the basis functions from this paper.

In 1960, A.A. Frost, R.E. Kellog and E.C. Curtiss¹⁰ published more detailed calculations on H_2^+ , using 9 basis functions and 40 points, giving an error of 0.1% in the energy. Both calculations selected essentially equidistant points.

In 1961, A.A. Frost, R.E. Kellog, B.M. Gimarc and J.D. Scargle¹¹ used Gaussian quadrature for calculations on H_2^+ , H_2 and He. The results of their most detailed calculations are shown in Table A1. Our calculations on He used their basis functions.

B.M. Gimarc and A.A. Frost performed more detailed (including excited states) calculations in 1963^{12,13} on He and Li. Their results are summarized in Table A2.

Table A1. Results of calculations on H_2^+ , H_2 and He.

	H_2^+	H_2	He
Energy (calculated)	-1.102629	-1.1728	-2.89705
Variance	0.00001	0.0056	0.0179
Energy (experimental)	-1.10263	-1.1744	-2.90372
Size of the basis	9	14	12
Number of quadrature points	2	288	60

Table A2. Results of calculations on He and Li.

	He(ground)	He(1st excited)	Li
Energy (calculated)	-2.2025	-2.1753	-7.4653
Variance	0.00016	0.000001	0.0506
Energy (experimental)	-2.90372	-2.1752	-7.4781
Size of the basis	30	30	51
Number of quadrature points	60	126	500

In their paper they studied the variation of the nonlinear parameters as well.

Based on their calculations on the Li atom the authors were sceptical about the possibility of extending the method to larger systems (in that version). In our opinion, the following possibilities exist for development:

1. Replacement of the Gaussian quadrature with a more general quadrature that is less dependent on the form of the function to be integrated. (The random quadrature used in this study has this property.)

2. Choose the $\varphi(\mathbf{x})$ in (1.1) in such a way that it is a seperable function of the internal coordinates. This would simplify the computation of the matrices \mathbf{S} , \mathbf{H} and \mathbf{G} .

3. Partition the integrand of (1.1) and evaluate the different terms on different sets of points (this would mean the modification of the idea of local energy).

In 1974, D.K. Harris and A.A. Frost¹⁷ repeated their calculations on H₂ with better precision. Their results are summarized in Table A3.

Table A3. Results of the detailed calculations on H₂.

Energy (calculated)	Variance	Basis size	Number of points
-1.1772	0.0049	12	144
-1.1725	0.0065	12	288
-1.17435	0.0032	12	960
-1.17442	0.0019	22	288
-1.17454	0.0020	22	440

Experimental energy= -1.17444

D.K. Harris and R.K. Roubal developed a procedure in 1968¹⁸ to optimize the nonlinear parameters as well. They performed calculations on HeH⁺⁺ in both ground and excited states. Their results are shown in Table A4.

Table A4. Results of calculations on HeH⁺⁺.

Energy (calculated)	-2.51204	-1.34502	-0.787064	-0.535764
Variance	0.00922	0.00362	0.000016	0.000341
Energy (experimental)	-2.51219	-1.34518	-0.78709	-0.571255
Size of the basis	10	10	10	10
Number of points	25	25	25	25

In 1968 the method was extended to perturbed systems by D.W. Davies⁴

A4.2. *Methods related or nearly equivalent to the local energy method.*

A4.2.1. *The method of Weinstein*³³.

D.H. Weinstein proved in 1934 that if $\varphi(\mathbf{x})$ is an approximate (normalized) solution of the SE belonging to an operator \hat{H} and $E = \langle \varphi(\mathbf{x}) | \hat{H} | \varphi(\mathbf{x}) \rangle$ then the maximum deviation of E from the eigenvalue of \hat{H} nearest to E is $\langle \varphi(\mathbf{x}) | (\hat{H} - E)^2 | \varphi(\mathbf{x}) \rangle$. Therefore, the smaller this deviation, the better approximation $\varphi(\mathbf{x})$ is to the eigenfunction of \hat{H} . Thus the combination of this estimate with a method that minimizes this deviation provides a method for finding the eigenfunctions of \hat{H} . It was L. Pauling⁷ who recognized that the LEM is just such method.

A4.2.2. *The method of T.A. Rourke and E.T. Steward*^{22,13}.

These authors introduced a variant of the LEM where the measure F1 of (1.1) is replaced by $F1/|E|$. This was done to decrease the fluctuation in the variance of the energy. (The

variance of the energy not only depends on the 'goodness' of the basis but on the error of the calculated energy and the authors intended to eliminate this latter dependency.)

They used a random quadrature, but restricted the points inside the domain of the orbitals of the atoms. (These domains were determined based on the paper of Waker and Cramer³².) This also made possible to ensure that the orbitals belonging to different shells will be orthogonal. As a result, they could ensure the antisymmetry of the wave function with much fewer terms. They also developed the statistical error estimate described in Sec. 2.3.2.1.

They repeated the Hartree-Fock calculations on systems containing 1-4 electrons, based on 25 points. The error in the energy was 1-10%, naturally larger than the errors of the corresponding Hartree-Fock energies (i.e., the 25 points are too few).

The coefficients were obtained with the gradient method and the gradient was obtained numerically. The differentiation in the kinetic energy operator was obtained by numerical differentiation to be able to easily use the subroutine calculating $\hat{H}\varphi_i(\mathbf{x})$ with different basis functions.

A4.2.3. *The method of H. Conroy*³.

H. Conroy minimized in 1964 the integral in (1.1) with a Monte Carlo method (i.e., used a random quadrature), but used the simplification of the FE described in Sec. 2.2.1.2. The value of the energy, kept fixed during the calculation, was obtained by extrapolating to zero. His results on H_2^+ and He are shown in Table A5.

Table A5. Results of calculations on H_2^+ and He.

Energy (calculated)	-1.1025	-2.89977
Variance	0.01	0.08
Energy (experimental)	-1.10263	-2.90372
Size of the basis	10	16
Number of points	300	600

A4.2.4. *The methods of J. Goodishman*^{14–16}.

J. Goodishman published a variational method in 1964¹⁵ where he calculated the elements of the \mathbf{H} matrix with a numerical quadrature and in a paper published with D. Secrest¹⁶ in 1964 they minimized the integral of (1.1) as approximated by a numerical quadrature and show that minimization of the integral in (1.1) is equivalent to minimizing the mean of the local energies, weighted by $\varphi^2(\mathbf{x})$. Error estimates are given based on the matrix elements of H^2 , a refinement of the estimate of Weinstein discussed in A4.2.1.

A4.2.5. *Relation to the method of energy moments.*

J.B. Delos and S.M. Blinder⁵ introduced a method in 1967 that seeks the solution of the SE using $\langle \varphi(\mathbf{x}) | \hat{H}^k | \varphi(\mathbf{x}) \rangle$ or $\hat{H}^k | \varphi(\mathbf{x})$. They showed that the LEM is nearly equivalent to a special case of the second version.

Also in 1967, H.J. Silverstone, Hoon-Lung Yi and L.R. Somorjai²⁴ introduced a variant of the method of moments that is based on the evaluation in the points of the configuration space of the integrands in the integrals $\langle \varphi(\mathbf{x}) | \hat{H}^k | \varphi(\mathbf{x}) \rangle / \langle \varphi(\mathbf{x}) | \varphi(\mathbf{x}) \rangle$. They called their method the method of local moments.

A4.2.6. *Relation to the integral Hellman-Feynmann theorem.*

L.P. Lowe and A. Mazziotti³⁴ showed in 1968 that the integral Hellman-Feynmann theorem can be considered a LEM where the weight function $\varphi^2(\mathbf{x})$ is replaced by the transition function $\varphi_1(\mathbf{x})\varphi_2(\mathbf{x}) / \langle \varphi_1(\mathbf{x}) | \varphi_2(\mathbf{x}) \rangle$ (the subscripts 1 and 2 refer to the two states). Similar relations can be found with the methods based on the partitioning of the energy operator.

A4.2.7. *Relation to the method of moments.*

The method of moments^{28,29} that approximates the solution of the SE with a function $\varphi(\mathbf{x})$ such that minimizes

$$\sum_{i=1}^n \langle w_i(\mathbf{x}) | (\hat{H} - E)^2 | \varphi(\mathbf{x}) \rangle \quad (\text{A4.1})$$

for a certain choice of functions $w_i(\mathbf{x})$ ($i = 1, 2, \dots, n$). Choosing the $w_i(\mathbf{x})$ s as Dirac deltas, we recover the LEM³⁰.

A4.2.7.1. *A development possibility³⁰.*

It is possible to develop a version of the LEM where, using the terminology of the method of moments, we use as weight functions Gaussian functions approximating Dirac deltas, instead of the Dirac deltas. This would replace discrete points with a small region around them. This would eliminate the drawback of the LEM related to it being restricted to single points without overly complicating the integration problems.

A4.2.8. *Relation to the method of T.A. Hoffmann.*

The method of T.A. Hoffmann, published in 1965, can be considered a generalized LEM. It calculates the elements of the density matrix of the system at various points in the configuration space as a function of a certain sets of parameters in such a way that the resulting matrix satisfies best the properties of the exact density matrix. This method thus uses a functional much different from $F1$.

Summary of the results.

The present work, besides presenting the variants of the LEM that can be found in the literature,

1. developed an estimate for the precision of the approximation (Sec. 2.3.2) that can be used to determine the number of points needed;
2. found a simple way to treat the weight function of (1.1) (Sec. 2.3.1.);
3. introduced a simplified FE (2.25) determining the optimal values of the parameters (Sec. 2.2.1.6) and showed a new method for its solution (Sec. 2.2.1.4);
4. studied the effect of changing the domain of integration (Sec. 3.4.1.);
5. proved two theorems that are required for the application of the method, but the studies in the literature neglected to do so (Appendices 1 and 3);
6. gave a proof for the equation (2.39) on which the random quadratures are based (Appendix 2).

* * *

Finally, I would like to thank chairman dr. Béla Csákváry, for making it possible for me to prepare this work in the department, to dr. Ferenc Török for the effective and cordial support in writing this thesis, to Miklos Preisich, director of MAVEMI for his support enabling the preparation of this work and for his permission to use the Gier computer at MAVEMI for the calculations and to my immediate supervisor at MAVEMI, dr. Tamás Szondy who brought the LEM to my attention and helped during the preparation of this work.

REFERENCES

1. Bartlett, J.W., *Phys. Rev.*, **98**, 1067 (1955).
2. Boys, S.F., *Proc. Roy. Soc. (London)*, **A309**, (1965).
3. Conroy, H.J., *Chem. Phys.*, **41**, 1327 (1964).
4. Davies, D.W., *Theoret. Chim. Acta*, **9**, 330 (1968).
5. Delos, J.B., Blinder, S.M., *J. Chem. Phys.*, **47**, 2784 (1967).
7. Frost, A.A., *J. Chem. Phys.*, **10**, 240 (1942)
8. Frost, A.A., *Theoret. Chim. Acta*, **1**, 36 (1962).
9. Frost, A.A., *J. Chem. Phys.*, **41**, 478 (1964).
10. Frost, A.A., Kellog, R.E., Curtis, E.C., *Rev. Mod. Phys.*, **32**, 313 (1960).
11. Frost, A.A., Kellog, R.E., Gimarc, B.M., Scargle, J.D., *J. Chem. Phys.*, **35**, 827 (1961).
12. Gimarc, B.M., Frost, A.A., *Theoret. Chim. Acta*, **1**, 87 (1962).
13. Gimarc, B.M., Frost, A.A., *J. Chem. Phys.*, **39**, 1968 (1963).
14. Goodishman, J., *J. Chem. Phys.*, **47**, 5247 (1967).
15. Goodishman, J., *J. Chem. Phys.*, **41**, 3889 (1964).
16. Goodishman, J., Secrest, D., *J. Chem. Phys.*, **45**, 1515 (1966).
17. Harris, D.K., Frost, A.A., *J. Chem. Phys.*, **40**, 204 (1964).
18. Harris, D.K., Roubal, R.K., *Theoret. Chim. Acta*, **9**, 303 (1968).
19. Kreko, B., *Mátrixszámítás* (book) Közgazdasági es Jogi Könyvkiadó, Budapest (1964).
20. Lyon, W.D., Matcha, R.L., Sanders, W.A., Meath, W.J., Hirschfelder, W.D., *J. Chem. Phys.*, **43**, 1095 (1965).
21. Rosenbrock, H.H., Storey, C., *Computational Techniques for Chemical Engineers* (book).
22. Rourke, T.A., Stewart, E.T., *Can. J. Phys.*, **45**, 2735 (1967).
23. Rourke, T.A., Stewart, E.T., *Can. J. Phys.*, **46**, 1603 (1968).
24. Silverstove, H.J., Moon-Lung Yin, Somorjai, R.L., *J. Chem. Phys.*, **47**, 4824 (1967).
25. Srejger, J.A., *Monte Carlo módszerek* (book) Műszaki Könyvkiadó. Budapest (1965).
26. Stanton, R.E., Taylor, R.L., *J. Chem. Phys.*, **45**, 565 (1966).
27. Szele, T., *Bevezetés az algebrába* (book) Tankönyvkiadó, Budapest (1963).
28. Szondy, E., Szondy, T., *Acta Physica, Acad. Sci. Hung.*, **20**, 253 (1966).
29. Szondy, T., *Naturforsch.*, **17A**, 552 (1962).
30. Szondy, T., *personal communication*.
31. Vince, I., *Matematikai Statisztika Ipari alkalmazása*, Műszaki Könyvkiadó. Budapest (1968).
32. Weber, J.T., Cramer, D.T., *J. Chem. Phys.*, **42**, 4116 (1965).
33. Weinstein, J.T., *Proc. Nat. Acad. Sci.*, **20**, 529 (1934).
34. Lowe, J.P., Mazziotti, A., *J. Chem. Phys.*, **48**, 877 (1968).
35. Hoffman, T.A., *Phil. Trans of the Soc. of London*, **257**, 309 (1965)

CONTENTS

INTRODUCTION	2
1. The foundations of the LEM.	2
2. Description of the method.....	3
2.1. The choice of the measure F	3
2.2. The choice of the function space.....	5
2.2.1 Searching for the c_k s minimizing $F1$	5
2.2.1.1. The role of the weight factor $\varphi^2(\mathbf{x})$ in $F2$	7
2.2.1.2. Simplifications of the Frost equation.	7
2.2.1.3. Solution of the Frost Equation with iteration.	8
2.2.1.4. An other possible iteration.	8
2.2.1.5. Solution of the Frost equation with relaxation.	8
2.2.1.6. Determination of the minimum of F1 with a gradient method.	8
2.2.2. Calculation of $\hat{H}\varphi_i(\mathbf{x})$	9
2.2.2.1. The form of the operator H in internal coordinates.	10
2.2.2.2. Numerical calculation of $\hat{H}\varphi_i(\mathbf{x})$	10
2.2.3. Optimizing the non-linear parameters.	11
2.2.3.1. Optimizing the nonlinear parameters using the linear parameters.	11
2.2.3.3. A special form of $\hat{H}\varphi_i(\mathbf{x})$ and $\varphi_i(\mathbf{x})$	11
2.3. The choice of the quadrature points.....	11
2.3.1. Random quadrature (Monte Carlo method).....	12
2.3.2. Estimation of the error of a random quadrature.....	13
2.3.2.1. The estimates of T.A. Rourke and E.T. Stewart.....	15
3. Calculations.....	15
3.1. The choice of $\varphi_i(\mathbf{x})$	15
3.2 The expressions for $\hat{H}\varphi_i(\mathbf{x})$	17
3.3. Solution of the Frost equation, round-off errors.....	18
3.4 The choice of the quadrature points.....	18
3.4.1. The domain of integration.....	20
3.5. The computer program.....	20
3.5.1. Major subroutines	21
3.5.2. The running of the program.....	21
3.5.3. Options to change the calculation	21
3.5.4. Printing of the partial results.....	22
3.5.6. Debugging the program.....	22
3.5.7. The dependence of the computation time on the size of the problem.....	23
4. Results.....	24
4.1. Energies.....	24
4.2. The dependence of the result on the choice of region.....	24
4.3. The stability of the results.....	24
4.4. The criterion for stopping the calculation.....	24
5. Evaluation of the method.....	24
Appendix 1: Proof for the existence of a non-trivial solution of the homogeneous sytem of linear equations obtained from the Frost equation.....	29
Appendix 2: Proof of the fundamental equation of the Monte Carlo method.....	30
Appendix 3: The linear independece of the basis functions used for He and H_2^+	31

Appendix 4: Historical overview	32
A4.1. Local energy method in the strict sense.....	32
A4.2. Methods related or nearly equivalent to the local energy method	34
A4.2.1. The method of Weinstein	34
A4.2.2. The method of T.A. Rourke and E.T. Steward	34
A4.2.3. The method of H. Conroy	35
A4.2.4. The methods of J. Goodishman.....	35
A4.2.5. Relation to the method of energy moments.....	35
A4.2.6. Relation to the integral Hellman-Feynmann theorem	36
A4.2.7. Relation to the method of moments.....	36
A4.2.7.1. A development possibility.....	36
A4.2.8. Relation to the method of T.A. Hoffmann.....	36
Summary of the results.....	37
REFERENCES	38
CONTENTS	39