JOURNAL OF COMPUTATIONAL PHYSICS 36, 198–208 (1980)

Theory and Method for Accelerating the Convergence of Self-Consistent Electronic Structure Calculations*

LUIZ G. FERREIRA

Instituto de Fisica, Universidade de São Paulo, Caixa Postal 20516, São Paulo, Brazil

Received January 10, 1979; revised April 11, 1979

In this paper, we present a method which accelerates the iterative process of self-consistent calculations of the electronic structure of atoms, molecules, and crystals. The theory is based on the behavior of potential parameters during the iterative process. A criterion of convergence is established. The convergence of the iterative process is fastest when, at the beginning of each iteration, the potential is a linear combination of the potentials of previous iterations. As many previous iterations should be combined as there are parameters describing the potential; the theory and method are applied in the test case of the molecules N_2 and CO with very good results.

1. STATEMENT OF THE PROBLEM

The calculation of the electronic structure of atoms, molecules, and crystals is an iterative procedure. In each iteration one uses an initial potential

$$v^{(i)}(\mathbf{r}).$$

The Schrödinger equation is solved for each occupied one-electron state, an electronic density is constructed from the wave functions, and the Poisson equation is solved to obtain a final potential

$$v^{(f)}(\mathbf{r}).$$

For the next iteration one starts from an initial potential which is a linear combination of the initial and final potentials of the previous iterations. Since the pioneer calculations of Hartree [1], has become common practice to generate the initial potential of iteration $i + 1$ according to

$$v_{i+1}^{(i)} = (1 - \alpha)\, v_i^{(i)} + \alpha v_i^{(f)}, \tag{1}$$

where α is a number greater than zero and less than one. The choice of α depends on the problem and on the experience of the researcher. For atoms and simple molecular

---

* Work supported by the Fundacão de Amparo à Pesquisa do Estado de São Paulo and the Instituto de Atividades Espaciais, São José dos Campos, Brazil.

0021-9991/80/080198-11$02.00/0

systems, values as large as 0.5 can be used [1]. For systems with many atoms, α has to be small, otherwise the iterations will not converge.

To speed up the iterative process, Pratt proposed a scheme to determine α, which has a simple graphical interpretation [2]. For each point $\mathbf{r}$, we plot the values $v^{(i)}$ and $v^{(f)}$ for iterations i and $i - 1$, as in Fig. 1. Through the points we pass a straight line. The crossing of this line with the dashed line of 45° inclination determines the initial potential for iteration $i + 1$. The Pratt scheme can be simply understood since the 45° line is where the final solution should be since $v^{(i)}$ and $v^{(f)}$ are equal.

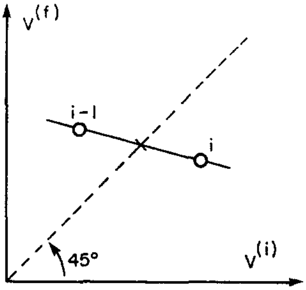

FIG. 1. The Pratt scheme.

Ultimately, the Pratt scheme is a linear and local theory of the iterative process since

$$v^{(f)}(\mathbf{r}) = v^{(f)}[v^{(i)}]$$

is approximated as

$$v^{(f)}(\mathbf{r}) = \int T(\mathbf{r}, \mathbf{r}')\, v^{(i)}(\mathbf{r}')\, d\mathbf{r}' + c(\mathbf{r}). \tag{2}$$

It is a local theory, in the sense that it assumes that $v^{(f)}(\mathbf{r})$ depends only on the value of $v^{(i)}$ at $\mathbf{r}$. Thus

$$T(\mathbf{r}, \mathbf{r}') = t(\mathbf{r})\, \delta(\mathbf{r} - \mathbf{r}') \tag{3}$$

or

$$v^{(f)}(\mathbf{r}) = t(\mathbf{r})\, v^{(i)}(\mathbf{r}) + c(\mathbf{r}). \tag{4}$$

Equation (4) combined with Eq. (1) and forcing

$$v_{i+1}^{(f)}(\mathbf{r}) = v_{i+1}^{(i)}(\mathbf{r}) \tag{5}$$

yields

$$\alpha(\mathbf{r}) = 1/(1 - t(\mathbf{r})). \tag{6}$$

This is the Pratt coefficient of mixing, which is position dependent.

The Pratt scheme has been widely used for atoms [3, 4], but has been avoided for molecules and crystals, when the local character of the scheme becomes inadequate. To understand the inadequacy of the Pratt scheme for molecules, consider the example of CO. If one raises the initial potential near the C atom, then electrons move toward the O atom, thus increasing the final potential there. In effect the kernel $T(r, r')$ in Eq. (2) may have long-range components for molecules. This invalidates the local approximation of Eq. (3). Achieving self-consistency in molecules is largely a problem of determining the right amount of charge transfer from one atom to another. This charge transfer is not taken care of by the Pratt scheme.

During the iterative process, one deals with potentials that increase smoothly from the nuclei and, for that reason, the potentials can be described by few parameters. We shall assume that at the beginning and at the end of one iteration, the potential can be written as

$$v(\mathbf{r}) = f_0(\mathbf{r}) + \sum_N V^{(N)} f_N(\mathbf{r}), \tag{7}$$

where $V^{(N)}$ are parameters and the f's are functions that do not change during the iterative process. Thus, looking for self-consistency in the potential becomes equivalent to the simpler problem of looking for self-consistency in the parameters $V^{(N)}$. Thus it is possible to choose a few parameters $V^{(N)}$ whose self-consistency, that is, the equality at the beginning and at the end of the iteration

$$V^{(N,i)} = V^{(N,f)} \tag{8}$$

leads to the self-consistency of the potential

$$v^{(i)}(\mathbf{r}) = v^{(f)}(\mathbf{r}).$$

Further, we intend to present a scheme of iterations that leads to the fastest self-consistent solution. Finally, we shall study the stability of the iterative process and when it is converging or diverging.

2. PARAMETERS DEFINING THE POTENTIAL

During the iterative process, the potential varies while some of its parameters remain fixed; for example, the atomic numbers Z, the ionicity of the molecule, and the interatomic distances. In addition to the fixed parameters, we want to find the free parameters that determine the potential as uniquely as possible. For practical reasons, the new parameters must be few in number and yet be able to determine $v(r)$ almost completely.

In the rest of this work, we choose the *potentials seen by each nucleus* as the free parameters. Thus, there are as many parameters as there are different atoms. We can present little justification for this choice except that it works.

For an atomic calculation, there would be just one parameter for there is just a single nucleus. The potential at the nucleus is

$$V = 8\pi \int_0^\infty rn(r)\, dr,$$

where $n(r)$ is the electronic density. The parameter reflects how the charge is distributed about the nucleus. Furthermore, the product ZV is the electron–nucleus energy, which is the most important term in the total energy of the atom. Achieving self-consistency is equivalent to finding the energy extreme; therefore it is not surprising that the potential seen by the nucleus determines the iterative process to a large extent. For multiatomic systems the arguments in favor of our choice remain valid, but, of course, the ultimate fact is that our parameters work.

3. LINEAR THEORY OF THE ITERATIVE PROCESS IN THE CASE OF ONE PARAMETER

Consider an atom, a molecule, or a crystal with just one atomic species. In this case, just one free parameter defines the potential. Then, instead of Eq. (2), we write

$$V^{(f)} = TV^{(i)} + C. \tag{9}$$

Knowing the parameters $V_i^{(i)}$, $V_{i-1}^{(i)}$ at the beginning of the iterations i and $i-1$, and knowing the corresponding parameters $V_i^{(f)}$ and $V_{i-1}^{(f)}$ at the end, we use Eq. (9) to establish the value of T:

$$T = \frac{V_i^{(f)} - V_{i-1}^{(f)}}{V_i^{(i)} - V_{i-1}^{(i)}}. \tag{10}$$

Assuming that V determines the potential $v(\mathbf{r})$ and that the latter is linear in V, Eq. (1) implies

$$V_{i+1}^{(i)} = (1-\alpha)\, V_i^{(i)} + \alpha V_i^{(f)}. \tag{11}$$

Then we use Eqs. (9), (10), and (11) to determine the difference

$$\Delta_{i+1} = V_{i+1}^{(f)} - V_{i+1}^{(i)} \tag{12}$$

between the parameters at the beginning and at the end of iteration $i+1$

$$\Delta_{i+1} = (1 - \alpha/\alpha_P)\, \Delta_i, \tag{13}$$

where

$$\alpha_P = 1/(1-T) \tag{14}$$

is the Pratt value. One sees from Eq. (13) that convergence is faster when

$$\alpha = \alpha_P.$$

In this case one would reach self-consistency in just one iteration. Of course, in an actual case, the parameter V does not determine the potential uniquely and the linearity of Eq. (9) is not exactly true. Therefore an exact converged value would not occur at iteration $i + 1$.

If one uses a value of α other than α_p, we obtain, after n iterations

$$\Delta_{i+n} = (1 - \alpha/\alpha_p)^n \Delta_i . \tag{15}$$

Assuming α_p is positive, the difference between the beginning and the end decreases if

$$0 < \alpha < 2\alpha_p .$$

If α is chosen outside this interval, the iterative process diverges.

This simple theory was used in connection with the self-consistent calculation of the N_2 molecule by the cellular method [5]. The atoms of this molecule being equal, one deals with only one parameter, the potential seen by either nuclei. In Fig. 2, we plot the pairs

$$V_i^{(i)}, \; V_i^{(f)}$$

for several runs with different α and different starting points. For instance, the triangles are successive iterations with $\alpha = 0.70$ starting from a Thomas–Fermi type of potential, and the squares were obtained from iterations with $\alpha = 0.50$. Figure 2 shows clearly two facts:

1. The points fall along a definite curve with very small scatter, independently of the history of the iteration, the starting point, and the value of α used for the sequence of iterations. Thus, the existence of a function

$$V^{(f)} = V^{(f)}(V^{(i)}) \tag{16}$$

is confirmed.

2. The curve is almost a straight line, showing that the linear theory is adequate.

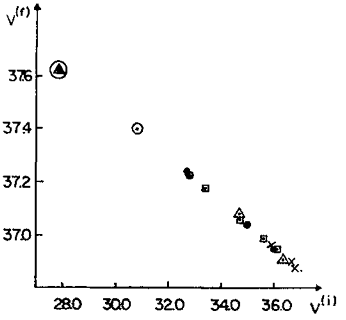

FIG. 2. Parameter at the end of iteration as function of parameter at beginning for the N_2 molecule and independent series of iterations.

4. A CONJECTURE ON THE STABILITY OF THE ITERATIVE PROCESS

In Figs. 3a and b, we consider two different types of crossing between the function

$$V^{(f)} = V^{(f)}(V^{(i)})$$

and the 45° dashed line. In Fig. 3a, the slope of the function at the crossing is less than one; in Fig. 3b, it is greater than one. In Fig. 3b, since

$$T > 1,$$

one has

$$\alpha_P < 0.$$

In order to reach the self-consistent solution, one needs a value of α in the interval

$$2\alpha_P < \alpha < 0,$$

that is, a negative value. We intend to show how the crossing in Fig. 3b can represent a point of instability.

Consider the molecule at the crossing of Fig. 3b and apply an external perturbing potential. Then define the parameter V so that it is zero at the crossing. The effect of the external perturbation is to shift the solution, or shift the value of the parameter from its zero value at the crossing. This shift can be calculated in the following way. Let $V^{(i)}$ be the shift of the potential at the beginning of the iteration. At the end of the iteration, the shift will be

$$TV^{(i)}$$

plus the parameter P, corresponding to the external perturbing potential

$$V^{(f)} = TV^{(i)} + P.$$

So, the new solution

$$V^{(f)} = V^{(i)} \equiv V$$

becomes

$$V = P/(1 - T).$$

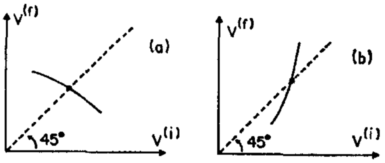

FIG. 3. Two types of self-consistent solutions.

Defining a proportionality ratio analogous to a dielectric constant

$$K = P/V,$$

one obtains

$$K = 1 - T.$$

As always, the stability requires a positive dielectric constant, or

$$T < 1,$$

which implies

$$\alpha_P > 0.$$

The argument above has a simple geometrical interpretation. The adding of an external perturbation is equivalent to displacing the $V^{(f)}$ curve vertically, as in Figs. 4a and b. This displacement moves the solution from 0 to 0′. In the case of Fig. 4a, a positive P implies a positive increase of the parameter of the solution. In the case of Fig. 4b, the molecule reacts in such a way that the parameter decreases.

Points of instability are probably the reason for the slowness of the convergence of some self-consistent calculations. A recent example was shown to us in the calculation of GaAs by the augmented-plane-wave method [6], when the starting potential, which was a sum of atomic potentials, fell near a point of instability. In cases like this, the iterative process may be long; from the point of instability to a stable point.

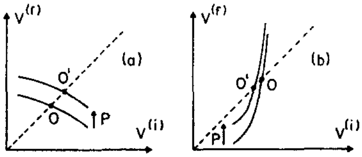

FIG. 4. Construction showing the stability of case (a) and instability in the case (b).

5. LINEAR THEORY IN THE CASE OF MANY PARAMETERS

We let $V^{(N)}$ represent the many parameters, and $V_i^{(N,i)}$ and $V_i^{(N,f)}$ the values at the beginning and at the end of the ith iteration. The linearity is now expressed as

$$V^{(N,f)} = \sum_M T_{NM} V^{(M,i)} + C^{(N)}. \tag{17}$$

At the end of iteration i, we know the values of the parameters at beginning and end for all the previous iterations. Using Eq. (17), we establish the relations

$$V_{i-n}^{(N,f)} - V_i^{(N,f)} = \sum T_{NM}(V_{i-n}^{(M,i)} - V_i^{(M,i)}). \tag{18}$$

From Eq. (18), we determine the matrix T_{NM}, which plays a role similar to that of T, for which we had just a single parameter.

To reach self-consistency rapidly, we will not only mix the potentials of the ith iteration but of previous iterations as well. So, instead of Eq. (1), we will prepare the initial potential for the $i+1$ iteration in the following way:

$$v^{(i)}_{i+1} = \sum_n \gamma_{i+1-n}[(1-\alpha)\, v^{(i)}_{i+1-n} + \alpha v^{(f)}_{i+1-n}]. \tag{19}$$

As many iterations will be summed as there are parameters. The γ's are weights that add up to one:

$$\sum_n \gamma_{i+1-n} = 1. \tag{20}$$

Equation (19) implies the following initial values of the parameters for the $i+1$ iteration

$$V^{(N,i)}_{i+1} = \sum_n \gamma_{i+1-n}[(1-\alpha)\, V^{(N,i)}_{i+1-n} + \alpha V^{(N,f)}_{i+1-n}]. \tag{21}$$

Inserting Eq. (21) in Eq. (17), we obtain $V^{(N,f)}_{i+1}$. Then, defining

$$\Delta^{(N)}_{i+1} = V^{(N,f)}_{i+1} - V^{(N,i)}_{i+1}, \tag{22}$$

we arrive at the equation

$$\begin{aligned}\Delta^{(N)}_{i+1} &= (1-\alpha) \sum_n \gamma_{i+1-n} \Delta^{(N)}_{i+1-n} \\ &\quad + \alpha \sum_M T_{NM} \sum_n \gamma_{i+1-n} \Delta^{(M)}_{i+1-n}.\end{aligned} \tag{23}$$

First of all, the iterative process will be faster when α and γ are chosen so that

$$\Delta^{(N)}_{i+1} = 0. \tag{24}$$

Let α_p be the value of α that satisfies Eq. (24). Then Eq. (24) implies

$$\sum_M \left(T_{NM} - \left(1 - \frac{1}{\alpha_p}\right) \delta_{NM}\right) A_M = 0, \tag{25}$$

where

$$A_N = \sum_n \gamma_{i+1-n} \Delta^{(N)}_{i+1-n}. \tag{26}$$

Then

$$1 - 1/\alpha_p$$

is one of the eigenvalues of the matrix T_{NM}. The size of the matrix is equal to the number of parameters, so there are as many possible values $\alpha^{(l)}_p$ as there are parameters.

If one of the eigenvalues $\alpha_p^{(l)}$ is negative, the iterative process is diverging because one is near an unstable point. Barring this circumstance, the $\alpha_p^{(l)}$ will all be positive. As will be seen shortly, in order to obtain a converging iterative process, one must choose the *smallest* eigenvalue $\alpha_p^{(l)}$. Once the value α_p is chosen, Eq. (26) is solved for the A_M. Then Eq. (25), together with Eq. (20), is solved for γ. The initial potential for the next iteration is then constructed according to Eq. (19).

Assuming we have chosen α and γ different from the eigenvalues $\alpha_p^{(l)}$ and the corresponding eigenvectors $\gamma_{i+1-n}^{(l)}$, the set of γ_{i+1-n} can be expanded in terms of the set of eigenvectors $\gamma_{i+1-n}^{(l)}$,

$$\gamma_{i+1-n} = \sum_l \beta_l \gamma_{i+1-n}^{(l)} . \tag{27}$$

According to Eq. (20),

$$\sum_l \beta_l = 1. \tag{28}$$

Then it is a simple matter to show that

$$\Delta_{i+1}^{(N)} = \sum_l \beta_l (1 - \alpha/\alpha_p^{(l)})\, A_N^{(l)}, \tag{29}$$

where $A_N^{(l)}$ are the eigenvectors of Eq. (26). Equation (29) shows that the iterative process is diverging if α is greater than twice the smallest eigenvalue $\alpha_p^{(l)}$, because repeated iterations would increase the deviations $\Delta^{(N)}$ due to the increasing powers of

$$1 - \alpha/\alpha_p^{(l)}.$$

The theory above was tested in a self-consistent calculation of the molecule CO by the cellular method. In this case there are two parameters, the potentials seen by each nucleus. We first made a run of 20 iterations according to Eq. (1) with $\alpha = 0.31$. The starting potential was of the Thomas–Fermi type. In Table I, we tabulate the deviations

$$\Delta^{(N)} = V^{(N,f)} - V^{(N,i)}$$

for each iteration, the total energy of the molecule E, and the energies ϵ_1, ϵ_2 of the two highest occupied one-electron states. For each three consecutive iterations of the first run, we used Eq. (18) to determine the matrix T_{NM}, Eq. (26) to determine the eigenvalues $\alpha_p^{(1)}$ and $\alpha_p^{(2)}$, and Eqs. (25) and (20) to determine the γ's corresponding to the smallest eigenvalue. Then a single iteration was made according to Eq. (19). In Table I, the iteration 4′ is based on iterations 1, 2, and 3; the iteration 5′ is based on 2, 3, and 4, but *not* 4′. Thus the primed iterations begin from the immediately anterior three unprimed iterations. Finally, we run the iterations 5″ and 6″ which are based on the iterations 2, 3, 4′ and 3, 4′, 5″, respectively. These double-primed iterations form the iterative sequence according to the present method.

TABLE I

Three Series of Iterations for the Molecule CO[a]

Interation	Sequence of iterations according to Eq. (1) with $\alpha = 0.31$					Single iterations according to Eq. (19)							Interation
	$\Delta^{(C)}$	$\Delta^{(0)}$	ϵ_1	ϵ_2	E	$\alpha_p^{(1)}$	$\alpha_p^{(2)}$	$\Delta^{(C)}$	$\Delta^{(0)}$	ϵ_1	ϵ_2	E	
1	6.9110	13.5689	−0.7455	−0.2440	−224.8820								
2	4.7067	8.0502	−0.7696	−0.3045	−225.3350								
3	3.2405	5.2841	−0.8425	−0.3680	−225.4801								
4	2.2076	3.4951	−0.8850	−0.4035	−225.5511	0.3941	1.0060	−0.1336	−0.4185	−0.9739	−0.4766	−225.6646	4′
5	1.4933	2.3182	−0.9090	−0.4228	−225.5887	0.5315	0.9400	−0.0043	−0.0245	−0.9484	−0.4526	−225.6454	5′
6	1.0050	1.5396	−0.9225	−0.4332	−225.6096	0.5609	0.9314	−0.0006	−0.0049	−0.9436	−0.4484	−225.6438	6′
7	0.6740	1.0231	−0.9301	−0.4388	−225.6217	0.5729	0.9280	−0.0002	−0.0009	−0.9416	−0.4466	−225.6430	7′
8	0.4507	0.6802	−0.9343	−0.4417	−225.6289	0.5672	0.9269	−0.0000	0.0004	−0.9407	−0.4458	−225.6428	8′
9	0.3008	0.4523	−0.9366	−0.4431	−225.6333	0.5519	0.9262	0.0001	0.0001	−0.9401	−0.4453	−225.6426	9′
10	0.2004	0.3308	−0.9378	−0.4438	−225.6361	0.5168	0.9260	0.0003	0.0000	−0.9399	−0.4450	−225.6426	10′
11	0.1340	0.2001	−0.9386	−0.4445	−225.6393	Sequence of iterations according to Eq. (19)							
						0.5293	0.9373	0.0021	−0.0187	−0.9475	−0.4519	−225.6450	5″
14	0.0391	0.0590	−0.9391	−0.4445	−225.6415								
						0.6032	0.9300	−0.0045	−0.0070	−0.9402	−0.4456	−225.6427	6″
17	0.0114	0.0174	−0.9390	−0.4443	−225.6421								
20	0.0033	0.0051	−0.9390	−0.4443	−225.6423								

[a] Energies are in Rydbergs.

Table I shows that the linear theory is adequate. Indeed, the fact that α_p remains approximately constant means that $V^{(C,f)}$ and $V^{(0,f)}$ are indeed functions of $V^{(C,i)}$ and $V^{(0,i)}$, and that these functions are linear. Second, Table I exhibits the improvements brought about by using Eq. (19) instead of Eq. (1).

6. CONCLUSION

Reaching self-consistency in atoms or small molecules is very easy. When one comes to the molecules with many atoms, the smallest eigenvalue $\alpha_p^{(l)}$ of Eq. (26) may be very small. In that case, the standard self-consistent iterative process, based on Eq. (1), may be unfeasible due to the necessity of using a very small value of α. Especially in this situation, one would use the present method and mix many iterations according to Eq. (19).

REFERENCES

1. D. R. HARTREE, "The Calculation of Atomic Structures," Wiley, New York, 1957.
2. G. W. PRATT, JR., *Phys. Rev.* **88** (1952), 1217.
3. F. HERMAN AND S. SKILLMAN, "Atomic Structure Calculations," Prentice-Hall, Englewood Cliffs, N.J., 1963.
4. J. P. DESCLAUX, D. F. MAYERS, AND F. O'BRIEN, *J. Phys. B* **4** (1971), 631.
5. L. G. FERREIRA AND J. R. LEITE, *Phys. Rev. A* **18** (1978), 335.
6. H. CLOSS, private communication.