Application of genetic algorithm to the calculation of bound states and local density approximations

Yehuda Zeiri
Department of Physics, Nuclear Research Center-Negev, Beer-Sheva 84190, P.O. Box 9001, Israel

Eyal Fattal and Ronnie Kosloff
Department of Physical Chemistry and The Fritz Haber Research Center for Molecular Dynamics, The Hebrew University, Jerusalem 91904, Israel

(Received 13 September 1994; accepted 16 November 1994)

A novel method, based on genetic algorithms, has been developed and applied to the solution of differential equations. The new approach is based on the use of real numbers to form the candidate solutions which are improved iteratively by a suitable breeding process. The algorithm was tested in the calculation of the bound states of a double well potential and in the nonlinear density functional calculation. Comparison of the results with those obtained using the direct relaxation method shows excellent agreement. © *1995 American Institute of Physics.*

I. INTRODUCTION

Genetic algorithms (GA) are global optimization methods based on several metaphors from biological evolution. The GA search for an optimal solution from a population of candidate solutions according to an objective function which is used to establish the fitness of each candidate as a solution. The governing process in the search is the application of appropriate breeding operators to candidate solutions in a given generation to form the candidates for the next generation. These operators are designed to preserve the most successful aspects of candidate fitness until the best possible solution is attained. Genetic algorithms have previously been applied in a wide variety of fields including engineering design, neural network synthesis, schedule optimization and conformational analysis of DNA.[1–3] In chemical physics these algorithms were applied to coherent control.[4,5]

In a recent publication a genetic algorithm was applied to the solution of ordinary differential equations (ODE).[6] The measure of how well does a particular candidate satisfy the ODE and its boundary conditions, after discretization of the differential operators, was used to define its fitness. The use of finite difference representation of the derivatives is a source of inefficiency since it involves only nearest neighbor interactions.

Recently, a new approach based on GA has been proposed for the search of lowest energy structure of molecular clusters.[7] This GA search procedure was based on the use of the control variables to represent the candidate solutions. The use of real number representation of the genotypes required the formulation of an appropriate set of operators which were used to form the $p+1$ generation when applied to candidates in the p'th generation.[7] In the method described below this approach was adapted, namely, continuous real numbers were used to form grid representations of candidate solutions to the ODEs. The second important feature of the new approach is the use of Fourier representation of the differential operators instead of the finite elements representation. Since the Fourier representation is a global one it inherently incorporates the boundary conditions and constitutes a more efficient description of the differential operators.[8] The GA operators were used in an adaptive fashion dynamically changing with the evolution, therefore the algorithm is termed real number adaptive genetic algorithm (RAGA). In the following, a brief outline of the new GA approach will be given. A detailed description of the method will be presented elsewhere.[9]

II. ALGORITHM DESCRIPTION

The initial set of N_{pop} candidate solutions where generated randomly, where a Gaussian random number generator was used to assign the components of the solutions at the grid points. Once the initial generation was formed, the fitness, f_i^p, values of the different candidate solutions were established (see definitions below). The next step in the algorithm was to construct a new generation by the application of various RAGA operators to candidate solutions of the present generation. This procedure was repeated until the convergence criteria were fulfilled. The transformation of the p'th generation to the $p+1$ generation was accomplished by the application of nine RAGA operators. These are listed as follows:

(1) (O_{best}): copy the k_{best} highest fitness candidate solutions to the new generation ($k_{\text{best}}/N_{\text{pop}} \simeq 0.05$). (2) ($O_{\text{rand}}$): formation of k_{rand} random candidates, to the new generation ($k_{\text{rand}}/N_{\text{pop}} \simeq 0.02$). (3) ($O_{\text{mute}}$)-$n$-mutation: one of the k_{best} genes is chosen randomly and some of its elements (randomly chosen) were modified by the addition of a random number evaluated from a normal Gaussian distribution. (4) (O_{inv}): inversion of the elements in a randomly chosen segment. (5) (O_{cros2}): two point cross-link between two parents to form two sons. (6) ($O_{\text{cros}n}$)-n-point cross link: the elements of two parent genes were copied to form two ''sons.'' The rearrangement of the parent's elements was accomplished by the following steps: a random number, ζ, was chosen from a uniform distribution in the range 0–1. If $\zeta \geqslant 0.5$ the element of parent 1 was copied to ''son 1'' and the corresponding element from parent 2 to ''son 2,'' while, if $\zeta < 0.5$ the element of parent 2 was copied to ''son 1'' and the corresponding element from parent 1 to ''son 2''. (7) (O_{av}): the arithmetic average of the elements of two parents

were used to form a son. (8) (O_{geom}): the geometric average of the elements of two parents were used to form a son.

In all cases, the probability to chose the ith individual as a parent was proportional to its fitness, f_i^p. It should be noted that all operations, except 2-point and n-point cross link, were performed in coordinate space, while O_{cros2} and $O_{\text{cros}n}$ were performed in coordinate or momentum space with 0.5 probability for each. An additional RAGA operator used was an inversion of the momentum space representation of candidate solution around its center, O_{invk}.

Fixed probabilities for the application of operations 1, 2 and 9 were used throughout the calculation, while the probabilities to use operators 3–8 were changed during the calculation according to their success to form high fitness candidates.[7] These probabilities were updated every N_p generations ($N_p \simeq 50$).

To complete the description of the new method the procedure of assigning the fitness functional f_i^p to the candidate solutions is described. The general form of the differential equations considered here is

$$\hat{\mathbf{H}}\Psi = (\hat{\mathbf{T}} + \hat{\mathbf{V}}) \cdot \Psi = \mathbf{E}\Psi, \tag{2.1}$$

where $\hat{\mathbf{H}}$ is the Hamiltonian, $\hat{\mathbf{T}} = -\hbar^2\nabla^2/2M$ is the kinetic energy operator and $\hat{\mathbf{V}}$ represents the potential energy operator. The kinetic energy operator is calculated by the Fourier method.[8] It should be noted that these general form of the Hamiltonian may include nonlinear terms, driving force terms, etc. If one is interested in the ground state of the system, the variational principle may be used to search for E_g and Ψ_g and these quantities could be used to evaluate the fitness of each candidate solution. However, if one is interested in the identification of an arbitrary eigenstate of the system which is closest to a given reference energy, E_{ref}, the evaluation of the f_i^p values could be done using the following procedure. Two quantities which are used in the search process are defined

$$\xi_1 = [\langle \Psi | (\hat{\mathbf{H}} - E_{\text{ref}})^2 | \Psi \rangle]^{1/2} \tag{2.2}$$

which is the expectation of the energy distance from a reference energy E_{ref}, and the energy dispersion:

$$\xi_2 = [\langle \Psi | \hat{\mathbf{H}}^2 | \Psi \rangle - (\langle \Psi | \hat{\mathbf{H}} | \Psi \rangle)^2]^{1/2}. \tag{2.3}$$

If Ψ_i represents the eigenstate whose energy is closest to E_{ref}, ξ_1 should converge to the energy difference between E_{ref} and E_i, while the dispersion ξ_2 should vanish. Hence, the fitness of a candidate solution was defined as $f_i^p = (\xi_1 \cdot \xi_2)^{1/2}$. This definition of the fitness allows us to obtain both eigenfunctions and eigenvalues closest to E_{ref} at any desired accuracy by the specification of the required value of f_i^p for convergence.

The last point which should be stressed is that one can use the orthogonality of the eigenstates of the system to enhance the rate of convergence. More specifically, during the search for the eigenstate which correspond to the energy value E_i (the closest to E_{ref}) one can project out from the candidate solutions components which belong to other eigenstates (which were already determined) by demanding the fulfillment of the orthogonality to these states.

TABLE I. Ground state eigenvalues for the double well system.

Mass	Iterations	Relaxation	GA
1	10000	0.886	0.887
10	10000	0.417	0.449
100	2000	0.139	0.139
1000	3000	0.0443	0.0443

III. APPLICATIONS

The examples chosen to demonstrate the RAGA algorithm are typical applications which have been explored by other methods such as quantum Monte Carlo[10,11] and direct methods.[12] The first example the double well potential is difficult to converge due to the tunneling splitting of the levels. The second example explores the electronic structure problem solved by the Hartree–Fock (HF) or the local density functional methods (LDA) which is difficult due to the nonlinear character of the equations.

A. Double well potential

The method was applied to the calculation of the eigenfunctions of an asymmetric double well potential recently used by Doll *et al.*[10] in the study of diffusional Monte Carlo algorithms. Table I shows the convergence of the method for the ground state eigenvalue together with the results of relaxation method[13] for various masses.

In both methods a grid of 32 points with $\Delta x = 0.16$ was sufficient to converge the results. From the table it can be concluded that it is harder to converge the cases with lighter mass in particular $M = 10$ for which the tunneling is most important. The power of the RAGA method to obtain directly the relevant eigenvalues is demonstrated in Figure 1.

First the ground state was obtained by searching for the eigenvalue closest to zero. This state was found to be localized in the deeper well. Then the target energy E_{ref} was again set to zero and the first state orthogonal to the ground state was obtained which was found to localize in the shallow

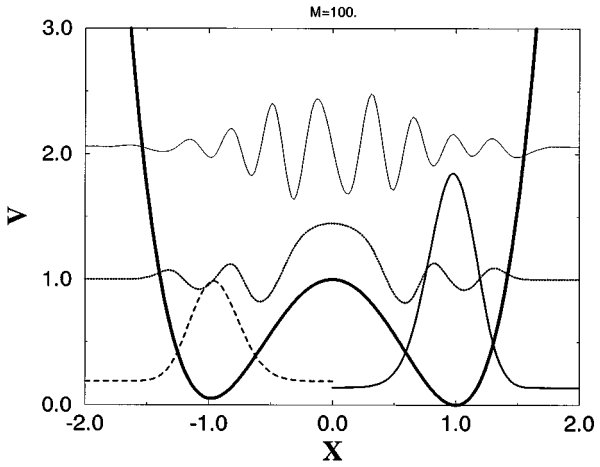

FIG. 1. The two lowest and two highly excited eigenfunction superimposed on their double well potential. Energies are 0.139 (solid) , 0.190 (dashed), 1.010 (dotted) 2.062 (solid). Potential parameters as in Ref. 10 with $M = 100$.

well. This procedure enables one to overcome also the problem of degenerate eigenstates. The two arbitrary chosen excited state eigenfunctions were obtained by setting $E_{\mathrm{ref}}=1.0$ which is the energy of the barrier and $E_{\mathrm{ref}}=2.2$.

In both cases the RAGA converged directly to the eigenstate closest to E_{ref} after 2000 iterations.

B. Nonlinear electronic structure example

The use of the Hartree–Fock and density functional approximations to solve the electronic structure problem leads to nonlinear differential equations. The nonlinear effective Hamiltonian for this problem becomes

$$\hat{\mathbf{H}}_{\mathbf{eff}}\psi_i=\left[-\frac{1}{2}\nabla^2+\hat{\mathbf{V}}_{\mathbf{eff}}\right]\psi_i=\epsilon_i\psi_i\,, \tag{3.1}$$

where i is the orbital index, and

$$\hat{\mathbf{V}}_{\mathbf{eff}}=\hat{\mathbf{v}}_{\mathbf{ext}}(\mathbf{r})+\int\frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\,d\mathbf{r}'+\hat{\mathbf{v}}_{xc}(\rho;\mathbf{r});$$

$$\rho(\mathbf{r})=\sum_i|\psi_i(\mathbf{r})|^2. \tag{3.2}$$

In the Hartree–Fock approximation in this example $\hat{\mathbf{v}}_{xc}=0$ and the second Coulombic term in the RHS of (3.2) is multiplied by half. When using the density functional approximation $\hat{\mathbf{v}}_{xc}$ is approximated by LDA.[11,12,14] The total energy functional becomes

$$E[\rho]=\sum_i\epsilon_i-\int\rho(\mathbf{r})\hat{\mathbf{v}}_{\mathbf{eff}}(\mathbf{r})d\mathbf{r}+J[\rho]$$

$$+\int\rho(\mathbf{r})\hat{\mathbf{v}}_{\mathrm{ext}}(\mathbf{r})d\mathbf{r}+E_{xc}[\rho]. \tag{3.3}$$

Where $J[\rho]$ is the Coulombic integral.[12] In the Hartree Fock approximation $\mathbf{E}_{xc}=0$ and $J[\rho]$ is multiplied by half.

To test the RAGA method a two electron problem was analysed. Analytic results can be obtained for harmonic binding potential between the electron and nucleus: $\hat{\mathbf{v}}_{\mathbf{ext}}=r^2/8$.[15] For this reason this problem has been chosen as a standard benchmark.[11,12] For this system the exact total energy becomes $E=2$.

The Hartree–Fock approximation[12] leads to $E^{\mathrm{HF}}=2.039\,33$ and the LDA result becomes $E^{\mathrm{LDA}}=2.026\,39$. The RAGA method was tested using a grid of 128 or 64 evenly spaced points with $N_{\mathrm{pop}}=200$. Both the Hartree-Fock and LDA cases were tested where the fitness criteria was the total energy $E[\rho]$ [Eq. (3.3)]. Figure 2 shows the convergence of the method for the first 2000 iterations. After 5000 iterations the results converge to the numerical accuracy of the numbers above.

It seems that the fact that the problem is nonlinear has no effect on the convergence. In Fig. 2 the very fast initial convergence of the method for the first 100 iterations is clearly seen. This phenomena has been found to be general in all applications tested.

The method has also been tested for a Helium atom where $\hat{\mathbf{v}}_{\mathbf{ext}}=-2/r$. The converged and experimental results of the He ground state energy are $E=-2.904$. The converged HF result for this problem is $E=-2.862$.[16] Using a grid of 128 points and $\Delta r=0.078\,125$ the method converged to $E=-2.8213$ for Hartree–Fock and to $E=-2.8338$ for LDA after 6000 generations. These numbers are converged in comparison to the ones obtained by the direct relaxation method using the same grid. It should be mentioned that the error is due to the use of an evenly spaced grid which only partially overcomes the Coulomb singularity. This grid leads to slow convergence and thus is not a result of the RAGA algorithm. This problem of grid representation can be overcome by using mapped grids[17] which can also be used in the RAGA algorithm. The extended range of energy in the Hamiltonian representation inherent in the Coulomb problem did not cause difficulties in the RAGA algorithm. This is contrary to the direct relaxation methods where the numerical effort increases with the energy range.

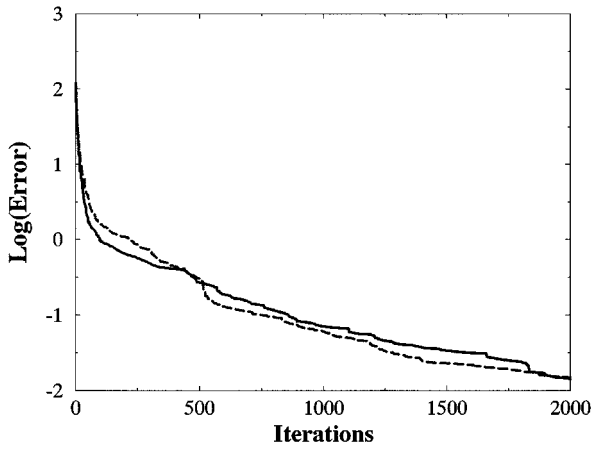

FIG. 2. Convergence of the genetic algorithm (where the error is defined as the difference between the converged result and the best energy functional of the current generation) as a function of the generation index. Shown are the density functional LDA results (solid) and HF results (dashed) for the two electron problem (Refs. 11 and 12). A grid of 64 points was used with $\Delta r=0.15625$ and $N_{\mathrm{pop}}=200$.

Direct method of solution of nonlinear problems require a very good initial guess in order to converge at all. The fast initial convergence of the RAGA and the fact that it is not dependent on an initial guess suggests a combined method where the RAGA prepares a good initial guess for the direct method. This has been tested by using the result after 50 generations for a direct method procedure based on a relaxation method.[17] This initial state then led to very fast convergence.

IV. CONCLUSIONS

Variational approaches allow one to convert differential equations into search problems. A new search method based on genetic algorithms has been described. The floating point representation of the genotypes requires the definition of a set of modified GA operators by which the search is advanced from one generation to the other. Convergence can be enhanced by the use of orthogonalization and symmetry operations. The probability to use the different GA rules defined in II was varied during the calculation. Examining the

importance of the these rules in the present examples shows that the cross link and averaging operators are the dominant ones during most of the calculation. These rules are expected to influence the rate of convergence more than the the ones that were used less frequently. This feature is currently under study. The direct search for an excited eigenfunction demonstrates the ability of the method to converge to the global minimum since each of the eigenvalues of the Hamiltonian constitutes a local minimum. The desired eigenvalue has only a larger cone of attraction. The RAGA is able to cope with solutions to nonlinear differential equations with strong singularities showing no sensitivity to large energy ranges. Other methods to solve these examples could be used such as Monte Carlo, and simulated annealing. However comparison with the GA approach showed that the later converged significantly faster.

To summarize the new search procedure has been shown to be an effective tool for a variety of quantum mechanical problems.

[1] D. E. Goldberg, *Genetic Algorithms in Search, Optimization, and Machine Learning* (Addison-Wesley, Reading, MA, 1989).
[2] L. Davis, *Handbook of Genetic Algorithms* (Van Nostrand Reinhold, New York, 1991).
[3] *Proceedings of the 3rd International Conference on Genetic Algorithms*, edited by J. D. Schaffer (Morgan Kaufman, San Mateo, 1989).
[4] R. S. Judson and H. Rabitz, Phys. Rev. Lett. **68**, 1500 (1992).
[5] G. J. Tóth and A. Lörincz, J. Chem. Phys. (unpublished).
[6] D. A. Diver, J. Phys. A **26**, 3503 (1993).
[7] (a) Y. Zeiri, Phys. Rev. E (submitted), for publication; (b) Y. Zeiri (in preparation).
[8] R. Kosloff, in *Numerical Grid Methods and Their Application to Schrodinger's Equation*, edited C. Cerjan (Kluwer Academic, The Netherlands, 1993), p. 175–194.
[9] R. Kosloff and Y. Zeiri (in preparation).
[10] J. Doll, Chem. Phys. Lett. **219**, 343 (1994).
[11] C. Filippi, C. J. Umrigar, and M. Taut, J. Chem. Phys. **100**, 1290 (1994).
[12] S. Kais, D. R. Herschbach, N. C. Handy, C. W. Murray, and G. J. Laming, J. Chem. Phys. **99**, 417 (1993).
[13] R. Kosloff and H. Tal-Ezer, Chem. Phys. Lett. **127**, 223 (1986).
[14] J. P. Perdew and Y. Wang, Phys. Rev. B **45**, 13244 (1992).
[15] N. R. Kestner and O. Sinanoğlu, Phys. Rev. **128**, 2687 (1962).
[16] F. L. Pilar, *Elementary Quantum Chemistry* (McGraw-Hill, New York, 1990), p. 279.
[17] E. Fattal and R. Kosloff (in preparation).