Comparison of Line Search Minimization Algorithms for Exploring the Topography of Multidimensional Potential Energy Surfaces: the Mg^+Ar_n case

J. Papadakis, G. S. Fanourgakis, S. C. Farantos, and M. Founargiotakis
Institute of Electronic Structure and Laser,
Foundation for Research and Technology - Hellas, and
Department of Chemistry
University of Crete,
Iraklion, Crete 711 10, Greece.

Abstract

The most robust numerical algorithms for unconstrained optimization which involve a line search are tested in the problem of locating stable structures and transition states of atomic microclusters. Specifically, the popular quenching technique is compared with conjugate gradient and variable metric algorithms in the Mg^+Ar_n clusters. It is found, that the variable metric method BFGS combined with an approximate line minimization routine is the most efficient and it shows global convergence properties. This technique is applied to find a few hundreds of stationary points of Mg^+Ar_{12} and to locate isomerization paths between the two most stable icosahedral structures found for Mg^+Ar_{12}. The latter correspond to a solvated and a non–solvated ion respectively.

I. Introduction

The Born–Oppenheimer approximation and the concept of molecular potential energy surface (pes) which results from this approximation, remain the basic approach in the theory of molecular systems. The location of the stationary points of the pes is the first step in globally understanding the dynamics of a molecule. Indeed, as we learn from non–linear mechanics[1] there is a hierarchical approach for the exploration of the geometry of phase space, and thus, the dynamics of a multidimensional system. This starts with the location of the stationary points of the potential function, and if one wants to cover in more detail the dynamical behavior of the non–linear system, he continues with the location of periodic orbits, tori, and stable and unstable manifolds that emerge from unstable periodic orbits.

The completion of such a program will give the global classical dynamics of the system at every energy. Unfortunately, this is a difficult task and it can be accomplished only for a few model systems, usually of two degrees of freedom.[2] For small polyatomic molecules the location of the stationary points of the pes, as well as of the most important periodic orbits and bifurcations of them it has been proved to be feasible.[3] A recent example is acetylene, a prototype molecule for studying quantum chaos in molecules.[4]

On the other hand, microclusters consisted of a few tens of atoms or molecules are too large for a detailed description with periodic orbits, and at the same time too small for a pure statistical mechanical treatment. The latter is manifested in molecular dynamics simulations where non–ergodic behavior is observed. For example, a classical dynamical study of small water clusters has shown,[5] that in the time scale of a few hundreds of picoseconds the system is far from distributing the energy among all degrees of freedom. Although, the trajectories of small clusters may show an early transition to chaos, still, the existence of approximate constants of motion may result in interesting dynamical phenomena, especially at energies where isomerization processes are possible.[5,6]

The location of stationary points of the pes is a zero order approximation to the global dynamics of the system.[3] By knowing the stationary points we can sketch the topography of the multidimensional pes, and thus, we can extract some general characteristics of the dynamics. In recent articles Berry and coworkers[7–9] exploit this methodology for investigating the cluster dynamics of inert gases and ionic substances.

Among the minimization algorithms the steepest descent method has been used in numerous studies of atomic and molecular clusters. Usually, this method is applied in conjunction to molecular dynamics techniques through the quenching of the trajectories to the nearby stationary point.[10] The advantage of the steepest descent or quenching method is its global convergence properties, i.e., it converges to a minimum from an arbitrary initial point, and that contrasts the Newton–Raphson method which always requires a good starting point in phase space for a successful convergence.[11] On the other hand, it is well known that the quenching method is considerably slow, whereas the Newton–Raphson method converges quadratically. The combination of the quenching method with the Newton–Raphson method tries to remedy the above drawback. The way that we have used this hybrid scheme in water clusters[6] is to apply first the quenching method to approach a stationary point with low accuracy, but close enough, such that the Newton–Raphson method that follows, converges to the stationary point.

The above deficiencies of these quite popular optimization schemes to cluster dynamics are well known to numerical analysts, and thus, more efficient algorithms have been proposed for finding stationary points in multidimensional functions with a large number of degrees of freedom. Particularly, the conjugate gradient methods and the variable metric methods are those which have been studied extensively, and it has been shown, that these methods behave well as far as the convergence criteria are concerned. Their common characteristic is that they are applied by using an iterative procedure for defining at every iteration an appropriate line along which the potential function is minimized. These algorithms require the evaluation of the gradient. Since, in studies of cluster dynamics we usually have an analytical function for the pes the latter is not a limitation.

The main purpose of the present article is to make a critical evaluation of the performance of the above minimization algorithms, quenching, conjugate gradient, and variable metric, in the specific case of atomic aggregates which contain inert gases. As an application we study the Mg^+Ar_n clusters.

Recently, we have investigated with molecular dynamics techniques the equilibrium structures and dynamics of small Mg^+Ar_n and $Mg^{++}Ar_n$ aggregates.[12] Particularly, for the Mg^+Ar_{12} it was found that two almost isoenergetic icosahedral structures exist. One with Mg^+ in the center of the icosahedron, and one with the cation on the surface of the cluster. Some authors are talking about solvated and non–solvated forms.[13] Such a finding immediately raises the question about the mechanism of isomerization from one conformation to the other. To answer such questions it is necessary to locate the intermediate minima and saddle points of the potential function. For this purpose we construct *stationary point diagrams*, that is, the distribution of the minima and saddle points along a reaction coordinate which in our case is the distance of magnesium cation from the center of mass of the cluster.

Although, we shall present *stationary point diagrams* for the doped argon clusters, however, we point out that this article focuses on the comparison of the different optimization techniques in finding local minima. We do not discuss methods for "walking" on the potential energy surface[14–17] or to find the global minimum of it.[18,19] We shall show, that the variable metric and conjugate gradient schemes are considerably more efficient than the traditional quenching techniques, and at the same time, they show the good convergence properties encountered in quenching methods.

II. Algorithms

The methods which we shall compare are, the steepest descent, the Newton–Raphson, from the conjugate gradient methods the Fletcher–Reeves (FR) and the Polak–Ribiere (PR), and from the variable metric methods the Broyden–Fletcher–Goldfarb–Shanno (BFGS) and the Davidon–Fletcher–Powell (DFP) algorithms. All of them are based on the location of an appropriate direction along which a new point is selected and at which the potential has a lower value than in the previous point. Both the conjugate gradient and the variable metric methods are based on the theorem which proves that at most n line minimizations are required for a quadratic function of n dimension to converge to the exact minimum. Since, the potential is minimized along the selected directions the choice of a good line minimization algorithm is also important.

An excellent discussion about the theory of algorithms for unconstrained optimization is presented by Nocedal.[20] In the following we give a brief description of these optimization techniques, but for details the reader is addressed to the books in refs [21–23].

The basic problem is to minimize a function of n variables, $\vec{q}$,

$$\min \quad V(\vec{q}),$$

where $V(\vec{q})$ is the pes and $\vec{f}(\vec{q}) = \nabla V(\vec{q})$ its gradient. The stationary points are located by using iterative schemes of the form

$$\vec{q}_{k+1} = \vec{q}_k + \lambda_k \cdot \vec{d}_k, \quad 0 < \lambda_k \leq 1, \tag{1}$$

where $\vec{d}_k$ is a search direction and λ_k determines the step–length which leads to a minimum of the function V along the direction $\vec{d}_k$. Usually, the starting point, $k = 0$, is a prespecified point in the configuration space, $\vec{q}_0$, and the initial search direction is given by the slope of the potential at that point (the steepest descent), $\vec{d}_0 = -\nabla V(\vec{q}_0)$.

The several line search optimization algorithms differ in the way that they choose the minimization lines and they can be classified by parameterizing the search direction with the matrices a_k and the scalar quantities b_k;

$$\vec{d}_k = -a_k \vec{f}_k + b_k \vec{d}_{k-1}. \tag{2}$$

The algorithms which we shall examine are described as:

1. Steepest descent

$$a_k = I, \quad b_k = 0, \tag{3}$$

where I is the unit matrix.

2. Newton–Raphson

$$a_k = H_k^{-1} = \left[\frac{\partial^2 V(\vec{q}_k)}{\partial \vec{q}_k^{\,2}} \right]^{-1}, \quad b_k = 0, \tag{4}$$

where H is the Hessian matrix.

3. Fletcher–Reeves

$$a_k = I, \quad b_k = \frac{\vec{f}_k^T \cdot \vec{f}_k}{\vec{f}_{k-1}^T \cdot \vec{f}_{k-1}}, \tag{5}$$

where $\vec{f}_k^T$ denotes the row vector of $\vec{f}_k$ (column vector).

4. Polak–Ribiere

$$a_k = I, \quad b_k = \frac{(\vec{f}_k - \vec{f}_{k-1})^T \cdot \vec{f}_k}{\vec{f}^T_{k-1} \cdot \vec{f}_{k-1}}. \tag{6}$$

5. Broyden–Fletcher–Goldfarb–Shanno

$$a_k = B_k^{BFGS}, \quad b_k = 0, \tag{7}$$

where B_k^{BFGS} is a nonsingular symmetric matrix that approximates the inverse of the Hessian of the potential function. The updates of B_k^{BFGS} are given by

$$B_{k+1}^{BFGS} = B_k^{BFGS} - \frac{B_k^{BFGS} \cdot y_k \cdot y_k^T \cdot B_k^{BFGS}}{y_k^T \cdot B_k^{BFGS} \cdot y_k} + \frac{s_k \cdot s_k^T}{s_k^T \cdot y_k} + X, \tag{8}$$

where

$$\vec{y}_k = \vec{f}_{k+1} - \vec{f}_k, \qquad \vec{s}_k = \vec{q}_{k+1} - \vec{q}_k, \tag{9}$$

$$X = \vec{y}_k^T \cdot B_k^{BFGS} \cdot \vec{y}_k \cdot u_k \cdot u_k^T, \tag{10}$$

and

$$\vec{u}_k = \left[\frac{\vec{s}_k}{\vec{s}_k^T \cdot \vec{y}_k} - \frac{B_k^{BFGS} \cdot \vec{y}_k}{\vec{y}_k^T \cdot B_k^{BFGS} \cdot \vec{y}_k} \right]. \tag{11}$$

The initial matrix B_0^{BFGS} is usually the unit matrix, but it can also be, and is recommended, the inverse of the Hessian at the specified initial point if it is available. Looking for a local minimum the initial Hessian should be a positive definite matrix.

6. Davidon–Fletcher–Powell

$$a_k = B_k^{DFP}, \quad b_k = 0, \tag{12}$$

where

$$B_k^{DFP} = B_k^{BFGS} - X. \tag{13}$$

Other updating formulae for B have been proposed, for example the symmetric rank–one methods,[20,24] but since the BFGS and the DFP algorithms are the most popular we have tested only them. The above formulae have been coded in Fortran programs and applied to the study of Mg^+Ar_{12} cluster. For the conjugate gradient and variable metric methods we have modified the codes given in the book "Numerical Recipes".[11]

III. Results and Comparisons

A. Potential Function

We have described the interaction potential of Mg^+Ar_n clusters as

$$V_{tot} = \sum_{k=2}^{n+1} V_n^+(1,k) + \sum_{j=2}^{n} \sum_{k=j+1}^{n+1} V_{nn}(j,k). \tag{14}$$

The first sum describes the interactions of Mg ion (labeled as atom 1) with the atoms of argon (labeled as atoms 2,3,...,$n+1$). The second term in Equation (14) denotes the interactions among argon atoms.

The first term is developed as

$$\begin{aligned} V_n^+(1,k) = V_0(r_{1k}) + V_{dis}(r_{1k}) - \frac{1}{2} a_k E_k^2 \Phi(r_{1k}) - \frac{1}{2} a_1 \vec{E}_{1k} \vec{E}_1 \\ + \frac{a_1 a_k E_1 E_k}{r_{1k}^3} [\hat{E}_1 \hat{E}_k - 3(\hat{E}_1 \hat{r}_{1k})(\hat{E}_k \hat{r}_{1k})], \end{aligned} \tag{15}$$

where

- $\vec{r_j}$ is the position vector of atom j in a space fixed cartesian coordinate system, $\vec{r}_{jk} = \vec{r}_j - \vec{r}_k$, and its norm $r_{jk} = |\vec{r}_j - \vec{r}_k|$.
- a_1 is the polarizability of magnesium cation.
- $a_k = a_{Ar}$ is the polarizability of argon atom k, $k = 2, 3, \ldots, n+1$.
- $\vec{E}_k$ is the intensity of the electric field at the position of atom k,

  $$\vec{E}_k = \frac{Q}{r_{k1}^2}(1 - \delta_{k1})\hat{r}_{k1} + \sum_{l=1, l\neq k}^{n+1} \frac{a_l E_l}{r_{kl}^3}[3(\hat{r}_{kl}\hat{E}_l)\hat{r}_{kl} - \hat{E}_l], \tag{16}$$

  where Q is the electric charge on atom 1, and δ_{k1} is Kronecker's delta function.
- $\vec{E}_{1k}$ is the contribution to the intensity of the electric field at the position of the ion which is due to the dipole moment of atom k.
- $\hat{E}_k$, and $\hat{r}_{jk}$ are the unit vectors in the directions of the intensity of the electric field on the atom k, and of the position of atom k relative to atom j respectively.
- $\Phi(r_{1k})$ is a switching function which effects a smooth interpolation between short range and charge–induced dipole interactions.

The first two terms in Equation (15) denote the repulsive and the dispersion potentials respectively. Both these terms are described with a diatomic Morse type function

$$V_0(r) \; + V_{dis}(r) \;=\; Ae^{2B(1-\frac{r}{r_m})} \;-\; 2Ce^{B(1-\frac{r}{r_m})}. \tag{17}$$

The third term in Equation (15) is the charge–induced dipole interaction, the forth term describes the effect of all other atoms on the ion, and the fifth term is due to the induced dipole–induced dipole interactions.

The part of the potential which yields the Ar–Ar interactions, is written as

$$V_{nn}(j,k) = V(r_{jk}) \;+\; \frac{a_j a_k E_j E_k}{r_{jk}^3}[\hat{E}_j \hat{E}_k \;-\; 3(\hat{E}_j \hat{r}_{jk})(\hat{E}_k \hat{r}_{jk})]. \tag{18}$$

This is the sum of a potential valid for a free argon dimer and an electrostatic induced dipole – induced dipole interaction. $V(r_{jk})$ is given by an empirical function proposed by Aziz and Slaman[25]

$$V(x) = \varepsilon\left(Ae^{-ax+bx^2} - \left\{\begin{array}{ll} e^{-(\frac{d}{x}-1)^2} & x < d \\ 1 & x \geq d \end{array}\right\}\left[\frac{c_6}{x^6} \;+\; \frac{c_8}{x^8} \;+\; \frac{c_{10}}{x^{10}}\right]\right), \tag{19}$$

where

$$x = r/r_0.$$

The polarizabilities of the atoms are $a_{Ar} = 11.13\ a_o^3$,[26] $a_{Mg^+} = 37.11\ a_o^3$.[26] It should be noted that Equation (14) is not a pairwise additive function. Many body interactions are introduced through the electrostatic terms. The induced dipole moment on atom k depends on the positions of all other atoms.

Results from ab initio calculations[26–29] for the dimer are used to fit the parameters of the Morse type potential (Equation (17)), and these are: the minimum energy, –3.25 kcal/mol, the harmonic frequency, 92 cm^{-1}, and the equilibrium bond length, $r_m = 5.47a_o$. The following switching function Φ (Equation(15)) was chosen

$$\Phi(r) = tanh\left(\frac{r \;-\; D}{E}\right). \tag{20}$$

We adjust the parameters D and E in such a way that the potential of the triatomic Mg^+Ar_2 reproduces the energy and geometry of the minimum obtained from the ab initio calculations of Bauschlicher et al..[26] The parameters of the above potential are given in ref. [12].

B. Comparison of Algorithms

In the following we shall examine the speed and the ability for convergence of the chosen optimization algorithms. We select initial conditions by using two methods. A systematic one by deforming the cluster from a known equilibrium geometry, and a random one by choosing an initial point with an arbitrary arrangement of the atoms. The second method usually gives high energy structures.

All runs were performed on a HP-735 computer with 64 Mbytes memory and with the operating system HP-UX v.9.01. The programs were compiled with the O2 optimizer.

1. The Newton–Raphson and the quenching methods

The Newton–Raphson approach requires the computation of the second derivatives of the potential. Furthermore, it is difficult to make a straightforward comparison of the Newton–Raphson method with the other techniques, since, it is always necessary to have a well thought starting point. This problem can be cured by applying damped Newton methods as we describe it later on, but then, the advantage of quadratic convergence is lost and the conjugate and variable metric algorithms become more efficient.

The quenching method, solution of the equations of motion with friction,[10] is the slowest when the location of the stationary points with high accuracy is required. Nevertheless, it is the most popular, especially in combination with the Newton–Raphson method. In the following we compare the quenching method applied without the Newton-Raphson step, and demanding the same convergence accuracy as in the other optimization techniques.

2. The Fletcher–Reeves and Polak–Ribiere conjugate gradient methods

These two methods are proved to be equivalent for quadratic functions but differences arise for general non quadratic functions. We have used a routine from Numerical Recipes (dbrent) for accurately minimizing the potential along the conjugate directions. Previous experience reveals that the Polak–Ribiere method is more efficient and converges quicker to the minimum.[11] In Fig. 1 we show representative convergence plots for two cases, one with a systematically selected point by deforming a known equilibrium geometry (Fig. 1a), and one case with a randomly chosen initial configuration (Fig. 1b). From Fig. 1 we can see the superiority of the Polak–Ribiere conjugate gradient method, particularly in the case when we start the minimization procedure from a random point. In Fig. 1b the Polak–Ribiere algorithm converges in 84 steps compared to 1500 steps needed by the Fletcher–Reeves method.

3. The BFGS and DFP variable metric methods

In these algorithms a subroutine which finds the minimum along the current direction is also required. The minimum can be found with an accurate line minimization algorithm or with an approximate one. As becomes apparent from the Equations (7) and (12) the variable metric methods differ from the Newton–Raphson method in the approximate way they calculate the inverse of the Hessian matrix. That is why the variable metric methods are also named *quasi–Newton methods.* It is well known that in the Newton–Raphson method there is no guarantee that moving in the descent direction a full Newton step the potential function will decrease. In situations where the function increases it is better to move a fraction of the Newton step but ensuring that the potential will decrease. Such methods are called *damped or under-relaxed Newton methods* and of course they do not converge quadratically but super–linearly.

Damped techniques can also be applied in quasi–Newton methods. We have tested the speed and convergence properties of the variable metric algorithms, BFGS and DFP, using an accurate line minimizer, the subroutine dbrent, and an approximate minimizer, the subroutine lnsrch, which are suggested in the book "Numerical Recipes".[11]

We did not find any substantial difference between BFGS and DFP (Figs 2a and 2b) in searching for stationary points in the cluster Mg^+Ar_{12}, and even for larger clusters with 14 argon atoms, when we used the accurate line minimization subroutine (dbrent).

This result changed dramatically when we used the approximate line minimization algorithm (lnsrch). In the example shown in Fig. 2d the minimum is achieved in

150 steps with the BFGS method compared to 3500 steps needed in the DFP. Overwhelmingly we conclude, that the BFGS method is much faster than the DFP method when an exact location of the minimum along a specified direction is not asked. Furthermore, tests with several initial conditions showed that the damped–BFGS routine always converges to a minimum with no particular problems.

In the following we compare the best conjugate gradient method, the Polak–Ribiere, with the best version of the variable metric algorithms, the damped–BFGS method, using the approximate line minimizer lnsrch.

4. Comparison of PR and damped–BFGS methods

In Fig. 3 we show convergence plots for two representative initial points; one selected from a prespecified minimum (Fig. 3a) and one random (Fig. 3b). The PR algorithm converges in less iteration steps in both cases. However, the amount of time required in the BFGS method is less than that of the PR method. Specifically, 9.62 seconds and 22.02 seconds for the case of Fig. 3a and 12.24 compared to 61.98 seconds in the case of Fig. 3b. The above differences are understood, since overall the potential and its derivatives are called more times in each iteration in the PR method than in the BFGS method.

In TABLE I we compare in four different cases the BFGS method with an exact line minimizer (dbrent) as well as with the approximate one (lnsrch), the Polak–Ribiere method and the quenching algorithm. We tabulate the total time, the time consumed by the calls of the potential and its derivatives, and the number of calls of the potential function. All studied cases classify the damped–BFGS as the fastest method.

It is worth mentioning that the properties of global convergence, i.e., to converge always to a minimum from every starting point, which is guaranteed in the quenching method,[20] is also valid for the PR and BFGS algorithms. However, we found that it is possible different algorithms to converge to different minima despite that they start from the same initial point in configuration space. For example, starting from a random initial point the Polak–Ribiere and the quenching methods as well as the BFGS and the quenching methods the probability to converge at the same minimum was about 50%. In a few cases we have also found the algorithms to converge not to a minimum but to a saddle point of the potential energy surface.

C. Isomerization Reaction Paths

Using mainly the damped–BFGS method we have located 95 minima. To find the saddle points among them we have applied the ideas of Berry and coworkers.[17] By running trajectories which connect two consecutive minima we can approach the saddle points. Then, the use of one of the optimization methods can accurately locate the saddle. An alternative method to find transition states is to follow the reaction path starting from one minimum.[14–16]

143 saddle points of Mg^+Ar_{12} in the energy range of about 2.5 kcal/mol have been found. The minima with the lowest energy are two icosahedral structures, one with the magnesium cation in the center of the icosahedron, isomer–(a) $(iso-(a))$, and the other with the cation on the surface, isomer–(b) $(iso-(b))$. In the previous article[12] we have shown that the absolute minimum is the $iso-(b)$ (-23.589 kcal/mol) with an energy 0.5 kcal/mol lower than that of $iso-(a)$, ($E = -23.056$ kcal/mol). The zero of the energy corresponds to a configuration with all atoms separated.

In Fig. 4 we have arranged energetically the minima (filled squares) and transition states (open circles) along an X-axis which represents the distance of magnesium cation from the center of mass of the cluster, R. Thus, the $iso-(a)$ corresponds to $R = 0\ a_o$ and stationary points with a non–solvated magnesium can reach configurations with R up to 6.5 a_o. The variable R may be considered as an effective reaction coordinate which describes the isomerization process.

Such *stationary point diagrams* can provide useful information about the mechanism of the isomerization processes in a cluster. From Fig. 4 we immediately note that the high symmetry conformations, $iso-(a)$ and $iso-(b)$, have a special stability because of the energy gaps that separate these minima from the relative minima. We can also remark that the density of the stationary points above the $iso-(b)$ is higher than that of the $iso-(a)$. This means that the cluster may be trapped in this region of configuration space for longer times before it reaches $iso-(b)$.

Having located all these minima and saddle points we now address the question about possible isomerization reaction paths of

$$iso-(a) \longleftrightarrow iso-(b). \tag{21}$$

The paths for the isomerization reaction are drawn by following the steepest descent directions from a specific transition state. In Fig. 5 we show a few cases of how we move from $iso-(a)$ to $iso-(b)$ or in reverse by connecting transition states with the neighboring minima. Some paths coincide in certain parts of the isomerization process. The thick solid and dashed lines on Fig. 5 are two paths for which we have drawn the structures of the visited stationary points and these are shown in Figs 6 and 7 respectively. M stands for a minimum and S for a saddle.

The computer can find millions of distinct isomerization paths. The mechanism that emerges from such diagrams is that the system may wander in configuration space passing through solvated and non–solvated structures before it reaches its icosahedral configurations. Specifically, in the stationary point diagram shown, we note, that all reaction paths pass through the highest saddle point (-21.2 kcal/mol). The geometry of this saddle point is marked with an asterisk in Figs 6 and 7. Its energy is in good agreement with previous molecular dynamics calculations where we estimated the transition from solid–like to liquid–like behavior of the $iso-(a)$ with a caloric curve and a root-mean-square diagram of the fluctuations of the bond lengths.[12] The idea of the existence of a unique transition state which dictates a reaction is particularly appealing, since we may use it in approximate rate theories such as RRKM.

However, at present it is difficult to make any final conclusion, and the above qualitative picture will be tested in the future with dynamical calculations where the isomerization rate constants will be computed.[30]

IV. Conclusions

We have carried out an evaluation of several unconstrained optimization algorithms in the problem of locating stationary points in multidimensional potential energy surfaces as those of atomic clusters. Particularly, we have tested the quenching method with the conjugate gradient methods of Polak–Ribiere and Fletcher–Reeves, and the variable metric methods of Broyden–Fletcher–Goldfarb–Shanno and Davidon–Fletcher–Powell.

We confirmed the superiority of both the conjugate gradient and the variable metric methods compared to the quenching method. All of them show good global convergence properties which means, that they converge to a minimum from arbitrary configuration initial points. An interesting result that we have found in this study is the larger speed of convergence of the BFGS method compared to DFP when the approximate line minimizer, lnsrch, is used. The BFGS-lnsrch routine is a damped quasi-Newton algorithm.

Although these results come out of the study of ion metal doped argon clusters the algorithms we have tested are general enough[20] to be applied to any smooth pes which describes other types of clusters. However, differences in the convergence rates may be found.

Using mainly the damped–BFGS method we have located 238 stationary points in the Mg^+Ar_{12} atomic aggregate. These minima and saddle points bridge the two icosahedral structures of the cluster which correspond to a solvated and a non–solvated cation. The arrangement of the stationary points in a plot of (Reaction Coordinate, Energy) provides useful information about the isomerization mechanism.

In our case we have chosen as reaction coordinate the distance of the Mg^+ from the center of mass of the cluster. From the constructed *stationary point diagrams* we have traced possible isomerization paths of $iso-(a) \longleftrightarrow iso-(b)$. It is interesting to find that the energy of the highest transition state in these specific paths is in proximity with the energy of a phase transition from solid–like to liquid–like determined in molecular dynamics calculations.[12]

As Berry and coworkers have shown[7–9] the *stationary point diagrams* may also be used for developing kinetic theories for the calculation of the reaction rate constants.

Such calculations will be presented in another publication for the Mg^+Ar_{12} cluster.[30]

Acknowledgments

This work was financially supported by the Greek Secretariat for Research and Technology, PENED-1994 (15774/296).

References

[1] S. Wiggins. *Introduction to Applied Nonlinear Dynamical Systems and Chaos.* Springer, New York, (1990).

[2] M. C. Gutzwiller. *Chaos in Classical and Quantum Mechanics*, volume 1. Springer-Verlag, (1990).

[3] S. C. Farantos, Int. Rev. Phys. Chem **15**, 345–374 (1996).

[4] R. Prosmiti, and S. C. Farantos, J. Chem. Phys. **103**, 3299 (1995).

[5] A. Vegiri and S. C. Farantos, J. Chem. Phys. **98**, 4059 (1993).

[6] S. C. Farantos, S. Kapetanakis and A. Vegiri, J. Phys. Chem. **97**, 12158 (1993).

[7] R. E. Kunz and R. S. Berry, J. Chem. Phys. **103**, 1904 (1995).

[8] J. P. Rose and R. S. Berry, J. Chem. Phys. **98**, 3246 (1993).

[9] J. P. Rose and R. S. Berry, J. Chem. Phys. **96**, 517 (1992).

[10] F. H. Stillinger and T. A. Weber, Phys. Rev. A **25**, 978 (1982).

[11] W. H. Press, B. P. Flannery, S. A. Teukolsky and W. T. Vetterling. *Numerical Recipes.* Cambridge Univ. Press, (1986).

[12] G. Fanourgakis and S. C. Farantos, J. Phys. Chem. **100**, 3900 (1996).

[13] D. J. Chartrand, J. C. Shelley and R. J. Le Roy, J. Phys. Chem. **95**, 8310 (1991).

[14] C. J. Cerjan and W. H. Miller, J. Chem. Phys. **75**, 2800 (1981).

[15] J. Simons, P. Jorgensen, H. Taylor and J. Ozment, J. Phys. Chem. **87**, 2745 (1983).

[16] D. J. Wales, J. Chem. Phys. **91**, 7002 (1989).

[17] R. S. Berry, H. L. Davis and T. L. Beck, Chem. Phys. Letters **147**, 13 (1988).

[18] C. D. Maranas and C. A. Floudas, J. Chem. Phys. **97**, 7667 (1992).

[19] S. K. Gregurick, M. H. Alexander and B. Hartke, J. Chem. Phys. **104**, 2684 (1996).

[20]J. Nocedal. In *Theory of algorithms for unconstrained optimization*, edited by A. Iserles, pages 199–242. Cambridge University Press, (1992).

[21]M. Hestenes. *Conjugate direction methods in optimization.* Springer-Verlag, New York Inc., (1980).

[22]R. Fletcher. *Practical methods of optimization*, volume 1. J. Wiley and sons, New York Inc., (1980).

[23]J. Stoer, and R. Bulirsch. *Introduction to Numerical Analysis.* New York, Springer, (1980).

[24]A. R. Conn, N. I. M. Gould and P. L. Toint, Math Prog. **2**, 177 (1991).

[25]R. A. Aziz and M. J. Slaman, Mol. Phys. **58**, 679 (1986).

[26]H. Partridge, C. W. Bauschlicher Jr. and S. R. Langhoff, J. Phys. Chem. **96**, 5350 (1992).

[27]C. W. Bauschlicher Jr., H. Partridge and S. R. Langhoff, Chem. Phys. Letters **165**, 272 (1990).

[28]C. W. Bauschlicher Jr., H. Partridge and S. R. Langhoff, J. Chem. Phys. **91**, 4733 (1989).

[29]C. W. Bauschlicher Jr. and H. Partridge, Chem. Phys. Letters **239**, 241 (1995).

[30]G. Fanourgakis, S. C. Farantos, P. Parneix and P. Bréchignac, J. Chem. Phys. **submitted** (1996).

Method	Total time	Total time for the function calls	Number of function calls	Name of the function
BFGS(lnsrch)	17.54	12.75	667	potder
		0.21	1	min
BFGS(dbrent)	95.81	71.98	3644	potder
		0.15	1	min
Polak-Ribiere	49.81	37.07	1870	potder
		0.00	1	min
Quenching	140.29	105.22	5026	potder
		0.83	2409	Step

Method	Total time	Total time for the function calls	Number of function calls	Name of the function
BFGS(lnsrch)	14.74	10.79	557	potder
		0.26	1	min
BFGS(dbrent)	65.97	48.98	2528	potder
		0.07	1	min
Polak-Ribiere	42.26	31.41	1591	potder
		0.01	1	min
Quenching	112.00	84.64	3969	potder
		0.42	1895	Step

Method	Total time	Total time for the function calls	Number of function calls	Name of the function
BFGS(lnsrch)	14.76	11.11	558	potder
		0.28	1	min
BFGS(dbrent)	71.52	53.72	2727	potder
		0.07	1	min
Polak-Ribiere	51.25	38.35	1950	potder
		0.00	1	min
Quenching	138.56	104.41	5160	potder
		0.70	2444	Step

Method	**Total time**	**Total time for the function calls**	**Number of function calls**	**Name of the function**
BFGS(lnsrch)	14.01	10.38	542	potder
		0.26	1	min
BFGS(dbrent)	69.04	51.47	2651	potder
		0.08	1	min
Polak-Ribiere	35.26	26.62	1322	potder
		0.00	1	min
Quenching	69.86	52.51	2721	potder
		0.31	1304	Step

TABLE I Distribution of the cpu time in the most time consuming subroutines. The routine potder calculates the potential function and its derivatives. The min is the line–minimization routine and step a routine used in the integration of the classical equations of motion. dbrent and lnsrch are the names of line minimization routines taken from ref. [11].

Figure Legends

1. Convergence plots for the conjugate gradient methods of Polak–Ribiere (PR) and Fletcher–Reeves (FR) with a systematically (a) and a randomly (b) selected initial point (see text for details). Note the difference in the starting energy in the two cases and the logarithmic scale in the x-axis.

2. Convergence plots for the variable metric methods BFGS and DFP. The accurate dbrent routine for the location of the minimum of the potential function along a line is used (see text for details) in the figures (a) and (b), whereas the routine lnsrch is used in figures (c) and (d). Figures (a) and (c) are produced with a systematic initial point, and figures (b) and (d) with a random initial point.

3. Comparison of the Polak–Ribiere conjugate gradient method with the damped–BFGS variable metric method using the approximate line minimizer lnsrch (see text for details). Starting energies (a) -10.0 kcal/mol and (b) +37.0 kcal/mol.

4. The energies of the minima (filled squares) and transition states (open circles) found for the cluster $Mg^{+}Ar_{12}$, and arranged along the distance of magnesium cation from the center of mass of the cluster, R. $iso-(a), (R=0)$ has an icosahedral geometry with Mg^{+} in the center.

5. Stationary point diagram and isomerization reaction paths for moving from the icosahedral structure with the magnesium in the center, $iso-(a)$, to the other icosahedral structure found with the magnesium on the surface, $iso-(b)$.

6. Geometries of the minima and transition states for the isomerization path shown in Fig. 5 with the solid lines. M denotes a minimum and S a saddle. The saddle marked with an asterisk has the highest energy for the specific isomerization path studied.

7. Geometries of the minima and transition states for the isomerization path shown in Fig. 5 with the dashed lines. M denotes a minimum and S a saddle. The saddle marked with an asterisk has the highest energy for the specific isomerization path studied.