Modeling of the Potential Energy Surface of Regrouping Reaction in Collinear Three-Atom Collision System Using Nonlinear Optimization

A. S. Gevorkyan¹, A. V. Ghulyan², and A.R. Barseghyan²

¹Institute of Informatics And Automation Problems NAS of Armenia

²State Engineering University of Armenia

g_ashot@sci.am, garthur@arm.hpl.com,

ast_prog@yahoo.com

Abstract. A two-dimensional analytical model with a set of adjusting parameters [1] is proposed for an interaction potential of three-atom collinear system. Different numerical methods for obtaining the optimal set of parameters are considered in the present work. These parameters are used to approximate a two-dimensional numerical array (as obtained from another quantum-chemical *ab initio* calculations for reaction surfaces) with model potential to given accuracy. Appropriate methods of numerical simulation have been analyzed and the optimal one was selected. Based on this model an algorithm for numerical solution of the problem of finding the adjusting parameters has been realized using successive iterations. It was implemented for two specific cases discussed below.

1 Formulation of the Problem

The absence of a universal analytic representation of the interaction potential in the three-atom system impedes the development of a sufficiently universal package of programs for numerical calculation of bimolecular chemical reactions. Note that the numerical solutions are also highly important for deeper understanding of the fundamentals of quantum mechanics, i.e., the relation of classical nonintegrability to the quantum chaos [1-3]. A modeled potential of two-dimensional surface for reactions (in curvilinear coordinates) of the following type:

$$U(u, v) = U_1(u) \left[A(u) + (C(u) + G(u)v + a(u)v^2) \exp(-2\alpha(u)v) - B(u) \exp(-\beta(u)v) \right], \quad (1)$$

was proposed in [1] after analyzing the geometrical and topological features of different reaction surfaces, where the functions $U_1(u)$, $A(u)$, $C(u)$, $G(u)$, $a(u)$, $\alpha(u)$, $B(u)$ and $\beta(u)$ were

$$F(u) = C_{i-} + \frac{C_{i+} - C_{i-}}{1 + \exp(-2\gamma_{C_i} u)} + \frac{C_{i0} \gamma_{C_i}^2}{\left[\exp(\gamma_{C_i} u) + \exp(-\gamma_{C_i} u) \right]^2}, \quad i = 1, \dots, 8, \quad (2)$$

and C_{i-} , C_{i+} , C_{i0} , γ_{C_i} were some adjusting parameters.

The present paper aims at the determination of an optimal algorithm and development on its basis of a program for computing the adjusting parameters of potential (1) by means of comparative analysis of numerical arrays of reaction surfaces of bimolecular chemical reactions with the potential (1).

2 Numerical Methods of the Problem Solution

The expressions (1)–(2) are seen to comprise 32 adjusting parameters. To find these, below we analyze and apply different methods for solution of systems of nonlinear equations. The most frequently used of these are the relaxation method, the Newton method, and the modified Newton method [4,5]. All the mentioned methods are based on the procedure for inverse Jacobian matrix computation with respect to the adjusting parameters. Note that the Newton method is one of the most efficient methods for minimization of convex functions in problems of unconstrained optimization.

The Newton method is based on the substitution of a function with the first two terms of the Taylor series expansion and following minimization of the obtained quadratic form;

$$f(x + \delta) \approx f(x) + f'(x)\delta + \frac{f''(x)}{2}\delta^2 + \dots \quad (3)$$

For sufficiently small δ and well-behaved functions, we may assume that $f(x + \delta) = 0$ and obtain from expression (3)

$$\delta = -\frac{f(x)}{f'(x)} \quad (4)$$

However, all these methods are efficient, i.e., readily converging, in case of successful selection of the initial approximation that is located close to the root of equation. It follows from model (1) that although the given function can have numerous extremums, they may prove local, and this does not provide global convergence, i.e., does not give reliable results. Even cursory analysis shows that these methods are inapt for solving the tasks.

The methods considered in [6] next to the above ones, belong to the class of the so-called “global methods”, which provide global convergence almost from any initial point. Considered here were the Newton method with backtracking, as well as the Broyden method (quasi-Newton method). One of the distinctive features of the Broyden method is that for evaluation of Jacobian the final differences are used, which is inexpedient here because the Jacobian may be evaluated analytically. For that reason, in what follows the method of Newton with backtracking will be analyzed in detail.

Here follows a brief description of the method in question for a specific task. We shall try now to solve a system of transcendental equations with 32 unknowns (the adjusting parameters)

$$F\mathbf{x} = 0 \quad (5)$$

where $\mathbf{x}$ is the vector of adjusting parameters.

The Newton step is computed from

$$\mathbf{x}_{new} = \mathbf{x}_{old} + \delta \mathbf{x} \quad (6)$$

where

$$\delta \mathbf{x} = -\mathbf{J}^{-1} \cdot \mathbf{F} \quad (7)$$

where $\mathbf{J}^{-1}$ is Jakobian.

The problem is minimizing the following expression

$$\mathbf{f} = \frac{1}{2}(\mathbf{F} \cdot \mathbf{F}) \quad (8)$$

The searching strategy is simple enough; at first the full Newton step is executed, if after that the expression (8) is minimized, then the procedure of minimization is continued in the same direction, otherwise the traversal along the Newton direction is reversed.

The backtracking procedure is as follows;

Expression (6) can be written as

$$\mathbf{x}_{new} = \mathbf{x}_{old} + \lambda \cdot \mathbf{p} \quad (9)$$

where $\mathbf{p} = \delta \mathbf{x}$, $0 < \lambda \leq 1$.

The execution of Newton step is suggested in [6] when the following condition is observed:

$$f(\mathbf{x}_{new}) \leq f(\mathbf{x}_{old}) + \alpha \cdot \nabla \mathbf{f} \cdot (\mathbf{x}_{new} - \mathbf{x}_{old}) \quad (10)$$

where $0 < \alpha < 1$.

The searching strategy in the backtracking algorithm implies:

The determination of the following function

$$g(\lambda) = f(\mathbf{x}_{old} + \lambda \cdot \mathbf{p}), \quad (11)$$

so as

$$g'(\lambda) = \nabla \mathbf{f} \cdot \mathbf{p}. \quad (12)$$

After execution of the first Newton step, an analysis of minimization procedure of expression (10) is carried out. If it was not minimized, then the function $g(\lambda)$ is calculated from the following equation

$$g(\lambda) = [g(1) - g(0) - g'(0)]\lambda^2 + g'(0)\lambda + g(0). \quad (13)$$

The solution of equation (13) regarding λ is

$$\lambda = -\frac{g'(0)}{2[g(1) - g(0) - g'(0)]} \quad (14)$$

At successive iterations g is modeled to the cubic form that may be written as

$$g(\lambda) = a\lambda^3 + b\lambda^2 + g'(0) + g(0). \quad (15)$$

To find correct results for λ , it is necessary to solve two equations with respect to a and b ;

$$\begin{bmatrix} a \\ b \end{bmatrix} = \frac{1}{\lambda_1 - \lambda_2} \begin{bmatrix} 1/\lambda_1^2 & -1/\lambda_2^2 \\ -\lambda_2/\lambda_1^2 & \lambda_1/\lambda_2^2 \end{bmatrix} \bullet \begin{bmatrix} g(\lambda_1) - g'(0)\lambda_1 - g(0) \\ g(\lambda_2) - g'(0)\lambda_2 - g(0) \end{bmatrix} \quad (16)$$

The minimum for the given function is reached in the vicinity of the value

$$\lambda = \frac{-b + \sqrt{b^2 - 3ag'(0)}}{3a}, \quad (17)$$

So, one is made certain that this method may be applied for calculation of the adjusting parameters. The next is the class of methods of variable metrics [6]. Now we shall describe these methods in more detail. Note, that if function (1) is minimized along some direction $\mathbf{u}$, then its gradient is to be normal to the vector $\mathbf{u}$, otherwise it will imply that a nonzero derivative along the direction $\mathbf{u}$ is found.

Suppose that $\mathbf{x}_0$ is the initial approximation. Hence, a desired function can be approximated by means of decomposition

$$\begin{aligned} f(\mathbf{x}) &= f(\mathbf{x}_0) + \sum_i \frac{\partial f}{\partial x_i} x_i + \frac{1}{2} \sum_i \frac{\partial^2 f}{\partial x_i \partial x_j} x_i x_j + \dots \\ &\approx \mathbf{c} - \mathbf{b} \cdot \mathbf{x} + \frac{1}{2} \mathbf{x} \cdot \mathbf{A} \cdot \mathbf{x} \end{aligned} \quad (18)$$

where $\mathbf{c} = f(\mathbf{x}_0)$, $\mathbf{b} = -\nabla f|_{\mathbf{x}_0}$, $|A|_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j}$ is the Hessian matrix.

It is easy to calculate ∇f from (18)

$$\nabla f = \mathbf{A} \cdot \mathbf{x} - \mathbf{b}. \quad (19)$$

The change in the direction of gradient is calculated as follows:

$$\delta(\nabla f) = \mathbf{A} \cdot \delta \mathbf{x} \quad (20)$$

If the direction of motion is along $\mathbf{u}$, in case of which the function is minimized and it is required to continue the motion along $\mathbf{v}$, then the condition ought to be fulfilled

$$\mathbf{0} = \mathbf{u} \cdot \delta(\nabla f) = \mathbf{u} \cdot \mathbf{A} \cdot \mathbf{v}. \quad (21)$$

The main difference between the methods of conjugate gradients and of the variable metrics consists in the fact that in the latter case the intermediate results, which have been accumulated at previous stages, are stored and updated. The main feature of methods of variable metrics is an iterative approximation of the inverse matrix $\mathbf{A}^{-1}$ with given accuracy. The methods of the last class under study refer to methods based on the model-trust region approach. Here the least-square method, and in particular, the method of Levenberg-Marquardt will be considered. Let us determine the function

$$\chi^2 \approx \gamma - \mathbf{d} \cdot \mathbf{a} + \frac{1}{2} \mathbf{a} \cdot \mathbf{D} \cdot \mathbf{a}, \quad (22)$$

where $\mathbf{d}$ - is the vector of unknown quantities (the adjusting parameters M) and $\mathbf{D}$ is the matrix of $M \times M$ dimensions:

$$\mathbf{a}_{\min} = \mathbf{a}_{\text{cur}} + \mathbf{D}^{-1} [\nabla \chi^2(\mathbf{a}_{\text{cur}})]. \quad (23)$$

The next approximation can be determined in the following way;

$$\mathbf{a}_{\text{next}} = \mathbf{a}_{\text{cur}} - \text{constant} \cdot \nabla \chi^2(\mathbf{a}_{\text{cur}}) \quad (24)$$

Here the *constant* was taken to be rather small not to infringe the descending trend. In expressions (23) and (24) it is necessary to compute the Gessian of function χ^2 for all unknowns. The main difference from the previous class of methods is that the direct calculation of the Gessian by means of minimization methods is impossible. The minimization methods permit the iterative evaluation of a function and of its gradients. The second essential difference from all previous methods is that all the above methods are based on the linear search or linear minimization. The methods in issue are based on nonlinear procedures with the use of least-squares formalism. The calculation procedure of the gradient and Gessian in the Levenberg-Marquardt method is described in [6].

The model under consideration can be represented as

$$U = U(\mathbf{u}, \mathbf{v}, \mathbf{C}) \quad (25)$$

where $\mathbf{C}$ is the vector of adjusting parameters having the dimension M . The function χ^2 can be determined in the following way

$$\chi^2 = \sum_{i=1}^N \frac{U(u_i, v_i) - U(u_i, v_i, \mathbf{C})}{\sigma_i^2} \quad (26)$$

The gradient of function χ^2 is

$$\frac{\partial \chi^2}{\partial C_k} = -2 \sum_{i=1}^N \frac{[U_i - U(u_i, v_i, \mathbf{C})]}{\sigma_i^2} \cdot \frac{\partial U(u_i, v_i, \mathbf{C})}{\partial C_k} \quad (27)$$

The second derivative is computed as follows:

$$\begin{aligned} \frac{\partial^2 \chi^2}{\partial C_k \partial C_l} = & 2 \sum_{i=1}^N \frac{1}{\sigma_i^2} \left[\frac{\partial U(u_i, v_i, \mathbf{C})}{\partial C_k} * \frac{\partial U(u_i, v_i, \mathbf{C})}{\partial C_l} - \right. \\ & \left. - [U_i - U(u_i, v_i, \mathbf{C})] \frac{\partial^2 U(u_i, v_i, \mathbf{C})}{\partial C_l \partial C_k} \right] \end{aligned} \quad (28)$$

Here we introduce some designations

$$\beta_k = \frac{1}{2} \frac{\partial \chi^2}{\partial C_k} \quad \alpha_{kl} = \frac{1}{2} \frac{\partial^2 \chi^2}{\partial C_k \partial C_l} \quad (29)$$

Using expressions (23) and (29) we obtain the relation

$$\sum_{l=1}^M \alpha_{kl} \delta C_l = \beta_k \quad (30)$$

Taking into account (24) a system of linear equations is obtained

$$\delta_{a_l} = \text{constant} \cdot \beta_l. \quad (31)$$

The second partial derivatives are computed using the formula

$$\alpha_{kl} = \sum_{i=1}^N \frac{1}{\sigma_i^2} \left[\frac{\partial U(u_i, v_i, \mathbf{C})}{\partial C_k} \frac{\partial U(u_i, v_i, \mathbf{C})}{\partial C_l} \right] \quad (32)$$

So, here all necessary formulas for closed computation of adjusting parameters are presented. All these methods are implemented in the system of NIQC. The adjusting parameters of energy surfaces for a number of chemical reactions have been computed by means of the Newton method with backtracking and the Levenberg-Marquardt method.

3 Analysis of Results and Visualization of Data

The text below witnesses the effectiveness of the numerical method used. The numerical array, by means of which the energy surface is built (Fig.1), may be generated by substitution of a selected set of adjusting parameters to formula (1).

Fig. 1. Three-dimensional and two-dimensional representations of the surface by substitution of adjusting parameters in the model (1).

The three-dimensional surface of interaction potential obtained after the substitution of a selected set of adjusting parameters in interaction potential (1) is shown in Fig.1a. The two-dimensional projection of interaction potential (the isoline of potential energy), in the plane $\{u, v\}$, is shown in Fig.1b.

On the basis of this array the program in the system NIQC has been implemented. For computation of adjusting parameters the following scheme is proposed:

1. By means of an appropriate test function the numerical data array is generated into the file. The file contains the values of appropriate coordinates u, v potential computed by the model (1) or another model.
2. Then the first approximation for 32 adjusting parameters is introduced.
3. At this stage the array written in file at the step 1 is read onto respective arrays $(u, v$ and $U(u, v))$.
4. The invocation of minimization procedure realized on the basis of the Levenberg-Marquardt method is made.
5. The values of the first approximation for adjusting parameters and an appropriate array with 32 zeros and units, by means of which the flag is set to allow or prohibit the changes of parameter with the corresponding index, are fed to the input of above function. The quadratic divergences are set to the value of unit.
6. Then the procedure with appropriate parameters is called in loop. The expression (26) is the minimization function. The preset minimum value gives the convergence criterion.
7. At the output of function the array of adjusting parameters providing the preset minimum value is obtained and saved in the file.
8. At the final stage, the substitution of appropriate adjusting parameters is performed and the difference between the initial values of function and those computed according to the above procedure is calculated for corresponding u and v . All the results are saved in the file.

Fig. 2. In 2a) and 2b) the two-dimensional surface for initial approximation and its differential graph obtained by computing the difference between the values of initial test function and those computed by model (1) based on the first approximation are shown respectively.

Fig. 3. In 3a) and 3b) the two-dimensional surface for initial approximation and its differential graph obtained by computing the difference between the values of initial test function and those computed by model (1) based on the first approximation are shown respectively.

Fig. 4. In 4a) and 4b) the two-dimensional surface for initial approximation and its differential graph obtained by computing the difference between the values of initial test function and those computed by model (1) based on the first approximation are shown respectively.

In Fig.5 the minimization process of expression (26) is illustrated.

The minimization process is illustrated by the graphs. In Figs. 2a) and 2b) the two-dimensional surface for initial approximation and its differential graph obtained by computing the difference between the values of initial test function and those computed by model (1) based on the first approximation are shown respectively. Figs. 3a) and 3b) contain the same graphs after 10 iterations, and, finally, Figs. 4a) and 4b) show the same graphs after the convergence of the process.

The following experiment was carried out with test function in the form

$$U(u, v) = \frac{\sin \sqrt{(u^2 + v^2)}}{\sqrt{(u^2 + v^2)}}. \quad (33)$$

For appropriate definition ranges of u and v the following three-dimensional 6a) and two-dimensional 6b) surfaces were obtained. Then, appropriate adjusting parameters of model (1) for the surface given by expression (33) were obtained. On the basis of these adjusting parameters and model (1) the numerical array has been generated, its difference graph found from the difference between the initial test function value and model (1) based on the initial approximation (Fig.6c). In Fig. 6d) the minimization process of expression (26) is shown, where the expression (33) is used as the test function.

Fig. 5. The minimization function according to expression (26) versus the number of iterations.

Fig. 6. For appropriate definition ranges of u and v the following three-dimensional 6a) and two-dimensional 6b) surfaces were obtained. Then, appropriate adjusting parameters of model (1) for the surface given by expression (33) were obtained. On the basis of these adjusting parameters and model (1) the numerical array has been generated, its difference graph found from the difference between the initial test function value and model (1) based on the initial approximation (Fig.6c). In Fig. 6d) the minimization process of expression (26) is shown, where the expression (33) is used as the test function.

4 Conclusion

The analysis of the analytical model (1) shows that:

1. It is proved that the energy surfaces of bimolecular reactions with different topological and geometrical features can be described by means of potential (1) with sets of adjusting parameters.

2. Some classes of computational methods and algorithms, used for finding adjusting parameters of model (1), have been analyzed in details. It is shown that the method most acceptable for solution of the task is the Levenberg-Marquardt method. The results obtained (Fig. 5 and Fig. 6d) show that with the help of this method the expression (26) is minimized rather quickly and a good approximation of function (1) is provided.

So, a numerical experiment proves the efficiency of the description of the two-dimensional reaction surface of various bimolecular chemical reactions by means of analytical potential (1).

Acknowledgement.

A. Gevorkyan thanks ISTC Grant A-823 for support and also gratefully acknowledges Prof. Yu. Shoukourian and Dr. V. Sahakyan for their kindness to this work.

References

1. Gevorkyan A.S. Dissertation of the Dr.Sci., 2000 St.Petersburg(S.Un.St-P.)
2. Bogdanov A.V., Gevorkyan A. S. "Three-body multichannel scattering as a model of irreversible quantum mechanics" Symposium of Non-Linear Theory and Its applications. Hilton Hawaiian Village, 1997.
3. A.V.Bogdanov, A.S.Gevorkyan, A.G.Grigoryan Internal Time Peculiarities as a Cause of Bifurcations Arising in Classical Trajectory Problem and Quantum Chaos Creation in Three-body System, AMS/IP Studies in Advanced Mathematics, V.13. p. 69-80, (1999).
4. Samarskiy A.A., "Numerical methods", 1997., Moscow, Nauka
5. Bakhvalov N.S., "Introduction to the numerical methods". Moscow, Nauka 1987.
6. William H. Press "Numerical Recipes in C" Cambridge University Press 1992.