LOCATION OF STATIONARY POINTS OF POTENTIAL ENERGY SURFACES II. A NEW PROCEDURE FOR FINDING SADDLE POINTS

MIKLÓS I. BÁN

Institute of Physical Chemistry, Attila József University

P.O.Box 105, H–6701 Szeged, Hungary

(Received November 12, 1989)

A NEW, QUASI—NEWTON—TYPE PROCEDURE HAS BEEN DEVELOPED FOR FINDING SADDLE POINTS. BY ASCRIBING THE INDEFINITE OPTIMIZATION PROBLEM TO A POSITIVE DEFINITE ONE THE HIGH NUMERICAL EFFICIENCY CHARACTERISTIC OF THIS CLASS OF ALGORITHMS (IN MINIMIZATIONS) HAS BEEN RETAINED WHILE THE INHERENT PROBLEMS OCCURING IN INDEFINITE OPTIMIZATIONS COULD BE AVOIDED.

Introduction

Theoretical (and in many respects also practical) investigations of chemical reactions require some knowledge of the geometrical features of the Born–Oppenheimer potential energy surfaces.

Among the many geometrical characteristics of the potential surface, the most important ones are the minima, the saddle points and the reaction path. The minima correspond to stable molecular configurations and are characterized by the matrix of second derivatives (Hessian) having only positive eigenvalues. The saddle point of the potential surface is a critical point where the Hessian also has one negative eigenvalue[1]. The reaction path is the unique curve of steepest descent, which starts at the saddle point and goes to the minima. The saddle point corresponds to the transition state of the chemical reaction.

[1]The existence of only one negative eigenvalue (first order saddle point) has only chemical reasons [22], *i.e.* it is not a mathematically motivated restriction.

and represents the maximum of the reaction path.

The main difference between the problems of finding a minimum and a saddle point is significant. Whereas the knowledge of the slope of the surface is enough to reach a minimum, to find a saddle point the curvature of the surface is also needed, the computational work thereby increasing substantially.

In a short survey on the multitude of methods used to determine saddle points, the pioneering work of McIVER and KOMORNICKI [1,2] should be mentioned first. Their algorithm is based on minimization of the square of gradient norm. Since their procedure is theoretically well established, it is widely used in spite of the relatively high numerical cost and the possible collapse to a minimum/maximum. During the past two decades, many other techniques have appeared for finding the transition state [3–8], together with excellent surface walking algorithms [9–13] which are suitable for controlling of the walking process to any kind of stationary points. Quite recently the quasi–Newton methods have begun to be used to locate saddle points [14–19]. Their power in finding minima justifies the significant efforts that have been made to solve the problems which arise when these algorithms are used for indefinite optimizations.

The aim of the present work was to establish a method which does not require the evaluation of the full matrix of second derivatives, but furnishes a high computational efficiency. This intention has been realized in a quasi–Newton–type algorithm using the first derivative of the function and a one–dimensional projection of the Hessian.

The paper is structured in several sections. In Section 1 the frame of quasi–Newton algorithms will be discussed briefly. Section 2A contains the basic idea, while 2B presents the general development of the method. Section 3 discusses the main properties of the algorithm, and Section 4 is devoted to the numerical illustrations. The Appendix gives a short account of the relation of the present method and the BELL-CRIGHTON-FLETCHER [17] algorithm.

1.) The quasi-Newton procedure

To outline the main properties of a quasi–Newton algorithm, a quadratic function will be considered:

$$f(x) = f(x_o) + g_o^t p + \frac{1}{2} p^t G p, \tag{1.1}$$

where f is a scalar, x_o and x are position vectors, g_o ($\equiv g(x_o)$) is the gradient, p ($\equiv dx$) is the change of coordinates, and G is the Hessian[2].

The gradient at point x is given by

$$g(x) \equiv g = g_o + Gp. \tag{1.2}$$

If a symmetric, positive definite H matrix is given (playing the role of an approximation to the inverse of the Hessian), the successive steps (from the k^{th} one) will be generated as follows [21,24–26].

a.) Generation of the step direction:

$$d_k = -H_k g_k, \quad (g^t d_k < 0) \tag{1.3}$$

b.) Estimation of the step-length:

$$x_{k+1} = x_k + \alpha_k d_k, \tag{1.4}$$

$$\alpha_k = \{ \alpha_k \,/\, g^t(x_k + \alpha_k d_k) d_k \equiv g_{k+1}^t d_k = 0 \} = -g_k^t d_k / d_k^t G d_k. \tag{1.5}$$

The parameter α_k is determined by the requirement that the acceptable point (x_{k+1}) must be a local minimum, implying that the gradient (g_{k+1}) has no component along the

[2]Small letters denote scalars and vectors, capital letters matrices, and superscript t refers to transposition.

step–vector.

c.) Evaluation of quantities which can be derived from primary data and *characterize the curvature of the surface:*

$$p_k = x_{k+1} - x_k = \alpha_k d_k, \tag{1.6}$$

$$\gamma_k = g_{k+1} - g_k = G p_k \tag{1.7}$$

(on a quadratic surface, g_{k+1} is given by (1.2)).

d.) The previously evaluated quantities are used to *update the H matrix* by absorbing the information collected on the curvature of the surface. This means that a correction term is added to H_k. Taking the DFP procedure [21,24–26] as an example:

$$H_{k+1}^{\gamma} = H_k \frac{p_k p_k^t}{\gamma_k^t p_k} - \frac{(H_k \gamma_k)(\gamma_k^t H_k)}{\gamma_k^t H_k \gamma_k}. \tag{1.8}$$

The correction ensures a perfect description of the portion of the surface explored by the actual step. This point distinguishes mainly the various quasi–Newton algorithms.

e.) After checking results *stop or return to step a.*

Some important characteristics of the whole quasi–Newton family are the following [21,24–26]:

- one dimensional searches are used successively;
- only the function values and gradients relating to the actual and immediate previous points are used;
- the location of the stationary point of a quadratic function requires (at most) as many steps as the number (n) of independent variables (quadratic convergence property).

For an indefinite matrix, non–zero (singular) vector(s) exist(s), satisfying the relation:

$$d_1^t G d_1 = 0. \tag{1.9}$$

If a quasi–Newton algorithm is applied to an indefinite objective function, such a singular vector may cause the total breakdown of the algorithm. Neither the step–length nor the update (1.8) of the H matrix can be defined [18].

When a native quasi–Newton algorithm is used, the presence of a negative eigenvalue prevents the approximation of the saddle point by an incorrect choice [11] of the search direction, even in the absence of a singular direction of search.

The existing techniques [14–18] developed for handling such functions explicitly modify a parent quasi–Newton method for adaption to identify optimization problems.

The framework of this very significant class of algorithms has been presented briefly in order to show the most important problems which occur when these methods are applied to indefinite objective functions.

In the following sections the theoretical background and technical aspects of the present algorithm will be discussed.

2.) Definition of the method

A. Qualitative considerations

The idea of the present algorithm was born from the intention to exploit the full computational efficiency of the quasi–Newton methods by avoiding somehow the problems of their application to indefinite functions.

Instead of modifying the algorithm itself, it seemed more attractive to define a new (associated) surface, which is convex and contacts the original surface at its saddle point.

This means that (in the sense of the idea) the two surfaces are in a first–order touching contact characterized by identical function values and first derivatives, but their

second derivatives differ in nature. While the critical point of the original surface is a saddle point, the "associated convex surface" has a minimum with the same coordinates.

The introduction of an associated convex surface would allow the application of quasi–Newton methods to find the minimum, which is located at the same point as the saddle point of the original surface .

B. The family of "associated convex surfaces"

Let the quadratic surface be determined by (1.1) and let its Hessian have one negative eigenvalue ("saddle surface"). Because of the quadratic nature of the function, the surface is unambiguously defined if the coordinates of the stationary point and its Hessian are known.

Let the eigenvectors and eigenvalues of G be

$$\{ u_i \} , \{ \lambda_i \} ; \qquad i \in [1,n] \tag{2.1}$$

and let a set of independent vectors be given:

$$\{ z_i \} , \qquad i \in [1,n] \tag{2.2}$$

which satisfy the following inequalities:

$$z_1^t \, G \, z_1 < 0, \tag{2.3}$$

$$z_j^t \, G \, z_j > 0 ; \qquad j \in [2,n] \tag{2.4}$$

and

$$z_i^t \, G \, z_j = z_i^t \, w_j = 0 ; \qquad \forall (i \neq j). \tag{2.5}$$

Relations (2.3) and (2.4) account for the indefiniteness of G, while (2.5) expresses the conjugated nature of the vectors (2.2) and the biorthogonality[3] of the dual vector pairs:

$$\{ z_i \} \text{ and } \{ w_i = Gz_i \} ; \qquad i \in [1,n]. \tag{2.6}$$

With the help of the specially chosen vector systems (2.1) and (2.2), projector matrices will first be constructed:

$$P_n^{(u)} = \overline{P}_n^{(u)} = \sum_{i=1}^{n} u_i u_i^t, \tag{2.7}$$

$$P_n^{(z)} = \sum_{i=1}^{n} \frac{w_i z_i^t}{w_i^t z_i}, \tag{2.8}$$

$$\overline{P}_n^{(z)} = \sum_{i=1}^{n} \frac{w_i z_i^t}{z_i^t w_i}, \tag{2.9}$$

where $P_n^{(u)}$ and $\overline{P}_n^{(z)}$ denote adjoint matrices. The projectors are of rank n, and yield the identity:

$$x = P_n^{(u)} x = P_n^{(z)} x = \overline{P}_n^{(z)} x \tag{2.10}$$

and the general projector property:

$$P_n^{(u)} = P_n^{(u)^2} ; \qquad P_n^{(z)} = P_n^{(z)^2} ; \qquad \overline{P}_n^{(z)} = \overline{P}_n^{(z)^2}. \tag{2.11}$$

(In general, our statements will be formulated with the help of the (orthogonal) eigenvectors (2.1) and the biorthogonal vector systems (2.6).)

Through the use of (2.7)–(2.9) the Hessian and its inverse will be represented in

[3]The upper/lower indices for contravariant/covariant vectors will not be used.

factorized forms:

$$G = P_n^{(z)} G\, P_n^{(z)} = \sum_{i=1}^{n} \frac{w_i w_i^t}{z_i^t w_i} = P_n^{(u)} G\, P_n^{(u)} = \sum_{i=1}^{n} \lambda_i u_i u_i^t, \tag{2.12}$$

$$G^{-1} = P_n^{(z)} G^{-1} P_n^{(z)} = \sum_{i=1}^{n} \frac{z_i z_i^t}{w_i^t z_i} = P_n^{(u)} G^{-n}\, P_n^{(u)} = \sum_{i=1}^{n} \lambda_i^{-1}\, u_i u_i^t. \tag{2.13}$$

The sign of the individual terms in (2.12) and (2.13) depend on the signs of the diagonal elements

$$\left\{\lambda_i\right\}; \qquad \left\{z_i^t\, w_i = z_i^t G z_i = w_i^t G^{-1} w_i\right\} \tag{2.14}$$

of the Hessians transformed by the (congruence) transformations

$$\left[U^t\right] G \left[U\right]; \qquad \left[Z^t\right] G \left[Z\right]; \qquad \left[W^t\right] G^{-1} \left[W\right] \tag{2.15}$$

into new (orthogonal and non–orthogonal) vector systems. ([U], [Z] and [W] are matrices whose columns are vectors { u_i }, { z_i }, { w_i } ; i$\in$ [1,n].) Since the number of diagonal elements having negative, zero and positive signs is conserved by the congruence transformation (SYLVESTER inertia theorem [23]). The number of terms with a given sign in (2.12) and (2.13) is equal to the number of eigenvalues with the same sign.

As our future aim is to manipulate the signs of the individual terms in (2.12) and (2.13), the (reflector) matrices

$$B^{(u)} = I - 2\, u_1 u_1^t, \tag{2.16}$$

$$B^{(z)} = I - 2\, \frac{w_1 z_1^t}{z_1^t w_1} \tag{2.17}$$

will be introduced (I is the identity matrix[4]).These matrices change the sign of that component of a vector which is parallel to u_1 or z_1.

With the help of (2.16) and (2.17) new matrices (of second derivatives) will be constructed by evaluating the following products:

$$B^{(u)}G = -\lambda_1 u_1 u_1^t + \sum_{j=2}^{n} \lambda_j u_j u_j^t, \tag{2.18}$$

$$B^{(z)}G = -\frac{w_1 w_1^t}{z_1^t w_1} + \sum_{j=2}^{n} \frac{w_j w_j^t}{z_j^t w_j}. \tag{2.19}$$

(It has been assumed that the Hessian has one negative eigenvalue ordered on the first place.) The negative signs (introduced only into the first terms) ensure the positive definiteness:

$$x^t(B^{(u)}G)x > 0; \qquad \forall\, t \in R^n, \tag{2.20}$$

$$x^t(B^{(z)}G)x > 0; \qquad \forall\, t \in R^n. \tag{2.21}$$

The first Hessian (2.18) differs from G in the reversion of the sign of the first eigenvalue. It therefore determines a convex surface the main curvature of which has been the same absolute values as G. The second new Hessian (2.19) represents a somewhat different case. To show this, the effect of multiplication by the projectors (from the left and right in (2.12) and (2.13)) has to be investigated. This is a two–step procedure, consisting of a congruence transformation (2.15) converting G into a non–orthogonal basis, and a back–transformation by the inverse matrices. The congruence transformation into a non–orthogonal basis changes the values of the diagonal elements. The multiplication by (2.17) reverses the sign of the first term, and therefore the back–transformation provides a

[4]They have also the following property:

$$I = B^{(u)^2} = B^{(z)^2}.$$

new Hessian (2.19) having main curvatures[5] (*even in absolute values*) other than G.

By this point, two Hessians (2.18) and (2.19) have been constructed, both of them determining convex surfaces. The first matrix (2.18) requires a knowledge of the eigenvector relating to the (one) negative eigenvalue and determines a surface of fully identical shape in the subspace of eigenvectors relating to the positive eigenvalues. The second matrix (2.19) requires exclusively a knowledge of a vector (z_1) satisfying the relation

$$z_1^t \, G \, z_1 = z_1^t \, w_1 < 0 \tag{2.22}$$

and determines a surface differing from G in its curvature properties, in the "positive subspace" too. Since the eigenvectors are usually not known, only this last case is of practical significance. This last surface will henceforward be referred to as the "associated convex surface".

Although the associated convex surface has different main curvatures (also in absolute values) from those of the original surface (determined by G), the position of the critical point (where the gradient is the zero–vector) remains unaffected. This implies that the critical points of the two surfaces are common first–order touching points, *i.e.* their function values and first derivatives are equal:

$$f(x_{cr}) = \tilde{f}(x_{cr}), \tag{2.23}$$

$$g(x_{cr}) = \tilde{g}(x_{cr}) = 0, \tag{2.24}$$

where x_{cr} is the coordinate of the critical point, and the designation $\sim$ will denote quantities defined on the convex surface.

[5]The difference between the convex surfaces determined by $B^{(u)}G$ (2.18) and $B^{(z)}G$ (2.19) depends on the non–orthogonality of the basis vectors $\{z_i\}$.

The associated convex surface depends on the choice of the vector (z_1) fulfilling the condition (2.22). Since an infinite number of such suitable vectors exists, an infinite numbers of surfaces contact at the critical point constitute the "family of associated convex surfaces".

The coincidence of the critical points (relating to the saddle–surface and the associated surface) allows a search for a minimum instead of a saddle point of identical location. This fact is of particular importance in establishing the new algorithm.

3.) The algorithm

The frame of the algorithm strictly follows the ideas described previously. Only the main properties will be discussed in a form which is rather 'conceptual' regarding that a detailed and neat program description will be presented (as a manual) elsewhere[6].

A. Generation of the convex surface

To show clearly the mechanism of the algorithm, a quadratic function will be considered with a Hessian having one negative eigenvalue[7].

There are two explicit requirements:

– an objective function to have a form which allows the evaluation of its first derivative is assumed,

– the existence of a vector (z_1) which fulfills condition (2.22) is assumed.

A suitable guess for z_1 is usually the line–segment connecting the minima at the "endpoints" of the reaction path (or an initial Hessian can provide a good choice). The vector w_1 will first be evaluated (by a finite difference):

$$w_1 = g_2 g_1 = G z_1 \tag{3.1}$$

[6]The FORTRAN code of the program is available, and it will also be submitted for distribution to the QCPE. On request copies will be sent by the author.

[7]The extension to higher–order saddle points (which have more than one negative eigenvalues) involves no new theoretical problem.

following afterwards the construction of (2.17).

If the reference frame is fixed to the stationary point, the gradient (at point x) of the saddle–function is given by

$$g(x) = G\,x \tag{3.2}$$

and the gradient of the associated convex function by

$$\tilde{g}(x) = B^{(z)}G\,x = B^{(z)}g(x). \tag{3.3}$$

The function value (at any general point) of the associated convex function can not be drawn from the data relating to the saddle–surface. The information on the associated convex function is therefore restricted (at any general point) to its derivative(s). This means that (in the course of searching) the associated convex surface is given only partially, without its function value. A full description of the surface becomes possible *a posteriori*, if the function value at the stationary point fixes the position of the graph of the function. Fortunately, the function value (of the saddle–function) is not necessary at all for the perfect working of the algorithm, and this substantially reduces the computational work.

B. The linear search

As the function value of the associated convex function is lacking, the line–search must be made without any reference to it. The point acceptable as a local stationary point has to be determined by the condition (see (1.5)):

$$\tilde{g}_{k+1}^{t}\,\tilde{d}_k = \epsilon \approx 0. \tag{3.4}$$

The important pecularities of the performance of the line–search do not influence its

structure and therefore they will not be discussed here.

C. Updating of the matrix–H

Again, an example (see (1.8)) by the DFP [21] procedure:

$$\tilde{H}_{k+1} = \tilde{H}_k + \frac{\tilde{p}_k \tilde{p}_k^t}{\tilde{\gamma}_k^t \tilde{p}_k} - \frac{(\tilde{H}_k \tilde{\gamma}_k)(\tilde{\gamma}_k^t \tilde{H}_k)}{\tilde{\gamma}_k^t \tilde{H}_k \tilde{\gamma}_k}, \tag{3.5}$$

$$\tilde{\gamma}_k = \tilde{g}_{k+1} - \tilde{g}_k, \tag{3.6}$$

$$\tilde{p}_k = x_{k+1} - x_k = \tilde{\alpha}_k \tilde{d}_k. \tag{3.7}$$

Although the formula (3.5) contains quantities relating to the associated convex surface the strict formal identity with the parent update scheme is obvious.

If the above conceptual modifications are introduced into a quasi–Newton algorithm, the new algorithm will be able to locate saddle points. Without going into details, the separate steps will be formally the same as in Section 1, but the designated quantities

$$\tilde{d}_k, \tilde{p}_k, \tilde{\alpha}_k, \tilde{g}_k, \tilde{\gamma}_k, \tilde{H}_k \tag{3.8}$$

in relation to the associated convex surface will appear everywhere.

The whole discussion so far referred to a quadratic objective function having a Hessian with one negative eigenvalue. For general functions, the structure of the algorithm remains the same as already defined, but minor changes and restrictions are indispensable to ensure the necessary conditions.

For non–quadratic functions, the Hessian has only a local meaning sense, so the reflector matrix will also be locally defined:

$$B^{(z)}(x) = I - 2\frac{w_j(x)z_j^t}{z_j^t w_j(x)} \tag{3.9}$$

and

$$\tilde{g}(x) = B^{(z)}(x)\, g(x). \tag{3.10}$$

All other quantities differ from those in the quadratic case by the appearance of (3.10). The vector z_j is constant but $w_j(x)$ will be reevaluated at any point. In consequence of the local character of the quantities the algorithm is defined in a region where the condition

$$z_j^t w_j(x) < 0 \tag{3.11}$$

is fulfilled.

A continuous approach to the stationary point is ensured only in that area where the Hessian has the assumed one negative eigenvalue.

To reach higher–order saddle points separate matrices (3.9) are needed for each independent direction of negative curvature.

The properties peculiar to the present quasi–Newton–type procedure have been discussed while other characteristics conform to the general behaviour of the whole quasi–Newton family.

To summarize the leading aspects:

i) The algorithm is defined in a region where a constant vector exists, pointing in the direction of negative curvature. It is assumed implicitely that the Hessian has one negative eigenvalue.

ii) On construction of the reflector matrix (3.9) the gradient of an "associated convex surface" can be obtained through (3.10). The convex surface is only partially defined by its derivative(s) and evaluation of the function value of the original (saddle–) function is not necessary.

iii) Through the use of the quantities (3.8) relating to the associated convex surface, any quasi–Newton algorithm (with the minor technical changes mentioned) will provide a minimum at the same coordinates as those of the saddle point of the original surface.

4.) Applications

The algorithm has been tested on various model surfaces and quantum chemical problems. Although the BFGS [21] and DFP [21] methods can equally be used, the examples will refer to the DFP algorithm.

A. The Adams model–potential [12]

$$E(x_1,x_2) = 2x_1^2(4 - x_1) + x_2^2(4 + x_2) - x_1x_2(6 - 17e^{-r^{0.5}}) \qquad (4.1)$$

has a minimum and two saddle points. The walk from the point (1.8,–0.2) to one of the saddle points is displayed in Table I.

The following quantum chemical examples have been chosen to test the efficiency of the algorithm in comparison with other procedures. BAKER's paper [27] presents convergence data of several rearrangement reactions. These computations have been done by the SIMON's algorithm [12] which was implemented and incorporated into the standard GAUSSIAN 82 program package. The tests of the present algorithm are relating to

Table I

	x_1	x_2	RMS gradient
1.	1.8000	–0.2000	11.5258
2.	1.9529	0.6489	3.9571
3.	2.2613	0.4625	0.5322
4.	2.2408	0.4415	0.4146E-02
5.	2.2410	0.4411	0.2236E-04
6.	2.2410	0.4411	0.4909E-14

CNDO/2 calculations. For the sake of comparison, the starting geometrical parameters were chosen to the same relative value with respect to the optimized parameters, as in the reference examples. The accuracy of the computations was also identical (RMS gradient 0.0003). The direction of negative curvature (z_1) was estimated for the vector pointing from the one minimum to the other.

It has to be mentioned, that SIMON's algorithm requires an analytically evaluated starting Hessian, while the present method requires only the direction/(and value) of the negative curvature.

B. The HCN ⟶ CNH reaction[8]

PM	G82	POP	SIM
7(12)	10(?)	9(?)	8(?)

(Final geometry: r(CN) = 1.221 Å, r(CH) = 1.121 Å, φ(HCN) = 87.14°)

C. The FCN ⟶ CNF reaction[9]

In BAKER's example [27] the start position was chosen very near to the minimum, and the Hessian has here only positive eigenvalues. This position is outside the region where the conditions of the present algorithm are fulfilled. In this case the midpoint of the line segment connecting the minima was used as a starting configuration

[8]The abbreviations used have the following meaning:
PM = present method
G82 = method of the standard GAUSSIAN 82 program package
POP = POPPINGER's method [3]
SIM = SIMON's method [12]
y(x) = y steps and x gradient evaluations

[9]see footnote 8

PM	G82	POP	SIM
8(13)	failed	15(?)	13(?)

(Final geometry: r(CN) = 1.263 Å, r(CF) = 1.333 Å, φ(FCN) = 93.30°)

The program system realizing the above ideas is able to locate also minima and points of (so called) gradient extremal curves [20,28]. Along a curve of this type, it is always possible to reach that region where the Hessian already has the required one negative eigenvalue.

D. The optimization of CH_3F[10]

The optimization of the structure of methyl fluoride is presented to show the working of the algorithm in minimizations. The efficiency of the program system is due to some particular modifications of the linear searching and updating steps respectively.

PM	G82	SIM
8(18)	11(?)	11(?)

(Final geometry: r(CF) = 1.344 Å, r(CH) = 1.118 Å, φ(FCH) = 109.30°)

Appendix

Under certain conditions the BELL–CHRIGHTON–FLETCHER (BCF) algorithm [17] and the present one generate an identical succession of points.

The procedure [17] is based on a separation of the n–dimensional space into a one–dimensional and (conjugated) (n–1)–dimensional subspaces. The functions will be

[10]see footnote 8

maximized in the one–dimensional subspace, while it will be minimized in the conjugated subspace. For general functions these steps will be repeated cyclically.

It will be shown that for a quadratic function (see (1.1)) identical steps will be produced by both algorithms, if the first step of the present procedure is a pure maximization. To illustrate the reason for this, those steps will be examined which are generated by the algorithms considered.

From the start point x, along the search direction z_i (see relations (2.2)–(2.6)) the BCF algorithm yields a step of length α_i:

$$0 = z_i^t g(x + \alpha_i z_i) = z_i^t g(x) + \alpha_i z_i^t G z_i, \tag{A.1}$$

$$\alpha_i = -\frac{z_i^t g(x)}{z_i^t G z_i}. \tag{A.2}$$

The first step is taken uphill in the BCF algorithm (*i.e.* $z_i \equiv z_1$), therefore

$$z_1^t g(x) > 0. \tag{A.3}$$

The vector z_1 satisfies condition (2.3) implying the relation:

$$\alpha_1 > 0. \tag{A.4}$$

The present algorithm determines a step of length from the same start point x, along the search direction c_i:

$$0 = c_i^t B^{(z)} g(x + \beta_i c_i) = c_i^t B^{(z)} g(x) + \beta_i c_i^t B^{(z)} G c_i, \tag{A.5}$$

$$\beta_i = - \frac{c_i^t B^{(z)} g(x)}{c_i^t B^{(z)} G c_i} \tag{A.6}$$

($B^{(z)}$ has been defined by (2.17).)

If the first step of the present procedure fulfills the identity

$$c_1 \equiv z_1 \tag{A.7}$$

then, because of the equality

$$B^{(z)} w_1 = - w_1, \tag{A.8}$$

the following relations are valid simultaneously:

$$c_1^t B^{(z)} g(x) < 0, \tag{A.9}$$

$$| c_1^t B^{(z)} g(x) | = | z_1^t g(x) | \tag{A.10}$$

and

$$c_1^t B^{(z)} G c_1 > 0, \tag{A.11}$$

$$| c_1^t B^{(z)} G c_1 | = | z_1^t G z_1 | \tag{A.12}$$

implying the identity

$$\beta_1 \equiv \alpha_1. \tag{A.13}$$

For those directions of search which are in the subspace conjugated to z_1, the matrix $B^{(z)}$ has no effect at all. The steps in the "positive subspace" are therefore identical.

While the above separation of the optimization problem (into pure maximization and pure minimizations) is an inherent basic assumption in the BCF algorithm, it is a special possible case in the present procedure.

Acknowledgement

In writing up this paper thanks are due to Dr. Imre Bálint for the consent of using many results of his work, consultations and advice during the period working in my group between January 1985 and July 1987. The support of this work is acknowledged to the Hungarian Ministry of Higher Education (grant Nos. P–437/84. and 746/86).

References

[1] *McIver,J.W., Komornicki,A:* J.Am.Chem.Soc.84,2625 (1972).

[2] *Komornicki,A., K.Morokuma, K.Ishida, R.Ditchfield, M.Conrad:* Chem.Phys.Lett.45,595 (1977).

[3] *Poppinger,D.:* Chem.Phys.Lett., 35, 550 (1975).

[4] *Halgren,T.A., W.N.Lipscomb:*Chem.Phys.Lett.43, 225 (1977).

[5] *Müller,K., L.D.Brown:* Theor.Chim.Acta 53,75 (1978).

[6] *Schlegel,H.B.:* J.Comp.Chem. 3,214 (1982).

[7] *Rothman,M.J., L.L.Lohr:* Chem.Phys.Lett. 70, 405 (1980).

[8] *Kato,S., K.Morokuma:* J.Chem.Phys. 73, 3800 (1980).

[9] *Crippen,G.M., H.A.Scheraga:* Arch.Biochem.Biophys. 144, 462 (1971).

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

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

[12] *Banarjee,A., N.Adams, J.Simons, R.Shepard:* J.Phys.Chem. 89, 52 (1985)

[13] *Taylor,H., J.Simons:* J.Phys.Chem. 89, 684 (1985).

[14] *Pulay,P., G.Fogarasi, F.Pang, J.E.Boggs:* J.Am.Chem.Soc. 101, 2550 (1979).

[15] *Sinclair,J.E., R.Fletcher:* J.Phys. C7, 864 (1974).

[16] *Fletcher,R.:* Conj.Grad.Meth.for Ind.Syst.;Proc.of the 1975 Dundee Num.Anal.Conf.; Springer–Verlag, Berlin, p73–89 (1975).

[17] *Bell,S., S.Crighton, R.Fletcher:* Chem.Phys.Lett. 82, 122 (1981).

[18] *Oren,S.S.:* J.Opt.Theor.Appl. 43, 167 (1984).

[19] *Bálint,I., M.Bán:* Theor.Chim.Acta 63, 255 (1983).

[20] *Bálint,I., M.Bán:* Int.J.Quant.Chem. 25, 667 (1984).

[21] *Fletcher,R.:* Practical Methods of Optimization;Vol.1, Wiley, New York, 1980.

[22] *Murrell,J.N., K.J.Laidler:* Trans.Far.Soc. 64, 371 (1978).

[23] *Parlett,B.N.:* The Symmetric Eigenvalue Problem, Prentice–Hall Ser.in Comp.Math., Englewood Cliffs, 1980.

[24] *Gill,P.E., W.Murray, M.H.Wright:* Practical Optimization, Academic Press, New York, 1981.

[25] *Powell,M.J.D. (ed):* Non–linear Optimization, Academic Press, New York, 1982.

[26] *Scales,L.E.:* Introduction to Non–linear Optimization, Macmillan, Basingstoke, 1985.

[27] *Baker,J.:* J.Comp.Chem. 7, 385 (1986).

[28] *Hoffmann,D.K.,* R.S.*Nord, K.Ruedenberg:* Theoret. Chim. Acta 69, 265 (1986).

РАСПОЛОЖЕНИЕ СТАЦИОНАРНЫХ ТОЧЕК ПОВЕРХНОСТЕЙ ПОТЕНЦИАЛЬНОЙ ЭНЕРГИИ II. НОВЫЙ МЕТОД ДЛЯ НАХОЖДЕНИЯ СЕДЛОВЫХ ТОЧЕК

М.И. БАН

Разработан новый квази–ньютоновский метод для нахождения седловых точек. При преобразовании проблемы неопределенной оптимизации в положительно определенную, сохранена высокая нумерическая эффективность (в минимализации) характерная для этого класса алгоритмов. В то же время органические проблемы, возникающие в неопределенных оптимизациях, могут быть избежены.