OPTIMIZATION OF MINIMA AND SADDLE POINTS

Trygve Helgaker
Department of Chemistry
University of Oslo
P.O.B. 1033 Blindern
N-0315 Oslo 3
Norway

The optimization of minima and saddle points of a function in many variables is reviewed. Emphasis is on methods applicable to the calculation of electronic wave functions (ground and excited states) and the optimization of minima and transition states of molecular potential energy surfaces.

I. INTRODUCTION

We discuss in this paper the unconstrained optimization of stationary points of a smooth function f(x) in many variables. The emphasis is on methods useful for calculating molecular electronic energies and for determining molecular equilibrium and transition state structures. The discussion is general and practical aspects concerning computer implementations are not treated.

The field of minimization is well developed, see for example the monographs by Fletcher[1], by Gill, Murray and Wright[2], and by Dennis and Schnabel[3]. The optimization of saddle points such as excited electronic states and molecular transition states has received little attention outside the field of quantum chemistry and is less developed. In addition less information - experimental or intuitive - is usually available about saddle points, compounding this problem. In some respects the determination of saddle points resembles art more than technique. For previous reviews of methods for optimization of equilibrium geometries and transition states, see Schlegel[4] and Head and Zerner[5].

---

[1] R. Fletcher, *Practical Methods of Optimization* (Wiley, Chichester, 1980), Vol. 1

[2] P. E. Gill, W. Murray, and M. H. Wright, *Practical Optimization* (Academic, London, 1981)

[3] J. E. Dennis and R. B. Schnabel, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations* (Prentice Hall, Englewood Cliffs, 1983)

[4] H. B. Schlegel, Adv. Chem. Phys. **69**, 249 (1987); H. B. Schlegel, in *New Theoretical Concepts for Understanding Organic Reactions*, edited by J. Bertrán and I. G. Csizmadia (Kluwer, Dordrecht, 1989)

B. O. Roos (ed.), *Lecture Notes in Quantum Chemistry*

II. STATIONARY POINTS

Before discussing methods for optimization of stationary points it is appropriate to state briefly the mathematical characterization of such points and lay out the basic strategies for their determination.

A. Characterization of Stationary Points

A stationary point x* of a smooth function f(x) in n variables may be characterized in terms of the derivatives at x*. Sufficient conditions for a *minimum* are

$$\left.\begin{aligned} \nabla f(x^*) &= 0 \\ \mathrm{In}\nabla^2 f(x^*) &= [n, 0, 0] \end{aligned}\right\} \Rightarrow \text{minimum} \tag{2.1}$$

where the gradient and Hessian are given by

$$\nabla f = \begin{bmatrix} \dfrac{\partial f}{\partial x_1} \\ \vdots \\ \dfrac{\partial f}{\partial x_n} \end{bmatrix}, \quad \nabla^2 f = \begin{bmatrix} \dfrac{\partial^2 f}{\partial x_1 \partial x_1} & \cdots & \dfrac{\partial^2 f}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial^2 f}{\partial x_n \partial x_1} & \cdots & \dfrac{\partial^2 f}{\partial x_n \partial x_n} \end{bmatrix} \tag{2.2}$$

The inertia InM of a real symmetric matrix M in Eq. (2.1) is a triple

$$\mathrm{InM} = [\, \pi(M), \ \nu(M), \ \delta(M)] \tag{2.3}$$

where $\pi(M)$ is the number of positive eigenvalues of M, $\nu(M)$ the number of negative eigenvalues (sometimes referred to as the index of M), and $\delta(M)$ the number of zero eigenvalues. Equation (2.1) means that at a minimum the function has a vanishing gradient (zero slope) and a positive definite Hessian (positive curvature in all directions). However, Eq. (2.1) is not a necessary condition for a minimum. A point with inertia [n-k,0,k] may still be a minimum but to find out we must examine higher derivatives.

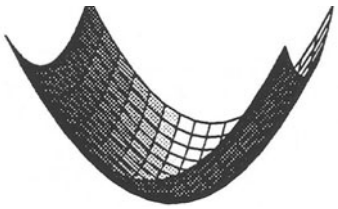
---

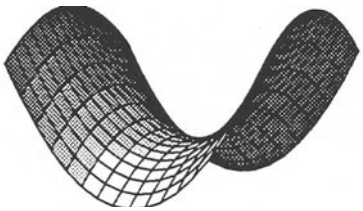
[5]J. D. Head and M. C. Zerner, Adv. Quantum Chem. **20**, 239 (1989)

Sufficient conditions for a k´th order *saddle point* are given by

$$\left.\begin{array}{l} \nabla f(x^*) = 0 \\ \mathrm{In}\nabla^2 f(x^*) = [n-k,\ k,\ 0] \end{array}\right\} \Rightarrow \text{k´th order saddle point} \qquad (2.4)$$

For example, a first-order saddle point has one and only one direction of negative curvature. Clearly, to characterize a stationary point we must be able to calculate the gradient and the Hessian eigenvalues of f(x). The figures below illustrate a minimum and a saddle point in two dimensions.





Both minima and saddle points are of interest. In the case of wave functions, the ground state is a minimum and the excited states are saddle points of the electronic energy function.[6] On potential surfaces minima and first-order saddle points correspond to equilibrium geometries and transition states. Higher-order saddle points on potential energy surfaces are of no interest.

B. Strategies for Optimization of Stationary Points

Optimizations may roughly be said to consist of two stages. In the first stage the purpose is to take us from some initial guess of the *optimizer* x* to a point in its neighborhood. In this *global* part of the search we may for example start with a Hessian with incorrect index and our goal is to locate an area with correct curvature. In the final *local* stage of the optimization our purpose is to determine the exact location of the optimizer

[6]However, not all minima and saddle points are satisfactory representations of electronic states. The decision as to whether a stationary point is a good approximation to an electronic state must be based on other criteria.

starting from a point in its immediate vicinity. A useful algorithm should handle both stages successfully. It should converge globally (i.e., from any starting point) and be fast in the local region.[7]

All practical methods start by constructing a *local model* of the function in the vicinity of the current estimate x_c of x^*. These models should accurately represent the function in some area around x_c, they should be easy to construct and yet flexible enough to guide us towards the stationary point. Such models are treated in Sec. III.[8]

In the *local* region we may expect the model to represent the function accurately in the neighborhood of x^*. Therefore, in this region the search is rather simple. We take a step to the optimizer of the local model, construct a new model and repeat until convergence. The different methods converge with a characteristic rate as discussed in Sec. IV.

In the *global* region the model should guide us in the right general direction towards x^*. This is relatively easy in minimizations since any step that reduces the function may be considered a step in the right direction. Global strategies for minimizations are treated in Sec. V.

In saddle point optimizations it is harder to judge the quality of a global step. Since a saddle point is a minimum in some directions and a maximum in others, the strategy is to identify these directions and take a step that increases and decreases the function accordingly. Such methods are discussed in Sec. VI.

III. LOCAL MODELS

We discuss in this section several models used in optimizations. Of these, the most successful are the quadratic model and its modifications, the restricted second-order model and the rational function model.

---

[7]It should be noted that there is no efficient way to ensure that the *global* minimum of a function is obtained in an optimization. In practice, we must be satisfied with methods which lead us to *local* minima.

[8]The parametrization of the function and the local models may significantly affect the performance of any optimization scheme. When choosing coordinates, we should reduce as much as possible coordinate couplings and higher-order dependencies.

A. The Linear Model

The simplest model is the *local linear or affine model*

$$m_A(x) = f(x_c) + \tilde{g}_c\, s \tag{3.1}$$

where

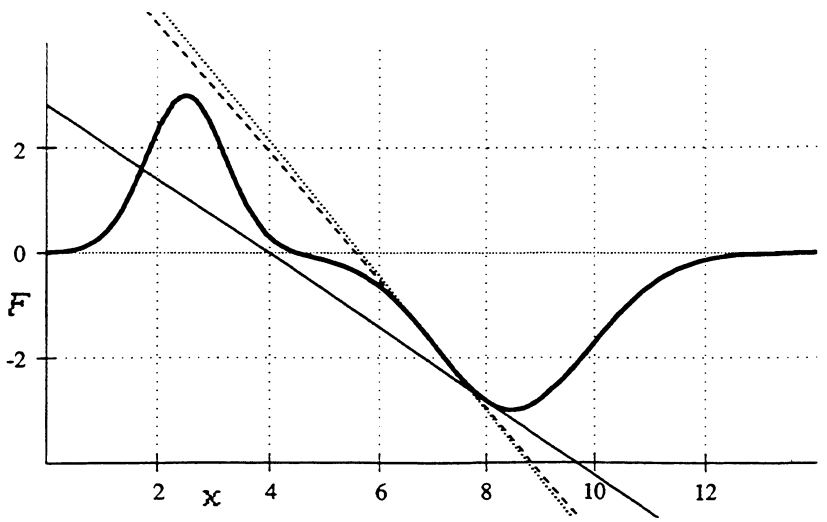
$$s = x - x_c \tag{3.2}$$

and g_c is the gradient of $f(x)$ at the current point

$$g_c = \nabla f(x_c) \tag{3.3}$$

The linear model, which may also be constructed from an approximate gradient, is simple but not particularly useful since it is unbounded and has no stationary point. It contains no information about the curvature of the function. It is the basis for the *steepest descent method* in which a step opposite the gradient is determined by line search (*vide infra*).

The figure below illustrates the linear model for a function in one variable. The function has a minimum at $x^* = 8.5$ and we have plotted models expanded around 7 (dotted line), 7.4 (dashed line), and 8 (full line). Clearly, these models provide little information about the location of the minimum.



B. The Quadratic Model

The problem with the linear model is that it gives no information about the curvature of f(x). This is provided by the more useful *local quadratic model*

$$m_Q(s) = f(x_c) + \tilde{g}_c s + \tfrac{1}{2}\tilde{s} B_c s \tag{3.4}$$

The symmetric matrix B_c is either the exact Hessian at x_c

$$G_c = \nabla^2 f(x_c) \tag{3.5}$$

or some approximation to it. If the exact Hessian is used the quadratic

$$m_{SO}(s) = f(x_c) + \tilde{g}_c s + \tfrac{1}{2}\tilde{s} G_c s \tag{3.6}$$

is referred to as the *second-order (SO) model* since it is the second-order Taylor expansion of the function around x_c. The SO model is expensive to construct since it requires the gradient as well as the Hessian but it gives complete information about slope as well as curvature at x_c.

To determine the stationary points of the quadratic model we differentiate the model and set the result equal to zero. We obtain a linear set of equations

$$B_c s = -g_c \tag{3.7}$$

which has a unique solution

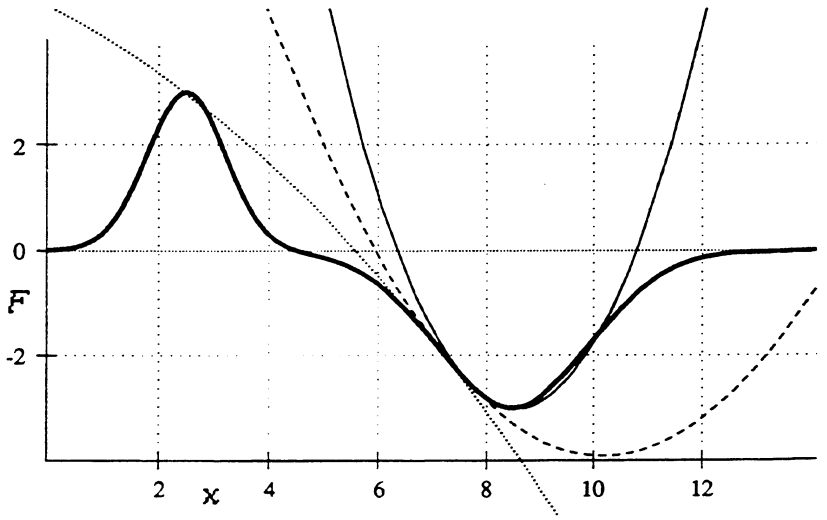
$$s = -B_c^{-1} g_c \tag{3.8}$$

provided B_c is nonsingular. The quadratic model therefore has one and only one stationary point. This is a minimum if B_c is positive definite. When B_c is the exact Hessian

$$s_N = -G_c^{-1} g_c \tag{3.9}$$

is called the *Newton step*. It forms the basis for *Newton's method* and its globally convergent modifications discussed later. When an approximate Hessian is used Eq. (3.8) is called the *quasi-Newton step*.

The figure below shows the SO models of the same function as for the linear model above. The model around 8 (full line) has a minimum close to the minimizer and the Newton step therefore provides a good estimate of x*. The model around 7.4 (dashed line) overshoots x*. The Newton step is not useful although the model has the correct convex shape. Finally, the model around 7 has the wrong shape and the Newton step takes us away from x*. Note that while there is little difference between the *linear* models at 7 and 7.4, the SO models differ strongly. We see that although the SO model gives more information about the function than does the linear model, its stationary point is not always a good estimate of x*.



C. The Restricted Second-Order Model

The quadratic model is an improvement on the linear model since it gives information about the curvature of the function and contains a stationary point. However, the model is still unbounded and it is a good approximation to f(x) only in some region around x_c. The region where we can trust the model to represent f(x) adequately is called the *trust region*. Usually it is impossible to specify this region in detail and for convenience we assume that it has the shape of a hypersphere $|s| \leq h$ where h is the *trust*

radius. This gives us the *restricted second-order (RSO) model*

$$m_{RSO}(s) = f(x_c) + \tilde{g}_c s + \tfrac{1}{2}\tilde{s} G_c s, \quad \tilde{s}s \le h^2 \tag{3.10}$$

using the exact Hessian.

The SO model has several stationary points. If the Newton step Eq. (3.9) is shorter than the trust radius $|s_N| < h$ then the RSO model has a stationary point in the interior. It also has at least two stationary points on the boundary $|s| = h$. To see this we introduce the Lagrangian

$$L(s,\mu) = m_{SO}(s) - \tfrac{1}{2}\mu(\tilde{s}s - h^2) \tag{3.11}$$

where μ is an undetermined multiplier. Differentiating this expression and setting the result equal to zero, we obtain

$$s(\mu) = -(G_c - \mu)^{-1} g_c \tag{3.12}$$

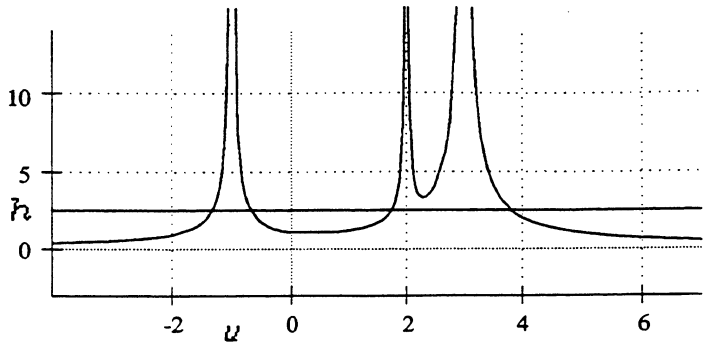
where the *level shift parameter* μ is chosen such that the step is to the boundary

$$\sqrt{\tilde{g}_c (G_c - \mu)^{-2} g_c} = h \tag{3.13}$$

To see the solutions to Eq. (3.13) more clearly we have plotted below its left- and right-hand sides as functions of μ for a function in three variables and with Hessian eigenvalues –1, 2, and 3. The step length function has poles at the eigenvalues as can be seen from Eq. (3.13) in the diagonal representation

$$\sum_i \frac{\phi_i^2}{(\lambda_i - \mu)^2} = h^2 \tag{3.14}$$

where λ_i are the eigenvalues and ϕ_i the components of the gradient along the eigenvectors. The eigenvalues determine the positions of the peaks and the gradient their width.



The solutions to Eq. (3.13) are found at the intersections of the two curves. Since the step goes to infinity at the eigenvalues and to zero at infinity there are at least two solutions: one with $\mu < \lambda_1$ (the smallest eigenvalue) and another with $\mu > \lambda_n$ (the largest eigenvalue). We may also have stationary points in each of the n–1 regions $\lambda_k < \mu < \lambda_{k+1}$. For large h we have two solutions in each region, for small h there may be no solution. If the model is constructed around a stationary point, the peaks are infinitely narrow and we have 2n stationary points on the boundary (along each eigenvector) in addition to the point in the interior.

It may be shown that if $\mu < \lambda_1$ then the solution to Eq. (3.13) is the global minimum on the boundary. In the diagonal representation the step may be written

$$\sigma_i = -\sum_i \frac{\phi_i}{\lambda_i - \mu} \tag{3.15}$$

where σ_i is the component of the step along the i'th eigenvector. We see that by selecting the solution $\mu < \lambda_1$ we take a step opposite the gradient in each mode.

The other solutions to Eq. (3.13) correspond to stationary points where the function is increased in some directions and reduced in others. For example, if we select a solution in the region $\lambda_2 < \mu < \lambda_3$ then the step is toward the gradient of the first two modes and opposite the gradient of all higher modes. The second-order change in the function may be written

$$\Delta m_{RSO} = \sum_i \frac{\phi_i^2 (\mu - \frac{1}{2}\lambda_i)}{(\lambda_i - \mu)^2} \tag{3.16}$$

and we see that the model decreases along all modes for which $\lambda_k > 2\mu$ and increases along all others.

The trust radius h reflects our confidence in the SO model. For highly anharmonic functions the trust region must be set small, for quadratic functions it is infinite. Clearly, during an optimization we must be prepared to modify h based on our experience with the function. We return to the problem of updating the trust radius later.

D. The Rational Function Model

The trust region was introduced since the SO model is a good approximation to the function only in some region around the expansion point. The resulting RSO model has stationary points on the boundary of the trust region and possibly a stationary point in the interior. These points are later used to construct globally convergent optimization algorithms.

There is another way to introduce restrictions on the step lengths in the global part of an optimization. The *rational function (RF) model* is given by[9]

$$m_{RF}(s) = f(x_c) + \frac{\tilde{g}_c s + \frac{1}{2}\tilde{s} G_c s}{1 + \tilde{s} S s} \tag{3.17}$$

which may be written in the form

$$m_{RF}(s) = f(x_c) + \frac{1}{2}\frac{[\tilde{s}\ 1]\begin{bmatrix} G_c & g_c \\ \tilde{g}_c & 0 \end{bmatrix}\begin{bmatrix} s \\ 1 \end{bmatrix}}{[\tilde{s}\ 1]\begin{bmatrix} S & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} s \\ 1 \end{bmatrix}} \tag{3.18}$$

where the metric S is symmetric matrix.[10] This model is bounded since large elements in the numerator are balanced by large elements in the denominator. Also, to second order the RF and SO models are identical since

---

[9]This method was used for optimization of electronic wave functions by A. Banerjee and F. Grein, Int. J. Quantum Chem. **10**, 123 (1976) and D. R. Yarkony, Chem. Phys. Lett. **77**, 634 (1981). It was applied to surface studies by A. Banerjee, N. Adams, J. Simons, and R. Shepard, J. Phys. Chem. **89**, 52 (1985). It has, to the author's knowledge, not been discussed in textbooks on optimization.

[10]We here use the exact Hessian although the RF model may also be constructed from an approximate (updated) Hessian.

$$[1 + \tilde{s}Ss]^{-1} = 1 - \tilde{s}Ss + O(s^4) \tag{3.19}$$

Therefore, in the RF model we have added higher order terms to the SO model to make it bounded. The explicit form of the RF model depends on the matrix S which should reflect the anharmonicity of the function. Usually we do not have information to specify S in detail and we simply take it to be the unit matrix multiplied by a scalar S.[11]

To find the stationary points of the RF model we differentiate Eq. (3.18) and set the result equal to zero. We arrive at the eigenvalue equations

$$\begin{bmatrix} G_c & g_c \\ \tilde{g}_c & 0 \end{bmatrix}\begin{bmatrix} s \\ 1 \end{bmatrix} = \nu\begin{bmatrix} S & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} s \\ 1 \end{bmatrix} \tag{3.20}$$

Solving these equations we obtain n+1 eigenvectors and eigenvalues

$$\nu = \frac{[\tilde{s}\ 1]\begin{bmatrix} G_c & g_c \\ \tilde{g}_c & 0 \end{bmatrix}\begin{bmatrix} s \\ 1 \end{bmatrix}}{[\tilde{s}\ 1]\begin{bmatrix} S & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} s \\ 1 \end{bmatrix}} = 2\Delta m_{RF}(s) \tag{3.21}$$

corresponding to the n+1 stationary points of the RF model. From Eq. (3.21) we see that the minimum belongs to the lowest eigenvalue. In general, the k´th eigenvalue belongs to a saddle point of index k–1. Notice that the eigenvalues Eq. (3.21) give the change in the model rather than in the function when the step is taken. The stationary points of the RF model do not necessarily represent stationary points of f(x) but they are useful for constructing globally convergent optimization algorithms.

Since the coefficient matrix of the eigenvalue equations Eq. (3.20)

$$G_c^+ = \begin{bmatrix} G_c & g_c \\ \tilde{g}_c & 0 \end{bmatrix} \tag{3.22}$$

has dimension n+1 and contains the Hessian in the upper left corner it is called the *augmented Hessian*. The n+1 eigenvalues of the augmented

[11]For the same reason we assumed that the trust region of the RSO model is a simple hypersphere with an adjustable radius h.

Hessian λ_k^+ bracket the Hessian eigenvalues

$$\lambda_1^+ \leq \lambda_1 \leq \lambda_2^+ \leq \lambda_2 \leq \cdots \leq \lambda_n \leq \lambda_{n+1}^+ \tag{3.23}$$

The eigenvalues of Eq. (3.20) coincide with the eigenvalues of the augmented Hessian only when S equals unity.

To compare the RSO and RF models we expand Eq. (3.20) and obtain

$$s = -(G_c - \mu)^{-1} g_c \tag{3.24}$$

$$\tilde{g}_c s = \frac{\mu}{S} \tag{3.25}$$

where

$$\mu = \nu S \tag{3.26}$$

Inserting Eq. (3.24) in Eq. (3.25) we find

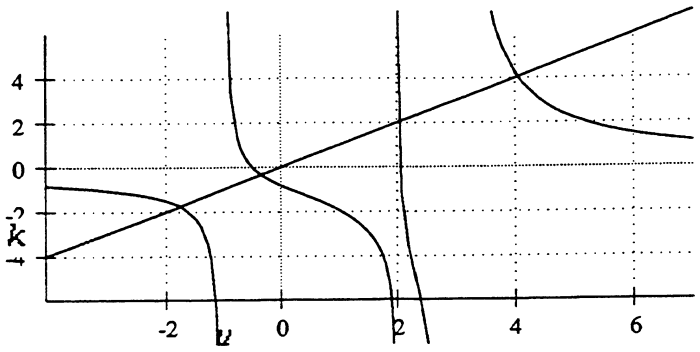
$$-\tilde{g}_c (G_c - \mu)^{-1} g_c = \frac{\mu}{S} \tag{3.27}$$

or in the diagonal representation

$$\sum_i \frac{\phi_i^2}{\mu - \lambda_i} = \frac{\mu}{S} \tag{3.28}$$

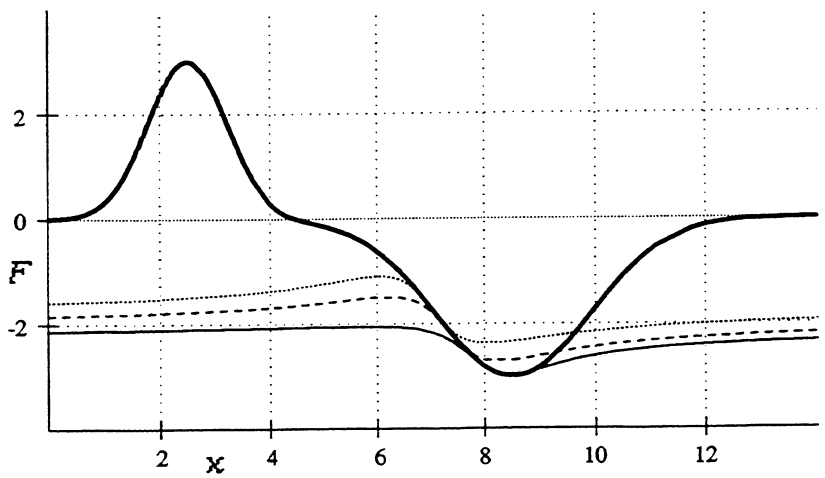
Plotting the left and right hand sides of this equation as functions of μ we obtain a figure such as the one below [using the same function as for Eq. (3.13)]. The left hand side goes to infinity at the Hessian eigenvalues. We have n+1 intersections, one for each stationary point of the rational model. Changing the metric S changes the slope of the straight line and therefore the intersections.

Once μ has been determined we calculate the step from the modified Newton equations Eq. (3.24). Therefore, the RF and RSO steps are calculated in the same way. The only difference is the prescription for determining the level shift. In the RSO approach μ reflects the trust radius h, in the RF model μ reflects the metric S. By varying h and S freely the same steps are obtained in the two models.



Close to a stationary point g_c vanishes. One of the eigenvalues of the augmented Hessian Eq. (3.22) then goes to zero and the rest approach those of G_c. The zero-eigenvalue step becomes the Newton step and the remaining n steps become infinite and parallel to the Hessian eigenvectors.

To summarize, in the RF approach we make the quadratic model bounded by adding higher-order terms. This introduces n+1 stationary points, which are obtained by diagonalizing the augmented Hessian Eq. (3.22). The figure below shows three RF models with S equal to unity, using the same function and expansion points as for the linear and quadratic models above. Each RF model has one maximum and one minimum in contrast to the SO models that have one stationary point only. The minima lie in the direction of the true minimizer.



E. Hessian Updates

If the exact Hessian is unavailable or computationally expensive, we may use an approximation. Approximate Hessians are usually obtained by one of several *Hessian update methods*. The update techniques are designed to determine an approximate Hessian B_+ at

$$x_+ = x_c + s_c \tag{3.29}$$

in terms of the Hessian B_c at x_c, the gradient difference

$$y_c = g_+ - g_c \tag{3.30}$$

and the step vector s_c. Expanding g_c around x_+ gives

$$y_c = B_+ s_c + O(s_c^2) \tag{3.31}$$

which shows that the gradient difference y_c contains a component of the finite-difference approximation to the exact Hessian along the direction s_c. This finite-difference information together with structural characteristics of the exact Hessian is used to form the Hessian updates described below.

Based on the finite-difference formula Eq. (3.31) all Hessian updates are required to fulfill the *quasi-Newton condition*

$$y_c = B_+ s_c \tag{3.32}$$

and to possess the property of *hereditary symmetry*, i.e., B_+ is symmetric if B_c is symmetric. These requirements are fulfilled by the *Powell-symmetric-Broyden (PSB) update* given by

$$B_+ = B_c + \frac{(\tilde{s}_c s_c) T_c \tilde{s}_c + (\tilde{s}_c s_c) s_c \tilde{T}_c - (\tilde{T}_c s_c) s_c \tilde{s}_c}{(\tilde{s}_c s_c)^2} \tag{3.33}$$

where

$$T_c = y_c - B_c s_c \tag{3.34}$$

Notice that the construction of the updated Hessian involves simple matrix and vector multiplications of gradient and step vectors.

It is often desirable that the approximate Hessian is positive definite so that the quadratic model has a minimum. To ensure this we may use the *Broyden-Fletcher-Goldfarb-Shanno (BFGS) update* given by

$$B_+ = B_c + \frac{y_c \tilde{y}_c}{\tilde{y}_c s_c} - \frac{B_c s_c \tilde{s}_c B_c}{\tilde{s}_c B_c s_c} \tag{3.35}$$

which under certain weak conditions on the step vector has the property of *hereditary positive definiteness*, i.e., if B_c is positive definite, B_+ is positive definite. As we shall see, this is useful for minimizations even when the exact Hessian has directions of negative curvature. It is then more appropriate to speak of B_c as an effective rather than approximate Hessian.

There are other Hessian updates but for minimizations the BFGS update is the most successful. Hessian update techniques are usually combined with line search (*vide infra*) and the resulting minimization algorithms are called *quasi-Newton methods*. In saddle point optimizations we must allow the approximate Hessian to become indefinite and the PSB update is therefore more appropriate.

IV THE LOCAL REGION

If the quadratic model or one of its modifications are used, the local region presents no difficulties. All methods converge rapidly since they effectively reduce to Newton´s method in the local region. In this section we briefly discuss local convergence rates and stopping criteria.

A. Convergence Rates

In the local region, optimization methods may be characterized by their rate of convergence. Let x_k be a sequence of points converging to x^*

$$\lim_{k\to\infty} x_k = x^* \tag{4.1}$$

and let e_k be the error in x_k

$$e_k = x_k - x^* \tag{4.2}$$

Convergence is said to be linear if

$$\lim_{k\to\infty} \frac{|e_{k+1}|}{|e_k|} = a \qquad \leftarrow \text{ linear convergence} \tag{4.3}$$

for some (preferably small) number a, superlinear if

$$\lim_{k\to\infty} \frac{|e_{k+1}|}{|e_k|} = 0 \qquad \leftarrow \text{ superlinear convergence} \tag{4.4}$$

and quadratic if

$$\lim_{k\to\infty} \frac{|e_{k+1}|}{|e_k|^2} = a \qquad \leftarrow \text{ quadratic convergence} \tag{4.5}$$

Note that quadratic convergence implies superlinear convergence.

Quadratic convergence means that eventually the number of correct figures in x_c doubles at each step, clearly a desirable property. Close to x^* Newton's method Eq. (3.9) shows quadratic convergence while quasi-Newton methods Eq. (3.8) show superlinear convergence. The RF step Eq. (3.20) converges quadratically when the exact Hessian is used. Steepest descent with exact line search converges linearly for minimization.

B. Stopping Criteria

Different stopping criteria may be used for optimizations. The most straightforward is to require the norm of the gradient to be smaller than some threshold

$$|g_c| \le \varepsilon \tag{4.6}$$

but we may also test on the predicted second-order change in the function

$$\tfrac{1}{2}\left|\tilde{g}_c B_c^{-1} g_c\right| \le \varepsilon \tag{4.7}$$

or on the size of the Newton step

$$\left|B_c^{-1} g_c\right| \le \varepsilon \tag{4.8}$$

In addition, it is important to check that the structure of the Hessian is correct. For example, if a geometry minimization has been carried out with constraints on the symmetry of the molecule, the solution may turn out to be a saddle point when symmetry breaking distortions are considered.

V. STRATEGIES FOR MINIMIZATION

Global strategies for minimization are needed whenever the current estimate of the minimizer is so far from x^* that the local model is not a good approximation to $f(x)$ in the neighborhood of x^*. Three methods are considered in this section: the quadratic model with line search, trust region (restricted second-order) minimization and rational function (augmented Hessian) minimization.

A. Line Searches

In the global region the Newton or quasi-Newton step may not be satisfactory. It may, for example, increase rather than decrease the function to be minimized. Although the step must then be rejected we may still use it to provide a direction for a one-dimensional minimization of the function. We then carry out a search along the Newton step until an acceptable reduction in the function is obtained and the result of this *line search* becomes our next step.

All line searches start by defining a *descent direction*. Consider all vectors z which fulfill the condition

$$\tilde{z}\, g_c < 0 \tag{5.1}$$

Since

$$\left.\frac{df(x_c + tz)}{dt}\right|_{t=0} = \tilde{z}\, g_c < 0 \tag{5.2}$$

there must be a positive number τ such that

$$f(x_c + tz) < f(x_c) \quad \text{for } 0 \le t \le \tau \tag{5.3}$$

and z is therefore said to be a descent direction. The negative gradient obviously is a descent direction (often referred to as the *steepest descent* direction) as is the Newton or quasi-Newton step

$$s_N = - B_c^{-1} g_c \tag{5.4}$$

provided the Hessian is positive definite

$$\tilde{s}_N g_c = - \tilde{g}_c B_c^{-1} g_c < 0 \tag{5.5}$$

It is for this reason the positive definite BFGS update Eq. (3.35) is preferred over the PSB update Eq. (3.33) for minimizations. The positive definite Newton step is usually a better direction than steepest descent since it takes into account features of the function further away from x_c than does steepest descent.

Given the direction of search z_c at x_c we must find a satisfactory step along this direction. It would seem that the best is to minimize $f(x_c + tz_c)$ with respect to t and take the step

$$s_c = x_c + t^* z_c \tag{5.6}$$

where t^* is the minimizer. However, such *exact line searches* are expensive and not used in practice. Instead *inexact or partial line searches* are used to generate an *acceptable* point s_c along z_c. By acceptable we mean for example a point that fulfills the condition

$$f(x_c + tz_c) < f(x_c) + \tfrac{1}{2} t \tilde{g}_c z_c \tag{5.7}$$

for some $0 < t \leq 1$. The parameter t may be determined by interpolation. We first try the full Newton step. If this is not acceptable a smaller step is obtained by interpolation and tested. This *backtracking* process is repeated until an acceptable step is found.

Line searches are often used in connection with Hessian update formulas and provide a relatively stable and efficient method for minimizations. However, line searches are not always successful. For example, if the Hessian is indefinite there is no natural way to choose the descent direction. We may then have to revert to steepest descent although this step makes no use of the information provided by the Hessian. It may also be

argued more generally that backtracking from the Newton step never makes full use of the available information since the Hessian is used only to generate the *direction* and not the *length* of the step. Alternative methods are provided by the RSO and RF models.

B. Trust Region Minimization

In the *trust region or restricted step method* we determine in each iteration the global minimizer of the RSO model Eq. (3.10). In the global region a step is taken to the boundary of the trust region

$$s(\mu) = -(G_c - \mu)^{-1} g_c \tag{5.8}$$

since this contains the minimizer of the model. In the local region the model has a minimizer inside the trust region and we take the Newton step

$$s(0) = -G_c^{-1} g_c \tag{5.9}$$

The method therefore reduces to Newton's method in the local region with its rapid rate of convergence.

When solving Eq. (5.8) the level shift parameter μ must be chosen such that $s(\mu)$ is the global minimizer on the boundary. From the discussion in Sec. III we know that μ must be smaller than the lowest Hessian eigenvalue. Also, μ must be negative since otherwise the step becomes longer than the Newton step. The exact value of μ may be found by bisection or interpolation.

The trust radius h is obtained by a feedback mechanism. In the first iteration some arbitrary but reasonable value of h is assumed. In the next iteration, h is modified based on a comparison between the predicted reduction in f(x) and the actual reduction. If the ratio between actual and predicted reductions

$$R_c = \frac{f_+ - f_c}{\tilde{g}_c s_c + \frac{1}{2} \tilde{s}_c G_c s_c} \tag{5.10}$$

is close to one h is increased. If the ratio is small the radius is reduced. If the ratio is negative the step is rejected, the trust region reduced and a new step is calculated from Eq. (5.8).

The trust region method is usually implemented with the exact Hessian. Updated Hessians may also be used but an approximate Hessian usually does not contain enough information about the function to make the trust region reliable in all directions. The trust region method provides us with the possibility to carry out an unbiased search in all directions at each step. An updated Hessian does not contain the information necessary for such a search.

It is interesting to note the difference between the RSO trust region method and Newton's method with line search. In the trust region method we first choose the size of the step (the trust radius), and then determine the direction of the step (constrained minimization within the trust region). In Newton's method with line search we first choose the direction of the step (the Newton direction) and next determine the size of the step (constrained minimization along the Newton direction). The trust region method is more robust (guaranteed convergence for smooth and bounded functions) and has no problems with indefinite Hessians. It is perhaps more conservative than line search since the size of the step is predetermined. However, line search requires additional energy calculations and is not equally well suited for handling indefinite Hessians. It is possible to combine features of trust region and line search minimizations by backtracking not along the Newton step but along the line generated by calculating trust region steps with different h (the Levenberg-Marquardt trajectory).

C. Rational Function Minimization

A global strategy for minimization may also be based on the RF model. In each iteration the eigenvalue equations

$$\begin{bmatrix} G_c & g_c \\ \tilde{g}_c & 0 \end{bmatrix} \begin{bmatrix} s \\ 1 \end{bmatrix} = \nu \begin{bmatrix} S & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} s \\ 1 \end{bmatrix} \tag{5.11}$$

are solved for the lowest eigenvalue ν and the corresponding eigenvector. This gives a step

$$s(\nu S) = -(G_c - \nu S)^{-1} g_c \tag{5.12}$$

to the minimizer of the rational function. This step is opposite the gradient in each mode if $\nu S < \lambda_1$. In the local region, ν goes to zero and the step approaches the Newton step. The parameter S may be used for step size control in the same way as h in trust region minimizations. However, S is usually set to one and the step is simply scaled down if it is unsatisfactory.

One advantage of the RF minimization over trust region RSO minimization is that we need only calculate the lowest eigenvalue and eigenvector of the augmented Hessian. In the trust region method we must first calculate the lowest eigenvalue of the Hessian and then solve a set of linear equations to obtain the step.

In conclusion, the trust region method is more intuitive than the RF model and provides a more natural step control. On the other hand, RF optimization avoids the solution of one set of linear equations, which is important when the number of variables is large.

VI. STRATEGIES FOR SADDLE POINT OPTIMIZATIONS

Optimizations of saddle points such as molecular transition states and excited electronic states are usually more difficult than minimizations. First, the methods to determine saddle points are less developed and less stable than methods for minimizations. Second, it is usually less clear where to start an optimization of a saddle point than a minimization.

In the immediate vicinity of a saddle point there is no problem. We proceed as for minimizations either by solving a set on linear equations to determine the Newton step Eq. (5.9), or by solving a set of eigenvalue equations to determine the near-zero eigenvalue solution Eq. (5.11). The difficulties are in the global part of the optimization.

The strategies for saddle point optimizations are different for electronic wave functions and for potential energy surfaces. First, in electronic structure calculations we are interested in saddle points of any order (although the first-order saddle points are the most important) whereas in surface studies we are interested in first-order saddle points only since these represent transition states. Second, the number of variables in electronic structure calculations is usually very large so that it is impossible to diagonalize the Hessian explicitly. In contrast, in surface studies the number of variables is usually quite small and we may easily trans-

form to the diagonal representation. Because of these differences we first briefly discuss methods useful for excited states and then describe methods for transition states.

A. Methods for Excited Electronic States

As mentioned in Sec. II the general strategy in saddle point calculations is to take a step that increases the function in some directions while reducing it in others. In the RSO and RF approaches the step may be written

(6.1)

or in the diagonal representation

$$\sigma_i = -\sum_i \frac{\phi_i}{\lambda_i - \mu} \tag{6.2}$$

By choosing a level shift in the range $\lambda_k < \mu < \lambda_{k+1}$ we take a step which initially at least increases the function along the k lowest modes and reduces it along all higher modes. Therefore, if at each step we select a level shift in this range we may eventually expect to enter the local region of the k´th excited state.

Within the RSO framework we first determine the correct range for μ by calculating the k+1 lowest eigenvalues of the Hessian.[12] Next we select an appropriate level shift in this range and finally solve a linear set of equations to obtain the step. For example, to move towards the first excited state we calculate two eigenvalues and solve one set of linear equations. The level shift may be adjusted to hit the boundary with little extra effort. But as noted in Sec. III there are either two or no solutions in the desired range. Therefore, the level shift cannot always be chosen unambiguously.

In the RF model the (k+1)´th eigenvector gives a step in the right direction. To obtain the step we must therefore calculate the k+1 lowest eigenvalues of the augmented Hessian. For example, when optimizing the first excited state we calculate two eigenvalues but do not solve a set linear

[12] In electronic structure calculations it is usually not possible to calculate all eigenvalues of the Hessian. Instead, methods have been developed to calculate selected eigenvalues. To obtain the k´th mode we must first calculate all lower modes.

equations as in the RSO approach. The RF model is therefore the preferred one for wave functions.

For a given S the RF model avoids the ambiguity in selecting the level shift. Usually S is set to unity and the step is scaled down if it becomes too large. In principle, however, S may be used for step control[13] and by allowing for all possible values of S we obtain the same steps as in the RSO model.

B. Gradient Extremals

In surface studies we are interested only in first-order saddle points since these represent transition states but unlike electronic calculations there may be more than one first-order saddle point of interest. We must therefore develop methods that allow us to guide the search in the more promising directions to catch the different transition states.

Using the same method as for the first excited electronic state, we select a level shift in the region $\lambda_1 < \mu < \lambda_2$. This procedure may indeed lead to a transition state but in this way we always increase the function along the lowest mode. However, if we wish to increase it along a higher mode this can only be accomplished in a somewhat unsatisfactory manner by coordinate scaling. Nevertheless, this method has been used by several authors with considerable success.[14] The problem of several first-order saddle points does not arise in electronic structure calculations since there is only one first excited state.[15]

To develop a method to locate saddle points by selectively following one eigenvector we note that at a stationary point all n components of the gradient in the diagonal representation are zero:

---

[13]A modification of this scheme has been used for optimizing excited states of multiconfigurational self-consistent field wave functions, see H. J. Aa. Jensen and H. Ågren, Chem. Phys. **104**, 229 (1986).

[14]C. J. Cerjan, W. H. Miller, J Chem. Phys. **75**, 2800 (1981); J. Simons, P. Jørgensen, H. Taylor, and J. Ozment, J. Phys. Chem. **87**, 2745 (1983); D. T. Nguyen and D. A. Case, J. Phys. Chem. **89**, 4020 (1985); H. J. Aa. Jensen, P. Jørgensen, and T. Helgaker, J. Chem. Phys. **85**, 3917 (1986); J. Nichols, H. Taylor, P. Schmidt, and J. Simons, J. Chem. Phys. **92**, 340 (1990).

[15]Our *approximate* electronic wave function may have more than one saddle point. Nevertheless, if we select our initial guess carefully we should be relatively close to the saddle point that most closely represents the true excited state and the level-shifted Newton step should guide us reliably to this point.

$$\phi(x_{SP1}) = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{6.3}$$

These n conditions define a point in n-dimensional space. We now move away from the stationary point in a controlled manner by relaxing only one of these conditions. For example, we may no longer require the second component of the gradient to be zero. We are then left with n–1 conditions, which define a line in n-dimensional space passing through the stationary point Eq. (6.3):

$$\phi[x_{GE}(t)] = \begin{bmatrix} 0 \\ \varphi(t) \\ \vdots \\ 0 \end{bmatrix} \tag{6.4}$$

We call this line a *gradient extremal (GE)*.[16] Unless the function increases indefinitely the gradient $\varphi(t)$ must eventually become zero or approach zero. It is therefore reasonable to expect that by following a GE we sooner or later hit a new stationary point

$$\phi(x_{SP2}) = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{6.5}$$

If we start at a minimum this must be a saddle point. This observation is the basis for the GE algorithm: Transition states are determined by carrying out one-dimensional searches along GEs starting at a minimum. Since n GEs pass through each stationary point, there are 2n directions along which we may carry out such line searches.

The condition that only one component of the gradient differs from zero means that the gradient is an eigenvector of the Hessian at GEs:

$$G(x)\, g(x) = \mu(x)\, g(x) \tag{6.6}$$

---

[16]The term gradient extremal was introduced by D. K. Hoffmann, R. S. Nord, and K. Ruedenberg, Theor. Chim. Acta. **69**, 265 (1986). Gradient extremals have also been discussed by J. Pancir, Collect. Czech. Chem. Commun. **40**, 1112 (1975) and by M. V. Basilevsky and A. G. Shamov, Chem. Phys. **60**, 347 (1981).

We obtain the same equations by optimizing the squared norm of the gradient in the contour subspace where f(x) is equal to a constant k. Differentiating the Lagrangian

$$L(x,\mu) = \tilde{g}(x)\, g(x) - 2\mu[f(x) - k] \tag{6.7}$$

and setting the result equal to zero we arrive at Eq. (6.6). This means, for example, that the gradient extremal belonging to the lowest eigenvalue may be interpreted as a valley floor.

To develop a practical method for tracing gradient extremals we return to the restricted second-order model Eq. (3.10).[17] In this model the Hessian is constant

$$G_{SO}(x) = G_c \tag{6.8}$$

and the gradient is given by

$$g_{SO}(x) = g_c + G_c s \tag{6.9}$$

Inserting Eqs. (6.8) and (6.9) in the gradient extremal equation Eq. (6.6) we obtain

$$(\mu - G_c)(g_c + G_c\, s) = 0 \tag{6.10}$$

If μ is different from all Hessian eigenvalues we obtain the Newton step

$$g_c + G_c\, s = 0 \tag{6.11}$$

which takes us to the stationary point on the SO model. This is trivially a gradient extremal point. We therefore set μ equal to the k´th Hessian eigenvalue and obtain

$$(\mu_k - G_c)(g_c + G_c\, s) = 0 \tag{6.12}$$

which has the solutions

---

[17] P. Jørgensen, H. J. Aa. Jensen, and T. Helgaker, Theor. Chim. Acta **73**, 55 (1988).

$$x_k(t) = -P_k G_c^{-1} g_c + t v_k \tag{6.13}$$

Here v_k is the eigenvector

$$G_c v_k = \mu_k v_k \tag{6.14}$$

and P_k the projector

$$P_k = 1 - v_k \tilde{v}_k \tag{6.15}$$

Therefore the k´th GE is a straight line passing through the stationary point of the model in the direction of the k´th eigenvector.

Equation (6.13) forms the basis for a second-order saddle point algorithm. In each iteration we first identify the mode to be followed (*the reaction mode*), then take a projected Newton step to minimize the function along all other modes (*the transverse modes*), and finally take a step along the reaction mode until we reach the boundary of the trust region. If the projected Newton step takes us out of the trust region, we minimize the transverse mode on the boundary instead and do not take a step along the reaction mode. In the local region x* lies within the trust region and we take the unprojected Newton step.

If the reaction mode becomes degenerate, this algorithm breaks down. We must then go to higher orders to take a step in the correct direction. Also, if there are many soft transverse modes the GE method may spend much time minimizing these.

C. Trust Region Image Minimization

In the GE method the minimization of the transverse modes comes first and the maximization of the reaction mode second. We now describe a method that weights minimization and maximization equally. We assume the existence of an *image* function with the following properties.[18] If the function to be optimized f(x) has the following gradient and eigenvalues at x

---

[18]The concept of image functions was introduced by C. M. Smith, Theor. Chim. Acta 74, 85 (1988).

$$\phi(x) = \begin{bmatrix} \phi_1(x) \\ \phi_2(x) \\ \vdots \\ \phi_n(x) \end{bmatrix} \qquad \lambda(x) = \begin{bmatrix} \lambda_1(x) \\ \lambda_2(x) \\ \vdots \\ \lambda_n(x) \end{bmatrix} \tag{6.16}$$

then the gradient and eigenvalues of the image function $\bar{f}(x)$ are

$$\bar{\phi}(x) = \begin{bmatrix} -\phi_1(x) \\ \phi_2(x) \\ \vdots \\ \phi_n(x) \end{bmatrix} \qquad \bar{\lambda}(x) = \begin{bmatrix} -\lambda_1(x) \\ \lambda_2(x) \\ \vdots \\ \lambda_n(x) \end{bmatrix} \tag{6.17}$$

Hence, in the diagonal representation the gradient and Hessian are identical except for opposite sign in the lowest mode.[19] Therefore, a first-order saddle point of the function coincides with a minimum of the image and we may determine the transition state by minimizing the image function.

To minimize the image function we use a second-order method since in each iteration the Hessian is needed anyway to identify the mode to be inverted (*the image mode*). Line search methods cannot be used since it is impossible to calculate the image function itself when carrying out the line search. However, the trust region RSO minimization requires only gradient and Hessian information and may therefore be used. In the diagonal representation the step Eq. (5.8) becomes

$$s(\mu) = -\frac{\phi_1}{(\lambda_1 + \mu)} v_1 - \sum_{i \neq 1} \frac{\phi_i}{(\lambda_i - \mu)} v_i \tag{6.18}$$

The only difference between the steps along the image mode and the transverse modes is the sign of the level shift. The level shift $\mu < -\lambda_1$ is determined such that the step is to the boundary of the trust region. Equation (6.18) forms the basis for the trust region image minimization (TRIM) method for calculating saddle points.[20]

---

[19]In the following we assume that the function and its image always differ in the first element (the lowest eigenvalue) although we may construct image functions by changing sign of any mode.

[20]T. Helgaker, Chem. Phys. Lett. **182** , 503 (1991).

To compare the image and GE methods, notice that in the diagonal representation the quadratic model may be written

$$m_{SO}(s) = f(x_c) + \sum_i m_i(\sigma_i) \tag{6.19}$$

where

$$m_i(\sigma_i) = \phi_i \sigma_i + \tfrac{1}{2} \lambda_i \sigma_i^2 \tag{6.20}$$

In RSO minimizations we minimize Eq. (6.19) within the trust region. If instead we wish to maximize the lowest mode and minimize the others we may use the model

$$\overline{m}_{SO}(s) = f(x_c) - m_1(\sigma_1) + \sum_{i \neq 1} m_i(\sigma_i) \tag{6.21}$$

since by reducing $-m_1(\sigma_1)$ we increase $m_1(\sigma_1)$. Equation (6.21) is the SO model of the image function.

We can now see the difference between the GE and TRIM methods. In the GE algorithm we first minimize the transverse modes and then maximize the reaction mode. In the TRIM method we minimize and maximize simultaneously by introducing an auxiliary function Eq. (6.21). The underlying idea of the GE method are lines connecting stationary points. The idea behind the TRIM method is an auxiliary function (the image function) whose minima coincide with the saddle points of the original function.

An image function does not exist for all functions. It exists by definition for all quadratic functions. It also exists trivially for functions in one variable since an image is obtained simply by changing the sign of the function. In any case, the image concept is useful for formulating an algorithm.

D. Rational Function Mode Following

In the GE algorithm we select one eigenvector as the reaction mode and follow this towards the transition state. A similar *mode following* technique has also been developed within the RF framework.[21] However, the RF

---

[21] J. Baker, J. Comput. Chem. 7, 385 (1986); A. Banerjee, N. Adams, J. Simons, and R. Shepard, J. Phys. Chem. 89, 52 (1985).

model Eq. (3.17) is not sufficiently flexible for mode following. Instead we use the model

$$m_{RF}(x) = f(x_c) + \frac{m_1(\sigma_1)}{1+\sigma_1^2} + \frac{\sum_{i\neq 1} m_i(\sigma_i)}{1+\sum_{i\neq k}\sigma_i^2} \tag{6.22}$$

where the first mode is the reaction mode. We have separated the model in two parts representing the reaction mode and the transverse modes. Each term is divided by a quadratic to make it bounded. The two metrics in Eq. (6.22) are set to unity. Note that two heuristic parameters appear in Eq. (6.22) rather than one as in the methods previously discussed.

The model Eq. (6.22) has 2n stationary points. To see this we differentiate the model and obtain two independent sets of equations

$$\begin{bmatrix} \lambda_1 & \phi_1 \\ \phi_1 & 0 \end{bmatrix} \begin{bmatrix} \sigma_1 \\ 1 \end{bmatrix} = \nu_R \begin{bmatrix} \sigma_1 \\ 1 \end{bmatrix} \tag{6.23}$$

$$\begin{bmatrix} \lambda_2 & & 0 & \phi_2 \\ & \ddots & & \vdots \\ 0 & & \lambda_n & \phi_n \\ \phi_2 & \cdots & \phi_n & 0 \end{bmatrix} \begin{bmatrix} \sigma_2 \\ \vdots \\ \sigma_n \\ 1 \end{bmatrix} = \nu_T \begin{bmatrix} \sigma_2 \\ \vdots \\ \sigma_n \\ 1 \end{bmatrix} \tag{6.24}$$

Equation (6.23) has two solutions for the reaction mode and Eq. (6.24) has n solutions for the transverse modes. To determine a transition state we choose the maximum of Eq. (6.23) and the minimum of Eq. (6.24). If the step becomes too big it is scaled down.

The three methods discussed here all work by the same principle. The Hessian is diagonalized and the reaction mode is identified. A step is then taken such the function is increased along the reaction mode and decreased along the transverse modes. The methods differ in the way this maximization and minimization is carried out.

VII. CONCLUSIONS

We have discussed the general features of methods for optimizing minima and saddle points. It should be remembered that the implementation of these methods involves adapting the general strategies to a specific

problem. It is therefore difficult to give hard and fast rules for optimizations. Also, the relative performance of the various methods is difficult to measure - it depends on the implementation of the methods and problem at hand. Therefore, no such comparisons have been given here.

Acknowledgments

I wish to thank Hans Jørgen Aa. Jensen, Poul Jørgensen, and Bernhard Schlegel for discussions.