A combined method for determining reaction paths, minima, and transition state geometries

Philippe Y. Ayala and H. Bernhard Schlegel
Wayne State University, Department of Chemistry, Detroit, Michigan 48202

(Received 4 February 1997; accepted 2 April 1997)

Mapping out a reaction mechanism involves optimizing the reactants and products, finding the transition state and following the reaction path connecting them. Transition states can be difficult to locate and reaction paths can be expensive to follow. We describe an efficient algorithm for determining the transition state, minima and reaction path in a single procedure. Starting with an approximate path represented by N points, the path is iteratively relaxed until one of the N points reached the transition state, the end points optimize to minima and the remaining points converged to a second order approximation of the steepest descent path. The method appears to be more reliable than conventional transition state optimization algorithms, and requires only energies and gradients, but not second derivative calculations. The procedure is illustrated by application to a number of model reactions. In most cases, the reaction mechanism can be described well using 5 to 7 points to represent the transition state, the minima and the path. The computational cost of relaxing the path is less than or comparable to the cost of standard techniques for finding the transition state and the minima, determining the transition vector and following the reaction path on both sides of the transition state. © *1997 American Institute of Physics.* [S0021-9606(97)01226-9]

I. INTRODUCTION

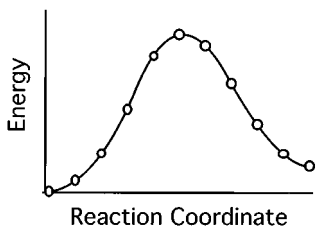
Equilibrium geometries, transition states, and reaction paths are central in the study of chemical reactions. Gradient-based methods have made the search for equilibrium geometries for small and medium size molecules almost routine. However, finding transition states can still be quite challenging. The use of redundant internal coordinates[1–5] has improved the efficiency of these methods, but following reaction paths can still take a large number of steps. Algorithms to find these features on potential energy surfaces have been reviewed recently.[6–11] In this paper, we present a combined algorithm that efficiently determines minima, transition state and the reaction path connecting them in a given region of an *ab initio* potential energy surface.

A transition state or first order saddle point is a stationary point that is an energy maximum in one direction and a minimum in all others. Quasi-Newton methods usually show a small radius of convergence because the specific direction for maximization is generally not known in advance and there is no guarantee that the optimization converges to the desired transition state. Combining synchronous transit and quasi-Newton methods to find transition states (TS) proves to be an efficient technique addressing some of these problems.[12,13] In this technique, the first few steps search for a maximum along an arc of circle connecting reactants and products and a minimum in all other directions. In the remaining optimization steps, a quasi-Newton based eigenvector following optimization[14] is guided by the tangent to the arc of circle passing through the putative TS and minima.

Once the transition state is located, mechanistic information about the chemical reaction can be obtained by following the steepest descent path[15] connecting reactants to products through the transition state. Various techniques for integrating the steepest descent path differential equation are available.[16–20] Even though some algorithms allow for reasonable step sizes and are very robust (e.g., the Gonzalez–Schlegel method[16]), following the reaction path requires many steps and is generally computationally expensive. In addition to the effort of determining the reaction path, the inherent drawback of following the steepest descent path is that it must start at the (sometimes very difficult to locate) transition state and requires the direction of the transition vector (usually necessitating a frequency calculation).

Jaisen and Shepard[21] describe a method for fitting the potential energy surface around an approximate reaction path using energies, first and second derivatives of the energy. The fitted surface is used to improve the reaction path and the process is repeated until it converges to the steepest descent path. Elber and Karplus[23–26] developed a method for obtaining a good approximation of the reaction path that does not require prior knowledge of the transition state and the potential energy surface curvature. This technique consists in minimizing the integral of the energy along the reaction path using an N-point discretization,

$$S=\frac{1}{L}\int_{\mathbf{q}_0}^{\mathbf{q}_{N+1}} E(\mathbf{q})\,ds \approx \frac{1}{L}\sum_{i=1}^{N} \frac{E(\mathbf{q}_i)+E(\mathbf{q}_{i-1})}{2}\,|\mathbf{q}_i-\mathbf{q}_{i-1}|. \tag{1}$$



Scheme 1: Elber–Karplus method.

At convergence, a good approximation for the location of the transition state can be obtained by interpolation which in turn can be refined by quasi-Newton methods. Typically 10 or more points are used to represent the path and special care (such as maintaining equal spacing and adding a self-avoidance potential[24]) must be taken so that the points do not collapse into minima during the optimization and to take into account rigid-body motions. The convergence of the integral optimization is very slow; McDouall *et al.*[27] noted that minimizing the two penalty functions to eliminate uneven path segments and rigid-body motions can be the most costly part of the algorithm. Because of the dimensionality of the problem ($3*n*N$, where n is the number of atoms in the molecular system and N is the number of points along the path) quasi-Newton optimization techniques other than conjugate-gradient algorithms[28] cannot be readily used to improve the convergence. Hence, the method in its unmodified form is only suitable for inexpensive levels of theory such as semi-empirical or molecular mechanics methods. A variant[25] that concentrates on the saddle point is somewhat more efficient and using a different approach Elber[26] also showed that a good representation of the steepest descent path can be obtained by representing the path by a series of straight lines defined by a set a points $\mathbf{q}_i$ and iteratively refining this approximation by minimizing the potential energy in the directions perpendicular to $(\mathbf{q}_{i-1} - \mathbf{q}_{i+1})$. The weakness of this method is the small radius of convergence,[26] possibly due mainly to the lack of spacing constraints.

Building upon these ideas, we have developed a combined procedure to find a transition state and follow the reaction path that is efficient enough for *ab initio* molecular orbital calculations. This procedure locates the transition state connecting two input structures and can represent the steepest descent path between them with as few as 3 to 5 points. The computational cost comparable to a regular transition state optimization plus reaction path following, but is intrinsically more reliable than the conventional approach. The algorithm is implemented in the development version of GAUSSIAN 94 (Ref. 29) and tested on a series of simple reactions that show a range of transition state geometries and reaction path properties: $H_2+F\rightarrow H+HF$; $CH_3O\rightarrow CH_2OH$; $SiH_2+H_2\rightarrow SiH_4$; $CH_2{=}CH_2+HF$; $CH_2{=}CH_2+CH_2{=}CH-CH{=}CH_2\rightarrow$cyclohexene (Diels–Alder reaction); $CH_2{=}CH_2+CH_2{=}CH-CH_3\rightarrow$pentene (ene reaction).

II. METHOD

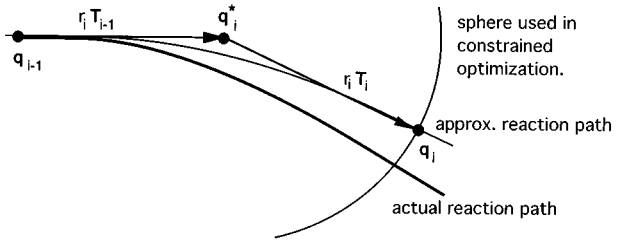
The present procedure refines an initial guess of the reaction path until the transition state and the steepest descent path are found. The initial guess of the points along the reaction path is obtained by interpolating between two (or three) input structures; the energy and gradient are calculated at each point and an empirical estimate of the Hessian is obtained for each point. The highest energy point on the path is chosen to optimize to the closest TS. This divides the path into two downhill segments. If the low energy end point of each segment is not already at a minimum, it can be optimized to the closest minimum. The remaining points on the path are optimized to lie on the steepest descent path. After each optimization step, the energies and gradients are calculated again, the Hessians are updated using the available gradients and the optimization is repeated until each point has converged.

A key feature of this approach is that the optimizations of the points on the path are guided by the neighboring points. For the putative transition state $\mathbf{q}_j$, the transition vector is assumed to be the tangent of the arc of circle passing through it and the neighboring point in each valley, $\mathbf{q}_{j-1}$, $\mathbf{q}_j$, and $\mathbf{q}_{j+1}$. This tangent is used to guide the quasi-Newton/eigenvector following optimization of the transition state (the conventional eigenvector following approach can have difficulties in maintaining the proper uphill search direction). For the steepest descent portion of the path, the second order Gonzalez–Schlegel path following method[16] is used; each segment the reaction path connecting a point $\mathbf{q}_i$ to its uphill neighboring point $\mathbf{q}_{i-1}$ is approximated by an arc of circle defined by the position of $\mathbf{q}_i$, $\mathbf{q}_{i-1}$, and the tangent to the path $\mathbf{T}_{i-1}$ (see Scheme 2). For each segment, a local quadratic surface about $\mathbf{q}_i$, $E(\mathbf{q}_i + \Delta\mathbf{q}_i) = E(\mathbf{q}_i) + \mathbf{g}_i^t\Delta\mathbf{q}_i + 1/2\Delta\mathbf{q}_i^t H_i\Delta\mathbf{q}_i$, is used to relax the point so that the steepest descent path differential Eq. (2) (Ref. 15) is satisfied, i.e., the tangent to the path $\mathbf{T}_i$ is the parallel to the gradient $\mathbf{g}_i$,

$$d\mathbf{q}/ds = -\mathbf{g}(\mathbf{q})/|\mathbf{g}(\mathbf{q})|. \tag{2}$$

The following steps provide an outline of the present procedure. Various aspects are discussed in more detail below.

(1) Using a two or three-structure input, construct an N-point approximation of the path by linear interpolation. Obtain an initial estimate of the Hessian at each point.
(2) Compute the energy and gradient for each point that moved significantly from its position in the previous path approximation. If allowed, update the trust radius for the optimizations and the path relaxation. On the first step, elect which points will be allowed to move toward the TS and minima.
(3) Update the Hessian for each point. First, update using gradients at neighboring points along the current path. For the first four relaxation steps, if the points along the path are too far apart (typically, separated by more than 0.6 a.u.), update the second derivative along the path



Scheme 2: Segment of the reaction path approximated by an arc of circle passing through $\mathbf{q}_{i-1}$ and $\mathbf{q}_i$ with tangents $\mathbf{T}_{i-1}$ and $\mathbf{T}_i$. The approximation can be refined by displacing $\mathbf{q}_i$ on the surface of a sphere centered at $\mathbf{q}_i^*$ and minimizing its energy.

using a quintic fit to the energies and gradients along the path. Second, update using the gradients from the current and previous position for each point.

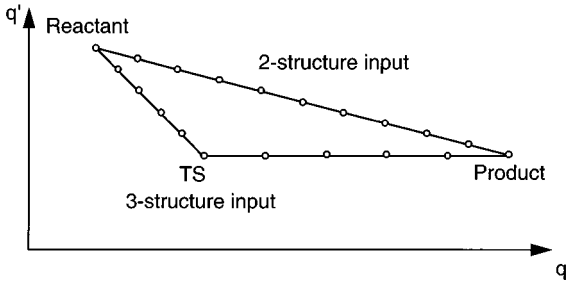
(4) Move the end points toward the minima, if allowed, using conventional quasi-Newton algorithm (the default algorithm in GAUSSIAN 94 (Ref. 29) is RFO (Ref. 30) preceded by a linear search whenever possible).
(5) Use the tangent to the arc of circle passing through the highest point and its two immediate neighbors to guide the eigenvector-following TS search.
(6) Perform microiterations for each point on the path using the local quadratic surface until convergence criteria are met or the trust radius exceeded.
(7) Check the convergence of the displacement and gradient for the optimizations and the displacement and perpendicular gradient reaction path refinement.
(8) If not converged, update the positions of the points on the path and go to step 2; if converged, print summary and stop.

A. Coordinates

To describe a reaction path, internal coordinates are preferred over Cartesian coordinates, since this avoids problems with rigid body rotations and translations along the path. Because a set of nonredundant (z-matrix) internal coordinates most appropriate for describing the bonding in one region of the potential energy surface can be a very poor set in another region, redundant internal coordinates appear to be better suited for treating the whole reaction path. The present method utilizes the redundant internal coordinates in GAUSSIAN 94.[29] These are based on the identification of bonds using covalent radii; valence angles and dihedral angles are generated for all appropriately bonded triplets and quadruplets of atoms. The reactantlike structure at one end of the path generates one set of redundant internal coordinates, and the productlike structure at the other end generates a second set. The union of these two sets defines the redundant internal coordinates used throughout the path relaxation process. The user can alter the final set of redundant internal coordinates by adding or removing coordinates. The final set of coordinates should be inspected closely for potential problems, e.g., valence angles and dihedrals incompatible with the symmetry of the reaction path, or two linear angles instead of a valance angle and a dihedral.

B. Input

Either two or three structures can be used to specify the initial approximation to the reaction path. If two structures are given, the initial guess of the reaction path is obtained by linear interpolation between the reactant and product in redundant internal coordinates. If a third structure is input as a guess for the transition state, the initial path is then approximated by two linear interpolations, first between reactant and TS and second, between TS and product (see Scheme 3) and an equal number of points is distributed on each leg of the path. The final set of redundant internal coordinates is the same in both the two-structure and three-structure input. Depending on the shape of the potential energy surface, the initial guess of the TS may not be the highest point along the initial path; either the highest energy point or the initial guess for the TS can be chosen to optimize to the TS.



Scheme 3: Initial approximation to the reaction path.

C. Path relaxation

The minima $\mathbf{q}_k$ are optimized using the standard quasi-Newton method in GAUSSIAN 94. The transition state $\mathbf{q}_j$ is optimized using the standard quasi-Newton/eigenvector following method, where the choice of the eigenvector is guided by the tangent at the transition state. This tangent is computed from the arc of a circle pasting through $\mathbf{q}_{j-1}$, $\mathbf{q}_j$, and $\mathbf{q}_{j+1}$,

$$\mathbf{T}'_j=(\mathbf{q}_{j+1}-\mathbf{q}_j)/|\mathbf{q}_j-\mathbf{q}_{j+1}|^2-(\mathbf{q}_{j-1}-\mathbf{q}_j)/|\mathbf{q}_j-\mathbf{q}_{j-1}|^2;$$

$$\mathbf{T}_j=\mathbf{T}'_j/|\mathbf{T}'_j|. \tag{3a}$$

After each optimization step, the path connecting two extrema is relaxed in a series of microiterations. Equation (2) is replaced by a set of finite difference equations, Eq. (3b), for the points between the TS $\mathbf{q}_j$ and minimum $\mathbf{q}_k$ ($j<k$),

$$\mathbf{T}_i=-\mathbf{g}_i/|\mathbf{g}_i|+\mathrm{Error}(i). \tag{3b}$$

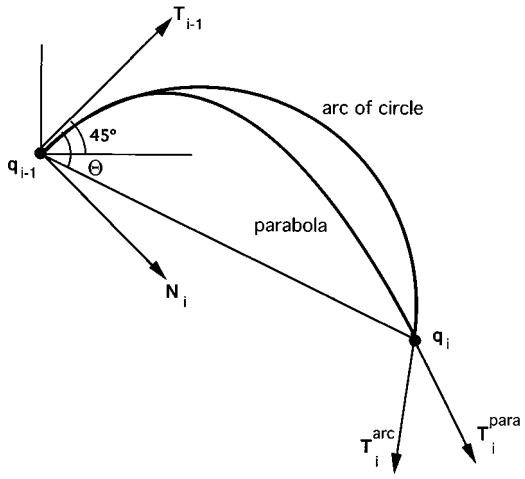
In the spirit of the Gonzalez–Schlegel reaction path following algorithm, the path between $\mathbf{q}_{i-1}$ and $\mathbf{q}_i$ is represented by an arc of a circle, as shown in Scheme 2 (note that $\mathbf{q}_{i-1}$ is the uphill point). Given $\mathbf{q}_{i-1}$, $\mathbf{q}_i$, and $\mathbf{T}_{i-1}$ (the normalized tangent at $\mathbf{q}_{i-1}$), one can readily construct $\mathbf{T}_i$ (the normalized tangent at $\mathbf{q}_i$), since two tangents to a circle form an isosceles triangle,

$$r_i\mathbf{T}_{i-1}+r_i\mathbf{T}_i=(\mathbf{q}_i-\mathbf{q}_{i-1});$$

$$r_i=(\mathbf{q}_i-\mathbf{q}_{i-1})^2/[2\mathbf{T}_{i-1}\cdot(\mathbf{q}_i-\mathbf{q}_{i-1})], \tag{3c}$$

$$\mathbf{T}_i^{\mathrm{arc}}=[(\mathbf{q}_i-\mathbf{q}_{i-1})-r_i\mathbf{T}_{i-1}]/r_i.$$

Whenever the angle formed by $(\mathbf{q}_i-\mathbf{q}_{i-1})$ and $\mathbf{T}_{i-1}$ exceeds 45°, an arc of a circle is probably a poor approximation to the reaction path. In this case, the path can be approximated by a parabola passing through $\mathbf{q}_{i-1}$ and $\mathbf{q}_i$ with a tangent $\mathbf{T}_{i-1}$ at $\mathbf{q}_{i-1}$, in the manner shown in Scheme 4.



Scheme 4: When the change in direction of the reaction path is too large, the segment of the reaction path is approximated by a parabola passing through $\mathbf{q}_{i-1}$ and $\mathbf{q}_i$ with tangents $\mathbf{T}_{i-1}$ and $\mathbf{T}_i$ rather than an arc of a circle.

$$\mathbf{T}_i^{\text{para}} = \mathbf{N}_i - \tan(\theta - 45)(\mathbf{T}_{i-1} - \mathbf{N}_i);$$

$$\mathbf{T}_1^{\text{para}} = \mathbf{T}_i'^{\text{para}} / |\mathbf{T}_i'^{\text{para}}|, \tag{3d}$$

where

$$\mathbf{N}_1' = (\mathbf{q}_i - \mathbf{q}_{i-1}) - (\mathbf{q}_i - \mathbf{q}_{i-1})^t \mathbf{T}_{i-1}\mathbf{T}_{i-1});$$

$$\mathbf{N}_i' = \mathbf{N}_i' / |\mathbf{N}_i'|, \tag{3e}$$

and

$$\cos(\theta) = (\mathbf{q}_i - \mathbf{q}_{i-1})^t \mathbf{T}_{i-1} / |\mathbf{q}_i - \mathbf{q}_{i-1}|. \tag{3f}$$

The construction in Scheme 4 allows for a smooth change from the tangent to the arc of a circle to the tangent of a parabola as the angle between $(\mathbf{q}_i - \mathbf{q}_{i-1})$ and $\mathbf{T}_{i-1}$ increases beyond 45°. If this construction is needed for more than one or two points of the converged reaction path, then more points should be used for a more accurate representation of the path.

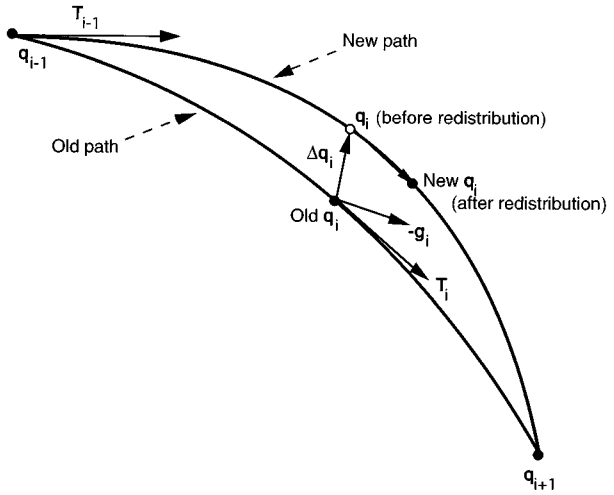
The constraints on the spacing of the points along the path are given by Eq. (3g), where the a_i's are the desired ratios of distances between the points in redundant internal coordinates ($a_i = 1$ for all i, corresponding to equal spacing is used here),

$$|\mathbf{q}_i - \mathbf{q}_{i-1}| / a_i = |\mathbf{q}_i - \mathbf{q}_{i+1}| / a_{i+1}. \tag{3g}$$

The convergence criteria used for the path refinement are similar to the regular criteria for the optimization; less than 0.000 45 and 0.000 30 a.u. for the maximum component and root mean square of the gradient perpendicular to the tangent, less than 0.001 8 and 0.001 2 a.u. for the maximum component and root mean square of the displacement, respectively.

D. Microiterations

The equations for the path relaxation (3b) and the spacing constraints (3g) form a strongly coupled set of equations. With care, these can be solved in a series of microiterations, as shown in Scheme 5. First, the constrained optimization for the reaction



Scheme 5: Microiteration step for one point showing relaxation of the path $\Delta\mathbf{q}_i$ and redistribution along the path.

path following is solved on the model quadratic surface about each point. Then the points along the approximate path are redistributed to satisfy the spacing constraints. This microiteration is repeated until all the displacements and the components of the gradient perpendicular to the tangent are below threshold (default, half of the corresponding criteria for the stationary point search), or until the total displacement for each point has exceeded the trust radius (default 0.3 a.u.).

The equation for the constrained optimization for the reaction path following is

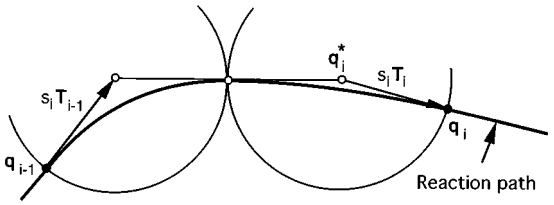
$$\mathbf{g}_i + \mathbf{H}_i \Delta\mathbf{q}_i = \lambda_i (s_i \mathbf{T}_i + \Delta\mathbf{q}_i) \tag{4a}$$

and corresponds to the search for a point lying on a sphere of radius s_i such that the normal and the gradient are parallel. The Lagrangian multiplier λ_i is chosen so that $|s_i\mathbf{T}_i + \Delta\mathbf{q}_i| = s_i$ (i.e., the point lies on the sphere) and λ_i is negative for the steepest descent direction.[26] Note that the Hessian for each point needs to be diagonalized only once, since it is constant during the microiterations.

In practice, the iteration of Eq. (4a) needs to be stabilized to prevent oscillations. First, a small maximum step size is used in each microiteration, typically two to four times the convergence criterion for displacements. Secondly, the iterations are damped by replacing the tangent $\mathbf{T}_i$ (calculated from the arc of a circle or a parabola) with the average of the tangent and the normalized gradient,

$$\tilde{\mathbf{T}}_i' = \mathbf{T}_i - \mathbf{g}_i / |\mathbf{g}_i|; \quad \tilde{\mathbf{T}}_i = \tilde{\mathbf{T}}_i' / |\tilde{\mathbf{T}}_i'|. \tag{4b}$$

When the microiterations have converged, the reaction path between $\mathbf{q}_{i-1}$ and $\mathbf{q}_i$ can be represented by an arc of a circle or by a parabola. However, during the course of the microiterations, it is not always possible to link $\mathbf{q}_{i-1}$ to $\mathbf{q}_i$ by a single arc of circle tangent to $\mathbf{T}_{i-1}$ and $\mathbf{T}_i$. The path can always be represented by two arcs, as shown in Scheme 6, where the radius s_i is the positive solution of $|\mathbf{q}_{i-1} + s_i\mathbf{T}_{i-1} - \mathbf{q}_i + s_i\mathbf{T}_i| = 2s_i$.



Scheme 6: Reaction path segment represented by two arcs during the microiterations.

Both the damping given by Eq. (4b) and the use of the two smaller arcs of circles constructed in Scheme 6 lead to more conservative steps and better stability in the microiterations. Despite these precautions, hysteris-type oscillations can still occur. To overcome these, the most downhill point involved in the oscillation needs to have its displacement damped heavily.[27] In the present implementation, whenever the angle formed by $\Delta\mathbf{q}_{i-1}$ and $\Delta\mathbf{q}_i$ is greater than 160°, $\Delta\mathbf{q}_i$ is zeroed.

After each microiteration relaxing the path, the points along the path need to be redistributed to maintain the requested spacing. First $(\mathbf{q}_{i-1} - \mathbf{q}_{i+1})$ is projected out of displacement $\Delta\mathbf{q}_i$. Then the position of each point is iteratively adjusted along the path by Eq. (5) until all the constraints are satisfied,

$$\mathbf{q}_{i\text{new}} = \mathbf{q}_i + \alpha_i(\mathbf{q}_{i-1} - \mathbf{q}_{i+1}) \quad j < i < k, \tag{5a}$$

where α_i is chosen such that

$$|\mathbf{q}_{i\text{new}} - \mathbf{q}_{i-1}| = |\mathbf{q}_{i\text{new}} - \mathbf{q}_{i+1}|. \tag{5b}$$

To avoid problems when the end point $\mathbf{q}_k$ is not optimized or not on the steepest descent path, the redistribution of the point $\mathbf{q}_{k-1}$ must be treated differently, and Eq. (5b) is replaced by

$$|\mathbf{q}_{k-1\text{new}} - \mathbf{q}_{k-2}| = (\mathbf{q}_k - \mathbf{q}_{k-1\text{new}})^t \cdot (\mathbf{q}_{k-1\text{new}} - \mathbf{q}_{k-2})/ |\mathbf{q}_{k-1\text{new}} - \mathbf{q}_{k-2}|. \tag{5c}$$

E. Trust radius update

Control of the step size for the optimization and the path relaxation is essential for proper performance. The trust radius can be fixed or can be updated before the Hessian update and the microiterations are performed. For the points along the path (excluding transition state and minima) the trust radius τ_i is updated according to Eq. (6),

$$\beta_i = (\mathbf{g}_{i\text{new}} - \mathbf{g}_{i\text{old}})^t \cdot [H_i(\mathbf{q}_{i\text{new}} - \mathbf{q}_{i\text{old}})]/ |H_i(\mathbf{q}_{i\text{new}} - \mathbf{q}_{i\text{old}})|^2, \tag{6a}$$

$$\tau_{i\text{new}} = 0.5\tau_{i\text{old}} \quad \text{if } (\beta_i < 0.7 \text{ or } \beta_i > 1.3), \tag{6b}$$

$$\tau_{i\text{new}} = \sqrt{2}\tau_{i\text{old}} \quad \text{if } (\beta_i > 0.85 \text{ and } \beta_i < 1.25). \tag{6c}$$

For the transition state and the minima the trust radius is updated in the same manner as for regular optimizations in GAUSSIAN 94. To keep a consistent and conservative trust radius update along the path, the root mean square of the updated trust radii is assigned to each point.

F. Hessian update

The success of the reaction path relaxation procedure depends critically on the quality of the Hessian at each point. However, calculation of the exact Hessian at each point would make the algorithm inappropriate for most levels of theory. The present method starts with an empirically estimated Hessian[31,32] for each point and updates it using the available gradients. In contrast to single-structure optimization, a lot of information about the shape of the potential energy surface along the path is available and the standard methods of updating of the Hessian can be improved. The updating is done in two stages; the first uses the gradients at neighboring points along the path, and the second uses the gradients for a given point at the current and previous iterations of relaxing the path. This provides better estimates of the Hessians and thus better convergence, especially for the TS search.

In the first stage of updating, the points along the path might be too far apart to insure a good update using the standard formulas. For the first four iterations only, a quintic polynomial is used to update the second derivative along the path. The path is approximated by an arc of a circle passing through $\mathbf{q}_{i-1}$, $\mathbf{q}_i$, and $\mathbf{q}_{i+1}$. The energies and gradients along the path at $\mathbf{q}_{i-1}$, $\mathbf{q}_i$, and $\mathbf{q}_{i+1}$ are fitted to a quintic polynomial and yield a second derivative γ_i along the path at $\mathbf{q}_i$. The updated Hessian at $\mathbf{q}_i$ is approximated by

$$H_{i\text{new}} = (1 - \mathbf{T}_i\mathbf{T}_i^t)^t H_{i\text{old}}(1 - \mathbf{T}_i\mathbf{T}_i^t) + \gamma_i \mathbf{T}_i\mathbf{T}_i^t, \tag{7a}$$

where the tangent at $\mathbf{q}_i$ is given by

$$\mathbf{T}_i' = (\mathbf{q}_{i+1} - \mathbf{q}_i)/|\mathbf{q}_i - \mathbf{q}_{i+1}|^2 - (\mathbf{q}_{i-1} - \mathbf{q}_i)/ |\mathbf{q}_i - \mathbf{q}_{i-1}|^2; \quad \mathbf{T}_i = \mathbf{T}_i'/|\mathbf{T}_i'|. \tag{7b}$$

In the second stage, the gradients from the current and previous iteration of each point are used to update the Hessian. For the points elected to optimize toward minima, a BFGS update[33–36] is employed. For points along the path, a Powell update could be used so that the Hessian is not forced to be positive definite. In the spirit of Bofill's update,[37] some improvement over the Powell update can be obtained using a BFGS/Powell mixture,

$$H_{\text{new}} = H_{\text{old}} + (1 - \phi)\Delta H_{(\text{BFGS})} + \phi\Delta H_{(\text{Powell})}, \tag{8a}$$

$$\Delta H_{(\text{BFGS})} = (\Delta\mathbf{g})(\Delta\mathbf{g})^t/\Delta\mathbf{g}^t\Delta\mathbf{q} - (H_{\text{old}}\Delta\mathbf{q}) \times (H_{\text{old}}\Delta\mathbf{q})^t/\Delta\mathbf{q}^t H_{\text{old}}\Delta\mathbf{q}, \tag{8b}$$

$$\Delta H_{(\text{Powell})} = [(\Delta\mathbf{g} - H_{\text{old}}\Delta\mathbf{q})(\Delta\mathbf{q})^t + (\Delta\mathbf{q})(\Delta\mathbf{g} - H_{\text{old}}\Delta\mathbf{q})^t]/\Delta\mathbf{q}^t\Delta\mathbf{q} - (\Delta\mathbf{g} - H_{\text{old}}\Delta\mathbf{q})^t\Delta\mathbf{q}(\Delta\mathbf{q})(\Delta\mathbf{q})^t/(\Delta\mathbf{q}^t\Delta\mathbf{q})^2, \tag{8c}$$

with

$$\phi=(\mathbf{g}^t_{\rm new}.\Delta\mathbf{q}/(|\mathbf{g}_{\rm new}||\Delta\mathbf{q}|))^2 \quad \text{if} \quad \Delta\mathbf{g}^t.\Delta\mathbf{q}>0$$

$$=1 \quad \text{if} \quad \Delta\mathbf{g}^t.\Delta\mathbf{q}<0 \quad \text{or} \quad \Delta\mathbf{q}^t H_{\rm old}\Delta\mathbf{q}<0. \tag{8d}$$

This Hessian updating mixture favors a BFGS update (positive semidefinite change) for displacements nearly perpendicular to the path and a Powell update along the path.

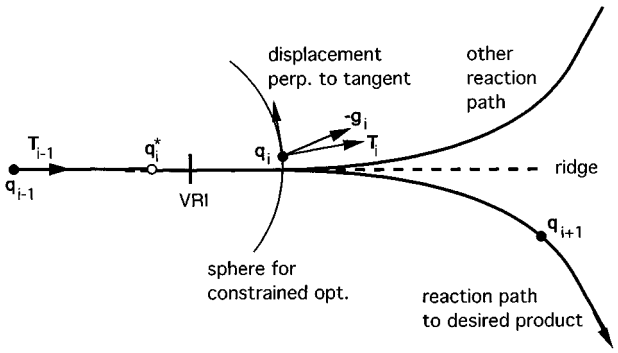
For the point elected to optimize toward TS, a Bofill update (a mixture of the Murtagh–Sargent update[38] and Powell update[39]) would be appropriate. Again, the update can be made more effective by using a BFGS-type update if the displacement is perpendicular to the transition vector. For the fifth and later steps the Hessian for the TS is updated using Eqs. (8a)–(8c) and Eq. (8e),

$$\phi=(\mathbf{T}^t_j.\Delta\mathbf{q}/|\Delta\mathbf{q}|)^2 \quad \text{if} \quad \Delta\mathbf{g}^t.\Delta\mathbf{q}>0$$

$$=1 \quad \text{if} \quad \Delta\mathbf{g}^t.\Delta\mathbf{q}<0 \quad \text{or} \quad \Delta\mathbf{q}^t H_{\rm old}\Delta\mathbf{q}<0. \tag{8e}$$

G. Branching of reaction paths

On some potential energy surfaces, the reaction path branches into two paths that descend into valleys on either side of a ridge. Branching along a steepest descent path is characterized by a valley ridge inflection (VRI) point, where one or more eigenvalues of the projected Hessian are zero.[40] Valley ridge inflection points are often encountered when the point group symmetry is lowered along the reaction path, such as in the $CH_2{=}CH_2{+}HF$ reaction, where the C_s transition state leads via C_1 structures to staggered fluoroethane. The reaction path bifurcation can be handled in a number of ways.[41–43] A steepest descent path near the ridge will follow the ridge for a considerable distance before descend into one of the two valleys. It is therefore important that the relaxation process keeps the points either on the ridge or on the product side of the ridge. As illustrated in Scheme 7, whenever the point being relaxed is on the side of the ridge opposite to the desired product, even slightly, then stepping in the direction perpendicular to the tangent takes it further down the wrong side, irrespective of the choice for the tangent ($\mathbf{T}_i$ parallel to $\mathbf{q}_i-\mathbf{q}_1^*$, $\mathbf{q}_i-\mathbf{q}_{i-1}$, or $\mathbf{q}_{i+1}-\mathbf{q}_{i-1}$, etc.). However, the sequence of microiterations described above can bring the point back on the ridge.



Scheme 7: Reaction path relaxation in the neighborhood of a valley-ridge inflection (VRI) point.

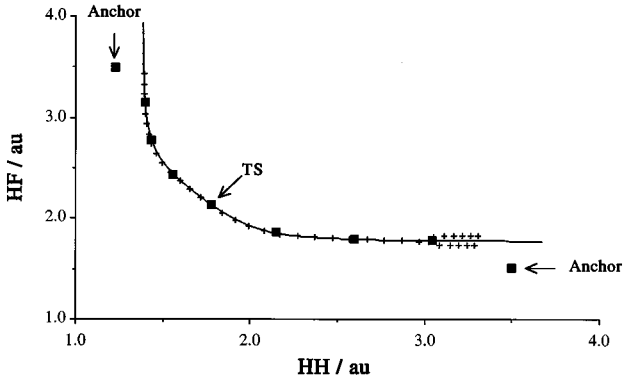


FIG. 1. Reaction path for $H_2+F\rightarrow H+HF$ in internal coordinates without mass weighting (solid line, GS algorithm; crosses, Euler algorithm; filled squares, 7 point relaxation).

III. APPLICATIONS

The procedure outlined above has been implemented in the development version of GAUSSIAN 94 (Ref. 29) and has been tested for a number of chemical reactions. The calculations were performed at the HF/3-21G level of theory. Unless stated otherwise, 7 points, including transition state and minima, were used to represent the reaction path.

The efficiency and accuracy of the algorithm was tested against the widely used Gonzalez–Schlegel reaction path following method using *Z*-matrix internal coordinates. Just as the steepest descent path in internal coordinates depends on the choice of internal coordinates, the present algorithm will converge to different path approximations depending on the choice of redundant internal coordinates. However, the stationary points along the path will be the same.

A. $F+H_2\rightarrow HF+H$

Figure 1 depicts the non-mass-weighted reaction path for this simple bimolecular reaction representative of atom abstraction, S_N2 reactions and group transfer reactions. The reaction coordinate was followed using both the Gonzalez–Schlegel (GS) second order algorithm and using the Euler

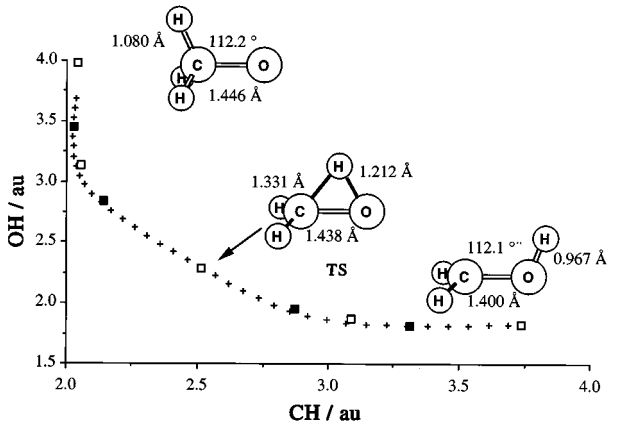


FIG. 2. Reaction path for $CH_2OH\rightarrow CH_3O$ in redundant internal coordinates without mass weighting (crosses, Euler algorithm; open squares, 5 point relaxation; filled squares, 7 point relaxation).

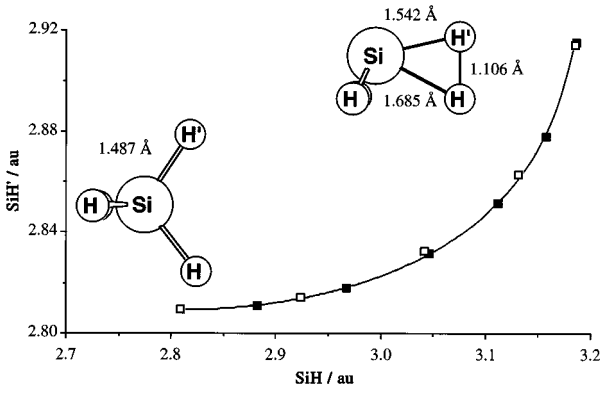


FIG. 3. Reaction path for $SiH_4 \rightarrow [SiH_2+H_2]^{\dagger}$ in nonredundant internal coordinates without mass weighting (solid line, GS algorithm; open squares, 5 point relaxation; filled squares, 7 point relaxation).

algorithm with a step size of 0.1 a.u. The two algorithms agree up to large HH distances where the Euler algorithm begins to oscillate. Starting from a linear synchronous transit (LST) approximation with anchor points away from the IRC path, the relaxation procedure converges quickly onto the steepest descent path in six iterations, the same number steps needed for the optimization of the TS. The agreement with the IRC obtained by the GS method is very good.

B. $CH_2OH \rightarrow CH_3O$

This 1,2 hydrogen shift reaction is a simple model for 1,*n* group shift reactions. The transition state is a 3-member ring which can be problematic in regular nonredundant internal (*Z*-matrix) TS optimizations.[1] As shown in Table I, using redundant internal coordinates and the 3-structure input found in Ref. 13, the TS is optimized in 8 steps, the reactant and product are optimized in 4 and 7 steps, respectively, and the path is fully converged after 8 steps. Also, there is little difference in the efficiency of a 5-point and 7-point relaxation, in both cases the path is converged in 8 steps. Figure 2 shows the relaxed reaction path and the IRC obtained by Euler integration in redundant internal coordinates. The agreement is very good for both the 7-point and 5-point paths.

C. $SiH_4 \rightarrow SiH_2+H_2$

This 1,1 elimination or insertion reaction is an other example for 3-member ring transition state. Just as for the 1,2 shift reaction, optimizing this TS can be challenging for regular *Z*-matrix optimizations.[1] The combined use of redundant internal coordinates and guided eigenvector following algorithm makes the TS search very efficient. Only 7 steps are needed to optimize the TS, 3 steps for SiH_4, and 8 steps for the reaction path. Nonredundant internal coordinates were also used to assess the quality of description given by the relaxation method. Figure 3 shows the agreement between the GS IRC following algorithm and the relaxed path in nonredundant internal coordinates on the reactant side of the reaction. The reaction path is fairly curved with a very late TS (i.e., close to the SiH_2+H_2 cluster minimum) and an equally good representation of the path is given using 5 or 7 points.

D. $CH_3-CH_2F \rightarrow CH_2=CH_2+HF$

This simple example of a reaction with a 4-membered ring transition state might be one of the more challenging reactions of our test suite. A loose C_{2v} cluster with the HF bond perpendicular to the CC bond is present on the product side. On the reactant side the actual IRC as obtained by the GS algorithm converges to a staggered conformation of fluoroethane. Starting from the three C_s structures of Ref. 13 and enforcing C_s symmetry throughout the relaxation procedure, the reactant and product are found after 28 and 5 steps, the TS is found after 11 steps, and the path is fully converged after 28 steps. The limiting factor for the performance of the algorithm appears to be the poor choice for the initial guess of the product (HF bond lies parallel to the CC bond). Using a better guess for product structure (the HF bond is perpendicular to the CC bond) yields much better results. The TS is found in 11 steps, the reactant is found in 5 steps, the product is found in 8 steps, and the path is fully converged in 14 steps.

Figure 4 shows the reaction path in redundant internal coordinates from the transition state descending to the staggered fluoroethane. The relaxed path successfully follows the ridge and then, past the valley ridge inflection point turns towards the correct product structure. Also shown in Fig. 4 is the reaction path for the methyl group rotation. Figure 5 compares the relaxed path with the reaction path in nonredundant internal coordinates as obtained by the GS algorithm using a 0.1 a.u. step size. Even though the path is very curved, a very good agreement is found. It should be noted that, with a step size of 0.5 a.u., the GS algorithm fails to converge at CF distances greater than 4.16 a.u. This is likely due to two facts. First, the GS algorithm carries the Hessian from one point to the next, and for large step sizes this causes problems. Second, the GS algorithm first steps in the direction of the gradient and this becomes inappropriate for

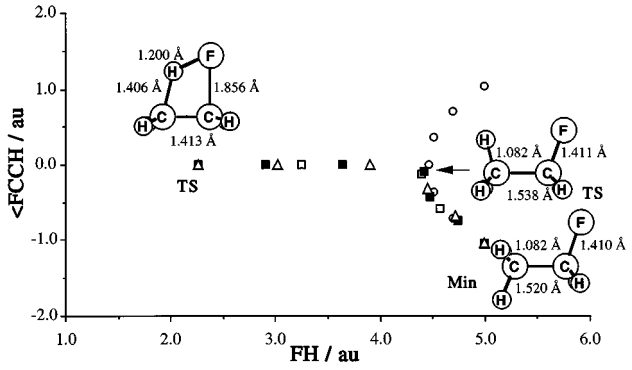


FIG. 4. Reaction path for $[C_2H_4+HF]^{\dagger} \rightarrow C_2H_5F$ (open squares, 5 point relaxation; triangles, 6 point relaxation; filled squares, 7 point relaxation) and reaction path for methyl rotation (open circles, 7 point relaxation).

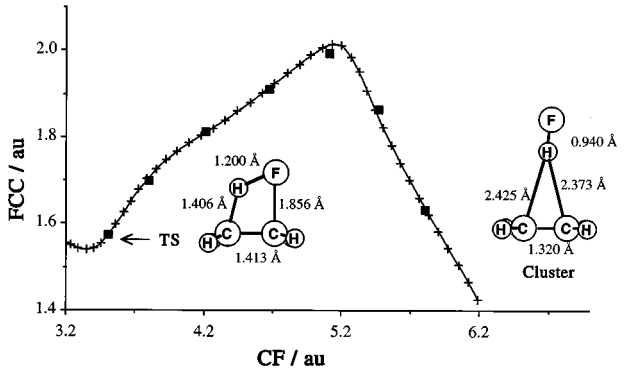


FIG. 5. Reaction path for $[C_2H_4+HF]^{\dagger} \rightarrow$cluster in nonredundant internal coordinates without mass weighting (crosses, GS algorithm; filled squares, 7 point relaxation).

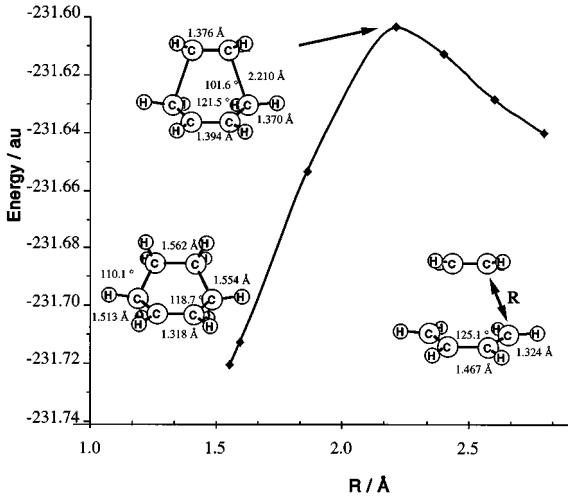


FIG. 6. Energy profile for the Diels–Alder reaction obtained by a 7 point path relaxation in redundant internal coordinates without mass weighting.

cases of large path curvature and large step size. This reaction path is clearly a challenging one. However, it is important to note that the use of the mixed Hessian update of Eq. (8) over the Powell update decreases the number of steps by 35% [17.3 steps per point using Powell update vs 11.3 steps per point using Eq. (8)].

E. Diels–Alder reaction

Reference 1 showed that the combined use of redundant internal coordinates and the tangent to the quadratic synchronous transit (QST) improves dramatically the transition state

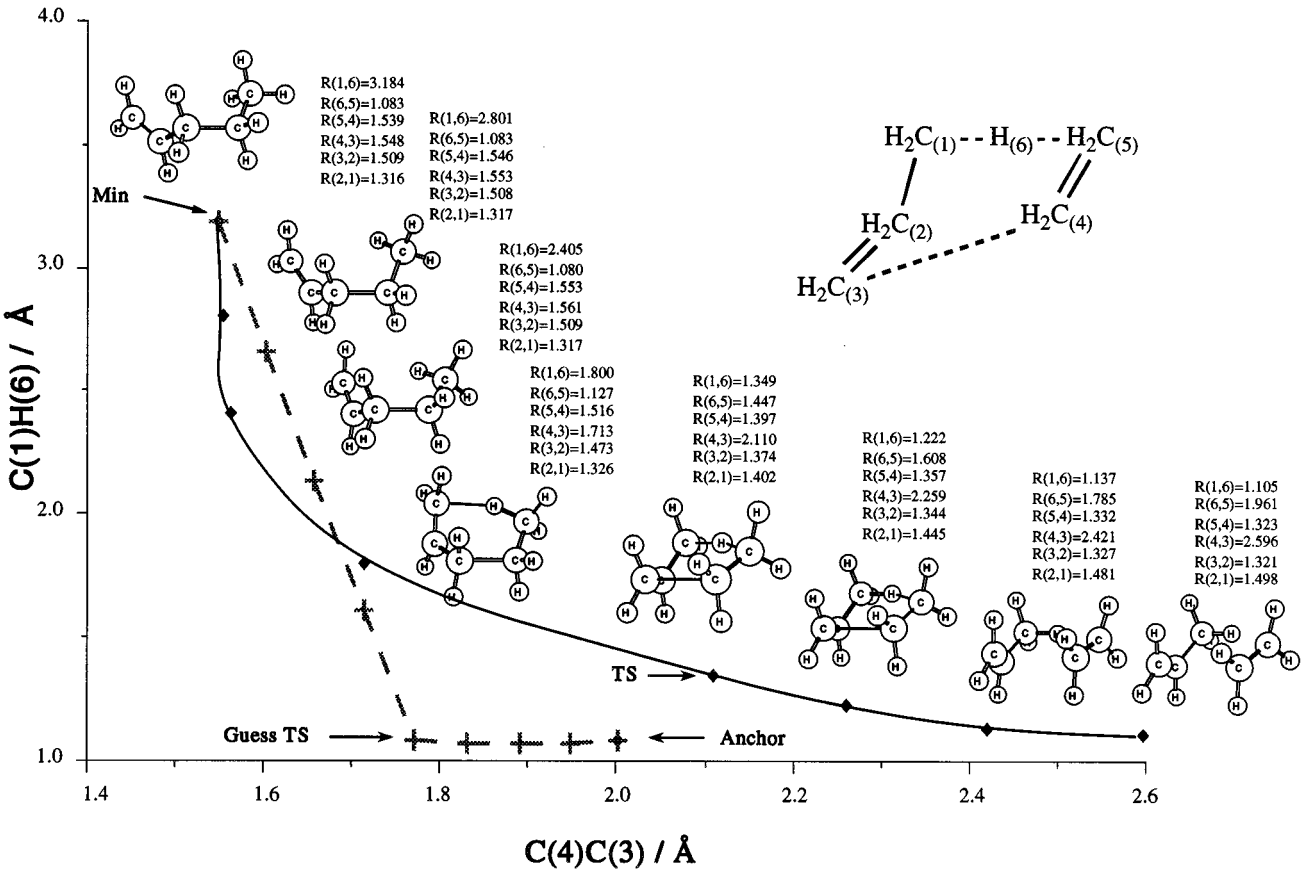


FIG. 7. Reaction path for the ene reaction in redundant internal coordinates without mass weighting (dashed line, interpolated starting guess for the path, filled diamonds, 7 point relaxation). Selected bond lengths (in Å) are listed for each converged structure.

TABLE I. Number of steps required to converge to the stationary points and to the reaction path.

Reaction[a]	Path relaxation[b]			Regular[c]	
	TS	R,P	Total	TS[d]	Total[e]
$CH_3O \rightarrow CH_2OH$	8	4,7	51	12	57 (15)
$SiH_4 \rightarrow SiH_2+H_2$	7	3	47	11	59 (16)
$C_2H_4+HF \rightarrow$ cluster[f]	11	5,8	73	16	94 (26)
Diels–Alder reaction	7	6	41	56	122 (22)
ene reaction[g]	14		101	fail	>126 (36)

[a]Structure given in Ref. 13; all calculations at the HF/3-21G level of theory.
[b]Path represented by 7 points, not including anchor poins for bimolecular reactions.
[c]Using one structure and nonredundant internal coordinates.
[d]Empirical estimate of the Hessian with two to four preliminary steps to calculate key rows and columns of the Hessian by numerical differentiation.
[e]Number of gradient calculations required to determine the path using the GS algorithm, estimated as three times the number of points along the path. The estimate of the number of points is based upon the distance between reactant and TS plus the distance between TS and product and a step size of 0.3 a.u.
[f]On the C_s potential energy surface. The guess structure for the cluster modified from Ref. 13 as discussed in the text.
[g]The optimized structure of pentene was used as an anchor point for the relaxation.

optimization for this reaction (56 steps required for a regular Z-matrix optimization, compared to 14 for redundant internal QST3). This good performance is improved further by the reaction path relaxation procedure; 7 steps are required to optimize the TS, 6 steps to optimize cyclohexene, and 8 steps to reach convergence on the steepest-descent path. The quicker TS optimization in the relaxation procedure compared to QST3 can be mainly attributed to a better Hessian in the first few steps and to a better search direction. Figure 6 shows the energy profile for the reaction as obtained by the 7-point relaxation procedure.

F. Ene reaction

The ene reaction proceeds through a six-membered ring transition state where one bond is broken and two are formed (Fig. 7). Not surprisingly, the search for the TS is very challenging using the regular procedure (see Table I). The search only converges when started with an exact Hessian. Using the present procedure the TS is found in 14 steps. The structures for the reactant and products found in Ref. 13 are not meant to be good approximations of the actual reactant and product structures, but rather are only intended to provide the TS search with good transition vector. Since the optimization of pentene from the structure in Ref. 13 takes in excess of 30 steps, the relaxation procedure was started with the converged reactant geometry. The path is fully converged after 15 steps. As for the Diels–Alder reaction an improvement is found in the TS search over the redundant internal QST3 procedure which takes 18 steps.[1]

IV. CONCLUSIONS

We have outlined an efficient procedure to find the transition state and reaction path connecting two structures. An approximate path represented by a small number of points is iteratively relaxed until it converges to the transition state and the steepest descent reaction path. The method combines the efficiency of redundant internal coordinates for geometry optimization, the robustness of the quadratic synchronous transit guided approach for finding transition states and the stability of the Gonzalez–Schlegel algorithm for reaction path following. Only gradients are used; analytical Hessians are not necessary. The procedure appears to be particularly well suited for studying reactions involving flexible and polycyclic transition states.

ACKNOWLEDGMENTS

This work was supported by grants from the National Science Foundation (CHE 94-00678) and Gaussian, Inc.

[1] C. Peng, P. Y. Ayala, H. B. Schlegel, and M. J. Frisch, J. Comp. Chem. **17**, 49 (1996).
[2] P. Pulay, G. Fogarasi, F. Pang, and J. E. Boggs, J. Am. Chem. Soc. **101**, 2550 (1979).
[3] P. Pulay and G. Fogarasi, J. Chem. Phys. **96**, 2856 (1992).
[4] G. Fogarasi, X. Zhou, P. Taylor, and P. Pulay, J. Am. Chem. Soc. **114**, 8191 (1992).
[5] J. Baker, J. Comput. Chem. **14**, 1085 (1993).
[6] H. B. Schlegel, in *Modern Electronic Structure Theory*, edited by D. R. Yarkony (World Scientific, Singapore, 1995).
[7] H. B. Schlegel, Adv. Chem. Phys. **67**, 249 (1987).
[8] J. D. Head and M. C. Zerner, Adv. Quantum Chem. **20**, 239 (1989).
[9] S. Bell and J. S. Crighton, J. Chem. Phys. **80**, 2464 (1984).
[10] H. B. Schlegel, J. Comput. Chem. **3**, 214 (1982).
[11] M. L. McKee and M. Page, Rev. Comp. Chem. **4**, 35 (1993).
[12] T. A. Halgren and W. N. Lipscomb, Chem. Phys. Lett. **49**, 225 (1977).
[13] C. Peng and H. B. Schlegel, Isr. J. Chem. **33**, 449 (1994).
[14] J. Simons and J. Nichols, Int. J. Quantum Chem. Symp. **24**, 263 (1990), and references therein; present implementation from J. Baker, J. Comput. Chem. **7**, 385 (1986).
[15] K. Fukui, Acc. Chem. Res. **14**, 363 (1981).
[16] C. Gonzalez and H. B. Schlegel, J. Chem. Phys. **94**, 5523 (1990).
[17] K. Muller and L. D. Brown, Theor. Chim. Acta **53**, 75 (1979).
[18] M. Page, C. Doubleday, and J. W. McIver, J. Chem. Phys. **93**, 5634 (1990).
[19] K. Ishida, K. Morokuma, and A. Komornicki, J. Chem. Phys. **66**, 2153 (1977).
[20] M. W. Schmidt, M. S. Gordon, and M. Dupuis, J. Am. Chem. Soc. **107**, 2585 (1985).
[21] C. Gonzalez and H. B. Schlegel, J. Chem. Phys. **95**, 5853 (1991).
[22] P. G. Jaisen and R. Shepard, Int. J. Quantum Chem. Symp. **22**, 183 (1988).
[23] R. Elber and M. Karplus, Chem. Phys. Lett. **139**, 375 (1987).
[24] R. Czerminski and R. Elber, Int. J. Quantum Chem. Symp. **24**, 167 (1990).
[25] R. Czerminski and R. Elber, J. Phys. Chem. **92**, 5580 (1990).
[26] C. Choi and R. Elber, J. Phys. Chem. **94**, 751 (1991).
[27] S. S-L. Chiu, J. W. McDouall, and I. H. Miller, J. Chem. Soc. Faraday Trans. **90**, 1575 (1994).
[28] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, in *Numerical Recipes* (Cambridge University Press, Cambridge, 1986).
[29] M. J. Frisch, G. W. Trucks, H. B. Schlegel, P. M. W. Gill, B. G. Johnson, M. A. Robb, J. R. Cheeseman, T. Keith, G. A. Petersson, J. A. Montgomery, K. Raghavachari, M. A. Al-Laham, V. G. Zakrzewski, J. V. Ortiz, J. B. Foresman, J. Cioslowski, B. B. Stefanov, A. Nanayakkara, M. Challacombe, C. Y. Peng, P. Y. Ayala, W. Chen, M. W. Wong, J. L. Andres, E. S. Replogle, R. Gomperts, R. L. Martin, D. J. Fox, J. S. Binkley, D. J. Defrees, J. Baker, J. P. Stewart, M. Head-Gordon, C. Gonzalez, and J. A.

Pople, GAUSSIAN 94, Gaussian, Inc., Pittsburgh, Pennsylvania, 1995.
[30] A. Banerjee, N. Adams, and J. Simons, J. Phys. Chem. **89**, 52 (1985).
[31] H. B. Schlegel, Theor. Chim. Acta **66**, 333 (1984).
[32] T. H. Fischer and J. Almlöf, J. Phys. Chem. **96**, 9768 (1992).
[33] C. G. Broyden, J. Inst. Math. Appl. **6**, 76 (1970).
[34] R. Fletcher, Comput. J. **13**, 317 (1970).
[35] D. Goldfarb, Math. Comput. **24**, 23 (1970).
[36] D. F. Shanno, Math. Comput. **24**, 647 (1970).
[37] J. M. Bofill, J. Comput. Chem. **15**, 1 (1994).
[38] B. A. Murtagh and R. W. H. Sargent, Comput. J. **13**, 185 (1972).
[39] M. J. D. Powell, Math. Prog. **1**, 26 (1971).
[40] P. Valtazanos and K. Ruedenberg, Theor. Chim. Acta **69**, 281 (1986).
[41] J. Simons, Int. J. Quantum Chem. **48**, 211 (1993).
[42] E. Bosch, M. Moreno, J. Lluch, and J. Bertran, Chem. Phys. Lett. **160**, 543 (1989).
[43] T. Taketsugu, N. Tajima, and K. Hirao, J. Chem. Phys. **105**, 1933 (1996).