CROATICA CHEMICA ACTA
CCACAA, ISSN-0011-1643, ISSN-1334-417X
Croat. Chem. Acta **82** (1) (2009) 283-290.
CCA-3318

Original Scientific Paper

The Importance of Step Control in Optimization Methods*

Jorge M. del Campo and Andreas M. Köster**

Departamento de Química, Cinvestav, Avenida Instituto Politécnico Nacional 2508, A.P. 14-740 México D.F. 07000, México

RECEIVED JUNE 2, 2008; REVISED AUGUST 8, 2008; ACCEPTED AUGUST 9, 2008

Abstract. The restricted step algorithm is described and applied to structure minimization. Robust solutions for the associated root-finding problem are derived. Structure minimizations employing the rational function optimization and the restricted step algorithm in the framework of the quasi-Newton method are compared. With approximated start Hessians both methods show a very similar performance. However, if the start Hessian matrix is calculated the restricted step algorithm is more reliable than the rational function optimization for the studied structure minimizations.

Keywords: restricted step algorithm, structure minimization, ADFT

INTRODUCTION

The improved availability of parallel computer architectures and the development of efficient parallel density functional theory (DFT) methods[1,2] have enabled reliable computational studies of molecules with several hundreds to thousands of atoms.[3–5] The recent development of auxiliary density functional theory (ADFT)[6] in the framework of linear combination of Gaussian-type orbital (LCGTO) DFT methodology provides an ideal theoretical basis for the implementation of highly efficient computer codes.[7–9] One example is the parallel LCGTO-DFT program deMon2k[10] which scales well up to several hundred CPUs. This permits first-principle frequency analysis of systems with more than 500 atoms. On the other hand, it also permits energy calculations and structure optimizations of systems with a few hundreds of atoms on small parallel architectures with less than 10 CPUs.[11] With the appearance of multi-core CPUs such studies have become increasingly popular in different research areas. They require robust and user friendly structure optimization methods. Depending on the application the freezing of atoms or particular degrees of freedom as well as the enforced equivalence of coordinates may be desirable.[12] In modern electronic structure theory programs the automatic generation of appropriate redundant[13] and delocalized[14] internal coordinates now greatly facilitates the optimization of larger molecular systems.

The majority of local optimization procedures are based on the quasi-Newton method. In this method analytic gradients and approximate second derivatives of the energy, that are collected in the Hessian matrix, are required. Both of these quantities can be efficiently calculated in the framework of ADFT. Besides the proper coordinate selection[15,16] the efficiency and reliability of quasi-Newton optimizations depend on various factors including the initial estimate for the Hessian matrix, the Hessian updating method and last but not least the control of the search direction and step size. For an excellent review on quasi-Newton optimization methods the following citation is recommended.[17]

Originally, the rational function optimization (RFO)[18,19] method was used for the determination of the optimization step during the structure minimization in deMon2k. More recently, we have implemented a restricted step algorithm (RSA)[20] based on the Levenberg-Marquardt method.[21,22] Even so, our main goal was the development of a hierarchical transition state finder[23] that combines the double ended saddle interpolation method[24] with the uphill trust region method,[25,26] we also applied the RSA to structure optimization problems. Most surprisingly, we found some remarkable differences between the RFO and RSA minimizations, in particular if a calculated start Hessian is employed. In this work we present a comparison between RFO and RSA minimizations and discuss the observed differences.

* Dedicated to Professor Zvonimir Maksić on the occasion of his 70th birthday.
** Author to whom correspondence should be addressed. (E-mail: akoster@cinvestav.mx)

The article is organized in the following manner. In the next section we describe the implementation of the restricted step algorithm in deMon2k. Detailed attention is paid to the solution of the root-finding procedure arising from the constrained local optimization problem in Levenberg-Marquardt methods. The computational details for the comparison of the RFO and RSA methods are given in the section following after that. The results of the RFO and RSA optimizations are presented and discussed in the next-to-last section. Concluding remarks are drawn in the last section.

RESTRICTED STEP ALGORITHM

In restricted step algorithms the maximum step length is enforced to lie inside a trust region h. Within the quadratic approximation the minimization problem then reads:

$$\min_{\boldsymbol{p}} q^{(k)}(\boldsymbol{p}) \quad \text{subject to} \quad |\boldsymbol{p}| \le h \tag{1}$$

The trust region is a hypersphere defined by $|\boldsymbol{p}| \le h$, where the scalar $h > 0$ is called the trust region radius. The model function $q^{(k)}$ is defined by the second order Taylor series expansion around the local origin, *i.e.* the point $\boldsymbol{x}^{(k)}$:

$$q^{(k)}(\boldsymbol{p}) = E^{(k)} + \tilde{\boldsymbol{g}}^{(k)}\boldsymbol{p} + \frac{1}{2}\tilde{\boldsymbol{p}}\boldsymbol{B}^{(k)}\boldsymbol{p} \tag{2}$$

$E^{(k)}$ is the value of the objective function, here the energy, $\boldsymbol{g}^{(k)}$ is the gradient vector and $\boldsymbol{B}^{(k)}$ the corresponding Hessian matrix or an approximation to it, each of them evaluated at the point $\boldsymbol{x}^{(k)}$. From now on we omit the superscript k for simplicity in the notation. If the solution $\boldsymbol{p}$ of Eq. (1) does not produce a sufficient decrease in the energy, the trust region is too large and needs to be reduced. After the reduction of the trust region radius Eq. (1) is solved again. This procedure is repeated until an acceptable decrease in the energy is achieved. In order to solve (1) the following Lagrange function is introduced:

$$L(\boldsymbol{p},\lambda) = E + \tilde{\boldsymbol{g}}\boldsymbol{p} + \frac{1}{2}\tilde{\boldsymbol{p}}\boldsymbol{B}\boldsymbol{P} + \frac{1}{2}\lambda\left(\tilde{\boldsymbol{p}}\boldsymbol{p} - h^2\right) \tag{3}$$

Here λ represents the undefined Lagrange multiplier. From the stationary condition of this Lagrange function we find the step direction as:

$$\boldsymbol{p} = -\left(\boldsymbol{B} + \lambda\boldsymbol{I}\right)^{-1}\boldsymbol{g} \tag{4}$$

In order to proceed further we now transform this step direction into the diagonal representation of the Hessian matrix. Since $\boldsymbol{B}$ is symmetric, there exists an orthogonal transformation matrix $\boldsymbol{U}$ and a corresponding diagonal matrix $\boldsymbol{B}_D$ such that $\tilde{\boldsymbol{U}}\boldsymbol{B}\boldsymbol{U}=\boldsymbol{B}_D$. In this diagonal representation of the Hessian matrix follows from (4):

$$\tilde{\boldsymbol{U}}\boldsymbol{B}\boldsymbol{U}\tilde{\boldsymbol{U}}\boldsymbol{p} + \lambda\tilde{\boldsymbol{U}}\boldsymbol{p} = -\tilde{\boldsymbol{U}}\boldsymbol{g} \tag{5}$$

We now introduce the projected step $\boldsymbol{s}$ and gradient $\boldsymbol{f}$ as:

$$\boldsymbol{s} = \tilde{\boldsymbol{U}}\boldsymbol{p} \tag{6}$$

$$\boldsymbol{f} = \tilde{\boldsymbol{U}}\boldsymbol{g} \tag{7}$$

With these quantities follows further:

$$\boldsymbol{B}_D\boldsymbol{s} + \lambda\boldsymbol{s} = -\boldsymbol{f} \tag{8}$$

Due to the diagonal nature of $\boldsymbol{B}_D$ the above equation system is decoupled and can be written in the form:

$$b_i s_i + \lambda s_i = -f_i \tag{9}$$

Here s_i and f_i are components of the corresponding vectors defined by equations (6) and (7) and b_i is the corresponding eigenvalue of the Hessian matrix. The projected step components are then given by:

$$s_i = -\frac{f_i}{b_i + \lambda} \tag{10}$$

The original step vector $\boldsymbol{p}$ can be calculated from these components as,

$$\boldsymbol{p} = -\sum_i \left(\frac{f_i}{b_i + \lambda}\right)\boldsymbol{u}_i \tag{11}$$

where $\boldsymbol{u}_i$ represents the eigenvector of the orthogonal transformation matrix $\boldsymbol{U}$ that corresponds to the Hessian eigenvalue b_i. In order to calculate the current step vector $\boldsymbol{p}$ the Lagrange multiplier λ remains to be determined. Because the diagonalization of the Hessian matrix represents an orthogonal transformation the norm of $\boldsymbol{s}$ equals the norm of $\boldsymbol{p}$ and, thus, it follows:

$$\left|\boldsymbol{s}(\lambda)\right|^2 = \left|\boldsymbol{p}(\lambda)\right|^2 = \sum_i \left(\frac{f_i}{b_i + \lambda}\right)^2 \le h^2 \tag{12}$$

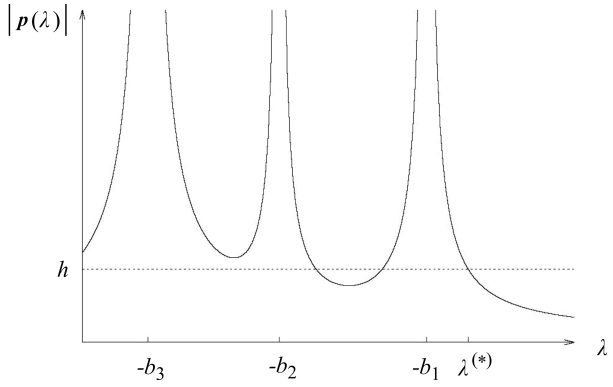

Figure 1. Schematic plot of the $|\boldsymbol{p}(\lambda)|$ function. Here are also shown the singularities at $-b_2$ and $-b_3$ which are usually occurring but are not necessary for the discussion in the text.

Figure 1 depicts the main features of the function $|\boldsymbol{p}(\lambda)|$ defined by Eq. (12). The dashed line represents the trust radius, h, b_i are the eigenvalues of the Hessian matrix and $\lambda^{(*)}$ is the optimal solution for (12). From this figure it can be seen that the $\lambda^{(*)}$ value fulfills the $b_i + \lambda^{(*)} > 0$ criterion. Therefore, the augmented Hessian matrix $\boldsymbol{B} + \lambda\boldsymbol{I}$ is positive definite and the search for a minimum is enforced. The following three cases may appear for the determination of λ:

Case I

The first case is defined by $\boldsymbol{B}$ being positive definite and $|\boldsymbol{B}^{-1}\boldsymbol{g}| \le h$. In this case the Newton step satisfies already (12), usually in the inequality sense, and $\lambda^{(*)} = 0$ is used.

Case II

This is the case depicted in Figure 1, where $f_1 \neq 0$ holds. It then follows:

$$\lim_{\lambda \to -b_1} \left|\boldsymbol{p}(\lambda)\right| = \infty \tag{13}$$

If further $b_1 + \lambda > 0$ holds, with b_1 denoting the lowest, maybe even negative eigenvalue of the Hessian matrix, we obtain:

$$\lim_{\lambda \to \infty} \left|\boldsymbol{p}(\lambda)\right| = 0 \tag{14}$$

Therefore, Eq. (12) is a monotonically decreasing function of λ on the interval $(-b_1,\infty)$. As Figure 1 shows, a solution to the constrain $|\boldsymbol{p}(\lambda)| = h$ exists always independent of the structure of the $\boldsymbol{B}$ matrix. Thus, a step $\boldsymbol{p}$ suitable for the minimization can be found under all circumstances. In order to obtain $\lambda^{(*)}$ a root-finding procedure must be applied. Due to the possible occurrence of several intersections of the $|\boldsymbol{p}(\lambda)|$ function with the h line (see Figure 1) the interval for the root-finding procedure must be adapted to the Hessian matrix structure. If $\boldsymbol{B}$ is positive definite and $|\boldsymbol{B}^{-1}\boldsymbol{g}| > h$, λ must be positive and, therefore, we search in the interval $(0, \infty)$. On the other hand, if $\boldsymbol{B}$ is not positive definite the search must be restricted to the interval $(-b_1,\infty)$.

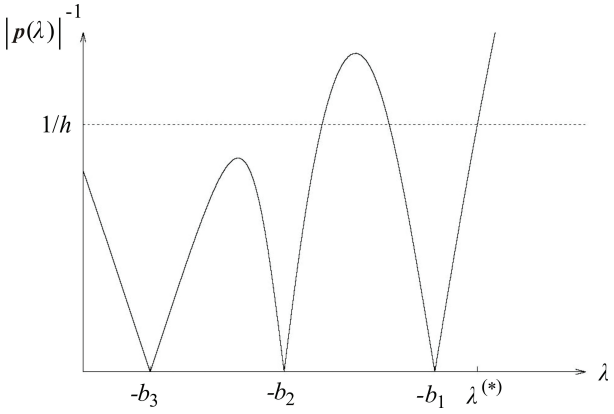

Figure 2. Schematic plot of the $|\boldsymbol{p}(\lambda)|^{-1}$ function. The linearization right from $-b_1$ is essential for the efficient root-finding, *i.e.* the determination of $\lambda^{(*)}$.

In principle, Newton's root-finding method could be used to find $\lambda^{(*)}$ that solves the constrain

$$\left|\boldsymbol{p}\left(\lambda^{(*)}\right)\right| - h = 0 \tag{15}$$

However, such an approach is numerically instable if $\lambda^{(*)}$ is close from the right hand side to b_1. In this situation $|\boldsymbol{p}(\lambda)|$ is highly nonlinear, and Newton's root-finding method may fail. Excellent results are obtained by reformulating the problem[27] in such a way that it becomes nearly linear close to the solution $\lambda^{(*)}$. To do so the following function is defined:

$$\varphi(\lambda) = \frac{1}{h} - \frac{1}{\left|\boldsymbol{p}(\lambda)\right|} = 0 \tag{16}$$

As Figure 2 shows this function is nearly linear in the solution range and, therefore, the root-finding with Newton's method works well. Applying this method to the φ function generates a sequence of iterates $\lambda^{(l)}$ of the form:

$$\lambda^{(l+1)} = \lambda^{(l)} - \frac{\varphi\left(\lambda^{(l)}\right)}{\varphi'\left(\lambda^{(l)}\right)} \tag{17}$$

here φ' denotes the derivative of φ.

Case III

In this case, also called the "hard case", $f_1 \neq 0$ does not hold. As a consequence the singularity at $-b_1$ disappears

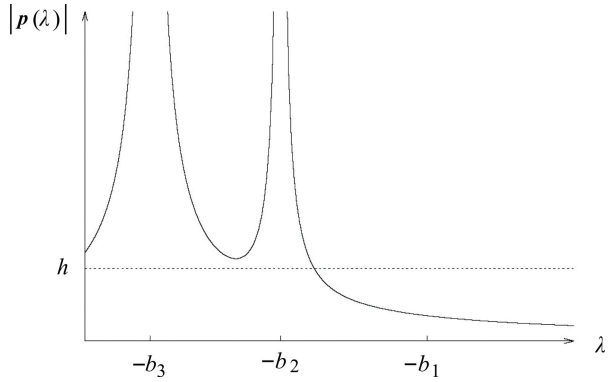

Figure 3. Schematic plot of the $|\boldsymbol{p}(\lambda)|$ function in the "hard case", *i.e.* without the singularity at $-b_1$.

(Figure 3). In the case a valid solution $\lambda^{(*)}$ exists in the open interval $(-b_1,\infty)$ Eq. (12) is applied. Otherwise, as depicted in Figure 3, the solution $\lambda^{(*)}$ is located in the domain $-b_2 < \lambda^{(*)} \leq -b_1$. If we now take the solution where $|\boldsymbol{p}(\lambda^{(*)})| = h$ holds, then the distance constrain is satisfied but the corresponding augmented Hessian matrix $\boldsymbol{B} + \lambda^{(*)}\boldsymbol{I}$ is not positive definite. To avoid this problem we follow the strategy proposed by Nocedal,[27] namely to set $\lambda^{(*)} = -b_1$ and redefine $\boldsymbol{p}(\lambda)$ as:

$$\boldsymbol{p}(\lambda) = -\sum_{i\neq 1}\left(\frac{f_i}{b_i - b_1}\right)\boldsymbol{u}_i + \tau\boldsymbol{u}_1 \tag{18}$$

Here $\boldsymbol{u}_1$ is the Hessian eigenvector that corresponds to the eigenvalue b_1 and τ is a scalar variable that remains to be fixed. We use the distance constrain to determine τ:

$$|\boldsymbol{p}(\lambda)|^2 = -\sum_{i\neq 1}\left(\frac{f_i}{b_i - b_1}\right)^2 + \tau^2 = h^2 \tag{19}$$

$$\tau = \sqrt{h^2 - \sum_{i\neq 1}\left(\frac{f_i}{b_i - b_1}\right)^2} \tag{20}$$

Therefore, it is always possible to choose τ such that the distance constrain $|\boldsymbol{p}(\lambda)| = h$ is fulfilled and that the corresponding augmented Hessian matrix is positive definite in order to ensure minimization.

COMPUTATIONAL METHOD

For the comparison of the RFO and RSA structure optimizations the test suite of Baker,[16] augmented with HCN and H_2O_2, is used. All calculations were performed with deMon2k[10] using the DZVP basis set[28] and the A2 auxiliary function set, which contains *s*, *p*, and *d* auxiliary functions and has been optimized for the DZVP basis set. The auxiliary density was used for the variational fitting of the Coulomb[29] and exchange-correlation[6] potential, *i.e.* all calculations were performed in the framework of auxiliary density functional theory. The generalized gradient approximation of Perdew, Burke and Ernzerhof (PBE)[30] was used for the exchange and correlation functional. For the numerical integration the standard adaptive grid (MEDIUM) from deMon2k was employed.[31] All molecules were fully optimized, using the here described RSA and the previous implemented[32] RFO method. The optimizations were performed in delocalized internal coordinates[14] based on an automatized selection of relevant primitive coordinates.[32] No symmetry constrains were employed. For the self-consistent field and structure optimization the default convergence criteria of deMon2k were used. The start Hessian matrix was either set to the unit matrix (UNIT) or approximated according to Fischer and Almlöf,[33] which is the default in deMon2k, Lindh *et al.*[34] or Baker.[14] The different approaches are labeled FISCHER, LINDH and BAKER, respectively. If the start Hessian matrix was calculated we label it EXACT. The BFGS update[35–38] was used during the structure optimization to actualize the Hessian matrix. All optimized structures were characterized by a frequency analysis.

RESULTS AND DISCUSSION

In Table 1 the RFO and RSA structure optimizations are compared using the molecules from the augmented Baker test suite. The comparison of the first two data columns of this table (FISCHER) shows that the performance of the RFO and RSA methods is nearly identical if an approximated start Hessian is used. In total the RFO and RSA methods needed 214 and 215 steps, respectively, for all structure optimizations. Both methods converged to the same equilibrium structures. However, the characterization of the optimized structures by frequency analysis revealed that both methods failed in locating minima for the highlighted entries in Table 1. The reason for these failures lies in the approximated start Hessian matrix, which does not reflect correctly the curvature of the potential energy surface at the starting point of the optimization. In simple words, the eigenvalue spectrum of the approximated start Hessian is wrong. This problem is not particular to the start Hessian approximation from Fischer and Almlöf as can be seen from Table 2. Here the number of optimization cycles for the RFO and RSA minimization of the highlighted molecules from Table 1 are presented employing different start Hessians. Again both methods show very similar performance. In particular, the number of

Table 1. Comparison of the RFO and RSA structure optimization cycles for the augmented Baker test suite. The energy differences, ΔE, between optimized structures obtained from an approximated (FISCHER) or calculated (EXACT) start Hessian matrix are given in kcal/mol. Significant differences are highlighted

	Start Hessian					
	FISCHER		EXACT		ΔE	
Molecule	RFO	RSA	RFO	RSA	RFO	RSA
Water	4	4	2	2	0.00	0.00
Hydrogen cyanide	9	8	9	9	0.00	0.00
Hydrogen peroxide	10	9	9	9	0.00	0.00
Ammonia	5	5	3	3	0.00	0.00
Ethane	4	4	3	3	0.00	0.00
Acetylene	5	5	4	4	0.00	0.00
Allene	4	4	2	2	0.00	0.00
Hydroxysulphane	7	7	5	6	0.00	0.00
Benzene	3	3	2	2	0.00	0.00
Methylamine	4	4	2	11	0.00	−5.22
Ethanol	5	5	3	3	0.00	0.00
Acetone	5	5	3	3	0.00	0.00
Disilyl ether	11	11	5	6	0.00	0.00
1,3,5-trisilacyclohexane	6	6	4	6	0.00	0.01
Benzaldehyde	6	6	10	10	0.00	0.00
1,3-difluorobenzene	4	7	5	6	0.00	0.00
1,3,5-trifluorobenzene	6	7	12	7	0.00	0.00
Neopentane	4	4	2	3	0.00	0.00
Furan	6	6	3	4	0.00	0.00
Naphtalene	5	6	3	4	0.00	0.00
1,5-difluoronaphtalene	5	5	4	4	0.00	0.00
2-Hydroxybicyclopentane	13	15	8	11	0.00	0.00
ACHTAR10	7	6	5	15	0.00	−0.35
ACANIL01	6	7	8	8	0.00	0.00
Benzidine	6	7	35	20	−2.13	−2.14
Pterin	7	8	6	15	0.12	−0.32
Difuropyrazine	8	9	11	12	0.00	0.00
Mesityl oxide	16	14	23	22	0.00	−1.86
Histidine	16	14	23	22	0.02	0.04
Dimethylpentane	5	5	5	5	0.00	0.00
Caffeine	7	8	8	12	0.00	0.00
Menthone	15	9	15	7	0.00	0.00
Sum	214	215	225	259		

optimization cycles for all three approximated start Hessians (FISCHER in Table 1, BAKER in Table 2 and LINDH in Table 2) are very similar. As expected the performance of the unit start Hessian matrix is much poorer. Nevertheless, all optimizations converged to the same structures which are for these molecules no minima.

If the initial Hessian matrix is calculated (EXACT; data column 3 and 4 in Table 1) the situation changes. At first glance it seems that the RFO method (in total 225 optimization steps) now out-performs the RSA method (in total 259 optimization steps). In fact, this is consistent with previous reports in the literature. However, this conclusion is wrong, at least for the comparison presented here. The RFO energy differences ΔE in Table 1 (data column 5) show that the RFO optimizations with calculated start Hessians, by and large, yield the same optimized structures as with the approximated start Hessians. The only exception is benzidine. The corresponding frequency analysis shows that the benzidine optimization with a calculated start Hessian converges to a minimum. In Figure 4 the results of the two optimizations with an approximated and calculated start Hessian are depicted. The relevant structural changes

Table 2. Comparison of the RFO and RSA structure optimization cycles employing different start Hessians. The entries BAKER and LINDH refer to start Hessian approximations from Baker[14] and Lindh *et al*.,[34] respectively. The entry UNIT refers to a unit matrix start Hessian

		RFO			RSA		
		START HESSIAN					
Molecule	Formula	BAKER	LINDH	UNIT	BAKER	LINDH	UNIT
Methylamine	CH_3NH_2	3	3	6	3	3	6
ACHTAR10	$CH_3COOCH_2CH_2NH_2$	6	7	13	6	7	13
Benzidine	$NH_2C_{12}H_8NH_2$	6	6	10	6	7	10
Pterin	$NH_2C_6N_4OH_3$	6	8	11	6	8	11
Mesityl oxide	$CH_3COCHC(CH_3)_2$	6	5	12	6	5	12
Sum		27	29	52	27	30	52

(see arrows in Figure 4) are the pyramidalization of the two amine substituents. Only if the start Hessian matrix is calculated, the RFO optimization of benzidine yields a minimum. However, the RFO optimizations of the other four highlighted molecules (methylamine, ACHTAR10, pterin and mesityl oxide) do not converge to minima, independently if the start Hessian matrix was calculated or not. For the RSA optimization the situation is different. In this case all optimizations in Table 1, except for mesityl oxide, yield minima if the start Hessian is calculated. This can be seen from data column 6 of Table 1, too. For all highlighted systems a significant lowering in the energy of the RSA optimized structures was found if the start Hessian was calculated. The corresponding frequency analysis confirmed that the lower energy structures, except for mesityl oxide, are indeed minima.

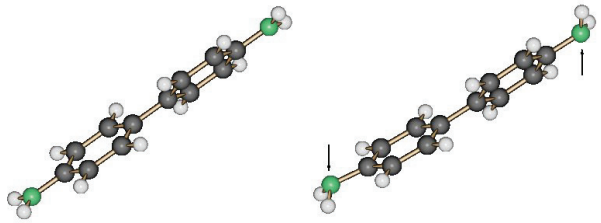

Figure 4. Optimized (RFO and RSA) benzidine structures. The left structure (2nd order saddle) is obtained from the optimization with an approximated start Hessian, whereas the right structure (minimum) is obtained from the optimization with a calculated start Hessian. The arrows indicate the relevant structure changes. Color code: C black, N green, H white.

Besides of the pyramidalization of amine groups, as discussed in the benzidine optimization and also found in the RSA optimization with calculated start Hessian matrix for methylamine and pterin, inner rotations are critical for some of the structure minimizations presented here. As an example the RSA optimized structures of ACHTAR10 are depicted in Figure 5. The indicated methyl group (arrow in Figure 5) represents a typical inner rotor. If the start Hessian is calculated the RSA quasi-Newton optimization yields a minimum for this structure (Figure 5, right), *i.e.* the correct orientation of the inner rotor is found. A similar situation applies for mesityl oxide. However, as depicted in Figure 6, this system possesses two inner rotors (arrow 1 and 2) that

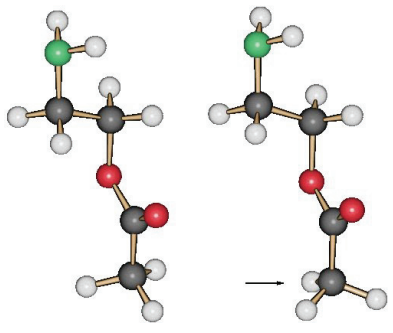

Figure 5. Optimized (RSA) ACHTAR10 structures. The left structure (transition state) is obtained from the optimization with an approximated start Hessian, whereas the right structure (minimum) is obtained from the optimization with a calculated start. The arrow indicates the relevant methyl group. Color code: C black, N green, O red, H white.

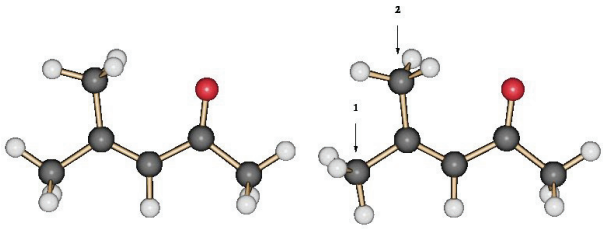

Figure 6. Optimized (RSA) mesityl oxide structures. The left structure is obtained from the optimization with an approximated start Hessian, whereas the right structure is obtained from the optimization with a calculated start Hessian. Both structures are transition states. The two coupled methyl groups 1 and 2 are indicated by arrows. Color code: C black, O red, H white.

are coupled. The RSA optimization with the calculated start Hessian correctly minimizes the configurational energy of rotor 1 in Figure 6. However, the BFGS update scheme employed in the quasi-Newton steps is unable to describe the coupling to rotor 2 correctly. As a consequence the methyl group of rotor 2 remains in its original configuration, which is not its minimum configuration after methyl group 1 has turned. This situation is confirmed by the frequency analysis, which yields one imaginary frequency associated to the rotation of methyl group 2. Of course, this situation can be resolved by substituting the update with the calculation of the Hessian matrix. In such Newton optimizations the RSA will always converge to a local minimum.[20] However, the corresponding computational effort increases considerably because the Hessian matrix must be calculated in each optimization step.

CONCLUSION

From our comparison of the RFO and RSA optimization follows that both methods perform almost identical if approximated start Hessians are used. If the start Hessian is calculated we find significant differences for molecules that possess negative curvatures in the Hessian at the starting structure. For these cases the RSA minimization has proven more reliable than the RFO minimization. Our analysis shows that the success of the RSA quasi-Newton optimization to locate minima, even in difficult situations, is mainly due to the simultaneous step size and step direction control. Therefore, a robust solver for the associated root-finding problem, as presented in this article, is mandatory. Despite the success of the RSA quasi-Newton minimization we found one case, mesityl oxide, where the location of a minimum failed. The analysis of this case shows that the common update schemes are usually not capable to describe correctly the coupling of inner rotors. This problem can be overcome with a RSA Newton minimization. However, the corresponding computational demand does not justify such an automatization at this time.

Thus, for the time being, care should be taken in the selection of starting geometries in order to avoid obvious transition state configurations that may trap the quasi-Newton optimization. On the other hand, if a frequency analysis has been performed and a higher order saddle point was identified we recommend to use the corresponding Hessian as a restart Hessian for the following structure optimization and to modify the critical point along the corresponding deformation frequency. This can be done in the same spirit as in an intrinsic reaction coordinate calculation (IRC).[39,40] If one is not really interested in the IRC path but only in a structure optimization then the use of a quasi-Newton method with a perturbed starting structure usually yields the desired result, similar to the examples shown here.

Acknowledgements. This work was financially supported by the CONACyT Project No. 60117-U. J.M.C gratefully acknowledges a CONACyT Ph.D. fellowship (180545).

REFERENCES

1. P. Hohenberg and W. Kohn, *Phys. Rev.* **136** (1964) B864
2. W. Kohn and L. J. Sham, *Phys. Rev.* **140** (1965) A1133.
3. R. R. Zope and B. I. Dunlap, *Phys. Rev. B* **71** (2005) 193104.
4. T. Inaba and F. Sato, *J. Comp. Chem.* **28** (2007) 984.
5. R. R. Zope, T. Baruah, M. R. Pederson, and B. I. Dunlap, *Phys. Rev. B* **77** (2008) 115452.
6. A. M. Köster, J. U. Reveles, and J. M. del Campo, *J. Chem. Phys.* **121** (2004) 3417.
7. U. Birkenheuer, A. B. Gordienko, V. A. Nasluzov, M. K. Fuchs-Rohr, and N. Rösch, *Int. J. Quantum Chem.* **102** (2005) 743.
8. L. Belpassi, F. Tarantelli, A. Sgamellotti, and H. M. Quiney, *J. Chem. Phys.* **124** (2006) 124104.
9. L. Belpassi, F. Tarantelli, A. Sgamellotti, and H. M. Quiney, *J. Chem. Phys.* **128** (2008) 124108.
10. A. M. Köster, P. Calaminici, M. E. Casida, R. Flores-Moreno, G. Geudtner, A. Goursot, T. Heine, A. Ipatov, F. Janetzko, J. M. del Campo, S. Patchkovskii, J. U. Reveles, D. R. Salahub, and A. Vela, deMon developers, 2006. http://www.demon-software.com
11. V. Dominguez-Soria, P. Calaminici, and A. Goursot, *J. Chem. Phys.* **127** (2007) 154710.
12. J. U. Reveles, S. N. Khanna, and A. M. Köster, *J. Mol. Struct. THEOCHEM* **762** (2006) 171.
13. C. Peng. P. Y. Ayala, and H. B. Schlegel, *J. Comp. Chem.* **17** (1996) 49.
14. J. Baker, A. Kessi, and B. Delley, *J. Chem. Phys.* **105** (1996) 192.
15. P. Pulay, G. Fogarasi, F. Pang, and J. E. Boggs, *J. Am. Chem. Soc.* **101** (1979) 2550.
16. J. Baker, *J. Comput. Chem.* **14** (1993) 1085.
17. V. Bakken and T. Helgaker, *J. Chem. Phys.* **117** (2002) 9160.
18. A. Banerjee, N. Adams, J. Simons, and R. Shepard, *J. Chem. Phys.* **89** (1985) 52.
19. J. Baker, *J. Comp. Chem.* **7** (1996) 385.
20. R. Fletcher, Practical Methods of Optimization, Vol.1, Wiley, New York,1981, pp. 77.
21. K. Levenberg, *Quart. Appl. Math.* **2** (1944) 164.
22. D. W. Marquardt, *SIAM J.* **11** (1963) 431.
23. J. M. del Campo and A. M. Köster, *J. Chem. Phys.* **129** (2008) 024108.
24. M. J. S. Dewar, E. F. Healy, and J. J. P. Stewart, *J. Chem. Soc., Faraday Trans. 2* **80** (1984) 227.
25. V. H. Nguyen, J. M. Ghuysen P. Culot, and G. Dive, *Theor. Chim. Acta* **82** (1992) 189.
26. J. M. Bofill, *J. Comp. Chem.* **15** (1994) 1.
27. J. Nocedal and S. J. Wright, *Numerical Optimization*, Springer-Verlag, New York 1999, pp. 78.
28. N. Godbout, D. R. Salahub, J. Andzelm, and E. Wimmer, *Can. J. Chem.* **70** (1992) 560.
29. B. I. Dunlap, J. W. D. Connolly, and J. R. Sabin, *J. Chem. Phys.* **71** (1979) 4993.
30. J. P. Perdew, K. Burke, and M. Ernzerhof, *Phys. Rev. Lett.* **77** (1996) 3865.
31. A. M. Köster, R. Flores-Moreno, and J. U. Reveles, *J. Chem. Phys.* **121** (2004) 681.
32. J. U. Reveles and A. M. Köster, *J. Comp. Chem.* **25** (2004) 1109.
33. T. H. Fischer and J. Almlöf, *J. Phys. Chem.* **96** (1992) 9768.

34. R, Lindh, A. Bernhardsson, G. Karlström, and P.-A. Malmqvist, *Chem. Phys. Lett.* **241** (1995) 423.
35. C. G. Broyden, *J. Inst. Math. Appl*. **6** (1970) 76.
36. R. Fletcher, *Comp. J.* **13** (1970) 317.
37. D. Goldfarb, *Math. Comput.* **24** (1970) 23.
38. D. F. Shanno, *Math. Comput.* **24** (1970) 647.
39. K. Fukui, *Acc. Chem. Res.* **14** (1981) 363.
40. C. Gonzalez and H. B. Schlegel, *J. Chem. Phys.* **90** (1989) 2154.

SAŽETAK

Važnost kontrole koraka u optimizacijskim metodama

Jorge M. del Campo i Andreas M. Köster

Departamento de Química, Cinvestav, Avenida Instituto Politécnico Nacional 2508, A.P. 14-740 México D.F. 07000, México

Opisan je algoritam s ograničenim korakom i primijenjen na optimizaciju strukturnih parametara. Razvijena su robusna rješenja za probleme nalaženja korijena. Uspoređene su minimizacije struktura primjenom racionalne optimizacije funkcije i algoritam s ograničenim korakom u sklopu kvazi-Newtonove metode. S aproksimativnim početnim Hessianima obje metode pokazuju slične performanse. No, ukoliko je početna Hessova matrica računana, algoritam s ograničenim korakom je pouzdaniji nego metoda racionalne optimizacije funkcije za proučavane minimizacije strukturnih parametara.