

Estimating the Hessian for gradient-type geometry optimizations

H. Bernhard Schlegel*

Department of Chemistry, Wayne State University, Detroit, MI, 48202, USA

Optimization methods that use gradients require initial estimates of the Hessian or second derivative matrix; the more accurate the estimate, the more rapid the convergence. For geometry optimization, an approximate Hessian or force constant matrix is constructed from a simple valence force field that takes into account the inherent connectivity and flexibility of the molecule. Empirical rules are used to estimate the diagonal force constants for a set of redundant internal coordinates consisting of all stretches, bends, torsions and out-of-plane deformations involving bonded atoms. The force constants are transformed from the redundant internal coordinates to Cartesian coordinates, and then from Cartesian coordinates to the non-redundant internal coordinates used in the specification of the geometry and optimization. This method is especially suitable for cyclic molecules. Problems associated with the choice of internal coordinates for geometry optimization are also discussed.

Key words: Optimization methods—geometry optimization—gradients—Hessian—force constants

1. Introduction

The analytical calculation of the energy gradient has greatly improved the efficiency and reliability of geometry optimizations with *ab initio* and semiempirical molecular orbital methods [1-3]. The equilibrium geometry of a molecule with N degrees of freedom can be found in ca. N steps using any of a family of gradient-type optimization methods [4, 5] such as conjugate gradient, quasi-Newton, variable metric, Murtagh-Sargent, etc. At the beginning of the optimization each of these methods requires an estimate of the Hessian (also called the

* Fellow of the Alfred P. Sloan Foundation, 1981-83.

second derivative matrix or the force constant matrix). The estimate need not be very accurate since the Hessian matrix is updated during the search for the minimum, and gradually approaches the correct second derivative matrix. The overall efficiency of the optimization, or the rate of convergence to the equilibrium geometry depends on the initial estimates of the Hessian as well as the starting geometry. Although the final optimized geometry is independent of the start geometry or Hessian, the closer these estimates are to the true second derivative matrix and the equilibrium geometry, the fewer steps will be required to complete the optimization.

Two limiting cases can be considered for the initial estimate of the Hessian. The simplest and most frequently used estimate is the unit matrix. Although this is an unbiased choice, all the useful structural information about the molecule is discarded, i.e. the nature of the atoms, the bonds between them, etc. Flexible coordinates (i.e. torsion and ring deformation) are not distinguished from stiff modes (i.e. bond stretching) and all coupling between coordinates is ignored. Such information must be accumulated during the course of the optimization at the expense of additional optimization steps. This is particularly detrimental for cyclic molecules whose coordinates are inherently strongly coupled.

In the other extreme, the full second derivative matrix can be calculated analytically [6]. If the energy surface for the molecule is approximately quadratic only one step is required to reach the optimum geometry. However, in many cases it is not feasible or not desirable to calculate the full Hessian directly, since this requires ca. N times the effort of a gradient calculation. Clearly, an intermediate method of estimating the Hessian is needed that incorporates the chemically important features of the molecule yet does not require a substantial computational effort.

The purpose of this paper is to outline a simple empirical procedure for estimating the Hessian that is especially suitable for cyclic molecules. In addition, some comments are made concerning the choice of internal coordinates to improve the rate of convergence of geometry optimization.

2. Procedure

Since the Hessian is the matrix of second derivatives of the energy surface with respect to the coordinates, its form will depend partly on the choice of coordinates being optimized. In principle, any non-redundant coordinate system is acceptable. However conventional optimization techniques as well as gradient methods function best if the coordinates are not strongly coupled (i.e. no large interaction force constants or off-diagonal Hessian matrix elements). Redundant coordinates must be avoided since the gradient cannot be determined uniquely for such coordinate systems. Independent of the coordinate system used for the optimization, the Hessian also contains information about the connectivity of the molecule, the bond strengths, conformational flexibility, etc. The separation of the two contributions can be exploited in the estimation of the Hessian.

An estimate of the Hessian for an arbitrary set of non-redundant internal coordinate system can be obtained by: (a) determining a complete set of redundant valence coordinates from the Cartesian coordinates of the atoms, (b) estimating the force constants for these coordinates, (c) transforming the force constants to Cartesian coordinates, (d) transforming the force constants into the final set of non-redundant internal coordinates used in the optimization.

(a) Redundant valence coordinates are commonly used in spectroscopy to describe molecular force fields [6]. These include bond stretch, valence angle bend, torsion about a bond, out-of-plane bend, and linear angle bend. Normally these are defined only for bonded atoms. Whether two atoms are bonded can be determined easily from the interatomic distances and a table of covalent radii. Thus a unique set of valence coordinates can be constructed readily from the Cartesian coordinates of the atoms. For clusters and loose complexes, some care must be taken to ensure that the coordinate system spans all degrees of freedom, and does not represent the complex as two disconnected fragments. The infinitesimal displacements of the valence coordinates, q' , are related to the Cartesian displacements, x , by the appropriate B matrix of Wilson [7]:

$$q' = B'x. \quad (1)$$

(b) Force constants in valence coordinates can be estimated from a variety of empirical rules developed by spectroscopists. However, the methods chosen for the present application must be widely applicable and depend only on the atom types and the molecular geometry. For stretching force constants, Badger's rule [8] falls in this category:

$$F_{\text{str}} = A/(r - B)^3 \quad (2)$$

where A and B are constants and r is the bond length. B depends only on the rows of the periodic table that the bonded atoms are in. Table 1 lists values for A and B calibrated to reproduce force constants for minimal and extended basis set ab initio calculations. An accuracy of ca. 10% is typical for H, C, N, O, F, Si, P, S and Cl compounds.

For angle bends a similarly general rule is lacking. Since the bending force constants do not vary greatly, a single constant is used for all three heavy atom bond angles (see Table 1). A smaller value is used if either terminal atom is hydrogen. If more data becomes available through analytical force constant calculations, a larger selection of values may be warranted. Dependence on the lengths of the bonds forming the angles could also be incorporated if necessary, via a bond-order/bond-energy expression [9].

The force constant for torsion about a double bond can be an order of magnitude larger than for torsion about a single bond. The following formula is used to take this into account crudely:

$$F_{\text{tors}} = A - B(r - r_{\text{cov}}); \quad r < r_{\text{cov}} + A/B \quad (3)$$

Table 1. Empirical force constants used to estimate the Hessian for *ab initio* geometry optimizations

| Parameters for force constants ^a | |
|--|--|
| <i>Bond Stretch</i> ($F_{\text{str}} = A/(r - B)^3$) | |
| $A = 1.734^b$ | |
| $B = -0.244$ | (1 st period—1 st period) ^c |
| 0.352 | (1 st period—2 nd period) |
| 1.085 | (2 nd period—2 nd period) |
| 0.660 | (1 st period—3 rd period) |
| 1.522 | (2 nd period—3 rd period) |
| 2.068 | (3 rd period—3 rd period) |
| <i>Angle Bend</i> ($F_{\text{bend}} = A$) | |
| $A = 0.160^b$ | (either or both terminal atoms hydrogen) |
| $A = 0.250^b$ | (all three heavy atom bends) |
| <i>Torsion</i> ($F_{\text{tors}} = A - B(r - r_{\text{cov}})$) | |
| $A = 0.0023^b$ | $B = 0.07^b$ |
| <i>Out-of-Plane</i> ($F_{\text{oop}} = Ad^4$) | |
| $A = 0.045^b$ | |

^a Force constants in hartree/bohr² or hartree/rad², bond lengths in bohr. Only force constants between bonded atoms are defined. Atoms are considered bonded if their internuclear distance is less than 1.35 times the sum of the covalent radii. The covalent radii use for the first 36 atoms are: 0.32, 0.60, 1.2, 1.05, 0.81, 0.77, 2*.74, 2*.72, 1.5, 1.4, 1.3, 1.17, 1.10, 1.04, 2*.99, 1.8, 1.6, 11*1.4, 1.3, 2*1.2, 2*1.1A.

^b Multiply by 1.3 for minimal basis set calculations.

^c I.e. H-H

where r is the length of the central bond and r_{cov} is the length of the corresponding single bond (i.e. sum of the covalent radii). Since a dihedral angle or torsional coordinate is defined for each distinct pair of atoms bonded to opposite ends of a central bond, C₂H₄ and C₂H₆ have 4 and 9 torsional coordinates, respectively.

For planar molecules, out-of-plane bending coordinates must be considered in addition to valence angle bends. However, as the molecule is made pyramidal, the force field is best described by valence angles alone. Thus the out-of-plane force constant is strongly attenuated by a function of the non-planarity

$$F_{\text{oop}} = Ad^4; \quad d = 1 - r_1 \cdot r_2 \times r_3 / |r_1| |r_2| |r_3| \quad (4)$$

where r_1 , r_2 , and r_3 are vectors pointing along the three bonds associated with the out-of-plane angle. Note that $0 \leq d \leq 1$ and that $d = 1$ for a planar center. An out-of-plane coordinate is defined for each distinct set of three atoms bonded to a central atom.

The expressions for the diagonal valence force constants are sufficient to account for the connectivity and the flexibility of the molecule. Because of the redundancy in the valence coordinates (i.e. 6 bonds and 6 angles in a planar ring; 6 bends

for a tetrahedral center; etc.), they also describe the most important coupling between the nonredundant coordinates used in the optimization (i.e. 5 bonds and 4 angles for the ring; 5 bends for tetrahedral coordination). For a simple valence force field, only the diagonal elements are non-zero. Additional coupling can be introduced by estimating some of the off-diagonal terms. The most important couplings are stretch-stretch and stretch-bend interactions. For two stretches sharing a common atom, -0.1 mdyn/Å is typical. For a stretch and a bend sharing a common bond, -0.5 mdyn/Å is an average value. Interaction force constants have not been included in the current implementation.

(c) Transformation of the force constants from valence coordinates to Cartesian space can be accomplished with the \mathbf{B}' matrix appropriate for the valence coordinates,

$$\mathbf{F}^{\text{cart}} = \mathbf{B}' \mathbf{F}^{\text{val}} \mathbf{B}'^T. \quad (5)$$

This transformation is especially simple if \mathbf{F}^{val} is diagonal, since \mathbf{B}' does not have to be stored.

(d) Once the force constant matrix is in Cartesian coordinates, it can be transformed to any non-redundant coordinate system. Typically the coordinates used to construct the molecule and to perform the optimization contain bond lengths, bond angles and dihedral angles. In addition to the actual atoms in the molecule, dummy atoms may be used to assist in describing the molecule or optimizing the geometry. The infinitesimal displacements of these non-redundant internal coordinates, q , can be written in terms of a Wilson \mathbf{B} matrix and the Cartesian displacements. Note that this \mathbf{B} matrix is different than \mathbf{B}' used to convert from the redundant valence coordinates.

$$q = \mathbf{B}x. \quad (6)$$

Transformation of the forces and the force constants requires the inverse of \mathbf{B} :

$$\begin{aligned} \mathbf{B}^{-1} &= (\mathbf{B}^T \mathbf{M} \mathbf{B})^{-1} \mathbf{B}^T \mathbf{M} \\ \mathbf{B}^{-1} q &= x. \end{aligned} \quad (7)$$

The auxiliary matrix \mathbf{M} depends on the conventions used for constructing the Cartesian coordinates from the bond lengths and angles. For the algorithm employed in the GAUSSIAN series of programs [10], \mathbf{M} is a unit matrix with additional zeros on the diagonal for M_{11} , M_{22} , M_{33} , M_{44} , M_{55} and M_{88} . These correspond to the Cartesian coordinates that are arbitrarily fixed (i.e. atom 1 is placed at the origin, atom 2 on the z axis and atom 3 in the xz plane).

The forces, f , and the force constants, F , are transformed according to:

$$\begin{aligned} f^{\text{int}} &= \mathbf{B}^{-1} f^{\text{cart}} \\ \mathbf{F}^{\text{int}} &= \mathbf{B}^{-1} \mathbf{F}^{\text{cart}} (\mathbf{B}^{-1})^T + (d\mathbf{B}^{-1}/dq) f^{\text{cart}}. \end{aligned} \quad (8)$$

The Cartesian forces and force constants that correspond to the derivatives with respect to the Cartesian coordinates of the dummy atoms are zero, since the

energy depends on the positions of only the real atoms. However, the internal forces and force constants for parameters involving dummy atoms are in general non-zero, because changes in these parameters can change the positions of the real atoms.

The second term in Eq. (9) arises from the curvilinear nature of the internal coordinate system. As in Eq. (2), this term is ignored because the force constant matrix is only approximate. However, if F^{cart} is calculated by analytical differentiation of the energy [6] (e.g. frequency calculations), this term must be taken into account (in the GAUSSIAN system $d\mathbf{B}^{-1}/dq$ is determined by numerical differentiation).

3. Discussion

In the algorithm outlined above, a simple valence force field is used to estimate the Hessian. There are a variety of other, more sophisticated methods used to represent force fields in spectroscopy [7], e.g. general valence force field, Urey-Bradley, etc. Usually these require a larger number of parameters and often must treat each new molecule as a special case. An alternate approach is to obtain a potential energy surface from a molecular mechanics program, such as MM2 [11] or BIGSTRAIN [12]. However, these methods are parameterized only for a limited set of interactions, and deal with the inherently more difficult problem of reproducing energies and geometries as well as force constants. If more accuracy is required than a simple force field can provide, an estimated Hessian for an *ab initio* geometry optimization can be obtained from second derivatives computed using semiempirical methods or smaller basis set *ab initio* calculations.

Even though the interactions between coordinates can be estimated, optimization methods work best if coordinate systems are chosen to reduce the coupling as much as possible. Several examples of different coordinate systems are shown in Fig. 1. The atoms in a planar 6 membered ring can be specified with 5 bond lengths and 4 angles. A change in any one of the angles will have a large effect on the ring closing bond. An alternate representation, II, reduces the strong coupling by letting each angle move only one atom rather than a chain of atoms. Another situation that mixes bond stretch with angle bend is indicated in III. If the transition state is specified by IV (distance to the H₂ midpoint, H₂ bond length and 2 angles specifying the orientation), the relation between the coordinates is much simpler. In both of these examples, a flexible coordinate, e.g. angle bend, was coupled to a stiff mode, bond stretch. A similar situation can occur with torsional modes coupled to angle bends. In *N*-chloromethylamine, the amino group can be specified by two dihedral angles, V. If the amino group rotates during an optimization, both angles must be changed synchronously. A better representation would separate the rotation and inversion coordinates, as in VI. The reader can no doubt think of many other examples.

A flexible mode coupled to a stiff mode corresponds to a long narrow valley running diagonally across the energy surface. In pathological cases, the valley is

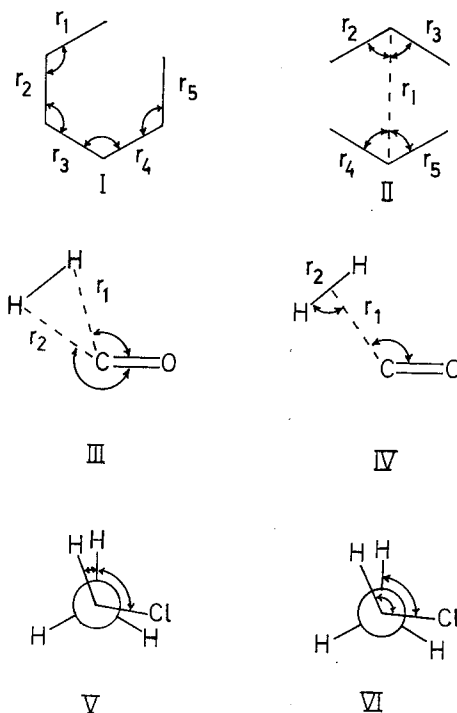


Fig. 1. Internal coordinates for geometry optimization of a planar 6 membered ring (I and II), $H_2 + CO$ transition structure (III and IV), and internal rotation-inversion in CH_3NCHl (V and VI). In each case, the second choice represents less coupling among the internal coordinates and should be less difficult to optimize

also curved. All gradient optimization methods will have a difficult time finding the minimum in this situation, since they will oscillate between the steep walls and make little progress along the valley floor. Such behavior is also seen in the solution of certain differential equations [13] (termed "stiff"). The best approach is to avoid or reduce the coupling by transforming to a new coordinate system. Additional areas that optimization methods that are known to have difficulties include in large shallow basins, such as those occurring in loose clusters and van der Waals complexes. Estimation of the Hessian near saddle points or transition states is also an unresolved problem.

Acknowledgments. Support from Research Corporation and the donors of the Petroleum Research Fund, administered by the American Chemical Society are gratefully acknowledged.

References

1. Pulay, P.: in *Modern theoretical chemistry*, Schaefer III, H. F. Ed., New York: Plenum 1977, Vol. 4
2. Schlegel, H. B.: *Computational theoretical organic chemistry*, Csizmadia, I. G., Daudel, R., Eds., Holland: Reidel 1981
3. Schlegel, H. B.: *J. Comput. Chem.* **3**, 214 (1982)
4. Fletcher, R.: in *Numerical methods for unconstrained optimization*, Murray, W. Ed., London: Academic Press 1972; Huang, H. Y.: *J. Opt. Theory Appl.* **5**, 405 (1970) and refs. cited therein.

5. Murtagh, B. A., Sargent, R. W.: *Comput. J.* **13**, 185 (1970)
6. Pople, J. A., Krishnan, R., Schlegel, H. B., Binkley, J. S.: *Int. J. Quantum Chem. Symp.* **13**, 225 (1979)
7. Wilson Jr., E. B., Decius, J. C., Cross, P. C.: *Molecular vibrations*, New York: McGraw-Hill 1955
8. Badger, R. M.: *J. Chem. Phys.* **2**, 128 (1934); **3**, 227 (1935)
9. See for example Chapter 4 in Johnston, H. S.: *Gas phase reaction rate theory*. New York: Ronald Press 1966
10. Binkley, J. S., Whiteside, R. A., Krishnan, R., Seeger, R., DeFrees, D. J., Schlegel, H. B., Topiol, S., Kahn, L. R., Pople, J. A.: *GAUSSIAN 80*, *QCPE* **13**, 406 (1981)
11. Allinger, N. L., Yuh, Y. H.: *MM2: Molecular mechanics II*, *QCPE* **13**, 395 (1981)
12. Andose, J. D., Engler, E. E., Collins, J. B., Hummel, J. P., Mislow, K., Schleyer, P. v. R.: *BIGSTRAIN: Empirical force field calculations*, *QCPE* **10**, 348 (1978)
13. See for example: *Stiff differential equations*, Willoughby, R. A. Ed., New York: Plenum Press 1974; or any more recent text on numerical solutions of differential equations

Received April 11, 1984