ON THE DEGENERACY OF THE HESSIAN MATRIX

István KOLOSSVÁRY*,☆ and Colin McMARTIN

Research Department, Pharmaceuticals Division, CIBA-GEIGY Corporation, Summit, NJ 07901, USA

Abstract

A widespread notion in the computational chemistry literature about the Hessian matrix has been revisited, namely, that the Hessian matrix over Cartesian space is sixfold degenerate due to the three translational and three rotational degrees of freedom. It has been shown that this is true only at critical points on the potential energy hypersurface, otherwise the Hessian matrix is only threefold degenerate. The rotational degrees of freedom generally do not cause degeneracy in the Hessian matrix away from critical points.

1. Introduction

The Hessian matrix plays a central role in computational chemistry. Numerous monographs provide a detailed introduction to the calculation and use of the Hessian matrix [1,2]. The Hessian matrix is the matrix of the second derivatives of the potential energy hypersurface [3] over a nuclear configuration space representing the atomic arrangements of a molecule in three-dimensional space. The dimensionality of the nuclear configuration space depends on the coordinate system defining the atomic coordinates. The simplest coordinate system is the Cartesian coordinate system, where the atomic coordinates are simply the X, Y, Z coordinates of each atom. In this case, for a molecule with N atoms, the nuclear configuration space is a $3*N$ dimensional Euclidean space, and the potential energy hypersurface over this $3N$ space is a $3*N$ dimensional surface embedded in the $3*N+1$ dimensional Euclidean space. Numerous other coordinate systems can be used, in particular so-called internal coordinates, such as bond lengths, bond angles, and torsion angles, which provide an internal coordinate frame, i.e. a coordinate system which is independent of the spatial orientation of the molecule. This means that the internal coordinates are invariant to translation or rotation of the molecule, whereas the Cartesian coordinates

---

*To whom all correspondence should be addressed.

☆On leave until January 1993 from the Department of General and Analytical Chemistry, Technical University Budapest, Szt. Gellért 4, H-1111 Budapest, Hungary.

are not. In other words, an internal nuclear configuration space is only $3 * N - 6$ ($3 * N - 5$ for linear molecules) dimensional, six less than a Cartesian nuclear configuration space, corresponding to the three translational and three rotational degrees of freedom in the three-dimensional space in which the molecule is represented.

There are two major areas in computational chemistry where the Hessian matrix is used. One is vibrational spectroscopy, and the other is energy minimization (geometry optimization). In vibrational analysis, the Hessian is used to calculate the vibrational frequencies and the atomic displacements associated with the so-called normal modes of vibration. The normal modes of vibration define directions in which the atoms vibrate in phase, i.e., there is no interference between any pairs of the normal modes. There are $3 * N - 6$ ($3 * N - 5$ for linear molecules) normal modes of vibration, which can be experimentally measured by vibrational spectroscopy. Within the harmonic approximation, i.e., when the potential energy is quadratic in the vicinity of the equilibrium arrangements of the atoms (for flexible molecules the number of such equilibrium arrangements is usually quite large), the normal modes of vibration can be calculated using the Hessian matrix. The eigenvalues of the (mass-weighted) Hessian matrix corresponding to a local minimum point of the potential energy hypersurface are proportional to the vibrational frequencies, and the eigenvectors define the directions of the normal modes of vibration for the corresponding equilibrium arrangement of the atoms. The Hessian matrix in Cartesian space at such a local minimum point (and other critical points) on the potential energy hypersurface has six zero eigenvalues (zero frequencies), corresponding to three translations and three rotations which do not generate any vibration because neither translation nor rotation alters the interatomic distances within the molecule. Since the normal modes of vibration are almost exclusively calculated for equilibrium atomic arrangements, the validity of these calculations, even in Cartesian space, certainly cannot be called into question. However, recently, the so-called extended normal modes of vibration have been introduced [4], and used primarily to calculate zero-point energies of nonequilibrium atomic arrangements, i.e., at non-critical points on the potential energy hypersurface [5]. It will be shown that the Hessian matrix in Cartesian space is only threefold degenerate at non-critical points, and therefore the correct formula given by Mezey [5] for zero-point energy calculations at non-critical points in generalized internal space, if unduly applied in Cartesian space, must be corrected to account for the generally non-zero rotational eigenvalues of the Hessian matrix.

The other important area where the Hessian matrix is used is energy minimization. The computational chemist devotes a large amount of time to find chemically stable conformations of molecules. According to theory, such stable conformations correspond to local minimum energy points on the potential energy hypersurface. It should be noted that another class of critical points besides minima, namely, saddle points, are also very important. The saddle points correspond to transition states interconnecting stable conformations. The Hessian matrix is also used to distinguish between local minima and saddle points [1,2].

The energy minimization starts with an arbitrary, non-minimum energy structure, and the potential energy function is minimized over the nuclear configuration space to give a local minimum. During minimization, the atoms of the molecule are rearranged step-by-step in an iterative process, yielding lower and lower energy structures until convergence is reached. Convergence is reached when the gradient (the vector of the first derivatives of the potential energy function) vanishes. It should be noted that the gradient not only vanishes at minima, but at saddle points and multiple maxima as well. One of the best known and frequently used minimization methods is based on the so-called Full-Matrix Newton–Raphson (FMNR) algorithm. The FMNR algorithm is an iterative process which updates the atomic arrangement of the molecule at each iteration step by producing a displacement vector. Atoms are moved simultaneously along this displacement vector from their previous position to a new one. The consecutive new atomic arrangements usually result in a lowering of the energy. The displacement vector of the FMNR algorithm is the product of the inverse Hessian matrix and the negative of the gradient vector. The problem arises when the inverse of the Hessian has to be calculated in Cartesian space. Due to the degeneracy of the Cartesian coordinate system, the Hessian matrix has zero eigenvalues, which means that the Hessian matrix is singular and therefore cannot be inverted. There are a number of ways to eliminate the degeneracy of the Hessian matrix to resolve its singularity [1]. However, all of these methods are designed to eliminate sixfold degeneracy (three translational and three rotational), whereas it will be shown that the Hessian matrix in Cartesian space, away from minima or other critical points, is only threefold degenerate.

It should be noted that the degeneracy of the Hessian matrix in Cartesian space would automatically be eliminated by the use of internal coordinates. However, the majority of calculations (molecular mechanics calculations in particular) involving the Hessian matrix are still carried out in Cartesian space. Therefore, the use of extended normal modes and the application of the FMNR algorithm for energy minimization, in Cartesian space, requires some precaution.

2. Discussion

The threefold degeneracy of the Hessian matrix in Cartesian space arising from the translational degrees of freedom is rather trivial. There are always three eigenvectors of the Hessian which correspond to three mutually orthogonal translations of the whole molecule. Any translation of the molecule is equivalent with an appropriate linear combination of those three orthogonal translations. Since a translation does not alter the interatomic distances within the molecule, the displacement of the atoms along the translational eigenvector does not generate any vibration. Therefore, the three associated eigenvalues, i.e., the pseudo-frequencies of the translations, are all zeroes.

For that matter, rotation of the whole molecule does not change the interatomic distances either. However, this situation is more complicated than translation. The

basic difference between translation and rotation is that the translational eigenvectors of the Hessian matrix do correspond to translations of the whole molecule, whereas the rotational eigenvectors do not correspond to true rotations of the molecule. The problem is that rotation is a nonlinear displacement, and although it can apparently be represented at the infinitesimal scale by a linear translation of each atom, this is only true to a first order of approximation. To a second order of approximation (implied by the use of the Hessian matrix), interatomic distances are not invariant to linear translation even on the infinitesimal scale. We will refer to this situation as a pseudo rotation. How then is pseudo rotation related to true rotation? The rotational eigenvector points to the direction where the atoms initially move when the molecule is rotated by a small angle about an appropriate axis falling through the centroid of the molecule. An orthogonal set of three such axes serves as a basis for any rotation of the molecule, i.e., any rotation of the molecule can be replaced by the linear combination of three rotations about the basis axes.

To further investigate the difference between true rotation and pseudo rotations, let us study the problem in two dimensions. Referring back to three dimensions is straightforward. In two dimensions, the Cartesian coordinate system is threefold degenerate. There are two translations and just one rotation about the centroid of the (two-dimensional) molecule, which can be set to be the origin of the Cartesian coordinate system. In this coordinate frame, pseudo rotation can be interpreted very easily. A graphical representation of the difference between true rotation and pseudo rotation of a diatomic molecule is shown in fig. 1. True rotation about the centroid of the molecule by a small angle $\delta\omega$ moves both atoms along the circle shown in fig. 1. The direction of the movement of an atom during rotation is always the tangent of the circle at the current position of the atom. This direction is always perpendicular to the current position vector of the atom, since the position vector corresponds to the radius of the circle. The initial direction then is perpendicular to the atom's initial position vector, i.e., pseudo rotation moves one atom along vector S and the other atom along vector $-S$. A small displacement of the atoms by pseudo rotation acts on the molecule something like a sheer, and no matter how small the displacement is, the distance between the two atoms does change. The change in distance is indicated by dashed line segments between the atomic positions after true rotation (filled knob on the circle) and after pseudo rotation (open knob on the tangent of the circle). Vectors S and $-S$ are the atomic displacement equivalents of the rotational eigenvector of the corresponding 4×4 Hessian matrix. T_1 and T_2 are the translational eigenvectors, and V is the only true vibrational eigenvector representing the single normal mode of vibration of a diatomic molecule.

Pseudo rotation alters the interatomic distances no matter whether or not the atomic arrangement corresponds to a critical or a non-critical point on the potential energy hypersurface. Why then is the eigenvalue of the rotational eigenvector non-zero only at non-critical points, but zero at critical points? The answer is best given by calculus. The eigenvalue of the rotational eigenvector of the Hessian matrix is the second derivative of the potential energy function along the direction of the
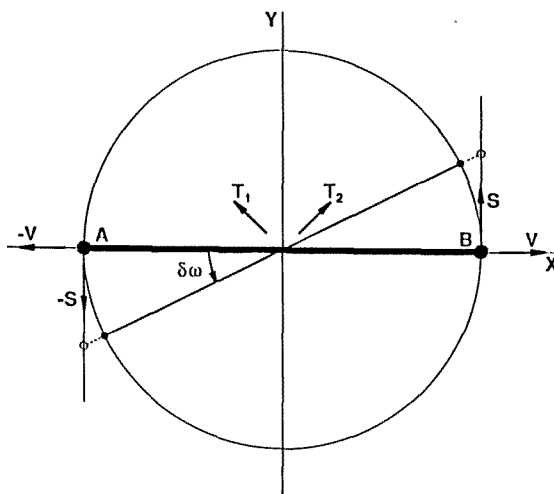
Fig. 1. Difference between true rotation and pseudo rotation. Molecule AB is represented in $X-Y$ Cartesian space. The potential energy of the molecule is a function of the interatomic distance, e.g., the familiar parabolic potential energy function $E = k(D_{AB} - D_0)^2$, where k is the force constant, D_{AB} is the actual interatomic distance, and D_0 is the equilibrium distance. The corresponding Hessian is a 4×4 matrix with four eigenvalues and eigenvectors: two translational, one rotational, and one true vibrational. The atomic displacement equivalents of the translational eigenvectors are denoted by vectors T_1 and T_2, indicating the translation of the whole molecule. The rotational eigenvector decomposes into vectors S and $-S$. Displacement of atom B along vector S and atom A along vector $-S$ is the pseudo rotation which might be called a sheer in this case. Pseudo rotation alters the interatomic distance, which is indicated by dashed line segments between the atomic positions after true rotation by a small angle $\delta\omega$ (filled knob on the circle) and after pseudo rotation (open knob on the tangent of the circle). The true vibrational eigenvector decomposes into vectors V and $-V$, indicating the single normal mode of vibration of molecule AB.

rotational eigenvector. We will refer to it as the pseudo rotational second derivative. The second derivative of the potential energy function along the true rotation can also be formulated; however, the true rotational second derivative, just like true rotational derivatives of any order, is always zero. The goal is to show that the pseudo rotational second derivative and the true rotational second derivative are different in general, but identical at critical points on the potential energy hypersurface.

Define the potential energy function E over the (X_i, Y_i) Cartesian coordinates of the N atoms of a molecule:

$$E = E(X_1, Y_1, X_2, \ldots, X_N, Y_N).$$ (1)

Also define a direction in the $2 * N$ dimensional space by a unit vector $D = (A_1, B_1, A_2, B_2, \ldots, A_N, B_N)$. The first derivative of E along D is:

$$\frac{\mathrm{d}E}{\mathrm{d}D} = \sum_{i=1}^{N} \left(\frac{\partial E}{\partial X_i} A_i + \frac{\partial E}{\partial Y_i} B_i \right). \tag{2}$$

The second derivative of E along D is as follows:

$$\frac{\mathrm{d}^2 E}{\mathrm{d}D^2} = \sum_{i=1}^{N} \sum_{j=1}^{N} \left(\frac{\partial^2 E}{\partial X_i \partial X_j} A_i A_j + 2 \frac{\partial^2 E}{\partial X_i \partial Y_j} A_i B_j + \frac{\partial^2 E}{\partial Y_i \partial Y_j} B_i B_j \right). \tag{3}$$

The direction R along the rotational eigenvector is:

$$R = (-Y_1, X_1, -Y_2, X_2, \ldots, -Y_N, X_N), \tag{4}$$

because for a position vector (X, Y), the perpendicular vector corresponding to rotation by $+90$ degrees, is $(-Y, X)$. With this, the second derivative of E along R, i.e., the eigenvalue of the rotational eigenvector, i.e., the pseudo rotational second derivative, can be formulated using eq. (3) as follows:

$$\frac{\mathrm{d}^2 E}{\mathrm{d}R^2} = C \sum_{i=1}^{N} \sum_{j=1}^{N} \left(\frac{\partial^2 E}{\partial X_i \partial X_j} Y_i Y_j - 2 \frac{\partial^2 E}{\partial X_i \partial Y_j} Y_i X_j + \frac{\partial^2 E}{\partial Y_i \partial Y_j} X_i X_j \right), \tag{5}$$

where C is a constant scaling R into a unit vector. It should be noted that in eq. (5) the X's and Y's are constants.

To calculate the true rotational derivatives of E, the chain rule is applied. The first derivative of E with respect to the rotational angle ω is:

$$\frac{\mathrm{d}E}{\mathrm{d}\omega} = \sum_{i=1}^{N} \left(\frac{\partial E}{\partial X_i} \frac{\mathrm{d}X_i}{\mathrm{d}\omega} + \frac{\partial E}{\partial Y_i} \frac{\mathrm{d}Y_i}{\mathrm{d}\omega} \right). \tag{6}$$

To calculate the derivatives of the (X, Y) coordinates with respect to ω, the rotational transform is needed. Rotation about the origin by a positive angle ω can be written:

$$X' = X \cos \omega - Y \sin \omega, \quad \text{and} \quad Y' = X \sin \omega + Y \cos \omega, \tag{7}$$

where X' and Y' are the new coordinates of the position vector (X, Y) after applying the rotation by a positive angle ω. The derivatives of these transforms are:

$$\frac{\mathrm{d}X'}{\mathrm{d}\omega} = -X \sin \omega - Y \cos \omega = -Y', \quad \text{and} \quad \frac{\mathrm{d}Y'}{\mathrm{d}\omega} = X \cos \omega - Y \sin \omega = X'. \tag{8}$$

With this, eq. (6) can be rewritten in an explicit form (X or X' is only a matter of preference):

$$\frac{dE}{d\omega} = \sum_{i=1}^{N} \left(\frac{\partial E}{\partial X_i}(-Y_i) + \frac{\partial E}{\partial Y_i} X_i \right) = 0. \tag{9}$$

It should be noted that in eq. (9), the X's and Y's are not constants. Equation (9) is equal to zero (see above), and it is the familiar form of the law of conservation of momentum.

The second derivative of E with respect to ω is also equal to zero and has the following form:

$$\frac{d^2E}{d\omega^2} = \sum_{i=1}^{N} \sum_{j=1}^{N} \left(\frac{\partial^2 E}{\partial X_i \partial X_j} Y_i Y_j - 2 \frac{\partial^2 E}{\partial X_i \partial Y_j} Y_i X_j + \frac{\partial^2 E}{\partial Y_i \partial Y_j} X_i X_j \right)$$
$$+ \sum_{i=1}^{N} \left(\frac{\partial E}{\partial X_i}(-1)X_i + \frac{\partial E}{\partial Y_i} 1(-Y_i) \right) = 0. \tag{10}$$

Comparison of eqs. (5) and (10) shows that the pseudo rotational second derivative of E and the true second rotational derivative of E behave exactly as expected, since eq. (10) is always equal to zero and eq. (5) is generally not. The only case when eq. (5) is also equal to zero is when the second term in eq. (10) vanishes. This term is the negative dot product of the atomic position vector and the gradient vector, and vanishes if and only if either the gradient vector is the zero vector (critical points on the potential energy hypersurface) or the position vector is perpendicular to the gradient vector. At first glance, this latter case might have physical significance. However, because the whole issue of rotational eigenvalues does not come up using internal coordinates, it cannot have any physical meaning, it is only a mathematical artifact of the degeneracy of the Cartesian coordinate system.

Having established the theorem in two dimensions, examine the situation in three dimensions. In three dimensions there is an orthogonal set of three rotations (see above). With appropriate coordinate transformation, the three orthogonal rotational axes can be superimposed with the X, Y, Z axes of the Cartesian coordinate system. With this, the rotation about an arbitrary axis is transformed into the much simpler case of rotation about the coordinate axes. Indeed, the rotational transform in eq. (7) and its three-dimensional version (rotation above the Z axis) look very much alike; in fact, the only difference is that in three dimensions there is an additional term ensuring that the Z coordinates are intact. Because the Z coordinates are intact, they do not appear in any differential terms, which means that in three dimensions, eqs. (5) and (10) do apply for the coordinate pair $X-Y$. Similarly, eqs. (5) and (10) (with the change of coordinates) also apply for the coordinate pairs $X-Z$ and $Y-Z$. Therefore, the rotational eigenvalues of the Hessian matrix in three-dimensional Cartesian space are generally not equal to zero. Nonetheless, they are equal to zero at critical points on the potential energy hypersurface where the gradient vanishes. A further look at eqs. (5) and (10) reveals that one or more rotational eigenvalues

may vanish at non-critical points as well, however, only with the fulfillment of rather special mathematical conditions. Again, these points do not appear to have any physical meaning.

3.   Conclusion

Any of the rotational eigenvalues of the Hessian matrix, i.e., eq. (5), which represents the pseudo rotational second derivative is not a true second derivative, since it is proportional to the second term in eq. (10) containing only first derivatives. For this reason, these eigenvalues do not contribute to the zero-point energy associated with the extended normal modes of vibration. Therefore, if the correct approximation to the zero-point energy at non-critical points introduced by Mezey [5] in generalized internal space should unduly be applied in Cartesian space, Mezey's formula formally includes the rotational eigenvalues of the Hessian matrix and those eigenvalues must be omitted from the summation, as implied in the formula [5].

As far as energy minimization is concerned, in order to resolve the singularity of the Hessian matrix in Cartesian space from a strictly mathematical point of view, one only has to eliminate the translational degree of freedom. However, it is indeed very much worthwhile to constrain the rotational degrees of freedom as well, not because the Hessian matrix could not be inverted without it, but because in this way there is no minimization time wasted in rotating the molecule back and forth. The reason why people almost exclusively use the rotational Eckart constraints [9] in FMNR minimizations is not mathematical necessity, but speed.

In summary, theoretical calculations performed in internal space involving the extended Hessian matrix cannot be performed in Cartesian space without considering the artifacts of Cartesian space on the extended Hessian matrix. Contrary to what has been implicitly or even explicitly [1, p. 68] suggested throughout the computational chemistry literature, only a handful of authors [6–8] have suggested that the rotational degeneracy of the Cartesian coordinate system is generally not reflected in zero eigenvalues associated with the rotational eigenvectors of the Hessian matrix at non-critical points on the potential energy hypersurface. Our analysis has shown that the rotational eigenvalues of the Hessian martix in Cartesian space are generally not equal to zero at non-critical points on the potential energy hypersurface; indeed, the absolute value of the rotational eigenvalues is proportional to the gradient of the potential energy function. It also means that in order to eliminate the rotational degeneracy of the Hessian matrix, one would have to extract the rotational eigenvectors themselves rather than simply looking for zero eigenvalues.

Acknowledgement

References

[1] U. Burkert and N.L. Allinger, *Molecular Mechanics*, ACS Monograph 177 (American Chemical Society, Washington, DC, 1982).

[2] P.G. Mezey, *Potential Energy Hypersurfaces* (Elsevier, 1987).

[3] Although the Hessian matrix as originally defined referred to critical points on the potential energy hypersurface, the second derivative matrix is commonly termed the Hessian matrix, or more precisely, extended Hessian matrix [4] throughout the entire potential energy hypersurface.

[4] A. Tachibana and K. Fukui, Theor. Chim. Acta (Berl.) 49(1978)321–347.

[5] P.G. Mezey, Theor. Chim. Acta (Berl.) 67(1985)115–136.

[6] D.H. Faber and C. Altona, Comput. Chem. 1(1977)203–213.

[7] W.J. Hehre, L. Radom, P. v. R. Schleyer and J.A. Pople, *Ab Initio Molecular Orbital Theory* (Wiley, 1986), p. 229.

[8] F.A.L. Anet, J. Amer. Chem. Soc. 112(1990)7172–7178.

[9] C. Eckart, Phys. Rev. 47(1935)552–558.