

Computational Method To Obtain the *F* Matrix for an Arbitrary Molecular Force Field and the *B* Matrix for Any Molecular Structure

MASARU IMAI

Computer Center, Gakushuin University, Mejiro, Tokyo 171, Japan

TAKESHI IIZUKA*

Department of Chemistry, Faculty of Education, Gunma University, Maebashi, Gunma 371, Japan

TADAYOSHI KAN

Department of Physics, Gakushuin University, Mejiro, Tokyo 171, Japan

Received July 19, 1982

A numerical method using the *F* matrix to calculate fundamental frequencies of polyatomic molecules by the *GF* method is proposed. The elements of the *F* matrix usually have been obtained by analytical approximation. Our new method is as follows. The coordinate vector *X* is set up to show the deviation from the stable conformation of a molecule and is called the sample point. The potential energies at the sample points are calculated by using an arbitrarily defined potential function. By use of the values obtained, the function is approximated to the second-order function by the least-squares method. In the same way as that used to obtain the *F* matrix, the *B* matrix is derived. This method is applied to H₂O, BF₃, and CF₄. About 100 sample points and a deviation by about 0.3 Å are found to be satisfactory for the numerical calculations of the *F* matrix and the *B* matrix, compared to the results obtained by the analytical method.

INTRODUCTION

The study of molecular vibrations by mathematical models is a basis for the analysis of infrared and Raman spectra of polyatomic molecules. The basic principles for the calculation of normal vibration frequencies were given by Wilson.¹ Shimanouchi's works^{2,3} have developed it to the remarkable achievements.

The outline of Wilson's method⁴ is as follows.

The internal potential *V* of a molecule is expressed in eq 1 by using the internal coordinate column vector *R*, the

$$V = \frac{1}{2} \mathbf{R}^T \mathbf{F} \mathbf{R} \quad (1)$$

transposed row vector *R*^T and the potential energy matrix *F*. The molecular kinetic energy *K* is expressed in eq 2, where

$$K = \frac{1}{2} \dot{\mathbf{X}}^T \mathbf{G}^{-1} \dot{\mathbf{X}} \quad (2)$$

$\dot{\mathbf{X}}$ is the derivative of *X* by time. *X* is the displacement column vector in Cartesian coordinates. *X*^T is its transposed row vector.

The kinetic energy matrix *G* is expressed as shown in eq 3,

$$\mathbf{G} = \mathbf{B} \mathbf{M}^{-1} \mathbf{B}^T \quad (3)$$

where *M* is a diagonal matrix whose diagonal elements are mass numbers. The relation between *X* and *R* is given as eq 4, where *B* is a transformation matrix. Eigenvalues λ₁, λ₂, ...,

$$\mathbf{R} = \mathbf{B} \mathbf{X} \quad (4)$$

and λ_n of *GF* can be obtained by solving the following secular equation (eq 5). The relation between λ and frequency ν is

$$|\mathbf{GF} - E\lambda| = 0 \quad (5)$$

as shown in eq 6. The following preparations are necessary

$$\nu_i = (1/2\pi)(\lambda_i)^{1/2} \quad (6)$$

for the calculation of vibrational frequencies. (a) Choose an internal coordinate system. (b) Set a molecular force field, such as the general force field, valence force field, Urey-Bradley force field, etc. (from ref 2, 4, 5, 7, and 8). (c) Construct an analytical *F* matrix from a and b. The *X* vector

and the *M* matrix are input to a computer after these preparations, and the *B* and the *G* matrices are calculated.

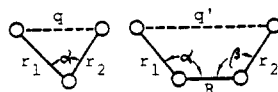
According to *GF* method (with the use of a computer), it is well-known that it is too troublesome to get the *F* matrix whose elements are analytically approximated for any molecular structure and for any force field. Especially, it is not easy to express nonbonded potential terms by other internal coordinates for a complex molecular structure. Equations 7 and 8 in Figure 1 show small displacements Δ*q* and Δ*q*' of the distances *q* and *q*', respectively, between nonbonded atoms. Thus it is difficult to calculate Δ*q* for a more complex molecular structure. Ordinarily nonlinear terms in Taylor expanded series are omitted, or a more simplified molecular model is adopted with respect to the structure or the force field.^{5,6} *F* matrices for simple molecules are well-known, but few matrices have been obtained for rather complex molecules.

As shown in the above example, generally it is very difficult to express elements of the *F* matrix in the form of a function. If it were possible to calculate the value of the *F* matrix without expressing its elements in the form of a function, the trouble cited above would be overcome. Here we introduce the terms "numerical procedure" and "analytical procedure" to clarify our present study: When the elements of *F* matrix are expressed by formulas as in Figure 2 and those values are computed by using the formulas, we call it an "analytical procedure". When the elements of the *F* matrix are not expressed explicitly by any formula and those values are directly computed as in our present study, we call it a "numerical procedure". A purpose of this study is to present our "numerical procedure" to overcome such difficulties in the usual "analytical procedure" mentioned above.

On the other hand, the elements of the *B* matrix can be generally derived by using the direction unit vector method given by Wilson.⁴ When the *B* matrix is derived by applying our "numerical procedure" which is used to derive the *F* matrix, Wilson's direction unit vectors are unnecessary.

FORMULATION

(1) **Procedure To Obtain the Elements of the *F* Matrix.** The internal potential *U* is defined by internal coordinates of atoms



$$\Delta q = (1/q) \{ (r_1 - r_2 \cos \alpha) \Delta r_1 + (r_2 - r_1 \cos \alpha) \Delta r_2 + r_1 r_2 \sin \alpha \Delta \alpha \} \quad (7)$$

$$\Delta q' = (1/q') \{ (R - r_1 \cos \alpha - r_2 \cos \beta) \Delta R + [r_1 - R \cos \alpha + r_2 \cos (\alpha + \beta)] \Delta r_1 \\ + [r_2 - R \cos \beta + r_1 \cos (\alpha + \beta)] \Delta r_2 + [R r_1 \sin \alpha - r_1 r_2 \sin (\alpha + \beta)] \Delta \alpha \\ + [R r_2 \sin \beta - r_1 r_2 \sin (\alpha + \beta)] \Delta \beta \} \quad (8)$$

Figure 1. Correlations between atomic distances from ref 7 and 8.

in a molecule (eq 9). \mathbf{R} is the internal coordinate vector of

$$U = U(\mathbf{R}, \mathbf{R}') \quad (9)$$

$$\mathbf{R} = \begin{bmatrix} R_1 \\ R_2 \\ \vdots \\ R_M \end{bmatrix} \quad (10)$$

an atom. \mathbf{R}' is the vector derived from \mathbf{R} . For an example, an element of \mathbf{R}' is the nonbonded interatomic distance if \mathbf{R} is a displacement of bond length Δr or bond angle $\Delta \alpha$ as shown in eq 7 and 8. Since the \mathbf{F} matrix is defined by eq 1, our problem is to find the \mathbf{F} matrix which satisfies the condition $U \approx V$ over an appropriate range of \mathbf{R} . This means that the function shown by eq 9 shall be approximated by a quadratic form of \mathbf{R} . The number of elements in the symmetric \mathbf{F} matrix is equal to M^2 . The number of elements to be calculated, J , is given by eq 11. (At this stage the symmetry of a molecular

$$J = (M^2 + M)/2 \quad (11)$$

structure is not yet considered.)

We use the least-squares method for adjustment so that U becomes as equal to V as possible, where U is as in eq 9 and V is as in eq 1. That is, Q will be as shown in eq 12, where

$$Q = \int_D (U - V)^2 d\mathbf{R} \quad (12)$$

D is an appropriate domain. We determine the \mathbf{F} matrix under the conditions for which Q becomes the minimum. This multiple integral is too difficult to carry out, because the number of variables M in eq 10 is too great and, furthermore, because they are not independent of each other. Therefore, we propose the following procedure. At first we choose some points from D in \mathbf{R} space as sample points (eq 13 where $\mathbf{R}^j \neq \mathbf{R}^k$ for $j \neq k$).

$$\mathbf{R}^i \text{ for } i = 1, 2, 3, \dots, L \quad (13)$$

The condition in eq 14 is, at least, necessary. It is discussed

$$L \geq J \quad (14)$$

later in detail how the sample points are chosen and a reasonable value of L is determined.

Now we use eq 15 instead of eq 12. The potential value

$$Q = \sum_{i=1}^L (U^i - V^i)^2 \quad (15)$$

for each \mathbf{R}^i can be calculated (eq 16) by use of eq 9.

$$U^i = U(\mathbf{R}^i, \mathbf{R}^i) \quad (16)$$

$$V = K' r_0 (\Delta r_1 + \Delta r_2) + \frac{1}{2} K (\Delta r_1)^2 + (\Delta r_2)^2 + H' r_0^2 \Delta \alpha + \frac{1}{2} H (r_0 \Delta \alpha)^2 + F' q_0 (\Delta q) + \frac{1}{2} F (\Delta q)^2$$

$$\mathbf{F} = \begin{bmatrix} K + t^2 F' + s^2 F & -t^2 F' + s^2 F & ts(F' + F)r_0 \\ -t^2 F' + s^2 F & K + t^2 F' + s^2 F & ts(F' + F)r_0 \\ ts(F' + F)r_0 & ts(F' + F)r_0 & (H - s^2 F' + t^2 F)r_0^2 \end{bmatrix}$$

$$s = \frac{r_0(1 - \cos \alpha)}{q_0}, \quad t = \frac{r_0 \sin \alpha}{q_0}$$

Figure 2. Potential energy and the expression of elements of the \mathbf{F} matrix for H_2O by UBFF.

Since

$$V^i = \sum_j \sum_k R_j^i F_{jk} R_k^i \quad (17)$$

for $j = 1, 2, 3, \dots, M$ and $k = 1, 2, 3, \dots, M$, our problem is to find all F_{jk} 's which make Q a minimum. That is

$$\partial Q / \partial F_{jk} = -2 \sum_{i=1}^L (U^i - V^i) R_j^i R_k^i = 0 \quad (18)$$

Rearranging expression 18, we obtain

$$\sum_i R_j^i R_k^i V^i = \sum_i U^i R_j^i R_k^i \quad (19)$$

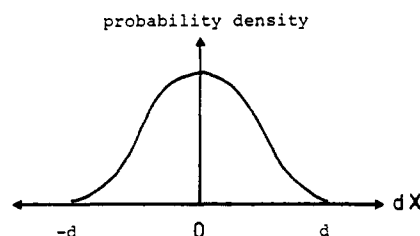


Figure 3. Normal distribution of probability density.

Substituting eq 17 into eq 19 and removing the nonindependent equations, we obtain linear equations of J variables (eq 20), for

$$\sum_i R_j^i R_k^i \sum_n \sum_m (2 - \delta_{n,m}) R_n^i R_m^i F_{nm} = \sum_i V^i R_j^i R_k^i \quad (20)$$

$j = 1, 2, 3, \dots, M; k = j, j+1, j+2, \dots, M$, where $\delta_{n,m}$ means Kronecker's δ (eq 21). Generally the number of unknown

$$\delta_{n,m} = 1 \text{ for } n = m \text{ and } 0 \text{ for } n \neq m \quad (21)$$

elements, J , in the \mathbf{F} matrix is equal to $(M^2 + M)/2$. When there may be some elements which have same values because of the symmetry of the molecular structure, J become less than that. In the case of H_2O , which has C_{2v} symmetry, the \mathbf{F} matrix has the following form

$$\begin{bmatrix} A & C & D \\ C & A & D \\ D & D & B \end{bmatrix}$$

The number of unknown elements in the \mathbf{F} matrix can be automatically reduced by our program to solve eq 20 according to the symmetry of the molecular structure.

We do not use the internal symmetry coordinates, because transforming the internal coordinates to the internal symmetry coordinates does not decrease the number of unknown elements in the \mathbf{F} matrix.

(2) How To Choose Sample Points. We ought to choose $J = {}_M C_2$ sample points at least, because we will express $U = U(\mathbf{R}, \mathbf{R})$ by only quadratic terms of M variables. The number

of sample points *L* should be large enough for high accuracy of the least-squares approximation over a wider range.

The normally distributed random numbers (see Figure 3) are respectively assigned to displacements Δx from the equilibrium positions of the atoms. In Figure 3 *d* shows a boundary of the range to generate random numbers.

We then calculate *F* and *B* matrices within the range of 0.01 Å ≤ *d* ≤ 1 Å and investigate a reasonable value of *d*.

(3) **Procedure To Obtain the Elements of the *B* Matrix.** In the same way as was used to obtain the *F* matrix, we choose some points from *D* in *X* space as sample points (eq 22). The

$$X^i \text{ for } i = 1, 2, 3, \dots, L \quad (22)$$

$$X^j \neq X^k \text{ for } j \neq k$$

internal coordinate vector *R*ⁱ can be calculated from *X*ⁱ. We assume the *Z* vector according to the definition of the *B* matrix (eq 23). Our procedure is to determine each row of the *B*

$$Z^i = BX^i \quad (23)$$

matrix under the condition in which the elements of the *R* vector and the *Z* vector are equal each to each. In order to obtain the *j*th row of the *B* matrix, we minimize the quantity *Q_j* in eq 24, where the value of *L* is restricted by the number

$$Q_j = \sum_{i=1}^L (R_j^i - Z_j^i)^2 \quad (24)$$

of atoms *N* (eq 25). Our problem is to find out all the *B_{jk}*'s

$$L \geq 3N \quad (25)$$

which make *Q_j* a minimum (eq 26 and 27). The normal

$$\partial Q_j / \partial B_{jk} = -2 \sum_{i=1}^L (R_j^i - Z_j^i) X_k^i = 0 \quad (26)$$

$$\sum_{i=1}^L \sum_{n=1}^{3N} X_n^i X_m^i B_{jn} = \sum_{i=1}^L R_j^i X_m^i \quad (27)$$

$$m = 1, 2, 3, \dots, 3N$$

equations to be solved have 3*N* variables. All of elements in the *B* matrix can be obtained in the same way for *j* = 1, 2, 3, ..., *M*. It should be mentioned that the coefficient matrices in eq 27 are constant and independent of the values of *j*.

RESULTS AND DISCUSSION

F_N, which Discussion calculated by our numerical procedure, can be evaluated according to the following $\delta_{N,F}$ (eq 28) and $\delta_{A,F}$ (eq 30). *U*ⁱ is calculated by the potential energy (eq

$$\delta_{N,F} = \sum_{i=1}^{L'} |U^i - V_N^i| / \sum_{i=1}^{L'} |U^i| \quad (28)$$

9), and *V_N* can be calculated by use of *F_N* (eq 29). Then let

$$V_N = R^T F_N R \quad (29)$$

$\delta_{A,F}$ and *V_A* be defined as follows

$$\delta_{A,F} = \sum_{i=1}^{L'} |U^i - V_A^i| / \sum_{i=1}^{L'} |U^i| \quad (30)$$

$$V_A = R^T F_A R \quad (31)$$

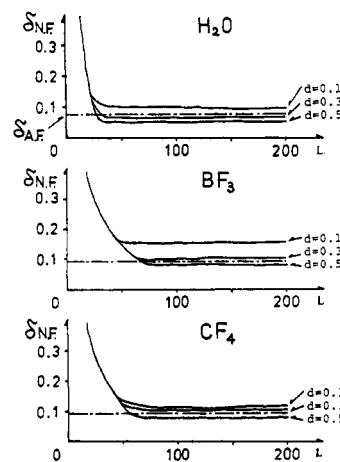


Figure 4. Values of $\delta_{N,F}$ and $\delta_{A,F}$ for the number of sample points, *L*, and the width of range, *d*, which contains them.

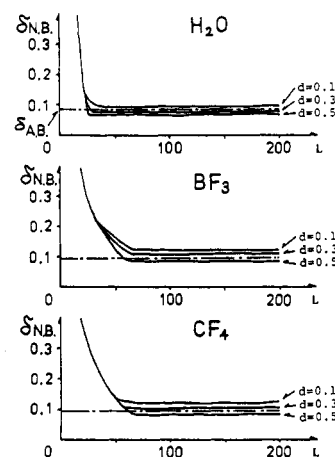


Figure 5. Values of $\delta_{N,B}$ and $\delta_{A,B}$ for the number of sample points, *L*, and the width of range, *d*, which contains them.

where *F_A* is the matrix calculated by the usual analytical procedure.

The sample points used in eq 28 and 30 must be chosen independently of those which are used in order to obtain *F_N*. The values of $\delta_{N,F}$ are nearly equal for values of *L*' = 10, 50, and 100. The case for *L*' = 50 is shown in Figure 4.

In a similar way, we can evaluate *B_N* by comparison to the values of $\delta_{N,B}$ and $\delta_{A,B}$ (eq 32 and 34). *B_N* is the *B* matrix

$$\delta_{N,B} = \sum_{i=1}^{L'} \sum_{j=1}^M |R_j^i - R_{Nj}^i| / \sum_{i=1}^{L'} \sum_{j=1}^M |R_j^i| \quad (32)$$

$$R_N = B_N X \quad (33)$$

$$\delta_{A,B} = \sum_{i=1}^{L'} \sum_{j=1}^M |R_j^i - R_{Aj}^i| / \sum_{i=1}^{L'} \sum_{j=1}^M |R_j^i| \quad (34)$$

$$R_A = B_A X \quad (35)$$

obtained by our numerical procedure. *B_A* is the *B* matrix obtained by the usual analytical procedure.

The sample points used in eq 32 and 34 must be chosen independently of those which are used in order to obtain *B_N*. The values of $\delta_{N,B}$ are nearly equal for values of *L*' = 10, 50, and 100. The case for *L*' = 50 is shown in Figure 5.

Figures 4 and 5 show results of $\delta_{N,F}$ and $\delta_{N,B}$ for the number of sample points *L* and the range $[-d, +d]$ in which the sample points are contained. These values of δ are calculated for three typical molecules, H₂O, BF₃, and CF₄. The reasonable value of *L* depends on the number of internal coordinates *M* and the number of atoms *N* under the conditions of eq 14 and 25. The values of *L* less than 100 might be insufficient, and even if *L* becomes more than 150, δ stands unchanged in each case.

Therefore, the reasonable value of L is about 100.

$\delta_{A,F}$ is between 0.07 and 0.1 which is calculated by using the F matrix and the B matrix derived by the usual analytical method. $\delta_{A,F}$ and $\delta_{A,B}$ derived from F_A and B_A , which have been mostly used to study molecular vibrations, can be the standards to evaluate F_N and B_N . As shown in Figures 3 and 4, δ becomes a minimum when d is 0.5. When d is smaller than 0.3, $\delta_{N,F}$ and $\delta_{N,B}$ become larger than $\delta_{A,F}$ and $\delta_{A,B}$, respectively. If the physical meaning of d is considered to be the maximum amplitude, it is better for d not to exceed 0.5. Therefore, a reasonable value of d is between 0.3 and 0.5 Å.

ACKNOWLEDGMENT

We express our hearty thanks to Prof. Mitsuo Tasumi, University of Tokyo, who made a critical reading of our manuscript.

Registry No. Water, 7732-18-5; boron trifluoride, 7637-07-2; tetrafluoromethane, 75-73-0.

REFERENCES AND NOTES

- (1) Wilson, E. B. "A Method of Obtaining the Expanded Secular Equation for the Vibration Frequencies of a Molecule". *J. Chem. Phys.* **1939**, *7*, 1047-1052.
- (2) Shimanouchi, T. *Nippon Kagaku Zasshi* **1965**, *86*, 261-281.
- (3) Shimanouchi, T. "Molecular and Crystal Data System". "Proceedings of the 6th International CODATA Conference"; Pergamon Press: Oxford, England, 1979; Vol. 6, pp 133-136.
- (4) Wilson, E. B.; Decius, J. C.; Cross, P. C. "Molecular Vibrations"; McGraw-Hill: New York, 1955.
- (5) Morino, Y.; Kuchitsu, K.; Shimanouchi, T. "The Mean Amplitudes of Thermal Vibrations in Polyatomic Molecules. 1. $CF_2=CF_2$ and $CH_2=CF_2$ ". *J. Chem. Phys.* **1952**, *20*, 726-733. Overend, J.; Scherer, J. R. "Transferability of Urey-Bradley Force Constants. 1. Calculation of Force Constants on a Digital Computer". *J. Chem. Phys.* **1960**, *32*, 1289-1295.
- (6) The redundancies of the gem and tree are shown by (2).
- (7) Shimanouchi, T.; Suzuki, I. "Method of Adjusting force Constants and Its Application to H_2O , H_2CO , CH_2Cl , and Their Deuterated Molecules". *J. Chem. Phys.* **1965**, *42*, 296-308.
- (8) Shimanouchi, T.; Nakagawa, I.; Hiraishi, J.; Ishii, M. "Force Constants of CF_4 , SiF_4 , BF_3 , CH_4 , SiH_4 , NH_3 , and PH_3 ". *J. Mol. Spectrosc.* **1966**, *19*, 78-107.