

Regularizing Procedure for Scaling *ab initio* Molecular Force Fields in Cartesian Coordinates

I. V. Kochikov*, A.V.Stepanova**, G.M.Kuramshina***

* Moscow State University,
Scientific Research
Computer Centre,
119992, Moscow, Russia
E-mail: igor@kochikov.ru

** Moscow State University,
Faculty of Chemistry,
Dept. Phys. Chemistry
119992, Moscow, Russia
E-mail: stepanna@hotmail.ru

*** Moscow State University
Faculty of Chemistry,
Dept. Phys. Chemistry
119992, Moscow Russia
E-mail: kuramshi@phys.chem.msu.ru

*The work of the first and third authors was supported by the RFBR-OB' (grant 05-07-98001_r),
the work of the third author was also supported by the RFBR (grant 05-03-35135_a)*

Inverse vibrational problem is a problem of finding the molecular force field parameters on a base of experimental data. In general case, this problem is written in a form of nonlinear operator equation

$$A_h F = \Lambda_\delta \quad (1)$$

where $F \in Z \in R^{n(n+1)/2}$ (Z is a set of possible solutions) is the unknown force constant matrix (real and symmetrical), $\Lambda \in R^m$ represents the set of available experimental data (vibrational frequencies, etc.) determined within δ error level: $\|\Lambda - \Lambda_\delta\| \leq \delta$. A is a nonlinear operator which maps matrix F on the Λ , h is a certain numerical estimate of the operator A uncertainty.

It is a common knowledge that the problem of calculating molecular force fields in the general approximation of small vibrations is an ill-posed problem (it does not satisfy any of the three well-posedness conditions by Hadamard). In most cases, the main difficulty arises from the fact that the problem does not have a unique solution. Indeed, for any molecular structure (except diatomic molecules), there may exist an infinite number of force field matrices that result in the same normal vibration frequencies. One of the ways to overcome the ambiguity of the solution lies in bringing in extra experimental data (e.g. frequencies of isotopic species of a molecule, or microwave spectra). However, for the large molecules it is quite common that the whole set of available experimental data is still insufficient to define the unique solution to the problem; on the other hand, for the wider set of experimental data the problem may have no solutions at all. Besides, even if the solution exists, it may be unstable with respect to the errors in operator A and the set of experimental data Λ .

There are a lot of publications describing application of numerical methods to the solving inverse vibrational problem on a base of the least-square procedure. Very often as criteria of minimization, the authors choose the "best" agreement between experimental and fitted vibrational frequencies. Regardless of the meaning of the "best agreement" criteria, we just note that it is insufficient due to the ill-posed nature of the inverse vibrational problem. Even in the case of a single molecule, it is a well-known fact that there exists infinite set of solutions, which exactly reproduce the given set of experimental frequencies. Inclusion of the expanded experimental information on isotopomers or related molecules frequencies might lead to the incompatibility of the mathematical problem and result in no solution at all. So, using any minimization procedure for solution of inverse vibrational problem it is necessary to include some additional criteria (that can be mathematically formulated) in the minimization procedure for selecting the unique solution.

The theory of regularization methods for nonlinear problems was developed in the last two decades [1] and applied to the inverse problems of vibrational spectroscopy [2].

We have proposed a principally new formulation of the problem of searching the molecular force field parameters using all available experimental data and quantum mechanical calculation results, and taking into account the a priori constraints for force constants. The essence of the approach is that using given experimental data and its accuracy we try to find by means of stable numerical methods the approximations to the so-called *normal pseudosolution*, i.e. matrix F which is the nearest by the chosen Euclidean norm to the given force constant matrix F^0 and satisfies to the set of a priori constraints D and experimental data Λ_δ with regard for the possible incompatibility of the problem [2]. The inclusion of some restrictions on the sought matrix of force constants (in our case it is a requirement of the closeness of solution to the matrix F^0) allows to obtain the unique solution from the variety of possible ones. The matrix F^0 may represent some model assumptions on the nature of intramolecular interactions; nowadays it is usually

obtained from quantum-chemical calculations and the optimized solution we define as *Regularized Quantum Mechanical Force Field* (RQMFF) [3].

The wide use of quantum mechanical calculations of vibrational spectra and harmonic force fields of polyatomic molecules induced the necessity of empirical corrections to the theoretical data for compatibility with experiment. The most popular approach is the so-called scaling procedure proposed by P. Pulay [4]. Though this scaling procedure often does not provide sufficient freedom to eliminate all discrepancies between calculated and observed data (in fact, it imposes rather strict limitations on the molecular force field [5, 6]), it has certain advantages that follow from the comparatively small number of adjustable parameters and so moderate computational resources are necessary to perform force-field refinement. Indeed, it is very attractive to find a limited number of scaling factors for a series of model molecules and (assuming their transferability) to use them for the correction of quantum mechanical force constants of more complicated molecular systems.

Pulay parameterization method is based on the representation of the force constant matrix as

$$F = BF^0B \quad (2)$$

where B is a diagonal matrix of scale factors, and F^0 is an *a priori* chosen matrix – here, again, matrix F^0 is usually a result of quantum-chemical calculations. Strictly speaking, such parameterization does not completely remove the ambiguity of the solution; however, this ambiguity may be resolved using ideas of regularization (by searching for the scale matrix B closest to unit matrix or by searching for the matrix F closest to F^0 , as we have shown in our previous studies [2]).

The scaling procedure (2) has initially been suggested for the force fields defined in the internal coordinate system. It has been shown that for many structural fragments of the molecules scale factors (for a given quantum-mechanical method) are approximately constant within a wide range of similar molecules. As a result, for many coordinates and quantum-mechanical methods, recommended force constants have been obtained which in most cases allow to approximate experimental frequencies with a reasonable degree of accuracy.

However, in the course of spectroscopic and structural research, introduction of the complete system of internal coordinates is the most tedious and time-consuming procedure, especially for the large molecular systems. From quantum chemistry we usually obtain force constant matrix in Cartesian coordinates. The idea of this paper is to avoid introducing internal coordinates but – at the same time – achieve compatibility of the Cartesian force fields with the experimental data in a frame of regularization approach.

This approach results in the introduction of scaling procedure for Cartesian force fields. Formally, scaling is still given by Eq. (2); however, as we will see later, matrix B cannot be assumed diagonal any more. The main peculiar feature of the force matrices in Cartesian coordinates is in the fact that they are not automatically independent of the molecular position and orientation. Physically meaningful force constant matrix should therefore satisfy a number of constraints. Let the force matrix be represented as an array of 3 by 3 submatrices corresponding to each atom:

$$F = \begin{pmatrix} f_{(11)} & f_{(12)} & \cdots & f_{(1N)} \\ f_{(21)} & f_{(22)} & \cdots & f_{(2N)} \\ \cdots & \cdots & \cdots & \cdots \\ f_{(N1)} & f_{(N2)} & \cdots & f_{(NN)} \end{pmatrix}$$

Here N is number of atoms in a molecule. Then the constraints are as follows (see our book [2]):

$$\sum_{i=1}^N f_{(ij)} = 0, \quad \sum_{i=1}^N V_i f_{(ij)} = 0, \quad j = 1, 2, \dots, N \quad (3)$$

where

$$V_i = \begin{pmatrix} 0 & -R_{iz}^0 & R_{iy}^0 \\ R_{iz}^0 & 0 & -R_{ix}^0 \\ -R_{iy}^0 & R_{ix}^0 & 0 \end{pmatrix},$$

and $R_{ix}^0, R_{iy}^0, R_{iz}^0$ are Cartesian components of the i -th atom equilibrium position. Equations (3) ensure that the force field represented by the matrix F is independent of displacement and rotation of a molecule as a whole.

In the scaling procedure (2) applied to the matrix F in Cartesian coordinates, we may assume that a priori matrix F^0 satisfies the requirements (3). However, this does not necessarily mean that the scaled matrix also satisfies these requirements. To ensure that scaled matrix is also physically meaningful, it is necessary to introduce constraints on the scale matrix B . It can be shown that matrix B should satisfy the following conditions:

1) Matrix B should consist of the unit 3×3 submatrices multiplied by some factors:

$$B = \begin{pmatrix} \beta_{11}E & \beta_{12}E & \dots & \beta_{1N}E \\ \beta_{21}E & \beta_{22}E & \dots & \beta_{2N}E \\ \dots & \dots & \dots & \dots \\ \beta_{N1}E & \beta_{N2}E & \dots & \beta_{NN}E \end{pmatrix}$$

2) The factors β_{ij} ($i, j, = 1, \dots, N$) should satisfy the conditions

$$\beta_{ij} = \beta_{ji};$$

$$\sum_{i=1}^N \beta_{1i} = \sum_{i=1}^N \beta_{2i} = \dots = \sum_{i=1}^N \beta_{Ni} = S = \text{const.} \quad (4)$$

If there are some extra constraints due to the symmetry or model assumptions, they should be used in addition to these constraints. Technically, matrix B contains $N(N-1)/2+1$ independent parameters, since all diagonal elements may be represented as

$$\beta_{ii} = S - \sum_{j \neq i} \beta_{ij}.$$

So we use the formulation of inverse vibrational problem in a form (1) when a set of a priori constraints on molecular force field D includes conditions (2) and (4). An example of solving an inverse problem by the scaling Cartesian force field for CCl_4 molecule is given below. We have started with *ab initio* calculation of Cartesian force field F^0 which gives the normal vibration frequencies shown in the first column of the following table.

Experimental and calculated frequencies of CCl_4 .

Symmetry	Ab initio (cm^{-1})	Fitted (cm^{-1})	Experimental (cm^{-1})
A_1	469	454	459
E	226	219	218
F_2	812	776	776
F_2	325	314	312

The scaling procedure (2) was then applied to Cartesian force field F^0 . The aim was to fit experimental frequencies shown in the last column of the table. The results of scaling are shown in the second column; it is easy to see that fitting has produced the results very close to the experimental. In this case, due to the symmetry restrictions, there were only 3 independent fitting parameters. The resulting matrix of β_{ij} coefficients is shown below.

$$[\beta] = \begin{pmatrix} & \text{C} & \text{Cl}_1 & \text{Cl}_2 & \text{Cl}_3 & \text{Cl}_4 \\ \text{C} & 0.9634 & 0.0084 & 0.0084 & 0.0084 & 0.0084 \\ \text{Cl}_1 & 0.0084 & 0.9723 & 0.0055 & 0.0055 & 0.0055 \\ \text{Cl}_2 & 0.0084 & 0.0055 & 0.9723 & 0.0055 & 0.0055 \\ \text{Cl}_3 & 0.0084 & 0.0055 & 0.0055 & 0.9723 & 0.0055 \\ \text{Cl}_4 & 0.0084 & 0.0055 & 0.0055 & 0.0055 & 0.9723 \end{pmatrix}.$$

In this matrix, each row and column correspond to one of the atoms of CCl₄ molecule.

The main goal of our further investigation is to make clear if the Cartesian scale factors reveal transferability properties similar to the properties of the scale factors in internal coordinates. These properties are very important for vibrational spectra calculations of relative compounds, especially among large biological molecules, associates and nanostructures.

References

1. A.N.Tikhonov, A.S.Leonov, A.G.Yagola. *Nonlinear Ill-posed Problems*. Chapman&Hall, London, 1998 (Original Russian language edition: *Nonlinear Ill-posed Problems*. Nauka, Moscow, 1993).
2. A.G.Yagola, I.V.Kochikov, G.M.Kuramshina, Yu.A.Pentin. *Inverse Problems of Vibrational Spectroscopy*. VSP, Zeist, The Netherlands, 1999.
3. G.M.Kuramshina, F.A.Weinhold, I.V.Kochikov, Yu.A.Pentin, A.G.Yagola. *J.Chem.Phys.* **100** (1994) 1414.
4. P. Pulay, G.Fogarasi, G.Pongor, J.E.Boggs, A.Vargha.. *J. Am. Chem. Soc.*, **105** (1983) 7037.
5. G. M.Kuramshina, A.G.Yagola, *Russian J. Struct. Chem.* **38** (1997) 181.
6. G. M.Kuramshina, F. Weinhold, *J. Mol. Struct.*, **410** (1997) 457.