

УДК 532.6

**ВЫЧИСЛИТЕЛЬНЫЕ АСПЕКТЫ РАСЧЕТА МАСШТАБИРУЮЩИХ МНОЖИТЕЛЕЙ
ПО ЭКСПЕРИМЕНТАЛЬНЫМ ДАННЫМ****И. В. Кочиков¹, Г. М. Курамшина², А. В. Степанова², А. Г. Ягола³**

Разработаны новые методы расчета набора масштабирующих множителей, определяющих силовое поле молекулы. Предложена новая постановка обратной колебательной задачи (так называемая обратная задача масштабирования). Приведены примеры расчета масштабирующих множителей для молекул метилсилана и перфторэтана в различных приближениях.

Ключевые слова: обратная колебательная задача, некорректные задачи, масштабирующие множители, регуляризация, метилсилан, математическое моделирование, численные методы.

1. Introduction. Many problems of physical chemistry belong to the class of inverse problems. In these problems, the properties of an object are determined from experimental data on the basis of a certain model connecting these properties with characteristics measured. Inverse problems typically lead to mathematical models that are not well-posed in the sense of Hadamard. This means that they may not have a solution in the strict sense; when solutions exist, they may not be unique and/or may not depend continuously on input data. Mathematical problems possessing such properties are called *ill-posed* problems, mostly due to instability of solutions with respect to data perturbations. Numerical methods that can cope with this problem are the so-called *regularization methods*. The theory of ill-posed problems (founded by A. N. Tikhonov and his scientific school in 1960s; see, for example, [1]) investigates and develops efficient stable numerical methods for the solution of ill-posed problems. An understanding of undetermined character of ill-posed problems and the concept of regularizing operators (algorithms) form the basis for this theory.

Molecular force fields provide important information on the molecular structure and dynamics. Molecular properties may be determined from experimental data of vibrational (infrared and Raman) spectroscopy as a result of solving the so-called inverse vibrational problems. The rapid progress in quantum mechanical calculations of theoretical harmonic force fields provides new ways for more accurate interpretation of experimental data and opens new opportunities in empirical force field calculations. The latter are particularly important for the large-size molecules when accurate *ab initio* calculations are impossible, so that empirical methods based on solving the inverse vibrational problems still remain the main source of data on the force field parameters.

The rapid progress in the investigation of rather large nanomolecules requires the development of special approaches for solving inverse vibrational problems lying far beyond the traditional methods based on the least-squares procedures. The analysis of large molecular systems (when a force constant matrix F is constructed from previously evaluated force constants of model compounds) runs across difficulties of possible incompatibility of the results obtained by different authors and by means of different numerical methods within different approximations (force field models). These difficulties are caused by nonuniqueness and instability of the solution to a inverse vibrational problem as well as by incompatibility of available experimental data with the harmonic model.

In this paper we demonstrate how *a priori* model assumptions and *ab initio* quantum mechanical calculations can be used for constructing the regularizing algorithms for the molecular force field calculations. We have proposed a principally new formulation of the problem of searching for 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 our approach is that we suggest (using given experimental data and its accuracy) to find the stable approximations to the so-called *normal pseudosolution* (i.e., to find a matrix F such that it is the nearest (in the chosen Euclidean norm) to a given force constant matrix F^0 , satisfies a set of *a priori* constraints D , and is compatible with experimental data Λ_δ with regard to the possible incompatibility of the problem [2]).

¹ Научно-исследовательский вычислительный центр, Московский государственный университет им. М. В. Ломоносова, 119992, Москва; e-mail: kochikov@tm-net.ru

² Химический факультет, Московский государственный университет им. М. В. Ломоносова, 119992, Москва; e-mail: kuramshi@phys.chem.msu.ru

³ Физический факультет, Московский государственный университет им. М. В. Ломоносова, 119992, Москва; e-mail: yagola@inverse.phys.msu.ru

Within this approach, the results tend to be as close to quantum mechanical data as an experiment allows. From the mathematical point of view, the corresponding algorithm provides approximations to the solution that tend to the exact solution when experimental data become more extensive and accurate. Imposing certain restrictions on the matrix of force constants (in our case, it is the requirement on the closeness of the solution to the matrix F^0) allows one to select a unique solution from the variety of possible choices.

2. Mathematical formulation of inverse vibrational problems. The idea of the force field arises from the treatment of a molecule as a mechanical system of nuclei where all the interactions due to electrons are included into the effective potential function $U(q_1, \dots, q_n)$. Here by q_1, \dots, q_n we denote $n = 3N - 6$ generalized coordinates of N -atomic nuclei of the molecule. The potential function minimum with respect to the nuclei coordinates defines the equilibrium geometry of the molecule; the second derivatives of the potential at the equilibrium

$$f_{ij} = \frac{\partial^2 U}{\partial q_i \partial q_j} \Big|_{eq} \quad (i, j = 1, \dots, n)$$

constitute a positive definite matrix F determining all the molecular characteristics related to small vibrations. The vibrational frequencies (obtained from the IR and Raman spectra) are the main type of experimental information on molecular vibrations. They are connected with the matrix of force constants by the eigenvalue equation

$$GFL = L\Lambda, \quad (1)$$

where Λ is the diagonal matrix composed of the squares of the molecular normal vibration frequencies $\omega_1, \dots, \omega_n$, $\Lambda = \text{diag} \{ \omega_1^2, \dots, \omega_n^2 \}$, G is the kinetic energy matrix in the momentum representation, and L is the matrix of normalized relative amplitudes.

If only the experimental frequencies of one molecular isotopomer are known, the inverse vibrational problem of finding the force constant matrix F is reduced to an inverse eigenvalue problem; hence, when G is not singular, the solution to equation (1) is represented by any matrix F such that

$$F = G^{-1/2} C * \Lambda C G^{-1/2}, \quad (2)$$

where C is an arbitrary orthogonal matrix (the asterisk indicates the transposed matrix).

Since equation (1) is the main source of data determining the force constants, it is evident that (except for diatomic molecules) the $n(n+1)/2$ parameters of F cannot be found uniquely from the n frequencies $\omega_1, \dots, \omega_n$. This has led, on the one hand, to attempts of using certain model assumptions on the structure of the matrix F and on the other hand, to introducing additional experimental data. Within the approximation considered, the force field of a molecule does not depend on the nuclei masses; for the spectra of m isotopic species we have, instead of equation (1), the system

$$(G_i F) L_i = L_i \Lambda_i, \quad i = 1, 2, \dots, m. \quad (3)$$

Additional information may also be extracted from ro-vibrational spectra (Coriolis constants), gas electron diffraction (mean-square amplitudes), and from other experimental evidence where the measured molecular constants depend on the force constant matrix F . It is important to note, however, that even if the number of available experimental measurements is equal to or exceeds the number of parameters to be determined, the uniqueness of the solution is not guaranteed.

The mathematical relations between the vibrational properties of a molecule (equations (1), (3), etc.) and their experimental representation can be summarized in the form of a single operator equation

$$AF = \Lambda_\delta, \quad (4)$$

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

This problem belongs to the class of nonlinear ill-posed problems, since it does not satisfy any of the three well-posedness conditions according to Hadamard. In the general case (except for the diatomic molecules), it may have a nonunique solution or no solutions at all (the incompatible problem); solutions may be unstable with respect to errors in the operator A and in the set of experimental data Λ . Regularization methods for nonlinear problems were developed in the last two decades [1] and were applied to the inverse problems of vibrational spectroscopy [2].

The most important idea of the regularization theory is that experimental data are insufficient for the unique and stable determination of some or all molecular parameters; hence, we should introduce some kind

of external knowledge or experience. In some cases, it is desirable to formulate these considerations as explicit additional restrictions on the set of possible solutions. When it is impossible, the more flexible approach is to choose the solution nearest (in a certain sense) to some *a priori* defined parameter set. This set may not necessarily conform to the experiment and should be based on data complementary to the experiment.

This external evidence may be derived from general ideas (for example, from molecular force field models or data on similar molecular structures) or, preferably, be based on *ab initio* calculations. An inverse vibrational problem is formulated as a problem of finding the so-called normal solution (or normal pseudo-(quasi-)solution in the case of incompatibility of input data) to the nonlinear operator equation (4).

The sought-for solution is a matrix $F^\alpha \in Z$ that reproduces experimental data within a given error level and is the nearest (in the Euclidean metrics) to some given matrix F^0 . All necessary model assumptions (explicit and implicit) concerning the form of the force field may be taken into account by the choice of some given *a priori* matrix of force constants F^0 and a pre-assigned set D of *a priori* constraints on the values of force constants. This set defines a form of the matrix F in the framework of the desired force field model (i.e., with specified zero elements, equality of some force constants, etc.). If no *a priori* data constrain the form of a solution, then D coincides with the set Z .

3. Constraints on the values of force constants based on *ab initio* calculations. The inverse scaling problem. It was stated earlier [3, 4] that in Tikhonov's regularizing procedure it is possible to increase the stability and accuracy of the calculated solution F^α by using

- a) an extended set of experimental data (including, for example, the Coriolis constants, mean-square amplitudes, frequencies of isotopomers or related compounds, etc.);
- b) an improved choice of the stabilizer matrix F^0 ;
- c) an improved choice of the constraint set D .

As a particularly effective choice of a stabilizer, we have proposed [5] to use the *ab initio* quantum mechanical matrix F^0 in the regularizing procedure. This leads to the concept of the regularized quantum mechanical force field (**RQM FF**) defined as the force constant matrix which is nearest to the corresponding quantum mechanical matrix F^0 and reproduces experimental frequencies within a given error level.

A correct choice of the constraint set D is also extremely important. Physically stipulated limitations may either decrease the range of acceptable matrices F or may provide criteria for selecting a unique solution from the set of tolerable ones. An incorrect choice of constraints may lead to an increased incompatibility of an inverse problem, eventually resulting in a pseudosolution with no physical meaning. A set of *a priori* constraints may arise from several types of limitations on force constant values (see, for example, [3, 4]):

1. some force constants may be assumed *a priori* to be zero;
2. some force constants may be assumed to satisfy the inequalities $a_{ij} \leq f_{ij} \leq b_{ij}$, where a_{ij} and b_{ij} are certain known values;
3. some force constants may be assumed to be equal in a series of related molecules (or conformers).

During the last 30 years, numerous attempts were undertaken in the development of quantum mechanical methods for evaluation of molecular structure and molecular dynamics of polyatomic molecules directly from Schrödinger's equation. At present, direct calculations of vibrational spectra by quantum mechanical calculations at different levels of theory are very important and widely used to interpret experimental data, especially in the case of large molecules. However, the use of very restrictive assumptions was (and still remains) necessary to make such calculations computationally feasible.

Most *ab initio* methods are based on the relatively simple Hartree–Fock approach (HF). At more advanced levels of theory, electron correlation neglected by the HF-approach is taken into account by means of the perturbation theory (the Moller–Plesset method (MP)) or by using the density functional theory (for example, the Becke method referred to as B3LYP). Account for electron correlation, though being computationally expensive, is proved to be very important for obtaining the correct equilibrium structure of many compounds. Another important feature of *ab initio* calculations is the size of the used basis set. The standard basis sets are composed of Gaussian functions and differ by the number of such functions (among the most popular sets, the basis set 6-31G* is relatively scarce, while 6-311++G** is more rich). Under certain circumstances, the standard basis sets have to be complemented by special functions; such extended sets have special notations (for example, AUG-cc-pVDZ, etc.). The basis extension is also computationally costly, so the usual practice is to find a reasonable compromise between the size of the basis set and the level of the theory. *Ab initio* calculations are usually denoted by the specifying level of a theory and a basis set, for example, the notation B3LYP/6-311++G** means that the DFT method with one of the Gaussian basis sets is used.

Among different approaches, the Hartree–Fock level calculations are routine and available even for very large systems consisting of up to hundreds of atoms. However, the quality of these calculations is insufficient

for direct comparison of theoretical and experimental vibrational frequencies. As a rule, the HF frequencies are overestimated (up to 15 %; these errors have a systematic character for the related compounds. The wide use of quantum mechanical calculations of vibrational spectra and harmonic force fields of polyatomic molecules induced the necessity of making empirical corrections to theoretical data for compatibility with an experiment. The most popular approach is the so-called scaling procedure proposed by P. Pulay [6], where the disagreement between experimental and theoretical frequencies is eliminated by introducing a finite (not very large) number of scaling factors.

This approach can also be formulated in the following form of *a priori* constraints: the final solution may be assumed to conform to Pulay's scaled force constant matrix [6], which may also be considered as a kind of constraint.

In this approach we can specify [3, 4] the set D as: $D = \{F: F = BF^0B\}$, $B = \text{diag}\{\beta_1, \dots, \beta_n\}$, where β_i are the scaling parameters.

Mathematically, the scaling procedure imposes rather strict limitations on the molecular force field [3] and often does not provide enough freedom to eliminate all discrepancies between the calculated and observed data. However, it has certain advantages that follow from the comparatively small number of adjustable parameters and, consequently, moderate computational resources required 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 to correct the quantum mechanical force constants of more complicated molecular systems. The most popular numerical procedure for calculation (optimization) of scaling factors is the least-squares procedure, but there are a few publications indicating the nonconvergence and instability of this numerical procedure while solving an inverse scaling problem. This is explained by the impossibility of using the traditional numerical methods for solving nonlinear ill-posed problems [1, 2, 7, 8].

In our works the following strict mathematical formulation of the inverse scaling problem has been proposed [9, 10]: the problem of finding scaling factors on the basis of experimental data is treated as an operator equation similar to (4):

$$AF(\beta) = \Lambda_\delta, \quad (5)$$

where β are scaling factors. Let the following norms in the Euclidean space be introduced:

$$\|F\| = \left[\sum_{ij} f_{ij}^2 \right]^{1/2}, \quad \|\Lambda\| = \left[\sum_{k=1}^l \lambda_k^2 \rho_k \right]^{1/2},$$

where $\rho_k > 0$ are certain positive weights, f_{ij} are the elements of the matrix F ; λ_k ($k = 1, \dots, l$) are the components of Λ .

Since problem (5) is also ill-posed, we have to regularize it. We formulate the problem as a requirement to find an approximation to such a solution $F_{n,\delta}$ to equation (5) that it is nearest (in Euclidean norm) to the quantum mechanical matrix F^0 , satisfies experimental data within a given error level δ ($\|A(F(\beta)) - \Lambda_\delta\| \leq \delta$), and has a special form proposed by Pulay. If we consider this problem taking into account its possible incompatibility, we come to the following formulation:

to find

$$F_{n,\delta} = \arg \min \|F - F^0\|$$

where

$$F \in \left\{ F : F \in D = \left\{ F : F = B^{1/2} F^0 B^{1/2} \right\}, \|AF - \Lambda_\delta\| \leq \mu + \delta \right\}.$$

Here B is a diagonal matrix of scaling factors β_i and μ is a measure of incompatibility of the problem [2]. The latter may appear due to the possible anharmonicity of experimental frequencies or the crudeness of the chosen model.

Finding the matrix $F_{n,\delta}$ may be performed by minimization of the Tikhonov functional

$$M^\alpha(\beta) = M^\alpha[F] = \|A_h F - \Lambda\|^2 + \alpha \|F - F^0\|^2,$$

where $F = F(\beta)$ and the regularization parameter α is chosen in accordance with the generalized discrepancy principle [1].

As a rule, the assumed limitations on the values of force constants of polyatomic molecules cannot be strictly proved. Nevertheless, numerical quantum mechanical results on molecular force fields can provide useful guidance in choosing realistic force field models for different types of molecules.

The least-squares procedure applied to solution of a inverse vibrational problem has been described in numerous publications. Very often, as a criterion of minimization, the authors choose the “best” agreement between experimental and fitted vibrational frequencies. However, there are situations when the “best” agreement criterion is of many meanings. We should like to note that this criterion 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 exist an infinite number of solutions which exactly satisfy any given set of the experimental frequencies. Multiple solutions also have been shown to exist when scaling procedure is applied. Addition of the expanded experimental information on frequencies of the isotopomers or related molecules may lead to incompatibility of the mathematical problem and may result in no solution at all within the conventional harmonic model. This means that, using any minimization procedure for solving the inverse vibrational problem, it is necessary to apply some additional criteria (that can be mathematically formulated) in the minimization procedure to select the unique solution.

4. Generalized inverse structural problem. A similar regularized procedure was proposed for solving the generalized inverse structural problem (GISP) [11] in the case of joint treatment of the experimental data obtained by different physical methods (the vibrational spectroscopy and the electron diffraction (ED) data and microwave (MW) spectroscopy). In general, solving the generalized inverse problem should lead to the simultaneous determination of molecular geometry (R) and the force field (F), each described by a finite set of certain parameters.

Let Λ be a set of all available experimental data and $A(R, F)$ be a procedure allowing to calculate these data from a set of molecular parameters R and F . Here Λ is a finite-dimensional vector from the normalized space R^m ; the parameters (R, F) may also be chosen so as to constitute a vector from R^n ; hence, A is an operator acting from R^n to R^m . We can write down the following nonlinear operator equation similar to (4):

$$A(R, F) = \Lambda_\delta. \tag{6}$$

This equation has some properties similar to those of equation (4): the solution of equation (6) may be not unique and/or may reveal instability with respect to experimental measurement errors. To solve this problem and obtain the solution, we again need to involve external assumptions based on some preliminary knowledge, analogous to those formulated for the inverse vibrational problem and, preferably, based on *ab initio* data.

In the case of GISP, this kind of “soft” constraints may be combined with a more rigid set of constraints imposed on the possible solutions in order to obtain the unique solution. For example, when there are several close (by magnitude) interatomic distances in a molecule, it is a common practice to determine only one of them from ED experiment, fixing all differences between the distances in *ab initio* magnitude. When it is impossible to find a solution within this model, a less restricting approach could be applied to find a solution that will have the same properties, unless this approach does not contradict to experiment. Otherwise, we will look for the solution nearest to the element possessing these properties.

Now let us give a strict formulation of the problem. Let (R^0, F^0) be an *a priori* given set of parameters (obtained, for example, from *ab initio* calculations), $\|\Lambda - \Lambda_\delta\| \leq \delta$, and δ be an experimental error level. Introduce a set of constraints D in R^n , which our solution should belong to. The *generalized inverse structural problem (GISP)* may be formulated as follows:

to find an approximation $(R, F)_\delta$ to the exact (normal) parameters (R, F) such that

$$(R, F)_\delta = \arg \min_{(R, F) \in Z_\delta} \|(R, F) - (R^0, F^0)\|,$$

where the solution is compatible with experimental data within the accuracy range, taking into account the possible measure of incompatibility μ

$$(R, F)_\delta \in Z_\delta \text{ where } Z_\delta = \{(R, F) \in D : \|A(R, F) - \Lambda\| \leq \mu + \delta\};$$

with an increasing accuracy, we get more accurate approximations to the exact (normal) (pseudo)solution:

$$(R, F)_\delta \rightarrow (R, F) \text{ when } \delta \rightarrow 0.$$

Similar to (1), one of the possible implementation of this procedure is obtaining such approximations on the basis of Tikhonov functional technique when we minimize

$$M^\alpha(R, F) = \|A(R, F) - \Lambda\|^2 + \alpha \|(R, F) - (R^0, F^0)\|^2$$

on the set D ; the regularization parameter α is chosen as a solution of the equation $\|A(R, F)_\alpha - \Lambda\| = \mu + \delta$, where $(R, F)_\alpha$ delivers a minimum to $M^\alpha(R, F)$. Obviously, when (R^0, F^0) is compatible with experimental data in itself, no further adjustment is necessary.

It is often the case that even in the absence of experimental errors an inverse problem remains incompatible. This may be caused by the roughness of a molecular model (for example, by ignoring some minor effects that cannot be properly accounted for at the selected level of model complexity). Under these circumstances, the approach described above should be reformulated in terms of (normal) pseudosolutions rather than in terms of exact (normal) solutions.

The details of mathematical properties of the above-mentioned equations, such as existence and uniqueness of the solution, depend on the properties of the operator $A(R, F)$. It is possible to ensure the desired properties for a wide range of inverse problems, including all the important problems under consideration.

The formulation given above is very general; in the implementation that follows we shall assume that R is a set of independent equilibrium geometry parameters and F is a set of harmonic force constants.

Now it is appropriate to give a brief summary of the features distinguishing the given approach from the various previously used attempts to solve a similar inverse problem [12].

Our approach is aimed at the simultaneous determination of the geometry and force field parameters of a molecule. It combines the techniques previously used in the IR spectroscopy and the ED data analysis. In particular, it allows one to use more flexible force field models when fitting ED data, far beyond the usually employed scaling of the *ab initio* force field.

Ab initio data (or any other external data) are automatically “weighed” so as to serve as an additional source of information when data supplied by an experiment are proved to be insufficient. There is no need to supply *ab initio* data with some kind of assumed errors, etc.

Molecular geometry is defined in terms of equilibrium distances, thus allowing compatibility with spectroscopic models and *ab initio* calculations. In addition, the self-consistency of a geometric configuration is automatically maintained at all stages of the analysis.

The complexity of the molecular models used in the analysis strongly depends on the availability and quality of experimental data. Since in most cases the vibrational spectra and ED patterns reveal the vibrational motion in a molecule resulting from small deviations of the atoms from their equilibrium positions, the molecular models are generally based on the assumption of small harmonic vibrations. In some cases of solving GISP within the scaling approximation, it is necessary to include the cubic part of the force field [13]. Similarly, in order to get a set of more reliable cubic force constants, it is undoubtedly beneficial to improve the *ab initio* values empirically (for simplicity, for example, using the Pulay harmonic scale factors). It has been our experience that two schemes of cubic constant scaling are generally feasible. Let the *ab initio* quadratic force constant f_{ij}^0 defined in natural internal coordinates be scaled as follows:

$$f_{ij}(\text{scaled}) = f_{ij}^0 \beta_i^{1/2} \beta_j^{1/2},$$

where β_i and β_j are the harmonic scale factors. Then, the cubic constant scaling mode can be formulated [14] as

$$f_{ijk}(\text{scaled}) = f_{ijk}^0 \beta_i^{1/2} \beta_j^{1/2} \beta_k^{1/2}$$

or, alternatively,

$$f_{ijk}(\text{scaled}) = f_{ijk}^0 \beta_i^{1/3} \beta_j^{1/3} \beta_k^{1/3},$$

where f_{ijk}^0 are the unscaled theoretical cubic constants. Both the scaling schemes reduce the vibrational problem to the determination of a much smaller number of parameters. The examples of applying the last procedure to different molecular systems including those with large amplitude motion are given in [15, 16].

The most important step in solving an inverse vibrational problem is formulating *a priori* constraints on the solution, which are taken from quantum mechanical calculations. Plausible constraints for the force field matrix followed recommendations discussed elsewhere [3,4].

5. Applications.

5.1. CH₃SiH₃. As an example of solving an inverse scaling problem within different scaling schemes, here we consider the calculation of the scaled force field of the methylsilane molecule CH₃SiH₃ at different theoretical levels. The quantum mechanical DFT (B3LYP) calculation of optimized structures and harmonic force fields for this molecule have been done using different standard basis sets (varying from 6-31G* to AUG-cc-pVDZ). The corresponding *ab initio* calculations were also performed at the HF/6-31G* and HF/AUG-cc-pVDZ levels of the theory. Quantum mechanical calculations were carried out using the GAUSSIAN-94 program package [17].

Optimized molecular structures of CH₃SiH₃ obtained at all applied theoretical levels belong to the C_{3v} symmetry point group. The theoretical and experimental parameters [18] are compared in Table 1. We introduced a redundant system of 20 internal coordinates (all bond distances, valence bond angles, and C-C torsion). Since the definition of the force field matrix in a redundant coordinate system is not unique, the convention was introduced that the matrix off-diagonal norm be a minimum. This choice is reasonable in accordance with the commonly used force field models [5].

Table 1. Theoretical and experimental structural parameters of methylsilane, CH₃SiH₃ (bond lengths in angstroms, valence angles in degrees)

Parameter	HF 6-31G*	B3LYP 6-31G*	B3LYP 6-311++ G**	B3LYP aug-cc- pVDZ	experiment [18]
<i>R</i> (C-Si)	1.888	1.889	1.885	1.893	1.867
<i>R</i> (C-H)	1.086	1.095	1.093	1.099	1.092
<i>R</i> (Si-H)	1.478	1.490	1.487	1.496	1.484
∠Si-C-H	111.1	111.0	111.0	111.0	108.44
∠C-Si-H	110.6	110.7	110.7	110.6	108.0
∠H-C-H	107.8	107.9	107.9	107.9	110.12
∠H-Si-H	108.3	108.1	108.2	108.3	
ϕ H-C-Si-H	179.9	179.9	179.9	180.0	

Using the B3LYP/6-31G* force constant matrix as F^0 and an expanded set of vibrational frequencies for the CH₃SiH₃, CH₃SiD₃, CD₃SiH₃, CD₃SiD₃ isotopomers [19–21], the inverse scaling problem ISP-I was solved by means of the *Spectrum* software package [2, 16]. In the first model (referred to as ISP-I) we introduced eight independent scaling factors (in accordance with the molecular symmetry). In the second model (ISP-II) only the frequencies of the “light” isotopomer were used in the minimization procedure. During this minimization procedure, the initial value of the regularization parameter was gradually reduced until the criterion of its choice has been met.

The comparison of experimental and fitted frequencies for calculations of ISP-I and II is given in Table 2. In this table we also present the results of solving the direct problem for three other isotopomers with the scaling factors obtained in calculation II.

Table 2. Comparison of theoretical, experimental, and fitted frequencies when solving the inverse scaling problem for CH₃SiH₃ (B3LYP/6-31G*) with frequencies of the four isotopomers. Experimental frequencies are taken from [20, 21]. Columns I and II show discrepancies between experimental values and those calculated in the variants ISP-I and ISP-II

Sym.	Assignment	CH ₃ SiH ₃				CH ₃ SiD ₃			
		exp.	theor.	I	II	exp.	theor.	I	II
A ₁	ν(CH)	2920	3052	13	-9	2924	3052	9	-12
	ν(SiH)	2179	2235	-11	-7	1561	1834	-17	-14
	δ(CH ₃)	1264	1332	1	-6	1264	1332	0	-6
	δ(SiH ₃)	933	950	12	6	652	802	-4	-7
	ν(C-Si)	701	687	4	0	741	674	4	0
A ₂	τ	183	197	3	0	169	185	0	-3
E	ν(CH)	2976	3129	30	9	2980	3129	26	5
	ν(SiH)	2168	2239	4	7	1577	1741	-7	-5
	δ(CH ₃)	1403	1497	16	3	1412	1497	6	-5
	δ(SiH ₃)	951	961	-7	-3	670	872	6	10
	ρ(CH ₃)	861	904	15	10	821	822	4	3
	ρ(SiH ₃)	528	522	-13	-19	421	439	-9	-15

The resulting scaling factors obtained in these two calculations as well as the factors obtained by solving the ISP-II for different theoretical levels with variation of a number of scaling factors are shown in Table 3.

We have also analyzed how the values of scaling factors depend on *a priori* constraints (the set *D*). In our case, the set *D* is determined by the number of the introduced parameters (scaling factors). In different calculations this number was taken equal to 8, 4, and 1. The results of solving the inverse scaling problems for four theoretical levels are also presented in Table 3. In this table, ρ/δ is a ratio of the minimum mean-square

Table 2 (continued)

Sym.	Assignment	CD ₃ SiH ₃				CD ₃ SiD ₃			
		exp.	theor.	I	II	exp.	theor.	I	II
A ₁	$\nu(\text{CH})$	2129	2199	39	43	2133	2199	-27	-42
	$\nu(\text{SiH})$	2181	2044	-75	-90	1562	1597	-18	-16
	$\delta(\text{CH}_3)$	1004	1045	-1	-6	1001	1044	-2	-7
	$\delta(\text{SiH}_3)$	930	874	9	0	717	720	0	-4
	$\nu(\text{C-Si})$	643	630	0	-4	617	606	-5	-8
A ₂	τ	154	155	-1	-3	134	140	-2	-4
	$\nu(\text{CH})$	2231	2323	-4	-20	2234	2323	-8	-24
	$\nu(\text{SiH})$	2170	2148	1	4	1577	1623	-7	-5
	$\delta(\text{CH}_3)$	1044	1085	-19	-28	1036	1085	-12	-21
	$\delta(\text{SiH}_3)$	950	852	-8	-3	693	712	-4	-4
	$\rho(\text{CH}_3)$	773	767	-1	-9	668	687	4	3
	$\rho(\text{SiH}_3)$	457	440	-19	-21	389	376	-19	-23

discrepancy value obtained in the minimization procedure to the experimental error (assumed to be $\pm 6 \text{ cm}^{-1}$ for each frequency).

Table 3. Comparison of scaling factors (β_i) for methylsilane obtained for different number of optimized parameters (ρ/δ is a ratio of resulting mean-square discrepancy to the error level). Only experimental frequencies of CH₃SiH₃ [19, 20] (except for the results of the first column for the B3LYP/6-31G* level) were used in calculations

Coordinate	HF 6-31G*	B3LYP 6-31G*		B3LYP 6-311++G**	B3LYP AUG-cc-pVDZ
		expanded set Λ			
8 scaling factors					
C-Si	0.9405	1.0545	1.0444	1.0376	1.0633
C-H	0.8316	0.9232	0.9100	0.9244	0.9185
Si-C-H	0.7639	0.9002	0.9037	0.9429	0.9782
H-C-H	0.7581	0.8977	0.8808	0.9200	0.9658
Si-H	0.8426	0.9414	0.9443	0.9555	0.9820
C-Si-H	0.8366	1.0165	0.9806	0.9782	1.0363
H-Si-H	0.8169	0.9627	0.9731	0.9729	0.9981
τ	0.8880	0.8925	0.8654	0.9117	0.9271
ρ/δ	1.10	2.23	1.10	0.90	1.17
4 scaling factors					
C-Si	0.9329	1.0443		1.0386	1.0699
C-H	0.8166	0.9059		0.9249	0.9273
Si-H	0.8380	0.9493		0.9591	0.9877
τ	0.8860	0.8650		0.9102	0.9255
ρ/δ	3.49	1.66		1.09	2.41
One scaling factor					
All cords.	0.8253	0.9217		0.9377	0.9483
ρ/δ	4.06	3.38		2.57	3.49

These results show that the quality of the ISP solutions depends very strongly on the set D . The strong narrowing of the set of minimized parameters leads to the poor reproduction of experimental data. The ill-posed character of the inverse vibrational problem, as a rule, cannot be avoided by increasing the set of experimental data because this could simultaneously lead to the incompatibility of input data.

5.2. C₂F₆. The molecule of perfluoroethane C₂F₆ belongs to the D_{3d} symmetry group. Quantum-mechanical calculations have been performed at the levels HF/6-31G*, MP2/6-31G*, and B3LYP/6-31G*. The optimized molecular geometries for these three cases are compared in Table 4. Here again a redundant system of 20 internal coordinates was introduced.

Table 4. Theoretical and experimental structural parameters of perfluoroethane, C₂F₆ (bond lengths in angstroms, valence angles in degrees)

Parameter	HF 6-31G*	B3LYP 6-31G*	MP2 6-31G*	Experiment[22]
R (C-C)	1.526	1.545	1.530	1.545
R (C-F)	1.311	1.338	1.340	1.326
∠C-C-F	109.8	109.9	109.7	109.8
∠F-C-F	109.1	109.1	109.2	

Due to symmetry considerations, five independent scale factors have been introduced for twenty internal coordinates. In table 5 we present the vibrational frequencies obtained in three different quantum-mechanical calculations. Initial values as well as those obtained after scaling are included.

Table 5. Experimental, theoretical and scaled values of vibrational frequencies for different levels of *ab initio* calculations. Error corresponds to the mean-square error (in cm⁻¹)

Sym.	Assign.	Exper. [23]	HF/6-31G*		MP2/6-31G*		B3LYP/6-31G*	
			Initial	Scaled	Initial	Scaled	Initial	Scaled
A _{1g}	ν(CC)	1417	1627	1426	1501	1428	1426	1421
	ν(CF)	807	887	781	813	783	803	797
	δ(CF ₃)	348	377	335	354	340	344	345
A _{1u}	τ	68	70	68	65	68	63	67
A _{2u}	ν(CF)	1116	1242	1110	1150	1124	1130	1127
	δ(CF ₃)	714	773	702	706	697	700	703
E _g	ν(CF)	1250	1426	1268	1292	1260	1261	1256
	δ(CF ₃)	619	673	614	621	615	611	615
	ρ(CF ₃)	372	415	383	378	378	371	376
E _u	ν(CF)	1251	1433	1263	1305	1265	1269	1257
	δ(CF ₃)	522	564	523	516	517	512	518
	ρ(CF ₃)	219	230	210	213	212	206	210
Error			116.1	12.3	34.0	11.3	10.5	6.9

Note that the B3LYP calculation has resulted in close correspondence of top experimental frequencies prior to any scaling, while the HF level of the theory gives systematically overestimated values. The scale factors for all the three variants of calculation are given in Table 6.

Table 6. Scale factors for perfluoroethane

	HF/6-31G*	MP2/6-31G*	B3LYP/6-31G*
R(C-C)	0.7130	0.8330	0.9793
R(C-F)	0.7675	0.9327	0.9766
∠CCF	0.8375	0.9850	1.0309
∠FCF	0.8706	1.0135	1.0266
τ	0.9459	1.0802	1.1594

6. CONCLUSIONS. The above discussion allows us to come to the following conclusions.

1. The strict mathematical formulation of the inverse scaling problem could provide a possibility of comparison of results obtained by different investigators. Obviously, it is necessary to work within the same physical and mathematical models and use the stable numerical methods for the unification of different calculations. Otherwise, the situation may be the same as in the empirical force field calculations when for the same molecule it is possible to find a lot of very different sets of force constants.

2. It is important that using the scaling scheme, as a rule, it is impossible to obtain a solution that reproduces experimental data within a given error level. We may search only for the *pseudosolution* of Problem I that satisfies the input frequencies (or the expanded set of experimental data) in the least-squares sense.

7. Acknowledgements. This work was partially supported by the RFBR-YUGRA (grant 03-07-96842) and by the RFBR (grant 02-01-00044).

СПИСОК ЛИТЕРАТУРЫ

1. A.N. Tikhonov, A.S. Leonov, A.G. Yagola. Nonlinear ill-posed problems. London: Chapman&Hall, 1998 (Original Russian language edition: Nonlinear ill-posed problems. Moscow: Nauka, 1993).
2. A.G. Yagola, I.V. Kochikov, G.M. Kuramshina, Yu.A. Pentin. Inverse problems of vibrational spectroscopy. Zeist (The Netherlands): VSP, 1999.
3. G.M. Kuramshina, A.G. Yagola. A priori constraints in the force field calculations of polyatomic molecules // J. Struct. Chem. 1997. **38**. 181–194.
4. G.M. Kuramshina, F. Weinhold. Constraints on the values of force constants for molecular force field models based on ab initio calculations // J. Mol. Struct. 1997. **410**, N 2. 457–462.
5. G.M. Kuramshina, F.A. Weinhold, I.V. Kochikov, Yu.A. Pentin, A.G. Yagola. Joint treatment of ab initio and experimental data in molecular force field calculations with Tikhonov's method of regularization // J. Chem. Phys. 1994. **100**, N 2. 1414–1424.
6. P. Pulay, G. Fogarasi, G. Pongor, J.E. Boggs, A. Vargha. Combination of theoretical ab initio and experimental information to obtain reliable harmonic force constants. Scaled quantum mechanical (SQM) force fields for glyoxal, acrolein, butadiene, formaldehyde, and ethylene // J. Am. Chem. Soc. 1983. **105**. 7037–7047.
7. T.I. Seidman, C.R. Vogel. Well-posedness and convergence of some regularization methods for nonlinear ill-posed problems // Inverse Problems. 1989. **5**. 227–238.
8. H.W. Engl, M. Hanke, A. Neubauer. Regularization of inverse problems. Dordrecht: Kluwer Academic Publishers, 1996.
9. A.V. Stepanova, I.V. Kochikov, G.M. Kuramshina, A.G. Yagola. Regularizing scale factor method for molecular force field calculations // Computer Assistance for Chemical Research. International Symposium CACR-96. Moscow, 1996. 52.
10. I.V. Kochikov, G.M. Kuramshina, A.V. Stepanova, A.G. Yagola. Regularizing scale factor method for molecular force field calculations // Moscow Univ. Bull. Ser. 3. Physics and Astronomy. 1997. **5**. 21–25 (in Russian).
11. I.V. Kochikov, Yu.I. Tarasov, V.P. Spiridonov, G.M. Kuramshina, A.G. Yagola, A.S. Saakyan, M.V. Popik, S. Samdal. Extension of a regularizing algorithm for the determination of equilibrium geometry and force field of free molecules from joint use of electron diffraction, molecular spectroscopy and ab initio data on systems with large-amplitude oscillatory motion // J. Mol. Struct. 1999. **485-486**. 421–443.
12. I.V. Kochikov. Inverse problems of molecular structure and force field calculation on the basis of vibrational spectroscopy and electron diffraction experimental data // ISIP2001. International Symposium on Inverse Problems in Engineering Mechanics. 6-9 February 2001. Nagano City (Japan), 2001. 121–124.
13. S. Kondo. Empirical improvement of the anharmonic ab initio force field of methyl fluoride // J. Chem. Phys. 1984. **81**, N 12. 5945–5951.
14. I.V. Kochikov, Yu.I. Tarasov, V.P. Spiridonov, G.M. Kuramshina, A.S. Saakyan, Yu.A. Pentin. The use of ab initio calculation results in electron diffraction studies of the equilibrium structure of molecules // Russian Journal of Physical Chemistry. 2001. **75**, N 3. 395–400.
15. I.V. Kochikov, Y.I. Tarasov, V.P. Spiridonov, G.M. Kuramshina, A.S. Saakyan, A.G. Yagola. The use of ab initio anharmonic force fields in experimental studies of equilibrium molecular geometry // J. Mol. Struct. 2000. **550-551**. 429–438.
16. I.V. Kochikov, Yu.I. Tarasov, V.P. Spiridonov, G.M. Kuramshina, et al. The equilibrium structure of thiophene by the combined use of electron diffraction, vibrational spectroscopy and microwave spectroscopy guided by theoretical calculations // J. Mol. Struct. 2001. **567-568**. 29–40.
17. M.J. Frisch, G.W. Trucks, H.B. Schlegel, P.M.W. Gill, B.G. Johnson, M.A. Robb, J.R. Cheeseman, T. Keith, G.A. Peterson, J.A. Montgomery, K. Raghavachari, M.A. Al-Laham, V.G. Zakrzewski, J.V. Ortiz, J.B. Foresman, J. Cioslowski, B.B. Stefanov, A. Nanayakhara, M. Chablacombe, C.Y. Peng, P.Y. Ayala, W. Chen, M.W. Wong, J.L. Andres, E.S. Replogie, R. Gomperts, R.L. Martin, D.J. Fox, J.S. Binkley, D.J. Deflees, J. Baker, J.P. Stewart, M. Head-Gordon, C. Gonzalez, J.A. Pople. Gaussian 94, Revision C2. Gaussian Inc. Pittsburgh, 1995.
18. R.W. Kilb, L. Pieree. Microwave spectrum, structure and internal barrier of methylsilane // J. Chem. Phys. 1957. **27**. 108–112.
19. R.E. Wilde. The infrared spectrum of CH₃SiH₃ and CH₃SiD₃ // J. Mol. Spectr. 1962. **8**, N 6. 427–454.
20. A.J.F. Clark, J.E. Drake. The vibrational spectra of germane and silane derivatives. Vibrational spectra and normal co-ordinate analysis of the methylsilanes // Can. J. Chem. 1977. **22**, N 4. 79–83.
21. S.V. Syn'ko, G.M. Kuramshina, A.I. L'vov, Yu.A. Pentin, G.S. Gol'din. Vibrational spectra of CH₃SiH₂D and CH₃SiHD₂ // Moscow Univ. Bull. Ser. 2. Chemistry. 1983. **24**, N 4. 360–364 (in Russian).
22. K.L. Gallaher, A. Yokozeki, S.H. Bauer. Reinvestigation of the structure of perfluoroethane by electron diffraction // J. Phys. Chem. 1974. **78**, N 23. 2389–2395.
23. I.M. Mills, W.B. Person, J.R. Scherer, B. Crawford, Jr. Vibrational intensities. IX. C₂F₆: extension and revision // J. Chem. Phys. 1958. **28**, N 5. 851–853.

Поступила в редакцию
08.10.2004