EURASIAN JOURNAL OF MATHEMATICAL
AND COMPUTER APPLICATIONS
ISSN 2306–6172
Volume 4 , Number4 (2016) 14–36

REGULARIZING ALGORITHMS FOR SOLVING NONLINEAR ILL-POSED PROBLEMS OF VIBRATIONAL SPECTROSCOPY

Kuramshina G.M., Yagola A.G.

Abstract Inverse problems of molecular force field calculation arising as a result of data processing in vibrational spectroscopy belong to the class of nonlinear ill-posed problems. In this paper we discuss the main mathematical results obtained within the theory of regularization for solving these problems. Different algorithms on the basis of regularizing methods were proposed for solving nonlinear ill-posed problems of molecular force field calculations and implemented in the software package SPECTRUM.

Key words: inverse problem, nonlinear ill-posed problem, theory of regularization, vibrational spectra, molecular force field.

AMS Mathematics Subject Classification: 49N45, 65F18, 92E10.

1 Introduction

We consider the inverse problems of vibrational spectroscopy related to calculations of the molecular force field parameters on the basis of experimental data (obtained from analysis of the vibrational, infrared and Raman, spectra). These problems lead to an operator equation of the standard form

$$A\ z = u \tag{1}$$

where $z \in Z$ is the set of properties of the object to be determined, $u \in U$ is the set of experimental data, Z and U are certain spaces (we assume that they are normed), and A is an operator, determined by the choice of the mathematical model of the phenomenon. Problems of the form of (1) are called *inverse problems* and are often ill-posed. A well-posed problem, as defined by Hadamard [1] should have the next properties:

1) equation (1) is solvable for any $u \in U$;

2) the solution is unique;

3) the solution is stable, i.e., small perturbations of u result in small perturbations of z.

If any of the conditions 1 - 3 is not satisfied, the problem is called *ill-posed*. In this case the classic methods are not suitable for solving (1) because errors of the experimental data (e.g. example, small changes in the data) may yield too large or arbitrary large perturbations in the solution. The goal of this review is to describe the stable methods of solving inverse problems in vibrational spectroscopy developed in our group. From the viewpoint of analyzing experimental data (obtained by means of infrared and Raman spectroscopy) a number of inverse problems arise. One of the most important is the so-called inverse vibrational problem of determining parameters of the molecular force field from given experimental data (vibrational frequencies, isotope frequency shifts, Coriolis constants, centrifugal distortion constants, etc.). The accumulation of data on molecular constants helps us to predict spectra and other properties of compounds not yet investigated, especially, in the last years for bulky molecular systems including biological molecules and assists the development of advances physical models in a theory of molecular structure. The molecular force field (the second derivatives of potential energy with respect to the nucleus coordinates) plays the key role in determining the properties of a molecule, in particular, its vibrational spectrum. Many approximations have been proposed for the calculation of a complete quadratic force field and have realized as computational programs used in practice [2-9]. Usually, the ill-posed character of inverse vibrational problems has led to some degree of subjectivity, related basically to constraints imposed on the solution to ensure physically meaningful results. In this way various models of the molecular force fields have been proposed and a great number of force field matrices have been calculated for various series of compounds. All these matrices were chosen to satisfy either experimental data or a priori known, but not explicitly formulated, criteria of the physically meaningful solutions. As a result, a situation existed (in particular, for complicated polyatomic molecules) when various so-called spontaneous methods for solving inverse problems lead to inconsistent force fields due to different criteria for the physical feasibility of solutions used by various investigators, and to the instability (with respect to small perturbations in the input information) of numerical methods used to solve the inverse problem. The investigations of the correctness of inverse problems of vibrational spectroscopy have been started in Lomonosov Moscow State University in early of 1980s [10]. In our publications [11 - 25] we have made an attempt to formulate and formalize all possible obvious (and not so obvious) model assumptions concerning the character of force fields which are widely used in vibrational spectroscopy and can be applied in empirical calculations of molecular force fields. On the basis of this formalization in the framework of Tikhonov's regularization theory we have constructed a principle for choosing a unique solution from the set of solutions. We have formulated such a principle in terms of closeness of the solution to the given matrix of force constants which satisfies all *a priori* assumptions concerning the model characteristics of the solution. A further development of investigations have been stimulated by the tremendous development of quantum mechanical calculations and their applications in a practical chemistry accompanying by an appearance of new very effective approaches of computational chemistry [26-33]. It was very natural approach to include the results of quantum mechanical calculations in our algorithms of solving inverse vibrational problems [34] and related fields of structural chemistry and to join the modern stable numerical methods with modern approaches of theoretical chemistry. In this paper we present our results on construction of stable numerical methods based on Tikhonov's regularization method for computing the force fields of polyatomic molecules on a base of experimental data and results of quantum chemical calculations and provide some numerical illustrations with real data.

2 Experimental sources of information and molecular force fields

Vibrational spectrum of a molecule appears is a result of change of the vibrational component of molecular total energy. This change arises due to interaction of infrared electromagnetic radiation with the substance, and in general vibrational spectrum is a complicated function of molecular geometry, nuclei masses and electron density distributions. Vibrational spectrum is a unique physical property, very important source of knowledge on molecular structure and can be used as an important tool for molecular identification. The positions and intensities of infrared (IR) bands and Raman lines are considered as indicators of the presence of specific functional groups in the molecule [6]. Important fields of vibrational spectroscopy applications are related to the identification of reactive intermediates, species in interstellar space, etc. Two approaches, classical and quantum mechanics, can be used for the consideration of vibrational spectra. In principle, one can describe a molecule as a quantum system but in certain cases consider the molecular properties within the classical model. To solve vibrational problem one can use approximations of classical mechanics. But e.g. the intensity of vibrational bands, appearance of overtones or combined bands can be appropriately explained only by quantum mechanics. The most important fact is that within the harmonic approximation both classical and quantum mechanical approaches result in the same equation of molecular vibrations. Theoretical vibrational frequencies are calculated with a help of so-called force constants (second derivatives of the molecular potential energy in respect to atomic displacements), either by direct diagonalization of the (mass-weighted) Cartesian force constant matrix or force constant matrix in some generalized coordinate system. The idea of the force field arises from the attempt to consider a molecule as a mechanical system of nuclei while all the interactions due to the electrons are included in an effective potential function $U(q_1, q_2, ..., q_n)$ where $\{q_1, q_2, ..., q_n\}$ denote $n = 3N$-6 generalized coordinates describing nuclei positions of N atomic molecule. Together with the nuclear masses, this function determines the most important properties of a molecule. It is important that the equilibrium configuration of the molecule should satisfy to the relation

$$\frac{\partial U}{\partial q} = 0,$$

and if we define coordinates as $(q_1 = q_2 = ... = q_n = 0)$ in the equilibrium configuration, the potential function can be performed as the series:

$$U\left(q_1, ..., q_n\right) = U_0 + \frac{1}{2}\sum_{i,j=1}^{n} f_{ij} q_i q_j + \mathrm{O}\left(\|q\|^3\right) \tag{2}$$

where U_0 is a certain constant, and parameters *fij* (the so-called force constants) $f_{ij} = \frac{\partial^2 U}{\partial q_i \partial q_j}$, *i, j = 1, ...,* n in the point of equilibrium constitute a positive definite matrix F which determines all molecular characteristics related to small vibrations. Mathematically, the concept of the force field can be obtained through the adiabatic theory of perturbations with the use of a small parameter related to the ratio of electron mass to the mass of nuclei. It can be shown that in a certain approximation the nuclei may be treated as particles moving in the force field determined by the potential energy function (2).

The main purposes of using a molecular force field can be determined as: (*a*) checking validity of various model assumptions commonly used by spectroscopists for approximation of the potential function; (*b*) predicting the vibrational properties of certain molecules (e.g. not yet observed) on a base of fundamental properties of the force field such as its isotopic invariance and the approximate transferability of force constants in a series of related compounds.

The measured infrared and Raman spectral frequencies ω_i are the main type of experimental data on molecular vibrations. They are connected with the matrix of force constants F by the eigenvalue equation

$$GFL = L\Lambda, \tag{3}$$

where $\Lambda = diag\ \{\omega_1^2,\ ...,\ \omega_n^2\}$ is a diagonal matrix consisting of the squares of the molecular normal vibrational frequencies, G is the kinetic energy matrix in the momentum representation which depends only on nuclear masses and the equilibrium configuration (assumed to be known within specified limits of error). The eigenvectors matrix L characterizes the relative amplitudes of vibrations in terms of classical mechanics (the form of each normal vibration).

While (3) 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 determined from n frequencies ω_1, ..., ω_n. We need to use certain model assumptions concerning the structure of the matrix F and additional experimental data, e.g. frequencies of molecular isotopomers. Within the approximations considered, the force field of a molecule is independent of the nuclear masses, and hence in a case of m molecular isotopic species, instead of (3), we have the system

$$(G_iF)L_i = L_i\Lambda_i \tag{4}$$

where $i = 1,2, ..., m$ is a number of molecule isotopomers. Usually, introduction of the additional data of isotopomers leads to a limited number of independent equations in system (4) thus leaving the inverse problem underdetermined. Important additional information on the molecular force field is provided by Coriolis constants ζ which characterize the vibrational-rotational interaction in the molecule possessed high symmetry. They are connected with matrix F in terms of the eigenvectors L of the problem (3):

$$\zeta = \frac{1}{M^2} L^* AMA^* L \tag{5}$$

where ζ is a matrix with vector elements consisting of the Coriolis constants, M is a diagonal matrix consisting of the nuclear masses, M is the sum of nuclear masses of the molecule, and **A** is a matrix with transpose **A*** connecting the Cartesian displacements of atoms with coordinates q, which can be found from the equilibrium configuration of the molecule. In a similar manner we can write the dependencies of other measured values on the matrix F, such as the mean-square amplitudes of the vibrations (determined from gas-phase electron diffraction) which may be calculated from the eigenvalues and eigenvectors of (3). Those additional experimental quantities (frequencies of isotopomers, Coriolis constants, mean-square amplitudes) while not so easily obtainable, are particularly important because these data do not depend on additional unknown parameters of any kind.

3 Quantum mechanical calculations of force fields

Methods of theoretical chemistry based on the Schrцdinger's equation allow us, at least in principle, to calculate all properties of a molecular system without using experimental data. Such methods are called *ab initio* (from the beginning or from first principles), they are based entirely on quantum mechanics and fundamental physical constants. The information necessary for such calculations includes initial geometry of a molecule (positions for the nuclei) and atomic masses, the electronic wave functions are described by a set of mathematical functions. Regretfully, to make such computations technically possible one needs to use very restrictive assumptions. These restrictions as well as an increasing number of atoms in a molecule (and, correspondingly, of the number of vibrational degrees of freedom for nuclear motion) have very strong effect on the accuracy of results. In relatively few cases it is feasible to carry out *ab initio* calculations which can predict frequencies within the limits of experimental error (say, e.g., within the 1% errors of the usual harmonic approximation) [9, 32]. Nevertheless, we expect further rapid progress of *ab initio* calculations. Very impressive results obtained in the last decades allow us to consider new ways of applications of *ab initio* calculations and we shall consider below some new ways of utilizing even the no-so-accurate results of quantum mechanical calculations that are routinely available at the present.

The development of *ab initio* force field calculations in the past two decades has been extremely rapid. The approaches in the modern theoretical chemistry are depended on the goal of investigation and the size of considered molecular system. Within the very popular Hartree-Fock (HF) self-consistent-field framework (representing a moderate level of nonrelativistic theory), calculations at this level of theory are routinely possible for very large molecules (containing hundreds and thousands atoms and more) at moderate basis set levels. But this method does not take into account an electron correlation, so the problem of obtaining improved theoretical results (with errors comparable to those of experiment) remains a major challenge for theoretical chemists. The accuracy of HF calculations is rather far from experimental limits, and these calculations often give rather appreciable errors in geometry (˜0.05 in bond lengths, ˜5◦ in bond angles) and in molecular force constants (˜10% error) determined by the shifts of calculated frequencies in a comparison with experimental frequencies.

Rapid progress in the extending force constant calculations to higher levels of theory has been associated with the introduction of analytic derivative methods (initially, for gradients, later for second order and higher derivatives [29] to replace the older numerical methods and with appearance of the so-called post-HF methods such as configuration interaction (CI), multi configuration self-consistent field (MCSF) and Mшller Plesset perturbation theory methods (MPn). Limitations are related to computational cost and rather strong restrictions on the size of a molecule. New possibilities for calculations of various molecular systems were opened with the development of approaches of Density Functional Theory (DFT) [30, 31]. Rapid progress in quantum mechanical calculations of harmonic force field calculations for moderate-size molecules with inclusion of electron correlation at the second order Mшller Plesset perturbation theory (MP2) and density-functional theory DFT levels [28-31] implemented in modern software packages [33] provides fundamental new possibilities for more accurate interpretation of experimental data as well as improved methods for empirical force field calculations.

Calculations of moderate size molecules performed at high theoretical level (taking into account electron correlation, etc.) can satisfactorily reproduce patterns of experimental data on molecular structure and vibrations. However, in a case of large molecular systems the best *ab initio* results significantly differ from observed data. The problem is that experimentally vibrational frequencies can be easily determined with wavenumber accuracy, which for the so-called fingerprint spectral region (specific for the given compound) implies an accuracy of 1:1000 - unattainable for modern quantum mechanical methods. This situation results in using special empirical corrections of theoretical results with a help of so-called scale factors [27] which lead *ab initio* results to the level of semi-empirical ones and cause the loss of important physical information. Despite significant progress in computational power as well as an existence of very effective quantum mechanical approaches, the accurate calculation of vibrational frequencies and intensities is not routine procedure for the most of molecular systems and is considered as a task for the expert. Obviously, the approach guaranteeing that solutions are obtained which reflect the most important properties of quantum mechanical potential surfaces is therefore preferable. The goal of our study was to use the modern stable numerical methods and to create an approach for data processing in vibrational spectroscopy and calculation of molecular force field which includes the results of quantum mechanical calculations and can guarantee solutions reflecting the most important properties of *ab initio* potential surfaces.

4 The mathematical formulation of the inverse vibrational problem

We consider equations (3) - (5), or some part of them, depending on the available experimental data, as a single operator equation

$$AF = \Lambda \tag{6}$$

(6) where the nonlinear operator A maps the real symmetrical matrix F to the set of eigenvalues of (3) (or (4)), the Coriolis constants, ζ (5), the mean square amplitudes, etc. This set of data may be represented as a vector in the finite-dimensional space Rl, where l is a number of known experimental data. The matrix F is also considered as a vector in the finite-dimensional space Z, consisting either of the elements of the matrix F or of the quantities by means of which this matrix can be parametrized. Note that any of the conditions of the well-posedness are not satisfied for operator equation (6):

1) Solvability.

It is easy to see that within harmonic model the system (4) (determined for different molecular isotopomers) is compatible only when the condition ${}^{DetG_i}/_{Det\Lambda_i} = const, i = 1, 2, m$ (m is the number of isotopomers) is satisfied.

This may be violated both by virtue of errors in measurements of Λi, and due to inaccurate specification of the molecular geometry Gi, or when exact experimental data are not available, because of the anharmonicity of the vibrations ignored by the operator of (6). Therefore, a matrix F reproducing the frequencies of all isotopomers within the limits of experimental errors does not exist.

2) Uniqueness of the solution of the problem.

If we know only the vibrational frequencies of one isotopomer of the molecule, equation (6) reduces to the inverse eigenvalue problem (2); and when G is not singular it follows that as solution of (6) we may take any matrix of the form

$$F = G^{-\frac{1}{2}} C^* \Lambda C G^{-\frac{1}{2}} \tag{7}$$

where C is an arbitrary orthogonal matrix. To choose a definite solution it is necessary to use additional information or to take into account some model assumptions.

3) Stability of the solution with respect to the perturbations of Λ and A.

For a system of the form (4), instability may be easily exemplified.

Therefore, the inverse vibrational problem is an example of an ill-posed problem since the three conditions of well-posedness are not satisfied in general. The difficulties which arise during the force field calculations (which are related to the non-uniqueness of solutions and the instability of the solution with respect to the perturbations of experimental data) have been discussed elsewhere and various special model assumptions have been proposed based on the intuition of the investigator. The absence of stability and uniqueness of the solution has often led to significant differences of force field parameters of the same molecules determined in different investigations. This has caused difficulties in the comparison and transferability of the force fields in series of related compounds and has made the physical interpretation of the experimental data much more obscure.

In our opinion, it is very important in the elucidation of any arbitrariness in the calculated force constants to use stable solutions of the inverse vibrational problem which have some specific properties. What is meant by “specific properties”? In practice chemists and spectroscopists often use certain model assumptions arising from the classical theory of chemical structure, involving the monotonic changing of the physico-chemical properties in a series of related molecules and the preservation of the properties of separate molecular fragments in the different compounds taking into account the nearest surrounding. It is the transferability of the force constants which related to these properties.

Now we consider possible ways to formalize these model considerations in order to use them in the force constant calculation. For example, model assumptions may be taken into account a priori by special choice of a given matrix F^0. In the framework of the so-called *valence force field* some off-diagonal elements of the matrix F may be taken to be equal to zero, reflecting the assumption on insignificance of some intramolecular interactions. For related molecules we can introduce some in-pair equalities of force constants for similar fragments, some elements of F may be known from preliminary calculations, etc. Therefore, we can formulate the inverse vibrational problem in the following way [10, 25].

We need to find, among the set of force constant matrices satisfying (6), the matrix which will be the nearest to some a priori given matrix F^0 (i.e., to find the so-called normal solution). In a case of an inconsistent problem (within harmonic approximation, this may happen in a joint treatment of the spectra of isotopomers, or by including additional experimental data), it is possible to find the matrix F for which the distance from AF to Λ is minimal, i.e., to find the so-called *pseudosolution* of the problem. When the pseudosolution is also non-unique, we must proceed as in the case of non-uniqueness of the solution - i.e., we choose from all possible pseudosolutions the one which is nearest to the given force field F^0 (the normal pseudosolution).

We now arrive at the following formulation of the inverse problem taking into account the principle formulated above for choosing the solution. Suppose we are given operator equation (6) and the operator A. which maps the symmetric, and positive-definite matrix F to the set of molecular vibrational frequencies squared (possibly including data for isotopomers of the molecule), and the mean vibrational amplitudes, Coriolis constants, etc., which may be taken from experiments. The dimension of the vector Λ is determined by the number of experimental data. Since the symmetric matrix F is determined by *n(n+1)/2* elements we can consider the unknown force constants as a vector of dimension *n(n+1)/2*. Then the operator A acts from the Euclidean space $R^{n(n+1)/2}$ into the Euclidean space R^l. In these spaces we can introduce the following norms:

$$\|F\| = \left(\sum_{ij}^{n} f_{ij}^2\right)^{\frac{1}{2}} ; \|\Lambda\| = \left(\sum_{k=1}^{l} \lambda_k^2 p_k\right)^{\frac{1}{2}} ,$$

where $\rho_k > 0$ are the positive weights; f_ij are the elements of matrix F; λ_k (k= 1,...l) are the components of Λ.

The operator A is continuous for all problems considered. However, equation (6) may have non-unique solution, or no solution at all, due to the anharmonicity of experimental frequencies. Suppose, we are given the matrix F^0 (vector of dimension *n(n+1)/2)*. It is necessary to find the normal pseudosolution of (6): that is, we find an element $F = \overline{F_n}$ for which min $||F - F^0||$is reached provided that $||AF - \Lambda|| = \mu$, where $\mu = \inf_{F \in D} \|AF - \Lambda\|$ and D is the closed set of a priori constraints on the values of the force constants. If no constraints are imposed, then $D = R^{n(n+1)/2}$.

The element F^0 should be specified from a priori considerations of the possible solutions, including both the approximate quantum mechanical calculations and other ideas (for example, the transferability of the force constants among similar fragments in a series of related compounds).

Let us denote the solution (vector) to be found as $\overline{F_n}$, if this vector is non-unique then a set of such vectors will be denoted as F_n . It is easy to see that if (6) is solvable then $\mu = 0$, and it remains to find the solution of (6) which is the nearest to the given vector F^0. But we do not know the exact form of either the vector Λ or operator A (the matrix G or matrices Gi). We only have the vector Λ_δ, determined from experimental data such that $||\Lambda_\delta - \Lambda|| \leq \delta$ (where $\delta > 0$ is the experimental error) and the operator A_h approximates the operator A; $h \geq 0$ is a parameter characterizing the proximity of Ah to A. The uncertainty in operator A is related to errors of the determination of matrix G (or Gi) calculated using experimental data on the geometrical parameters of the equilibrium configuration quoted with errors. Therefore, the problem of the solution of (6) arises when we do not know exact forms of A and Λ, but only their approximations A_h and Λ_δ and their errors (h, δ) and it is necessary to find the vector

$F_{h,\delta}$ approximating the exact solution $-Fn$. The difficulties in solving this problem are related to its ill-posed character.

When designing stable methods for solving ill-posed problems there is the understanding that ill-posed problems are generally underdetermined. To solve such problems one has to use additional information and formulate the criteria for choosing approximate solutions. If such criteria are formulated and mathematically formalized, it is possible to construct stable methods of solving ill-posed problems - so-called *regularizing algorithms* [35-43].

5 Nonlinear ill-posed problems

In this section we consider ill-posed problems in (for simplicity) the finite-dimensional case. And then we give the formal statement of inverse vibrational problem as nonlinear ill-posed problem. At first consider the next example of operator equation

$$Az = u, z \in D \subseteq Z, u \in U \tag{8}$$

where D is a set of constraints, Z and U are metric spaces. Let $D = Z = U = R$, where R is the space of real numbers with the norm $|| \; z \; || = |z \; |$. Fix positive numbers α_0, z_0, u_0 , and define the continuous nonlinear operator A: $Z \rightarrow U$:

$$Az = \begin{cases} \left[a_0 \left(z_0^2 - z^2 + z \left(z_0 - z\right)\right)\right]^{1/2} + u_0, & \text{if} \quad 0 \leq z \leq z_0; \\ \left[a_0 \left(z_0^2 - z^2\right)\right]^{1/2} + u_0, & \text{if} \quad z > z_0; \\ a_0^{1/2} z_0 + z^2 + u_0, & \text{if} \quad z < z_0. \end{cases} \tag{9}$$

Let's consider (8) with the exact data (A, u). It has the unique solution . We introduce Tikhonov's functional

$$M^a \left[z\right] = \left|Az - u\right|^2 + a \left\|z\right\|^2, z \in D \subseteq Z, a > 0, \tag{10}$$

and consider the variational problem: for fixed α>0, find an element such that

$$M^a \left[z^a\right] = \inf\{M^a \left[z\right] : z \in D\} \tag{11}$$

Then problem (11) with data (A, u_0) from (9) has the solution z^α =\{z_0 for $0 < \alpha < \alpha_0$; $[0, z_0]$ for $\alpha = \alpha_0$; 0 for $\alpha > \alpha_0$

Hence, the discrepancy

$$\beta\left(\alpha\right) = \left\|Az^\alpha - u_0\right\| = \begin{cases} 0 & \text{for } 0 < \alpha < \alpha_0; \\ u_0 & \text{for } \alpha > \alpha_0. \end{cases}$$

And the equation

$$\beta(\alpha) = \delta \tag{12}$$

for choosing the regularization parameter α by the discrepancy principle [40] has no solution in the ordinary sense for any δ, satisfying $0 < \delta < u_0$.

This example shows that the discrepancy principle generally gives no possibility of choosing the value of the regularization parameter for nonlinear problems. Similar difficulties arise in other strategies involving an a posteriori choice of the regularization parameter. These difficulties are connected with the non-uniqueness of the solution of (11) for the nonlinear case. This non-uniqueness leads to another difficulty. Let us assume that the regularization parameter $\alpha(\delta) > 0$ is chosen and problem (11) has more than one minimizer for $\alpha = \alpha(\delta)$. The question is whether it is possible to take an arbitrary element $z^{\alpha(\delta)}$ as an approximate solution or if a special choice of such an element is needed. Our example shows that if we choose $\alpha(\delta)$ as a solution of (12) in a general sense, i.e., as a "jump" point of a monotone function $\beta(\alpha)$ over δ, then the "unsuccessful" choice of $z^{\alpha(\delta)} : z^{\alpha(\delta)} = 0$, from the set of the extremals corresponding to this $\alpha(\delta)$ can result in the approximate solutions $z^{\alpha(\delta)}$ not converging to the exact solution $\overline{z} = z_0$. Thus, the algorithm of Tikhonov's regularization method of solving nonlinear problem (8) should in in general contain, along with the rule for choosing the parameter $\alpha(\delta)$, also a rule for selecting $z^{\alpha(\delta)}$ from the set of solutions of (11) for $\alpha = \alpha(\delta)$

Let us give the general formulation of Tikhonov's scheme for constructing a regularizing algorithm to solve the main problem: for the operator (8) on D find $z^* \in D$ for which

$$\rho(Az^*, u) = \inf\{\rho(Az, u) : z \in D\} \equiv \mu \tag{13}$$

(ρ is the distance in the metric space U). We call such z^* a *quasi-solution* for problem (8). In the case of $D = Z$, problem (13) gives a pseudo-solution of (8). If the measure of incompatibility is equal to zero, then the solutions of (13) are the solutions of (8) on D. The quasi-solution problem (13) may be ill-posed. Namely, problem (13) may not be solvable, and if a solution of (13) exists, it may be non-unique or unstable (with respect to perturbations of the data $(A,\ u)$) in the metric of Z.

We assume that to some element$u = \overline{u}$ there corresponds a nonempty set $Z^* \subseteq D$ of quasi-solutions and that Z^* may consist of more than one element. Furthermore, we suppose that a functional $\Omega(z)$ is defined on D and bounded below:

$$\Omega(z) \geq \Omega^* \equiv \inf\{\Omega(z) : z \in D\} \geq O.$$

The Ω-optimal quasisolution problem for (8) is as follows: find a such that

$$\Omega(z) = \inf\{\Omega(z) : z \in Z^*\} \equiv \bar{\Omega}. \tag{14}$$

We denote the set of Ω-optimal quasisolutions of (8) by $\overline{Z}$. If $D = Z$, then $\overline{Z}$ is the set of the Ω-optimal pseudosolutions of (8). For simplicity we use only the term "pseudo-solution" below.

We suppose that, instead of the unknown exact data$(A, \bar{u})$, we are given approximate data (A_h, u_δ) which satisfy the following conditions:

$$u_\delta \in U,\, \rho\left(\bar{u}, u_\delta\right) \leq \delta, \quad A_h \in \mathbf{A}, \rho\left(Az, A_h z\right) \leq \psi\left(h, \Omega\left(z\right)\right) \forall z \in D.$$

Here the function ψ represents the known measure of approximation of the precise operator A by the approximate operatorA_h, and $A_h \in \mathbf{A}$, a class of admissible operators. We are given also numerical characterizations $h,\ \delta \geq 0$ of the "closeness" of $(Ah,\ u\delta)$ to $(A,\ u)$. The main problem is to construct from the approximate data

$(A_h, u_\delta, h, \delta)$ in (8) an element $z_\eta = z_\eta(A_h, u_\delta, \psi, h, \delta) \in D$ which converges to a member of the set $\bar{Z}$ of Ω-optimal solutions as $\eta = (h, \delta) \to 0$.

Let us formulate our basic assumptions:

1) The class **A** consists of the operators A continuous from D to U.

2) The functional $\Omega(z)$ is lower semicontinuous (e.g., continuous) on D:

$$\forall z_0 \in D, \forall \{z_n\} \in D : z_n \to z_0 \Rightarrow \lim_{n\to\infty} \inf \Omega(z_n) \geq \Omega(z_0).$$

3) If K is an arbitrary number such that $K \geq \Omega^*$, then the set $\Omega_K = \{z \in D : \Omega(z) \leq K\}$ is compact in Z.

4) The measure of approximation $\psi(h, \Omega)$ is assumed to be defined for $h \geq 0, \Omega \geq \Omega^*$, to depend continuously on all its arguments, to be monotonically increasing with respect to Ω for any $h > 0$, and to satisfy the equality $\psi(0, \Omega) = 0, \forall \Omega \geq \Omega^*$.

Conditions 1-3 guarantee that $\overline{Z}$ is nonempty. Tikhonov's scheme for constructing of a regularizing algorithm is based on using the smoothing functional [36]

$$M^\alpha[z] = f[\rho(A_h z, u_\delta)] + \alpha \Omega(z), z \in D, \alpha > 0 \tag{15}$$

in the conditional extreme problem (11). Here *f(x)* is an auxiliary function. A common choice is $f(x) = x^m, m \geq 2$.

We denote the set of extremals of (11) which correspond to a given $\alpha > 0$ by Z^α. Again, conditions 1 - 3 imply that $Z^\alpha$is nonempty. The scheme for constructing an approximation to the set $\overline{Z}$ includes: (a) the choice of the regularization parameter $\alpha_\eta = \alpha_\eta(A_h, u_\delta, \psi, h, \delta)$; (b) the specification of the set Z^{α_η} corresponding to α_η , and a special selection of an element z^{α_η} in this set. We take the element z^{α_η} chosen in this way as a solution of the main problem. Procedures a and b must be accomplished so as to guarantee the convergence $z^{\alpha_\eta} \to \overline{Z}$ as $\eta \to 0$. Thus, Tikhonov regularizing algorithms differ from each other by the method of choosing α_η and by the method of selecting z^{α_η} . Sometimes the latter procedure is not necessary, and one can take an arbitrary element z^{α_η} from Z^{α_η} as an approximate solution.

It is in this way that the generalized analogs of a posteriori parameter choice strategies are used. They were introduced in [40], their descriptions may also be found in [36]. For their formulations we define some auxiliary functions and functionals for $\alpha > 0$, and $z^\alpha \in Z^\alpha$:

$$\begin{aligned}
&\gamma(\alpha) = \Omega(z^\alpha), \beta(\alpha) = \mathrm{f}[\rho(A_h z^\alpha, u_\delta)] \equiv I(z^\alpha) \ , \\
&\phi(\alpha) = M^\alpha[z^\alpha], \\
&\pi(\alpha) = \mathrm{f}[\psi(h, \gamma(\alpha)) + \delta + \mu_\eta] \equiv \Pi(z^\alpha) \ , \\
&\gamma(\alpha) = \Omega(z^\alpha), \beta(\alpha) = \mathrm{f}[\rho(A_h z^\alpha, u_\delta)] \equiv I(z^\alpha) \ , \\
&\phi(\alpha) = M^\alpha[z^\alpha], \pi(\alpha) = \mathrm{f}[\psi(h, \gamma(\alpha)) + \delta + \mu_\eta] \equiv \Pi(z^\alpha) \ , \\
&\rho(\alpha) = \beta(\alpha) - \pi(\alpha) \equiv \mathrm{P}(z^\alpha), \\
&\varepsilon(\alpha) = \phi(\alpha) - f\left\{\mu_\eta + [\psi(h, \gamma(\alpha)) + \delta]^P\right\} \equiv E(z^\alpha), \\
&p = const, 0 < p < 1; \forall z^\alpha \in Z^\alpha; \alpha > 0.
\end{aligned} \tag{16}$$

Here $\mu_\eta = \inf\{\rho(A_h z, u_\delta) + \psi(h, \Omega(z)) + \delta : z \in D\}$ is a generalized measure of incompatibility for nonlinear problems having the properties:$\mu_\eta > \mu, \mu_\eta \to \mu$ as $\eta \to$

0$ [36]. All the functions (9) (except ϕ) are generally many-valued. They have the following properties.

Lemma [36]. If the functions $\gamma, \beta, \pi, \rho, \varepsilon$ are single-valued and continuous everywhere for $\alpha > 0$ except perhaps for a countable set of their common points of discontinuity of the first kind, which are points of multiple-valuedness, then there exists at least two elements z_+^α and z_-^αin the set Z^α such that

$$\gamma\left(\alpha \pm 0\right) = \Omega\left(z_\pm^\alpha\right), \ \rho\left(\alpha \pm 0\right) = \Pi\left(z_\pm^\alpha\right), \ \varepsilon\left(\alpha \pm 0\right) = \mathrm{E}\left(z_\pm^\alpha\right)$$

The functions β, ρ, ε are monotonically non-decreasing and γ, π are non-increasing. The function φ is single-valued, continuous, and non-decreasing.

The generalized discrepancy principle (GDP) for nonlinear problems consists of the following steps.

1. The choice of the regularization parameter as a generalized solution $\alpha > 0$ of the equation

$$\rho(\alpha) = 0 \tag{17}$$

Here and afterwards we say that α is the generalized solution of (17) for a monotone function ρ if α is the ordinary solution or if is the "jump"-point of this function over 0.

2. An approximate solution z^{α_η} from the set Z^{α_η} is selected by means of the following rule: let q >1 and C>1 be fixed constants $\alpha_{1=}{}^{\alpha_\eta}/_q$, and $\alpha_2 = \alpha_\eta \cdot q$ are auxiliary regularization parameters, and let z^{α_1} and z^{α_2} be extremals of (11) for $\alpha = \alpha_{1,2}$. If the inequality

$$\mathrm{I}\left(z^{\alpha_2}\right) \geq C\Pi\left(z^{\alpha_1}\right) - \left(C-1\right)f\left(\mu_\eta\right) \tag{18}$$

holds for z^{α_1} and z^{α_2} , then any elements $z^{\alpha_1} \in Z^{\alpha_\eta}$, subject to the condition $\mathrm{P}\left(z^{\alpha_\eta}\right) \leq 0$ can be taken as the approximate solution. For instance we can take $z^{\alpha_\eta} = z_-^{\alpha_\eta}$. But if

$$\mathrm{I}\left(z^{\alpha_2}\right) < C\Pi\left(z^{\alpha_1}\right) - \left(C-1\right)f\left(\mu_\eta\right) \tag{19}$$

then we choose z^{α_η} so as to have $\mathrm{P}\left(z^{\alpha_\eta}\right) \geq 0$, for example $z^{\alpha_\eta} = z_+^{\alpha_\eta}$,

Note that we do not need any selection rule if α_η is an ordinary solution of (10). In this case the equality $\mathrm{P}\left(z^{\alpha_\eta}\right) = \rho\left(\alpha_\eta\right) = 0$ holds, and an arbitrary element $z^{\alpha_\eta} \in Z^{\alpha_\eta}$ can be taken as an approximate solution.

Theorem 1 [36]. Suppose that for any quasisolution of (8) the inequality

$$\Omega\left(z^*\right) > \Omega^* = \inf\left\{\Omega\left(z\right) z \in D\right\}$$

holds. Then (a) (17) has a positive generalized solution; (b) for any sequence $\eta_n = (h_n, \delta_n)$ such that $\eta_n \to 0$ as $n \to \infty$, the corresponding sequence $\{z_n\}$ of approximate solutions, which is found by GDP has the following properties: $z_n \to \bar{Z}, \Omega\left(z^n\right) \to \bar{\Omega}$ as $n \to \infty$.

In many practical cases it is very convenient to take $\Omega\left(z\right) = \|z\|^r$($r$ is a constant, r >1). If it is known in addition that (8) has a solution on D, then the value μ_η can be omitted. It can be shown that the GDP in linear and nonlinear cases has a number of optimal properties [36].

6 Regularizing algorithms for solving the inverse vibrational problem

The inverse vibrational problem under investigation is nonlinear. Let us consider again (6) in the standard form

$$AF = \Lambda, F \in R^{n(n+1)/2}, \Lambda \in R^{l}. \tag{20}$$

The existence of the normal (relative to a given a priory estimate F^0) pseudosolution $\bar{F}_n$ of the exact problem (20) may be guaranteed if the operator A includes the operator of the direct vibrational problem for a single molecule. The uniqueness of $\bar{F}_n$ cannot be guaranteed.How may the error of the approximate operator Ah be estimated? The estimation $\|A - A_h\| \leq h$ is impossible because the nonlinear operators have no norm. It is obvious that for the operator of the inverse vibrational problem this error is related to errors of the G matrix. It is possible to obtain an estimate in the form [15, 16]

$\|AF - A_hF\| \leq \psi\left(h, F\right),$

where ψ is a known continuous function, which approaches 0 as the errors of the equilibrium geometry of the molecule decrease to zero. In particular, the estimates may be obtained in the form

$$\psi\left(h, F\right) = \psi\left(h, \|F\|\right); \psi\left(h, F\right) = h(A_hF)$$

The error in the right part of (20) is determined as previously.

We arrive at the following formulation of the inverse problem.

Problem I. Suppose we are given (20) where $F \in D \subseteq Z, \Lambda \in U$, Z and U are finite-dimensional spaces, D is a closed set of a priori constraints of the problem, and A is a nonlinear operator continuous in D. It is required to find an approximate solution of (20) when instead of A and Λ, we are given their approximations Ah and $\Lambda\delta$ such that $\|\Lambda - \Lambda_\delta\| \leq \delta, \|AF - A_hF\| \leq \psi\left(h, F\right)$ for $F \in D$; here ψ *(h, F)* is a known continuous function, which approaches zero as $h \rightarrow 0$ uniformly for all $F \in D \cap \bar{S}\left(0, R\right)$, where $\bar{S}\left(0, R\right)$ is a closed ball with center at $F = 0$ and an arbitrary radius R. The error in specifying the operator A involves an error in determining the equilibrium configuration of the molecule, the parameters of which can be found experimentally. Note that Problem I satisfies none of the conditions of correctness of the problem.

We shall consider now the problem of constructing a normal pseudosolution of Problem I with exact right side and operator.

Problem II. It is required to obtain

$$\bar{F}_n = \text{argmin } \left\|F - F^0\right\| \ , \ F : F \in \ D \ , \ \|AF - \Lambda\| = \mu \ ,$$

where $\mu = \|AF - \Lambda\|, F \in D$.

The element $F^0 \in Z$ should be specified from a priori requirements on the solution, using both approximate quantum mechanical calculations and other ideas (for example, the transferability of the force constants to similar fragments of molecules). In the case when a unique solution of (20) exists, it is clear that its normal pseudosolution is identical with the solution itself. Taking all the above mentioned into account we can formulate the following problem.

Problem III. Suppose we are given (20); it is required from the approximate data $\{A_h, \Lambda_\delta, h, \delta\}$ to obtain approximations $F_\eta \in D$ to the solution $\bar{F}_n$ of Problem II such that $F_\eta \underset{\eta \to 0}{\to} \bar{F}_n$ i.e., the algorithm for finding F should be Tikhonov regularizing.

Now we shall consider the simplest formulation of Problem I.

Problem I'. The vibrational spectrum of a single molecule is known, and in (20) the operator A corresponds to the vector $F \in R^{n(n+1)/2}$ which is made up of the elements of the symmetric matrix F of order n, the ordered set of eigenvalues of the matrix GF. We shall use the ordered set of squares of the molecular vibrational frequencies as the right-hand side $\Lambda \in R^l$.

Problem II'. It is required to find the normal solution

$$\bar{F}_n = \arg\min \left\| F - F^0 \right\|, F \in \{F : AF \in \Lambda\},$$

since Problem I is always solvable and as well as solutions are non-unique (except the case when $n = 1$).

Since the operator A in (20) is completely defined by the specification of the matrix G, we shall estimate the deviation of the approximately specified operator A_h (corresponding to certain G_ξ) from the exact operator A (corresponding to G) by the error in specifying matrix G. We suppose that in a certain matrix norm $\|G - G_\xi\| \le \xi$.

In the space R^l of the right-hand sides we shall introduce the Euclidean norm with positive weights, while in $R^{n(n+1)/2}$ we use the Euclidean norm. Suppose that instead of the accurate value of the right-hand side Λ we specify Λ_δ , such that $\|\Lambda - \Lambda_\delta\| \le \delta$

The following theorems on the stability of Problems I and II hold [15].

Theorem 2. Problem I is stable in the Hausdorff metric with respect to the perturbations of the operator and the right-hand side.

Here the Hausdorf distance (metrics), *ρ(A,B)*, is determined in the following way: for any closed set A and B in normed space

$$\rho(A, B) = \sup_{x \in A} \inf_{y \in B} \|x - y\| + \sup_{y \in A} \inf_{x \in B} \|x - y\|.$$

Theorem 3. If Problem II has a unique solution, then it is stable to perturbations of the operator and the right-hand side.

The proofs of these theorems are given in [15].

Most existing methods of solving inverse problems in vibrational spectroscopy are intended for (3) and (4), i.e., for the cases when only the frequencies of the vibrations of the molecules are used as experimental data. As regards the inverse vibrational problem for one molecule, i.e., solutions of (3), it follows from (7) that the choice of the unique solution from their set corresponds to the choice of a certain orthogonal matrix C. Such selection can be made from a priori considerations (e.g., we can use as C the unit matrix or the matrix of eigenvectors of G). This approach suffers from the following drawbacks:

it is impossible to take into account the a priori limitations imposed on the elements of the matrix F;

the results of the calculations are almost never transferable to isotopic varieties of the molecule.

The least-squares method which consists of minimizing on a set of a priori constraints is the one most often encountered. However, in the view of the ill-posed nature of the problem, this method cannot be applied directly to solving problems with approximate data. Rather, the method must be regularized by taking into account the possible inconsistency of the problem and also the non-uniqueness of its solutions. If we attempt to find the normal pseudosolutions (with respect to a certain F_0) we arrive at the formulation of the problem described above.

Which regularizing algorithms can be proposed for finding a normal pseudosolution? One of them is the generalized discrepancy principle [36]. Many other approaches one can find in [35-40]. We have also proposed a modification of the GDP based on the possibility of estimating the error of the operator in the form

$$\|AF - A_h F\| \le h \|A_h F\|, h < 1, \tag{21}$$

which corresponds to specification of the relative error AF and is a more convenient estimate for the problem considered than the monotone function ψ .

Suppose F_η^α is an extremum (possibly non-unique) of Tikhonov's functional

$$M^\alpha [F] = \|A_h F - \Lambda\|^2 + \alpha \left\|F - F^0\right\|^2 \tag{22}$$

in the set D. The existence of an extreme can be proved (see [15]). We shall introduce the function

$$\rho_\alpha(\alpha) = \left\|A_h F_\eta^\alpha - \Lambda_\delta\right\| - \frac{1}{1-h}\left[\hat{\mu} + k\left(\delta + h\|\Lambda_\delta\|\right)\right]$$

where $k > 1$ is a constant and

$$\hat{\mu} = \inf_{F \in D}\left\{\|A_h F - \Lambda_\delta\| + \delta + h\|A_h F\|\right\}$$

If the condition

$$\left\|A_h F^0 - \Lambda_\delta\right\| > \tfrac{1}{1-h}\left[\hat{\mu} + k\left(\delta + h\|\Lambda_\delta\|\right)\right] \tag{23}$$

is satisfied, then the equation $\rho_\eta(\alpha) = 0$ has a generalized solution $\alpha_\eta > 0$ (i.e., α_η is such that $\rho_\eta(\alpha) > 0$ when $\alpha > \alpha_\eta; \rho_\eta(\alpha) < 0$ when $\alpha < \alpha_\eta$). If α_η is a point of continuity of $\rho_\eta(\alpha)$, we have $\rho_\eta(\alpha_\eta) = 0$. This assertion follows from the monotonicity of $\rho_\eta(\alpha)$ and the limit relations (as $\alpha \to 0$ and $\alpha \to +\propto$).

Now we shall formulate the algorithm for finding approximations to the normal pseudosolutions of (20) . If condition (23) is not satisfied, we take $F_\eta = F^0$ as an approximate solution; in the contrary case we find $\alpha_\eta > 0$ (the generalized solution of the equation $\rho_\eta(\alpha) = 0$), and assume $F_\eta = F^{\alpha_\eta}$. If the extreme of the functional (23) is non-unique, we choose the one for which $\|A_h F^{\alpha_\eta} - \Lambda_\delta\| \le \frac{1}{1-h}\left[\hat{\mu} + k\left(\delta + h\|\Lambda_\delta\|\right)\right]$ (the possibility of the making such a choice was shown in [15]).

Theorem 4. The algorithm formulated above is Tikhonov regularizing.

The proof is given in [15].

For versions of Problem I in which the estimate of the error of the operator cannot be written in the form (21), but the requirements of Problem I are satisfied we can use the following version of the generalized discrepancy method.

Problem 4. It is required to obtain

$$\begin{aligned}&\inf \|F - F^0\|\,,\\ &F \in F_\eta = \{F : F \in D, \quad \|A_h F - \Lambda_\delta\| \leq \delta + \psi\,(h, F) + \hat{\mu}\,(A_h, \Lambda_\delta)\}\end{aligned}$$

where the estimate of the measure of incompatibility of the exact problem from above is

$$\hat{\mu} = \inf_{F \in D} \{\|A_h F - \Lambda_\delta\| + \psi\,(h, F) + \delta\}\,.$$

Lemma 2. Suppose the conditions of Problem I are satisfied.

Then $\hat{\mu}_\eta > \mu$, and $\hat{\mu}_\eta \to \mu = \inf\limits_{F \in D} \|AF - \Lambda\|$ as $\eta \to 0$.

Lemma 3. Problem IV is solvable for any $\Lambda_\delta \in U$ such that $\|\Lambda - \Lambda_\delta\| \leq \delta$ and for a continuous operator A such that $\|A_h F - AF\| < \psi\,(h, F)$.

The proof of this lemma relies on the fact that for all $\eta > 0$, the set F_η is nonempty (since $\bar{F}_\eta \in F_\eta$), closed and bounded.

Theorem 5. The algorithm defined by the extremal Problem IV, is Tikhonov regularizing for Problem I.

Details of proofs, estimates of the error of the operator and some details of the numerical realization of the methods are given in [15, 24, 25].

7 Computer implementation

The software package SPECTRUM [18, 24, 25] for processing spectroscopic data includes all regularizing algorithms described above. The structure of SPECTRUM is shown below. The package allows calculation of normal vibrations in the arbitrary system of internal coordinates as well as in Cartesian coordinates. Redundant coordinates are also allowed. Input force constant matrices may be represented in different ways (in Cartesian, internal or symmetry coordinates) and different units ($10^6 cm^{-2}, mdyn/\dot{A}, Hartree/Bohr^2$) ; the package provides automatic transformation of matrices in various coordinate systems and units. Force constants may be fitted to experiment in the standard way (when matrix elements are directly defined) and using Pulay scaling approach, when the set of scale factors is defined for an a priori given force constant matrix. It is possible to impose constraints fixing certain force constants at the pre-defined values. The software allows to introduce any coordinates that are represented as linear combinations of the internal coordinates. The user can use different systems of coordinates including internal, symmetry and some special coordinate system - arbitrary user coordinates.

It is also possible to simultaneously simulate several isotopic species of a molecule for calculation of the single molecular force field. Different molecules can also be treated simultaneously under condition that certain sets of force constants in both molecules are assumed equivalent. The package also provides experimental frequencies fitting when some of the frequencies are not available from the experiment.

The package contains three distinct software modules. The first module (molgraph) is developed for *graphic visualization* of molecular 3D model; the second module (symm) is the *symmetry analyzing routine* which includes the symmetry point groups up to Oh and is very useful in a case of large molecules with high symmetry. It provides the allows the automatic construction of symmetry and local symmetry coordinates, The third module is related to the solving direct or inverse vibrational problems. The current version of a package is implemented advantages of 64-bit processing and virtual memory management with increasing both the processing rate and the limit size of molecules (more than 1000 atoms) possible in a reasonable time of calculations. The package allows the input of different kind of data and the imposing different constraints on the solution of the inverse problem. Various examples of practical applications of regularizing algorithms to data processing in vibrational spectroscopy and structural chemistry are presented in our publications [41-54]. Now the software SPECTRUM is a part of information system ISMOL [55-57], the general scheme of this database is performed in Figure 1.

Input data

General structural data	**Force constants data**
1. Equilibrium geometry (Cartesian coordinates and masses of all atoms) 2. Experimental data on frequencies, including isotopomers, Coriolis constants, mean amplitudes, etc 3. Description of internal coordinates 4. Definition of symmetry coordinates	Possible choices: 1. Matrix in symmetry coordinates. 2. Matrix in internal coordinates. 3. Matrix in Cartesian coordinates.
Choice of numeric method	**Model constraints on matrix F**
Possible choices: 1. Conjugate gradient method of minimizing Tikhonov's functional 2. Monte-Carlo method	Possible choices: 1.Preserving values of certain elements 2.Setting zero values of certain elements 3. (Auto) Setting zero values of elements corresponding to distant coordinates 4. Maintaining equivalence between certain elements in force field matrices for different molecules processed simultaneously

Data processing sequence

1. Conversion of input matrices given in Cartesian or internal coordinates to matrices in symmetry coordinates. Verification of matrix symmetry (adjusting, if necessary).

2. Creating, preprocessing and applying specified constraints on matrix F.

3 (Inverse problem only). Minimizing Tikhonov functional for the problem with the specified numerical method and appropriate choice of regularization parameter.

4. Solving direct problem for resulting matrix in symmetry coordinates.

5. Constructing force field matrix in internal coordinates.

6. Solving direct problem in internal coordinates.

7. Calculating additional characteristics of interest, such as mean square amplitudes of vibrations and potential energy distribution.

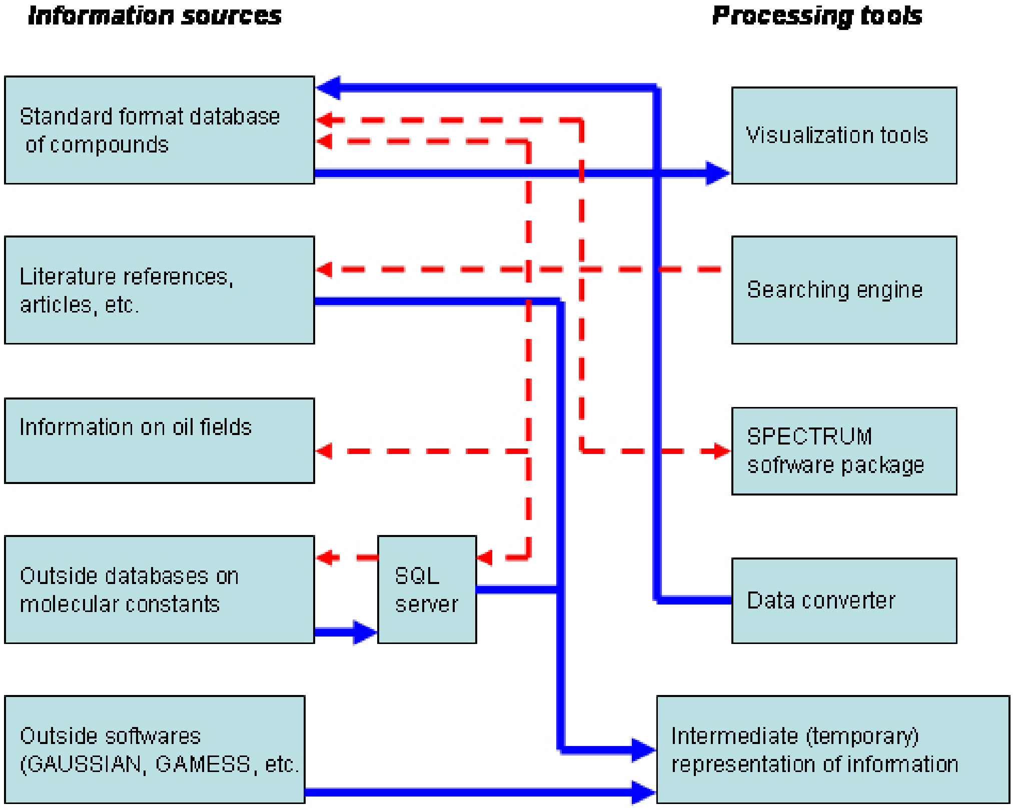

Figure 1:Hybrid information system ISMOL

It should be noted that the package allows the processing of more than one molecule simultaneously (each of them possessing several isotopic species) in a case when the force field model includes equivalencies of certain force matrix elements throughout the optimization process. This option is of special value for verifying transferability properties of force constants. Additional features of the package include the following: possible model constraints applied to matrices in internal coordinates; using redundant internal coordinates in all implemented algorithms; automatic inclusion of redundancy conditions; choice of the regularization parameter in accordance with the generalized discrepancy principle.

Note also that conversion from Cartesian to internal coordinates may be not unique if coordinates are redundant. In this case, the software package allows two ways [34]:

a) to generate a canonical matrix F (that is, with minimal rank sufficient to represent all vibrational degrees of freedom, $3N$-6 where N is number of atoms);

b) to generate a matrix F with the minimal off-diagonal norm (this is a commonly used model assumption).

Results of practical applications for molecular systems including biological ones will be published in the next paper.

Acknowledgement

Authors are grateful to the Russian Foundation on Basic Research for partial financial support (Projects No 14-03-00929-a, 16-01-00450-a).

References

[1] H.J. Hadamard, *Le probleme de Cauchy et les equations aux derivees partielles lineaires hyperbolique*, Hermann, Paris, 1932.

[2] E. Bright Wilson Jr., J.C. Decius, P.C. Cross, Molecular Vibrations, *The Theory of Infrared and Raman Vibrational Spectra, Dover Publications*, New York, 1980.

[3] M.V. Volkenstein, L.A. Gribov, M.A. El'yashevich, B.I. Stepanov, *Vibrations of molecules*, Nauka, Moscow, 1972.

[4] L.M. Sverdlov, M.A. Kovner, E.P. Krainov, *Vibrational spectra of polyatomic molecules,* Wiley, New York, 1973.

[5] S. Califano, Vibrational states, Wiley, New York, 1976.

[6] Handbook of Vibrational Spectroscopy, J. M. Chalmers and P. R. Griffiths, Eds., John Wiley & Sons, 2002.

[7] P.C. Peinter, M.M. Coleman, J.L. Koenig, *The theory of vibrational spectroscopy and its application to polymeric materials*, John Wiley and Sons, Inc., New York, 1982.

[8] L.A. Gribov, W.J. Orwille-Thomas, *Theory and methods of calculation of molecular spectra.* Wiley, Chichester, New York, 1988.

[9] R.J. Meier, *Calculating the vibrational spectra of molecules: An introduction for experimentalists with contemporary examples*, Vibrational Spectroscopy, Vol. 43 (2007), pp. 26-37.

[10] I.V. Kochikov, G.M. Kuramshina, I.A. Pentin, A.G. Iagola, *Regularizing algorithm of the inverse vibrational problem solution*, Doklady Akademii Nauk SSSR, Vol. 261, no. 5 (1981), pp. 1104-1106.

[11] I.V. Kochikov, G.M. Kuramshina, Yu.A. Pentin, A.G. Yagola, *A stable method for the calculation of the force fields of polyatomic molecules in dependent coordinates, Theoretical and Experimental Chemistry*, Vol. 20, no. 1 (1984), pp. 66-72.

[12] I.V. Kochikov, A.G. Yagola, G.M. Kuramshina, V.M. Kovba, Y.A. Pentin, *Force-fields and mean amplitudes of vivration of chromium, molybdenum and tungsten oxotetrafluorides.* Theochem – Journal of Molecular Structure, Vol. 15, no. 3-4 (1984), pp. 355-360.

[13] I.V. Kochikov, G.M. Kuramshina, Yu.A. Pentin, A.G. Yagola, *Calculation of force fields of polyatomic molecules by the Tikhonov regularization method*, Doklady Akademii Nauk SSSR, Vol. 283, no. 4 (1985), pp. 850-854.

[14] I.V. Kochikov, G.M. Kuramshina, S.I.Chernik, A.G. Yagola, *Algorithm of the detection of the regular solution of a wave reverse problem, based on the Monte-Carlo method*, Vestnik Moskovskogo universiteta seriya 2 Khimiya, Vol. 27, no. 6 (1986), pp. 597-601.

[15] I.V. Kochikov, G.M. Kuramshina, A.G. Yagola*, Stable numerical methods of solving certain inverse problems of vibrational spectroscopy*, USSR Computational Mathematics and Mathematical Physics, Vol. 27, no. 11-12 (1987), pp. 33-40.

[16] I.V. Kochikov, G.M. Kuramshina, Yu.A. Pentin, A.G. *Yagola, Regularizing algorithms for molecular force field calculations*, Journal of Molecular Structure (Theochem), Vol. 272 (1992), pp. 13-33.

[17] I.V. Kochikov, G.M. Kuramshina, Yu.A. Pentin, *Application of Tikhonov's regularization method to the calculation of force fields of polyatomic molecules*, Russian Journal of Physical Chemistry, Vol. 61, no.4 (1987), pp. 451-457.

[18] I.V. Kochikov, G.M. Kuramshina, *A complex of programs for the force field calculations of polyatomic molecules by the Tikhonov regularization method.* Vestnik Moskovskogo Universiteta,

Seria 2, Khimia. Vol. 26, no. 4 (1985), pp. 354-358.

[19]. I.V. Kochikov, G.M. Kuramshina, Yu.A. Pentin, V.M. Senyavin, A.G. Yagola, *Realization of program complex for computation of multiatomic molecules using regulating algorithms on IBM PC/AT.* Zhurnal Fizicheskoi Khimii, Vol. 64, no. 12 (1990), pp. 3393-3395.

[20] A.S. Dostavalova, I.V. Kochikov, G.M. Kuramshina, A.G. Yagola, *Simultaneous calculation of force fields for series of polyatomic molecules*, Doklady Akademii Nauk SSSR, Vol. 315, no. 6 (1990), pp.1368-1373.

[21] I.V. Kochikov, A.G. Yagola, G.M. Kuramshina, V.M. Kovba, Y.A. *Pentin, Calculation of force fields of chromium, molybdenum and tungsten hexafluorides and dioxofluorides by means of Tikhonov's regularization method.* Spectrochimica Acta, Ser. A., Vol. 41, no. 1/2 (1985), pp.185-189.

[22] I.V. Kochikov, G.M. Kuramshina, S.V.Syn'ko, Yu.A. Pentin, *Normal coordinate analysis of CH2ClSiX3, CHCl2lSiX3, CCl3lSiX3 (X = H, D, Cl) molecules using Tikhonov's regularization method*, Journal of Molecular Structure, Vol. 172 (1988), pp. 299-308.

[23] G.M. Kuramshina, Yu.A. Pentin, *Modeling of vibrational spectra of polyatomic molecules using stable numerical methods.* Russian Journal of Physical Chemistry. 67: 95-101 (1993).

[24] I.V. Kochikov, G.M. Kuramshina, Yu.A. Pentin, A.G. Yagola, *Inverse problems of vibrational spectroscopy.* Moscow State Univ. Publ., Moscow (1993)

[25] A.G. Yagola, I.V.Kochikov, G.M. Kuramshina, Yu.A. Pentin,. *Inverse problems of vibrational spectroscopy. (Inverse and III-Posed Problems)* VSP, Zeist, The Netherlands, 1999, pp 297.

[26] J.A. Pople, R. Krishnan, H.B. Schlegel, J.S. Binkley, *Electron correlation theories and their application to the study of simple reaction potential surfaces*, International Journal of Quantum Chemistry Symp. XIV (1978), pp. 545-560.

[27] G. Fogarashi, P. Pulay, *Ab initio calculation of force fields and vibrational spectra. In: Durig, J.R. (ed.): Vibrational spectra and structure.* Elsevier, Amsterdam, Vol. 1, (1985) pp. 125- 217.

[28] W.J. Hehre, L. Radom, L., P.V.R.Schlegel, *Ab initio molecular orbital theory.* Wiley, New York, 1985.

[29] P. Pulay, *Second and third derivatives of variational energy expressions: application to multiconfigurational self-consistent field wave functions.* Journal of Chemical Physics, Vol. 78 (1983), pp. 5043-5051.

[30]. R.G. Parr, W.Yang, *Density-functional theory of atoms and molecules.* Oxford Univ. Press, Oxford, 1992.

[31] P. Carbonniere, V. Barone, *Performances of different density functionals in the computation of vibrational spectra beyond the harmonic approximation*, Chemical Physics Letters, Vol. 399 (2004), 226–229.

[32] J.E. Boggs, *Molecular quantum mechanics - what now and where next.* Journal of Molecular Structure , Vol. 130 (1985), pp. 31-42.

[33] J.B. Foresman, A.E. Frish, A.E.: *Exploring chemistry with electronic structure methods 2nd ed.* - Gaussian, Inc. Pittsburgh, PA, 1996.

[34] G.M. Kuramshina, F. Weinhold, I.V. Kochikov, A.G. Yagola, A.G., Yu.A. *Pentin, Joint treatment of ab initio and experimental data in molecular force field calculations with Tikhonov's method of regularization.* Journal of Chemical Physics, Vol. 100, no.2 (1994), pp. 1414-1424.

[35]. A.N Tikhonov, A.V Goncharsky, V.V. Stepanov, A.G. Yagola, *Numerical methods for the solution of ill-posed problems.* Kluwer, Dordrecht, 1995.

[36] A.N. Tikhonov, A.S. Leonov, A.G. Yagola, *Nonlinear ill-posed problems.* Chapman & Hall , London, 1997.

[37] A.G. Yagola, *Choice of the regularization parameter using the generalized discrepancy principle,* Doklady Akademii Nauk SSSR, Vol. 245 (1979), pp. 37-39.

[38]. A.G. Yagola, *The generalized discrepancy method in reflexive spaces.* Doklady Akademii Nauk SSSR, Vol. 249 (1979), pp. 71-73.

[39] A.G. Yagola, *Choice of the regularization parameter when solving ill-posed problems in reflexive spaces*, USSR Computational Mathematics and Mathematical Physics, Vol. 20 (1980), pp. 586-596.

[40] A.V. Goncharsky, A.S. Leonov, A.G. Yagola*, The discrepancy principle for solving nonlinear ill-posed problems.* Soviet. Math. Doklady, Vol. 15 (1974), pp. 166-168.

[41] G.M. Kuramshina, F. Weinhold, I.V. Kochikov, Yu.A. Pentin, A.G. Yagola, *Regularized force molecule fields based on nonempirical quantum chemical calculations.* Zhurnal Fizicheskoi Khimii, Vol. 68, no. 3 (1994), pp. 401-414.

[42] Yu.A. Pentin, G.M. Kuramshina, *Molecular force fields obtained using the nonempirical stabilizer of the regularizing functional*, Journal of Structural Chemistry, Vol 36, no.2 (1995), pp. 180-190.

[43] G.M. Kuramshina, A.V. Lobanov, A.G. Yagola, *Searching normal solution of inverse vibrational problem by the method of Monte-Carlo Journal of Molecular Structure*, Vol. 348 (1995), pp. 147-150.

[44] G.M. Kuramshina, A.G. Yagola, *A priori constraints in the force field calculations of polyatomic molecules.* Journal of Structural Chemistry, Vol. 38, no. 2 (1997), pp. 181-194.

[45] G.M. Kuramshina, F.Weinhold, *Constraints on the values of force constants for molecular force field models based on ab initio calculations*, Journal of Molecular Structure, Vol. 410-411 (1997), pp. 457-462.

[46]. I.V. Kochikov, G.M. Kuramshina, A.V. Stepanova, A.G. Yagola, *Regularized method of scaling factors for calculations of molecular force fields*, Moscow University Physics, Bulletine, no. 5 (1997), pp. 21-25.

[47]. I.V. Kochikov, Yu.I. Tarasov, G.M. Kuramshina, V.P. Spiridonov, A.G. Yagola, T.G. Strand, *Regularizing algorithm for determination of equilibrium geometry and harmonic force field of free molecules from joint use of electron diffraction, vibrational spectroscopy and ab initio data with application to benzene*, Journal of Molecular Structure, Vol. 445 (1998), pp. 243-258.

[48] G.M. Kuramshina, Yu.A. Pentin, A.G. Yagola, *New approaches for molecular conformer force field analysis in combination with ab initio results*, Journal of Molecular Structure, Vol. 509 (1999), pp. 255-263.

[49] G.M.Kuramshina, Yu.A.Pentin, *DFT force fields, vibrational spectra and potential functions for hindered internal rotation of CF3CH2CH2Cl and CCl3CH2CH2Cl*, Journal of Molecular Structure, Vol. 481 (1999), pp. 161-168.

[50]. I.V. Kochikov, Yu.I. Tarasov, V.P. Spiridonov, G.M. Kuramshina, A.G. Yagola, A.S. Saakjan, 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*, Journal of Molecular Structure, Vol. 485-486 (1999), pp. 421-443.

[51] I.V. Kochikov, Yu.I. Tarasov, V.P. Spiridonov, G.M. Kuramshina, A.S. Saakjan, A.G. Yagola, *The use of ab initio anharmonic force fields in experimental studies of equilibrium molecular geometry*, Journal of Molecular Structure, Vol. 550-551 (2000), pp. 429-438.

[52] I.V. Kochikov, Yu. I. Tarasov, V. P. Spiridonov, G. M. Kuramshina, D. W. H. Rankin, A.S. Saakjan, A.G. Yagola. *The equilibrium structure of thiophene by the combined use of electron diffraction, vibrational spectroscopy and microwave spectroscopy guided by theoretical calculations.* Journal of Molecular Structure, Vol. 567-568 (2001), pp. 29-40.

[53] I.V. Kochikov, G.M. Kuramshina, A.V. Stepanova, A.G. Yagola, *Numerical aspects of the calculation of scaling factors from experimental data*, Numerical methods and programming, Vol. 5 (2004), pp. 281-290.

[54] I.V. Kochikov, G.M. Kuramshina, A.V. Stepanova, *New approach for the correction of Ab initio molecular force fields in Cartesian coordinates*, International Journal of Quantum Chemistry, Vol. 109, no. 1 (2009), pp. 28-33.

[55] I.V. Kochikov, G.M. Kuramshina, L.M. Samkov, D.A. Sharapov, S.A.Sharapova, *A hybrid database (information system) on the molecular spectral constants*, Numerical methods and programming, Vol. 6 (2005), pp.83-87.

[56] I.V. Kochikov, G.M. Kuramshina, L.M. Samkov, D.A. Sharapov, S.A. Sharapova *Data structuring and formalization in the hybrid database on molecular spectral constants (ISMOL).* Numerical methods and programming. Vol. 7 (2006), pp. 111-116.

[57] I.V. Kochikov, G.M. Kuramshina, L.M. Samkov, D.A. Sharapov, S.A.Sharapova*. Identifica-*

tion algorithms for substances in the hybrid database on molecular spectral constants (ISMOL). Numerical methods and programming, Vol 8 (2007), pp. 70-73.

Gulnara M. Kuramshina
Department of Physical Chemistry
Faculty of Chemistry
Lomonosov Moscow State University
Moscow 119991, Russia
Fax: (7)(495)932-8846
Email: kuramshi@phys.chem.msu.ru
Anatoly G. Yagola
Department of Mathematics
Faculty of Physics
Lomonosov Moscow State University
Moscow 119991, Russia
Fax: (7)(495)932-8820
Email: yagola@physics.msu.ru

Received 01.12.2016, Accepted 20.12.2016