

# Theoretical ROVibrational Energies (TROVE): A robust numerical approach to the calculation of rovibrational energies for polyatomic molecules

Sergei N. Yurchenko<sup>a,\*</sup>, Walter Thiel<sup>b</sup>, Per Jensen<sup>c</sup>

<sup>a</sup> Institut für Physikalische Chemie und Elektrochemie, TU Dresden, D-01062 Dresden, Germany

<sup>b</sup> Max-Planck-Institut für Kohlenforschung, Kaiser-Wilhelm-Platz 1, D-45470 Mülheim an der Ruhr, Germany

<sup>c</sup> FB C—Mathematik und Naturwissenschaften, Fachgruppe Chemie, Bergische Universität Wuppertal, D-42097 Wuppertal, Germany

Received 18 May 2007; in revised form 12 July 2007

Available online 17 August 2007

## Abstract

We present a new computational method with associated computer program TROVE (Theoretical ROVibrational Energies) to perform variational calculations of rovibrational energies for general polyatomic molecules of arbitrary structure in isolated electronic states. The (approximate) nuclear kinetic energy operator is represented as an expansion in terms of internal coordinates. The main feature of the computational scheme is a numerical construction of the kinetic energy operator, which is an integral part of the computation process. Thus the scheme is self-contained, i.e., it requires no analytical pre-derivation of the kinetic energy operator. It is also general, since it can be used in connection with any internal coordinates. The method represents an extension of our model for pyramidal  $XY_3$  molecules reported previously [S.N. Yurchenko, M. Carvajal, P. Jensen, H. Lin, J.J. Zheng, W. Thiel, *Mol. Phys.* 103 (2005) 359]. Non-rigid molecules are treated in the Hougen–Bunker–Johns approach [J.T. Hougen, P.R. Bunker, J.W.C. Johns, *J. Mol. Spectrosc.* 34 (1970) 136]. In this case, the variational calculations employ a numerical finite basis representation for the large-amplitude motion using basis functions that are generated by Numerov–Cooley integration of the appropriate one-dimensional Schrödinger equation.

© 2007 Elsevier Inc. All rights reserved.

**Keywords:** Theoretical rovibrational energies; Polyatomic molecules

## 1. Introduction

Expressed in terms of the Cartesian coordinates of the nuclei in a laboratory-fixed axis system, the Schrödinger equation for the translation, rotation, and vibration of a polyatomic molecule, obtained in the Born–Oppenheimer approximation, has a very simple form

$$\left( -\frac{\hbar^2}{2} \sum_{i=1}^N \frac{1}{m_i} \nabla_i^2 + V \right) \Psi_{\text{trv}} = E_{\text{trv}} \Psi_{\text{trv}}. \quad (1)$$

Here, nucleus  $i$  ( $=1, 2, 3, \dots, N$ ) has the mass  $m_i$  and the coordinates  $(R_{iX}, R_{iY}, R_{iZ})$  in the chosen laboratory-fixed

axis system  $XYZ$ , and  $\nabla_i^2 = \partial^2/\partial R_{iX}^2 + \partial^2/\partial R_{iY}^2 + \partial^2/\partial R_{iZ}^2$ . In Eq. (1),  $V$  is the Born–Oppenheimer potential energy function and  $\Psi_{\text{trv}}$  is the translation–rotation–vibration wavefunction with the associated translation–rotation–vibration energy  $E_{\text{trv}}$ . Even though the Schrödinger equation in Eq. (1) has a mathematically simple form, it is very badly suited for actual solution: the coordinate choice does not reflect the fact that the translation can be exactly separated from the rotation and vibration motion, and that these latter types of motion are often fairly weakly coupled. Therefore, it is necessary to introduce more suitable coordinates, and this usually involves the introduction of a molecule-fixed axis system  $xyz$  (sometimes referred to as a *body-fixed* axis system) with origin at the nuclear center of mass. The molecule-fixed axis system is defined so that molecular rotation can be viewed as the motion of this axis

\* Corresponding author. Fax: +49 351 463 35953.

E-mail address: [s.yurchenko@chemie.tu-dresden.de](mailto:s.yurchenko@chemie.tu-dresden.de) (S.N. Yurchenko).

system (the molecule-fixed axis system follows the rotation of the molecule), whilst vibration is the motion of the nuclei relative to the molecule-fixed axis system, described by suitable internal coordinates.

Molecules differing by their number of nuclei and/or by their structure, however, require different internal coordinates to describe their vibrational motion. Therefore, such molecules have very different rotation–vibration Hamiltonian operators. One can hardly imagine a universal set of internal coordinates capable of properly representing arbitrary vibrational motions. As a result, the spectroscopic literature describes numerous, quite complex nuclear kinetic energy operators for different kinds of molecules such as diatomic molecules, triatomic chain molecules, triatomic ‘triangular’ molecules, four-atomic chain molecules, four-atomic pyramidal molecules, etc. Each theoretical model is normally accompanied by a computer program for obtaining the eigenvalues and eigenfunctions of the resulting rotation–vibration Hamiltonian, but each of these programs can, of course, only be used for the type of molecule for which it is designed. Each program involves a kinetic energy based on a coordinate choice which cannot be changed without profound changes to the program. It would seem that each type of molecule requires a specific coordinate choice and, consequently, a specific kinetic energy operator and, finally, a specific computer program for computing its rotation–vibration energies. The only coordinates that can be defined in a general manner are the normal coordinates, but they are mostly suited for describing small-amplitude vibrations around a single potential energy minimum and so their applicability is limited.

We report here the development of a general computational paradigm, which can be used in a “black box” fashion, for calculating rotation–vibration energies for molecules in isolated electronic states. This computational scheme is applicable for a general molecule of arbitrary structure. It allows the user to make an arbitrary choice of the vibrational coordinates to be used in the kinetic energy operator. This generality is achieved by including the coordinate transformation into the computational process as a numerical procedure.

The advantages of numerically determined kinetic energy operators have already been recognized [1–9], especially by the discrete variable representation (DVR) community [2,8,9]. A numerical pointwise definition of the kinetic energy operator perfectly fits the basic idea of DVR. Besides it can provide an exact (within the Born–Oppenheimer approximation) evaluation of the corresponding matrix elements [9]. Alternatively, the kinetic energy operator can be defined as a Taylor series in (or in functions of) internal coordinates with expansion parameters obtained numerically [3–5]. We follow this latter strategy, and our Hamiltonian operator is expressed as a Taylor series in terms of the internal coordinates, and its eigenvalues and eigenfunctions are obtained in the finite basis representation (FBR) in terms of a basis set that can be arbitrarily chosen. The coordinate transformation,

or rather the momentum transformation, from Cartesian coordinates to the chosen set of vibrational coordinates is made with the help of Taylor series whose expansion coefficients are computed numerically in a recursive procedure. The scheme is self-contained, i.e., it requires no analytical pre-derivation of the kinetic energy operator. The method reported here is an extension of our model for  $XY_3$  molecules [10] to an arbitrary non-rigid molecule.

An important application of the energies and wavefunctions resulting from solutions of the molecular rotation–vibration Schrödinger equation is the simulation of rotation–vibration spectra; such simulations aid the assignments of experimental spectra and facilitate the interpretation of remote-sensing experiments that produce spectra of molecules in environments such as the upper layers of the Earth’s atmosphere and interstellar space. To provide this type of theoretical support for experimental high-resolution molecular spectroscopy it is obviously not sufficient to calculate only the molecular vibrational energies. Rotation–vibration energies and wavefunctions must be calculated and, furthermore, these calculations must be carried out for a rotational excitation (given by the angular momentum quantum number  $J$ ) sufficiently high that experimental spectra can be simulated. It should be mentioned that in the theoretical approach of the present paper, maximum separation between rotation and vibration is achieved through the use of Eckart conditions [11,12]. The approach is designed with the aim of it being able to access high  $J$  values and thus provide energies and wavefunctions for the simulations of realistic rotation–vibration spectra. Alternative variational treatments of the rotation–vibration motion of small molecules (see, for example, Refs. [13,14] for  $XY_3$  pyramidal molecules) typically make use of a nuclear kinetic energy operator exact within the Born–Oppenheimer approximation. This kinetic energy operator is expressed in terms of geometrically defined coordinates and no measures are taken to minimize rotation–vibration interaction. This usually limits the range of  $J$  values accessible as exemplified by the fact that Ref. [14] reports  $NH_3$  energy calculations for  $J \leq 2$  whereas, with the theoretical model described in Ref. [10], of which the model of the present work is an extension, we have simulated spectra involving states with  $J \leq 18$  for this molecule [15] and calculated the energies of states with  $J \leq 80$  for  $PH_3$  [16].

The paper is structured as follows. In Section 2, we introduce the method for the recursive construction of the coordinate transformation and the corresponding Hamiltonian operator. The constructed Hamiltonian is given as a Taylor expansion suited for the computation of matrix elements. In Section 3, we show how arbitrary vibrational coordinates can be generated in terms of predefined, basic coordinate types. The potential expansion is discussed in Section 4, while the numerically defined vibrational basis sets and matrix element calculations are elucidated in Section 5. The rotational basis functions are defined in Section 6. Our “black box” algorithm for

calculating rotation–vibration energies which makes use of all these ingredients is discussed in detail in Section 7. Applications of this algorithm as implemented in the computer program TROVE are presented in Section 8 for three example molecules,  $\text{H}_2\text{S}$ ,  $\text{CH}_3^+$ , and  $\text{CO}$ . The results are verified against experimental data (where available) and the results of other theoretical calculations. In Section 9, we discuss the convergence properties of the variational calculations and in Section 10 we offer conclusions.

## 2. The construction of the kinetic energy operator through a recursive numerical scheme

We aim at expressing the Hamiltonian contained in Eq. (1) in terms of  $3N$  new generalized coordinates

$$\Xi = (R_X^{\text{CM}}, R_Y^{\text{CM}}, R_Z^{\text{CM}}, \theta, \phi, \chi, \xi_1, \xi_2, \dots, \xi_{3N-6}). \quad (2)$$

Here  $R_F^{\text{CM}} = \sum_{i=1}^N m_i R_{iF} / \sum_{j=1}^N m_j$  ( $F = X, Y, Z$ ) is the  $F$ -coordinate of the nuclear center of mass; these three coordinates describe the translational motion. The Euler angles  $(\theta, \phi, \chi)$  define the orientation of the  $xyz$  axis system relative to the  $XYZ$  system (see, for example, Section 10.1.1 of Ref. [12]) and describe the over-all rotation, and the vibrational motion is described by the internal coordinates  $\xi_n$ . The derivation given here follows closely the general scheme proposed by Sørensen [17].

It is well known (see, for example, Ref. [12]) that the transformed kinetic energy operator  $\hat{T}$  is essentially a quadratic form in the generalized momenta

$$\hat{\Pi} = (\hat{P}_X^{\text{CM}}, \hat{P}_Y^{\text{CM}}, \hat{P}_Z^{\text{CM}}, \hat{J}_x, \hat{J}_y, \hat{J}_z, \hat{p}_1, \hat{p}_2, \dots, \hat{p}_{3N-6}), \quad (3)$$

where  $\hat{P}_F^{\text{CM}}$  ( $F = X, Y, Z$ ) is the momentum conjugate to the translational coordinate  $R_F^{\text{CM}}$ ,  $(\hat{J}_x, \hat{J}_y, \hat{J}_z)$  are the  $xyz$  components of the total angular momentum [12], and  $\hat{p}_n = -i\hbar\partial/\partial\xi_n$  ( $n = 1, \dots, 3N - 6$ ) is the momentum conjugate to the vibrational coordinate  $\xi_n$ . That is, we can write  $\hat{T}$  as

$$\begin{aligned} \hat{T} = & \frac{1}{2} \sum_{F=X,Y,Z} \hat{P}_F^{\text{CM}} G_{FF} \hat{P}_F^{\text{CM}} \\ & + \frac{1}{2} \sum_{\alpha=x,y,z} \sum_{\alpha'=x,y,z} \hat{J}_\alpha G_{\alpha,\alpha'}(\xi) \hat{J}_{\alpha'} \\ & - \frac{i\hbar}{2} \sum_{\alpha=x,y,z} \sum_{n=1}^{3N-6} \left[ \hat{J}_\alpha G_{\alpha,n}(\xi) \frac{\partial}{\partial \xi_n} + \frac{\partial}{\partial \xi_n} G_{\alpha,n}(\xi) \hat{J}_\alpha \right] \\ & - \frac{\hbar^2}{2} \sum_{n=1}^{3N-6} \sum_{n'=1}^{3N-6} \frac{\partial}{\partial \xi_n} G_{n,n'}(\xi) \frac{\partial}{\partial \xi_{n'}} + U(\xi), \end{aligned} \quad (4)$$

where  $G_{XX} = G_{YY} = G_{ZZ} = 1/\sum_{j=1}^N m_j$ , and the pseudopotential term  $U(\xi)$  and the ‘expansion coefficients’  $G_{\alpha,\alpha'}(\xi)$ ,  $G_{\alpha,n}(\xi)$  and  $G_{n,n'}(\xi)$  depend on the vibrational coordinates  $\xi_n$ . We use the notation ‘ $(\xi)$ ’ to signify that the function depends on the complete set of vibrational coordinates  $\xi_1, \xi_2, \dots, \xi_{3N-6}$ . The first term on the right-hand side of Eq. (4) is the Hamiltonian associated with the translational

motion; it is exactly separable from the remaining terms which describe the ‘internal’ motion of the nuclei.

Introducing the ‘generalized’ kinetic energy coefficient  $G_{\lambda,\lambda'}(\xi)$  and denoting by  $\Xi_\lambda$  and  $\hat{\Pi}_\lambda$  ( $\lambda = 1, \dots, 3N$ ) general elements of  $\Xi$  and  $\hat{\Pi}$ , respectively, we rewrite the kinetic energy operator in Eq. (4) in the compact form

$$\hat{T} = \frac{1}{2} \sum_{\lambda=1}^{3N} \sum_{\lambda'=1}^{3N} \hat{\Pi}_\lambda G_{\lambda,\lambda'}(\xi) \hat{\Pi}_{\lambda'} + U(\xi). \quad (5)$$

As vibrational coordinates  $\xi_n$  we can choose any coordinates that represent the internal degrees of freedom and unambiguously define the instantaneous relative positions of the nuclei. Examples are normal coordinates, internal displacement coordinates (i.e., the displacements of bond lengths, bond angles, and dihedral angles from their respective equilibrium values), linearized internal coordinates,  $\dots$ , and symmetrized linear combinations of such coordinates. We must express the quantities  $G_{\lambda,\lambda'}(\xi)$  [Eq. (5)], the pseudopotential term  $U(\xi)$  [Eq. (5)], and the Born–Oppenheimer potential energy function  $V$  [Eq. (1)] in terms of the  $\xi_n$  coordinates. As discussed above, we aim at expressing  $G_{\lambda,\lambda'}(\xi)$  and  $U(\xi)$  as series expansions. In practice, we choose convenient functions

$$g_n = g_n(\xi_n), \quad (6)$$

where each  $g_n$  only depends on *one* coordinate  $\xi_n$ , and expand

$$G_{\lambda,\lambda'} = \sum_{l_1,l_2,l_3,\dots} G_{l_1 l_2 l_3 \dots}^{\lambda,\lambda'} g_1^{l_1} g_2^{l_2} g_3^{l_3} \dots, \quad (7)$$

$$U = \sum_{l_1,l_2,l_3,\dots} U_{l_1 l_2 l_3 \dots} g_1^{l_1} g_2^{l_2} g_3^{l_3} \dots, \quad (8)$$

where  $G_{l_1 l_2 l_3 \dots}^{\lambda,\lambda'}$  and  $U_{l_1 l_2 l_3 \dots}$  are constant expansion coefficients. We also express  $V$  as an expansion

$$V = \sum_{l_1,l_2,l_3,\dots} V_{l_1 l_2 l_3 \dots} f_1^{l_1} f_2^{l_2} f_3^{l_3} \dots, \quad (9)$$

where  $V_{l_1 l_2 l_3 \dots}$  are constant expansion coefficients, in terms of convenient expansion functions

$$f_n = f_n(\xi_n). \quad (10)$$

An example of a function  $f_n \neq \xi_n$  is  $f_n = 1 - \exp(-a\xi_n) = 1 - \exp(-a(r - r_e))$ , where  $a$  is a Morse parameter,  $r_n$  is a bond length, and  $r_e$  is its equilibrium value. Another example is  $f_n = \cos(\xi_n) = \cos(\alpha_n)$ , where  $\alpha_n$  is an inter-bond angle. For the kinetic energy expansions in Eqs. (7) and (8), we usually choose  $g_n = f_n$  or  $g_n = \xi_n$ . In the former case, the kinetic energy expansion variable  $g_n$  coincides with the potential energy variable  $f_n$ . This is convenient for certain one-dimensional basis functions, such as harmonic-oscillator functions and Morse-oscillator eigenfunctions, since then many required matrix elements can be computed from analytical expressions. Alternatively, when the matrix elements are computed by numerical quadrature, and we

are free to choose any expansion variables  $g_n$ , we prefer the simplest choice of  $g_n = \xi_n$ .

Sørensen's strategy [17] for converting the kinetic energy operator contained in Eq. (1) to that given in Eq. (4) is based on the fact that the initial form of the operator can be straightforwardly expressed as

$$\hat{T} = \sum_{F=X,Y,Z} \sum_{i=1}^N \frac{\hat{P}_{iF}^2}{2m_i} = \sum_{i=1}^N \frac{\hat{\mathbf{P}}_i^2}{2m_i}, \quad (11)$$

where  $\hat{P}_{iF}$  ( $i = 1, \dots, N$ ,  $F = X, Y, Z$ ) is the momentum conjugate to the Cartesian coordinate  $R_{iF}$  and the vector  $\hat{\mathbf{P}}_i$  has the  $XYZ$  coordinates  $(\hat{P}_{iX}, \hat{P}_{iY}, \hat{P}_{iZ})$ . Sørensen [17] now inserts in Eq. (11) the chain-rule transformation in Hermitian form

$$\hat{P}_{iF} = \frac{1}{2} \sum_{\lambda=1}^{3N} (s_{\lambda,iF} \hat{\Pi}_\lambda + \hat{\Pi}_\lambda s_{\lambda,iF}) \quad (12)$$

with

$$s_{\lambda,iF} = \frac{\partial \Xi_\lambda}{\partial R_{iF}}. \quad (13)$$

The Jacobian-matrix elements  $s_{\lambda,iF}$  define vectors  $\mathbf{s}_{\lambda,i}$  with  $XYZ$  coordinates  $(s_{\lambda,iX}, s_{\lambda,iY}, s_{\lambda,iZ})$ , and we have from Eq. (12) and from the definition of  $\hat{\mathbf{P}}_i$  that

$$\hat{\mathbf{P}}_i = \frac{1}{2} \sum_{\lambda=1}^{3N} (\mathbf{s}_{\lambda,i} \hat{\Pi}_\lambda + \hat{\Pi}_\lambda \mathbf{s}_{\lambda,i}). \quad (14)$$

When this result is inserted in Eq. (11) we obtain fairly straightforwardly

$$G_{\lambda,\lambda'} = \sum_{i=1}^N \frac{\mathbf{s}_{\lambda,i} \cdot \mathbf{s}'_{\lambda',i}}{m_i}, \quad (15)$$

$$U = \sum_{\lambda=1}^{3N} \sum_{\lambda'=1}^{3N} \sum_{i=1}^N \left\{ \frac{1}{8} m_i^{-1} [\hat{\Pi}_\lambda, \mathbf{s}_{\lambda,i}] \cdot [\hat{\Pi}_{\lambda'}, \mathbf{s}'_{\lambda',i}] + \frac{1}{4} m_i^{-1} \mathbf{s}_{\lambda,i} \cdot [\hat{\Pi}_\lambda, [\hat{\Pi}_{\lambda'}, \mathbf{s}'_{\lambda',i}]] \right\}. \quad (16)$$

With the definition

$$t_{iF,\lambda} = \frac{\partial R_{iF}}{\partial \Xi_\lambda}, \quad (17)$$

application of the chain rule yields

$$\sum_{i=1}^N \sum_{F=X,Y,Z} \frac{\partial \Xi_\lambda}{\partial R_{iF}} \frac{\partial R_{iF}}{\partial \Xi_{\lambda'}} = \mathbf{s}_{\lambda,i} \cdot \mathbf{t}_{i,\lambda'} = \delta_{\lambda,\lambda'}, \quad (18)$$

where the vectors  $\mathbf{t}_{i,\lambda'}$  have the  $XYZ$  coordinates  $(t_{iX,\lambda'}, t_{iY,\lambda'}, t_{iZ,\lambda'})$ . If the  $\mathbf{t}_{i,\lambda'}$ -vectors are known, we can obtain the  $s_{\lambda,i}$ -vectors by solving Eq. (18). Sørensen [17] showed that the  $xyz$  coordinates of the  $\mathbf{t}_{i,\lambda'}$ -vectors are given as

$$\begin{aligned} t_{i\alpha,\alpha'} &= \delta_{\alpha,\alpha'} \quad (\text{translation}), \\ t_{i\alpha,\beta} &= \sum_{\gamma} \epsilon_{\alpha\beta\gamma} R_{i\gamma}^{\text{MS}} \quad (\text{rotation}), \\ t_{i\alpha,n} &= \partial R_{i\alpha}^{\text{MS}} / \partial \xi_n \quad (\text{vibration}). \end{aligned} \quad (19)$$

In Eq. (19), the indices  $\alpha$ ,  $\alpha'$ ,  $\beta$ , and  $\gamma$  all assume the values  $x$ ,  $y$ , or  $z$ ;  $\epsilon_{\alpha\beta\gamma}$  is an element of the fully antisymmetric tensor;  $n = 1, 2, \dots, 3N - 6$ ; and  $R_{i\alpha}^{\text{MS}}$ ,  $\alpha = x, y, \text{ or } z$ , is the  $\xi$ -dependent  $\alpha$ -coordinate of the position vector of nucleus  $i$  in the  $xyz$  axis system. The expression for  $t_{i\alpha,\beta}$  is consistent with the  $xyz$  axis system being defined by Eckart conditions [11] as described, for example, in Ref. [12]. The results in Eq. (19), in conjunction with the fact that the scalar product of Eq. (18) yields the same value when calculated in the  $XYZ$  and  $xyz$  axis systems, allow us to consider Eq. (18) as a closed equation system fully defined in the  $xyz$  axis system. We solve Eq. (18) numerically by utilizing the perturbation-type, recursive approach developed in Ref. [10]. For the sake of completeness we reproduce in the following paragraphs a few results from Ref. [10].

Assuming that all  $g_n$  are small and of the same order of magnitude  $\sim O(\epsilon)$ , we expand both  $s_{\lambda,i\alpha}$  and  $t_{i\alpha,\mu}$  as power series in  $g_n(\xi_n)$

$$s_{\lambda,i\alpha}(g_n) = \sum_{L \geq 0} \sum_{L[l]} s_{L[l]}^{\lambda,i\alpha}(g) L[l], \quad (20)$$

$$t_{i\alpha,\mu}(g_n) = \sum_{L \geq 0} \sum_{L[l]} t_{L[l]}^{i\alpha,\mu}(g) L[l], \quad (21)$$

where we have introduced the notation

$$\begin{aligned} &\sum_{L[l]} f_{L[l]}(x) L[l] \\ &\equiv \sum_{l_1=0}^L \sum_{l_2=0}^{(L-l_1)} \sum_{l_3=0}^{(L-l_1-l_2)} \dots \sum_{l_{(3N-6)-1}=0}^{(L-l_1-l_2-\dots-l_{(3N-6)-2})} f_{l_1 l_2 l_3 \dots l_{(3N-6)}}^L \prod_i x_i^{l_i}. \end{aligned} \quad (22)$$

For each set of  $\{L, l_1, l_2, \dots, l_{3N-6}\}$ , the index  $l_{3N-6}$  in Eq. (22) is redundant and constrained to  $L - l_1 - l_2 - \dots - l_{(3N-6)-1}$ . This definition causes the expansion terms in Eq. (22) to be arranged according to their total power  $L = \sum_{n=1}^{3N-6} l_n$ , which defines the perturbation orders  $O(\epsilon^L)$  in Eqs. (20) and (21). Substituting Eqs. (20) and (21) into Eq. (18) and collecting coefficients of  $g_1^{l_1} g_2^{l_2} g_3^{l_3} \dots$  at each perturbation order  $L = l_1 + l_2 + \dots + l_{3N-6}$ , we obtain systems of linear equations for  $s_{L[l]}^{\lambda,i\alpha}$ , which are defined recursively as (see also Ref. [10])

$$\sum_{i\alpha} t_{0[0]}^{i\alpha,\mu} s_{0[0]}^{\lambda,i\alpha} = \delta_{\lambda,\mu}, \quad (23)$$

$$\sum_{i\alpha} t_{0[0]}^{i\alpha,\mu} s_{L[l]}^{\lambda,i\alpha} = b_{L[l]}^{\mu,\lambda}, \quad L > 0, \quad (24)$$

where

$$b_{L[l]}^{\mu,\lambda} = - \sum_{K=0}^{L-1} \sum_{K[k]} \sum_{i\beta} t_{(L-K)[l-k]}^{i\beta,\mu} s_{K[k]}^{\lambda,i\beta}. \quad (25)$$

For each combination of indices ( $L[l], \lambda$ ), Eqs. (23) and (24) represent a set of linear equations of type  $\mathbf{T}\mathbf{x} = \mathbf{b}$  for the vector  $\mathbf{x}$  with elements  $x_{i\alpha} = s_{L[l]}^{\lambda, i\alpha}$  ( $i = 1, \dots, N$ ,  $\alpha = x, y, z$ ,  $\lambda = 1, \dots, 3N$ ). The  $3N$ -dimensional ‘right-hand side’ vector  $\mathbf{b}$  with the elements  $b_\mu = b_{L[l]}^{\mu, \lambda}$  is recursively determined at each perturbation order  $L$  by means of Eq. (25). The corresponding  $3N \times 3N$   $\mathbf{T}$ -matrix is given by the elements  $T_{\mu, i\alpha} = t_{0[0]}^{\mu, i\alpha}$  ( $\mu = 1, \dots, 3N$ ), which are independent of  $\lambda$ . The linear equation systems in Eqs. (23) and (24) can be prepared and solved numerically. The order of the system is small ( $3N$ ), i.e., the solution  $s_{L[l]}^{\lambda, i\alpha}$  can be easily obtained by means of linear-algebra routines from standard computer libraries.

All coefficients required in Eqs. (23)–(25) are constructed using Eq. (19): we expand  $R_{ix}^{\text{MS}}$  and  $\partial R_{ix}^{\text{MS}}/\partial \xi_n$  as power series in the vibrational variables  $g_n$  using the notation of Eq. (22)

$$R_{ix}^{\text{MS}}(\xi) = \sum_{L \geq 0} \sum_{L[l]} R_{L[l]}^{ix}(g)^{L[l]}, \quad (26)$$

$$\partial R_{ix}^{\text{MS}}/\partial \xi_n = \sum_{L \geq 0} \sum_{L[l]} R_{L[l]}^{ix, n}(g)^{L[l]}, \quad (27)$$

where we have assumed that all necessary derivatives of  $R_{ix}^{\text{MS}}$  with respect to  $g_n$  and  $\xi_n$  are defined. Insertion into Eqs. (19) and (21) yields

$$t_{L[l]}^{ix, \alpha'} = \delta_{\alpha, \alpha'} \delta_{L, 0}, \quad (28)$$

$$t_{L[l]}^{ix, \beta} = \sum_{\gamma} \epsilon_{\alpha\beta\gamma} R_{L[l]}^{i\gamma}, \quad (29)$$

$$t_{L[l]}^{ix, n} = R_{L[l]}^{ix, n}. \quad (30)$$

In our scheme all quantities are given as polynomials in the  $g_n$  variables. This includes the kinetic energy coefficients  $G_{\lambda, \lambda'}$  and the pseudopotential  $U$  in Eqs. (15) and (16), respectively:  $G_{\lambda, \lambda'}$  is a product of two polynomials  $s_{\lambda, i\alpha}$ , while  $U$  can be constructed from products and derivatives of  $s_{\lambda, i\alpha}$  functions [17]

$$U = \frac{\hbar^2}{4} \sum_i \frac{1}{m_i} \left[ \frac{1}{2} \sum_{\alpha\beta\gamma} \epsilon_{\alpha\beta\gamma} \epsilon_{\alpha\beta'\gamma'} s_{\beta, i\gamma} s_{\beta', i\gamma'} + \sum_{\alpha\beta\gamma} \sum_n \epsilon_{\alpha\beta\gamma} s_{n, i\gamma} \frac{\partial s_{\beta, i\alpha}}{\partial \xi_n} - \sum_{\alpha} \sum_{nm} s_{n, i\alpha} \frac{\partial^2 s_{m, i\alpha}}{\partial \xi_n \partial \xi_m} - \sum_{\alpha} \sum_{nm} \frac{1}{2} \frac{1}{m_i} \frac{\partial s_{n, i\alpha}}{\partial \xi_n} \frac{\partial s_{m, i\alpha}}{\partial \xi_m} \right], \quad (31)$$

where  $\alpha, \beta, \gamma, \beta', \gamma'$  assume the values  $x, y$ , or  $z$ , and  $n, m = 1, 2, 3, \dots, 3N - 6$ . Eq. (31) is obtained from Eq. (16) after applying Eq. (3). In order to evaluate Eqs. (15) and (31) we translate all analytical operations on  $s_{\lambda, i\alpha}$  into numerical operations on their expansion coefficients. The product of two polynomials  $h = f \times g$  can be defined by means of its expansion coefficients as

$$h_{L[l]} = \sum_{K=0}^L \sum_{K[k]} f_{(L-K)[l-k]} g_{K[k]}, \quad (32)$$

where  $f_{L[l]}$ ,  $g_{L[l]}$ , and  $h_{L[l]}$  are the expansion coefficients of  $f$ ,  $g$ , and  $h$ , respectively. The order of the expansion  $h$  is taken as  $\min\{\text{order}(f), \text{order}(g)\}$ . The derivative  $\partial/\partial \xi_n$  of the

polynomial  $f$  of order  $L + 1$  is a polynomial  $h$  of order  $L$  with the coefficients

$$h_{L[l]}^L = (L_n + 1) f_{L[l], i_1, i_2, \dots, i_{L_n+1}, \dots, i_{3N-6}}^{L+1}, \quad (33)$$

which can also be written using the compact form of Eq. (22)

$$h_{L[l]} = (L_n + 1) f_{(L+1)[L_n+1]}. \quad (34)$$

For example, the kinetic energy factor  $G_{\lambda, \lambda'}(\xi)$  is a product of two polynomials  $s_{\lambda, i\beta}$  and  $s_{\lambda', i\beta}$  expressed as power series in the  $g_n$  variables

$$G_{\lambda, \lambda'}(\xi) = \sum_{L=0}^{N_{\text{kin}}} \sum_{L[l]} G_{L[l]}^{\lambda, \lambda'}(g_n)^{L[l]} \quad (35)$$

with expansion coefficients defined by

$$G_{L[l]}^{\lambda, \lambda'} = \sum_{i=1}^N \sum_{\beta=x,y,z} \frac{1}{m_i} \sum_{K=0}^L \sum_{K[k]} s_{(L-K)[l-k]}^{\lambda, i\beta} s_{K[k]}^{\lambda', i\beta}. \quad (36)$$

The expansion of the pseudopotential function  $U$

$$U(\xi) = \sum_{L=0}^{N_{\text{kin}}} \sum_{L[l]} U_{L[l]}(g)^{L[l]} \quad (37)$$

is obtained from Eq. (31) in the same manner, employing the operations of Eqs. (32) and (34).

The construction of the kinetic energy coefficients  $G_{\lambda, \lambda'}$  and the pseudopotential  $U$  by means of the recurrence relations in Eqs. (23) and (24) is purely numerical and an integral part of the energy calculation process. The kinetic energy coefficients  $G_{\lambda, \lambda'}$  and the pseudopotential function  $U$  in Eqs. (7) and (8), respectively, are truncated after the same perturbation order  $L = N_{\text{kin}}$ . The convergence issues are discussed in detail in Section 9.

### 3. Vibrational coordinates

The computational scheme defined by Eqs. (23) and (24) is general; it can be used with any choice of physically satisfactory vibrational coordinates  $\xi_n$ , as long as we can provide the associated coefficients  $t_{ix, \mu}$ . Towards this end, we have defined three basic types of coordinate representations which can be used to construct more sophisticated coordinates. These are (A) linearized coordinates, (B) geometrically defined coordinates, and (C) coordinate representations for non-rigid molecules (i.e., molecules with large-amplitude vibrations [12]).

#### 3.1. Linearized coordinates

The linearized coordinates are introduced in terms of the Cartesian displacements  $d_{ix}$  of the nuclei from their equilibrium positions  $a_{ix}$  in the  $xyz$  axis system [12]

$$R_{ix}^{\text{MS}} = a_{ix} + d_{ix}. \quad (38)$$

In general, the  $3N - 6$  independent internal displacement coordinates  $\xi_n$  are non-linear functions of the displacements

$d_{ix}$ . A set of  $3N - 6$  linearized coordinates  $\xi_n \equiv \xi_n^\ell$  are defined to be linear combinations of  $d_{ix}$  and to coincide with the  $3N - 6$  coordinates  $\xi_n$  in the linear approximation [12]

$$\xi_n^\ell = \sum_{i=1}^N \sum_{\alpha=x,y,z} B_{n,ix} d_{ix}; \quad n = 1, 2, \dots, 3N - 6, \quad (39)$$

where  $B_{n,ix} = \partial \xi_n / \partial d_{ix}$ , derived at equilibrium. The elements  $B_{n,ix}$  can be obtained from purely geometrical considerations (see, for example, Ref. [12]).

The constant equilibrium coordinates  $a_{ix}$  satisfy

$$\sum_{i=1}^N m_i a_{ix} = 0 \quad (40)$$

because the  $xyz$  axis system has its origin at the nuclear center of mass, and traditionally the  $xyz$  axis system is further chosen such that it is a principal axis system at equilibrium [12]

$$\sum_{i=1}^N m_i a_{ix} a_{i\beta} = 0, \quad \alpha \neq \beta. \quad (41)$$

It is often trivial to determine the quantities  $a_{ix}$ ; they depend on the equilibrium values for the structural parameters of the molecule. However, in general we can obtain them numerically as part of the computational procedure: we provide the equilibrium structure in a  $Z$ -matrix format (see, for example, Ref. [18]) and thus generate the equilibrium nuclear coordinates  $a'_{ix}$  in an arbitrarily chosen molecule-fixed axis system  $x'y'z'$ , which is then transformed to the molecular principal axis system  $xyz$ . This latter operation is purely numerical and involves the diagonalization of the inertial matrix.

For linearized coordinates, the expansion in Eq. (26) is linear

$$d_{ix} = \sum_n A_{ix,n} \xi_n^\ell. \quad (42)$$

The  $\mathbf{A}$ -matrix elements  $A_{ix,n}$  can be determined numerically; they satisfy the three center-of-mass conditions

$$\sum_{i=1}^N m_i A_{ix,n} = 0, \quad \alpha = x, y, z, \quad (43)$$

the three Eckart conditions [11]

$$\sum_{i=1}^N \sum_{\beta,\gamma=x,y,z} m_i (\epsilon_{\alpha\beta\gamma} A_{i\beta,n} a_{i\gamma}) = 0, \quad \alpha = x, y, z, \quad (44)$$

and the  $3N - 6$  orthogonality conditions

$$\sum_{i=1}^N \sum_{\alpha=x,y,z} A_{ix,n} B_{m,ix} = \delta_{n,m}. \quad (45)$$

Eqs. (43)–(45) are linear in  $A_{ix,n}$ , and so these equations can always be solved numerically.

In order to construct the kinetic energy operator expansions in Eqs. (7) and (8), we need to specify the variables  $g_n$ . If we simply choose  $g_n = \xi_n^\ell$  for all  $n$ , the corresponding

elements  $t_{\alpha,ix}^{L[l]}$  ( $L > 0$ ) in Eqs. (23)–(25) take a very simple form: the only non-vanishing  $R_{L[l]}^{ix}$  and  $R_{L[l]}^{ix,n}$  are

$$R_{0[0]}^{ix} = a_{ix}, \quad R_{0[0]}^{ix,n} = A_{ix,n}, \quad R_{1[l_n=1]}^{ix} = A_{ix,n}. \quad (46)$$

Here,  $R_{1[l_n=1]}^{ix}$  refers to the linear term  $\xi_n^\ell$  in the expansion of Eq. (26). That is, the only non-zero  $t_{ix,\mu}$  values in Eqs. (28)–(30) correspond to  $L \leq 1$ .

In the case of  $g_n \neq \xi_n^\ell$  the corresponding elements  $t_{ix,\mu}$  are constructed in two steps: first we numerically invert the one variable (1D) function  $g_n = g_n(\xi_n^\ell)$  defined by Eq. (6) and then expand  $\xi_n^\ell = \xi_n^\ell(g_n)$  in terms of  $g_n$  to obtain  $R_{L[l]}^{ix}$ . This is also implemented as a numerical procedure in the program based on the theory described here.

The linearized coordinates represent our first basic coordinate type. They can be used to define other equivalent variables, such as the normal coordinates  $q_n$ . The Cartesian displacement coordinates  $d_{ix}$  are linear in the normal coordinates

$$d_{ix} = \sum_{n=1}^{3N-6} \frac{l_{ix,n}}{\sqrt{m_i}} q_n \quad (47)$$

with expansion coefficients  $l_{ix,n}/\sqrt{m_i}$ . The coefficients  $l_{ix,n}$  are subject to center-of-mass and Eckart conditions (see, for example, Ref. [12]) obtained by substituting by  $l_{ix,n}/\sqrt{m_i}$  for  $A_{ix,n}$  in Eqs. (43) and (44), to the normalization conditions

$$\sum_{i=1}^N \sum_{\alpha=x,y,z} l_{ix,m} l_{ix,n} = \delta_{m,n}, \quad (48)$$

and to the requirement that in the harmonic approximation, the potential energy is expressed as

$$V = \frac{1}{2} \sum_{n=1}^{3N-6} \omega_n q_n^2. \quad (49)$$

In the computational scheme described here, the transformation matrix elements  $l_{ix,n}$  are computed for any molecule as follows. We prepare the ‘linearized’ matrix  $A_{ix,m}$  as described above; it is then transformed so that it satisfies Eqs. (48) and (49). The transformations involved are standard linear algebra operations, which are implemented in standard computer libraries.

The non-zero coefficients  $R_{L[l]}^{ix}$  and  $R_{L[l]}^{ix,n}$  are given by Eq. (46) with  $l_{ix,n}/\sqrt{m_i}$  substituted for  $A_{ix,n}$ , i.e.,

$$R_{0[0]}^{ix} = a_{ix}, \quad R_{0[0]}^{ix,n} = l_{ix,n}/\sqrt{m_i}, \quad R_{1[l_n=1]}^{ix} = l_{ix,n}/\sqrt{m_i}, \quad (50)$$

where we show the non-vanishing coefficients only.

### 3.2. Geometrically defined coordinates

We use the  $Z$ -matrix technique (see, for example, Ref. [18]) to set up the geometrically defined coordinates. Although the linearized and normal coordinates provide a rather simple form for the kinetic energy operator, they often fail to describe the potential energy properly. The geometrically defined coordinates are generally better

suiting for describing the potential energy function, and they constitute the second basic coordinate type treated here. The Cartesian displacements  $d_{ix}$  in Eq. (38) are nonlinear functions of the geometrically defined coordinates  $\xi_n$ , and the expansion coefficients  $R_{L[l]}^{ix}$  must thus be obtained also for  $L > 1$ . We do this as follows. We use a numerical procedure, which computes  $3N - 6$  values of geometrically defined coordinates  $\xi_n$  for a given set of the  $3N$  instantaneous values of the Cartesian displacements  $d_{ix}$ . Eckart conditions [12] are introduced to define the molecule-fixed axis system, and the  $3N$  Cartesian displacement coordinates must satisfy both these conditions and the center-of-mass conditions. The coordinate transformation  $R_{ix}^{\text{CM}} = R_{ix}^{\text{CM}}(\xi_n)$  established in this manner is then differentiated with respect to the variables  $g_n$  by finite-difference methods. This produces expansion coefficients  $R_{L[l]}^{ix}$  as derivatives at equilibrium. It should be noted that it is crucial to maintain high computational accuracy when employing the finite-difference methods. We use ‘quadruple’ precision (i.e., 16-byte representation of real variables) in this step of the computation (see also the discussion in Section 4).

In order to construct the expansion coefficients  $R_{L[l]}^{ix,n}$  we employ the chain rule

$$\frac{\partial R_{ix}^{\text{MS}}}{\partial \xi_l} = \frac{\partial R_{ix}^{\text{MS}}}{\partial g_l} \frac{\partial g_l}{\partial \xi_l}. \quad (51)$$

We first calculate  $g_l(\xi_l)$ , expand it in terms of  $\xi_l$ , and finally multiply by the derivative of the polynomial  $R_{ix}^{\text{MS}}$  in Eq. (26) with respect to  $g_l$ . The required operations have been defined above in terms of polynomial expansion coefficients. The coefficients  $R_{L[l]}^{ix}$  are then obtained using Eqs. (28)–(30).

### 3.3. Coordinates describing large-amplitude vibrations

Thus far, we have assumed that both the kinetic and the potential energy operators can be represented as converged, Taylor-type power series as expressed in Eqs. (7)–(9). We now extend the treatment to the situation of a molecule with one vibrational mode of large amplitude, such as inversion, torsion or other types of motion along a minimum energy ‘reaction path.’ We denote the coordinate describing this vibrational mode as  $\rho$  with the corresponding index as  $r$  and arrange the set of internal coordinates  $\{\xi_1, \xi_2, \dots, \xi_{3N-6}\}$  such as to have  $\rho$  last, i.e.,  $\xi_{3N-6} = \rho$  and  $r = 3N - 6$ . We follow the Hougen–Bunker–Johns (HBJ) approach [19]. The  $3N - 7$  small-amplitude vibrational modes are defined as displacements from a flexible reference configuration that follows the large-amplitude motion. When the molecule is in the reference configuration, nucleus  $i$  has the  $\alpha$ -coordinate ( $\alpha = x, y, \text{ or } z$ )  $a_{i\alpha}(\rho)$  in the  $xyz$  axis system. In practice, we modify the treatment given in the preceding sections by introducing a grid of equidistant  $\rho$ -values  $\rho_k, k = 0, \dots, n_\rho$ , and expand all relevant functions in terms of the small-amplitude coordinates  $\{\xi_1, \xi_2, \dots, \xi_{3N-7}\}$  (or in terms of the  $g_n(\xi_n)$  or  $f_n(\xi_n)$  as

appropriate) at each such point. Let us first consider the linearized coordinates  $\xi_n^\ell$  as the basic coordinates type utilized for the small-amplitude coordinates (in analogy to the treatment introduced in Ref. [10], which is extended here from  $\text{XY}_3$  to arbitrary molecules). In the HBJ approach [19], the Cartesian coordinates  $R_{ix}^{\text{MS}}$  introduced in Eq. (38) are defined as

$$R_{ix}^{\text{MS}} = a_{ix}(\rho_k) + \sum_n A_{ix,n}(\rho_k) \xi_n^\ell. \quad (52)$$

The quantities  $a_{ix}(\rho)$  and  $A_{ix,n}(\rho)$  are constructed so that they satisfy the center-of-mass conditions [Eq. (40)], the Eckart conditions [12], and the Sayvetz [20] condition

$$\sum_{i=1}^N \sum_{\alpha=x,y,z} m_i \frac{\partial R_{ix}^{\text{MS}}(\rho)}{\partial \rho} a_{i\alpha} = 0. \quad (53)$$

The procedure to compute the coefficients  $A_{ix,n}(\rho_k)$  at a  $\rho$  grid point is equivalent to the scheme described in Section 3.1, where the last row  $\lambda = 3N - 6$  of the matrix  $T_{\lambda,ix}$  is now substituted by

$$T_{r,ix}(\rho_k) = \sum_{i=1}^N \sum_{\beta=x,y,z} m_i \frac{\partial a_{i\beta}(\rho_k)}{\partial \rho}, \quad (54)$$

and the corresponding linear equations  $\mathbf{T}\mathbf{x} = \mathbf{b}$  are solved at each grid point  $\rho_k$ . The same applies also to  $R_{L[l]}^{ix}(\rho_k)$  and  $R_{L[l]}^{ix,n}(\rho_k)$  in Eqs. (28)–(30); these quantities are given by

$$R_{0[0]}^{ix}(\rho_k) = a_{ix}(\rho_k), \quad R_{1[l_n=1]}^{ix}(\rho_k) = A_{ix,n}(\rho_k), \quad (55)$$

$$R_{0[0]}^{ix,n}(\rho_k) = A_{ix,n}(\rho_k), \quad R_{0[0]}^{ix,r}(\rho_k) = \frac{\partial a_{ix}}{\partial \rho}(\rho_k), \quad (56)$$

$$R_{1[l_n=1]}^{ix,r}(\rho_k) = \frac{\partial A_{ix,n}}{\partial \rho}(\rho_k). \quad (57)$$

The expansion coefficients  $G_{L[l]}^{\lambda,\lambda'}(\rho_k)$  and  $U_{L[l]}$  are also computed on the grid of  $\rho$ -values.

The flexible reference configuration can also be combined with geometrically defined coordinates describing the small-amplitude vibrations: in this case, we carry out the construction procedure described in Section 3.2 at each  $\rho_k$  point, using Eq. (52).

## 4. Expansion of the potential energy function

In order to complete the construction of the rovibrational Hamiltonian to be used in the variational solution of the corresponding Schrödinger equation, the potential energy function should also be expressed as an expansion, Eq. (22), in terms of the variables  $f_n(\xi_n)$

$$V(\xi_n) = \sum_{L=0}^{N_{\text{pot}}} \sum_{L[l]} V_{L[l]}(f_n)^{L[l]}. \quad (58)$$

It should be noted that an arbitrary ‘input’ potential energy function may be defined in terms of coordinates other than those used here, say  $\xi_n' \neq \xi_n$ . We then employ finite-difference methods for computing the expansion coefficients

required to expand the coordinates  $\xi'_n$  in terms of the  $\xi_n$  variables and, subsequently, to expand  $V(\xi')$  in terms of the  $\xi_n$  variables. This finite-difference scheme provides a high degree of flexibility; we can in principle use arbitrary input potentials. However, in practice we require that these input potentials be compatible with one of our basic coordinate types.

The drawback of the numerical derivation is the accumulation of round-off errors. For high-order derivatives the finite-difference method will eventually break down. We have found, for example, that numerical derivation with the two-point central formula works well only for derivatives up to 6th order when computed in double precision (i.e., with 8-byte representation of real variables). Employing ‘quadruple’ precision (i.e., 16-byte representation) allowed us to calculate accurate values of 12th–16th order derivatives by means of the two-point central formula. This is sufficient for most applications since in practice, we require that the expansion in Eq. (58) has converged for  $L = 10$  at the most. If higher-order expansions are necessary, the calculations become too expensive and it is not feasible to use the approach described here. Obviously, we can investigate in convergence tests whether the requirement is satisfied.

## 5. Vibrational basis functions and matrix elements

We define the vibrational basis set functions as eigenfunctions of suitable model Hamiltonian operators, constructed by simplifying the total Hamiltonian. Our rotation–vibration Hamiltonian  $\hat{H}_{rv}$  is given as an expansion

$$\hat{H}_{rv} = \frac{1}{2} \sum_{L \geq 0} \sum_{L[l]} \sum_{\lambda, \lambda'} \hat{\Pi}_\lambda G_{L[l]}^{\lambda \lambda'}(g)^{L[l]} \hat{\Pi}_{\lambda'} + \sum_{L \geq 0} \sum_{L[l]} U_{L[l]}(g)^{L[l]} + \sum_{L \geq 0} \sum_{L[l]} V_{L[l]}(f)^{L[l]} \quad (59)$$

in terms of the variables  $g_n(\xi_n)$  and  $f_n(\xi_n)$ . All expansion coefficients are computed numerically as outlined above. When we use coordinates describing a large-amplitude vibration (Section 3.3), expansions are made at each point in a grid of equidistantly spaced  $\rho_k$ -values.

The vibrational basis set functions  $|v\rangle$  are constructed as products of 1D basis functions

$$|v\rangle = \prod_v |v_v\rangle = \phi_{v_1}(\xi_1) \phi_{v_2}(\xi_2) \dots \phi_{v_{3N-6}}(\xi_{3N-6}), \quad (60)$$

where  $\phi_{v_n}$  is a standard analytical function (we have implemented harmonic-oscillator and Morse-oscillator eigenfunctions) or a function obtained numerically in a Numerov-Cooley integration [21]. In both scenarios, these functions are eigenfunctions of a model Hamiltonian obtained by introducing simplifications in Eq. (59). When we use Numerov-Cooley integration to obtain the basis functions, we solve the Schrödinger equation

$$\hat{H}_n^{(1D)} |v_n\rangle = E_{v_n} |v_n\rangle \quad (61)$$

for the Hamiltonian

$$\hat{H}_n = -\frac{\hbar^2}{2} \frac{\partial}{\partial \xi_n} G_{n,n}^{(1D)}(\xi_n) \frac{\partial}{\partial \xi_n} + V^{(1D)}(\xi_n) + U^{(1D)}(\xi_n). \quad (62)$$

In Eq. (62), we obtain  $G_{n,n}^{(1D)}(\xi_n)$ ,  $V^{(1D)}(\xi_n)$ , and  $U^{(1D)}(\xi_n)$  by constraining the  $3N - 7$  coordinates  $\xi_k$ ,  $k \neq n$ , to their equilibrium values in the  $\xi$ -dependent functions of Eq. (59). The subscript ‘1D’ is used here to indicate this ‘1D-reduction.’ The one-dimensional Schrödinger equation in Eq. (61) is solved by the numerical Numerov-Cooley method [21] on a grid of equidistantly spaced  $\xi_n$ -values as detailed in Ref. [22].

The vibrational matrix elements of the Hamiltonian in Eq. (59) are expressed in terms of 1D matrix elements involving integration over the coordinate  $\xi_n$

$$V_{v_n, v'_n}^l(n) = \langle v_n | f_n^l(\xi_n) | v'_n \rangle, \quad (63)$$

$$T_{v_n, v'_n}^{(0),l}(n) = \langle v_n | g_n^l(\xi_n) | v'_n \rangle, \quad (64)$$

$$T_{v_n, v'_n}^{(1),l}(n) = \left\langle v_n \left| g_n^l(\xi_n) \frac{\partial}{\partial \xi_n} \right| v'_n \right\rangle = - \left\langle v'_n \left| \frac{\partial}{\partial \xi_n} g_n^l(\xi_n) \right| v_n \right\rangle, \quad (65)$$

$$T_{v_n, v'_n}^{(2),l}(n) = \left\langle v_n \left| \frac{\partial}{\partial \xi_n} g_n^l(\xi_n) \frac{\partial}{\partial \xi_n} \right| v'_n \right\rangle, \quad (66)$$

where  $f_n^l(\xi_n) = [f_n(\xi_n)]^l$  and  $g_n^l(\xi_n) = [g_n(\xi_n)]^l$ . When we use the HBJ approach to describe a molecule with large-amplitude vibration, the index  $n$  in Eqs. (63)–(66) runs over the values  $1, 2, \dots, 3N - 7$ ; in the other cases considered it runs over the values  $1, 2, \dots, 3N - 6$ . The 1D integrals in Eqs. (63)–(66) are computed with Simpson’s rule, and the first derivatives  $\frac{\partial}{\partial \xi_n} |v'_n\rangle$  are computed with finite-difference methods. Although the kinetic and potential energy operators are expanded in terms of the different variables  $g_n$  and  $f_n$ , respectively [see Eq. (59)], for the calculation of the derivatives they are transformed back to depend on the variable  $\xi_n$ . The vibrational matrix elements of the Hamiltonian  $\hat{H}_{rv}$  in Eq. (59) are now given by

$$\begin{aligned} \langle v | \hat{H}_{rv} | v' \rangle &= \frac{1}{2} \sum_{\alpha\beta} \hat{J}_\alpha \prod_j T_{v_j, v'_j}^{(0),l_j}(j) G_{L[l]}^{\alpha\beta} \hat{J}_\beta + \prod_j V_{v_j, v'_j}^{l_j}(j) V_{L[l]} \\ &\quad - \frac{i\hbar}{2} \sum_{\alpha n} \hat{J}_\alpha \prod_j \left( T_{v_j, v'_j}^{(\delta_{n,j}),l_j}(j) G_{L[l]}^{\alpha n} - T_{v'_j, v_j}^{(\delta_{n,j}),l_j}(j) G_{L[l]}^{\alpha n} \right) \\ &\quad - \frac{\hbar^2}{2} \sum_{nm} \prod_j T_{v_j, v'_j}^{(p_j),l_j}(j) G_{L[l]}^{nm} + \prod_j T_{v_j, v'_j}^{(0),l_j}(j) U_{L[l]}, \end{aligned} \quad (67)$$

where  $p_j$  in the last line is used as a short-hand notation for  $\delta_{n,j} + \delta_{m,j}$ . In the HBJ approach, the large-amplitude coordinate  $\rho$  is processed in a slightly different manner. In this case, the  $\rho$  dependence of all functions entering into the Hamiltonian of Eq. (59) is given explicitly on a grid of  $\rho_k$  values through the expansion coefficients  $V_{L[l]}(\rho)$ ,  $U_{L[l]}(\rho)$ , and  $G_{L[l]}^{\lambda, \mu}(\rho)$ . The Numerov-Cooley eigenfunctions of Eq. (62) are used for computing (by means of Simpson’s rule) all necessary matrix elements



$$V_{v_r, v_r'}^{L[l]}(r) = \langle v_r | V_{L[l]}(\rho) | v_r' \rangle, \quad (68)$$

$$U_{v_r, v_r'}^{L[l]}(r) = \langle v_r | U_{L[l]}(\rho) | v_r' \rangle, \quad (69)$$

$$G_{v_r, v_r'}^{\lambda, \lambda', L[l]}(r) = \langle v_r | G_{L[l]}^{\lambda, \lambda'}(\rho) | v_r' \rangle, \quad (70)$$

$$G_{v_r, v_r'}^{r, L[l]}(r) = -G_{v_r, v_r'}^{r, \lambda, L[l]} = \left\langle v_r \left| G_{L[l]}^{\lambda, r}(\rho) \frac{\partial}{\partial \rho} \right| v_r' \right\rangle, \quad (71)$$

$$G_{v_r, v_r'}^{r, r, L[l]}(r) = \left\langle v_r \left| \frac{\partial}{\partial \rho} G_{L[l]}^{r, r}(\rho) \frac{\partial}{\partial \rho} \right| v_r' \right\rangle. \quad (72)$$

We denote by  $v_r$  the principal quantum number associated with the coordinate  $\rho$ ; according to our convention  $r = 3N - 6$ . In the HBJ case, we must replace in the Hamiltonian matrix of Eq. (67) all the quantities  $G_{L[l]}^{\lambda, \mu}$ ,  $U_{L[l]}$ , and  $V_{L[l]}$  by the corresponding matrix elements in Eqs. (68)–(72).

When we use analytical 1D basis functions  $\phi_{v_n}$  (at present, we have implemented the use of harmonic-oscillator and Morse-oscillator eigenfunctions), the 1D Hamiltonian in Eq. (62) is defined in accordance with these functions (i.e., as the Hamiltonian of the 1D harmonic-oscillator or the Morse-oscillator). The analytical expressions for the appropriate 1D matrix elements are implemented in the program.

## 6. Rotational basis functions

The rotational basis set is built from the rigid-rotor functions given as linear combinations  $|J, K, m, \pm\rangle$

$$|J, 0, m, +\rangle = |J, 0, m\rangle, \quad (73)$$

$$|J, K, m, \pm\rangle = \frac{p(J, K, \pm)}{\sqrt{2}} (|J, K, m\rangle \pm |J, -K, m\rangle), \quad (74)$$

where  $p(J, K, \pm)$  is a phase factor chosen to provide a real matrix representation of the kinetic energy operator. The functions  $|J, K, m\rangle$  in Eqs. (73) and (74) are rigid-rotor eigenfunctions [12] with phases chosen as in Ref. [23]. The matrix elements  $\langle J, K', m | \hat{J}_x | J, K'', m \rangle$  and  $\langle J, K', m | \hat{J}_x \hat{J}_\beta | J, K'', m \rangle$  are obtained analytically [12,23]. These matrix elements are required, together with the vibrational matrix elements of Eq. (67), to generate the matrix representation of the Hamiltonian  $\hat{H}_{rv}$ .

## 7. Computational scheme

In the preceding sections, we have outlined how we obtain, for an arbitrary molecule, the rotation–vibration Hamiltonian in the form of the expansion in Eq. (59) and how we calculate the matrix elements of this Hamiltonian in a basis of functions

$$|v, J, K, m, \pm\rangle = \prod_v |v_v\rangle \times |J, K, m, \pm\rangle, \quad (75)$$

where the vibrational basis functions  $|v_v\rangle$  are defined in connection with Eq. (60) and the rotational basis functions  $|J, K, m, \pm\rangle$  are defined in Eqs. (73) and (74). The rotation–vibration energies and wavefunctions can now be

computed in a standard variational calculation, i.e., by diagonalizing the (truncated) matrix representation of the Hamiltonian that we have constructed.

The Hamiltonian matrix block is diagonal in  $J$ , and so each  $J$  block can be diagonalized separately. The truncation of the matrix blocks defines an active space for the calculation. This can be done different ways:

- (1) The simplest way is to set an upper limit for each of the principal quantum numbers defining the vibrational basis set [Eq. (60)]

$$v_n \leq v_n^{\max}, \quad n = 1, 2, \dots, 3N - 6. \quad (76)$$

This produces a  $(3N - 6)$ D ‘hyperbox’ in the space of the quantum numbers  $(v_1, v_2, \dots, v_{3N-6})$  with the number of basis functions as a volume  $v_1^{\max} \times v_2^{\max} \times \dots \times v_{3N-6}^{\max}$ . In most applications this value is too large for the computer capacity available, and it has to be reduced.

- (2) An alternative truncation employs the polyad number (see, for example, [24–26] and the references cited therein)

$$P = \sum_n a_n v_n, \quad (77)$$

where  $n$  runs over all modes,  $v_n$  are the quantum numbers, and  $a_n$  are the ‘resonance’ coefficients. The vibrational energy spectra of many molecules exhibit a polyad structure (see, for example, Ref. [12]), and the idea is to use this fact for defining the basis set truncation. The condition

$$P = \sum_n a_n v_n \leq P_{\max} \quad (78)$$

cuts the corners off the  $(3N - 6)$ D hyperbox mentioned above. This is a simple and straightforward way to control the size of the basis set, which we adopted in several applications [10,15,27–29]. The resonance coefficients  $a_i$  can be estimated from the harmonic frequencies by

$$a_n \approx \omega_n / \omega_0, \quad n = 1, 2, \dots, 3N - 6,$$

where  $\omega_0$  is smallest  $\omega_n$ -value for the molecule. We describe below an application to the  $\text{H}_2\text{S}$  molecule, for which  $\omega_1 \approx \omega_3 \approx 2\omega_2$  so that  $a_1 = a_3 = 2$  and  $a_2 = 1$ . Thus, for  $\text{H}_2\text{S}$  we can truncate the basis set by the condition  $P = 2(v_1 + v_2) + v_3 \leq P_{\max}$ . Non-integer  $a_i$  values can be also used.

- (3) There are molecules whose vibrational energy spectra do not have polyad structures. For such molecules, we may employ a third truncation mechanism which limits the basis set size using energy considerations [30,31]. We require that the active space provides the energy values up to a limit of  $E_{\max}$ . The problem with determining the corresponding active space is, of course, that the energies and their quantum number assignment are not known until we have done the

calculation. However this information can be ‘guessed’ by means of the following simple arguments. All basis functions that we employ for matrix element calculations are eigenfunctions of 1D Hamiltonians [see Eq. (62)], one for each vibrational mode. Consequently, we can approximate the solution of the final Schrödinger equation by

$$E_{v_1, v_2, \dots, v_{(3N-6)}}^{\text{approx.}} = E_{v_1} + E_{v_2} + E_{v_3} + \dots + E_{v_{(3N-6)}}, \quad (79)$$

where  $E_{v_n}$  is an eigenvalue of the 1D Hamiltonian  $\hat{H}_n$  from Eq. (62). Thus, we can construct the active space by selecting only those quantum basis functions that provide  $E_{v_1, v_2, \dots, v_{(3N-6)}}^{\text{approx.}} \leq E_{\text{max}}$ . This is a fairly sophisticated way of choosing the basis set, which however can be fully automated.

In actual applications, we can apply all three truncation schemes simultaneously and choose the smallest resulting active space. Obviously, the chosen basis sets must be subjected to convergence tests: the final energies must always be checked against the basis set size for convergence.

A technical detail should still be mentioned in connection with basis set truncation. For computational efficiency, we should ensure that the sizes of the Hamiltonian matrix blocks to be diagonalized are such that they, and their eigenvectors, can be stored in the available computer memory. For the H<sub>2</sub>S example discussed below, this requirement is easily met: Truncation by the polyad condition  $2(v_1 + v_2) + v_3 \leq 16$  produces a  $J = 10$  matrix block of dimension 4431. It takes up about 150 MB at double precision, which is perfectly affordable on present-day computers.

We have written the FORTRAN program TROVE (Theoretical ROVibrational Energies) to carry out the computations described above. The program is designed to calculate the rovibrational energies for an arbitrary polyatomic molecule in an isolated electronic state. Arbitrary internal coordinates can be used to describe the vibrational motion, and arbitrary vibrational basis functions can be defined for setting up the matrix representation of the rovibrational Hamiltonian. This generality is achieved by using the methods of object-orienting programming. Logically the program can be viewed as having two levels: the abstract level and the level of the specifications. The abstract level forms the core of the program and operates the coordinate transformations, the basis sets, the matrix elements, and the kinetic and potential energy expansions as abstract objects. Their specifications are hidden in exchangeable program units that are completely separate from the program core. The program can be thought as a computer language allowing one to create an arbitrary ‘project’ using predefined program modules. The specifications of the project are defined in a separate unit; they can be changed without changing the program core.

## 8. Applications

As an example of a rotation–vibration calculation, we apply the TROVE program to H<sub>2</sub>S in its electronic ground state. H<sub>2</sub>S is a relatively simple molecule for which both a large amount of experimental data [32,33] and high-level potential energy surfaces (e.g., from Ref. [34]) are available. We compare our results with those from the program DVR3D [35], designed for simulating spectra of triatomic molecules. DVR3D employs a nuclear kinetic energy operator exact within the Born–Oppenheimer approximation (customarily called an ‘Exact Kinetic Energy,’ or EKE, operator) and obtains the rotation–vibration energies in a discrete variable representation (DVR) approach. Thus, the philosophy behind DVR3D is rather different from that behind TROVE, and so a comparison between the results obtained with the two programs provides a stringent test of the validity of the approximations made in the TROVE model. In the comparisons, we can consider the DVR3D results to be exact within the Born–Oppenheimer approximation (see, for example, discussion in Ref. [34]). The potential energy function for H<sub>2</sub>S used in the present work is taken from Ref. [34].

For the TROVE calculations, the following steps are necessary:

- (1) Define the molecular structure in  $Z$ -matrix form, see Ref. [18]. When we do not employ HBJ theory, the  $Z$  matrix provides the equilibrium geometry (in the case of H<sub>2</sub>S, the equilibrium bond length and bond angle). When using the HBJ approach, the  $Z$  matrix defines a reference structure that also appears in the definition of the internal coordinates  $\xi_n$ .
- (2) Choose the basic coordinate type, i.e., choose between linearized and geometrically defined coordinates. Specify whether HBJ theory is to be used, and for which vibrational mode. For standard calculations on H<sub>2</sub>S, we choose the linearized coordinates  $r_1^\ell$ ,  $r_2^\ell$ , and  $\alpha^\ell$ , corresponding to the two H–S bond lengths and the bond angle  $\alpha = \angle(\text{H–S–H})$ , as defined in Section 3.1. In an HBJ-type calculation, the large-amplitude vibrational coordinate is  $\rho = \alpha$ .
- (3) Specify the internal coordinates  $\xi_n$ . The default choice corresponds to displacements of the basic coordinates from their equilibrium values. For H<sub>2</sub>S, this gives  $\xi_1 = r_1^\ell - r_e$ ,  $\xi_2 = r_2^\ell - r_e$ , and  $\xi_3 = \alpha^\ell - \alpha_e$ . Another obvious choice is to use the symmetrized combinations

$$\begin{aligned} S_1 &= \frac{1}{2} ([r_1^\ell - r_e] + [r_2^\ell - r_e]), \\ S_3 &= \frac{1}{2} (r_1^\ell - r_2^\ell), \\ S_2 &= \alpha^\ell - \alpha_e. \end{aligned} \quad (80)$$

In general, we can use any coordinates that can be related analytically to the basic ones. The options

- mentioned here are already implemented in the program, but any exotic coordinates can be readily defined.
- (4) Choose the variables  $g_n$  and  $f_n$  for the expansions of the kinetic and potential operators, respectively. For  $\text{H}_2\text{S}$  we employ  $g_n = \xi_n$  ( $n = 1, 2, 3$ ),  $f_n(\xi_n) = 1 - \exp(-a(r_n^\ell - r_e))$  ( $n = 1, 2$ ) and  $f_3(\xi_3) = \cos(\alpha^\ell) - \cos(\alpha_e)$ . In the present implementation, we can use different expansion coordinates (i.e.,  $g_n \neq f_n$ ) only in conjunction with numerically generated basis functions. In this case, all matrix elements are computed numerically, and so the expansion variables  $g_n$  and  $f_n$  can be arbitrarily chosen. With analytical (harmonic-oscillator or Morse-oscillator) basis functions and analytically derived matrix elements, we must use  $g_n = f_n$ .
  - (5) Choose the basis set. As described above, we can either use numerically generated 1D basis functions or one of three types of analytical 1D functions: (a) harmonic-oscillator functions, (b) Morse-oscillator functions, or (c) spherical harmonic functions. Additional analytical basis functions can be defined by adding to the program corresponding modules; such extensions do not require the core part of the program to be changed. The numerical basis functions are more flexible and capable of adapting to the molecule considered: these functions are obtained by solving a 1D Schrödinger equation for the degree of freedom in question. For  $\text{H}_2\text{S}$ , we employ numerically generated basis functions for all three vibrational modes. The two stretching vibrations are equivalent, therefore the corresponding basis functions must be generated only once.
  - (6) Choose the values of the parameters  $N_{\text{kin}}$  and  $N_{\text{pot}}$  (i.e., the orders of the polynomial expansions representing the kinetic and potential energy operators, respectively) entering into Eqs. (35), (37), and (58). The simplest zero-order approximation corresponds to  $N_{\text{kin}} = 0$  and  $N_{\text{pot}} = 2$ . For  $\text{H}_2\text{S}$ , we use  $N_{\text{kin}} = 8$  and  $N_{\text{pot}} = 10$ .
  - (7) Define the active space by any, or all, of the three schemes described in Section 7. For  $\text{H}_2\text{S}$  we use the polyad condition  $P = 2(v_1 + v_2) + v_3 \leq 20$ .
  - (8) Define the potential energy function. The potential energy function used in a given calculation is implemented in the form of a FORTRAN subroutine providing potential energy values at input values of geometrically defined coordinates. Geometrically defined coordinates are required as input to ensure that the potential energy subroutines are compatible with the core part of the program. The potential energy values are computed at points required by the finite-difference scheme applied. They are used to construct the expansion coefficients  $V_{l_1 l_2 l_3 \dots}$  from Eq. (58). For  $\text{H}_2\text{S}$ , we have implemented the MOR-BID expansion [36] of the potential energy function, which we use in conjunction with the parameter values from Ref. [34].

Table 1 reports vibrational term values of  $\text{H}_2\text{S}$  obtained with TROVE as described above. In the TROVE calculations, all finite-difference differentiations are done at quadruple precision with a two-point central formula and a step size of 0.001 Å. The results of a TROVE calculation made without the use of HBJ theory are given in the column headed ‘Standard.’ These results are compared to experimental values collected in Ref. [37], and to results obtained with the DVR3D program suite [35] with the same potential energy function (column headed ‘DVR3D’). The TROVE results are in very close agreement with those obtained with DVR3D, the root-mean-square (rms) deviation is only 0.041  $\text{cm}^{-1}$  for the 29 term values collected in Table 1. We have repeated the  $\text{H}_2\text{S}$  calculation using HBJ theory. The bond angle  $\alpha$  is taken as the large-amplitude vibrational coordinate called  $\rho$  in Section 3.3. We use Numerov-Cooley-generated bending basis functions. The results are given in the column headed ‘HBJ’ in Table 1. In relation to the DVR3D results, there is an improvement over the ‘Standard’ results, the rms deviation is now 0.027  $\text{cm}^{-1}$ .

We have also tested for  $\text{H}_2\text{S}$  the normal-coordinate Hamiltonian representation in conjunction with a harmonic-oscillator basis set. In this case  $g_n$  and  $f_n$  are defined automatically as normal coordinates (see Section 3). All other options correspond to the calculations already described. The calculated vibrational term values are shown in the column headed ‘Harm.’ in Table 1. Here, the agreement with DVR3D is good only for the lowest levels, it deteriorates with increasing excitation. Apparently the harmonic-oscillator basis functions require a larger basis set to reach convergence for all term values listed in Table 1.

As a further test calculation, we have used TROVE to calculate the vibrational energies of  $\text{CH}_3^+$ . Here, we also used HBJ theory, taking the inversion motion of  $\text{CH}_3^+$  to be the large-amplitude vibration. We employed linearized coordinates as done in our previous work on  $\text{NH}_3$  [10,15] with three stretching variables

$$\xi_k = r_k^\ell - r_e, \quad k = 1, 2, 3, \quad (81)$$

two symmetrized bending variables

$$\xi_4 = S_{4a} = \frac{1}{\sqrt{6}}(2\alpha_1^\ell - \alpha_2^\ell - \alpha_3^\ell), \quad (82)$$

$$\xi_5 = S_{4b} = \frac{1}{\sqrt{2}}(\alpha_2^\ell - \alpha_3^\ell), \quad (83)$$

and an out-of-plane-vibration coordinate

$$\xi_6 = \rho \equiv \frac{\mathbf{r}_1 \cdot (\mathbf{r}_2 \times \mathbf{r}_3)}{r_1 r_2 r_3}, \quad (84)$$

where  $\mathbf{r}_i$ ,  $i = 1, 2, 3$ , is the vector connecting the C nucleus with the proton  $i$ ,  $r_i = |\mathbf{r}_i|$ , and the bond angle  $\alpha_i = \angle(\text{H}_j\text{--C--H}_k)$ , where  $(i, j, k)$  is a permutation of the numbers (1, 2, 3). In the potential function expansion we used the Morse coordinates  $f_n = 1 - \exp(-a\xi_n)$  for the three stretch-

Table 1

Vibrational term values of H<sub>2</sub>S (in cm<sup>-1</sup>) corresponding to the eight lowest polyads in the H<sub>2</sub>S vibrational energy spectrum

$v_1v_2v_3$	Obs. <sup>a</sup>	DVR3D	Standard	$\Delta^b$	HBJ	$\Delta^c$	Harm.	$\Delta^d$
010	1182.57	1182.893	1182.894	-0.001	1182.893	0.000	1182.902	-0.009
020	2353.96	2354.126	2354.129	-0.004	2354.126	-0.001	2354.150	-0.024
100	2614.14	2614.359	2614.360	-0.001	2614.363	-0.005	2614.366	-0.002
001	2628.45	2628.764	2628.766	-0.001	2628.770	-0.005	2628.773	-0.003
030	3513.79	3513.801	3513.811	-0.010	3513.802	-0.001	3513.847	-0.045
110	3779.17	3778.998	3779.003	-0.005	3779.007	-0.009	3779.028	-0.021
011	3789.27	3789.137	3789.142	-0.005	3789.146	-0.009	3789.168	-0.022
040	4661.68	4661.522	4661.543	-0.021	4661.524	-0.002	4661.595	-0.071
120	4932.70	4932.395	4932.409	-0.014	4932.409	-0.014	4932.447	-0.037
021	4939.10	4938.603	4938.621	-0.018	4938.618	-0.015	4938.656	-0.038
200	5144.99	5145.243	5145.246	-0.003	5145.261	-0.019	5145.366	-0.104
101	5147.22	5147.216	5147.217	-0.001	5147.235	-0.019	5147.344	-0.109
002	5243.10	5243.476	5243.482	-0.005	5243.487	-0.011	5243.513	-0.026
050	5797.24	5796.561	5796.592	-0.030	5796.565	-0.004	5796.647	-0.082
130	6074.58	6074.625	6074.678	-0.053	6074.647	-0.021	6074.729	-0.083
031	6077.60	6077.196	6077.259	-0.063	6077.219	-0.022	6077.296	-0.078
210	6288.15	6287.832	6287.852	-0.020	6287.879	-0.047	6288.197	-0.318
111	6289.17	6288.627	6288.640	-0.013	6288.674	-0.047	6289.006	-0.332
012	6388.10	6384.725	6384.744	-0.019	6384.744	-0.019	6384.838	-0.094
060		6917.964	6917.902	0.062	6917.969	-0.005	6917.935	0.034
140		7204.448	7204.630	-0.182	7204.479	-0.032	7204.586	-0.107
041		7205.251	7205.434	-0.183	7205.282	-0.030	7205.426	-0.145
121	7420.09	7419.543	7419.599	-0.056	7419.625	-0.082	7419.969	-0.345
022		7419.555	7419.607	-0.052	7419.637	-0.082	7420.043	-0.406
220		7515.895	7515.952	-0.057	7515.925	-0.030	7516.053	-0.128
102	7576.38	7576.367	7576.350	0.017	7576.408	-0.041	7579.332	-2.925
201	7576.54	7576.426	7576.407	0.019	7576.467	-0.041	7579.377	-2.910
300	7752.26	7752.564	7752.578	-0.015	7752.587	-0.023	7753.118	-0.531
003	7779.32	7779.713	7779.726	-0.013	7779.741	-0.028	7780.069	-0.328

<sup>a</sup> Experimentally derived values collected in Ref. [37].<sup>b</sup>  $E(\text{DVR3D}) - E(\text{Standard})$  in cm<sup>-1</sup>.<sup>c</sup>  $E(\text{DVR3D}) - E(\text{HBJ})$  in cm<sup>-1</sup>.<sup>d</sup>  $E(\text{DVR3D}) - E(\text{Harm.})$  in cm<sup>-1</sup>.

ing modes and  $f_4 = S_{4a}$ ,  $f_5 = S_{4b}$  for the two bending modes. The sixth coordinate  $\xi_6 = \rho$  is a large-amplitude coordinate, which defines the starting point of the expansions of the kinetic and potential energy operators. For all modes we employed numerically defined basis sets, and so we could use a kinetic energy expansion different from the potential energy expansion. All five expansion variables  $g_n$  were taken to be equal to the respective  $\xi_n$  coordinate. The kinetic and potential energy expansion were truncated at  $N_{\text{kin}} = 6$  and  $N_{\text{pot}} = 6$ , respectively. The basis set was controlled by the polyad truncation scheme with  $P = 1.5(v_1 + v_2 + v_3) + v_4 + v_5 + v_6 \leq 18$ , where each of the principal quantum numbers  $v_1$ ,  $v_2$ ,  $v_3$ ,  $v_4$ ,  $v_5$ , and  $v_6$  corresponds to a  $\xi_n$  coordinate. The final Hamiltonian matrix was separated into two blocks according to the parity of the inversion basis functions [12]. This led to two matrices of dimensions 10478 and 8033, respectively, that had to be diagonalized. The potential energy function was taken from Ref. [38]. In Table 2, we compare our results (column headed ‘TROVE’) with theoretical term values from Ref. [38] computed with the TetraVib program [39] (column headed ‘TetraVib’) utilizing the same potential energy function. The agreement is reasonable with an rms deviation of only 0.1 cm<sup>-1</sup> for the term values

collected in Table 2, which is in accord with the accuracy of the TetraVib method (see Ref. [39]).

Finally, we consider the rather trivial example of a diatomic molecule. We have calculated the vibrational term values of the <sup>12</sup>C<sup>16</sup>O molecule from the empirical Simons–Parr–Finlan potential curve from Ref. [40]. We construct and solve the Schrödinger equation using HBJ theory. The large-amplitude coordinate  $\rho$  is the internuclear distance  $r$  which is the only extant vibrational coordinate. The corresponding 1D Schrödinger equation is solved by the Numerov–Cooley technique, and we generate the 30 lowest eigenvalues and list them in the column headed ‘HBJ’ in Table 3. The agreement with the RKR energies from Ref. [41] is excellent for all reported decimal digits. This documents the quality of our 1D Numerov–Cooley solver, i.e., the tool that we use for constructing the numerically defined basis sets.

The CO problem can be solved in many other ways, for example, using a Dunham expansion of the potential energy (i.e., a standard Taylor series in  $r - r_e$ ) in conjunction with a variational solution of the Schrödinger equation by means of a harmonic-oscillator basis set. In this case, the coordinate expansions are derived by means of the recursive solution of Eqs. (23) and (24). In the diago-

Table 2  
Vibrational term values of CH<sub>3</sub><sup>+</sup> (in cm<sup>-1</sup>)

State	± <sup>a</sup>	TetraVib	TROVE	Δ <sup>b</sup>
v <sub>2</sub>	–	1377.82	1377.85	–0.03
v <sub>4</sub> <sup>1</sup>	+	1387.01	1387.04	–0.03
2v <sub>2</sub>	+	2723.02	2723.11	–0.09
2v <sub>4</sub> <sup>2</sup>	+	2770.85	2770.94	–0.09
v <sub>2</sub> + v <sub>4</sub> <sup>1</sup>	–	2772.77	2772.82	–0.05
2v <sub>4</sub> <sup>0</sup>	+	2780.18	2780.26	–0.08
v <sub>1</sub>	+	2942.27	2942.37	–0.10
v <sub>3</sub> <sup>1</sup>	+	3108.29	3108.38	–0.09
3v <sub>2</sub>	–	4071.30	4071.40	–0.10
2v <sub>2</sub> + v <sub>4</sub> <sup>1</sup>	+	4101.18	4101.57	–0.39
3v <sub>4</sub> <sup>3</sup>	+	4150.94	4151.26	–0.32
3v <sub>4</sub> <sup>3</sup>	+	4151.60	4151.88	–0.28
v <sub>2</sub> + 2v <sub>4</sub> <sup>2</sup>	–	4164.44	4164.53	–0.09
3v <sub>4</sub> <sup>1</sup>	+	4177.99	4178.19	–0.20
v <sub>2</sub> + 2v <sub>4</sub> <sup>0</sup>	–	4180.55	4180.62	–0.07
v <sub>1</sub> + v <sub>2</sub>	–	4324.30	4324.45	–0.15
v <sub>1</sub> + v <sub>4</sub> <sup>1</sup>	+	4342.41	4342.56	–0.15
v <sub>2</sub> + v <sub>3</sub> <sup>1</sup>	–	4470.42	4470.56	–0.14
v <sub>3</sub> <sup>1</sup> + v <sub>4</sub> <sup>1</sup>	+	4490.61	4490.79	–0.18
v <sub>3</sub> <sup>1</sup> + v <sub>4</sub> <sup>1</sup>	+	4503.18	4503.34	–0.16
v <sub>3</sub> <sup>1</sup> + v <sub>4</sub> <sup>1</sup>	+	4523.42	4523.57	–0.15
4v <sub>2</sub>	+	5401.36	5401.71	–0.35
3v <sub>2</sub> + v <sub>4</sub> <sup>1</sup>	–	5456.46	5456.75	–0.29
2v <sub>2</sub> + 2v <sub>4</sub> <sup>2</sup>	+	5474.22	5475.59	–1.37
2v <sub>2</sub> + 2v <sub>4</sub> <sup>0</sup>	+	5478.49	5480.11	–1.62
4v <sub>4</sub> <sup>4</sup>	+	5528.01	5528.84	–0.83
v <sub>2</sub> + 3v <sub>4</sub> <sup>3</sup>	–	5552.23	5552.50	–0.27
v <sub>2</sub> + 3v <sub>4</sub> <sup>3</sup>	–	5553.01	5553.19	–0.18
4v <sub>4</sub> <sup>2</sup>	+	5575.37	5575.78	–0.41
v <sub>2</sub> + 3v <sub>4</sub> <sup>1</sup>	–	5588.36	5588.49	–0.13

<sup>a</sup> Parity of the vibrational state [12].

<sup>b</sup>  $E(\text{TetraVib}) - E(\text{TROVE})$  in cm<sup>-1</sup>.

nalization step, we used 50 harmonic-oscillator eigenfunctions as basis. The calculated term values are listed in the column headed ‘Dunham’ in Table 3. The agreement with experiment is good for the lower states only; this is a consequence of the poor convergence of the re-expansion of the SPF potential energy function as a 12th order Taylor series (we used  $N_{\text{pot}} = 12$  in Eq. (58)) which diverges at larger values of  $|r - r_{\text{el}}|$ . A more accurate variational solution can be obtained in terms of Morse variables and Morse-oscillator basis functions (see the column headed ‘Morse’ of Table 3); the Morse expansion obviously provides a better representation of the CO potential energy function.

Thus far, we have presented TROVE results only for vibrational energies. However, as indicated in Section 6, TROVE also treats rotation and can generate rotation–vibration energies. In the program code, the rotational degrees of freedom are treated in parallel to the vibrational degrees; the basis functions and the matrix elements associated with them (i.e., the rigid-rotor eigenfunctions and matrix elements of Section 6) are specified in exchangeable program modules. The program automatically generates the expansion coefficients of the rotational and Coriolis–interaction terms in the kinetic energy operator  $G_{\lambda,\lambda'}$  in the recursive procedure, along with the expansion coefficients for the vibrational terms. We have used TROVE to calculate the rotational energies of H<sub>2</sub>S for  $J \leq 10$  and compare the

results with those from the program DVR3D. The differences for selected term values (belonging to states with  $J = 1, 2, 3, 4, 10$  below 4500 cm<sup>-1</sup>) are plotted against the term values in Fig. 1. The rms deviation between these two data sets is 0.005 cm<sup>-1</sup>, comparable to the agreement obtained in the comparison between the vibrational energies.

## 9. Convergence properties

In the present work, we approximate the  $\xi_n$ -dependent functions  $G_{\lambda,\lambda'}(\xi)$  and  $U(\xi)$  [Eq. (5)], and the Born–Oppenheimer potential energy function  $V$  [Eq. (1)], as finite-order expansions in the  $\xi_n$  coordinates. Owing to the approximate expressions used for  $G_{\lambda,\lambda'}(\xi)$  and  $U(\xi)$ , the resulting kinetic energy operator is not exact within the Born–Oppenheimer approximation. For example, with the parameter  $N_{\text{kin}} = 2$ , the expansions of  $G_{\lambda,\lambda'}$  are truncated after the quadratic terms  $\xi_i \xi_j$ . Obviously it is important to investigate the error introduced by the truncation of the kinetic energy expansion. Towards this end, we have performed a number of tests for H<sub>2</sub>S, calculating the vibrational energies for different values of  $N_{\text{kin}}$ . The results for  $N_{\text{kin}} = 4, 6, 8$ , and 10 are shown in Fig. 2. The differences between the results for  $N_{\text{kin}} = 8$  and 10 indicate that the calculations are basically completely converged for  $N_{\text{kin}} = 8$ .

One may ask whether the convergence of the energies implies that the approximate kinetic energy operator used in the present work actually converges to such an extent that it becomes ‘exact.’ The agreement of our results with the EKE ones (Table 1) suggests that this could indeed be the case.

Other convergence tests pertain to the basis set size. We have stated above that the polyad number truncation, as defined by Eq. (78), provides a reasonable way of controlling the size of the active space. We show in Fig. 3 the residuals ( $E(\text{DVR3D}) - E(\text{Standard})$ ) for the vibrational states of H<sub>2</sub>S included in Table 1, determined for different values of  $P_{\text{max}}$ . The almost linear dependence of the residuals on  $P_{\text{max}}$  in the logarithmic plot of Fig. 3 suggests that the convergence is exponential. Thus, we can use an extrapolation technique known from *ab initio* electronic structure calculations to reach the ‘Complete Basis Set’ (CBS) limit. For example, an exponential expression

$$E_i(P_{\text{max}}) = E_i^\infty + b_i \exp(-P_{\text{max}}/t_i) \quad (85)$$

seems to be appropriate for this purpose. In Eq. (85),  $P_{\text{max}}$  denotes the polyad truncation number while  $E_i^\infty$ ,  $b_i$ , and  $t_i$  are fitting parameters, and  $i$  is a short-hand notation for the vibrational quantum numbers, i.e.,  $(v_1, v_2, v_3)$  in the case of H<sub>2</sub>S. Clearly  $\lim_{P_{\text{max}} \rightarrow \infty} E_i(P_{\text{max}}) = E_i^\infty$  and so  $E_i^\infty$  is the CBS-extrapolated value.

## 10. Summary and conclusions

We report here a theoretical model, with associated computer program TROVE, for calculating the rotation–vibration energies of di- and polyatomic molecules in

Table 3  
Vibrational term values of  $^{12}\text{C}^{16}\text{O}$  (in  $\text{cm}^{-1}$ )

$v$	Obs. <sup>a</sup>	HBJ	O–C <sup>b</sup>	Dunham	O–C <sup>c</sup>	Morse	O–C <sup>d</sup>
0	1081.7764	1081.7764	0.0000	1081.7764	0.0000	1081.7778	–0.0014
1	3225.0478	3225.0478	0.0000	3225.0478	0.0000	3225.0490	–0.0012
2	5341.8389	5341.8390	–0.0001	5341.8390	–0.0001	5341.8400	–0.0010
3	7432.2158	7432.2160	–0.0002	7432.2160	–0.0002	7432.2168	–0.0008
4	9496.2462	9496.2464	–0.0002	9496.2466	–0.0004	9496.2470	–0.0006
5	11533.9992	11533.9994	–0.0002	11534.0003	–0.0011	11534.0000	–0.0006
6	13545.5458	13545.5460	–0.0002	13545.5494	–0.0036	13545.5468	–0.0008
7	15530.9588	15530.9589	–0.0001	15530.9699	–0.0111	15530.9605	–0.0016
8	17490.3123	17490.3124	–0.0001	17490.3448	–0.0325	17490.3160	–0.0036
9	19423.6823	19423.6824	–0.0001	19423.7691	–0.0868	19423.6904	–0.0080
10	21331.1464	21331.1464	0.0000	21331.3612	–0.2148	21331.1629	–0.0165
11	23212.7833	23212.7834	–0.0001	23213.2820	–0.4987	23212.8155	–0.0321
12	25068.6734	25068.6736	–0.0002	25069.7682	–1.0948	25068.7328	–0.0592
13	26898.8983	26898.8985	–0.0002	26901.1878	–2.2895	26899.0030	–0.1045
14	28703.5405	28703.5409	–0.0004	28708.1334	–4.5929	28703.7181	–0.1773
15	30482.6835	30482.6840	–0.0005	30491.6172	–8.9337	30482.9750	–0.2909
16	32236.4115	32236.4122	–0.0007	32253.7340	–17.3225	32236.8759	–0.4637
17	33964.8091	33964.8101	–0.0010	34000.2735	–35.4644	33965.5303	–0.7202
18	35667.9611	35667.9626	–0.0015	35747.7314	–79.7703	35669.0558	–1.0932
19	37345.9524	37345.9545	–0.0021	37530.3879	–184.4355	37347.5801	–1.6256
20	38998.8673	38998.8704	–0.0031	39391.4064	–392.5391	39001.2434	–2.3730
21	40626.7898	40626.7942	–0.0044	41361.3491	–734.5593	40630.2004	–3.4062
22	42229.8026	42229.8087	–0.0061	43451.2911	–1221.488	42234.6247	–4.8160

<sup>a</sup> Experimentally derived term values [41].

<sup>b</sup>  $E(\text{Obs.}) - E(\text{HBJ})$  in  $\text{cm}^{-1}$ .

<sup>c</sup>  $E(\text{Obs.}) - E(\text{Dunham})$  in  $\text{cm}^{-1}$ .

<sup>d</sup>  $E(\text{Obs.}) - E(\text{Morse})$  in  $\text{cm}^{-1}$ .

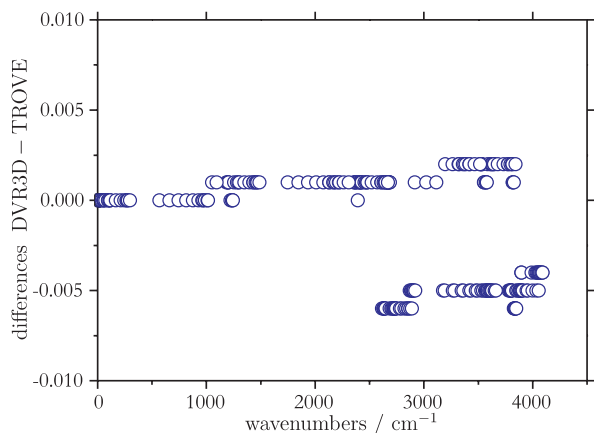


Fig. 1. The differences ( $\text{cm}^{-1}$ ) between the rotation–vibration term values of  $\text{H}_2\text{S}$  calculated using DVR3D and TROVE, plotted against the term values themselves (see text).

isolated electronic states. TROVE is conceived as an ‘expert-system’ program: when we specify the coordinates that we wish to use for describing the vibrational motion and define suitable basis functions depending on these coordinates, TROVE will construct the rotation–vibrational Hamiltonian for the molecule, expressed as a truncated power series in the vibrational coordinates (or in suitably chosen functions of these parameters), generate the matrix representation of the Hamiltonian in terms of the chosen basis functions, and diagonalize the resulting matrix blocks to obtain the rotation–vibration energies

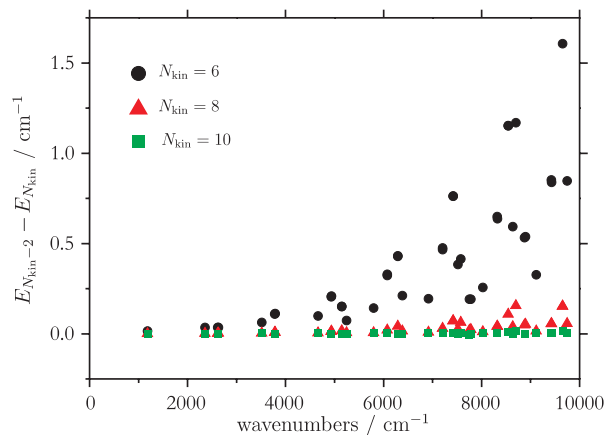


Fig. 2. The effect of the kinetic energy truncation on the vibrational term values of  $\text{H}_2\text{S}$  corresponding to the six lowest polyads in the  $\text{H}_2\text{S}$  vibrational energy spectrum, plotted against the term values themselves. The circles represent the differences  $E_4 - E_6$  between term values calculated for  $N_{\text{kin}} = 4$  and 6, respectively. The triangles and squares show the analogous differences  $E_6 - E_8$  and  $E_8 - E_{10}$ .

and wavefunctions. The high flexibility of TROVE is achieved by using numerical methods, in particular finite-difference differentiation, rather than analytical derivations, whenever possible. We have compared TROVE-calculated energies with experimental data and energies calculated with other theoretical methods for the molecules  $\text{H}_2\text{S}$ ,  $\text{CH}_3^+$ , and  $\text{CO}$ , and the results are extremely encouraging. In particular the comparison, for  $\text{H}_2\text{S}$ , with energies

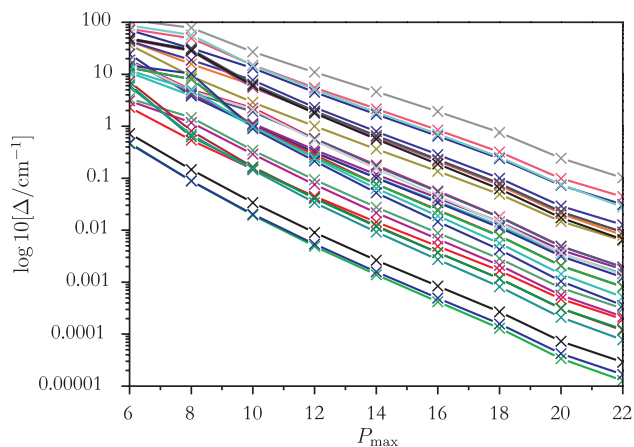


Fig. 3. Residuals ( $\Delta = E(\text{DVR3D}) - E_i^{P_{\max}}$ ) for the vibrational states of  $\text{H}_2\text{S}$  included in Table 1, determined with increasing basis sets defined by  $P_{\max} = 8, 10, \dots, 22$  (see text). The lowest line corresponds to the (0, 1, 0) vibrational level, the second lowest to (0, 2, 0) and so on.

computed with the program DVR3D [35] is very satisfactory in view of the fact that DVR3D is based on ideas very different from those behind TROVE: the kinetic energy operator used in DVR3D is exact within the Born–Oppenheimer approximation, and the rotation–vibration energies are obtained in the DVR approach.

There are still a few odds and ends in TROVE that need attending to. TROVE cannot, at the moment, deal correctly with singularities of the rotation–vibration Hamiltonian. Such singularities are important for molecules that are linear at equilibrium, such as HCN and HCCH, or have a low barrier to linearity, such as  $\tilde{X}^3B_1 \text{CH}_2$ . It is discussed at great length in Ref. [22] how to treat these singularities in the Hougen–Bunker–Johns theory now implemented in TROVE, and the appropriate modifications of TROVE can be made. Also, at the moment TROVE does not make use of permutation–inversion symmetry [12] to block-diagonalize the matrix representation of the rotation–vibration Hamiltonian but work is in progress to implement it.

### Acknowledgments

We acknowledge support from the European Commission through contract no. MRTN-CT-2004-512202 “Quantitative Spectroscopy for Atmospheric and Astrophysical Research” (QUASAAR). The work of P.J. is supported in part by the Fonds der Chemischen Industrie.

### References

[1] A. Niño, C. Muñoz-Caro, *Comput. Chem.* 18 (1994) 27–32.

[2] D. Luckhaus, *J. Chem. Phys.* 106 (1997) 8409–8426.  
 [3] M.L. Senent, *Chem. Phys. Lett.* 296 (1998) 299–306.  
 [4] M.A. Harthcock, J. Laane, *J. Phys. Chem.* 89 (1985) 4231–4240.  
 [5] J.R. Durig, W. Zhao, *J. Phys. Chem.* 98 (1994) 9202–9206.  
 [6] D.J. Rush, K.B. Wiberg, *J. Phys. Chem. A* 101 (1997) 3143–3151.  
 [7] Y.G. Smeyers, F.J. Meléndez, M.L. Senent, *J. Chem. Phys.* 106 (1997) 1709–1717.  
 [8] D. Luckhaus, *J. Chem. Phys.* 113 (2000) 1329–1347.  
 [9] D. Lauvergnat, A. Nauts, *J. Chem. Phys.* 116 (2002) 8560–8570.  
 [10] S.N. Yurchenko, M. Carvajal, P. Jensen, H. Lin, J.J. Zheng, W. Thiel, *Mol. Phys.* 103 (2005) 359–378.  
 [11] C. Eckart, *Phys. Rev.* 47 (1935) 552–558.  
 [12] P.R. Bunker, P. Jensen, *Molecular Symmetry and Spectroscopy*, 2nd ed., NRC Research Press, Ottawa, 1998.  
 [13] C. Léonard, N.C. Handy, S. Carter, J.M. Bowman, *Spectrochim. Acta A* 58 (2002) 825–838.  
 [14] S.M. Colwell, S. Carter, N.C. Handy, *Mol. Phys.* 101 (2003) 523–544.  
 [15] S.N. Yurchenko, M. Carvajal, H. Lin, J.J. Zheng, W. Thiel, P. Jensen, *J. Chem. Phys.* 122 (2005) 104317.  
 [16] S.N. Yurchenko, W. Thiel, S. Patchkovskii, P. Jensen, *Phys. Chem. Chem. Phys.* 7 (2005) 573–582.  
 [17] G.O. Sørensen, in: M.J.S. Dewar et al. (Eds.), *Topics in Current Chemistry*, vol. 82, Springer-Verlag, Heidelberg, 1979, p. 99.  
 [18] W.J. Hehre, L. Radom, P.V.R. Schleyer, J.A. Pople, *Ab Initio Molecular Orbital Theory*, Wiley, New York, 1986.  
 [19] J.T. Hougen, P.R. Bunker, J.W.C. Johns, *J. Mol. Spectrosc.* 34 (1970) 136–172.  
 [20] A. Sayvetz, *J. Chem. Phys.* 7 (1939) 383–389.  
 [21] B. Numerov, *Publ. of the Central Astrophys. Observatory (Rus)* 2 (1923) 188–259;  
 J.W. Cooley, *Math. Comp.* 15 (1961) 363–374.  
 [22] P. Jensen, *Comp. Phys. Rep.* 1 (1983) 1–55.  
 [23] D. Papoušek, M.R. Aliev, *Molecular Vibrational/Rotational Spectra*, Elsevier, Amsterdam, 1982.  
 [24] M.E. Kellman, V. Tyng, *Acc. Chem. Res.* 40 (2007) 243–250.  
 [25] M.M. Sprague, S.G. Ramesh, E.L. Sibert, *J. Chem. Phys.* 124 (2006) 114307.  
 [26] S. Albert, H. Hollenstein, M. Quack, M. Willeke, *Mol. Phys.* 104 (2006) 2719–2735.  
 [27] S.N. Yurchenko, M. Carvajal, W. Thiel, P. Jensen, *J. Mol. Spectrosc.* 239 (2006) 71–87.  
 [28] S.N. Yurchenko, J. Breidung, W. Thiel, *Theor. Chem. Acc.* 114 (2005) 333–340.  
 [29] S.N. Yurchenko, M. Carvajal, W. Thiel, H. Lin, P. Jensen, *Adv. Quant. Chem.* 48 (2005) 209–238.  
 [30] H.-S. Lee, J.C. Light, *J. Chem. Phys.* 118 (2003) 3458–3469.  
 [31] R. Dawes, T. Carrington, *J. Chem. Phys.* 122 (2005) 134101.  
 [32] I.N. Kozin, P. Jensen, *J. Mol. Spectrosc.* 163 (1994) 483–509.  
 [33] J.-M. Flaud, O. Vaittinen, A. Campargue, *J. Mol. Spectrosc.* 190 (1998) 262–268.  
 [34] O.L. Polyansky, P. Jensen, J. Tennyson, *J. Mol. Spectrosc.* 178 (1996) 184–188.  
 [35] J. Tennyson, J.R. Henderson, N.G. Fulton, *Comp. Phys. Commun.* 86 (1995) 175–198.  
 [36] P. Jensen, *J. Mol. Spectrosc.* 128 (1988) 478–501.  
 [37] P. Barletta, A.G. Császár, H.M. Quiney, J. Tennyson, *Chem. Phys. Lett.* 361 (2002) 121–128.  
 [38] H.-G. Yu, T.J. Sears, *J. Chem. Phys.* 117 (2002) 666–669.  
 [39] H.-G. Yu, J.T. Muckerman, *J. Mol. Spectrosc.* 214 (2002) 1–20.  
 [40] K. Nakagawa, M. Akiyama, *Chem. Phys. Lett.* 190 (1992) 91–96.  
 [41] H. Telle, U. Tell, *J. Mol. Spectrosc.* 85 (1981) 248–252.