# Method Development in Ro-Vibrational Calculations of Polyatomic Molecules


**Thomas M. Mellor**

University College London

**Physics and Astronomy**


Submitted to University College London (UCL) in partial fulfilment of the requirements for the award of the degree of Doctor of Philosophy.

Primary supervisor: **Prof. Sergey Yurchenko**

Secondary supervisor: **Prof. Jonathan Tennyson**

Examining committee: **Dr. David Benoit and Prof. Graham Worth**

Thesis submission date: **January 20, 2023**

# Declaration

I, **Thomas M. Mellor**, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

. . . . . .
**Thomas M. Mellor**
**January 20, 2023**

# Publications

Publications relevant to this thesis:

- **Chapter 4**: T. M. Mellor, A. Owens, J. Tennyson, and S. N. Yurchenko. "MARVEL analysis of high-resolution spectra of thioformaldehyde ($H_2CS$)". In: *J. Mol. Spectrosc.* 391 (2023), p. 111732. DOI: 10.1016/j.jms.2022.111732
- **Chapter 4**: T. Mellor, A. Owens, J. Tennyson, and S. N. Yurchenko. "ExoMol line lists – XLVIII. High-temperature line list of thioformaldehyde ($H_2CS$)". In: *Mon. Not. R. Astron. Soc.* (2023). DOI: 10.1093/mnras/stad111
- **Chapter 5**: T. M. Mellor, S. N. Yurchenko, and P. Jensen. "Artificial Symmetries for Calculating Vibrational Energies of Linear Molecules". In: *Symmetry* 13.4 (2021). DOI: 10.3390/sym13040548
- **Chapter 6**: T. M. Mellor. "Molecular frames for a symmetry-adapted rotational basis set". In: *Mol. Phys.* 120.18 (2022), e2118638. DOI: 10.1080/00268976.2022.2118638
- **Chapter 7**: T. M. Mellor, S. N. Yurchenko, B. P. Mant, and P. Jensen. "Transformation Properties under the Operations of the Molecular Symmetry Groups $\boldsymbol{G}_{36}$ and $\boldsymbol{G}_{36}$(EM) of Ethane $H_3CCH_3$". In: *Symmetry* 11.862 (2019). DOI: 10.3390/sym11070862

Other publications:

- S. N. Yurchenko, T. M. Mellor, R. S. Freedman, and J. Tennyson. "ExoMol line lists – XXXIX. Ro-vibrational molecular line list for $CO_2$". In: *Mon. Not. R. Astron. Soc.* 496.4 (2020), pp. 5282–5291. DOI: 10.1093/mnras/staa1874
- S. N. Yurchenko and T. M. Mellor. "Treating linear molecules in calculations of rotation-vibration spectra". In: *J. Chem. Phys.* 153.15 (2020), p. 154106. DOI: 10.1063/5.0019546
- J. Tennyson, S. N. Yurchenko, A. F. Al-Refaie, V. H. J. Clark, K. L. Chubb, E. K. Conway, A. Dewan, M. N. Gorman, C. Hill, A. E. Lynas-Gray, T. Mellor, L. K. McKemmish, A. Owens, O. L. Polyansky, M. Semenov, W. Somogyi, G. Tinetti, A. Upadhyay, I. Waldmann, Y. Wang, S. Wright, and O. P. Yurchenko. "The 2020 release of the ExoMol database: Molecular line lists for exoplanet and other hot atmospheres". In: *J. Quant. Spectrosc. Radiat. Transf.* 255 (2020), p. 107228. DOI: 10.1016/j.jqsrt.2020.107228

# Abstract

Successfully computing the ro-vibrational spectrum of a polyatomic molecule requires consideration of several factors. Among them is the representation of the kinetic energy operator (KEO), the choice of body-fixed (BF) frame, and the use of the molecular symmetry group. The work detailed in this thesis develops all three and forms a part of the ExoMol project which is concerned with the calculation of molecular line lists of astronomical significance.

For the first factor, I enhanced one of the main programs of the ExoMol project, Trove, by enabling the use of externally programmed and analytic KEOs in variational calculations. I utilised this approach in the Marvelised line list calculation of $H_2CS$ which covers the $0\,cm^{-1}$ to $8000\,cm^{-1}$ range for states up to $J = 120$. I also collated all experimentally available transitions and, with the Marvel program, converted them to highly accurate experimental energy levels. The energies of the calculated line list were then replaced by Marvel energies when available.

The states of all such line lists in Trove have an assigned symmmetry label according to their appropriate molecular symmetry group. A robust symmetrisation procedure is one of the main features in Trove. I further exploited group theory by contructing and implementing an artificial group for use in Trove's $3N - 6$ approach for the variational calculations of linear molecules. This allowed much of the pre-existing infrastructure to be used with minimal changes.

An analytic KEO complicates matters by necessitating a BF-frame alternative to the usual Eckart frame. I elucidate the choice of BF-frames which permit rotational symmetrisation and suggest example alternative frames.

Finally, one of the suggested frames is used for the analytic KEO of $C_2H_6$. The preliminary work on this molecule is described, with a focus of my implementation of its molecular symmetry group in Trove.

# Acknowledgements

First and foremost I would like to thank my supervisor Sergey Yurchenko. Without his guidance and ever-present help I wouldn't have been able to finish my PhD. I would also like to thank Jonathan Tennyson and everyone in the ExoMol group for making the day-to-day experience truly delightful. I'm only sorry that half of it was marred by the pandemic.

I thank Barry Mant and Alec Owens for all their help on various projeccts throughout the entire degree and Bridgette Cooper whose programming skills unsurpassed in the group. The spectroscopy "Bible" Molecular symmetry and spectroscopy came in handy countless times and I was fortunate to work with late Per Jensen which I am grateful for. Mikhail Semenov and Jaya Chand were a pleasure to work with during a time when everyone was stuck at home and they helped make the ethane MARVEL work much more bearable.

Finally, I'd like to thank my family for their support and everyone else in UCL during my time there.

# Contents

# List of Figures

# List of Tables

# Chapter 1

# Introduction

Since the launch of the Kepler space telescope in 2009 [9], there has been an explosion in the number of verified planets beyond our solar system ("exoplanets") [10]. Characterising the atmosphere of an exoplanet – the molecular composition, the temperature, and the pressure, ideally as a function of the altitude – is of enormous interest to astronomers. This information can be inferred only from the spectrum produced when the light from the star interacts with the planet's atmosphere: roughly speaking, the proportion of light absorbed by the atmosphere as a function of the light's frequency. Astronomers use statistical methods to find the most likely atmospheric properties from the observed data [11].

The confidence in the predicted properties largely depends on the quality of the underlying molecular spectra. While laboratory measurements are immensely valuable, experiments could never provide all the necessary data due to the sheer quantity needed. They must be heavily supplemented by theoretical calculations. Although less accurate than experiment, theoretical calculations are easier and safer to perform.

Molecular spectroscopy – the study of the interaction of matter with electromagnetic (EM) radiation – is what allows us to simulate the spectra. As a molecule is a quantum mechanical system, its bound states have discrete energy levels. Transitions between energy levels are achieved through the absorption or the emission of a photon. The energy of this photon is given by $h\nu$ where $h$ is Planck's constant and $\nu$ the photon's frequency.

The complexity of the quantum systems, consisting of several interacting bodies, precludes any sort of analytic solution to the associated Schrödinger equation; numerical approaches are necessary. Still, performing the necessary calculations is a daunting task as one must solve a set of coupled second order differential equations. Thankfully, well-chosen assumptions can simplify things considerably. Of chief importance is to approximate a molecular state as a product of an electronic state, a rotational state of the nuclei, and a vibrational state of the nuclei. Transitions between molecular states are then combinations of transitions between the sub-states. The energy of the sub-state transition follows the inequality

$$E_{\text{rotational}} < E_{\text{vibrational}} < E_{\text{electronic}} \tag{1.0.1}$$

which, for polyatomic molecules, are in the microwave, infrared (IR), and ultraviolet regions of the EM spectrum, respectively. The work in this thesis is limited to IR energies

or below, so only ro-vibrational (rotational vibrational) transitions will be discussed. Thus, only the lowest lying electronic state is required. The rotational and vibrational states are further sub-classified by various labels (quantum numbers), both rigorous and approximate.

This thesis forms part of the ExoMol project, which is concerned with the production of calculated line lists of molecules of astronomical importance [12]. A line list contains information about the ro-vibrational states of a molecule, as well as transition information. Figure 1.1 shows the work flow for the generation of a generic line list in ExoMol. The electronic problem is solved first and the electronic state is represented by the potential energy surface (PES). Initially, this is purely theoretical, or *ab initio*, and, when added to the kinetic energy operator (KEO), is the ro-vibrational Hamiltonian is used to calculate the ro-vibrational states. For the ro-vibrational energies to be of the required spectroscopic accuracy, the PES is then further "refined" from the available experimental data which feeds back into the ro-vibrational state calculations. Calculating the transitions requires both the ro-vibrational wavefunctions and a dipole moment surface (DMS). An *ab initio* DMS is adequate in this case.

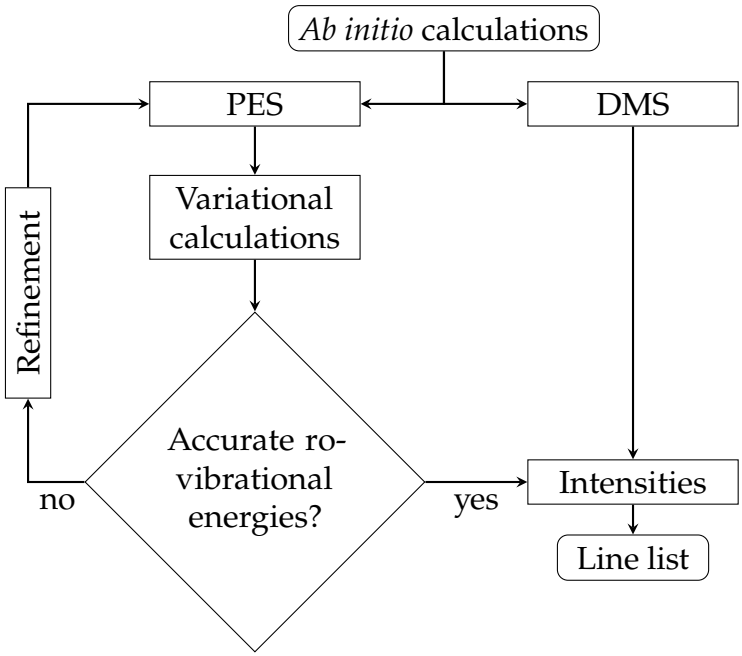



Figure 1.1: The generic work flow for the creation of a line list in ExoMol.

As any line list is a large project, there are many factors to consider. Among them is the representation of the KEO – in terms of the choice of coordinates and the axis used – and the use of symmetry to reduce the computational load and to more precisely classify the ro-vibrational states. These will be the focal points of this thesis.

The first way I advanced the process of ro-vibrational calculations – and the first broad aim of this thesis – was through the development of an analytic KEO approach. An analytic KEO is more compact than the standard Taylor expanded KEO of (the ExoMol program) TROVE. Moreover, as the Hamiltonian and the eigenfunctions are expressed in valance coordinates, the eigenenergies should be expected to converge with a smaller basis set compared to TROVE's usual linearised coordinates. The analytic

KEO implementation was tested in the calculation of an $H_2CS$ line list.

The second broad aim of this thesis is to further exploit molecular symmetry in ro-vibrational calculations. Generally speaking, molecular symmetry is invaluable. This assisted in the $3N-6$ implementation for calculating line lists of triatomic molecules. It also allowed me to determine how molecular symmetry restricts the choice of a body-fixed (BF) frame.

Finally, I completed much of the necessary preliminary work for the line list calculation of $C_2H_6$. It distinguishes itself from other ExoMol molecules in two ways: first, as it contains eight atoms, it is much larger than a typical ExoMol molecule and therefore prone to the effects of the "curse of dimensionality". Second, and somewhat related to the first, is that $C_2H_6$ has many symmetries. For ro-vibrational calculations, this is a two way sword: although the symmetries can be exploited, and thus a greater number of them is advantageous, they also come with idiosyncrasies which must be properly addressed.

With these difficulties in mind, I derived an analytic KEO for $C_2H_6$ which should go some way towards combatting the dimensionality issue. I also utilised the symmetries of $C_2H_6$ substantially in TROVE's symmetrisation procedure.

The structure of this thesis is as follows: Chapter 2 reviews the necessary background knowledge required for the other Chapters. It contains a more thorough description of the work flow of Figure 1.1 and, in particular, the unique steps in TROVE. There is also a focus on the calculation of a KEO in general, with the analytic KEO implementation in TROVE detailed in Chapter 3. Finally, basic results from the theory of finite groups and how they pertain to ro-vibrational calculations are provided. Following those chapters, Chapter 4 discusses the MARVELised $H_2CS$ line list recently computed. The $H_2CS$ MARVEL project and the line list were covered in Refs. [1] and [2], respectively, and form the basis of the chapter. An application of group theory for TROVE's triatomic molecule implementation is described in Chapter 5. It is predominately based on the work in Ref. [3]. Chapter 6, taken primarily from Ref. [4], examines the possible BF-frames needed to utilise the symmetry of the rotational coordinates, with the example of $CH_3Cl$. Next, Chapter 7 centres on the $C_2H_6$ molecule and describes the application of molecular symmetry onto it. It also builds on the work of the previous chapter by stating alternative BF-frames for $C_2H_6$. The work is mainly derived from Ref. [5]. Finally, Chapter 8 concludes this thesis and outlines the next steps in the line list calculation of $C_2H_6$.

# Chapter 2

# Background

## 2.1 The Born-Oppenheimer approximation

The first step in any line list calculation is the determination of the eigenfunctions and their eigenenergies. These are found by solving the ro-vibronic Schrödinger equation

$$\hat{H}_{\mathrm{rve}}\Psi_{\mathrm{rve}} = E_{\mathrm{rve}}\Psi_{\mathrm{rve}} \tag{2.1.1}$$

where $\hat{H}_{\mathrm{rve}}$ is the ro-vibronic molecular Hamiltonian, that is, one that ignores the interactions arising from the spins of the particles. In the nuclear centre of mass, the ro-vibronic molecular Hamiltonian for $l$ electrons of mass $m_{\mathrm{e}}$ and charge $e$ with coordinates $(R_{ix}, R_{iy}, R_{iz})$ and $N$ nuclei of charge $C_\alpha e$ and total mass $M$ with coordinates $(R_{\alpha x}, R_{\alpha y}, R_{\alpha z})$ is given by

$$\begin{aligned}\hat{H}_{\mathrm{rve}} = &\underbrace{-\frac{\hbar^2}{2m_{\mathrm{e}}}\sum_{i}^{l}\nabla_i^2 - \frac{\hbar^2}{2M}\sum_{i,j}^{l}\nabla_i\cdot\nabla_j}_{\hat{T}_{\mathrm{e}}} \\ &\underbrace{-\frac{\hbar^2}{2}\sum_{\alpha=2}^{N}\frac{\nabla_\alpha^2}{m_\alpha} + \frac{\hbar^2}{2M}\sum_{\alpha,\alpha'=2}^{N}\nabla_\alpha\cdot\nabla_{\alpha'}}_{\hat{T}_{\mathrm{nu}}} \\ &\underbrace{+\sum_{i<j}^{l}\frac{C_iC_je^2}{4\pi\epsilon_0\mathcal{R}_{ij}} + \sum_{i}^{l}\sum_{\alpha=2}^{N}\frac{C_\alpha C_ie^2}{4\pi\epsilon_0\mathcal{R}_{\alpha i}}}_{V_{\mathrm{e,nu}}} + \underbrace{\sum_{i}^{l}\sum_{\substack{\alpha'=2\\ \alpha>\alpha'}}^{N}\frac{C_\alpha C_{\alpha'}e^2}{4\pi\epsilon_0\mathcal{R}_{\alpha\alpha'}}}_{V_{\mathrm{nu}}}\end{aligned} \tag{2.1.2}$$

where $\nabla_i$ is

$$\nabla_i = (\partial/\partial R_{ix}\ ,\ \partial/\partial R_{iy}\ ,\ \partial/\partial R_{iz}) \tag{2.1.3}$$

and $\mathcal{R}_{ij}$ is the distance between electrons $i$ and $j$ and similarly for $\nabla_\alpha$ and $\mathcal{R}_{\alpha\alpha'}$. The underbraced terms $\hat{T}_{\mathrm{e}}$, $\hat{T}_{\mathrm{nu}}$, $V_{\mathrm{e,\,nu}}$, and $V_{\mathrm{nu}}$ are the electronic kinetic energy, nuclear kinetic energy, electron-electron and electron-nuclear Coulomb interactions, and the nuclear-nuclear Coulomb interactions, respectively.

As expected from the translational invariance of free space, the Hamiltonian components due to the nuclear centre of mass coordinates are completely decoupled from the other coordinates and thus are omitted. The current Schrödinger equation, however, is intractable being a coupled set of differential equations for $3N + 3l - 3$ coordinates and therefore further transformations and approximations are necessary. First note that the $\hat{T}_{\mathrm{nu}}$ and $\hat{T}_{\mathrm{e}}$ are decoupled and there are no terms which contain derivatives for both types of coordinates. The only electronic and nuclear coupling is through $V_{\mathrm{e,\,nu}}$.

Since the electron masses are of the order of $10^{-3}$ to $10^{-4}$ times the nuclear masses, the electronic times scales of motion should be $10^{-3}$ to $10^{-4}$ that of nuclei with comparable momentum to the electrons. This means that on the electronic time scales of motion, the nuclei can be considered stationary. This is the essence of the *Born-Oppenheimer (BO) approximation* [13]. Quantitatively, ro-vibronic Hamiltonian's eigenfunction $\Psi_{\mathrm{rve}}$ is expressed as a product of two terms $\Psi_{\mathrm{rve}} = \Psi_{\mathrm{e}}(R_{\mathrm{e}}; R_{\mathrm{nu}})\Psi_{\mathrm{nu}}(R_{\mathrm{nu}})$ where the electronic component $\Psi_{\mathrm{e}}$ is a function of the electronic coordinates $R_{\mathrm{e}}$ explicitly and a function of the nuclear coordinates $R_{\mathrm{nu}}$ only parametrically.

Applying the ro-vibronic Hamiltonian to this function, we obtain

$$\Psi_{\mathrm{nu}}(\hat{T}_{\mathrm{e}} + V_{\mathrm{e}} + V_{\mathrm{e,\,nu}})\Psi_{\mathrm{e}} + \Psi_{\mathrm{e}}(\hat{T}_{\mathrm{nu}} + V_{\mathrm{nu}})\Psi_{\mathrm{nu}} = E_{\mathrm{rve}}\Psi_{\mathrm{rve}}. \tag{2.1.4}$$

The action of $\hat{T}_{\mathrm{nu}}$ on $\Psi_{\mathrm{e}}$ is omitted as this is the term arising due to the nuclear motion on the electronic solution of the Schrödinger equation and by the above discussion should be zero. Moreover, the second term in $\hat{T}_{\mathrm{e}} = \hat{T}_{\mathrm{e}}^{0} + \hat{T}_{\mathrm{e}}'$ contains a factor $1/M$ compared to the first whose is $1/m_{\mathrm{e}}$. There ratio is the same as the ratio of electronic and nuclear time scales of motion, so the second term can be neglected under the BO approximation ordinarily.

From the nuclei's perspective, a change in their position results in a virtually instantaneous adjustment of the electronic position. Because the only coupling they have with the electrons is through the Coulomb interaction, this effectively results in a potential experienced by the nuclei. The first term in brackets in Eq. (2.1.4) is a Hamiltonian which can be solved for different nuclear positions to find the electronic eigenenergy $V_{\mathrm{e}}(R_{\mathrm{nu}})$ as the function of nuclear positions.

Many methods exist for solving these equations, such as coupled-cluster [14], configuration interaction [15], multi-reference configuration interaction [16], and Møller-Plesset perturbation theory [17]. All expand on the Hatree-Fock method [18, 19], and are iterative solutions, known as self-consistent field approaches, which take an initial guess for the eigenfunctions and use a derived equation to create a new solution. This solution is then the new starting function and the process is repeated until it converges to a stable solution. Further details for these methods can be found in the provided references.

Once a sufficiently large set of energies arising from the nuclear-nuclear interaction plus the electronic potential is found on the grid of nuclear positions, the energies must be fitted to an analytic function. This fit is known as the *ab initio* potential energy surface (PES) which is denoted here by $V_{\mathrm{e}}(R_{\mathrm{nu}}) + V_{\mathrm{nu}}$. After substituting this into Eq. (2.1.4), it becomes the nuclear Schrödinger equation:

$$\Psi_{\mathrm{e}}[\hat{T}_{\mathrm{nu}}\Psi_{\mathrm{nu}} + (V_{\mathrm{e}}(R_{\mathrm{nu}}) + V_{\mathrm{nu}})\Psi_{\mathrm{nu}}] = E_{\mathrm{rve}}\Psi_{\mathrm{rve}}. \tag{2.1.5}$$

TROVE uses the variational approach to solve this equation. Briefly, a Hamiltonian matrix is built from a suitable (symmetry-adapted) basis set and then diagonalised. The matrix' eigenvectors are the ro-vibrational states and the eigenvalues are their energies. Section 2.6 discusses this in more detail. For the molecules in question in this thesis, the ground state potential energy surfaces and also the dipole moment surfaces, defined in Eq. (2.2.8), were calculated with the coupled cluster method.

## 2.2 Intensity calculations

### 2.2.1 Line intensity

In the presence of a weak external electromagnetic (EM) field, there is a probability that a molecule in state $\Psi_i$ with energy $E_i$ will absorb a photon of frequency $\nu_{if}$ and transition to state $\Psi_f$ with energy $E_f$. This section will derive the equations for various quantities related to the interaction of a weak EM field with matter, most crucially the well defined Einstein A coefficients which give the probability of spontaneous decay per unit time from $\Psi_f$ to $\Psi_i$.

To begin [20], in thermodynamic equilibrium, the upper state occupancy is static, i.e. $\mathrm{d}N_f/\mathrm{d}t = 0$, where $N_f$ is the number of molecules in state $\Psi_f$. For the transition between $\Psi_i$ and $\Psi_f$ there are three possibilities: the EM field stimulates a transition from $\Psi_i$ to $\Psi_f$, or vice versa, and the molecule in state $\Psi_f$ spontaneously decays to $\Psi_i$. We assume that the rate of stimulated emission/absorption is proportional to the field's energy density at the frequency $\nu_{if}$ of the emitted/absorbed photon. The above is expressed by the equation

$$\frac{\mathrm{d}N_f}{\mathrm{d}t} = \rho_\nu(\nu_{fi})N_iB_{if} - \rho_\nu(\nu_{fi})N_fB_{fi} - N_fA_{fi} = 0 \tag{2.2.1}$$

where $\rho_\nu(\nu_{fi})$ is the radiation's spectral energy density (energy density per frequency), and $A_{fi}$ is the aforementioned Einstein A coefficient while $B_{if}$ and $B_{fi}$ are the Einstein B coefficients for stimulated absorption and emission, respectively. The spectral energy density is assumed to be due to a black body. Solving for $\rho_\nu(\nu_{fi})$ and noting that $N_f$ and $N_i$ follow the Boltzmann distribution, we obtain

$$\rho_\nu(\nu_{fi}) = \frac{A_{fi}}{B_{if}e^{h\nu_{fi}/kT} - B_{fi}} \tag{2.2.2}$$

which, when compared to Planck's radiation law, provides us with the relations $B_{if} = B_{fi}$ and

$$A_{fi} = \frac{8\pi h\nu_{fi}^3}{c^3}B_{fi}. \tag{2.2.3}$$

Next we will determine the Einstein B coefficients. In a weak electromagnetic field [21], so that the vector potential $\boldsymbol{A}$ is small and so $A^2$ can be neglected, the perturbing Hamiltonian on a particle of charge $Ce$ is given by (assuming the Coulomb gauge)

$$\hat{H}_1 = -\frac{Ce}{m}\hat{\boldsymbol{P}}\cdot\boldsymbol{A} \tag{2.2.4}$$

where $\hat{\boldsymbol{P}}$ is the momentum operator. We set the vector potential $\boldsymbol{A}$ to the form

$$\boldsymbol{A}(\boldsymbol{r},t) = \sum_j \boldsymbol{A}_0(\omega_j)\cos\left(\boldsymbol{k}_j\cdot\boldsymbol{r}-\omega_j t\right) = \sum_j \frac{\boldsymbol{A}_0(\omega_j)}{2}(e^{i(\boldsymbol{k}_j\cdot\boldsymbol{r}-\omega_j t)}+e^{-i(\boldsymbol{k}_j\cdot\boldsymbol{r}-\omega_j t)})). \quad (2.2.5)$$

In the so-called electric dipole approximation, the transition wavelength is assumed to be much larger than the atom's size so that $\boldsymbol{k}\cdot\boldsymbol{r}$ is approximately constant along the atomic size and can be set to zero. Coupled with Fermi's Golden Rule, the average transition rate $R_{if}$ is given by

$$R_{if} = \sum_j \left(\frac{\pi^2}{h^2}\left|\frac{Ce}{m}\left\langle\Psi_f\right|\boldsymbol{A}_0(\omega_j)\cdot\hat{\boldsymbol{P}}\left|\Psi_i\right\rangle\right|^2(\delta(\nu_{if}-\nu_j)+\delta(\nu_{if}+\nu_j))\right). \quad (2.2.6)$$

where, as the light is incoherent, the intensities are summed rather than the amplitudes. The squared amplitudes are known as transition moments.

Consider the first half of term $j$ of Eq. (2.2.6). With $\hat{\boldsymbol{P}} = -im[\hat{H}_0,\boldsymbol{r}]/\hbar$ we can write it as

$$\frac{\pi^2}{h^2}\left|\left\langle\psi_f\right|\omega_j\boldsymbol{A}_0\cdot Ce\boldsymbol{r}\left|\psi_i\right\rangle\right|^2\delta(\nu_{fi}-\nu_j). \quad (2.2.7)$$

where, due to the delta function, we made the substitution $2\pi\nu_{fi} = \omega_{fi} = \omega_j$. Using that $\boldsymbol{E} = -\partial\boldsymbol{A}/\partial t$ we have $\omega_j\boldsymbol{A}_0(\omega_j) = \boldsymbol{E}_0(\omega_j)$. The interaction Hamiltonian becomes $\boldsymbol{E}_0(\omega_j)\cdot Ce\boldsymbol{r}$. Generalised to a molecule, the interaction is

$$\boldsymbol{E}_0(\omega_j)\cdot\sum_i C_i e\boldsymbol{r}_i = \boldsymbol{E}_0(\omega_j)\cdot\boldsymbol{\mu} \quad (2.2.8)$$

where $\boldsymbol{\mu}$ has the structure of a classical dipole moment which justifies its name. The sum is over all electrons and nuclei.

As the light is incoherent, $\boldsymbol{E}_0(\omega_j)$ has a random polarisation vector $\boldsymbol{P}$. The transition moment will contain factors such as $P_X P_Y$ whose average over a uniform distribution is zero and $P_X P_X$ whose average is 1/3. The average transition rate for this term therefore becomes

$$\frac{\pi^2 E_0(\omega_j)^2}{3h^2}\sum_A|\left\langle\Psi_f\right|\mu_A\left|\Psi_i\right\rangle|^2\delta(\nu_{if}-\nu_j) \quad (2.2.9)$$

where the sum $A$ is over the three Cartesian coordinate components of the dipole moment. We desire to substitute the squared electric field amplitude for the spectral energy density $\rho_\omega(\omega_j)$. The energy density of the full field, given by $\langle\varepsilon_0 E^2\rangle$ [22], where

the average is over time, is

$$\begin{aligned}
\langle \varepsilon_0 E^2 \rangle &= \left\langle \sum_{jk} \varepsilon_0 \frac{\boldsymbol{E}_0(\omega_j)\boldsymbol{E}_0(\omega_k)}{2} \left(\cos\big(\omega_j t - \omega_k t\big) + \cos\big(\omega_j t + \omega_k t\big)\right) \right\rangle \\
&= \sum_{jk} \varepsilon_0 \frac{E_0(\omega_j)E_0(\omega_k)}{2} \left(\tfrac{1}{2}\delta_{j,k} + \tfrac{1}{2}\delta_{j,-k}\right) \\
&= \sum_{j} \varepsilon_0 \frac{E_0(\omega_j)^2}{2} \quad \text{using } E_0(\omega_j) = E_0(-\omega_j) \\
&= \sum_{\omega_j > 0} \varepsilon_0 E_0(\omega_j)^2 \\
&= \sum_{\omega_j > 0} \frac{\varepsilon_0 E_0(\omega_j)^2}{\Delta\omega} \Delta\omega.
\end{aligned} \tag{2.2.10}$$

The spectral energy density is thus $\varepsilon_0 E_0(\omega_j)^2/\Delta\omega$. Substituting this into Eq. (2.2.9) and restricting the sum to $\omega_j > 0$ by doubling the value, we obtain

$$\begin{aligned}
R_{if} &= \sum_{\omega_j > 0} \frac{2\pi^2 \rho_\omega(\omega_j)\Delta\omega}{3h^2\varepsilon_0} \sum_A |\left\langle \Psi_f \right| \mu_A \left| \Psi_i \right\rangle|^2 (\delta(\nu_{if} - \nu_j) + \delta(\nu_{if} + \nu_j)) \\
&= \sum_{\nu_j > 0} \frac{2\pi^2 \rho_\nu(\nu_j)\Delta\nu}{3h^2\varepsilon_0} \sum_A |\left\langle \Psi_f \right| \mu_A \left| \Psi_i \right\rangle|^2 \delta(\nu_{if} - \nu_j) + \delta(\nu_{if} + \nu_j)).
\end{aligned} \tag{2.2.11}$$

Taking the limit $\lim_{\Delta v \to 0} \sum_i \Delta\nu = \int \mathrm{d}\nu$ and performing the integral, the rate $R_{if}$ is

$$R_{if} = \frac{2\pi^2 \rho_\nu(\nu_{if})}{3h^2\varepsilon_0} \sum_A |\left\langle \Psi_f \right| \mu_A \left| \Psi_i \right\rangle|^2. \tag{2.2.12}$$

This is the rate of absorption. The rate of emission is the same except $\omega_{if}$ is replaced by $\omega_{fi}$, and we therefore must take the second term of Eq. (2.2.6). We deduce the Einstein A coefficient [23]

$$A_{if} = \frac{16\pi^3 \nu_{if}}{3h\varepsilon_0 c^3} \sum_A |\left\langle \Psi_f \right| \mu_A \left| \Psi_i \right\rangle|^2. \tag{2.2.13}$$

The term $\sum_A |\left\langle \Psi_f \right| \mu_A \left| \Psi_i \right\rangle|^2$ is known as the *line strength*.

Now we can determine the line intensity. According to the theory of radiative transfer [24], for the isotropic field the reduction in the spectral intensity $I_\nu(\nu)$ (the energy per unit area per unit time per unit frequency) when the radiation crosses a distance $\mathrm{d}z$ through a molecular gas is given by

$$\mathrm{d}I_\nu(\nu) = -\sigma(\nu) I_\nu(\nu) \mathrm{d}z \tag{2.2.14}$$

where $\sigma(\nu)$ is the cross section with units $\mathrm{m}^2$ and has the interpretation as the "area" of molecules which impede the radiation transfer. Solving this equation for $I_\nu(\nu)$ gives

$$I_\nu(\nu) = I_0 e^{-\sigma(\nu)z} = I_0 e^{-c^* \alpha(\nu) z}. \tag{2.2.15}$$

where $c^*$ is the molecular concentration and $\alpha(\nu)$ is the absorption coefficient.

All that remains is to find the function $\alpha(\nu)$. $\mathrm{d}I_\nu(\nu)$ has units energy per area and is the amount of energy lost when the radiation crosses a distance $\mathrm{d}z$. The amount lost from the $i$ to $f$ transition will be the energy per transition $h\nu_{if}$ multiplied by the number of molecules in the initial state and the probability of a transition. The latter will be $\rho_\nu(\nu)B_{if}g(\nu-\nu_{if})$ where $g(\nu-\nu_{if})$ is a lineshape function peaked at $\nu_{if}$ and is unity under integration. It specifies the "spread" of the line width. The origin of this probability is explained in Section 2.2.2. The spectral energy density can be replaced by the spectral intensity with $\rho_\nu(\nu) = I_\nu(\nu)/c$ [25] so that

$$\mathrm{d}I_\nu(\nu) = \frac{1}{A}\frac{h\nu_{if}}{c}I_\nu(\nu)(N_f - N_i)B_{if}g(\nu-\nu_{if}) \tag{2.2.16}$$

where $A$ is the area under consideration. $N_f$ can be written as $c^*Ae^{-E_f/kT}/Q(T)\mathrm{d}z$ where $Q(T)$ is the partition function; likewise for $N_i$. The line intensity for the transition is defined as

$$I(i\to f) = \int \alpha(\nu)\mathrm{d}\nu = \frac{2\pi^2\nu_{if}e^{-E_i/k_BT}}{3hc\,\varepsilon_0 Q(T)}\sum_A |\left\langle\Psi_f\right|\mu_A\left|\Psi_i\right\rangle|^2\left[1-\exp\left(\frac{-h\nu_{if}}{k_BT}\right)\right]. \tag{2.2.17}$$

### 2.2.2 Line broadening

As mentioned above, absorption lines include the lineshape profile. Lines are never sharp as there are various mechanisms which broaden them in reality. Intrinsic to the system is the lifetime or natural broadening from the spontaneous decay. The lines take on the characteristic Lorentz profile. The origin of this can be glimpsed by modifying the derivation of Fermi's golden rule by using the equation

$$\dot{c}_f(t) = -\frac{i}{\hbar}\left\langle\Psi_f\right|\hat{H}'\left|\Psi_i\right\rangle e^{i(\omega_{fi}-\omega)t} - \frac{\Gamma}{2}c_f(t) \tag{2.2.18}$$

where $c_f(t)$ is the (time dependant) coefficient of $\Psi_f$ for the eigenfunction $\Psi(t)$. A term $-\frac{\Gamma}{2}c_f(t)$ is added to represent spontaneous decay of rate $\Gamma/2$ proportional to the upper state's "occupancy" $c_f(t)$. This can be solved for the probability $P_f$ for $t\to\infty$ where we obtain

$$P_f = \frac{1}{\hbar^2}|\left\langle\Psi_f\right|\hat{H}_1(t)\left|\Psi_i\right\rangle|^2\frac{1}{(\Gamma/2)^2+(\Delta\omega)^2} \tag{2.2.19}$$

where $\Delta\omega = (\omega_{fi}-\omega)$ and the probability distribution has the expected Lorentz profile. In fact, $\Gamma = A_{fi}$ and the division by two in Eq. (2.2.19) was because we anticipated the necessity to square an exponential to obtain the probability from the coefficient.

In practice, lifetime broadening's effect on the profile can be neglected as it is dwarfed by two other effects, the first of which is collisional broadening. The perturbing interactions between neighbouring molecules can reduce the upper state lifetime (increasing $A_{fi}$) and thus further broaden the line profile. Moreover, the interactions may also shift the line peak by $\Delta$. Both effects can be expressed by the (normalised profile)

$$L(\omega') = \frac{1}{\pi}\frac{\Gamma_C/2}{(\Gamma_C/2)^2+(\omega')^2} \tag{2.2.20}$$

where $\omega' = \omega - \omega_{fi} - \Delta$ and $\Gamma_C$ is the collisionally broadened decay rate.

The other major type, Doppler broadening, is of a different origin and has a different concomitant profile. This arises as the frequency of the incident light onto the molecule is modulated by the molecule's motion relative to the radiation source. If we consider only the effect in one dimension, the frequency and velocity distributions are related by

$$P_\nu(\nu)\mathrm{d}\nu = P_v(v)\frac{\mathrm{d}v}{\mathrm{d}\nu}\mathrm{d}\nu \tag{2.2.21}$$

where $v$ is the molecule's velocity. As the molecular velocities are non-relativistic, the incident and observed frequencies are related to the velocity via

$$\nu = \nu_{fi}\left(1 + \frac{v}{c}\right) \tag{2.2.22}$$

which, when solved for $v$ and substituted into Eq. (2.2.21), and noting that the velocity distribution $P_v(v)$ is given by the Maxwell distribution, results in

$$P_\nu(\nu - \nu_{fi}) = D(\nu - \nu_{fi}) = \frac{c}{\nu_{fi}}\sqrt{\frac{m}{2\pi kT}}\exp\left(-\frac{mc^2(\nu - \nu_{fi})^2}{2kT\nu_{fi}^2}\right) \tag{2.2.23}$$

which is a Gaussian of full width at half maximum

$$\Gamma_D = \sqrt{\frac{8kT\ln 2}{mc^2}}\nu_{fi}. \tag{2.2.24}$$

To combine the two effects, they are approximated as independent processes. In this way, once the perturbing interactions which cause the collisional broadening are accounted for, further broadening due to the Doppler shift is applied. Mathematically, this is a convolution of the two profiles, defined by

$$V(\nu) = \int_{-\infty}^{\infty} D(\nu')L(\nu - \nu')\mathrm{d}\nu' \tag{2.2.25}$$

called a Voigt profile. To derive this, consider transitions at frequencies $\nu_i'$ to $\nu_i' + \Delta\nu'$ from the band centre without Doppler broadening. These have the probability $L(\nu_i')\Delta\nu'$. The effect of Doppler broadening is to spread out the these transitions so that the probability $V_i(\nu)$ of a transition at $\nu$ becomes

$$V_i(\nu) = D(\nu - \nu_i')L(\nu_i')\Delta\nu'. \tag{2.2.26}$$

Summing over all $i$ and taking the limit $\lim_{\Delta\nu'\to 0}\sum_i \Delta\nu' = \int \mathrm{d}\nu'$ we obtain

$$V(\nu) = \int_{-\infty}^{\infty} D(\nu - \nu')L(\nu')\mathrm{d}\nu'. \tag{2.2.27}$$

Changing the integration variable to $\nu - \nu'$ and using that $L$ is symmetric about 0, one obtains Eq. (2.2.25).

## 2.3 The kinetic energy operator in curvilinear coordinates

### 2.3.1 General form

As explained in Section 2.1, within the BO approximation, deriving the Hamiltonian involves building both the PES and the kinetic energy operator (KEO). For the molecules considered in this work, the construction of the PESs was performed by others, and they will not be discussed in any significant detail. As part of my PhD project, TROVE was significantly updated to enable ro-vibrational calculations in other coordinates types. In particular, valence coordinates were utilised. Thus, it is instructive to review the steps involved in transforming the KEO for $N$ atoms expressed in Cartesian coordinates to the desired form.

To begin, some notation. The position vector of nucleus $\alpha$ from the nuclear centre of mass will be denoted $\boldsymbol{R}_\alpha$. We use a space-fixed (SF) frame whose origin is at the nuclear centre of mass (COM) and whose axes have a fixed direction in space and a body-fixed (BF) frame whose axes rotate with the molecule in some predetermined fashion. The SF and BF unit vectors will be denoted by $\hat{\boldsymbol{e}}_F$ and $\hat{\boldsymbol{e}}_g$, that is, uppercase subscripts are SF unit vectors and lowercase are BF unit vectors. Then, the coordinates of the position vector in each frame are

$$R_{\alpha F} = \boldsymbol{R}_\alpha \cdot \hat{\boldsymbol{e}}_F, \quad r_{\alpha g} = \boldsymbol{R}_\alpha \cdot \hat{\boldsymbol{e}}_g. \tag{2.3.1}$$

and $R_\alpha$ is the set of Cartesian coordinates $(R_{\alpha x}, R_{\alpha y}, R_{\alpha z})$ for atom $\alpha$. A list of components will be denoted as $(R_1, \ldots, R_N) = R$.

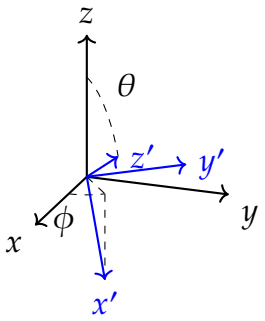
Typically, the position vectors are expressed in terms of the vibrational and angular coordinates, e.g. $\boldsymbol{R}_\alpha(q_1, \ldots, q_n, \phi, \theta, \chi) = \boldsymbol{R}_\alpha(q, \phi, \theta, \chi)$ where the latter 3 coordinates are the Euler angles which specify the orientation of the BF-frame with respect to the SF-frame [26]. We use the standard definition of the Euler angles which involves first rotating the BF-frame by $\phi$ about the BF $z$-axis (which is also the SF $z$-axis initially), then rotating the BF-frame by $\theta$ about the BF $y$-axis, and finally by $\chi$ about the BF $z$-axis. All rotations are counter-clockwise (right handed). See Figure 2.1. With this definition the coordinates $R_\alpha$ in the SF-frame are given by

$$R_\alpha(q_1, \ldots, q_n, \phi, \theta, \chi) = M_z(\phi) M_y(\theta) M_z(\chi) r_\alpha(q_1, \ldots, q_n) = M r_\alpha \tag{2.3.2}$$

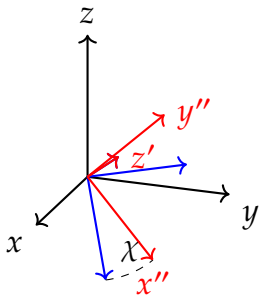
where

$$\begin{aligned} M_z(\phi) &= \begin{pmatrix} \cos\phi & -\sin\phi & 0 \\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix}, \\ M_y(\theta) &= \begin{pmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{pmatrix}, \\ M_z(\chi) &= \begin{pmatrix} \cos\chi & -\sin\chi & 0 \\ \sin\chi & \cos\chi & 0 \\ 0 & 0 & 1 \end{pmatrix}, \end{aligned} \tag{2.3.3}$$

and $r_\alpha$ are the coordinates of $\boldsymbol{R}_\alpha$ in the BF-frame.




(a) The first two Euler angles $\phi$ and $\theta$ which are essentially angles defining the spherical coordinates. The first rotation of $\phi$ is about the $z$-axis of both the SF- and BF-frame which initially coincide. The BF-frame is then rotated by $\theta$ about the BF $y$-axis or the $y'$-axis of the figure.




(b) The final Euler angle $\chi$ is due to a rotation about the BF $z$-axis or the $z'$-axis of the figure.

Figure 2.1: A visual representation of the Euler angles.

With this parametrisation [27], the KEO has the general structure

$$\hat{T} = \frac{1}{2}\hat{\pi}_a^\dagger g_{ab}\hat{\pi}_b \tag{2.3.4}$$

where a repeated index implies summation over all the coordinates and $\hat{\pi}_a$ is the conjugate momentum for $\xi_a$, given by

$$\hat{\pi}_a = -i\hbar\frac{\partial}{\partial\xi_a} \tag{2.3.5}$$

and the matrix $g_{ab}$ is

$$g_{ab} = \frac{\partial\xi_a}{\partial x_c}\frac{\partial\xi_b}{\partial x_c} \equiv B_{ac}B_{bc}. \tag{2.3.6}$$

where we define the matrix $B$. These can be further reformulated in terms of new operators $\hat{\Pi}_a$ defined by

$$\hat{\Pi}_a = (A^{-1})_{ab}\hat{\pi}_b \tag{2.3.7}$$

for some matrix $A$. The KEO then becomes

$$\hat{T} = \frac{1}{2}\hat{\Pi}_a^\dagger G_{ab}\hat{\Pi}_b \tag{2.3.8}$$

where $G_{ab} = A_{ac}A_{bd}g_{cd} = A_{ac}A_{bd}B_{ce}B_{df} = (AB)_{ae}(AB)_{bf} = C_{ae}C_{bf}$, with the definition $C = AB$. We retain the conjugate momenta for the COM and vibrational coordinates, while the conjugate momenta of the Euler angles are transformed to quasi-momentum

operators $(\hat{J}_x, \hat{J}_y, \hat{J}_z)$ defined by

$$\begin{pmatrix} \hat{J}_x \\ \hat{J}_y \\ \hat{J}_z \end{pmatrix} = -i\hbar \begin{pmatrix} \sin\chi & -\dfrac{\cos\chi}{\sin\theta} & \cot\theta\cos\chi \\ \cos\chi & \dfrac{\sin\chi}{\sin\theta} & -\cot\theta\sin\chi \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \dfrac{\partial}{\partial\theta} \\ \dfrac{\partial}{\partial\phi} \\ \dfrac{\partial}{\partial\chi} \end{pmatrix} \tag{2.3.9}$$

where the left hand side operators are in fact the BF components of the SF angular momentum. Finally, when using the integration volume element

$$\mathrm{d}\tau = \sin\theta\,\mathrm{d}\theta\,\mathrm{d}\phi\,\mathrm{d}\chi\,\mathrm{d}q_1\ldots\mathrm{d}q_N\,,$$

the KEO can be expressed as

$$\hat{T} = \tfrac{1}{2}(\hat{\Pi}_a + \tfrac{1}{2}\Lambda_a)G_{ab}(\hat{\Pi}_b - \tfrac{1}{2}\Lambda_b)$$

where $\Lambda_a = (C_{bd}\hat{\Pi}_b C_{da})$ and the brackets signify that this is purely a multiplicative operator, that is, the $\hat{\Pi}$s only act on $C$. Expanding the brackets, we obtain

$$\tfrac{1}{2}\hat{\Pi}_a G_{ab}\hat{\Pi}_b - \tfrac{1}{2}(\hat{\Pi}_a G_{ab}\Lambda_b) - \tfrac{1}{4}\Lambda_a G_{ab}\Lambda_b \tag{2.3.10}$$

while Appendix 2 of Ref. [27] shows that the purely multiplicative terms (known as the pseudo potential $U$) can be written as

$$U = \tfrac{1}{4}C_{ab}[\hat{\Pi}_a,[\hat{\Pi}_d,C_{db}]] + \tfrac{1}{4}[\hat{\Pi}_a,C_{ab}][\hat{\Pi}_d,C_{db}] \tag{2.3.11}$$

which is how Ref. [28] expresses $U$. The formulas for $G_{ab}$, Eq. 5.15 of Ref. [27], as well those for $U$ in 2.85 to 2.87 of Ref. [28], reveals that they are functions only of the vibrational coordinates. Henceforth the vibrational momentum operators will be written with lowercase $\pi$.

### 2.3.2 Sørensen's approach to derive KEOs

From the expression of the KEO, it is apparent that the main difficulty is calculating the $C$ matrix. To derive this, TROVE follows the method of Ref. [28], which we will briefly outline. Adopting his notation, the $G$ matrix and pseudo potential $U$ are given by

$$\begin{aligned} G_{ab} &= \sum_\alpha \frac{1}{m_\alpha}\boldsymbol{s}_{a,\alpha}\cdot\boldsymbol{s}_{b,\alpha} \\ U &= \frac{1}{4}\sum_{ab\alpha i}\frac{1}{m_\alpha}(s_{a,\alpha i}[\hat{\pi}_a,[\hat{\pi}_b,s_{b,\alpha i}]] + \frac{1}{2}[\hat{\pi}_a,s_{a,\alpha i}][\hat{\pi}_b,s_{b,\alpha i}]) \end{aligned} \tag{2.3.12}$$

where $\alpha \in \{1,\ldots,N\}$ ranges over the atoms, $m_\alpha$ is the mass of the $\alpha$th atom, and $i \in \{x,y,z\}$ ranges over the three components of $\boldsymbol{s}_{a,\alpha}$. Here, $s_{a,\alpha i}$ relates to $C_{ab}$ with $s_{a,\alpha i} = \sqrt{m_\alpha}C_{ab}$, where $b$ is matched to $\alpha i$. The $s$-vectors are extracted from the $s$-matrix, the inverse of the $t$-matrix: the $a$th row of $s$, say, is given by $\boldsymbol{s}_a$. This is then further

separated into groups of three (by atom) so that $\boldsymbol{s}_a$ is the $a$th row of $s$ and $\boldsymbol{s}_{a,\alpha}$ the $\alpha$th group of $\boldsymbol{s}_a$. A similar convention applies for the $t$-vectors, with the difference that $\boldsymbol{t}_a$ is the $a$th *column* of $t$. A note on notation: rotational, vibrational, and translation $s$ and $t$ vectors are labelled with $\{g, h\}$, $k$, and $F$, respectively. General components of these vectors are labelled with other letters, and atoms are labelled by $\{\alpha, \beta\}$.

The $t$-matrix itself is constructed from three types of vectors,

$$\boldsymbol{t}_{F,\alpha} = \hat{\boldsymbol{e}}_F \quad \text{(translations)}, \tag{2.3.13}$$
$$\boldsymbol{t}_{g,\alpha} = \hat{\boldsymbol{e}}_g \times \boldsymbol{R}_\alpha \quad \text{(rotations)}, \tag{2.3.14}$$
$$\boldsymbol{t}_{k,\alpha} = \partial_{q_k} \boldsymbol{R}_\alpha \quad \text{(vibrations)}. \tag{2.3.15}$$

Since we use the conjugate momenta for the translational and vibrational coordinates, one can directly calculate their $s$-vectors without the $t$-vectors,

$$\boldsymbol{s}_{F,\alpha} = \frac{m_\alpha}{M} \hat{\boldsymbol{e}}_F \quad \text{(translations)}, \tag{2.3.16}$$
$$\boldsymbol{s}_{k,\alpha} = \partial_{\alpha F} q_k \hat{\boldsymbol{e}}_F \quad \text{(vibrations)}. \tag{2.3.17}$$

Obtaining the rotational $s$-vectors requires further work. One first creates three constraints defined by the orientation of the axes. These are functions of the BF coordinates $r_\alpha$ and are always zero:

$$C^{(g)}(r_{1x}, r_{1y}, \ldots, r_{Nz}) = 0 \tag{2.3.18}$$

where $g \in \{x, y, z\}$, although the ordering of the conditions is irrelevant. We shall see simple examples later. One then creates a new set of vectors given by

$$\boldsymbol{c}_{g,\alpha} = \sum_f \hat{\boldsymbol{e}}_f \frac{\partial C^{(g)}}{\partial r_{\alpha f}}. \tag{2.3.19}$$

Using these and the rotational $t$-vectors, one can construct the matrix:

$$J_{gg'} = \sum_\alpha \boldsymbol{c}_{g,\alpha} \cdot \boldsymbol{t}_{g',\alpha}. \tag{2.3.20}$$

The rotational $s$-vectors can then finally be determined and are given by

$$\boldsymbol{s}_{g,\alpha} = \sum_{g'} J^{-1}_{gg'} \boldsymbol{c}_{g',\alpha}. \tag{2.3.21}$$

One further point to note is that the $c$-vectors must satisfy the requirement that

$$\sum_\alpha \boldsymbol{c}_{g,\alpha} = 0. \tag{2.3.22}$$

This arises because of the conditions on the rotational $s$-vectors that the translational $t$-vectors impose:

$$\sum_\alpha \boldsymbol{t}_{F,\alpha} \cdot \boldsymbol{s}_{g,\alpha} = \sum_\alpha \hat{\boldsymbol{e}}_F \cdot \boldsymbol{s}_{g,\alpha} = 0. \tag{2.3.23}$$

An analogous condition exists for the vibrational $s$-vectors. This implies

$$\sum_{\alpha} \boldsymbol{s}_{g,\alpha} = 0 \tag{2.3.24}$$

which, because $J$ is invertible, imposes it on the $\boldsymbol{c}_{g,\alpha}$ vectors also. If they do not satisfy this condition, they can be readily adjusted to $\boldsymbol{c}'$ via

$$\boldsymbol{c}'_{g,\alpha} = \boldsymbol{c}_{g,\alpha} - \frac{m_\alpha}{M} \sum_{\alpha'} \boldsymbol{c}_{g,\alpha'} \tag{2.3.25}$$

where $M$ is the sum of the nuclear masses. This does not affect the other necessary conditions.

## 2.4 Group theory and Molecular Symmetry groups

A running theme throughout this thesis is exploiting a molecule's symmetry to assist in the calculation of its spectrum. Indeed, the so-called irreducible representation (irrep) of an eigenfunction is essential in determining its spin-statistical weight $g_{\text{ns}}$ which affects the intensity of a transition it is involved in. In particular, $g_{\text{ns}}$ of zero means the one can skip the calculation of the line strength. And, while not being essential for ro-vibrational energy calculations, a symmetry-adapted basis set $\{\Psi\}$ (that is, one where each basis function belong to an irrep of the group) greatly assists in the diagonalisation of the ro-vibrational Hamiltonian matrix; it is block diagonal with the form

$$\left\langle \Psi^{\Gamma_s,J}_{\mu,n_s} \middle| \hat{H}_{\text{rv}} \middle| \Psi^{\Gamma_t,J'}_{\mu',n_t} \right\rangle = H_{\mu,\mu'} \delta_{t,s} \delta_{n_s,n_t} \delta_{J,J'} \tag{2.4.1}$$

where the indices $\mu$ and $\mu'$ label the basis functions, $J$ and $J'$ are the rotational quantum numbers, $\Gamma_s$ and $\Gamma_t$ denote irreps of the symmetry group, and $n_s(n_t)$ labels the components of the irrep $\Gamma_s(\Gamma_t)$.

In this section, we shall review the basic aspects of finite groups and the representation of such groups to properly understand the meaning of Eq. (2.4.1) and how molecular symmetry emerges as the Molecular Symmetry group.

### 2.4.1 The definition of a group

In mathematics, a group [29] is defined as a set of elements $G$ equipped with a binary operation $*$ such that the set is closed under $*$ and that $*$ satisfies the following

- Associativity: for all $a$, $b$, and $c$ in $G$, $(a * b) * c = a * (b * c)$.
- Identity element: there exists an element $e \in G$ such that, for every $a \in G$, one has $e * a = a * e = a$.
- Inverse element: for each $a \in G$, there exists a unique element $a^{-1} \in G$ such that $a * a^{-1} = a^{-1} * a = e$.

Collectively, these are the group axioms.

For brevity, the notation $*$ for the operation will be dropped as it is implicit when two elements in $G$ are written as a pair. Groups naturally arise when considering the symmetries of an object. Colloquially, the elements of a group are operations on an object which leave the object unchanged or invariant. For example, the rotation of an equililateral triangle by $2\pi/3$ with an axis of rotation through its centre and perpendicular to the triangular plane leaves it indistinguishable from before the rotation. Combining this rotation (and its inverse which is a rotation by $-2\pi/3$ or $4\pi/3$) with the rotations and reflections through an axes parallel to the triangular plane and running from a vertex of the triangle to the midpoint of the opposite side, one forms the so-called *point group* of the triangle, $\boldsymbol{D}_3$. Although the idea of a point group can be extended to three dimensions and applied to molecules in equilibrium, with a particularly useful application for linear molecules, these are not true symmetry groups of a molecule.

### 2.4.2 Molecular Symmetry group

Lougett-Higgins [30] elucidated the concept of molecular symmetry in his construction of the Molecular Symmetry (MS) groups. There are two types of MS operations. The first type permutes the labels of identical nuclei. Consider the atoms 1, 2, and 3 which have position vectors $(\boldsymbol{R}_1, \boldsymbol{R}_2, \boldsymbol{R}_3)$. The MS operations that permute the atomic labels in turn permute the position vectors. The operation (123), for example, which is shown by Figure 2.2, and which relabels atom 1 as atom 2, atom 2 as atom 3, and so on, changes the position vectors by

$$(\boldsymbol{R}_1, \boldsymbol{R}_2, \boldsymbol{R}_3) \rightarrow (\boldsymbol{R}_3, \boldsymbol{R}_1, \boldsymbol{R}_2) \tag{2.4.2}$$

as now atom 2 has position vector $\boldsymbol{R}_1$, and so on. The other type of operation is the

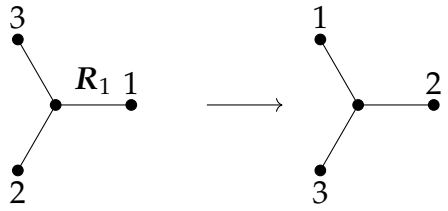



Figure 2.2: The transformation (123) on a hypothetical molecule.

inversion which changes the position vector $\boldsymbol{R}_\alpha$ to $-\boldsymbol{R}_\alpha$.

Any operation in the MS group is formed by a feasible (in the language of Longuet-Higgins) combination of these operations. To construct the irreps of Eq. (2.4.1), we must understand the more general theory of representations of finite groups, the topic of Section 2.4.4. We must also understand how the MS group is manifested through the transformation of the Hamiltonian and its eigenfunctions, which means, in effect, the transformation of functions (and derivatives) of the coordinates. This question is addressed in the next two sections and will prove important later to correctly transform functions of the $C_2H_6$ vibrational coordinates under MS operations. It will also be used to rigorously define how the Hamiltonian is transformed under MS operations as well, and therefore demonstrate its invariance.

### MS operations on functions

We'll only need to consider functions of the Cartesian coordinates, for the moment in the SF-frame. The behaviour of the functions under MS operations is then essentially the same as that of the position vectors. The operation (123) on $f(R_1, R_2, R_3)$ transforms it to

$$\begin{aligned} f(R_1, R_2, R_3) \to f(R_3, R_1, R_2) &= f(P(R_1, R_2, R_3)) \\ &= f \circ P(R_1, R_2, R_3) \end{aligned} \tag{2.4.3}$$

where $P$ represents the permutation operation. The meaning of Eq. (2.4.3) is that the function $f$ is evaluated at the transformed Cartesian coordinates.

A vibrational coordinate is just a specific type of function of the Cartesian coordinates, again in the SF-frame for the time being. For example, $q_1(R_1, R_2, R_3)$ is transformed to $q_1' = q_1(R_3, R_1, R_2)$ for coordinate $q_1$. Typically, one can express $q_1'$ in terms of the original coordinates, for example $q_1(R_3, R_1, R_2) = q_2(R_1, R_2, R_3)$ for coordinate $q_2$ [5]. For the inversion operation, the vibrational coordinate may have positive parity, as in the case of bond lengths, or negative parity, in the case dihedral angles. Thus, many results for Cartesian coordinates can be translated to curvilinear coordinates (and functions thereof) which have an effective definition for their change under MS operations.

### MS operations on derivatives

In either representation of the Hamiltonian, differential operators are present. In our approach, they have no meaning in isolation – that is, without a function being differentiated – and instead we consider the derivative of the function as itself a function which is changed by MS operations in the same way as any other function defined above. In other words, when applying MS operations to functions being differentiated, first the derivative is performed, and then the permutation is applied. In mathematical language this means that $\partial_{\alpha i} f(R_1, R_2, R_3)$ is transformed to $\partial_{\alpha i} f(R_3, R_1, R_2)$, where the derivative is with respect to the $\alpha i$th parameter. In this approach, the derivative $\partial_{\alpha i}$ is just a label to partially differentiate $f$ by its $\alpha i$th parameter. It is only the coordinates at which the derivative is evaluated which are transformed by the MS operation.

If one desires to keep the operator form of the Hamiltonian where the derivative is manifestly present, the order of the steps has to be reversed such that the expression is of the form

$$\partial_{\alpha j}(f \circ P)(R_1, \ldots, R_n). \tag{2.4.4}$$

In this form, one transforms the function $f$ to $f \circ P$ instead of evaluating the derivative of $f$ at the new coordinate. The term $\partial_{\alpha' j}(f \circ P)(R_1, \ldots, R_n)$ is given by

$$\partial_{\alpha' j}(f \circ P)(R_1, \ldots, R_n) = \sum_{\beta} \partial_{\alpha' j} P^{\beta i}(R_1, \ldots, R_n) \partial_{\beta i} f(P(R_1, \ldots, R_n)) \tag{2.4.5}$$

due to the chain rule, where $\partial_{\alpha' j} P^{\beta i}$ is the $\alpha' j$th derivative of the $\beta i$th component of $P$. Summation over $i$ is assumed. If $R_\alpha \to R_{\alpha'}$ then,

$$\begin{aligned} \partial_{\alpha' i}(f \circ P)(R_1, \ldots, R_n) &= \sum_{\beta} \partial_{\alpha' i} P^{\beta j}(R_1, \ldots, R_n) \partial_{\beta j} f(P(R_1, \ldots, R_n)) \\ &= \partial_{\alpha i} f(P(R_1, \ldots, R_n)). \end{aligned} \tag{2.4.6}$$

A more succinct form of this equation is

$$\partial_{\alpha' i}(f \circ P) = (\partial_{\alpha i} f) \circ P. \tag{2.4.7}$$

Thus, one can maintain the operator form by permuting the derivative parameter with the same rule as the coordinate permutation. The second derivative works the same way:

$$\partial^2_{\alpha' i}(f \circ P) = (\partial^2_{\alpha i} f) \circ P. \tag{2.4.8}$$

For the inversion operation $E^*$, we apply the same reasoning

$$\begin{aligned} \partial_{\alpha i}(f \circ E^*)(R_1, \ldots, R_n) &= \sum_\beta \partial_{\alpha i} E^{*\beta j}(R_1, \ldots, R_n) \partial_{\beta j} f(E^*(R_1, \ldots, R_n)) \\ &= -\partial_{\alpha i} f(E^*(R_1, \ldots, R_n)) \\ &= -\partial_{\alpha i} f(-R_1, \ldots, -R_n) \end{aligned} \tag{2.4.9}$$

so that in this case also the operators transform the same way.

There may be an added complication in transforming derivatives of functions of the vibrational coordinates when the transformation's effect on the coordinates is not a permutation but instead an affine transformation. As a typical example, consider the transformation of $(q_1, q_2)$ to be of the form

$$\begin{pmatrix} q_1' \\ q_2' \end{pmatrix} = \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \begin{pmatrix} q_1 \\ q_2 \end{pmatrix} \tag{2.4.10}$$

where $a^2 + b^2 = 1$. Then we have

$$\begin{aligned} \partial_1(f \circ P)(q_1, q_2) &= a\partial_1 f(P(q_1, q_2)) - b\partial_2 f(P(q_1, q_2)), \\ \partial_2(f \circ P)(q_1, q_2) &= b\partial_1 f(P(q_1, q_2)) + a\partial_2 f(P(q_1, q_2)) \end{aligned} \tag{2.4.11}$$

which can be rearranged to give

$$\begin{aligned} \partial_1 f(P(q_1, q_2)) &= a\partial_1(f \circ P)(q_1, q_2) + b\partial_2(f \circ P)(q_1, q_2), \\ \partial_2 f(P(q_1, q_2)) &= -b\partial_1(f \circ P)(q_1, q_2) + a\partial_2(f \circ P)(q_1, q_2). \end{aligned} \tag{2.4.12}$$

This can also be recovered using the chain rule where we find

$$\frac{\partial}{\partial q_1'} = a\frac{\partial}{\partial q_1} + b\frac{\partial}{\partial q_2} \tag{2.4.13}$$

and likewise for $\partial / \partial q_2'$. So, again, the derivatives transform as the coordinates.

### 2.4.3 Hamiltonian invariance

The MS operations are based on the indistinguishability of identical nuclei and arise when looking at the structure of the BO Hamiltonian. In general, a symmetry of a system is exhibited by the Hamiltonian's invariance under the group operation. Using the rules of Section 2.4.2, it is evident from the structure of the Hamiltonian of Eq. (2.1.2) that it is invariant under MS operations. In the context of the ro-vibrational calculations,

the invariance of the symmetry operation $P$ on the nuclear Hamiltonian $\hat{H}$'s action on the wavefunction $\Psi$ can be expressed as

$$P\hat{H}\Psi = \hat{H}P\Psi. \tag{2.4.14}$$

Since the Hamiltonian contains differential operators, the meaning of the left hand side of Eq. (2.4.14) is that the derivatives are evaluated first, then the permutation is applied. On the right hand side, the original Hamiltonian is applied to the transformed wavefunction. When the order of the operations on the left hand side is swapped, one obtains an expression of the form $\hat{H}'P\psi$ where $\hat{H}'$ is a transformed Hamiltonian. Because $\hat{H}$ is invariant, $\hat{H}' = \hat{H}$.

### 2.4.4 Representation of MS groups

Consider the action of $O_g$ on $\hat{H}\Psi$ where $\Psi$ is an eigenfunction of $\hat{H}$. The result is

$$\begin{aligned} O_g\hat{H}\Psi &= O_gE\Psi \\ \hat{H}O_g\Psi &= EO_g\Psi \end{aligned} \tag{2.4.15}$$

where $E$ is $\Psi$'s eigenvalue. Therefore, $O_g\Psi$ is also an eigenfunction of $\hat{H}$ with the same eigenvalue. Since the eigenfunctions with eigenvalue $E$ form an $n$-dimensional subspace of the full eigenspace, $O_g\Psi$ must be a linear combination of the degenerate eigenfunctions $(\Psi_1, \ldots, \Psi_n)$. Thus, when $O_g$ is applied to the eigenfunctions, the result is

$$\begin{pmatrix} O_g\Psi_1 \\ \vdots \\ O_g\Psi_n \end{pmatrix} = M(g)\begin{pmatrix} \Psi_1 \\ \vdots \\ \Psi_n \end{pmatrix} \tag{2.4.16}$$

where $M(g)$ is a matrix for the degenerate subspace. The set of all such matrices along with the eigenfunctions form an $n$-dimensional *representation* $\Gamma$ of $G$ for the degenerate subspace.

#### Equivalent, direct sum, and irreducible representations

If we transform $M(g)$ by the unitary operator $U$ to $UM(g)U^{-1}$ and also transform the degenerate eigenfunctions $\Psi = (\Psi_1, \ldots, \Psi_n)$ to $U\Psi$ we obtain a new representation. Representations connected in this way are known as *equivalent*.

For any representation, which in general is reducible, we want to transform to an equivalent representation where each operation matrix is block diagonal and each block cannot be decomposed into smaller blocks. The representation formed by those blocks are then known as irreducible representations (irreps) and a representation $\Gamma$ is decomposed into

$$\Gamma = a_1\Gamma_1 \oplus \ldots \oplus a_m\Gamma_m \tag{2.4.17}$$

where each irrep $\Gamma_i$ occurs $a_i$ times in the representation $\Gamma$. That is, each matrix for $\Gamma_i$ is repeated $a_i$ times as blocks in each matrix of $\Gamma$. The decomposition of Eq. (2.4.17) is known as a *direct sum*.

### Orthogonality relations and characters

Arguably the most important results for our purposes are Schur's orthogonality relations. If $\Gamma_i$ and $\Gamma_j$ are irreps of dimensions $l_i$ and $l_j$, respectively, with matrix representations $M^{\Gamma_i}(g)$ and $M^{\Gamma_j}(g)$ for operation $g \in G$, respectively, then the following identity holds

$$\sum_{g \in G} M^{\Gamma_i}(g)^*_{mn} M^{\Gamma_j}(g)_{m'n'} = \frac{|G|}{l_i} \delta_{ij} \delta_{mm'} \delta_{nn'} \tag{2.4.18}$$

where $|G|$ is the size of $G$.

A special case of this relation is found by taking $n = m$ and $n' = m'$ and summing over $m$ and $m'$:

$$\sum_{g \in G} \mathrm{Tr}(M^{\Gamma_i}(g))^* \mathrm{Tr}(M^{\Gamma_j}(g)) = \sum_{g \in G} \chi^{\Gamma_i}(g)^* \chi^{\Gamma_j}(g) = |G| \delta_{ij} \tag{2.4.19}$$

where $\chi^{\Gamma_i}(g)$ is the trace of $M^{\Gamma_i}(g)$ and is known as the irrep's *character*. The property $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$ dictates that equivalent irreps have the same characters. Because of this, and because of the decomposition of Eq. (2.4.17), we also have

$$\chi^{\Gamma}(g) = a_1 \chi^{\Gamma_1}(g) + \ldots + a_m \chi^{\Gamma_m}(g). \tag{2.4.20}$$

Moreover, the coefficients $a_i$ can be found with the orthogonality relation:

$$a_i = \frac{1}{|G|} \sum_{g \in G} \chi^{\Gamma}(g) \chi^{\Gamma_i}(g)^*. \tag{2.4.21}$$

Using the decomposition of Eq. (2.4.20), we have

$$\frac{1}{|G|} \sum_{g \in G} \chi^{\Gamma}(g)^* \chi^{\Gamma}(g) = a_1^2 + \ldots + a_m^2 \tag{2.4.22}$$

so that $\Gamma$ is irreducible if and only if the above sum equals 1.

### Conjugacy classes and character tables

Group elements can be partitioned according to *conjugacy classes*. Two elements $g$ and $h$ are in the same conjugacy class if there exists an element $c$ such that $g = chc^{-1}$. This is an equivalence relation and hence is a partition on the set of group elements. The matrices $M(g)$ have the same characters for all elements of an equivalence class. One can prove the second orthogonality relations: if $g$ and $h$ are elements in the conjugacy classes $C$ and $C'$, respectively, then we have

$$\sum_{i}^{s} \chi^{\Gamma_i}(g) \chi^{\Gamma_i}(h)^* = \begin{cases} |G|/|C| & C = C' \\ 0 & C \neq C' \end{cases} \tag{2.4.23}$$

where the sum is over all possible irreps $s$ of the group which is also the number of the group's conjugacy classes. This also satisfies

$$l_1^2 + \ldots + l_s^2 = |G| \tag{2.4.24}$$

where the $l_i$ is the dimension of irrep $\Gamma_i$.

The above discussion can be summarised in a *character table* where the rows are the irreps while the columns represent the conjugacy classes of the group. As the number of the two is equal, it is a square table. A $ij$th entry is then the character for the $j$th conjugacy class of the $i$th irrep. The rows and columns also are orthogonal as per Eq. (2.4.19) and Eq. (2.4.23). Table 2.1 shows an example character table for the $\boldsymbol{C}_{3v}$(M) group whose elements are $\{E, (123), (132), (23)^*, (13)^*, (12)^*\}$.

Table 2.1: Character table of the $\boldsymbol{C}_{3v}$(M) group. The rows labels are the irreps while the columns label representative members of the conjugacy classes. The characters of these elements for the given irrep is shown in the table.

| $\Gamma$ | $E$ | $(123)$ | $(23)^*$ |
|---|---|---|---|
| $A_1$ | 1 | 1 | 1 |
| $A_2$ | 1 | 1 | $-1$ |
| $E$ | 2 | $-1$ | 0 |

**Projection operators**

Once the irreps have be found, we can obtain, from a set of degenerate eigenfunctions, the decomposition into the irreps by applying so-called projection operators onto a degenerate eigenfunction [31]. The projection onto the $m$th component of the $i$th irrep is given by

$$P^{\Gamma_i}_{mm} = \frac{d_i}{|G|} \sum_{g \in G} M^{\Gamma_i}(g)^*_{mm} O_g. \tag{2.4.25}$$

The function $\psi_n$ is related to a decomposed set of functions via $\psi_n = U_{n,jk} \phi^{\Gamma_j}_k$, where $U_{n,jk}$ is the $n, jk$ component of the unitary $U$ and $jk$ correponds to the $k$th component of the irrep $\Gamma_j$. Applying the projection operator to $\psi_n$ in this form and using that $\phi^{\Gamma_j}_k$ transforms irreducibly, along with the orthogonality relation of Eq. (2.4.18), we can obtain

$$P^{\Gamma_i}_{mm} \psi_n = U_{n,im} \phi^{\Gamma_i}_m \tag{2.4.26}$$

which is the desired function up to the multiplicative factor $U_{n,im}$. The result can be normalised, and the remaining functions can be obtained with the transfer operators

$$P^{\Gamma_i}_{mt} = \frac{d_i}{|G|} \sum_{g \in G} M^{\Gamma_i}(g)^*_{mt} O_g. \tag{2.4.27}$$

Through similar steps as in the case of the projection operator, we find that $P^{\Gamma_i}_{mt} \phi^{\Gamma_i}_m = \phi^{\Gamma_i}_t$.

**Product representations**

It is commonplace in ro-vibrational calculations to multiply functions which have known transformation properties into a new function. The problem then is in determining the transformation property combined function's transformation properties. If $\phi_k^{\Gamma_i}$ and $\psi_l^{\Gamma_j}$ transform as the $k$th and $l$th components of the $\Gamma_i$ and $\Gamma_j$ irreps, respectively, then the function $\phi_k^{\Gamma_i}\psi_l^{\Gamma_j}$ transforms as follows

$$O_g\phi_k^{\Gamma_i}\psi_l^{\Gamma_j} = M^{\Gamma_i}(g)_{km}M^{\Gamma_j}(g)_{ln}\phi_m^{\Gamma_i}\psi_n^{\Gamma_j} \tag{2.4.28}$$

so that the $kl, mn$ component of the transformation matrix for group element $g$ is given by $M^{\Gamma_i}(g)_{km}M^{\Gamma_j}(g)_{ln}$. This way of multiplying matrices is known as the *outer product* and the generated representation is known as the product representation. The product representation is denoted $\Gamma_i \otimes \Gamma_j$. One can also straightforwardly see that the trace of the combined representation (given by summing over the $km, km$ components) is equal to the product of the characters of the individual representation. In general, this representation will be reducible and can be reduced through the projection operator technique described in Section 2.4.4.

## 2.4.5 Vanishing matrix elements and selection rules

The orthogonality rules have further uses. One example is establishing when the Hamiltonian matrix elements between symmetrised product functions are zero. Eq. (2.4.1) told us that the Hamiltonian matrix element is only non-zero if the irreps and the components of those irreps are the same for each function. In the following we prove this result.

First, let us note that the definite integral $I(f)$ of a representation function $f$ should be invariant under MS operations as the result is a number. Also, as long as the integration range is symmetric for coordinates which are transformed into each other, the integral of the transformed function should be the same as the untransformed function. Thus, $I(f) = O_g I(f) = I(O_g f)$. We may apply the sum over all symmetry operations to find

$$I(f) = I\left(\frac{1}{|G|}\sum_{g\in G} O_g f\right) = I(P^{\Gamma_s} f) \tag{2.4.29}$$

where $\Gamma_s$ is the fully symmetric representation. The integral is non-zero only if $f$ contains the symmetric representation.

For the example of the Hamiltonian matrix integrand, given by $\Psi_k^{\Gamma_i *}\hat{H}_{\mathrm{rv}}\Psi_l^{\Gamma_j}$, we have

$$\begin{aligned}\frac{1}{|G|}\sum_{g\in G} O_g\Psi_k^{\Gamma_i *}\hat{H}_{\mathrm{rv}}\Psi_l^{\Gamma_j} &= \frac{1}{|G|}\sum_{g\in G} M^{\Gamma_i}(g)^*_{km}M^{\Gamma_j}(g)_{ln}\Psi_m^{\Gamma_i *}\hat{H}_{\mathrm{rv}}\Psi_n^{\Gamma_j} \\ &= \left(\frac{1}{|G|}\sum_{g\in G} M^{\Gamma_i}(g)^*_{km}M^{\Gamma_j}(g)_{ln}\right)\Psi_m^{\Gamma_i *}\hat{H}_{\mathrm{rv}}\Psi_n^{\Gamma_j}\end{aligned} \tag{2.4.30}$$

where in the first line is due to $\hat{H}_{\text{rv}}$'s invariance under symmetry operations. The term in brackets is only non-zero if $i = j$ and if $m = n$. Thus, the irrep and the component of the irrep must be the same for both functions, which is precisely the result stated at the outset of Section 2.4.

The other example to consider is that of the dipole moment transition elements, which contains the matrix elements

$$\left\langle \Psi_k^i \right| \mu_F \left| \Psi_l^j \right\rangle \tag{2.4.31}$$

where $\mu_F$ is the dipole moment operator whose components are given by

$$\mu_F = \sum_i C_i e R_{iF} \tag{2.4.32}$$

which is a sum over the particles and $C_i e$ is the charge of particle $i$. The dipole moment is invariant under permutations but changes sign under permutation-inversion operations. This is a one dimensional irrep $\Gamma^*$. The multiplication of $\Gamma^* \otimes \Gamma_j$ is another irrep (due to Eq. (2.4.22)). Using the same argument as for the Hamiltonian matrix element, the dipole moment matrix element is non-zero only if $\Gamma_i = \Gamma^* \otimes \Gamma_j$ and $k = l$. This means that the irrep $\Gamma_i$ must have the same behaviour as the irrep $\Gamma_j$ except for a sign change due to MS operations containing the inversion operation. This is known as the *parity selection rule* which must change for dipole allowed transitions. If the character $\Gamma_j$ for an operation is zero the character of $\Gamma_i$ for the same operation will also be zero.

Another selection rule can be found by considering the rotation group where it can be shown that the principle rotational quantum number $J$ (for non-hyperfine transitions) satisfies $\Delta J = 0$ (for $J \neq 0$) or $\Delta J = \pm 1$ [32].

## 2.5 Spin-statistical weights

The discussion so far has omitted the nuclear spins. Instead, the focus has been on the nuclear and electronic coordinates as well electronic spins. An internal molecular wavefunction must be a function of all degrees of freedom. This segues into the *spin-statistics theorem*, which states that the internal wavefunction is symmetric under the exchange of identical integer-spin particles (bosons) and antisymmetric under the exchange of half-integer-spin particles (fermions) [33]. Thus, the internal wavefunction is an irrep which is one-dimensional and whose behaviour under permutations is fully specified by the nuclear spins. Under the behaviour of a permutation-inversion operation, the inversion part of the operation can also change the wavefunction's sign (or not): wavefunction's parity.

The above implies that there are two possible one-dimensional irreps the internal molecular wavefunction can be, which we label as $\Gamma_+$ and $\Gamma_-$ following the notation of Ref. [31]. $\Gamma_+$ has positive parity due to the inversion operation while $\Gamma_-$ has negative parity. The internal wavefunction $\Psi_{\text{int}}$ is a product of the ro-vibronic $\Psi_{\text{rve}}$ and nucleur spin $\Psi_{\text{ns}}$ wavefunctions. The combined symmetry $\Gamma_{\text{int}}$ is given by $\Gamma_{\text{int}} = \Gamma_{\text{rve}} \otimes \Gamma_{\text{ns}} \supset \Gamma_{\pm}$, that is, either of the symmetries $\Gamma_+$ or $\Gamma_-$ must be contained in $\Gamma_{\text{rve}} \otimes \Gamma_{\text{ns}}$.

To establish which irreps are contained in the spin functions, consider the product basis for each nucleus $\alpha$ with spin $S_\alpha$ and $z$-axis projection $m_{\alpha_i}$, where $-S_\alpha \leq m_{\alpha_i} \leq S_\alpha$ in steps of 1 ($2S_\alpha + 1$ functions), given by

$$|S_1, m_1\rangle \ldots |S_N, m_N\rangle \,. \tag{2.5.1}$$

The nuclear-spin basis contains all such functions. The inversion operation has no effect on the spin functions. The permutation operation swaps spins of identical nuclei. One obtains an invariant function (which would be on the diagonal of the transformation matrix, and thus contribute to the character) only if each $m_i$ for the nuclei permuted into each other are all identical. Thus, for each set of $a$ permuted nuclei, there are $(2S_a + 1)$ possible functions. The set may be of size one. The character of the operation is thus

$$\prod_a (2S_a + 1) \tag{2.5.2}$$

where the product is over all the sets of separately permuted nuclei. Performing this for all operations, where the permutation-inversion characters are the same as the counterpart permutation characters, results in the representation of the entire product basis set. The irreps can be found with the same standard reduction techniques. The result is therefore

$$\Gamma_{\text{tot ns}} = a_1\Gamma_1 \oplus \ldots \oplus a_m\Gamma_m \tag{2.5.3}$$

as usual. Multiplying each nuclear-spin irrep $\Gamma_{\text{ns}}$ with every ro-vibronic irrep $\Gamma_{\text{rve}}$ and determining how many contain the $\Gamma_\pm$ representation ($\pm$ means + or − here) provides the spin-statistical weights. For example, if $\Gamma_{\text{ns}} = \Gamma_i$ and $\Gamma_i \otimes \Gamma_{\text{rve}} \supset \Gamma_\pm$ then the spin-statistical weight of $\Gamma_{\text{rve}}$ is $a_i$. In general, for each $\Gamma_{\text{ns}}$ there will be only one irrep $\Gamma_{\text{rve}}$ such that $\Gamma_i \otimes \Gamma_{\text{rve}} \supset \Gamma_\pm$, as we shall now show. This will also provide a more straightforward method to find the spin statistical weight. Assume the decomposition of Eq. (2.5.3). Then, for each $\Gamma_k$ we want to know the value of

$$\begin{aligned} a_\pm &= \frac{1}{|G|}\sum_{g\in G} \chi^{\Gamma_i\otimes\Gamma_k}(g)\chi^{\Gamma_\pm}(g)^* \\ &= \frac{1}{|G|}\sum_{g\in G} \chi^{\Gamma_i}(g)\chi^{\Gamma_k}(g)\chi^{\Gamma_\pm}(g) \end{aligned} \tag{2.5.4}$$

where in the second line we used the character of a product and that the characters of $\Gamma_\pm$ are real. The representation generated by $\Gamma_i \otimes \Gamma_\pm$ is an irrep, and thus there is only one irrep $\Gamma_k$ which ensures that $a_\pm = 1$ for each $\Gamma_\pm$. Moreover, $\Gamma_k$ must have characters $\chi^{\Gamma_i}(g)^*\chi^{\Gamma_\pm}(g)$, so that $\Gamma_k = {\Gamma_i}^* \otimes \Gamma_\pm$.

The spin-statistical weights $a_i$ can be written in terms of the ro-vibronic irreps $\Gamma_k = {\Gamma_i}^* \otimes \Gamma_\pm$. That is, they can be written in the decomposition

$$a_1({\Gamma_1}^* \otimes \Gamma_+) \oplus a_1({\Gamma_1}^* \otimes \Gamma_-) \oplus \ldots \oplus a_m({\Gamma_m}^* \otimes \Gamma_+) \oplus a_m({\Gamma_m}^* \otimes \Gamma_-). \tag{2.5.5}$$

The characters of this decomposition structure can be found as follows. For a permutation $g$ (with no inversion), we must have $\chi^{\Gamma_k}(g) = \chi^{\Gamma_i}(g)^*$ for even permutations of fermions as in these cases $\chi_\pm(g) = 1$ and $\chi^{\Gamma_k}(g) = -\chi^{\Gamma_i}(g)^*$ for odd permutations as in these

cases $\chi^{\Gamma_\pm}(g) = -1$. Thus, there are two irreps for each $\Gamma_\pm$ each contributing $\chi^{\Gamma_i}(g)^*$ and $-\chi^{\Gamma_i}(g)^*$ to the decomposition character for the even and odd permutations, respectively. For operations with inversion $g^*$, we have $\chi^{\Gamma_k}(g^*) = \chi^{\Gamma_k}(g)$ for the irrep $\Gamma_k$ which is combined with $\Gamma_i$ to make $\Gamma_+$ and $\chi^{\Gamma_k}(g^*) = -\chi^{\Gamma_k}(g)$ for the irrep $\Gamma_k$ which is combined with $\Gamma_i$ to make $\Gamma_-$.

The character of Eq. (2.5.5) must be zero for operations containing inversion as the characters for the ${\Gamma_i}^* \otimes \Gamma_\pm$ pairs cancel. The permutation operations will have the character

$$2\prod(2S_a + 1)(-1)^{(2S_a)(n_a-1)} \tag{2.5.6}$$

where $n_a$ is the number of nuclei in the set [31]. A product of two numbers is only odd if the two constituents are odd. So, $n_a$ must be even (an odd permutation) and $2S_a$ must be odd. The latter occurs for fermions. There is therefore an overall negative factor for an odd permutation of fermions. The factor 2 at the front is from the two contributions of the $\Gamma_i^* \otimes \Gamma_\pm$ pairs. Since $\Gamma_{\text{tot ns}}$ is real as are the decomposition coefficients, the $a_i$s are the same for complex conjugated representations, and therefore complex conjugation can be dropped from Eq. (2.5.5).

In the expression for the line strength, given by

$$S(\text{f} \leftarrow \text{i}) = \sum_A \left| \left\langle \Psi^f_{\text{int}} \right| \mu_A \left| \Psi^i_{\text{int}} \right\rangle \right|^2, \tag{2.5.7}$$

we see that the nuclear spin part $\Psi_{\text{ns}}$ of the eigenfunction $\Psi_{\text{int}} = \Psi_{\text{ns}}\Psi_{\text{rve}}$ is not affected by the dipole moment operator and can be factored out of the integral. Since the nuclear spins functions are orthogonal, the line strength will only be zero if the nuclear spin functions are the same for the initial and final states. Due to Eq. (2.5.5), and the parity selection rule, for each pair of ro-vibrational eigenfunctions that produce a non-zero line strength, there will be only $g_{\text{ns}}$ nuclear spin functions of the correct symmetry. This is the nuclear spin degeneracy.

## 2.6 Ro-vibrational calculations using TROVE

### 2.6.1 Introduction

The ExoMol program TROVE (Theoretical RO-Vibrational Energies) [34, 35] was employed for the work described in this thesis. It is a general variational program for computing ro-vibrational spectra and properties for small- to medium-sized polyatomic molecules of arbitrary structure and it has been applied to a large number of polyatomic species [36–55], most of which have considerable symmetry (e.g., with molecular symmetry groups [31] such as $\boldsymbol{C}_{3\text{v}}$(M), $\boldsymbol{D}_{n\text{h}}$(M) and $\boldsymbol{T}_{\text{d}}$(M)). In this section we shall detail TROVE's automatic approach for constructing a symmetry-adapted basis set which solves the ro-vibrational Hamiltonian.

The philosophy for TROVE is that it be a black-box program capable of generating the ro-vibronic states and Einstein A coefficients of any polyatomic molecule, and many strides have been made in that regard. Currently, TROVE requires an externally inputted PES, molecular structure, the nuclear masses, the parametrisation of the vibrational

coordinates, the MS group of the molecule, and the coordinate transformation under the MS group operations. The MS group itself, including the irreps, is programmed separately. TROVE uses an automatic procedure with the provided input to generate an expansion of the KEO in linearised coordinates to a chosen order. However, since the KEOs for the molecules in this thesis were generated externally and were analytic, the details of this procedure will not be discussed here. More information can be found in Ref. [34].

TROVE's method to build a symmetry-adapted basis set happens in stages, with each stage incorporating more coordinates. This procedure is summarised below.

### 2.6.2 One-dimensional primitives

The utility of a symmetry-adapted basis is clear from Eq. (2.4.1). In TROVE, vibrational eigenfunctions are built up from a sum-of-products of one-dimensional primitive functions of the vibrational coordinates $\phi_{n_k}(q_k)$. To construct these, one-dimensional vibrational Hamiltonians, given by,

$$\frac{1}{2}\hat{\pi}_k G_{kk}^{\mathrm{1D}} \hat{\pi}_k + U(q_k)^{\mathrm{1D}} + V(q_k)^{\mathrm{1D}} \tag{2.6.1}$$

where $G_{kk}^{\mathrm{1D}}$, $U(q_k)^{\mathrm{1D}}$, and $V(q_k)^{\mathrm{1D}}$ are formed by setting the all coordinates except $q_k$ in $G_{kk}$, the pseudo potential, and the PES to their equilibrium values. The eigenfunctions of this Hamiltonian are then found either through the Numerov-Cooley method, or by diagonalisation of the matrix constructed from a specific set of basis functions, such as Hermite polynomials or a Fourier basis set.

Recently, TROVE was updated to be able to perform ro-vibrational calculations on triatomic molecules whose KEO was analytic and expressed in terms of the valence coordinates. To deal with the singularity which occurs at linearity, specially designed basis sets were used based on the associated Laguerre or associated Legendre polynomials which contained a prefactor of $\sqrt{\sin\rho}$, where $\rho$ is the angular deviation from linearity, that killed the singularity. Details can be found in Ref. [7].

Regardless of the method used to generate the 1D eigenfunctions, the result is, for each coordinate $q_k$, a set of $N_k$ functions $\phi_{n_k}(q_k)$, where $N_k$ is an input parameter. It may vary for different vibrational coordinates but should be the same for subsets of the vibrational coordinates which transform into each other under MS operations. These subsets are called coordinate "classes" in TROVE.

### 2.6.3 Reduced hamiltonians

The next step is to produce reduced Hamiltonians each of which are only functions of a coordinate class. By averaging the Hamiltonian over the "ground state" ($n_i = 0$; $\phi_0(q_i) \equiv |0\rangle_{q_i}$) primitives associated with the coordinates in all subsets not under consideration, i.e.

$$\langle 0|_{q_1} \ldots \langle 0|_{q_{k-1}} \langle 0|_{q_{l+1}} \ldots \hat{H}_{\mathrm{rv}} \ldots |0\rangle_{q_{l+1}} |0\rangle_{q_{k-1}} \ldots |0\rangle_{q_1} \tag{2.6.2}$$

where $\{q_k \ldots q_l\}$ are the coordinates in the class of interest, one obtains the reduced Hamiltonian which depends only on these coordinates; it commutes with the operations

in the MS group. The products $\phi_{n_k}(q_k)\ldots\phi_{n_l}(q_l)$ are the basis functions for the diagonalisation of this reduced Hamiltonian.

### 2.6.4 Rigid-symmetric-rotor function representations

For the rotational coordinates, TROVE uses the rigid-rotor functions $|J,k,m\rangle$ [31] as primitive basis functions. For $K \neq 0$ (where $K = |k|$), the symmetrised functions are defined as

$$|JKm\eta\rangle = \frac{i^{\eta}(-1)^{\sigma}}{\sqrt{2}}(|J,k,m\rangle + (-1)^{J+K+\eta}|J,-k,m\rangle) \tag{2.6.3}$$

where

$$\begin{cases} \sigma = K \bmod 3 & \eta = 1 \\ \sigma = 0 & \eta = 0, \end{cases} \tag{2.6.4}$$

while for $K = 0$ it is $|J00\eta\rangle = i^{\eta}|J,0,0\rangle$ where $\eta = J \bmod 2$.

### 2.6.5 Symmetrisation of the vibrational basis set

Once a set of eigenfunctions of the reduced Hamiltonian is constructed for a subset mentioned above, the degenerate eigenfunctions (which in general transform reducibly) are symmetrised by a sampling and projection procedure elaborated in Sections 4 and 5 of Ref. [56]. Briefly, a set of geometries is sampled (typically around 50) and each MS operation is applied to the eigenfunctions to obtain an overdetermined set of linear equations which are solved for the transformation matrix effecting the operation on the degenerate eigenfunctions. Symbolically, if $\Psi_{il}$ is the $l$th geometry for eigenfunction $i$, then it transforms as ${\Psi_{il}}' = M_{ij}(g)\Psi_{jl}$ for operation $g$. The matrix equation for $A_{il} = \Psi_{il}$ is $A' = M(g)A$ and one can find the Moore-Penrose right inverse $A^{+}$ [57] such that $A'A^{+} = M(g)$.

Using these matrices' characters and the irrep matrices, one can obtain the irreducible coefficients of the reducible matrix as well as projection operators onto the irreps. These operators are then applied to the eigenfunctions to generate the symmetrised basis functions.

One then produces the basis functions for the complete TROVE calculation as all products of the reduced Hamiltonian eigenfunctions, one eigenfunction from each class. In general, these total basis functions do not transform irreducibly, and so a further application of the symmetrisation procedure is required, although the sampling step can be omitted as the transformation properties of the "factor" functions are already known.

### 2.6.6 PES symmetrisation

The structure of the analytical PES must be ensured to be invariant under MS operations before being fit to the *ab initio* energies. Each potential term has the structure

$$V_k^{\text{initial}} = \prod_{i=1}^{3N-6} \xi_i^{k_i}, \tag{2.6.5}$$

where $\xi_i$ is, in general, a function of $q_i$, and the maximum expansion order $\sum_i k_i$ is an input parameter. Each MS group operation $O_g$ is independently applied to $V_k^{\text{initial}}$, i.e.,

$$V_k^{O_g} = O_g\, V_k^{\text{initial}}(\xi) = O_g \left( \prod_{i=1}^{3N-6} \xi_i^{k_i} \right) \tag{2.6.6}$$

to create $|G|$ new terms. Here, $k$ denotes the $3N-6$-dimensional hyper-index $\{k_1, \ldots, k_{3N-6}\}$ and $(\xi)$ denotes $\{\xi_1, \ldots, \xi_{3N-6}\}$. The results are summed up to produce a final term

$$V_k^{\text{final}} = \sum_{g \in G} V_k^{O_g}, \tag{2.6.7}$$

which is itself subjected to the $|G|$ MS group symmetry operations $O_g$ to check its invariance. The total potential function is then given by the expression

$$V_{\text{total}}(\xi) = \sum_k f_k V_k^{\text{final}}(\xi) \tag{2.6.8}$$

where $f_k$ are the expansion coefficients determined through a least-squares fitting to the generated *ab initio* data

### 2.6.7 Ro-vibrational basis set

The symmetrised vibrational basis set is then used to calculate the ground rotational ($J = 0$) eigenfunctions $\Psi_{i_{\text{vib}}}^{J=0,\Gamma_{\text{vib}}}$ with symmetry label $\Gamma_{\text{vib}}$. The final ro-vibrational basis set for $J > 0$ computations is then formed as a contracted product of these vibrational functions and symmetrised rigid-rotor wavefunctions $|J, K, \Gamma_{\text{rot}}\rangle$ as given by

$$\Psi_{i_{\text{vib}},K}^{J,\Gamma} = \{\Psi_{i_{\text{vib}}}^{J=0,\Gamma_{\text{vib}}} \times |J, k, \Gamma_{\text{rot}}, m\rangle\}^{\Gamma}, \tag{2.6.9}$$

where $\Gamma$, $\Gamma_{\text{vib}}$ and $\Gamma_{\text{rot}}$ are the total, vibrational, and rotational symmetries, $k$ and $m$ is the projection of the angular momentum on the BF $z$-axis and SF $Z$-axis, respectively, and $i_{\text{vib}}$ is a TROVE vibrational index to count the $\Psi_{i_{\text{vib}}}^{J=0,\Gamma_{\text{vib}}}$ functions regardless of their symmetry. For further details of the TROVE symmetry-adaptation and contraction procedure, see Ref. [56].

## 2.7 PES refinement

Purely *ab initio* calculations are insufficient to calculate spectra with line positions of the required accuracy for high resolution applications. In particular, the *ab initio* PES needs to be updated or "refined" to experimental data in practice [58]. The TROVE approach to accomplish this is to use the *ab initio* PES as the starting point and assume a refined structure given by

$$V' = V_0 + \sum_i \beta_i f_i(q) \tag{2.7.1}$$

where the $\beta_i$ are the refinement parameters and $f_i$ are various (totally symmetric) functions of the vibrational coordinates. To compute the newer refined energies one must diagonalise this Hamiltonian. The *ab initio* Hamiltonian eigenfunctions $\{\Psi^0\}$ are an obvious choice for the basis set. Then, the refined Hamiltonian's eigenfunctions can be found by solving the matrix equation

$$\left\langle \Psi_m^0 \middle| \hat{H}' \middle| \Psi_n^0 \right\rangle = \left\langle \Psi_m^0 \middle| \hat{H}_0 \middle| \Psi_n^0 \right\rangle + \sum_i \beta_i \left\langle \Psi_m^0 \middle| f_i(q) \middle| \Psi_n^0 \right\rangle = \delta_{mn} E_m^0 + \sum_i \beta_i \left\langle \Psi_m^0 \middle| f_i(q) \middle| \Psi_n^0 \right\rangle . \tag{2.7.2}$$

It is only necessary to calculate the integrals $\left\langle \Psi_m^0 \middle| f_i(q) \middle| \Psi_n^0 \right\rangle$ once. Eigenfunctions for updated parameters values can be found by diagonalising the updated matrix.

The aim in refinement is to construct a PES which produces eigenenergies as close to experiment as possible: a minimisation problem. One straightforward way to accomplish this is to minimise the quadratic residuals given by

$$S(\beta^n) = \sum_i w_i (E_i^n(\beta^n) - E_i^m)^2 \tag{2.7.3}$$

where $w_i$ are the weights, typically based on the uncertainty of the measured energy $E_i^m$, and $E_i^n(\beta^n)$ are the $n$th calculated energies for the parameter set $\beta^n$. The updated set $\beta^{n+1}$ is found from $\beta^n$ using the Newton-Guess algorithm as follows. First, the eigenenergies around the parameter values are approximated by the linear expansion

$$E_i^n(\beta^n + \Delta\beta) \simeq E_i^n(\beta^n) + \sum_j \frac{\partial E_i^n}{\partial \beta_j} \Delta\beta_j. \tag{2.7.4}$$

Inserting this approximation into Eq. (2.7.3), differentiating with respect to $\Delta\beta_k$, and equating to zero, we obtain

$$-\sum_i w_i (E_i^n - E_i^m) \frac{\partial E_i^n}{\partial \beta_k} = \sum_{ij} w_i \frac{\partial E_i^n}{\partial \beta_k} \frac{\partial E_i^n}{\partial \beta_j} \Delta\beta_j \tag{2.7.5}$$

which has the structure of the matrix equation $A_{kj} x_j = b_k$ and can be solved for $\Delta\beta$ to obtain $\beta^{n+1} = \beta^n + \Delta\beta$.

The remaining problem is to find the derivatives $\partial E_i^n / \partial \beta_j$. Ordinarily, one would need to calculate the energies at at least the neighbouring values of $\beta_j$ to estimate the derivatives. Fortunately there is a useful result, known as the Hellmann-Feynman theorem [59], which states that

$$\left. \frac{\partial E_i^n}{\partial \beta_j} \right|_{\beta^n} = \left\langle \Psi_i^n \middle| \left. \frac{\partial \hat{H}'}{\partial \beta_j} \right|_{\beta^n} \middle| \Psi_i^n \right\rangle \tag{2.7.6}$$

where $\Psi_i^n$ are the eigenfunctions with eigenenergies $E_i^n$. The right hand side is simply

$$\left\langle \Psi_i^n \middle| f_j(q) \middle| \Psi_i^n \right\rangle \tag{2.7.7}$$

and can be found from the $\left\langle \Psi^0 \middle| f_j(q) \middle| \Psi^0 \right\rangle$ integrals and the eigenfunction coefficients.

Although we could use Eq. (2.7.3) as the objective function, it is potentially not ideal. First, we can only invert Eq. (2.7.5) if the number of measured energies is greater than the number of parameters. Second, we may want to maintain the overall structure of the *ab initio* PES which should be relatively close to the actual potential; there is a danger of overfitting with Eq. (2.7.3). To address both these problems, we use an objective function of the form

$$\sum_i w_i (E_i^m - E_i^n)^2 + \sum_j z (E_j^0 - E_j^n)^2 \tag{2.7.8}$$

where the second term is the sum of squares of the difference between the *ab initio* energies and the $n$th refined energies, weighted by $z$, which essentially controls the bias of the refinement toward the *ab initio* or experimental energies. The ratios $z/w_i$ can be varied, starting from a large value, and decreased until an optimal result is found. It prevents the refined PES to deviate too severely from the *ab initio* PES.

# Appendix 2.A Derivation of the transformation of the valence coordinates

In this section the transformation properties of the valence coordinates under MS operations will be derived. The example used will be based on the ethane ($H_3{}^{12}C{}^{12}CH_3$) molecule (shown in Figure 7.2) and the generators of its MS group $\boldsymbol{G}_{36}$. However, the precise structure of this molecule is not relevant to the results obtained, which are merely illustrative examples as to the procedure in obtaining these transformation properties.

## 2.A.1 Bond lengths

The bond length between nuclei 1 and 2, $\mathcal{R}_{12}$ (with position vectors $\boldsymbol{R}_1$ and $\boldsymbol{R}_2$, respectively), is defined by $|\boldsymbol{R}_1 - \boldsymbol{R}_2| = \mathcal{R}_{12}$. We see that this value is unaffected by $E^*$ and the other operations simply result in the relabelling of the bond lengths. For example, the $H_1{-}C_a$ bond length is given by $\mathcal{R}_{1a} = |\boldsymbol{R}_1 - \boldsymbol{R}_a|$ and the operation (123) transforms $\mathcal{R}_{1a}$ to $\mathcal{R}'_{1a} = |\boldsymbol{R}_3 - \boldsymbol{R}_a| = \mathcal{R}_{3a}$, and we obtain, once again, an obvious result. The operation $(14)(26)(35)(ab)^*$ yields $\mathcal{R}'_{1a} = |-\boldsymbol{R}_4 + \boldsymbol{R}_b| = |\boldsymbol{R}_4 - \boldsymbol{R}_b| = \mathcal{R}_{4b}$

## 2.A.2 Bond angles

To obtain the angle between nuclei 1 and $b$ via nucleus $a$ we define $\boldsymbol{R}_{1a} = \boldsymbol{R}_1 - \boldsymbol{R}_a$ and likewise for $\boldsymbol{R}_{ba}$. Then, the sought angle is

$$\alpha_1 = \arccos\left(\frac{\boldsymbol{R}_{1a} \cdot \boldsymbol{R}_{ba}}{|\boldsymbol{R}_{1a}||\boldsymbol{R}_{ba}|}\right). \tag{2.A.1}$$

Despite the expression for $\alpha_1$ being more complex than the expression for $\mathcal{R}_{1a}$, the effect of MS group operations is analogous. $E^*$ produces no effect and the other operations relabel. For (123), we have

$$\alpha'_1 = \arccos\left(\frac{\boldsymbol{R}_{3a} \cdot \boldsymbol{R}_{ba}}{|\boldsymbol{R}_{3a}||\boldsymbol{R}_{ba}|}\right) = \alpha_3, \tag{2.A.2}$$

while for (14)(26)(35)$(ab)^*$

$$\alpha_1' = \arccos\left(\frac{(-\boldsymbol{R}_{4b})\cdot(-\boldsymbol{R}_{ab})}{|-\boldsymbol{R}_{4b}||-\boldsymbol{R}_{ab}|}\right) = \arccos\left(\frac{\boldsymbol{R}_{4b}\cdot\boldsymbol{R}_{ab}}{|\boldsymbol{R}_{4b}||\boldsymbol{R}_{ab}|}\right) = \alpha_4. \tag{2.A.3}$$

### 2.A.3 Dihedral angles

The expressions involving the dihedral angles are the most complicated. Consider the four nuclei labelled 1, $a$, $b$, and 4 in Figure 2.3. The dihedral angle between the plane spanned by 1, $a$, and $b$ and that spanned by $a$, $b$, and 4 is shown in Figures 2.3 and 2.4. We define $\langle v\rangle$ as the unit vector obtained by normalisation of $v$, $\langle v\rangle = v/|v|$. The orientation of the $z$-axis is defined by $\hat{\boldsymbol{e}}_z = \langle \boldsymbol{R}_{ab}\rangle$. We need two further axes $x$ and $y$ (whose orientations are defined by the unit vectors $\hat{\boldsymbol{e}}_x$ and $\hat{\boldsymbol{e}}_y$, respectively) which, together with $z$, form a right-handed axis system.

To simplify the discussion, we take the $xy$ plane to be horizontal and the $z$-axis to be vertical. We aim at obtaining the $x$ and $y$ components of $\boldsymbol{R}_{1a}$ in order to find the "horizontal" angle $\theta$ it makes with $\boldsymbol{R}_{4b}$, as shown in Figure 2.4. We require the $x$-axis to be directed along the horizontal component (i.e., the component perpendicular to $\hat{\boldsymbol{e}}_z$) of $\boldsymbol{R}_{4b}$. Thus, the $y$-axis is perpendicular to the plane defined by $\hat{\boldsymbol{e}}_z$ and $\boldsymbol{R}_{4b}$, so that we have $\hat{\boldsymbol{e}}_y = \langle \hat{\boldsymbol{e}}_z \times \boldsymbol{R}_{4b}\rangle$ and therefore $\hat{\boldsymbol{e}}_x = \hat{\boldsymbol{e}}_z \times \hat{\boldsymbol{e}}_y = \hat{\boldsymbol{e}}_z \times \langle \hat{\boldsymbol{e}}_z \times \boldsymbol{R}_{4b}\rangle$, as shown in Figure 2.5. The unit vector $\hat{\boldsymbol{e}}_{x'} = \hat{\boldsymbol{e}}_z \times \langle \hat{\boldsymbol{e}}_z \times \boldsymbol{R}_{1a}\rangle$ defines an $x'$-axis in the 1–$a$–$b$ plane; this axis is analogous to the $x$-axis in the 4–$a$–$b$ plane. The dihedral angle between the two planes is the angle between the $x$ and $x'$ axes.

The $x$ and $y$ components of $\hat{\boldsymbol{e}}_{x'}$ are $\hat{\boldsymbol{e}}_x \cdot \hat{\boldsymbol{e}}_{x'}$ and $\hat{\boldsymbol{e}}_y \cdot \hat{\boldsymbol{e}}_{x'}$, respectively. To obtain the dihedral angle in the range $[-\pi, \pi]$ with the correct sign, we use the standard trigonometric function arctan2 (For $(x, y) = (r\ \cos\varphi, r\ \sin\varphi)$, the function $\text{arctan2}(y, x) = \varphi \in [-\pi, \pi]$) to obtain

$$\begin{aligned} \theta = \tau_{41} &= \text{arctan2}(\hat{\boldsymbol{e}}_y \cdot \hat{\boldsymbol{e}}_{x'}, \hat{\boldsymbol{e}}_x \cdot \hat{\boldsymbol{e}}_{x'}) \\ &= \text{arctan2}(\hat{\boldsymbol{e}}_y \cdot \hat{\boldsymbol{e}}_z \times \langle \hat{\boldsymbol{e}}_z \times \boldsymbol{R}_{1a}\rangle, \hat{\boldsymbol{e}}_x \cdot \hat{\boldsymbol{e}}_z \times \langle \hat{\boldsymbol{e}}_z \times \boldsymbol{R}_{1a}\rangle) \end{aligned} \tag{2.A.4}$$

which, written more explicitly, is

$$\begin{aligned} \theta = \tau_{41} = \text{arctan2}[&\langle \hat{\boldsymbol{e}}_z \times \boldsymbol{R}_{4b}\rangle \cdot (\hat{\boldsymbol{e}}_z \times \langle \hat{\boldsymbol{e}}_z \times \boldsymbol{R}_{1a}\rangle), \\ &(\hat{\boldsymbol{e}}_z \times \langle \hat{\boldsymbol{e}}_z \times \boldsymbol{R}_{4b}\rangle) \cdot (\hat{\boldsymbol{e}}_z \times \langle \hat{\boldsymbol{e}}_z \times \boldsymbol{R}_{1a}\rangle)] \\ = \text{arctan2}[&\hat{\boldsymbol{e}}_z \cdot (\langle \hat{\boldsymbol{e}}_z \times \boldsymbol{R}_{1a}\rangle \times \langle \hat{\boldsymbol{e}}_z \times \boldsymbol{R}_{4b}\rangle), \\ &(\hat{\boldsymbol{e}}_z \times \langle \hat{\boldsymbol{e}}_z \times \boldsymbol{R}_{1a}\rangle) \cdot (\hat{\boldsymbol{e}}_z \times \langle \hat{\boldsymbol{e}}_z \times \boldsymbol{R}_{4b}\rangle)] \end{aligned} \tag{2.A.5}$$

where the last expression emphasises the equivalence of nuclei 1 and 4.

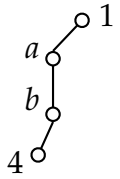



Figure 2.3: Four nuclei 1, 4, $a$, and $b$. The dihedral angle $\theta$ is the angle between the 1–$a$–$b$ and 4–$a$–$b$ planes. We define the positive direction of the angle by the right hand rule with the thumb pointing in the $\boldsymbol{R}_{ab}$ direction.

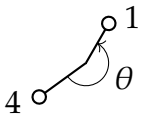



Figure 2.4: Top down view of Figure 2.3 with the dihedral angle (with the defined direction) marked as $\theta$.

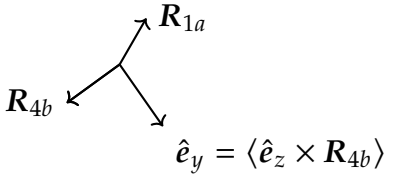



Figure 2.5: The $y$-axis.

We define the dihedral angles $\theta_{ij}$ and $\tau_{ij}$ in terms of Equation (2.A.5) and, by means of this equation, we can determine their transformation properties under the generating operations of $\boldsymbol{G}_{36}$. We note that generally $E^*\text{arctan2}(y,x) = \text{arctan2}(-y,x) = 2\pi - \text{arctan2}(y,x)$. The dihedral angle $\theta_{23}$, for example, is defined by

$$\begin{aligned}\theta_{23} = \text{arctan2}[&\hat{\boldsymbol{e}}_z \cdot (\langle \hat{\boldsymbol{e}}_z \times \boldsymbol{R}_{3a} \rangle \times \langle \hat{\boldsymbol{e}}_z \times \boldsymbol{R}_{2a} \rangle), \\ &(\hat{\boldsymbol{e}}_z \times \langle \hat{\boldsymbol{e}}_z \times \boldsymbol{R}_{3a} \rangle) \cdot (\hat{\boldsymbol{e}}_z \times \langle \hat{\boldsymbol{e}}_z \times \boldsymbol{R}_{2a} \rangle)]\end{aligned} \tag{2.A.6}$$

Under (123), we obtain

$$\begin{aligned}\theta'_{23} = \text{arctan2}[&\hat{\boldsymbol{e}}_z \cdot (\langle \hat{\boldsymbol{e}}_z \times \boldsymbol{R}_{2a} \rangle \times \langle \hat{\boldsymbol{e}}_z \times \boldsymbol{R}_{1a} \rangle), \\ &(\hat{\boldsymbol{e}}_z \times \langle \hat{\boldsymbol{e}}_z \times \boldsymbol{R}_{2a} \rangle) \cdot (\hat{\boldsymbol{e}}_z \times \langle \hat{\boldsymbol{e}}_z \times \boldsymbol{R}_{1a} \rangle)] = \theta_{12}.\end{aligned} \tag{2.A.7}$$

To obtain the effect of (14)(26)(35)$(ab)^*$, we initially apply (14)(26)(35)$(ab)$ with the result

$$\begin{aligned}\text{arctan2}[&-\hat{\boldsymbol{e}}_z \cdot (\langle -\hat{\boldsymbol{e}}_z \times \boldsymbol{R}_{5b} \rangle \times \langle -\hat{\boldsymbol{e}}_z \times \boldsymbol{R}_{6b} \rangle), \\ &(-\hat{\boldsymbol{e}}_z \times \langle -\hat{\boldsymbol{e}}_z \times \boldsymbol{R}_{5b} \rangle) \cdot (-\hat{\boldsymbol{e}}_z \times \langle -\hat{\boldsymbol{e}}_z \times \boldsymbol{R}_{6b} \rangle)].\end{aligned} \tag{2.A.8}$$

Applying $E^*$ reverses $\hat{\boldsymbol{e}}_z$, and by swapping $\boldsymbol{R}_{5b}$ and $\boldsymbol{R}_{6b}$ we obtain the final result:

$$\begin{aligned}\theta'_{23} = \text{arctan2}[&-\hat{\boldsymbol{e}}_z \cdot (\langle -\hat{\boldsymbol{e}}_z \times \boldsymbol{R}_{6b} \rangle \times \langle -\hat{\boldsymbol{e}}_z \times \boldsymbol{R}_{5b} \rangle), \\ &(-\hat{\boldsymbol{e}}_z \times \langle -\hat{\boldsymbol{e}}_z \times \boldsymbol{R}_{6b} \rangle) \cdot (-\hat{\boldsymbol{e}}_z \times \langle -\hat{\boldsymbol{e}}_z \times \boldsymbol{R}_{5b} \rangle)] = \theta_{56}.\end{aligned} \tag{2.A.9}$$

where the positive direction of the dihedral angle is in the sense of the proton numbering.

Finally, applying (14)(25)(36)$(ab)$ to $\theta_{23}$ gives

$$\begin{aligned}\theta'_{23} = \text{arctan2}[&-\hat{\boldsymbol{e}}_z \cdot (\langle -\hat{\boldsymbol{e}}_z \times \boldsymbol{R}_{6b} \rangle \times \langle -\hat{\boldsymbol{e}}_z \times \boldsymbol{R}_{5b} \rangle), \\ &(-\hat{\boldsymbol{e}}_z \times \langle \hat{\boldsymbol{e}}_z \times \boldsymbol{R}_{6b} \rangle) \cdot (-\hat{\boldsymbol{e}}_z \times \langle -\hat{\boldsymbol{e}}_z \times \boldsymbol{R}_{5b} \rangle)] = \theta_{56}.\end{aligned} \tag{2.A.10}$$

For $\tau_{41}$, which is defined by going counterclockwise from 4 to 1, the equation is

$$\begin{aligned}\text{arctan2}[&\hat{\boldsymbol{e}}_z \cdot (\langle \hat{\boldsymbol{e}}_z \times \boldsymbol{R}_{1a} \rangle \times \langle \hat{\boldsymbol{e}}_z \times \boldsymbol{R}_{4b} \rangle), \\ &(\hat{\boldsymbol{e}}_z \times \langle \hat{\boldsymbol{e}}_z \times \boldsymbol{R}_{1a} \rangle) \cdot (\hat{\boldsymbol{e}}_z \times \langle \hat{\boldsymbol{e}}_z \times \boldsymbol{R}_{4b} \rangle)] = \tau_{41}.\end{aligned} \tag{2.A.11}$$

Under operation (123)(456), this becomes

$$\begin{aligned}\tau'_{41} = \operatorname{arctan2}[&\hat{\boldsymbol{e}}_z \cdot (\langle \hat{\boldsymbol{e}}_z \times \boldsymbol{R}_{3a} \rangle \times \langle \hat{\boldsymbol{e}}_z \times \boldsymbol{R}_{6b} \rangle), \\ &(\hat{\boldsymbol{e}}_z \times \langle \hat{\boldsymbol{e}}_z \times \boldsymbol{R}_{3a} \rangle) \cdot (\hat{\boldsymbol{e}}_z \times \langle \hat{\boldsymbol{e}}_z \times \boldsymbol{R}_{6b} \rangle)] = \tau_{63}\end{aligned} \tag{2.A.12}$$

To determine the effect of (14)(26)(35)$(ab)^*$, we first apply (14)(26)(35)$(ab)$ to $\tau_{41}$ and obtain

$$\begin{aligned}\operatorname{arctan2}[&-\hat{\boldsymbol{e}}_z \cdot (\langle -\hat{\boldsymbol{e}}_z \times \boldsymbol{R}_{4b} \rangle \times \langle -\hat{\boldsymbol{e}}_z \times \boldsymbol{R}_{1a} \rangle), \\ &(-\hat{\boldsymbol{e}}_z \times \langle -\hat{\boldsymbol{e}}_z \times \boldsymbol{R}_{4b} \rangle) \cdot (-\hat{\boldsymbol{e}}_z \times \langle -\hat{\boldsymbol{e}}_z \times \boldsymbol{R}_{1a} \rangle)].\end{aligned} \tag{2.A.13}$$

Applying $E^*$ reverses $\hat{\boldsymbol{e}}_z$. After swapping the order of $\boldsymbol{R}_{1a}$ and $\boldsymbol{R}_{4b}$, we have

$$\begin{aligned}\tau'_{41} = \operatorname{arctan2}[&-\hat{\boldsymbol{e}}_z \cdot (\langle \hat{\boldsymbol{e}}_z \times \boldsymbol{R}_{1a} \rangle \times \langle \hat{\boldsymbol{e}}_z \times \boldsymbol{R}_{4b} \rangle), \\ &(\hat{\boldsymbol{e}}_z \times \langle \hat{\boldsymbol{e}}_z \times \boldsymbol{R}_{1a} \rangle) \cdot (\hat{\boldsymbol{e}}_z \times \langle \hat{\boldsymbol{e}}_z \times \boldsymbol{R}_{4b} \rangle)] = 2\pi - \tau_{41}.\end{aligned} \tag{2.A.14}$$

Finally, applying (14)(25)(36)$(ab)$ to $\tau_{41}$ gives

$$\begin{aligned}\tau'_{41} = \operatorname{arctan2}[&-\hat{\boldsymbol{e}}_z \cdot (\langle -\hat{\boldsymbol{e}}_z \times \boldsymbol{R}_{4b} \rangle \times \langle -\hat{\boldsymbol{e}}_z \times \boldsymbol{R}_{1a} \rangle), \\ &(-\hat{\boldsymbol{e}}_z \times \langle -\hat{\boldsymbol{e}}_z \times \boldsymbol{R}_{4b} \rangle) \cdot (-\hat{\boldsymbol{e}}_z \times \langle -\hat{\boldsymbol{e}}_z \times \boldsymbol{R}_{1a} \rangle)] = \tau_{41}.\end{aligned} \tag{2.A.15}$$

# Chapter 3

# Analytic KEOs in TROVE

## 3.1 Introduction

The construction of compact and, often, exact representations of KEOs is well established. Prior to the widespread application of variational methods in diagonalisation, Watson derived the exact kinetic energy operator in normal coordinates [60]. To enable practical calculations of high energy states, molecules with large amplitude motion, and complexes, exploration of curvilinear coordinate types also proved necessary [61]. Ref. [62] determined a general form of the KEO for triatomic molecules. Handy provided an exact kinetic energy operator for valence coordinates [63]. The valence, as well as Jacobi [64] and Radau, coordinates can be considered to be a subset of the more general polyspherical coordinates, pursued in Ref. [65] and emulated in Section 3.7.

This chapter describes a new TROVE implementation which generalises the Taylor expansion of the KEO to a sum-of-product "expansion" in terms of arbitrary functions of the valence coordinates, as opposed to being simply powers of the coordinates. A special case for linear molecules has already been applied to $CO_2$ [6] with the theoretical aspects of the approach elaborated in Ref. [7].

## 3.2 The form of the KEO

We wish to express the components of $G$ and $U$ as a sum-of-product of 1D functions of the coordinates, i.e. in the form

$$\sum_p a_p f_{i_1}(q_1) f_{i_2}(q_2) \dots f_{i_{3N-6}}(q_{3N-6}) \tag{3.2.1}$$

where the $f$s are functions and $a_p$ is the coefficient with $p \in \{i_1, \dots, i_{3N-6}\}$. In this expression, each $i_j$ denotes the $i$th function for the $j$th coordinate. The sum is over all such possibilities.

This form generalises the current TROVE approach where the components have the structure

$$\sum_p a_p (f_1(q_1))^{i_1} (f_2(q_2))^{i_2} \dots (f_{3N-6}(q_{3N-6}))^{i_{3N-6}} \tag{3.2.2}$$

where the sum is over all $p$ and $p \in \{i_1, \ldots, i_{3N-6} | i_1 + \ldots + i_{3N-6} \leq p_{\max}\}$ for some integer $p_{\max}$, i.e. Taylor expanding the $G$ components and $U$ in functions of each of the coordinates with a maximum overall power governed by $p_{\max}$. Typically, the $q$s are linearised coordinates constructed from geometrically defined coordinates. Using the notation of the more general expression, one can see that $f_{i_j}(q_i) = (f_i(q_i))^j$.

In the new TROVE implementation, no restrictions are made on the form of the $f$s and the same machinery is applicable in diagonalising the Hamiltonian when using the Taylor expansion or the arbitrary functions. For the practical implementation, however, we have considered valence coordinates in particular, and we are able with relative simplicity to generate exact and separable KEOs.

There are two advantages with this particular implementation. Evidently, the analytic kinetic energy operator is exact and therefore should be expected to give more accurate results. For larger molecules this becomes increasingly pronounced as the number of components in the KEO grows exponentially with the molecular size at a given $p_{\max}$ – therefore limiting the value of a practical expansion order.

Related to this, and mentioned briefly above, is that valence coordinates, in comparison to linearised coordinates, are in a sense the more "natural" choice of coordinates as their meaning is more directly physically relevant. They are more suitable for calculating high energy states and for non rigid systems. Eigenfunctions in terms of these coordinates should converge faster, i.e. for a smaller basis set, than those defined using linearised coordinates [35].

To diagonalise the Hamiltonian [34], we must evaluate matrix elements $\langle \psi_i | G | \psi_j \rangle$ and $\langle \psi_i | U | \psi_j \rangle$, where $|\psi_i\rangle$ is the $i$th basis function. These have the general form

$$\left( \prod_a |n_a\rangle \right) \times |J, K, m, \pm\rangle \tag{3.2.3}$$

where $|n_a\rangle$ is the $n$th excitation of the $a$th vibrational coordinate. The former TROVE approach stores certain elementary integrals, namely

$$\begin{aligned} &\langle n_i | (f_i(q_i))^j | n_i' \rangle , \\ &\langle n_i | (f_i(q_i))^j \hat{\pi}_i | n_i' \rangle , \\ &\langle n_i | \hat{\pi}_i (f_i(q_i))^j | n_i' \rangle = - \langle n_i | (f_i(q_i))^j \hat{\pi}_i | n_i' \rangle , \\ &\langle n_i | \hat{\pi}_i (f_i(q_i))^j \hat{\pi}_i | n_i' \rangle . \end{aligned} \tag{3.2.4}$$

This allows the matrix elements to be calculated through only algebraic manipulations of the elementary integrals. Thus, we see that it is crucial for the terms in the newer method to be separable so that all integrals reduce to a sum-of-products of 1D integrals.

In the next section we will see that, for a KEO in valence coordinates, there is a finite list of possible functions that are present in $G$ and $U$ and that this list is independent of the molecule, though it does depend on the BF-frame in a systematic way. The proof uses the intermediate result detailing the relation between $s$-vector components under a change of BF-frame (valid for any vibrational coordinate). This will prove useful in Chapter 6. Finally, a description of the implementation of this method in *Mathematica* will be given and the necessary modifications to the TROVE input files needed to utilise this KEO structure.

## 3.3 $s$-vector components under a change of BF-frame

To show that the $G$ matrix and $U$ have a sum-of-product form, one only needs to show that the $s$-vectors have this structure. To demonstrate this result, we shall state how the $s$-vector components relate when the BF-frame definition is changed, with the proof given in Appendix 3.B. Note that this is not the same as a rotation of the original BF-frame. Instead, only the BF coordinates in terms of the vibrational coordinates $r(q)$ are changed, as well as the conditions $C^{(g)}(r)$ that define the frame. With those relations, we will then calculate the $s$-vectors under the simplest frame conditions and use the relations to re-express them when in the desired frame.

The translational $s$-vectors are invariant under a change of the BF-frame. The BF coordinates $r_\alpha$ become $Sr_\alpha$ in a different BF-frame, where $S$ is a rotation matrix. The vibrational $s$-vectors are (Eq. (2.3.17))

$$\boldsymbol{s}_{k,\alpha} = \hat{\boldsymbol{e}}_F \partial_{\alpha F} q_k(R_1, \ldots, R_N) = \hat{\boldsymbol{e}}_g \partial_{\alpha g} q_k(r_1, \ldots, r_N) \tag{3.3.1}$$

where the latter is proved in Appendix 3.A and states that the vibrational $s$-vector components in the BF-frame are found by evaluating $\partial_{\alpha g} q_k$ at the BF coordinates. Appendix 3.B shows that in a different BF-frame the vibrational $s$-vectors components are related to the original components $s_{k,\alpha}$ via the $S$ rotation matrix as

$$s_{k,\alpha} \rightarrow S s_{k,\alpha} \tag{3.3.2}$$

which means they are related as the BF coordinates. Finally, the rotational $s$-vectors, whose BF components can also be found by evaluating their function at the BF coordinates, relate as

$$s_{\mathrm{r},\alpha} \rightarrow S \left( s_{\mathrm{r},\alpha} - \sum_k \omega_k {s_{k,\alpha}}^T \right) S^T \tag{3.3.3}$$

where $s_{\mathrm{r},\alpha}$ is the matrix with $g$th row $s_{g,\alpha}$. Here $\omega_k$ is defined as

$$\omega_k = \begin{pmatrix} \omega_{k,x} \\ \omega_{k,y} \\ \omega_{k,z} \end{pmatrix} \tag{3.3.4}$$

and also

$$S^T \frac{\partial S}{\partial q_k} = \begin{pmatrix} 0 & -\omega_{k,z} & \omega_{k,y} \\ \omega_{k,z} & 0 & -\omega_{k,x} \\ -\omega_{k,y} & \omega_{k,x} & 0 \end{pmatrix}. \tag{3.3.5}$$

An arbitrary change in the BF-frame can be expressed as a combination of three rotations about three axis – here labelled $e$, $g$, and $h$ – i.e. $S = M_e M_g M_h$, with the corresponding angles $\theta_e$, $\theta_g$, and $\theta_h$. The rotational $s$-matrix components are related by

$$s_{\mathrm{r},\alpha} \rightarrow S \left( s_{\mathrm{r},\alpha} - \sum_k \omega_k {s_{k,\alpha}}^T \right) S^T \tag{3.3.6}$$

where $\omega_{k,l}$ is given by

$$\omega_{k,l} = \frac{\partial \theta_e}{\partial q_k} (M_g M_h)_{el} + \frac{\partial \theta_g}{\partial q_k} (M_h)_{gl} + \frac{\partial \theta_h}{\partial q_k} \delta_{hl}. \tag{3.3.7}$$

For example, if the rotation is $S = M_x M_y M_z$, then $\omega_k$ for the new rotational $s$-matrix components are

$$\omega_{k,l} = \frac{\partial \theta_x}{\partial q_k}(M_y M_z)_{xl} + \frac{\partial \theta_y}{\partial q_k}(M_z)_{yl} + \frac{\partial \theta_z}{\partial q_k}\delta_{zl}. \tag{3.3.8}$$

The rotational $s$-matrix relation can be expressed as

$$s_{\mathrm{r},\alpha} \rightarrow S s_{\mathrm{r},\alpha} S^T - \sum_k \left( \left( \frac{\partial \theta_z}{\partial q_k} M_x M_y e_z + \frac{\partial \theta_y}{\partial q_k} M_x e_y + \frac{\partial \theta_x}{\partial q_k} e_x \right) {s_{k,\alpha}}^T \right) S^T. \tag{3.3.9}$$

where

$$e_x = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad e_y = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \text{ and } e_z = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \tag{3.3.10}$$

Ref. [66] provides the relationship between the original $G$ matrix and the $G$ matrix in a new BF-frame. To compare those relations to our results, we must use our relations to calculate the $G$ matrix components in a new BF-frame. The translational and vibrational components are invariant, as expected. For the Coriolis components, we apply $1/m_\alpha S s_{k'\alpha}$ to the right of both sides of Eq. (3.3.9) and sum over $\alpha$, to obtain

$$G_{\mathrm{cor},k'} \rightarrow S G_{\mathrm{cor},k'} - \sum_k \left( \frac{\partial \theta_z}{\partial q_k} M_x M_y e_z + \frac{\partial \theta_y}{\partial q_k} M_x e_y + \frac{\partial \theta_x}{\partial q_k} e_x \right) G_{\mathrm{vib},kk'} \tag{3.3.11}$$

where $G_{\mathrm{cor},k}$ are the Coriolis components (as a column vector) of the $G$ matrix corresponding to the $k$th vibrational mode and $G_{\mathrm{vib},kk'}$ is the $kk'$ component of the vibrational part of the $G$ vector. This can be written as the matrix equation

$$G_{\mathrm{cor}} \rightarrow S G_{\mathrm{cor}} - R D G_{\mathrm{vib}} \tag{3.3.12}$$

where $R$ and $D$ are given by,

$$R = M_x M_y \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} + M_x \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{3.3.13}$$

and

$$D = \begin{pmatrix} \frac{\partial \theta_x}{\partial q_1} & \cdots & \frac{\partial \theta_x}{\partial q_{3N-6}} \\ \frac{\partial \theta_y}{\partial q_1} & \cdots & \frac{\partial \theta_y}{\partial q_{3N-6}} \\ \frac{\partial \theta_z}{\partial q_1} & \cdots & \frac{\partial \theta_z}{\partial q_{3N-6}} \end{pmatrix}. \tag{3.3.14}$$

The transformation of Ref. [66] is

$$G_{\mathrm{cor}} \rightarrow S^T G_{\mathrm{cor}} - \mathcal{R}^T D G_{\mathrm{vib}} \tag{3.3.15}$$

where

$$\mathcal{R} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} M_z + \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} M_y M_z \tag{3.3.16}$$

and $S = M_x M_y M_z$. Our transformation compares to Ref. [66] when it is taken to account that their $S$ on the operators corresponds to our $S^T$ on the coordinates. Therefore, $S = M_z(-\theta_z)M_y(-\theta_y)M_x(-\theta_x) = M_z(\Theta_z)M_y(\Theta_y)M_x(\Theta_x)$ is applied to Eq. (3.3.7) instead, where $\Theta = -\theta$. Following the same steps as before, Eq. (3.3.15) is obtained by bearing in mind that $M(-\theta) = M^T(\theta)$ and that the matrix $D$ contains derivatives of $\Theta$, not $\theta$.

Finally, the transformation of the rotational $G$ matrix is given by

$$G'_{\text{rot}} = SG_{\text{rot}}S^T - RDG^T_{\text{cor}}S^T - SG_{\text{cor}}D^TR^T + RDG_{\text{vib}}D^TR^T. \tag{3.3.17}$$

which again compares with Ref. [66] with the aforementioned difference in the rotation. The pseudo potential is known to be invariant, although this is not immediately obvious from the form of the pseudo potential given in Eq. (3.4.12). It is shown explicitly in Appendix 3.C.

## 3.4 Complete set of functions of the coordinates

As explained in Ref. Section 3.2, we wish to express the KEO as a sum-of-products of 1D functions of the coordinates. We first must calculate the expansion functions $f_{i_p}$ for the $s$-vectors for the simplest choice of BF-frame. The functions that appear in the KEO in an arbitrary frame can be obtained from those $s$-vectors by the $s$-vector relations between different choices of BF-frame. In this section the bond vector frame of Figure 3.1 will be used in obtaining the simplest form of example $s$-vectors. These are the $a$-$b$ bond length, the $c$-$a$-$b$ angle, and the dihedral angle between the planes formed by $c$-$a$-$b$ and $a$-$b$-$d$. The $d$ atom could have also been connected to $a$, but this does not change the conclusion of the results.

For any set of four (or fewer, if it is a triatomic or diatomic) atoms, one can make this choice of frame. In this frame, the $s$-vector components will look identical aside from a substitution in the coordinate names.

Once a vibrational $s$-vector components are calculated, one can apply a string of rotations $S_1 \dots S_n$ as

$$s_{k,\alpha} \rightarrow S_1 \dots S_n s_{k,\alpha} \tag{3.4.1}$$

and rotate the axes by $S_n^T \dots S_1^T$ from the initial frame to the desired frame. The desired frame is the frame in which the rotational $s$-vector components are calculated. In valence coordinates, these rotations only contain sine and cosine functions of the angles and dihedral angles. Therefore, as long as, in the simplest frame, the $s$-vector components have the product form, the $s$-vector components in the desired frame will also have the product form. In the following, the bond length, angle, and dihedral angle $s$-vector components are explicitly shown in the simplest frame.

### 3.4.1 Bond length

The two atoms $a$ and $b$, with $a$ at the origin and $b$ lying along the $z$-axis, have BF coordinates $r_a = (0, 0, 0)$ and $r_b = (0, 0, \mathcal{R}_{ab})$ where $\mathcal{R}_{ab} = |\boldsymbol{R}_{ab}| = |\boldsymbol{R}_b - \boldsymbol{R}_a|$. The equation for the bond length is given by

$$\mathcal{R}_{ab} = \sqrt{(R_{bx} - R_{ax})^2 + (R_{by} - R_{ay})^2 + (R_{bz} - R_{az})^2}. \tag{3.4.2}$$

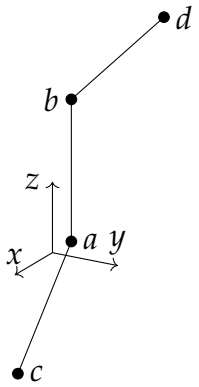



Figure 3.1: The initial bond vector frame from which the simplest form of the $s$-vectors of the atoms $a$ to $d$ are obtained. The origin is at the $a$ atom and the axes are offset for clarity.

The required derivatives are given by, once we substitute the BF coordinates expressed in terms of the vibrational coordinates,

$$\begin{aligned} \nabla_a \mathcal{R}_{ab} &= (0, 0, -1), \\ \nabla_b \mathcal{R}_{ab} &= (0, 0, 1) \end{aligned} \tag{3.4.3}$$

with the rest being zero, so in this case there is no function of the vibrational coordinates in the vibrational part of $s$.

### 3.4.2 Planar angles

The formula for the angle between atoms $a$, $b$, and $c$ is

$$\theta_a = \arccos\left(\frac{\boldsymbol{R}_{ba} \cdot \boldsymbol{R}_{ca}}{\mathcal{R}_{ba}\mathcal{R}_{ac}}\right) \tag{3.4.4}$$

where $\mathcal{R}_{ca} = |\boldsymbol{R}_{ca}|$. We shall assume the same coordinate values for these atoms as for the bond length case. The substituted derivatives are

$$\begin{aligned} \nabla_a \theta_a &= \left(\frac{1}{\mathcal{R}_{ba}} - \frac{\cos\theta_a}{\mathcal{R}_{ca}}, 0, \frac{\sin\theta_a}{\mathcal{R}_{ca}}\right), \\ \nabla_b \theta_a &= \left(-\frac{1}{\mathcal{R}_{ba}}, 0, 0\right), \\ \nabla_c \theta_a &= \left(\frac{\cos\theta_a}{\mathcal{R}_{ca}}, 0, -\frac{\sin\theta_a}{\mathcal{R}_{ca}}\right) \end{aligned} \tag{3.4.5}$$

with the rest being zero, so that $1/\mathcal{R}_{ba}$ and $1/\mathcal{R}_{ca}$ have to be added to the list of possibilities.

### 3.4.3 Dihedral angles

The coordinates are

$$\begin{aligned} r_a &= (0,0,0),\\ r_b &= (0,0,\mathcal{R}_{ab}),\\ r_c &= (\mathcal{R}_{ca}\sin\theta_a, 0, \mathcal{R}_{ca}\cos\theta_a),\\ r_d &= (\mathcal{R}_{db}\cos\phi\sin\theta_b, \mathcal{R}_{db}\sin\phi\sin\theta_b, \mathcal{R}_{db}-\mathcal{R}_{db}\cos\theta_b) \end{aligned} \tag{3.4.6}$$

where, $\mathcal{R}_{db} = |\boldsymbol{R}_{db}|$, $\theta_b$ is the planar angle between $a$, $b$, and $d$, and $\phi$ is the dihedral angle. The equation for the dihedral angle is given by

$$\begin{aligned} \phi = \arctan2[&(\boldsymbol{R}_{ba}\times\boldsymbol{R}_{ca})\cdot(\boldsymbol{R}_{ba}\times\boldsymbol{R}_{db})\mathcal{R}_{ab}, (\boldsymbol{R}_{ba}\times(\boldsymbol{R}_{ba}\times\boldsymbol{R}_{ca}))\cdot(\boldsymbol{R}_{ba}\times\boldsymbol{R}_{db})] =\\ \arctan2[&(\boldsymbol{R}_{ba}\cdot\boldsymbol{R}_{ba})(\boldsymbol{R}_{db}\cdot\boldsymbol{R}_{ca}) - (\boldsymbol{R}_{ba}\cdot\boldsymbol{R}_{ca})(\boldsymbol{R}_{ba}\cdot\boldsymbol{R}_{db}),\\ &\mathcal{R}_{ba}(\boldsymbol{R}_{ba}\cdot(\boldsymbol{R}_{ca}\times\boldsymbol{R}_{db}))] \end{aligned} \tag{3.4.7}$$

and the substituted derivatives are given by

$$\begin{aligned} \nabla_a\phi &= \left(\frac{\sin\phi}{\mathcal{R}_{ab}\tan\theta_c}, -\frac{1}{\mathcal{R}_{ab}\tan\theta_a} - \frac{\cos\phi}{\mathcal{R}_{ab}\tan\theta_c} + \frac{1}{\mathcal{R}_{ca}\sin\theta_a}, 0\right),\\ \nabla_b\phi &= \left(-\frac{\sin\phi}{\mathcal{R}_{ab}\tan\theta_c} + \frac{\sin\phi}{\mathcal{R}_{db}\sin\theta_c}, \frac{1}{\mathcal{R}_{ab}\tan\theta_a} + \frac{\cos\phi}{\mathcal{R}_{ab}\tan\theta_c} - \frac{\cos\phi}{\mathcal{R}_{db}\sin\theta_c}, 0\right),\\ \nabla_c\phi &= \left(0, -\frac{1}{\mathcal{R}_{ca}\sin\theta_a}, 0\right),\\ \nabla_d\phi &= \left(-\frac{\sin\phi}{\mathcal{R}_{db}\sin\theta_c}, \frac{\cos\phi}{\mathcal{R}_{db}\sin\theta_c}, 0\right). \end{aligned} \tag{3.4.8}$$

If $d$ was attached to $a$ rather than $b$, then $\theta_c$ in this expression is replaced by $\pi - \theta_c$. It is apparent that the additional functions the dihedral angle generates are $\csc\theta$ and $\cot\theta$ of the planar angles (but not dihedral angles). In the general case, to arrive at the coordinate system this was calculated in, we have to transform the coordinates via a matrix containing $\sin\theta$ and $\cos\theta$, for the analogous $\theta$ in the general case, which could in principle generate new functions. However, all multiplications can be expressed in terms of pre-existing functions, the most complex of which is $\cos\theta\cot\theta = \csc\theta - \sin\theta$.

### 3.4.4 Rotational $s$-vector components

Following Sørensen's method, we must define three conditions on the frame that are always zero. Using our initial frame, as shown in Figure 3.1, the conditions are

$$\begin{aligned} C^{(x)} &= r_{bx} - r_{ax} = 0,\\ C^{(y)} &= r_{by} - r_{ay} = 0,\\ C^{(z)} &= r_{cy} - r_{ay} = 0 \end{aligned} \tag{3.4.9}$$

where the first two conditions signify that the $z$-axis lies along $a$ to $b$ and the third that $x$-axis lies along the plane of $b$, $a$, and $c$. Using the method in Section 2.3.2, the rotational

$s$-vector components are, as matrices, given by

$$
\begin{aligned}
s_{\mathrm{r},a} &= \begin{pmatrix} 0 & 1/\mathcal{R}_{ab} & 0 \\ -1/\mathcal{R}_{ab} & 0 & 0 \\ 0 & \cot\theta_a/\mathcal{R}_{ab} - \csc\theta_a\mathcal{R}_{ca} & 0 \end{pmatrix}, \\
s_{\mathrm{r},b} &= \begin{pmatrix} 0 & -1/\mathcal{R}_{ab} & 0 \\ 1/\mathcal{R}_{ab} & 0 & 0 \\ 0 & -\cot\theta_a/\mathcal{R}_{ab} & 0 \end{pmatrix}, \\
s_{\mathrm{r},c} &= \begin{pmatrix} 0 & 0 & 0 \\ & 0 & 00 \\ 0 & -\cot\theta_a/\mathcal{R}_{ca} & 0 \end{pmatrix}, \\
s_{\mathrm{r},d} &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}
\end{aligned}
\tag{3.4.10}
$$

and thus we see that the rotational $s$-vectors do not introduce new functions of the coordinates in this frame. For an arbitrary number of atoms, all other rotational $s_{\mathrm{r},\alpha}$ matrices are zero as beyond $\alpha = 3$ the vectors $\boldsymbol{c}_{g,\alpha}$ are zero.

### 3.4.5 All valence coordinates

In summary, for the $s$-vectors, the functions for the bond lengths, planar angles, and dihedral angles $\mathcal{R}$, $\theta$, and $\phi$ in the initial frame are:

- Bond Lengths: $1/\mathcal{R}$
- Planar Angles: $\cos\theta$, $\sin\theta$, $\cot\theta$, $\csc\theta$
- Dihedral Angles: $\sin\phi$, $\cos\phi$.

It is only possible to maintain the sum-of-product form with a linear combination of coordinates for the dihedral angles, a result stated but not shown in Ref. [67]. This is used in Chapter 7 for $C_2H_6$.

In constructing the $G$ matrix and pseudo potential $U$, we must combine expressions of these functions. The $G$ matrix is given by

$$
G_{ab} = \sum_{\alpha} \frac{1}{m_\alpha} \boldsymbol{s}_{a,\alpha} \cdot \boldsymbol{s}_{b,\alpha} \tag{3.4.11}
$$

so we see we must multiply each function in a set with every function of the set. For the

pseudo potential $U$, it can be shown [34] to consist of four terms:

$$\begin{aligned} U_1 &= \frac{1}{8}\sum_{\alpha,g,g'}\frac{1}{m_\alpha}(\boldsymbol{e}_g\times\boldsymbol{s}_{g,\alpha})\cdot(\boldsymbol{e}_{g'}\times\boldsymbol{s}_{g',\alpha}),\\ U_2 &= -\frac{1}{4}\sum_{\alpha,g,k}\frac{1}{m_\alpha}\left(\boldsymbol{e}_g\times\boldsymbol{s}_{k,\alpha}\right)\cdot\frac{\partial\boldsymbol{s}_{g,\alpha}}{\partial q_k},\\ U_3 &= -\frac{1}{4}\sum_{\alpha,k,k'}\frac{1}{m_\alpha}\boldsymbol{s}_{k,\alpha}\cdot\frac{\partial^2\boldsymbol{s}_{k',\alpha}}{\partial q_k\partial q_{k'}},\\ U_4 &= -\frac{1}{8}\sum_{\alpha,k,k'}\frac{1}{m_\alpha}\frac{\partial\boldsymbol{s}_{k,\alpha}}{\partial q_k}\cdot\frac{\partial\boldsymbol{s}_{k',\alpha}}{\partial q_{k'}} \end{aligned} \tag{3.4.12}$$

where $U$ is the sum of these terms. Thus, we must also multiply first and second derivatives of each function in a set with every function of the set. We also must multiply a derivative of each function with the derivative of every function of the set. However, in practice, not all these combinations will actually be present in the Hamiltonian.

In the previous consideration the frame was oriented as in Figure 3.1. If a more complex choice is required, the sum-of-product form is maintained as long as the rotation angle $\theta$ is in the form

$$\theta = f_1(q_1) + \ldots + f_k(q_k). \tag{3.4.13}$$

With this choice, the $s$-vector components also remain in the some-of-product form, albeit with potentially extra or alternative functions of the coordinates. We will now show an example of an alternative frame and how this changes the $s$-vector components.

## 3.5 Example frame

The first case is the $x$-axis bisecting a dihedral angle $\phi$, shown in Figure 3.2. Using this frame requires rotating the bond vector frame to the bisector frame with the following matrix:

$$S = M_z(-\phi/2) = \begin{pmatrix} \cos\phi/2 & \sin\phi/2 & 0\\ -\sin\phi/2 & \cos\phi/2 & 0\\ 0 & 0 & 1 \end{pmatrix}. \tag{3.5.1}$$

For the atoms whose positions do not depend on $\phi$, namely the first three, the vibrational $s$-vectors also do not contain $\phi$, so that applying the rotation matrix $S$ merely introduces $\sin(\phi/2)$ and $\cos(\phi/2)$. For other atoms, the matrix $S_1$ in the list $S_1 \ldots S_n$ is $S_1 = M_z(\phi)$, and therefore, in the product $SS_1$, $\phi$ is replaced by $\phi/2$. Finally, when we apply the matrix on the $\phi$ dihedral $s$-vectors, we find again that the expressions are combined in such a way that the $s$-vectors only involve $\phi/2$ instead of $\phi$.

In practice, it is simpler to generate the $s$-vectors in this frame by defining the BF coordinates of this frame in terms of the vibrational coordinates and using conditions

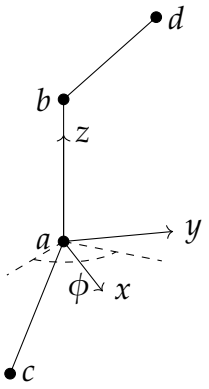



Figure 3.2: The bisecting frame where the $x$-axis bisects the $\phi$ dihedral angle illustrated by the dashed arc in the figure. The origin of the axes is set to the $a$ atom for clarity.

appropriate to this frame. In that case, the conditions are

$$\begin{aligned} C^{(x)} &= r_{by} - r_{ay} = 0, \\ C^{(y)} &= r_{bx} - r_{ax} = 0, \\ C^{(z)} &= (r_{cy} - r_{ay})(r_{dx} - r_{ax}) + (r_{cx} - r_{ax})(r_{dy} - r_{ay}) \end{aligned} \tag{3.5.2}$$

where the (different) third condition stems from the observation that the tangent of the angle between the $x$-axis to the bond $\boldsymbol{R}_c - \boldsymbol{R}_a$ projected onto the $xy$ plane (given by $(r_{cy} - r_{ay})/(r_{cx} - r_{ax})$) is negative the angle between the $x$-axis and the bond $\boldsymbol{R}_c - \boldsymbol{R}_a$ projected onto the $xy$ plane. Thus

$$\frac{r_{cy} - r_{ay}}{r_{cx} - r_{ax}} + \frac{r_{dy} - r_{ay}}{r_{dx} - r_{dx}} = 0 \tag{3.5.3}$$

and we multiply by the denominators to simplify the condition. With this, the rotational $s$-vectors we obtain depend on $\phi/2$ but not $\phi$, in other words the KEO no longer depends on $\phi$ at all. This reduces the number of 1D functions present and the KEO remains compact, a desirable property.

A similar frame, that of the $z$-axis bisecting the first planar angle, can also be used. The results for the $s$-vectors are analogous to the bisecting $x$-axis example.

## 3.6 Example of formaldehyde

The new implementation and underlying methodology was tested on the formaldehyde ($H_2CO$) molecule to compare to pre-existing TROVE calculations which used an expanded KEO, detailed in Ref. [41]. Only the vibrational energies for the $J = 0$ states are compared as this was deemed sufficient to verify the program's correctness. The structure is shown in Figure 3.3. The six coordinates used are the C−O bond, the two C−H bonds, the two H−C−O angles, and the dihedral angle of the two H−C−O planes. A bisecting frame was chosen, with the $x$-axis pointing at the midpoint of the dihedral angle, and the $z$-axis along the C−O bond towards O. Aside from the change in the KEO's representation, the TROVE parameters were identical to those of Ref. [41].

In Ref. [41], the KEO has an expansion order of 6. Both calculations expanded the potential energy to order 8, although the expansions in Ref. [41] and in the present work were performed in linearised and valence coordinates, respectively. The basis functions (see Eq. (3.2.3)) were contracted according to the polyad rule of the 1D vibrational basis functions given by:

$$P = 2(n_2 + n_3) + n_1 + n_4 + n_5 + n_6 \leq P_{\max}$$

where $n_i$ is the $n$th excitation of the $i$th vibrational coordinate in the above order. The trucation is set at $P_{\max} = 16$. Table 3.1 shows a comparison of the first 17 $A_1$ state energies, using the parameter values of Ref. [41]. The close agreement provides confidence that the method was program correctly .

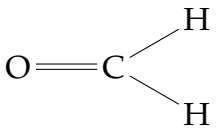



Figure 3.3: The structure of formaldehyde ($H_2CO$).

Table 3.1: Comparison of the first 17 $A_1$ state energies of Ref. [41] and the new approach.

| $I$ | $E(\text{cm}^{-1})$ (this work) | $E(\text{cm}^{-1})$[41] |
|---|---|---|
| 1 | 0000.000000 | 0000.000000 |
| 2 | 1500.100565 | 1500.120955 |
| 3 | 1746.026899 | 1746.045388 |
| 4 | 2327.438439 | 2327.497142 |
| 5 | 2494.277373 | 2494.322937 |
| 6 | 2782.368637 | 2782.410921 |
| 7 | 2998.921856 | 2999.006647 |
| 8 | 3238.844004 | 3238.937891 |
| 9 | 3471.649488 | 3471.719306 |
| 10 | 3825.530391 | 3825.967015 |
| 11 | 3935.924687 | 3936.435541 |
| 12 | 4057.853105 | 4058.101422 |
| 13 | 4083.461950 | 4083.490190 |
| 14 | 4247.422778 | 4247.609826 |
| 15 | 4256.128549 | 4256.314862 |
| 16 | 4495.183906 | 4495.499848 |
| 17 | 4529.485437 | 4529.635737 |

## 3.7 $s$-vectors in polyspherical coordinates

As a further application, the KEO in polyspherical coordinates of Ref. [65] will be re-derived, generating the $s$-vectors that produce the $G$ matrix. As usual for the Sørensen

method, the $s$-vectors are first derived, so this approach perhaps has the advantage of being less demanding to program. It also demonstrates the flexibility of the Sørensen method for deriving KEOs expressed in terms of conjugate momenta.

The polyspherical approach involves parametrising the $N$ position vectors of an $N$ atom system into $N-1$ vectors (the polyspherical vectors) which are linear combinations of the position vectors. The overall translation is factored out in the polyspherical vectors. For simplicity of notation we also denote the polyspherical vectors as $\tilde{\boldsymbol{R}}_\beta$ and the transformation matrix between the position and polyspherical vectors as $A$. Then components $\tilde{R}_{\beta F}$, $\beta \in \{1,\dots,N-1\}$ and $F \in \{X,Y,Z\}$ are defined by

$$\tilde{R}_{\beta F} = \sum_\alpha A_{\beta\alpha} R_{\alpha F}$$

where $R_{\alpha F}$ are the Cartesian coordinates. The standard approach is to align the BF $z$-axis along $\tilde{\boldsymbol{R}}_{N-1}$ and set the BF $x$-axis to be in the plane of $\tilde{\boldsymbol{R}}_{N-1}$ and $\tilde{\boldsymbol{R}}_{N-2}$. The components of each $\tilde{\boldsymbol{R}}_\beta$ are then expressed as spherical coordinates relative to these axes, i.e.

$$\tilde{R}_\beta = (\mathcal{R}_\beta \sin\theta_\beta \cos\phi_\beta, \mathcal{R}_\beta \sin\theta_\beta \sin\phi_\beta, \mathcal{R}_\beta \cos\theta_\beta) \tag{3.7.1}$$

where $\mathcal{R}_\beta$ is the magnitude for $\tilde{\boldsymbol{R}}_\beta$, $\theta_\beta$ is the planar angle between $\tilde{\boldsymbol{R}}_{N-1}$ and $\tilde{\boldsymbol{R}}_\beta$, and $\phi$ is the dihedral angle between the plane generated by $\tilde{\boldsymbol{R}}_{N-1}$ and $\tilde{\boldsymbol{R}}_{N-2}$ and the plane generated by $\tilde{\boldsymbol{R}}_{N-1}$ and $\tilde{\boldsymbol{R}}_\beta$.

We shall now derive the expressions for the $s$-vectors with these coordinates. First, $s_{\mathcal{R}_\beta,\alpha i}$ is

$$s_{\mathcal{R}_\beta,\alpha i} = \frac{\partial \mathcal{R}_\beta}{\partial R_{\alpha i}} = \sum_{\gamma j} \frac{\partial \sqrt{\tilde{\boldsymbol{R}}_\beta \cdot \tilde{\boldsymbol{R}}_\beta}}{\partial \tilde{R}_{\gamma j}} \frac{\partial \tilde{R}_{\gamma j}}{\partial R_{\alpha i}} = \frac{\tilde{R}_{\beta i}}{\mathcal{R}_\beta} A_{\beta\alpha} \tag{3.7.2}$$

or $\boldsymbol{s}_{\mathcal{R}_\beta,\alpha} = A_{\beta\alpha}\tilde{\boldsymbol{R}}_\beta/\mathcal{R}_\beta$. Next, $s_{\theta_\beta,\alpha i}$ is given by

$$\begin{aligned}
s_{\theta_\beta,\alpha i} &= \frac{\partial \theta_\beta}{\partial R_{\alpha i}} = \sum_{\gamma j} \frac{\partial}{\partial \tilde{R}_{\gamma j}} \arccos\left(\frac{\tilde{\boldsymbol{R}}_\beta \cdot \tilde{\boldsymbol{R}}_{N-1}}{\mathcal{R}_\beta \mathcal{R}_{N-1}}\right) \frac{\partial \tilde{R}_{\gamma j}}{\partial R_{\alpha i}} \\
&= -A_{\beta\alpha}\left(\frac{\tilde{R}_{N-1i}}{\mathcal{R}_\beta \mathcal{R}_{N-1}} - \frac{(\tilde{\boldsymbol{R}}_\beta \cdot \tilde{\boldsymbol{R}}_{N-1})\tilde{R}_{\beta i}}{\mathcal{R}_\beta^3 \mathcal{R}_{N-1}}\right)\frac{1}{\sin\theta_\beta} \\
&\quad - A_{N-1\alpha}\left(\frac{\tilde{R}_{\beta i}}{\mathcal{R}_\beta \mathcal{R}_{N-1}} - \frac{(\tilde{\boldsymbol{R}}_\beta \cdot \tilde{\boldsymbol{R}}_{N-1})\tilde{R}_{N-1i}}{\mathcal{R}_\beta \mathcal{R}_{N-1}^3}\right)\frac{1}{\sin\theta_\beta} \\
&= -A_{\beta\alpha}\left(\frac{\tilde{R}_{N-1i}}{\mathcal{R}_\beta \mathcal{R}_{N-1}} - \frac{\cos\theta_\beta \tilde{R}_{\beta i}}{\mathcal{R}_\beta^2}\right)\frac{1}{\sin\theta_\beta} - A_{N-1\alpha}\left(\frac{\tilde{R}_{\beta i}}{\mathcal{R}_\beta \mathcal{R}_{N-1}} - \frac{\cos\theta_\beta \tilde{R}_{N-1i}}{\mathcal{R}_{N-1}^2}\right)\frac{1}{\sin\theta_\beta}
\end{aligned} \tag{3.7.3}$$

so $\boldsymbol{s}_{\theta_\beta,\alpha}$ is given by

$$\boldsymbol{s}_{\theta_\beta\alpha} = A_{\beta\alpha}\frac{1}{\mathcal{R}_\beta}(\cos\theta_\beta\cos\phi_\beta, \cos\theta_\beta\sin\phi_\beta, -\sin\theta_\beta) - A_{N-1\alpha}\frac{1}{\mathcal{R}_{N-1}}(\cos\phi_\beta, \sin\phi_\beta, 0). \tag{3.7.4}$$

The dihedral $s$-vectors $s_{\phi_\beta \alpha i}$ are

$$\begin{aligned} s_{\phi_\beta,\alpha i} &= \frac{\partial \phi_\beta}{\partial R_{\alpha i}} \\ &= \sum_{\gamma j} \frac{\partial}{\partial \tilde{R}_{\gamma j}} \text{arctan2}[(\tilde{\boldsymbol{R}}_{N-1} \cdot \tilde{\boldsymbol{R}}_{N-1})(\tilde{\boldsymbol{R}}_\beta \cdot \tilde{\boldsymbol{R}}_{N-2}) - (\tilde{\boldsymbol{R}}_{N-1} \cdot \tilde{\boldsymbol{R}}_\beta)(\tilde{\boldsymbol{R}}_{N-1} \cdot \tilde{\boldsymbol{R}}_{N-2}), \\ &\qquad \mathcal{R}_{N-1}(\tilde{\boldsymbol{R}}_{N-1} \cdot (\tilde{\boldsymbol{R}}_{N-2} \times \tilde{\boldsymbol{R}}_\beta))] \frac{\partial R_{\gamma j}}{\partial R_{\alpha i}} \end{aligned} \tag{3.7.5}$$

which can be expressed as

$$\begin{aligned} s_{\phi_\beta,\alpha i} = \frac{1}{\mathcal{R}_{N-1}^2 \mathcal{R}_\beta \mathcal{R}_{N-2} \sin\theta_\beta \sin\theta_{N-2}} [ \\ - A_{\beta\alpha} \sin\phi_\beta (\mathcal{R}_{N-1}^2 \tilde{R}_{N-2i} - \tilde{R}_{N-1i} \mathcal{R}_{N-1} \mathcal{R}_{N-2} \cos\theta_{N-2}) \\ + A_{\beta\alpha} \mathcal{R}_{N-1} \cos\phi_\beta (\tilde{\boldsymbol{R}}_{N-1} \times \tilde{\boldsymbol{R}}_{N-2})_i \\ - 2A_{N-1\alpha} \sin\phi_\beta \tilde{R}_{N-1i} \mathcal{R}_\beta \mathcal{R}_{N-2} (\cos\theta_\beta \cos\theta_{N-2} + \sin\theta_\beta \sin\theta_{N-2} \cos\phi_\beta) \\ + A_{N-1\alpha} \sin\phi_\beta (\tilde{R}_{\beta i} \mathcal{R}_{N-1} \mathcal{R}_{N-2} \cos\theta_{N-2} + \tilde{R}_{N-2i} \mathcal{R}_{N-1} \mathcal{R}_\beta \cos\theta_\beta) \\ + A_{N-1\alpha} \cos\phi_\beta (\mathcal{R}_{N-1} + \tilde{R}_{N-1i}) (\tilde{\boldsymbol{R}}_{N-2} \times \tilde{\boldsymbol{R}}_\beta)_i) \\ + A_{N-2\alpha} (-\sin\phi_\beta (\mathcal{R}_{N-1}^2 \tilde{R}_{\beta i} - \tilde{R}_{N-1i} \mathcal{R}_{N-1} \mathcal{R}_\beta \cos\theta_\beta) + \cos\phi_\beta (\mathcal{R}_{N-1} \tilde{\boldsymbol{R}}_\beta \times \tilde{\boldsymbol{R}}_{N-1})_i]. \end{aligned} \tag{3.7.6}$$

Thus, $\boldsymbol{s}_{\phi_\beta,\alpha}$ can be written as

$$\begin{aligned} \boldsymbol{s}_{\phi_\beta,\alpha} = A_{\beta,\alpha} \frac{1}{\mathcal{R}_\beta \sin\theta_\beta} (-\sin\phi_\beta, \cos\phi_\beta, 0), \\ + A_{N-1\alpha} \frac{1}{\mathcal{R}_{N-1}} (\sin\phi_\beta \cot\theta_\beta, \cot\theta_{N-2} - \cos\phi_\beta \cot\theta_\beta, 0), \\ - A_{N-2\alpha} \frac{1}{\mathcal{R}_{N-2} \sin\theta_{N-2}} \hat{\boldsymbol{e}}_Y. \end{aligned} \tag{3.7.7}$$

For the rotational $s$-vectors we have to define three conditions as before. The $z$-axis is aligned with $\tilde{\boldsymbol{R}}_{N-1}$ and the $x$-axis is in the plane generated by $\tilde{\boldsymbol{R}}_{N-1}$ and $\tilde{\boldsymbol{R}}_{N-2}$, which is the same choice of axes as in Section 3.4.4, $\tilde{\boldsymbol{R}}_{N-1}$ takes the place of $\boldsymbol{R}_{ab}$, $\tilde{\boldsymbol{R}}_{N-2}$ of $\boldsymbol{R}_{ca}$. The appropriate Sørensen conditions are thus

$$\begin{aligned} C^{(x)} &= R_{N-1x}, \\ C^{(y)} &= R_{N-1y}, \\ C^{(z)} &= R_{N-2y} \end{aligned} \tag{3.7.8}$$

and therefore the $c$-vectors are

$$\begin{aligned} \boldsymbol{c}_{x,\alpha} &= A_{N-1\alpha} \hat{\boldsymbol{e}}_X, \\ \boldsymbol{c}_{y,\alpha} &= A_{N-1\alpha} \hat{\boldsymbol{e}}_Y, \\ \boldsymbol{c}_{z,\alpha} &= A_{N-2\alpha} \hat{\boldsymbol{e}}_Y. \end{aligned} \tag{3.7.9}$$

The condition Eq. (2.3.22) for $\boldsymbol{c}_{x,\alpha}$, given by $\sum_\alpha A_{N-1\alpha} = 0$, is automatically satisfied as the $\tilde{\boldsymbol{R}}$s are invariant under the translation of $\boldsymbol{R}$s, and likewise for the other $\boldsymbol{c}$s. The $J$ matrix turns out to be

$$J = \begin{pmatrix} 0 & \tilde{R}_{N-1z} & 0 \\ -\tilde{R}_{N-1z} & 0 & \tilde{R}_{N-1x} \\ -\tilde{R}_{N-2z} & 0 & \tilde{R}_{N-2x} \end{pmatrix} = \begin{pmatrix} 0 & \mathcal{R}_{N-1} & 0 \\ -\mathcal{R}_{N-1} & 0 & 0 \\ -\mathcal{R}_{N-2}\cos\theta_{N-2} & 0 & \mathcal{R}_{N-2}\sin\theta_{N-2} \end{pmatrix} \tag{3.7.10}$$

and the rotational $s$-vectors become

$$\begin{aligned} \boldsymbol{s}_{x,\alpha} &= -A_{N-1\alpha}\frac{1}{\mathcal{R}_{N-1}}\hat{\boldsymbol{e}}_X, \\ \boldsymbol{s}_{y,\alpha} &= A_{N-1\alpha}\frac{1}{\mathcal{R}_{N-1}}\hat{\boldsymbol{e}}_X, \\ \boldsymbol{s}_{z,\alpha} &= -A_{N-1\alpha}\frac{\cot\theta_{N-1}}{\mathcal{R}_{N-1}}\hat{\boldsymbol{e}}_Y + A_{N-2\alpha}\frac{1}{\mathcal{R}_{N-2}\sin\theta_{N-2}}\hat{\boldsymbol{e}}_Y. \end{aligned} \tag{3.7.11}$$

In summary, we can write the $s$-vectors as

$$\begin{aligned} \boldsymbol{s}_{R_\beta,\alpha} &= A_{\beta\alpha}(\cos\phi_\beta\sin\theta_\beta, \sin\phi_\beta\sin\theta_\beta, \cos\theta_\beta), \\ \boldsymbol{s}_{\theta_\beta,\alpha} &= A_{\beta\alpha}\frac{1}{\mathcal{R}_\beta}(\cos\theta_\beta\cos\phi_\beta, \cos\theta_\beta\sin\phi_\beta, -\sin\theta_\beta) - A_{N-1\alpha}\frac{1}{\mathcal{R}_{N-1}}(\cos\phi_\beta, \sin\phi_\beta, 0), \\ \boldsymbol{s}_{\phi_\beta,\alpha} &= A_{\beta,\alpha}\frac{1}{\mathcal{R}_\beta\sin\theta_\beta}(-\sin\phi_\beta, \cos\phi_\beta, 0), \\ &\quad + A_{N-1\alpha}\frac{1}{\mathcal{R}_{N-1}}(\sin\phi_\beta\cot\theta_\beta, \cot\theta_{N-2} - \cos\phi_\beta\cot\theta_\beta, 0), \\ &\quad - A_{N-2\alpha}\frac{1}{\mathcal{R}_{N-2}\sin\theta_{N-2}}\hat{\boldsymbol{e}}_Y, \\ \boldsymbol{s}_{x,\alpha} &= -A_{N-1\alpha}\frac{1}{\mathcal{R}_{N-1}}\hat{\boldsymbol{e}}_X, \\ \boldsymbol{s}_{y,\alpha} &= A_{N-1\alpha}\frac{1}{\mathcal{R}_{N-1}}\hat{\boldsymbol{e}}_Y, \\ \boldsymbol{s}_{z,\alpha} &= -A_{N-1\alpha}\frac{\cot\theta_{N-1}}{R_{N-1}}\hat{\boldsymbol{e}}_Y + A_{N-2\alpha}\frac{1}{R_{N-2}\sin\theta_{N-2}}\hat{\boldsymbol{e}}_Y. \end{aligned} \tag{3.7.12}$$

To generate the $G$ matrix elements from these vectors, we must compute

$$G_{ab} = \sum_\alpha \frac{1}{m_\alpha}\boldsymbol{s}_{a,\alpha}\cdot\boldsymbol{s}_{b,\alpha} \tag{3.7.13}$$

where we see that the only factor that varies with $\alpha$ is the element of the $A$ matrix. Borrowing the notation of Ref. [65], we define $M_{\beta\gamma} = \sum_\alpha A_{\beta\alpha}m_\alpha^{-1}A_{\gamma\alpha} = (AmA^{-1})_{\beta\alpha}$, where $m$ is a diagonal matrix and $m_{\alpha\alpha} = 1/m_\alpha$. With the caveat that $s_{\theta_\beta,\alpha}$ has to be multiplied by $-\sin\theta_\beta$ before evaluating the $G$ matrix in order to fit the coordinate definitions of Ref. [65], we see that the results of Appendix B of Ref. [65] are matched.

## 3.8 Practical implementation in TROVE

*Fortran*, which lacks the capability of symbolic manipulation, necessitated the usage of an external program – using the Sørensen method – to derive analytic KEOs in the sum-of-products form of 1D elementary functions. *Python* was tested, but its symbolic manipulation packages were deemed not as powerful as *Mathematica*, which was ultimately used. The *Mathematica* script which generates the KEO was written to be as general as possible, though each molecule has its own peculiarities. A version of the script is available in the supplementary material of Ref. [2].

In the Sørensen approach of Section 2.3.2, one must define (for a given molecule) the masses of the nuclei, the vibrational coordinates $q(R) = q(r)$ as functions of the Cartesian coordinates, the three conditions $C^{(g)}(r)$ in terms of the BF coordinates, and, finally, the BF coordinates $r(q)$ in terms of the vibrational coordinates. Here, the $r$ in $q(r)$ and $C^{(g)}(r)$ represents the labels for the BF coordinates and likewise the $q$ in $r(q)$ represents the labels for the vibrational coordinates.

At this point in the calculation it is worthwhile to check that the $C^{(g)}(r)$ conditions are zero (Eq. (2.3.18)) by substituting the Cartesian coordinates $r$ in $C^{(g)}(r)$ with $r(q)$ to form $C^{(g)}(r(q))$. While $C^{(g)}(r)$ is only a function of the $r$ labels and therefore not identically zero, we recall that the BF coordinates $r(q)$ assume the rotational conditions and therefore $C^{(g)}(r(q))$ must be zero.

In a same vein, the coordinates $q(r)$ should be checked to be correctly defined by also substituting $r$ with $r(q)$. Since $r(q)$ and $q(r)$ are effectively inverse operations, $r(q(r))$, after simplification, should be the identity operation. It may happen that the expression is too complicated for the analytic software to fully simplify the expression such that the result is evidently an identity operation. In that case, for the function representing $q_i(q)$, one may replace all other vibrational coordinates with random values and then plot the resulting function (of $q_i$) as a function of $q_i$. The result again should be the identity operation. However, as the $q$ contain trigonometric functions, which are periodic and hence do not have a well defined inverse across their entire domain, there may be discontinuities in this plot.

Once these are obtained, the remainder of the script should in theory be practically identical for all molecules. First the $s$-vector $s_{k,\alpha}(r)$ components as functions of the BF coordinates are determined. The Cartesian coordinates $r$ must be substituted with $r(q)$ to form $s_{k,\alpha}(r(q))$. The resulting expressions are typically not in their simplest form, or even in the sum-of-product form. They must be simplified by the software.

In *Mathematica* there is an added complication of what constitutes "simplification." If it can combine trigonometric functions using product-to-sum identities, such as

$$\sin\alpha\cos\beta + \sin\beta\cos\alpha = \sin(\alpha+\beta), \tag{3.8.1}$$

it will, but these are not helpful in this application as the expressions must be separated again to force the sum-of-product form on the $s$-vectors. Moreover, comprehensive simplification in *Mathematica* is time consuming. The identities actually applied should be manually – and wisely – chosen. The essential identity is the Pythagorean identity $\sin^2\alpha + \cos^2\alpha = 1$. Other useful identities are the trigonometric identities

$$\begin{aligned} 2\sin\alpha\cos\alpha &= \sin(2\alpha), \\ \cos^2\alpha - \sin^2\alpha &= \cos(2\alpha) \end{aligned} \tag{3.8.2}$$

and *Mathematica* can be set to limit its trigonometric simplification to just those. Sometimes one encounters the need for the identities

$$\begin{aligned} 1 - \tan^2\alpha &= \sec^2\alpha, \\ 1 - \cot^2\alpha &= \csc^2\alpha. \end{aligned} \tag{3.8.3}$$

Finally, to prevent unwanted simplification present when including the `Automatic` keyword, *Mathematica* may be forced to simplify expressions such as $\sqrt{f(x)^2 f(y)^2}$ to $\sqrt{f(x)^2}\sqrt{f(y)^2}$ and then $\sqrt{f(x)^2}$ to $f(x)$ (i.e. positive roots). There may be trial and error involved in choosing the correct simplification rules to apply to maximise the speed but also to guarantee a sum-of-product form while minimising the size of the $s$-vectors.

For some molecules, a trigonometric function may be a function of a linear combination of the dihedral angles rather than a single dihedral angle. Despite the above recommendation, it is preferable to also not expand any trigonometric functions using the sum-to-product rules until a later stage to maximise the trigonometric simplifications first: these are trickier for the software to find after the sum-to-product identities are applied.

If the initial BF-frame is not the desired frame, the $s$-vector components of the latter can be calculated. After they are found, the above simplification procedure can be applied. Then the $s$-vector components of the initial BF frame can be used to find the vibrational $G$ matrix components and the pseudo potential – these are frame independent – and the new $s$-vector components should be used to find the Coriolis and rotational $G$ matrix components. Once again, they can be simplified with the above procedure.

At this stage, the trigonometric sum-to-product identities should be applied so that the KEO is in the sum-of-product form, but it can still be simplified. This step concludes the calculation of the KEO in the software. The subsequent steps are specific to TROVE, which must be able to read the KEO correctly. The result of these steps is a list of 1D elementary functions and the expansion coefficients. Currently, they are numerical and depend on the initial masses of the nuclei.

In TROVE, each term of Eq. (3.2.2) is numbered by

$$(i_1, \ldots, i_{3N-6}) \rightarrow n_{i_1,\ldots,i_{3N-6}} \tag{3.8.4}$$

and this $3N-6$ dimensional index is converted to a single index $n_{i_1,\ldots,i_{3N-6}}$. Actually, TROVE iterates through each possible value of the index (for a given polyad number) and converts it to the $3N-6$ dimensional index. How this number works in practice is, to find the $n+1$ term with the $n$th term $(i_j, \ldots, i_{3N-6})$, one finds the first non-zero $i_j$ starting from the right, adds 1 to $i_{j-1}$, and then sets $i_{3N-6} = i_j - 1$ and $i_j = 0$ if $j \neq 3N-6$. An example of this is shown in Table 3.2. Once $i_1 = p < p_{\max}$ and the rest zero, the next term is $i_{3N-6} = p + 1$ are the rest zero. In this way, a lookup table is created from the single index to the $3N-6$ dimensional index.

In the analytic KEO approach presented here, term numbers $(i_1, \ldots, i_{3N-6})$ are taken from Eq. (3.2.1). Therefore, the reverse procedure of obtaining the single index from the $3N-6$ dimensional index is required. TROVE's previous implementation of this used an inefficient recursive function which was rarely needed, so an alternative approach was

Table 3.2: The TROVE number on the left column for each term on the right column

| Number | Term |
|---|---|
| 1120 | (2, 1, 2, 3, 0) |
| 1121 | (2, 1, 3, 0, 2) |
| 1122 | (2, 1, 3, 1, 1) |
| 1123 | (2, 1, 3, 2, 0) |
| 1124 | (2, 1, 4, 0, 1) |
| 1225 | (2, 1, 4, 1, 0) |
| 1226 | (2, 1, 5, 0, 0) |
| 1227 | (2, 2, 0, 0, 4) |

necessary. A direct mapping can be accomplished through the formula

$$n_{i_1,\ldots,i_{3N-6}} = 1 + \binom{p_{3N-6} + 3N - 6 - 1}{3N-6} + \sum_{\substack{j=1 \\ i_j>0}}^{3N-6} \sum_{k=0}^{i_j} \binom{p_{3N-6} - p_{j-1} + 3N - 6 - j - 1 - k}{3N-6-p-j-1} \tag{3.8.5}$$

where the left-most sum is over all $i_j$ greater than zero and

$$p_k = \sum_{j=1}^{k} i_j. \tag{3.8.6}$$

For TROVE to correctly read the analytic 1D functions $f_{i_j}(q_i)$, each function of $q_i$ in the KEO in *Mathematica* is converted to the form $fi[j]$ and TROVE is told which functions the numbering corresponds to. For the first task, the *Mathematica* script determines every 1D function present in the KEO. These are sorted according to the "total-power" of each "fundamental" function which is the sum of the exponents of the fundamental functions in a 1D function. For example, the 1D function of the form $\sin^2 q_i \cos q_i$ has total power $3 = 2 + 1$ from the fundamental functions $\sin q_i$ and $\cos q_i$, respectively. The 1D function mapping for all functions of $q_i$ could be

$$(\sin^2 q_i \cos q_i \to fi[4], \cos^2 q_i \to fi[3], \cos q_i \to fi[2], \sin q_i \to fi[1]) \tag{3.8.7}$$

where the mappings are ordered from the highest total-power to lowest. This is the order with which the substitutions are made. This would prevent, for example, $\cos q_i$ being used in the substitution of $\sin^2 q_i \cos q_i$ to $\sin^2 q_i fi[3]$, which is not correct.

With the substituted KEO, each expansion term can be represented in the form $(a_p, i_1, \ldots, i_{3N-6})$ where $a_p$ is the coefficient. For the problem of TROVE reading this input, there are a few approaches. The terms can be numbered by Eq. (3.8.5) so TROVE only needs to be told which function each number corresponds. One approach, which was effectively taken for programming the KEO for the homo- and heteronuclear triatomics, is hard coding the functions into TROVE. This is not an ideal solution and is untenable for a large KEO. It also goes against the black-box philosophy of TROVE.

An alternative approach, one used for $H_2CS$, uses an updated TROVE which reads the mapping (specified in the input) on-the-fly. Each fundamental function can effectively be expressed in the form $g(bq^n)^m$. The numbering $n$ and $m$ may not be unique, but the mapping need only be correct. The function $g$ is typically the identity or a trigonometric function. A 1D function of the form

$$g_1(b_1 q^{n_1})^{m_1} \dots g_r(b_r q^{n_r})^{m_r} \tag{3.8.8}$$

is expressed, in a text format, as

```
r m1 g1 b1 n1 m2 g2 b2 n2 ... m2 gr br nr
```
(3.8.9)

which is part of the `BASIC-FUNCTION` block in the input file read by TROVE. Pointer functions in *Fortran* are used to specify the fundamental function $g_i$ on-the-fly.

This concludes the steps necessary to construct a KEO for a molecule such as $H_2CS$. For larger molecules, such as $C_2H_6$, numbering by Eq. (3.8.5) is no longer tenable. Ordinarily, the expansion order (the polyad $p_{\max}$) is specified and is typically around 6. For a molecule with a small to medium number of modes, the number of possible terms $(i_1, \dots, i_{3N-6})$ where $i_1 + \dots + i_{3N-6} \leq p_{\max}$ is reasonable. In TROVE, an array of this size is created before the KEO is read (or created) and then its values are assigned according to the coefficients. In the case of the analytic KEO the polyad has no physical meaning and is just a convenient way to inform TROVE of the size of the necessary array. For $C_2H_6$ this array becomes impractically large as there are 18 modes and up to 13 functions for a single mode. Another method is needed.

To combat this issue, a modification was made to TROVE which now can read the input that includes the numbers $(i_1, \dots, i_{3N-6})$. Since TROVE was not designed with this in mind, the redundant polyad and number of terms are still required in the input, although these numbers are not utilised. Modifying TROVE in this way was a significant challenge as much of the code rests upon the assumption that the numbering of the terms is according to Eq. (3.8.5).

## 3.9 Chapter summary

This chapter delved into the procedure implemented in *Mathematica* and TROVE to use analytic KEOs. The general Sørensen methodology as well as how to implement it practically was described. The relation between the $s$-vectors components for different BF-frames was proven and used to deduce the functions of the valence coordinates present in the $s$-vectors and by extension the KEO. This approach was taken to re-derive the KEO for polyspherical coordinates stated by Ref. [65]. Finally, the implementation was tested for $H_2CO$ with the results compared to previously published values.

## Appendix 3.A BF-frame $s$- and $t$-vector components

This section will demonstrate that BF-frame components of the $s$- and $t$-vectors can be found by evaluating their functions using the BF coordinates $r(q)$.

First, the vibrational $s$-vectors in any frame are found by evaluating $\partial_{\alpha i} q_k$ in the Cartesian coordinates of that frame. To show this, we use the property that the value of $q_k$ does not depend on which frame it is evaluated in, i.e.

$$q_k(MR_1 + d, \dots, MR_n + d) = q_k(R_1, \dots, R_n). \tag{3.A.1}$$

where $MR_\alpha + d$ are the coordinates of $\boldsymbol{R}_\alpha$ when the frame is rotated and translated. Calling $T(R) = MR + d$, we want to determine $\partial_{\alpha i}(q_k)(T(R_1), \dots, T(R_n))$ in terms of $\partial_{\alpha i}(q_k)(R_1, \dots, R_n)$. Let us define the function $\tilde{T}$ as

$$\tilde{T}(R_1, \dots R_n) = (T(R_1), \dots, T(R_n)). \tag{3.A.2}$$

Then, the translational and rotational invariance can be expressed as

$$q_k \circ \tilde{T}(R_1, \dots, R_n) = q_k(R_1, \dots, R_n). \tag{3.A.3}$$

Differentiating both sides, and using the chain rule, we obtain

$$\partial_{\alpha F}(q_k)(R_1, \dots, R_n) = \partial_{\alpha F}(q_k \circ \tilde{T})(R_1, \dots, R_n) \tag{3.A.4}$$

$$= \sum_\beta \partial_{\alpha F}(\tilde{T})^{\beta G}(R_1, \dots, R_n) \partial_{\beta G}(q_k)(\tilde{T}(R_1, \dots, R_N)) \tag{3.A.5}$$

Here $\partial_{\alpha F}(\tilde{T})^{\beta G}$ is the $\beta G$ component of the derivative of $\tilde{T}$, which is non-zero only when $\beta = \alpha$ and is given by $M_{GF}$. We thus have

$$\partial_{\alpha F}(q_k)(R_1, \dots, R_N) = M_{GF} \partial_{\alpha G}(q_k)(T(R_1), \dots, T(R_n)). \tag{3.A.6}$$

Thus

$$\begin{aligned} \boldsymbol{s}_{k,\alpha} = \hat{\boldsymbol{e}}_F \partial_{\alpha F}(q_k)(R_1, \dots, R_N) &= M_{GF} \hat{\boldsymbol{e}}_F \partial_{\alpha G}(q_k)(T(R_1), \dots, T(R_n)) \\ &= \hat{\boldsymbol{e}}_G \partial_{\alpha G}(q_k)(T(R_1), \dots, T(R_n)) \end{aligned} \tag{3.A.7}$$

where $\hat{\boldsymbol{e}}_G = M_{GF} \hat{\boldsymbol{e}}_F$. Since the Cartesian components in the new frame are given by $M_{GF} R_{\alpha F} + d$, $\hat{\boldsymbol{e}}_G$ are the unit vectors in the same frame, which tells us the component of $\boldsymbol{s}_{k,\alpha}$ is the derivative of $q_k$ evaluated at the Cartesian coordinates of that frame. This gives a computational way to determine the components of $\boldsymbol{s}_{k,\alpha}$ in any frame that does not involve first evaluating the components in the SF-frame and applying the rotation matrix. In particular, it provides a way to find the BF components.

The remaining vectors are easier. The translation $t$-vector is

$$\boldsymbol{t}_{F,\alpha} = \hat{\boldsymbol{e}}_F = M_{Fg} \hat{\boldsymbol{e}}_g \tag{3.A.8}$$

so that the BF-frame components are $M_F = (M_{Fx}, M_{Fy}, M_{Fz})$. The translational $s$-vector is similar.

For the rotational $t$-vector, we have

$$\begin{aligned} \boldsymbol{t}_{g,\alpha} &= \hat{\boldsymbol{e}}_g \times \boldsymbol{R}_\alpha \\ &= \hat{\boldsymbol{e}}_g \times r_{\alpha h} \hat{\boldsymbol{e}}_h \\ &= r_{\alpha h} \varepsilon_{ghf} \hat{\boldsymbol{e}}_f \end{aligned} \tag{3.A.9}$$

so $t_{g,\alpha f} = \varepsilon_{ghf} r_{\alpha h}$. Finally, the rotational $s$-vector components are defined in the BF-frame.

## Appendix 3.B Derivation of the $s$-vector relations

In this section we shall derive the relations of the $s$-vectors under a change in BF-frame, as stated in the main text. All vector components are assumed to be in the BF-frame. First, we note that the translational $s$- and $t$-vector components are invariant under a change of BF-frame. The coordinates $r_\alpha$ of the original BF-frame become $Sr_\alpha$ under a change of frame, where $S$ is the rotation matrix. Here we ignore translation since $r_\alpha$ are assumed to be calculated with the origin at the centre of nuclear mass. Then, the vibrational $t$-vector components $t_{k,\alpha}$ relate as

$$t_{k,\alpha} = \partial_k r_\alpha \rightarrow \partial_k (S r_\alpha) = S t_{k,\alpha} + \frac{\partial S}{\partial q_k} r_\alpha = S \left( t_{k,\alpha} + \sum_g \omega_{k,g} t_{\alpha,g} \right). \tag{3.B.1}$$

where $\omega_k = (\omega_{k,x}, \omega_{k,y}, \omega_{k,z})$ is defined by

$$S^T \frac{\partial S}{\partial q_k} s = (\omega_k \times s). \tag{3.B.2}$$

Next, we consider the vibrational $s$-vectors $s_{k,\alpha}$. We may again use the argument of Appendix 3.A that all vibrational coordinates $q_k$ are invariant under $r_\alpha \rightarrow S r_\alpha + d$, where $d$ are displacement coordinates. We find

$$\partial_{\alpha i}(q_k)(r_1, \ldots, r_N) = S_{ji} \partial_{\alpha j}(q_k)(\tilde{T}(r_1, \ldots, r_N)) \tag{3.B.3}$$

which can be inverted to give

$$\partial_{\alpha m}(q_k)(\tilde{T}(r_1, \ldots, r_N)) = S_{mi} \partial_{\alpha i}(q_k)(r_1, \ldots, r_N). \tag{3.B.4}$$

This is the required result. Thus

$$s_{k,\alpha} \rightarrow S s_{k,\alpha} \tag{3.B.5}$$

which means they are related as the Cartesian coordinates.

We must also check the vibrational $t$- and $s$- vectors behave properly as inverses:

$$\begin{aligned} \sum_\alpha {s_{k,\alpha}}^T t_{k',\alpha} &\rightarrow \underbrace{\sum_\alpha {s_{k,\alpha}}^T t_{k',\alpha}}_{\delta_{kk'}} + \sum_g \omega_{k',g} \underbrace{\sum_\alpha {s_{k,\alpha}}^T t_{g,\alpha}}_{0} \\ &= \delta_{kk'}. \end{aligned} \tag{3.B.6}$$

We also note that the invariance of the vibrational $G$ elements is satisfied, as expected.

To determine the relation of the rotational $s$-vectors, we first calculate the relation of the rotational $t$-vectors, which are

$$t_{g,\alpha i} = \varepsilon_{igj} r_{\alpha j} \rightarrow \varepsilon_{igj} S_{jm} r_{\alpha m}. \tag{3.B.7}$$

This can be expressed in terms of the original $t$-vectors components:

$$\begin{aligned} S_{gh} t_{h,\alpha j} S_{ij} &= S_{gh} S_{ij} \varepsilon_{jhm} r_{\alpha m} \\ &= S_{gh} S_{ij} S_{pn} \varepsilon_{ihn} S_{pm} r_{\alpha m} = \det(S) \varepsilon_{igp} S_{pm} r_{\alpha m} = \varepsilon_{igp} S_{pm} r_{\alpha m}. \end{aligned} \tag{3.B.8}$$

Viewing $t_{g,\alpha}$ as the matrix $t_{\mathrm{r},\alpha}$, where $g$ signifies the columns, this can be expressed as

$$t_{\mathrm{r},\alpha} \rightarrow S t_{\mathrm{r},\alpha} S^T. \tag{3.B.9}$$

For the transformation of the rotational $s$-vectors, we assert that they become

$$s_{g,\alpha i} \rightarrow s'_{g,\alpha i} = S_{gh} S_{ij} s_{h,\alpha j} - \sum_{\beta} \frac{\partial S_{mn}}{\partial q_k} s_{k,\alpha j} s_{h,\beta p} S_{mp} r_{\beta n} S_{ij} S_{gh} \tag{3.B.10}$$

where $k \in \{1, \ldots, 3N-6\}$ ranges over the vibrational coordinates and $\beta \in \{1, \ldots, N\}$ ranges over the atoms. This can be simplified to and rewritten for the rotational $s$-matrix as

$$s_{\mathrm{r},\alpha} \rightarrow S \left( s_{\mathrm{r},\alpha} - \sum_{k'} \omega_{k'} {s_{k',\alpha}}^T \right) S^T. \tag{3.B.11}$$

To show that this is the correct transformation, we will show that this is the inverse of the transformed $t$-matrix. First, with the translational $t$-vectors, we have

$$\begin{aligned} \sum_{\alpha} s_{\mathrm{r},\alpha} t_{F,\alpha} &\rightarrow \sum_{\alpha} S \left( s_{\mathrm{r},\alpha} - \sum_{k'} \omega_{k'} {s_{k',\alpha}}^T \right) S^T M_F \\ &= S \left( \underbrace{\sum_{\alpha} s_{\mathrm{r},\alpha}}_{0} - \sum_{k'} \omega_{k'} \left( \underbrace{\sum_{\alpha} {s_{k',\alpha}}^T}_{0} \right) \right) S^T M_F \end{aligned} \tag{3.B.12}$$

where $M_F = (M_{Fx}, M_{Fy}, M_{Fz})$ and $M$ is Euler matrix of Eq. (2.3.2) and both terms are zero due to condition Eq. (2.3.24) on the rotational and vibrational $s$-vectors. For the rotational $t$-vectors, have have

$$\begin{aligned} \sum_{\alpha} s_{\mathrm{r},\alpha} t_{\mathrm{r},\alpha} &\rightarrow \sum_{\alpha} S \left( s_{\mathrm{r},\alpha} - \sum_{k'} \omega_{k'} {s_{k',\alpha}}^T \right) t_{\mathrm{r},\alpha} S^T \\ &= S \left( \underbrace{\sum_{\alpha} s_{\mathrm{r},\alpha} t_{\mathrm{r},\alpha}}_{I} - \sum_{k'} \omega_{k'} \left( \underbrace{\sum_{\alpha} {s_{k',\alpha}}^T t_{\mathrm{r},\alpha}}_{0} \right) \right) S^T \\ &= S S^T = I. \end{aligned} \tag{3.B.13}$$

Finally, for the vibrational $t$-vectors, we have

$$\begin{aligned}\sum_\alpha s_{\mathrm{r},\alpha} t_{k,\alpha} &\to \sum_\alpha S\left(s_{\mathrm{r},\alpha} - \sum_{k'} \omega_{k'} {s_{k',\alpha}}^T\right)\left(t_{k,\alpha} + \sum_g \omega_{k,g} t_{g,\alpha}\right)\\ &= S\left(\underbrace{\sum_\alpha s_{\mathrm{r},\alpha} t_{k,\alpha}}_{0} + \underbrace{\sum_{g,\alpha} \omega_{k,g} s_{r,\alpha} t_{\alpha,g}}_{\omega_k} - \underbrace{\sum_{\alpha,k'} \omega_{k'} {s_{k',\alpha}}^T t_{k,\alpha}}_{\omega_k} + \underbrace{\sum_{\alpha,g,k'} \omega_{k,g}\omega_{k'} {s_{k',\alpha}}^T t_{g,\alpha}}_{0}\right)\\ &= S(\omega_k - \omega_k) = 0\end{aligned} \tag{3.B.14}$$

which concludes the proof.

## Appendix 3.C Invariance of the pseudo potential

The pseudo potential $U$ consists of the four terms:

$$\begin{aligned}U_1 &= \frac{1}{8}\sum_{\alpha,g,g'} \frac{1}{m_\alpha}(\hat{\boldsymbol{e}}_g \times \boldsymbol{s}_{g,\alpha}) \cdot (\hat{\boldsymbol{e}}_{g'} \times \boldsymbol{s}_{g',\alpha}),\\ U_2 &= \frac{1}{4}\sum_{\alpha,g,k} \frac{1}{m_\alpha}\left(\hat{\boldsymbol{e}}_g \times \boldsymbol{s}_{k,\alpha}\right) \cdot \frac{\partial \boldsymbol{s}_{g,\alpha}}{\partial q_k},\\ U_3 &= -\frac{1}{4}\sum_{\alpha,k,k'} \frac{1}{m_\alpha} \boldsymbol{s}_{k,\alpha} \cdot \frac{\partial^2 \boldsymbol{s}_{k',\alpha}}{\partial q_k \partial q_{k'}},\\ U_4 &= -\frac{1}{8}\sum_{\alpha,k,k'} \frac{1}{m_\alpha} \frac{\partial \boldsymbol{s}_{k,\alpha}}{\partial q_k} \cdot \frac{\partial \boldsymbol{s}_{k',\alpha}}{\partial q_{k'}}\end{aligned} \tag{3.C.1}$$

where $U$ is the sum of these terms. The first term can be rewritten as (in Einstein summation notation)

$$U_1 = \frac{1}{8}\sum_\alpha \frac{1}{m_\alpha}(s_{g,\alpha i} s_{g,\alpha i} - s_{g,\alpha h} s_{h,\alpha g}) \tag{3.C.2}$$

or, in matrix notation,

$$U_1 = \frac{1}{8}\sum_\alpha \frac{1}{m_\alpha} \mathrm{Tr}(s_{\mathrm{r},\alpha} s_{\mathrm{r},\alpha}^T - s_{\mathrm{r},\alpha} s_{\mathrm{r},\alpha}). \tag{3.C.3}$$

The second term can be rewritten as (in Einstein summation notation)

$$U_2 = \frac{1}{4}\sum_\alpha \frac{1}{m_\alpha} \varepsilon_{igj} s_{k,\alpha j} \frac{\partial s_{g,\alpha i}}{\partial q_k}. \tag{3.C.4}$$

We will briefly outline the methods use to obtain the changes $\Delta U_i$ of each $U_i$ – where $\Delta U_i$ is the difference between the $U_i$ before and after changing the BF-frame – before writing them all out. For $U_1$, we have

$$\begin{aligned}\Delta U_1 &= \frac{1}{8}\sum_\alpha \frac{1}{m_\alpha}\left(\mathrm{Tr}(S\tilde{s}_{\mathrm{r},\alpha}\tilde{s}_{\mathrm{r},\alpha}^T S^T) - \mathrm{Tr}(s_{\mathrm{r},\alpha}{s_{\mathrm{r},\alpha}}^T) - \mathrm{Tr}(S\tilde{s}_{\mathrm{r},\alpha}\tilde{s}_{\mathrm{r},\alpha}S^T) + \mathrm{Tr}(s_{\mathrm{r},\alpha}s_{\mathrm{r},\alpha})\right)\\ &= \frac{1}{8}\sum_\alpha \frac{1}{m_\alpha}\left(\mathrm{Tr}(\tilde{s}_{\mathrm{r},\alpha}\tilde{s}_{\mathrm{r},\alpha}^T) - \mathrm{Tr}(s_{\mathrm{r},\alpha}{s_{\mathrm{r},\alpha}}^T) - \mathrm{Tr}(\tilde{s}_{\mathrm{r},\alpha}\tilde{s}_{\mathrm{r},\alpha}) + \mathrm{Tr}(s_{\mathrm{r},\alpha}s_{\mathrm{r},\alpha})\right)\end{aligned} \tag{3.C.5}$$

where

$$\tilde{s}_{\mathrm{r},\alpha} = s_{\mathrm{r},\alpha} - \sum_k \omega_k {s_{k,\alpha}}^T. \tag{3.C.6}$$

An example term would be

$$\mathrm{Tr}(\omega_k {s_{k,\alpha}}^T s_{k',\alpha}{\omega_{k'}}^T) = (\omega_k\cdot\omega_{k'})(s_{k,\alpha}\cdot s_{k',\alpha}). \tag{3.C.7}$$

For $U_2$, we have

$$\Delta U_2 = \frac{1}{4}\sum_\alpha \frac{1}{m_\alpha}\left(\varepsilon_{igj}S_{jl}s_{k,\alpha l}\frac{\partial}{\partial q_k}(S_{gh}S_{im}\tilde{s}_{h,\alpha m}) - \varepsilon_{igj}s_{k,\alpha j}\frac{\partial s_{g,\alpha i}}{\partial q_k}\right). \tag{3.C.8}$$

A term such as

$$\varepsilon_{igj}S_{jl}s_{k,\alpha l}\frac{\partial S_{gh}}{\partial q_k}S_{im}\tilde{s}_{h,\alpha m} \tag{3.C.9}$$

becomes

$$\begin{aligned}\varepsilon_{ig'j}S_{im}S_{g'n}S_{jl}s_{k,\alpha l}S_{gn}\frac{\partial S_{gh}}{\partial q_k}\tilde{s}_{h,\alpha m} &= \varepsilon_{mnl}\varepsilon_{ngh}s_{k,\alpha l}\omega_{k,g}\tilde{s}_{h,\alpha m}\\ &= \mathrm{Tr}(\tilde{s}_{\mathrm{r},\alpha})(s_{k,\alpha}\cdot\omega_k) - {s_{k,\alpha}}^T\tilde{s}_{\mathrm{r},\alpha}\omega_k.\end{aligned} \tag{3.C.10}$$

For $U_3$, we have

$$\Delta U_3 = -\frac{1}{4}\sum_{\alpha,k,k'}\frac{1}{m_\alpha}\left(Ss_{k,\alpha}\cdot\frac{\partial^2 Ss_{k',\alpha}}{\partial q_k\partial q_{k'}} - s_{k,\alpha}\cdot\frac{\partial^2 s_{k',\alpha}}{\partial q_k\partial q_{k'}}\right). \tag{3.C.11}$$

A term such as

$$Ss_{k,\alpha}\cdot\frac{\partial}{\partial q_k}\left(\frac{\partial S}{\partial q_{k'}}s_{k',\alpha}\right) \tag{3.C.12}$$

becomes

$$\begin{aligned}Ss_{k,\alpha}\cdot\frac{\partial}{\partial q_k}\left(\frac{\partial S}{\partial q_{k'}}s_{k',\alpha}\right) = Ss_{k,\alpha}\cdot\frac{\partial}{\partial q_k}\left(SS^T\frac{\partial S}{\partial q_{k'}}s_{k',\alpha}\right) &= Ss_{k,\alpha}\cdot\frac{\partial}{\partial q_k}\left(S\omega_{k'}\times s_{k',\alpha}\right)\\ &= (s_{k,\alpha}\cdot(\omega_k\times(\omega_{k'}\times s_{k',\alpha})) + (s_{k,\alpha}\cdot\left(\frac{\partial}{\partial q_k}(\omega_{k'}\times s_{k',\alpha})\right).\end{aligned} \tag{3.C.13}$$

A similar procedure can be used for $U_4$.

With these expressions the differences $\Delta U_i$s are given by

$$
\begin{aligned}
\Delta U_1 =& + \frac{1}{8} \sum_{\alpha k k'} ((\omega_k \cdot \omega_{k'})(s_{k,\alpha} s_{k',\alpha}) - (\omega_k \cdot s_{k',\alpha})(s_{k,\alpha} \cdot \omega_{k',\alpha})) \\
& + \frac{1}{4} \sum_{\alpha k} \left( s_{k,\alpha}^T s_{\mathrm{r},\alpha} \omega_k - {\omega_k}^T s_{\mathrm{r},\alpha} s_{k,\alpha} \right), \\
\Delta U_2 =& + \frac{1}{4} \sum_{\alpha k k'} (\omega_k \cdot s_{k',\alpha})(s_{k,\alpha} \cdot \omega_{k'}) - (\omega_k \cdot \omega_{k'})(s_{k,\alpha} \cdot s_{k',\alpha})) \\
& + \frac{1}{4} \sum_{\alpha k} \left( {\omega_k}^T s_{\mathrm{r},\alpha} s_{k,\alpha} - {s_{k,\alpha}}^T s_{\mathrm{r},\alpha} \omega_k \right) \\
& + \frac{1}{4} \sum_{\alpha k k'} \left( s_{k,\alpha} \cdot \frac{\partial}{\partial q_k} (\omega_{k'} \times s_{k',\alpha}) \right), \\
\Delta U_3 =& - \frac{1}{4} \sum_{\alpha k k'} \left( s_{k,\alpha} \cdot \frac{\partial}{\partial q_k} (\omega_{k'} \times s_{k',\alpha}) + s_{k,\alpha} \cdot \left( \omega_{k'} \times \frac{\partial s_{k'}}{\partial q_{k'}} \right) \right) \\
& - \frac{1}{4} \sum_{\alpha k k'} ((\omega_k \cdot s_{k',\alpha})(s_{k,\alpha} \cdot \omega_{k'}) - (\omega_k \cdot \omega_{k'})(s_{k,\alpha} \cdot s_{k',\alpha})), \\
\Delta U_4 =& - \frac{1}{8} \sum_{\alpha k k'} (\omega_k \cdot s_{k',\alpha})(s_{k,\alpha} \cdot \omega_{k'}) - (\omega_k \cdot \omega_{k'})(s_{k,\alpha} \cdot s_{k',\alpha})) \\
& - \frac{1}{8} \sum_{\alpha k k'} \left( -2 s_{k,\alpha} \cdot \left( \omega_{k'} \times \frac{\partial s_{k'}}{\partial q_{k'}} \right) \right).
\end{aligned}
\tag{3.C.14}
$$

An inspection of these expressions show that $\Delta U = \Delta U_1 + \Delta U_2 + \Delta U_3 + \Delta U_4 = 0$, as expected.

# Chapter 4

# MARVELised $H_2CS$ Line List

## 4.1 Introduction

Since its original interstellar detection [68], thioformaldehyde (main isotopologue $^1H_2{}^{12}C^{32}S$, henceforth referred to as $H_2CS$) has been observed in a variety of astronomical environments. For example, in massive star-forming regions such as Sagittarius B2 [69] where it can be efficiently formed through bimolecular reactions [70], in nearby [71, 72] and more distant galaxies [73], molecular clouds [74], and in the comet Hale-Bopp [75] which was the first detection of thioformaldehyde in a comet. Given that it is a simple organosulphur molecule, thioformaldehyde can be expected in the atmospheres of exoplanets where sulphur chemistry is known to play a significant role in atmospheric composition [76, 77]. The need for accurate and complete molecular opacity data of $H_2CS$ is the motivation for the production of a hot line list for the ExoMol database [8] and is the subject of this chapter.

It is broadly split into two sections, based on the publications Refs. [1, 2]. This former was an extensive spectroscopic literature search of $H_2CS$ which extracted all meaningful transitions into a consistent, labelled dataset with measurement uncertainties. This dataset was processed using the MARVEL (Measured Active Rotational-Vibrational Energy Levels) algorithm [78–81] which produced an extensive list of ro-vibrational energy levels of $H_2CS$ in its ground electronic state.

The MARVEL dataset was used in the computation of a new comprehensive line list of $H_2CS$, which is the subject of the second half of the chapter. This calculation improved upon previous line lists by empirically-refining the PES with the MARVEL values. Moreover, after the line list was computed, the calculated energy levels were replaced with the empirically-derived MARVEL values. The variational ro-vibrational calculations also employed an analytic KEO discussed in Chapter 3.

## 4.2 MARVEL procedure

### 4.2.1 The approach and input structure

Based on the theory of spectroscopic networks [80, 82–84], MARVEL takes a user-constructed dataset of assigned spectroscopic transitions and converts them into a consistent set of labelled energy levels with the measurement uncertainties propagated

from the input transitions to the output energies. To convert the transitions, it essentially minimises the weighted quadratic sum

$$\sum_i \frac{1}{\sigma_i^2}\left(t_i - \sum_j X_{ij}E_j\right)^2 \tag{4.2.1}$$

where $t_i$ and $\sigma_i$ are the $i$th transition and associated uncertainty, respectively, $E_j$ is the $j$th energy level, and $X_{ij}$ is

$$X_{ij} = \begin{cases} 1 & j \text{ is the upper state for transition } i \\ -1 & j \text{ is the lower state for transition } i \\ 0 & \text{otherwise.} \end{cases} \tag{4.2.2}$$

Differentiating with respect to $E_k$, equating to zero, and rearranging , we obtain

$$X_{ki}^T \sum_i \frac{1}{\sigma_i^2} t_i = X_{ki}^T \sum_i \frac{1}{\sigma_i^2} X_{ij} E_j \tag{4.2.3}$$

which is a set of linear equations which can be solved for the energy levels.

From a user-perspective, the MARVEL input dataset of transitions has the general structure

$$\nu \quad \text{in. unc} \quad \text{up. unc} \quad \text{QN}' \quad \text{QN}'' \quad \text{source}.i$$

where $\nu$ is the transition wavenumber ($cm^{-1}$), in. unc is the initial measurement uncertainty ($cm^{-1}$), up. unc is the updated uncertainty ($cm^{-1}$) which may differ from the initial uncertainty, QN′ and QN″ are the quantum numbers of the upper and lower states involved in the transition, respectively, and source.$i$ is the literature source tag concatenated with a counting number $i$ of the datum from this source. source.$i$ is a unique ID of the transition in the dataset and is required by MARVEL. The source tag, for example 72BeKlKiJo, is based on the notation employed by the IUPAC task group on water [85, 86] and is generally speaking formed from the year published and the first two letters of, up to, the initial four authors' surnames. An extract of the MARVEL input file of $H_2CS$ containing the labelled transitions is shown in Table 4.1.

Table 4.1: Extract from the MARVEL transition file. The MARVEL frequency wavenumber $\nu$ and uncertainties are in $cm^{-1}$. There are two uncertainty columns to allow the input uncertainty to be updated while retaining the original uncertainty of the source.

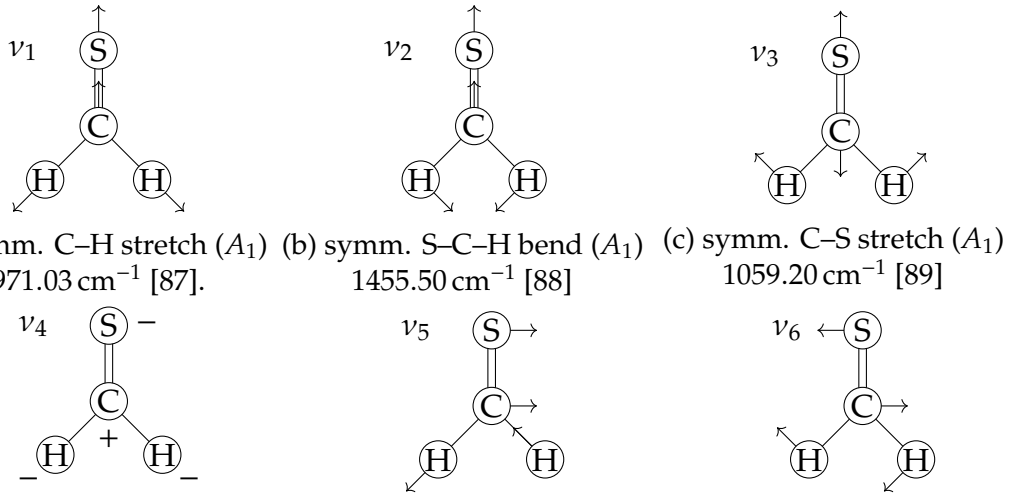
| $\nu$ | unc. ($cm^{-1}$) | unc. ($cm^{-1}$) | Quantum 'numbers' of upper states | | | | | | | | | | Quantum 'numbers' of lower states | | | | | | | | Source |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | $J'$ | $K_a'$ | $K_c'$ | $n_1'$ | $n_2'$ | $n_3'$ | $n_4'$ | $n_5'$ | $n_6'$ | $J''$ | $K_a''$ | $K_c''$ | $n_1''$ | $n_2''$ | $n_3''$ | $n_4''$ | $n_5''$ | $n_6''$ | |
| 8.01661895 | 0.00000167 | 0.00000167 | 7 | 4 | 4 | 0 | 0 | 0 | 0 | 0 | 0 | 6 | 4 | 3 | 0 | 0 | 0 | 0 | 0 | 0 | 08MaMeWiDe.35 |

### 4.2.2 $H_2CS$ quantum numbers

$H_2CS$ is a semi-rigid, asymmetric top molecule that can be classified according to the $\boldsymbol{C}_{2v}$(M) molecular symmetry group with the irreducible representations $A_1$, $A_2$, $B_1$, and $B_2$ [31]. Transitions follow the parity selection rules $A_1 \leftrightarrow A_2$ and $B_1 \leftrightarrow B_2$; and the standard rotational selection rules, $J' - J'' = 0, \pm 1,\ J' + J'' \neq 0$; where $'$ and $''$ denote the upper and lower state, respectively.

Nine quantum numbers were required to uniquely label a rotation-vibration energy level of $H_2CS$ in the ground electronic state: the six vibrational normal mode quantum numbers $\nu_1$ to $\nu_6$, detailed in Figure 4.1, and the three standard rigid-rotor quantum numbers $J$, $K_a$, and $K_c$. Here, $J$ is the total angular momentum quantum number, $K_a$ and $K_c$ are the oblate and prolate quantum numbers, respectively. These are restricted by the convention $J = K_a + K_c$ or $J = K_a + K_c + 1$ (the latter only when $K_a \neq 0$ and $K_c \neq 0$). This gives $2J + 1$ states for a given $J$.

The chosen set of quantum numbers has to be consistent across the whole dataset and care must be taken when extracting data from literature sources as different labelling conventions are occasionally encountered.




(a) symm. C–H stretch ($A_1$) 2971.03 $cm^{-1}$ [87]. (b) symm. S–C–H bend ($A_1$) 1455.50 $cm^{-1}$ [88] (c) symm. C–S stretch ($A_1$) 1059.20 $cm^{-1}$ [89]

(d) out-of-plane bend ($B_1$) 990.18 $cm^{-1}$ [89] (e) assym. C–H stretch ($B_2$) 3024.61 $cm^{-1}$ [87] (f) assym. S–C–H bend ($B_2$) 991.02 $cm^{-1}$ [89]

Figure 4.1: Vibrational modes and measured band centers of $H_2CS$.

Table 4.2 shows the relationship between the rotational quantum numbers and overall rotational symmetry. This needs to combined with the symmetry of the vibrational modes given above to give the total symmetry of a given state.

## 4.3 Experimental data sources

Spectroscopic data was extracted from 11 published sources and these are summarised in Table 4.3. This contains the energy and $J$ range of each source, the bands covered, the number of available transitions (A), the number of validated transitions (V) by the MARVEL procedure, and, from those validated transitions, the mean and maximum

Table 4.2: Symmetry species of $J_{K_a,K_c}$ levels of $H_2CS$ in the $C_{2v}$(M) group [31].

| $K_a$ | $K_c$ | $\Gamma_{rot}$ |
|---|---|---|
| even | even | $A_1$ |
| even | odd | $A_2$ |
| odd | even | $B_2$ |
| odd | odd | $B_1$ |

uncertainty obtained from the uncertainties quoted in each study. Four sources provided their data in digital format, while the other literature sources were processed using *ABBYY FineReader*. The number of available transitions for a given source does not include transitions taken from prior sources.

There were two primary reasons for the exclusion of a transition. The first is when its provenance was not clear so the line could not be verified to be purely experimental. The second is when the difference between the energy of a state determined by the MARVEL procedure compared to the energy as predicted by a transition was greater than the uncertainty of the transition. Typically, these states were well supported by several other transitions (from another source) with the outlier transition biasing the MARVEL energy of the state towards the transition predicted energy even if the transition uncertainty was increased. Thus, the removal of the transition improved the final MARVEL energy. A negative sign is added to any transition wavenumber to remove it from the MARVEL procedure.

A number of changes were made to the extracted transition data. The energies of states with $K_a$ close to $J$ may depend only very weakly on $K_c$ and often give rise to two very-closely-spaced transitions, a phenomenon known as $K$-doubling (see Section 4.3.2). In sources where the prolate ($K_c$) quantum numbers were missing, degenerate transitions with both possible values were added into the input transitions file which respected the symmetry rule. The number of available transitions in Table 4.3 reflects this change.

Due to the close band centres of $\nu_4$ and $\nu_6$, many states had to be reassigned to ensure consistency in the input transitions file. The basis and method of reassignment is explained in Section 4.3.2, with the changes made to the individual sources detailed below.

**70JoPo** [90]: A ground vibrational study whose data was included in 19MuMaTh.

**71JoPoKi** [98]: A ground vibrational study whose data was included in 72BeKlKiJo.

**71JoOl** [87]: An infrared study with the $\nu_1$, $\nu_5$, and $2\nu_2$ bands. The original assignment for the latter was $2\nu_6$ but this was incorrect and has since been updated. The transitions from this source were challenging to scan due to the print quality and were not always unambiguous. The transitions were also of relatively low resolution as illustrated in Figure 4.2. This paper had the largest number of non-validated transitions. These were the asterisk-marked transitions in their publication, which, in certain cases, were unobserved but expected transitions, however the precise transitions where this applied was not specified. Thus, all had to be removed from the validated set. These were labelled with a "_a" after the source number but before the label number. In total, 368 transitions were removed for this reason or because the energy they predicted deviated from that predicted by MARVEL beyond their uncertainty.

Table 4.3: Experimental literature sources of $H_2CS$ spectra and their coverage. For the MARVEL analysis, V is the number of verified transitions and A is the number of available transitions. The mean and maximum uncertainties of each dataset reflects only the verified transitions. $\nu_0$ means ground vibrational.

| Tag | Energy range ($cm^{-1}$) | Band | $J$ range | V/A | Unc. Mean/Max ($cm^{-1}$) |
|---|---|---|---|---|---|
| 70JoPo [90] | 0.209 to 0.349 | $\nu_0$ | 3 to 4 | 2/2 | $3.34\times10^{-7}$ / $3.34\times10^{-7}$ |
| 71JoOl [87] | 2843 to 3129 | $\nu_1$, $2\nu_2$, $\nu_5$ | 0 to 30, 1 to 30, 2 to 24 | 1110/1478 | $7.559\times10^{-3}$ / $2.000\times10^{-2}$ |
| 72BeKlKiJo [91] | 0.105 to 8.141 | $\nu_0$ | 0 to 27 | 56/56 | $1.167\times10^{-4}$ / $2.715\times10^{-4}$ |
| 77FaKrMu [92] | 0.0349 | $\nu_0$ | 0 to 1 | 1/1 | $2.23\times10^{-9}$ / $2.23\times10^{-9}$ |
| 80BeDu [93] | 924.9 to 1059 | $\nu_3$, $\nu_4$, $\nu_6$ | 1 to 6, 3 to 9, 3 to 9 | 63/82 | $2.000\times10^{-3}$ / $2.000\times10^{-3}$ |
| 81TuHaMi [94] | 818.8 to 1221 | $\nu_4$, $\nu_6$ | 0 to 22, 2 to 23 | 102/386 | $3.839\times10^{-3}$ / $2.800\times10^{-2}$ |
| 93McBr [88] | 31.46 to 1499 | $\nu_0$, $\nu_2$ | 21 to 54, 1 to 37 | 922/935 | $3.256\times10^{-2}$ / $5.000\times10^{-2}$ |
| 94ClHuAdMe [95] | 2.222 to 402.1 | $\nu_0$ | 0 to 36 | 302/355 | $4.000\times10^{-4}$ / $4.000\times10^{-4}$ |
| 08MaMeWiDe [96] | 4.161 to 12.40 | $\nu_0$ | 4 to 41 | 113/113 | $1.980\times10^{-6}$ / $3.340\times10^{-6}$ |
| 19MuMaTh [97] | 4.161 to 46.23 | $\nu_0$ | 4 to 42 | 317/317 | $1.020\times10^{-6}$ / $3.336\times10^{-6}$ |
| 08FlLaPeKi [89] | 690.7 to 1338 | $\nu_3$, $\nu_4$, $\nu_6$ | 0 to 50, 0 to 50, 1 to 51 | 8649/8655 | $4.748\times10^{-4}$ / $1.000\times10^{-3}$ |

**72BeKlKiJo** [91]: A ground vibrational study up to $J = 27$. No difficulties in validating the transitions.

**77FaKrMu** [92]: One measured line from this study was included in the 19MuMaTh data set.

**80BeDu** [93]: Infrared study involving the $\nu_3$, $\nu_4$, and $\nu_6$ bands. 19 transitions were removed as the energy they predicted deviated from that predicted by MARVEL beyond their uncertainty.

**08FlLaPeKi** [89]: An infrared study of the $\nu_3$, $\nu_4$, and $\nu_6$ bands. The authors of this study were contacted and provided a more complete set of data which was used in the MARVEL analysis. Several quantum numbers were adjusted. 18 lines had the $K'_a$ quantum number increased by 2 with the corresponding $K'_c$ quantum number decreased by 2 to maintain the symmetry and $K'_a + K'_c$ requirement. 2919 lines had their $K'_c$ value adjusted by 1 or $-1$ so that the symmetry selection rule was satisfied. Finally, 1625 lines had their $\nu_4$ and $\nu_6$ assignment swapped (discussed in Sec. 4.3.2). Six transitions were removed as their optimal uncertainty was too high.

**81TuHaMi** [94]: Infrared study involving the $\nu_3$, $\nu_4$, and $\nu_6$ bands. Some listed transitions were taken from 80BeDu so we ignored these and used the original data source. These were labelled with a "_o" after the source name but before the label number. Eight transitions had their $\nu_4$ and $\nu_6$ assignment swapped. Ten transitions had their $K'_c$ values changed by 1, three had their $K'_c$ values changed by 2, and one had its $K'_c$ value changed by 3. One transition had its $K'_a$ value changed by 1 and four had their $K'_a$ value changed by 2. Eight transitions had their $K''_c$ values changed by 1. One transition had its $K''_a$ value changed by 1 and one had its $K''_a$ value changed by 2. 284 transitions were removed as the energy they predicted deviated from that predicted by MARVEL beyond their uncertainty.

**93McBr** [88]: The authors of this study were contacted and they provided us with the transitions from the $\nu_2$ band. It is the only source containing the $\nu_2$ band; the transitions are relatively low resolution, as illustrated in Figure 4.2 where the highest circle for 93McBr corresponds to the $\nu_2$ transitions. No subsequent high resolution studies of this band have been undertaken. Ground vibrational transitions from this source are

contained in 19MuMaTh and 08MaMeWiDe.

**94ClHuAdMe** [95]: An electronic study of the $\tilde{A}$–$\tilde{X}$ transition of thioformaldehyde. The ground state combination differences reported in 19MuMaTh and taken from this study were used in the MARVEL dataset. 53 transitions were removed as the energy they predicted deviated from that predicted by MARVEL beyond their uncertainty.

**08MaMeWiDe** [96]: A ground vibrational study which included transition data from multiple sources including 94ClHuAdMe, 71JoPoKi, 80BeDu, 72BeKlKiJo, 93McBr as well as CDMS (Cologne Database for Molecular Spectroscopy) [99] data and the FASSST [100] data. Transitions in the MARVEL input file from 93McBr numbered 113 and above are from this source.

**19MuMaTh** [97]: A ground vibrational study which incorporated data from multiple sources including 77FaKrMu, 70PoJo, 71JoPoKi, 72BeKlKiJo, 94ClHuAdMe, and 93McBr. It also contained Koln and FASSST data. Transitions from 93McBr numbered 1 to 112 are from this source.

There is an overlap in data coverage from the sources 08FlLaPeKi, 80BeDu, and 81TuHaMi as, collectively, they contained transitions involving the $\nu_3$, $\nu_4$, and $\nu_6$ bands. Figure 4.2 illustrates the uncertainties of the data by source. The uncertainties of 08BeDu and 81TuHaMi were consistently higher than the ones from 08FlLaPeKi. 08FlLaPeKi's dataset was also significantly larger and, consequently, each state was well represented by several transitions. Due to the lower uncertainties, transitions to a given upper state were much more internally consistent. Because of these factors, 08FlLaPeKi was favoured over the other two sources. In total, 19 transitions from 80BeDu and 284 transitions from 81TuHaMi were removed because they were not in agreement with the other data.

### 4.3.1 Artificial transitions

The complete internal states of $H_2CS$ are either ortho (nuclear spin functions of symmetry $A_1$) or para (nuclear spin functions of symmetry $B_1$). There are no transitions between those states. As a result, the MARVEL analysis produces two separate networks of energy levels, an ortho-network and a para-network, which are disconnected. To ensure the energies of the all para-states relative to the ro-vibrational ground state of the molecule are correct, it is necessary to link the ortho-network and para-network through an artificial transition. This was done using the pseudo-experimental transition wavenumbers reconstructed with the PGOPHER program [101] where the spectroscopic constants were taken from 08FlLaPeKi. The ground vibrational transition $1_{11} \leftarrow 0_{00}$ (states labelled as $J_{K_aK_c}$) with a value of $10.281\,826(100)\,\mathrm{cm}^{-1}$ was added to the input MARVEL transition file to link the networks.

### 4.3.2 Reassignment of transition data

The out-of-plane bending mode $\nu_4$ and the asymmetric S–C–H bending mode $\nu_6$ have very close band centres and interact through a massive A-type Coriolis resonance [89]. To ensure consistency across all literature sources, a number of reassignments involving these levels had to be performed using the assignment from the TROVE calculations. As this assignment was based on the valence coordinates, it did not produce normal mode

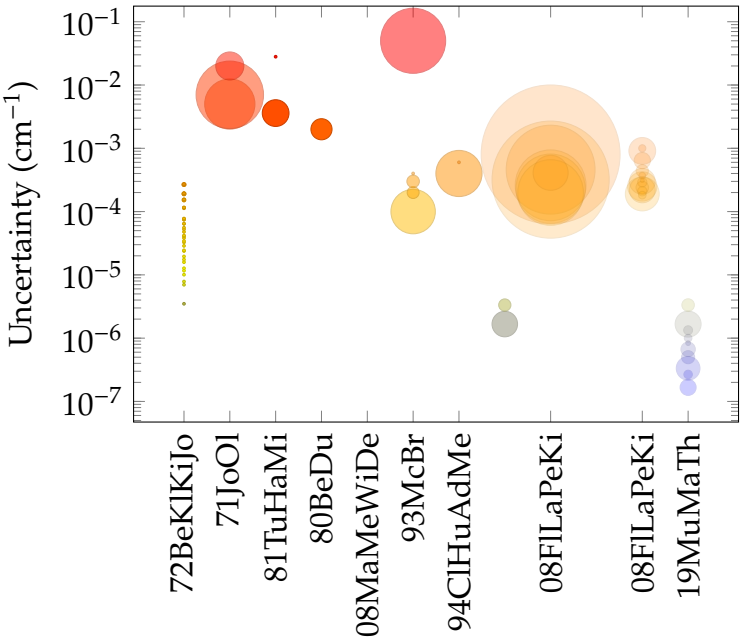



Figure 4.2: Uncertainties in the measured data (vertical axis, log scale) from each literature source (horizontal axis). Each circle corresponds to a given uncertainty. The vertical coordinate of the centre of the circle is the uncertainty, and the area of the circle is proportional to the number of transitions with that uncertainty. The smallest and largest circles correspond to 1 and 2729 transitions, respectively.

quantum numbers. Thus, the TROVE assignment had to be translated to the normal mode assignment.

The valence coordinates "classes" TROVE used (labelled $q_1$ to $q_6$) are the C–S stretch, the two C–H stretches, the two S–C–H bends, and the dihedral angle between the two S–C–H planes, with the BF frame of Section 3.6. The ro-vibrational states have the form

$$\psi = \sum_p a_p \, |J, K, \sigma\rangle_p \, \phi^p_{v_1}(q_1) \ldots \phi^p_{v_6}(q_6). \tag{4.3.1}$$

Quantum numbers are assigned based on the largest coefficient $a_p$ and are given by $(J, K, \sigma, v_1, \ldots, v_6)$ as well as rotational, vibrational, and total symmetries of the state. The resulting assignments are internally consistent and can be unambiguously mapped to normal mode assignments. Because the available TROVE ro-vibrational states greatly outnumber those made available experimentally, reassigning the TROVE states purely on the basis of the experimental assignments would not be possible. Further assignments would be necessary to ensure internal consistency without the guidance of the experimental assignments. Thus, for practical reasons in the refinement and MARVELisation of the computed line list, it was more straightforward to reassign the experimental data on the more complete TROVE data.

For the rotational quantum numbers, we must understand the meaning of $K_a$. The rigid-rotor Hamiltonian is given by

$$\hat{H}_{\text{rot}} = \hbar^{-2}(A_e \hat{J}_a^2 + B_e \hat{J}_b^2 + C_e \hat{J}_c^2) \tag{4.3.2}$$

Table 4.4: Symmetry species of the $|J, K, \sigma\rangle$ functions of $H_2CS$ in the $\boldsymbol{C}_{2v}$(M) group [31]. $\sigma$ can take the values 0 and 1.

| $K$ | $\sigma$ | $\Gamma_{\rm rot}$ |
|---|---|---|
| even | 0 | $A_1$ |
| even | 1 | $A_2$ |
| odd | 1 | $B_2$ |
| odd | 0 | $B_1$ |

where the rotational constants are defined by

$$D_{\rm e} = \frac{\hbar^2 \mu^{\rm e}_{dd}}{2hc}, \quad D_{\rm e}, d = \quad A_{\rm e}, a, \quad B_{\rm e}, b \quad C_{\rm e}, c \tag{4.3.3}$$

and $\mu^{\rm e}_{dd} = (I^{\rm e})^{-1}_{dd}$ where $I^{\rm e}$ equilibrium moment of inertia matrix. $H_2CS$ is a near-prolate top as $A_{\rm e} > B_{\rm e} \sim C_{\rm e}$ ($A_{\rm e} = 9.727\,{\rm cm}^{-1}$, $B_{\rm e} = 0.5904\,{\rm cm}^{-1}$, $C_{\rm e} = 0.5555\,{\rm cm}^{-1}$). If one sets the $z$-axis to be along the $a$-axis, then $\hat{H}_{\rm rot}$ can be expressed as

$$\hat{H}_{\rm rot} = \hbar^{-2}[(B_{\rm e} + C_{\rm e})/2\hat{J}^2 + (A_{\rm e} - (B_{\rm e} + C_{\rm e})/2)J_z^2 + (B_{\rm e} - C_{\rm e})/4((J^+)^2 + (J^-)^2)] \tag{4.3.4}$$

where $J^+$ and $J^-$ are raising and lowering operators, respectively. The close value of $B_{\rm e}$ and $C_{\rm e}$ explains the near-degeneracy of states with the same value of $K_a$, the projection of the angular momentum onto the $z$- or $a$-axis, as then $B_{\rm e} - C_{\rm e}$ is close to zero.

The $a$-axis in this case is along the C−S bond and this was the direction of the $z$-axis in the Trove calculations. It is interesting to note that the equilibrium rotational kinetic energy operator in the TROVE representation is given by Eq. (4.3.2) but with constants $\hbar^{-2}A_{\rm e} = 19.0256\,{\rm cm}^{-1}$ and $\hbar^{-2}B_{\rm e} = \hbar^{-2}C_{\rm e} = 2.6684\,{\rm cm}^{-1}$. The difference is presumably due to the different choice of BF-frame and vibrational coordinates.

The above discussion implies $K_a$ should be set to $K$. $K_c$ is chosen so that the rotational symmetry matches the $|J, K, \sigma\rangle$ function (given by Table 4.4) according to Table 4.2.

Mapping the TROVE vibrational quantum numbers to the normal mode ones is accomplished by first determining which normal modes the TROVE coordinates correspond to. For example, varying the C–H stretches (excitations $v_2$ and $v_3$) must be a combination of the symmetric or asymmetric C–H stretches (excitations $n_1$ and $n_5$). Moreover, the total degree of excitation must match. That is, we have

$$n_1 + n_5 = v_2 + v_4. \tag{4.3.5}$$

The precise values of the quantum numbers are found by matching the symmetries of the states. The full mapping between the two for the states of interest is given in Table 4.5.

For the MARVEL input file, the reassignment process involved swapping the upper state $v_4'$ and $v_6'$ quantum numbers in transitions from the literature sources 08FlLaPeKi and 81TuHaMi (see Table 4.3). Also, the upper state $K_c'$ value was changed by 1 or −1 to maintain the symmetry selection rule of the transition while also ensuring that the convention $J' = K_a' + K_c'$ or $J' = K_a' + K_c' + 1$ was satisfied. To generate the lower

Table 4.5: The mapping between the TROVE quantum numbers and the standard normal mode notation with the associated irrep of the states. In TROVE, the assignment within a "class" is somewhat unpredictable as only the sum is of significance as illustrated in the table.

| Normal mode | TROVE quantum numbers | | | | Symmetry |
|---|---|---|---|---|---|
| | $v_1$ | $v_2 + v_3$ | $v_4 + v_5$ | $v_6$ | |
| $\nu_1$ | 0 | 1 | 0 | 0 | $A_1$ |
| $\nu_2$ | 0 | 0 | 1 | 0 | $A_1$ |
| $2\nu_2$ | 0 | 0 | 2 | 0 | $A_1$ |
| $\nu_3$ | 1 | 0 | 0 | 0 | $A_1$ |
| $\nu_4$ | 0 | 0 | 0 | 1 | $B_1$ |
| $\nu_5$ | 0 | 1 | 0 | 0 | $B_2$ |
| $\nu_6$ | 0 | 0 | 1 | 0 | $B_2$ |

state energies, the computer program PGOPHER [101] was used with the spectroscopic constants given in 08FlLaPeKi. Since the lower states were purely rotational, there was no ambiguity in their assignment.

## 4.4 MARVEL analysis

The MARVEL analysis was performed using the compiled MARVEL version 3 with the Cholesky (analytic) method. An extract from the output MARVEL energy level file is shown in Table 4.6. In total, 11 638 non-unique transitions out of 12 380 transitions up to $J = 54$ were processed, resulting in 4254 ro-vibrational energy levels up to 3729.07 $cm^{-1}$. Using the new ExoMol $H_2CS$ line list (Section 4.5) shows the number of transitions between MARVEL states stronger than $1 \times 10^{-30}$ cm/molecule (the HITRAN threshold) is 74 425.

Table 4.6: Extract from the MARVEL energy file. The quantum numbers/labels are described in the text and are followed by the MARVEL energy term value ($cm^{-1}$), the uncertainty of the state ($cm^{-1}$), and the number of transitions supporting the state in question.

| Quantum 'numbers' | | | | | | | | | $E$ | Unc. | no. of |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $J$ | $K_a$ | $K_c$ | $n_1$ | $n_2$ | $n_3$ | $n_4$ | $n_5$ | $n_6$ | $cm^{-1}$ | $cm^{-1}$ | trans. |
| 5 | 2 | 3 | 0 | 0 | 0 | 0 | 0 | 0 | 53.792652349 | 0.000316979 | 27 |
| 9 | 1 | 9 | 0 | 0 | 0 | 0 | 0 | 0 | 59.916981444 | 0.000135763 | 28 |
| 6 | 2 | 5 | 0 | 0 | 0 | 0 | 0 | 0 | 60.662392583 | 0.000316979 | 28 |
| 6 | 2 | 4 | 0 | 0 | 0 | 0 | 0 | 0 | 60.669363102 | 0.000316975 | 25 |
| 9 | 1 | 8 | 0 | 0 | 0 | 0 | 0 | 0 | 61.486995429 | 0.000135763 | 28 |

In Figure 4.3, the energy level coverage of the MARVEL ro-vibrational states is plotted as a function of the angular momentum quantum number, $J$, for the different vibrational

bands. Each vibrational band is illustrated by colour but within each band the different $K_a$ and $K_c$ states have the same colour. As can be seen, the energy separation between the $\nu_4$ and $\nu_6$ states is very small. TROVE consistently assigned states of higher energy as $\nu_6$ states and this provided further justification for basing the experimental MARVEL assignment on the TROVE assignment.

Table 4.7 shows the lowest energy state of each fundamental band as determined by MARVEL, as well as the predicted band centre, which was obtained by subtracting the energy of the pure rotational state from the ro-vibrational state. These values can be compared to the experimentally measured band centres quoted in Figure 4.1, and we see that for the $\nu_3$, $\nu_4$, and $\nu_6$ modes the MARVEL band centres are the same to within two decimal places. For the other fundamentals, the MARVEL band centres deviate from the measured band centres at two decimal places. We can see what the deviation is by comparing the minimum (or maximum) possible value of the experimental band centres (those quoted in Figure 4.1) and the predicted MARVEL band centres. For example, the minimum possible value for $\nu_1$ is $2971.025\,\mathrm{cm}^{-1}$ and the MARVEL band centre is $2971.018\,16\,\mathrm{cm}^{-1}$. Thus, the deviation is $0.006\,84\,\mathrm{cm}^{-1}$. For the bands $\nu_2$ and $\nu_5$, the MARVEL energies are different by $0.0273\,\mathrm{cm}^{-1}$ and $0.0006\,\mathrm{cm}^{-1}$, respectively. This is not wholly unexpected, however, as the associated experiments were of lower accuracy. Consequently, the MARVEL uncertainties on the lowest energy states for these bands are $0.007\,\mathrm{cm}^{-1}$, $0.05\,\mathrm{cm}^{-1}$, and $0.007\,\mathrm{cm}^{-1}$, respectively.

Table 4.4 shows the difference between the available CDMS [102] and MARVEL $\nu_0$ energies. The CDMS data was generated by taking the union of the lower state energies and the lower state energy plus the transition wavenumber. The sources were essentially the same as the ground vibrational sources used for the MARVEL analysis. The majority are clustered around 0 and those have a maximum difference of $\sim 1 \times 10^{-3}\,\mathrm{cm}^{-1}$. The average uncertainty of the MARVEL ground vibrational energies is $7.42 \times 10^{-4}\,\mathrm{cm}^{-1}$. Thus, the agreement is good, and the lack of any significant outlier implies that no state has an incorrect energy or assignment.

Figure 4.5a shows examples of the original 08FlLaPeKi assignment for the $\nu_4$ and $\nu_6$ for both $K_a = 1$ and $K_a = 2$ where the plots show the 08FlLaPeKi energy minus $BJ(J+1)$, with $B$ being the rotational constant – a measure of the approximate $J$ energy dependence. This reduces the energy range and thus increases the readability of the plot. There is an apparent misassignment of the $\nu_4$ and $K_a = 2$ energies in the region $20 < J < 30$. The TROVE assignment, shown in Figure 4.5b, partially rectifies this issue.

However, there are still graphically-incorrect assignments for $\nu_4$, $K_a = 2$ and $\nu_6$, $K_a = 2$. Swapping the assignment for these states for $J > 37$ resulted in the assignment shown in Figure 4.6. For the $\nu_6$ and $K_a = 2$ state, one notes that the peculiar behaviour of the energy difference of the two $K_a$ states. This is due to a resonance interaction with the nearby $\nu_3$, $K_a = 1$ states, also shown in Figure 4.6. The lower lying $\nu_3$, $K_a = 1$ states have symmetries which alternate between $B_1$ and $B_2$ in step with the upper lying $\nu_6$, $K_a = 2$ states.

A further seemingly problematic set of assignments are for the $K_a = 0$ and $K_a = 1$ states of the $\nu_4$ and $\nu_6$ bands, the $\nu_4$ band being shown in Figure 4.7. In these cases, rather than the two near-degenerate states being those with the same $K_a$, one is assigned as the $K_a = 0$ state and the other the $K_a = 1$ state. The other $K_a = 1$ state is assigned to the energy further away. This is because of the strong interaction, and hence mixing,

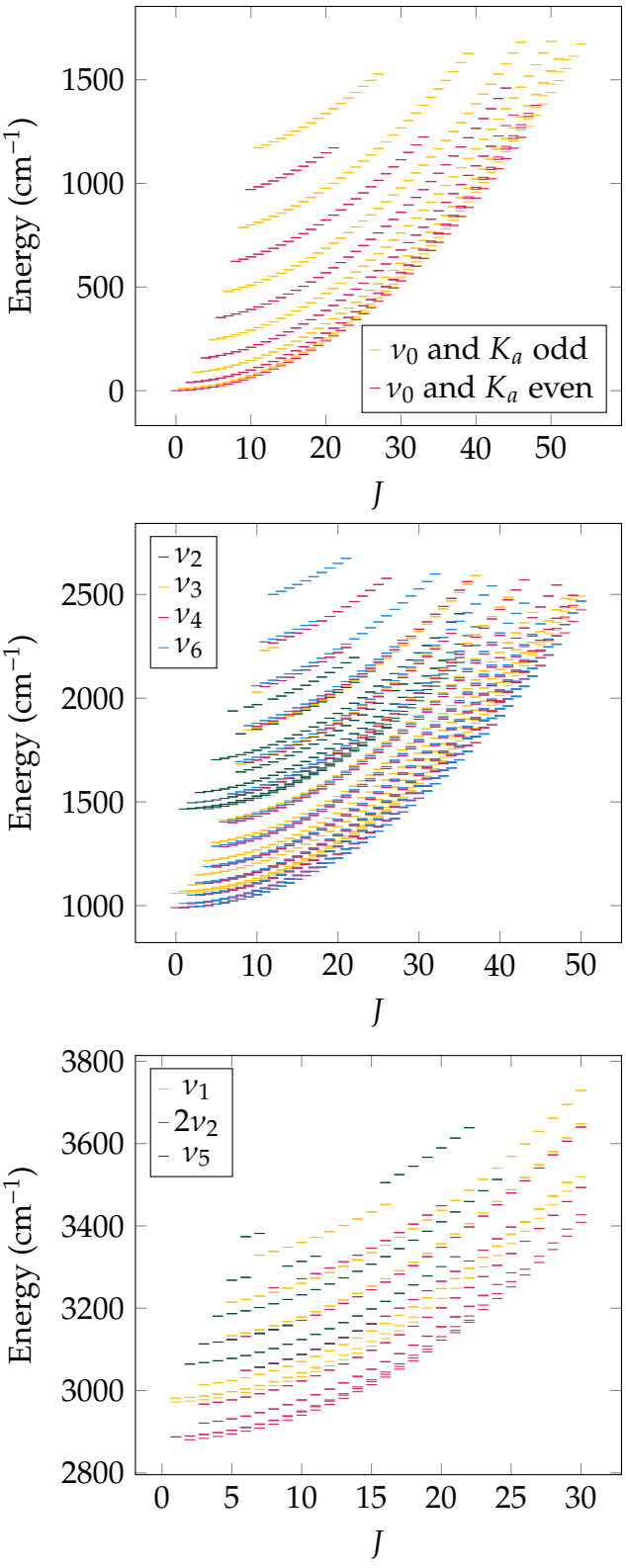



Figure 4.3: The MARVEL energy of each state as a function of the angular momentum quantum number, $J$. Each band has a different colour. Multiple states from the same band with the same $J$ have different $K_a$ and $K_c$ numbers (not labelled). The $\nu_0$ states are separated by ortho ($K_a$ odd) or para ($K_a$ even).

Table 4.7: MARVEL energies of states from each band with the associated quantum numbers and symmetry. The MARVEL uncertainty and the number of transitions which supported a state (No. trans.) are also quoted. Above dividing line: The MARVEL energies from the lowest lying state from each band as well as the band centre which is obtained by subtracting the MARVEL energy of the purely rotational state from the ro-vibrational energy. Below dividing line: The subsequent states from each band.

| $(J_{K_aK_c}n_1n_2n_3n_4n_5n_6)$ | Irrep | MARVEL energy ($cm^{-1}$) | Uncertainty ($cm^{-1}$) | No. trans. | Band centre ($cm^{-1}$) |
|---|---|---|---|---|---|
| $(0_{00}000000)$ | $A_1$ | 0.000 00 | 0.000 00 | 8 | 0.000 00 |
| $(1_{01}100000)$ | $B_1$ | 2972.164 00 | 0.007 00 | 1 | 2971.018 16 |
| $(1_{11}010000)$ | $A_2$ | 1465.859 13 | 0.050 00 | 1 | 1455.577 31 |
| $(2_{02}020000)$ | $A_1$ | 2880.555 12 | 0.007 00 | 2 | 2877.117 72 |
| $(0_{00}001000)$ | $A_1$ | 1059.204 76 | 0.000 46 | 1 | 1059.204 76 |
| $(0_{00}000100)$ | $B_1$ | 990.182 14 | 0.000 47 | 2 | 990.182 14 |
| $(7_{07}000010)$ | $A_2$ | 3056.673 37 | 0.007 00 | 1 | 3024.604 40 |
| $(1_{01}000001)$ | $A_2$ | 992.164 06 | 0.000 81 | 2 | 991.018 22 |
| $(1_{01}000000)$ | $A_2$ | 1.145 84 | 0.000 04 | 10 | |
| $(2_{02}000000)$ | $A_1$ | 3.437 39 | 0.000 09 | 15 | |
| $(3_{03}000000)$ | $A_2$ | 6.874 40 | 0.000 14 | 16 | |
| $(2_{02}100000)$ | $A_1$ | 2974.459 84 | 0.005 00 | 1 | |
| $(3_{03}100000)$ | $B_1$ | 2977.888 23 | 0.005 00 | 2 | |
| $(1_{11}100000)$ | $B_1$ | 2981.234 12 | 0.005 00 | 2 | |
| $(4_{04}010000)$ | $A_2$ | 1466.957 07 | 0.050 00 | 1 | |
| $(2_{12}010000)$ | $B_1$ | 1468.113 37 | 0.050 00 | 2 | |
| $(2_{11}010000)$ | $B_2$ | 1468.227 40 | 0.050 00 | 2 | |
| $(3_{03}020000)$ | $A_2$ | 2884.022 50 | 0.007 00 | 2 | |
| $(1_{11}020000)$ | $B_1$ | 2887.486 73 | 0.007 00 | 1 | |
| $(1_{10}020000)$ | $B_2$ | 2887.524 03 | 0.007 00 | 2 | |
| $(1_{01}001000)$ | $A_2$ | 1060.345 03 | 0.000 91 | 2 | |
| $(2_{02}001000)$ | $A_1$ | 1062.625 43 | 0.000 46 | 2 | |
| $(3_{03}001000)$ | $A_2$ | 1066.045 82 | 0.000 47 | 2 | |
| $(1_{11}000100)$ | $A_1$ | 990.884 14 | 0.002 00 | 1 | |
| $(1_{10}000100)$ | $A_2$ | 990.885 04 | 0.000 80 | 3 | |
| $(1_{01}000100)$ | $B_2$ | 991.325 47 | 0.000 34 | 3 | |
| $(2_{20}000010)$ | $B_2$ | 3064.413 53 | 0.007 01 | 1 | |
| $(2_{21}000010)$ | $B_1$ | 3064.413 53 | 0.007 01 | 1 | |
| $(8_{08}000010)$ | $B_2$ | 3065.845 98 | 0.007 00 | 1 | |
| $(2_{02}000001)$ | $B_2$ | 994.451 57 | 0.000 81 | 3 | |
| $(3_{03}000001)$ | $B_1$ | 997.882 37 | 0.000 81 | 3 | |
| $(4_{04}000001)$ | $B_2$ | 1002.455 90 | 0.000 82 | 3 | |

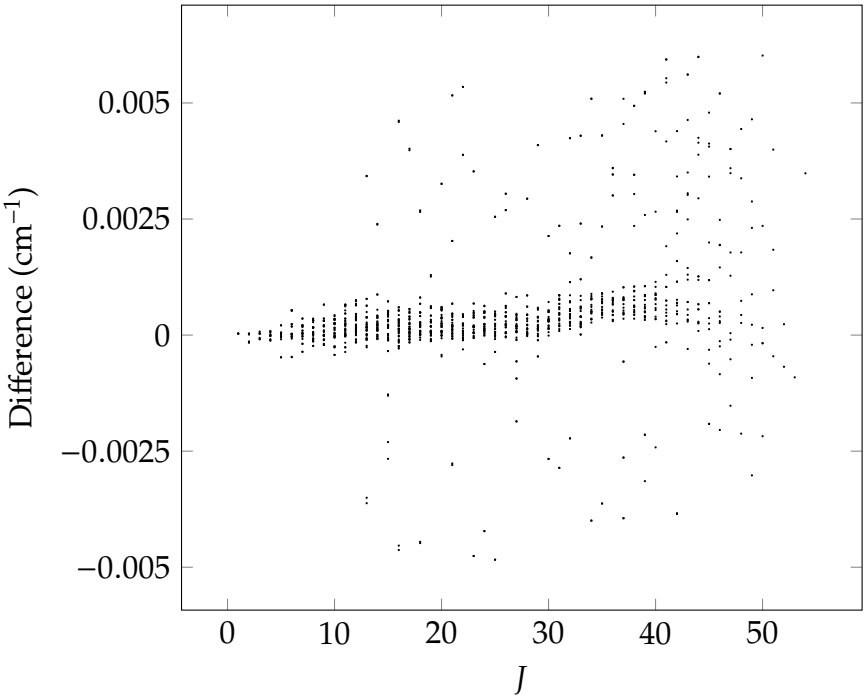
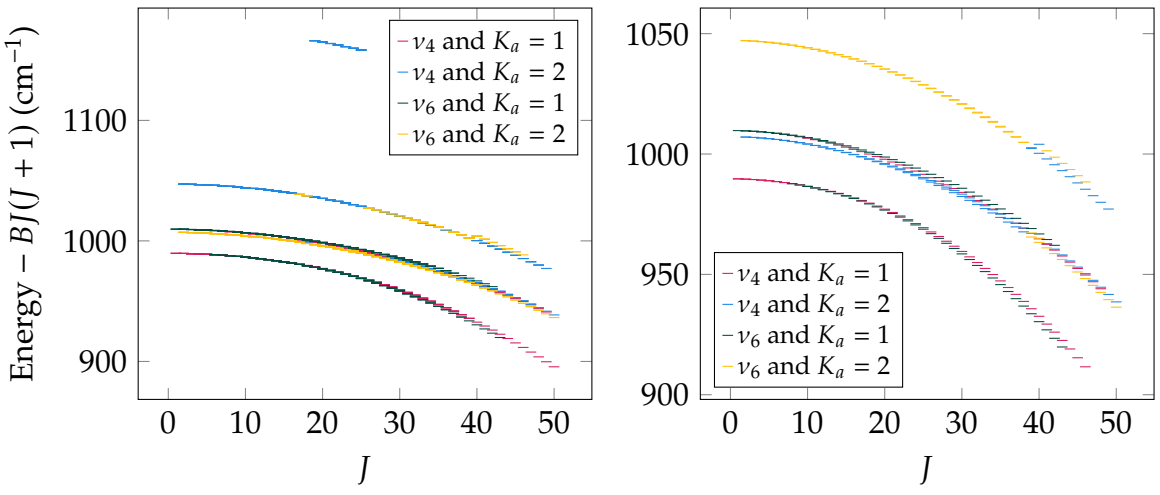



Figure 4.4: Difference between vibrational ground state energies from MARVEL and CDMS as a function of angular momentum quantum number, $J$.




(a) 08FlLaPeKi assignment.

(b) TROVE assignment.

Figure 4.5: A plot of the MARVEL Energy $- BJ(J+1)$ of the $K_a = 1$ and $K_a = 2$ states for the $\nu_4$ and $\nu_6$ bands as a function of the angular momentum quantum number, $J$, where $B$ is the rotational constant.

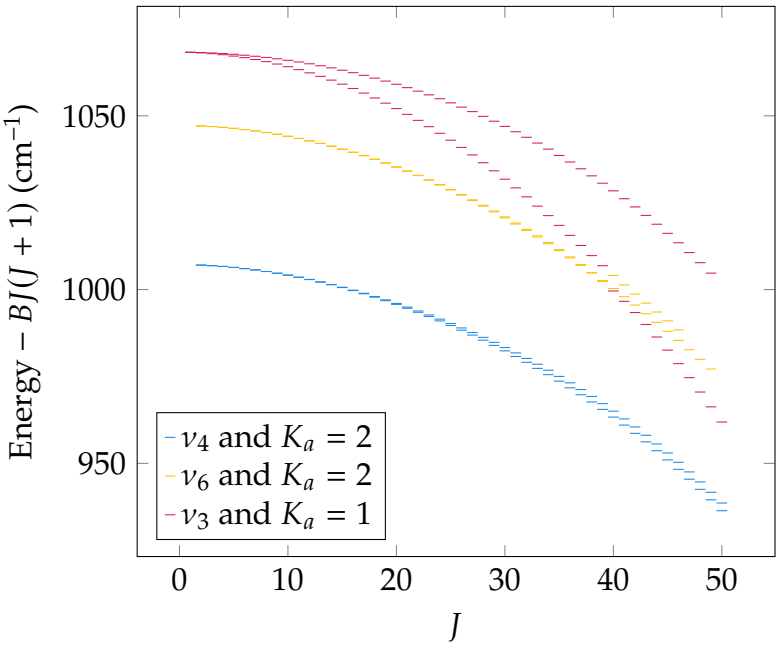



Figure 4.6: A plot of the Energy − $BJ(J+1)$ of the $K_a = 1$ and $K_a = 2$ states for the $\nu_4$ and $\nu_6$ bands as a function of the angular momentum quantum number, $J$, where $B$ is the rotational constant. This is after the $J > 37$ states of $\nu_4$ and $K_a = 2$ and $\nu_6$ and $K_a = 2$ were swapped. The states $\nu_3$ and $K_a = 1$ are also shown.

between these two levels.

It is not possible to manually reassign these states as it would involve changing the $K_a$ value from odd to even (or vice versa) which changes the symmetry from $A$ type to $B$ type (or vice versa) (see Table 4.2). Changing the $K_c$ value by one or reassigning $\nu_4$ states to $\nu_6$ states (or vice versa) does not do this (see Table 4.1), so it is impossible to return to the original irreducible representation with this reassignment.

The uncertainties for each state by band are shown in Figure 4.8. One notes the expected behaviour for the vibrational ground state ($\nu_0$), $\nu_3$, $\nu_4$, and $\nu_6$ bands. The input data for these bands was primarily from 08MaMeWiDe, 08FlLaPeKi, and 19MuMaTh, which were of higher quality than the sources for the other bands. The energies for the other bands are sourced exclusively from 93McBr and 71JoOl. Figure 4.8 shows that their uncertainty is flat rather than growing with $J$. This is a result of both large input uncertainties, as shown on Figure 4.2, with a mean of $7.559 \times 10^{-3}$ cm$^{-1}$ and $2.000 \times 10^{-3}$ cm$^{-1}$ for 71JoOl and 93McBr, respectively, and for uniform uncertainties for all transitions within a band.

Finally, Figure 4.9 shows the energy of each state minus its band centre split by band. The size of each point is proportional to the number of transitions that support that state. The $\nu_0$, $\nu_3$, $\nu_4$, and $\nu_6$ states are measured up to a higher rotational excitation with states up to $J = 51$. The maximum rotational excitation has a range of $J = 24$ to $J = 37$ for the other bands.

The number of transitions supporting a state varies between bands. The strongest is the ground vibrational whose maximum support is 35 transitions. $\nu_4$ and $\nu_6$ are also well supported, though not as well as $\nu_0$, with a maximum of 9. The remaining states

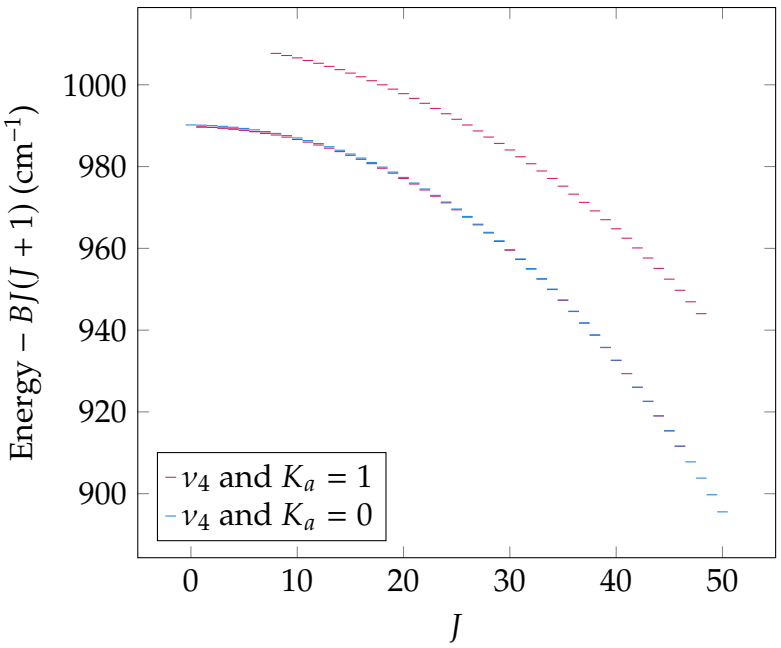



Figure 4.7: A plot of the Energy − $BJ(J+1)$ of the $K_a = 1$ and $K_a = 0$ states for the $\nu_4$ band as a function of the angular momentum quantum number, $J$, where $B$ is the rotational constant. These were sourced from the energy levels file of the MARVEL output after the transitions were reassigned using TROVE.

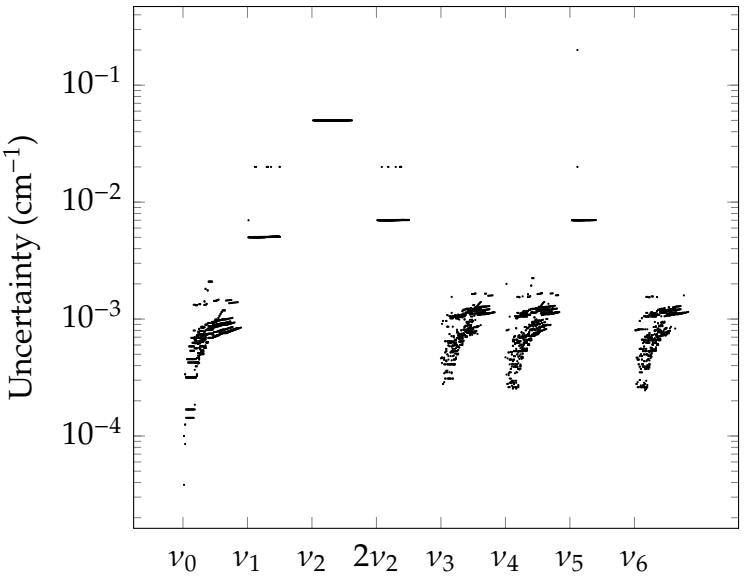



Figure 4.8: Uncertainty of the MARVEL energies levels (vertical axis, log scale) split by band (horizontal axis). The uncertainties are plotted as a function of the angular momentum quantum number, $J$. The left-most uncertainty corresponds to $J = 0$ states while the crossover point to the next band corresponds to $J = 60$. Multiple states from the same band with the same $J$ number have different $K_a$ and $K_c$ numbers.

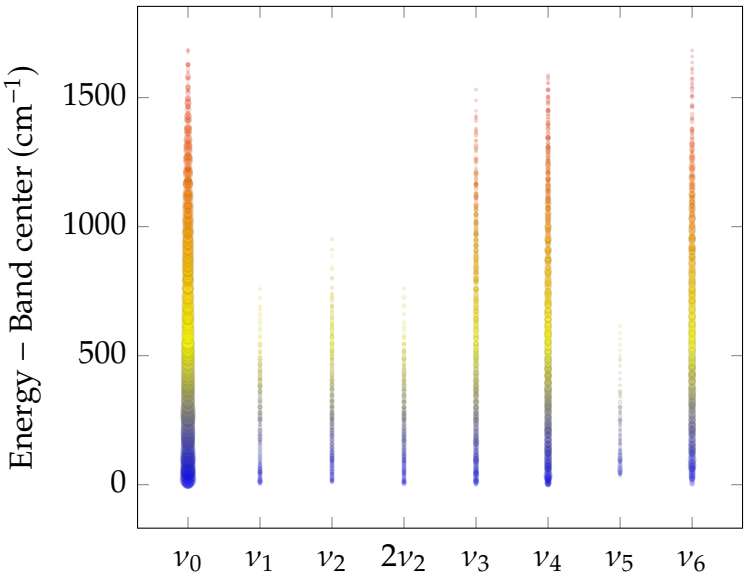



Figure 4.9: The energy of each state minus the band centre of each state (the vertical axis) split by band (the horizontal axis). Each circle corresponds to a given state. The vertical coordinate of the centre of the circle is the energy of the state minus the band centre, and the area of the circle is proportional to the number of transitions that supports that state. The maximum number of supporting transitions is 35.

are much less supported, with several higher rotationally excited states in $\nu_1$, $\nu_2$, $2\nu_2$, $\nu_6$, and, to a lesser extent, $\nu_3$ being only supported by a single transition. $\nu_3$'s maximum is 6 while the rest have a maximum of 3. The MARVEL output energies for states only supported by a single transition are much less trustworthy and depend largely on the accuracy of the stated uncertainty of the experiment; transitions to these states cannot be compared to ensure consistency and accuracy.

## 4.5 Variational nuclear motion calculations

This section summarises the steps in the ro-vibrational calculations of $H_2CS$ in TROVE, with the general methodology already outlined in Section 2.6 and Chapter 3. The Hamiltonian and basis functions used to construct the eigenfunctions were expressed in terms of the previously specified valence coordinates: $\mathcal{R}_{CS}$ (the C–S bond), $\mathcal{R}_{CH_1}$ and $\mathcal{R}_{CH_2}$ (the C–H bonds), $\alpha_{SCH_1}$ and $\alpha_{SCH_2}$ (the H–C–S angles), and $\tau$ (the dihedral angle between the H–C–S planes). The dipole moment surface (DMS) was also expressed in terms of these coordinates, and along with the eigenfunctions, was used to compute ro-vibrational line intensities using the GPU-code MPI-GAIN [103].

For the KEO and DMS, a (right-handed) body-fixed frame was chosen such that the $x$-axis bisected the dihedral angle $\tau$, the $z$-axis pointed along the C–S bond, and the frame's origin was at the centre-of-mass of the nuclei. For each class the finite set of functions in the KEO is as follows. For the bond lengths, they are

$$\frac{1}{\mathcal{R}}, \frac{1}{\mathcal{R}^2}, \frac{1}{\mathcal{R}^3}, \text{ and } \frac{1}{\mathcal{R}^4}. \tag{4.5.1}$$

For the bond angles, they are

$$\cos\alpha,\ \cot\alpha,\ \csc\alpha,\ \sin\alpha,\ \cot^2\alpha,\ \cot\alpha\csc\alpha,\ \text{and}\ \csc^2\alpha. \tag{4.5.2}$$

Finally, for the dihedral angle, they are

$$\cos(\tau/2),\ \sin(\tau/2),\ \cos^2(\tau/2),\ \cos(\tau/2)\sin(\tau/2),\ \text{and}\ \sin^2(\tau/2). \tag{4.5.3}$$

The full KEO is provided in Appendix A of Ref. [2].

The vibrational basis set cut-off applied, based on a polyad number truncation scheme, was

$$P = n_{\mathrm{CS}} + 2(n_{\mathrm{CH_1}} + n_{\mathrm{CH_2}}) + n_{\mathrm{H_1CS}} + n_{\mathrm{H_2CS}} + n_\tau \leq P_{\max}, \tag{4.5.4}$$

which used the polyad cutoff $P_{\max} = 14$, chosen to ensure that full ro-vibrational calculations were computationally tractable when generating a line list up to very high $J$ (discussed in Ref. [104]).

The PES used in this work is based on an *ab initio* surface originally generated by Ref. [105], which utilised state-of-the-art electronic structure methods in the calculation process. The PES was represented as an expansion in terms of six coordinates,

$$\xi_i = 1 - \exp\left[-b_i(\mathcal{R}_i - \mathcal{R}_i^{\mathrm{e}})\right], \quad i = \mathrm{CS}, \mathrm{CH_1}, \mathrm{CH_2}, \tag{4.5.5}$$

$$\xi_j = \alpha_j - \alpha_j^{\mathrm{e}}, \quad j = \mathrm{SCH_1}, \mathrm{SCH_2}, \tag{4.5.6}$$

$$\xi_6 = 1 + \cos\tau. \tag{4.5.7}$$

Here $b_i$ is a Morse-oscillator parameter. The PES was expressed analytically as,

$$V = \sum_{ijklmn} a_{ijklmn} \xi_{\mathrm{CS}}^i \xi_{\mathrm{CH_1}}^j \xi_{\mathrm{CH_2}}^k \xi_{\mathrm{SCH_1}}^l \xi_{\mathrm{SCH_2}}^m \xi_\tau^n, \tag{4.5.8}$$

where $a_{ijklmn}$ are the expansion parameters with maximum expansion order $i + j + k + l + m + n = 6$ with the linear expansion parameters fixed to zero. A total of 413 parameters were utilised including the three equilibrium ($\mathcal{R}_{\mathrm{CS}}^{\mathrm{e}}$, $\mathcal{R}_{\mathrm{CH}}^{\mathrm{e}}$, $\alpha_{\mathrm{SCH}}^{\mathrm{e}}$) and two Morse parameters ($b_{\mathrm{CH}}$, $b_{\mathrm{CS}}$). The expansion parameters assumed the values of the *ab initio* PES from Ref. [105]. Sergey Yurchenko refined the PES by varying the quadratic expansion parameters (14 in total) varied in the refinement to the empirically-derived MARVEL energy levels. The refinement used $H_2CS$ MARVEL energies for $J = 0, 1, 2, 3, 4,$ and 5 (218 in total) covering the ground, $\nu_1$, $\nu_2$, $\nu_3$, $\nu_4$, $\nu_5$, $\nu_6$, and $2\nu_2$ vibrational bands. For the equilibrium structural parameters the optimised geometry values from [104]: $\mathcal{R}_{\mathrm{CS}}^{\mathrm{e}} = 1.608\,952$ Å, $\mathcal{R}_{\mathrm{CH}}^{\mathrm{e}} = 1.086\,848$ Å, and $\alpha_{\mathrm{SCH}}^{\mathrm{e}} = 121.750°$ were used. The PES was constrained to the *ab initio* PES of Ref. [105] in the fitting.

Because an exact KEO was employed in the variational calculations, the refinement showed very quick convergence. It can be expected that the valance coordinate representation improved the convergence of the variationally computed ro-vibrational energies of $H_2CS$, especially at high $J$s, as discussed in Ref. [35]. The quality of the fit is demonstrated in Figure 4.10, where the fitting residuals, i.e. the energy difference (in $\mathrm{cm}^{-1}$) between the MARVEL and computed TROVE $H_2CS$ values, are plotted for the seven vibrational states used in the refinement. The fitting residuals are all below 1 $\mathrm{cm}^{-1}$

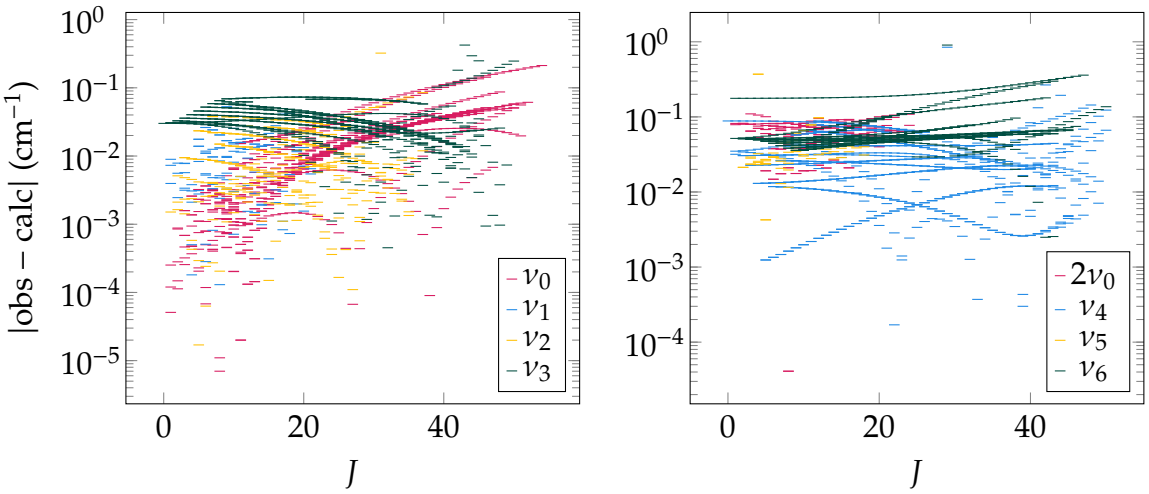



Figure 4.10: The fitting residual errors (obs-calc), i.e. the energy difference (in cm$^{-1}$) between the empirically-derived MARVEL energies and computed TROVE values, as a function of the total angular momentum quantum number $J$. Residual errors are shown for seven vibrational bands and have been computed using the newly refined PES of $H_2CS$.

with the majority of bands possessing errors orders-of-magnitude smaller, notably for the ground state which is to be expected since it corresponds to the lowest part of the PES. The errors are substantially smaller than those of the original *ab initio* $H_2CS$ PES of Ref. [105], which would be in the region of 1 cm$^{-1}$ to 5 cm$^{-1}$, demonstrating the improvements in accuracy that can be achieved with a refined PES.

### 4.5.1 Dipole moment surfaces

A previously published *ab initio* DMS was utilised for intensity calculations. The DMS of Ref. [104] was computed using coupled cluster theory CCSD(T) in conjunction with the augmented correlation-consistent basis set aug-cc-pVQZ(+d for S) [106, 107]. The three dipole components making up the DMS were represented analytically using a sum-of-products expansion in terms of linear expansion variables for the six vibrational coordinates (further details can be found in Ref. [104]).

## 4.6 The MOTY line list of $H_2CS$

The newly computed line list for $H_2CS$, called MOTY, contains 43 561 116 660 transitions between 52 292 454 states and covers the 0 cm$^{-1}$ to 8000 cm$^{-1}$ range for ro-vibrational states with rotational excitation up to $J = 120$. The lower and upper state energy thresholds were chosen to be 8000 cm$^{-1}$ and 18 000 cm$^{-1}$, respectively.

As is standard now for ExoMol line lists, all molecular states possess an uncertainty. The uncertainties were defined either as the MARVEL uncertainty if available, or estimated via the expression:

$$\text{unc} = 0.4n_1 + 0.2(n_2 + n_3 + n_4 + n_5 + n_6) + 0.002J(J+1), \tag{4.6.1}$$

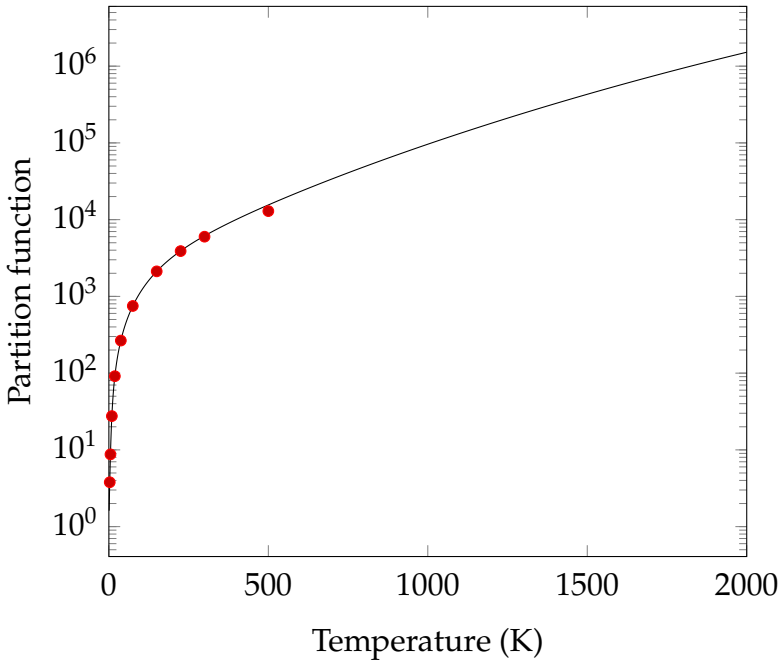



Figure 4.11: Temperature-dependence of the partition function $Q(T)$ of $H_2CS$ computed using the MOTY line list and compared to the CDMS values [102].

where $n_1$ to $n_6$ are the TROVE quantum numbers. These uncertainties are only approximate and grow steadily with increasing rotational and vibrational excitation. States with an estimated uncertainty may well be highly accurate, however, without experimental verification it is difficult to verify.

The temperature-dependent partition function $Q(T)$ of $H_2CS$ was been computed on a 1 K grid in the 1 K to 2000 K range. Partition function values were computed by summing over all the computed ro-vibrational states of the MOTY line list and used nuclear spin statistical weights of $g_{\rm ns} = 1, 1, 3$, and 3 for states of symmetry $A_1$, $A_2$, $B_1$, and $B_2$, respectively. The dependence of $Q(T)$ on temperature is illustrated in Figure 4.11, which shows the partition function values from the CDMS [102] database. The latter only sums over rotational states which explains the lower value at $T = 500$ K.

Following the ExoMolHD strategy [108], in order to improve the accuracy of the MOTY line list and tailor it to high-resolution applications, the calculated energy levels are replaced with the empirically-derived MARVEL values if available.

## 4.7 Spectra simulations

All spectra simulations used the program ExoCross [109]. In Figure 4.12, a general overview of the $H_2CS$ spectrum at three different temperatures (296 K, 1000 K, 2000 K) is plotted. Absolute absorption cross-sections were computed at a resolution of 1 $cm^{-1}$ using a Gaussian line profile with a half width at half maximum (HWHM) of 1 $cm^{-1}$. As expected, the higher temperature spectra exhibit spectral flattening as the previously weaker features gain intensity.

A more detailed band-by-band illustration of the room temperature spectrum of

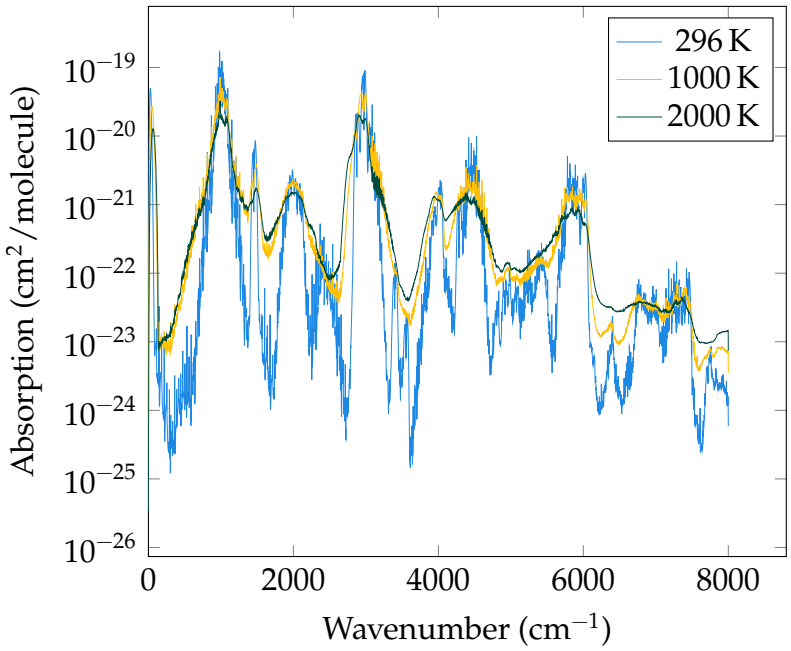



Figure 4.12: Absorption cross-sections of $H_2CS$ (at $T$ = 296 K, 1000 K, and 2000 K). A Gaussian line profile with a half-width-half-maximum (HWHM) of 1 cm$^{-1}$ was used on a grid of resolution 1 cm$^{-1}$.

$H_2CS$ is shown in Figure 4.13. In Figure 4.14 (left panel), a comparison of a synthetic microwave spectrum of $H_2CS$ between the MOTY ExoMol and the CDMS [102, 110] line list is given. The good agreement of the intensities indicate that the equilibrium dipole moments used in CDMS and the MOTY line list agree well. The CDMS line list is based on a dipole moment value of 1.6491 D by [92]. The ExoMol *ab initio* dipole moment gives an expectation value of 1.648 72 D for the ground vibrational state. These values can also be compared to an experimental laser-Stark value of 1.6483 D by [111].

The right-hand side panel of Figure 4.14 shows a comparison with a high-resolution spectrum of $H_2CS$ by Ref. [89] for the $\nu_3$, $\nu_4$, and $\nu_6$ region, although the experimental data only gives relative intensities. The agreement is excellent. Further confirmation of the reliability of the MOTY line list intensities can be seen when comparing to recent theoretical calculations by Ref. [112], which gave an intensity of $1.01\times10^{-20}$ cm/molecule [112] (using coupled cluster theory, CCSD(T)/cc-pVTZ) for the $\nu_4$ transition $3_{03} \leftarrow 3_{13}$, compared to our value of $1.02 \times 10^{-20}$ cm/molecule at $T$ = 296 K.

Figure 4.15 illustrates the impact of the MARVELisation procedure for the MOTY line list of $H_2CS$. It shows a comparison between room-temperature spectra of $H_2CS$ computed using the entire "unMARVELised" MOTY line list and using the MARVELised (upper and lower) states only. The total number of $H_2CS$ transitions at $T$ = 296 K above the HITRAN threshold (between 0 cm$^{-1}$ to 8000 cm$^{-1}$) is 32 094 935. For example, the weak transitions in the region of 2000 cm$^{-1}$ belongs to the hot bands (transitions between excited vibrational states) formed from $\nu_1/\nu_5$ (upper) to $\nu_4/\nu_6$ (lower) states, which has not been experimentally characterised. This figure can provide some indication of the $H_2CS$ spectral regions that are suitable for high-resolution applications.

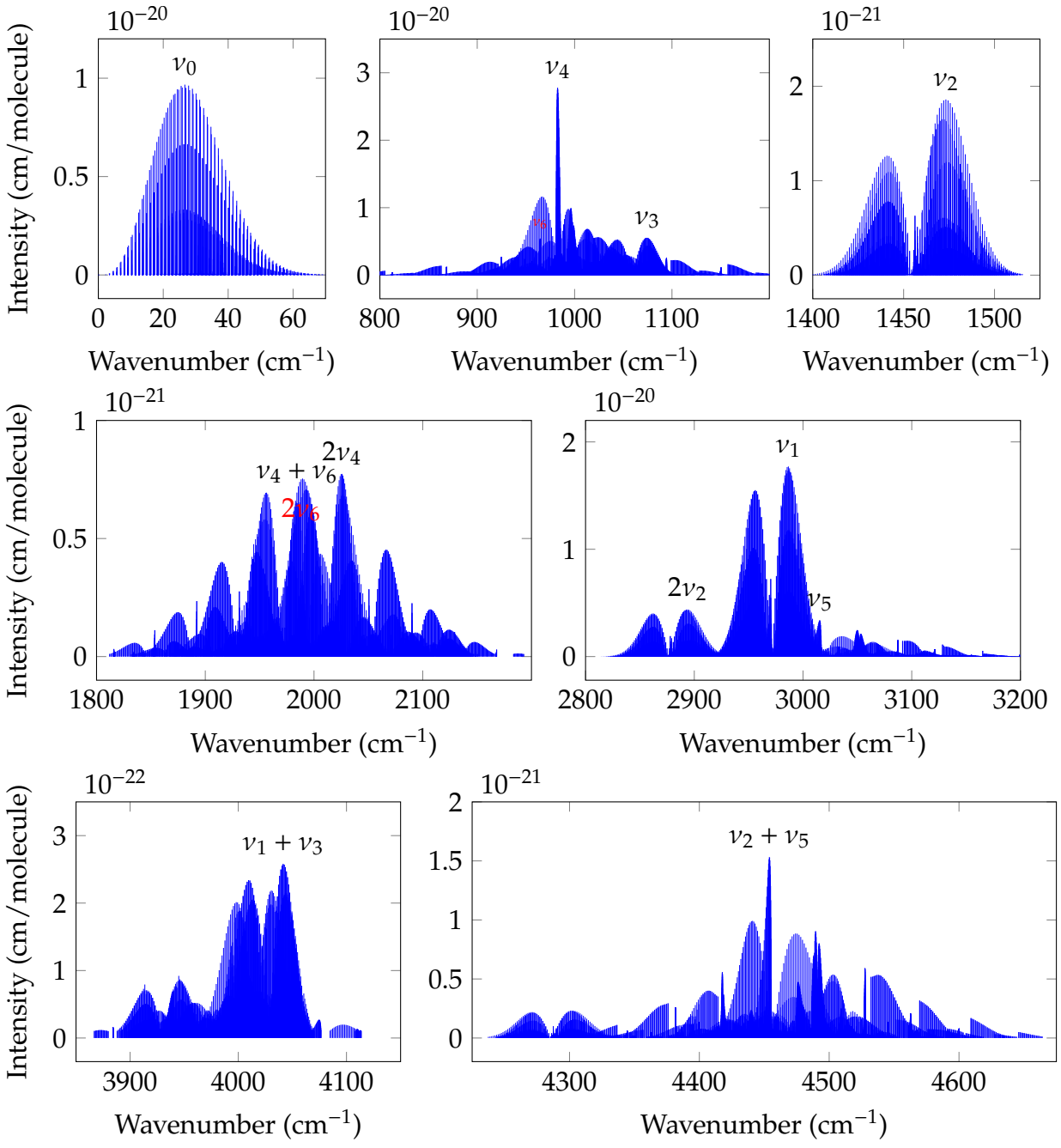



Figure 4.13: Overview of the strongest bands of $H_2CS$ at $T = 296\,K$.

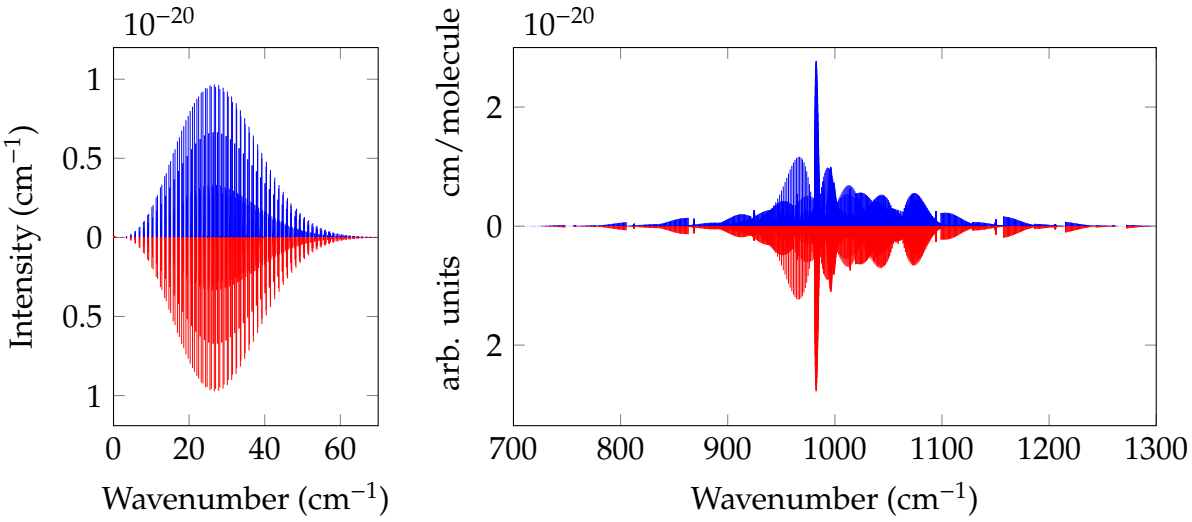



Figure 4.14: Left: Comparison of ExoMol and CDMS [102] microwave absorption cross-sections of $H_2CS$ at $T = 296$ K. Right: Comparison of ExoMol (cm/molecule) and experimental (arb.units) [89] absorption cross-sections of the $\nu_3$, $\nu_4$, and $\nu_6$ region of $H_2CS$ at $T = 296$ K.

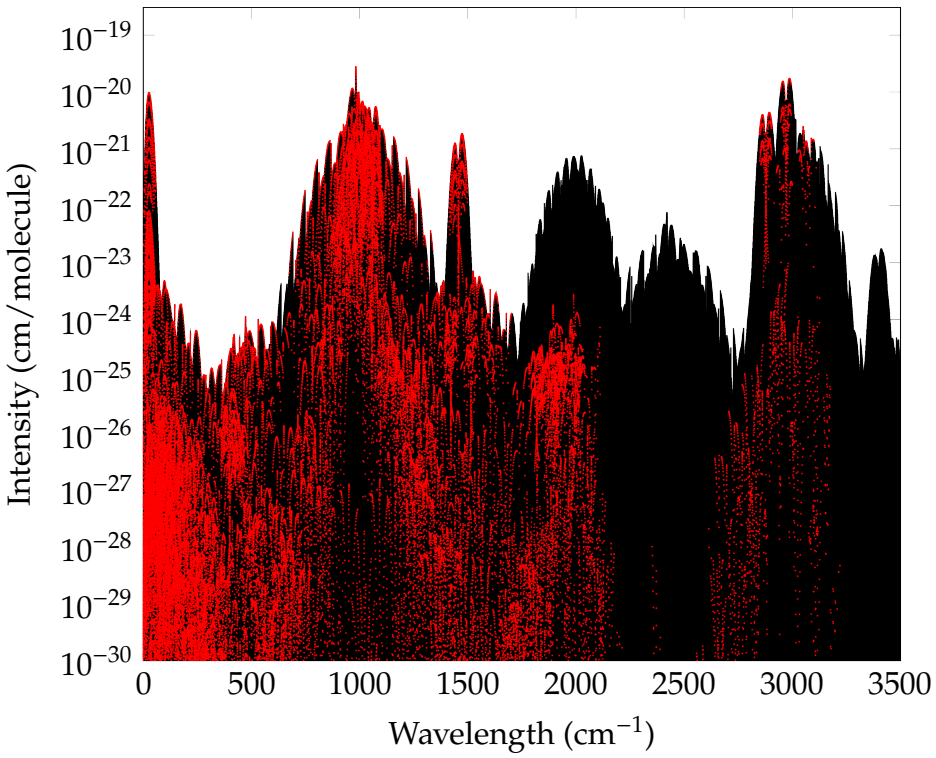



Figure 4.15: Room temperature ($T = 296$ K) spectra of $H_2CS$ showing the coverage of the MARVELised transitions (red points) compared to the total MOTY spectrum.

## 4.8 Chapter summary

This chapter described the comprehensive analysis of the published spectroscopic literature of the main isotopologue of thioformaldehyde ($^1H_2{}^{12}C^{32}S$). All meaningful transitions from 11 literature sources were extracted and verified, resulting in 11 638 validated transitions using the MARVEL algorithm. The data covered the ground, $\nu_1$, $\nu_2$, $\nu_3$, $\nu_4$, $\nu_5$, $\nu_6$, and $2\nu_2$ vibrational bands and produced 4254 ro-vibrational energy levels up to $J = 54$ below 3729 $cm^{-1}$.

Morever, a hot comprehensive ro-vibrational line list of $H_2CS$ was presented. The MOTY line list contains over 43.5 billion transitions and covers the 0 $cm^{-1}$ to 8000 $cm^{-1}$ range (wavelengths $\lambda > 1.3$ µm) for states with rotational excitation up to $J = 120$. The calculations used a new empirically-refined PES and also utilised an exact KEO. The MOTY line list was also MARVELised by replacing the calculated energy levels with more accurate empirically-derived MARVEL values, therefore tailoring the line list to high-resolution applications in certain spectral windows.

# Chapter 5

# Artificial Symmetry for Linear Molecules

## 5.1 Introduction

The description of the ro-vibrational motion of a molecule has so far been limited to the "$3N-6$" vibrational coordinate methodology. For triatomic molecules with a large, non-accessible barrier to linearity, such as $H_2S$ ($24\,423(75)\,cm^{-1}$ [113]) a KEO can be expressed in terms of two bonds, one angle, and three rotational coordinates [31]. For convenience, the $\rho$ angle used measures the deviation from linearity, i.e. $\rho = \pi - \alpha$ where $\alpha$ is the inter-bond angle. The rotational part of an analytic KEO in these coordinates contains terms with a $1/\sin\rho$ factor. If the linear configuration is feasible, such as for $CO_2$ or $H_2O$, then the KEO is singular at linearity.

The singularity is a result of the separation of the coordinates into vibrational and rotational terms. For the latter, the Euler angles depend solely on the positions of the nuclei. At linearity, one can specify the molecular orientation in terms of only two angles. In particular, a rotation about the molecular axis is redundant. Typically, this is the $\chi$ rotation about the BF-frame $z$-axis, and therefore one observes the singular terms in the $G_{iz}$ elements of the $G$ matrix (if non-zero). Moreover, if a non-Euclidean integration volume is used, as is the case for the Sørensen method, then the pseudo-potential may contain singular terms also.

In the literature, there are two main approaches to resolve the singularity issue: the "$3N-5$" or "$3N-6$" methodologies, i.e. the number of vibrational coordinates involved. Both essentially abandon solving the vibrational part of the Hamiltonian separately (and first) from the rotational part of the Hamiltonian. Rather, the part of the rotational KEO which gives rise to the singularity is treated simultaneously with the vibrational part.

The traditional textbook approach is to "transfer" one rotational coordinate to the vibrational coordinates (hence $3N-5$) leaving only two rotational degrees of freedom. This approach is usually associated with the normal modes as vibrational coordinates. The KEO is not singular in such a representation, which is the main point of the $3N-5$ methodology [114]. This approach was previously implemented in TROVE for $C_2H_2$ [115]. The second approach was implemented more recently in TROVE [7] and was utilised in $CO_2$ [6]. This chapter will cover the symmetry aspects of this implementation based on the work published in [3]. In this approach, the singularity is killed with specially designed basis functions.

## 5.2 $3N-5$ and $3N-6$ basis functions

When the $3N-5$ normal mode approach is taken for triatomic molecules, the vibrational basis is constructed as

$$\left|v_1, v_2, v_3^{\ell_3}\right\rangle = \Phi_{v_1}(Q_1)\Phi_{v_2}(Q_2)\,\Phi_{v_3,\ell_3}(Q_{3a}, Q_{3b}). \tag{5.2.1}$$

where $\Phi_{v_i}(Q_i)$ are Harmonic oscillator functions for the $Q_i$ normal modes describing the two stretches and $\Phi_{v_3,\ell_3}(Q_{3a}, Q_{3b})$ is a doubly-degenerate Harmonic oscillator function for modes describing the bend. The bending coordinates are often represented by polar coordinates $(Q_{3a}, Q_{3b}) = (Q_3\cos\chi, Q_3\sin\chi)$ and we can identify $\chi$ as the rotation about the $z$-axis. The vibrational basis function is combined with the rotational function $S_{Jkm}(\theta,\phi)$ where

$$S_{Jkm}(\theta,\phi) = \sqrt{2\pi}e^{-ik\chi}\,|J,k,m\rangle \tag{5.2.2}$$

or, in other words, the $\chi$ independent part of $|J,k,m\rangle$. This recognises that $\Phi_{v_3,\ell_3}$ contains the factor $\exp(i\ell_3\chi)$, already associated with $\exp(ik\chi)$. To transform correctly under the group $\boldsymbol{C}_{2\mathrm{v}}(\mathrm{M})$, only combinations $\ell_3 = k$ are allowed.

In the $3N-6$ framework applied to a linear molecule, a vibrational basis function is given by

$$\left|v_1, v_2, v_3^{\ell}\right\rangle = \varphi_{v_1}(\mathcal{R}_1)\,\varphi_{v_2}(\mathcal{R}_2)\,\varphi_{v_3,\ell}(\rho), \tag{5.2.3}$$

where the stretching vibrations are parametrised by the bond lengths $\mathcal{R}_1$ and $\mathcal{R}_2$. The bending function $\varphi_{v_3,\ell}(\rho)$ in Eq. (5.2.3) plays the central role in the $3N-6$ formalism by being constructed to possess the correct dependence at $\rho\longrightarrow 0$ to resolve the singularity of the KEO [7, 116]:

$$\varphi_{v_3,\ell}(\rho) \sim \sqrt{\rho^{2\ell+1}}. \tag{5.2.4}$$

Examples of functions used are the Laguerre and associated Legendre polynomials.

In this formalism, there is no analogous $\ell = k$ restriction when the vibrational functions are combined with rotational basis functions. However, orthogonality of the complete ro-vibrational basis set is still a desirable property. While $\ell$ can no longer be interpreted as $k$, applying the condition for the $3N-6$ approach is a natural way to maintain orthogonality. The resulting complete basis functions

$$\left|v_1, v_2, v_3^{\ell_3}; J, k=\ell_3, m\right\rangle = \varphi_{v_1}(\mathcal{R}_1)\,\varphi_{v_2}(\mathcal{R}_2)\,\varphi_{v_3,\ell_3}(\rho)\;|J,k=\ell_3,m\rangle \tag{5.2.5}$$

are generally orthogonal, the orthogonality being manifested by the integration over $\chi$ for functions with different values of $\ell_3$ and over $\rho$ for functions with the same value of $\ell_3$ but different values of $v_3$.

The remainder of this chapter will discuss on a practical level how this condition was applied to the existing structure of TROVE. In particular, the symmetrisation scheme is exploited by assigning the vibrational and rotational basis functions to irreps of an "artificial" symmetry group (one not based on the molecular structure) which is designed such that the $\ell_3 = k$ irreps can be straightforwardly selected. The group will first be detailed and an illustrating example will then be provided. This relatively unconventional approach is opposite to the usual practice, where the symmetry group to a large extent is dictated by the physics of the molecule itself, which in turn defines the computational approach.

## 5.3 "Artificial" group $\boldsymbol{D}_{n\mathrm{h}}$(AEM)

The finite artificial extended MS group $\boldsymbol{D}_{n\mathrm{h}}$(AEM) is defined as follows

$$\boldsymbol{D}_{n\mathrm{h}}(\mathrm{AEM}) = \boldsymbol{Z}_2 \underbrace{\otimes \ldots \otimes}_{n-1} \boldsymbol{Z}_2 \tag{5.3.1}$$

where $\boldsymbol{Z}_2$ is the cyclic group of order 2 and therefore consists of the set $\{0, 1\}$ with addition modulo 2. The integer $n$ depends on the value of $\ell_{\max}$. The vibrational $\left|v_1, v_2, v_3^{\ell}\right\rangle$ and rigid-rotor $|J, k, m\rangle$ basis set functions of differing $\ell$ (or $k$) values are assigned to different irreps of this group.

The group $\boldsymbol{D}_{n\mathrm{h}}$(AEM) must fulfil certain conditions necessary for our purposes. First, all irreps should be one-dimensional and real for simplicity. This is to ensure the bending functions all transform independently under $\boldsymbol{D}_{n\mathrm{h}}$(AEM). The irreps of $\boldsymbol{D}_{n\mathrm{h}}$(AEM) will be labelled as $\boldsymbol{C}_{2\mathrm{v}}$(M) irreps with an extra superscript (see Table 5.1), e.g. $A_1^4$. The bending function $\psi_{v_3,\ell}(\rho)$ or rigid rotor function $|J, k, m\rangle$, if they transform as irrep $\Gamma$ in $\boldsymbol{C}_{2\mathrm{v}}$, would be assigned to $\Gamma^{\ell}$ or $\Gamma^{k}$, respectively. For example, a vibrational function with $\ell = 4$ and transforming as $A_1$ in $\boldsymbol{C}_{2\mathrm{v}}$(M) would be assigned the symmetry $A_1^4$ in the $\boldsymbol{D}_{n\mathrm{h}}$(AEM).

Table 5.1: Character table for the MS group $\boldsymbol{C}_{2\mathrm{v}}$(M). The last four columns show the group operations, with two labels for each operation.

| $\Gamma$ | $E$ | (12) | $(12)^*$ | $E^*$ |
|---|---|---|---|---|
| $A_1$ | 1 | 1 | 1 | 1 |
| $B_1$ | 1 | −1 | 1 | −1 |
| $A_2$ | 1 | 1 | −1 | −1 |
| $B_2$ | 1 | −1 | −1 | 1 |

From $\boldsymbol{D}_{n\mathrm{h}}$(AEM), four elements are selected and matched with a $\boldsymbol{C}_{2\mathrm{v}}$(M) element. Then, the characters of those elements for each 0-superscripted irrep should be the same as the corresponding $\boldsymbol{C}_{2\mathrm{v}}$ irrep. When combining a bending function which transforms $\Gamma_1^{\ell}$ with a rigid rotor function which transforms as $\Gamma_2^{k}$, their product should transform as $\Gamma_1^{\ell} \otimes \Gamma_2^{k} = (\Gamma_1 \otimes \Gamma_2)^m$ for some $m \neq 0$ if $\ell \neq k$. If $\ell = k$, then they should transform as $\Gamma_1^{\ell} \otimes \Gamma_2^{\ell} = (\Gamma_1 \otimes \Gamma_2)^0$. For example, $A_2^4 \otimes B_1^4$ should be $B_2^0$.

These rules are illustrated with the group $\boldsymbol{D}_{4\mathrm{h}}$(AEM) in Table 5.2 for the characters and Table 5.3 for the multiplication table of the irreps.

As discussed above, these properties are intended to allow us to assign each vibrational basis function to an irrep which depends on $\ell$. It will also combine $k$ and $\ell$ basis set functions correctly so that the condition of $k = \ell$ can be imposed by giving an irrep with superscript $\neq 0$ a statistical weight factor of 0.

To show that $\boldsymbol{D}_{n\mathrm{h}}$(AEM) has the required properties, and is effectively the only group that does, first note that since the irreps are one dimensional, the group is abelian, and thus, by the fundamental theorem of finite abelian groups [117], can be expressed as a

Table 5.2: Character table for the $\boldsymbol{D}_{4\mathrm{h}}$(AEM) group. Note that the characters of the 0-superscripted irreps for these operations are the same as those of the corresponding irreps for the $\boldsymbol{C}_{2\mathrm{v}}$(M) group.

| $\boldsymbol{D}_{4\mathrm{h}}$(AEM) | $E^0$ | $C_2^0$ | $\sigma^0$ | $\sigma_v^0$ | $E^1$ | $C_2^1$ | $\sigma^1$ | $\sigma_v^1$ | $E^2$ | $C_2^2$ | $\sigma^2$ | $\sigma_v^2$ | $E^3$ | $C_2^3$ | $\sigma^3$ | $\sigma_v^3$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $A_1^0$ | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| $B_1^0$ | 1 | −1 | 1 | −1 | 1 | −1 | 1 | −1 | 1 | −1 | 1 | −1 | 1 | −1 | 1 | −1 |
| $A_2^0$ | 1 | 1 | −1 | −1 | 1 | 1 | −1 | −1 | 1 | 1 | −1 | −1 | 1 | 1 | −1 | −1 |
| $B_2^0$ | 1 | −1 | −1 | 1 | 1 | −1 | −1 | 1 | 1 | −1 | −1 | 1 | 1 | −1 | −1 | 1 |
| $A_1^1$ | 1 | 1 | 1 | 1 | −1 | −1 | −1 | −1 | 1 | 1 | 1 | 1 | −1 | −1 | −1 | −1 |
| $B_1^1$ | 1 | −1 | 1 | −1 | −1 | 1 | −1 | 1 | 1 | −1 | 1 | −1 | −1 | 1 | −1 | 1 |
| $A_2^1$ | 1 | 1 | −1 | −1 | −1 | −1 | 1 | 1 | 1 | 1 | −1 | −1 | −1 | −1 | −1 | 1 |
| $B_2^1$ | 1 | −1 | −1 | 1 | −1 | 1 | 1 | −1 | 1 | −1 | −1 | 1 | −1 | 1 | 1 | −1 |
| $A_1^2$ | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | −1 | −1 | −1 | −1 | −1 | −1 | −1 | −1 |
| $B_1^2$ | 1 | −1 | 1 | −1 | 1 | −1 | 1 | −1 | −1 | 1 | −1 | 1 | −1 | 1 | −1 | 1 |
| $A_2^2$ | 1 | 1 | −1 | −1 | 1 | 1 | −1 | −1 | −1 | −1 | 1 | 1 | −1 | −1 | 1 | 1 |
| $B_2^2$ | 1 | −1 | −1 | 1 | 1 | −1 | −1 | 1 | −1 | 1 | 1 | −1 | 1 | 1 | 1 | −1 |
| $A_1^3$ | 1 | 1 | 1 | 1 | −1 | −1 | −1 | −1 | −1 | −1 | −1 | −1 | 1 | 1 | 1 | 1 |
| $B_1^3$ | 1 | −1 | 1 | −1 | −1 | 1 | −1 | 1 | −1 | 1 | −1 | 1 | 1 | −1 | 1 | −1 |
| $A_2^3$ | 1 | 1 | −1 | −1 | −1 | −1 | 1 | 1 | −1 | −1 | 1 | 1 | 1 | 1 | −1 | −1 |
| $B_2^3$ | 1 | −1 | −1 | 1 | −1 | 1 | 1 | −1 | −1 | 1 | 1 | −1 | 1 | −1 | −1 | 1 |

Table 5.3: Multiplication table for the irreps of the $\boldsymbol{D}_{4\mathrm{h}}$(AEM) group. The vertical and horizontal lines demarcate the blocks of different superscript values. Note that the diagonal blocks are all 0-superscripted, while off diagonal ones are non-0-superscripted.

| $\otimes$ | $A_1^0$ | $B_1^0$ | $A_2^0$ | $B_2^0$ | $A_1^1$ | $B_1^1$ | $A_2^1$ | $B_2^1$ | $A_1^2$ | $B_1^2$ | $A_2^2$ | $B_2^2$ | $A_1^3$ | $B_1^3$ | $A_2^3$ | $B_2^3$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $A_1^0$ | $A_1^0$ | $B_1^0$ | $A_2^0$ | $B_2^0$ | $A_1^1$ | $B_1^1$ | $A_2^1$ | $B_2^1$ | $A_1^2$ | $B_1^2$ | $A_2^2$ | $B_2^2$ | $A_1^3$ | $B_1^3$ | $A_2^3$ | $B_2^3$ |
| $B_1^0$ | $B_1^0$ | $A_1^0$ | $B_2^0$ | $A_2^0$ | $B_1^1$ | $A_1^1$ | $B_2^1$ | $A_2^1$ | $B_1^2$ | $A_1^2$ | $B_2^2$ | $A_2^2$ | $B_1^3$ | $A_1^3$ | $B_2^3$ | $A_2^3$ |
| $A_2^0$ | $A_2^0$ | $B_2^0$ | $A_1^0$ | $B_1^0$ | $A_2^1$ | $B_2^1$ | $A_1^1$ | $B_1^1$ | $A_2^2$ | $B_2^2$ | $A_1^2$ | $B_1^2$ | $A_2^3$ | $B_2^3$ | $A_1^3$ | $B_1^3$ |
| $B_2^0$ | $B_2^0$ | $A_2^0$ | $B_1^0$ | $A_1^0$ | $B_2^1$ | $A_2^1$ | $B_1^1$ | $A_1^1$ | $B_2^2$ | $A_2^2$ | $B_1^2$ | $A_1^2$ | $B_2^3$ | $A_2^3$ | $B_1^3$ | $A_1^3$ |
| $A_1^1$ | $A_1^1$ | $B_1^1$ | $A_2^1$ | $B_2^1$ | $A_1^0$ | $B_1^0$ | $A_2^0$ | $B_2^0$ | $A_1^3$ | $B_1^3$ | $A_2^3$ | $B_2^3$ | $A_1^2$ | $B_1^2$ | $A_2^2$ | $B_2^2$ |
| $B_1^1$ | $B_1^1$ | $A_1^1$ | $B_2^1$ | $A_2^1$ | $B_1^0$ | $A_1^0$ | $B_2^0$ | $A_2^0$ | $B_1^3$ | $A_1^3$ | $B_2^3$ | $A_2^3$ | $B_1^2$ | $A_1^2$ | $B_2^2$ | $A_2^2$ |
| $A_2^1$ | $A_2^1$ | $B_2^1$ | $A_1^1$ | $B_1^1$ | $A_2^0$ | $B_2^0$ | $A_1^0$ | $B_1^0$ | $A_2^3$ | $B_2^3$ | $A_1^3$ | $B_1^3$ | $A_2^2$ | $B_2^2$ | $A_1^2$ | $B_1^2$ |
| $B_2^1$ | $B_2^1$ | $A_2^1$ | $B_1^1$ | $A_1^1$ | $B_2^0$ | $A_2^0$ | $B_1^0$ | $A_1^0$ | $B_2^3$ | $A_2^3$ | $B_1^3$ | $A_1^3$ | $B_2^2$ | $A_2^2$ | $B_1^2$ | $A_1^2$ |
| $A_1^2$ | $A_1^2$ | $B_1^2$ | $A_2^2$ | $B_2^2$ | $A_1^3$ | $B_1^3$ | $A_2^3$ | $B_2^3$ | $A_1^0$ | $B_1^0$ | $A_2^0$ | $B_2^0$ | $A_1^1$ | $B_1^1$ | $A_2^1$ | $B_2^1$ |
| $B_1^2$ | $B_1^2$ | $A_1^2$ | $B_2^2$ | $A_2^2$ | $B_1^3$ | $A_1^3$ | $B_2^3$ | $A_2^3$ | $B_1^0$ | $A_1^0$ | $B_2^0$ | $A_2^0$ | $B_1^1$ | $A_1^1$ | $B_2^1$ | $A_2^1$ |
| $A_2^2$ | $A_2^2$ | $B_2^2$ | $A_1^2$ | $B_1^2$ | $A_2^3$ | $B_2^3$ | $A_1^3$ | $B_1^3$ | $A_2^0$ | $B_2^0$ | $A_1^0$ | $B_1^0$ | $A_2^1$ | $B_2^1$ | $A_1^1$ | $B_1^1$ |
| $B_2^2$ | $B_2^2$ | $A_2^2$ | $B_1^2$ | $A_1^2$ | $B_2^3$ | $A_2^3$ | $B_1^3$ | $A_1^3$ | $B_2^0$ | $A_2^0$ | $B_1^0$ | $A_1^0$ | $B_2^1$ | $A_2^1$ | $B_1^1$ | $A_1^1$ |
| $A_1^3$ | $A_1^3$ | $B_1^3$ | $A_2^3$ | $B_2^3$ | $A_1^2$ | $B_1^2$ | $A_2^2$ | $B_2^2$ | $A_1^1$ | $B_1^1$ | $A_2^1$ | $B_2^1$ | $A_1^0$ | $B_1^0$ | $A_2^0$ | $B_2^0$ |
| $B_1^3$ | $B_1^3$ | $A_1^3$ | $B_2^3$ | $A_2^3$ | $B_1^2$ | $A_1^2$ | $B_2^2$ | $A_2^2$ | $B_1^1$ | $A_1^1$ | $B_2^1$ | $A_2^1$ | $B_1^0$ | $A_1^0$ | $B_2^0$ | $A_2^0$ |
| $A_2^3$ | $A_2^3$ | $B_2^3$ | $A_1^3$ | $B_1^3$ | $A_2^2$ | $B_2^2$ | $A_1^2$ | $B_1^2$ | $A_2^1$ | $B_2^1$ | $A_1^1$ | $B_1^1$ | $A_2^0$ | $B_2^0$ | $A_1^0$ | $B_1^0$ |
| $B_2^3$ | $B_2^3$ | $A_2^3$ | $B_1^3$ | $A_1^3$ | $B_2^2$ | $A_2^2$ | $B_1^2$ | $A_1^2$ | $B_2^1$ | $A_2^1$ | $B_1^1$ | $A_1^1$ | $B_2^0$ | $A_2^0$ | $B_1^0$ | $A_1^0$ |

direct product of cyclic groups of prime order, i.e.

$$\boldsymbol{Z}_{m_1} \otimes \ldots \otimes \boldsymbol{Z}_{m_n} \tag{5.3.2}$$

where $m_i$ is prime and the $\boldsymbol{Z}_{m_i}$ consists of the set $\{0, 1, \ldots, m_i - 1\}$ with addition modulo $m_i$ for $i \in \{0, \ldots, n\}$. However, the cyclic group $\boldsymbol{Z}_m$ has a representation where element $i$ is associated with $e^{2\pi i/m}$, which is only real for all $i$ if $m = 2$, thus our group can only be of the form

$$\boldsymbol{D}_{n\mathrm{h}}(\mathrm{AEM}) = \boldsymbol{Z}_2 \underbrace{\otimes \ldots \otimes}_{n-1} \boldsymbol{Z}_2. \tag{5.3.3}$$

for some $n$. For a given $\ell_{\max}$ (or the same $k_{\max}$), we must have at least $4(\ell_{\max} + 1)$ representations (4 for each $\ell$ including $\ell = 0$). $\boldsymbol{D}_{n\mathrm{h}}$(AEM) has $2^n$ representations, hence

$$n = \lceil \log_2 4(l_{\max} + 1) \rceil. \tag{5.3.4}$$

where $\lceil\rceil$ rounds up to the nearest integer. To show that the irrep product properties hold, first consider $\boldsymbol{Z}_2$. It has character table shown in Table 5.4. In the following we

Table 5.4: Character table for the $\boldsymbol{Z}_2$ group.

| $\boldsymbol{Z}_2$ | 0 | 1 |
|---|---|---|
| $A$ | 1 | 1 |
| $B$ | 1 | −1 |

omit labelling the elements and the irreps and simply write the character table as a matrix. Then $\boldsymbol{Z}_2$ is written

$$\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}. \tag{5.3.5}$$

The character table of a direct product of two groups can then be found by taking an outer product of a matrix. This is defined by

$$\begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix} \otimes \begin{pmatrix} b_{11} & \ldots & b_{1s} \\ \vdots & & \vdots \\ b_{r1} & \ldots & b_{rs} \end{pmatrix} = \tag{5.3.6}$$

$$\begin{pmatrix} a_{11}b_{11} & \ldots & a_{1n}b_{11} & \ldots & a_{11}b_{1s} & \ldots & a_{1n}b_{1s} \\ \vdots & & \vdots & & \vdots & & \vdots \\ a_{m1}b_{11} & \ldots & a_{mn}b_{11} & \ldots & a_{m1}b_{1s} & \ldots & a_{mn}b_{1s} \\ \vdots & & \vdots & & \vdots & & \vdots \\ a_{11}b_{r1} & \ldots & a_{1n}b_{r1} & \ldots & a_{11}b_{rs} & \ldots & a_{1n}b_{rs} \\ \vdots & & \vdots & & \vdots & & \vdots \\ a_{m1}b_{r1} & \ldots & a_{mn}b_{r1} & \ldots & a_{m1}b_{rs} & \ldots & a_{mn}b_{rs} \end{pmatrix}. \tag{5.3.7}$$

Now, $\boldsymbol{C}_{2\mathrm{v}}(\mathrm{M}) = \boldsymbol{Z}_2 \otimes \boldsymbol{Z}_2$. So $\boldsymbol{D}_{2\mathrm{h}}$(AEM) works for $k_{\max} = 0$ where we label the 0-superscripted irreps of $\boldsymbol{D}_{2\mathrm{h}}$(AEM) in the same order as Table 5.2. We call the $\boldsymbol{D}_{2\mathrm{h}}$(AEM) character table matrix $G$, i.e.

$$G = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix} \tag{5.3.8}$$

and say that this has the property that $G \times G = G$ in that for any irreps of $\Gamma_1$ and $\Gamma_2$ of $\boldsymbol{D}_{2\mathrm{h}}$(AEM), the result of $\Gamma_1 \times \Gamma_2$ is another irrep in $\boldsymbol{D}_{2\mathrm{h}}$(AEM). If we now consider the case $\ell_{\max} = 3$, we have the character matrix

$$\begin{pmatrix} G & G & G & G \\ G & -G & G & -G \\ G & G & -G & -G \\ G & -G & -G & G \end{pmatrix}. \tag{5.3.9}$$

This is a more succinct way to write Table 5.2. The first row containing four irreps is 0-superscript labelled as in Table 5.2. The next four are 1-superscripted, in the same order as the corresponding 0-superscripted irreps, and so on. With this construction, for any two irreps of $\Gamma_1$ and $\Gamma_2$ of $\boldsymbol{D}_{4\mathrm{h}}$(AEM), if they have the same superscript, $\Gamma_1 \otimes \Gamma_2$ will cancel the minus signs of the $G$s and the result will be an irrep of the 0-superscripted block. Moreover, this will be the same irrep as one would obtain had we used the $\boldsymbol{C}_{2\mathrm{v}}(\mathrm{M})$ group. If the irreps do not have the same superscript, we will have a $-G \times G = -G$ type multiplication whose result will be an irrep not in the 0-superscripted block. We can see this easily generalises to any $\ell_{\max}$. This concludes the definition of $\boldsymbol{D}_{n\mathrm{h}}$(AEM). Table 5.5 states how the characters for $\boldsymbol{D}_{n\mathrm{h}}$(AEM) are defined for an arbitrary $n$, though in practice one would use the outer product formulation to build the character table.

Table 5.5: The character table for the $\boldsymbol{D}_{n\mathrm{h}}$(AEM) group for some $n$. The character for the corresponding to the $i$th row and $j$th column is given by $f(i,j) = f(j,i)$. Here, $k$ is given by $k = 2^{n-2} - 1$. Starting the row and column number from zero, the output of the function $f$ is as follows: first $i$ and $j$ are converted into binary numbers and their bitwise sum is calculated. If the number of 1s in the result is odd, character is $-1$; if it is even, the character is 1. For example, 7 and 4 would be 111&100 = 100, so the number of 1s is 1 (odd) and thus the character is $-1$.

| $\boldsymbol{D}_{n\mathrm{h}}$(AEM) | $E^0$ | $C_2^0$ | $\sigma^0$ | $\sigma_v^0$ | ... | $\sigma_v^k$ |
|---|---|---|---|---|---|---|
| $A_1^0$ | 1 | 1 | 1 | 1 | | $f(4k+3,0)$ |
| $B_1^0$ | 1 | −1 | 1 | −1 | | $f(4k+3,1)$ |
| $A_2^0$ | 1 | 1 | −1 | −1 | | $f(4k+3,2)$ |
| $B_2^0$ | 1 | −1 | −1 | 1 | | $f(4k+3,3)$ |
| ⋮ | | | | | | |
| $B_2^k$ | | | | | | $f(4k+3,4k+3)$ |

The structure introduced is similar to that of Extended molecular symmetry groups $\boldsymbol{G}$(EM), which are used in the ro-vibrational problems of non-rigid molecules such as hydrogen peroxide ($H_2O_2$) [48] and ethane ($C_2H_6$) [5].

In these cases, the group structure is

$$\boldsymbol{G}(\mathrm{EM}) = \boldsymbol{G}(\mathrm{M}) \otimes \{E, E'\} \cong \boldsymbol{G}(\mathrm{M}) \otimes \boldsymbol{Z}_2 \tag{5.3.10}$$

while $\boldsymbol{D}_{n\mathrm{h}}$(AEM) is given by

$$\boldsymbol{Z}_2 \underbrace{\otimes \ldots \otimes}_{n-1} \boldsymbol{Z}_2 = \boldsymbol{C}_{2\mathrm{v}}(\mathrm{M}) \otimes \boldsymbol{Z}_2 \underbrace{\otimes \ldots \otimes}_{n-3} \boldsymbol{Z}_2 \tag{5.3.11}$$

which clarifies how this sort of construction would work for any arbitrary group $G$:

$$\boldsymbol{G} \otimes \boldsymbol{Z}_2 \otimes \ldots \otimes \boldsymbol{Z}_2. \tag{5.3.12}$$

The basis functions for these types of molecules also contain artificial, unphysical irreps. For example, the rotational or vibrational wavefunctions of $\boldsymbol{G}_{36}$(EM), the MS group of ethane $C_2H_6$, can transform as odd or even under $E'$, labelled as $s$ and $d$, respectively. Likewise, the full rotation-vibrational wavefunction can only be of $s$ type and the artificial $d$ types must be eliminated [5]. Chapter 7 will discuss this further.

The artificial symmetry group introduced here has been implemented in the existing structure of the TROVE program, as illustrated below.

## 5.4 Example application of $D_{n\mathrm{h}}$(AEM)

Let us consider a ro-vibrational basis set of a centrosymmetric linear triatomic ($CO_2$) in a symmetry adapted form as (compare to Eq. (5.2.3)) and a bisector frame [7]:

$$\Phi^{\Gamma} = \left\{ [\phi_{i_1}(\mathcal{R}_1, \mathcal{R}_2)]^{\Gamma_1} [\phi_{i_2,\ell}(\rho)]^{\Gamma_2} |J, K, m, \Gamma_{\mathrm{rot}}\rangle \right\}^{\Gamma_{\mathrm{tot}}},$$

where $\Gamma$s are the irreps, $K = |k|$, $i$s are function labels, $\mathcal{R}_1$ and $\mathcal{R}_2$ are the equivalent stretches, and $\rho$ is the inter-bond angle. Assuming a small rotational problem with $J_{\max} = 4$, we need $k_{\max} = 4$, and hence $n = \lceil \log_2 20 \rceil = 5$ (Eq. (5.3.4)). There are $2^5 = 32$ irreps in the group $\boldsymbol{D}_{5\mathrm{h}}$(AEM), but our basis functions can be only one of (for $k_{\max} = 4$) $4(4 + 1) = 20$ irreps, and the remainder is not utilised. However, they are still necessary when combining basis functions as was illustrated in Table 5.3.

The procedure works as follows: initially the reduced Hamiltonian matrix for the bending motion is manually made block diagonal in $\ell$ and the whole matrix is diagonalized. In principle, TROVE could treat these blocks separately and diagnolise each individually. However, in practice, the dimension is small enough that diagnolizing the entire matrix is computationally acceptable.

To obtain symmetrised (in $\boldsymbol{D}_{5\mathrm{h}}$(AEM)) eigenfunctions for both the bending and stretching Hamiltonians, the standard TROVE symmetrisation procedure is applied. The resulting irreps are the 0-superscripted ones of $\boldsymbol{D}_{5\mathrm{h}}$(AEM) and they correspond to the $\boldsymbol{C}_{2\mathrm{v}}$(M) ones. To correctly symmetrise this way, the operation $O^m$, using the notation of Table 5.2, has the same behaviour on the coordinates as operation $O^0$.

Once 0-superscripted irreps are obtained, the rotational and bending wavefunctions are reassigned to the appropriate $k$- or $\ell$-superscripted irrep.

Table 5.6 lists a few vibrational energies with their symmetry assignment in $\boldsymbol{D}_{5h}$(AEM). The irreps generated by the bending basis functions can only be $A_1$, and they are assigned to $A_1^\ell$ for a given $\ell$. In this case $\ell \leq 4$. The stretching functions are all assigned as a 0-superscripted irrep. This ensures that when they are combined with the bending function, the "base" irrep obtained (i.e. the letter and subscript of the irrep label) would have been the same had we used the $\boldsymbol{C}_{2v}$(M) group. The only 0-superscripted irrep is the one where $\ell = k$, as expected, and only this would be retained from this set.

Table 5.7 shows the $A_1^0$ and $A_2^0$ states for the $J = 2$ case. Here we see that $k$ ranges from 0 to 2, and the rotational irreps are superscripted accordingly. The combined stretching and bending irreps are the same as their corresponding rotational irrep, and thus the full functions are of type $A_1^0$ or $A_2^0$.

Table 5.6: The $J = 0$ vibrational states, including the symmetry $\Gamma$ in $\boldsymbol{D}_{5h}$(AEM) of the full state. $v_1$, $v_2$, $\ell$ and $v_3$ are the linear molecule quantum numbers of $CO_2$.

| $\Gamma$ | $E(\text{cm}^{-1})$ | $v_1$ | $v_2$ | $\ell$ | $v_3$ |
|---|---|---|---|---|---|
| $A_1^0$ | 0.00 | 0 | 0 | 0 | 0 |
| $A_1^1$ | 667.75 | 0 | 1 | 1 | 0 |
| $A_1^0$ | 1285.40 | 0 | 2 | 0 | 0 |
| $A_1^2$ | 1336.67 | 0 | 2 | 2 | 0 |
| $A_1^0$ | 1388.21 | 1 | 0 | 0 | 0 |
| $A_1^1$ | 1932.82 | 0 | 3 | 1 | 0 |
| $A_1^3$ | 2006.73 | 0 | 3 | 3 | 0 |
| $A_1^1$ | 2077.23 | 1 | 1 | 1 | 0 |
| $B_2^0$ | 2349.17 | 0 | 0 | 0 | 1 |
| $A_1^0$ | 2548.34 | 1 | 2 | 0 | 0 |
| $A_1^2$ | 2586.55 | 0 | 4 | 2 | 0 |
| $A_1^0$ | 2671.14 | 2 | 0 | 0 | 0 |
| $A_1^4$ | 2677.94 | 0 | 4 | 4 | 0 |
| $A_1^2$ | 2762.27 | 1 | 2 | 2 | 0 |
| $A_1^0$ | 2797.16 | 1 | 2 | 0 | 0 |
| $B_2^1$ | 3004.45 | 0 | 1 | 1 | 1 |
| $A_1^1$ | 3181.79 | 1 | 3 | 1 | 0 |

## 5.5 Chapter Summary

The artificial symmetry group $\boldsymbol{D}_{nh}$(AEM) introduced in this chapter was designed so that imposing a $\ell = k$ condition within the framework TROVE was straightforward. This ensured that the ro-vibrational basis was orthogonal. The construction of $\boldsymbol{D}_{nh}$(AEM) is based on the cyclic group of order 2. $\boldsymbol{D}_{nh}$(AEM) is shown to have all the properties required: basis functions with different $\ell$ or $k$ functions (vibrational or rotational,

Table 5.7: The $J = 2$ ro-vibrational states including the symmetry of the full state and the symmetry of the rotational, stretching, and bending parts. The assignment of $v_1, v_2, \ell, v_3$ and $K$ is approximate and based on the largest contribution to the eigenfunction. Only states with non-zero nuclear statistical weighs are shown.

| $\Gamma_{\text{tot}}$ | $E(\text{cm}^{-1})$ | $\Gamma^{\text{rot}}$ | $K$ | $\Gamma^{\text{stretch}}$ | $\Gamma^{\text{bend}}$ | $v_1$ | $v_2$ | $\ell$ | $v_3$ |
|---|---|---|---|---|---|---|---|---|---|
| $A_1^0$ | 2.34 | $A_1^0$ | 0 | $A_1^0$ | $A_1^0$ | 0 | 0 | 0 | 0 |
| $A_2^0$ | 669.71 | $A_2^1$ | 1 | $A_1^0$ | $A_1^1$ | 0 | 1 | 1 | 0 |
| $A_1^0$ | 1287.75 | $A_1^0$ | 0 | $A_1^0$ | $A_1^0$ | 0 | 2 | 0 | 0 |
| $A_1^0$ | 1337.46 | $A_1^2$ | 2 | $A_1^0$ | $A_1^2$ | 0 | 2 | 2 | 0 |
| $A_1^0$ | 1390.55 | $A_1^0$ | 0 | $A_1^0$ | $A_1^0$ | 1 | 0 | 0 | 0 |
| $A_2^0$ | 1934.78 | $A_2^1$ | 1 | $A_1^0$ | $A_1^1$ | 0 | 3 | 1 | 0 |
| $A_2^0$ | 2079.19 | $A_2^1$ | 1 | $A_1^0$ | $A_1^1$ | 1 | 1 | 1 | 0 |
| $A_1^0$ | 2550.69 | $A_1^0$ | 0 | $A_1^0$ | $A_1^0$ | 1 | 2 | 0 | 0 |
| $A_1^0$ | 2587.33 | $A_1^2$ | 2 | $A_1^0$ | $A_1^2$ | 0 | 4 | 2 | 0 |
| $A_1^0$ | 2673.48 | $A_1^0$ | 0 | $A_1^0$ | $A_1^0$ | 2 | 0 | 0 | 0 |
| $A_1^0$ | 2763.06 | $A_1^2$ | 2 | $A_1^0$ | $A_1^2$ | 1 | 2 | 2 | 0 |
| $A_1^0$ | 2799.50 | $A_1^0$ | 0 | $A_1^0$ | $A_1^0$ | 1 | 2 | 0 | 0 |
| $A_1^0$ | 3006.39 | $B_2^1$ | 1 | $B_2^0$ | $A_1^1$ | 0 | 1 | 1 | 1 |
| $A_2^0$ | 3183.76 | $A_2^1$ | 1 | $A_1^0$ | $A_1^1$ | 1 | 3 | 1 | 0 |

respectively) can be assigned to different irreducible representations of $\boldsymbol{D}_{n\text{h}}$(AEM). As opposed to the ordinary method of using MS groups dictated by the physics of the molecule, here the group was defined with the practical application of being able to exploit TROVE's symmetrisation procedure such that a $k = \ell$ constraint could be applied to the basis set.

# Chapter 6

# Molecular Frames for Symmetry Adaptation

## 6.1 Introduction

The bisector frame was used both for the $H_2CS$ and $CO_2$ ro-vibrational calculations described in this thesis. This is fairly atypical of TROVE. When TROVE Taylor expands a KEO, instead of using an analytic one, the default choice is the Eckart or Sayvetz frame. Crucially, they allow for the classification of the rotational functions as irreps of the MS group. To do so, one needs to know how the Euler angles transform under the MS operations, or, in other words, how the BF-frames rotate under these operations. There are so-called equivalent rotations [31] of the BF-frames which occur during an MS operation.

For general frames, however, the rotation associated with an MS operation depends on the vibrational coordinates. In particular, the $\chi$ Euler angle transforms as [31]

$$\chi' = \pm\chi + \gamma(q_1, \ldots, q_n). \tag{6.1.1}$$

Because the rotational and vibrational coordinates are no longer separable, such a dependency will lead to a breakdown of the rotational basis set symmetrisation and prevent the construction of a symmetry-adapted rotational basis set [67].

For $H_2CS$ and $CO_2$ it was quite obvious how to choose a frame which "respects the symmetry," a vague way to say that the $\gamma$ angle above is constant for all MS operations. Not all molecules will be this obvious. One could stick to the Eckart and Sayvetz frames, but these can only be solved analytically for triatomics [118] and planar molecules [119]. These are also suboptimal for molecules without a well-defined equilibrium. Our goal is to find geometric frames–ones that can be expressed as an analytic function of the vibrational coordinates–for other molecules where the frame's rotations are constants under symmetry operations.

This Chapter attempts to develop a systematic procedure find these alternative frames and is based on the publication [4]. Its results will be used in the next Chapter for the ethane molecule. The structure of the remainder of this Chapter is as follows: Section 6.2 describes the action of the MS operations on the rotational and vibrational coordinates. The example of an ABA molecule illustrates frames where the rotation under MS operations depends on the vibrational coordinates. The transformation and

invariance of the kinetic energy operator is stated and proved in Section 6.3. With the same ABA illustrative example, Section 6.4 demonstrates why rotational basis set symmetrisation is not possible for certain frames. Section 6.5 shows that the correct transformation of the kinetic energy operator was derived for an ABA molecule for both a bisector frame and a bond vector frame. It further proves that the rotations are constant under symmetry operations when using the Eckart frames. Section 6.6 largely follows the pattern of Section 6.5, but for the $CH_3Cl$ molecule. It also provides a geometric frame for which the rotations are constant under symmetry operations. Section 6.7 shows how one can find this frame. Finally, Section 6.8 mentions the possible avenues for further work.

## 6.2 Coordinate transformations

The relation between the SF and BF coordinates is

$$R_\alpha = M r_\alpha. \tag{6.2.1}$$

Under MS operations, the BF-frame rotation is found by ensuring that the transformation of the left hand side is equal to the transformed Euler matrix times the transformed BF coordinates. The latter is found by replacing the vibrational coordinates in $r_\alpha(q)$ with their transformed counterparts $r_\alpha(q') = \tilde{r}_\alpha(q)$.

As an example, consider a triatomic molecule of type ABA and let the A-type atoms (called $A_1$ and $A_2$) have coordinates $\boldsymbol{R}_{A_1}(q_1, q_2, q_3, \phi, \theta, \chi)$ and $\boldsymbol{R}_{A_2}(q_1, q_2, q_3, \phi, \theta, \chi)$. In this representation, $q_1$ and $q_2$ are the AB bonds while $q_3$ is the ABA angle.

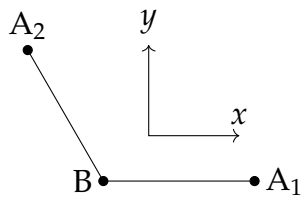



Figure 6.1: The bond vector BF-frame for the ABA molecule with the $x$-axis parallel to the $A_1$–B bond and the $y$-axis in the plane of the three atoms. The axes are offset for clarity.

For the BF-frame, one possible choice is with the $x$-axis parallel to the bond from atom B to $A_1$ (i.e. parallel to the $\boldsymbol{R}_{A_1} - \boldsymbol{R}_B$ vector) and the $z$-axis perpendicular to the plane of ABA. With this choice, the BF coordinates for the ABA molecule are

$$\begin{aligned} r_{A_1} &= (q_1, 0, 0), \\ r_{A_2} &= (q_2 \cos q_3, q_2 \sin q_3, 0), \\ r_B &= (0, 0, 0). \end{aligned} \tag{6.2.2}$$

The operation (12) permutes $q_1$ and $q_2$ resulting in new BF coordinates given by

$$\begin{aligned} \tilde{r}_{A_1} &= (q_2, 0, 0), \\ \tilde{r}_{A_2} &= (q_1 \cos q_3, q_1 \sin q_3, 0), \\ \tilde{r}_B &= (0, 0, 0) \end{aligned} \tag{6.2.3}$$

where $\tilde{r}_\alpha$ is the result after transforming the vibrational coordinates in the expression for $r_\alpha$. In general, the Euler angles also change, and we call the Euler matrix with the transformed angles as $\tilde{M}$. As we have ${R_\alpha}' = \tilde{M}\tilde{r}_\alpha$ and $\tilde{M}$ now defines a rotated BF-fame, we see that the transformed $\tilde{r}_\alpha$ are the BF coordinates of $r_\alpha$ in the rotated BF-frame and after the transformation. The BF-frame is rotated in such a way so as to ensure that ${R_{\mathrm{A}_1}}' = R_{\mathrm{A}_2}$ and vice versa. In this case, the new SF coordinates are

$$ {R_\alpha}' = M_x(\phi + \pi)M_y(\pi - \theta)M_z(2\pi - \chi - q_3)\tilde{r}_\alpha = R_{\alpha'} = \tilde{M}\tilde{r}_\alpha. \tag{6.2.4} $$

In performing this operation, we now have 3 frames: the SF-frame and the two BF-frames. Denoting the new BF-frame unit vectors as $\hat{\boldsymbol{e}}'_g$, i.e. with an asterisk, the components of $\boldsymbol{R}_\alpha$ are then

$$ \boldsymbol{R}_\alpha \cdot \hat{\boldsymbol{e}}'_g = \tilde{r}_{\alpha,g}. \tag{6.2.5} $$

It is of interest to note what the transformed Cartesian coordinates in the original BF-frame are (denoted by ${r_\alpha}'$). Since the Euler matrix of this frame is unchanged, if ${R_\alpha}' = R_{\alpha'}$, we have

$$ {R_\alpha}' = M{r_\alpha}' = R_{\alpha'} = Mr_{\alpha'} = \tilde{M}\tilde{r}_\alpha \tag{6.2.6} $$

and so ${r_\alpha}' = r_{\alpha'}$. We also have

$$ {r_\alpha}' = M^T\tilde{M}\tilde{r}_\alpha \equiv N\tilde{r}_\alpha \tag{6.2.7} $$

where $N \equiv M^T\tilde{M}$ is the rotation matrix relating the coordinates of $\boldsymbol{R}_\alpha$ in the new BF-frame to that of the old BF-frame. In this case, it is

$$ N_{(12)} = \begin{pmatrix} \cos q_3 & \sin q_3 & 0 \\ \sin q_3 & -\cos q_3 & 0 \\ 0 & 0 & -1 \end{pmatrix}. \tag{6.2.8} $$

For completeness, the MS transformation properties of the $s$- and $t$-vectors are stated in Section 6.3, with the proofs in Appendix 6.A.

## 6.3 KEO transformation

### 6.3.1 $s$- and $t$-vector and operator transformations

For the transformation $R_\alpha \rightarrow R_{\alpha'}$ and $q_k \rightarrow q_{k'}$, the definition and transformation of the vectors and operators are summarised in Table 6.1. Table 6.2 shows the transformation for the vibrational $s$-vectors and vibrational momentum operators for more complex vibrational transformations. The transformation of the rotational $s$-vectors and angular momentum operators is unchanged.

In Appendix 6.A it is shown that the transformed vibrational $s$-vectors can also be expressed in the new BF-frame by performing the vibrational coordinate transformation (which we shall denote as $\tilde{s}_{k,\alpha g}$) on the original BF components $s_{k,\alpha g}$. The same procedure also applies to the other vectors.

Table 6.1: The transformation of the $t$- and $s$-vectors. This assumes that ${R_\alpha}' = R_{\alpha'}$ and that ${q_k}' = q_{k'}$. The matrix is $N$ is that of Eq. (6.2.7). The term $\dot{\gamma}_k = \partial_k \gamma$ where $\gamma$ is the angle in the $\chi$ transformation $\chi' = \pm\chi + \gamma$ of Table 6.3. The sign in front $\dot{\gamma}_k$ corresponds to the left and right transformations of Table 6.3. Summation over repeated indices is assumed.

| Object | Equation | Transformation |
|---|---|---|
| $\boldsymbol{s}_{k,\alpha}$ | $\hat{\boldsymbol{e}}_F \partial_{\alpha F} q_k$ | $\boldsymbol{s}_{k',\alpha'}$ |
| $\boldsymbol{s}_{g,\alpha}$ | Eq. (2.3.21) | $N_{hg}\boldsymbol{s}_{h,\alpha'} \mp \dot{\gamma}_k \boldsymbol{s}_{k,\alpha} N_{zg}$ |
| $\boldsymbol{t}_{k,\alpha}$ | $\partial_k \boldsymbol{R}_\alpha$ | $\boldsymbol{t}_{k',\alpha'} \pm \dot{\gamma}_{k'} \boldsymbol{t}_{z,\alpha'}$ |
| $\boldsymbol{t}_{g,\alpha}$ | $\hat{\boldsymbol{e}}_g \times \boldsymbol{R}_\alpha$ | $N_{hg}\boldsymbol{t}_{h,\alpha'}$ |
| $\hat{J}_g$ | Eq. (2.3.9) | $N_{hg}\hat{J}_h$ [31] |
| $\hat{\pi}_k$ | $-i\hbar\, \partial/\partial q_k$ | $\hat{\pi}_{k'} \pm \dot{\gamma}_{k'} \hat{J}_z$ |
| $a^2 + b^2 = 1$ | | |

Table 6.2: The transformation of $q_k$ in the left column and the corresponding transformed vibrational $s$-vector and vibrational momentum operator. The transformations of the Cartesian coordinates, the rotational $s$-vectors, and the angular momentum operators remain the same as Table 6.1.

| Coordinate | $s$-vector | Momentum operator |
|---|---|---|
| $q_k \rightarrow -q_{k'}$ | $-\boldsymbol{s}_{k',\alpha'}$ | $-(\hat{\pi}_{k'} \pm \dot{\gamma}_{k'} \hat{J}_z)$ |
| $q_k \rightarrow a q_k + b q_{k'}$ | $a\boldsymbol{s}_{k,\alpha'} + b\boldsymbol{s}_{k',\alpha'}$ | $a(\hat{\pi}_k \pm \dot{\gamma}_k \hat{J}_z) + b(\hat{\pi}_{k'} \pm \dot{\gamma}_{k'} \hat{J}_z)$ |

Table 6.3: The change in the Euler angles for a given rotation of the BF-frame. $R_\alpha^\pi$ and $R_z^\beta$ are rotations of the BF-frame. $R_\alpha^\pi$ is a rotation of $\pi$ radians about an axis in the $xy$ plane making an angle $\alpha$ about with the $x$-axis ($\alpha$ is measured in the right handed sense about the $z$-axis) and $R_z^\beta$ is a rotation of $\beta$ radians about the $z$-axis ($\beta$ measured in the right handed sense about the $z$-axis) [31].

| | $R_\alpha^\pi$ | $R_z^\beta$ |
|---|---|---|
| $\theta$ | $\pi - \theta$ | $\theta$ |
| $\phi$ | $\phi + \pi$ | $\phi$ |
| $\chi$ | $2\pi - 2\alpha - \chi$ | $\chi + \beta$ |

### 6.3.2 $G$-matrix transformation

From the transformation properties of the $s$-vectors, we can obtain the transformation of the $G$-matrix. Assuming the vibrational coordinates $q_i$ and $q_{i'}$ transform to $q_j$ and $q_{j'}$, respectively, the vibrational, rotational, and Coriolis components of the $G$-matrix transform as

$$\begin{aligned} G_{ii'}' &= G_{jj'}, \\ G_{gg'}' &= N_{hg}N_{h'g'}G_{hh'} + \dot{\gamma}_k\dot{\gamma}_{k'}G_{kk'}N_{zg}N_{zg'} - \dot{\gamma}_k G_{kh}(N_{hg'}N_{zg} + N_{hg}N_{zg'}), \\ G_{gi}' &= N_{hg}G_{hj} - \dot{\gamma}_k G_{kj}N_{zg} \end{aligned} \tag{6.3.1}$$

where the rotation matrix $N$ is that of Eq. (6.2.7) and relates the directions of the original and rotated BF-frames and $\dot{\gamma}_k = \partial_k\gamma$ defined in Table 6.1. The left hand side of the above equations are of the form $G_{ab} \circ P(q_1, \ldots, q_n)$, where $P$ is the MS operation which transforms the coordinates first, after which the $G$ matrix element $G_{ab}$ is applied to the transformed coordinates. The right hand side is in terms of $G_{ab}(q_1, \ldots, q_n)$ which is the matrix element $G_{ab}$ applied to the un-transformed coordinates.

From the definition of the $G$ matrix, we have

$$G_{ab} = \sum_\alpha \frac{1}{m_\alpha} s_{\alpha,aF}s_{\alpha,bF} = \sum_\alpha \frac{1}{m_\alpha} s_{\alpha,ag}s_{\alpha,bg} = \sum_\alpha \frac{1}{m_\alpha} s_{\alpha,ag'}s_{\alpha,bg'} \tag{6.3.2}$$

where the $s$-vector components are in the SF-, old BF-, and new BF-frame, respectively. Since the new BF-frame components of the transformed $s$-vectors are given by $\tilde{s}_{\alpha,a}$, to transform the $G$ matrix, one can also perform the vibrational coordinate transformation on the original $G$ matrix.

Action of the full Hamiltonian on a function $\psi(q_1, \ldots, q_n, \phi, \theta, \chi)$ is of the form

$$\begin{aligned} \hat{H}\psi = \tfrac{1}{2}\hat{\pi}_a G_{ab}\hat{\pi}_b\psi + U\psi + V\psi = \; & \tfrac{1}{2}G_{kk'}\hat{\pi}_k\hat{\pi}_{k'}\psi + \tfrac{1}{2}(\hat{\pi}_k G_{kk'})(\hat{\pi}_{k'}\psi) + \tfrac{1}{2}G_{gh}\hat{J}_g\hat{J}_h\psi \\ & + \tfrac{1}{2}G_{gk}\hat{J}_g\hat{\pi}_k\psi + (\hat{\pi}_k G_{kg})\hat{J}_g\psi + \tfrac{1}{2}G_{kg}\hat{\pi}_k\hat{J}_g\psi + U\psi + V\psi. \end{aligned} \tag{6.3.3}$$

where implicit summation of repeated indices is assumed. In Appendix 6.B it is shown that the Hamiltonian is invariant under the transformations of Table 6.1.

## 6.4 Rotational symmetrisation

TROVE eigenfunctions have the form

$$\sum_{ab} c_{ab} \, |J, |k|\rangle_a \, \varphi_b(q_1, \ldots, q_n). \tag{6.4.1}$$

If one applies the full Hamiltonian to the function $\varphi_b(q_1, \ldots, q_n)$, then the rotational and Coriolis parts drop out, and we can find the transformation of $\hat{H}\varphi$ under MS operations. The vibrational operators then transform as $\hat{\pi}_k' = \hat{\pi}_{k'}$, i.e. without the second term, and the Hamiltonian is invariant, showing the vibrational eigenfunctions also are irreps of the same group [31].

To symmetrise the rotational part, however, one uses the transformation properties of the rigid rotors under BF-frame rotations. These are defined in Table 6.3 and given by

$$R_z^\beta \left|J,k,m\right\rangle = e^{ik\beta}\left|J,k,m\right\rangle \tag{6.4.2}$$

and

$$R_\alpha^\pi \left|J,k,m\right\rangle = (-1)^J e^{-2ik\alpha}\left|J,-k,m\right\rangle . \tag{6.4.3}$$

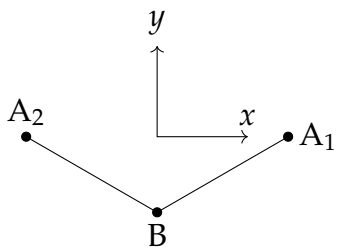



Figure 6.2: The principal axis system for the ABA molecule in vibrational equilibrium with the $z$-axis pointing out of the plane and the $y$-axis bisecting the $A_1$–B–$A_2$ bond. The axes are offset for clarity.

If we refer back to our ABA example and instead use the principal axis system at the vibrational equilibrium, shown in Figure 6.2, with the $z$-axis pointing out of the plane and the $y$-axis bisecting the ABA angle, then one can follow the prescription of Section. 12.1 of [31] and find the "equivalent rotations" of the axes during an MS operation. There are four symmetry operations for the MS group $\boldsymbol{C}_{2v}$(M): $\{E, E^*, (12), (12)^*\}$, and they have the equivalent rotations of $R_z^0, R_z^\pi, R_{\pi/2}^\pi, R_0^\pi$, respectively. This results in the transformation properties of $(\left|J,k,m\right\rangle, \left|J,-k,m\right\rangle)$ being

$$\begin{aligned}
R^0\begin{pmatrix}\left|J,k,m\right\rangle\\ \left|J,-k,m\right\rangle\end{pmatrix} &= \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}\begin{pmatrix}\left|J,k,m\right\rangle\\ \left|J,-k,m\right\rangle\end{pmatrix},\\
R_z^\pi\begin{pmatrix}\left|J,k,m\right\rangle\\ \left|J,-k,m\right\rangle\end{pmatrix} &= (-1)^k\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}\begin{pmatrix}\left|J,k,m\right\rangle\\ \left|J,-k,m\right\rangle\end{pmatrix},\\
R_{\pi/2}^\pi\begin{pmatrix}\left|J,k,m\right\rangle\\ \left|J,-k,m\right\rangle\end{pmatrix} &= (-1)^{J+k}\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}\begin{pmatrix}\left|J,k,m\right\rangle\\ \left|J,-k,m\right\rangle\end{pmatrix},\\
R_0^\pi\begin{pmatrix}\left|J,k,m\right\rangle\\ \left|J,-k,m\right\rangle\end{pmatrix} &= (-1)^J\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}\begin{pmatrix}\left|J,k,m\right\rangle\\ \left|J,-k,m\right\rangle\end{pmatrix}
\end{aligned} \tag{6.4.4}$$

which is a representation of the MS group $\boldsymbol{C}_{2v}$(M) of the ABA molecule. Irreducible representations can be formed with projection operators. The result is that the symmetrised functions $\left|J,|k|,m,p\right\rangle$ have the form

$$\left|J,|k|,m,p\right\rangle = \frac{1}{\sqrt{2}}(\left|J,k,m\right\rangle + p(J,|k|)\left|J,-k,m\right\rangle) \tag{6.4.5}$$

where $p(J,|k|) = \pm 1$. The values of $p$ and the symmetry are summarised in Table 6.4.

Table 6.4: The irrep of the function $|J, |k|, m, p\rangle$ for the MS group $\boldsymbol{C}_{2v}$(M) depending on the values of $J$, $|k|$, and $p$.

| $J$ | $\lvert k\rvert$ | $p(J, \lvert k\rvert)$ | $\Gamma$ |
|---|---|---|---|
| even | even | 1 | $A_1$ |
| even | even | −1 | $A_2$ |
| odd | even | −1 | $A_1$ |
| odd | even | 1 | $A_2$ |
| even | odd | −1 | $B_1$ |
| even | odd | 1 | $B_2$ |
| odd | odd | 1 | $B_1$ |
| odd | odd | −1 | $B_2$ |

The key aspect of these transformations that enable this is that the rotation matrices are constant (up to $J$ and $k$) which allow standard combinations of the rigid rotors (dependant on $J$ and $k$) that transform irreducibly under the MS group transformations [67].

In our original example, however, we saw that the rotations associated with the MS group were not the equivalent rotations, nor were they constant. The operation (12), for example, included the rotations $R_0^\pi R_z^{q_3}$ whose effect on the rigid rotors is

$$R_0^\pi R_z^{q_3} \begin{pmatrix} |J,k,m\rangle \\ |J,-k,m\rangle \end{pmatrix} = (-1)^J \begin{pmatrix} 0 & e^{-ikq_3} \\ e^{ikq_3} & 0 \end{pmatrix} \begin{pmatrix} |J,k,m\rangle \\ |J,-k,m\rangle \end{pmatrix}. \tag{6.4.6}$$

With this transformation, the representation of the group now depends on the angle $q_3$, and one can no longer form $q_3$ independent combinations of the rigid rotors that transform as the irreps of the group. For example, if $J$ and $k$ are both even, the $A_1$ irrep has the form

$$|J, |k|\rangle^{A_1} = \frac{1}{\sqrt{2}}(|J,k,m\rangle + e^{-ikq_3}\,|J,-k,m\rangle). \tag{6.4.7}$$

Thus, one cannot have separate symmetry-adapted rotational functions used in Eq. (6.4.1) using this choice of BF-frame. A separate rotational basis set can still be used, but it cannot be assigned a symmetry label of the group in question. To do this, a different BF-frame is required.

The upshot of this result is that to keep the rotational and vibrational coordinates separate, and thus for the rotational functions to transform with matrices that do not depend on the vibrational coordinates, the rotation angle must be constant and its axis of rotation relative to the BF-frame must also be constant. This will ensure $N$ is constant and $\dot{\gamma}_k = 0$. Whatever the rotation axis is, the angle must be $2\pi/n$ where $n$ is the order of the MS operation.

## 6.5 ABA example

### 6.5.1 Bond vector frame

To demonstrate the MS transformation of the KEO and its invariance, we return to the ABA molecule in more detail. For our original BF-frame we know that matrix $N$ for the operation (12) is Eq. (6.2.8). The components of the $G$ matrix are as follows [7]: the rotational part is

$$G_{\text{rot}} = \begin{pmatrix} \dfrac{\cot q_3^2}{\mu_{\text{AB}} q_1^2} + \dfrac{\csc q_3^2}{\mu_{\text{AB}} q_2^2} - \dfrac{2\cot q_3 \csc q_3}{m_{\text{B}} q_1 q_2} & \dfrac{\cot q_3}{\mu_{\text{AB}} q_1^2} - \dfrac{\csc q_3}{m_{\text{B}} q_1 q_2} & 0 \\ \dfrac{\cot q_3}{\mu_{\text{AB}} q_1^2} - \dfrac{\csc q_3}{m_{\text{B}} q_1 q_2} & \dfrac{1}{\mu_{\text{AB}} q_1^2} & 0 \\ 0 & 0 & \dfrac{1}{\mu_{\text{AB}} q_1^2} \end{pmatrix} \tag{6.5.1}$$

where $\mu_{\text{AB}}$ is the reduced mass

$$\frac{1}{\mu_{\text{AB}}} = \frac{1}{m_{\text{A}}} + \frac{1}{m_{\text{B}}} \tag{6.5.2}$$

and $m_{\text{A}}$ and $m_{\text{B}}$ are the masses of A and B, respectively; the Coriolis part is

$$G_{\text{cor}} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & \dfrac{\sin q_3}{m_{\text{B}} q_1} \\ 0 & 0 & -\dfrac{1}{\mu_{\text{AB}} q_1^2} + \dfrac{\cos q_3}{m_{\text{B}} q_1 q_2}. \end{pmatrix} \tag{6.5.3}$$

where the rows correspond to the vibrational coordinates and the columns to the axes; the vibrational part of the $G$ matrix

$$G_{\text{vib}} = \begin{pmatrix} \dfrac{1}{\mu_{\text{AB}}} & \dfrac{\cos q_3}{m_{\text{B}}} & -\dfrac{\sin q_3}{m_{\text{B}} q_2} \\ \dfrac{\cos q_3}{m_{\text{B}}} & \dfrac{1}{\mu_{\text{AB}}} & -\dfrac{\sin q_3}{m_{\text{B}} q_1} \\ -\dfrac{\sin q_3}{m_{\text{B}} q_2} & -\dfrac{\sin q_3}{m_{\text{B}} q_1} & \dfrac{1}{\mu_{\text{AB}}}\left(\dfrac{1}{q_1^2} + \dfrac{1}{q_2^2}\right) - \dfrac{2\cos q_3}{M_B q_1 q_2} \end{pmatrix}. \tag{6.5.4}$$

If we list the vibrational momentum and angular momentum operators as

$$\hat{\Pi} = \begin{pmatrix} \hat{\pi}_1 \\ \hat{\pi}_2 \\ \hat{\pi}_3 \end{pmatrix}, \quad \hat{J} = \begin{pmatrix} \hat{J}_x \\ \hat{J}_y \\ \hat{J}_z \end{pmatrix} \tag{6.5.5}$$

then the full kinetic energy operator (besides the pseudo-potential) is given by

$$\frac{1}{2}\hat{\Pi}^T G_{\text{vib}} \hat{\Pi} + \frac{1}{2}\hat{\Pi}^T G_{\text{cor}} \hat{J} + \frac{1}{2}\hat{J}^T G_{\text{cor}}^T \hat{\Pi} + \frac{1}{2}\hat{J}^T G_{\text{rot}} \hat{J}.$$

During the MS operation, we transform $q_1 \leftrightarrow q_2$ in each of the $G$ matrix elements. The operator $\hat{\Pi}$ transforms to

$$\hat{\Pi}' = C\hat{\Pi} - D\hat{J} \tag{6.5.6}$$

where $C$ and $D$ are given by

$$C = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad D = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{6.5.7}$$

The rotational operators transform as $\hat{J}' = N_{(12)}^T \hat{J} \equiv N^T \hat{J}$. The transformed kinetic energy operator is then

$$\begin{aligned} &\frac{1}{2}\hat{\Pi}^T \underbrace{(C^T \tilde{G}_{\text{vib}} C)}_{G_{\text{vib}}} \hat{\Pi} \\ &+ \frac{1}{2}\hat{\Pi}^T \underbrace{(C^T \tilde{G}_{\text{cor}} N^T - C^T \tilde{G}_{\text{vib}} D)}_{G_{\text{cor}}} \hat{J} \\ &+ \frac{1}{2}\hat{J}^T \underbrace{(N\tilde{G}_{\text{cor}}^T C - D^T \tilde{G}_{\text{vib}} C)}_{G_{\text{cor}}^T} \hat{\Pi} \\ &+ \frac{1}{2}\hat{J}^T \underbrace{(N\tilde{G}_{\text{rot}} N^T + D^T \tilde{G}_{\text{vib}} D - D^T \tilde{G}_{\text{cor}} N^T - N\tilde{G}_{\text{cor}}^T D)}_{G_{\text{rot}}} \hat{J} \end{aligned} \tag{6.5.8}$$

where the expressions in brackets can be confirmed to be the original $G$ matrix elements. The other generator for the MS group of the ABA molecule is $E^*$. For this operation, the vibrational coordinates are unchanged so that $\tilde{r}_\alpha = r_\alpha$. For ${R_\alpha}' = -R_\alpha$ to be satisfied, the BF-frame rotates by $\pi$ about the $z$-axis. The rotation $N_{E^*}$ is then

$$N_{E^*} = M_z(\pi) = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{6.5.9}$$

and we have ${R_\alpha}' = MN_{E^*}\tilde{r}_\alpha = -R_\alpha$, as required. The corresponding operation $R_z^\pi$ changes the rigid rotors by the second line of Eq. (6.4.4) and thus an irreducible representation for the subgroup $\{E, E^*\}$ can be formed. In other words, we can use a product basis set with this BF-frame to symmetrise the rotation basis functions by parity (the subgroup $\{E, E^*\}$), but not of the full MS group $\boldsymbol{C}_{2\text{v}}$(M).

### 6.5.2 Bisector frame

Another obvious BF-frame choice for the ABA molecule is the $y$-axis bisecting the angle and the $z$-axis pointing out of the plane. The equivalent rotations for this frame were given in Eq. (6.4.4). There is a subtle difference, however. In that frame, the $y$-axis bisects ABA angle as this is the orientation of the principal axis system at vibrational

equilibrium. In this case, the $y$-axis bisects the ABA angle for all geometries. The BF coordinates in this frame are given by

$$\begin{aligned} r_{A_1} &= (q_1 \sin(q_3/2), q_1 \cos(q_3/2), 0), \\ r_{A_2} &= (-q_2 \sin(q_3/2), q_2 \cos(q_3/2), 0), \\ r_B &= (0, 0, 0). \end{aligned} \tag{6.5.10}$$

This time, when we apply the (12) operation, the matrix $N_{(12)}$ given by

$$N_{(12)} = M_y(\pi) = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \tag{6.5.11}$$

has exactly the property that $R_1{}' = MN_{(12)}\tilde{r}_1 = R_2$, and likewise for $R_2$. In this case we can build a symmetry adapted rotational basis set because the rotation angle is $\pi = 2\pi/2$ and the axis of rotation is constant relative to the BF-frame for this operation.

For this frame, the vibrational $G$ matrix is [7] Eq. (6.5.4).

The rotational $G$ matrix is

$$G_{\text{rot}} = \begin{pmatrix} \dfrac{\sec^2(q_3/2)}{4\mu_{AB}}\left(\dfrac{1}{q_2^2}+\dfrac{1}{q_1^2}\right)+\dfrac{\sec^2(q_3/2)}{2m_B q_1 q_2} & \dfrac{\csc q_3}{2\mu_{AB}}\left(\dfrac{1}{q_2^2}-\dfrac{1}{q_1^2}\right) & 0 \\ \dfrac{\csc q_3}{2\mu_{AB}}\left(\dfrac{1}{q_2^2}-\dfrac{1}{q_1^2}\right) & \dfrac{\csc^2(q_3/2)}{4\mu_{AB}}\left(\dfrac{1}{q_2^2}+\dfrac{1}{q_1^2}\right)-\dfrac{\csc^2(q_3/2)}{2m_B q_1 q_2} & 0 \\ 0 & 0 & \dfrac{1}{4\mu_{AB}}\left(\dfrac{1}{q_2^2}+\dfrac{1}{q_1^2}\right)+\dfrac{\cos q_3}{4m_B q_1 q_2} \end{pmatrix} \tag{6.5.12}$$

The Coriolis $G$ matrix is

$$G_{\text{cor}} = \begin{pmatrix} 0 & 0 & -\dfrac{\sin q_3}{2m_B q_2} \\ 0 & 0 & \dfrac{\sin q_3}{2m_B q_1} \\ 0 & 0 & \dfrac{1}{2\mu_{AB}}\left(\dfrac{1}{q_2^2}-\dfrac{1}{q_1^2}\right) \end{pmatrix}. \tag{6.5.13}$$

In this BF-frame, $D = 0$ for operation (12). One can again use Eq. (6.5.8) to confirm the invariance of the $G$ matrix.

### 6.5.3 Transformation through the frame condition

A more qualitative way to understand these results, which will prove important when studying the Eckart frames, is to examine the three conditions which define the BF-frame

in the Sørensen approach. In the case of the bisector frame of the ABA molecule

$$\begin{aligned} C^1 &= z_{\mathrm{A}_1} - z_B = 0, \\ C^2 &= z_{\mathrm{A}_2} - z_B = 0, \\ C^3 &= (x_{\mathrm{A}_1} - x_B)(y_{\mathrm{A}_2} - y_B) + (x_{\mathrm{A}_2} - x_B)(y_{\mathrm{A}_1} - y_B) = 0 \end{aligned} \tag{6.5.14}$$

where the subscript refers to the atom the coordinate refers to. It should be pointed out that these do not totally define the BF-frame orientation, as we can rotate the BF-frame by $\pi$ about any of the BF-frame axes and the conditions would still hold. In general the conditions also do not define the handedness of the axes as one can change all the BF coordinates $r_\alpha$ to $-r_\alpha$ (i.e. make the axes left handed) and the conditions would still hold.

An MS operation applied to each condition relabels the BF coordinates. The result is the transformed BF coordinates in the original BF-frame. For example, for the (12) operation, they become

$$\begin{aligned} C^{1'} &= z_{\mathrm{A}_2} - z_B = 0, \\ C^{2'} &= z_{\mathrm{A}_1} - z_B = 0, \\ C^{3'} &= (x_{\mathrm{A}_2} - x_B)(y_{\mathrm{A}_1} - y_B) + (x_{\mathrm{A}_1} - x_B)(y_{\mathrm{A}_2} - y_B) = 0 \end{aligned} \tag{6.5.15}$$

which are actually still zero, a manifestation of the possible choice we had in the BF-frame that satisfies these conditions. We know that we must apply a $\pi$ rotation about the $z$-axis, after which they become

$$\begin{aligned} C^{1'} &= z_{\mathrm{A}_1} - z_B = 0, \\ C^{2'} &= z_{\mathrm{A}_2} - z_B = 0, \\ C^{3'} &= (x_{\mathrm{A}_1} - x_B)(y_{\mathrm{A}_2} - y_B) + (x_{\mathrm{A}_2} - x_B)(y_{\mathrm{A}_1} - y_B) = 0 \end{aligned} \tag{6.5.16}$$

which are the original conditions and thus zero.

In the bond vector frame choice for the molecule, the conditions were

$$\begin{aligned} C^1 &= z_{\mathrm{A}_2} - z_B = 0, \\ C^2 &= z_{\mathrm{A}_1} - z_B = 0, \\ C^3 &= y_{\mathrm{A}_1} - y_B = 0 \end{aligned} \tag{6.5.17}$$

which after a transformation of (12) and a rotation by $\pi$ are

$$\begin{aligned} C^{1'} &= z_{\mathrm{A}_1} - z_B = 0, \\ C^{2'} &= z_{\mathrm{A}_2} - z_B = 0, \\ C^{3'} &= -y_{\mathrm{A}_2} + y_B. \end{aligned} \tag{6.5.18}$$

The last condition is in general not zero.

### 6.5.4 Eckart frame

Approaching the problem from this point of view provides us with an alternative way to check if the BF-frame undergoes a constant rotation under MS operations. This is especially useful when the BF coordinates as functions of the vibrational coordinates are not available analytically. In particular, we can check that the Eckart frame also satisfies the aforementioned rotation condition.

The conditions for the Eckart frame are defined as [120]

$$C = \sum_{\alpha} m_\alpha a_\alpha \times r_\alpha \tag{6.5.19}$$

where the equilibrium frame is defined to be the principal axis system (PAS). The $a_{\alpha g}$ are the BF Cartesian components of the nuclei at equilibrium and enter the condition as constants. As stated above, the PAS for the ABA molecule in vibrational equilibrium is the bisector. We thus have

$$(m_\mathrm{A}[a_{\mathrm{A}_1} \times r_{\mathrm{A}_1} + a_{\mathrm{A}_2} \times r_{\mathrm{A}_2}] + m_\mathrm{B} a_\mathrm{B} \times r_\mathrm{B}) = 0 \tag{6.5.20}$$

where $a_{\mathrm{A}_1}$, $a_{\mathrm{A}_2}$, and $a_\mathrm{B}$ are of the form

$$a_{\mathrm{A}_1} = \begin{pmatrix} b \\ c \\ 0 \end{pmatrix}, \quad a_{\mathrm{A}_2} = \begin{pmatrix} -b \\ c \\ 0 \end{pmatrix}, \quad a_\mathrm{B} = \begin{pmatrix} 0 \\ d \\ 0 \end{pmatrix}. \tag{6.5.21}$$

In particular, if $N_{(12)} \equiv N = M_y(\pi)$, then $a_{\mathrm{A}_2} = N^T a_{\mathrm{A}_1}$ and $a_\mathrm{B} = N^T a_\mathrm{B}$. Once the operation (12) is applied, we have

$$(m_\mathrm{A}[a_{\mathrm{A}_1} \times r_{\mathrm{A}_2} + a_{\mathrm{A}_2} \times r_{\mathrm{A}_1}] + m_\mathrm{B} a_\mathrm{B} \times r_\mathrm{B}) \tag{6.5.22}$$

where it should be reiterated that the components $r_\alpha$ are in the original BF-frame. If we then rotate the frame the new conditions are

$$\begin{aligned} C' &= (m_\mathrm{A}[a_{\mathrm{A}_1} \times N r_{\mathrm{A}_2} + a_{\mathrm{A}_2} \times N r_{\mathrm{A}_1}] + m_\mathrm{B} a_\mathrm{B} \times N r_\mathrm{B}) \\ &= NN^T(m_\mathrm{A}[a_{\mathrm{A}_1} \times N r_{\mathrm{A}_2} + a_{\mathrm{A}_2} \times N r_{\mathrm{A}_1}] + m_\mathrm{B} a_\mathrm{B} \times N r_\mathrm{B}) \\ &= N(m_\mathrm{A}[N^T a_{\mathrm{A}_1} \times r_{\mathrm{A}_2} + N^T a_{\mathrm{A}_2} \times r_{\mathrm{A}_1}] + m_\mathrm{B} N^T a_\mathrm{B} \times r_\mathrm{B}) \\ &= N(m_\mathrm{A}[a_{\mathrm{A}_2} \times r_{\mathrm{A}_2} + a_{\mathrm{A}_1} \times r_{\mathrm{A}_1}] + m_\mathrm{B} a_\mathrm{B} \times r_\mathrm{B}) = 0 \end{aligned} \tag{6.5.23}$$

where in the second line we applied $I = NN^T$ to the conditions and in the third line used the property that, if $N$ is a rotation, $N(a \times b) = Na \times Nb$. The last line is $N$ applied to the original conditions and therefore zero.

The operation $E^*$ is even easier and can be trivially shown to work as the operation changes the $r_i$ to $-r_i$ while $M_z(\pi)^T a \equiv N^T a = -a$. We have therefore shown that the Eckart frames are applicable for ABA molecules if one desires a symmetry-adapted rotational basis set.

## 6.6 $CH_3Cl$ example

### 6.6.1 Eckart frame

For ABA molecules, it was relatively straightforward to choose alternative (to Eckart) axes with which one could symmetrise the rotational basis functions. For other molecules the choice is not so simple. The example of $CH_3Cl$ (shown in Figure 6.3) is used to demonstrate that the Eckart frame (whose equilibrium orientation is shown in Figure 6.4) can be employed to build a symmetry-adapted rotational basis set. An alternative and geometric frame is then stated which can also be used for rotational symmetrisation.

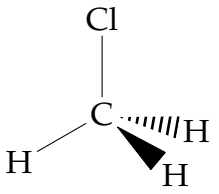



Figure 6.3: The structure of the $CH_3Cl$ molecule.

The Eckart conditions for $CH_3Cl$ are

$$([m_C a_C \times r_C + m_{Cl} a_{Cl} \times r_{Cl}] + m_H[a_{H_1} \times r_{H_1} + a_{H_2} \times r_{H_2} + a_{H_3} \times r_{H_3}]) = 0 \qquad (6.6.1)$$

where the PAS frame for the equilibrium geometry is shown in Figure 6.4. Because of this, both $a_{Cl}$ and $a_C$ only have a $z$ component.

The MS group of $CH_3Cl$ is $\boldsymbol{C}_{3v}$(M) and can be constructed from two generating operations (123) and $(23)^*$ with equivalent rotations of $R_z^{2\pi/3}$ and $R_{\pi/2}^{\pi}$, respectively.

For the operation (123), we can relate the equilibrium coordinates as $a_{H_2} = A^2 a_{H_1}$ and $a_{H_3} = A a_{H_1}$ where $A = M_z(2\pi/3)$ from Eq. (2.3.3). The Eckart conditions can thus be can be expressed as

$$([m_C a_C \times r_C + m_{Cl} a_{Cl} \times r_{Cl}] + m_H[a_{H_1} \times r_{H_1} + A^2 a_{H_1} \times r_{H_2} + A a_{H_1} \times r_{H_3}]) = 0. \quad (6.6.2)$$

Applying the operation (123), the Eckart conditions become

$$([m_C a_C \times r_C + m_{Cl} a_{Cl} \times r_{Cl}] + m_H[a_{H_1} \times r_{H_3} + A^2 a_{H_1} \times r_{H_1} + A a_{H_1} \times r_{H_2}]). \qquad (6.6.3)$$

As usual, we rotate the frame, in this case by $R_z^{2\pi/3}$. This transforms $r_i$ to $A^T r_i$. Using the same procedure as for the ABA case we can rewrite this as

$$A^T([m_C a_C \times r_C + m_{Cl} a_{Cl} \times r_{Cl}] + m_H[A a_{H_1} \times r_{H_3} + a_{H_1} \times r_{H_1} + A^2 a_{H_1} \times r_{H_2}]) = 0 \quad (6.6.4)$$

where we used that $A$ applied to the equilibrium coordinates of Cl or C does not change them. The final result is precisely the same as in the case of ABA.

For the operation $(23)^*$, the rotation matrix is $N = M_y(\pi)$. We can again relate the equilibrium coordinates by $N a_{H_2} = -a_{H_3}$ and vice versa. Also, $N a_{H_1} = -a_{H_1}$ and likewise for C and Cl. The Eckart conditions can thus be written

$$([-m_C N a_C \times r_C - m_{Cl} N a_{Cl} \times r_{Cl}] - m_H[N a_{H_1} \times r_{H_1} - N a_{H_3} \times r_{H_2} - N a_{H_2} \times r_{H_3}]) = 0. \quad (6.6.5)$$

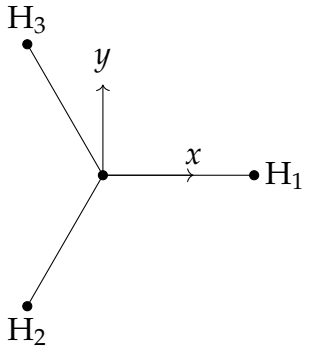



Figure 6.4: The equilibrium Eckart frame for $CH_3Cl$. The $z$-axis is along the C–Cl bond and the $x$-axis in the plane formed by Cl–C–$H_1$.

The result after applying $(23)^*$ is

$$([m_C N a_C \times r_C + m_{Cl} N a_{Cl} \times r_{Cl}] + m_H [N a_{H_1} \times r_{H_1} + N a_{H_3} \times r_{H_3} + N a_{H_2} \times r_{H_2}]). \quad (6.6.6)$$

Rotating the frame by $N$, multiplying the conditions by $NN^T$, and using the same steps as before, we have

$$N([m_C a_C \times r_C + m_{Cl} a_{Cl} \times r_{Cl}] + m_H [a_{H_1} \times r_{H_1} + a_{H_3} \times r_{H_3} + a_{H_3} \times r_{H_3}]) = 0. \quad (6.6.7)$$

The technique used in this proof can be applied to many different molecules where the Eckart frame is used as it relied on relations between the equilibrium Cartesian coordinates of the molecule. In vibrational equilibrium, the Cartesian coordinates of the nuclei are related to each other by constant rotations. These rotations can be found, for example, through graphical means, which is the technique used to find the equivalent rotations explained in Ref. [31]. Alternatively, although the form of the BF coordinates as a function of vibrational coordinates for all geometries is not known, their equilibrium "structure" (à la Eq. (6.5.21)) is. The required rotation for an MS operation can then by found by satisfying $a_{\alpha'} = N\tilde{a}_\alpha = Na_\alpha$, where $a_\alpha$ are the equilibrium BF coordinates. These do not change under the vibrational coordinate transformation because the vibrational coordinates are effectively frozen. Since the matrix $N$ should be the same regardless of the geometry, if it is found in vibrational equilibrium, it is found in general. As demonstrated above, if one uses the PAS system at the vibrational equilibrium to define the reference geometry, then the equivalent rotations are the correct rotations of the BF-frame undergoing MS operations.

### 6.6.2 Geometric $CH_3Cl$ frame

The following introduces a frame definition which can be solved analytically. In this case, the $z$-axis always points parallel the C–Cl bond and the direction of the $x$-axis is defined by $\phi$, given by

$$\phi_1 = \frac{1}{3}(\theta_{21} - \theta_{13}) \quad (6.6.8)$$

where $\theta_{21}$ is the dihedral angle from the plane formed by Cl–C–$H_2$ to the plane Cl–C–$H_1$ and likewise for $\theta_{13}$. This is illustrated in Figure 6.5. Had we decided to start from $H_2$,

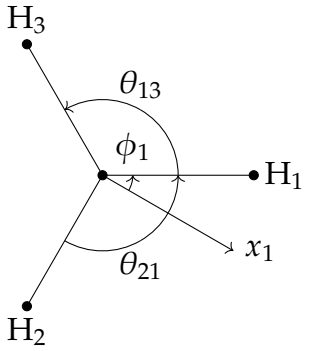



Figure 6.5: The dihedrals that determine the value of $\phi_1 = 1/3(\theta_{21} - \theta_{13})$.

the choice of $x$-axis direction ($x_2$) would be determined by

$$\phi_2 = \frac{1}{3}(\theta_{32} - \theta_{21}) \tag{6.6.9}$$

where Figure 6.6 displays this choice. A feature of these choices is the angle between

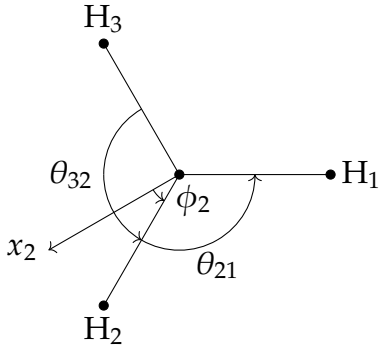



Figure 6.6: The dihedrals that determine the value of $\phi_2 = 1/3(\theta_{32} - \theta_{21})$.

the $x$-axes, as shown in Figure 6.7, can be found as follows:

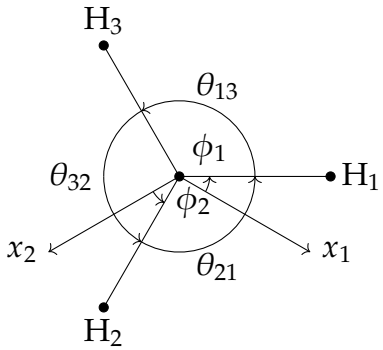



Figure 6.7: All the angles relevant angles to determine the angle between $x_1$ and $x_2$.

$$\begin{aligned}\theta_{21} - \phi_1 + \phi_2 &= \theta_{21} - \frac{1}{3}(\theta_{21} - \theta_{13}) + \frac{1}{3}(\theta_{32} - \theta_{21}) \\ &= \frac{1}{3}(\theta_{21} + \theta_{13} + \theta_{32}) = \frac{2\pi}{3}.\end{aligned} \tag{6.6.10}$$

Thus the angle between them is always $2\pi/3$, irrespective of the coordinates of the atoms. It is this property that allows the frame to be used when symmetrising the rotational basis set.

To demonstrate that, with this frame, the Cartesian coordinates transform correctly under MS operations if the BF-frame rotates by constant angles, let us first parametrise the vibrational coordinates by the Cl–C stretch ($q_1$), the 3 C–H stretches ($q_2$ to $q_4$), the three Cl–C–H bends ($q_5$ to $q_7$), and the symmetrised coordinates defined by

$$\begin{aligned} q_8 &= \frac{1}{\sqrt{6}}(2\theta_{32} - \theta_{21} - \theta_{13}), \\ q_9 &= \frac{1}{\sqrt{2}}(\theta_{21} - \theta_{13}). \end{aligned} \tag{6.6.11}$$

Note that $q_9$ is the angle $3/\sqrt{2}\phi_1$. The BF coordinates in the bond vector frame are given by (with the origin at the C atom)

$$\begin{aligned} r_{\mathrm{C}} &= (0,0,0), \\ r_{\mathrm{Cl}} &= (0,0,q_1), \\ r_{\mathrm{H}_1} &= (q_2 \sin q_5, 0, q_2 \cos q_5)\,, \\ r_{\mathrm{H}_2} &= (q_3 \cos\theta_{21} \sin q_6, -q_3 \sin\theta_{21} \sin q_6, q_3 \cos q_6)\,, \\ r_{\mathrm{H}_3} &= (q_4 \cos\theta_{13} \sin q_7, q_4 \sin\theta_{13} \sin q_7, q_4 \cos q_7)\,. \end{aligned} \tag{6.6.12}$$

Solving Eq. (6.6.11) for the angles $\theta_{21}$ and $\theta_{13}$ by using the substitution $\theta_{32} = 2\pi - \theta_{21} - \theta_{13}$, we obtain

$$\begin{aligned} \theta_{21} &= \frac{4\pi}{6} - \frac{q_8}{\sqrt{6}} + \frac{q_9}{\sqrt{2}}, \\ \theta_{13} &= \frac{4\pi}{6} - \frac{q_8}{\sqrt{6}} - \frac{q_9}{\sqrt{2}} \end{aligned} \tag{6.6.13}$$

which can be substituted into Eq. (6.6.12). Applying the rotation matrix $M_z(\sqrt{2}q_9/3)$ (the matrix $S$ of Eq. (6.7.1)) to the BF coordinates and using the relations $\cos(x + \pi/2) = -\sin x$ and $\sin(x + \pi/2) = \cos x$, the final result is

$$\begin{aligned} r_{\mathrm{C}} &= (0,0,0), \\ r_{\mathrm{Cl}} &= (0,0,q_1), \\ r_{\mathrm{H}_1} &= \left(q_2 \cos\left(\frac{\sqrt{2}q_9}{3}\right) \sin q_5, q_2 \sin\left(\frac{\sqrt{2}q_9}{3}\right) \sin q_5, q_2 \cos q_5\right), \\ r_{\mathrm{H}_2} &= \left(-q_3 \sin\left(\frac{\pi}{6} - \frac{q_8}{\sqrt{6}} + \frac{q_9}{3\sqrt{2}}\right) \sin q_6, -q_3 \cos\left(\frac{\pi}{6} - \frac{q_8}{\sqrt{6}} + \frac{q_9}{3\sqrt{2}}\right) \sin q_6, q_3 \cos q_6\right), \\ r_{\mathrm{H}_3} &= \left(-q_4 \sin\left(\frac{\pi}{6} - \frac{q_8}{\sqrt{6}} - \frac{q_9}{3\sqrt{2}}\right) \sin q_7, q_4 \cos\left(\frac{\pi}{6} - \frac{q_8}{\sqrt{6}} - \frac{q_9}{3\sqrt{2}}\right) \sin q_7, q_4 \cos q_7\right). \end{aligned} \tag{6.6.14}$$

The operation (123) changes the stretches and angles in the usual manner and changes the dihedrals according to

$$\begin{aligned} q_8' &= \frac{1}{\sqrt{6}}(2\theta_{21} - \theta_{13} - \theta_{32}), \\ q_9' &= \frac{1}{\sqrt{2}}(\theta_{13} - \theta_{32}) \end{aligned} \tag{6.6.15}$$

or

$$\begin{pmatrix} q_8' \\ q_9' \end{pmatrix} = \begin{pmatrix} -1/2 & \sqrt{3}/2 \\ -\sqrt{3}/2 & -1/2 \end{pmatrix} \begin{pmatrix} q_8 \\ q_9 \end{pmatrix}. \tag{6.6.16}$$

The transformed coordinates are then

$$\begin{aligned} \tilde{r}_C &= (0,0,0), \\ \tilde{r}_{Cl} &= (0,0,q_1), \\ \tilde{r}_{H_1} &= \left( q_4 \cos\left( -\frac{q_8}{\sqrt{6}} - \frac{q_9}{3\sqrt{2}} \right) \sin q_7, q_4 \sin\left( -\frac{q_8}{\sqrt{6}} - \frac{q_9}{3\sqrt{2}} \right) \sin q_7, q_4 \cos q_7 \right), \\ \tilde{r}_{H_2} &= \left( q_2 \sin\left( -\frac{\pi}{6} + \frac{\sqrt{2}q_9}{3} \right) \sin q_5, -q_2 \cos\left( -\frac{\pi}{6} + \frac{\sqrt{2}q_9}{3} \right) \sin q_5, q_2 \cos q_5 \right), \\ \tilde{r}_{H_3} &= \left( q_3 \sin\left( -\frac{\pi}{6} - \frac{q_8}{\sqrt{6}} + \frac{q_9}{3\sqrt{2}} \right) \sin q_6, q_3 \cos\left( -\frac{\pi}{6} - \frac{q_8}{\sqrt{6}} + \frac{q_9}{3\sqrt{2}} \right) \sin q_6, q_3 \cos q_6 \right). \end{aligned} \tag{6.6.17}$$

These are in the new BF-frame. If we now consider them in the old BF-frame via the rotation matrix

$$N_{(123)} = \begin{pmatrix} -\sin(\pi/6) & -\cos(\pi/6) & 0 \\ \cos(\pi/6) & -\sin(\pi/6) & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} -\cos(\pi/3) & -\sin(\pi/3) & 0 \\ \sin(\pi/3) & -\cos(\pi/3) & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{6.6.18}$$

and apply the first form to $\tilde{r}_{H_1}$ and $\tilde{r}_{H_2}$ and the second form to $\tilde{r}_{H_3}$, we obtain the original coordinates but permuted by (123).

For the operation $(23)^*$, the stretches and angles transform in an obvious way while the dihedrals transform as

$$\begin{pmatrix} q_8' \\ q_9' \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} q_8 \\ q_9 \end{pmatrix}. \tag{6.6.19}$$

Performing this transformation and also applying the matrix $M_y(\pi)$ to the Cartesian

coordinates, we obtain

$$\begin{aligned}
M_y(\pi)\tilde{r}_{\mathrm{C}} &= (0,0,0),\\
M_y(\pi)\tilde{r}_{\mathrm{Cl}} &= (0,0,-q_1),\\
M_y(\pi)\tilde{r}_{\mathrm{H}_1} &= \left(-q_2\cos\left(\frac{\sqrt{2}q_9}{3}\right)\sin q_5, -q_2\sin\left(\frac{\sqrt{2}q_9}{3}\right)\sin q_5, -q_2\cos q_5\right),\\
M_y(\pi)\tilde{r}_{\mathrm{H}_2} &= \left(q_3\sin\left(\frac{\pi}{6}-\frac{q_8}{\sqrt{6}}-\frac{q_9}{3\sqrt{2}}\right)\sin q_6, -q_3\cos\left(\frac{\pi}{6}-\frac{q_8}{\sqrt{6}}-\frac{q_9}{3\sqrt{2}}\right)\sin q_6, -q_3\cos q_6\right),\\
M_y(\pi)\tilde{r}_{\mathrm{H}_3} &= \left(q_4\sin\left(\frac{\pi}{6}-\frac{q_8}{\sqrt{6}}+\frac{q_9}{3\sqrt{2}}\right)\sin q_7, q_4\cos\left(\frac{\pi}{6}-\frac{q_8}{\sqrt{6}}+\frac{q_9}{3\sqrt{2}}\right)\sin q_7, -q_4\cos q_7\right)
\end{aligned} \tag{6.6.20}$$

which are the original Cartesian coordinates with 2 and 3 permuted and then inverted.

## 6.7 Finding the geometric BF-frames

This section explains how to find the angle $\phi_1$ in the case of $CH_3Cl$ and more generally introduces the methodology for any molecule.

For a general symmetry operation, if $R_\alpha \to R_{\alpha'}$, we required that $r_{\alpha'} = N\tilde{r}_\alpha$ or $N^T r_{\alpha'} = \tilde{r}_\alpha$. We then stipulate that $N$ is a constant matrix, for example $N_{(123)} = M_z(2\pi/3)$ for $CH_3Cl$. The Cartesian coordinates $r_\alpha$ are in the old BF-frame. In the new BF-frame, they will be of the form $S(q_1,\ldots,q_n)r_\alpha$ for some rotation matrix $S$ which is a function of the vibrational coordinates. We thus want

$$N^T S r_{\alpha'} = \tilde{S}\tilde{r}_\alpha \tag{6.7.1}$$

for every operation. In the case of $CH_3Cl$ we see that the $z$-axis along the Cl–C bond should be maintained. Only a rotation about the $z$-axis is required. Since motion of the atoms in this plane is parametrised by the dihedral angles, we will assume the simplest rotation of the form $M_z(aq_8 + bq_9)$ where $a$ and $b$ are constants. The position of $H_1$ in this, yet, undetermined frame is given by

$$r_{\mathrm{H}_1} = (q_2\cos(aq_8+bq_9)\sin q_5, q_2\sin(aq_8+bq_9)\sin q_5, q_2\cos q_5) \tag{6.7.2}$$

while $N^T r_{\mathrm{H}_3}$ is given by

$$\begin{aligned}
N^T r_{\mathrm{H}_3} = \Bigg(&q_4\cos\left(-\frac{q_8}{\sqrt{6}}-\frac{q_9}{\sqrt{2}}+aq_8+bq_9\right)\sin q_7,\\
&q_4\sin\left(-\frac{q_8}{\sqrt{6}}-\frac{q_9}{\sqrt{2}}+aq_8+bq_9\right), q_4\cos q_7\Bigg).
\end{aligned} \tag{6.7.3}$$

Thus we want to equate the transformed $aq_8 + bq_9$ to $-\frac{q_8}{\sqrt{6}} - \frac{q_9}{\sqrt{2}} + aq_8 + bq_9$, or

$$-a\frac{q_8}{2} + a\frac{\sqrt{3}q_9}{2} - b\frac{\sqrt{3}}{2}q_8 - b\frac{q_9}{2} = -\frac{q_8}{\sqrt{6}} - \frac{q_9}{\sqrt{2}} + aq_8 + bq_9 \tag{6.7.4}$$

from which we obtain

$$b = \frac{\sqrt{2}}{3} + \frac{(-3q_8 + \sqrt{3}q_9)a}{\sqrt{3}q_8 + 3q_9}. \tag{6.7.5}$$

For the operation $(23)^*$, $r_{H_2}$ in this frame is given by

$$r_{H_2} = \left(q_3 \cos\left(\frac{4\pi}{6} - \frac{q_8}{\sqrt{6}} + \frac{q_9}{\sqrt{2}} - aq_8 - bq_9\right) \sin q_6, \right.$$
$$\left. -q_3 \sin\left(\frac{4\pi}{6} - \frac{q_8}{\sqrt{6}} + \frac{q_9}{\sqrt{2}} - aq_8 - bq_9\right) \sin q_6, q_3 \cos q_6\right) \tag{6.7.6}$$

and $-Nr_{H_3} = -M_y(\pi)r_{H_3}$ is

$$r_{H_3} = \left(q_3 \cos\left(\frac{4\pi}{6} - \frac{q_8}{\sqrt{6}} - \frac{q_9}{\sqrt{2}} + aq_8 + bq_9\right) \sin q_6, \right.$$
$$\left. -q_3 \sin\left(\frac{4\pi}{6} - \frac{q_8}{\sqrt{6}} + \frac{q_9}{\sqrt{2}} + aq_8 + bq_9\right) \sin q_6, q_3 \cos q_6\right). \tag{6.7.7}$$

Equating the transformed Eq. (6.7.6) to Eq. (6.7.7), we have

$$\frac{4\pi}{6} - \frac{q_8}{\sqrt{6}} - \frac{q_9}{\sqrt{2}} - aq_8 + bq_9 = \frac{4\pi}{6} - \frac{q_8}{\sqrt{6}} - \frac{q_9}{\sqrt{2}} + aq_8 + bq_9. \tag{6.7.8}$$

From these we obtain $a = 0$ and $b = \sqrt{2}/3$ as stated previously.

## 6.8 Further applications

In the above sections, the focus was on valence coordinates as these are the simplest to define and use to determine the appropriate frames. However, this does not preclude the use of other coordinates. For example, the Jacobi [121] and Radau [122] coordinates offer alternative parametrisations. These have advantages due to the simplification of the Hamiltonian.

It is feasible that a similar procedure could be applied for these coordinates for rotational basis symmetrisation. Indeed, the $s$-vector approach stipulates that the vibrational coordinates are independent of the translation and rotation of the molecule. It does not specify the form of the coordinates. For $CH_3F$, for example, the kinetic energy operator was derived in [123] where the canonical point was the centre of mass of $CH_3$ group and the vectors were the distances from that point to the hydrogen and fluorine nuclei. The frame chosen was the bond vector using vectors from the canonical point to two of the hydrogens. Using vibrational coordinates analogous to the valence ones for $CH_3Cl$ of Section 6.6.2, it is clear that the structure of the Cartesian coordinates as a function of the vibrational is identical, and hence the same process should work.

The size of the Coriolis coupling was not considered, but it is arguably the most important factor when comparing these frames. Previously research into this has been undertaken; see, for example, Ref. [124] who compared the Radau bisector, the valence bisector, and the Eckart frame for symmetric triatomics and found that the Radau

bisector is superior in reducing this coupling and should lead to better ro-vibrational separation. Comparing the BF frames of this chapter is a possible avenue for further research.

Moving beyond the molecules chosen, the question of the possibility of finding non-Eckart frames where rotational symmetrisation is possible still remains. For certain molecules, such as ammonia, a trisector can define the $z$-axis. Then, once again, two symmetrised coordinates based on the dihedrals could be used to rotate about this $z$-axis as the structure is the same as the $CH_3Cl$ molecule. For other molecules, such as $CH_4$, it is remarkably more challenging. For Eq. (6.7.1) one can replace $r_{\alpha'}$ with $B\tilde{r}_\alpha$, where $B$ is the rotation matrix relating the old and rotated bond vector frame (i.e. it is the matrix $N$ of Eq. (6.2.7) for the bond vector frame rather than the desired frame). Then, the equation that must be solved is

$$N\tilde{S} = SB \tag{6.8.1}$$

for each generator. For $CH_4$, the matrix $S$ would be parametrised by five symmetrised angles and the coefficient values could only be found numerically, if they could be found at all. In general, finding the frame for a given molecular structure and coordinate choice would have to be done on a case-by-case basis.

## 6.9 Chapter summary

The main result in this chapter was the determination of geometric BF-frames for a certain class of molecules – with the $CH_3Cl$ molecule used as a defining example – where rotational symmetrisation is possible. The Eckart frame was also shown to be suitable for this molecule, with the form of the proofs extendable to other molecules as it relied on the correspondence between the MS group of the molecule and the point group of the molecule when in vibrational equilibrium.

Furthermore, one approach to find such frames is described for $CH_3Cl$, and a method to check frames is also given, provided one has the BF coordinates as a function of the vibrational coordinates.

# Appendix 6.A Derivation of the transformations

### 6.A.1 Vibrational $s$-vector transformation

We want to determine transformation of the vibrational $s$-vectors due to an MS transformation. As before, we first differentiate, then evaluate the $s$-vector at the transformed coordinates. Suppose $\boldsymbol{R}_\alpha$ is transformed to $\boldsymbol{R}_{\alpha'}$. Then, if $q_k$ is transformed to $q'_k$, we have

$$\begin{aligned} \partial_{\alpha' F}(q_k \circ P)(R_1, \ldots, R_n) &= \sum_\beta \partial_{\alpha' F} P^{\beta G}(R_1, \ldots, R_n) \partial_{\beta G} q_k(P(R_1, \ldots, R_n)) \\ \partial_{\alpha' F} q_{k'}(R_1, \ldots, R_n) &= \partial_{\alpha F} q_k(P(R_1, \ldots, R_n)) \end{aligned} \tag{6.A.1}$$

or

$$\boldsymbol{s}'_{k,\alpha} = \boldsymbol{s}_{k',\alpha'}. \tag{6.A.2}$$

In short, the first index transforms like the vibrational coordinates and the second like the Cartesian coordinates.

If $q_k$ and $q_{k'}$ transformed as in Eq. (6.B.6), then the transformation of $\boldsymbol{s}_{k,\alpha}$ would be

$$\boldsymbol{s}'_{k,\alpha} = a\boldsymbol{s}_{k,\alpha'} + b\boldsymbol{s}_{k',\alpha'} \tag{6.A.3}$$

and similarly for the other vibrational $s$-vectors.

From a computational standpoint, the BF $s$-vector components are typically used. We have

$$\begin{aligned} \boldsymbol{s}'_{k,\alpha} &= \partial_{\alpha I} q_k(P(R_1, \ldots, R_n))\hat{\boldsymbol{e}}_I \\ &= \partial_{\alpha I} q_k(R_{p_1}, \ldots, R_{p_n}))\hat{\boldsymbol{e}}_I \\ &= \partial_{\alpha g} q_k(T(R_{p_1}), \ldots, T(R_{p_n}))\hat{\boldsymbol{e}}_g \\ &= \partial_{\alpha g} q_k(\tilde{r}_1, \ldots, \tilde{r}_n)\hat{\boldsymbol{e}}_g \end{aligned} \tag{6.A.4}$$

as $T(R_{p_\alpha})$ are the coordinates of atom $\alpha$ after the MS operation is performed and in the new BF-frame. As we've seen, this is given by $\tilde{r}_\alpha$, so that to transform the BF components $s_{k,\alpha i}$ we simply perform the vibrational coordinate transformation (which we shall denote as $\tilde{s}_{k,\alpha g}$). This will give us the transformed $s$-vector components in the new BF-frame.

## 6.A.2 Rotational $t$-vector transformation

The second set of vectors to consider are the rotational $t$-vectors as

$$\begin{aligned} \boldsymbol{t}_{g,\alpha} &= \hat{\boldsymbol{e}}_g \times \boldsymbol{R}_\alpha \\ &= M^T_{gG}\hat{\boldsymbol{e}}_G \times \boldsymbol{R}_\alpha \\ &= M_{Gg}\varepsilon_{IGF}R_{\alpha F}\hat{\boldsymbol{e}}_I \end{aligned} \tag{6.A.5}$$

where $M^T$ is the inverse Euler matrix. The SF components are therefore given by $t_{g,\alpha I} = M_{Gg}\varepsilon_{IGF}R_{\alpha F}$ and the BF components are given by $t_{g,\alpha h} = \varepsilon_{hgf}r_{\alpha f}$.

An MS transformation changes $R_\alpha$ to $R_{\alpha'}$ and $M$ to the new matrix relating components in the new BF-frame to the SF-frame, denoted by $\tilde{M}$. We can thus write

$$\begin{aligned} \boldsymbol{t}_{g,\alpha}{}' &= \tilde{M}_{Gg}\varepsilon_{IGF}R_{\alpha' F}\hat{\boldsymbol{e}}_I \\ &= \tilde{M}_{Gg}M_{Gl}M_{Hl}\varepsilon_{IHF}R_{\alpha' F}\hat{\boldsymbol{e}}_I \\ &= N_{lg}\boldsymbol{t}_{l,\alpha'} \end{aligned} \tag{6.A.6}$$

where $N$ is the rotation matrix relating the coordinates in the new BF-frame to the old BF-frame and is Eq. (6.2.7). We also have

$$\boldsymbol{t}_{g,\alpha}{}' = \hat{\boldsymbol{e}}'_g \times \boldsymbol{R}_{\alpha'} \tag{6.A.7}$$

and thus has components in the new BF-frame given by $\tilde{t}_{g,\alpha h} = \varepsilon_{hgf}\tilde{r}_{\alpha f}$ which can be obtained by performing the vibrational coordinate transformation on the old $t_{g,\alpha h}$ components.

### 6.A.3 Vibrational $t$-vector transformation

The third set of vectors are the vibrational $t$-vectors, defined by

$$\begin{aligned} \boldsymbol{t}_{k,\alpha} &= \partial_k \boldsymbol{R}_\alpha \\ &= \partial_k R_{\alpha F} \hat{\boldsymbol{e}}_F \\ &= \partial_k r_{\alpha g} \hat{\boldsymbol{e}}_g \end{aligned} \tag{6.A.8}$$

where now the position vectors are wholly defined by the vibrational and rotational coordinates. The SF components thus transform as

$$\partial_k R_{\alpha F}(P(q_1, \ldots, q_n, \phi, \theta, \chi)). \tag{6.A.9}$$

We use the standard procedure to write this in terms of the original components

$$\begin{aligned} t_{k',\alpha' F} &= \partial_{k'}(R_{\alpha F} \circ P)(q_1, \ldots, q_n, \phi, \theta, \chi) \\ &= \partial_{k'} P^m(q_1, \ldots) \partial_m R_{\alpha F}(P(q_1, \ldots)) \\ &= \partial_k R_{\alpha F}(P(q_1, \ldots)) + \sum_{g=n+1}^{g=n+3} \partial_{k'} P^g(q_1, \ldots) \partial_g R_{\alpha F}(P(q_1, \ldots)) \\ &= t_{k,\alpha F}{}' + \sum_{g=n+1}^{g=n+3} \partial_{k'} P^g(q_1, \ldots) \partial_g R_{\alpha F}(P(q_1, \ldots)). \end{aligned} \tag{6.A.10}$$

To simplify the last term, we turn to the expressions for the change in the Euler angles when the BF-frame is rotated (see Table 6.3). We can immediately see that, in general, only the transformation of $\chi$ might contain a vibrational coordinate dependence, due to the angles $\beta$ and $\alpha$. The combined effect will be in the form

$$\begin{aligned} \chi' &= \chi + \gamma(q_1, \ldots, q_n) \quad \text{or} \\ \chi' &= -\chi + \gamma(q_1, \ldots, q_n) \end{aligned} \tag{6.A.11}$$

for some function $\gamma$. Therefore, $g = n + 3$ in the last line of Eq. (6.A.10). We thus have

$$\begin{aligned} t_{k',\alpha' F} &= t_{k,\alpha F}{}' + \partial_{k'} \gamma(q_1, \ldots, q_n) \partial_{n+3} R_{\alpha F}(P(q_1, \ldots)) \\ &= t_{k,\alpha F}{}' + \partial_{k'} \gamma(q_1, \ldots)(\partial_{n+3} M_{Fg}) r_{\alpha g}(P(q_1, \ldots)) \\ &= t_{k,\alpha F}{}' + \partial_{k'} \gamma(q_1, \ldots)(\partial_{n+3} M_{Fg}) M_{Hg} M_{Hh} r_{\alpha h}(P(q_1, \ldots)) \\ &= t_{k,\alpha F}{}' + \Omega_{k',FH} R_{\alpha H}(P(q_1, \ldots)) \end{aligned} \tag{6.A.12}$$

where $\Omega_{k'}$ is the antisymmetric matrix

$$\Omega_{k'} = \partial_{k'} \gamma(q_1, \ldots) \begin{pmatrix} 0 & -\cos\theta & \sin\theta \sin\phi \\ \cos\theta & 0 & -\cos\phi \sin\theta \\ -\sin\theta \sin\phi & \cos\phi \sin\theta & 0 \end{pmatrix}. \tag{6.A.13}$$

We also call $\partial_{k'} \gamma(q_1, \ldots)$ as $\dot{\gamma}_{k'}$ for brevity.

One can rewrite $\Omega_{k',FH} R_{\alpha H}$ as $\varepsilon_{FGH} \omega_{k',G} R_{\alpha H}$ where $\omega_{k'}$ is defined as

$$\omega_{k'} = \dot{\gamma}_{k'}(\cos\phi \sin\theta, \sin\theta \sin\phi, \cos\theta) \tag{6.A.14}$$

Actually, $\omega_{k',G} = \dot{\gamma}_{k'} M_{Gz}$ and $\tilde{\omega}_{k',G} = \mp\omega_{k',G}$ with the sign change corresponding to the left and right columns of Table 6.3, respectively. Assuming a change of the form of the left column, we have

$$\begin{aligned} t_{k',\alpha'F} &= t_{k,\alpha F}{}' + \omega_{k',G}\varepsilon_{FGH}R_{\alpha H}(P(q_1,\ldots)) \\ &= t_{k,\alpha F}{}' + \tilde{\omega}_{k',G}\varepsilon_{FGH}R_{\alpha' H} \\ &= t_{k,\alpha F}{}' + \tilde{\omega}_{k',I}M_{Ig}M_{Gg}\varepsilon_{FGH}R_{\alpha' H} \\ &= t_{k,\alpha F}{}' + \tilde{\omega}_{k',I}M_{Ig}t_{g,\alpha' F} \\ &= t_{k,\alpha F}{}' - \dot{\gamma}_{k'}\delta_{gz}t_{g,\alpha' F} \\ &= t_{k,\alpha F}{}' - \dot{\gamma}_{k'}t_{z,\alpha' F} \end{aligned} \tag{6.A.15}$$

where the penultimate line is due to the fact that $\tilde{\omega}_{k',I}M_{Ig} = -\dot{\gamma}_{k'}M_{Iz}M_{Ig} = -\dot{\gamma}_{k'}\delta_{gz}$. In vector form, then, the final transformation of the vibrational $t$-vectors is given by

$$\boldsymbol{t}_{k,\alpha}{}' = \boldsymbol{t}_{k',\alpha'} + \dot{\gamma}_{k'}\boldsymbol{t}_{z,\alpha'}. \tag{6.A.16}$$

As before, since $\boldsymbol{t}_{k,\alpha} = \partial_k r_{\alpha g}\hat{\boldsymbol{e}}_g$, in the new BF-frame the components of the transformed $\boldsymbol{t}_{k,\alpha}$ can be obtained by performing the vibrational coordinate transformation on the old BF components of $\boldsymbol{t}_{k,\alpha}$.

### 6.A.4 Rotational $s$-vector transformation

Collecting the other transformations, we thus have so far

$$\begin{aligned} \boldsymbol{s}_{k,\alpha}{}' &= \boldsymbol{s}_{k',\alpha'} \\ \boldsymbol{t}_{g,\alpha}{}' &= N_{lg}\boldsymbol{t}_{l,\alpha'} \\ \boldsymbol{t}_{k,\alpha}{}' &= \boldsymbol{t}_{k',\alpha'} + \dot{\gamma}_{k'}\boldsymbol{t}_{z,\alpha'}. \end{aligned} \tag{6.A.17}$$

From these we can find the transformation property of the rotational $s$-vectors. They transform as

$$\boldsymbol{s}_{g,\alpha}{}' = N_{lg}\boldsymbol{s}_{l,\alpha'} - \dot{\gamma}_k \boldsymbol{s}_{k,\alpha'}N_{zg}. \tag{6.A.18}$$

One can show that the transformed $s$- and $t$-vectors are then inverses. The components of the transformed vector in the new BF-frame are again the components in the old BF-frame with the vibrational coordinate transformation applied, which are denoted as $\tilde{s}_{g,\alpha h}$.

### 6.A.5 Momentum and angular momentum operator transformation

First, the angular momentum operator is defined as [65]

$$\begin{pmatrix} \hat{J}_x \\ \hat{J}_y \\ \hat{J}_z \end{pmatrix} = -i\hbar \begin{pmatrix} \sin\chi & -\dfrac{\cos\chi}{\sin\theta} & \cot\theta\cos\chi \\ \cos\chi & \dfrac{\sin\chi}{\sin\theta} & -\cot\theta\sin\chi \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{\partial}{\partial\theta} \\ \frac{\partial}{\partial\phi} \\ \frac{\partial}{\partial\chi} \end{pmatrix}. \tag{6.A.19}$$

If we apply the symmetry operation $P$ on $(\hat{J}_x\psi, \hat{J}_y\psi, \hat{J}_z\psi)$, we have

$$
\begin{aligned}
&\begin{pmatrix} \hat{J}_x\psi(P(q_1,\ldots)) \\ \hat{J}_y\psi(P(q_1,\ldots)) \\ \hat{J}_z\psi(P(q_1,\ldots)) \end{pmatrix} \\
&= -i\hbar \begin{pmatrix} \sin(-\chi+\gamma) & -\dfrac{\cos(-\chi+\gamma)}{\sin(\pi-\theta)} & \cot(\pi-\theta)\cos(-\chi+\gamma) \\ \cos(-\chi+\gamma) & \dfrac{\sin(-\chi+\gamma)}{\sin(\pi-\theta)} & -\cot(\pi-\theta)\sin(-\chi+\gamma) \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \partial_{n+1}\psi(P(q_1,\ldots)) \\ \partial_{n+2}\psi(P(q_1,\ldots)) \\ \partial_{n+3}\psi(P(q_1,\ldots)) \end{pmatrix} \\
&= -i\hbar \begin{pmatrix} \sin(-\chi+\gamma) & -\dfrac{\cos(-\chi+\gamma)}{\sin\theta} & -\cot\theta\cos(-\chi+\gamma) \\ \cos(-\chi+\gamma) & \dfrac{\sin(-\chi+\gamma)}{\sin\theta} & \cot\theta\sin(-\chi+\gamma) \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -\partial_{n+1}(\psi\circ P)(q_1,\ldots) \\ \partial_{n+2}(\psi\circ P)(q_1,\ldots) \\ -\partial_{n+3}(\psi\circ P)(P(q_1,\ldots) \end{pmatrix} \\
&= -i\hbar \begin{pmatrix} \cos\gamma & -\sin\gamma & 0 \\ -\sin\gamma & -\cos\gamma & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} \sin\chi & -\dfrac{\cos\chi}{\sin\theta} & \cot\theta\cos\chi \\ \cos\chi & \dfrac{\sin\chi}{\sin\theta} & -\cot\theta\sin\chi \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \partial_{n+1}(\psi\circ P)(q_1,\ldots) \\ \partial_{n+2}(\psi\circ P)(q_1,\ldots) \\ \partial_{n+3}(\psi\circ P)(P(q_1,\ldots) \end{pmatrix}
\end{aligned}
\tag{6.A.20}
$$

where the left hand side is the $N$ matrix of Eq. (6.A.6). The operation $P$ was assumed to cause a rotation of the form of the left column of Table 6.3. In the second line, we used the chain rule as

$$
\partial_{n+1}(\psi\circ P)(q_1,\ldots) = \partial_{n+1}P^m(q_1,\ldots)\partial_m\psi(P(q_1,\ldots)) = -\partial_{n+1}\psi(P(q_1,\ldots)). \tag{6.A.21}
$$

Thus, $\hat{J}'_g = N_{lg}\hat{J}_l$.

For the vibrational momentum operators, we have, assuming $q_k$ transforms to $q_{k'}$,

$$
\begin{aligned}
\partial_{k'}(\psi\circ P)(q_1,\ldots) &= \partial_{k'}P^m(q_1,\ldots)\partial_m\psi(P(q_1,\ldots)) \\
&= \partial_k\psi(P(q_1,\ldots)) + \dot{\gamma}_{k'}\partial_{n+3}\psi(P(q_1,\ldots)).
\end{aligned}
\tag{6.A.22}
$$

Using $(\partial_{n+3}\psi)\circ P = -\partial_{n+3}(\psi\circ P)$ we thus have

$$
\partial_k\psi(P(q_1,\ldots)) = (\partial_{k'} + \dot{\gamma}_{k'}\partial_{n+3})(\psi\circ P)(q_1,\ldots). \tag{6.A.23}
$$

If we apply the right hand side operator on the $G$-matrix elements, only the first term is non-zero. Thus, we can use the right hand side operator in general. The momentum operators are defined as $\hat{\pi}_k = -i\hbar\,\partial/\partial q_k$ and transform as

$$
\hat{\pi}'_k = \hat{\pi}_{k'} + \dot{\gamma}_{k'}\hat{J}_z. \tag{6.A.24}
$$

# Appendix 6.B KEO invariance

If $q_i{}' = q_j$ and $q_{i'}{}' = q_{j'}$ then the KEO (without the pseudo-potential) transforms as

$$\begin{aligned}
\hat{\pi}'_i G_{ii'}{}' \hat{\pi}'_{i'} &= (\hat{\pi}_j + \dot{\gamma}_j J_z) G_{jj'} (\hat{\pi}_{j'} + \dot{\gamma}_{j'} J_z) \\
&= \hat{\pi}_j G_{jj'} \hat{\pi}_{j'} + \underbrace{\hat{\pi}_j \dot{\gamma}_{j'} G_{jj'} \hat{J}_z}_{①} + \underbrace{\dot{\gamma}_j \hat{J}_z G_{jj'} \hat{\pi}_{j'}}_{②} + \underbrace{\dot{\gamma}_j \dot{\gamma}_{j'} \hat{J}_z G_{jj'} \hat{J}_z}_{③}, \\
\hat{\pi}'_i G_{ig}{}' \hat{J}'_g &= (\hat{\pi}_j + \dot{\gamma}_j J_z)(N_{hg} G_{jh} - \dot{\gamma}_k G_{jk} N_{zg}) N_{lg} \hat{J}_l \\
&= \hat{\pi}_j G_{jl} \hat{J}_l + \underbrace{\dot{\gamma}_j \hat{J}_z G_{jl} \hat{J}_l}_{④} - \underbrace{\hat{\pi}_j \dot{\gamma}_k G_{jk} \hat{J}_z}_{①} - \underbrace{\dot{\gamma}_k \dot{\gamma}_j \hat{J}_z G_{jk} \hat{J}_z}_{③}, \\
\hat{J}'_g G_{gi}{}' \hat{\pi}'_i &= N_{lg} \hat{J}_l (N_{hg} G_{hj} - \dot{\gamma}_k G_{kj} N_{zg})(\hat{\pi}_j + \dot{\gamma}_j J_z) \\
&= \hat{J}_l G_{lj} \hat{\pi}_j + \underbrace{\dot{\gamma}_j \hat{J}_l G_{lj} \hat{J}_z}_{⑤} - \underbrace{\dot{\gamma}_k \hat{J}_z G_{kj} \hat{\pi}_j}_{②} - \underbrace{\dot{\gamma}_k \dot{\gamma}_j \hat{J}_z G_{kj} \hat{J}_z}_{⑥}, \\
\hat{J}'_g G_{gg'}{}' \hat{J}'_{g'} &= N_{lg} \hat{J}_l [N_{hg} N_{h'g'} G_{hh'} + \dot{\gamma}_k \dot{\gamma}_{k'} G_{kk'} N_{zg'} N_{zg} - \dot{\gamma}_k G_{kh} (N_{hg'} N_{zg} + N_{hg} N_{zg'})] N_{l'g'} \hat{J}_{l'} \\
&= \hat{J}_h G_{hh'} \hat{J}_{h'} + \underbrace{\dot{\gamma}_k \dot{\gamma}_{k'} \hat{J}_z G_{kk'} \hat{J}_z}_{⑥} - \underbrace{\dot{\gamma}_k \hat{J}_z G_{kh} \hat{J}_h}_{④} - \underbrace{\dot{\gamma}_k \hat{J}_h G_{hk} \hat{J}_z}_{⑤}.
\end{aligned} \tag{6.B.1}$$

Terms under-braced with the same number cancel. What remains when all is summed together is the original Hamiltonian, showing invariance.

This derivation assumed that the change in Euler angles was of the form of the left column of Table 6.3. If the transformation was as of the right column, then the rotational $s$-vectors change as

$$\boldsymbol{s}_{g,\alpha}{}' = N_{lg} \boldsymbol{s}_{l,\alpha'} + \dot{\gamma}_k \boldsymbol{s}_{k,\alpha'} N_{zg} \tag{6.B.2}$$

and the vibrational momentum operators change as

$$\hat{\pi}'_k = \hat{\pi}_{k'} - \dot{\gamma}_{k'} \hat{J}_z \tag{6.B.3}$$

so only the sign in front of $\dot{\gamma}_k$ changes and can be absorbed into its definition, meaning invariance is maintained. Likewise, if, during an MS operation, $q'_k = -q_{k'}$, then the rotational $s$-vector change is the same, the vibrational $s$-vectors change as

$$\boldsymbol{s}_{k,\alpha}{}' = -\boldsymbol{s}_{k',\alpha'} \tag{6.B.4}$$

and the vibrational momentum operators change as

$$\hat{\pi}'_k = -(\hat{\pi}_{k'} + \dot{\gamma}_{k'} \hat{J}_z) \tag{6.B.5}$$

so that, in any Hamiltonian term $\hat{\pi}_k G_{ka} \hat{\pi}_a$, the two minus signs cancel out and again the invariance is maintained. Finally, if the transformation of $q_k$ and $q_{k'}$ is of the form

$$\begin{pmatrix} q_k{}' \\ q_{k'}{}' \end{pmatrix} = \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \begin{pmatrix} q_k \\ q_{k'} \end{pmatrix}, \tag{6.B.6}$$

where $a^2 + b^2 = 1$, then the vibrational $s$-vectors transform as

$$\boldsymbol{s}_{k,\alpha}{}' = a\boldsymbol{s}_{k,\alpha'} + b\boldsymbol{s}_{k',\alpha'}, \tag{6.B.7}$$

and likewise for $\boldsymbol{s}_{k',\alpha}$, the rotational operators and rotational $s$-vectors change in the same way as previously, and the vibrational operators transform as

$$\hat{\pi}'_k = a(\hat{\pi}_k + \dot{\gamma}_k \hat{J}_z) + b(\hat{\pi}_{k'} + \dot{\gamma}_{k'} \hat{J}_z) \tag{6.B.8}$$

and similarly for the other operator. To show invariance, we express the kinetic energy operator as

$$\sum_{cd} \sum_{\alpha} \frac{1}{m_\alpha} \hat{\pi}_c \boldsymbol{s}_{c,\alpha} \cdot \boldsymbol{s}_{d,\alpha} \hat{\pi}_d. \tag{6.B.9}$$

Let us show only the relevant terms in the $c$ summation:

$$\sum_{d} \sum_{\alpha} \frac{1}{m_\alpha} (\hat{\pi}_k \boldsymbol{s}_{k,\alpha} + \hat{\pi}_{k'} \boldsymbol{s}_{k',\alpha}) \cdot \boldsymbol{s}_{d,\alpha} \hat{\pi}_d \tag{6.B.10}$$

Then, when the kinetic energy is transformed and the relevant terms simplified, we have

$$\begin{aligned} &\sum_{d} \sum_{\alpha} \frac{1}{m_\alpha} (\hat{\pi}'_k \boldsymbol{s}_{k,\alpha}{}' + \hat{\pi}'_{k'} \boldsymbol{s}_{k',\alpha}{}') \cdot \boldsymbol{s}_{d,\alpha}{}' \hat{\pi}'_d \\ &= \sum_{d} \sum_{\alpha} \frac{1}{m_\alpha} ((\hat{\pi}_k + \dot{\gamma}_k \hat{J}_z) \boldsymbol{s}_{k,\alpha} + (\hat{\pi}_{k'} + \dot{\gamma}_{k'} \hat{J}_z) \boldsymbol{s}_{k',\alpha}) \cdot \boldsymbol{s}_{d,\alpha}{}' \hat{\pi}'_d \end{aligned} \tag{6.B.11}$$

so the expression in brackets is the same as that of the prior transformed $s$-vectors and operations and thus the result is an invariant Hamiltonian. More complicated examples work the same way. The proof the the invariance of the pseudo-potential is given in Ref. [4].

# Chapter 7

# $G_{36}$(EM) MS Group for $C_2H_6$

## 7.1 Introduction

The final core chapter of this thesis combines the knowledge of previous chapters, namely group theory, analytic KEOs, and geometric frames, to the study of the ethane molecule ($H_3{}^{12}C{}^{12}CH_3$), shown in Figure 7.1. The bulk of it is based on the publication Ref. [5]. Unfortunately, while some progress has been made to the ultimate goal of a line list calculation, there are no adequate numerical results to speak of, yet. Due to the molecule's size and the complexity of the symmetry group, significant challenge is to be expected. The next chapter briefly describes the current (as of writing this thesis) issues.

The remainder of this chapter primarily discusses the MS group of $G_{36}$ and its extension for ro-vibrational calculations $G_{36}$(EM). As explained in Section 2.6, TROVE requires, for each irrep of the symmetry group in question [31], a group of matrices constituting that irrep. However, for $G_{36}$ or $G_{36}$(EM), these matrices have not been available in the literature thus far. This chapter aims at describing fully the groups $G_{36}$ and $G_{36}$(EM) and, in particular, determining matrix groups that define the irreps. The matrix groups obtained are used for symmetrising the PES and the ro-vibrational basis functions for ethane.

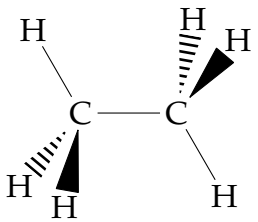



Figure 7.1: The structure of ethane in the staggered configuration.

## 7.2 The structure of the $G_{36}$ group

Longuet-Higgins [30] showed that the group $G_{36}$ can be written as a direct product of two smaller groups $C_{3v}^{(-)}$ and $C_{3v}^{(+)}$

$$G_{36} = C_{3v}^{(-)} \otimes C_{3v}^{(+)}; \tag{7.2.1}$$

both of these subgroups are of order 6 and isomorphic to the $C_{3v}$ point group. The top row and leftmost column of Table 7.1 define the elements of these two groups. For convenience, we label the elements of $C_{3v}^{(\pm)}$ as $g_j^{(\pm)}$, $j = 1, 2, \ldots, 6$; this notation is also defined in Table 7.1. The nuclei are labelled as in Figure 7.2.

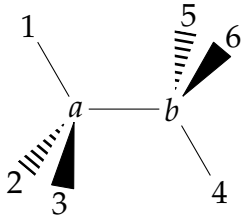



Figure 7.2: The labelling of the ethane nuclei.

In Table 7.1, the products $g_j^{(-)}\, g_k^{(+)} = g_k^{(+)}\, g_j^{(-)}$, where $g_j^{(-)} \in C_{3v}^{(-)}$ and $g_k^{(+)} \in C_{3v}^{(+)}$ are listed. Since $G_{36} = C_{3v}^{(-)} \otimes C_{3v}^{(+)}$, the 36 possible products $g_j^{(-)}\, g_k^{(+)} = g_k^{(+)}\, g_j^{(-)}$ constitute the complete group $G_{36}$. Due to the product group structure, we may also label elements in $G_{36}$ by the ordered pair $(g_j^{(-)}, g_k^{(+)})$.

Each of the groups $C_{3v}^{(-)}$ and $C_{3v}^{(+)}$ has three conjugacy classes, $\mathcal{C}_1^{(\pm)} = \{E\} = \{g_1^{(\pm)}\}$, $\mathcal{C}_2^{(\pm)} = \{g_2^{(\pm)}, g_3^{(\pm)}\}$, and $\mathcal{C}_3^{(\pm)} = \{g_4^{(\pm)}, g_5^{(\pm)}, g_6^{(\pm)}\}$. With the direct product structure of $G_{36}$, it is clear that the conjugacy clases of $G_{36}$ are obtained as $\mathcal{C}_i^{(-)} \times \mathcal{C}_j^{(+)}$, that is, a conjugacy class of $G_{36}$ contains all elements $RS$ where $R \in \mathcal{C}_i^{(-)}$ and $S \in \mathcal{C}_j^{(+)}$. The top row and leftmost column of Table 7.1 indicates the conjugacy class structures of $C_{3v}^{(+)}$ and $C_{3v}^{(-)}$, respectively. A conjugacy class $\mathcal{C}_i^{(+)}$ of $C_{3v}^{(+)}$ is simultaneously, in the form $\mathcal{C}_1^{(-)} \times \mathcal{C}_i^{(+)}$, a conjugacy class of $G_{36}$. Similarly, the conjugacy classes of $C_{3v}^{(-)}$ are simultaneously conjugacy classes of $G_{36}$. In Table 7.1 indicates the complete set of $G_{36}$ conjugacy classes $\mathcal{C}_i^{(-)} \times \mathcal{C}_j^{(+)}$.

Table 7.1: The conjugacy class structure of $G_{36}$[a].

| $g_1^{(-)} = g_1^{(+)} = E$ | $g_2^{(-)} = (132)(456)$ | $g_3^{(-)} = (123)(465)$ | $g_4^{(-)} = (14)(25)(36)(ab)$ | $g_5^{(-)} = (16)(24)(35)(ab)$ | $g_6^{(-)} = (15)(26)(34)(ab)$ |
|---|---|---|---|---|---|
| $g_2^{(+)} = (123)(456)$ | (465) | (132) | (153426)$(ab)$ | (143625)$(ab)$ | (163524)$(ab)$ |
| $g_3^{(+)} = (132)(465)$ | (123) | (456) | (162435)$(ab)$ | (152634)$(ab)$ | (142536)$(ab)$ |
| $g_4^{(+)} = (14)(26)(35)(ab)^*$ | (152436)$(ab)^*$ | (163425)$(ab)^*$ | (23)(56)$^*$ | (12)(46)$^*$ | (13)(45)$^*$ |
| $g_5^{(+)} = (16)(25)(34)(ab)^*$ | (142635)$(ab)^*$ | (153624)$(ab)^*$ | (13)(46)$^*$ | (23)(45)$^*$ | (12)(56)$^*$ |
| $g_6^{(+)} = (15)(24)(36)(ab)^*$ | (162534)$(ab)^*$ | (143526)$(ab)^*$ | (12)(45)$^*$ | (13)(56)$^*$ | (23)(46)$^*$ |

[a] The top row and leftmost column contain the $G_{36}$ elements that also belong to the $C_{3v}^{(-)}$ or $C_{3v}^{(+)}$ group, respectively. The remaining entries are the products $g_j^{(-)}\, g_k^{(+)} = g_k^{(+)}\, g_j^{(-)}$, where $g_j^{(-)} \in C_{3v}^{(-)}$ is at the top of the column and $g_k^{(+)} \in C_{3v}^{(+)}$ is at the left end of the row. The horizontal and vertical lines denote the separation of the conjugacy classes.

## 7.3 Irreducible representations of $G_{36}$

The sum over group elements of the square of the trace of the representation matrix $M^{(\Gamma^{(-)},\Gamma^{(+)})}[(g^{(-)}, g^{(+)})]$, defined by direct product irrep matrices $M^{\Gamma^{(-)}}[g^{(-)}]$ and $M^{\Gamma^{(+)}}[g^{(+)}]$, is

$$\frac{1}{|G_{36}|}\sum_{g\in G_{36}} \chi^{(\Gamma^{(-)},\Gamma^{(+)})}(g)\chi^{(\Gamma^{(-)},\Gamma^{(+)})}(g)^* = \left(\frac{1}{|{C_{3v}}^{(+)}|}\sum_{g^{(-)}\in C_{3v}^{(-)}} \chi^{\Gamma^{(-)}}(g^{(-)})\chi^{\Gamma^{(-)}}(g^{(-)})^*\right)\times \left(\frac{1}{|{C_{3v}}^{(-)}|}\sum_{g^{(+)}\in C_{3v}^{(+)}} \chi^{\Gamma^{(+)}}(g^{(+)})\chi^{\Gamma^{(+)}}(g^{(+)})^*\right) = 1 \tag{7.3.1}$$

so that the product representation is an irreducible representation. Since there are nine conjugacy classes for $G_{36}$ and nine such irreps, every irrep is of this form. The $G_{36}$ irreps could therefore in principle be labelled as $(\Gamma^{(-)}, \Gamma^{(+)})$, where $\Gamma^{(-)}$ is an irrep of $C_{3v}^{(-)}$ and $\Gamma^{(+)}$ is an irrep of $C_{3v}^{(+)}$. $C_{3v}^{(-)}$ and $C_{3v}^{(+)}$ both have the one-dimensional irreps $A_1$ and $A_2$ together with the two-dimensional irrep $E$ (Table 7.7 in Appendix 7.A). Customarily, the irreps of $G_{36}$ are labelled as in Table 12 of Longuet-Higgins [30]. Table 7.2 is the character table for $G_{36}$, obtained from Table 12 of Longuet-Higgins [30], with the irreducible representations labelled by their customary labels $\Gamma_{36}$ and the combination labels $(\Gamma^{(-)}, \Gamma^{(+)})$. The bottom row of the table indicates the $\mathcal{C}_i^{(-)} \times \mathcal{C}_j^{(+)}$ label of each class. For each conjugacy class of $G_{36}$ , Table 7.2 gives a representative element together with

the number of group elements in the conjugacy class. The complete conjugacy classes are obtained from Table 7.1.

Of particular concern are the doubly-degenerate irreducible representations $E_1 = (E, A_1)$, $E_2 = (E, A_2)$, $E_3 = (A_1, E)$, and $E_4 = (A_2, E)$ together with the four-dimensional irreducible representation $G = (E, E)$.

Table 7.2: Character table of $G_{36}$.

| $\Gamma_{36}$ | $(\Gamma^{(-)}, \Gamma^{(+)})$ | $E$ | $(123)(456)$ | $(14)(26)(35)(ab)^*$ | $(123)(465)$ | $(123)$ | $(142635)(ab)^*$ | $(14)(25)(36)(ab)$ | $(142536)(ab)$ | $(12)(45)^*$ |
|---|---|---|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 2 | 4 | 6 | 3 | 6 | 9 |
| $A_1$ | $(A_1, A_1)$ | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| $A_2$ | $(A_2, A_1)$ | 1 | 1 | 1 | 1 | 1 | 1 | −1 | −1 | −1 |
| $A_3$ | $(A_1, A_2)$ | 1 | 1 | −1 | 1 | 1 | −1 | 1 | 1 | −1 |
| $A_4$ | $(A_2, A_2)$ | 1 | 1 | −1 | 1 | 1 | −1 | −1 | −1 | 1 |
| $E_1$ | $(E, A_1)$ | 2 | 2 | 2 | −1 | −1 | −1 | 0 | 0 | 0 |
| $E_2$ | $(E, A_2)$ | 2 | 2 | −2 | −1 | −1 | 1 | 0 | 0 | 0 |
| $E_3$ | $(A_1, E)$ | 2 | −1 | 0 | 2 | −1 | 0 | 2 | −1 | 0 |
| $E_4$ | $(A_2, E)$ | 2 | −1 | 0 | 2 | −1 | 0 | −2 | 1 | 0 |
| $G$ | $(E, E)$ | 4 | −2 | 0 | −2 | 1 | 0 | 0 | 0 | 0 |
| | | $\mathcal{C}_1^{(-)} \times \mathcal{C}_1^{(+)}$ | $\mathcal{C}_1^{(-)} \times \mathcal{C}_2^{(+)}$ | $\mathcal{C}_1^{(-)} \times \mathcal{C}_3^{(+)}$ | $\mathcal{C}_2^{(-)} \times \mathcal{C}_1^{(+)}$ | $\mathcal{C}_2^{(-)} \times \mathcal{C}_2^{(+)}$ | $\mathcal{C}_2^{(-)} \times \mathcal{C}_3^{(+)}$ | $\mathcal{C}_3^{(-)} \times \mathcal{C}_1^{(+)}$ | $\mathcal{C}_3^{(-)} \times \mathcal{C}_2^{(+)}$ | $\mathcal{C}_3^{(-)} \times \mathcal{C}_3^{(+)}$ |

In Section 12.4 of Ref. [31], and in Section 3.1 of Ref. [115], it is discussed how the point group $\boldsymbol{C}_{3v}$ and any group isomorphic to it can be defined in terms of two generating operations, one of which belongs to the two-member class $\mathcal{C}_2^{(\pm)}$ of $\boldsymbol{C}_{3v}$ and the other to the three-member class $\mathcal{C}_3^{(\pm)}$. For the group $\boldsymbol{C}_{3v}^{(-)}$, the generating operations selected are $g_2^{(-)} = (123)(465)$ and $g_4^{(-)} = (14)(25)(36)(ab)$, whereas for $\boldsymbol{C}_{3v}^{(+)}$, the operations $g_2^{(+)} = (123)(456)$ and $g_4^{(+)} = (14)(26)(35)(ab)^*$ are chosen. We then have $g_3^{(\pm)} = (g_2^{(\pm)})^2$, $g_6^{(\pm)} = g_2^{(\pm)} g_4^{(\pm)}$, and $g_5^{(\pm)} = g_2^{(\pm)} g_6^{(\pm)}$.

## 7.4 $G_{36}$ irrep matrices

The representation matrices for the non-degenerate irreps $A_1$, $A_2$, $A_3$, and $A_4$ of $\boldsymbol{G}_{36}$ are uniquely defined as equal to the representation characters; these can be found in

Table 7.2. The $2 \times 2$ representation matrices of the irreps $E_1$, $E_2$, $E_3$, and $E_4$ are, however, are defined up to equivalent representation. Having found one such set of matrices $M^{E_i}[g]$, where $g_g \in \boldsymbol{G}_{36}$, we can generate infinitely many equivalent representations with representation matrices $UM_{E_n}[g]U^{-1}$, where $U$ is an arbitrary and invertible $2 \times 2$ matrix. In practice, representation matrices in TROVE are real and orthogonal, and that drastically limits the possible choices. One particular set of representation matrices is selected here for $E_1$ by initially choosing the representation matrix for $g_2^{(-)} = (123)(465)$, one of the generating operations of $\boldsymbol{C}_{3v}^{(-)}$. It is set to

$$M^E[(123)(465)] = \begin{pmatrix} \cos\left(\frac{2\pi}{3}\right) & -\sin\left(\frac{2\pi}{3}\right) \\ \sin\left(\frac{2\pi}{3}\right) & \cos\left(\frac{2\pi}{3}\right) \end{pmatrix} = \begin{pmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}. \tag{7.4.1}$$

The $2 \times 2$ orthogonal matrix $M^E[(123)(465)]$ satisfies the relation $M^E[(123)(465)]^3 = I$, the $2 \times 2$ unit matrix, imposed by the fact that $[(123)(465)]^3 = E$. An alternative choice would be the matrix with the signs of the $\sin(2\pi/3)$ terms reversed. The other generating operation of $\boldsymbol{C}_{3v}^{(-)}$, $g_4^{(-)} = (14)(25)(36)(ab)$, is self-inverse: $(g_4^{(-)})^2 = [(14)(25)(36)(ab)]^2 = E$. The $2 \times 2$ orthogonal matrix representing $g_4^{(-)}$ is also self-inverse and we can choose it as

$$M^E[(14)(25)(36)(ab)] = \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix} \tag{7.4.2}$$

where $\theta$ is arbitrary. All such matrices satisfy $M^E[(14)(25)(36)(ab)]^2 = I$. $\theta = 0$ is chosen, so that

$$M^E[(14)(25)(36)(ab)] = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{7.4.3}$$

The two matrices $M^E[(123)(465)]$ and $M^E[(14)(25)(36)(ab)]$ have traces of $-1$ and 0, respectively, and it is seen in Table 7.7 that they generate the irrep $E$ of $\boldsymbol{C}_{3v}^{(-)}$. We can now use the relations in Section 12.4 of Ref. [31] or, equivalently, in Section 3.1 of Ref. [115] to determine, by matrix multiplication involving $M^E[(123)(465)]$ and $M^E[(14)(25)(36)(ab)]$, the representation matrices for all operations in $\boldsymbol{C}_{3v}^{(-)}$.

It was discussed above how the irrep $E_1$ of $\boldsymbol{G}_{36}$ can be classified as $(\Gamma^{(-)}, \Gamma^{(+)}) = (E, A_1)$, where $\Gamma^{(-)}$ and $\Gamma^{(+)}$ are irreps of $\boldsymbol{C}_{3v}^{(-)}$ and $\boldsymbol{C}_{3v}^{(+)}$, respectively. We have already determined a group of representation matrices belonging to the $E$ irrep of $\boldsymbol{C}_{3v}^{(-)}$, and to obtain one for the $A_1$ (totally symmetric) irrep of $\boldsymbol{C}_{3v}^{(+)}$, the $1 \times 1$ representation matrices

$$M^{A_1}[(123)(456)] = M^{A_1}[(14)(26)(35)(ab)^*] = 1 \tag{7.4.4}$$

are introduced for its generating operations $g_2^{(+)} = (123)(456)$ and $g_4^{(+)} = (14)(26)(35)(ab)^*$. Again, we can use the relations in Section 12.4 of Ref. [31] or in Section 3.1 of Ref. [115] to determine the representation matrices for all operations on $\boldsymbol{C}_{3v}^{(+)}$. It is rather trivial here since these representation matrices all are the $1 \times 1$ matrix 1.

We now have $E$ representation matrices for the six operations in $\boldsymbol{C}_{3v}^{(-)}$ and $A_1$ representation matrices for the six operations in $\boldsymbol{C}_{3v}^{(+)}$, and we can form $E_1$ representation

matrices for the 36 operations in $\boldsymbol{G}_{36}$ by forming the 36 products of $M^E[g^{(-)}]$, $g^{(-)} \in \boldsymbol{C}_{3v}^{(-)}$, with the constant (=1 always in this case) $M^{A_1}[g^{(+)}]$, $g^{(+)} \in \boldsymbol{C}_{3v}^{(+)}$.

For $E_2 = (E, A_2)$, we can obtain the representation matrices in the same vein as for $E_1$. The only difference is that we replace the $\boldsymbol{C}_{3v}^{(+)}$ representation matrices by

$$M^{A_2}[(123)(456)] = 1 \text{ and } M^{A_2}[(14)(26)(35)(ab)^*] = -1. \tag{7.4.5}$$

For $E_3 = (A_1, E)$ and $E_4 = (A_2, E)$, the representation matrices are determined by interchanging the $\boldsymbol{C}_{3v}^{(-)}$ and $\boldsymbol{C}_{3v}^{(+)}$ representation matrices in the determination made for $E_1$ and $E_2$, respectively.

Finally, the direct product of the representation matrices of the two $E$ representations gives the $G = (E, E)$ representation.

## 7.5 Ro-vibrational coordinates used for ethane

For ethane, there are $3 \times 8 - 7 = 17$ small-amplitude vibrational coordinates. These will be referred to as the rigid coordinates as they have an equilibrium value. The torsional coordinate (describing the independent rotations of the $CH_3$ groups) does not have an equilibrium. Instead, it has three. Figure 7.3 shows representative members of three of the vibrational coordinate classes: the C−C bond length denoted by $\mathcal{R}^C$, one of six C−$H_k$ bonds denoted by $\mathcal{R}$, and one of six bond angles ($H_k$–C–C) denoted by $\alpha$. The small-amplitude vibrational coordinates $\mathcal{R}^C$, $\mathcal{R}_k$, and $\alpha_k$ ($k = 1 \ldots 6$) measure the displacements of the respective internal coordinates from their equilibrium values, that is the coordinates actually used are $\mathcal{R}^C - \mathcal{R}_e^C$, $\mathcal{R}_i - \mathcal{R}_e$ ($i = 0 \ldots 6$) and $\alpha_k - \alpha_e$ ($k = 1 \ldots 6$). The $\mathcal{R}_k/\alpha_k$ coordinates are equivalent and so they have a common equilibrium value $\mathcal{R}_e/\alpha_e$.

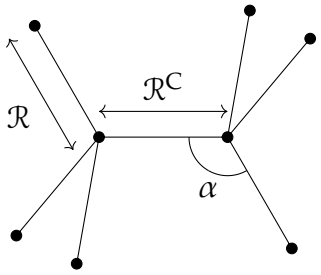



Figure 7.3: Representative members of three of the vibrational coordinate classes. Here, $\mathcal{R}^C$ is the C−C bond length, $\mathcal{R}$ is one of the six C–$H_k$ bond lengths $r_k$, and $\alpha$ is one of the six ($H_k$-C-C) bond angles $\alpha_k$.

The last vibrational class is obtained from six dihedral angles $\theta_{12}$, $\theta_{23}$, $\theta_{31}$, $\theta_{45}$, $\theta_{56}$, $\theta_{64}$, one of which is labelled $\theta$ in Figure 7.4. $\theta_{ij}$ is the angle between the $H_i$−C−C and $H_j$−C−C planes, where protons $i$ and $j$ belong to the same $CH_3$ group. Only four of the six angles are linearly independent due to the constraints $\theta_{12} + \theta_{23} + \theta_{31} = \theta_{45} + \theta_{56} + \theta_{64} = 2\pi$. The positive directions of rotation for the $\theta_{ij}$ angles are from proton $1 \to 2 \to 3$ for the one $CH_3$ group, and from proton $4 \to 5 \to 6$ for the other. The independent

coordinates constructed from the six $\theta_{ij}$ angles are

$$\begin{aligned}\gamma_1 &= \frac{1}{\sqrt{6}}(2\theta_{23} - \theta_{31} - \theta_{12}),\\ \gamma_2 &= \frac{1}{\sqrt{2}}(\theta_{31} - \theta_{12}),\\ \delta_1 &= \frac{1}{\sqrt{6}}(2\theta_{56} - \theta_{64} - \theta_{45}),\\ \delta_2 &= \frac{1}{\sqrt{2}}(\theta_{64} - \theta_{45}),\end{aligned} \tag{7.5.1}$$

which transform as the $G$ representation of $\boldsymbol{G}_{36}$.

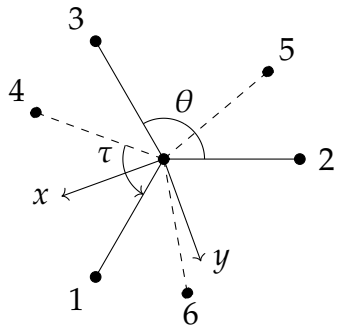



Figure 7.4: A projection of ethane, with the $CH_3$ group containing protons 1, 2, and 3, indicated by solid C−H bonds, being closest to the viewer. One of the dihedral angles used in the vibrational classes is labelled by $\theta$ and the torsional angle is labelled by $\tau$ and is measured in the counterclockwise direction. The $x$-axis halves the dihedral angle between the $H_1-C-C$ and $H_4-C-C$ planes.

To describe the orientation of the BF-frame when the rigid coordinates are in equilibrium, one approach is to attach coordinate axes on each $CH_3$ group. Here, the $z$-axis is the same for both and points from $C_b$ to $C_a$, while the $x_a$- and $x_b$-axes point in the direction of $C_a-H_1$ and $C_b-H_4$, respectively, when the molecule is viewed in the projection of Figure 7.5. The $y$-axes ensure that the Cartesian axes are right handed. With this construction, the $\theta$ and $\phi$ Euler angles describing the direction of the $z$-axis for $CH_3$ group are the same while the $\chi$ angles describing the rotation about the $z$-axis are different and are denoted by $\chi_a$ and $\chi_b$. These increase in the counterclockwise direction due to the right hand rule.

As explained in Ref. [31], to achieve maximum separation of the torsional, and rotational motion, it is expedient to define two new coordinates from $\chi_a$ and $\chi_b$. The rotational coordinate $\chi$ is set to to

$$\chi = \frac{1}{2}(\chi_a + \chi_b) \tag{7.5.2}$$

and hence the $x$-axis shown in Figure 7.4 halves the angle between $H_1-C-C$ and $H_4-C-C$ and increases in the counterclockwise direction. The torsional angle $\tau$ could be defined as

$$\tau = \chi_a - \chi_b \tag{7.5.3}$$

and hence, as indicated in Figure 7.4, is the angle from $H_4-C-C$ to $H_1-C-C$ in the counterclockwise direction. TROVE uses a different choice and defines $\tau$ as the average of three dihedral angles. To define these, three pairs of protons $(4,1)$, $(6,2)$, and $(5,3)$ are formed with the two members belonging to different $CH_3$ groups. The pairs are chosen such that the protons $i$ and $j$ in each $(i,j)$ pair form a dihedral angle $\tau_{ij}$ of $\pi$ in the staggered equilibrium geometry of Figure 7.1 and $\tau_{ij} = 0$ for the eclipsed geometry (see Figure 7.6), using the labelling of Figure 7.2. In a general instantaneous geometry, each $(i,j)$ pair defines a dihedral angle $\tau_{ij}$ (where the positive direction of rotation for the $\tau_{ij}$ angles is from proton $1 \rightarrow 2 \rightarrow 3$) and the torsional angle is then given by a symmetric combination

$$\tau = \frac{1}{3}(\tau_{41} + \tau_{62} + \tau_{53}). \tag{7.5.4}$$

With this definition, $\tau = 0$, $2\pi/3$, and $4\pi/3$ correspond to eclipsed configurations, while, at $\tau = \pi/3$, $\pi$, and $5\pi/3$, the molecule is in one of its three rigid coordinate equilibrium geometries. The torsional angle $\tau$ has definite transformation properties under the operations of $\boldsymbol{G}_{36}$ (see also Appendix 7.B). As discussed in Section 7.5.2, the two coordinate-pair values $(\tau, \chi)$ and $(\tau + 2\pi, \chi + \pi)$ describe the same physical situation. However, the coordinate which $\tau$ is based on, $\chi_a - \chi_b$, has a range of $4\pi$. Although a given geometry can be specified by a value of $\tau$ in the interval $[0, 2\pi]$, we must allow $\tau$ to range over $[0, 4\pi]$ to obtain a correct correlation with $\chi_a - \chi_b$ (see Section 7.5.2 below).

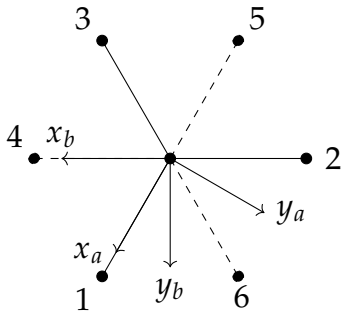



Figure 7.5: A projection of ethane, with the $CH_3$ group containing protons 1, 2, and 3, indicated by solid C−H bonds, being closest to the viewer. The $x$ and $y$ components of the coordinate axes attached to each $CH_3$ group is shown, the subscript $a$ signifying that the coordinate axes are for the $C_aH_3$ group. To ensure the coordinate system is right handed, the $z$-axis (the same for both groups) points from $C_b$ to $C_a$. With this construction, the $\theta$ and $\phi$ Euler angles describing the direction of the $z$-axis are the same for each $CH_3$ group while the $\chi$ angle describing the rotation about the $z$-axis are different and are denoted by $\chi_a$ and $\chi_b$. These increase in the counterclockwise direction due to the right hand rule.

In conclusion, the coordinate classes, used to diagonalise the reduced Hamiltonians, are

1. the C-C bond length $\mathcal{R}^C$;
2. six C-H bond lengths $\mathcal{R}_k$, $k = 1, 2, \ldots, 6$;
3. six bond angles ($H_k$-C-C) = $\alpha_k$, $k = 1, 2, \ldots, 6$;

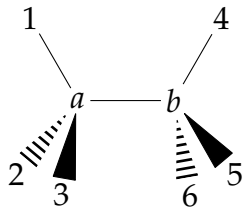



Figure 7.6: Ethane in the eclipsed configuration.

4. four dihedral-angle coordinates $\gamma_1$, $\gamma_2$, $\delta_1$, and $\delta_2$;

5. the torsional angle $\tau$; and

6. the three rotational angles $(\theta, \phi, \chi)$.

The primitive basis functions for classes 1, 2, and 3 are obtained using the Numerov–Cooley approach [34, 125, 126], while, harmonic-oscillator eigenfunctions [31] are used for class 4. The primative basis set for class 5 is constructed from the normalised, $4\pi$-periodic Fourier series functions $\sqrt{1/2\pi}\cos(k\tau/2)$ and $\sqrt{1/2\pi}\sin(k\tau/2)$. Both the KEO and PES must be functions of $\tau$ in the extended interval $[0, 4\pi]$ (see Section 7.5.2).

### 7.5.1 Transformation of the vibrational coordinates under $G_{36}$

As detailed in Section 7.3, one can construct every operation of $\boldsymbol{G}_{36}$ with only four generating operations. In the TROVE calculations, the generators $g_2^{(-)} = (132)(456)$, $g_2^{(+)} = (123)(456)$, $g_4^{(+)} = (14)(26)(35)(ab)^*$ and $g_4^{(-)} = (14)(25)(36)(ab)$ are used. Then, to describe the transformation properties of a coordinate for all elements in $\boldsymbol{G}_{36}$, one only needs to determine these properties for the four generating operations. In TROVE, a procedure to generate the irreducible representations of $\boldsymbol{G}_{36}$ based on these four group generators and the multiplication rules in Table 7.1 was implemented. The latter can be conveniently represented as the following recursive rule:

$$T_i = T_j\, T_k, \tag{7.5.5}$$

where the operations $T_i$, $T_j$ and $T_k$ $(i, j, k = 1, \ldots, 36)$ are as organised in Table 7.3.

Each of the vibrational coordinates and the torsional coordinate can be expressed as a function of the nuclear Cartesian coordinates. The procedure of Appendix 2.A was used to determine the transformation of these coordinates under the $\boldsymbol{G}_{36}$ operations in a systematic way, although in the majority of cases the result is intuitive.

The C−C bond length $\mathcal{R}^{\mathrm{C}}$ is invariant under all $\boldsymbol{G}_{36}$ operations. The six C−$H_k$ bond lengths $\mathcal{R}_k$ and the six bond angles $H_k$-C-C = $\alpha_k$ are given the generic labels $\beta_k$, as the two classes transform identically. ($\beta_1$, $\beta_2$, $\beta_3$) have the following transformation

Table 7.3: The recursive rules to generate the elements of $G_{36}$ using the four generators $T_2 = g_2^{(+)} = (123)(465)$, $T_4 = g_4^{(-)} = (14)(25)(36)(ab)$, $T_7 = g_2^{(-)} = (132)(456)$, and $T_{19} = g_4^{(+)} = (14)(26)(35)(ab)^*$. See also Table 7.1 for the class structure of $G_{36}$ and Figure 7.7 for an illustration of the effects of the generators.

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| $T_1$ | = | $E$ | | $T_{19}$ | = | $(14)(25)(36)(ab)$ | $= g_4^{(-)}$ |
| $T_2$ | = | $(123)(456)$ | $= g_2^{(+)}$ | $T_{20}$ | = | $(16)(24)(34)(ab)$ | $= T_7\,T_{21}$ |
| $T_3$ | = | $(132)(465)$ | $= T_2{}^2$ | $T_{21}$ | = | $(15)(26)(34)(ab)$ | $= T_7\,T_{19}$ |
| $T_4$ | = | $(14)(26)(35)(ab)^*$ | $= g_4^{(+)}$ | $T_{22}$ | = | $(153426)(ab)$ | $= T_{19}\,T_2$ |
| $T_5$ | = | $(16)(25)(34)(ab)^*$ | $= T_2\,T_6$ | $T_{23}$ | = | $(162435)(ab)$ | $= T_{19}\,T_3$ |
| $T_6$ | = | $(15)(24)(36)(ab)^*$ | $= T_2\,T_4$ | $T_{24}$ | = | $(143624)(ab)$ | $= T_{20}\,T_2$ |
| $T_7$ | = | $(132)(456)$ | $= g_2^{(-)}$ | $T_{25}$ | = | $(152634)(ab)$ | $= T_{20}\,T_3$ |
| $T_8$ | = | $(123)(465)$ | $= T_7{}^2$ | $T_{26}$ | = | $(163524)(ab)$ | $= T_{21}\,T_2$ |
| $T_9$ | = | $(465)$ | $= T_7\,T_2$ | $T_{27}$ | = | $(142536)(ab)$ | $= T_{21}\,T_3$ |
| $T_{10}$ | = | $(123)$ | $= T_7\,T_3$ | $T_{28}$ | = | $(23)(56)^*$ | $= T_{19}\,T_4$ |
| $T_{11}$ | = | $(132)$ | $= T_8\,T_2$ | $T_{29}$ | = | $(13)(46)^*$ | $= T_{19}\,T_5$ |
| $T_{12}$ | = | $(456)$ | $= T_8\,T_3$ | $T_{30}$ | = | $(12)(45)^*$ | $= T_{19}\,T_6$ |
| $T_{13}$ | = | $(152436)(ab)^*$ | $= T_7\,T_4$ | $T_{31}$ | = | $(12)(46)^*$ | $= T_{20}\,T_4$ |
| $T_{14}$ | = | $(142635)(ab)^*$ | $= T_7\,T_5$ | $T_{32}$ | = | $(23)(45)^*$ | $= T_{20}\,T_5$ |
| $T_{15}$ | = | $(162534)(ab)^*$ | $= T_7\,T_6$ | $T_{33}$ | = | $(13)(56)^*$ | $= T_{20}\,T_6$ |
| $T_{16}$ | = | $(163425)(ab)^*$ | $= T_8\,T_4$ | $T_{34}$ | = | $(13)(45)^*$ | $= T_{21}\,T_4$ |
| $T_{17}$ | = | $(153624)(ab)^*$ | $= T_8\,T_5$ | $T_{35}$ | = | $(12)(56)^*$ | $= T_{21}\,T_5$ |
| $T_{18}$ | = | $(143526)(ab)^*$ | $= T_8\,T_6$ | $T_{36}$ | = | $(23)(46)^*$ | $= T_{21}\,T_6$ |

properties under the generating operations:

$$\begin{aligned}
\begin{pmatrix}\beta'_1\\ \beta'_2\\ \beta'_3\end{pmatrix} &= \begin{pmatrix}0 & 0 & 1\\ 1 & 0 & 0\\ 0 & 1 & 0\end{pmatrix}\begin{pmatrix}\beta_1\\ \beta_2\\ \beta_3\end{pmatrix} \text{ for } (123)(456),\\
\begin{pmatrix}\beta'_1\\ \beta'_2\\ \beta'_3\end{pmatrix} &= \begin{pmatrix}1 & 0 & 0\\ 0 & 0 & 1\\ 0 & 1 & 0\end{pmatrix}\begin{pmatrix}\beta_4\\ \beta_5\\ \beta_6\end{pmatrix} \text{ for } (14)(26)(35)(ab)^*,\\
\begin{pmatrix}\beta'_1\\ \beta'_2\\ \beta'_3\end{pmatrix} &= \begin{pmatrix}0 & 1 & 0\\ 0 & 0 & 1\\ 1 & 0 & 0\end{pmatrix}\begin{pmatrix}\beta_1\\ \beta_2\\ \beta_3\end{pmatrix} \text{ for } (132)(456),\\
\begin{pmatrix}\beta'_1\\ \beta'_2\\ \beta'_3\end{pmatrix} &= \begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}\beta_4\\ \beta_5\\ \beta_6\end{pmatrix} \text{ for } (14)(25)(36)(ab)
\end{aligned} \tag{7.5.6}$$

while for $(\beta_4, \beta_5, \beta_6)$ they are given by

$$\begin{aligned}
\begin{pmatrix}\beta'_4\\ \beta'_5\\ \beta'_6\end{pmatrix} &= \begin{pmatrix}0 & 0 & 1\\ 1 & 0 & 0\\ 0 & 1 & 0\end{pmatrix}\begin{pmatrix}\beta_4\\ \beta_5\\ \beta_6\end{pmatrix} \text{ for } (123)(456),\\
\begin{pmatrix}\beta'_4\\ \beta'_5\\ \beta'_6\end{pmatrix} &= \begin{pmatrix}1 & 0 & 0\\ 0 & 0 & 1\\ 0 & 1 & 0\end{pmatrix}\begin{pmatrix}\beta_1\\ \beta_2\\ \beta_3\end{pmatrix} \text{ for } (14)(26)(35)(ab)^*,\\
\begin{pmatrix}\beta'_4\\ \beta'_5\\ \beta'_6\end{pmatrix} &= \begin{pmatrix}0 & 0 & 1\\ 1 & 0 & 0\\ 0 & 1 & 0\end{pmatrix}\begin{pmatrix}\beta_4\\ \beta_5\\ \beta_6\end{pmatrix} \text{ for } (132)(456),\\
\begin{pmatrix}\beta'_4\\ \beta'_5\\ \beta'_6\end{pmatrix} &= \begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}\beta_1\\ \beta_2\\ \beta_3\end{pmatrix} \text{ for } (14)(25)(36)(ab).
\end{aligned} \tag{7.5.7}$$

The dihedral angle coordinates $(\gamma_1, \gamma_2, \delta_1, \delta_2)$ are mixed by the $G_{36}$ operations. The transformation properties are most easily determined by using the correspondence with the transformation of the $r_k$ and $\alpha_k$ coordinates. For $(\gamma_1, \gamma_2)$, we write

$$\begin{pmatrix}\gamma_1\\ \gamma_2\\ 2\pi\end{pmatrix} = Z\begin{pmatrix}\theta_{12}\\ \theta_{23}\\ \theta_{31}\end{pmatrix} \tag{7.5.8}$$

with

$$Z = \begin{pmatrix}-\frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} & -\frac{1}{\sqrt{6}}\\ -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}}\\ 1 & 1 & 1\end{pmatrix} \tag{7.5.9}$$

where the third row of $Z$ takes into account the constraint $\theta_{12} + \theta_{23} + \theta_{31} = 2\pi$.

After the operation (123)(456), protons 1, 2, and 3 are found at the positions initially occupied by protons 3, 1, and 2, respectively, and so the angles $\theta_{12}$, $\theta_{23}$, $\theta_{31}$ are permuted

as follows

$$\begin{pmatrix}\theta'_{12}\\ \theta'_{23}\\ \theta'_{31}\end{pmatrix} = \begin{pmatrix}\theta_{31}\\ \theta_{12}\\ \theta_{23}\end{pmatrix} = S\begin{pmatrix}\theta_{12}\\ \theta_{23}\\ \theta_{31}\end{pmatrix} = SZ^{-1}\begin{pmatrix}\gamma_1\\ \gamma_2\\ 2\pi\end{pmatrix} \tag{7.5.10}$$

with

$$S = \begin{pmatrix}0 & 0 & 1\\ 1 & 0 & 0\\ 0 & 1 & 0\end{pmatrix} \tag{7.5.11}$$

and so the transformed values of $(\gamma_1, \gamma_2)$ are

$$\begin{pmatrix}\gamma'_1\\ \gamma'_2\\ 2\pi\end{pmatrix} = Z\begin{pmatrix}\theta'_{12}\\ \theta'_{23}\\ \theta'_{31}\end{pmatrix} = ZSZ^{-1}\begin{pmatrix}\gamma_1\\ \gamma_2\\ 2\pi\end{pmatrix} \tag{7.5.12}$$

where

$$Z\,S\,Z^{-1} = \begin{pmatrix}-\frac{1}{2} & -\frac{\sqrt{3}}{2} & 0\\ \frac{\sqrt{3}}{2} & -\frac{1}{2} & 0\\ 0 & 0 & 1\end{pmatrix}. \tag{7.5.13}$$

With the upper $2\times 2$ corner of this matrix, we can express the transformed values $(\gamma'_1, \gamma'_2)$ in terms of $(\gamma_1, \gamma_2)$. The transformation matrix for $(\gamma_1, \gamma_2)$ for (123)(456), and the ones for the other generating operations, obtained in a similar manner, are

$$\begin{aligned}
\begin{pmatrix}\gamma'_1\\ \gamma'_2\end{pmatrix} &= \begin{pmatrix}-\frac{1}{2} & -\frac{\sqrt{3}}{2}\\ \frac{\sqrt{3}}{2} & -\frac{1}{2}\end{pmatrix}\begin{pmatrix}\gamma_1\\ \gamma_2\end{pmatrix} \text{ for (123)(456)},\\
\begin{pmatrix}\gamma'_1\\ \gamma'_2\end{pmatrix} &= \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\begin{pmatrix}\delta_1\\ \delta_2\end{pmatrix} \text{ for (14)(26)(35)}(ab)^*,\\
\begin{pmatrix}\gamma'_1\\ \gamma'_2\end{pmatrix} &= \begin{pmatrix}-\frac{1}{2} & \frac{\sqrt{3}}{2}\\ -\frac{\sqrt{3}}{2} & -\frac{1}{2}\end{pmatrix}\begin{pmatrix}\gamma_1\\ \gamma_2\end{pmatrix} \text{ for (132)(456)},\\
\begin{pmatrix}\gamma'_1\\ \gamma'_2\end{pmatrix} &= \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}\begin{pmatrix}\delta_1\\ \delta_2\end{pmatrix} \text{ for (14)(25)(36)}(ab),
\end{aligned} \tag{7.5.14}$$

and those for $(\delta_1, \delta_2)$ are given by

$$\begin{aligned}
\begin{pmatrix}\delta'_1\\ \delta'_2\end{pmatrix} &= \begin{pmatrix}-\frac{1}{2} & -\frac{\sqrt{3}}{2}\\ \frac{\sqrt{3}}{2} & -\frac{1}{2}\end{pmatrix}\begin{pmatrix}\delta_1\\ \delta_2\end{pmatrix} \text{ for (123)(456)},\\
\begin{pmatrix}\delta'_1\\ \delta'_2\end{pmatrix} &= \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\begin{pmatrix}\gamma_1\\ \gamma_2\end{pmatrix} \text{ for (14)(26)(35)}(ab)^*,\\
\begin{pmatrix}\delta'_1\\ \delta'_2\end{pmatrix} &= \begin{pmatrix}-\frac{1}{2} & -\frac{\sqrt{3}}{2}\\ \frac{\sqrt{3}}{2} & -\frac{1}{2}\end{pmatrix}\begin{pmatrix}\delta_1\\ \delta_2\end{pmatrix} \text{ for (132)(456)},\\
\begin{pmatrix}\delta'_1\\ \delta'_2\end{pmatrix} &= \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}\begin{pmatrix}\gamma_1\\ \gamma_2\end{pmatrix} \text{ for (14)(25)(36)}(ab).
\end{aligned} \tag{7.5.15}$$

The transformation of the torsional angle $\tau$ due to $G_{36}$ operations is detailed in Appendix 7.B. The results are listed in Table 7.4, which also provides the equivalent rotation of the BF-frame which will be used in the next section.

### 7.5.2 The extended molecular symmetry group $G_{36}$(EM)

As described in Section 15.4.4 of [31], the separation of the rotational and torsional degrees of freedom has led to $\chi$ and $\tau$ being double-valued. That is, there are two sets of $(\chi, \tau)$ values associated with the same physical situation. This is most straightforwardly seen by considering Eq. (7.5.2) and the three angles $\chi_a$, $\chi_b$, and $\chi$ appearing in it. The angle $\chi_a$ is determined entirely by the positions in space of protons 1, 2, 3 and their carbon nucleus $C_a$; $\chi_b$ is determined by the positions of protons 4, 5, 6 and their carbon nucleus $C_b$; and $\chi$ is the average of the two. Due to the $2\pi$ periodicity of $\chi_a$, increasing it by $2\pi$ does not change the positions in space of the nuclei in $C_aH_3$, however, in this case, $\chi \rightarrow \chi + \pi$ and $\tau \rightarrow \tau + 2\pi$. Thus, the two coordinate pairs ($\chi$,$\tau$) and ($\chi + \pi$,$\tau + 2\pi$) describe identical physical situations.

One way of avoiding this ambiguity would be to use a BF-frame system with, for example, $\chi = \chi_b$. This BF-frame is attached to the $CH_3$ group with protons 4, 5, and 6, and not influenced by the other $CH_3$ group. A $\chi$ coordinate chosen in this manner has no ambiguity. However, this choice precludes the separate symmetrisation of the rotational basis set, a requirement of TROVE and thoroughly explained in Chapter 6.

For non-rigid molecules, the Eckart conditions are amended by the Sayvetz condition [127]

$$\sum_{\alpha} m_\alpha a'_\alpha(\tau) \cdot (r_\alpha - a_\alpha) = 0 \tag{7.5.16}$$

where now the reference Cartesian coordinates $a(\tau)$ are functions of a non-rigid vibrational coordinate $\tau$ and the BF-frame is set to the PAS frame for every value of $\tau$ while all other coordinates are in equilibrium. The prime on $a'(\tau)$ signifies that it is a derivative. The frame in this case is the bisector of Section 7.5. If a reference frame based on $\chi = \chi_b$ was chosen with the $z$-axis parallel to the C–C bond and the origin in the middle of the C–C bond, the reference coordinates would be

$$a_{H_1}(\tau) = \begin{pmatrix} c\cos(\tau) \\ c\sin(\tau) \\ d \end{pmatrix}, \quad a_{H_4}(\tau) = \begin{pmatrix} c \\ 0 \\ -d \end{pmatrix} \tag{7.5.17}$$

for some $c$ and $d$. The coordinates of the other atoms would be

$$a_{H_2} = Aa_{H_1}, \quad a_{H_3} = A^2 a_{H_1}, \quad a_{H_6} = Aa_{H_4}, \quad a_{H_5} = A^2 a_{H_4}, \quad a_{C_b} = -a_{C_a} \tag{7.5.18}$$

where $A = M_z(2\pi/3)$. The carbons would only have a $z$-component and would be constants. The Eckart conditions would be

$$\begin{aligned} C = m_C(a_{C_a} \times r_{C_a} - a_{C_a} \times r_{C_b}) \\ + m_H(a_{H_1} \times r_{H_1} + Aa_{H_1} \times r_{H_2} + A^2 a_{H_1} \times r_{H_3} \\ + a_{H_4} \times r_{H_4} + A^2 a_{H_4} \times r_{H_5} + Aa_{H_4} \times r_{H_6}) = 0. \end{aligned} \tag{7.5.19}$$

The operation (123)(456) changes the hydrogen labels in the usual way. However, due to the change in the $\tau$ angle, the PAS equilibrium coordinates would change according to $a \rightarrow Aa$ for hydrogens 1, 2, and 3. The hydrogens 4, 5, and 6 and the carbons would

be unchanged. We would have

$$\begin{aligned} C' = m_{\mathrm{C}}(a_{\mathrm{C}_a} \times r_{\mathrm{C}_a} - a_{\mathrm{C}_a} \times r_{\mathrm{C}_b}) \\ + m_{\mathrm{H}}(Aa_{\mathrm{H}_1} \times r_{\mathrm{H}_3} + A^2 a_{\mathrm{H}_1} \times r_{\mathrm{H}_1} + a_{\mathrm{H}_1} \times r_{\mathrm{H}_2} \\ + a_{\mathrm{H}_4} \times r_{\mathrm{H}_6} + A^2 a_{\mathrm{H}_4} \times r_{\mathrm{H}_4} + Aa_{\mathrm{H}_4} \times r_{\mathrm{H}_5}). \end{aligned} \tag{7.5.20}$$

It is not possible to apply a rotation of $2\pi/3$ or $4\pi/3$ about the $z$-axis for $C' = 0$. Thus, we cannot symmetrise the rotational basis set with this frame.

On the other hand, the bisector reference frame for the Eckart-Sayvetz conditions can be used for this purpose, and in the process of proving this we will determine the equivalent rotations of the BF-frame for the $\boldsymbol{G}_{36}$ operations. Later, a geometric frame will also be introduced.

For the transformation of the torsional angle $\tau$ due to $\boldsymbol{G}_{36}$ operations, we use the results of Section 7.B. The results are listed in Table 7.4, which also provides the equivalent rotation of the BF-frame.

For the bisector PAS, the coordinates of $H_1$ and $H_4$ are given by

$$a_{\mathrm{H}_1}(\tau) = \begin{pmatrix} c\cos(\tau/2) \\ c\sin(\tau/2) \\ d \end{pmatrix}, \quad a_{\mathrm{H}_4}(\tau) = \begin{pmatrix} c\cos(\tau/2) \\ -c\sin(\tau/2) \\ -d \end{pmatrix} \tag{7.5.21}$$

for some $c$ and $d$. The coordinates of the other atoms are Eq. (7.5.18) for the new $a_{\mathrm{H}_1}$ and $a_{\mathrm{H}_4}$ where $A = M_z(2\pi/3)$. The carbons only have a $z$-component and are constants. The Eckart conditions are the same as Eq. (7.5.19) (except the $a$s depend on $\tau$) and the Sayvetz condition is given by

$$\begin{aligned} S = m_{\mathrm{C}}(a'_{\mathrm{C}_a} \cdot r_{\mathrm{C}_a} - a'_{\mathrm{C}_a} \cdot r_{\mathrm{C}_b}) - m_{\mathrm{C}}(a'_{\mathrm{C}_a} \cdot a_{\mathrm{C}_a} - a'_{\mathrm{C}_a} \cdot a_{\mathrm{C}_b}) \\ + m_{\mathrm{H}}(a'_{\mathrm{H}_1} \cdot r_{\mathrm{H}_1} + Aa'_{\mathrm{H}_1} \cdot r_{\mathrm{H}_2} + A^2 a'_{\mathrm{H}_1} \cdot r_{\mathrm{H}_1} + a'_{\mathrm{H}_4} \cdot r_{\mathrm{H}_4} + A^2 a'_{\mathrm{H}_4} \cdot r_{\mathrm{H}_5} + a'_{\mathrm{H}_4} \cdot r_{\mathrm{H}_6}) \\ - m_{\mathrm{H}}(a'_{\mathrm{H}_1} \cdot a_{\mathrm{H}_1} + Aa'_{\mathrm{H}_1} \cdot a_{\mathrm{H}_2} + A^2 a'_{\mathrm{H}_1} \cdot a_{\mathrm{H}_1} + a'_{\mathrm{H}_4} \cdot a_{\mathrm{H}_4} + A^2 a'_{\mathrm{H}_4} \cdot a_{\mathrm{H}_5} + a'_{\mathrm{H}_4} \cdot a_{\mathrm{H}_6}) = 0. \end{aligned} \tag{7.5.22}$$

Table 7.4: The generators of the extended group $\boldsymbol{G}_{36}$ of $C_2H_6$ and their effect on the torsional angle ($\tau$) and the equivalent rotation of the generator.

| Transformed $\tau$ | Equivalent rotation | $\boldsymbol{G}_{36}$ generator |
|---|---|---|
| $\tau - 4\pi/3$ | $E$ | (123)(456) |
| $\tau$ | $R_z^{2\pi/3}$ | (132)(456) |
| $2\pi - \tau$ | $E$ | (14)(26)(35)$(ab)^*$ |
| $\tau$ | $R_0^{\pi}$ | (14)(25)(36)$(ab)$ |

The operation (123)(456) changes the PAS equilibrium coordinates according to $a \to A^2 a$ for hydrogens 1, 2, and 3 and $a \to Aa$ for hydrogens 4, 5, and 6. The carbons

are unchanged. We thus have

$$C' = m_C(a_{C_a} \times r_{C_a} - a_{C_a} \times r_{C_b}) + m_H(A^2 a_{H_1} \times r_{H_3} + a_{H_1} \times r_{H_1} + A a_{H_1} \times r_{H_2} + A a_{H_4} \times r_{H_6} + a_{H_4} \times r_{H_4} + A^2 a_{H_4} \times r_{H_5}) = 0. \quad (7.5.23)$$

which is the original equation and therefore there is no equivalent rotation. The derivatives of the PAS equilibrium coordinates transform the same way under a change in $\tau$ and thus the Sayvetz condition is trivially satisfied.

For the operation (132)(456), the angle $\tau$ does not change, but the rotation $R_z^{2\pi/3}$ is applied and thus the instantaneous Cartesian coordinates change by $r \to A^T r = A^2 r$. We thus have

$$C' = m_C(a_{C_a} \times A^2 r_{C_a} - a_{C_a} \times A^2 r_{C_b}) + m_H(a_{H_1} \times A^2 r_{H_2} + A a_{H_1} \times A^2 r_{H_3} + A^2 a_{H_1} \times A^2 r_{H_1} + a_{H_4} \times A^2 r_{H_6} + A^2 a_{H_4} \times A^2 r_{H_4} + A a_{H_4} \times A^2 r_{H_5}). \quad (7.5.24)$$

We apply $A^T A$ to the conditions and using the same technique as before, the result is $A^T$ applied to the original conditions, and this is zero. For the Sayvetz condition, we use that $Aa \cdot Ab = a \cdot b$ and pre-apply $A$ to both terms of all dot products. The original condition is re-obtained.

The operation (14)(26)(35)$(ab)^*$ has no rotation and changes $a_{H_1} \to -a_{H_4}$ and vice versa. The result is

$$C' = -m_C(a_{C_a} \times r_{C_b} - a_{C_a} \times r_{C_a}) + m_H(a_{H_1} \times r_{H_1} + A a_{H_4} \times r_{H_6} + A^2 a_{H_4} \times r_{H_5} + a_{H_1} \times r_{H_1} + A^2 a_{H_1} \times r_{H_3} + A a_{H_1} \times r_{H_2})). \quad (7.5.25)$$

which is the original set of conditions and thus is zero. The Sayvetz condition is satisfied in the same way.

The element (14)(25)(36)$(ab)$ does not transform $\tau$ but has the equivalent rotation $R_0^\pi$. Thus it transforms the instantaneous coordinates to $r \to Nr$, where $N = M_x(\pi)$, and changes the conditions to

$$C' = m_C(a_{C_a} \times N r_{C_b} - a_{C_a} \times N r_{C_a}) + m_H(a_{H_1} \times N r_{H_4} + A a_{H_1} \times N r_{H_5} + A^2 a_{H_1} \times N r_{H_6} + a_{H_4} \times N r_{H_1} + A^2 a_{H_4} \times N r_{H_2} + A a_{H_4} \times N r_{H_1}). \quad (7.5.26)$$

The matrix $N$ changes the equilibrium Cartesian coordinates of the carbons by $Na = -a$ and $H_1$'s by $N a_{H_1} = a_{H_4}$. Also, $NA = A^2 N$ and $NA^2 = AN$. Thus, in applying $I = NN$ to the transformed condition, we have

$$C' = N(m_C(-a_{C_a} \times r_{C_b} + a_{C_a} \times r_{C_a}) + m_H(a_{H_4} \times r_{H_4} + A^2 a_{H_4} \times r_{H_5} + A a_{H_4} \times r_{H_6} + a_{H_1} \times r_{H_1} + A a_{H_1} \times r_{H_2} + A^2 a_{H_1} \times r_{H_1})) = 0. \quad (7.5.27)$$

which is zero. The Sayvetz condition works the same way as in the previous operations.

Before completing this proof, we must deal with the double-valuedness of $(\tau, \chi)$. This involves extending $G_{36}$ to the extended molucular symmetry group $G_{36}$(EM) in the manner first introduced by Hougen [128], as detailed in Section 15.4.4 of Ref. [31]. This extension involves the introduction of a fictitious operation $E'$ (taken to be different from the identity $E$) which, for ethane, we can think of as letting the $C_aH_3$ do a full torsional revolution relative to the $C_bH_3$. That is, $E'$ has the effect of transforming $\chi \rightarrow \chi + \pi$ and $\tau \rightarrow \tau + 2\pi$. After the application of $(E')^2$ the BF-frame is back where it started, and so we take $(E')^2 = E$. $E'$ does not affect the rigid coordinates; it has the same effect as the identity operation on the complete rotation-torsion-vibration wavefunction of ethane.

Table 7.5 shows the transformations of the $G_{36}$ operations on $\chi$ using the equivalent rotations of Table 7.4, as well as repeating the effect on $\tau$. It also includes the effect of $E'$ to show all generators of $G_{36}$(EM). As $E'$ does not effect any other coordinate, it commutes (for those coordinates) with $G_{36}$ operations. Actually, $E'$ also commutes with the $G_{36}$ operations for $\tau$ and $\chi$. This can be checked with Table 7.5. Therefore, $G_{36}$(EM) has the structure of a direct product

$$G_{36}(\mathrm{EM}) = G_{36} \times \{E, E'\}. \tag{7.5.28}$$

The two-element group $\{E, E'\}$ is cyclic of order 2 and has two irreps $A'$ and $A''$, both one-dimensional, with the representation matrices 1 or $-1$, respectively, under $E'$. The irreps of $G_{36}$(EM) are straightforwardly constructed from those of $G_{36}$. Each irrep $\Gamma$ of $G_{36}$ in Table 7.2 gives rise to two irreps of $G_{36}$(EM), $\Gamma_s = (\Gamma, A')$ and $\Gamma_d = (\Gamma, A'')$ as given in Table A-33 of [31]. An irrep $\Gamma_s$ has identical characters for the operations $E$ and $E'$, $\chi_s[E'] = \chi_s[E]$, whereas for the irrep $\Gamma_d$, $\chi_d[E'] = -\chi_d[E]$. As long as we pretend that $E' \neq E$, we must also pretend that the coordinate values $(\tau, \chi)$ and $E'(\tau, \chi) = (\tau + 2\pi, \chi + \pi)$ describe different physical situations. As a consequence, we must allow $\tau$ to be periodic with a period of $4\pi$ as already mentioned in connection with Eq. (7.5.4). The torsional potential energy function $V^{1D}(\tau)$ is periodic with period $2\pi$, $V^{1D}(\tau) = V^{1D}(\tau + 2\pi)$ for $\tau \in [0, 2\pi]$, and this symmetry causes the torsional wavefunctions to be either symmetric (of $\Gamma_s$ symmetry) or antisymmetric (of $\Gamma_d$ symmetry) under $E'$.

Table 7.5: Transformation of the torsion angle $\tau$ and the rotation angle $\chi$ under the generators of $G_{36}$(EM).

| Transformed $\tau$ | Transformed $\chi$ | $G_{36}$(EM) generator |
|---|---|---|
| $\tau - 4\pi/3$ | $\chi$ | (123)(456) |
| $\tau$ | $\chi + 2\pi/3$ | (132)(456) |
| $2\pi - \tau$ | $\chi + \pi$ | $(14)(26)(35)(ab)^*$ |
| $\tau$ | $\chi$ | $(14)(25)(36)(ab)$ |
| $\tau + 2\pi$ | $\chi + \pi$ | $E'$ |

We know that, in reality, $E' = E$, and so only functions and coordinates with $\Gamma_s$ symmetries occur in nature. Since we can form, for example, basis functions of an allowed $\Gamma_s$ symmetry as products of an even number of factors, each with a forbidden symmetry $\Gamma'_d$, say, we need to consider also the $\Gamma_d$ symmetries initially for the torsional

and rotational basis functions. The final wavefunctions resulting from our theoretical calculations should be subjected to a "reality check": they must necessarily have a $\Gamma_s$ symmetry in $G_{36}$(EM). In particular, torsional basis functions of $d$ symmetry must be combined with rotational basis functions of $d$ symmetry to produce a torsion-rotation basis function of an allowed $s$ symmetry.

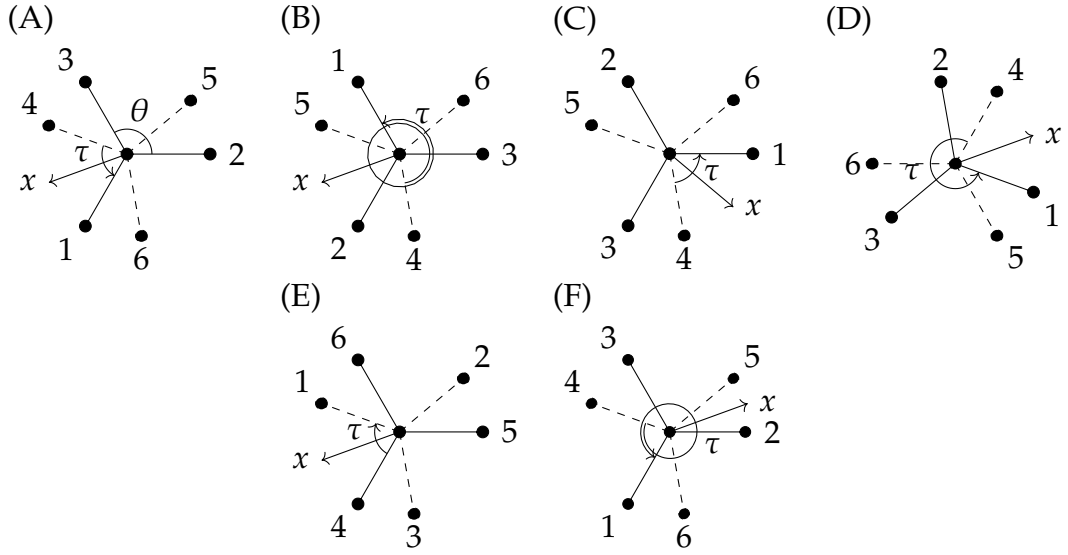



Figure 7.7: Projections of ethane showing the effects of the $G_{36}$(EM) generators (and their $G_{36}$ partners; see Table 7.5) on $\tau$ and $\chi$; the change of $\tau$ is represented by the curved arrow encircling the C–C axis and the change of $\chi$ is illustrated by the change in $x$-axis orientation. (A) Starting configuration with $\tau = 4\pi/9$ and the $x$-axis forming the clockwise angle $8\pi/9$ with the horizontal. (B) The effect of (123)(465) which causes $\tau$ to decrease by $4\pi/3$ (equivalent to an increase of $8\pi/3$ ) and $\chi$ to remain constant. (C) The effect of (132)(456) with $\tau$ remaining constant and $\chi$ changing by $+2\pi/3$. (D) The effect of (14)(26)(35)$(ab)^*$ with $\tau \rightarrow 2\pi - \tau$ and $\chi \rightarrow \chi + \pi$. (E) The effect of (14)(25)(36)$(ab)$ under which $\tau$ and $\chi$ are both invariant. (F) The effect of $E'$ which has no $G_{36}$ partner, $E'\tau = \tau + 2\pi$ and $E'\chi = \chi + \pi$.

We may now return to the rotational basis symmetrisation with a Sayvetz frame. Once again, $E'$ only changes $\tau$ by $\tau \rightarrow \tau + 2\pi$. The Euler angle $\chi$ changes by $\chi \rightarrow \chi + \pi$ with an associated equivalent rotation $R_z^{\pi}$. Due to the $\tau$ change, the equilibrium $a$s transform as $a \rightarrow M_z(\pi)a$; the frame rotation changes the $r_\alpha$s by $r_\alpha \rightarrow M_z(\pi)r_\alpha$. Applying $I = NN$, where $N = M_z(\pi)$, to the transformed condition then results in $M_z(\pi)$ applied to the original condition and is thus zero. The Sayvetz condition again works the same way. Therefore we may use the Sayvetz frame with a PES bisector to symmetrise the rotational basis set.

As for the geometric frame, it has the $x$-axis pointing at an angle

$$\phi = \frac{(\gamma_2 + \delta_2)}{3\sqrt{2}} \tag{7.5.29}$$

from the bisector between $H_1$–$H_4$. See Figure 7.8. The proof that this frame can be used to symmetrise a rotational basis set is given in Appendix 3 of [4].

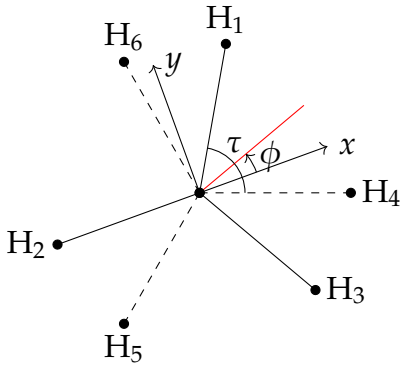



Figure 7.8: The geometric frame for $C_2H_6$. The $z$-axis is parallel to the C–C bond. The $x$-axis is at an angle $\phi = (\gamma_{15} + \delta_{17})/(3\sqrt{2})$ from the bisector of the dihedral angle between the planes formed by C–C–$H_1$ and C–C–$H_4$ which is denoted by the red line.

### 7.5.3 Rigid-symmetric-rotor function representations

As explained in Section 2.6.4, TROVE uses rigid-symmetric-rotor eigenfunctions $|J, k, m\rangle$ [31] as primitive rotational basis functions. For $K \neq 0$ (where $K = |k|$), the corresponding symmetrised functions are defined as

$$|JKm\eta\rangle = \frac{i^\eta(-1)^\sigma}{\sqrt{2}}(|J, k, m\rangle + (-1)^{J+K+\eta}\,|J, -k, m\rangle) \tag{7.5.30}$$

where

$$\begin{cases} \sigma = K \bmod 3 & \eta = 1 \\ \sigma = 0 & \eta = 0, \end{cases} \tag{7.5.31}$$

while for $K = 0$ it is $|J00\eta\rangle = i^\eta\,|J, 0, 0\rangle$ where $\eta = J \bmod 2$. With the rigid rotor transformation properties given in Equations (12-46) and (12-47) of [31], along with Tables 12-1 and 15-6 of [31], we can determine the irreps of the rigid rotor basis functions by using the characters of the transformation; the results are summarised in Table 7.6. While these are the correct irreps up to equivalence, the functions as defined in Eq. (7.5.30) do not have the correct transformation properties using the choice made in Eq. (7.4.3). In particular, for $K = 3n + 1$ and $K$ odd the transformation for operation $g = (14)(25)(36)(ab)$ is

$$O_g\begin{pmatrix} |JKm0\rangle \\ |JKm1\rangle \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} |JKm0\rangle \\ |JKm1\rangle \end{pmatrix} \tag{7.5.32}$$

which is an equivalent representation of the operation.

The function combinations in Eq. (7.5.30) are desirable because their associated Hamiltonian matrix elements are real; TROVE assumes real rotational matrix elements. To continue using these combinations, the representation of Eq. (7.5.32) is used in $\boldsymbol{C}_{3v}^{(-)}$, that is, the irreps of $\boldsymbol{C}_{3v}^{(-)}$ are $(A_1, A_2, \tilde{E})$ where the matrix for the $\tilde{E}$ representation of $(14)(25)(36)(ab)$ is the reflection matrix of Eq. (7.5.32). The representations of $\boldsymbol{C}_{3v}^{(+)}$ are as provided in Section 7.3.

Table 7.6: The irreps of the rigid rotor wavefunctions. For $K > 0$, the two functions with $\eta = 0, 1$ generate a two-dimensional irrep or the direct sum of two one-dimensional irreps. For a given $K$ value, we list first the $\eta = 0$ function and then the $\eta = 1$ one. $n$ is a positive integer.

| $K$ | $\Gamma$ |
|---|---|
| 0 ($J$ even) | $A_{1s}$ |
| 0 ($J$ odd) | $A_{2s}$ |
| $K = 3n$ ($K$ even) | $A_{1s} \oplus A_{2s}$ |
| $K = 3n$ ($K$ odd) | $A_{4d} \oplus A_{3d}$ |
| $K = 3n + 1$ ($K$ even) | $E_{1s}$ |
| $K = 3n + 1$ ($K$ odd) | $E_{2d}$ |
| $K = 3n + 2$ ($K$ even) | $E_{1s}$ |
| $K = 3n + 2$ ($K$ odd) | $E_{2d}$ |

## 7.6 PES coordinates

Originally, the PES was represented in the coordinates

$$\xi_1 = 1 - \exp\left(-a(\mathcal{R}^{\mathrm{C}} - \mathcal{R}^{\mathrm{C}}_{\mathrm{e}})\right) \tag{7.6.1}$$

$$\xi_j = 1 - \exp\left(-b(\mathcal{R}_i - \mathcal{R}_{\mathrm{e}})\right)\ ;\ \ j = 2,3,4,5,6,7\,,\ \ i = j - 1, \tag{7.6.2}$$

($a$ is used for the C–C internal coordinate $\mathcal{R}^{\mathrm{C}}$, and $b$ is used for the six C–H internal coordinates $\mathcal{R}_1$, $\mathcal{R}_2$, $\mathcal{R}_3$, $\mathcal{R}_4$, $\mathcal{R}_5$ and $\mathcal{R}_6$) the bending angular coordinates

$$\xi_k = (\alpha_i - \alpha_{\mathrm{e}})\,;\ \ k = 8,9,10,11,12,13\,,\ \ i = k - 7, \tag{7.6.3}$$

the dihedral coordinates from Equation (7.5.1)

$$(\xi_{14}, \xi_{15}, \xi_{16}, \xi_{17}) = (\gamma_1, \gamma_2, \delta_1, \delta_2) \tag{7.6.4}$$

and, finally, the torsional term

$$\xi_{18} = 1 + \cos 3\tau \tag{7.6.5}$$

where $\tau$ is defined in Eq. (7.5.4). The quantities $g_{\mathrm{e}}$, $r_{\mathrm{e}}$, and $\alpha_{\mathrm{e}}$ are the equilibrium structural parameter values of $C_2H_6$. Unfortunately, a PES expanded in these coordinates does not correctly fit the *ab initio* points. As of writing this thesis, this is an ongoing problem and is discussed in the next chapter.

## 7.7 Class 5 numerical example

This section applies the symmetrisation procedure on a small test basis set for the torsional coordinate. For ethane, I modified the sampling procedure in TROVE by applying it to the five group generators only (out of 72). This simple modification led to

a significant speedup of the sampling part of the code and is now the standard part of TROVE.

Figure 7.9 shows the torsional potential energy as a function of $\tau$. To solve this Schrödinger equation in a Fourier-series basis, $\tau$ is let to vary from 0 to $4\pi$ so that the potential energy curve in Figure 7.9 has six minima. The figure indicates the lowest allowed torsional energies. Some of these energies are degenerate, i.e., associated with more than one eigenfunction of a given irreducible representation. In the limit of an infinite barrier height, the energies should be six-fold degenerate. With the actual, finite height of the barrier, the energies form near-degenerate clusters with a total multiplicity of 6 as shown in Figure 7.10. However, since the description of the torsional angle $\tau \in [0, 4\pi]$ is unphysical, only three of the states in the cluster (of $s$ symmetry) exist in nature if they are combined with rigid-rotor basis functions of $s$ symmetry. The other three states (of $d$ symmetry) must be combined with rigid-rotor basis functions of $d$ symmetry in order that the total rotation-torsion state can exist in nature. For $J = 0$, only the $s$-type torsional states will exist since only $s$-type rigid-rotor basis functions are available. The $d$-type ro-vibrational functions are disregarded in TROVE as they are assigned a nuclear spin statistical weight of zero. The $s$-type nuclear spin statistical weights are 6, 10, 6, 10, 4, 4, 2, 6, and 12 for the $\boldsymbol{G}_{36}$(EM) symmetries $A_{1s}$, $A_{2s}$, $A_{3s}$, $A_{4s}$, $E_{1s}$, $E_{2s}$, $E_{3s}$, $E_{4s}$, and $G_s$, respectively.

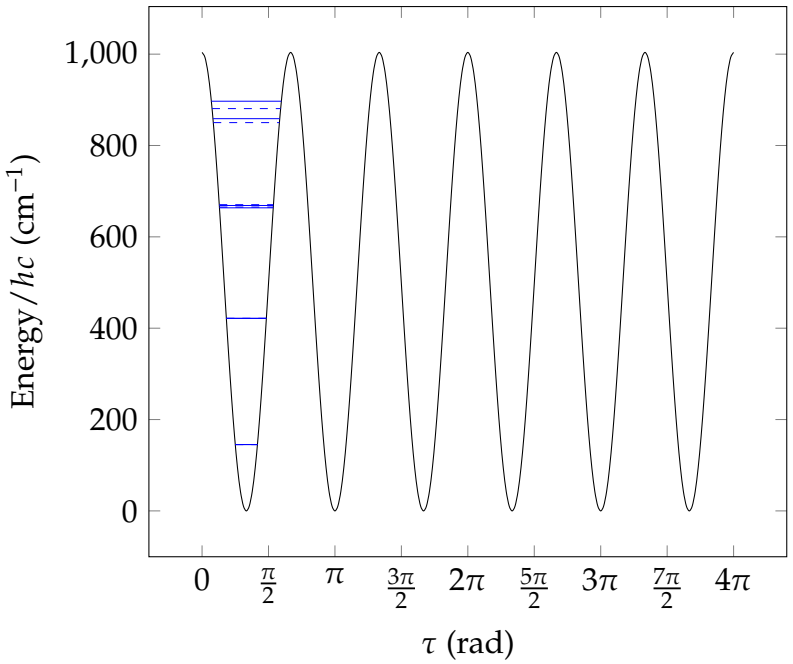



Figure 7.9: The torsional potential energy as a function of the torsion angle $\tau$. The allowed energy values are marked by blue horizontal lines. Each energy may correspond to more than one eigenfunction of a given irreducible representation. Dashed lines indicate states of $d$-type symmetry.

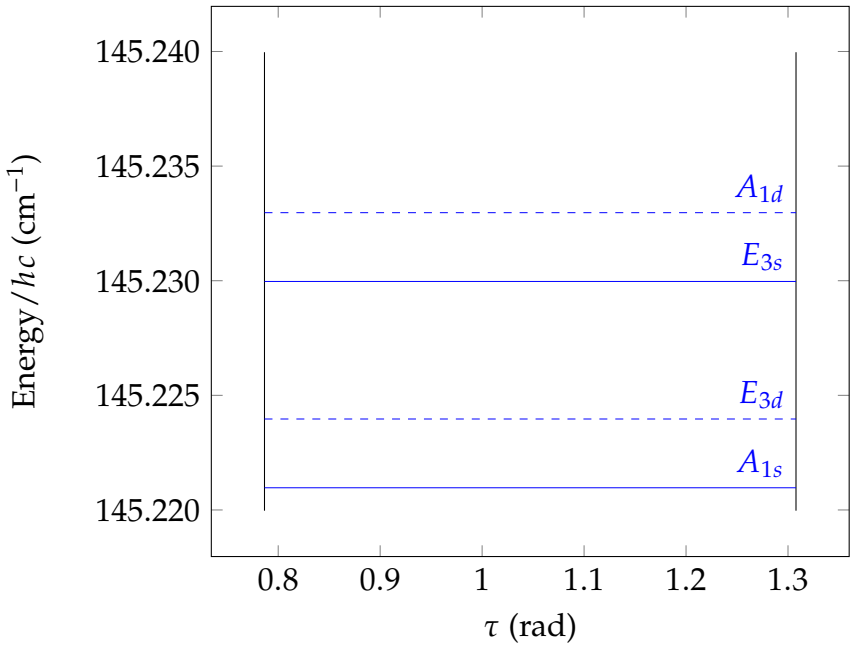



Figure 7.10: A enlarged detail of Figure 7.9, showing the lowest energy cluster with the $G_{36}$(EM) symmetry labels indicated.

## 7.8 Chapter summary

This chapter has presented a detailed description of the $G_{36}$ and $G_{36}$(EM) molecular symmetry groups. A full set of irreducible representation matrices have been derived and tested for constructing symmetry-adapted potential energy functions and basis functions of ethane $C_2H_6$. Both the construction of the transformation matrices and the symmetry adaption can be implemented as numerical procedures as part of computational approaches to the solution of the ro-vibrational Schrödinger equation. A self-consistent choice of the vibrational, torsional and rotational coordinates for ethane, satisfying the $G_{36}$(EM) symmetry requirements, has been introduced and analysed in full detail. The irreducible representation matrices as well as the coordinate choice made for ethane in this chapter have been implemented in TROVE. These results will be important in the future line-list calculation of ethane.

# Appendix 7.A Character Table of $C_{3v}^{(-)}$ and $C_{3v}^{(+)}$.

$G_{36}$ is the direct product of $C_{3v}^{(-)}$ and $C_{3v}^{(+)}$, $G_{36} = C_{3v}^{(-)} \times C_{3v}^{(+)}$, and so the irreps of $G_{36}$ are obtained from those of $C_{3v}^{(-)}$ and $C_{3v}^{(+)}$. These latter irreps are given in Table 7.7.

Table 7.7: Common character tables of $C_{3v}^{(-)}$ and $C_{3v}^{(+)}$.

| $\Gamma$ | $\mathcal{C}_1^{(\pm)}$ | $\mathcal{C}_2^{(\pm)}$ | $\mathcal{C}_3^{(\pm)}$ |
|---|---|---|---|
| | 1 | 2 | 3 |
| $A_1$ | 1 | 1 | 1 |
| $A_2$ | 1 | 1 | −1 |
| $E$ | 2 | −1 | 0 |

We label the elements of $C_{3v}^{(\pm)}$ as $g_j^{(\pm)}$, $j = 1, 2, \ldots, 6$ (see Table 7.1). The nuclei are labelled as in Figure 7.2. $C_{3v}^{(\pm)}$ has three classes, $\mathcal{C}_1^{(\pm)} = \{E\} = \{g_1^{(\pm)}\}$, $\mathcal{C}_2^{(\pm)} = \{g_2^{(\pm)}, g_3^{(\pm)}\}$, and $\mathcal{C}_3^{(\pm)} = \{g_4^{(\pm)}, g_5^{(\pm)}, g_6^{(\pm)}\}$.

# Appendix 7.B Transformation of $\tau$

The torsional coordinate $\tau$ is defined by Equation (7.5.4):

$$\tau = \frac{1}{3}(\tau_{41} + \tau_{62} + \tau_{53}), \tag{7.B.1}$$

and we investigate how this coordinate transforms under the generating operations of $G_{36}$, $g_2^{(+)} = (123)(456)$, $g_2^{(-)} = (132)(456)$, $g_4^{(+)} = (14)(26)(35)(ab)^*$, and $g_4^{(-)} = (14)(25)(36)(ab)$ used in the TROVE calculations. The transformation properties of the dihedral angles $\tau_{ij}$ are derived as outlined in Appendix 2.A.

The operation (123)(456) permutes the protons 1, 2, 3, 4, 5, and 6 to the positions previously occupied by the protons labelled 3, 1, 2, 6, 4, and 5, and the transformed value of $\tau$ is given by

$$\tau' = \frac{1}{3}(\tau_{63} + \tau_{51} + \tau_{42}) \tag{7.B.2}$$

where

$$\tau_{63} = \tau_{62} + \theta_{23}, \tag{7.B.3}$$
$$\tau_{51} = \tau_{53} + \theta_{31}, \tag{7.B.4}$$
$$\tau_{42} = \tau_{41} + \theta_{12}. \tag{7.B.5}$$

For example, $\tau_{51} = \tau_{53} + \theta_{31}$, the plus sign coming about because the positive direction of rotation for the $\tau_{ij}$ dihedral angles (proton $1 \rightarrow 2 \rightarrow 3$) is the same as that of $\theta_{12}$, $\theta_{23}$, and $\theta_{31}$. Hence

$$\tau' = \frac{1}{3}\left(\tau_{41} + \tau_{62} + \tau_{53} + [\theta_{12} + \theta_{23} + \theta_{31}]\right) \tag{7.B.6}$$
$$= \frac{1}{3}\left(\tau_{41} + \tau_{62} + \tau_{53} + 2\pi\right) = \tau + 2\pi/3 \tag{7.B.7}$$

or, equivalently, $\tau' = \tau - 4\pi/3$ as given in Table 7.5, to ensure that $(123)^3(456)^3 = E$.

After carrying out the operation (132)(456), the protons 1, 2, 3, 4, 5, and 6 are found at the positions previously occupied by the protons labelled 2, 3, 1, 6, 4, and 5, respectively. Consequently, the transformed value of $\tau$ is given by

$$\tau' = \frac{1}{3}(\tau_{62} + \tau_{53} + \tau_{41}) = \tau. \tag{7.B.8}$$

For (14)(26)(35)$(ab)^*$,

$$\tau' = \frac{1}{3}(\tau_{14} + \tau_{26} + \tau_{35}) \tag{7.B.9}$$

$$= \frac{1}{3}(-\tau_{41} - \tau_{62} - \tau_{53}) = -\tau \tag{7.B.10}$$

or, equivalently, $2\pi - \tau$ as given in Table 7.5.

Finally, for (14)(25)(36)$(ab)$,

$$\tau' = \frac{1}{3}(-\tau_{14} - \tau_{26} - \tau_{35}) \tag{7.B.11}$$

$$= \frac{1}{3}(\tau_{41} + \tau_{62} + \tau_{53}) = \tau. \tag{7.B.12}$$

The transformation properties derived here are summarised in Table 7.5.

# Chapter 8

# Summary and Future Work

## 8.1 Summary

The two main focal points in this thesis were an analytic KEO approach and the exploitation of symmetry groups, both of which facilitated ro-vibrational line list calculations in TROVE.

Chapter 3 explored an analytic representation of the KEO, with a focus on valence coordinates. On the theoretical side, I found the fundamental functions present in the $s$-vectors expressed in valence coordinates. This also established the necessary conditions for the KEO to maintain a sum-of-product form, a prerequisite for TROVE calculations and a highly desirable property in general. Moreover, I derived the relationship between $s$-vector components in different BF-frames. This is a generalisation of a previous result of Ref. [66] which applied to the $G$-matrix. The advantage in relating the $s$-vector components is that they have a more compact structure than the $G$ matrix and therefore are easier to transform. Finally, the flexibility of the Sørensen method was demonstrated from the derivation of the $s$-vectors expressed in polyspherical coordinates. Here again, the more compact $s$-vector representation is more favourable to work with.

On the practical side, I wrote a general *Mathematica* script that generates analytic KEOs. With it, relatively few input changes are needed for each molecule. I also modified TROVE such that it could read the *Mathematica* output in multiple formats and perform the calculations. The first (and worst) approach was hard-coding the 1D functions present in the KEO. I also developed an on-the-fly procedure to read the 1D functions in the TROVE input. In both formats, the numbering of the functions was based on and exploited TROVE's bijection between a $3N-6$ dimensional index and a single index. For larger molecules this was no longer tenable and I modified TROVE so that it would read the full $3N-6$ dimensional index.

Utilising the results of the previous chapter, Chapter 4 detailed the line list calculation of $H_2CS$. This built upon the previous work of Ref. [104] by using the *ab initio* PES of Ref. [105]. The $H_2CS$ Hamiltonian was expressed in valence coordinates and employed an analytic KEO derived with the aforementioned *Mathematica* script. The accuracy of the PES was insufficient so this was refined by Sergey Yurchenko. The refinement parameters taken from the results of the $H_2CS$ MARVEL project. Here, I collated all experimentally available transitions and used the MARVEL program to convert it to a set of 4254 experimental energy levels. The very close band centres between the $\nu_4$ and $\nu_6$

modes meant a large proportion of transitions involving those states had to be relabelled to ensure consistency. Using a refined PES and analytic KEO proved worthwhile as the resulting line list was of high quality. The line list covered the $0\,\mathrm{cm}^{-1}$ to $8000\,\mathrm{cm}^{-1}$ range for states up to $J = 120$ was calculated. By MARVELising this line list, the calculated ro-vibrational energy levels were replaced with MARVEL energy levels where possible.

Chapter 5 was a change of pace where I described a group whose motivation was divorced from physical reality but nevertheless still proved useful for TROVE's $3N - 6$ implementation for linear molecules. Here, I created (and programmed in TROVE) an artificial group in such a way to set the condition $k = \ell$ neatly within TROVE's pre-existing infrastructure. It is an unconventional way to use groups, not being motivated by the physics but rather for as a convenient solution to a practical issue.

For $C_2H_6$, I eventually realised that a bisector frame could not be used to symmertrise a rotational basis set. In Chapter 6 I provided a systematic approach to the transformation of the coordinates and the frame under different BF-frames. This resulted in a condition which dictated whether rotational-symmetrisation is possible. Using the method described to find symmetrisable frames, I defined geometric frames for the $CH_3Cl$ and $C_2H_6$ that were not Eckart and Sayvetz frames. This knowledge may hopefully prove useful for related molecules and be expanded for others. It also suggests a possible link between symmetry-separated rotational and vibrational coordinates and the magnitude of the coupling between them. This would have to be studied further.

The final chapter, and the work for $C_2H_6$ in general, combines much of the knowledge from the previous chapters. Being a much larger molecule, the disadvantage of Taylor expanding a KEO begins to rear its head. Practically, it is only possible to expand the KEO up to second order. The value in using an analytic KEO then becomes clear, both due to its accuracy but also because it is much more compact. A KEO in valence coordinates is also ideally suited for a non-rigid molecule such as $C_2H_6$. The requirement of a (non-bisector) geometric frame for rotational-symmetrisation also makes use of the $s$-vector relations between frames of Chapter 3 as calculating the KEO in the final frame by using the (complicated) rotational conditions is far more difficult than using a bond-vector frame and applying a subsequent frame rotation and transforming the $s$-vectors accordingly.

Chapter 7 also made substantial use of the group theory of Chapter 2. Previously it was found that $\boldsymbol{G}_{36}$ can be written as a direct product of two $\boldsymbol{C}_{3v}$ groups and there are only four generators. This heavily reduced the workload required in the implementation of $\boldsymbol{G}_{36}$and its manifestation in $C_2H_6$ in TROVE. I used this structure for the generation of the irreps and the transformation of the coordinates: only the transformations for the generator operations were programmed. I also exploited it in TROVE's symmetrisation step where the sampling procedure, used to determine the representation matrices, was only applied to the generators. In all cases, the remaining transformations/matrices were derived from the $\boldsymbol{G}_{36}$ product table.

The extension to the $\boldsymbol{G}_{36}$(EM) group and the use of the $\tau$ dihedral angle had a few subtleties that stalled progress. First, I did not immediately understand that there is an ambiguity in the transformation of $\tau$ using the method of Appendix 7.B . For all transformations, there are two possible choices. Transformations with an order of 3 only had one choice which respected the group properties, and must be used. For transformations with an order of two, there is freedom, but either choice locks in a

specific transformation of the rotational coordinates. They must be internally consistent and satisfy the rules laid out in Section 6.2. Section 7.5.2 debunks the notation that one may use a bond-vector frame attached to one of the $CH_3$ groups in equilibrium, even for the Sayvetz frame. Finally, the particularly-defined rigid-rotor functions of Section 7.5.3 did not transform as the irreps programmed in TROVE but rather by an equivalent representation. The irreps for the $C_{3v}$ groups needed to be modified to respect this.

In summary, while there is much work to do with the line list calculation of $C_2H_6$, a lot of groundwork has already been laid.

## 8.2 Future work

Looking further, there are currently two issues. The first relates to the analytic representation, and the second more general. For the former, the externally calculated analytic KEO is not presently compatible with the PES internally expanded in TROVE. Specifically, for basis sets beyond the minimal size, Eq. (2.4.1) is non-zero between certain symmetries. The magnitude of the error is of the order of $1\,\mathrm{cm}^{-1}$ and does not appear to grow with the size of the basis set. The number of non-zero terms also is much smaller than the number of basis functions. The final vibrational energies are comparable to those calculated with the Taylor expanded KEO, but of course the final calculations should not have these errors.

The source of the error is not clear. It is fairly difficult to match precisely the coordinate definition in TROVE to those in *Mathematica*, specifically the dihedral coordinates, owning to the slightly different way they are defined in both. In TROVE, a Z-matrix is used to define the initial vibrational coordinates of each atom, which does use the equation for the dihedrals in terms of the Cartesian coordinates. The equation for the dihedrals has a "sign" attached to it in that the angle from a first plane to a second plane is negative the angle from the second to the first. The symmetry adapted coordinates in TROVE, which are linear combinations of dihedrals, are just algebraic expressions of the coordinate variables. In fact, the Z-matrix dihedrals (of which there are five) are not the un-symmetrised dihedrals, of which there are nine. The latter are also linear combinations of the former.

The analytic KEO must have the symmertised dihedrals expressed as equations of the Cartesian coordinates. The symmetrised variables must also be expressed in terms of the un-symmetrised. The two must be internally compatible but they must also be compatible with TROVE. There are many correct ways for compatibility within *Mathematica*, but they will not necessarily be compatible with TROVE. These incompatibilities boil down to sign differences and are hard to pinpoint, although this may not be the cause at all.

The KEO which worked the furthest was tested by equating the KEO at a random point and then comparing the transformed KEO under MS operations, using the transformation properties of the coordinates programmed in TROVE, and then was equated at the same point. The result was an invariant Hamiltonian. Nevertheless, the issue remains.

The other more general problem concerns fitting of the KEO. As explained in

Section 2.6.6, the functional form of the *ab initio* PES must be invariant under MS operations. The highly symmetric nature of $C_2H_6$ invariably means that many of the expansion parameters are identical. In general, the more symmetric (and therefore fewer expansion parameters for a given expansion order), the higher the expansion order must be to fully fit the *ab initio* energies while maintaining the symmetry. Related to this is the torsional coordinate. It was mentioned that the coordinate expansions of Eq. (7.6) are not sufficient. One example of this is the second order bend expansion, modulated by a function of the torsional coordinate. The torsional function $\cos 3\tau$ does not respect the symmetry. Instead terms of the form

$$\begin{aligned} &\cos\tau(\alpha_1\alpha_4 + \alpha_2\alpha_6 + \alpha_3\alpha_5) \\ &+ \cos(\tau - 4\pi/3)(\alpha_3\alpha_6 + \alpha_1\alpha_5 + \alpha_2\alpha_4) \\ &+ \cos(\tau + 4\pi/3)(\alpha_2\alpha_5 + \alpha_3\alpha_4 + \alpha_1\alpha_5) \end{aligned} \tag{8.2.1}$$

are needed. In general, combinations of $\cos\tau$, $\sin\tau$, $\cos 3\tau$, and $\sin 3\tau$ will be present. An adequate model is a work in progress.

For the experimental side of $C_2H_6$, a MARVEL project has started, with help from Mikhail Semenov and Jaya Chand, but this has paused after the data was compiled as the states are assigned in the $\boldsymbol{D}_{3h}$ group. There is no one-to-one correlation from $\boldsymbol{D}_{3h}$ to $\boldsymbol{G}_{36}$. Accurate calculated energies in the $\boldsymbol{G}_{36}$ group are required to match to the experimental energies and thus reassign the latter.

We hope that in surmounting the PES fitting problem, and, ideally, in ensuring the analytic KEO is compatible with TROVE, that proper ro-vibrational calculations for $C_2H_6$ can commence. In reality, other obstacles probably remain. For example, because of the sheer number of vibrational modes and subsequent basis set size, more effective basis set pruning techniques will likely be required.

# Bibliography


[1] T. M. Mellor et al. "MARVEL analysis of high-resolution spectra of thioformaldehyde ($H_2CS$)". In: *J. Mol. Spectrosc.* 391 (2023), p. 111732. DOI: 10.1016/j.jms.2022.111732.

[2] T. Mellor et al. "ExoMol line lists – XLVIII. High-temperature line list of thioformaldehyde ($H_2CS$)". In: *Mon. Not. R. Astron. Soc.* (2023). DOI: 10.1093/mnras/stad111.

[3] T. M. Mellor, S. N. Yurchenko, and P. Jensen. "Artificial Symmetries for Calculating Vibrational Energies of Linear Molecules". In: *Symmetry* 13.4 (2021). DOI: 10.3390/sym13040548.

[4] T. M. Mellor. "Molecular frames for a symmetry-adapted rotational basis set". In: *Mol. Phys.* 120.18 (2022), e2118638. DOI: 10.1080/00268976.2022.2118638.

[5] T. M. Mellor et al. "Transformation Properties under the Operations of the Molecular Symmetry Groups $\boldsymbol{G}_{36}$ and $\boldsymbol{G}_{36}$(EM) of Ethane $H_3CCH_3$". In: *Symmetry* 11.862 (2019). DOI: 10.3390/sym11070862.

[6] S. N. Yurchenko et al. "ExoMol line lists – XXXIX. Ro-vibrational molecular line list for $CO_2$". In: *Mon. Not. R. Astron. Soc.* 496.4 (2020), pp. 5282–5291. DOI: 10.1093/mnras/staa1874.

[7] S. N. Yurchenko and T. M. Mellor. "Treating linear molecules in calculations of rotation-vibration spectra". In: *J. Chem. Phys.* 153.15 (2020), p. 154106. DOI: 10.1063/5.0019546.

[8] J. Tennyson et al. "The 2020 release of the ExoMol database: Molecular line lists for exoplanet and other hot atmospheres". In: *J. Quant. Spectrosc. Radiat. Transf.* 255 (2020), p. 107228. DOI: 10.1016/j.jqsrt.2020.107228.

[9] W. Borucki et al. "KEPLER: Search for Earth-Size Planets in the Habitable Zone". In: *Proc. Int. Astron. Union* 4.S253 (2008), pp. 289–299. DOI: 10.1017/S1743921308026513.

[10] R. L. Akeson et al. "The NASA Exoplanet Archive: Data and Tools for Exoplanet Research". In: *Publ. Astron. Soc. Pac.* 125.930 (2013), p. 989. DOI: 10.1086/672273.

[11] N. Madhusudhan. "Atmospheric Retrieval of Exoplanets". In: *Handbook of Exoplanets*. Chan: Springer International Publishing, 2018, pp. 1–30. ISBN: 978-3-319-30648-3. DOI: 10.1007/978-3-319-30648-3_104-1.

[12] J. Tennyson and S. N. Yurchenko. "The ExoMol project: Software for computing large molecular line lists". In: *Intern. J. Quantum Chem.* 117.2 (2017), pp. 92–103. DOI: 10.1002/qua.25190.

[13] M. Born and R. Oppenheimer. "Zur Quantentheorie der Molekeln". In: *Ann. Phys.* 389.20 (1927), pp. 457–484. DOI: 10.1002/andp.19273892002.

[14] R. J. Bartlett and M. Musiał. "Coupled-cluster theory in quantum chemistry". In: *Rev. Mod. Phys.* 79.1 (2007), pp. 291–352. DOI: 10.1103/RevModPhys.79.291.

[15] C. David Sherrill and H. F. Schaefer III. "The Configuration Interaction Method: Advances in Highly Correlated Approaches". In: *Adv. Quantum Chem.* 34 (1999), pp. 143–269. DOI: 10.1016/S0065-3276(08)60532-8.

[16] T. Shiozaki, G. Knizia, and H.-J. Werner. "Explicitly correlated multireference configuration interaction: MRCI-F12". In: *J. Chem. Phys.* 134.3 (2011), p. 034113. DOI: 10.1063/1.3528720.

[17] W. Kutzelnigg. "How many-body perturbation theory (MBPT) has changed quantum chemistry". In: *Intern. J. Quantum Chem.* 109.15 (2009), pp. 3858–3884. DOI: 10.1002/qua.22384.

[18] D. R. Hartree. "The Wave Mechanics of an Atom with a Non-Coulomb Central Field. Part I. Theory and Methods". In: *Math. Proc. Camb. Philos. Soc* 24.1 (1928), pp. 89–110. DOI: 10.1017/S0305004100011919.

[19] V. Fock. "Näherungsmethode zur Lösung des quantenmechanischen Mehrkörperproblems". In: *Z. Phys* 61.1-2 (1930), pp. 126–148. DOI: 10.1007/BF01340294.

[20] A. Einstein. "Strahlungs-Emission und Absorption nach der Quantentheorie". In: *Phys. J.* 18 (1916), pp. 318–323.

[21] R. Shankar. *Principles of quantum mechanics*. New York, NY: Plenum, 1980.

[22] D. J. Griffiths. *Introduction to electrodynamics; 4th ed.* Re-published by Cambridge University Press in 2017. Boston, MA: Pearson, 2013. DOI: 1108420419.

[23] P. F. Bernath. "MoLLIST: Molecular Line Lists, Intensities and Spectra". In: *J. Quant. Spectrosc. Radiat. Transf.* 240 (2020), p. 106687. DOI: 10.1016/j.jqsrt.2019.106687.

[24] S. Chandrasekhar. *Radiative transfer*. Dover Publications, 2016.

[25] *ISO 9288:1989*. 2022.

[26] H. Goldstein, C. P. Poole, and J. Safko. *Classical mechanics*. Addison Wesley, 2002.

[27] A. Nauts and X. Chapuisat. "Momentum, quasi-momentum and hamiltonian operators in terms of arbitrary curvilinear coordinates, with special emphasis on molecular hamiltonians". In: *Mol. Phys.* 55.6 (1985), pp. 1287–1318. DOI: 10.1080/00268978500102031.

[28] G. O. Sørensen. "A new approach to the Hamiltonian of nonrigid molecules". In: *Large Amplitude Motion in Molecules II*. Vol. 82. Topics in Current Chemistry. Heidelberg: Springer Berlin Heidelberg, 1979, pp. 97–175.

[29] B. Steinberg. *Representation theory of finite groups: An introductory approach*. Springer Science &; Business Media, 2011.

[30] H. C. Longuet-Higgins. "The symmetry groups of non-rigid molecules". In: *Mol. Phys.* 6.5 (1963), pp. 445–460. DOI: 10.1080/00268976300100501.

[31] P. R. Bunker and P. Jensen. *Molecular Symmetry and Spectroscopy*. 2nd ed. Ottawa: NRC Research Press, 1998.

[32] R. N. Zare. *Angular Momentum: Understanding Spatial Aspects in Chemistry and Physics*. 1st. Wiley, 1988.

[33] W. Pauli. "The Connection Between Spin and Statistics". In: *Phys. Rev.* 58.8 (1940), pp. 716–722. DOI: 10.1103/PhysRev.58.716.

[34] S. N. Yurchenko, W. Thiel, and P. Jensen. "Theoretical ROVibrational Energies (TROVE): A robust numerical approach to the calculation of rovibrational energies for polyatomic molecules". In: *J. Mol. Spectrosc.* 245.2 (2007), pp. 126–140. DOI: 10.1016/j.jms.2007.07.009.

[35] A. Yachmenev and S. N. Yurchenko. "Automatic differentiation method for numerical construction of the rotational-vibrational Hamiltonian as a power series in the curvilinear internal coordinates using the Eckart frame". In: *J. Chem. Phys.* 143 (2015), p. 014105. DOI: 10.1063/1.4923039.

[36] S. N. Yurchenko et al. "A Variationally Computed 300 K Line List for $NH_3$". In: *J. Phys. Chem. A* 113 (2009), pp. 11845–11855. DOI: 10.1021/jp9029425.

[37] A. Yachmenev et al. "A new "spectroscopic" potential energy surface for formaldehyde in its ground electronic state". In: *J. Chem. Phys.* 134 (2011), p. 244307. DOI: 10.1063/1.3599927.

[38] C. Sousa-Silva et al. "High temperature partition functions and thermodynamic data for ammonia and phosphine". In: *J. Quant. Spectrosc. Radiat. Transf.* 142 (2014), pp. 66–74. DOI: 10.1016/j.jqsrt.2014.03.012.

[39] C. Sousa-Silva et al. "ExoMol line lists - VII. The rotation-vibration spectrum of phosphine up to 1500 K". In: *Mon. Not. R. Astron. Soc.* 446 (2015), pp. 2337–2347. DOI: 10.1093/mnras/stu2246.

[40] D. S. Underwood et al. "Rotational spectrum of $SO_3$ and theoretical evidence for the formation of sixfold rotational energy-level clusters in its vibrational ground state". In: *J. Chem. Phys.* 140.24 (2014), p. 244316. DOI: 10.1063/1.4882865.

[41] A. F. Al-Refaie et al. "ExoMol line lists – VIII. A variationally computed line list for hot formaldehyde". In: *Mon. Not. R. Astron. Soc.* 448.2 (2015), pp. 1704–1714. DOI: 10.1093/mnras/stv091.

[42] S. N. Yurchenko and J. Tennyson. "ExoMol line lists - IV. The rotation-vibration spectrum of methane up to 1500 K". In: *Mon. Not. R. Astron. Soc.* 440.2 (2014), pp. 1649–1661. DOI: 10.1093/mnras/stu326.

[43] A. F. Al-Refaie et al. "A variationally calculated room temperature line-list for $H_2O_2$". In: *J. Mol. Spectrosc.* 318 (2015), pp. 84–90. DOI: 10.1016/j.jms.2015.10.004.

[44] A. Owens et al. "Accurate *ab initio* vibrational energies of methyl chloride". In: *J. Chem. Phys.* 142 (2015), p. 244306. DOI: 10.1063/1.4922890.

[45] A. Owens et al. "A global potential energy surface and dipole moment surface for silane". In: *J. Chem. Phys.* 143.24 (2015), p. 244317. DOI: 10.1063/1.4938563.

[46] A. Y. Adam et al. "Ro-vibrational averaging of the isotropic hyperfine coupling constant for the methyl radical". In: *J. Chem. Phys.* 143.24 (2015), p. 244306. DOI: 10.1063/1.4938253.

[47] A. Owens et al. "Accurate prediction of the ammonia probes of a variable proton-to-electron mass ratio". In: *Mon. Not. R. Astron. Soc.* 450.3 (2015), pp. 3191–3200. DOI: 10.1093/mnras/stv869.

[48] A. F. Al-Refaie et al. "ExoMol line lists - XV. A new hot line list for hydrogen peroxide". In: *Mon. Not. R. Astron. Soc.* 461.1 (2016), pp. 1012–1022. DOI: 10.1093/mnras/stw1295.

[49] D. S. Underwood et al. "ExoMol molecular line lists - XIV. The rotation-vibration spectrum of hot $SO_2$". In: *Mon. Not. R. Astron. Soc.* 459.4 (2016), pp. 3890–3899. DOI: 10.1093/mnras/stw849.

[50] A. Owens et al. "A global *ab initio* dipole moment surface for methyl chloride". In: *J. Quant. Spectrosc. Radiat. Transf.* 184 (2016), pp. 100–110. DOI: 10.1016/j.jqsrt.2016.06.037.

[51] A. Owens et al. "A highly accurate *ab initio* potential energy surface for methane". In: *J. Chem. Phys.* 145.10 (2016), p. 104305. DOI: http://dx.doi.org/10.1063/1.4962261.

[52] A. Owens et al. "ExoMol line lists - XXIX. The rotation-vibration spectrum of methyl chloride up to 1200 K". In: *Mon. Not. R. Astron. Soc.* 479.3 (2018), pp. 3002–3010. DOI: 10.1093/mnras/sty1542.

[53] B. P. Mant et al. "ExoMol molecular line lists - XXVII. Spectra of $C_2H_4$". In: *Mon. Not. R. Astron. Soc.* 478.3 (2018), pp. 3220–3232. DOI: 10.1093/mnras/sty1239.

[54] P. A. Coles et al. "A variationally computed room temperature line list for $AsH_3$". In: *Phys. Chem. Chem. Phys.* 21 (2019), pp. 3264–3277. DOI: 10.1039/C8CP07110A.

[55] B. P. Mant et al. "The infrared spectrum of $PF_3$ and analysis of rotational energy clustering effect". In: *Mol. Phys.* 118.1 (2019), e1581951. DOI: 10.1080/00268976.2019.1581951.

[56] S. N. Yurchenko, A. Yachmenev, and R. I. Ovsyannikov. "Symmetry Adapted Ro-vibrational Basis Functions for Variational Nuclear Motion: TROVE Approach". In: *J. Chem. Theory Comput.* 13.9 (2017), pp. 4368–4381. DOI: 10.1021/acs.jctc.7b00506.

[57] R. Penrose. "A generalized inverse for matrices". In: *Math. Proc. Camb. Philos. Soc.* 51.3 (1955), pp. 406–413. DOI: 10.1017/S0305004100030401.

[58] S. N. Yurchenko et al. "Towards efficient refinement of molecular potential energy surfaces: Ammonia as a case study". In: *J. Mol. Spectrosc.* 268.1-2 (2011), pp. 123–129. DOI: 10.1016/j.jms.2011.04.005.

[59] R. P. Feynman. "Forces in Molecules". In: *Phys. Rev.* 56.4 (1939), pp. 340–343. DOI: 10.1103/PhysRev.56.340.

[60] J. K. G. Watson. "Simplification of the molecular vibration-rotation hamiltonian". In: *Mol. Phys.* 15.5 (1968), pp. 479–490. DOI: 10.1080/00268976800101381.

[61] A. G. Császár and N. C. Handy. "Exact quantum mechanical vibrational kinetic energy operator of sequentially bonded molecules in valence internal coordinates". In: *J. Chem. Phys.* 102.10 (1995), pp. 3962–3967. DOI: 10.1063/1.468524.

[62] B. T. Sutcliffe and J. Tennyson. "A general treatment of vibration-rotation coordinates for triatomic molecules". In: *Intern. J. Quantum Chem.* 39.2 (1991), pp. 183–196. DOI: 10.1002/qua.560390208.

[63] N. C. Handy. "The derivation of vibration-rotation kinetic-energy operators, in internal coordinates". In: *Mol. Phys.* 61.1 (1987), pp. 207–223. DOI: 10.1080/00268978700101081.

[64] H. Meyer. "The Molecular Hamiltonian". In: *Annu. Rev. Phys. Chem.* 53.1 (2002), pp. 141–172. DOI: 10.1146/annurev.physchem.53.082201.124330.

[65] F. Gatti and C. Iung. "Exact and constrained kinetic energy operators for polyatomic molecules: The polyspherical approach". In: *Phys. Rep.* 484.1-2 (2009), pp. 1–69. DOI: 10.1016/j.physrep.2009.05.003.

[66] K. L. Mardis and E. L. Sibert III. "Derivation of rotation-vibration Hamiltonians that satisfy the Casimir condition". In: *J. Chem. Phys.* 106.16 (1997), pp. 6618–6621. DOI: 10.1063/1.473658.

[67] M. J. Bramley, W. H. Green Jr., and N. C. Handy. "Vibration-rotation coordinates and kinetic energy operators for polyatomic molecules". In: *Mol. Phys.* 73.6 (1991), pp. 1183–1208. DOI: 10.1080/00268979100101871.

[68] M. W. Sinclair et al. "Detection of Interstellar Thioformaldehyde". In: *Aust. J. Phys.* 26.1 (1973), pp. 85–92. DOI: 10.1071/ph730085.

[69] F. F. Gardner, J. B. Whiteoak, and B. Hoglund. "Observations of the 3-GHz transition of $H_2CS$". In: *Mon. Not. R. Astron. Soc.* 191.1 (1980), 19P–24P. DOI: 10.1093/mnras/191.1.19P.

[70] S. Doddipatla et al. "A chemical dynamics study on the gas phase formation of thioformaldehyde ($H_2CS$) and its thiohydroxycarbene isomer (HCSH)". In: *Proc. Nat. Acad. Sci.* 117.37 (2020), pp. 22712–22719. DOI: 10.1073/pnas.2004881117.

[71] S. Martín et al. "Sulfur Chemistry and Isotopic Ratios in the Starburst Galaxy NGC 253". In: *Astrophys. J.* 620.1 (2005), pp. 210–216. DOI: 10.1086/426888.

[72] A. Heikkilä, L. E. B. Johansson, and H. Olofsson. "Molecular abundance variations in the Magellanic Clouds". In: *Astron. Astrophys.* 344 (1999), pp. 817–847.

[73] S. Muller et al. "Molecules at $z = 0.89$". In: *Astron. Astrophys.* 535 (2011), A103. DOI: 10.1051/0004-6361/201117096.

[74] G. Esplugues et al. "Gas phase Elemental abundances in Molecular cloudS (GEMS)". In: *Astron. Astrophys.* 662 (2022), A52. DOI: 10.1051/0004-6361/202142936.

[75] L. M. Woodney et al. "Sulfur Chemistry at Millimeter Wavelengths in C/Hale-Bopp". In: *Earth Moon Planets* 78.1/3 (1997), pp. 69–70. DOI: 10.1023/a:1006275412491.

[76] C. He et al. "Sulfur-driven haze formation in warm $CO_2$-rich exoplanet atmospheres". In: *Nat. Astron.* 4.10 (2020), pp. 986–993. DOI: 10.1038/s41550-020-1072-9.

[77] R. Hobbs et al. "Sulfur chemistry in the atmospheres of warm and hot Jupiters". In: *Mon. Not. R. Astron. Soc.* 506.3 (2021), pp. 3186–3204. DOI: 10.1093/mnras/stab1839.

[78] T. Furtenbacher, A. G. Császár, and J. Tennyson. "MARVEL: measured active rotational-vibrational energy levels". In: *J. Mol. Spectrosc.* 245.2 (2007), pp. 115–125. DOI: 10.1016/j.jms.2007.07.005.

[79] A. G. Császár et al. "An Active Database Approach to Complete Rotational-Vibrational Spectra of Small Molecules". In: *Annu. Rep. Comput. Chem.* 3 (2007), pp. 155–176. DOI: 10.1016/s1574-1400(07)03009-5.

[80] T. Furtenbacher and A. G. Császár. "The role of intensities in determining characteristics of spectroscopic networks". In: *J. Molec. Struct. (THEOCHEM)* 1009 (2012), pp. 123–129. DOI: 10.1016/j.molstruc.2011.10.057.

[81] R. Tóbiás et al. "Accurate empirical rovibrational energies and transitions of $H_2{}^{16}O$". In: *Phys. Chem. Chem. Phys.* 21.7 (2019), pp. 3473–3495. DOI: 10.1039/c8cp05169k.

[82] A. G. Császár and T. Furtenbacher. "Spectroscopic networks". In: *J. Mol. Spectrosc.* 266.2 (2011), pp. 99–103. DOI: 10.1016/j.jms.2011.03.031.

[83] T. Furtenbacher et al. "Simple molecules as complex systems". In: *Sci. Rep.* 4 (2014), p. 4654. DOI: 10.1038/srep04654.

[84] P. Árendás, T. Furtenbacher, and A. G. Császár. "On spectra of spectra". In: *J. Math. Chem.* 54.3 (2016), pp. 806–822. DOI: 10.1007/s10910-016-0591-1.

[85] J. Tennyson et al. "IUPAC critical Evaluation of the rotational-vibrational spectra of water vapor. Part II. Energy levels and transition wavenumbers for $HD^{16}O$, $HD^{17}O$, and $HD^{18}O$". In: *J. Quant. Spectrosc. Radiat. Transf.* 111.2 (2010), pp. 2160–2184. DOI: 10.1016/j.jqsrt.2010.06.012.

[86] J. Tennyson et al. "A database of water transitions from experiment and theory (IUPAC Technical Report)". In: *Pure Appl. Chem.* 86.1 (2014), pp. 71–83. DOI: 10.1515/pac-2014-5012.

[87] J. Johns and W. Olson. "The infrared spectrum of thioformaldehyde". In: *J. Mol. Spectrosc.* 39.3 (1971), pp. 479–505. DOI: 10.1016/0022-2852(71)90219-0.

[88] D. Mcnaughton and D. Bruget. "Far-Infrared and $\nu_2$ Vibration-Rotation Spectrum of Thioformaldehyde and Infrared Spectrum of Thioglyoxal". In: *J. Mol. Spectrosc.* 159.2 (1993), pp. 340–349. DOI: 10.1006/jmsp.1993.1132.

[89] J.-M. Flaud et al. "The first high-resolution analysis of the 10 µm absorption of thioformaldehyde". In: *J. Quant. Spectrosc. Radiat. Transf.* 109.6 (2008), pp. 995–1003. DOI: 10.1016/j.jqsrt.2007.11.004.

[90] D. R. Johnson and F. X. Powell. "Microwave Detection of Thioformaldehyde". In: *Science* 169.3946 (1970), pp. 679–680. DOI: 10.1126/science.169.3946.679.

[91] Y. Beers et al. "Millimeter wave spectrum of thioformaldehyde". In: *J. Mol. Spectrosc.* 44.3 (1972), pp. 553–557. DOI: 10.1016/0022-2852(72)90263-9.

[92] B. Fabricant, D. Krieger, and J. S. Muenter. "Molecular beam electric resonance study of formaldehyde, thioformaldehyde, and ketene". In: *J. Chem. Phys.* 67.4 (1977), pp. 1576–1586. DOI: 10.1063/1.434988.

[93] D. Bedwell and G. Duxbury. "Laser Stark spectroscopy of thioformaldehyde in the 10 µm region: The $\nu_3$, $\nu_4$, and the $\nu_6$ fundamentals". In: *J. Mol. Spectrosc.* 84.2 (1980), pp. 531–558. DOI: 10.1016/0022-2852(80)90042-9.

[94] P. Turner, L. Halonen, and I. Mills. "Fourier transform infrared spectra of $H_2CS$ and $D_2CS$". In: *J. Mol. Spectrosc.* 88.2 (1981), pp. 402–419. DOI: 10.1016/0022-2852(81)90190-9.

[95] D. J. Clouthier et al. "Sub-Doppler spectroscopy of thioformaldehyde: Excited state perturbations and evidence for rotation-induced vibrational mixing in the ground state". In: *J. Chem. Phys.* 101.9 (1994), pp. 7300–7310. DOI: 10.1063/1.468287.

[96] A. Maeda et al. "High-Frequency Rotational Spectrum of Thioformaldehyde, $H_2CS$, in the Ground Vibrational State". In: *Astron. J.* 176.2 (2008), pp. 543–550. DOI: 10.1086/528684.

[97] H. S. P. Müller et al. "Laboratory spectroscopic study of isotopic thioformaldehyde, $H_2CS$, and determination of its equilibrium structure". In: *Astron. Astrophys.* 621 (2019), A143. DOI: 10.1051/0004-6361/201834517.

[98] D. R. Johnson, F. X. Powell, and W. H. Kirchhoff. "Microwave spectrum, ground state structure, and dipole moment of thioformaldehyde". In: *J. Mol. Spectrosc.* 39.1 (1971), pp. 136–145. DOI: 10.1016/0022-2852(71)90284-0.

[99] G. Winnewisser et al. "Precision broadband Spectroscopy in the Terahertz Region". In: *J. Mol. Spectrosc.* 165.1 (1994), pp. 294–300. DOI: 10.1006/jmsp.1994.1132.

[100] D. T. Petkie et al. "A fast scan submillimeter spectroscopic technique". In: *Rev. Sci. Instrum.* 68.4 (1997), pp. 1675–1683. DOI: 10.1063/1.1147970.

[101] C. M. Western. "PGOPHER: A program for simulating rotational, vibrational and electronic spectra ". In: *J. Quant. Spectrosc. Radiat. Transf.* 186 (2017), pp. 221–242. DOI: 10.1016/j.jqsrt.2016.04.010.

[102] Müller, H. S. P. et al. "The Cologne Database for Molecular Spectroscopy, CDMS". In: *Astron. Astrophys.* 370.3 (2001), pp. L49–L52. DOI: 10.1051/0004-6361:20010367.

[103] A. F. Al-Refaie, S. N. Yurchenko, and J. Tennyson. "GPU Accelerated INtensities MPI (GAIN-MPI): A new method of computing Einstein-A coefficients". In: *Comput. Phys. Commun.* 214 (2017), pp. 216–224. DOI: 10.1016/j.cpc.2017.01.013.

[104] A. Yachmenev, I. Polyak, and W. Thiel. "Theoretical rotation-vibration spectrum of thioformaldehyde". In: *J. Chem. Phys.* 139.20 (2013), p. 204308. DOI: 10.1063/1.4832322.

[105] A. Yachmenev et al. "High-level *ab initio* potential energy surfaces and vibrational energies of $H_2CS$". In: *J. Chem. Phys.* 135.2 (2011), p. 074302. DOI: 10.1063/1.3624570.

[106] T. H. Dunning. "Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen ". In: *J. Chem. Phys.* 90.2 (1989), pp. 1007–1023. DOI: 10.1063/1.456153.

[107] T. H. Dunning, K. A. Peterson, and A. K. Wilson. "Gaussian basis sets for use in correlated molecular calculations. X. The atoms aluminum through argon revisited". In: *J. Chem. Phys.* 114 (2001), p. 9244. DOI: 10.1063/1.1367373.

[108] S. N. Yurchenko and J. Tennyson. "Atomic and Molecular Line Data". In: *ExoFrontiers*. 2514-3433. IOP Publishing, 2021, 21-1 to 21–8. ISBN: 978-0-7503-1472-5. DOI: 10.1088/2514-3433/abfa8fch21.

[109] S. N. Yurchenko, A. F. Al-Refaie, and J. Tennyson. "ExoCross: a general program for generating spectra from molecular line lists". In: *Astron. Astrophys.* 614 (2018), A131. DOI: 10.1051/0004-6361/201732531.

[110] H. S. P. Müller et al. "The Cologne Database for Molecular Spectroscopy, CDMS: a useful tool for astronomers and spectroscopists". In: *J. Molec. Struct. (THEOCHEM)* 742.1-3 (2005), pp. 215–227. DOI: 10.1016/j.molstruc.2005.01.027.

[111] A. P. Cox, S. D. Hubbard, and H. Kato. "The microwave spectrum of thioformaldehyde, $CD_2S$, and $CH_2S$: Average structure, dipole moments, and $^{33}S$ quadrupole coupling". In: *J. Mol. Spectrosc.* 93.1 (1982), pp. 196–208. DOI: 10.1016/0022-2852(82)90283-1.

[112] S. Erfort, M. Tschöpe, and G. Rauhut. "Toward a fully automated calculation of rovibrational infrared intensities for semi-rigid polyatomic molecules". In: *J. Chem. Phys.* 152.24 (2020), p. 244104. DOI: 10.1063/5.0011832.

[113] G. Tarczay et al. "The barrier to linearity of hydrogen sulphide". In: *Chem. Phys. Lett.* 322.1-2 (2000), pp. 119–128. DOI: 10.1016/S0009-2614(00)00407-3.

[114] J. Watson. "Vibration-rotation Hamiltonians of linear molecules". In: *Mol. Phys.* 79.5 (1993), pp. 943–951. DOI: 10.1080/00268979300101741.

[115] K. L. Chubb, P. Jensen, and S. N. Yurchenko. "Symmetry Adaptation of the Rotation-Vibration Theory for Linear Molecules". In: *Symmetry* 10.5 (2018), p. 137. DOI: 10.3390/sym10050137.

[116] P. Jensen. "The nonrigid bender Hamiltonian for calculating the rotation-vibration energy levels of a triatomic molecule". In: *Comp. Phys. Rep.* 1 (1983), pp. 1–55. DOI: 10.1016/0167-7977(83)90003-5.

[117] H. Kurzweil and B. Stellmacher. *The Theory of Finite Group*. New York: Springer-Verlag, 2004.

[118] H. Wei and T. Carrington Jr. "Explicit expressions for triatomic Eckart frames in Jacobi, Radau, and bond coordinates". In: *J. Chem. Phys.* 107.8 (1997), pp. 2813–2818. DOI: 10.1063/1.474639.

[119] H. Wei. "Eckart frames for planar molecules". In: *J. Chem. Phys.* 118.16 (2003), pp. 7202–7207. DOI: 10.1063/1.1562624.

[120] C. Eckart. "Some Studies Concerning Rotating Axes and Polyatomic Molecules". In: *Phys. Rev.* 47.7 (1935), pp. 552–558. DOI: 10.1103/PhysRev.47.552.

[121] V. Aquilanti and S. Cavalli. "Coordinates for molecular dynamics: Orthogonal local systems". In: *J. Chem. Phys.* 85.3 (1986), pp. 1355–1361. DOI: 10.1063/1.451223.

[122] B. R. Johnson and W. P. Reinhardt. "Adiabatic separations of stretching and bending vibrations: Application to $H_2O$". In: *J. Chem. Phys.* 85.8 (1986), pp. 4538–4556. DOI: 10.1063/1.451775.

[123] Z. Zhao et al. "Full-dimensional vibrational calculations of five-atom molecules using a combination of Radau and Jacobi coordinates: Applications to methane and fluoromethane". In: *J. Chem. Phys.* 144.20 (2016), p. 204302. DOI: 10.1063/1.4950028.

[124] J. Sarka et al. "On neglecting Coriolis and related couplings in first-principles rovibrational spectroscopy: Considerations of symmetry, accuracy, and simplicity. II. Case studies for $H_2O$ isotopologues, $H_3^+$, $O_3$, and $NH_3$". In: *Spectrochim. Acta A* 250 (2021), p. 119164. DOI: 10.1016/j.saa.2020.119164.

[125] B. V. Noumerov. "A Method of Extrapolation of Perturbations". In: *Mon. Not. R. Astron. Soc.* 84.8 (1924), pp. 592–602. DOI: 10.1093/mnras/84.8.592.

[126] J. W. Cooley. "An Improved Eigenvalue Corrector Formula for Solving the Schrödinger Equation for Central Fields". In: *Math. Comput.* 15.76 (1961), pp. 363–374. DOI: 10.2307/2003025.

[127] A. Sayvetz. "The Kinetic Energy of Polyatomic Molecules". In: *J. Chem. Phys.* 7.6 (1939), pp. 383–389. DOI: 10.1063/1.1750455.

[128] J. T. Hougen. "A Group-Theoretical Treatment of Electronic, Vibrational, Torsional, and Rotational Motions in The Dimethylacetylene Molecule". In: *Can. J. Phys.* 42.10 (1964), pp. 1920–1937. DOI: 10.1139/p64-182.