

# PHONANA

John Kendrick

School of Chemistry, University of Leeds, Leeds, LS2 9JT United Kingdom.  
j.kendrick@leeds.ac.uk

Andrew Burnett

School of Chemistry, University of Leeds, Leeds, LS2 9JT United Kingdom  
a.d.burnett@leeds.ac.uk

## Contents

<b>1. INTRODUCTION</b>	<b>1</b>
<b>2. THEORY</b>	<b>2</b>
2.1. Molecular Systems . . . . .	2
2.2. Periodic Systems . . . . .	3
<b>3. IMPLEMENTATION</b>	<b>4</b>
<b>4. QM/MM PROGRAMS</b>	<b>5</b>
<b>5. OPTIONS</b>	<b>6</b>
<b>6. EXAMPLES</b>	<b>6</b>
6.1. Usage . . . . .	6
6.2. Output . . . . .	7
<b>References</b>	<b>8</b>

## 1. INTRODUCTION

phonana is a ‘helper’ program which uses the underlying modules of PDielec to read an output file from an MM or QM calculation of the phonon modes in a crystal. The phonon modes are analysed and the percentage of molecular centre-of-mass and rigid body rotational motion in each mode is calculated. Using the results of the analysis it is possible to differentiate between inter- and intra-molecular phonon modes (see for instance Jepsen et al [1].

## 2. THEORY

### 2.1. Molecular Systems

It is common in molecular calculations of vibrational properties to construct a force constant matrix which enforces the requirements of zero energy change for centre-of-mass motion and rigid-body rotation. This can be achieved by using projection operators to transform the second derivative matrix to a set of coordinates which no longer include centre-of-mass motion or rigid-body rotation.

Defining a projection operator as;

$$\bar{\bar{P}} = \bar{\bar{1}} - \bar{V}\bar{V}^T \quad (1)$$

where  $\bar{V}$  is an orthonormal column vector with length  $3N$  ( $N$  is the number of atoms), the centre-of-mass motion can be described in Cartesian space as all atoms moving along the x, y or z axis with the same displacement. So considering the projection of the centre-of-mass motion along the x-axis, for each atom  $a$  we can write;

$$\bar{V}_a^x = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad (2)$$

For the projection operator which will project out all components of translation along the x-axis for every atom in the molecule we have;

$$\bar{V}^x = \begin{pmatrix} \bar{V}_1^x \\ \bar{V}_2^x \\ \vdots \end{pmatrix} \quad (3)$$

There are two other projection operators describing translation along the y- and z-axis. In a similar fashion it is possible to describe an infinitesimal molecular rotation using a vector  $V$  and therefore constructing a projection operator to remove rigid-body rotation.

If the coordinates of atom  $a$  relative to the centre-of-mass of the molecule are  $x, y$  and  $z$ , the component of the projection vector,  $\bar{V}$  representing rotation about the x-axis in the yz-plane is;

$$\bar{V}_a^{yz} = \begin{pmatrix} 0 \\ -z \\ +y \end{pmatrix} \quad (4)$$

In a similar fashion rotations about the y- and z-axis are respectively;

$$\bar{V}_a^{xz} = \begin{pmatrix} +z \\ 0 \\ -x \end{pmatrix} \quad (5)$$

and

$$\bar{V}_a^{xy} = \begin{pmatrix} -y \\ +x \\ 0 \end{pmatrix} \quad (6)$$

In a similar fashion to Equation (3) the complete projection operators for the 3 rotational modes can be assembled (using rotation about the x-axis as an example)

$$\bar{\mathbf{V}}^{yz} = \begin{pmatrix} \bar{\mathbf{V}}_1^{yz} \\ \bar{\mathbf{V}}_2^{yz} \\ \vdots \end{pmatrix} \quad (7)$$

In practice the projection operators will be defined using mass-weighted cartesian coordinates as this simplifies the expressions used later. As an example the transformation to mass-weighted coordinates for atom  $a$  translating along the x-axis is;

$$\sqrt{m_a} \bar{\mathbf{V}}_a^x \quad (8)$$

## 2.2. Periodic Systems

The invariants of the energy in a periodic system are the three translational modes corresponding to motion of all the atoms in the same direction. The rotation of an infinite lattice is not invariant. However, it is proposed to construct projection operators for each molecule in the unit cell, which will project out the motion of each *molecular* translation and rotation. Such projection operators will be used to separate the external from the internal modes and there will therefore be a projection operator for each external mode in the system.

Hug and Haesler have shown [2] that the vibrational kinetic energy can be decomposed into single centre atomic contributions by considering the kinetic energy,  $T_p$ , of a normal mode  $Q_p$ ;

$$T_p = \frac{1}{2} \dot{Q}_p^2 = \sum_a T_{a,p} = \frac{1}{2} \sum_a m_a \bar{\Delta \dot{\mathbf{x}}}_{a,p}^T \bar{\Delta \dot{\mathbf{x}}}_{a,p} = \frac{1}{2} \dot{Q}_p^2 \sum_a \bar{\mathbf{L}}_{a,p}^T \bar{\mathbf{L}}_{a,p} \quad (9)$$

Where the sum is over atoms,  $a$ , with mass  $m_a$ ,  $\bar{\Delta \dot{\mathbf{x}}}_{a,p}$  is the time derivative of the Cartesian displacement vector of atom  $a$  along the  $p^{th}$  normal mode and  $\bar{\mathbf{L}}_p$  are the components of the orthogonal transformation relating the  $p^{th}$  normal mode and mass-weighted Cartesian displacements. Thus  $\bar{\mathbf{L}}$  are the eigenvectors of the mass-weighted second derivative matrix. The atom subscript is used to indicate only those atomic coordinates involving atom  $a$  are being considered. Derivation of this expression has made use of the following relationship between the mass weighted normal mode;

$$\sqrt{m_a} \bar{\Delta \mathbf{x}}_{a,p} = \bar{\mathbf{L}}_{a,p} Q_p \quad (10)$$

As proposed by Hug and Hasler [2], since the kinetic energy can be expressed as atomic contributions, consideration of the virial theorem indicates that this is true also for the potential energy. Using mass weighted Cartesian coordinates for convenience, an analysis of the phonon modes may therefore be constructed

in the following way. A given phonon mode,  $p$ , will have its total kinetic energy partitioned between the molecules in the unit cell according to;

$$E_p^{total} = \sum_{mol} E_p^{mol} \quad (11)$$

$$E_p^{mol} = \dot{Q}_p^2 \sum_{a \in mol} \bar{\mathbf{L}}_{a,p}^T \bar{\mathbf{L}}_{a,p} \quad (12)$$

To calculate the contribution a particular mode has to the centre-of-mass kinetic energy we use the projection operators given in Equation (3) (now in mass-weighted Cartesian coordinates) where only the atoms in that molecule are used to construct the operator;

$$\bar{\bar{\mathbf{P}}}_{mol}^x = \bar{\mathbf{1}} - \bar{\mathbf{V}}_{mol}^x (\bar{\mathbf{V}}_{mol}^x)^T \quad (13)$$

This approach is used to define 6 projection operators for each molecule. The centre-of-mass energy contained in the  $p_{th}$  mode can be written;

$$E_p^{cm} = \dot{Q}_p^2 \sum_{i=x,y,z} \sum_{mol} (\bar{\bar{\mathbf{P}}}_{mol}^i \bar{\mathbf{L}}_p)^T (\bar{\bar{\mathbf{P}}}_{mol}^i \bar{\mathbf{L}}_p) \quad (14)$$

and the rigid-body rotational energy associated with the mode is;

$$E_p^{rot} = \dot{Q}_p^2 \sum_{i=xy,yz,xz} \sum_{mol} (\bar{\bar{\mathbf{P}}}_{mol}^i \bar{\mathbf{L}}_p)^T (\bar{\bar{\mathbf{P}}}_{mol}^i \bar{\mathbf{L}}_p) \quad (15)$$

The molecular vibrational contribution, which can be used to classify the internal modes of the system, can be obtained by subtracting the centre-of-mass and rigid-body rotational; energies from the total.

$$E_p^{vib} = E_p^{total} - E_p^{cm} - E_p^{rot} \quad (16)$$

Since all of the energy terms depend in the same way on  $\dot{Q}_p$ , it is not required in the calculation of the *relative* contributions to the energy coming from the external (molecular centre-of-mass and rigid-body rotation) and the internal modes (vibrational contributions)

### 3. IMPLEMENTATION

In order to calculate the relative internal and external contributions to each phonon mode it is first necessary to identify the molecules in the crystal. In many cases the unit cell is packed with atoms in such a way that the cell is filled, rather than in a way reflecting the bondedness of the molecules. Phonana therefore first replicates the atoms in all cells neighbouring the central unit cell. Within this supercell the bonds are determined by calculating the distances between all atoms in the supercell. In practice an order N method is used whereby instead of searching all the supercell only the space around each atom

is searched for potentially bonded partners. The criterion of the presence of a bond between atoms  $i$  and  $j$  is given by the requirement that the distance between the atoms  $r_{ij}$  is less than the bonding requirement;

$$r_{ij} < scale(radius_i + radius_j) + toler \quad (17)$$

Here  $radius_i$  is the covalent radius of atom  $i$ ,  $scale$  and  $toler$  are factors which can be altered when running the program.

Once the bonding in the supercell has been determined the program starts with the first atom in the central cell and determines all atoms which are connected to the molecule that it is in. If there are any remaining atoms in the central cell which are not bonded yet then further molecules are added until all atoms in the central cell have been allocated to a molecule.

Finally a new cell with the same dimensions as the original is constructed. Where necessary each molecule is shifted into the cell so that the molecules centre-of-mass lies within the it.

The projection operators, Equation (13), can now be constructed and the relative energies calculated using Equations (14), (15) and (16).

## 4. QM/MM PROGRAMS

The MM / QM packages supported are summarised below.

**VASP** -program vasp OUTCAR The name provided on the command line is an OUTCAR file. The OUTCAR is read by phonana to determine the unit cell, atomic masses, frequencies, normal modes, Born charge tensors and optical permittivity. The VASP run can be a DFPT or numerical calculation of the response.

**CASTEP** -program castep seedname The name provided on the command line is the seedname for the calculation. The corresponding seedname.castep file in the current directory is read and processed to determine the unit cell, atomic masses, optical permittivity and born charge tensors. The normal modes and their frequencies are determined from the seedname.phonon file. The CASTEP run needs to be a DFPT (phonon+efield) task.

**CRYSTAL** -program crystal outputfilename The name on the command line is a file ending in .out, containing the output of a CRYSTAL14 run. The contents of this file alone are sufficient to provide the unit cell, atomic masses, frequencies, normal modes and Born charge tensors. However, the number of significant figures for the normal modes is not sufficient for an accurate calculation and it is therefore recommended that the HESSFREQ.DAT and BORN.DAT files are also made available. If they are present in the directory where phonana is run from, it uses these files to calculate the Born charge tensors, frequencies and normal modes. By default CRYSTAL projects out the pure translational modes of the system before calculating the frequencies, this can also be done by the phonana package. The CRYSTAL calculation needs to be a frequency calculation (FREQCALC) with the infrared intensity (INTENS) selected. The default algorithm does not calculate the optical permittivity, so

this needs to be provided on the command line. However, if the CPHF or CPKS algorithm is used for the frequency calculation, the optical permittivity is calculated and PDielec will automatically read it from the output file. By default CRYSTAL projects out the pure translational modes of the system before calculating the frequencies, this can also be done by the PDielec package. Small differences in the calculated frequencies between the CRYSTAL program and phonana have been observed. These have been found to be due to a slightly different method for symmetrising the 2<sup>nd</sup> derivative matrix, because of this an optional directive “-hessian crystal” can be used to indicate that phonana should use the same symmetrisation as CRYSTAL14.

**ABINIT** -program abinit outputfilename The output file should come from a run containing three datasets. One to calculate the wavefunction at the optimized geometry, one to calculate the field perturbations and one to calculate the second derivatives. Examples of input files and output files are available with the distribution.

**QE** -program qe outputfilename The output file is the dynamical matrix file, specified by “filedyn” in a run of the quantum espresso phonon package. Examples of input and output files are given in the phonana distribution

**GULP** -program gulp outputfilename The name on the command line is a file ending in .gout, containing the output of a GULP run. The contents of this file alone are sufficient to provide the unit cell, atomic masses, frequencies, normal modes, Born charge tensors and optical permittivity. Because GULP only writes out the Born charge matrices for the asymmetric unit, it is necessary to run a frequency calculation using P1 symmetry and a complete unit cell. The key words; nosymm, phonon, intensity, eigen and cart are recommended for the GULP calculation. In the case that no shells are used in the calculation the optical permittivity is not available in the output and it is necessary to provide it on the command line (see -optical and -optical\_tensor options below).

## 5. OPTIONS

There are several command options and these are summarized below. Some of the options may be repeated.

## 6. EXAMPLES

### 6.1. Usage

```
phonana -program vasp OUTCAR
```

This reads the results of the VASP calculation and prints a summary of the molecular centre-of-mass and molecular rotation contributions to each phonon mode.

```
phonana -program vasp OUTCAR -excel results.xlsx
```

Option	Default	Purpose
-program string		string can be “abinit”, “castep”, “crystal”, “gulp”, “qe”, “experiment” or “vasp”
-vmin real	0.0	real is the starting wavenumber ( $\text{cm}^{-1}$ ) for the frequency range
-vmax real	5000.0	real is the final wavenumber ( $\text{cm}^{-1}$ ) for the frequency range
-radius element real		The covalent radius of ‘element’ is set to real
-toler real	0.1	Set the tolerance value used for calculating bonds. A bond is created between two atoms if the distance is less than or equal to the tolerance value
-scale real	1.1	Set the scale factor for calculating bonds
-excel s		Writes the results to an excel spread sheet with the name specified by the string s.
-csv s		Output is sent to a comma separated file specified by the string s.
-print		Additional output is provided from the QM or MM calculation
-ignore k		Ignore the kth mode
-mode k		Only use the kth mode in the analysis
-eckart		The translational modes are projected out of the hessian before diagonalisation
-hessian string		string may be either “crystal” or “symm”. In the case of “crystal” the hessian is calculated in the crystal basis
-masses string		string can be either “program”, “average” or “isotopic”, meaning that the masses are taken from the program, averaged or isotopic
-mass string real		The atomic mass of string is set to real. This can be used to explore the effect of mass on the phonon frequencies

**Table 1.** Command line options for phonana <sup>1</sup>This column indicates if a command line option can be used more than once

As above but with the results written to a spreadsheet

```
phonana -program vasp OUTCAR -excel results.xlsx -eckart -radius S 1.0
```

As above but with the eckardt projection of the crystal centre-of-mass motion. The covalent radius of the sulphur atom is set to 1 Angstrom.

## 6.2. Output

The example of a calculation on isoleucine using Castep is available in the Examples/Castep/Isoleucine directory of the distribution. The analysis was performed using;

```
phonana -program castep phonon.castep -excel results.xlsx
```

The program first finds the four molecules of isoleucine in the unit cell of the crystal structure. After summarising the the unit cell information, the elements of the atoms are given. The ordering of the atoms at this point is different to the ordering in the original cell and in the QM calculation. The program lists the cartesian and fractional coordinates. Again these will not be the same as in the original calculation. For each molecule the program lists the index numbers of the atoms in the molecule, the molecular mass and the centres of mass of the molecules. The results are summarised in Table 2.

Table 3 shows that results of the analysis of the phonon modes with frequencies below  $100 \text{ cm}^{-1}$ . The first 3 modes should have zero frequency as they are the translationally invariant modes of the lattice. The negative number shown here actually indicates that the mode has an imaginary frequency. As

Molecule	Mass	CM <sub>a</sub>	CM <sub>b</sub>	CM <sub>c</sub>
0	131.172	0.293600	0.051683	0.304906
1	131.172	0.706400	0.551683	0.695094
2	131.172	0.793231	0.588919	0.287125
3	131.172	0.206769	0.088919	0.712875

**Table 2.** Molecular constituents of the isoleucine unit cell (centre-of-masses are in fractional coordinates)

can be seen under the column %col-cme they are almost completely associated with molecular centre-of-mass motion. This is true also of the lowest non-zero frequency at 32,81  $\text{cm}^{-1}$  and to some extent of the next mode at 39.73  $\text{cm}^{-1}$ . Although for this latter mode there is considerable contribution from rigid-body rotational motion (see the %mol-rot column). As the frequency of the phonon mode increases the contribution from rigid-body motion generally decreases and contribution from vibrational mode (see the %vib column) increases. The total contribution from molecular motion is summarised in the last four columns.

## References

- [1] P. U. Jepsen and S. J. Clark, “Precise ab-initio prediction of terahertz vibrational modes in crystalline systems,” *Chemical Physics Letters*, vol. 442, no. 4–6, pp. 275–280, Jul. 2007.
- [2] W. Hug and J. Haesler, “Is the Vibrational Optical Activity of,” vol. 104, pp. 695–715, 2005.



Freq(cm-1)	%mol-cme	%mol-rot	%vib	%mol-0	%mol-1	%mol-2	%mol-3
-0.07	99.5	0.5	0.0	26.9	26.9	23.1	23.1
-0.06	98.6	1.1	0.3	23.6	23.6	26.4	26.4
-0.04	99.7	0.2	0.1	23.8	23.8	26.2	26.2
32.81	92.0	5.9	2.1	26.8	26.8	23.2	23.2
39.73	77.9	19.6	2.5	22.1	22.1	27.9	27.9
49.42	10.4	81.9	7.7	19.0	19.0	31.0	31.0
52.12	80.7	17.3	2.0	17.6	17.6	32.4	32.4
54.15	73.4	19.0	7.5	21.2	21.2	28.8	28.8
59.59	37.7	56.6	5.7	23.9	23.9	26.1	26.1
62.05	77.3	17.2	5.5	23.4	23.4	26.6	26.6
69.91	9.6	69.4	21.0	17.6	17.6	32.4	32.4
73.19	7.3	61.1	31.6	6.6	6.6	43.4	43.4
76.00	13.6	61.7	24.7	31.6	31.6	18.4	18.4
76.25	3.2	48.1	48.7	10.6	10.6	39.4	39.4
80.49	31.5	43.8	24.7	19.3	19.3	30.7	30.7
87.54	37.9	35.4	26.7	26.8	26.8	23.2	23.2
89.31	12.2	58.1	29.6	27.5	27.5	22.5	22.5
90.49	16.9	53.9	29.2	22.4	22.4	27.6	27.6
94.51	49.1	36.9	14.0	39.4	39.4	10.6	10.6
94.79	47.0	16.9	36.1	23.0	23.0	27.0	27.0

**Table 3.** Percentage contributions of the centre-of-mass and the rotational molecular modes to each phonon mode