CHEM-279 FINAL PROJECT

Visualization Module for Normal Modes

By: Festo Muhire & Paul Graggs

Project Goals
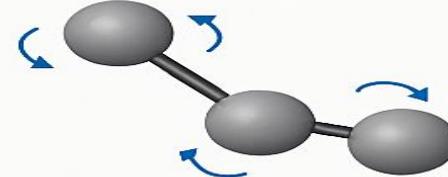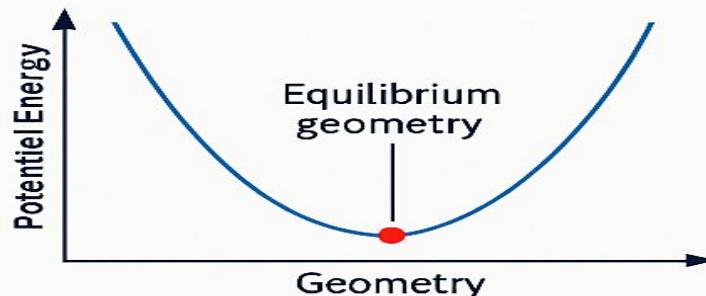
The goal of this project was to develop a program for obtaining the vibrational frequencies and visualizing the normal modes of an optimized molecular system.

We hoped to be able to reproduce both the expected (experimental) values and the expected visual representations.

Obtaining the frequencies would allow us to confirm the geometry was correctly optimized, predict spectroscopy signatures, and build intuition for how molecules behave.

Introduction to Molecular Normal Modes
How Small Vibrations Reveal Deep Physical Insights

- Normal modes arise from smll oscillations around the equilibrium geometry
- They provide the foundation for vibrational analysis, spectroscopy and reaction*path theore
- This project links electronic structure, gradients, the Hessian, and vibrational frequencies

How This Builds on What We've Learned

- You have already computed energies and gradients
- The Hessian extends this to second derivatives
- Diagonalizing the Hessian reveals fundamental vibrational motions

Normal Modes: Vibrational Analysis

Normal Modes = Independent Vibrational Patterns of a Molecule

- Or simple solutions of a system of oscillators where all the degrees of freedom oscillate with the same frequency.

For a Given Molecule with N atoms there are 3N degrees of freedom

The number of normal modes equals:

- For Nonlinear Molecules
  - 3N - 6
- For Linear Molecules
  - 3N - 5

How Do We Obtain The Normal Modes?

Process to Compute the Normal Modes:

1. Optimize the molecular geometry

2. Compute the Hessian matrix

3. Convert to mass-weighted Hessian

4. Solve the eigenvalue problem

5. The eigenvalues are the frequencies

6. The eigenvectors are the normal mode displacement patterns

Geometry Optimization: Homework 5 extra credit

- To optimize the geometry, we refactored the optimization code from homework 1 to be compatible with our CNDO/2 code.

- Normal modes can only be computed once the molecule is fully optimized.

- At the optimized geometry, moving any atom slightly in any direction increases the energy.

Linear Algebra Review

What are Eigenvectors and Eigenvalues?

Eigenvectors are nonzero vectors that, when a linear transformation is applied to it, results in a new vector that has the same direction as the original vector but a different magnitude.

Eigenvalues are scalars that scale the original vector into the new vector.

The Hessian Matrix

The second derivatives of the energy (in our case).

H for E = f(x,y)   H for E = f(x,y,z)

$$H = \begin{bmatrix} \dfrac{\partial^2 f}{\partial x^2} & \dfrac{\partial^2 f}{\partial x \partial y} \\[2ex] \dfrac{\partial^2 f}{\partial y \partial x} & \dfrac{\partial^2 f}{\partial y^2} \end{bmatrix}$$

$$\mathbf{H}(f) = \begin{pmatrix} \dfrac{\partial^2 f}{\partial x^2} & \dfrac{\partial^2 f}{\partial x \partial y} & \dfrac{\partial^2 f}{\partial x \partial z} \\[2ex] \dfrac{\partial^2 f}{\partial y \partial x} & \dfrac{\partial^2 f}{\partial y^2} & \dfrac{\partial^2 f}{\partial y \partial z} \\[2ex] \dfrac{\partial^2 f}{\partial z \partial x} & \dfrac{\partial^2 f}{\partial z \partial y} & \dfrac{\partial^2 f}{\partial z^2} \end{pmatrix}$$

What does the Hessian Do

- The Hessian Matrix is a linear transformation that maps atomic displacements to restoring forces
- In other words, the Hessian tells you how hard the energy pushes back if you displace an atom
- The Hessian Matrix is a 3N x 3N matrix, where N is the number of Atoms
- By diagonalizing (Solving the Eigenvalue Problem) we can determine the independent vibrational motions and their frequencies

Why is the Hessian 3N x 3N?

- The Gradient Matrix could be represented as a 3N vector, where each atom has its x coordinate y coordinate, and z coordinate adjacent to each other as such:
  - $[X_0, Y_0, Z_0, X_1, Y_1, Z_1, ..., X_{n-1}, Y_{n-1}, Z_{n-1}]$
- To build the Hessian we take the second derivative of each atom with respect to each coordinate. Therefore we take the second derivative with respect to X_0 with respect to X, Y, and Z.
- As a result for every entry in the 3N vector we compute the second derivative with respect to all 3N coordinates, which form a 3N x 3N matrix

What do the normal modes of the Hessian represent

- Each Eigenvector is a specific pattern of motion in a given direction, where the eigenvalue is the magnitude of the displacement in that direction
- The eigenvalues of translational and rotational motion map those displacement to the zero vector
- If all the atoms move in a way where the bond lengths remain constant the energy is does not change (rotational and translational invariance)
- If an atom moves in a way that changes a bond length the energy increase and the intermolecular forces act to restore the molecule to the relaxed position

Central Difference Approximation

In Homework five we computed the first derivative of the Energy and stored it in our Gradient Matrix.

Since this matrix contained already contained the first derivatives of the energy, we could use finite differences to approximate the second derivative.

$$\frac{\partial F_i}{\partial x_j} \approx \frac{F_i(x_j + h) - F_i(x_j - h)}{2h}$$

Convert to Mass Weighted Hessian

```cpp
// Function to calculate the weighted Hessian matrix:
mat calculate_weighted_hessian(const mat& hessian, const Molecule& molecule)
{
    // 1. Create a weighted Hessian Matrix: (H_mw(i, j) = H(i, j) * sqrt(m_i * m_j))
    int N = molecule.atoms.size();
    mat weighted_hessian(3*N, 3*N, fill::zeros);
    for (int i = 0; i < 3*N; ++i)
    {
        for (int j = 0; j < 3*N; ++j)
        {
            int atom_i_idx = i / 3;
            int atom_j_idx = j / 3;
            double mass_i = atomic_masses.at(molecule.atoms[atom_i_idx].atomic_number) * AMU_TO_AU;
            double mass_j = atomic_masses.at(molecule.atoms[atom_j_idx].atomic_number) * AMU_TO_AU;
            weighted_hessian(i, j) = hessian(i, j)/ std::sqrt(mass_i * mass_j);
        }
    }
    // 2. Return the weighted Hessian matrix:
    return weighted_hessian;
}
```

VERIFICATION STRATEGY: psi4

The Hessian Matrix H

Units: Hartree / Bohr^2

$$H = \begin{bmatrix} -0.0192949 & 0 & 0 & 0.0192949 & 0 & 0 \\ 0 & 0.0398945 & 0 & 0 & -0.0398945 & 0 \\ 0 & 0 & 0.0398945 & 0 & 0 & -0.0398945 \\ 0.0192949 & 0 & 0 & -0.0192949 & 0 & 0 \\ 0 & -0.0398945 & 0 & 0 & 0.0398945 & 0 \\ 0 & 0 & -0.0398945 & 0 & 0 & 0.0398945 \end{bmatrix}$$

Versus Our Result

Mass-Weighted Hessian $F = M^{-1/2}HM^{-1/2}$

Units: Hartree / (Bohr^2·amu)

$$F = \begin{bmatrix} 19.319 & 0 & 0 & -19.319 & 0 & 0 \\ 0 & -1.2642 \times 10^{-6} & 0 & 0 & 1.2642 \times 10^{-6} & 0 \\ 0 & 0 & -1.2642 \times 10^{-6} & 0 & 0 & 1.2642 \times 10^{-6} \\ -19.319 & 0 & 0 & 19.319 & 0 & 0 \\ 0 & 1.2642 \times 10^{-6} & 0 & 0 & -1.2642 \times 10^{-6} & 0 \\ 0 & 0 & 1.2642 \times 10^{-6} & 0 & 0 & -1.2642 \times 10^{-6} \end{bmatrix}$$

Solving The Eigenvalue Problem
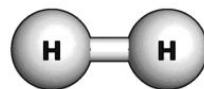
- Since the eigenvector represent the modes we are trying to visualize, we can leverage a linear algebra library to compute the eigenvectors.
- The modes we are interested in will have none zero eigenvalues
- If the geometry was not perfectly optimized we may get imaginary modes, which we will be able to tell because their eigenvalues will be negative.
- Any eigenvectors with negative eigenvalues we will ignore.

Results: Normal Modes & Frequencies

Molecules – LINEAR	Frequencies (cm^{-1})	Experimental values (cm^{-1})	Error (%)
H$_2$	Mode 1: 0 Mode 2: 0 Mode 3: 0 Mode 4: 0 Mode 5: 0 Mode 6: 6126.37	4401	39.2
HF	Mode 1: 0 Mode 2: 0 Mode 3: 0 Mode 4: 0 Mode 5: 0 Mode 6: 6094.76	4138	47.3
HO	Mode 1: 0 Mode 2: 0 Mode 3: 0 Mode 4: 0 Mode 5: 0 Mode 6: 5703.66	3737	52.6
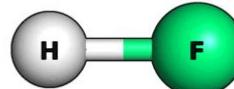
Visualization: Linear Molecules
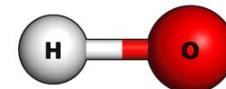
Molecule: H$_2$ (N = 2)

Molecule: HF (N = 2)

Molecule: HO (N = 2)

H-H Bond Stretch
Mode 6: 6126.37 cm^{-1}
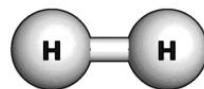
H-F Bond Stretch
Mode 6: 6094.76 cm^{-1}

H-O Bond Stretch
Mode 6: 5703.66 cm^{-1}
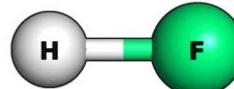
Visualization: Linear Molecules
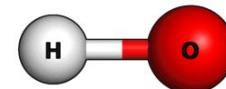
Molecule: H$_2$ (N = 2)

**Molecule: HF (N = 2)**

Molecule: HO (N = 2)

H-H Bond Stretch
Mode 6: 6126.37 cm^{-1}

H-F Bond Stretch
Mode 6: 6094.76 cm^{-1}
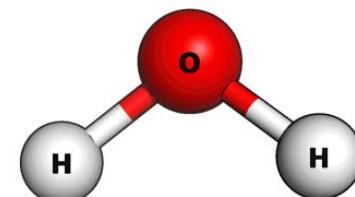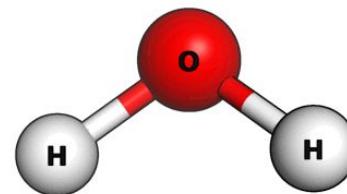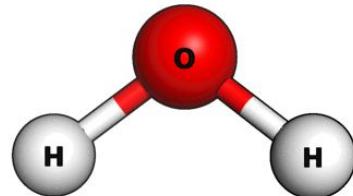
H-O Bond Stretch
Mode 6: 5703.66 cm^{-1}

Results: Normal Modes & Frequencies

| Molecules – NON LINEAR	Frequencies (cm^{-1})	Experimental values (cm^{-1})	Error(%)
H$_2$O **[3 Atom Molecule]**	Mode 1: 0 Mode 2: 0 Mode 3: 0 Mode 4: 0 Mode 5: 0 Mode 6: 0 Mode 7: 1766.42 Mode 8: 5679.51 Mode 9: 5694.04	1595 3657 3756	10.74 55.3 51.6

Visualizations: Non-linear Molecules

Molecule: H_2O (N = 3)

Symmetric Bend
Mode 7: 1766.42 cm-1

Symmetric Stretch
Mode 8: 5679.51 cm^{-1}
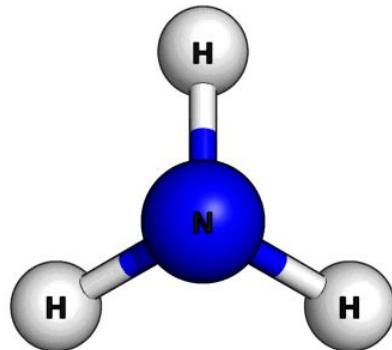
Asymmetric Stretch
Mode 9: 5694.04 cm^{-1}

Results: Normal Modes & Frequencies
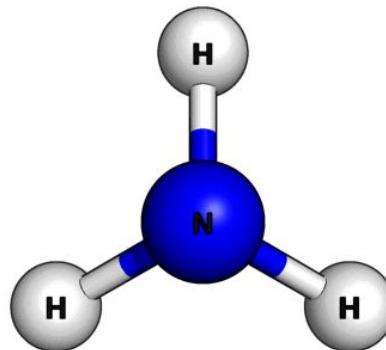
| Molecules – NON LINEAR	Frequencies (cm^{-1})	Experimental values (cm^{-1})	Error(%)
NH$_3$ **[4 Atom Molecule]**	Mode 1: 0 Mode 2: 0 Mode 3: 0 Mode 4: 0 Mode 5: 0 Mode 6: 0 Mode 7: 1506.12 Mode 8: 1778.95 Mode 9: 1779.13 Mode 10: 5225.6 Mode 11: 5264.05 Mode 12: 5264.13	950 1627 3337 3414	58.5 9.33 9.35 56.6 54.19 54.19

Visualizations: Non-linear Molecules
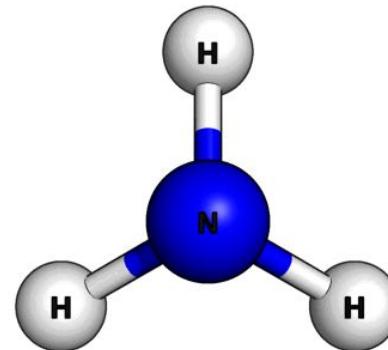
Molecule: NH_3 (N = 4)


Symmetric Bend
Mode 7: 1506.12 cm-1
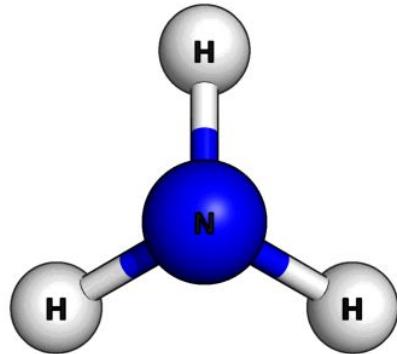
Asymmetric Bend
Mode 8: 1778.95 cm-1
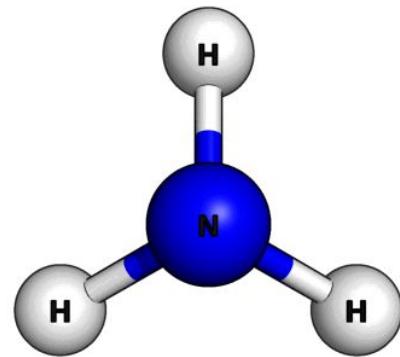
Asymmetric Bend
Mode 9: 1779.13 cm-1

Visualizations: Non-linear Molecules
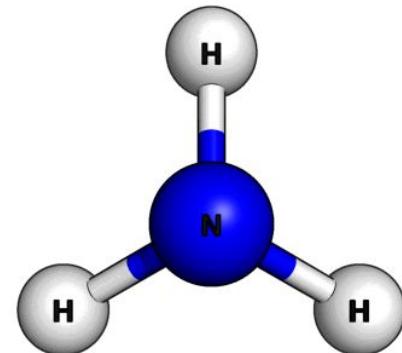
**Molecule: NH_3 (N = 4)**

Symmetric Stretch
Mode 10: 5225.6 cm-1

Asymmetric Stretch
Mode 11: 5264.05 cm-1

Asymmetric Stretch
Mode 12: 5264.13 cm-1
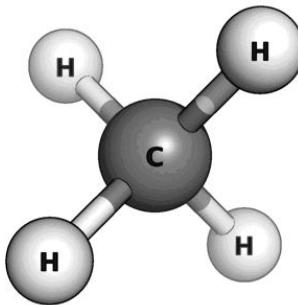
Results: Normal Modes & Frequencies

| Molecules – NON LINEAR	Frequencies (cm^{-1})	Experimental values (cm^{-1})	Error(%)
CH$_4$ **[5 Atom Molecule]**	Mode 1: 0 Mode 2: 0 Mode 3: 0 Mode 4: 0 Mode 5: 0 Mode 6: 0 Mode 7: 1648.43 Mode 8: 1648.43 Mode 9: 1648.43 Mode 10: 1760.53 Mode 11: 1760.53 Mode 12: 4837.61 Mode 13: 4950.33 Mode 14: 4950.33 Mode 15: 4950.33	1380 1412 3215 3104	19.45 24.68 50.46 59.48
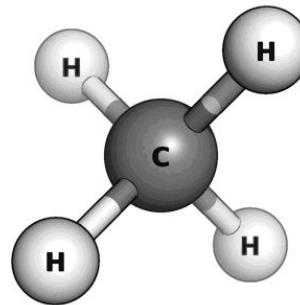
# Visualization continued:
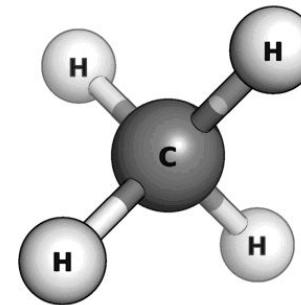
Molecule: CH$_4$ (N = 5)

A triply degenerate bend due to geometry

Asymmetric bend
Mode 7: 1648.43 cm-1
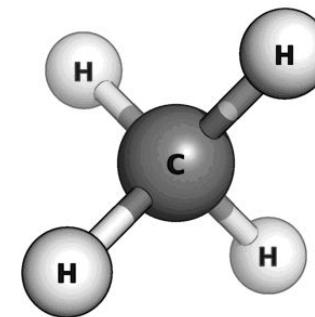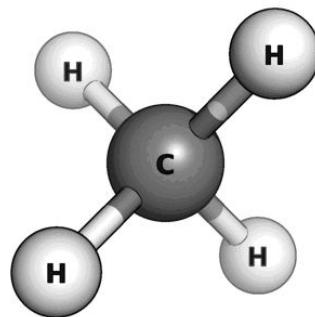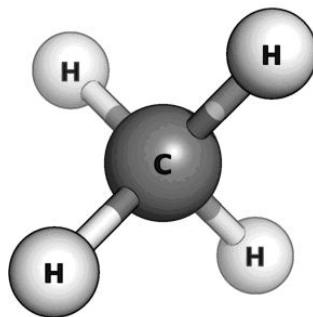
Asymmetric bend
Mode 8: 1648.43 cm-1

Asymmetric bend
Mode 9: 1648.43 cm-1

# Visualization continued:
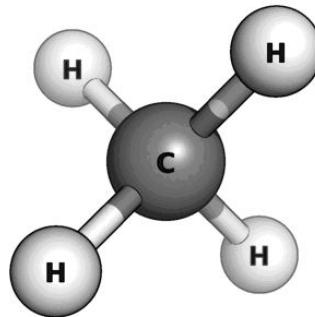
Molecule: CH$_4$ (N = 5)

Symmetric bend
Mode 10: 1760.53 cm-1

Symmetric bend
Mode 11: 1760.53 cm-1
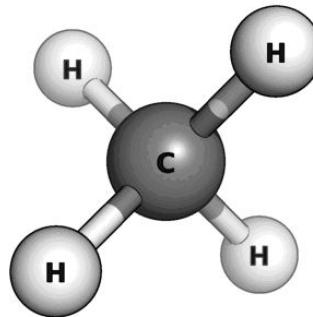
Symmetric Stretch
Mode 12: 4837.61 cm-1

# Visualization continued:

Molecule: CH$_4$ (N = 5)

A triply degenerate stretch due to geometry
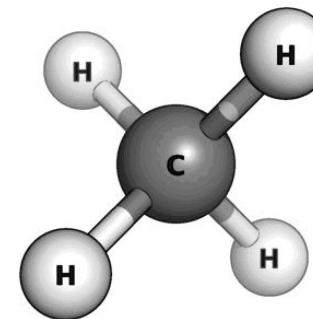
Asymmetric stretch
Mode 13: 4950.33 cm-1

Asymmetric stretch
Mode 14: 4950.33 cm-1

Asymmetric Stretch
Mode 15: 4950.33 cm-1

Observations & Interpretations

- Observations: <u>Zero Modes in Linear vs. Non-Linear Molecules</u>
    - Linear molecules (H_2, HF), the first **5 modes** are consistently 0 cm^{-1}
    - Non-linear molecules (H_2O, NH_3), the first **6 modes** are consistently 0 cm^{-1}
- Interpretation:
    - Zero values represent Translational (moving through space) & Rotational (spinning) modes
    - When molecules rotate or drifts freely in space, bond lengths and angles do not change. Hence, if there's no change in potential energy, the force constant is zero, resulting in zero frequency
    - Fundamental degrees of freedom: linear molecules cannot rotate around their own axis, losing 1 rotational degree of freedom ($3N - 5$) compared to non-linear molecules ($3N-6$)

Observations & Interpretations

- Observation: <u>Distinction Between Low and High Frequency Modes</u>
  - The bending modes, such as H_2O [**~1766 cm^{-1}**], NH_3 [**~1506 cm^{-1}**], and CH_4 [**~1648 cm^{-1}**] are significantly lower than their stretching counterparts.
- Interpretation:
  - These modes correspond to the bending or deformation motion ("umbrella" inversion of Ammonia, or the "scissoring" of water).
  - Geometrically, it requires less energy to change a bond angle than to stretch a bond length.
  - The high values indicate "stiff" chemical bonds. To stretch atoms and break these bonds, it requires significant energy, resulting in high oscillations frequencies
  - In physical terms, the lower frequencies indicate that "the springs" controlling the angles are looser than the "springs" controlling bond distances

Observations & Interpretations

- Observation: <u>Symmetry in Ammonia and Methane</u>
  - For Ammonia, we observe distinct pairs of identical frequencies (e.g., Modes 7 & 8, as well as Mode 10 & 11)
  - For Methane, we observe sets of similar frequencies [Asymmetric motions], as well as distinct frequencies mode 12 [symmetric motion]
- Interpretation:
  - Ammonia & methane are interesting molecules, which involves symmetry fingerprint → These closely identical values represent degeneracy. Due to Ammonia's triangular symmetry & methane tetrahedral geometry, the motion of "Hydrogen A" is indistinguishable from "Hydrogen B". The code correctly realizes that these distinct motion require the same exact amount of energy for their motion.

Results vs. Experimental values → Error Analysis

- Interpretation of frequency accuracy:
  - All calculated frequencies consistently overestimated experimental values for both linear and non-linear molecules
  - Linear molecules deviations ranging from **39.2%** to **52.6%**
  - Non-linear stretching modes had higher deviations ~**51 - 55%**
- "Stiffness" of the model:
  - The overestimation is an indication of the use of a "stiff" theoretical model such as Hartree-Fock with a minimal basis set
  - The model restricts electron flexibility, hence creating steeper potential energy well than reality. A steeper well corresponds to a stiffer spring, resulting in higher calculated frequencies
- Bending vs. Stretching:
  - The bending mode of H_2O had only a **10.74%** error and NH_3 asymmetric bends had ~**9.3%** error → the basis set captures angular changes well, but fails to describe bond elongation for stretching modes

Limitations

Vibrational Analysis Relies On Optimized Structures.

One of the limitations we ran into is our geometry optimizer is not perfect and doesn't always find the true minimum energy structure

In those cases our normal mode calculations are not accurate

Another issues is that our upstream code only generates basis functions for certain atoms. Our code therefore will not compute the normal modes (or do any of the other things) for molecules contain sulfur or phosphorus for instance

Conclusion

- Project success
  - Developed a high-performance C++ module that automates the pipeline from molecular geometry optimization to normal mode analysis & visualization
- Mathematical robustness
  - Module able to distinguish and handle correctly linear & non-linear degrees of freedom
  - Eigenvalues were overestimated due to theory limitations, the eigenvectors (the shapes of vibrations shown in the slides) are physically correct and maintained degeneracy
- Computational Efficiency
  - By integrating the backend calculations with visualization file generation, the module facilitates the analysis workflow compared to manual parsing of the log files
- Future work
  - As seen, frequency overestimation is a byproduct of limited quantum mechanical approximations in Hartree-Fock. Future implementations could involve using Density Functional Theory (DFT) to bring the frequency values closer to experimental data

References

1. Experimental frequencies: https://www.chem.purdue.edu/jmol/vibs/
2. CHEM-279 Lecture materials
3. Code Review & Troubleshooting: Google Gemini Pro