Research Notes

Alex Kerr

1 Introduction

A collection of notes in regards to improving thermal conductivity across molecular interfaces, computationally.

2 Optimization

Optimization methods are needed in order to find the coordinates of atoms that minimized the potential energy in molecules. In this paper we will only attempt to address optimization methods that are utilized in our program. We will not address the theory or many alternative optimization methods. *Practical Methods of Optimization* by *Fletcher* is a great resource for optimization methods.

2.1 Constrained vs. Unconstrained

We are presented with a problem where we want to minimize some function $F(\boldsymbol{x})$ that is a function of n variables. This is equivalent to finding the *maximum* of $-F(\boldsymbol{x})$, so we understand the methods of maximization and minimization to be equivalent. Collectively we call this optimization. There are two types of optimization: constrained and unconstrained. In the constrained problem we are trying to minimize $F(\boldsymbol{x})$ while subjected to a series of m constraints of the form:

$$C_i(\boldsymbol{x}) = 0; \quad i = 1, 2, ..., m \tag{1}$$

Examples of constraints in our molecular geometry problem might be keeping the position of atom i constant ($\boldsymbol{r}_i - \boldsymbol{r}_0 = 0$), keeping the relative distance between atoms i and j constant ($R_{ij} - R_0 = 0$), keeping the bond angle formed by atoms i,j, and k constant, etc. The unconstrained problem is simpler; minimize $F(\boldsymbol{x})$ by altering $\boldsymbol{x}$ however you need to do so.

For now we will focus on unconstrained techniques as they are the simpler and more general concepts.

2.2 Gradient Methods

Since we could be dealing with molecules of $\sim 10^3$ atoms, we have decided to focus on first derivative, or gradient, methods to find local minima/maxima. Higher order methods require the calculation of the Hessian matrix, which has dimensions of $3N \times 3N$.

2.2.1 Steepest Descent

$$\boldsymbol{q}_{i+1} = \boldsymbol{q}_i - s_i \boldsymbol{g}_i \tag{2}$$

where $\boldsymbol{q}_i$ is the set of all coordinates at step i, $\boldsymbol{g}$ is the gradient of the potential energy function with respect to the coordinates, and s is a step-size. Steepset Descent can converge slowly; one of the reasons is that adjacent steps are orthogonal to each other which can cause the minimizer to "zig zag" down valleys". Usually the Conjugate Gradient method is preferable.

2.2.2 Conjugate Gradient

Similar to the Steepest Descent method, we update our positions like so:

$$\boldsymbol{q}_{i+1} = \boldsymbol{q}_i + s_i \boldsymbol{v}_i \tag{3}$$

$$\boldsymbol{v}_i = -\boldsymbol{g}_i + \gamma_i \boldsymbol{v}_{i-1} \tag{4}$$

$$\gamma_i = \frac{\boldsymbol{g}_i \cdot \boldsymbol{g}_i}{\boldsymbol{g}_{i-1} \cdot \boldsymbol{g}_{i-1}} \tag{5}$$

In effect what you're doing in this method is taking in account the history of the previous steps to guide your current one. This typically leads to faster convergence.

2.2.3 Line Search

Determining the step size for minimizer iteration can be problematic. Finding the exact best step size can be very expensive and can bottleneck the entire minimization process. Some methods use first derivation information while others just use the main function evaluated at different points. One can sample points along the search direction but this requires many function evaluations. One technique is to bracket the step size along the search direction, then assume the function is shaped like a parabola within that bracket. The step size is calculated to be the minimum of the would be parabola. See Press's *Numerical Recipes: The art of scientific computing* [5]

2.3 Second Order Methods

As stated earlier, when molecules are larger their related computations become more expensive at an accelerated rate. However it may be worthwhile to explore these methods for smaller molecules, such as single functional groups.

2.3.1 Newton-Raphson Method

Higher order methods usually involve calculating the Hessian matrix ($\boldsymbol{H}$) corresponding to the function to be minimized. You may remember from your first calculus course that the local minimum of a function $f(x)$ can be found iteratively through the Netwon-Raphson method:

$$x_{i+1} = x_i - \frac{f'(x_i)}{f''(x_i)} \tag{6}$$

If f is quadratic in x then this method finds the exact minimum in one step. We can generalize this method to n dimensions:

$$\boldsymbol{q}_{i+1} = \boldsymbol{q}_i - \boldsymbol{H}^{-1}\boldsymbol{g} \tag{7}$$

If each of the n coordinates contributes to F quadratically, then again this process only takes one step to be solved exactly. This method has nice features including generally requiring a smaller number of steps to reach precision and no step size needs to be calculated. However each step is more computationally expensive than gradient methods. For larger molecules the Hessian and its inverse can be prohibitively expensive to calculate.

2.4 The Constrained Case

Fortunately many of the concepts from unconstrained optimization carry over to the more special constrained problems. However while dealing in Cartesian coordinates our constraints are "very" nonlinear (constant bond length, bond angle, etc.) While the deficiencies in using Cartesian coordinates are very real and have been discussed, efficient algorithms have been presented for the constrained problem [1].

2.4.1 Penalty Function Method

One way of enforcing "soft" constraints is to add extra terms to $F(\boldsymbol{x})$ that are zero when the constraints are met and are large otherwise. Each extra term "penalizes" the coordinates for straying too far from their constraints in a way to effectively alter the force, or gradient, calculations. A typical example is the quadratic penalty function:

$$F'(\boldsymbol{x}) = F(\boldsymbol{x}) + \sum_{i=1}^{m} \frac{1}{2}\sigma_i \left(C_i(\boldsymbol{x}_i)\right)^2 \tag{8}$$

The big advantage of using penalty functions is that you can treat your problem as an unconstrained one. You calculate your energies the same way but your new gradients are calculated from Equation 8. The idea is that you reach precision in a run, increase σ_i then run until precision is met again. σ_i is increased sequentially in this manner for harder and harder enforcement of your constraints.

However convergence time can be quite slow with this method and the extent to which your constraints are met can vary.

2.4.2 Lagrange Multipliers Method

The classical Lagrange multiplier method has been shown to have the same general applicability for Cartesian coordinate constraints as they do for internal coordinates. One thing to keep in mind is that each constraint adds a dimension to the problem in Cartesian coordinates, so the Lagrange multipliers will have to be altered in the same way the coordinates are. The Lagrangian function takes the place of $F(\boldsymbol{x})$ in the unconstrained problem:

$$L\left(\boldsymbol{x}, \boldsymbol{\lambda}\right) = F(\boldsymbol{x}) - \sum_{i=1}^{m} \lambda_i C_i(\boldsymbol{x}) \tag{9}$$

$$\frac{\partial L}{\partial x_j} = \frac{\partial F(\boldsymbol{x})}{\partial x_j} - \sum_{i=1}^{m} \lambda_i \frac{\partial C_i(\boldsymbol{x})}{\partial x_j} \tag{10}$$

$$\frac{\partial L}{\partial \lambda_i} = -C_i(\boldsymbol{x}) \tag{11}$$

Here $\nabla L = 0$ ($\frac{\partial L}{\partial x_j} = \frac{\partial L}{\partial \lambda_i} = 0$) gives a possible solution to the constrained problem like $\boldsymbol{g} = 0$ for the unconstrained case. Note that the del operator here refers to the partial derivatives with respect to the coordinates *and* the Lagrange multipliers. If using Cartesian coordinates, constraints actually add dimensionality to the problem. This is just one of the differences between using internal coordinates and Cartesians as laid out in [1]

2.4.3 Mode Following

This algorithm can be incorporated with the penalty function or Lagrange multiplier methods to reach convergence faster. The method involves minimizing your molecule along its lowest energy normal mode and maximizing all others. Presented in [1].

3 Common Forcefield Interactions

We will explore calculating potential energies and their derivatives using Cartesian coordinates.

3.1 Bond Stretching

There are usually potential terms for bonded atoms that are quadratic in relative atomic distance ala Hooke's law:

$$U_{ij} = \frac{1}{2} k_{ij} (R_{ij} - R_{ij0})^2 \tag{12}$$

where R_{ij} is the distance between atoms i and j:

$$R_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2} \tag{13}$$

k_{ij} is the force constant that is specific to i,j and R_{ij0} is the equilibrium distance, also specific to i, j.

3.2 Bond Bending

It is also common to see potential terms quadratic in bond angle:

$$U_{ijk} = \frac{1}{2} k_{ijk} (\theta_{ijk} - \theta_{ijk0})^2 \tag{14}$$

where θ_{ijk} is the angle formed by drawing a line from i to j, then from j to k:

$$\theta_{ijk} = \cos^{-1} \left(\frac{R_{ij}^2 + R_{jk}^2 - R_{ik}^2}{2 R_{ij} R_{jk}} \right) \tag{15}$$

by the Law of Cosines. Similar to Equation 12, k_{ijk} and θ_{ijk0} are force constants.

3.3   Dihedral Motion

In 3D structures we may have contributions to the potential energy from the dihedral angles (torsions) created by sets of 4 atoms:

$$U_{ijkl} = \sum_n V_n \left[1 + \cos\left(n\omega_{ijkl} - g_n\right)\right] \tag{16}$$

where ω_{ijkl} is the angle between the planes formed by i,j and k,l:

$$\omega_{ijkl} = \cos^{-1}\left(\frac{(\boldsymbol{r}_{jk} \times \boldsymbol{r}_{ij}) \cdot (\boldsymbol{r}_{kl} \times \boldsymbol{r}_{jk})}{R_{ij} R_{jk}^2 R_{kl} \sin\theta_{ijk} \sin\theta_{jkl}}\right) \tag{17}$$

What we have in Equation 16 is an effective Fourier series of terms that depend on the torsions. The number of terms in the sum (n) and the phase(s) g_n will depend on atoms i, j, k, l.

3.4   Nonbonded Interactions

In addition to the potential energy involved in bonded atoms, there are energies that are associated among atoms that are *not* bonded. Usually these terms are associated with some kind of cutoff function as these terms go to zero at sufficiently large separations.

3.4.1   Lennard-Jones Potential

Force fields will often have a Lennard-Jones 12-6 type interaction between nonbonded atoms:

$$U_{ij} = \frac{C_{12}}{R_{ij}^{12}} - \frac{C_6}{R_{ij}^6} \tag{18}$$

Again it should be stressed that this interaction takes place if atoms i and j are NOT bonded.

3.4.2   Electrostatics

Electrostatic energy can be accounted for in a number of different ways. Some force fields elect to calculate the Coulombic energy between point charges, dipole-dipole interactions, higher order multipole interactions, or some combination. The potential energy associated with two point charges follows from Coulomb's law:

$$U_{ij} = \frac{1}{4\pi\varepsilon_0} \frac{Q_i Q_j}{R_{ij}} \tag{19}$$

4   Program Forcefields

Our program will use specific force fields that we will address here.

4.1   Amber ff99

We will incorporate the Amber force field with the ff99 parameter set. The force field is given by[2]:

$$
\begin{aligned}
U_{total} = &\sum_{bonds} K_r (r - r_{eq})^2 + \sum_{angles} K_\theta (\theta - \theta_{eq})^2 \\
&+ \sum_{dihedrals} \sum_n \frac{V_n}{2} \left[1 + \cos(n\omega - \gamma)\right] + \sum_{i<j} \left[\frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^6} + \frac{q_i q_j}{\epsilon R_{ij}}\right]
\end{aligned} \tag{20}
$$

Here the bond lengths are given by r. The last summation is understood to be among nonbonded i and j and will undoubtedly use a cutoff function.

4.2   GAFF

The General Amber Forcefield Field, supposedly viable across many different molecule types. Is functionally identical to Amber, but with different parameters.

4.3 Tersoff

We may use the Tersoff model for molecules composed purely of carbon. For two carbon atoms i and j the potential energy between them is:

$$U_{ij} = f_C(r_{ij}) \left[a_{ij} f_A(r_{ij}) - b_{ij} f_R(r_{ij}) \right] \tag{21}$$

$$f_A(r_{ij}) = A \exp\left(-\lambda_1 r_{ij}\right) \tag{22}$$

$$f_R(r_{ij}) = B \exp\left(-\lambda_2 r_{ij}\right) \tag{23}$$

where r_{ij} is the distance between i,j and $f_A(r_{ij})$, $f_R(r_{ij})$ are competing attractive and repulsive pairwise terms. $f_C(r_{ij})$ is a cutoff term that ensures only nearest-neighbor interactions. $a_{i}j$ is normally set to unity while the bond angle term b_{ij} models covalent bonding:

$$b_{ij} = \left(1 + \beta^n \zeta_{ij}^n\right)^{-1/2n} \tag{24}$$

$$\zeta_{ij} = \sum_{k \neq i,j} f_C(r_{ik}) g_{ijk} \exp\left[\lambda_3^3 \left(r_{ij} - r_{ik}\right)^3\right] \tag{25}$$

$$g_{ijk} = 1 + \frac{c^2}{d^2} - \frac{c^2}{d^2 + \left[h - \cos\left(\theta_{ijk}\right)\right]^2} \tag{26}$$

where θ_{ijk} is the angle between atoms i, j, and k. The various parameters are given for carbon.[3]

5 Derivatives

For many reasons we will want to calculate derivatives of the potential energies. First derivatives are used to calculate gradient vectors, while second derivatives are used to calculate the Hessian matrix.

We aim to calculate derivatives with respect to the x,y,z coordinates of the atoms.

5.1 Numerical Derivatives

In general derivatives can be approximated by the centered finite-difference equation:

$$\frac{df(x)}{dx} \simeq \frac{f(x + \Delta) - f(x - \Delta x)}{2\Delta x} \tag{27}$$

Partial derivatives of the potential energy function can be found similarly:

$$\frac{\partial U(\boldsymbol{q})}{\partial x_i} \simeq \frac{U(\boldsymbol{q} + \Delta x_i) - U(\boldsymbol{q} - \Delta x_i)}{2\Delta x_i} \tag{28}$$

where x_i is the x-position of atom i and $\boldsymbol{q}$ is the set of x,y,z coordinates for all of the atoms. In practice this derivative is found by moving atom i some distance s in the positive x-direction and calculating the potential energy, then doing the same after shifting it in the negative direction (just remember to shift it back to its original position!). Since only the difference in energies is required to calculate the derivative, one should only calculate the potential that is pertinent to atom i for this purpose in order to save resources.

In general numerical derivatives are more computationally expensive than analytical derivatives, so the latter is more desirable when the option is available.

5.2 Analytical Derivatives

When potential energy functions are known and well-defined, it is usually worthwhile to hard-code their analytical derivatives for the sake of gradient and Hessian calculations. We will attempt to list derivatives of the potential terms from Section 3. Finding these derivatives is a cumbersome task but is manageable by breaking it down into steps thanks to the chain rule. Refer to Appendix A for explicit derivations.

5.2.1 Summary

For now the gradient vector contributions for the atoms can be summarized below.

Bond length:

$$\frac{\partial U_{ij}}{\partial \boldsymbol{r}_i} = k_{ij} \left(R_{ij} - R_{ij0} \right) \frac{\boldsymbol{r}_i - \boldsymbol{r}_j}{R_{ij}} \tag{29}$$

Bond angle:

$$\frac{\partial U_{ijk}}{\partial \boldsymbol{r}_i} = k_{ijk} \frac{\theta_{ijk} - \theta_{ijk0}}{\sqrt{1 - \cos^2 \theta_{ijk}}} \frac{1}{R_{ij}} \left(\frac{\boldsymbol{r}_i - \boldsymbol{r}_j}{R_{ij}} \cos \theta_{ijk} - \frac{\boldsymbol{r}_k - \boldsymbol{r}_j}{R_{kj}} \right) \tag{30}$$

And similarly for k which makes sense for its relative role in the angle formation:

$$\frac{\partial U_{ijk}}{\partial \boldsymbol{r}_k} = k_{ijk} \frac{\theta_{ijk} - \theta_{ijk0}}{\sqrt{1 - \cos^2 \theta_{ijk}}} \frac{1}{R_{kj}} \left(\frac{\boldsymbol{r}_k - \boldsymbol{r}_j}{R_{kj}} \cos \theta_{ijk} - \frac{\boldsymbol{r}_i - \boldsymbol{r}_j}{R_{ij}} \right) \tag{31}$$

We cannot forget the effective gradient on the middle atom in each bond angle:

$$\frac{\partial U_{ijk}}{\partial \boldsymbol{r}_j} = -k_{ijk} \left(\theta_{ijk} - \theta_{ijk0} \right) \left(\frac{\partial \theta_{ijk}}{\partial \boldsymbol{r}_i} + \frac{\partial \theta_{ijk}}{\partial \boldsymbol{r}_k} \right) \tag{32}$$

Dihedral angle:

$$\frac{\partial \omega_{ijkl}}{\partial \boldsymbol{r}_i} = \frac{R_{kj}}{|\boldsymbol{r}_{ij} \times \boldsymbol{r}_{kj}|^2} \left(\boldsymbol{r}_{ij} \times \boldsymbol{r}_{kj} \right) \tag{33}$$

$$\frac{\partial \omega_{ijkl}}{\partial \boldsymbol{r}_l} = \frac{-R_{kj}}{|\boldsymbol{r}_{kj} \times \boldsymbol{r}_{kl}|^2} \left(\boldsymbol{r}_{kj} \times \boldsymbol{r}_{kl} \right) \tag{34}$$

$$\frac{\partial \omega_{ijkl}}{\partial \boldsymbol{r}_j} = \left(\frac{\boldsymbol{r}_{ij} \cdot \boldsymbol{r}_{kj}}{R_{kj}^2} - 1 \right) \frac{\partial \omega_{ijkl}}{\partial \boldsymbol{r}_i} - \frac{\boldsymbol{r}_{kl} \cdot \boldsymbol{r}_{kj}}{R_{kj}^2} \frac{\partial \omega_{ijkl}}{\partial \boldsymbol{r}_l} \tag{35}$$

$$\frac{\partial \omega_{ijkl}}{\partial \boldsymbol{r}_k} = \left(\frac{\boldsymbol{r}_{kl} \cdot \boldsymbol{r}_{kj}}{R_{kj}^2} - 1 \right) \frac{\partial \omega_{ijkl}}{\partial \boldsymbol{r}_l} - \frac{\boldsymbol{r}_{ij} \cdot \boldsymbol{r}_{kj}}{R_{kj}^2} \frac{\partial \omega_{ijkl}}{\partial \boldsymbol{r}_i} \tag{36}$$

The resulting contribution towards the gradient on atom a would then be:

$$\frac{\partial U_{ijkl}}{\partial \boldsymbol{r}_a} = -\sum_n n V_n \sin(n \omega_{ijkl} - \gamma) \frac{\partial \omega_{ijkl}}{\partial \boldsymbol{r}_a} \tag{37}$$

Lennard-Jones:

$$\frac{\partial U_{ij}}{\partial \boldsymbol{r}_i} = \left(\frac{6 C_6}{R_{ij}^7} - \frac{12 C_{12}}{R_{ij}^{13}} \right) \frac{\boldsymbol{r}_{ij}}{R_{ij}} \tag{38}$$

Electrostatic (point charge interaction):

$$\frac{\partial U_{ij}}{\partial \boldsymbol{r}_i} = -\frac{Q_i Q_j}{4\pi \varepsilon_0} \frac{(\boldsymbol{r}_i - \boldsymbol{r}_j)}{|\boldsymbol{r}_i - \boldsymbol{r}_j|^3} \tag{39}$$

6 Thermal Conductivity Calculations

6.1 Atomic Motion

In order to determine the thermal conductivity in a molecule, we must understand the motion of its atoms. If we assume a classical system of particles that interact linearly, we are presented with N coupled differential equations:

$$\tilde{M} \ddot{\mathbf{q}} + \tilde{\Gamma} \dot{\mathbf{q}} + \tilde{K} \mathbf{q} = \mathbf{F} \tag{40}$$

where $\tilde{M}$ and $\tilde{\Gamma}$ are the mass and drag coefficient diagonal matrices. $\tilde{K}$ is the familiar Hessian matrix filled with the respective spring constants of the system. In a 3D system, the $1st$ index refers the the x-component of the $1st$ atom while the $2nd$ and $3rd$ indices refer to the y and z portions.

If we assume a harmonic solution

$$\mathbf{q}(t) = \mathbf{a}e^{\lambda t} \tag{41}$$

we have a quadratic eigenvalue equation to solve (considering the 2nd order DE). If we define

$$\mathbf{b} = \lambda \mathbf{a} \tag{42}$$

we can retrieve a linear eigenvalue equation by doubling the dimensionality of the problem. The new eigensystem looks like

$$\begin{bmatrix} \tilde{0} & \tilde{1} \\ \tilde{K} & \tilde{\Gamma} \end{bmatrix} \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix} = \lambda \begin{bmatrix} \tilde{1} & \tilde{0} \\ \tilde{0} & -\tilde{M} \end{bmatrix} \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix} \tag{43}$$

6.2 Green's Function

We seek the Green's function for this differential operator:

$$\tilde{M}\frac{\partial^2}{\partial t^2} + \tilde{\Gamma}\frac{\partial}{\partial t} + \tilde{K} \tag{44}$$

The solution to the inhomogenous DE in Equation 40 is determined in the traditional way from the Green's function:

$$\mathbf{q}(t) = \int_0^t \tilde{G}(t, t')\mathbf{F}(t')dt' \tag{45}$$

The idea is that if we know the response of an impulse source, we can find the response to any source. Green's functions can be expanded in terms of their eigenfunctions.

$$G_{ij}(t - t') = \sum_\sigma c_{\sigma,j} a_{k,\sigma} \exp\left[\lambda_\sigma (t - t')\right] \tag{46}$$

Here $a_{k,\sigma}$ corresponds to the kth component of the σth eigenvector from Equation 43. Note that j and k will cycle 1 through N (so only the top half of the eigenvectors will be used) while σ goes from 1 to $2N$.

It can be shown* that the coefficients in Equation 46 solve the following set of $2N^2$ equations:

$$\sum_\sigma (m_i\lambda_\sigma + \gamma_i)a_{i,\sigma}c_{\sigma,k} = \delta_{ik} \tag{47}$$

$$\sum_\sigma a_{j,\sigma}c_{\sigma,k} = 0 \tag{48}$$

These equations come from the fact that $\tilde{L}\tilde{G}$ must return $\tilde{1}\delta(t - t')$ where $\tilde{L}$ is the differential operator in Equation 44. In practice these equations are incorporated in a numerical linear solver via matrices (after all these equations are dependent on each other!). A solver may ask for the A and B in the $AX = B$ problem which looks like:

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1,2N} \\ a_{21} & a_{22} & \dots & a_{2,2N} \\ \vdots & \vdots & \ddots & \vdots \\ a_{N1} & a_{N2} & \dots & a_{N,2N} \\ (m_1\lambda_1 + \gamma_1)a_{11} & (m_1\lambda_2 + \gamma_1)a_{12} & \dots & (m_1\lambda_{2N} + \gamma_1)a_{1,2N} \\ (m_2\lambda_1 + \gamma_2)a_{21} & (m_2\lambda_2 + \gamma_2)a_{22} & \dots & (m_2\lambda_{2N} + \gamma_2)a_{2,2N} \\ \vdots & \vdots & \ddots & \vdots \\ (m_N\lambda_1 + \gamma_N)a_{N1} & (m_N\lambda_2 + \gamma_N)a_{N2} & \dots & (m_N\lambda_{2N} + \gamma_N)a_{N,2N} \end{bmatrix} \tag{49}$$

$$B = \begin{bmatrix} \tilde{0} \\ \tilde{1} \end{bmatrix} \tag{50}$$

and the collection of coefficients comes out as

$$X = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1,N} \\ c_{21} & c_{22} & \cdots & c_{2,N} \\ \vdots & \vdots & \ddots & \vdots \\ c_{2N,1} & c_{2N,2} & \cdots & c_{2N,N} \end{bmatrix} \tag{51}$$

Notice how the dimensionality shifts in X: indices that corresponded to rows/components of eigenvectors in A correspond to columns in X.

6.3 Power Driving

Thermal conductivity will be related to the power driven to one atom by another atom. It can be shown* that in a steady state $(t \to \infty)$ limit this power can be calculated from the same Green's functions coefficients from above. The power driven from atom i to atom j is:

$$P_{i \to j} = K_{ij} \sum_{\sigma, \tau} \left[c_{\sigma,1} a_{i,\sigma} c_{\tau,1} a_{j,\tau} T_H + c_{\sigma,N} a_{i,\sigma} c_{\tau,N} a_{j,\tau} T_C \right] 2\gamma_0 k_B \frac{\lambda_\sigma - \lambda_\tau}{\lambda_\sigma + \lambda_\tau} \tag{52}$$

for the 1st and Nth atoms to be driven by a hot bath and cold bath respectively, with identical drag coefficients. Extra care must be taken in multiple dimensions. Understanding that this power must be zero when $T_H = T_C$, the terms on both sides should cancel out. A simplified answer therefore is

$$P_{i \to j} = 2\gamma_0 k_B (T_h - T_c) K_{ij} \sum_{\sigma, \tau} \left[c_{\sigma,1} c_{\tau,1} a_{i,\sigma} a_{j,\tau} \frac{\lambda_\sigma - \lambda_\tau}{\lambda_\sigma + \lambda_\tau} \right] \tag{53}$$

This power only corresponds to atoms that directly interact (otherwise K_{ij} would be zero). The thermal conductivity of molecules needs to be interpreted from this.

References

[1] Jon Baker. Geomerty optimization in cartesian coordinates: Constrained optimization.

[2] Wendy D. Cornell. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules.

[3] L. Lindsay. Optimized tersoff and brenner empirical potential parameters for lattice dynamics and phonon thermal transport in carbon nanotubes and graphene.

[4] Ross Moore. Features and their derivatives.

[5] William H. Press. *Numerical recipes : the art of scientific computing*. Cambridge University Press, 3 edition.

A Analytical Derivatives

Here are slightly explicit derivations of common derivatives.

A.1 Bond Length Terms

Given Equation 12, we find the first derivative wrt to x_i to be:

$$\frac{\partial U_{ij}}{\partial x_i} = \frac{\partial U_{ij}}{\partial R_{ij}} \frac{\partial R_{ij}}{\partial x_i} \tag{54}$$

$$\frac{\partial U_{ij}}{\partial R_{ij}} = k_{ij} \left(R_{ij} - R_{ij0} \right) \tag{55}$$

$$\frac{\partial R_{ij}}{\partial x_i} = \frac{x_i - x_j}{R_{ij}} = -\frac{\partial R_{ij}}{\partial x_j} \tag{56}$$

A.2 Bond Angle Terms

Considering Equation 14:

$$\frac{\partial U_{ijk}}{\partial x_i} = \frac{\partial U_{ijk}}{\partial \theta_{ijk}} \left[\frac{\partial \theta_{ijk}}{\partial R_{ij}} \frac{\partial R_{ij}}{\partial x_i} + \frac{\partial \theta_{ijk}}{\partial R_{ik}} \frac{\partial R_{ik}}{\partial x_i} \right] \tag{57}$$

$$\frac{\partial U_{ijk}}{\partial \theta_{ijk}} = k_{ijk} \left(\theta_{ijk} - \theta_{ijk0} \right) \tag{58}$$

$$\frac{\partial \theta_{ijk}}{\partial R_{ij}} = \frac{1}{\sqrt{1 - \cos^2(\theta_{ijk})}} \left[\frac{\cos(\theta_{ijk})}{R_{ij}} - \frac{1}{R_{jk}} \right] \tag{59}$$

A.3 Dihedral Angle Terms

Due to numerical instability as the torsions go to zero, we elect to use these derivatives for the dihedral angles [4]:

$$\frac{\partial \omega_{ijkl}}{\partial \boldsymbol{r}_i} = \frac{R_{kj}}{|\boldsymbol{r}_{ij} \times \boldsymbol{r}_{kj}|^2} \left(\boldsymbol{r}_{ij} \times \boldsymbol{r}_{kj} \right) \tag{60}$$

$$\frac{\partial \omega_{ijkl}}{\partial \boldsymbol{r}_l} = -\frac{R_{kj}}{|\boldsymbol{r}_{kj} \times \boldsymbol{r}_{kl}|^2} \left(\boldsymbol{r}_{kj} \times \boldsymbol{r}_{kl} \right) \tag{61}$$

A.4 Nonbonded Terms

These terms are simple enough to calculate, but there must be some consideration on how to approach the cutoff function. Nevertheless, for L-J potentials:

$$\frac{\partial U_{ij}}{\partial x_i} = \left(\frac{6C_6}{R_{ij}^7} - \frac{12C_{12}}{R_{ij}^{13}} \right) \frac{\partial R_{ij}}{\partial x_i} \tag{62}$$

We have the (negative) Coulomb's law for the positional derivative of the electrostatic energy between two point charges:

$$\frac{\partial U_{ij}}{\partial \boldsymbol{r}_i} = -\frac{Q_i Q_j}{4\pi\varepsilon_0} \frac{(\boldsymbol{r}_i - \boldsymbol{r}_j)}{|\boldsymbol{r}_i - \boldsymbol{r}_j|^3} \tag{63}$$

B Hessian Matrix

B.1 General

A Hessian matrix in general is a square matrix of second-order partial derivatives of some function. In our normal-mode analysis we are concerned with the potential function of our system. Therefore our Hessian matrix is as follows:

$$
\boldsymbol{K} = \begin{bmatrix} k_{11} & k_{12} & \dots & k_{1n} \\ k_{21} & k_{22} & \dots & k_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ k_{n1} & k_{n2} & \dots & k_{nn} \end{bmatrix} = \begin{bmatrix} \frac{\partial^2 V}{\partial x_1^2} & \frac{\partial^2 V}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 V}{\partial x_1 \partial x_n} \\ \frac{\partial^2 V}{\partial x_2 \partial x_1} & \frac{\partial^2 V}{\partial x_2^2} & \cdots & \frac{\partial^2 V}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 V}{\partial x_n \partial x_1} & \frac{\partial^2 V}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 V}{\partial x_n^2} \end{bmatrix} \tag{64}
$$

Each of these elements will act as our effective spring constants when we approximate our true potential to be that of a system of balls and springs obeying Hooke's law near the energy minimum.

B.2 Existence of Negative Elements

For an interatomic potential that is exactly parabolic is in all 3 degrees of freedom we have:

$$
V_{ij} = \frac{1}{2} k_x (x_i - x_j - x_0)^2 + \frac{1}{2} k_y (y_i - y_j - y_0)^2 + \frac{1}{2} k_z (z_i - z_j - z_0)^2 \tag{65}
$$

where k_x, k_y, k_z are positive real numbers. If our system is just 2 atoms, then this reflects the total potential in which case we can easily relate the effective spring constants to the second derivatives of the potential. Let us first take the first derivative in terms of a specifc coordinate x_i:

$$
\frac{\partial V}{\partial x_i} = k_x (x_i - x_j - x_0) \tag{66}
$$

If we take another derivative with respect to x_i we get what we expect. I am more concerned about the 'mixed' second derivative:

$$
\frac{\partial}{\partial x_j} \left(\frac{\partial V}{\partial x_i} \right) = \frac{\partial^2 V}{\partial x_j \partial x_i} = -k_x \tag{67}
$$

Even in this simple example we see that the Hessian matrix is allowed to have negative non-diagonal terms.

B.3 Relations Between Elements

B.3.1 Symmetry

We assert that our matrix is symmetric:

$$
\frac{\partial^2 V}{\partial x_i \partial x_j} = \frac{\partial^2 V}{\partial x_j \partial x_i} \tag{68}
$$

for any two coordinates x_i and x_j.

B.3.2 Self-Interaction Term

We also assert space homogeneity:

$$
k_{ij}^{aa} = -\sum_{b \neq a} k_{ij}^{ab} \tag{69}
$$

where k_{ij}^{ab} is the change of the ith component of the force on atom a when you move atom b in the jth direction. In mathematical terms:

$$
k_{ij}^{ab} = \frac{\partial^2 V}{\partial x_i^a \partial x_j^b} \tag{70}
$$

which is what we understand to be a spring-constant in the simple harmonic approximation.

The fact of Equation 69 allows one to run simple checks for spring constant matrices. After inspecting the layout of the Hessian, one see's that the sum of every element separated by 2 other elements in each row would add to 0 (approximately, ignoring floating point error). This also means the sum of each element in each row is 0. Likewise the sum of every element in the matrix should return 0. If these checks are valid (within floating point error) it does not necessarily mean that the spring constants were calculated correctly, but if they are not valid then you know you have a bug in your code!

B.4 Dispersion Relation

In an infinite lattice we can pick any unit cell and build a dynamical matrix using the simple harmonic approximation:

$$\boldsymbol{D} = \begin{bmatrix} k_{11} & k_{12} & \dots & k_{1n} \\ k_{21} & k_{22} & \dots & k_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ k_{n1} & k_{n2} & \dots & k_{nn} \end{bmatrix} \tag{71}$$

C Higher-Order Derivatives

If second-order derivatives are obtained, Hessian matrices can be found analytically.

C.1 Harmonic Bond Stretching

$$\frac{\partial^2 V}{\partial q_i \partial p_i} = \sum_{j \sim i} k_{ij} \left[\delta_{q,p} + \frac{b_0}{r_{ij}} \left(\frac{q_{ij} p_{ij}}{r_{ij}^2} - \delta_{q,p} \right) \right] \tag{72}$$

Where $q, p \in \{x, y, z\}$ and $j \sim i$ means j bonded to i. For the off-diagonal blocks:

$$\frac{\partial^2 V}{\partial q_i \partial p_j} = -k_{ij} \left[\delta_{q,p} + \frac{b_0}{r_{ij}} \left(\frac{q_{ij} p_{ij}}{r_{ij}^2} - \delta_{q,p} \right) \right] \tag{73}$$

where $i \neq j$.