

### Assignment 8: A Bit of Programming: Nonbonded Versus Bonded Energy Computations

This is a small programming assignment. At least one assignment in this modeling course should give you such first-hand experience! If you are a novice in programming, NYU staff, the TA, and the course assistant can help you, so set up an appointment with them early. You will have *two weeks* for this assignment.

#### 1. Programming Nonbonded Energy Computations.

We will begin with the nonbonded energy computations since they are most straightforward (but most expensive!)

Write a simple program to compute the nonbonded energy of a system of 1000 atoms. The nonbonded energy, Lennard Jones and Coulomb terms, should have the form:

$$E_{NONB} = \sum_{i < j} \left[ \frac{-A_{ij}}{R_{ij}^3} + \frac{B_{ij}}{R_{ij}^6} + \frac{q_i q_j}{\sqrt{R_{ij}}} \right]. \quad (\text{D.1})$$

Here  $R_{ij}$  is an interatomic distance *squared*, and  $A_{ij}$ ,  $B_{ij}$ ,  $q_i$ , and  $q_j$  are the familiar energy parameters.

For an atom  $\mathbf{x}_k$  of Cartesian components  $\{x_{k1}, x_{k2}, x_{k3}\}$ ,

$$R_{ij} = (x_{j1} - x_{i1})^2 + (x_{j2} - x_{i2})^2 + (x_{j3} - x_{i3})^2, \quad (\text{D.2})$$

and the interatomic distance  $r_{ij}$  is

$$r_{ij} = \sqrt{R_{ij}}.$$

Set up your program to read in arbitrary atomic coordinate data — a file will be sent to you electronically in case you want to use it<sup>5</sup> — and *repeat the nonbonded energy calculation* 10,000 times for all  $\{i, j\}$  pairs with  $i < j$  for  $j = 1, \dots, 1000$ . The  $R_{ij}$  calculations can be placed in some inline function.

Perform the calculations for the nonbonded energy evaluations for both single and double precision, and record the total CPU time in each case.

Also report how much CPU time and CPU percentage the square-root ( $\sqrt{\quad}$ ) operation consumes. There are special timing functions that describe the distribution of CPU time among the various program parts.

Describe the machine you are using, the precision, and attach the subroutine and program output, along with the above results.

---

<sup>5</sup>The file can be obtained through the link to the course web site or directly from the author.

## 2. Programming The Bonded Energy Computations.

Next we will write three additional subprograms to compute the bond energy, bond-angle energy, and dihedral-angle energy of a molecular system. For now, we will assume there are 1000 bonds, 1000 bond angles, and 1000 dihedral angles. We will again perform 10,000 energy evaluations of each energy term, with each sweep here involving 1000 internal variables. This large number is necessary to get reliable timing values for the bonded interactions.

You can choose in each case any representative potential form. For example, you may use:

$$E_{BOND} = \sum_{i,j \in S_B} S_{ij} (r_{ij} - \bar{r}_{ij})^2, \quad (D.3)$$

$$E_{BANG} = \sum_{i,j,k \in S_{BA}} K_{ijk} (\cos \theta_{ijk} - \cos \bar{\theta}_{ijk})^2, \quad (D.4)$$

$$E_{TOR} = \sum_{ijkl \in S_{DA}} \left( \frac{V_{ijkl}}{2} [1 + \cos(3\tau_{ijkl})] \right), \quad (D.5)$$

where the sets  $S_B$ ,  $S_{BA}$ , and  $S_{DA}$  contain all bonds, bond angles, and dihedral angles, respectively. Here  $\theta$  and  $\tau$  denote a bond angle and dihedral angle, respectively, of a given triplet or quadruplet of atoms. The values with overhead bar symbols indicate reference values.

Since we are interested only in timings for now, you can use any pairs, triplets, or quadruplets in your sample energy routines — even the same sequence — repeatedly, as long as the total number of interactions used to obtain each energy term is 1000.

Some program segments which you may find helpful are posted on the website. The derivative components are present in the angle routines, *but you do not need them* for this assignment. You can find details of the `cosba` and `cosda` subroutines in an article<sup>6</sup>.

For your convenience, an addendum to this assignment also summarizes the basic geometric relations involved in defining internal variables.

Report the CPU time required for each routine in a table, including absolute time as well as percentage of the total time of *bonded* energy components. Again, attach your programs and output to the report of the results.

---

<sup>6</sup>“A Recipe for Evaluating and Differentiating  $\cos\phi$  Expressions”, *J. Comp. Chem.* **10**, 951–956, 1989.

3. **Setting up A Polymer Model.** For obtaining realistic CPU estimates, we will now consider a simple  $n$ -alkane chain with the chemical formula  $\text{CH}_3-(\text{CH}_2)_m-\text{CH}_3$ , where  $m$  is an integer. For large  $m$ , this is polyethylene. For  $m = 2$ , for example, we have butane, chemical formula  $\text{C}_4\text{H}_{10}$ . To have about 1000 atoms, we will use  $m = 330$  for our model calculations.

Determine the number of bonds, bond angles, dihedral angles, and unique interatomic distances (atom pairs) that polyethylene has as a function of  $m$ . Consider all the distinct possibilities for the bonds and angles. Report these expressions.

Then report how many bonds, bond angles, dihedral angles, and unique atom pairs the polymer has for the case  $m = 330$ .

4. **Bonded Versus Nonbonded Energy Computations.**

Now we will combine the timings above to estimate the CPU time spent in bonded versus nonbonded energy computations for 10,000 iterations (of energy evaluations) for our polymer of 998 atoms.

Scale the timings you obtained above (10,000 iterations for 1000 atoms for the nonbonded terms, and 10,000 iterations for 1000 bonds, bond angles, and dihedral angles) so that they correspond to the numbers relevant for our polymer with  $m = 330$ , as determined in item 3 above.

Collect the data in one table which reports the CPU time and percentage required for each of the four subroutines.

What can you conclude? What can you suggest to speed up the nonbonded computations, especially if derivatives are also required?

5. **Extra Credit!**

For extra credit (the grade on this will replace your lowest homework grade), write the four subroutines above specifically for polyethylene. This means that you should use realistic coordinates, as well as correct data structures so that you consider all relevant bonds and angles for this polymer. Similarly, for energy parameters, associate values according to atom, bond, and angle types (e.g., C-C and C-H bonds, C-C-C, H-C-H, and H-C-C bond angles, and rotations about C-C bonds). You can use any resources on *Insight* to help you.

Hand in all programs and results as requested above.

---

**Addendum to Assignment 8:  
Definitions of Internal Variables in Molecules**

A bond angle  $\theta_{ijk}$  formed by a bonded triplet of atoms  $i$ - $j$ - $k$  is expressed as an inner product:

$$\cos \theta_{ijk} = \frac{(\mathbf{x}_k - \mathbf{x}_j) \bullet (\mathbf{x}_i - \mathbf{x}_j)}{r_{jk} r_{ji}}, \quad (\text{D.6})$$

or

$$\cos \theta_{ijk} = (\mathbf{r}_{jk} \bullet \mathbf{r}_{ji}) / r_{jk} r_{ji},$$

where the distance *vector* from atom  $j$  to  $i$  is given by

$$\mathbf{r}_{ji} = \mathbf{x}_i - \mathbf{x}_j = [x_{i1} - x_{j1}, x_{i2} - x_{j2}, x_{i3} - x_{j3}]^T. \quad (\text{D.7})$$

A dihedral angle  $\tau_{ijkl}$ , defining the rotation of bond  $i$ - $j$  about bond  $j$ - $k$  with respect to  $k$ - $l$ , is expressed as

$$\cos \tau_{ijkl} = \mathbf{n}_{ab} \bullet \mathbf{n}_{bc}. \quad (\text{D.8})$$

The vectors  $\mathbf{n}_{ab}$  and  $\mathbf{n}_{bc}$  denote unit normals to planes spanned by vectors  $\{\mathbf{a}, \mathbf{b}\}$  and  $\{\mathbf{b}, \mathbf{c}\}$ , respectively, where  $\mathbf{a} = \mathbf{r}_{ij}$ ,  $\mathbf{b} = \mathbf{r}_{jk}$ , and  $\mathbf{c} = \mathbf{r}_{kl}$ . Denoting  $\theta_{ab}$  and  $\theta_{bc}$  as angles  $\theta_{ijk}$  and  $\theta_{jkl}$ , respectively, we write:

$$\cos \tau_{ijkl} = \frac{\mathbf{a} \times \mathbf{b}}{\|\mathbf{a}\| \|\mathbf{b}\| \sin \theta_{ab}} \bullet \frac{\mathbf{b} \times \mathbf{c}}{\|\mathbf{b}\| \|\mathbf{c}\| \sin \theta_{bc}}. \quad (\text{D.9})$$

The sign of  $\tau_{ijkl}$  is determined by the sign of the triple scalar product  $\mathbf{a} \bullet (\mathbf{b} \times \mathbf{c})$ .

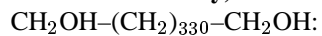
To simplify potential energy equations (and differentiation when needed) it is convenient to work with inner product expressions and use Lagrange's identity  $(\mathbf{a} \times \mathbf{b}) \bullet (\mathbf{c} \times \mathbf{d}) = (\mathbf{a} \bullet \mathbf{c})(\mathbf{b} \bullet \mathbf{d}) - (\mathbf{b} \bullet \mathbf{c})(\mathbf{a} \bullet \mathbf{d})$ . This produces the alternative expression:

$$\begin{aligned} \cos \tau_{ijkl} &= \frac{(\mathbf{a} \times \mathbf{b}) \bullet (\mathbf{b} \times \mathbf{c})}{[(\mathbf{a} \times \mathbf{b}) \bullet (\mathbf{a} \times \mathbf{b}) (\mathbf{b} \times \mathbf{c}) \bullet (\mathbf{b} \times \mathbf{c})]^{1/2}} \\ &= \frac{(\mathbf{a} \bullet \mathbf{b})(\mathbf{b} \bullet \mathbf{c}) - (\mathbf{a} \bullet \mathbf{c})(\mathbf{b} \bullet \mathbf{b})}{\left\{ [(\mathbf{a} \bullet \mathbf{a})(\mathbf{b} \bullet \mathbf{b}) - (\mathbf{a} \bullet \mathbf{b})^2] [(\mathbf{b} \bullet \mathbf{b})(\mathbf{c} \bullet \mathbf{c}) - (\mathbf{b} \bullet \mathbf{c})^2] \right\}^{1/2}}. \end{aligned} \quad (\text{D.10})$$

According to this convention,  $\tau = 0^\circ$  defines a *cis* coplanar orientation for atoms  $i$ - $j$ - $k$ - $l$ ,  $\tau = 180^\circ$  defines a *trans* coplanar orientation, and a positive sign corresponds to a clockwise rotation of the far bond with respect to the near bond (when viewed along the  $j$ - $k$  bond).

---

(See code segments on website)

**Coordinate file (available electronically) for the 1000-atom molecule**

Atom	X	Y	Z	ID
1	4.988000	2.012136	-7.818089	O
2	5.915912	2.025234	-7.572307	H
3	4.596542	0.674197	-8.131964	C
4	5.198226	0.305558	-8.962735	H
5	4.751310	0.037135	-7.261160	H
6	3.108016	0.653187	-8.526237	C
7	2.517394	1.015322	-7.710808	H
8	2.956134	1.278341	-9.381230	H
9	2.686321	-0.787809	-8.863891	C
10	2.838205	-1.412964	-8.008898	H
..	.....	.....	.....	.

Etc.