

# Some Useful Math for Molecular Mechanics

Britt H. Park

June 17, 2002

## 1 Introduction

In theory molecular mechanics is not that complicated: solve Newton's equations of motion for a molecular system (molecular dynamics), or minimize the energy of a molecular system. The energy of a molecular system is expressed as a function composed of the following terms (or any of numerous variants):

### Bond Stretching

$$E_b = K_b(r - r_0)^2 \quad (1)$$

where  $r$  is the distance between two bonded atoms.

### Dihedral Angle Bending

$$E_a = K_a(\theta - \theta_0)^2 \quad (2)$$

where  $\theta$  is the angle formed by three atoms.

### Torsion Angle Twisting

$$E_t = K_t(1 - \cos(n(\tau - \tau_0))) \quad (3)$$

where  $\tau$  is the torsion angle formed by four atoms bonded linearly and  $n$  is an integer.

### Electrostatic Potential

$$E_q = K_q \frac{q_i q_j}{r_{ij}} \quad (4)$$

where  $K_q$  is the coulombic constant,  $q_i$  and  $q_j$  are the partial charges on atoms  $i$  and  $j$  and  $r_{ij}$  is the distance between them.

### Van der Waals Potential

$$E_v = \frac{A}{r_{ij}^{12}} - \frac{B}{r_{ij}^6} \quad (5)$$

With a little analytic geometry the above energy terms are easy enough to calculate. However, in order to do something useful with a simulated molecular system one needs to be able to calculate the derivatives of the potential energy with respect to the Cartesian coordinates of the system's atoms. There have

been numerous reviews and several books published that discuss molecular mechanics. As far as I've been able to tell, however, no one has ever published the derivatives of the functions used in molecular mechanics. Therefore, in this paper I show how to calculate the derivatives of the above typical energy terms, and then go on to describe some useful details for real world molecular mechanics.

Throughout this paper I make use of the following conventions. I use  $\mathbf{a}$  to designate a vector. I use  $a_x$ ,  $a_y$  or  $a_z$  to designate the  $x$ ,  $y$  or  $z$  components of vector  $\mathbf{a}$ . I use  $r_a$  to designate the length of the vector  $\mathbf{a}$ .

## 2 Basic Strategy

The basic strategy for calculating  $\partial E/\partial x_i$ , where  $x_i$  is a single Cartesian coordinate of an atom in a system of interest for some particular function  $E$  is to calculate  $\partial E/\partial \lambda$  where  $\lambda$  is the internal coordinate for the given energy function—bond length for bond stretching terms, bond angles for angle bending terms *etc.* Then we use the chain rule

$$\frac{\partial E}{\partial x_i} = \frac{\partial E}{\partial \lambda} \frac{\partial \lambda}{\partial x_i} \quad (6)$$

to find each partial Cartesian derivative.

The pioneers of molecular mechanics had to work out their derivatives by hand which must have been time consuming considering how complicated the derivatives can get. For this paper I've chosen an easier route and used *Mathematica* to do the calculations. I had to do considerable coaxing to get *Mathematica* to produce readable expressions. Pencil and paper derivations would have taken much longer, however. If you ever need to calculate other complex derivatives I suggest that you use a program like *Mathematica*, at the very least to check your pencil and paper calculations.

## 3 Bond Stretching

Let's start by calculating the derivatives for bond stretching terms. To repeat Equation 1 we start with:

$$E_b = K_b(r - r_0)^2 \quad (7)$$

First we get

$$\frac{\partial E_b}{\partial r} = 2K_b(r - r_0) \quad (8)$$

which is about as simple as it gets. Next we want to calculate  $\partial r/\partial x_i$ . To do that we need an expression for  $r$ . Let  $\mathbf{a}$  and  $\mathbf{b}$  be the coordinates of two bonded atoms. We get

$$r = |\mathbf{a} - \mathbf{b}| \quad (9)$$

which can be expanded in terms of the individual components of  $\mathbf{a}$  and  $\mathbf{b}$  as

$$r = \sqrt{(a_x - b_x)^2 + (a_y - b_y)^2 + (a_z - b_z)^2} \quad (10)$$

Hence with some judicious substitutions we can get all 6 partial derivatives

$$\frac{\partial r}{\partial a_x} = \frac{a_x - b_x}{r} \quad (11)$$

$$\frac{\partial r}{\partial a_y} = \frac{a_y - b_y}{r} \quad (12)$$

$$\frac{\partial r}{\partial a_z} = \frac{a_z - b_z}{r} \quad (13)$$

$$\frac{\partial r}{\partial b_x} = \frac{b_x - a_x}{r} \quad (14)$$

$$\frac{\partial r}{\partial b_y} = \frac{b_y - a_y}{r} \quad (15)$$

$$\frac{\partial r}{\partial b_z} = \frac{b_z - a_z}{r} \quad (16)$$

Easy, no? Applying the chain rule to Equations 8 and 11 we get

$$\frac{\partial E_b}{\partial a_x} = \frac{2K_b(r - r_0)(a_x - b_x)}{r} \quad (17)$$

The others follow just as easily.

## 4 Angle Bending

Now things start to get a little involved. Starting from Equation 2:

$$E_a = K_a(\theta - \theta_0)^2 \quad (18)$$

we get

$$\frac{\partial E_a}{\partial \theta} = 2K_a(\theta - \theta_0) \quad (19)$$

Next we want to calculate  $\partial\theta/\partial x_i$  for each of the Cartesian coordinates. First we need an expression for  $\theta$ . Start with  $\mathbf{a}$ ,  $\mathbf{b}$  and  $\mathbf{c}$  the positions of the three atoms that form the angle. We get

$$\theta = \arccos\left(\frac{(\mathbf{a} - \mathbf{b}) \cdot (\mathbf{c} - \mathbf{b})}{|\mathbf{a} - \mathbf{b}||\mathbf{c} - \mathbf{b}|}\right) \quad (20)$$

from which we can get the 9 partial derivatives of  $\theta$  relative to the Cartesian coordinates of  $\mathbf{a}$ ,  $\mathbf{b}$  and  $\mathbf{c}$ .

$$\frac{\partial\theta}{\partial a_x} = \frac{\cos(\theta)r_{bc}(a_x - b_x) + r_{ba}(b_x - c_x)}{\sin(\theta)r_{ba}^2 r_{bc}} \quad (21)$$

$$\frac{\partial\theta}{\partial a_y} = \frac{\cos(\theta)r_{bc}(a_y - b_y) + r_{ba}(b_y - c_y)}{\sin(\theta)r_{ba}^2 r_{bc}} \quad (22)$$

$$\frac{\partial\theta}{\partial a_z} = \frac{\cos(\theta)r_{bc}(a_z - b_z) + r_{ba}(b_z - c_z)}{\sin(\theta)r_{ba}^2 r_{bc}} \quad (23)$$

Now they get a little trickier for  $\mathbf{b}$ .

$$\frac{\partial\theta}{\partial b_x} = \frac{\cos(\theta)(r_{ba}^2(b_x - c_x) + r_{bc}^2(b_x - a_x))}{\sin(\theta)r_{ba}^2 r_{bc}^2} + \frac{a_x - 2b_x + c_x}{\sin(\theta)r_{ba}r_{bc}} \quad (24)$$

$$\frac{\partial\theta}{\partial b_y} = \frac{\cos(\theta)(r_{ba}^2(b_y - c_y) + r_{bc}^2(b_y - a_y))}{\sin(\theta)r_{ba}^2 r_{bc}^2} + \frac{a_y - 2b_y + c_y}{\sin(\theta)r_{ba}r_{bc}} \quad (25)$$

$$\frac{\partial\theta}{\partial b_z} = \frac{\cos(\theta)(r_{ba}^2(b_z - c_z) + r_{bc}^2(b_z - a_z))}{\sin(\theta)r_{ba}^2 r_{bc}^2} + \frac{a_z - 2b_z + c_z}{\sin(\theta)r_{ba}r_{bc}} \quad (26)$$

The derivatives for  $\mathbf{c}$  are once again simpler.

$$\frac{\partial\theta}{\partial c_x} = \frac{\cos(\theta)r_{ba}(c_x - b_x) + r_{bc}(b_x - a_x)}{\sin(\theta)r_{ba}r_{bc}^2} \quad (27)$$

$$\frac{\partial\theta}{\partial c_y} = \frac{\cos(\theta)r_{ba}(c_y - b_y) + r_{bc}(b_y - a_y)}{\sin(\theta)r_{ba}r_{bc}^2} \quad (28)$$

$$\frac{\partial\theta}{\partial c_z} = \frac{\cos(\theta)r_{ba}(c_z - b_z) + r_{bc}(b_z - a_z)}{\sin(\theta)r_{ba}r_{bc}^2} \quad (29)$$

As for Bond Stretching the chain rule is easily applied to get expressions for the various  $\partial E_a / \partial x_i$ .

## 5 Torsion Angle Twisting

Starting from Equation 3

$$E_t = K_t(1 - \cos(n(\tau - \tau_0))) \quad (30)$$

we get

$$\frac{\partial E_t}{\partial \tau} = nK_t \sin(n(\tau - \tau_0)) \quad (31)$$

How do we calculate  $\tau$ ? Let's start with  $\mathbf{a}$ ,  $\mathbf{b}$ ,  $\mathbf{c}$  and  $\mathbf{d}$ , the coordinates of the four atoms that form the torsion angle. Let's define two vectors,  $\mathbf{e}$  and  $\mathbf{f}$  as

$$\mathbf{e} = (\mathbf{c} - \mathbf{b}) \times (\mathbf{a} - \mathbf{b}) \quad (32)$$

and

$$\mathbf{f} = (\mathbf{d} - \mathbf{c}) \times (\mathbf{b} - \mathbf{c}) \quad (33)$$

The vector  $\mathbf{e}$  is perpendicular to the plane formed by  $\mathbf{a}$ ,  $\mathbf{b}$  and  $\mathbf{c}$ . The vector  $\mathbf{f}$  is perpendicular to  $\mathbf{b}$ ,  $\mathbf{c}$  and  $\mathbf{d}$ . These give us

$$\cos(\tau) = \frac{\mathbf{e} \cdot \mathbf{f}}{|\mathbf{e}||\mathbf{f}|} \quad (34)$$

Now  $\tau$  can either be  $\arccos(\cos(\tau))$  or  $2\pi - \arccos(\cos(\tau))$ , the former if  $\mathbf{e} \cdot (\mathbf{d} - \mathbf{c})$  is positive and the latter if negative. It turns out that one gets the same values for the various  $\partial\tau/\partial x_i$  for both cases. So we'll calculate the derivatives for the first case only. Starting from

$$\tau = \arccos\left(\frac{\mathbf{e} \cdot \mathbf{f}}{|\mathbf{e}||\mathbf{f}|}\right) \quad (35)$$

we can get, through much algebraic manipulation, the 12 partial derivatives of  $\tau$  with respect to the Cartesian coordinates of the atoms that make up the coordinates. First with respect to atom  $\mathbf{a}$

$$\frac{\partial\tau}{\partial a_x} = \frac{(c_z - b_z)(\cos(\tau)r_f e_y - r_e f_y) + (b_y - c_y)(\cos(\tau)r_f e_z - r_e f_z)}{\sin(\tau)r_e^2 r_f} \quad (36)$$

$$\frac{\partial\tau}{\partial a_y} = \frac{(c_x - b_x)(\cos(\tau)r_f e_z - r_e f_z) + (b_z - c_z)(\cos(\tau)r_f e_x - r_e f_x)}{\sin(\tau)r_e^2 r_f} \quad (37)$$

$$\frac{\partial\tau}{\partial a_z} = \frac{(c_y - b_y)(\cos(\tau)r_f e_x - r_e f_x) + (b_x - c_x)(\cos(\tau)r_f e_y - r_e f_y)}{\sin(\tau)r_e^2 r_f} \quad (38)$$

Next we do the same for  $\mathbf{b}$

$$\begin{aligned} \frac{\partial\tau}{\partial b_x} &= \frac{(a_z - c_z)(\cos(\tau)r_e^2 f_y - r_e r_f f_y) + (d_z - c_z)(\cos(\tau)r_e^2 f_y - r_e r_f f_y)}{\sin(\tau)r_e^2 r_f^2} \\ &+ \frac{(c_y - d_y)(\cos(\tau)r_e^2 f_z - r_e r_f f_z) + (c_y - a_y)(\cos(\tau)r_e^2 f_z - r_e r_f f_z)}{\sin(\tau)r_e^2 r_f^2} \end{aligned} \quad (39)$$

$$\begin{aligned} \frac{\partial\tau}{\partial b_y} &= \frac{(a_x - c_x)(\cos(\tau)r_e^2 f_z - r_e r_f f_z) + (d_x - c_x)(\cos(\tau)r_e^2 f_z - r_e r_f f_z)}{\sin(\tau)r_e^2 r_f^2} \\ &+ \frac{(c_z - d_z)(\cos(\tau)r_e^2 f_x - r_e r_f f_x) + (c_z - a_z)(\cos(\tau)r_e^2 f_x - r_e r_f f_x)}{\sin(\tau)r_e^2 r_f^2} \end{aligned} \quad (40)$$

$$\begin{aligned} \frac{\partial\tau}{\partial b_z} &= \frac{(a_y - c_y)(\cos(\tau)r_e^2 f_x - r_e r_f f_x) + (d_y - c_y)(\cos(\tau)r_e^2 f_x - r_e r_f f_x)}{\sin(\tau)r_e^2 r_f^2} \\ &+ \frac{(c_x - d_x)(\cos(\tau)r_e^2 f_y - r_e r_f f_y) + (c_x - a_x)(\cos(\tau)r_e^2 f_y - r_e r_f f_y)}{\sin(\tau)r_e^2 r_f^2} \end{aligned} \quad (41)$$

Kind of hairy. So are the derivatives for **c**.

$$\begin{aligned} \frac{\partial \tau}{\partial c_x} &= \frac{(b_z - a_z)(\cos(\tau)r_f^2 e_y - r_e r_f f_y) + (b_z - d_z)(\cos(\tau)r_e^2 f_y - r_e r_f e_y)}{\sin(\tau)r_e^2 r_f^2} \\ &+ \frac{(d_y - b_y)(\cos(\tau)r_e^2 f_z - r_e r_f e_z) + (a_y - b_y)(\cos(\tau)r_f^2 e_z - r_e r_f f_z)}{\sin(\tau)r_e^2 r_f^2} \end{aligned} \quad (42)$$

$$\begin{aligned} \frac{\partial \tau}{\partial c_y} &= \frac{(b_x - a_x)(\cos(\tau)r_f^2 e_z - r_e r_f f_z) + (b_x - d_x)(\cos(\tau)r_e^2 f_z - r_e r_f e_z)}{\sin(\tau)r_e^2 r_f^2} \\ &+ \frac{(d_z - b_z)(\cos(\tau)r_e^2 f_x - r_e r_f e_x) + (a_z - b_z)(\cos(\tau)r_f^2 e_x - r_e r_f f_x)}{\sin(\tau)r_e^2 r_f^2} \end{aligned} \quad (43)$$

$$\begin{aligned} \frac{\partial \tau}{\partial c_z} &= \frac{(b_y - a_y)(\cos(\tau)r_f^2 e_x - r_e r_f f_x) + (b_y - d_y)(\cos(\tau)r_e^2 f_x - r_e r_f e_x)}{\sin(\tau)r_e^2 r_f^2} \\ &+ \frac{(d_x - b_x)(\cos(\tau)r_e^2 f_y - r_e r_f e_y) + (a_x - b_x)(\cos(\tau)r_f^2 e_y - r_e r_f f_y)}{\sin(\tau)r_e^2 r_f^2} \end{aligned} \quad (44)$$

The derivatives for **d** are simpler again.

$$\frac{\partial \tau}{\partial d_x} = \frac{(c_z - b_z)(\cos(\tau)r_e f_y - r_f e_y) + (b_y - c_y)(\cos(\tau)r_e f_z - r_f e_z)}{\sin(\tau)r_e r_f^2} \quad (45)$$

$$\frac{\partial \tau}{\partial d_y} = \frac{(c_x - b_x)(\cos(\tau)r_e f_z - r_f e_z) + (b_z - c_z)(\cos(\tau)r_e f_x - r_f e_x)}{\sin(\tau)r_e r_f^2} \quad (46)$$

$$\frac{\partial \tau}{\partial d_z} = \frac{(c_y - b_y)(\cos(\tau)r_e f_x - r_f e_x) + (b_x - c_x)(\cos(\tau)r_e f_y - r_f e_y)}{\sin(\tau)r_e r_f^2} \quad (47)$$

The chain rule applies here, too.

## 6 Electrostatic Potential

Starting with Equation 4

$$E_q = K_q \frac{q_i q_j}{r_{ij}} \quad (48)$$

we get

$$\frac{\partial E_q}{\partial r_{ij}} = -K_q \frac{q_i q_j}{r_{ij}^2} \quad (49)$$

The derivatives of  $r_{ij}$  are the same as for Bond Stretching and don't need repeating.

## 7 Van der Waals Potential

Starting with Equation 5

$$E_v = \frac{A}{r_{ij}^{12}} - \frac{B}{r_{ij}^6} \quad (50)$$

we get

$$\frac{\partial E_v}{\partial r_{ij}} = \frac{-12A}{r_{ij}^{13}} + \frac{6B}{r_{ij}^7} \quad (51)$$

The derivatives of  $r_{ij}$  are the same as for Bond Stretching and don't need repeating.

## 8 Truncating Non-Bonded Interactions

When modeling a molecule *in vacuo* one can use the energy functions described above as they stand. In more complex simulations—proteins in water for example—periodic boundary conditions are commonly used. In that case all non-bonded interactions need to be calculated for every image of every atom, in other words an infinite number of times. One solution to this problem is to truncate non-bonded energies at a cutoff distance. Naïvely truncating by simply setting energy and force to 0 for all interatomic distances more than some  $r_c$ , results in discontinuities, which, in turn, causes various distortions of the simulation. Luckily there are methods of truncating non-bonded interactions which do not introduce discontinuities and therefore distort the simulation much less. The generally accepted best way of truncating non-bonded interactions uses a switching function. Start with some radially symmetric potential,  $U_0(r)$  and a cutoff distance  $r_c$  one can construct the function:

$$U(r) = U_0(r) - (U_0(r_c) + (r - r_c)U_0'(r_c)) \quad (52)$$

This new function has the pleasing property of going to 0 at  $r = r_c$ . It has the additional pleasing property that its derivative

$$U'(r) = U_0'(r) - U_0'(r_c) \quad (53)$$

also goes to zero smoothly at  $r = r_c$ . This gives us the energy of a pair of interacting atoms as Equation 52 for all  $r < r_c$  and 0 for  $r \geq r_c$ .

The next two subsections apply this method to Van der Waals and Electrostatic potentials.

### 8.1 Truncated Van der Waals interactions

We'll repeat Equation 5 again

$$E_v = \frac{A}{r^{12}} - \frac{B}{r^6} \quad (54)$$

Normally the constants  $A$  and  $B$  are computed from  $r_0$ , the equilibrium distance between two atoms and  $\epsilon_0$  the depth of the potential energy at  $r_0$ .  $A$  and  $B$  can be expressed in terms of  $r_0$  and  $\epsilon_0$

$$A = \epsilon_0 r_0^{12} \quad (55)$$

$$B = 2\epsilon_0 r_0^6 \quad (56)$$

Now let's plug Equation 5 into Equation 52 to get

$$E_{vt} = \frac{A}{r^{12}} - \frac{B}{r^6} - \left( \frac{A}{r_c^{12}} - \frac{B}{r_c^6} + (r - r_c) \left( \frac{6B}{r_c^7} - \frac{12A}{r_c^{13}} \right) \right) \quad (57)$$

Its derivative with respect to  $r$  is

$$\frac{\partial E_{vt}}{\partial r} = -\frac{12A}{r^{13}} + \frac{6B}{r^7} + \frac{12A}{r_c^{13}} - \frac{6B}{r_c^7} \quad (58)$$

The next question to ask is what are  $A$  and  $B$  in terms of  $r_0$  and  $\epsilon_0$ . We ask *Mathematica* and get the answers:

$$A = \frac{\epsilon_0 (r_c^7 - r_0^7) r_0^{12} r_c^{12}}{r_0^{19} - 14r_0^{13} r_c^6 + 13r_0^{12} r_c^7 + 13r_0^7 r_c^{12} - 14r_0^6 r_c^{13} + r_c^{19}} \quad (59)$$

$$B = \frac{2\epsilon_0 (r_c^{13} - r_0^{13}) r_0^6 r_c^6}{r_0^{19} - 14r_0^{13} r_c^6 + 13r_0^{12} r_c^7 + 13r_0^7 r_c^{12} - 14r_0^6 r_c^{13} + r_c^{19}} \quad (60)$$

considerably less elegant than for the non-truncated Van der Waals potential.

I should introduce a word of caution here. The above expressions for  $A$  and  $B$  are correct but not in line with the way actual parameter sets are modified to include smooth truncation. For example I have used the above formulas with the parameters from ENCAD and found to my chagrin that the calculated properties of water turn out rather badly (diffusion coefficient too high, third peak in the O-O radial distribution too high). The way that ENCAD adds in smooth truncation is by multiplying the original  $A$  value by a scaling factor (0.84 for an 8Å cutoff). This has the effect of softening the repulsive term in the Van der Waals equation.

## 8.2 Truncated Electrostatic Potential

By plugging into Equation 4

$$E_q = K_q \frac{q_i q_j}{r} \quad (61)$$

and simplifying we get this following

$$E_{qt} = K_q \frac{q_i q_j (r - r_c)^2}{r r_c^2} \quad (62)$$

From this with some further simplification we get the derivative:

$$\frac{\partial E_{qt}}{\partial r} = -K_q \frac{q_i q_j}{r^2} + K_q \frac{q_i q_j}{r_c^2} \quad (63)$$

For the coulombic potential the force is shifted quite significantly, depending on the value of  $r_c$ , smaller values shifted more and larger values shifted less. This makes, for instance, bulk water behave more like a solution of high ionic strength.