Brief Article

The Author

December 18, 2023

1 Setting up the model

The challenge of finding a hard sphere that fits between three atoms requires considering the van der Waals radius of each of the atoms. We could ignore this in the calculation if the van der Waals radii are all the same and simply subtract it from the nuclear distance to the center of the calculated sphere. However, if the radii are different then they must be explicitly considered in the algorithm.

We first present the graphical model in two dimensions as it is easier to visualize and set up as compared to the three-dimensional model. Further, in two dimensions, the model is rather straightforward. We consider three, non-collinear atoms **1**, **2**, and **3** at nuclear coordinates (x_1, y_1), (x_2, y_2), and (x_3, y_3), with respective van der Waals radii of r_1, r_2, and r_3.

Obviously we have to consider instances where mathematical solutions have no chemical relevance, or cases where our chemical input may confuse a mathematical model. Since we are feeding input data from the output of a molecular dynamics simulation, these three atoms would likely not be collinear since our present input data does not include collinear groups such at an atom with a triple bond and a single bond or at an atom with two double bonds. Practical examples of such collinearity include $\mathrm{R{-}N^{-}{-}N^{+}{\equiv}N}$ azides, acetylene $\mathrm{H{-}C{\equiv}C{-}H}$, magnesium hydride $\mathrm{H{-}Mg{-}H}$, and oth-

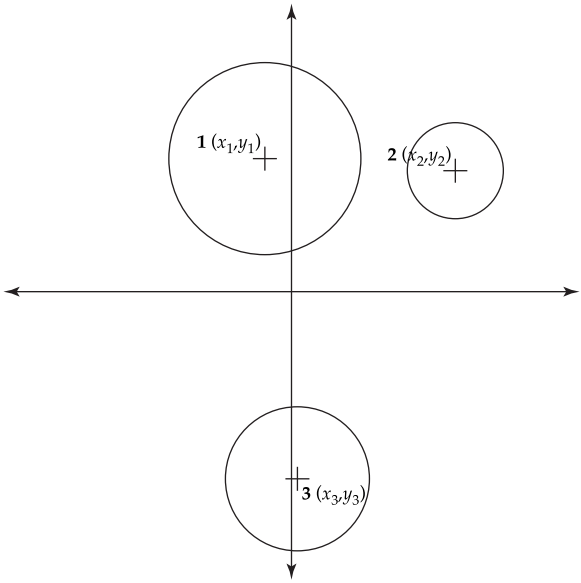

Figure 1: Three non-collinear atoms in the xy plane.

ers. Presently, we do not include such moieties, so we do not include checks for co-linearity. A future, more general version of this code should include such checks. We further recognize that there should be no atoms from our simulation that are closer to each other than 0.074 nm (0.74 Å), a dihydrogen bond length. Examples below show non-bonded atoms that are further away from each other than the sums of their van der Waals radii. Considering only non-collinear, non-bonded atoms, Fig. 1 graphically demonstrates the setup.

Again, the idea is to calculate the biggest sphere (3D) or circle (2D) that fits within three given spheres/circles. In other words, we are trying to calculate the location and size of the sphere/circle that touches our given spheres/circles *exactly once*. For our 2D case as

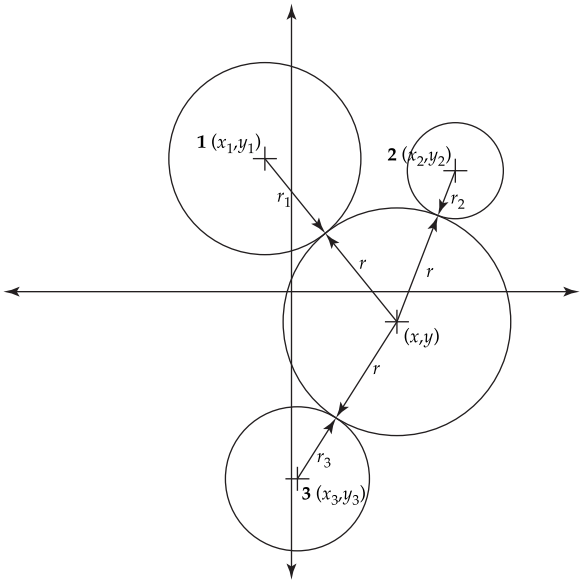

Figure 2: A circle that touches our three non-collinear atoms in the xy plane at exactly one point. Here the radius of the circle solution r is smaller than the distance between the solution center and any of the given points, **1**, **2**, or **3**.

in Fig. 1, we would be looking for a circle of radius r that is exactly $r + r_1$ away from the nucleus of **1**, $r + r_2$ away from the nucleus of **2**, and $r + r_3$ away from the nucleus of **3**. Figure 2 shows this condition graphically, and we'd want to solve for the x,y coordinates of this circle and its radius, r. One challenge is that there are *two* solutions to the problem of a circle touching our three other non-collinear circles at exactly one point each, where Fig. 3 shows this second solution. As a practical matter, we discount the solution shown in Fig. 3 for its non-physical chemical relevance. Choosing between the two solutions for r is straightforward by remembering

that a calculation may erroneously yield a negative sign for r, and simply choosing the value with the smallest $|r|$ should be the correct answer. Another check is to compare the distance between the given atoms and the center of the calculated sphere: in the Fig. 2 case, r is a smaller value than the distance between x,y and each of the three given nuclear coordinates x_1,y_1 and x_2,y_2 and x_3,y_3. However, any the non-physical solution such as in Fig. 3, the value of r is clearly greater than distance between x,y and any of the three given nuclear coordinates x_1,y_1 or x_2,y_2 or x_3,y_3. This can be determined with checks following the calculation.

The solution for the circle in Figs. 2 and 3 involves solving the three simultaneous equations in eq 1. Throughout the 2D and 3D models, we use italicized x, y, z, and r for the variables, italicized x_1, y_1, etc. for the given atom positions, italicized r_1, r_2 for the van der Waals radii, and r for the calculated radius of the circle/sphere for which we are solving. The remainder of the constants are set in Roman type.

$$(x - x_1)^2 + (y - y_1)^2 = (r + r_1)^2 \tag{1a}$$

$$(x - x_2)^2 + (y - y_2)^2 = (r + r_2)^2 \tag{1b}$$

$$(x - x_3)^2 + (y - y_3)^2 = (r + r_3)^2 \tag{1c}$$

We will not belabor the solution to the simultaneous equations in eq 1 as the two-dimensional model was only presented for visualization of the model and its impact on the equations of interest. Suffice it to say that the results of a molecular dynamics simulation would provide the "constants" for (x_1, y_1), (x_2, y_2), (x_3, y_3), and the hard-sphere van der Waals radii, r_1, r_2, and r_3 as we would know what elements correspond to each atom. We are more interested in to getting to the three-dimensional version of the simultaneous equations in eq 1. In the three-dimensional case, an extra variable exists in z, so an equation is necessary relative to the two-dimensional case. As the two-dimensional case was restricted by

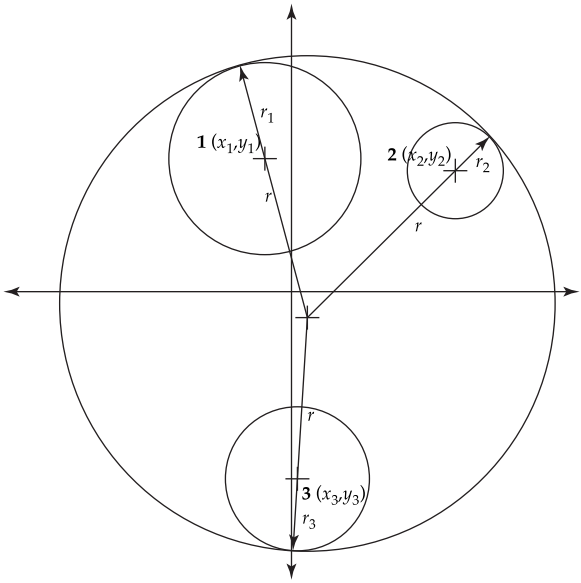

Figure 3: Another circle that touches our three non-collinear atoms in the xy plane at exactly one point. Here the radius of the circle solution r is greater than the distance between the solution center and any of the given points, **1**, **2**, or **3**.

construction to the xy plane, the additional equation in the three-dimensional case is the definition of the plane containing the points (x_1, y_1, z_1), (x_2, y_2, z_2), and (x_3, y_3, z_3). That is to say, the center of the spherical solution must lie in the plane of the three atomic nuclei. Thus eq 2 gives the three-dimensional model for a sphere that touches three non-collinear atoms at exactly one point.

$$(x - x_1)^2 + (y - y_1)^2 + (z - z_1)^2 = (r + r_1)^2 \tag{2a}$$
$$(x - x_2)^2 + (y - y_2)^2 + (z - z_2)^2 = (r + r_2)^2 \tag{2b}$$
$$(x - x_3)^2 + (y - y_3)^2 + (z - z_3)^2 = (r + r_3)^2 \tag{2c}$$
$$\mathrm{a}x + \mathrm{b}y + \mathrm{c}z - \mathrm{d} = 0 \tag{2d}$$

2 Solving the mathematical model

We attack first the solution to the family of simultaneous equations in eq 2 by determining a, b, c, and d in eq 2d from the definition of a plane given by three points, and using that to define z in terms of a, b, c, d, x, and y. Specifically, a, b, and c come from the cross product of the vectors between two pairs of the three given atomic nuclear coordinates. We define vectors **a** between points 1 and 2 as $\mathbf{a} = (x_2 - x_1, y_2 - y_1, z_2 - z_1)$ and vector **b** between points 1 and 3 as $\mathbf{b} = (x_3 - x_1, y_3 - y_1, z_3 - z_1)$. Representing unit vectors along the x, y, and z axes respectively with $\hat{i}$, $\hat{j}$, $\hat{j}$, eq 3 gives the cross product and eq 4 defines a, b, and c.

$$\mathbf{a} \times \mathbf{b} = \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ x_2 - x_1 & y_2 - y_1 & z_2 - z_1 \\ x_3 - x_1 & y_3 - y_1 & z_3 - z_1 \end{vmatrix} = (\mathrm{a}, \mathrm{b}, \mathrm{c}) \tag{3}$$

$$\mathrm{a} \equiv (y_2 - y_1)(z_3 - z_1) - (z_2 - z_1)(y_3 - y_1) \tag{4a}$$
$$\mathrm{b} \equiv (z_2 - z_1)(x_3 - x_1) - (x_2 - x_1)(z_3 - z_1) \tag{4b}$$
$$\mathrm{c} \equiv (x_2 - x_1)(y_3 - y_1) - (y_2 - y_1)(x_3 - x_1) \tag{4c}$$

With a, b, and c defined, we can use any of the points to solve eq 2d for d in eq 6. We arbitrarily choose the values from point **1** at (x_1, y_1, z_1).

$$\mathrm{a}x + \mathrm{b}y + \mathrm{c}z - \mathrm{d} = 0 \tag{2d}$$

$$\mathrm{d} = \mathrm{a}x + \mathrm{b}y + \mathrm{c}z \tag{5}$$

$$\mathrm{d} \equiv \mathrm{a}x_1 + \mathrm{b}y_1 + \mathrm{c}z_1 \tag{6}$$

Now, with a, b, c, and d being known, we can rewrite eq 2d for z in eq 7 that we will use later to take z out of eqs 2a–2c and their subsequently manipulated forms.

$$z = \frac{\mathrm{d}}{\mathrm{c}} - \frac{\mathrm{b}}{\mathrm{c}}y - \frac{\mathrm{a}}{\mathrm{c}}x \tag{7}$$

With eq 7 we have z expressed in terms of x, y, defined constants, and the coordinates of one of the given atomic nuclei. We would like to further do this with x and y to reframe eqs 2a–2c to define r. Equations 2a–2c quadratically depend on x, y, z, and r, so we can subtract these equations from each other to establish a linear dependence on x, y, z, and r. Equation 8 subtracts eq 2b from eq 2a.

$$\begin{aligned} &2x(x_2 - x_1) + x_1^2 - x_2^2 + \\ &2y(y_2 - y_1) + y_1^2 - y_2^2 + \\ &2z(z_2 - z_1) + z_1^2 - z_2^2 = 2r(r_1 - r_2) + r_1^2 - r_2^2 \end{aligned} \tag{8}$$

Equation 9 collects constants in eq 8 as p' that is defined in eq 10, and eq 11 removes z from eq 9 per eq 7.

$$2x(x_2 - x_1) + 2y(y_2 - y_1) + 2z(z_2 - z_1) = 2r(r_1 - r_2) + \mathrm{p}' \tag{9}$$

where

$$\mathrm{p}' \equiv r_1^2 - r_2^2 + x_2^2 - x_1^2 + y_2^2 - y_1^2 + z_2^2 - z_1^2 \tag{10}$$

Plugging in our definition for z from eq 7 into eq 9 yields eq 11.

$$2x(x_2 - x_1) + 2y(y_2 - y_1) + 2\left(\frac{d}{c} - \frac{b}{c}y - \frac{a}{c}x\right)(z_2 - z_1) = 2r(r_1 - r_2) + p' \quad (11)$$

Equation 12 rearranges eq 11 in terms of new constants l, m, n, and p that are respectively defined in eqs 13–16.

$$lx + my = nr + p \quad (12)$$

$$l \equiv x_2 - x_1 + \frac{a}{c}z_1 - \frac{a}{c}z_2 \quad (13)$$

$$m \equiv y_2 - y_1 + \frac{b}{c}z_1 - \frac{b}{c}z_2 \quad (14)$$

$$n \equiv r_1 - r_2 \quad (15)$$

$$p \equiv \frac{1}{2}p' + \frac{d}{c}z_1 - \frac{d}{c}z_2 \quad (16)$$

Similarly to 8, eq 17 subtracts eq 2c from eq 2a.

$$\begin{aligned} &2x(x_3 - x_1) - x_1^2 + x_3^2 + \\ &2y(y_3 - y_1) - y_1^2 + y_3^2 + \\ &2z(z_3 - z_1) - z_1^2 + z_3^2 = 2r(r_1 - r_3) + r_1^2 - r_3^2 \end{aligned} \quad (17)$$

Similar manipulations and simplifications as above can remove z from eq 17 to yield eq 18 and additional collection of constants in eqs 19–23.

$$qx + sy = tr + w \quad (18)$$

$$\mathrm{q} \equiv x_3 - x_1 + \frac{\mathrm{a}}{\mathrm{c}} z_1 - \frac{\mathrm{a}}{\mathrm{c}} z_3 \tag{19}$$

$$\mathrm{s} \equiv y_3 - y_1 + \frac{\mathrm{b}}{\mathrm{c}} z_1 - \frac{\mathrm{b}}{\mathrm{c}} z_3 \tag{20}$$

$$\mathrm{t} \equiv r_1 - r_3 \tag{21}$$

$$\mathrm{w}' \equiv r_1^2 - r_3^2 + x_3^2 - x_1^2 + y_3^2 - y_1^2 + z_3^2 - z_1^2 \tag{22}$$

$$\mathrm{w} \equiv \frac{1}{2}\mathrm{w}' + \frac{\mathrm{d}}{\mathrm{c}} z_1 - \frac{\mathrm{d}}{\mathrm{c}} z_3 \tag{23}$$

Now, with eqs 12 and 18 that depend on x, y, and r, we can combine and rearrange them to get an equation that represents x in terms of r and an equation that represents y in terms of r. To eliminate y, we multiply all terms in eq 12 by m, multiply all terms in eq 18 by s, and subtract one from the other to yield eqs 24 and 25.

$$\begin{aligned} & \mathrm{sl}x + \mathrm{sm}y = \mathrm{sn}r + \mathrm{sp} \\ - \quad & \mathrm{mq}x + \mathrm{ms}y = \mathrm{mt}r + \mathrm{mw} \\ \hline & (\mathrm{sl} - \mathrm{mq})x = (\mathrm{sn} - \mathrm{mt})r + \mathrm{sp} - \mathrm{mw} \\ & x = \frac{\mathrm{sn} - \mathrm{mt}}{\mathrm{sl} - \mathrm{mq}} r + \frac{\mathrm{sp} - \mathrm{mw}}{\mathrm{sl} - \mathrm{mq}} \end{aligned} \tag{24}$$

or

$$x = \mathrm{A}r + \mathrm{B}' \tag{25}$$

where

$$\mathrm{A} \equiv \frac{\mathrm{sn} - \mathrm{mt}}{\mathrm{sl} - \mathrm{mq}} \tag{26}$$

$$\mathrm{B}' \equiv \frac{\mathrm{sp} - \mathrm{mw}}{\mathrm{sl} - \mathrm{mq}} \tag{27}$$

To eliminate x we perform similar algebra on eqs 12 and 18 to yield eqs 28 and 29.

$$\begin{aligned} \mathrm{ql}x + \mathrm{qm}y &= \mathrm{qn}r + \mathrm{qp} \\ - \quad \mathrm{lq}x + \mathrm{ls}y &= \mathrm{lt}r + \mathrm{lw} \\ \hline (\mathrm{qm} - \mathrm{ls})y &= (\mathrm{qn} - \mathrm{lt})r + \mathrm{qp} - \mathrm{lw} \\ y &= \frac{\mathrm{qn} - \mathrm{lt}}{\mathrm{qm} - \mathrm{ls}} r + \frac{\mathrm{qp} - \mathrm{lw}}{\mathrm{qm} - \mathrm{ls}} \end{aligned} \tag{28}$$

or

$$y = \mathrm{C}r + \mathrm{D}' \tag{29}$$

where

$$\mathrm{C} \equiv \frac{\mathrm{qn} - \mathrm{lt}}{\mathrm{qm} - \mathrm{ls}} \tag{30}$$

$$\mathrm{D}' \equiv \frac{\mathrm{qp} - \mathrm{lw}}{\mathrm{qm} - \mathrm{ls}} \tag{31}$$

With eqs 25 and 29 in hand, we now have representations for x and y that we can can combine with our earlier representation of z from eq 7 in terms of the given atomic coordinates and the van der Waals radii. Substituting each of these into one of the equations between a nucleus and the radius of interest, say, eq 2a, we can work through to get r.

$$(x - x_1)^2 + (y - y_1)^2 + (z - z_1)^2 = (r + r_1)^2 \tag{2a}$$

when

$$\mathrm{B} \equiv \mathrm{B}' - x_1 \tag{32}$$

$$\mathrm{D} \equiv \mathrm{D}' - y_1 \tag{33}$$

then

$$(Ar + B)^2 + (Cr + D)^2 + \left[\left(\frac{-b}{c}C - \frac{-a}{c}A\right) + \frac{d}{c} - \frac{b}{c}D' - \frac{a}{c}B' - z_1\right]^2 = (r + r_1)^2 \quad (34)$$

or

$$(Ar + B)^2 + (Cr + D)^2 + (Er + F)^2 = (r + r_1)^2 \quad (35)$$

with

$$E \equiv \frac{-b}{c}C - \frac{a}{c}A \quad (36)$$

$$F \equiv \frac{d}{c} - \frac{b}{c}D' - \frac{a}{c}B' - z_1 \quad (37)$$

From eq 35 one simply collects terms that depend on r^2, r, and constants.

$$A^2r^2 + 2ABr + B^2 + C^2r^2 + 2CDr + D^2 + E^2r^2 + 2EFr + F^2 = r^2 + 2rr_1 + r_1^2 \quad (38)$$

$$\left(A^2 + C^2 + E^2 - 1\right)r^2 + 2\left(AB + CD + EF - r_1\right)r + \left(B^2 + D^2 + F^2 - r_1^2\right) = 0 \quad (39)$$

From eq 39, applying the quadratic formula yields a solution for r.

$$r \equiv \frac{-\left(AB+CD+EF-r_1\right)}{A^2+C^2+E^2-1} \pm \frac{\sqrt{\left(AB+CD+EF-r_1\right)^2-\left(A^2+C^2+E^2-1\right)\left(B^2+D^2+F^2-r_1^2\right)}}{A^2+C^2+E^2-1} \quad (40)$$

Equation 40 defines two possible solutions for r as shown for the 2D case in Figs. 2 and 3. Importantly, a practical implementation in a programming language that does not provide solving functions to yield eq 40 requires constants defined *in order*: a in eq 4a, b in eq 4b, c in eq 4c, d in eq 6, l in eq 13, m in eq 14, n in eq 15, p′ in eq 10, p in eq 16, q in eq 19, s in eq 20, t in eq 21, w′ in eq 22, w in eq 23, A in eq 26, B′ in eq 27, B in eq 32, C in eq 30, D′ in eq 31, D in eq 33, E in eq 36, and F in eq 37. With r in hand, spatial coordinates of the sphere come from x from eq 25, y from eq 29, and z in eq 7. Bear in mind that if the calculation yields a negative value of r and its absolute value is the r of interest, then that negative value must still be used in eqs. 25 and 29. Such implementation would obviously be quite moot in *Mathematica*, *Matlab*, *R*, or other programming languages that solve simultaneous equations, but it may remain faster even in these languages to implement this algebraic form rather than ask such programs to repeatedly solve eq 2. Comparing relative implementation speeds is beyond the scope of this study.

3 Practical pseudocode for the mathematical model

Moar words!