Polynomial Equations Arising from Apollonius Problems in Biochemistry and Pharmacology

Robert H. Lewis

Fordham University
http://www.bway.net/~lewis/
lewis@bway.net

Stephen Bridgett

Queens University, Belfast
s.bridgett@qub.ac.uk

The Apollonius Circle Problem dates to Greek antiquity, circa 250 BC. Given three circles in the plane, find or construct a circle tangent to all three. This was generalized by replacing some circles with straight lines. Descartes (and many later people) considered a special case in which all four circles are mutually tangent to each other (i.e. pairwise). In this paper we consider the general case in two and three dimensions, and further generalizations with ellipsoids and lines. We believe we are the first to explicitly find the polynomial equations describing the parameters of the solution sphere in these generalized cases. Doing so is quite a challenge for the best computer algebra systems. We report below some comparative times using various computer algebra systems on some of these problems.

Perhaps the best introduction to the Apollonius problem is in Courant and Robbins [CR], p. 125 and p. 161. However, many people worked on these questions in the twentieth century, including Boyd [B], Coxeter[C], Kasner and Supnick [KS], and Pedoe [P]. Frederick Soddy, a Nobel prize winner in Chemistry in 1921, expressed a solution to the special case as a theorem in the form of a poem, “The Kiss Precise,” which was published in the journal *Nature* [S]. Soddy proved that for four mutually tangent circles the curvatures are related by

$$2(\kappa_1^2 + \kappa_2^2 + \kappa_3^2 + \kappa_4^2) = (\kappa_1 + \kappa_2 + \kappa_3 + \kappa_4)^2$$

(which was known to Descartes) and similarly for $n+2$ mutually tangent n-spheres in n-space

$$n\left(\sum_{i=1}^{n+2}\kappa_i^2\right)=\left(\sum_{i=1}^{n+2}\kappa_i\right)^2$$

Recently Lagarias, Mallows, and Wilks [LMW] have written a paper describing packings of such circles/spheres in hyperbolic n-space and other geometries. Roanes-Lozano and Roanes-Macias [RR] have a different approach to some of the questions we ask here.

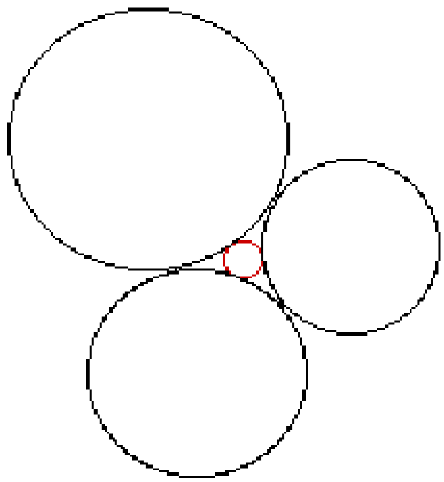

Figure 1. Special case, two dimensions, where all circles are mutually tangent, solution circle in red.

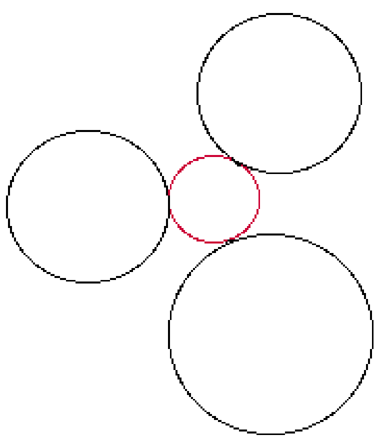

Figure 2. General case, two dimensions, solution circle is tangent to the original three, which are arbitrarily placed. In general there are eight solutions, based on each of the originals being inside or outside the solution.

Biochemical Motivation

Apollonius problems are of interest in their own right. However, the motivation for this work came originally from medical research, specifically the problem of computing the medial axis of the space around a molecule: obtaining the position and radius of a sphere which touches four known spheres or ellipsoids.

All life forms depend on interactions between molecules. X-ray crytallography and Nuclear Magnetic Resonance (NMR) yield the three dimensional structure of many large and important proteins. The coordinates of atoms within these molecules are now available from public databases and can be visualised on computer displays.

Within these molecules there are receptor sites, where a correctly shaped and correctly charged hormone or drug (often referred to as a "ligand") can attach, distorting the overall shape of the receptor molecule to cause some important action within the cell. This is similar to the concept of the correct key fitting into a lock. However, in biochemistry, the "lock" often changes shape slightly to accommodate the key. Moreover, there are often several slightly differently shaped keys that fit the lock, some binding to the receptor strongly and others binding weakly.

For instance, the shape of the blood's haemoglobin molecule means an oxygen molecule binds weakly to it, so the oxygen can be released where needed, whereas carbon monoxide, being a slightly different shape, binds very strongly to the same site on haemoglobin. This can be fatal.

Automated docking algorithms

Identifying which naturally occurring ligand molecules fit into which receptor is a very important task in biochemistry and pharmacology. Moreover, in *rational drug design* the aim is often to design drugs which better fit the receptor than the natural hormone, so blocking the action of the natural hormone. The design of the ACE inhibitor used to reduce blood pressure is a good example of this.

The use of computer algorithms to select a potential drug from thousands of known ligands for a particular receptor is receiving much research interest by several groups. Some aims for such an algorithm include:

1. To help identify potential binding sites in proteins

2. To estimate if a ligand can reach the binding site

3. To predict the ligand's most suitable orientation in the site

4. To determine how good a fit the ligand would be, compared to other known ligands

Software such as DOCK [E] is being used by pharmacheutical companies. For larger molecules, methods have been investigated such as Ackermann [A] which use a numerical best-fit estimate.

We present here another approach, the "medial axis". The medial axis is defined as the locus of the center of a maximal disc (in 2D), or sphere (in 3D) as it rolls around the interior or exterior of an object. It is effectively a skeleton of the original object, and has been found helpful in computer aided engineering for meshing, simplifying and dimensionally reducing components.

In the medial axis of a molecule the medial edges or surfaces are equidistant from the atoms which the sphere touches, and the radius of the medial sphere varies to be maximal. This is different from the Conelly surfaces, where a probe sphere of constant size is used. As the images below demonstrate, the medial axis approach leads to the Apollonius type questions of spheres touching other spheres, ellipsoids, or lines.

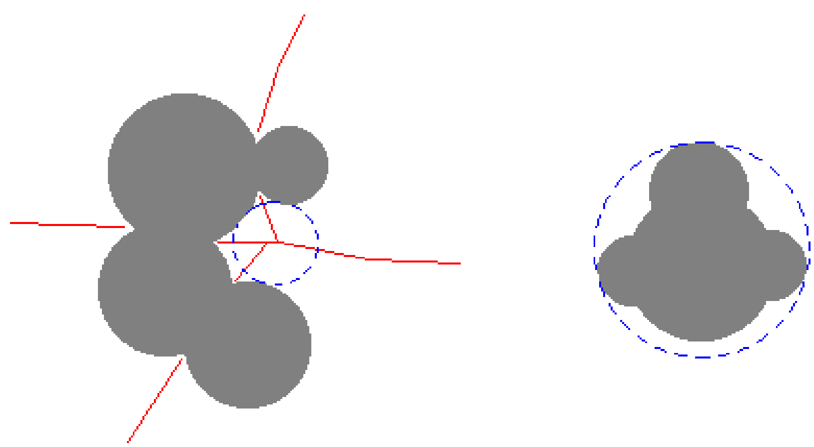

Figure 3: Left: Medial axis (lines) for a very simple 2D molecule (gray atoms) with one medial disc shown (dashed circle). Right: The diameter of the enclosing (dashed) circle around this molecule quickly confirms that it is too big to fit into the pocket in the molecule on the left.

Mathematical Approach

We are interested in the general Apollonius problem, a solution circle (or sphere) to be tangent to the original ones, which are arbitrarily placed. We want to "know" the solution symbolically and algebraically. That is, for the radius and for each coordinate of the center, we want an equation expressing it in terms of the symbolic parameters of the three (four) given circles, ellipses, or lines (spheres, ellipsoids). Ideally the equation will be of low degree so that if numerical values are plugged in for the symbolic parameters of the given shapes, a simple one variable equation will result. We will see below that in some cases we can attain this ideal, but in others we compromise by assuming that the lines or ellipsoids are in certain important but special orientations.

We begin with a system of polynomial equations

$$f_1(x, y, z, a, b, \ldots) = 0$$

$$f_2(x, y, z, a, b, \ldots) = 0$$

$$f_3(x, y, z, a, b, \ldots) = 0$$

$$\ldots\ldots$$

expressing the intersection and tangency conditions. Suppose that x, y, and z are the desired variables of the solution circle (sphere) and that $a, b, \ldots$ are the parameters of the given shapes. For each of x, y, z in turn we want to derive from the system above a single equation containing only it and the parameters $a, b, \ldots$. To accompish this elimination, we use either *Gröbner Bases* or *resultants.* Gröbner Bases are well known, resultants less so, especially the apparent method of choice, the Cayley-Bezout-Dixon-KSY method [KSY], [LS].

We shall compare solutions done with computer algebra systems Maple, Mathematica, CoCoA, and Fermat.

The Cayley-Dixon-Bezout-KSY Resultant Method

Since this method is not well known, we include a brief description. More details are in [KSY].

To decide if there is a common root of n polynomial equations in $n-1$ variables $x, y, z, \ldots$ and k parameters a, b, ...

$$f_1(x, y, z, \ldots, a, b, \ldots) = 0$$

$$f_2(x, y, z, \ldots, a, b, \ldots) = 0$$

$$f_3(x, y, z, \ldots, a, b, \ldots) = 0$$

$$\ldots\ldots$$

- Create the Cayley-Dixon matrix, $n \times n$, by substituting some new vars $t_1, \ldots, t_{n-1}$ into the equations in a certain way.

- Compute cd = determinant of the Cayley-Dixon matrix (a function of the new variables, variables, and parameters).
- Form a second matrix by extracting the coefficients relative to the new variables from cd. This matrix can be large, its size depends on the degrees of the polynomials f_i.
- Ideally, let dx = the determinant of the second matrix. If the system has a common solution, then $dx = 0$. dx involves only the parameters.
- Problem: the second matrix need not be square, or might have det $= 0$ identically. Then the method appears to fail.
- However, we may continue [KSY]: Find <u>any</u> maximal rank submatrix; let ksy = its determinant. Existence of a common solution implies $ksy = 0$.
- $ksy = 0$ is the desired equation, but one must be aware of spurious factors.

Two Dimensional Results

Problem 1: Three circles given, find a fourth tangent to all three.

First step: Given one circle A with center (ax, ay) and radius ar find equation(s) that a second "solution" circle S, (sx, sy, sr) must satisfy to be tangent. We could begin by letting (x, y) be the point of intersection. Then (x, y) is on the first circle iff $(x - ax)^2 + (y - ay)^2 = ar^2$, and on the second iff $(x - sx)^2 + (y - sy)^2 = sr^2$. Using derivatives, tangency gives a third equation, so we may eliminate x, y. This *first step* gives the equation that the six parameters must satisfy for tangency.

However the desired *first step* equation is geometrically obvious without going through the process in the above paragraph:

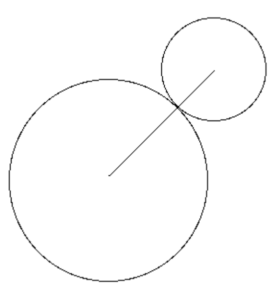

Figure 4. Two tangent circles.

From figure 4, the point of tangency is on the line connecting the centers. Each center is on the circle with the other's center and the sum of the two radii. Thus, the equation we want is obviously

$$(sx - ax)^2 + (sy - ay)^2 = (sr + ar)^2$$

So: given three circles A, B, and C, the solution circle parameters sx, sy, sr must satisfy three equations:

$$(sx - ax)^2 + (sy - ay)^2 = (sr + ar)^2$$

$$(sx - bx)^2 + (sy - by)^2 = (sr + br)^2$$

$$(sx - cx)^2 + (sy - cy)^2 = (sr + cr)^2$$

(These are the equations for the case of interest to us, when each of the original circles is outside the solution circle. Other possibilities involve changing some pluses and minuses above.)

Second Step, the Solution: There are several ways to proceed. As pointed out in [CR], one may expand the above equations and subtract to eliminate square terms, then solve for some variables and substitute back. Then too, one may simplify the equations by assuming that one of the circles, say A, is centered at the origin, as this is simply a translation, and furthermore that $ar = 1$, as this is just a change of scale. However, for consistency and comparison of later more difficult cases, let us proceed with the fully symbolic, Dixon-KSY method without substituting any constants. We used the computer algebra system Fermat [F] running on a Macintosh Blue and White G3 at 400 mhz. Think of the above as three equations in the "variables" sx and sr and the "parameters" sy plus the nine a, b, c parameters. Eliminate sx, sr and obtain the resultant in sy, i.e. the equation sy must satisfy in terms of the original data. It takes about 3 meg of RAM, 0.6 seconds. The answer, an irreducible polynomial, has 593 terms, is degree 2 in sy. Similar results obtain for sx and sr.

Problem 2: Three ellipses given, find a circle tangent to all three.

Assume the ellipses are all parallel to the axes.

Write equation of ellipse A and circle S, then tangency equations:

$$a_2^2(x - ax)^2 + a_1^2(y - ay)^2 = a_1^2 a_2^2$$

$$(x - sx)^2 + (y - sy)^2 = sr^2$$

$$a_1^2(y - ay)(x - sx) - a_2^2(x - ax)(y - sy) = 0$$

First step: Just as in problem 1, we must eliminate x, y from the above. But it is not so easy now! The point of tangency need not lie on the line connecting the centers.

Using again Fermat/Dixon/Mac, the elimination is completed using 2.1 seconds and about 3 meg RAM. The irreducible answer has 696 terms in the parameters ax, ay, a_1, a_2, sx, sy, sr. Recall that this first step was trivial in the three circles case and fit easily on one line.

Second step: For a fully symbolic solution, we would now produce three such 696 term polynomials, one for each of three given ellipses A, B, C. Then we would get an equation for, say, sr by eliminating sx and sy. However, already in the Dixon-KSY algorithm at the early step of computing cd, the determinant of the Cayley-Dixon matrix, Fermat exhausted all 700 meg of RAM on the machine. Therefore, instead of a fully symbolic solution, we made up numerical examples of three ellipses. So each of the three equations now had only the three vars sr, sx, sy, between 36 - 92 terms. Apply Dixon-KSY to get equation for sr. The second matrix was 88×88, with each entry a polynomial over the integers in sr only. The rank was 72. The determinant computation (by one method) took 5500 seconds, and the final answer had 169 terms.

Alternatively: forgo the Dixon-KSY method. We took the three equations in sr, sx, sy with 36 - 92 terms and solved numerically with a multivariate Newton's method. This works in seconds. Here is one example, showing all eight solution circles:

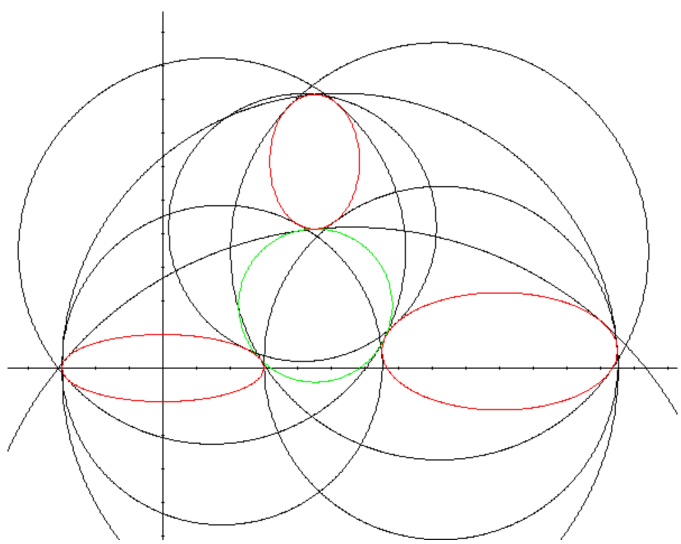

Figure 5. All eight solution circles (black and green) tangent to three given ellipses (red).

Three Dimensional Results

Problem 3: Four spheres given, find a sphere tangent to all.

Analagous to the circles in 2-space:

$$
\begin{aligned}
&(sx - ax)^2 + (sy - ay)^2 + (sz - az)^2 = (sr + ar)^2 \\
&(sx - bx)^2 + (sy - by)^2 + (sz - bz)^2 = (sr + br)^2 \\
&(sx - cx)^2 + (sy - cy)^2 + (sz - cz)^2 = (sr + cr)^2 \\
&(sx - dx)^2 + (sy - dy)^2 + (sz - dz)^2 = (sr + dr)^2
\end{aligned}
$$

By elimination, get an equation for sr in terms of the a, b, c, d vars, etc. The answer has 18366 terms and is quadartic in sr.

Several people used their favorite CAS to work on this problem. Henry Cejtin [HC] used a Groebner basis library written in Scheme. He used an elimination-ordering. To get the equation for sr took just over 1 hour and about 50 meg of RAM. Doing it for the special case where ax, ay and az are all $= 0$ took 46 seconds and used a bit under 8 meg of RAM. Solving for sx took 1.75 hours CPU time and used 60 meg RAM. In the $ax = ay = az = 0$ special case it took 300 seconds CPU time and used under 13 meg of RAM. All of these times are on an 400 MHz Pentium II machine.

Coauthor Stephen Bridgett ran Mathematica with the simplified version of the equations (ie. $ax = ay = az = 0$) with the *Solve* command. Computing sy took 4.6 hours and 41 meg RAM for the simplified case of $ax = ay = az = 0$. It could not do the full problem.

Robert Israel [RI] ran Maple 6 *with(Groebner)* on a Sun SPARC Solaris. Maple gave up after 42 minutes, using 647 meg RAM. On the simplified problem $ax = ay = az = 0$ it completed in 14 minutes using 574 meg RAM.

Coauthor Lewis tried the full problem with CoCoA 3.6 on the same Macintosh as above. The full problem was solved for sr in 95 minutess and 29 meg RAM.

Coauthor Lewis tried the full problem with Fermat/Dixon. Computing sz took 11.3 seconds, 79 meg RAM. For sx, 9.4 seconds and 5 meg RAM.

Problem 4: Four ellipsoids given, find a sphere tangent to all.

Similar to 2-D case of three ellipses. Define the ellipsoid by semi-axes a_1, a_2, a_3, center (ax, ay, az). Let (x, y, z) be the point of common tangency with a sphere sx, sy,

sz, sr. As before, the first step is to get two equations saying (x, y, z) is on the ellipsoid and sphere, and two more from partial derivatives. Try to eliminate (x, y, z). As before, we could consider first plugging in $ax = ay = az = 0$ (the reduced case).

Stephen Bridgett and Henry Cejtin both reported complete failure after many hours, even with the reduced case. Fermat/Dixon solves the reduced case rather easily, 130 seconds using 113 meg of RAM. As a challenge, Lewis had Fermat/Dixon solve the full case. It took about 400 meg of RAM and 438 minutes. Determinant cd had 1.8 million terms, and the final answer 80372 terms.

Recall this is all just step 1.

Next, as with the ellipse case in two dimensions, we made up an example and ran a multivariate Newton's method. This finished rather easily and yields, for example:

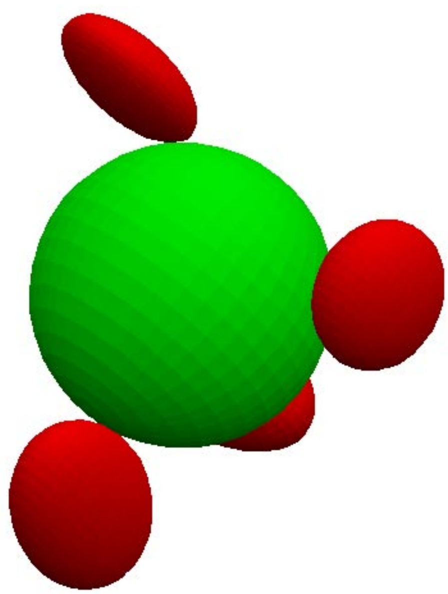

Figure 6. Solution sphere tangent to four given ellipsoids.

Problem 5: Four lines given, find a sphere tangent to all.

A line b is defined by a point on the line (bx, by, bz) and a vector (through the origin) parallel to the line, (bxn, byn, bzn).

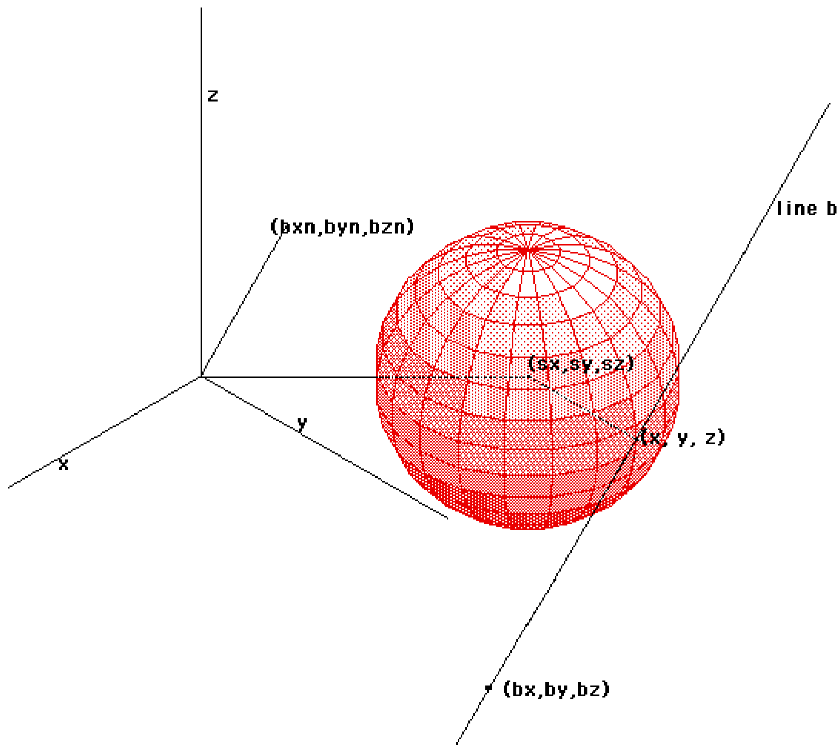

Figure 7: Parameterization of line tangent to sphere in three-space.

First step: The equations that a point (x, y, z) must satisfy if it lies on the sphere sx, sy, sz, r and the line b and is the point of tangency of the line to the sphere:

$$(x - sx)^2 + (y - sy)^2 + (z - sz)^2 - r^2 = 0$$

$$bxn(x - sx) + byn(y - sy) + bzn(z - sz) = 0$$

$$byn(z - bz) - bzn(y - by) = 0$$

$$bxn(z - bz) - bzn(x - bx) = 0$$

The first equation says (x, y, z) is on the sphere. The second equation says that the vector $(x - sx, y - sy, z - sz)$ is perpendicular to the line b. The third and fourth come from the fact that the cross product of the vectors $(x - bx, y - by, z - bz)$ and (bxn, byn, bzn) is 0.

We applied Dixon-KSY to that system, easily eliminating (x, y, z) and got the resultant $bres$:

$(-bzn^2 - byn^2 - bxn^2)r^2 + (byn^2 + bxn^2)sz^2 + (((-2byn)bzn)sy + ((-2bxn)bzn)sx + ((2by)byn + (2bx)bxn)bzn + (-2bz)byn^2 + (-2bz)bxn^2)sz + (bzn^2 + bxn^2)sy^2 + (((-2bxn)byn)sx + (-2by)bzn^2 + ((2bz)byn)bzn + ((2bx)bxn)byn + (-2by)bxn^2)sy + (bzn^2+byn^2)sx^2+((-2bx)bzn^2+((2bz)bxn)bzn+(-2bx)byn^2+((2by)bxn)byn)sx+(by^2+bx^2)bzn^2+(((-2by)bz)byn+((-2bx)bz)bxn)bzn+(bz^2+bx^2)byn^2+(((-2bx)by)bxn)byn+(bz^2 + by^2)bxn^2$

Second step: Now we need four lines a, b, c, d each with an equation like the above, $ares$, $bres$, $cres$, $dres$. For a full symbolic solution, we want four resultants, in which one of each variable sx, sy, sz, r is retained, along with all the parameters $bxn, byn, bzn, bx, by, \ldots, ax, ay, axn, \ldots, cx, \ldots$. Theoretically, we could feed all four equations into the Dixon-KSY method to get the four equations one-by-one. This seemed to be hopeless, given the 750 meg of RAM available. We added the reasonable assumption that the first line a is the x-axis, $ax = ay = az = 0$, $axn = 1, ayn = 0, azn = 0$. But this was still too large a problem. We therefore considered in addition several special cases of actual biochemical interest.

Case 1: $bx = by = 0$, $cz = dz = 0$, $bzn = 0$. line b passes through a point on the z-axis, is in a plane parallel to the xy-plane. Assume also line c and line d pass through xy-plane. Fermat/Dixon completed. The answer for sx has 394477 terms and occupies a text file of 9.5 meg.

Case 2: Three lines touch at a point, so $ax = ay = az = 0$, $bx = by = bz = 0$, $cx = cy = cz = 0$. The fourth line is oriented parallel to the x-axis but could be located anywhere, so $dx = 0$, $dzn = dyn = 0, dxn = 1$.

The second matrix in the Dixon method was 5×5. The determinant took 140 seconds to compute and had 607704 terms. It took 20 minutes to compute its content, which had 6 terms. Dividing the content yields the answer for sy. In total this took 168 meg RAM. The answer for sy has 282741 terms and occupies a text file of 5.8 meg.

Summary

We were able to find equations for two- and three-dimensional problems of the Apollonius type, i.e., finding circles or spheres tangent to some given circles, ellipses, spheres, ellipsoids, or lines. We found that the combination of Fermat/Dixon-KSY resultant is the method of choice. For several problems our solutions are fully symbolic, for others we had to add simplifying assumptions to get an answer within RAM constraints. These problems arise from the biochemistry and pharmacology of molecules fitting against other molecules. Our large polynomial solutions may be downloaded from Robert H. Lewis by writing to the email address above.

References

[A] F. Ackermann, G. Herrmann, S. Posch, and G. Sagerer, (1998) Estimation and filtering of potential protein-protein docking positions, Bioinformatics, **14**, No 2, p. 196-205.

[B] D. W. Boyd, The osculatory packing of a three-dimensional sphere. Canadian J. Math. **25** (1973), p. 303-322.

[C] H. S. M. Coxeter, The problem of Apollonius, Amer. Math. Monthly **75** (1968), p. 5-15.

[CR] R. Courant and H. Robbins. *What is Mathematics?* Oxford University Press, 1941.

[E] T. Ewing, S. Makino, A. Skillman, I. Kuntz, DOCK 4.0: Search strategies for automated molecular docking of flexible molecule databases, J. Comput. Aided Mol. Des. May 2001, 15(5) p. 411-28.

[F] Robert H. Lewis, Computer Algebra System *Fermat*. http://www.bway.net/~lewis/

[HC] Henry Cejtin, personal communication.

[KS] E. Kasner and F. Supnick, The Apollonian packing of circles, Proc. Nat. Acad. Sci. USA **29** (1943), p. 378-384.

[KSY] D. Kapur, T. Saxena, and L. Yang, Algebraic and Geometric Reasoning Using Dixon Resultants, in: *Proc. of the International Symposium on Symbolic and Algebraic Computation* (A.C.M. Press, New York, 1994).

[LMW] Lagarias, Mallows, and Wilks (in prep.). Beyond the Descartes Circle Theorem.

[LS] Robert H. Lewis and Peter F. Stiller, Solving the Recognition Problem for Six Lines Using the Dixon Resultant. Mathematics and Computers in Simulation **49** (1999) p. 203-219.

[P] D. Pedoe, On a Theorem in Geometry, Amer. Math. Monthly **74** (1967), p. 627-640.

[RI] Robert Israel, personal communication.

[RR] E. Roanes-Lozano and E. Roanes-Macias, personal communication.

[S] F. Soddy, *Nature*, June 20, 1936, p. 1021.