A System of Polynomial Equations and a Solution by an Unusual Method

Ken Rimey

University of California, Berkeley

Abstract

A system of three third order polynomial equations with parameters is presented, along with its solution by the construction of a single carefully chosen determinant. This system displays some characteristics which may be common in practice, but are not dealt with effectively by most automatic solvers.

1. Introduction

Recently, I considered the problem of solving the following system of three algebraic equations in three unknowns, x, y, and z.

$$f = 0 \ ; \ \ g = 0 \ ; \ \ h = 0 \tag{1a}$$

where

$$\begin{aligned} f &= \lambda x + \varepsilon y z - x\,(x^2 + \alpha y^2 + \beta z^2) \\ g &= \lambda y + \varepsilon z x - y\,(y^2 + \alpha z^2 + \beta x^2) \\ h &= \lambda z + \varepsilon x y - z\,(z^2 + \alpha x^2 + \beta y^2) \end{aligned} \tag{1b}$$

The solutions for any given parameter values are easily obtained, but the general solution is desired. This problem is apparently much too complex for present, exact, algebraic system solvers such as Macsyma's ‘‘algsys.’’ It is also too big to be handled by direct manipulation with Macsyma's resultant facility.

These equations originate in the work of Jim Swift at the physics department at U.C. Berkeley [1]. At the onset of convection, the motion of a rotating fluid which is being heated from below can be described by the system of three, first order, ordinary differential equations:

$$\dot{x} = f \ \ ; \ \ \dot{y} = g \ \ ; \ \ \dot{z} = h$$

Equations (1) describe the stationary points of this system.

I found an approach which easily provides the solution. Unfortunately, it seems to be applicable to few systems, and I do not have a characterization of that class of systems. Nevertheless, I present the method here because it might be used to solve some otherwise insoluble real-life problems.

A secondary purpose of this paper is to illustrate the complexity that can arise in seemingly innocent systems when they involve unspecified parameters. An important point is that the specification of a solution in terms of an irreducible polynomial is often preferable to an explicit solution involving radicals.

2. The symmetries of the system

Note that the second and third equations are obtained from the first by cyclically permuting x, y, and z. The system as a whole is invariant under the following four transformations.

$$x \to y \ , \ \ y \to z \ , \ \ z \to x \ ;$$

$$y \to -y \ , \ \ z \to -z \ ;$$

$$x \to -x \ , \ \ z \to -z \ ;$$

$$x \to -x \ , \ \ y \to -y \ ;$$

In other words, if (x, y, z) is a solution, then we obtain other solutions by cyclically permuting the components, or by flipping the signs of any two components.

3. The easy solutions

Obviously one solution is

$$x = y = z = 0$$

Consider the case where y and z vanish, but x does not. We easily find the two solutions

$$x^2 = \lambda , \ \ y = z = 0 \tag{2}$$

We deduce four more solutions from these two by symmetry, for a total of seven.

Now consider the case where x, y, and z are equal but do not vanish. We find two more solutions.

$$(1 + \alpha + \beta)\, x^2 - \varepsilon x - \lambda = 0 \, , \ \ x = y = z$$

In fact, by symmetry we have found eight more, for a total of fifteen.

A system of three equations of third order in three variables will generally have

$$3^3 = 27$$

complex solutions. Therefore, there are 12 that remain to be found. These 12 will be equivalent under the 12-fold symmetry of the system. So there actually remains only one more solution to be found, but its three components will all be different.

4. The attempt using resultants

The most well-known method for solving such systems is to eliminate variables one at a time by taking resultants. We might take the resultant of f and h with respect to z by typing the following to Macsyma.

$$res1{:}resultant(f,h,z)\$$$

This yields a ninth order polynomial in x and y which vanishes if and only if f and h have a common solution for z (in terms of x, y, and the parameters). Then we take the resultant of g and h with respect to z:

$$res2{:}resultant(g,h,z)\$$$

Finally, we eliminate y from the two resultant to obtain a polynomial in only x:

$$bigres{:}resultant(res1,res2,y)\$$$

The procedure would be to find the roots of this polynomial, and then proceed in the obvious way to find the corresponding values of y and z. Actually, it is necessary to consider separately the case where the leading coefficients of $res1$ and $res2$, viewed as polynomials in y, vanish simultaneously, but we will not discuss this.

Not all x satisfying

$$bigres = 0$$

will correspond to solutions of the system. The fact that x and y satisfy

$$res1 = 0, \quad res2 = 0$$

implies only that f and h have a common root for z, and that g and h have a common root, but not that f, g, and h have a common root. This fact is easily forgotten by users of this resultant technique.

This is the approach taken by algsys. A basic problem is that $bigres$ will contain an irreducible factor of tenth order with coefficients involving parameters. There is no way to find the roots of such a thing in general, and therefore, though these ten roots are in fact spurious, algsys is doomed to fail.

Actually, algsys will not get this far. Bigres is far too big to be computed.

We can do better by noting that we can remove a factor of x from $res1$, and a factor of y from $res2$, and that then $res1$ and $res2$ contain only even powers of x and y. If we invent new variables

$$u = x^2, \quad v = y^2$$

it becomes feasible to take the resultant of the new $res1$ and $res2$ with respect to v. The result is a huge 16th order polynomial in u. It contains several factors corresponding to the solutions which we have already found. It is a simple matter to divide them out.

What remains is the product of a third order and a tenth order polynomial in u. The first gives us the three components of the final solution. The second is entirely spurious. Unfortunately, we only have the product of these two polynomials, not the two polynomials themselves; and the product is too large to factor.

In fact, in a relatively compact internal form it is over one megabyte in size. The difficulty of attempting to compute $bigres$ becomes clear when we realize that it would involve the square of this polynomial. Squaring a polynomial in 5 variables would be expected to increase its size by about 2^5.

5. The alternate approach

Let us view f, g, and h as polynomials in y and z:

$$f = [x(\lambda - x^2)] - [\alpha x]y^2 + [\varepsilon]yz - [\beta x]z^2$$

$$g = [\lambda - \beta x^2]y + [\varepsilon x]z - [1]y^3 - [\alpha]yz^2$$

$$h = [\varepsilon x]y + [\lambda - \alpha x^2]z - [\beta]y^2 z - [1]z^3$$

Note that f involves only zeroth and second order terms, while g and h involve only first and third order terms. Consider the following system of ten equations:

$$y^3 f = 0, \quad y^2 z f = 0, \quad y z^2 f = 0, \quad z^3 f = 0$$

$$y^2 g = 0, \quad y z g = 0, \quad z^2 g = 0$$

$$y^2 h = 0, \quad y z h = 0, \quad z^2 h = 0$$

These equations involve only third and fifth order terms. Think of them as a system of ten **linear** equations for the ten unknowns:

$$y^5, \; y^4 z, \; y^3 z^2, \; y^2 z^3, \; y z^4, \; z^5$$

$$y^3, \; y^2 z, \; y z^2, \; z^3$$

This system has a solution for y and z, not both zero, only if the determinant of the 10 by 10 coefficient matrix vanishes. This determinant is a 12th order polynomial in x. It has some factors which correspond to the solutions we have already found. What remains is the following irreducible cubic in x^2:

$$(\alpha\beta - 1)^2 \gamma^2 x^6 - (\alpha\beta - 1)\gamma^2 [\varepsilon^2 + (\alpha + \beta - 2)\lambda] x^4$$
$$+ \gamma\delta[(1 + \alpha + \beta)\varepsilon^2 + \gamma\lambda] x^2 - \varepsilon^2 \delta^2 = 0$$

where

$$\gamma = \frac{(\alpha - \beta)^2 + (\alpha - 1)^2 + (\beta - 1)^2}{2}$$

$$\delta = (\alpha - 1)(\beta - 1)\lambda - \varepsilon^2$$

The values of x^2, y^2, and z^2 for the remaining solutions are given by the three roots of this cubic. The correct choice of signs and ordering can be deduced by an analysis of the bifurcation of these solutions from the six solutions derived from equations (2).

The quantity $x^2 + y^2 + z^2$, which is the same for all twelve solutions, is of particular physical interest. It is the sum of the three roots of our cubic in x^2. This is just the negative of the coefficient of x^4 when the leading coefficient is normalized to one.

$$x^2 + y^2 + z^2 = \frac{\varepsilon^2 + (\alpha + \beta - 2)\lambda}{\alpha\beta - 1}$$

Despite considerable effort, Jim Swift was unable to derive this result by hand, and subsequently asked for my assistance. We later discovered that A. M. Soward, in similar but independent work on the rotating convection problem [2], had also solved equations (1). He expressed some surprise on hearing that we had also succeeded, as he had accomplished this by some rather lengthy and difficult hand-calculations.

6. Is this approach useful for other systems?

The first problem one faces is in trying to find a set of multipliers that will produce a linear system of n equations in the same number n of ''unknowns''. These unknowns might be monomials, or any set of n linearly independent polynomials in the variables one wishes to eliminate. First, it is not obvious that such a set of multipliers exists. Second, this method will not be worthwhile unless n is small.

The second problem is that it is entirely possible that the determinant will vanish identically.

These determinants bear a resemblance to those that occur in the generalization of the theory of resultants to more than two polynomials [3]. In order to compute the generalized resultant of our three polynomials in y and z, one would first construct a certain 35 by 28 matrix analogous to our 10 by 10 matrix. The generalized resultant is the greatest common divisor of the determinants of the 28 by 28 submatrices of this matrix. In other words, the generalized resultant is the condition that this matrix have submaximal rank.

Clearly our method is in effect an adaptation of this procedure to take advantage of the fortunate nature of our particular system of equations. However, it may be that systems of such a fortunate nature are common.

I welcome suggestions from others who might provide general methods that can solve systems of this complexity, perhaps by taking into account symmetries or known solutions. I also welcome reports of successes or failures of this approach on other problems.

7. Acknowledgments

This work was supported in part by DOE Contract DE-AM03-76SF00034, P.A. DE-AT03-76E10358, and in part by the System Development Foundation. Discussions with Jim Swift, David R. Barton, and Profs. Fateman and Kahan were useful in clarifying some of the issues in this note.

8. References

[1] J. W. Swift, ''Convection in a Rotating Fluid Layer'', in ''Contemporary Mathematics Series'', vol. 28, J. Marsden, ed., American Mathematical Society, 1984.

[2] A. M. Soward, ''Bifurcation and Stability of Finite Amplitude Convection in a Rotating Layer'', submitted to Physica D.

[3] F. S. Macaulay, ''The Algebraic Theory of Modular Systems'', Cambridge University Press, 1916, reprinted as ''Cambridge Tracts in Mathematics and Mathematical Physics'', No. 19, 1964.