Formulae, Algorithms, and Quartic Extrema

D. J. Jeffrey
Department of Applied Mathematics
The University of Western Ontario
London, Ontario, Canada N6A 5B7

Introduction

The reader probably thinks that finding the turning points of a quartic polynomial is a 'solved problem', so the first thing to establish is the possibility of improving on the standard method, which is taken to be the calculus one. Thus, given a real polynomial

$$P(x) = a_4x^4 + a_3x^3 + a_2x^2 + a_1x + a_0 \ , \tag{1}$$

whose turning points are required, one sets the derivative to zero and solves the cubic

$$P'(x) = 4a_4x^3 + 3a_3x^2 + 2a_2x + a_1 = 0 \ . \tag{2}$$

This will have 1 or 3 real roots; if there are 3 real roots, one must test each one separately to ascertain the nature of the turning point, and hence finally obtain the global extremum. The process of finding extrema has been automated by several computer algebra systems (CAS), to a greater or lesser extent.

So what is there to improve? If the coefficients are known numerically, then not very much, but suppose that some of the coefficients a_i are known only symbolically. Consider the ways in which mathematicians give solutions to problems containing symbolic parameters. The first way is to give a formula, meaning an explicit function of the parameters. This is the method fixed in the popular imagination: spies in espionage stories have always chased after *the formula.* For example, the infimum of a quadratic polynomial, a result used below, is given by the following formula. If $a_2 > 0$, then

$$\inf(a_2x^2 + a_1x + a_0) = a_0 - \frac{a_1^2}{4a_2} \ , \tag{3}$$

and, moreover, the value of x that gives this infimum is $x_\mathrm{f} = -\frac{1}{2}a_1/a_2$, another formula. We recall that this well-known result can be deduced without calculus by completing the square. Among CAS, Maple is able to return (3) through its `minimize` command, although it has no syntax for returning the position x_f.

In contrast to the solution of the quadratic problem, the solution of the quartic problem was given above in the form of a *procedure* or *algorithm*, not a formula.

Theorem 1 *Formulae are better than algorithms.*

Proof: There has never been a story or film in which rival groups of spies chase, explode and kill each other in order to gain possession of an algorithm. For a formula, on the other hand, they will "stop at nothing". Q.E.D.

Of course, many think that spy stories rely too much on formulas, but that is a different topic. Certainly computer algebra systems prefer formulae, and their syntax has to be stretched in order to accommodate algorithms. Maple does this when the command `minimize`$(P(x), x)$ returns

$$\frac{1}{16a_4}\left[(8a_4a_2 - 3a_3^2)X^2 + (12a_4a_1 - 2a_3a_2)X - a_3a_1\right] + a_0 \ ,$$

with

$$X = \mathtt{RootOf}(4a_4Z^3 + 3a_3Z^2 + 2a_2Z + a_1, Z) \ .$$

The first argument of the Maple function `RootOf` is the equation to solve, and the second argument is the variable to solve for; then X can be any one of the roots of the given equation. Notice, in passing, that Maple has used the properties of X to reduce the quartic (1) to a quadratic expression.

We can now pose the problem to be considered: can the algorithm (1) and (2) be replaced by a formula, or at least some formulae. Not only is the answer *yes*, but you will not have to brave a single scorpion to learn it (just perhaps a few cobwebs).

Formula one

As an opening, we can reduce the number of unknowns we have to face by taking out a few coefficients. Dividing through by a_4 makes the polynomial in (1) *monic*, and one term can be disposed of by shifting to the variable $y = x + \frac{1}{4}a_3/a_4$. Actually, a_3 and a_4 are like minor spies at the start of the movie: they were included so that you would be impressed by their swift elimination. Any constant terms simply raise or lower everything and can be added on at the end. Cast into formal language, these transformations become a lemma.

Lemma 2 *Provided $a_4 > 0$, the infimum of a general quartic is given by*

$$\inf(a_4x^4 + a_3x^3 + a_2x^2 + a_1x + a_0) = a_4\inf(y^4 + 3b_2y^2 + 2b_1y) + a_0 + \beta_0 \ ,$$

where the factors 3 and 2 simplify later formulae, and

$$b_2 = \frac{a_2}{3a_4} - \frac{a_3^2}{8a_4^2} \ , \qquad b_1 = \frac{a_1}{2a_4} + \frac{a_3^3}{16a_4^3} - \frac{a_2a_3}{4a_4^2} \ , \qquad \beta_0 = \frac{a_2a_3^2}{16a_4^2} - \frac{a_1a_3}{4a_4} - \frac{3a_3^4}{256a_4^3} \ .$$

The derivation of the minimum of the reduced polynomial $y^4 + 3b_2y^2 + 2b_1y$ has to treat the cases $b_1 = 0$ and $b_1 \neq 0$ separately. So starting with the more general case $b_1 \neq 0$, we can plant our first formula in a theorem.

Theorem 3 *If the coefficient $b_1 \neq 0$, the quartic polynomial*

$$P_4(y) = y^4 + 3b_2y^2 + 2b_1y \tag{4}$$

has an infimum on the real line given by

$$\inf P_4 = M(b_1, b_2) = -\tfrac{3}{4}(k_\mathrm{f} - b_2)(k_\mathrm{f} - 3b_2) \ , \tag{5}$$

where

$$k_\mathrm{f} = s^{1/3} + b_2^2 s^{-1/3} + b_2 \ , \tag{6}$$

$$s = b_1^2 + b_2^3 + \sqrt{b_1^4 + 2b_1^2b_2^3} \ , \tag{7}$$

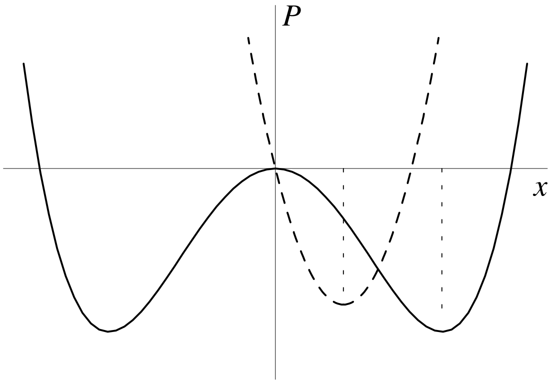

Figure 1: The splitting of the quartic into two parts. The fourth-degree part (solid line) is symmetric. The positions of the minima vary with k. As k decreases, the parabola (dashed) moves out. For the value of k that makes the two minima coincide, the minimum of the original quartic is obtained.

and $s^{1/3}$ and $s^{-1/3}$ are always interpreted as the principal values of the powers. Moreover, the infimum of P_4 is located at $y = y_{\rm f} = -b_1/k_{\rm f}$.

Proof: The idea is to break $P_4(y)$ into two pieces in such a way that each piece can be minimized easily. Then one uses the fact that any two polynomials $f(y)$ and $g(y)$, both bounded below, obey

$$\inf(f(y)+g(y)) \geq \inf f(y) + \inf g(y) \ ,$$

with equality holding when the same value of y minimizes both f and g. When we split P_4 into two polynomials, denoted $P_4^{(1)}$ and $P_4^{(2)}$, we also introduce a parameter k, which is assumed to satisfy $k > 0$.

$$P_4 = [y^4 + (3b_2 - k)y^2] + [ky^2 + 2b_1 y] = P_4^{(1)} + P_4^{(2)} \ .$$

So long as k is positive, the parabola $P_4^{(2)}$ has a minimum. Figure 1 illustrates the splitting of P_4. The turning points of the two parts can now be found without calculus. Completing the square can be used to rewrite $P_4^{(1)}$ as

$$P_4^{(1)} = y^4 - (k - 3b_2)y^2 = \left[y^2 - \tfrac{1}{2}(k - 3b_2)\right]^2 - \tfrac{1}{4}(k - 3b_2)^2 \ .$$

If it is additionally assumed that $k - 3b_2 > 0$, then $P_4^{(1)}$ has the minimum $-\frac{1}{4}(k - 3b_2)^2$ at the points where $y^2 = \frac{1}{2}(k - 3b_2)$. The infimum of $P_4^{(2)}$ is $-b_1^2/k$, by (3), and therefore

$$\inf(P_4) \geq -b_1^2/k - \tfrac{1}{4}(k - 3b_2)^2 \ . \tag{8}$$

If a value of k can be found that forces the two points where $P_4^{(1)}$ and $P_4^{(2)}$ have minima to coincide, then we have also located the infimum of P_4. Of course, any such value for k will also have to satisfy our accumulated assumptions $k > 0$ and $k > 3b_2$. The turning point for $P_4^{(2)}$, being at $y = -b_1/k$, moves closer to the origin as k increases, while that of $P_4^{(1)}$, being at $y^2 = \frac{1}{2}(k - 3b_2)$ moves away. So they will coincide when k takes a value such that

$$\tfrac{1}{2}(k - 3b_2) = (-b_1/k)^2 \ .$$

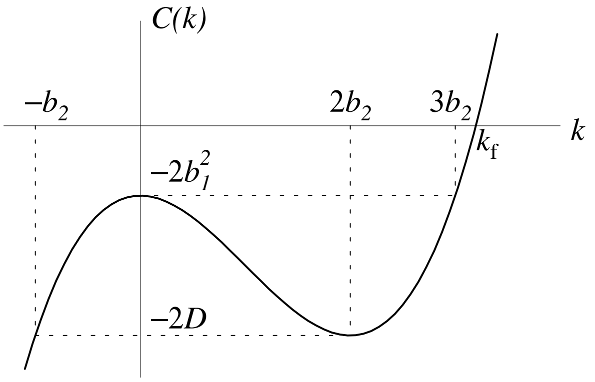

Figure 2: A graph of the auxiliary cubic for $b_2 > 0$. It can be seen that the root $k_{\rm f}$ is positive and greater than $3b_2$ as required. The quantity $D = b_1^2 + 2b_2^3$ is positive.

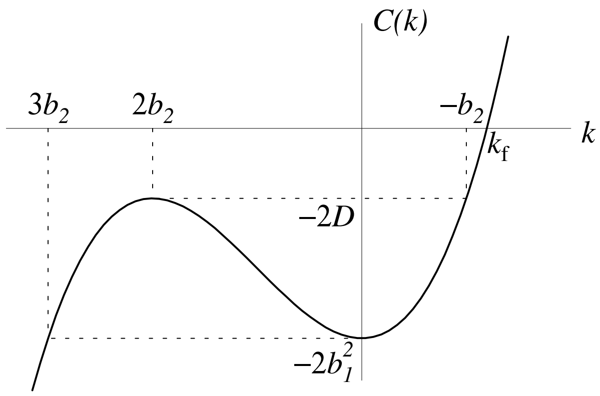

Figure 3: A graph of the auxiliary cubic for $b_2 < 0$. The quantity $D = b_1^2 + 2b_2^3$ is positive, so there is only one real root.

This is equivalent to the cubic equation,

$$k^3 - 3b_2k^2 - 2b_1^2 = 0 \ . \tag{9}$$

I shall call $C(k) = k^3 - 3b_2k^2 - 2b_1^2$ the auxiliary cubic. Another way to get this equation is to use calculus to maximize the right side of (8) directly. The cubic equation (9) has a unique positive solution. The intermediate value theorem could be used to show this analytically, but the equation plays a central role in what follows, so it is better to understand its properties graphically. Figures 2 – 4 show plots of $C(k)$ for the different ranges of the parameters. The twists in the plots can be understood by noticing that $C(0) = C(3b_2) = -2b_1^2$ and $C'(0) = C'(2b_2) = 0$. Also $C(2b_2) = C(-b_2) = -2b_1^2 - 4b_2^3 = -2D$, after introducing the abbreviation $D = b_1^2 + 2b_2^3$.

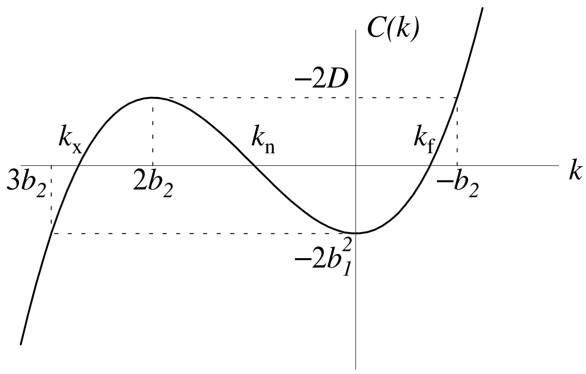

Figure 4: A graph of the auxiliary cubic for $b_2 < 0$. The quantity $D = b_1^2 + 2b_2^3$ is negative, so there are 3 real roots. They are marked on the figure as k_x, k_n and k_f.

Denote the positive solution of (9) by k_f. Since $k_\mathrm{f} > 3b_2$, the minimum of (4) is obtained from (8) as

$$\inf(P_4) = -b_1^2/k_\mathrm{f} - \tfrac{1}{4}(k_\mathrm{f} - 3b_2)^2 \ .$$

The first term of this expression is better transformed so that b_1 does not appear explicitly, for reasons that will be given below. The transformation is made by rewriting (9) in the form

$$\tfrac{1}{2}k^2 - \tfrac{3}{2}b_2 k = b_1^2/k \ , \tag{10}$$

and hence (5) is obtained.

It remains to find an explicit formula for k_f. The expression (6) is a standard solution of (9), but the fact that it gives the positive solution of (9) must be verified. Rewrite (7), introducing D, as $s = b_1^2 + b_2^3 + \sqrt{b_1^2 D}$. First, consider the case $b_2 \geq 0$ and $D \geq 0$; all terms in (6) are real and positive. Second, consider $b_2 < 0$ and $D \geq 0$, meaning $b_1^2 \geq -2b_2^3$. Then $s > -2b_2^3 + b_2^3 + \sqrt{b_1^2 D} > -b_2^3$, and therefore $s^{1/3} > -b_2 > 0$, and $k_\mathrm{f} > 0$. Finally, if $b_2 < 0$ and $D < 0$, then s will be complex, explicitly $s = b_1^2 + b_2^3 + i\sqrt{-b_1^2 D}$. Squaring and adding the real and imaginary parts shows $|s|^2 = b_2^6$, and so in polar form $s = -b_2^3 e^{i\theta}$ with $0 < \theta < \pi$, since the imaginary part is positive and therefore in the upper half-plane. Then

$$s^{1/3} + b_2^2 s^{-1/3} = -b_2 e^{i\theta/3} - b_2 e^{-i\theta/3} = -2b_2 \cos\tfrac{1}{3}\theta \ , \tag{11}$$

and since $2\cos\frac{1}{3}\theta > 1$, the value of k_f is real and positive.

The special case $b_1 = 0$ must now be taken up, but it is an easy one because there are only two terms in the polynomial.

Theorem 4 *For the case $b_1 = 0$, the polynomial $P_4(y) = y^4 + 3b_2 y^2$ has the infimum*

$$\inf P_4 = -\tfrac{9}{4}\min(0, b_2)^2 \ ,$$

at the points $y^2 = -\frac{3}{2}\min(0, b_2)$.

Proof: If $b_2 \geq 0$, then clearly the minimum is 0 when $y = 0$. If $b_2 < 0$, then completing the square can be used again to give the minimum as $-\frac{9}{4}b_2^2$ at $y^2 = -\frac{3}{2}b_2$. The theorem uses the minimum function to combine these cases.

Now there is an interesting development. One takes formula (5) for $M(b_1, b_2)$, forgets that it was derived for $b_1 \neq 0$, and substitutes $b_1 = 0$. For the case $b_2 > 0$, one computes $k_{\mathrm{f}} = 3b_2$ and $M(0, b_2) = 0$. For $b_2 < 0$, equation (11) can be reused with $\theta = \pi$, making $s = -b_2$ and $k_{\mathrm{f}} = 0$. Then $M(0, b_2) = -\frac{9}{4}b_2^2$. Thus for these cases, M continues to give the correct result. This is because of the transformation (10). For $b_2 = 0$, k_{f} contains a term $0/0$ and this prevents a simple substitution from obtaining $M(0, 0) = 0$, or in other words, $M(b_1, b_2)$ has a removeable singularity at $b_1 = b_2 = 0$. Unfortunately, this trick cannot be repeated for the position of the infimum y_{f}, and that has to remain having a piecewise definition.

$$y_{\mathrm{f}}(b_1, b_2) = \begin{cases} -b_1/k_{\mathrm{f}} \ , & b_1 \neq 0, \\ \sqrt{-\min(0, 3b_2/2)} \ , & \text{otherwise.} \end{cases}$$

The positive root has been arbitrarily chosen for definiteness in the case $b_1 = 0$.

A secondary formula.

A quartic polynomial can have 3 turning points, corresponding to (2) having 3 real roots. It is also possible for (9) to have 3 real roots. Is there a connection? At first sight, it seems not, because k was assumed to be positive, and there is only one positive solution of (9). In spite of this doubt, negative values of k do indeed give the other turning points. First we give a formula for the secondary minimum.

Theorem 5 *If the coefficient $b_1 \neq 0$, and $D = b_1^2 + 2b_2^3 < 0$, the quartic polynomial $P_4(y)$ defined in (4) has a secondary minimum $N(b_1, b_2)$ equal to*

$$N(b_1, b_2) = 3b_2 k_{\mathrm{n}} - \tfrac{3}{4}k_{\mathrm{n}}^2 - \tfrac{9}{4}b_2^2 \ ,$$

where

$$k_{\mathrm{n}} = s^{1/3}e^{-2\pi i/3} + b_2^2 s^{-1/3}e^{2\pi i/3} + b_2 \ ,$$

and s is unchanged from equation (7). Moreover, the secondary minimum is located at $y = y_{\mathrm{n}} = -b_1/k_{\mathrm{n}}$.

Proof: The only quantity that is different from theorem 3 is k_{n}, which is a different solution of (9). Figure 4 illustrates that it is the root satisfying $2b_2 < k_{\mathrm{n}} < 0$, as we now show. For the given range of parameters, $s = -b_2^3 e^{i\theta}$, where $0 < \theta < \pi$, as in (11). Therefore $k_{\mathrm{n}} = b_2(1 - 2\cos\phi)$, with $-\frac{2}{3}\pi < \phi \leq -\frac{1}{3}\pi$. Even if the parameter k is negative, it is still true that $P_4 = P_4^{(1)} + P_4^{(2)}$, and the derivatives of P_4 can be calculated by adding those of $P_4^{(1)}$ and $P_4^{(2)}$. The derivatives of P_4 at $y = y_{\mathrm{n}}$ are thus computed to be

$$\frac{dP_4}{dy}(y_{\mathrm{n}}) = 0 \qquad \text{and} \qquad \frac{d^2P_4}{dy^2}(y_{\mathrm{n}}) = 6k_{\mathrm{n}} - 12b_2 \geq 0 \ .$$

Therefore, the point y_{n} is a local minimum, but not the infimum, which corresponds to a positive value of k.

By now, it is clear that the third root of $C(k)$ gives the relative maximum between the two minima.

Theorem 6 *With the notation already defined, P_4 has a relative maximum when $D < 0$ and k_{x} is the root of $C(k)$ satisfying $3b_2 < k_{\mathrm{x}} < 2b_2$. The formula for k_{x} is*

$$k_{\mathrm{x}} = s^{1/3}e^{2\pi i/3} + b_2^2 s^{-1/3}e^{-2\pi i/3} + b_2 \ .$$

Properties of the solutions

In spy stories, everyone seems to think that possessing the formula is all that is required — a bit like a weak student facing a mathematics exam. However, one must be able to use it. The formulae just derived can be used to show that the turning points of P_4 have some interesting properties. For example, the sign of b_1 is all that decides the side of the origin on which the infimum lies; if there is a secondary minimum, then both it and the local maximum are always on the same side of the origin, and that is the opposite side from the infimum; the infimum is always further away from the origin than the secondary minimum.

After several pages of algebra, it is always comforting to try a few numerical examples and see that everything works out. No one wants a repeat of the last scenes of *The Maltese Falcon*. In addition, the examples here carry some useful lessons of their own. So consider $y^4 - 14y^2 - 24y$, which of course has been carefully rigged to have integer turning points. Substituting $b_2 = -14/3$ and $b_1 = -12$ into (6) and asking Maple to simplify the result gives

$$k_{\mathrm{f}} = \frac{2\left(143 + 180\sqrt{3}i\right)^{2/3} + 98 - 14\left(143 + 180\sqrt{3}i\right)^{1/3}}{3\left(143 + 180\sqrt{3}i\right)^{1/3}} \ .$$

All computer systems can approximate this to 4.0000000, but none can automatically simplify it to the exact number 4. This is an unavoidable difficulty associated with solving a cubic using the standard formulae. The simplification $\left(143 + 180\sqrt{3}\, i\right)^{1/3} = \frac{1}{2}(13 + 3\sqrt{3}\, i)$, which is needed to obtain the exact result, is not implemented in any present computer system; perhaps not many humans would make the simplification spontaneously either. Of course most of the time, no simplification is possible. In any event, the infimum is at $x_{\mathrm{f}} = 3$, and equals -117. In the same way, and with the same difficulties, $k_n = -6$ and $k_x = -12$.

Simply using the formula M given in (5) to compute a numerical minimum is not a very interesting application. A more challenging question is to find the values of p that make the polynomial $x^4 + 3px^2 + 2x + 2$ positive for all x. This type of problem is a simple example of quantifier elimination [1]. The condition is simply $M(1,p) + 2 > 0$, which becomes a long messy inequality when written out explicitly. Plotting the expression numerically shows that the answer is $p > -1/3$, but an analytic proof is a real challenge.

The final example does not aspire to present a general method for a class of problems, but the following challenge arose at the time of writing this paper. Given the points (x_i, y_i) equal to $(0,4)$, $(1,2)$, $(3,1)$, $(4,2)$, $(6,5)$, find a convex polynomial that passes through them. The Lagrange interpolating polynomial is a quartic:

$$y_L = \sum_{i=1}^{5} y_i \prod_{j \neq i} \frac{x - x_j}{x_i - x_j} = -\tfrac{1}{90}x^4 + \tfrac{4}{45}x^3 + \tfrac{13}{45}x^2 - \tfrac{71}{30}x + 4 \ .$$

This does not satisfy $y'' > 0$ everywhere. Therefore we investigate whether a sixth-degree polynomial can be found. For unknown coefficients a and b, we write

$$y_S = y_L + (ax + b)\prod_{i=1}^{5}(x - x_i) \ .$$

For all a and b, this passes through the given points. Two derivatives of this give a quartic inequality $y_S'' > 0$, which will be satisfied if $\inf y_S'' > 0$. This reduces to an inequality in the two variables a and b after using (5). Plotting contours shows that a region exists that satisfies the constraint, and in particular it includes the rectangle $1/2000 < a \leq 1/1000$, $0 < b \leq 1/1000$.

A computer epilogue

Just when you think it is time to roll the references, a familiar character reappears: the computer. The introduction mentioned that many CAS have routines to minimize functions, so the programming aspects of Theorem 3 are of interest. It was seen above that if (6) is used with numerical coefficients, the system might obtain a result that is correct, but not in the simplest form. In a similar way, if the evaluation of (6) generates intermediate quantities that are complex, a small nonzero imaginary part can appear in the final result because of rounding errors.

An alternative to the explicit formula (6) is to use a token such as the `RootOf` offered in Maple V release 4. To the description above, we can add that it also accepts a third argument, in the form of an interval that brackets the required root. The interval could be specified using (6), of course, but it is better to surrender precision to gain algebraic simplicity. Now $k_\mathrm{f} \to 3b_2$ as $b_2 \to \infty$, and $k_\mathrm{f} \to (2b_1^2)^{1/3}$ as $b_2 \to 0$, and $k_\mathrm{f} \to \sqrt{-2b_1^2/3b_2}$ as $b_2 \to -\infty$. An estimate that takes these limits into account, while staying with integral powers is $k_\mathrm{a} = 1 + 3|b_2| + \frac{2}{3}b_1^2$, which is an upper bound on k_f because $C(k_\mathrm{a}) > 0$. Thus, the expression (6) can be replaced by

$$k_\mathrm{f} = \mathtt{RootOf}(k^3 - 3b_2k^2 - 2b_1^2, k, 0 \mathrel{..} 1 + 3|b_2| + \tfrac{2}{3}b_1^2) \ ,$$

where some artistic licence has been taken with Maple's input language. Similar constructions exist in other systems. The advantages of this approach are that the system has the possibility of obtaining the best representation of the root directly, and that the case $b_1 = b_2 = 0$ is no longer a removable singularity. The disadvantages are that the representation is unfamiliar, and it may not allow the further analysis that is possible with the explicit form; also the simplification routines existing for this type of construction are not yet at all strong.

A final comment repeats what has been achieved from a slightly different perspective. The paper opened with a cubic equation (2) whose roots gave turning points. The main theorem replaced this with a different cubic. One cannot avoid solving a cubic sooner or later, but the auxiliary cubic in (9) has the advantage that one knows in advance which root to select and where it will be.

Acknowledgement. The last example came from a discussion with S. Watt, J. Grimm and A. Galligo; in particular, I thank J. Grimm for showing me some related results obtained from a more general point of view. This article was written while on leave at INRIA, Sophia Antipolis, France, and the hospitality of S.M. Watt is gratefully acknowledged.

One of the referees proposed a counterexample to theorem 1: unfortunately my agents have been unable to locate the movie *Sneakers* to verify this intelligence report.

References

[1] Lazard, D. Quantifier elimination: optimal solution for two classical problems, *J. Symbolic Comp.* 5 (1988), 261–266.