J. Symbolic Computation (1988) **5**, 29–35

Real Quantifier Elimination is Doubly Exponential

JAMES H. DAVENPORT AND JOOS HEINTZ

School of Mathematical Sciences, University of Bath, Claverton Down, Bath BA2 7AY, England and

F.B. Mathematik, J.W. Goethe Universität, D-6000 Frankfurt/Main F.R.G. and Universidad Nacional de La Plata, La Plata, Argentina.

(Received 7 March 1987)

We show that quantifier elimination over real closed fields can require doubly exponential space (and hence time). This is done by explicitly constructing a sequence of expressions whose length is linear in the number of quantifiers, but whose quantifier-free expression has length doubly exponential in the number of quantifiers. The results can be applied to cylindrical algebraic decomposition, showing that this can be doubly exponential. The double exponents of our lower bounds are about one fifth of the double exponents of the best-known upper bounds.

1. Introduction.

Tarski [1951] was the first to show that quantifier elimination was possible for the first-order theory of real-closed fields. His method was totally impractical, and was substantially modified and improved by Collins [1975], who gave a doubly-exponential algorithm for *cylindrical algebraic decomposition*, a process which divided $\mathbf{R}^n$ into a series of regions, on each of which the members of a given family of polynomials were identically zero, identically positive or identically negative. Here, and throughout, the expression *doubly exponential* refers to the dependence of the running time on n, the number of dimensions. The running time is polynomially dependent on all the other measures of size involved: the number of polynomials, their degree and the length of their coefficients.

It is the purpose of this paper to show that this doubly-exponential behaviour is intrinsic to the problem. This result has already been proved by Weispfenning [1985] by completely different methods, but we feel that our method is of independent interest. The method proceeds by adapting a construction of Heintz [1983] which was used to show that complex-valued quantifier elimination was doubly exponential. In the next section we construct the formulae, and explain the method, which is essentially a technique for writing $x_1^{2^{2^{k+1}}} = x_2$ in a form whose length, when considered as a dense polynomial, is only linear in k. The following section proves that any quantifier-free expression for this requires at least 2^{2^k} symbols if written densely. It should be noted that our formulation is different from that of Ben-Or *et al.* [1986], who take satisfiability as their criterion, i.e. they are considering formulae with no free variables. It seems that there is still much to be understood on the relationship between their view and ours.

0747–7171/88/010029+07 $03.00/0

This can be applied to cylindrical algebraic decomposition, since the corresponding decomposition of space requires at least one region for each zero of this polynomial. The result for cylindrical decomposition is independent of any considerations of density or sparsity, since we can show that the number of regions that have to be described is doubly exponential.

2. The Formulae.

We consider a sequence of formulae $\phi_0, \phi_1, \ldots, \phi_k$, each with four free variables. The even-numbered formulae will have x_{1R}, x_{1I}, x_{2R} and x_{2I} as free variables, while the odd-numbered formulae will have z_{1R}, z_{1I}, z_{2R} and z_{2I} as free variables. The basic formula ϕ_0 is given as

$$\left(x_{1R}^4 - 6x_{1R}^2x_{1I}^2 + x_{1I}^4 = x_{2R}\right) \wedge \left(4x_{1R}^3x_{1I} - 4x_{1R}x_{1I}^3 = x_{2I}\right).$$

While this formula may appear mysterious, it is simply an expression of the real and imaginary parts of the equality

$$(x_{1R} + ix_{1I})^4 = x_{2R} + ix_{2I}.$$

Hence we can regard ϕ_0 as being the *complex* equation

$$x_1^4 = x_2. \tag{1}$$

The rule for computing ϕ_{j+1} from ϕ_j depends on the parity of j: here we give the rule for even j, while that for odd j is obtained by interchanging the roles of x and z.

$$\begin{aligned}\phi_{j+1}(z_{1R}, z_{1I}, z_{2R}, z_{2I}) = \exists y_R \exists y_I \forall x_{1R} \forall x_{1I} \forall x_{2R} \forall x_{2I} \\ \left(\left((x_{1R} = z_{1R} \wedge x_{1I} = z_{1I} \wedge x_{2R} = y_R \wedge x_{2I} = y_I\right)\right. \\ \vee \left(x_{1R} = y_R \wedge x_{1I} = y_I \wedge x_{2R} = z_{2R} \wedge x_{2I} = z_{2I}\right)\Big) \\ \Rightarrow \phi_j(x_{1R}, x_{1I}, x_{2R}, x_{2I})\Big).\end{aligned}$$

In its complex form, this equation reduces to

$$\phi_{j+1}(z_1, z_2) = \exists y \forall x_1 \forall x_2 \left(\left(\left(x_1 = z_1 \wedge x_2 = y\right) \vee \left(x_1 = y \wedge x_2 = z_2\right)\right) \Rightarrow \phi_j(x_1, x_2)\right).$$

This formula is logically equivalent to the formula

$$\exists y \forall x_1 \forall x_2 \left(\left(\left(x_1 = z_1 \wedge x_2 = y\right) \Rightarrow \phi_j(x_1, x_2)\right) \wedge \left(\left(x_1 = y \wedge x_2 = z_2\right) \Rightarrow \phi_j(x_1, x_2)\right)\right).$$

In turn, we can simplify this, since the implications are trivially true unless x_1 and x_2 have the values given in the hypotheses of the implications. Hence $\phi_{j+1}(z_1, z_2)$ is logically equivalent to

$$\exists y \left(\phi_j(z_1, y) \wedge \phi_j(y, z_2)\right). \tag{2}$$

While we have only made these transformations on the complex version of ϕ_{j+1}, the same transformations could clearly be made on the real version.

Proposition 1. *The complex version of ϕ_j is equivalent to the equation*

$$x_1^{2^{2j+1}} = x_2$$

(or the same equation in terms of z_1 and z_2 if j is odd).

The proof is a trivial induction on equations (1) and (2).

Corollary. *ϕ_j is equivalent to the following logical formula, where $\mathcal{R}$ and $\mathcal{I}$ stand for the real part and the imaginary part respectively:*

$$\mathcal{R}\left((x_{1R}+ix_{1I})^{2^{2^{j+1}}}\right)=x_{2R}\wedge\mathcal{I}\left((x_{1R}+ix_{1I})^{2^{2^{j+1}}}\right)=x_{2I}$$

(or the same equations written in terms of z_{1R}, z_{1I}, z_{2R} and z_{2I} if j is odd).

Although these equations are phrased in terms of complex variables, they do in fact have a purely real expression, as

$$\left(\sum_{l=0}^{2^{2^{j+1}-1}}(-1)^l\binom{2^{2^{j+1}}}{2l}x_{1R}^{2^{2^{j+1}}-2l}x_{1I}^{2l}\right)=x_{2R}$$
$$\wedge\left(\sum_{l=0}^{2^{2^{j+1}-1}-1}(-1)^l\binom{2^{2^{j+1}}}{2l+1}x_{1R}^{2^{2^{j+1}}-2l-1}x_{1I}^{2l+1}\right)=x_{2I}.$$

Proposition 2. *$\phi_k(x_{1R},x_{1I},1,0)$ (or $\phi_k(z_{1R},z_{1I},1,0)$ if k is odd) defines a semi-algebraic set in $\mathbf{R}^2$ consisting of $2^{2^{k+1}}$ isolated points.*

Each point of the set defined by ϕ_k corresponds to a $2^{2^{k+1}}$-th complex root of unity.

It is worth noting that the alphabet required to define ϕ_k is independent of k, and hence the length of ϕ_k is linear in k — $44k+37$ symbols the way we have written it.

3. The size of a quantifier-free expression.

In the complex case, Heintz [1983] was able to argue that the only way to define a subset of $\mathbf{C}^1$ consisting of $2^{2^{k+1}}$ points is via a polynomial of that degree (or a set of polynomials of which that is the sum of the degrees). This is not so obvious in the real case, as a quantifier-free expression may well contain inequalities as well as equalities, and several levels of logical conjunctions and disjunctions, such as $A\wedge(B\vee C\vee(D\wedge E))$.

If $p(x,y)$ and $q(x,y)$ are two polynomials in $\mathbf{R}[x,y]$ we define $D(p)$ to be the set of isolated points of $p=0$ and $D(p,q)$ to be the set of isolated points of the intersection of the curves $p=0$ and $q=0$ which are not isolated points of the curves considered separately. We note that $D(p_1p_2)\subset D(p_1)\cup D(p_2)$ and that $D(p_1p_2,q)\subset D(p_1,q)\cup D(p_2,q)$.

Proposition 3. *Given any quantifier-free polynomial expression in two real variables x and y, involving polynomials $p_1(x,y),\ldots,p_n(x,y)$, the set of isolated points of the subset of $\mathbf{R}^2$ defined by the formula is a subset of*

$$\left(\bigcup_{i=1}^{m}D(q_i)\right)\cup\left(\bigcup_{i=1}^{m}\bigcup_{j=i+1}^{m}D(q_i,q_j)\right)$$

where the q_i are the irreducible factors of the p_i (in fact, it suffices to ensure that the q_i are without multiple factors, and relatively prime).

Proof. We can re-write the expression in terms of the q_i, using the facts that $fg=0$ is equivalent to $f=0\vee g=0$, and $fg>0$ is equivalent to $(f>0\wedge g>$

$0) \vee (f < 0 \wedge g < 0)$. After applying the distributive laws to our expression, we can assume that each constituent of the expression is given by the conjunction of a number of elementary formulae of the form $q_i > 0$, $q_i = 0$ or their negations. Furthermore, the negation of $q_i > 0$ can be written as $q_i = 0 \vee q_i < 0$, and this term can be expanded with the other disjunctions, so that we need not consider such expressions. Hence each constituent of the expression is of the form

$$a_1 = 0 \wedge \ldots \wedge a_k = 0 \wedge b_1 > 0 \wedge \ldots \wedge b_l > 0 \wedge c_1 \neq 0 \wedge \ldots \wedge c_m \neq 0,$$

where the a_i, b_i and c_i are some of the q_i or their negatives. It is, of course, perfectly concievable that some of k, l or m could be zero. However, if

$$a_1 = 0 \wedge \cdots \wedge a_k = 0$$

does not define an isolated point (and possibly some other connected components), then the expression as a whole can not define an isolated point, since the inequalities are open conditions, while isolated points are closed. Hence the isolated points defined by the original expression are defined by the combination of certain equalities in the transformed expression (which may have come from inequalities in the original: for example

$$x^2 + y^2 - 1 \leq 0 \wedge y \geq 1$$

defines an isolated point $x = 0 \wedge y = 1$, but this will be transformed into the two equalities, as well as into three other sets that are actually void). In particular $k > 0$ for an isolated point. There are then two possibilities: either $a_1 = 0$ defines this isolated point, or it defines a curve which passes through this point. In the former case, the point is an element of $D(a_1)$, which may well contain other components as well. In the latter case, it would be natural to assume that the point was an element of $D(a_1, a_2)$, but this is not necessarily the case, since $a_2 = 0$ may define the same curve. However, there has to exist an i such that

$$a_1 = 0 \wedge \ldots \wedge a_{i-1} = 0$$

defines a curve passing through the point in question, but

$$a_1 = 0 \wedge \ldots \wedge a_i = 0$$

defines the point precisely. Since a_1 and a_i are relatively prime, $\{(x, y) : a_1(x, y) = a_i(x, y) = 0\}$ is a variety of dimension zero, and so consists only of isolated points. Thus the point belongs to $D(a_1, a_i)$ (as well as possibly to several other such sets).

Proposition 4. *If p_i is a polynomial of total degree d_i, the total number of isolated points is at most*

$$\left(\sum_{i=1}^{n} d_i\right)^2.$$

Proof. After the previous proposition, it suffices to show that

$$\left|\left(\bigcup_{i=1}^{m} D(q_i)\right) \cup \left(\bigcup_{i=1}^{m} \bigcup_{j=i+1}^{m} D(q_i, q_j)\right)\right| < \left(\sum_{i=1}^{n} d_i\right)^2.$$

If q_i is a polynomial of degree e_i, then we have that $\sum_{i=1}^{n} d_i \geq \sum_{i=1}^{m} e_i$. But $D(q_i)$ is the set of isolated real points of one equation, and any such point has to be a multiple point (since in the neighbourhood of the point we must have complex y values corresponding to real x values, and these must come in conjugate pairs). Hence the x-coordinate of any such point is a root of the discriminant of q_i. Furthermore, the number of y values of multiple points with a given x value is at most half the multiplicity of this root of the discriminant. Since the discriminant of a polynomial of total degree d has degree at most $2d^2$, we deduce that $|D(q_i)| \leq e_i^2$. An alternative approach proceeds via the Bezout inequality (Theorem 1 of Heintz [1983]). The multiple point is a common root of q_i and dq_i/dy, and hence belongs to a set of degree bounded by the products of the total degrees of q_i and dq_i/dy, i.e. $e_i(e_i - 1) \leq e_i^2$.

Similarly $D(q_i, q_j)$ is the set of isolated points of two equations, and any such point has an x coordinate which is a root of the resultant of q_i and q_j (which is non-zero by the irreducibility of the q_i). Furthermore, the number of common points with this x coordinate is at most the multiplicity of this root of the resultant. Since the degree of a resultant is at most twice the product of the degrees, we deduce that $|D(q_i, q_j)| \leq 2e_ie_j$. Again, it would be possible to proceed via Bezout's Inequality and an argument on total degree. Hence

$$\begin{aligned}\left|\left(\bigcup_{i=1}^{m} D(q_i)\right) \cup \left(\bigcup_{i=1}^{m}\bigcup_{j=i+1}^{m} D(q_i, q_j)\right)\right| &\leq \left(\sum_{i=1}^{m} |D(q_i)|\right) + \left(\sum_{i=1}^{m}\sum_{j=i+1}^{m} |D(q_i, q_j)|\right) \\ &\leq \left(\sum_{i=1}^{m} e_i^2\right) + \left(\sum_{i=1}^{m}\sum_{j=i+1}^{m} 2e_ie_j\right) \\ &= \left(\sum_{i=1}^{m} e_i\right)^2.\end{aligned}$$

Remarking that, in a dense representation of polynomials, a polynomial in two variables of total degree n requires at least n symbols to write it, we can combine propositions 2 and 4 to deduce

Theorem 1. *There exist formulae, of length linear in k, containing $6k$ quantifiers and two free variables, such that the real quantifier-free expressions corresponding to them require at least 2^{2^k} symbols to write down.*

We started ϕ_0 as a quartic equation in order to obtain an extra factor of two which would cancel with the square-root. If we had made ϕ_0 into a quadratic, the current result would have been $2^{2^{k-1}}$. Since every data item must be written, we deduce the following result.

Corollary. *The time required to eliminate n quantifiers over* $\mathbf{R}$ *is doubly exponential in n.*

4. Applications to Cylindrical Algebraic Decomposition.

It is well-known that quantifier elimination is an easy consequence of a cylindrical algebraic decomposition, and hence it would be possible to deduce that cylindrical algebraic decomposition requires doubly exponential time. In fact, we can proceed

directly, since the cylindrical algebraic decomposition of $\mathbf{R}^{6k+2}$ induced by ϕ_k contains at least $2^{2^{k+1}}$ 0-dimensional regions — one corresponding to each root of unity. Hence we have proved

Theorem 2. *The time required to decompose* $\mathbf{R}^{6k+2}$ *cylindrically according to* $8k+2$ *polynomials of degree at most* 4 *is at least* $2^{2^{k+1}}$.

In terms of the dimensionality of the space, this becomes $2^{2^{(n+4)/6}}$. The exponent $(n+4)/6$ in this lower bound should be compared with Collins' [1975] upper bound of $2n+8$, since improved by McCallum [1985b] (see also McCallum [1985a]) to $n+\log n+7$, and by Davenport [1985] to $n+\log n+5$.

It is possible to do slightly better than the previous results would indicate. For example, if we replace the induction rule for the ϕ_j (in its complex form) by

$$\phi_{j+1}(z_1,z_2) = \exists y \exists w \forall x_1 \forall x_2$$
$$(((x_1 = z_1 \wedge x_2 = y) \vee (x_1 = y \wedge x_2 = w) \vee (x_1 = w \wedge x_2 = z_2)) \Rightarrow \phi_j(x_1,x_2)),$$

we get a formula logically equivalent to

$$\exists y \exists w\, (\phi_j(z_1,y) \wedge \phi_j(y,w) \wedge \phi_j(w,z_2)),$$

and Proposition 2 would contain the number 4^{3^k} instead of $2^{2^{k+1}}$ (at the cost of needing two more symbols w_R and w_I and two more quantifiers at each step). Adding another two variables and quantifiers gives us $4^{4^k} = 2^{2^{2k+1}}$. Since we are now in dimension $10k+2$, this gives us the following variant of Theorem 2:

Theorem 2′. *The time required to decompose* $\mathbf{R}^{10k+2}$ *cylindrically according to* $16k+2$ *polynomials of degree at most* 4 *is at least* $2^{2^{2k+1}}$.

In terms of the dimensionality of the space, this becomes $2^{2^{(n+3)/5}}$, which seems to be about the limit of this method.

5. Conclusions

There seems little to add to Theorems one and two. Clearly, there is still quite a gap between the exponents of these lower bounds and the best-known upper bounds, and it would be interesting to know how to narrow this. It would also be interesting to know whether quantifier elimination is still doubly-exponential if the number of bound variables is constant. Weispfenning [1985] has shown that, for *linear* formulae, bounding the number of quantifiers makes the problem polynomial, and bounding the number of alternations makes it simply exponential.

Note added in proof. Lars Langemyr (Stockholm) and the first author have applied McCallum's method to a special case, viz.

$$\exists c \forall b \forall a(((a = d \wedge b = c) \vee (a = c \wedge b = 1)) \Rightarrow a^2 = b),$$

which reduces to the formula $d^4 = 1$, whose *real* solution consists of $d = 1$ and $d = -1$. The equations (four linear and one quadratic) induce a cellular algebraic decomposition of $\mathbf{R}^4$ into 3837 cells. The computation took 2 hours (7202.8 seconds) on an otherwise unloaded discless SUN 3/160 (with 8 Mb real memory). SAC-2 needed a 4 Mb heap

to compute the decomposition and the formulae for the cells. Further details of this calculation are to appear in the SIGSAM Bulletin.

The authors wish to thank the Centre de Calcul de l'Esplanade, Université Louis Pasteur, Strasbourg, where this work was done, and in particular Maurice Mignotte, who introduced the authors to each other. The first author was visiting the Centre de Mathématiques de l'École Polytechnique, Palaiseau, France at the time.

References

Ben-Or, M., Kozen, D. & Reif, J. (1986). The Complexity of Elementary Algebra and Geometry. J. Comput. Syst. Sci. **32** 251–264.

Collins, G. E. (1975). Quantifier Elimination for Real Closed Fields by Cylindrical Algebraic Decomposition. Proc. 2nd. GI Conf. Automata Theory & Formal Languages (Springer Lecture Notes in Computer Science 33) pp. 134–183. MR 55 #771.

Davenport, J. H. (1985). Computer Algebra for Cylindrical Algebraic Decomposition. TRITA-NA-8511, NADA, KTH, Stockholm, Sept. 1985.

Heintz, J. (1983). Definability and Fast Quantifier Elimination in Algebraically Closed Fields. Theor. Comp. Sci. **26** (1983) pp. 239–277. MR 85a:68061.

McCallum, S. (1985a). An Improved Projection Operation for Cylindrical Algebraic Decomposition. Proc. EUROCAL 85, Vol. 2 (Springer Lecture Notes in Computer Science 204) pp. 277–278.

McCallum, S. (1985b). An Improved Projection Operation for Cylindrical Algebraic Decomposition. Computer Science Tech. Report 548, Univ. Wisconsin at Madison, Feb., 1985.

Tarski, A. (1951). A Decision Method for Elementary Algebra and Geometry. 2nd. ed., Univ. Cal. Press, Berkeley, 1951. MR 10 #499.

Weispfenning, V. (1985). The Complexity of Linear Problems in Fields. Manuscript, 1985 (revised March 1986).