On Propagation of Equational Constraints in CAD-Based Quantifier Elimination

Scott McCallum
Department of Computing
Division of Information and Communication Sciences
Macquarie University NSW 2109
Australia
scott@ics.mq.edu.au

ABSTRACT

Collins [4] observed that quantifier elimination problems often have equational constraints, and he asserted that such constraints can be used to reduce the projection sets required for cylindrical algebraic decomposition (cad) based quantifier elimination. This paper follows on from [11], and validates the use of a semi-restricted equational projection scheme throughout the projection phase of cad. The fully restricted projection scheme as originally proposed in [4] is proved valid for four variable problems under certain conditions.

1. INTRODUCTION

We begin by recalling some terminology introduced by Collins [4]. A *constraint* of a quantifier elimination problem is an atomic formula which is logically implied by the quantifier-free matrix of the prenex input formula. If the atomic formula is an equation it is called an *equational constraint*, and the polynomial in the equation is an *equational constraint polynomial*. Collins [4] observed that quantifier elimination problems often have equational constraints, and he asserted that such constraints can be used to reduce the projection sets required for cylindrical algebraic decomposition (cad) based quantifier elimination. Collins [4] stated that, since the matrix of the input formula can only be true in the sections of an equational constraint polynomial, other polynomials occurring in the formula need only be sign-invariant in those sections and do not need to be delineable. He asserted therefore that the projection operation of McCallum [8, 9, 10] could be restricted in the manner indicated in the following slightly modified excerpt:

> if f is an equational constraint polynomial and $g_1, \ldots, g_n$ are the remaining polynomials then the projection set may consist of only
>
> - the discriminant and (enough of) the coefficients of f, and
> - for each i, the resultant of f and g_i.
>
> . . .
>
> Since polynomials which are to be projected are always made irreducible, it is important to realize that the equational constraint polynomial f may be a product of several irreducible polynomials $f_1, \ldots, f_m$. Then the discriminant of f has as factors
>
> - the discriminant of each f_i, and
> - the square of each resultant $\mathrm{res}(f_i, f_j)$, $i < j$,
>
> and in the projection set the coefficients of f may be replaced by appropriate coefficients of the f_i.

Referring to the program QEPCAD for quantifier elimination by partial cylindrical algebraic decomposition [5], Collins continues:

> If f_1 and f_2 are both equational constraint polynomials then so is their resultant, $\mathrm{res}(f_1, f_2)$, since
>
> $$f_1 = 0 \wedge f_2 = 0 \Rightarrow \mathrm{res}(f_1, f_2) = 0.$$
>
> Currently the user must declare the equational constraint polynomials which occur in the input formula, but the program will then propagate these equational resultants by application of this 'resultant rule'.

Noting that the validity of the restricted projection scheme informally described in the exercpts above is not obvious in general, McCallum [11] proved that the use of the restricted equational projection operation is valid for the first projection. The subject of the present paper relates to the issue of equational constraint propagation, mentioned in the second excerpt above. Our first result is straightforward. We define a projection operation which we call the *semi-restricted*

ISSAC 2001, 7/01, Ontario, Canada
©2001 ACM 1-58113-417-7/ 01/ 0007 $5.00

projection, which is not quite as restricted as that proposed above. We show that one can use the (fully) restricted projection operation for the first and last projections, and the semi-restricted projection operation for every other projection required. Equational constraints, whether propagated or present in the input formula, can thus be used at every level of projection to reduce the size of the projection sets. The reduction in projection set size is greatest for the first and last projections.

Our second result relates to problems involving at most four variables. It states that, subject to certain technical conditions, the equational projection method exactly as proposed in the excerpts above (that is, the fully restricted projection operation) can be used for each of the three projections required.

In Section 2 we make precise the semi-restricted equational projection operation to which we refer above. We prove the validity of the use of this operation in QEPCAD. Section 3 provides a detailed proof of the key lemma for the main theorem of Section 2. Section 4 focuses on four variable problems. We describe the technical conditions under which the fully restricted equational projection can be used for the second (hence every) projection, for problems involving four variables. We present theorems which justify the use of the fully restricted projection operation for such problems. In Section 5 we report the results of some computational experiments using program QEPCAD. Section 6 discusses some issues related to the possible extension of the validity arguments presented in this paper.

2. SEMI-RESTRICTED PROJECTION

Background material describing the cad algorithm and its theory can be found in [2, 3, 5, 9, 10]. We will make use of terminology and results from [9, 10]. A summary of the definitions of the key technical terms (such as 'submanifold of $\mathbf{R}^n$', 'analytic delineability', 'order-invariance' and 'section') used in this paper can be found in [11].

Let A be a set of pairwise relatively prime irreducible integral polynomials in r variables $x_1, \ldots, x_r$ of positive degrees in x_r, where $r \geq 2$. We recall from [9, 10] that the *improved projection* $P(A)$ of A is the union of the set of all nonzero coefficients with respect to x_r of the elements of A, the set of all discriminants $\mathrm{discr}_{x_r}(f)$ of elements of A and the set of all resultants $\mathrm{res}_{x_r}(f, g)$ of pairs f, g of distinct elements of A. The main theorem (for general $r \geq 2$) about the projection operator P is stated and proved in [10] (and recalled in [11]). It says that, for a given connected submanifold S of $\mathbf{R}^{r-1}$ in which every element of $P(A)$ is order-invariant, we have that each element of A either vanishes identically on S or is analytic delineable on S, the sections of elements of A over S are pairwise disjoint, and each element of A is order-invariant in every such section (or vanishes identically on S). This theorem makes possible the repeated application of the projection operator P throughout the projection phase of cad, provided that a certain technical condition on A called *well-orientedness* holds.

Let E be a subset of A. We recall from [11] that the *restricted projection of A relative to E*, $P_E(A)$, is defined as follows:

$$P_E(A) = P(E) \cup \{\mathrm{res}_{x_r}(f, g) \mid f \in E, g \in A, g \notin E\}.$$

The main theorem about $P_E(A)$ is stated and proved in [11]. It says that, for a given connected submanifold S of $\mathbf{R}^{r-1}$ in which every element of $P_E(A)$ is order-invariant, we have that each element of E either vanishes identically on S or is analytic delineable on S, the sections of elements of E on S are pairwise disjoint, and each element of $A - E$ is sign-invariant in every such section. Because this theorem falls short of guaranteeing the *order-invariance* of the elements of $A - E$ in the sections of E, the theorem is unfortunately not strong enough to validate iterated application of the restricted projection operation: the use of the restricted projection is guaranteed to be valid only for the *first* projection operation performed (provided that a suitable equational constraint is present).

We now define the *semi-restricted projection of A relative to E, $P_E^*(A)$*, as follows:

$$P_E^*(A) = P_E(A) \cup \{\mathrm{discr}_{x_r}(g) \mid g \in A, g \notin E\}.$$

Notice that we have

$$P_E(A) \subseteq P_E^*(A) \subseteq P(A)$$

and that the containment relations can be strict. (If, for example, $A = \{f, g, h\}$ and $E = \{f\}$, then $P_E^*(A) - P_E(A)$ contains $\mathrm{discr}_{x_r}(g)$ and $\mathrm{discr}_{x_r}(h)$, while $P(A) - P_E^*(A)$ contains $\mathrm{res}_{x_r}(g, h)$, amongst some other polynomials.)

The key lemma for the main result of this section - which is the main theorem about $P_E^*(A)$, to be stated shortly - could be viewed as an analogue of the key lemma (Theorem 2.2) of [11]. Both the hypotheses and the conclusions of the result stated below are stronger than their counterparts in Theorem 2.2 of [11].

THEOREM 2.1. *Let $r \geq 2$, let $f(x_1, \ldots, x_r)$ and $g(x_1, \ldots, x_r)$ be real polynomials of positive degrees in the main variable x_r, let $R(f, g)$ and $D(g)$ be the resultant with respect to x_r of f and g and the discriminant with respect to x_r of g, respectively, and suppose that both $R(f, g)$ and $D(g)$ are nonzero. Let S be a connected submanifold of $\mathbf{R}^{r-1}$ on which f is analytic delineable and g does not vanish identically, and in which both $R(f, g)$ and $D(g)$ are order-invariant. Then g is order-invariant in each section of f over S.*

Theorem 2.1 is proved in Section 3. (It "almost" follows immediately from Theorem 2 of [10]. But strictly speaking Theorem 2 of [10] cannot be applied, because g is not assumed to be degree-invariant on S. Nevertheless, a straightforward reduction argument along the lines of Section 4 of [9] allows one to apply the Zariski theorem itself - Theorem 3 of [10] - to give the desired conclusion. The details are provided in Section 3.)

Our main result pertaining to $P_E^*(A)$ can now be stated (cf. [11], Theorem 2.3). This result will permit the valid repeated application of the semi-restricted projection operator, where applicable, throughout the projection phase of cad, as will shortly be explained in detail.

THEOREM 2.2. *Let A be a set of pairwise relatively prime irreducible integral polynomials in $x_1, \ldots, x_r$ of positive degrees in x_r, where $r \geq 2$. Let $E \subseteq A$. Let S be a connected submanifold of $\mathbf{R}^{r-1}$. Suppose that each element of $P_E^*(A)$ is order-invariant in S. Then each element of E either vanishes identically on S or is analytic delineable on S, the sections over S of the elements of E which don't vanish identically on S are pairwise disjoint, and each element of A which doesn't vanish identically on S is order-invariant in every such section.*

Proof. All of the conclusions except for the order-invariance of each element of $A - E$ which doesn't vanish identically on S in the "E-sections" over S follow immediately by Theorem 2.3 of [11]. Take an arbitrary element g of $A - E$ which doesn't vanish identically on S and take an arbitrary section σ of some element f of E which doesn't vanish identically on S. Now $P_E^*(A)$ contains $\mathrm{res}_{x_r}(f, g)$ and $\mathrm{discr}_{x_r}(g)$, by definition, and these two polynomials are order-invariant in S, by hypothesis. Hence, by Theorem 2.1, g is order-invariant in σ. □

Now let A be an arbitrary set of integral polynomials in $x_1, \ldots, x_r$, and let E be a subset of A. We define the semi-restricted projection $P_E^*(A)$ of A relative to E as follows:

$$P_E^*(A) = \mathrm{cont}(A) \cup P_F^*(B),$$

where $\mathrm{cont}(A)$ is the set of nonzero contents with respect to x_r of the elements of A, B is the finest squarefree basis for the set $\mathrm{prim}(A)$ of all primitive parts with respect to x_r of the elements of A which have positive degree in x_r and $F \subseteq B$ is the finest squarefree basis for $\mathrm{prim}(E)$. One sees that $P_E^*(A)$, for arbitrary A, is defined by analogy with the definition of $P_E(A)$, for A arbitrary, given in Section 5 of [11], and by analogy with the definition of $P(A)$, for A arbitrary, given in Section 6 of [10].

We describe a variant of the algorithm QEPCAD of [5]. The algorithm accepts as input a quantified formula

$$\phi^* = (Q_{f+1}x_{f+1}) \ldots (Q_r x_r)\phi(x_1, \ldots, x_r),$$

where $0 \leq f < r$ and ϕ is a quantifier-free formula. The algorithm will produce as output a boolean value w and a formula $\phi' = \phi'(x_1, \ldots, x_f)$ such that if the set A of polynomials occurring in ϕ^* is well-oriented ([9], [10]) then $w = true$ and ϕ' is a quantifier-free formula equivalent to ϕ^*, and otherwise $w = false$.

Step 1 of QEPCAD is the projection phase of the algorithm in which we compute projection polynomials of all orders. We define $J_r = A$. Suppose first that there is some equational constraint $e_r = 0$ present in ϕ^* such that e_r is primitive and has positive degree with respect to x_r. It follows from Theorem 2.3 of [11] that QEPCAD can use the restricted projection operator P_{E_r} to compute the first projection $J_{r-1} = P_{E_r}(J_r)$, where $E_r = \{e_r\}$. If there is more than one such equational constraint polynomial then any one of them may be chosen as e_r in this definition of the projection set. The polynomial selected as e_r is called the *pivot*. On the other hand, if there is no such suitable equational constraint at level r then we compute $J_{r-1} = P(J_r)$. Now let $1 < k < r$ and suppose that J_k has been defined (as the set of projection polynomials at level k, that is, the set of polynomials in the variables $x_1, \ldots, x_k$ obtained by suitably projecting the set J_{k+1}). Suppose first that there is some suitable (that is, primitive and of positive degree in x_k) equational constraint e_k at level k. Note that such an equational constraint at level k could occur in the input formula ϕ^*, or could be a constraint *propagated* from higher level constraints by application of the resultant rule mentioned in Section 1. We put E_k equal to the set of elements of J_k which are divisors of e_k. It follows from Theorem 2.2 above that QEPCAD can use the semi-restricted projection operator $P_{E_k}^*$ to compute the next projection $J_{k-1} = P_{E_k}^*(J_k)$. Again, if there is more than one such equational constraint polynomial at level k then any one of them may be chosen as the pivot e_k. On the other hand, if there is no such suitable equational constraint at level k then we compute $J_{k-1} = P(J_k)$. For the last projection (that is, $k = 2$), if there is a suitable equational constraint e_2 at level 2, then one can actually compute the fully restricted projection $J_1 = P_{E_2}(J_2)$ (by the remarks at the end of Section 5 of [11]).

Step 5 of QEPCAD, a key part of the stack construction phase, prescribes the construction of the stack over a chosen cell c. Algorithm CCHILD ([5], page 311) actually accomplishes this construction. Algorithm CCHILD must be modified so that, in case c has positive dimension, it will detect whether some basis polynomial vanishes identically on c and, if so, set w equal to $false$. Actually this modification to CCHILD is already described in [10].

3. PROOF OF THEOREM 2.1

We shall prove the theorem for the case $r = 3$. The proof readily generalizes to arbitrary $r \geq 2$. We treat the case $r = 3$ in order to keep the notation as simple as possible (and since the case $r = 2$ is easy enough to be disposed of without using the argument required for $r \geq 3$.) Let σ be a section of f over S, determined by the analytic function $\theta : S \to \mathbf{R}$. By Theorem 2.2 of [11], g is sign-invariant in σ. If $g \neq 0$ throughout σ, then g is order-invariant (of order 0) in σ. So henceforth assume that $g = 0$ throughout σ.

By connectedness of σ, it suffices to show that g is order-invariant in σ near an arbitrary point (a, b, c) of σ. That is, it is enough to show that for every point (a, b, c) of σ, there exists a neighbourhood N of (a, b, c) such that g is order-invariant in $\sigma \cap N$. Let (a, b, c) be a point of σ and let m be the multiplicity of the root $c = \theta(a, b)$ of $f(a, b, z)$.

We assume that S has positive dimension. (The dimension 0 case is trivial.) That g is order-invariant in σ near (a, b, c) is quite easy to see in case the dimension of S equals 2. For in this case, S is a connected open subset of the plane. Hence, as $D(g)$, a nonzero polynomial, is order-invariant in S, $D(g)$ vanishes nowhere in S. Therefore $m = 1$ and $\partial g/\partial z(a, b, c) \neq 0$ (by Lemma 7.1). Hence, by continuity of $\partial g/\partial z$, g is order-invariant (of order 1) in σ near (a, b, c).

The remaining case to consider is that in which the dimension of S is 1, that is, S is a smooth curve in the plane. There is no loss of generality in assuming that $c = 0$. By Theorem 2.2 of [9], we choose coordinates (u, v) about the point (a, b) of S such that S is defined locally by the equation $v = 0$ in the new coordinate system. By Hensel's Lemma (Theorem 3.1 of [11], in which one should replace the real field $\mathbf{R}$ by the complex field $\mathbf{C}$), the zero set in $\mathbf{C}^3$ of g near the origin is identical with the zero set of a Weierstrass polynomial ([6], [9])

$$h(u,v,z) = z^m + c_1(u,v)z^{m-1} + \cdots + c_m(u,v)$$

near the origin. It is straightforward to verify that the discriminant $F(u,v)$ of $h(u,v,z)$ does not vanish identically, and that F is order-invariant in the complex u-axis T^*, near the origin (the complex u-axis is the subset of the complex u,v-plane $\mathbf{C}^2$ defined by the equation $v = 0$).

Hence, by Theorem 3 of [10], for every fixed $(u,0)$ in T^* near the origin, $h(u,v,z)$, as a polynomial in z, has exactly one root (necessarily of multiplicity m), and h is order-invariant in the locus (graph) of this root over T^*. Therefore, for every fixed (x,y) in S near (a,b), $g(x,y,z)$, as a polynomial in z, has exactly one root (necessarily of multiplicity m and necessarily real) near $c = \theta(a,b)$, and g is order-invariant in the locus (graph) of this root over S (near (a,b)). Since $g = 0$ throughout σ, σ must contain the locus of the root of g near c over S (near (a,b)). Therefore g is order-invariant in σ near (a,b,c). □

4. IMPROVEMENTS TO PROJECTION FOR FOUR VARIABLE PROBLEMS

In this section we describe a further improvement to the projection phase of QEPCAD which can be used for four variable quantifier elimination problems. Let (x,y,z,w) denote (x_1,x_2,x_3,x_4). Suppose that algorithm QEPCAD is presented with the quantified formula ϕ^* involving x,y,z and w, ordered in this way. Let A be the set of polynomials occurring in ϕ^*. Suppose first that there is some equational constraint $f(x,y,z,w) = 0$ present in ϕ^* such that f is primitive and of positive degree with respect to w. (We use the symbol f instead of e_4 for notational convenience.) We shall also assume for simplicity that f is squarefree. By Theorem 2.3 of [11] we can compute $J_3 = P_{\{f\}}(A)$. If, on the other hand, there is no such constraint present, then we compute $J_3 = P(A)$.

Next suppose that $e(x,y,z) = 0$ is an equational constraint of ϕ^* which is primitive and of positive degree with respect to z. (We use the symbol e instead of e_3 for notational convenience.) As remarked in the previous section, such a constraint could occur in ϕ^*, or e could be the resultant of f and some other equational constraint polynomial, primitive, squarefree and of positive degree in w. We shall also assume for simplicity that e is squarefree. We define E_3 to be the set of elements of J_3 which are divisors of e. Let $d(x,y,z)$ denote the discriminant of f with respect to w. We say that f and e are *well-placed* if $\mathrm{discr}_z(e)$ and $\mathrm{res}_z(e,d)$ are relatively prime. (Note that if f and e are well-placed then e and d are relatively prime.) Let $g(x,y,z,w)$ be an integral polynomial which is of positive degree in w and is prime to f, and let $r(x,y,z)$ denote the resultant of f and g with respect to w. We say that $g(x,y,z,w)$ is *well-positioned with respect to f and e* if $\mathrm{discr}_z(e)$ and $\mathrm{res}_z(e,r)$ are relatively prime. (Note that if g is well-positioned with respect to f and e then e and r are relatively prime.) We say that g is a *parent* of e if r is a divisor of e.

If f and e are well-placed and, for every irreducible factor g of positive degree in w of A prime to f and not a parent of e, g is well-positioned with respect to f and e, then we compute $J_2 = P_{E_3}(J_3)$ using the fully restricted projection operator P_{E_3}. If these conditions are not met then we compute $J_2 = P^*_{E_3}(J_3)$ using the semi-restricted projection operator $P^*_{E_3}$, just as described in the previous section. If there is no such constraint $e(x,y,z) = 0$, then we compute $J_2 = P(J_3)$.

Suppose finally that there is a suitable constraint $e_2(x,y) = 0$. Then, as remarked in the previous section, we can compute for the third and final projection $J_1 = P_{E_2}(J_2)$ using the fully restricted projection, where E_2 is the set of all elements of J_2 which are divisors of e_2. On the other hand, if there is no such constraint, then we compute $J_1 = P(J_2)$.

The validity of this improvement of the projection process described in Section 2 rests upon the following two theorems.

THEOREM 4.1. *Let A be a set of integral polynomials in x,y,z,w and let $f(x,y,z,w)$ be an element of A which is primitive, squarefree and of positive degree with respect to w. Let $J_3 = P_{\{f\}}(A)$, as above. Let $e(x,y,z)$ be an integral polynomial which is primitive, squarefree and of positive degree in z, such that e is the product of some elements of J_3. Suppose that f and e are well-placed, define E_3 to be the set of elements of J_3 which are divisors of e, and set $J_2 = P_{E_3}(J_3)$. Let S be a connected submanifold of the plane of positive dimension and suppose that each element of J_2 is order-invariant in S. Then e is analytic delineable on S, and is order-invariant in each of its sections over S. Moreover, for every section σ of e over S, f either vanishes identically or is analytic delineable on σ.*

Proof. That e is analytic delineable on S and is order-invariant in each of its sections over S follows by Theorem 2.3 of [11] and Lemma A.3 of [9]. Moreover every element of J_3 is sign-invariant in each section of e over S, by Theorem 2.3 of [11]. Let σ be a section of e over S, and suppose that f does not vanish identically on σ. Now f is degree-invariant on σ since every element of J_3 is sign-invariant in σ. Let $d(x,y,z)$ denote the discriminant of f with respect to w. In case d vanishes nowhere in σ, the analytic delineability of f on σ follows immediately by Theorem 2 of [10]. So henceforth assume that d vanishes at some point of σ, and hence throughout σ, since d is sign-invariant in σ (because d is the product of certain elements of J_3, and every element of J_3 is sign-invariant in σ.)

We claim that $\partial e/\partial z$ vanishes nowhere in σ. Suppose that this is not the case. Then there is a point $x_0, y_0, z_0)$ in σ such that $\partial e/\partial z(x_0,y_0,z_0) = 0$. Put $\delta(x,y) = \mathrm{discr}_z(e)$. Since also $e(x_0,y_0,z_0) = 0$, we must have $\delta(x_0,y_0) = 0$ (by Lemma 7.1). Since δ is order-invariant in S (because every element of $J_2 = P_{E_3}(J_3)$ is order-invariant in S), δ vanishes throughout S. Put $\rho(x,y) = \mathrm{res}_z(e,d)$. Now ρ also vanishes throughout S, because e and d both vanish throughout σ. Since S has positive dimension, ρ and δ have infinitely many common zeros. Hence, by Bezout's theorem (Theorem III-3.1 of [12], ρ and δ have a common factor which has positive degree in either x or y (or both). This contradicts the assumption that f and e are well-placed. The claim is proved.

By connectedness of S (hence σ), it suffices to show that f is analytic delineable on σ near an arbitrary point of σ. Let (x_0, y_0, z_0) be such a point: we've shown above that $\partial e/\partial z(x_0, y_0, z_0) \neq 0$. There is no loss of generality in assuming that $(x_0, y_0, z_0) = (0, 0, 0)$. Let

$$e(x, y, z) = a_0(x, y)z^n + a_1(x, y)z^{n-1} + \cdots + a_n(x, y).$$

Now $e(0, 0, z) = \bar{q}(z)z$, for some integral polynomial $\bar{q}(z)$, with $\bar{q}(0) \neq 0$, since $\partial e/\partial z(0, 0, 0) \neq 0 = e(0, 0, 0)$. Therefore, by Hensel's Lemma (Theorem 3.1 of [11]), there is an open box B_1 about the origin in $\mathbf{R}^2$ and polynomials

$$q(x, y, z) = b_0(x, y)z^{n-1} + b_1(x, y)z^{n-2} + \cdots + b_{n-1}(x, y),$$

$$h(x, y, z) = z - c(x, y)$$

whose coefficients $b_j(x, y)$ and $c(x, y)$ are real power series in x and y, absolutely convergent in B_1, such that $q(0, 0, z) = \bar{q}(z)$, $h(0, 0, z) = z$ and

$$e(x, y, z) = q(x, y, z)h(x, y, z).$$

Since a function defined as the sum of a convergent power series is analytic (by Theorems 55 and 56 of [6]), the $b_j(x, y)$ and $c(x, y)$ are analytic in B_1. Since $q(0, 0, 0) = \bar{q}(0) \neq 0$ and q is analytic - hence continuous - near $(0, 0, 0)$, there exists $\epsilon > 0$ and an open box $B_2 \subseteq B_1$ about the origin such that $q(x, y, z) \neq 0$ for all $(x, y, z) \in B_2 \times (-\epsilon, +\epsilon)$. In the open box $B_2 \times (-\epsilon, +\epsilon)$, therefore, the real variety of e is identical with the graph of the real analytic function $z = c(x, y)$.

For $(x, y) \in B_2$, put

$$f^*(x, y, w) = f(x, y, c(x, y), w).$$

Then $f^*(x, y, w)$ is a polynomial in w whose coefficients are real analytic functions defined in B_2. Let $d^*(x, y)$ be the discriminant of $f^*(x, y, w)$ with respect to w. Then $d^*(x, y) = d(x, y, c(x, y))$, for all $(x, y) \in B_2$. Now we have

$$d^*(x, y) = \mathrm{res}_z(h, d)$$

for all $(x, y) \in B_2$, by Theorem 1 of [7]. Put $\rho(x, y) = \mathrm{res}_z(e, d)$. Since d and e are well-placed, by assumption, $\rho(x, y)$ is a nonzero polynomial. By Theorem 3 of [7],

$$\rho(x, y) = \mathrm{res}_z(q, d)\mathrm{res}_z(h, d)$$

for all $(x, y) \in B_2$.

Now Lemma A.3 of [9] applied here says that ρ is order-invariant in $S \cap B_2$ if and only if both $\mathrm{res}(q, d)$ and $\mathrm{res}_z(h, d)$ are order-invariant in $S \cap B_2$. We claim that ρ is order-invariant in $S \cap B_2$. The claim is proved as follows. We have

$$d = \mathrm{cont}(d)d_1 \ldots d_m$$

with $\mathrm{cont}(d)$ denoting the content of d as a polynomial in z and each d_j an irreducible factor of some element of $J_3 - E_3$, and

$$e = e_1 \ldots e_l$$

with each e_i an irreducible factor of some element of E_3. So

$$\rho(x, y) = \mathrm{cont}(d)^n \prod_{i,j} \mathrm{res}_z(e_i, d_j),$$

by repeated application of Theorem 3 of [7]. Now each $\mathrm{res}_z(e_i, d_j)$ is an element of $J_2 = P_{E_3}(J_3)$ by definition, and so is order-invariant in $S \cap B_2$, by hypothesis. Also $\mathrm{cont}(d)$ is a product of elements of J_2. Therefore, by Lemma A.3 of [9], $\rho(x, y)$ is order-invariant in $S \cap B_2$. This proves the claim. Hence, by Lemma A.3 of [9], $\mathrm{res}_z(h, d) = d^*$ is order-invariant in $S \cap B_2$.

It was assumed previously that f does not vanish identically on σ. Therefore f^* does not vanish identically on $S \cap B_2$. It was noted previously that f is degree-invariant on σ. Therefore f^* is degree-invariant on $S \cap B_2$. Hence, by Theorem 7.1 (which is a slight generalisation of Theorem 2 of [10]) applied to f^*, f^* is analytic delineable on $S \cap B_2$. Therefore f is analytic delineable on σ, near the origin. By connectedness of S (hence σ), f is analytic delineable on σ. □

THEOREM 4.2. *Assume the hypotheses of the previous theorem. Assume in addition that, for each irreducible factor g of positive degree in w of A prime to f and not a parent of e, g is well-positioned with respect to f and e. Then, in addition to the conclusions of the previous theorem, we have that each irreducible factor g of positive degree in w of A prime to f is sign-invariant in every section of f over each section σ of e over S on which f is analytic delineable.*

Proof. Since we have assumed the hypotheses of Theorem 3.1, we can immediately deduce the conclusions of that theorem. That is, e is analytic delineable on S and is order-invariant in each of its sections over S. Moreover, for every section σ of e over S, f either vanishes identically or is analytic delineable on σ. Let g be an irreducible factor of positive degree in w of A which is prime to f and let σ be a section of e over S on which f is analytic delineable. We must prove that g is sign-invariant in each section of f over σ. Let $r(x, y, z)$ denote $\mathrm{res}_w(f, g)$.

Suppose first that g is a parent of e, that is, $r(x, y, z)$ is a divisor of $e(x, y, z)$. Then, since e is order-invariant in σ, r is also order-invariant in σ, by Lemma A.3 of [9]. Therefore g is sign-invariant in each section of f over σ, by Theorem 2.2 of [11]. The proof is finished in this case.

So henceforth we shall assume that g is not a parent of e. Recall that a crucial assumption of the theorem being proved is that g is then well-positioned woth respect to f and e. In case r vanishes nowhere in σ, the sign-invariance of g in each section of f over σ again follows immediately by Theorem 2.2 of [11]. The proof is finished in this case also.

So henceforth we shall also assume that r vanishes at some point of σ, and hence throughout σ, since r is sign-invariant in σ (because r is the product of certain elements of J_3, and every element of J_3 is sign-invariant in σ, by Theorem 2.3 of [11]).

We claim that $\partial e/\partial z$ vanishes nowhere in σ. Suppose not. Then there is a point (x_0, y_0, z_0) in σ such that $\partial e/\partial z(x_0, y_0, z_0) = 0$. Put $\delta(x, y) = \mathrm{discr}_z(e)$. Since also $e(x_0, y_0, z_0) = 0$, we must have $\delta(x_0, y_0) = 0$ (by Lemma 7.1). Since δ is order-invariant in S (because every element of $J_2 = P_{E_3}(J_3)$ is order-invariant in S), δ vanishes throughout S. Put $\rho(x, y) = \mathrm{res}_z(e, r)$. Now ρ also vanishes throughout S, because e and

r both vanish throughout σ. Since S has positive dimension, δ and ρ have infinitely many common zeros. Hence, by Bezout's theorem, δ and ρ have a common factor which has positive degree in either x or y (or both). This contradicts the assumption that g is well-positioned with respect to f and e. The claim is proved.

By connectedness of S (hence σ), it suffices to show that g is sign-invariant in each section of f over σ near an arbitrary point of σ. Let (x_0, y_0, z_0) be such a point: we've shown above that $\partial e/\partial z(x_0, y_0, z_0) \neq 0$. There is no loss of generality in assuming that $(x_0, y_0, z_0) = (0, 0, 0)$. As in the proof of Theorem 4.1, by Hensel's Lemma (Theorem 3.1 of [11]) there exists an open box B_2 about the origin in $\mathbf{R}^2$, a real number $\epsilon > 0$, and polynomials in z $q(x, y, z)$ and $h(x, y, z) = z - c(x, y)$, whose coefficients are real power series in x and y, absolutely convergent in B_2, with $q \neq 0$ throughout $B_2 \times (-\epsilon, +\epsilon)$, such that

$$e(x, y, z) = q(x, y, z)h(x, y, z).$$

For $(x, y) \in B_2$, put

$$f^*(x, y, w) = f(x, y, c(x, y), w),$$

$$g^*(x, y, w) = g(x, y, c(x, y), w).$$

Then f^* and g^* are polynomials in w whose coefficients are real analytic functions defined in B_2. Let $r^*(x, y)$ be the resultant of $f^*(x, y, w)$ and $g^*(x, y, w)$ with respect to w. Then

$$r^*(x, y) = r(x, y, c(x, y))$$

for all $(x, y) \in B_2$. Now we have

$$r^*(x, y) = \mathrm{res}_z(h, r)$$

for all $(x, y) \in B_2$, by Theorem 1 of [7]. Put $\rho(x, y) = \mathrm{res}_z(e, r)$. Since g is well-positioned with respect to f and e by assumption, $\rho(x, y)$ is a nonzero polynomial. By Theorem 3 of [7],

$$\rho(x, y) = \mathrm{res}_z(q, r)\mathrm{res}_z(h, r)$$

for all $(x, y) \in B_2$.

We claim that ρ is order-invariant in $S \cap B_2$. The claim is proved as follows. We have

$$r = \mathrm{cont}(r) r_1 \ldots r_m$$

with each r_j an irreducible factor of some element of $J_3 - E_3$, and

$$e = e_1 \ldots e_l$$

with each e_i an irreducible factor of some element of E_3. So, where $n = \deg_z e$,

$$\rho(x, y) = \mathrm{cont}(r)^n \prod_{i,j} \mathrm{res}_z(e_i, r_j),$$

by repeated application of Theorem 3 of [7]. Now each $\mathrm{res}_z(e_i, r_j)$ is an element of $J_2 = P_{E_3}(J_3)$ by definition, and so is order-invariant in $S \cap B_2$, by hypothesis. Also $\mathrm{cont}(r)$ is a product of elements of J_2. Therefore, by Lemma A.3 of [9], ρ is order-invariant in $S \cap B_2$, proving the claim.

Hence, by Lemma A.3 of [9], $\mathrm{res}(h, r) = r^*$ is order-invariant in $S \cap B_2$. Now f^* is delineable on $S \cap B_2$ since f is delineable on σ. Therefore, by Theorem 7.2 applied to f^* and g^*, g^* is sign-invariant in each section of f^* over $S \cap B_2$. Therefore g is sign-invariant in each section of f over σ, near the origin. By connectedness of S (hence σ), g is sign-invariant in each section of f over σ. □

5. EXAMPLES

Hong and Collins have implemented algorithm QEPCAD [5] using the saclib computer algebra system. The most recent version of the program QEPCAD incorporates a number of improvements, contributed by Hong, Collins, Johnson, Encarnacion and Brown, including the application of equational constraints.

We applied the program QEPCAD to two different sample problems, the results of which are discussed below. Computing times were measured on a Sun server having a 292 MHz ultraSPARC risc processor. Unless otherwise specified, four megabytes of memory were made available for each experiment.

5.1 Solotareff's problem

We consider a special case of the Solotareff approximation problem [1]. This problem is discussed in [4]. The input formula is:

$$(\exists b)(\exists u)(\exists v)[-1 < u \wedge u < v \wedge v < 1 \wedge$$

$$u^4 + 2u^3 - au^2 - bu - (1 - a) = 2 - b \wedge$$

$$v^4 + 2v^3 - av^2 - bv - (1 - a) = -2 + b \wedge$$

$$4u^3 + 6u^2 - 2au - b = 0 \wedge$$

$$4v^3 + 6v^2 - 2av - b = 0 \wedge 2 - b > 0].$$

This formula is obtained from the one given in Section 7.3 of [10] by setting r equal to 2. One can see that the formula contains four equational constraint polynomials, two of them having v as the main variable, and two with u as the main variable. Program QEPCAD was run for this problem both with and without the use of equational constraints in the projection phase. In both cases, the program solved the problem, producing the solution formula

$$81a^3 - 180a^2 + 448a - 432 = 0.$$

When the program is run without using equational constraints, 39 projection polynomials altogether, having a total of 32 distinct irreducible factors are produced. The total time, incorporating the projection, stack construction and solution formula phases, was 1360 milliseconds.

The first time the program was run using equational constraints as described in Section 4 the program made unfelicitous choices of the pivot constraints f and e at levels 4 and 3, respectively. The choices of pivots were unfelicitous because $f = 4v^3 + 6v^2 - 2av - b$ and $e = 4u^3 + 6u^2 - 2au - b$ are not well-placed, as is easily seen. Fortunately the program allows the user to override the (somewhat arbitrary) choices of pivot, and with

$$f = v^4 + 2v^3 - av^2 - bv - b + a + 1,$$

$$e = 4u^3 + 6u^2 - 2au - b$$

f and e are well-placed. Moreover, for each irreducible factor g at level 4 prime to f, g is well-positioned with respect to f and e. So the use of equational constraints as described in Section 4 is valid with these choices of pivots.

The program produced 22 projection polynomials altogether, having a total of 26 distinct irreducible factors. The total execution time was only 110 milliseconds, which is about a tenfold speedup.

5.2 System of equations

Four integral polynomials A, B, C, D in the variables x, y, z, w, each of total degree 2, and with coefficients in the range $[-7, +7]$, were generated randomly. The following formula was constructed:

$$(\exists x)(\exists y)(\exists z)(\exists w)[A = 0 \wedge B = 0 \wedge$$

$$C = 0 \wedge D = 0]$$

and presented as input to QEPCAD. The program was first run using equational constraints, as per Section 4. The program's choices of pivot constraints ($f = A$, $e = \text{res}_w(A, D)$) was seen to be felicitous, in that f and e are well-placed, etc.

The program produced 21 projection polynomials altogether, having a total of 39 distinct irreducible factors. The program proceeded to solve the problem, producing the formula $0 = 0$ (that is, *true*) as output, in a total time of 216 seconds.

The program was also run without using equational constraints. Eight megabytes of memory were made available. The program ran out of memory during computation of the third (and last) projection set J_1, after 530 seconds. The program had constructed 65 projection polynomials - at levels 3 and 2 only - having a total of 79 distinct irreducible factors (at level 2 or above). Some of the bivariate projection factors have total degree as high as 16 and integer coefficients of length approximately 25 decimal digits. This compares with only 8 projection polynomials at levels 3 and 2 having a total of 15 irreducible factors (at level 2 or above), produced when equational constraints were used. The most complex bivariate projection factor has total degree 8 and integer coefficients of length approximately 11 decimal digits, when equational constraints were used. It thus seems unlikely that, without using equational constraints, the program could solve the problem in a reasonable amount of time, even if enough memory space could be provided.

6. DISCUSSION

The technical conditions for the validity of Theorems 4.1 and 4.2 - namely, that the equational constraint polynomials $f(x, y, z, w)$ and $e(x, y, z)$ are well-placed, etc. - seem often to hold in practice. However it is not ideal to have such conditions in these theorems. While these conditions allow the validity proofs for use of equational constraints as originally proposed for four variable problems to go through, these conditions don't seem strong enough to allow corresponding validity proofs for problems involving more than four variables to proceed. We have made some effort to attempt to remove these technical conditions from Theorems 4.1 and 4.2. However we haven't yet found a way to do this. But neither have we found a counter-example to the conjecture obtained by removing the condition that f and e be well-placed from Theorem 4.1.

It was suggested in [11] that, in order to validate the use of the restricted equational projection scheme originally proposed throughout the entire projection phase of QEPCAD in case $r > 3$, one would require stronger versions of Theorems 2.2 and 2.3 of [11]. In particular, one would need the conclusion of Theorem 2.2 of [11] to be that g is *order-invariant* in each section of f over S. However, sadly, we have since found a counter-example to such a hoped-for strengthening of Theorem 2.2 of [11]. Take $f(x, y, z) = z - x$ and $g(x, y, z) = z^2 - y^2 - x^2$. Put $R(x, y) = \text{res}_z(f, g) = -y^2$. Take S to be the x-axis in the plane $\mathbf{R}^2$. Now f is delineable on S and R is order-invariant in S (the order of $R(x, y)$ at each point of S is 2). But consider the unique section of f over S. We have $\text{ord}_{(0,0,0)} g = 2$ but $\text{ord}_{(a,0,a)} g = 1$ for every $a \neq 0$. So g is not order-invariant in σ.

Notwithstanding the counter-example given above, it remains a possibility - and an intriguing one at that - that the restricted equational projection scheme originally proposed is valid (with no technical conditions attached) for $r > 3$. No "direct" counter-example to the validity of this scheme has yet been found. Indeed preliminary experimental evidence obtained by Collins points to the scheme's general validity.

7. APPENDIX

We first state a straightforward lemma concerning the discriminant of a multivariate polynomial.

LEMMA 7.1. *Let $r \geq 2$, let x denote $(x_1, \ldots, x_{r-1})$ and let $f(x, x_r) = f_n(x)x_r^n + f_{n-1}(x)x_r^{n-1} + \cdots + f_0(x)$ be a polynomial of degree $n \geq 2$ in x_r over some field k. Let $D(x)$ be the discriminant* $\text{discr}(f)$ *with respect to x_r of f. Let $a \in k^{r-1}$ and let $d =$* $\deg f(a, x_r)$. *Then $D(a)$ equals* $\text{discr}(f(a, x_r))$, *if $d = n$, $D(a)$ equals $f_{n-1}(a)^2 \text{discr}(f(a, x_r))$, if $d = n - 1$, and $D(a)$ equals 0, otherwise.*

The proof of the lemma is quite straightforward (and would be a pleasant and instructive exercise for the reader).

THEOREM 7.1. *Let $r \geq 2$. Let U be an open subset of $\mathbf{R}^{r-1}$ and let $f(x, x_r)$ be a polynomial in x_r of positive degree whose coefficients are real-valued analytic functions defined in U. Let $D(x)$ be the discriminant of f and suppose that $D(x) \neq 0$. Let S be a connected submanifild of $\mathbf{R}^{r-1}$ such that f is degree-invariant and not identically vanishing on $S \cap U$, and D is order-invariant in $S \cap U$. Then f is analytic delineable on $S \cap U$ and is order-invariant in each of its sections over $S \cap U$.*

This is a slight generalisation of Theorem 2 of [10]: but the proof carries over with only very slight changes.

THEOREM 7.2. *Let $r \geq 2$. Let U be an open subset of $\mathbf{R}^{r-1}$ and let $f(x, x_r)$ and $g(x, x_r)$ be polynomials in x_r of positive degree whose coefficients are real-valued analytic functions defined in U. Let $R(x)$ be the resultant of f and g, and suppose that $R(x) \neq 0$. Let S be a connected submanifold* *of* $\mathbf{R}^{r-1}$ *such that* f *is delineable on* $S \cap U$ *and* R *is order-invariant in* $S \cap U$*. Then* g *is sign-invariant in each section of* f *over* $S \cap U$*.*

This is a slight generalisation of Theorem 2.2 of [11]: but the proof carries over with only very slight changes.

8. ACKNOWLEDGEMENT

The author is pleased to acknowledge fruitful discussions on the subject of this paper with George Collins, without whom the paper would not have been possible. The research was supported by a Macquarie University Research Grant, which provided partial support for a short visit to Macquarie by George Collins early in 2000.

9. REFERENCES

[1] ACHIESER, N. I. *Theory of Approximation.* Ungar, New York, 1956.

[2] ARNON, D. S., COLLINS, G. E., AND MCCALLUM, S. Cylindrical algebraic decomposition i: The basic algorithm. *SIAM Journal on Computing 13* (1984), 865–877.

[3] COLLINS, G. E. Quantifier elimination for real closed fields by cylindrical algebraic decomposition. In *Proceedings of 2nd GI Conference on Automata Theory and Formal Languages, Lecture Notes in Computer Science 33* (Berlin, 1975), Springer-Verlag, pp. 134–183.

[4] COLLINS, G. E. Quantifier elimination by cylindrical algebraic decomposition – twenty years of progress. In *Quantifier Elimination and Cylindrical Algebraic Decomposition* (Vienna, 1998), B. F. Caviness and J. R. Johnson, Eds., Springer-Verlag, pp. 8–23.

[5] COLLINS, G. E., AND HONG, H. Partial cylindrical algebraic decomposition for quantifier elimination. *Journal of Symbolic Computation 12* (1991), 299–328.

[6] KAPLAN, W. *Introduction to Analytic Functions.* Addison-Wesley, Reading, 1966.

[7] LOOS, R. Computing in algebraic extensions. In *Computing, Supplementum 4* (Vienna, 1982), Springer-Verlag, pp. 173–187.

[8] MCCALLUM, S. *An Improved Projection Operation for Cylindrical Algebraic Decomposition.* PhD thesis, University of Wisconsin–Madison, 1984.

[9] MCCALLUM, S. An improved projection operation for cylindrical algebraic decomposition of three-dimensional space. *Journal of Symbolic Computation 5* (1988), 141–161.

[10] MCCALLUM, S. An improved projection operation for cylindrical algebraic decomposition. In *Quantifier Elimination and Cylindrical Algebraic Decomposition* (Vienna, 1998), B. F. Caviness and J. R. Johnson, Eds., Springer-Verlag, pp. 242–268.

[11] MCCALLUM, S. On projection in cad-based quantifier elimination with equational constraint. In *Proceedings of the 1999 International Symposium on Symbolic and Algebraic Computation* (New York, 1999), S. Dooley, Ed., ACM Press, pp. 145–149.

[12] WALKER, R. J. *Algebraic Curves.* Springer-Verlag, New York, 1978.