On Projection in CAD-Based Quantifier Elimination with Equational Constraint

Scott McCallum

Department of Computing
Division of Information and Communication Sciences
Macquaric 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 provides a detailed partial validity proof for the restricted equational projection method outlined by Collins [4]. The proof is sufficient to validate the use of the method for the first projection of the projection phase. A further consequence is that the method can be used for both the first and second projections in the case of a trivariate quantifier elimination problem having sufficient equational constraints.

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 climination. 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 signinvariant in those sections and do not need to be delineable. (The concepts of 'section', 'sign-invariance' and 'delineability' are a standard part of the theory of cad, and are reviewed in Section 2.) He asserted therefore that the projection operation of McCallum [7, 8, 9] 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

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage, and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. ISSAC '99, Vancouver, British Columbia, Canada. © 1999 ACM 1-58113-073-2/99/07 \$ 5.00

- 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 $res(f_i, f_j)$, i < j,

and in the projection set the coefficients of f may be replaced by appropriate coefficients of the f_i .

The purpose of this paper is twofold. Our first aim is to make precise the restricted equational projection operation which is somewhat informally described in the above excerpt. Our second aim is to examine carefully the extent to which such a restricted equational projection operation can be used with complete confidence for quantifier elimination problems in which equational constraints are present. Such an examination is necessary because the validity of the restricted projection scheme informally described above is not obvious in general. The validity of the scheme is certainly clear in the bivariate case. For a general validity argument, however, one would need to pay close attention to the hypotheses and the conclusions of the main theorem on the validity of the improved projection operation for cylindrical algebraic decomposition ([9], Theorem 1). In particular one would want to assume that the elements of the restricted equational projection set are order-invariant in a cell S of (r-1)-space, and one would want to prove that each g_i is order-invariant in the sections of f over S. (The concept of 'order-invariance' is reviewed in Section 2.) While we conjecture that a result along these lines is true, we can at present prove only a somewhat weaker result (Theorem 2.3), one conclusion of which is that each g_i is only sign-invariant in the sections of f over S.

The main result of the present paper (Theorem 2.3) validates the use of the restricted equational projection operation for the *first* projection only. A relatively easy consequence of this result is that, for quantifier elimination problems involving three variables, the restricted equational projection operation can be used for both the first and second projections, where appropriate.

It will be important to establish the complete validity of the restricted equational projection method, because the method is very effective in reducing the amount of time taken by the algorithm QEPCAD for cad-based quantifier elimination described in [5] and [4], for problem instances in which equational constraints are present. The present paper represents progress toward this goal.

In Section 2 we make precise the restricted equational projection operation the partial use of which we can validate, and state the main result. Sections 3 and 4 provide the proof of the main lemma for the main validity result. Section 5 briefly sketches the (partial) use of the restricted equational projection operation in quantifier elimination.

2 Statement of Main Result

Background material describing the cad algorithm and its theory can be found in [2, 3, 4, 5, 8, 9]. We will make use of terminology and results from McCallum [8, 9]. Let f be a real analytic function defined in some open subset U of real n-space \mathbb{R}^n and let p be a point of U. We say that f has order k at p, and write $\operatorname{ord}_p f = k$, provided that k is the least non-negative integer such that some partial derivative of total order k does not vanish at p; we say that f has infinite order at p if there is no such k. We say that f is order-invariant (respectively, $\operatorname{sign-invariant}$) in a subset S of U provided that the order (respectively, sign) of f is the same at every point of S. We remark that if f is order-invariant in a connected subset S of U, then f is sign-invariant in S. However the converse is not true, as the example f(x, y) = xy, $U = \mathbb{R}^2$ and $S = \operatorname{the} x$ -axis shows.

An analytic submanifold of \mathbb{R}^n of dimension s is a nonempty set S in \mathbb{R}^n which "looks locally like real s-space \mathbb{R}^{sn} ; that is, for every point p of S, there is an analytic coordinate system about p with respect to which S is locally the intersection of n-s coordinate hyperplanes. Since the only kind of submanifold we shall consider in this paper is the analytic kind, we shall henceforth omit the term "analytic" when referring to submanifolds. McCallum [8, 9] provides a brief summary of some basic properties of submanifolds, while the reader can consult [11] for precise definitions. Actually the concept of submanifold is really only peripheral to the main content of the present paper, though it is mentioned here.

We next reproduce for the reader's convenience the definition from [8] of the important notion of delineability. An r-variate real polynomial $f(x_1, \ldots, x_r)$ is said to be delineable on a subset S of \mathbb{R}^{r-1} if

- 1. the portion of the real variety of f that lies in the cylinder $S \times \mathbb{R}$ over S consists of the union of the graphs (called *sections*) of some $k \geq 0$ continuous functions $\theta_1 < \ldots < \theta_k$ from S to \mathbb{R} ; and
- 2. there exist integers $m_1, \ldots, m_k \geq 1$ such that for every point $(a_1, \ldots, a_{r-1}) \in S$, the multiplicity of the root $\theta_i(a_1, \ldots, a_{r-1})$ of $f(a_1, \ldots, a_{r-1}, x_r)$ (considered as a polynomial in x_r alone) is m_i .

We would like to point out that the definition of delineability which we use is slightly different from the one found in [2], which defines a weaker notion of delineability. One of the theorems stated below uses the notion of analytic delineability: f is termed analytic delineable on a submanifold S of \mathbb{R}^{r-1} if f is delineable on S by means of analytic functions $\theta_1, \ldots, \theta_k$ defined in S.

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 [8, 9] 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 $\operatorname{discr}_{x_r}(f)$ with respect to x_r of elements f of A and the set of all resultants $\operatorname{res}_{x_r}(f,g)$ with respect to x_r of pairs f, g of distinct elements of A.

An r-variate real polynomial $f(x_1, \ldots, x_r)$ is said to vanish identically on a subset S of \mathbb{R}^{r-1} if $f(p_1, \ldots, p_{r-1}, x_r) = 0$ for every point (p_1, \ldots, p_{r-1}) of S, while f is said to be degree-invariant on S if the degree of $f(p_1, \ldots, p_{r-1}, x_r)$ (as a polynomial in x_r) is the same for every point (p_1, \ldots, p_{r-1}) of S.

We now recall the main theorem of [9].

Theorem 2.1 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 $r \geq 2$. Let S be a connected submanifold of \mathbb{R}^{r-1} . Suppose that each element of P(A) is order-invariant in S. Then each element of A either vanishes identically on S or is analytic delineable on S, the sections over S of the elements of A which do not vanish identically on S are pairwise disjoint, and each element of A which does not vanish identically on S is order-invariant in every such section.

We now state the main lemma for the main result (Theorem 2.3) of this paper.

Theorem 2.2 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 $R(x_1, \ldots, x_{r-1})$ be the resultant of f and g, and suppose that $R \neq 0$. Let S be a connected subset of \mathbb{R}^{r-1} on which f is delineable and in which R is order-invariant. Then g is sign-invariant in each section of f over S.

A proof of Theorem 2.2 is presented in Section 4. We now proceed to make precise the notion of the restricted equational projection operation which was described intuitively in the previous section. 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 $r \geq 2$. Let E be a subset of A. Define the restricted projection of A relative to E, $P_E(A)$, as follows:

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

Here now is our main result.

Theorem 2.3 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 $r \geq 2$. Let E be a subset of A. Let S be a connected submanifold of \mathbb{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 do not vanish identically on S are pairwise disjoint, each element of E is order-invariant in every such section, and each element of A not in E is sign-invariant in every such section.

Proof. That each element of E either vanishes identically on S or is delineable on S, that the sections over S of the elements of E which do not vanish identically on S are pairwise

disjoint and that each element of E is order-invariant in every such section immediately follow from Theorem 2.1, since $P(E) \subseteq P_E(A)$, by definition of $P_E(A)$. Let f be an element of E which does not vanish identically on E and let E be a section of E over E. Let E be an element of E. If E is clearly sign-invariant in E. Suppose that E is clearly sign-invariant in E. Suppose that E is order-invariant in E is in E in E. By hypothesis, E is order-invariant in E. Hence, by Theorem 2.2, E is sign-invariant in E.

3 Hensel's Lemma

In the next section we shall present a proof of Theorem 2.2 for the case r=3. The proof readily generalizes to arbitrary $r\geq 2$. We have treated 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 $r\geq 3$).

Our proof of Theorem 2.2 will make essential use of a classical reducibility criterion – Hensel's Lemma – for a polynomial $f(x_1, \ldots, x_{r-1}, x_r)$ in x_r whose coefficients are elements of the formal power series ring $k[[x_1, \ldots, x_{r-1}]]$, where k is any field. Since we shall be treating the case r=3, we henceforth assume that r=3 and put $(x,y,z)=(x_1,x_2,x_3)$. Briefly, Hensel's Lemma says that if f(x,y,z) factors nontrivially over k after putting x and y both equal to 0, then f factors non-trivially over k[[x,y]]. We shall state precisely a special case of the result for which $k=\mathbb{R}$, the field of all real numbers.

Theorem 3.1 Let

$$f(x, y, z) = a_0(x, y)z^n + a_1(x, y)z^{n-1} + \dots + a_n(x, y)$$

be a polynomial in z of degree $n \geq 0$ whose coefficients $a_i(x,y)$ are elements of $\mathbb{R}[[x,y]]$. Suppose that $f(0,0,z) = \bar{q}(z)z^m$, where $m \geq 0$, $\bar{q}(z) \in \mathbb{R}[z]$ and $\bar{q}(0) \neq 0$. Then there exist unique polynomials

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

$$h(x, y, z) = z^m + c_1(x, y)z^{m-1} + \dots + c_m(x, y)$$

whose coefficients $b_j(x,y)$ and $c_k(x,y)$ are elements of $\mathbb{R}[[x,y]]$, such that $q(0,0,z) = \overline{q}(z)$, $h(0,0,z) = z^m$ and f(x,y,z) = q(x,y,z)h(x,y,z). If the coefficients of f are all absolutely convergent in some neighbourhood N_1 of the origin, then there exists a neighbourhood $N_2 \subseteq N_1$ of the origin in which all of the coefficients of q and h are absolutely convergent.

Abhyankar ([1], Lecture 12) provides a proof of Hensel's Lemma for the case in which r=2 and k is an arbitrary field. When k is an arbitrary field the issue of convergence doesn't arise. However in the exercises on pages 92 and 95 he indicates the validity of Hensel's Lemma for polynomials whose coefficients are absolutely convergent real power series in one or more variables. Furthermore, Abhyankar assumes that f is a monic polynomial in z – that is, $a_0 = 1$ – and concludes that q, as well as h, is monic – that is, $b_0 = 1$. However it is quite straightforward to slightly modify the argument given by Abhyankar to yield precisely the theorem stated above.

4 Proof of Theorem 2.2

As indicated above, we shall treat here the case r=3. The proof immediately generalizes to arbitrary $r\geq 2$. Let σ be a section of f over S, determined by the continuous function $\theta:S\to\mathbb{R}$.

By connectedness of σ , it suffices to show that g is signinvariant 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 sign-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). That g is sign-invariant in σ near (a,b,c) follows easily by continuity of g in case $g(a,b,c)\neq 0$. So henceforth we will assume that g(a,b,c)=0.

There is no loss of generality in assuming that (a, b, c) = (0, 0, 0). We shall focus attention on the zero set of f near the origin using the special case of Hensel's Lemma (Theorem 3.1) stated in the previous section. Let n be the degree of f with respect to z. Since $f(0, 0, z) = \bar{q}(z)z^m$, for some real polynomial $\bar{q}(z)$ with $\bar{q}(0) \neq 0$, by Theorem 3.1 there exists a box B_1 about the origin in \mathbb{R}^2 and polynomials

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

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

whose coefficients $b_j(x,y)$ and $c_k(x,y)$ are real power series in x and y, absolutely convergent in B_1 , such that $q(0,0,z)=\bar{q}(z),\ h(0,0,z)=z^m$ and f(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 $c_k(x,y)$ 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 a box $B_2\subseteq B_1$ about the origin such that $q(x,y,z)\neq 0$ for all $(x,y,z)\in B_2\times (-\epsilon,+\epsilon)$. By continuity of θ , after refining B_2 to a smaller box about the origin, if necessary, we may assume further that for all $(x,y)\in S\cap B_2$, $\theta(x,y)\in (-\epsilon,+\epsilon)$.

We claim that for all $(x,y) \in S \cap B_2$, $\theta(x,y)$ is a root of h(x,y,z) of multiplicity m, hence the unique root of h(x,y,z). The proof is as follows. Let (x,y) be a fixed element of $S \cap B_2$. By delineability of f on S, $\theta(x,y)$ is a root of f(x,y,z) of multiplicity m. Now we have

$$f(x, y, z) = q(x, y, z)h(x, y, z),$$

and $q(x, y, \theta(x, y)) \neq 0$, by construction of B_2 and ϵ . Hence $\theta(x, y)$ is a root of h(x, y, z) of multiplicity m, as claimed.

Let P(x,y) be the resultant of h(x,y,z) and g(x,y,z). We claim that there is an analytic function Q(x,y) in B_2 such that

$$R(x,y) = Q(x,y)P(x,y)$$

for all $(x,y) \in B_2$. To prove this claim, note first that we can put $Q(x,y) = \operatorname{res}(q,g)$ in case n > m. If, on the other hand, n = m, then we can put $Q(x,y) = a_0(x,y)^p$, where $a_0(x,y)$ is the leading coefficient of f and p is the degree of g. Now Lemma A.3 of [8] applied here says that R is order-invariant in $S \cap B_2$ if and only if both Q and P are order-invariant in $S \cap B_2$. Hence by this lemma, P is order-invariant in $S \cap B_2$, since R is order-invariant in S, by hypothesis. But P(0,0) = 0, by virtue of the vanishing of h and g at (0,0,0). Therefore P(x,y) = 0 for all $(x,y) \in S \cap B_2$. Let (x,y) be a fixed point of $S \cap B_2$. Then h(x,y,z) and g(x,y,z) have a common root, since P(x,y) = 0. Hence $g(x,y,\theta(x,y)) = 0$, since $\theta(x,y)$ is

the unique root of h(x,y,z), as shown previously. We've shown that, where $N=B_2\times {\rm I\!R},\ g$ is sign-invariant in $\sigma\cap N$. That is, g is sign-invariant in σ near the origin, as required. \square

5 Quantifier Elimination Using Restricted Equational Projection

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

$$P_E(A) = \operatorname{cont}(A) \cup P_F(B),$$

where cont(A) is the set of nonzero contents with respect to x_r of the elements of A, B is the finest squarefree basis [8, 9] for the set 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 is the finest squarefree basis for prim(E).

We shall describe a variant of the algorithm QEPCAD of Collins and Hong [5]. The algorithm will accept as input a quantified formula

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

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 [8, 9] 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 the projection polynomials of all orders. In our modified version of QEPCAD the first projection operation is to be performed differently if an equational constraint of positive degree in x_r is present in ϕ . Suppose indeed that $e(x_1, \ldots, x_r)$ is an equational constraint polynomial of positive degree in x_r of ϕ . Then for the first projection we compute the restricted projection $P_E(A)$, where $E = \{e\}$. But for subsequent projections, we must use the unrestricted projection operator P. If no equational constraint of positive degree in x_r is present then we use P for every projection.

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 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 of CCHILD, which essentially pertains to the use of the projection P, is already described in [9].

The validity of the modified version of QEPCAD follows readily from Theorems 2.1 and 2.3.

The projection phase of algorithm QEPCAD could be modified still further in the special case r=3. Suppose that $e(x_1,x_2,x_3)$ is an equational constraint polynomial of positive degree in x_3 . Then, as described above, for the first projection we compute the restricted projection $J=P_E(A)$, where $E=\{e\}$. Suppose further that there is an equational constraint $r(x_1,x_2)=0$ such that r has positive degree in x_2 . (Such a constraint polynomial could occur in the input formula ϕ^* , or could be the resultant of two equational constraint polynomials of positive degrees in x_3 , as pointed out by Collins [4].) Then, for the second projection, we can compute the restricted projection $P_R(J)$, where $R=\{r\}$. The

reason is as follows. Since J is a set of bivariate polynomials, Theorem 2.3 applied to the finest squarefree basis B for prim(J) implies that each polynomial in B, and hence each polynomial in J, is order-invariant in every section of r over a cell in \mathbb{R}^1 in which every element of $P_R(J)$ is order-invariant. Thus Theorem 2.3 can be applied again to validate the first projection set $J = P_E(A)$, with respect to the sections of r in the plane.

6 Discussion

In order to validate the use of the restricted projection scheme proposed by Collins [4] throughout the entire projection phase of QEPCAD in case r > 3 one would require stronger versions of Theorems 2.2 and 2.3. In particular, one would need the conclusion of Theorem 2.2 to be that qis order-invariant in each section of f over S and one would need the fourth conclusion of Theorem 2.3 to be that each element of A is order-invariant in every section over S of each element of E which does not vanish identically on S. There appears to be a close connection between the "intersection multiplicity" of two hypersurfaces f = 0 and g = 0at a point (a_1, \ldots, a_r) of r-space and the order of the resultant R of f and g at the point (a_1, \ldots, a_{r-1}) (see Lectures 16 and 17 of [1], and Theorem 5.3 in Chapter IV of [10]). We are hopeful that a study of this relationship will yield the required strengthening of Theorems 2.2 and 2.3.

For examples illustrating the use of equational constraints in quantifier elimination the reader is referred to [4].

Acknowledgements

This research was undertaken while the author was a visiting Research Associate Professor in the Department of Computer and Information Sciences at the University of Delaware. The author was supported by NSF Grant CCR-9712246. The author is grateful to George Collins for his hospitality during his stay at the University of Delaware.

The author also wishes to acknowledge helpful comments provided by the anonymous referees.

References

- [1] ABHYANKAR, S. S. Algebraic Geometry for Scientists and Engineers. American Mathematical Society, Providence, 1990.
- [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. *Jour*nal of Symbolic Computation 12 (1991), 299–328.
- [6] KAPLAN, W. Introduction to Analytic Functions. Addison-Wesley, Reading, 1966.
- [7] McCallum, S. An Improved Projection Operation for Cylindrical Algebraic Decomposition. PhD thesis, University of Wisconsin-Madison, 1984.
- [8] McCallum, S. An improved projection operation for cylindrical algebraic decomposition of threedimensional space. *Journal of Symbolic Computation* 5 (1988), 141–161.
- [9] 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.
- [10] WALKER, R. J. Algebraic Curves. Springer-Verlag, New York, 1978.
- [11] WHITNEY, H. Complex Analytic Varieties. Addison-Wesley, Menlo Park, 1972.