1982

Cylindrical Algebraic Decomposition I: The Basic Algorithm

Dennis S. Arnon

George E. Collins

Scot McCallum

Report Number:
82-427A

CYLINDRICAL ALGEBRAIC DECOMPOSITION I:

THE BASIC ALGORITHM

by

Dennis S. Arnon
Computer Science Department
Purdue University
West Lafayette, Indiana, USA 47907

George E. Collins
Computer Science Department
University of Wisconsin - Madison
Madison, Wisconsin, USA 53706

Scott McCallum
Computer Science Department
University of Wisconsin - Madison
Madison, Wisconsin, USA 53706

CSD TR-427A
Department of Computer Sciences
Purdue University
December 5, 1983
(Revised and abridged edition of CSD TR-427, issued December 22, 1982)

ABSTRACT

Given a set of r-variate integral polynomials, a *cylindrical algebraic decomposition (cad)* of euclidean r-space E^r partitions E^r into connected subsets compatible with the zeros of the polynomials. Collins gave a cad construction algorithm in 1975, as part of a quantifier elimination procedure for real closed fields. The algorithm has subsequently found diverse applications (optimization, curve display); new applications have been proposed (term rewriting systems, motion planning). In the present two-part paper, we give an algorithm for determining the pairs of adjacent cells in a cad of E^2. This capability is often needed in applications. In Part I we describe the essential features of the r-space cad algorithm, to provide a framework for the adjacency algorithm in Part II.

Keywords: polynomial zeros, computer algebra, computational geometry, semi-algebraic geometry, real closed fields, decision procedures, real algebraic geometry.

1. **Introduction.** Given a set of r-variate integral polynomials, a *cylindrical algebraic decomposition (cad)* of euclidean r-space E^r partitions E^r into connected subsets compatible with the zeros of the polynomials. By "compatible with the zeros of the polynomials" we mean that on each subset of the cad, each of the polynomials either vanishes everywhere or nowhere. For example, consider the bivariate polynomial

$$y^4 - 2y^3 + y^2 - 3x^2y + 2x^4.$$

Its zeros comprise the curve shown in Figure 1. Figure 2 shows a cad of the plane compatible with its zeros. The cad consists of the distinct "dots", "arcs", and "patches of white space" of the Figure. (A rigorous definition of cad is given in Section 2).

Cad's were introduced by Collins in 1973 (see [COL75]) as part of a new quantifier elimination, and hence decision, method for elementary algebra

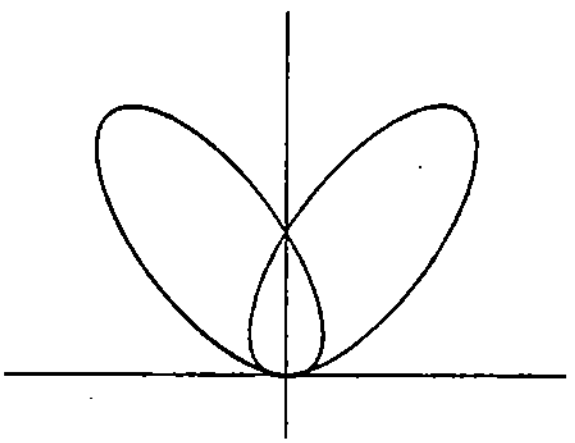

Figure 1

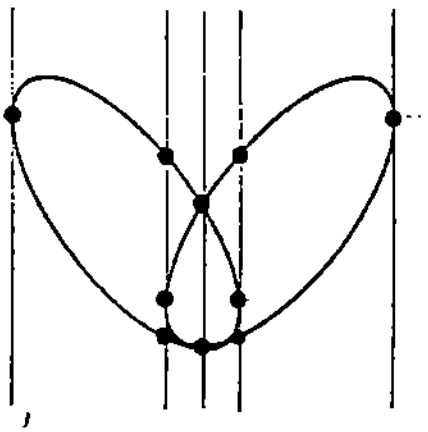

Figure 2

and geometry. He gave an algorithm for cad construction, and proved that for any fixed number of variables, its computing time is a polynomial function of the remaining parameters of input size. As can be seen in the example above, cad's are closely related to the classical simplicial and CW-complexes of algebraic topology. In fact, the essential strategy of Collins' cad algorithm, induction on dimension, can be found in van der Waerden's 1929 argument ([WAE29], pp. 360-361) that real algebraic varieties are triangulable.

Collins' cad-based decision procedure for elementary algebra and geometry is the best known (see [FER79]; very little besides a cad is needed for the decision procedure). J. Schwartz and M. Sharir used the cad algorithm to solve a motion planning problem ([SCH83a],[SCH83b]). D. Lankford [LAN78] and N. Dershowitz [DER79] pointed out that a decision procedure for elementary algebra and geometry could be used to test the termination of

term-rewriting systems. P. Kahn used cad's to solve a problem on rigid frameworks in algebraic topology ([KAH79]). Kahn also observed ([KAH78]) that a cad algorithm provides a basis for a constructive proof that real algebraic varieties are triangulable, and thus for computing the homology groups of a real algebraic variety.

Implementation of Collins' cad algorithm began soon after its introduction, culminating in the first complete program in 1981 [ARN81]. The program has begun to find use; in May, 1982 the termination of the term-rewriting system for group theory in the Appendix of [HUE80] was verified using it. It has also been utilized for display of algebraic curves [ARN83]. In 1977, Müller implemented certain subalgorithms of the cad algorithm and used them to solve algebraic optimization problems [MUE77].

We use a somewhat different (but equivalent) definition of cad than that in [COL75]; we devote Section 2 to it. We then take up the cad algorithm. Its intuitive strategy can be described by means of an example. Consider the curve of Figures 1 and 2. Given the bivariate polynomial which defines it, we will compute univariate polynomials whose roots contain a "silhouette" of the curve. By a "silhouette", we mean the projection onto the x-axis of the "significant points" of the curve. Its "significant points" are its singularities (e.g. self-crossings, cusps, isolated points), and the points at which its tangent is vertical. At any of its "non-significant" points, it is "well-behaved", i.e. locally the graph of a continuous real-valued function of a real argument. Suppose we decompose E^1 into the points of the silhouette and their complementary open intervals. Then the portion of the curve "over" each of these points (intervals) consists of finitely many disjoint, well-behaved "dots" ("arcs"). Our cad of the plane is made by decomposing the

portion of the plane "over" each point (interval) in E^1 into the "dots" ("arcs") of the curve, and the "arcs" ("patches") of the complement of the curve, that it contains.

For our sample curve, we compute a single univariate polynomial (its discriminant):

$$2048x^{12} - 4608x^{10} + 37x^8 + 12x^6.$$

This polynomial has five roots, whose approximate values are -1.49, -0.23, 0.0, 0.23, and 1.49. All roots but the third are projections of points with vertical tangent The third is the projection of the two singularities (self-crossings). Using the roots, we decompose the real line into points and open intervals (Figure 3). The Cartesian products of each of the eleven elements of this decomposition with a line, give us eleven vertical lines and "strips" in the plane. As we see in Figure 2, each "significant point" of the curve lies on one of the vertical lines of this decomposition, and within each "strip", the curve has finitely many disjoint, simply behaved, "arcs". The "dots" and "arcs" which make up each line, and the "arcs" and "patches of white space" which make up each "strip", constitute a cad of the plane, compatible with the curve.

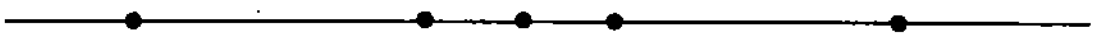

Figure 3

The general algorithm consists of three phases: projection (computing successive sets of polynomials in $I_{r-1}, I_{r-2}, \ldots, I_1$; the zeros of each set contain a "silhouette" of the "significant points" of the zeros in the next higher dimensional space), base (constructing a decomposition of E^1), and extension (successive extension of the decomposition of E^1 to a decomposition of E^2, E^2 to E^3, ..., E^{r-1} to E^r). In Sections 3,4, and 5 we describe each of these phases in turn. In the interests of succintness, we will at various times specify simple but inefficient methods of performing computations (for example, isolating the roots of a product of polynomials, rather than isolating the roots of each of the factors separately). In Section 6, we give a detailed example of the algorithm.

2. Definition of cylindrical algebraic decomposition. Connectivity plays an important role in the theory of cad's. It is convenient to have a term for a nonempty connected subset of E^r; we will call such sets *regions*. For a region R, the *cylinder over* R, written $Z(R)$, is $R \times E$. A *section* of $Z(R)$ is a set s of points $< \alpha, f(\alpha) >$, where α ranges over R, and f is a continuous, real-valued function on R. s, in other words, is the graph of f. We say such an s is the *f-section* of $Z(R)$. A *sector* of $Z(R)$ is a set $\hat{s}$ of all points $< \alpha, b >$, where α ranges over R and $f_1(\alpha) < b < f_2(\alpha)$ for (continuous, real-valued) functions $f_1 < f_2$. The constant functions $f_1 = -\infty$, and $f_2 = +\infty$, are allowed. Such an $\hat{s}$ is the (f_1, f_2)*–sector* of $Z(R)$. Clearly sections and sectors of cylinders are regions. Note that if $r = 0$ and $R = E^0 =$ a point, then $Z(R) = E^1$, any point of E^1 is a section of $Z(R)$, and any open interval in E^1 is a sector of $Z(R)$.

For any subset X of E^r, a *decomposition* of X is a finite collection of disjoint regions whose union is X. Continuous, real-valued functions $f_1 < f_2 < \cdots < f_k$, $k \geq 0$, defined on R, naturally determine a decomposition of $Z(R)$ consisting of the following regions: (1) the (f_i, f_{i+1})–sectors of $Z(R)$ for $0 \leq i \leq k$, where $f_0 = -\infty$ and $f_{k+1} = +\infty$, and (2) the f_i-sections of $Z(R)$ for $1 \leq i \leq k$. We call such a decomposition a *stack over* R (determined by $f_1, \ldots, f_k$).

A decomposition D of E^r is *cylindrical* if either (1) $r = 1$ and D is a stack over E^0, or (2) $r > 1$, and there is a cylindrical decomposition D' of E^{r-1} such that for each region R of D', some subset of D is a stack over R. It is clear that D' is unique for D, and thus associated with any cylindrical decomposition D of E^r are unique *induced* cylindrical decompositions of E^i for $i = r-1, r-2, \ldots, 1$. Conversely, given a cad $\hat{D}$ of E^i, $i < r$, a cad D of E^r is an *extension* of $\hat{D}$ if D induces $\hat{D}$.

For $0 \leq i \leq r$, an *i-cell* in E^r is a subset of E^r which is homeomorphic to E^i. It is not difficult to see that if c is an i-cell, then any section of $Z(c)$ is an i-cell, and any sector of $Z(c)$ is an $(i+1)$-cell (these observations are due to P. Kahn [KAH78]). It follows by induction that every element of a cylindrical decomposition is an i-cell for some i.

The decomposition of E^2 in Figure 2 is cylindrical. Figure 3 shows the induced decomposition of E^1, consisting of five 0-cells and six 1-cells. The decomposition in Figure 2 consists of eleven stacks. The first, or leftmost, stack consists of a single 2-dimensional sector; the next stack consists of two 1-dimensional sectors and one 0-dimensional section; and so forth.

A subset of E^r is *semi-algebraic* if it can be constructed by finitely many applications of the union, intersection, and complementation operations, starting from sets of the form

$$\{x \in E^r \mid F(x) \geq 0\},$$

where F is an element of $\mathbf{Z}[x_1, \ldots, x_r]$, the ring of integral polynomials in r variables. We write I_r to denote $\mathbf{Z}[x_1, \ldots, x_r]$. As we shall now see, a different definition of semi-algebraic set is possible, from which one obtains a useful characterization of such sets. By a *formula* we will mean a well-formed formula of the first order theory of real closed fields. (The "first order theory of real closed fields" is a precise name for what we referred to above as "elementary algebra and geometry"; see [KRE67]). The formulas of the theory of real closed fields involve elements of I_r. A *definable set* in E^k is a set S such that for some formula $\Psi(x_1, \ldots, x_k)$, S is the set of points in E^k satisfying Ψ. Ψ is a *defining formula* for S. (We follow the convention that $\varphi(x_1, \ldots, x_k)$ denotes a formula φ in which all occurrences of $x_1, \ldots, x_k$ are free, each x_i may or may not occur in φ, and no variables besides $x_1, \ldots, x_k$ occur free in φ.) A definable set is *semi-algebraic* if it has a defining formula which is quantifier-free. It is well-known that there exists a quantifier elimination method for real closed fields ([TAR48]). Hence a subset of E^r is semi-algebraic if and only if it is definable.

A decomposition is *algebraic* if each of its regions is a semi-algebraic set. A *cylindrical algebraic decomposition* of E^r is a decomposition which is both cylindrical and algebraic.

Let X be a subset of E^r, and let F be an element of I_r. F is *invariant* on X (and X is *F-invariant)*, if one of the following three conditions holds:

(1) $F(\alpha) > 0$ for all α in X. ("F has positive sign on X").
(2) $F(\alpha) = 0$ for all α in X. ("F has zero sign on X").
(3) $F(\alpha) < 0$ for all α in X. ("F has negative sign on X").

Let $A = \{A_1, \ldots, A_n\}$, be a subset of I_r ("subset of I_r" will always mean "finite subset"). X is *A-invariant* if each A_i is invariant on X. A collection of subsets of E^r is A-invariant if each element of the collection is.

The decomposition in Figure 2 is an A-invariant cad of E^2 for $A = \{y^4 - 2y^3 + y^2 - 3x^2y + 2x^4\}$. Note that a set $A \subset I_r$ does not uniquely determine an A-invariant cad D of E^r. Since any subset of an A-invariant region is also A-invariant, we can subdivide one or more regions of D to obtain another, "finer", A-invariant cad.

3. The cylindrical algebraic decomposition algorithm: projection phase. Let us begin with a more precise version of the cad algorithm outline at the end Section 1. Let $A \subset I_r$ denote the set of input polynomials, and suppose $r \geq 2$. The algorithm begins by computing a set $PROJ(A) \subset I_{r-1}$ ("*PROJ*" stands for "projection"), such that for any $PROJ(A)$-invariant cad D' of E^{r-1}, there is an A-invariant cad D of E^r which induces D'. Then the algorithm calls itself recursively on $PROJ(A)$ to get such a D'. Finally D' is extended to D. If $r = 1$, an A-invariant cad of E^1 is constructed directly.

Thus for $r \geq 2$, if we trace the algorithm, we see it compute $PROJ(A)$, then $PROJ(PROJ(A)) = PROJ^2(A)$, and so on, until $PROJ^{r-1}(A)$ has been computed. This is the projection phase. The construction of a $PROJ^{r-1}(A)$-invariant cad of E^1 is the base phase. The successive extensions of the cad of E^1 to a cad of E^2, the cad of E^2 to a cad of E^3, and so on, until an A-invariant cad of E^r is obtained, are the extension phase. We remark that for the example of Section 1, where $A = \{y^4 - 2y^3 + y^2 - 3x^2y + 2x^4\}$, $PROJ(A) =$

$\{2048x^{12} - 4608x^{10} + 37x^8 + 12x^6\}$.

The key to the projection phase is to define the map *PROJ* (which takes a subset of I_r to a subset of I_{r-1}), and to prove that it has the desired property. We stated this property above as: any $PROJ(A)$-invariant cad of E^{r-1} is induced by some A-invariant cad of E^r. To establish this, clearly it suffices to show that over any semi-algebraic, $PROJ(A)$-invariant region in E^{r-1}, there exists an A-invariant algebraic stack. In this section, we will define *PROJ* and outline the proof that it has this latter property.

Central to our definition of *PROJ* will be the notion of delineability, which we alluded to in Section 1 with the term "well-behaved". For $F \in I_r$, $r \geq 1$, let $V(F)$ denote the real variety of F, i.e. the zero set of F. Let R be a region in E^{r-1}. F is *delineable* on R if the portion of $V(F)$ lying in $Z(R)$ consists of k disjoint sections of $Z(R)$, for some $k \geq 0$. Clearly when F is delineable on R, it gives rise to a stack over R, namely the stack determined by the continuous functions whose graphs make up $V(F) \cap Z(R)$. We write $S(F,R)$ to denote this stack, and speak of the *F-sections* of $Z(R)$. One easily sees that $S(F,R)$ is F-invariant.

For example, consider again $F(x,y) = y^4 - 2y^3 + y^2 - 3x^2y + 2x^4$. F is delineable on each of the eleven cells shown in Figure 1, and in fact the stacks which comprise the cad of Figure 2 are just the stacks determined by F over these eleven cells.

Our tentative strategy for defining *PROJ* is: insure that for any $PROJ(A)$-invariant region R, the following two conditions hold: (1) each $A_i \in A$ is delineable on R, and (2) the sections of $Z(R)$ belonging to different A_i and A_j are either disjoint or identical. If these conditions are met, then clearly

we have an A-invariant stack over R, namely the stack determined by the functions whose graphs are the sections of the A_i's.

The lefthand drawing in Figure 4 illustrates a region R and hypothetical bivariate polynomials A_1, A_2, and A_3 for which these conditions do not hold. A_1 and A_2 are delineable on R, but the A_1-section meets the A_2-section. A_3 is not delineable on R. The righthand drawing in the Figure illustrates a partition of R into five regions, on each of which the conditions are satisfied.

The following example points out a difficulty with our tentative strategy. Let $A \subset I_2$ be the set $\{A_1(x,y),A_2(x,y)\} = \{x, y^2 + x^2 - 1\}$. $A_1(0,y)$ is the zero

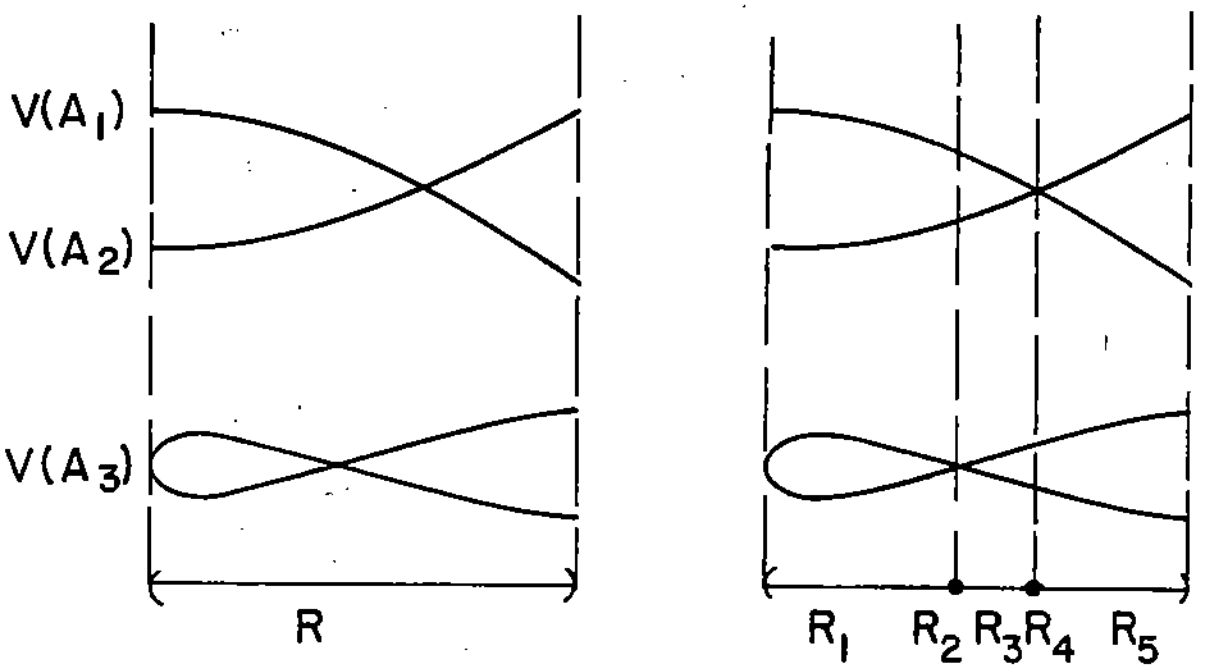

Figure 4

polynomial.

hence A_1 vanishes everywhere on $Z(\{0\})$, hence A_1 is not delineable on the set $\{0\} \subset E^1$, nor on any superset of it. We resolve this difficulty as follows. We say $F \in I_r$ is *identically zero* on $X \subset E^{r-1}$ if $F(\alpha, x_r)$ is the zero polynomial for every $\alpha \in X$. If F is identically zero on X, then any decomposition of $Z(X)$ will be F-invariant. Hence we may simply ignore F in decomposing $Z(X)$. In particular, in our example, we need only take account of the sections of A_2 in decomposing $Z(\{0\})$. Thus, we modify condition (1) above to read "(1') each $A_i \in A$ is either delineable or identically zero on R".

$PROJ(A)$ will consist of two kinds of elements: those designed to attend to condition (1'), and those to attend to condition (2). Elements of both kinds are formed from the coefficients of the polynomials of A by addition, subtraction, and multiplication (remark: I_r consists of polynomials in x_r whose coefficients are elements of I_{r-1}). We now specify exactly how this is done.

Let J be any unique factorization domain, and let F and G be nonzero elements of $J[x]$. We write $deg(F)$ to denote the degree of F (the zero polynomial has degree $-\infty$). Let $n = min(deg(F), deg(G))$. For $0 \leq j < n$, let $S_j(F,G)$ denote the j^{th} subresultant of F and G, an element of $J[x]$ of degree $\leq j$. (each coefficient of $S_j(F,G)$ is the determinant of a certain matrix of F and G coefficients; see [LOO82b], [BRT71], or [COL75] for the exact definition). For $0 \leq j < n$, the *j^{th} principal subresultant coefficient of F and G*, written $psc_j(F,G)$, is the coefficient of x^j in $S_j(F,G)$. We define $psc_n(F,G)$ to be $1 \in J$.

The following theorem is the basis for definition of the first class of elements of $PROJ(A)$. Some notation: suppose F is an element of I_r. The *derivative* of F, written F', is the partial derivative of F with respect to x_r. $deg(F)$ is the degree of F in x_r. For $\alpha \in E^{r-1}$, we write $F_\alpha(x_r)$ or F_α to denote $F(\alpha, x_r)$.

THEOREM 3.1. *Let $F \in I_r$, $r \geq 2$, and let R be a region in E^{r-1}. Suppose that $deg(F_\alpha)$ is constant and nonnegative for $\alpha \in R$, and that if positive, then the least k such that $psc_k(F_\alpha, F'_\alpha) \neq 0$ is constant for $\alpha \in R$. Then F is delineable on R.*

A proof is given in [ACM82] (Theorem 3.6). The essential ideas are contained in the proof of Theorem 4 of [COL75]. [COL75] uses a definition of delineability stronger than ours, but a polynomial delineable by that definition is delineable by ours.

The theorem suggests that for a $PROJ(A)$-invariant region R, and for each $A_i \in A$, we should have $deg((A_i)_\alpha)$ constant for $\alpha \in R$. This may be a nontrivial requirement. Consider, for example, the polynomial

$$F(x,y,z) = (y^2 + x^2 - 1)z^3 + (x - 1)z^2 + (x - 1)^2 + y^2.$$

If R is a region in the plane disjoint from the unit circle, then F_α has degree 3. If R is a subset of the unit circle which does not contain the point $< 1,0 >$, then F_α has degree 2. If R is the point $<1,0>$, then F_α is the zero polynomial. *PROJ* must separate these cases. The theorem also suggests that if R is a $PROJ(A)$-invariant region which does not contain the point $<1,0>$, then we should also insure that the least k such that $psc_k(F_\alpha, F'_\alpha) \neq 0$ is constant for $\alpha \in R$.

We introduce the notion of the reductum of a polynomial. For any nonzero $F \in I_r = I_{r-1}[x_r]$, $ldcf(F)$ denotes the leading coefficient of F. The *leading term* of F, written $ldt(F)$, is

$$ldcf(F) \cdot x_r^{deg(F)}.$$

The *reductum* of F, written $red(F)$, is $F - ldt(F)$. If $F = 0$, we define $red(F) = 0$. For any $k \geq 0$, the *kth reductum of* F, written $red^k(F)$, is defined by induction on k:

$$red^0(F) = F.$$

$$red^{k+1}(F) = red(red^k(F)).$$

For any $F \in I_r$, the *reducta set of* F, written $RED(F)$, is

$$\{red^k(F) \mid 0 \leq k \leq deg(F) \ \ \& \ red^k(F) \neq 0\}.$$

Thus the reducta set of our sample $F(x,y,z)$ above is

$$\{F(x,y,z),\ (x-1)z^2 + (x-1)^2 + y^2,\ (x-1)^2 + y^2\}.$$

We can now incorporate the notion of reductum into a specification of the (first) desired property of a $PROJ(A)$-invariant region R. For each $F \in A$, there should exist an m such that $deg(F_\alpha) = m$ for all $\alpha \in R$. Furthermore, if m is positive, then where i is such that $deg(red^i(F)) = m$, and $Q = red^i(F)$, the least k such that $psc_k(Q_\alpha, Q'_\alpha) \neq 0$ should be constant for $\alpha \in R$.

Let F and G be nonzero elements of $I_r[x]$. Let $n = min(deg(F), deg(G))$. The *psc set of* F *and* G, written $PSC(F,G)$, is

$$\{psc_j(F,G) \mid 0 \leq j \leq n \ \ \& \ psc_j(F,G) \neq 0\}$$

If either $F = 0$ or $G = 0$, then $PSC(F,G)$ is defined to be the empty set. Let $A = \{A_1, \ldots, A_n\}$, $n \geq 1$, be a set of polynomials in I_r, $r \geq 2$. $PROJ_1(A) \subset I_{r-1}$.

the first class of polynomials in $PROJ(A)$, is defined as follows. For each $1 \leq i \leq n$, let $R_i = RED(A_i)$. Then

$$PROJ_1(A) = \bigcup_{i=1}^{n} \bigcup_{G_i \in R_i} (\{ldcf(G_i)\} \cup PSC(G_i, G'_i))$$

With the following simple observation, we can prove that $PROJ_1$ behaves as we want. Suppose F and G are nonzero elements of I_r, and suppose that for some $\alpha \in E^{r-1}$, $deg(F) = deg(F_\alpha) \geq 0$, and $deg(G) = deg(G_\alpha) \geq 0$. Let $n = min(deg(F), deg(G))$. Then for every j, $0 \leq j \leq n$, it is the case that $(psc_j(F,G))_\alpha = psc_j(F_\alpha, G_\alpha)$. We see this as follows. For $j < n$, since $deg(F) = deg(F_\alpha)$ and $deg(G) = deg(G_\alpha)$, the matrix obtained by evaluating the entries of the Sylvester matrix of F and G at α is just the Sylvester matrix of F_α and G_α, hence if $j < n$ then $(S_j(F,G))_\alpha$ is equal to $S_j(F_\alpha, G_\alpha)$, and so $(psc_j(F,G))_\alpha = psc_j(F_\alpha, G_\alpha)$. If $j = n$, then $(psc_j(F,G))_\alpha = psc_j(F_\alpha, G_\alpha) = 1$.

THEOREM 3.2. *For $A \subset I_r$, $r \geq 2$, if R is a $PROJ_1(A)$-invariant region in E^{r-1}, then every element of A is either delineable or identically zero on R.*

Proof. Consider any $F \in A$. If $F = 0$, then F is identically zero on R. Suppose $F \neq 0$. By definition, $PROJ_1(A)$ includes every nonzero coefficient of F, so each coefficient of F either vanishes everywhere or nowhere on R. Hence $deg(F_\alpha)$ is constant for $\alpha \in R$. Let $deg_R(F)$ denote this constant value. If $deg_R(F) = -\infty$, then F is identically zero on R. If $deg_R(F) = 0$, then obviously F is delineable on R. Suppose $deg_R(F) \geq 1$. Then there is a unique reductum Q of F such that $deg(Q) = deg_R(Q) = deg_R(F)$. Then $F_\alpha = Q_\alpha$ for all $\alpha \in R$, hence if Q is delineable on R, then F is delineable on R. Since $PSC(Q, Q') \subset PROJ(A)$, the least k such that $(psc_k(Q, Q'))_\alpha \neq 0$ is constant for

$\alpha \in R$. Hence by our observation above, the least k such that $psc_k(Q_\alpha, Q'_\alpha) \neq 0$ is constant for $\alpha \in R$. Hence by Theorem 3.1, Q is delineable on R, hence F is delineable on R. Thus every element of A is either identically zero or delineable on R. ▪

The following theorem is the basis for definition of the second class of elements of $PROJ(A)$.

THEOREM 3.3. *Let $A \subset I_r$, $r \geq 2$, and let R be a region in E^{r-1}. Suppose that for every $F \in A$, the hypotheses of Theorem 3.1 are satisfied. Suppose also that for every $F, G \in A$, $F \neq G$, the least k such that $psc_k(F_\alpha, G_\alpha) \neq 0$ is constant for $\alpha \in R$. Then every $F \in A$ is delineable on R, and for every $F, G \in A$, any F-section and any G-section of $Z(R)$ are either disjoint or identical.*

A proof is given in [ACM82] (Theorem 3.7). The essential ideas are contained in the proof of Theorem 5 of [COL75].

Let A and R_i be as in the definition of $PROJ_1$. Let

$$PROJ_2(A) = \bigcup_{1 \leq i < j \leq n} \bigcup_{G_i \in R_i \ \& \ G_j \in R_j} PSC(G_i, G_j)$$

$PROJ(A)$ is the union of $PROJ_1(A)$ and $PROJ_2(A)$.

We prove that $PROJ$ works, i.e. that conditions (1') and (2) are satisfied for a $PROJ(A)$-invariant region:

THEOREM 3.4. *For $A \subset I_r$, $r \geq 2$, if R is a $PROJ(A)$-invariant region in E^{r-1}, then every element of A is either delineable or identically zero on R, and for every $F, G \in A$, any F-section and any G-section of $Z(R)$ are either disjoint or identical.*

Proof. By Theorem 3.2, every element of A is either delineable or identically zero on R. Let B be the set of elements of A which are delineable on R. Now by an argument similar to that used in the proof of Theorem 3.2, using Theorem 3.3 in place of Theorem 3.1, it follows that for every $F, G \in A$, any F-section and any G-section of $Z(R)$ are either disjoint or identical. •

This completes the proof that if R is a $PROJ(A)$-invariant region in E^{r-1}, then there exists an A-invariant stack over R, namely the stack whose sections are the sections of those A_i's in A which are delineable on R. We write $S(A,R)$ to denote this stack. Our agenda for this Section will be completed by showing that if also R is semi-algebraic, then $S(A,R)$ is algebraic. By our remarks in Section 2, it suffices to show that each region of $S(A,R)$ is definable. Let x denote $< x_1, \ldots, x_{r-1} >$ and y denote x_r. Any section of S is an F-section of $Z(R)$ for some $F \in A$ which is delineable on R; say that it is the j^{th} section of $S(F,R)$ (where sections are numbered from bottom to top). Then we can define it as the set of $<x,y>$ satisfying a formula "$x \in R$ and y is the j^{th} real root of $F(x,y)$". If φ is a defining formula for R, then the following is such a formula:

$$\varphi(x) \;\&\; (\exists y_1)(\exists y_2) \cdots (\exists y_{j-1})[\, y_1 < y_2 < \cdots < y_{j-1} < y$$
$$\&\; F(x,y_1) = 0 \;\&\; F(x,y_2) = 0 \;\&\; \cdots \;\&\; F(x,y_{j-1}) = 0 \;\&\; F(x,y) = 0$$
$$\&\; (\forall y_{j+1})\{\, (\, y_{j+1} \neq y_1 \;\&\; y_{j+1} \neq y_2 \;\&\; \cdots \;\&\; y_{j+1} \neq y_{j-1} \;\&$$
$$y_{j+1} \neq y \;\&\; F(x,y_{j+1}) = 0\,) \Rightarrow y_{j+1} > y \;\}\;].$$

The sectors of S can now be defined using the defining formulas for the sections: any sector of S is either the set of $<x,y>$ between two sections of S, or the set of $<x,y>$ above the topmost section of S, or the set of $<x,y>$

below the bottommost section of S. This concludes the proof.

4. The cylindrical algebraic decomposition algorithm: base phase. Let us use the precise definition of cad given in Section 2 to give precise specifications for a cad algorithm. Its input is a set $A \subset I_r$, $r \geq 1$. Its output is a description of an A-invariant cad D of E^r. This description should inform one of the number of cells in the cad, how they are arranged into stacks, and the sign of each element of A on each cell. In this section, we define the index of a cell in a cad; the cad algorithm meets the first two of the above requirements by producing a list of indices of the cells of the cad of E^r that it constructs. We also define in this section an exact representation for algebraic points in E^r, that is, points whose coordinates are all real algebraic numbers. The cad algorithm constructs, for each cell of the cad of E^r, an exact representation of a particular algebraic point belonging to that cell (we call this a *sample point* for the cell). The sign of $A_i \in A$ on a particular cell can then be determined by evaluating A_i (exactly) at the cell's sample point, and in this way we meet the third requirement above.

Where $A \subset I_r$ is the input to the cad algorithm, in the projection phase we computed $PROJ(A)$, $PROJ^2(A)$, and finally $K = PROJ^{r-1}(A) \subset I_1$. It is the task of the base phase to construct a K-invariant cad D^* of E^1, that is, to construct cell indices and sample points for the cells of such a cad. Let us now define cell indices.

In a cad of E^1, the index of the leftmost 1-cell (the 1-cell with left endpoint $-\infty$), is (1). The index of the 0-cell (if any) immediately to its right is (2), the index of the 1-cell to the right of that 0-cell (if any) is (3), etc. Suppose that cell indices have been defined for cad's of E^{r-1}, $r \geq 2$. Let D be a

cad of E^r. D induces a cad D' of E^{r-1}. Any cell d of D is an element of a stack $S(c)$ over a cell c of D'. Let $(i_1, \ldots, i_{r-1})$ be the index of c. The cells of $S(c)$ may be numbered from bottom to top, with the bottommost sector being called cell 1, the section above it (if any) cell 2, the sector above that (if any) cell 3, etc. If d is the j^{th} cell of the stack by this numbering, then its cell index is $(i_1, \ldots, i_{r-1}, j)$.

The sum of the parities of the components of a cell index is the dimension of the cell (even parity = 0, odd parity = 1). In a cad of E^2, for example, cell (2,4) is a 0-cell, (2,5) is a 1-cell.

We begin the base phase by constructing the set of all distinct (i.e. relatively prime) irreducible factors of nonzero elements of K (see [KAL82] for polynomial factorization algorithms). Let $M = \{M_1, \ldots, M_k\} \subset I_1$ be the set of these factors. The real roots $\alpha_1 < \cdots < \alpha_n$, $n \geq 0$, of $\prod M$ will be the 0-cells of D^* (if $n = 0$ then D^* consists of the single 1-cell E^1). We determine the α_j's by isolating the real roots of the individual M_i's [CLO82]. By their relative primeness, no two elements of M have a common root. Hence by refining the isolating intervals for the α_j's, we obtain a collection of disjoint left-open and right-closed intervals $(r_1, s_1]$, $(r_2, s_2]$ ..... $(r_n, s_n]$ with rational endpoints, each containing exactly one α_j, and with $r_1 < s_1 \leq r_2 < \cdots$

As soon as we know n, we can trivially write down the indices of the $2n+1$ cells of D^*. To describe sample point construction, we first define a representation for an algebraic point in E^i, $i \geq 1$. Loos ([LOO82a], Section 1) describes the representation of a real algebraic number γ by its minimal polynomial $M(x)$, and an isolating interval for a particular root of $M(x)$. With γ so represented, and letting $m = deg(M)$, one can represent any ele-

ment of $Q(\gamma)$ as an element of $Q[x]$ of degree $\leq m$ - 1 (as Loos describes). For an algebraic point in E^i, there exists a real algebraic γ such that each coordinate of the point is in $Q(\gamma)$; γ is a *primitive element* for the point. Our representation for the point is: a primitive element γ and an i-tuple of elements of $Q(\gamma)$, all represented as described by Loos.

For the 1-cells of D^* we primarily use appropriately chosen (rational) endpoints from the isolating intervals above as sample points. However, if $s_i = r_{i+1}$ is a 0-cell, we find (by bisecton) a positive rational ε, such that $(r_{i+1} + \varepsilon, s_{i+1}]$ isolates α_{i+1}, and use $r_{i+1} + \varepsilon$ as sample point for cell $(2i + 1)$. Also, we use $s_n + 1$ as a sample point for cell $(2n + 1)$. If $D^* = \{E^1\}$, we use an arbitrary rational number. Obviously the only point in a 0-cell is the cell itself. Its value is an algebraic number. Thus all our sample points for D^* are algebraic numbers, and hence can be trivially expressed in our just-defined algebraic point representation. Examples of sample points for a cad of E^1 can be found in Section 6.

5. The cylindrical algebraic decomposition algorithm: extension phase. First, consider the extension of the cad D^* of E^1 to a cad of E^2. In the projection phase, we computed a set $J = PROJ^{r-2}(A) \subset I_2$. Let c be a cell of D^*. The stack $S(J,c)$ is a subset of the cad of E^2 we are building. Let α be the sample point for c, and let J_c be the product of all elements G of J for which $G(\alpha,x_2)\neq 0$. Using algorithms for exact arithmetic in $Q(\alpha)$ [L0082a], we construct $J_c(\alpha,x_2)\in Q(\alpha)[x_2]$. We isolate the real roots of $J_c(\alpha,x_2)$ ([L0082a], Section 2). This determines $S(J,c)$: β is a root of $J_c(\alpha,x_2)$ if and only if $<\alpha,\beta>$ lies on a section of $S(J,c)$. For each such β, we use the representation for α, the isolating interval for β, and the algorithms NORMAL

and SIMPLE of [LO082a] to construct a primitive element γ for $Q(\alpha,\beta)$; we use γ to construct a representation of the form we require for $<\alpha,\beta>$. We get sector sample points for $S(J,c)$ from α and the (rational) endpoints of the isolating intervals for the roots of $J_c(\alpha,x_2)$, much as was done in Section 4 for E^1. Thus sector sample points are of the form $<\alpha,r>$, r rational, so we can take $\gamma = \alpha$. Given the cell index for c, and the isolated roots of $J_c(\alpha,x_2)$, we can trivially write down the indices for the cells of $S(J,c)$ (as for E^1 in Section 4).

After processing each cell c of D^* in this fashion, we have determined a cad of E^2 and constructed a sample point for each cell.

Extension from E^{i-1} to E^i for $3 \le i \le r$ is essentially the same as from E^1 to E^2. The only difference is that a sample point in E^{i-1} has $i-1$, instead of just one, coordinates. But where α is the primitive element of E^{i-1} sample point, and $F = F(x_1,\ldots,x_i)$ an element of I_i, arithmetic in $Q(\alpha)$ still suffices for constructing the univariate polynomial over $Q(\alpha)$ that results from substituting the coordinates $<a_1,\ldots,a_{i-1}>$ for $<x_1,\ldots,x_{i-1}>$ in F.

The following abstract algorithm summarizes our discussion of the cad algorithm.

CAD(r,A;I,S)

Inputs: r is a positive inteeger. A is a list of $n \ge 0$ integral polynomials in r variables.

Outputs: I is a list of the indices of the cells comprising an A-invariant cad D of E^r. S is a list of sample points for D.

(1) [$r = 1$.] If $r > 1$ then go to 2. Set $I \leftarrow$ the empty list. Set $S \leftarrow$ the empty list. Isolate the real roots of the irreducible factors of the nonzero elements of A. Construct the indices of the cells of D and add them to I. Construct sample points for the cells of D and add them to S. Exit.

(2) [$r > 1$.] Set $P \leftarrow PROJ(A)$. Call CAD recursively with inputs $r-1$ and P to obtain outputs I' and S' that specify a cad D' of E^{r-1}. Set $I \leftarrow$ the empty list. Set $S \leftarrow$ the empty list. For each cell c of D', let i denote the index of c, let α denote the sample point for c, and carry out the following four steps: first, set $h(x_r) \leftarrow \prod\{A_j(\alpha, x_r) \mid A_j \in A \ \& \ A_j(\alpha, x_r) \neq 0\}$; second, isolate the real roots of $h(x_r)$; third, use i, α, and the isolating intervals for the roots of h to construct cell indices and sample points for the sections and sectors of a stack over c; fourth, add the new indices to I and the new sample points to S. Exit ▪

6. An example. We now show what algorithm CAD does for a particular example in E^2. Let

$$A_1(x,y) = 144y^2 + 96x^2y + 9x^4 + 105x^2 + 70x - 98,$$

$$A_2(x,y) = xy^2 + 6xy + x^3 + 9x,$$

and $A = \{A_1, A_2\}$. CAD is called with input A. We compute $PROJ(A)$:

$ldcf(A_1) = 144,$

$psc_0(A_1, A_1') = -580608(x^4 - 15x^2 - 10x + 14),$

$psc_1(A_1, A_1') = 1,$

$ldcf(red(A_1)) = 96x^2,$

$psc_0(red(A_1), [red(A_1)]') = 1,$

$ldcf(red^2(A_1)) = 9x^4 + 105x^2 + 70x - 98,$

$ldcf(A_2) = x,$

$psc_0(A_2, A_2') = 4x^5,$

$psc_1(A_2, A_2') = 1,$

$ldcf(red(A_2)) = 6x,$

$psc_0(red(A_2), [red(A_2)]') = 1,$

$ldcf(red^2(A_2)) = x(x^2 + 9),$

$psc_0(A_1\ A_2) =$

$x^2(81x^8 + 3330x^6 + 1260x^5 - 37395x^4 - 45780x^3 - 32096x^2 + 167720x + 1435204),$

$psc_1(A_1, A_2) = 96x(x^2 - 9),$

$psc_2(A_1, A_2) = 1,$

$psc_0(red(A_1), A_2) =$

$x(81x^8 + 5922x^6 + 1260x^5 + 31725x^4 - 25620x^3 + 40768x^2 - 13720x + 9604),$

$psc_1(red(A_1), A_2) = 1,$

$psc_0(A_1, red(A_2)) = -36x(3x^4 - 33x^2 - 70x - 226),$

$psc_1(A_1, red(A_2)) = 1.$

$psc_0(red(A_1), red(A_2)) = 1.$

It turns out that the roots of $p_1(x) = x^4 - 15x^2 - 10x + 14$ and $p_2(x) = x$ give us a "silhouette" of $V(A_1) \cup V(A_2)$, hence for simplicity in this example, let us set $PROJ(A) = \{p_1(x), p_2(x)\}$ (in general, $PROJ(A)$ may contain superfluous elements; [COL75] and [ARN81] describe techniques for detecting and eliminating such elements).

p_1 and p_2 are both irreducible, so we have $M_1 = p_1$ and $M_2 = p_2$ in the notation of Section 5. M_1 has four real roots with approximate values -3.26,

-1.51, 0.7, and 4.08; M_2 has the unique root $x = 0$. The following collection of isolating intervals for these roots satisfies the conditions set out in Section 5:

$$(-4,-3],\ (-2,-1],\ (-1,0],\ (\tfrac{1}{2},1],\ (4,8].$$

Since there are five 0-cells, the cell indices for the cad are (1), (2), ..., (11).

We now construct representations for the sample points of the induced cad of E^1. Each 1-cell will have a rational sample point, hence any rational γ will be a primitive element. We arbitrarily choose $\gamma = 0$. $(-1,0]$ is an isolating interval for γ as a root of its mimimal polynomial. We may take the 1-cell sample points to be -4, -2, -1, $\frac{1}{2}$, 4, and 9.

The four irrational 0-cells have as their primitive elements the four roots of $M_1(x)$. The representation for the leftmost 0-cell, for example, consists of $M_1(x)$, the isolating interval $(-4,3]$, and the 1-tuple $<x>$, where x corresponds to the element γ of $Q(\gamma)$. The 0-cell $x = 0$ is represented in the same fashion as the rational 1-cell sample points.

We now come to the extension phase of the algorithm. Let c be the leftmost 1-cell of the cad D' of E^1. $A_1(-4,y)\neq 0$ and $A_2(-4,y)\neq 0$, hence

$$A_1A_2(-4,y) = 24\,(y^2 + 6y + 25)\,(24y^2 + 256y + 601).$$

$y^2 + 6y + 24$ has no real roots, but $24y^2 + 256y + 601$ has two real roots, which can be isolated by the intervals $(-8,-7]$ and $(-4,-2]$. Thus the stack $S(c)$ has two sections and three sectors; the indices for these cells are (1,1), (1,2),..., (1,5). From the endpoints of the isolating intervals we obtain sector sample points of $<-4,-8>$, $<-4,-4>$, and $<-4,-1>$ (which will be represented in the customary fashion). The two roots γ_1 and γ_2 of

$24y^2 + 256y + 601$ are both y-coordinates for the section sample points and primitive elements for these sample points. Thus the (representations for the) section sample points are

$$\{24y^2 + 256y + 601,\ (-8,-7]\ ,<-4,y>\}$$

and

$$\{24y^2 + 256y + 601,\ (-4,-2]\ ,<-4,y>\}.$$

Now let c be the leftmost 0-cell of D'; let α also denote this point. $A_1(\alpha,y)\neq 0$ and $A_2(\alpha,y)\neq 0$; we have

$$A_1A_2(\alpha,y) = (y^2 + 6y + \alpha^2 + 9)(y + \tfrac{1}{3}\alpha^2)^2.$$

$y^2 + 6y + \alpha^2 + 9 \in Q(\alpha)[y]$ has no real roots, but obviously $y + \frac{1}{3}\alpha^2$ has exactly one; $(-8,8]$ is an isolating interval for it. Hence $S(c)$ has one section and two sectors; the indices of these cells are (2,1), (2,2), and (2,3). The appropriate representations for $< -\alpha,-8>$ and $< -\alpha,9>$ are the sector sample points. Since $y + \frac{1}{3}\alpha^2$ is linear in y, its root is an element of $Q(\alpha)$. Hence

$$\{M_1(x),\ (-4,3],\ < x, -\tfrac{1}{3}x^2 > \}$$

is the representation of the section sample point.

Thus in this particular case it was not necessary to apply the NORMAL and SIMPLE algorithms of [LOO82a] to find primitive elements for the sections of $S(c)$, and it is also not necessary for the other sample points of this example. In general, however, for a 0-cell α, $A_c(\alpha,y)$ will have nonlinear factors with real roots, and it will be necessary to apply NORMAL and SIMPLE. Saying this another way, where α is a 0-cell of D' and $< \alpha,\beta >$ is a section

sample point of D, we had in our example above $Q(\alpha,\beta) = Q(\alpha)$, but in general, $Q(\alpha)$ will be a proper subfield of $Q(\alpha,\beta)$.

The steps we have gone through above for a 1-cell and a 0-cell are carried out for the remaining cells of D' to complete the determination of the A-invariant cad D of E^2.

Although information of the sort we have described is all that would actually be produced by CAD, it may be useful to show a picture of the decomposition of the plane to which the information corresponds. The curve defined by $A_1(x,y) = 0$ has three connected components which are easily identified in Figure 5 below. The curve defined by $A_2(x,y) = 0$ is just the y-axis, i.e. the same curve as defined by $x = 0$, and cuts through the middle of the second component of $V(A_1)$. The A-invariant cad of E^2 which CAD determines is shown in Figure 5. We remark that the curve $A_1(x,y)$ is

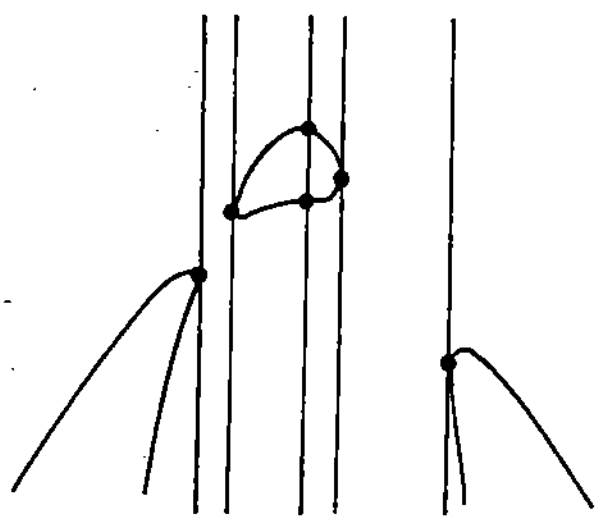

Figure 5

from ([HIL32], p. 329).

7. References

Reference for both Parts I and II are collected at the end of Part II.