Parallel Resultant Computation

Doug Ierardi*
Dexter Kozen**

TR 90-1087
January 1990

Department of Computer Science
Cornell University
Ithaca, NY 14853-7501

*Supported in part by NSF grant CCR-8901061 and in part by the NSF center for Discrete Mathematics and Theoretical Computer Science (DIMACS).
**Supported by NSF grant CCR-8901061.

Parallel Resultant Computation*

Doug Ierardi[†]
Department of Computer Science
University of Southern California
Los Angeles, California 90089

Dexter Kozen[‡]
Department of Computer Science
Cornell University
Ithaca, New York 14853

January 24, 1990

Abstract

A *resultant* is a purely algebraic criterion for determining when a finite collection of polynomials have a common zero. It has been shown to be a useful tool in the design of efficient parallel and sequential algorithms in symbolic algebra, computational geometry, computational number theory, and robotics.

We begin with a brief history of resultants and a discussion of some of their important applications. Next we review some of the mathematical background in commutative algebra that will be used in subsequent sections. The Nullstellensatz of Hilbert is presented in both its strong and weak forms. We also discuss briefly the necessary background on graded algebras, and define affine and projective spaces over arbitrary fields. We next present a detailed account of the resultant of a pair of univariate polynomials, and present efficient parallel algorithms for its computation. The theory of *subresultants* is developed in detail, and the computation of polynomial remainder sequences is derived. A resultant system for several univariate polynomials and algorithms for the gcd of several polynomials are given. Finally, we develop the theory of multivariate resultants as a natural extension of the univariate case. Here we treat both classical results on the projective (homogeneous) case, as well as more recent results on the affine (inhomogeneous) case. The *u-resultant* of a set of multivariate polynomials is defined and a parallel algorithm is presented. We discuss the computation of *generalized characteristic polynomials* and relate them to the decision problem for the theories of real closed and algebraically closed fields.

*To appear in: J. Reif (ed.), *Synthesis of Parallel Algorithms*, Morgan Kaufmann, 1990.

[†]Supported in part by NSF grant CCR-8901061 and in part by the NSF Center for Discrete Mathematics and Theoretical Computer Science (DIMACS).

[‡]Supported by NSF grant CCR-8901061.

Contents

1 Introduction

The subject of this chapter is the computation of *resultants.* A resultant is a purely algebraic criterion for determining when a finite collection of polynomial equations has a common solution, expressed in terms of the coefficients of these polynomials. The investigation of such criteria belongs historically to the branch of mathematics known as Elimination Theory, the goal of which was to solve systems of polynomial equations by successive elimination of variables. Fundamental aspects of this project were developed by Hermann, Hilbert, Kronecker, Lasker, Macaulay and Noether at the turn of this century, marking the beginning of a fusion of algebra and geometry which later found fuller expression in the development of algebraic geometry.

Much of the fundamental work in Elimination Theory was pursued at a time when constructive methods in mathematics prevailed. In fact, the ostensible goal of the theory—solving systems of polynomial equations—had such obvious practical significance that the efficiency of algorithms was already a concern. Macaulay expressed this point of view in his 1914 tract *The Algebraic Theory of Modular Systems* when he wrote that the current body of knowledge in Elimination Theory

> might be regarded as in some measure complete if it were admitted that a problem is finished with when its solution has been reduced to a finite number of feasible operations. If however the operations are too numerous or too involved to be carried out in practice the solution is only a theoretical one; and its importance then lies not in itself, but in the theorems with which it is associated and to which it leads. Such a theoretical solution must be regarded primarily as a preliminary and not the final stage in the consideration of the problem.

The study of algorithms in elimination theory has not yet reached its "final stage": provably optimal algorithms are still lacking. Nevertheless, contemporary ideas in algorithm design and complexity are continually being brought to bear, and the most efficient sequential and parallel algorithms for these problems have been discovered during the last decade.

The theory of resultants rests on the well known Nullstellensatz of Hilbert, which relates the algebra of the ring of polynomials over indeterminates $x_1, \ldots, x_n$ with coefficients in an algebraically closed field k (denoted $k[x_1, \ldots, x_n]$) and the geometry of the space k^n of n-tuples of elements of k. For example, over the complex numbers $\mathbf{C}$, the Nullstellensatz asserts that m polynomials $f_1, \ldots, f_m$ with complex coefficients have no common solutions in $\mathbf{C}^n$ exactly when there are m additional polynomials $g_1, \ldots, g_m$ such that

$$f_1 g_1 + f_2 g_2 + \cdots + f_m g_m \quad = \quad 1 \, .$$

The foundation of Elimination Theory lies in the fact that the existence or nonexistence of these g_i's can be determined solely by examining the coefficients of the f_i's, and that an algebraic criterion can be constructed uniformly once the degrees of these polynomials are specified. From this observation the notion of a *resultant* arose—a polynomial in the coefficients of the given polynomials which vanishes exactly when a common solution exists.

Stated in this way, it might seem that the resultant yields no more than a decision procedure for the existence of solutions; but the fact that it provides a *purely algebraic criterion* extends its usefulness significantly. Resultants can be used in constructing solutions to systems of equations, both symbolically and numerically (by approximation). They have been employed successfully in the design of efficient parallel and sequential algorithms in symbolic algebra, computational geometry, and computational number theory, and have found important practical applications in solid modeling and robotics.

1.1 Outline of this Chapter

This chapter presents parallel algorithms to compute the resultants of both univariate and multivariate polynomials. We begin in §2 with a review of some of the mathematical background in commutative algebra that will be used in subsequent sections. The Nullstellensatz of Hilbert is presented in both its strong and weak forms. We also discuss briefly the necessary background on graded algebras, and define affine and projective spaces over arbitrary fields.

In §3, we give a detailed account of the construction of the resultant of a pair of univariate polynomials. The treatment is also extended to deal also with several polynomials in a single variable. In exploring properties of these calculations, the theory of *subresultants* is developed in detail, and an efficient parallel algorithm for the computation of *polynomial remainder sequences* is derived in a natural way. We discuss the applications of subresultants in parallel greatest common divisor (gcd) algorithms and in computing the extended Euclidean scheme. These algorithms have played a major role in the recent development of parallel methods in real geometry. For example, the algorithm of Ben-Or, Kozen and Reif [1], which gives an efficient parallel decision procedure for questions in the theory of real closed fields, employs a variety of parallel resultant-based techniques. Although a complete discussion of these applications is beyond the scope of this article, our presentation should be sufficient to enable the interested reader to pursue more advanced topics in the literature. References to such applications are provided at the end of this chapter.

In §4, the theory of *multivariate* or *higher-order resultants* is developed as a natural extension of the univariate case. Here we treat both classical results on the projective (homogeneous) case, as well as more recent results on the affine (inhomogeneous) case. In this section our presentation is necessarily more abbreviated. Some statements whose proofs rely on deep algebraic or geometric arguments are not explored in detail. However, the general strategy for obtaining the desired algorithms from these results is explored fully.

Constructions involving multivariate resultants have played a large role in the development of efficient sequential algorithms during the past decade, but only within the last few years have special cases of these results contributed significantly to the improvement of parallel algorithm design. The *u-resultant* of a set of n polynomials in n variables is perhaps the most important tool here, and we present a simple parallel algorithm for its computation. We also discuss the computation of so-called *generalized characteristic polynomials* and demonstrate how they aid in adapting resultants for the homogeneous case to the inhomogeneous case. These techniques have been developed recently in the design of parallel algorithms for deciding questions in the theories of real closed and algebraically closed fields and for eliminating quantifiers in these theories. Although an exposition of such applications is beyond the scope of this article, we have tried to gather together the fundamental results of the modern and classical theories and present them in sufficient detail to enable the interested reader to pursue more recent work in this area.

Many of the problems considered in this chapter are ultimately reduced to computations in linear algebra, such as the computation of determinants or characterisitic polynomials. A variety of efficient parallel algorithms are known for these problems and are discussed in detail elsewhere in this volume [24].

2 Mathematical Preliminaries

This section presents much of the necessary mathematical background from commutative algebra and ring theory necessary for the following sections. We assume some familiarity with linear algebra and parallel algorithms in linear algebra as presented in [24]. The later parts of this section, on graded rings and projective spaces, are used only in §4 to develop the theory of multivariate resultants.

Unless otherwise noted, k will denote an algebraically closed field of arbitrary characteristic. In general, lower case Roman letters will denote variables and lower case Greek letters will denote elements of the field k. We write $\overline{x}$ for a sequence of elements or variables $x_1, \ldots, x_n$. If $\overline{x}$ is a sequence of n variables and $E = (e_1, \ldots, e_n)$ is a multi-index, we write $\overline{x}^E$ for the monomial $x_1^{e_1} x_2^{e_2} \cdots x_n^{e_n}$. If R is a ring, then $R[\overline{x}]$ denotes the ring of polynomials in the variables $\overline{x}$ with coefficients in R.

2.1 Polynomial Rings and Ideals

Let $R = k[x_1, \ldots, x_n]$. An *ideal* of R is a subset I of R closed under addition and closed under multiplication by elements of R. A *basis* for an ideal I is a set of polynomials B which generates I in the sense that every $f \in I$ can be written

$$f \;=\; g_1 f_1 + \cdots + g_m f_m$$

for some $f_1, \ldots, f_m \in B$ and $g_1, \ldots, g_m \in R$. When an ideal has a finite basis, we say that it is *finitely generated* and write $(f_1, \ldots, f_m)$ for the ideal generated by the polynomials $f_1, \ldots, f_m$.

Theorem 1 (Hilbert Basis Theorem) *Every ideal $I \subseteq R$ is finitely generated. In other words, every I is of the form $(f_1, \ldots, f_m)$ for some polynomials $f_1, \ldots, f_m \in R$.*

Define the *(total) degree* of a monomial $\prod_{i=1}^{n} x_i^{e_i} \in R$ to be $\sum_{i=1}^{n} e_i$. The degree of a polynomial $f \in R$ is the maximum degree of any term of f. We write R_e for the subset of all polynomials in R of degree at most $e \geq 0$. Each R_e is in fact a *vector space* over k of dimension $\binom{e+n}{n}$. We take as a basis for R_e the set of all monic[1] monomials of degree at most e:

$$\{x_1^{e_1} x_2^{e_2} \cdots x_n^{e_n} \mid e_1 + e_2 + \cdots + e_n \leq e\}\ .$$

2.2 Geometric Background

In this section we assume that k is an algebraically closed field. For example, k may be $\mathbf{C}$, the field of complex numbers.

[1] *I.e.*, with leading coefficient 1.

We denote by $\mathbf{A}_{\overline{x}}^n$ the space of n-tuples elements of k, called the *n-dimensional affine space over k with coordinate functions $\overline{x} = x_1, \ldots, x_n$.* We write $\mathbf{A}^n$ and omit the subscript when the coordinates are understood. For $f_1, \ldots, f_m \in R$, define

$$V(f_1, \ldots, f_m) \quad = \quad \{\overline{\xi} \in k^n \mid f_1(\overline{\xi}) = \cdots = f_m(\overline{\xi}) = 0\} \ ,$$

the set of common zeros of these polynomials in $\mathbf{A}^n$. A set of this form is called *algebraic.* A principle link between the geometry of algebraic sets in $\mathbf{A}^n$ and the ideal structure of the polynomial ring R is given by Hilbert's Nullstellensatz, or "theorem of the zeros". To state the Nullstellensatz in its so-called weak form, we let K be an arbitrary subfield of the algebraically closed field k and consider the polynomial ring $R' = K[x_1, \ldots, x_n] \subseteq R$. Let us begin with the case $n = 1$, where a proof of the weak form of Hilbert's theorem is more familiar. We know that any pair of univariate polynomials f_1 and f_2 have a common solution in k exactly when they have a nontrivial greatest common divisor (gcd), *i.e.* a gcd which is a non-constant polynomial. Since $K[x_1]$ is a Euclidean ring, the Euclidean algorithm for computing gcds works here and implies that there exist additional polynomials $g_1, g_2 \in R'$ such that

$$g_1 f_1 + g_2 f_2 \quad = \quad \gcd(f_1, f_2) \ .$$

So a necessary and sufficient condition for the existence of a common zero of f_1 and f_2 is that there are no polynomials g_1 and $g_2 \in R'$ such that $g_1 f_1 + g_2 f_2 = 1$. The Euclidean algorithm itself provides a sequential decision procedure in this case.

The Nullstellensatz provides an analogue of this result for the case of multivariate polynomials, in which the polynomial ring R is not Euclidean and the existence of a nontrivial gcd is not a necessary condition for the existence of common zeros.

Theorem 2 (Hilbert's Nullstellensatz, weak form) *Let $n \geq 1$ and let $R = K[x_1, \ldots, x_n]$. For $f_1, \ldots, f_m \in R$, there exist elements $\xi_1, \ldots, \xi_n \in k$ algebraic over K such that*

$$f_1(\xi_1, \ldots, \xi_n) \quad = \quad \cdots \quad = \quad f_m(\xi_1, \ldots, \xi_n) \quad = \quad 0$$

if and only if $(f_1, \ldots, f_m) \neq R$.

From the definition of ideals we know that $I = R$ just when $1 \in I$, so as an immediate corollary of the Nullstellensatz we obtain the following criterion for deciding when a set of polynomials in R has no common zero.

Corollary 3 *Let $f_1, \ldots, f_m \in R$. Then $V(f_1, \ldots, f_m) = \emptyset$ if and only if $1 \in (f_1, \ldots, f_m)$.*

Now if $f_1, \ldots, f_m$ are polynomials in R, then we know that they have no common algebraic zeros exactly when $1 \in (f_1, \ldots, f_m)$, or in other words when there exist additional polynomials $g_1, \ldots, g_m \in R$ such that

$$g_1 f_1 + g_2 f_2 + \cdots + g_m f_m \quad = \quad 1 \ .$$

The strong form of the Nullstellensatz tells us more about the relation between the ideal I and its zero set $V(I)$.

Theorem 4 (Hilbert's Nullstellensatz, strong form) *Let $I \subseteq R$ be an ideal and let $f \in R$. Then f vanishes on every point in $V(I)$ if and only if $f^m \in I$ for some $m \geq 1$.*

This theorem will be useful in defining a homogeneous analogue of the Nullstellensatz in the next section.

2.3 Homogeneous Polynomials

The following facts are used only in the development of multivariate resultants in §4 below. Our presentation parallels that of the previous sections, defining graded rings of homogeneous polynomials and a homogeneous version of the Nullstellensatz.

2.3.1 Graded Algebras

A *graded ring* is a ring S together with a collection $\{S_e \mid e \geq 0\}$ of subgroups of the additive group of S such that

- $S = \bigoplus_{e=0}^{\infty} S_e$, and[2]
- $S_d S_e \subseteq S_{d+e}$ for all $d, e \geq 0$.

An element $f \in S$ is called *homogeneous* if $f \in S_e$ for some $e \geq 0$. For $I \subseteq S$ an ideal, define $I_e = I \cap S_e$. The ideal I is called a *homogeneous ideal* if it is generated by its homogeneous elements,

$$I = \bigoplus_{d=0}^{\infty} I_d ,$$

or equivalently, if I has a basis of homogeneous elements [16, §5.13].

A typical example of a graded ring—and the one on which we focus in §4—is the polynomial ring $S = R[x_0, \ldots, x_n]$, which can be graded in the following way. A polynomial $f \in R[x_0, \ldots, x_n]$ is said to be *homogeneous of degree e* if every term of f has total degree e and *homogeneous* if it is homogeneous of some degree. We define S_e to be the collection of all polynomials in S which are homogeneous of degree e. Since any polynomial $f \in S$ of degree e can be written uniquely as a sum

$$f = f_0 + f_1 + f_2 + \cdots + f_e$$

[2]The direct sum ($\bigoplus$) signifies that every element of S can be written uniquely as a sum of the form $\sum_{i=0}^{\infty} f_i$ where each $f_i \in S_i$ and all but a finite number of these f_i's are zero.

where each f_i is homogeneous of degree i, it follows that S is a graded ring. When R is a field, it is a simple exercise to show that each subgroup S_e is in fact a *vector space* over R of dimension $\binom{e+n}{n}$. In this case, we take the set M_e of all monic monomials of degree e,

$$M_e = \{x_0^{e_0} x_1^{e_1} \cdots x_n^{e_n} \mid e_0 + e_1 + \cdots + e_n = e\} ,$$

as a basis for the R-vector space S_e. We write $n_e = \binom{e+n}{n} = |M_e|$.

2.3.2 Projective Spaces

We define the n-dimensional projective space $\mathbf{P}_x^n$ to be the set of $\sim$-equivalence classes of points in $\mathbf{A}_x^{n+1} - \{(0, \ldots, 0)\}$, where

$$(\xi_0, \ldots, \xi_n) \sim (\eta_0, \ldots, \eta_n)$$

if there exists a nonzero $\kappa \in k$ such that $\xi_i = \kappa\eta_i$, $0 \leq i \leq n$. Note that each point in $\mathbf{P}^n$ is just a line through (but excluding) the origin in $\mathbf{A}^{n+1}$. To more easily distinguish between points in affine and projective spaces, we write $(\xi_0 : \xi_1 : \cdots : \xi_n)$ for the $\sim$-equivalence class of affine points $\{(\kappa\xi_0, \ldots, \kappa\xi_n) \mid \kappa \in k \text{ and } \kappa \neq 0\}$.

Let $S = k[x_0, \ldots, x_n]$ be graded as in §2.3.1. Then $f \in S$ is homogeneous of degree e exactly when

$$f(\kappa\xi_0, \ldots, \kappa\xi_n) = \kappa^e f(\xi_0, \ldots, \xi_n)$$

for all $\xi_0, \ldots, \xi_n$ and $\kappa \in k$. This implies that the zero sets of homogeneous polynomials respect $\sim$-equivalence classes, and it is meaningful to speak of the points in projective space that are the zeros of a homogeneous polynomial. Note that the affine point $(0, \ldots, 0)$ is a zero of every homogeneous polynomial and consequently has no counterpart in projective space.

The affine space $\mathbf{A}^n$ is embedded in the projective space $\mathbf{P}^n$ under the *standard embedding*

$$(\xi_1, \ldots, \xi_n) \mapsto (1 : \xi_1 : \cdots : \xi_n) .$$

The only points of $\mathbf{P}^n$ not in the image of this map are the points of the so-called *hyperplane at infinity*:

$$\{(0 : \xi_1 : \cdots : \xi_n) \mid \text{not all } \xi_i = 0, 1 \leq i \leq n\} .$$

For any homogeneous ideal $I \subseteq S$, we will define

$$V(I) = \{\overline{\xi} \in \mathbf{P}^n \mid f(\overline{\xi}) = 0 \text{ for all } f \in I\} ,$$

the zero set of I in $\mathbf{P}^n$. As in the affine case, the geometry of $\mathbf{P}^n$ is related to the ideal structure of the graded ring S by the Nullstellensatz, now in a homogeneous form.

Theorem 5 (Homogeneous Nullstellensatz) *Let $I \subseteq S$ be a homogeneous ideal and let $f \in S_e$ for some $e > 0$. Then f vanishes on every point in $V(I)$ if and only if $f^m \in I$ for some $m > 0$.*

As in the affine case, we can use this theorem to define necessary and sufficient conditions for the emptiness of $V(I)$ for any homogeneous ideal I of S. The previous criterion that $1 \in I$ is still sufficient but no longer necessary. The difference arises because there is no projective counterpart of the affine point $(0, \ldots, 0)$.

Corollary 6 *Let $f_1, \ldots, f_m$ be homogeneous polynomials in S and let $I = (f_1, \ldots, f_m)$. Then $V(I) = \emptyset$ if and only if $I_d = S_d$ for some $d \geq 1$.*

In particular, since S_d is generated by the monomial basis M_d as a vector space over k, it suffices to show that all monomials of degree d are in I, *i.e.* that $M_d \subseteq I_d$.

3 Univariate Resultants

In this section we develop the classical *univariate* or *Sylvester resultant*, an algebraic condition on the coefficients of a pair of univariate polynomials that determines whether they have a common root.

3.1 The Sylvester Matrix and the Resultant of Two Univariate Polynomials

Let k be an algebraically closed field and x an indeterminate. Consider two univariate polynomials $f, g \in k[x]$ of degree $\deg f$ and $\deg g$, respectively:

$$\begin{aligned} f(x) &= \sum_{i=0}^{\deg f} \alpha_i x^i \\ g(x) &= \sum_{i=0}^{\deg g} \beta_i x^i \ . \end{aligned}$$

Arrange the coefficients of f and g in staggered columns to form a square matrix Φ as in the following figure, with $\deg g$ columns of coefficients of f and $\deg f$ columns of coefficients of g. The figure illustrates the case $\deg f = 5$ and $\deg g = 4$.

$$\Phi = \begin{bmatrix} \alpha_5 & 0 & 0 & 0 & \beta_4 & 0 & 0 & 0 & 0 \\ \alpha_4 & \alpha_5 & 0 & 0 & \beta_3 & \beta_4 & 0 & 0 & 0 \\ \alpha_3 & \alpha_4 & \alpha_5 & 0 & \beta_2 & \beta_3 & \beta_4 & 0 & 0 \\ \alpha_2 & \alpha_3 & \alpha_4 & \alpha_5 & \beta_1 & \beta_2 & \beta_3 & \beta_4 & 0 \\ \alpha_1 & \alpha_2 & \alpha_3 & \alpha_4 & \beta_0 & \beta_1 & \beta_2 & \beta_3 & \beta_4 \\ \alpha_0 & \alpha_1 & \alpha_2 & \alpha_3 & 0 & \beta_0 & \beta_1 & \beta_2 & \beta_3 \\ 0 & \alpha_0 & \alpha_1 & \alpha_2 & 0 & 0 & \beta_0 & \beta_1 & \beta_2 \\ 0 & 0 & \alpha_0 & \alpha_1 & 0 & 0 & 0 & \beta_0 & \beta_1 \\ 0 & 0 & 0 & \alpha_0 & 0 & 0 & 0 & 0 & \beta_0 \end{bmatrix} \qquad (1)$$
$$\underbrace{\qquad\qquad\qquad}_{\deg g} \quad \underbrace{\qquad\qquad\qquad\qquad}_{\deg f}$$

Definition 7 *The matrix* Φ *is called the* Sylvester *or* resultant matrix *of f and g, and its determinant* $\det \Phi$ *is called the* resultant *of f and g.*

Theorem 8 *The univariate polynomials f and g have a common root in k if and only if Φ is singular, i.e. if and only if the resultant of f and g vanishes.*

Proof. Equivalently, we must show that Φ is singular iff f and g have a nontrivial gcd, or in other words iff the degree of the lcm ℓ of f and g is strictly less than $\deg fg = \deg f + \deg g$.

Let $k[x]_d$ denote the space of polynomials in $k[x]$ of degree at most d. This is a vector space of dimension $d+1$ over k with standard basis $1, x, \ldots, x^d$. The spaces $k[x]_{\deg g - 1} \times k[x]_{\deg f - 1}$ and

$k[x]_{\deg f+\deg g-1}$ are both vector spaces of dimension $\deg f + \deg g$, and under the standard basis the matrix Φ denotes the linear map

$$\begin{aligned} \varphi \ &: \ k[x]_{\deg g-1} \times k[x]_{\deg f-1} \ \to \ k[x]_{\deg f+\deg g-1} \\ &: \ (s,t) \ \mapsto \ sf + tg \ . \end{aligned}$$

If Φ is singular, then the kernel of φ is nontrivial. Thus there exist nonzero s, t with $\deg s < \deg g$ and $\deg t < \deg f$ such that $sf = -tg$. Then ℓ divides $sf = -tg$, so its degree is strictly less than $\deg f + \deg g$. Conversely, if $\deg \ell < \deg f + \deg g$, then $(\frac{\ell}{f}, -\frac{\ell}{g}) \in \ker \varphi$, thus Φ is singular. $\square$

The resultant of two univariate polynomials can be computed in *NC* using Csanky's algorithm [10] in characteristic 0 or Berkowitz' [2] or Chistov's [9] algorithm in arbitrary characteristic; see [24].

Definition 9 *Consider the coefficients of f and g as indeterminates $\overline{a} = a_{\deg f}, \ldots, a_0$, $\overline{b} = b_{\deg g}, \ldots, b_0$. Then the determinant of Φ is a polynomial in $k[\overline{a}, \overline{b}]$ of degree* $\deg f + \deg g$. *This polynomial is called the* resultant polynomial.

For any specialization $\overline{\alpha}, \overline{\beta}$ of the indeterminates $\overline{a}, \overline{b}$ with $\alpha_{\deg f} \neq 0$ and $\beta_{\deg g} \neq 0$ giving polynomials $f, g \in k[x]$, the value of the resultant polynomial on $\overline{\alpha}, \overline{\beta}$ is the resultant of f and g.

3.2 Subresultants, Polynomial Remainder Sequences, and the Extended Euclidean Scheme

An important application of resultants is in the calculation of the *polynomial remainder sequences* (PRS) that accrue from the execution of the Euclidean algorithm. Coefficients of elements of the PRS can be expressed as signed quotients of products of minors of the Sylvester matrix [4]. This holds as well for the elements of the *extended Euclidean scheme* (EES), of which the PRS forms a part. Thus all coefficients of elements of the PRS and EES can be computed in *NC* [3,25]. In this section, we describe this algorithm and prove its correctness.

The basic fact on which the Euclidean algorithm rests is that for any pair of polynomials f and $g \neq 0$, there exist a unique quotient $q \in k[x]$ and remainder $r \in k[x]$ such that $\deg r < \deg g$ and $f = qg + r$. The Euclidean algorithm calculates the sequence $f_0 = f$, $f_1 = g$, f_2, f_3, ..., f_n, $f_{n+1} = 0$, where for $2 \leq m \leq n+1$, f_m is the remainder obtained by dividing f_{m-2} by f_{m-1}. The polynomial f_n is the last nonzero polynomial in the sequence and is the gcd of f and g. This sequence is known as the *polynomial remainder sequence* (PRS) of f and g.

Definition 10 *For $2 \leq m \leq n+1$, let $q_m \in k[x]$ be the quotient obtained in the division of f_{m-2} by f_{m-1}; thus*

$$f_m \ = \ f_{m-2} - q_m f_{m-1} \ .$$

Consider the polynomials $s_0, s_1, \ldots, s_n, s_{n+1}$ and $t_0, t_1, \ldots, t_n, t_{n+1}$ defined by

$$\begin{array}{rclrcl} s_0 &=& 1 & t_0 &=& 0 \\ s_1 &=& 0 & t_1 &=& 1 \\ s_m &=& s_{m-2} - q_m s_{m-1} & t_m &=& t_{m-2} - q_m t_{m-1} \ , \quad 2 \le m \le n+1 \ . \end{array}$$

The collection of all these polynomials f_m, s_m, t_m, and q_m is known as the extended Euclidean scheme *of f and g.*

The significance of s_m and t_m is given in the following lemma.

Lemma 11 *Assume that* $\deg g \le \deg f$. *For* $2 \le m \le n+1$,

(i) $\deg s_m = \deg g - \deg f_{m-1}$

(ii) $\deg t_m = \deg f - \deg f_{m-1}$

(iii) $s_m f + t_m g = f_m$.

Moreover, s_m and t_m are the unique polynomials s and t such that

(a) $\deg s < \deg g - \deg f_m$

(b) $\deg t < \deg f - \deg f_m$

(c) $sf + tg$ and f_m have the same degree and leading coefficient.

Proof. Statements (i–iii) are easily proved by induction on m. These properties also imply that s_m and t_m satisfy (a–c). Thus it remains to show uniqueness.

Let s and t be any polynomials satisfying (a–c). (Under assumption (c), the statements (a) and (b) are equivalent, so we will only need to use one.) We have

$$\begin{array}{rcl} \deg s_m(sf+tg) &<& \deg s_m + \deg f_{m-1} \text{ by (c)} \\ &=& \deg g \text{ by (i)} \ , \end{array}$$

$$\begin{array}{rcl} \deg s(s_m f + t_m g) &=& \deg s + \deg f_m \text{ by (iii)} \\ &<& \deg g \text{ by (a)} \ . \end{array}$$

Subtracting, we get

$$\begin{array}{rcl} \deg(s_m(sf+tg) - s(s_m f + t_m g)) &=& \deg(s_m t - s t_m) g \\ &<& \deg g \ , \end{array}$$

which is possible only if $st_m = s_m t$. But an easy inductive argument shows that s_m and t_m are relatively prime—in fact

$$s_m t_{m-1} - t_m s_{m-1} = (-1)^m$$

—therefore there exists a polynomial u such that $s = us_m$ and $t = ut_m$. Then

$$\begin{aligned} sf + tg &= us_m f + ut_m g \\ &= uf_m \ . \end{aligned}$$

But by (c), it must be that $u = 1$, therefore $s = s_m$ and $t = t_m$. $\square$

At this point we are ready to define subresultants. For $0 \leq d \leq \deg g - 1$, let Φ_d be the $(\deg f + \deg g - 2d) \times (\deg f + \deg g - 2d)$ submatrix of Φ obtained by deleting the last d columns of coefficients of f, the last d columns of coefficients of g, and the last $2d$ rows. The following figure illustrates Φ_d for $\deg f = 5$, $\deg g = 4$, and $d = 2$:

$$\Phi_d = \begin{bmatrix} \alpha_5 & 0 & \beta_4 & 0 & 0 \\ \alpha_4 & \alpha_5 & \beta_3 & \beta_4 & 0 \\ \alpha_3 & \alpha_4 & \beta_2 & \beta_3 & \beta_4 \\ \alpha_2 & \alpha_3 & \beta_1 & \beta_2 & \beta_3 \\ \alpha_1 & \alpha_2 & \beta_0 & \beta_1 & \beta_2 \end{bmatrix} \tag{2}$$
$$\underbrace{\qquad\quad}_{\deg g - d} \quad \underbrace{\qquad\qquad\quad}_{\deg f - d}$$

Under the standard basis, Φ_d represents the linear map

$$\begin{aligned} \varphi_d \ : \ & k[x]_{\deg g-d-1} \times k[x]_{\deg f-d-1} \ \to \ k[x]_{\deg f+\deg g-2d-1} \\ : \ & (s,t) \ \mapsto \ \text{the quotient obtained in the division of } sf+tg \text{ by } x^d. \end{aligned}$$

Definition 12 *The matrix Φ_d is called the d^{th}* subresultant matrix *of f and g, and its determinant* $\det \Phi_d$ *is called the d^{th}* subresultant *of f and g.*

Theorem 13 *The matrix Φ_d is nonsingular if and only if $d = \deg f_m$ for some f_m in the PRS. In this case, the vector of coefficients of s_m and t_m forms the unique solution x of the nonsingular system*

$$\Phi_d x = (0,\ldots,0,a_m)^T \ , \tag{3}$$

where a_m is the leading coefficient of f_m.

Proof. If $d \neq \deg f_m$ for any m, let m be the largest number such that $\deg f_{m-1} > d$. Then $2 \leq m \leq n+1$, and by Lemma 11,

$$\begin{array}{rcccl}\deg s_m & = & \deg g - \deg f_{m-1} & < & \deg g - d \\ \deg t_m & = & \deg f - \deg f_{m-1} & < & \deg f - d \\ \deg(s_m f + t_m g) & = & \deg f_m & < & d\ ,\end{array}$$

so $(s_m, t_m) \in \ker \varphi_d$, therefore Φ_d is singular.

Now suppose that $d = \deg f_m$. Then $\varphi_d(s_m, t_m) = a_m$, the leading coefficient of f_m, therefore the vector x of coefficients of s_m and t_m satisfies (3). Moreover, by Lemma 11, s_m and t_m are unique, therefore Φ_d is nonsingular. □

This theorem gives rise to an *NC* algorithm for calculating all elements of the EES.

Algorithm 14 (Extended Euclidean Scheme)

Given polynomials f and g:

1. *For each $d < \deg g$, compute the d^{th} subresultant $\det \Phi_d$. The d for which Φ_d is nonsingular are exactly the degrees of the f_m in the PRS.*

2. *For each $d = \deg f_m$, $m \geq 2$, solve the nonsingular system*

$$\Phi_d x \quad = \quad (0, \ldots, 0, 1)^T\ . \tag{4}$$

This gives the coefficients of $s'_m = \frac{s_m}{a_m}$ and $t'_m = \frac{t_m}{a_m}$. (We do not yet know a_m.)

3. *Compute $f'_m = s'_m f + t'_m g$. This is the monic associate $\frac{f_m}{a_m}$ of f_m.*

4. *Divide f'_{m-2} by f'_{m-1} using Algorithm 16 below. (Alternatively, solve equation (4) using the d^{th} subresultant matrix of f'_{m-2} and f'_{m-1}.) This gives a constant b_m and polynomial p_m such that*

$$b_m f'_m \quad = \quad f'_{m-2} - p_m f'_{m-1}\ .$$

5. *Compute*

$$a_m \quad = \quad \begin{cases} a_0 b_2 b_4 b_6 \cdots b_m\ , & m \text{ even} \\ a_1 b_3 b_5 b_7 \cdots b_m\ , & m \text{ odd} \end{cases} \tag{5}$$

$$q_m \quad = \quad \frac{a_{m-2}}{a_{m-1}} p_m \tag{6}$$

$$\begin{aligned} f_m &= a_m f'_m \\ s_m &= a_m s'_m \\ t_m &= a_m t'_m\ . \end{aligned}$$

The computations in Step 5 are justified by the following argument. In Step 4, we computed b_m and p_m such that

$$\frac{b_m}{a_m} f_m = \frac{f_{m-2}}{a_{m-2}} - p_m \frac{f_{m-1}}{a_{m-1}} ,$$

thus by the uniqueness of quotient and remainder,

$$\begin{aligned} \frac{b_m a_{m-2}}{a_m} f_m &= f_{m-2} - \frac{a_{m-2}}{a_{m-1}} p_m f_{m-1} \\ &= f_{m-2} - q_m f_{m-1} \\ &= f_m , \end{aligned}$$

whence follow (6) and the recurrence

$$a_m = b_m a_{m-2}$$

with solution (5).

The solution vector to (4) computed in Step 2 is the last column of Φ_d^{-1}, which by Cramer's rule is the last column of the adjoint of Φ_d divided by $\det \Phi_d$. This indicates that all the coefficients of s'_m and t'_m are signed quotients of minors of the Sylvester matrix.

Algorithm 14 can be implemented in *NC* using standard tools for linear algebra (see [24,25]).

3.3 Polynomial Division with Remainder

Let f, g be polynomials, $\deg g \leq \deg f$, and let q and r be the quotient and remainder respectively obtained in the division of f by g. Algorithm 14 suggests an *NC* algorithm for polynomial division with remainder: compute the subresultants to find $d = \deg r$, then solve (4) with Φ_d.

However, this algorithm has two serious liabilities:

- It requires the computation of all the subresultants.
- It requires divisions in k.

The latter becomes a major problem when the coefficients of f and g are indeterminates. Algorithm 14 expresses the coefficients of q and r as quotients of polynomials in the coefficients of f and g. This is so even when the divisor g is monic, in which case the coefficients of q and r are polynomial functions of the coefficients of f and g rather than rational functions. In the computation of the multivariate resultant to be presented in §4, it will be essential to have a polynomial division algorithm for monic g that does not use any divisions in k, but only the ring operations $\cdot$ and $+$.

Here we give a resultant-style algorithm used by Canny [6] that alleviates these problems. The algorithm is based on the following theorem.

Theorem 15 *Let $k = \deg f - \deg g + 2$, the dimension of $\Phi_{\deg g-1}$. If g is monic, then the coefficient of x^{k-i} in q is*

$$(-1)^i \det \Phi^{(i-1)}_{\deg g-1} \; , \quad 2 \le i \le k \; ,$$

where $\det \Phi^{(i)}_{\deg g-1}$ is the i^{th} principal minor (determinant of the upper left $i \times i$ submatrix) of $\Phi_{\deg g-1}$.

Proof. Let

$$\begin{aligned} f &= \sum_{i=0}^{\deg f} \alpha_i x^i & \qquad q &= \sum_{i=0}^{\deg q} \gamma_i x^i \\ g &= \sum_{i=0}^{\deg g} \beta_i x^i & \qquad r &= \sum_{i=0}^{\deg r} \delta_i x^i \end{aligned}$$

Note that $\deg q = k - 2$. The equation $r = f - qg$ is expressed in the $(\deg f + 1) \times k$ linear system

$$\begin{bmatrix} \alpha_9 & 1 & 0 & 0 & 0 \\ \alpha_8 & \beta_5 & 1 & 0 & 0 \\ \alpha_7 & \beta_4 & \beta_5 & 1 & 0 \\ \alpha_6 & \beta_3 & \beta_4 & \beta_5 & 1 \\ \alpha_5 & \beta_2 & \beta_3 & \beta_4 & \beta_5 \\ \alpha_4 & \beta_1 & \beta_2 & \beta_3 & \beta_4 \\ \alpha_3 & \beta_0 & \beta_1 & \beta_2 & \beta_3 \\ \alpha_2 & 0 & \beta_0 & \beta_1 & \beta_2 \\ \alpha_1 & 0 & 0 & \beta_0 & \beta_1 \\ \alpha_0 & 0 & 0 & 0 & \beta_0 \end{bmatrix} \begin{bmatrix} 1 \\ -\gamma_3 \\ -\gamma_2 \\ -\gamma_1 \\ -\gamma_0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ \delta_4 \\ \delta_3 \\ \delta_2 \\ \delta_1 \\ \delta_0 \end{bmatrix} \tag{7}$$

here illustrated for the case $\deg f = 9$, $\deg g = 6$, and $\deg r = 4$. The first k rows of this matrix comprise $\Phi_{\deg g-1}$.

Now consider the $k \times k$ system obtained by taking the first $k-1$ rows of (7) and last row $(1, 0, \ldots, 0)$:

$$\begin{bmatrix} \alpha_9 & 1 & 0 & 0 & 0 \\ \alpha_8 & \beta_5 & 1 & 0 & 0 \\ \alpha_7 & \beta_4 & \beta_5 & 1 & 0 \\ \alpha_6 & \beta_3 & \beta_4 & \beta_5 & 1 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ -\gamma_3 \\ -\gamma_2 \\ -\gamma_1 \\ -\gamma_0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \tag{8}$$

and let A be the $k \times k$ matrix in (8). Certainly A is nonsingular, since its determinant is $(-1)^{k+1}$. The inverse of A is given by Cramer's rule: the i,j^{th} entry of A^{-1} is

$$(-1)^{i+j} \frac{\det A_{j,i}}{\det A} \quad = \quad (-1)^{i+j+k+1} \det A_{j,i} \; ,$$

where $A_{j,i}$ is the $(k-1) \times (k-1)$ submatrix obtained from A by dropping the j^{th} row and i^{th} column. In particular, the last column of A^{-1}, which by (8) is the vector $(1, -\gamma_{k-2}, \ldots, -\gamma_0)^T$, contains

$$(-1)^{i+1} \det A_{k,i} \ , \quad 1 \leq i \leq k \ .$$

But note that for this particular matrix,

$$\begin{aligned} \det A_{k,i} &= \det A^{(i-1)} \\ &= \det \Phi^{(i-1)}_{\deg g - 1} \ , \quad 1 \leq i \leq k \ . \end{aligned}$$

Thus for $2 \leq i \leq k$,

$$\begin{aligned} \gamma_{k-i} &= -(-1)^{i+1} \det \Phi^{(i-1)}_{\deg g - 1} \\ &= (-1)^i \det \Phi^{(i-1)}_{\deg g - 1} \ . \end{aligned}$$

□

Using Theorem 15, we can give the following simple division-free algorithm for polynomial division with remainder when the divisor is monic:

Algorithm 16 (Polynomial Division with Remainder)

Given polynomials f and g, $\deg f \geq \deg g$, g monic:

1. *Compute the principal minors of $\Phi_{\deg g - 1}$.*

2. *Set*

$$\gamma_{k-i} = (-1)^i \det \Phi^{(i-1)}_{\deg g - 1} \ , \quad 2 \leq i \leq k \ ,$$

where $k = \deg f - \deg g + 2$. Then γ_j is the coefficient of x^j in q.

3. *Set $r = f - qg$.*

If g is not monic, then divisions are inevitable. However, we can apply Algorithm 16 to the monic associate of g, then adjust q afterward by dividing by the leading coefficient of g.

All operations can be implemented in NC. The principal minors of $\Phi_{\deg g - 1}$ can be computed without division using Berkowitz' [2] or Chistov's [9] algorithm.

3.4 A Resultant System for Several Univariate Polynomials

The constructions of §3.1 and §3.2 can be modified to yield *NC* algorithms for testing whether a set of univariate polynomials has a common root and for computing their gcd. We reduce these problems to the case of two univariate polynomials over a larger field. In §3.6 below, we will show that in the presence of a source of randomness, these algorithms have implementations that are no less efficient than the algorithms of §3.1 and §3.2 for two polynomials.

Let $f_0, \ldots, f_n \in k[x]$. Let y be a new indeterminate, and consider the bivariate polynomial

$$f(x,y) \quad = \quad f_0(x) + f_1(x)y + f_2(x)y^2 + \cdots + f_{n-1}(x)y^{n-1} \ . \tag{9}$$

We regard f as a polynomial in the indeterminate x with coefficients in the transcendental extension $k(y)$ of k. Thus it makes sense to consider the gcd of f and f_n over $k(y)[x]$.

Lemma 17 *The* gcd *of f and f_n is the same as the* gcd *of $f_0, \ldots, f_n$. In other words, f, f_n and $f_0, \ldots, f_n$ generate the same principal ideal in $k(y)[x]$.*

Proof. Let g be the monic gcd of $f_0, \ldots, f_n$ in $k[x]$ and let h be the monic gcd of f and f_n in $k(y)[x]$. Certainly g divides h, since g divides f and f_n. To show that h divides g, it suffices to show that h divides $f_0, \ldots, f_n$. Since h divides f_n and h is monic, $h \in k[x]$. For $0 \leq i \leq n$, let q_i and r_i be the quotient and remainder, respectively, obtained in the division of f_i by h in $k[x]$. Then $f = qh + r$, where

$$\begin{aligned} q &= \sum_{i=0}^{n-1} q_i y^i \\ r &= \sum_{i=0}^{n-1} r_i y^i \ , \end{aligned}$$

and the degree of r as a polynomial in $k(y)[x]$ is less than $\deg h$. Since h divides f in $k(y)[x]$, r must be identically zero as a polynomial in $k(y)[x]$. Since y is transcendental, r is also identically zero as a polynomial in $k[x,y]$. Thus $r_i = 0$ and h divides f_i, $0 \leq i \leq n$. □

Now form the Sylvester matrix Φ of f and f_n. The entries of Φ are polynomials in $k[y]$ of degree at most $n-1$, and the resultant $\det \Phi$ is a polynomial in $k[y]$ of degree at most $(n-1) \cdot \deg f_n$.

Theorem 18 *The polynomials $f_0, \ldots, f_n$ have a common root in k if and only if* $\det \Phi$ *vanishes identically.*

Proof. The polynomials $f_0, \ldots, f_n$ have a common root in k iff they have nontrivial gcd. By Lemma 17, this occurs iff f and f_n have a nontrivial gcd in $k(y)[x]$, *i.e.* if f and f_n have a common root in the algebraic closure of $k(y)$. By Theorem 8, this occurs iff the resultant of f and f_n vanishes identically. □

This gives rise to the following NC algorithm:

Algorithm 19 (Resultant of Several Polynomials)

For polynomials $f_0, \ldots, f_n \in k[x]$:

1. *Let y be a new indeterminate. Form the polynomial*

$$f(x,y) = \sum_{i=0}^{n-1} f_i y^i .$$

(For efficiency, it makes sense to take f_n of minimum degree among $f_0, \ldots, f_n$.)

2. *Form the Sylvester matrix Φ of f and f_n. The entries are polynomials in $k[y]$ of degree at most $n-1$.*

3. *Calculate the resultant* $\det \Phi$ *of f and f_n using Berkowitz' [2] or Chistov's [9] algorithm. This is a polynomial in $k[y]$ of degree at most $(n-1) \cdot \deg f_n$.*

4. *Check whether* $\det \Phi$ *vanishes identically. By Theorem 18, this occurs iff $f_0, \ldots, f_n$ have a common root.*

As in §3.1, if the coefficients of $f_0, \ldots, f_n$ are indeterminates, then Algorithm 19 can be carried out symbolically. The entries of the Sylvester matrix of f and f_n are then polynomials in y and the indeterminate coefficients $\overline{a}$ of the f_i. The resultant is a polynomial $r(\overline{a}, y)$ of degree at most $(n-1) \cdot \deg f_n$ in y and $\deg f_n + \max_{i=0}^{n-1} \deg f_i$ in the $\overline{a}$. Considering $r(\overline{a}, y)$ as a polynomial in y with coefficients in $k[\overline{a}]$, Theorem 18 implies that all these coefficients vanish under a specialization $\overline{\alpha}$ of $\overline{a}$ iff the polynomials $f_0, \ldots, f_n$ resulting from the specialization $\overline{\alpha}$ have a common root in k. The coefficients of $r(\overline{a}, y)$ thus form a *resultant system* for $f_0, \ldots, f_n$.

It is possible to work in the polynomial ring $k[x, y, \overline{a}]$ explicitly and compute the symbolic resultant $r(\overline{a}, y)$ in NC using Berkowitz' [2] or Chistov's [9] algorithm. These algorithms produce a polylog-depth, polynomial-size circuit C for $r(\overline{a}, y)$ over $+$, $\cdot$, constants, and inputs $\overline{a}$ and y. For any specialization $\overline{\alpha}$ of $\overline{a}$, since $r(\overline{\alpha}, y)$ is of degree at most $(n-1) \cdot \deg f_n$, we can test in NC whether $r(\overline{\alpha}, y)$ vanishes identically by evaluating it at $(n-1) \cdot \deg f_n + 1$ sample elements of k using the circuit C.

3.5 The GCD of Several Univariate Polynomials

The construction of §3.4 can be extended to give a deterministic NC algorithm for computing the gcd of several polynomials. We will show in §3.6 below how to improve the efficiency in the presence of a source of randomness.

Algorithm 20 (GCD of Several Polynomials)

For polynomials $f_0, \ldots, f_n \in k[x]$:

1. *Let y be a new indeterminate, and form the polynomial*

$$f(x,y) \quad = \quad \sum_{i=0}^{n-1} f_i y^i \ .$$

(For efficiency, take f_n of minimum degree among $f_0, \ldots, f_n$.)

2. *Form the subresultant matrices Φ_d of f and f_n. The entries of Φ_d are polynomials in y of degree at most $n-1$.*

3. *Compute the subresultants over $k(y)[x]$ using Berkowitz' [2] or Chistov's [9] algorithm. The d^{th} subresultant $\det \Phi_d$ is a polynomial in y of degree at most $(n-1) \cdot (\deg f_n - d)$.*

4. *Let d be the smallest number such that $\det \Phi_d$ does not vanish identically. This is the degree of the $\gcd$ of f_n and f.*

5. *Compute the monic $\gcd$ of f_n and f as in Algorithm 14 using polynomial arithmetic on the coefficients. By Lemma 17, this is also the $\gcd$ of $f_0, \ldots, f_n$.*

6. *Reduce the coefficients. As computed in Step 5, they are rational functions of y, i.e. quotients of polynomials in y, but they reduce to elements of k because the monic $\gcd$ is in $k[x]$.*

3.6 Improving Efficiency with Randomness

If we have access to a source of randomness, then we can obtain significantly more efficient algorithms than those of §3.4 and §3.5 by using a randomly chosen element of k in place of the indeterminate y. This is in fact the method of choice in most implementations. This approach is based on the observation that a nonzero polynomial is not likely to vanish when evaluated on a random input chosen from a sufficiently large sample set.

This idea is made concrete in the following lemma, proven independently by Zippel [26] and Schwartz [21]:

Lemma 21 *Let $p(x_1, \ldots, x_n)$ be a nonzero polynomial of total degree d with coefficients in k, and let S be a finite subset of k. If p is evaluated on a random element $(s_1, \ldots, s_n) \in S^n$, then the probability that $p(s_1, \ldots, s_n) = 0$ is at most $\frac{d}{|S|}$.*

This gives rise to the following probabilistic *NC* algorithm:

Algorithm 22 (Resultant of Several Polynomials (Probabilistic Version))

For polynomials $f_0, \ldots, f_n \in k[x]$:

1. *Select a random element β uniformly from a large finite set $S \subseteq k$.*

2. *Form the polynomial*

$$f_0(x) + f_1(x)\beta + f_2(x)\beta^2 + \cdots + f_{n-1}(x)\beta^{n-1} .$$

This is $f(x, \beta)$, where f is the polynomial (9).

3. *Calculate the resultant r of $f(x, \beta)$ and $f_n(x)$. If $f_0, \ldots, f_n$ have a common root, then $r = 0$ with probability 1. If $f_0, \ldots, f_n$ do not have a common root, then by Lemma 21, $r = 0$ with probability at most*

$$\frac{(n-1) \cdot \deg f_n}{|S|} .$$

4. *Reduce the probability of error by repeated trials.*

This algorithm does not provide a way to check the accuracy of the output. This liability is corrected in the gcd algorithm below.

As shown in Lemma 17, if g is the gcd of $f_0, \ldots, f_n$ in $k[x]$, then g is also the gcd of f and f_n in $k(y)[x]$, where f is the polynomial (9). With high probability, g will also be the gcd of f_n and $f(x, \beta)$ for a random β chosen uniformly from a sufficiently large subset of k. This is because g always divides the gcd of f_n and $f(x, \beta)$, and by Theorem 13, g is not itself the gcd of f_n and $f(x, \beta)$ iff the d^{th} subresultant of f_n and $f(x, \beta)$ vanishes, where $d = \deg g$. This is the d^{th} subresultant of f_n and $f(x, y)$ evaluated at β, and again by Theorem 13, the d^{th} subresultant of f_n and $f(x, y)$ does not vanish identically, thus Lemma 21 applies. This gives the following *RNC* algorithm:

Algorithm 23 (GCD of Several Polynomials (Probabilistic Version))

For polynomials $f_0, \ldots, f_n \in k[x]$ with gcd *g of degree d:*

1. *Select a random element β uniformly from a large set $S \subseteq k$.*

2. *Form the polynomial*

$$f_0(x) + f_1(x)\beta + f_2(x)\beta^2 + \cdots + f_{n-1}(x)\beta^{n-1} .$$

This is $f(x, \beta)$, where f is the polynomial (9).

3. *Calculate the* gcd *h of $f(x, \beta)$ and f_n as in Algorithm 14. Then g divides h, and h does not divide g iff the d^{th} subresultant of $f(x, \beta)$ and f_n vanishes. By Lemma 21, this happens with probability at most*

$$p = \frac{(n-1) \cdot (\deg f_n - d)}{|S|} . \tag{10}$$

4. *Check whether h divides g by checking whether h divides f_i, $0 \leq i \leq n-1$, using Algorithm 16. If so, h is the desired* gcd. *If not, go back to Step 1 and repeat with a new random β.*

Using the value p in (10) as a bound on the probability of failure in each trial, and assuming the trials are independent, the expected number of trials before successfully obtaining the gcd of $f_0, \ldots, f_n$ is at most $\frac{1}{1-p}$.

See von zur Gathen [25] for another approach, attributed therein to Steve Cook.

4 Multivariate Resultants

The resultant of univariate polynomials is a classical tool that has played a considerable role in modern algorithms in symbolic algebra and computational geometry. It provides an effectively computable algebraic criterion for deciding when two or more univariate polynomials have a common root. Quite early in this century, effective means were developed for generalizing the resultant to the case of multivariate polynomials. Here, however, there must be a somewhat different approach. Chevalley had noted that there can be no strictly algebraic criterion for the existence of common solutions to a system of multivariate polynomials in the sense that the projection of an algebraic set is not necessarily algebraic.

Because there is no purely algebraic criterion, classical attempts at algorithms for the multivariate case diverge. One direction saw the development of *resolvents* through iterated application of the classical univariate resultant to eliminate variables one at a time. In the work of Hermann, Kronecker, Macaulay and others, an algebraic criterion was found for the special case where the given polynomials are all homogeneous. Here one is dealing with the existence of common zeros in projective space. The solution, which is both elegant and reasonably efficient, parallels and generalizes the basic theory developed in §3 for the univariate case.

In this section we review some of the basic properties of multivariate resultants and their computation. We do not present detailed proofs, but instead state the properties of various algebraic objects and give constructions of such objects which have been found useful in the design of algebraic algorithms. As in §2, $S = k[x_0, \ldots, x_n]$ denotes the graded ring of polynomials in $n+1$ variables. We let S_d denote the $\binom{n+d}{n}$-dimensional k-vector space of homogeneous polynomials of degree d in S.

4.1 Resultant Systems

Classical elimination theory considered both necessary and sufficient conditions for homogeneous polynomials $f_1, \ldots, f_m$ to have no common projective zeros. Recall that the Homogeneous Nullstellensatz asserts that this occurs exactly when, for some d, every monomial $\overline{x}^E$ of degree d can be written

$$\overline{x}^E = g_1 f_1 + g_2 f_2 + \cdots + g_m f_m$$

for some polynomials $g_1, \ldots, g_m \in S$. An effective criterion for deciding whether the f_i's have a common projective zero can be derived by finding a bound on this degree d and on the degrees of the g_j's which must exist when there are no solutions.

An initial solution to this problem is provided by the next theorem. It is Lazard's [14] modern generalization of a classical result of Kronecker, proved using homological methods.

Theorem 24 (Effective Homogeneous Nullstellensatz) *Let $f_1, \ldots, f_m \in S = k[x_0, \ldots, x_n]$ be homogeneous polynomials generating the ideal I. Assume that $m \geq n+1$ and that* $\deg f_1 \geq$

$\deg f_2 \geq \cdots \geq \deg f_m$. *Let* $d = 1 + \sum_{i=1}^{n+1}(\deg f_i - 1)$. *Then* $V(f_1, \ldots, f_m) = \emptyset$ *if and only if* I *contains every monomial of degree* d.

When $m \leq n$ it is known that $V(f_1, \ldots, f_m)$ is never empty, a fact which follows from Krull's Hauptidealsatz and the Projective Dimension Theorem [11, §1.5]. When $m \geq n+1$, this degree bound is tight. For example, the polynomials $x_0^{d_0}, x_1^{d_1}, \ldots, x_n^{d_n}$ have no common projective zeros, although the ideal generated by them does not contain the monomial $x_0^{d_0-1}x_1^{d_0-1}\cdots x_n^{d_n-1}$ of degree $\sum_{i=0}^{n}(d_i - 1) = d - 1$. It does, however, contain every monomial of degree d.

We derive a parallel algorithm to verify this condition by reduction to a problem of linear algebra. Recall that the additive subgroup S_d of S is a vector space over k with basis M_d. We need to show that every monomial $\overline{x}^E \in M_d$ is in the ideal generated by the f_i's. This task is simplified by the next lemma.

Lemma 25 *Let* $I = (f_1, \ldots, f_m)$ *and let* $d_i = \deg f_i$, $1 \leq i \leq m$. *If* $h \in I_d$, *then there are homogeneous polynomials* $g_1, \ldots, g_m$ *with* $g_i \in S_{d-d_i}$ *such that* $h = g_1 f_1 + g_2 f_2 + \cdots + g_m f_m$.

The lemma is proved by noting that if $h \in I_d$, then there are polynomials $g_1', \ldots, g_m' \in k[\overline{x}]$ such that $h = \sum_{i=1}^{m} g_i' f_i$. One then observes that all terms of g_i' which are not of of degree $d - d_i$ give rise to terms that are cancelled in the summation, so if we take g_i to be the part of g_i' that is homogeneous of degree $d - d_i$, then $h = \sum_{i=1}^{m} g_i f_i$ as well.

It is easy to see that the map φ which takes every vector of m homogeneous polynomials $g_1, \ldots, g_m$, $g_i \in S_{d-d_i}$, to the sum $\sum_{i=1}^{m} g_i f_i$,

$$\begin{aligned} \varphi \;:\; \prod_{i=1}^{m} S_{d-d_i} &\;\rightarrow\; S_d \\ :\; (g_1, \ldots, g_m) &\;\mapsto\; \sum_{i=1}^{m} g_i f_i \;, \end{aligned} \tag{11}$$

is a linear map of vector spaces over k. From Lemma 25 it also follows that the image of φ is exactly I_d. We can now define the matrix Φ of the map φ with respect to the bases M_d and M_{d_i}, $1 \leq i \leq m$, analogous to the matrix Φ of §3.1. We index the rows of this matrix by the elements of M_d and the columns by pairs $(i, \overline{x}^B)$ where $1 \leq i \leq m$ and $\overline{x}^B \in M_{d-d_i}$. The column indices comprise a basis for $\prod_{i=1}^{m} S_{d-d_i}$ of size $\sum_{i=1}^{m} n_{d-d_i}$. The entry of Φ in row $\overline{x}^A$ and column $(i, \overline{x}^B)$ is just the coefficient of the term $\overline{x}^A$ in the polynomial $\overline{x}^B f_i(\overline{x})$. When $n = 1$ and $m = 2$, Φ gives the Sylvester matrix (1) of the two univariate polynomials $f_1(1, x_1)$ and $f_2(1, x_1)$ defined in §3.1.

The Effective Homogeneous Nullstellensatz (Theorem 24) now implies that $V(f_1, \ldots, f_m) = \emptyset$ if and only if φ is surjective. This happens exactly when the matrix Φ has full rank $|M_d| = n_d$. So one way to verify that the given system of polynomials has no solution is to find a nonzero $n_d \times n_d$ minor of Φ, which can exist only if Φ has rank n_d.

The multivariate resultant allows us to eliminate variables from some collections of multivariate polynomials. Suppose that $f_1(\overline{x},\overline{y}),\ldots,f_m(\overline{x},\overline{y})$ are polynomials in two sets of variables $\overline{x} = x_0,\ldots,x_n$ and $\overline{y} = y_1,\ldots,y_{n'}$, and in addition that they are homogeneous as polynomials in the variables $\overline{x}$. Construct the matrix $\Phi(\overline{y})$ with respect to $\overline{x}$, so that the entries of $\Phi(\overline{y})$ are polynomials in $\overline{y}$. Then the matrix $\Phi(\overline{y})$ has the following property: for any point $\overline{\gamma} \in \mathbf{A}^{n'}$, the system $f_1(\overline{x},\overline{\gamma}),\ldots,f_m(\overline{x},\overline{\gamma})$ has a solution in the variables $\overline{x}$ if and only if $\Phi(\overline{\gamma})$ has rank strictly less than n_d [23, §19]. Thus

$$\begin{aligned}
&\{\,\overline{\gamma} \in \mathbf{A}^{n'} \mid \exists \xi \in \mathbf{P}^n \; f_1(\overline{\xi},\overline{\gamma}) = \cdots = f_m(\overline{\xi},\overline{\gamma}) = 0\,\} \\
&\quad = \{\,\overline{\gamma} \in \mathbf{A}^{n'} \mid \text{rank } \Phi(\overline{\gamma}) < n^d\,\} \\
&\quad = \{\,\overline{\gamma} \in \mathbf{A}^{n'} \mid \text{all } n_d \times n_d \text{ minors of } \Phi(\overline{\gamma}) \text{ are zero}\,\}.
\end{aligned}$$

As in §3, instead of concentrating on a single set of polynomial equations, we can go one step further and regard the coefficients as parameters. This allows us to compute a general algebraic criterion expressed in terms of the indeterminate coefficients. One can in principle compute a *resultant system* consisting of a set of polynomials in the indeterminate coefficients, then simply evaluate them on the coefficients of any given system and immediately determine the whether or not a solution exists. Unfortunately, the number of polynomials which arise and their degree make the computation unrealistically complex in all but the most trivial cases. Nevertheless, we continue to discuss the computation of resultant systems in these terms for several reasons. First, there are in fact some essential results from classical elimination theory that require this form. In addition, this form allows us emphasize that the algorithms which we present can be expressed solely in terms of the basic operations of the ring of coefficients, and hence commute with substitution. Finally, it permits us to express the complexity of the quantities that arise in this computation in a purely algebraic manner in terms of the complexity of each coefficient.

Let $f_1,\ldots,f_m$ be a set of homogeneous polynomials in the variables $x_0,\ldots,x_n$ with indeterminate coefficients among the variables $\overline{c}$. In other words, we are considering polynomials f_i of the form $f_i(\overline{c},\overline{x}) = \sum_A c_{i,A}\overline{x}^A$ where A ranges over multi-indices of some fixed degree d_i, each $\overline{x}^A$ is a monomial of degree d_i, and each coefficient $c_{i,A}$ is a distinct indeterminate.

Definition 26 *A* resultant system *for $f_1,\ldots,f_m$ is a collection of polynomials $g_1,\ldots,g_r \in k[\overline{c}]$ with the following property: for every specialization $\overline{c} \mapsto \overline{\gamma}$ of the coefficients to elements of k, the polynomials $f_1(\overline{\gamma},\overline{x}),\ldots,f_m(\overline{\gamma},\overline{x})$ have a common solution in $\mathbf{P}^n$ if and only if $g_i(\overline{\gamma}) = 0$, $1 \le i \le r$. In other words*

$$V(g_1,\ldots,g_r) \;=\; \{\,\overline{\gamma} \mid \exists \overline{\xi} \in \mathbf{P}^n \; f_1(\overline{\gamma},\overline{\xi}) = \cdots = f_m(\overline{\gamma},\overline{\xi}) = 0\,\}\ .$$

If $r = 1$, this polynomial is called a resultant *for the system $f_1,\ldots,f_m$.*

Additional details can be found in the texts of Macaulay [15, Ch. 1] and van der Waerden [22, Ch. 7].

Under current technology, the direct calculation of such resultant systems by computing all minors of the corresponding matrix $\Phi(\overline{c})$ is infeasible because of the large number of polynomials and their high degree. To improve this situation somewhat, we will use the parallel algebraic matrix rank algorithm of Mulmuley [17]. We summarize the relevant aspects of the construction in the statement of the next lemma. A more complete treatment can be found in [24].

Lemma 27 (Mulmuley) *Let A be an $m \times n$ matrix over an arbitrary field k, $m \geq n$, and let z, w be indeterminates. It is possible to compute a polynomial $p \in k[w, z]$ of degree at most $(2m-1)n$ in w and $2n$ in z such that the rank of A is $\frac{2n-j}{2}$, where z^j is the highest power of z that divides p. In particular, A is of full rank if and only if $j = 0$, i.e. if and only if $p(w, 0)$ does not vanish identically. Moreover, the computation uses only the ring operations of the field k and can be implemented by an arithmetic circuit of size polynomial in $m + n$ and depth $O(\log^2(m + n))$.*

Since Mulmuley's algorithm uses only the ring operations of k, it can also be performed on matrices containing indeterminate entries. As a consequence, specialization of these indeterminates to elements of k give the same result as first performing the substitution and then applying the algorithm.

Theorem 28 (Solving systems of homogeneous equations) *Let $f_1, \ldots, f_m \in k[\overline{y}][\overline{x}]$ be polynomials that are homogeneous in the variables $\overline{x}$ of degree $d_1, \ldots, d_m$, respectively, with coefficients in $k[\overline{y}]$. A necessary and sufficient algebraic condition for the existence of a common zero $\overline{\xi} \in \mathbf{P}^n$ expressed in terms of the parameters $\overline{y}$ can be computed in parallel polynomial time with respect to the elementary operations of the ring $k[\overline{y}]$. In other words, we can compute a set of polynomials $g_1, \ldots, g_r \in k[\overline{y}]$ such that*

$$g_1(\overline{\gamma}) = \cdots = g_r(\overline{\gamma}) = 0 \quad \Leftrightarrow \quad \exists \overline{\xi} \in \mathbf{P}^n \ f_1(\overline{\xi}, \overline{\gamma}) = \cdots = f_m(\overline{\xi}, \overline{\gamma}) = 0 \ .$$

If each $d_i \leq d$, then the entire computation requires $O(d^{3n})$ processors and time $O((n \log d)^2)$. The number of polynomials in the system, and their degree, are at most exponential in n and $\max_{i=1}^m d_i$.

Proof. Suppose $f_1, \ldots, f_m$ are as described in the statement of the theorem. Let Φ be the matrix of the linear map φ of (11). Then Φ is an $n_d \times (\sum_{i=1}^m n_{d_i})$ matrix over the ring $k[\overline{y}]$. To compute a resultant system, we will apply the parallel matrix rank algorithm of Mulmuley to Φ. From this algorithm, we obtain a polynomial $q(\overline{y}, w) = \sum_{j=0}^e q_j(\overline{y}) w^j$ that for every substitution of field elements for the variables $\overline{y}$ is identically zero as a polynomial in w if and only if Φ does not have full rank n_d. As a polynomial in w, $q(\overline{y}, w)$ is identically zero just when all of its coefficients are zero. The collection of coefficients $\{ q_i(\overline{y}) \mid 0 \leq i \leq e \}$ thus comprises a resultant system for the given set of polynomials. Since the computation involves only the operations of the coefficient ring, the computation also commutes with specialization of coefficients. □

4.2 The Resultant of n Polynomials in n Variables

When the number of homogeneous polynomials equals the number of variables, there is also a single resultant polynomial. The presentation in this section follows along classical lines. For a complete development when $k = \mathbf{C}$ which uses only elementary arguments, see Renegar [20].

Lemma 29 *For $n+1$ homogeneous polynomials $f_0, \ldots, f_n$ in the $n+1$ variables $x_0, \ldots, x_n$ with indeterminate coefficients $\overline{c}$, there exists a single resultant polynomial $r(\overline{c})$, which can be effectively constructed from the matrix Φ.*

Let $\deg f_i = d_i$, $0 \leq i \leq n$, and define $d = 1 + (\sum_{i=0}^{n} d_i - 1)$ and $n_d = \binom{n+d}{n} = |M_d|$. Recall that the matrix Φ has n_d rows, each indexed by a monomial in M_d. To construct the resultant polynomial, we first partition M_d into $n+1$ sets. Say that a monomial $\overline{x}^A \in M_d$ is *reduced in x_i* if $\overline{x}^A$ is not divisible by $x_i^{d_i}$. For each i, $0 \leq i \leq n$, define M_d^i to be the set of monomials in M_d that are divisible by $x_i^{d_i}$ and reduced in the variables $x_0, \ldots, x_{i-1}$. Then the sets $M_d^0, \ldots, M_d^n$ comprise a partition of M_d.

Let A be the square submatrix of Φ obtained by selecting the columns of Φ labeled by $(i, \overline{x}^E)$ for $0 \leq i \leq n$ and $\overline{x}^E \in M_d^i$. It can be shown that, as a polynomial in the indeterminate coefficients $\overline{c}$, $\det A$ is divisible by the resultant of the f_i's. Moreover the quotient of the determinant by the resultant is itself the determinant of a submatrix B of A, obtained by

1. eliminating those columns in A corresponding to indices $(i, \overline{x}^E)$ where the monomial $\overline{x}^E$ is reduced in n of the variables $\overline{x}$; and

2. for each such column $(i, \overline{x}^E)$ removed in Step 1, eliminating the row containing the coefficient of $x_i^{d_i}$ in this column.

This fact is proven by Macaulay [15]. The desired resultant can be computed as a polynomial in the indeterminates $\overline{c}$ by constructing the matrices A and B from Φ and finding the quotient $r(\overline{c}) = \det A / \det B$. When constructed in this manner, the submatrix B depends only on the coefficients of the polynomials $f_1 |_{x_0=0}, \ldots, f_n |_{x_0=0}$, and for each i, $r(\overline{c})$ is homogeneous of degree $\prod_{j \neq i} d_j$ in the coefficients of f_i. Macaulay also shows that r is irreducible as a polynomial in $k[\overline{c}]$ [15, Ch. 1]. See [22, Ch. 7] for further details.

Calculation of the resultant as presented above requires a computation in which all of the coefficients are indeterminates. In most cases, this computation is prohibitively expensive, since there are $\sum_{i=0}^{n} n_{d_i}$ such coefficients. Most often we do not need to know r as a polynomial in the $\overline{c}$'s, but are interested only in the image of r under some substitution for these coefficients. For example, let us assume that we wish to compute the resultant of the homogeneous polynomials $f_0, \ldots, f_n \in k[x_0, \ldots, x_n]$, where all of the coefficients are constants in the field k. Note that all of the above calculations that require only ring operations commute with substitution for the indeterminates [6, Ch. 3], since substitution determines a ring homomorphism. The determinant, for example, is defined and computed using only ring operations; so we can either compute $\det A$ as a polynomial in $k[\overline{c}]$ and then specialize the variables $\overline{c}$ to the coefficients of the f_i's, or specialize the coefficients and then compute the determinant [2]. In both cases the result will be the same.

The only other operation required in the construction is the division of these determinants. Unfortunately this operation does not commute with specialization. For example, where all coefficients are elements of the field k as above, it is possible that $\det A = \det B = 0$ under the given substitution, while the resultant itself is nonzero. To overcome this obstacle, we modify the construction as follows. Observe that the coefficients c_i of the terms $x_i^{d_i}$ in f_i lie only on the diagonals of the matrices A and B, and that these comprise all diagonal entries. Instead of computing the determinants of A and B directly, we compute the *characteristic polynomials* of these matrices, *i.e.* the determinants

$$\begin{aligned} a(\overline{c},t) &= \det(tI - A) \\ b(\overline{c},t) &= \det(tI - B) \end{aligned}$$

where t is a new indeterminate. These determinants are obtained from the determinants of A and B by substituting $t - c_i$ for the coefficient c_i of the term $c_i x_i^{d_i}$ in each f_i and negating all other coefficients. It follows that b divides a, and the quotient $r(\overline{c},t)$ is just the resultant of the new system

$$tx_0^{d_0} - f_0(\overline{c},\overline{x}), \quad \ldots, \quad tx_n^{d_n} - f_n(\overline{c},\overline{x}) \; . \tag{12}$$

Moreover, it is easy to see that neither $a(\overline{\gamma},t)$ nor $b(\overline{\gamma},t)$ vanishes identically for any $\overline{\gamma} \in k^n$. This implies that

$$b(\overline{\gamma},t) \cdot r(\overline{\gamma},t) \;=\; a(\overline{\gamma},t) \; ,$$

so $r(\overline{\gamma},t)$ is the resultant of the system (12) under the specified substitution, and $r(\overline{\gamma},0)$ the resultant of the original system $f_0,\ldots,f_n$. (It is easy to see that negation of the coefficients does not change the resultant.) Hence we can compute the resultant of $n+1$ homogeneous polynomials in $n+1$ variables with coefficients in a field k by computing the constant term of the quotient of the characteristic polynomials of A and B.

As noted in §3.3, computation of the quotient of two univariate polynomials such as $a(\overline{\gamma},t)$ and $b(\overline{\gamma},t)$ can be performed in parallel, using only the ring operations of the coefficient field when the divisor is monic. But we claim further that the quotient of the characteristic polynomials $a(\overline{c},t)$ and $b(\overline{c},t)$ *as polynomials in* t is in fact the resultant $r(\overline{c},t)$ of the polynomials in (12). For suppose this were not the case, and let r' be the quotient of a and b as polynomials in t. It must be that the actual resultant r divides r'. But if $r \neq r'$, then r' must have an additional factor $p(\overline{c})$ such that $r'(\overline{c},t) = r(\overline{c},t)p(\overline{c})$. So

$$a(\overline{c},t) \;=\; b(\overline{c},t)r(\overline{c},t)p(\overline{c}) \; ,$$

and since

$$a(\overline{c}, t) \quad = \quad t^d + \sum_{i=0}^{d-1} a_i(\overline{c}) t^i$$

it is clear that $p = 1$.

Hence $r(\overline{c}, t)$ must be the the resultant of (12), and the constant term $r(\overline{c}, 0)$ of this polynomial is always the resultant of the original system $f_0, \ldots, f_n$. Since the constant term of this quotient can be computed using only ring operations, the entire computation described above now commutes with specialization of the indeterminate coefficients. In other words, we can first substitute elements of any division ring for the coefficients of the f_i's and then perform the indicated computation. The result is guaranteed to be the same as would be obtained by constructing the actual resultant polynomial and then performing the same substitution.

This discussion suggests the following parallel algorithm for computing resultants. Let $f_0, \ldots, f_n$ be polynomials in the variables $x_0, \ldots, x_n$ with coefficients in a ring R'. Construct the matrices A and B and compute their characteristic polynomials $a, b \in R'[t]$, as discussed in [24]. For example, R' may be the polynomial ring $k[\overline{y}]$. Write

$$\begin{aligned} a(t) &= t^{n_d} + a_{n_d-1} t^{n_d-1} + \cdots + a_1 t + a_0 \\ b(t) &= t^e + b_{e-1} t^{e-1} + \cdots + b_1 t + b_0 \end{aligned}$$

where $e = n_d - \prod_{i=0}^{n} d_i$, and compute the quotient a/b using Algorithm 16. Since b is monic, only the ring operations are needed. This gives the desired resultant.

See also [6, §3.1.3] for another method of computing these classical resultants.

Algorithm 30 (Resultant of n polynomials in n variables) *The resultant of a set of $n+1$ homogeneous polynomials $f_0, \ldots, f_n$ in $n+1$ variables $x_0, \ldots, x_n$ with coefficients in a ring $R' = K[\overline{y}]$ can be computed in parallel polynomial time relative to the elementary operations of the ring R'. Moreover, since the computation uses only ring operations on the coefficients, the calculation commutes with substitution.*

The computation of the characteristic polynomials a and b require operations on matrices of size n_d. Using Chistov's algorithm (see [24]), this can be done in parallel time $O(\log^2 n_d)$ in the elementary operations of the coefficient ring R'. The quotient of these two polynomials requires computing the determinant of a matrix of size $n_d - e = \prod_{i=0}^{n} d_i$, and so can be executed in parallel time $O(\log^2 \prod_{i=0}^{n} d_i) = O((\sum_{i=0}^{n} \log d_i)^2)$, again measured in terms of the elementary operations of the ring of coefficients R'. We can conservatively bound the size of elements in R' that arise in this computation by noting that the coefficients of a are polynomials of degree $\prod_{j \neq i} \deg d_i$ in the coefficients of each f_i. For example, assume that $f_0, \ldots, f_n$ are integral polynomials with maximum degree d, and let c be a bound on the number of bits necessary to express any one coefficient. Then the coefficients of a require fewer than $(n+1)d^n c$ bits, and those of the resultant no more than $(n+1)d^{2n}c$ bits. On the other hand, if the coefficients of the f_i's are polynomials in the ring $k[y]$ and have maximum degree d in $\overline{x}$ and e in y, then the the coefficients of a are polynomials of degree no more than $(n+1)d^n e$ in y, and those of the resultant of degree at most $(n+1)d^{2n}e$ in y.

4.3 The u-Resultant

The u-resultant is a classical tool for solving systems of homogeneous equations which have only a finite number of projective solutions. Its use has been revived by computational applications, as in the zero-finding algorithm of Lazard [14], the approximation algorithm of Renegar [18], and the work of Canny [6] in theoretical robotics and of Renegar [19,20] in real algebraic geometry.

Behind the construction of u-resultants lies the following idea. Suppose that we have n homogeneous polynomial equations $f_1, \ldots, f_n$ in the $n+1$ variables $x_0, \ldots, x_n$, with only a finite number of projective solutions $\overline{\xi}^{(1)}, \ldots, \overline{\xi}^{(s)} \in \mathbf{P}^n$. Then for almost every additional polynomial f_0 which we might add to this system, the enlarged system will have no common zeros. This is true even if the degree of f_0 is constrained to be 1.

More concretely, let $u_0, \ldots, u_n$ be new indeterminates and let f be the linear form $\overline{u} \cdot \overline{x} = \sum_{i=0}^{n} u_i x_i$. We show now that for most assignments of values $\overline{v} \in \mathbf{P}^n$ to $\overline{u}$, the system $\overline{v} \cdot \overline{x}, f_1, \ldots, f_n$ has no common zeros. Construct the resultant $r(\overline{u})$ of these $n+1$ polynomials as a polynomial in the variables $\overline{u}$. By the characterization of the resultant given in §4.2, we know that r is homogeneous in $\overline{u}$, and that for any point $\overline{v} \in \mathbf{P}^n$, $r(\overline{v}) = 0$ if and only if $f_1, \ldots, f_n$ and $\overline{v} \cdot \overline{x}$ have a common solution. Equivalently, $r(\overline{v}) = 0$ if and only if $\overline{v} \cdot \overline{\xi}^{(j)} = 0$ for some j. This means that

$$V(r) \;=\; V\left(\prod_{j=1}^{s} \overline{\xi}^{(j)} \cdot \overline{u}\right) .$$

The Homogeneous Nullstellensatz now says that each of r and $\prod_{j=1}^{s} \overline{\xi}^{(j)} \cdot \overline{u}$ divides some power of the other, therefore r factors as a product of linear forms, each of the form $\overline{\xi}^{(j)} \cdot \overline{u}$ for some zero $\xi^{(j)}$ of the f_i's. Hence all of the zeros of the f_i's can be recovered by computing this resultant r and factoring it over $k[\overline{u}]$. The coefficients of these factors are the coordinates of the common zeros.

Definition 31 (The u-Resultant) *Let $f_1, \ldots, f_n \in k[x_0, \ldots, x_n]$ be homogeneous polynomials generating a zero-dimensional ideal, and let $\overline{u} = u_0, \ldots, u_n$ be a set of new indeterminates. Then the resultant of the f_i's and the polynomial $\overline{u} \cdot \overline{x}$ with respect to the variables $\overline{x}$ is called the* u-resultant *of the f_i's.*

The classical theorem on the u-resultant asserts that, when the set $V(f_1, \ldots, f_n)$ is finite, these points and their multiplicities can be recovered from a factorization of the polynomial $r(\overline{u})$.

Lemma 32 *Let $f_1, \ldots, f_n \in k[x_0, \ldots, x_n]$ be homogeneous polynomials, and assume that $V = V(f_1, \ldots, f_n)$ is finite, where each $\overline{\xi} \in V$ has multiplicity $\mu(\overline{\xi})$. Then the u-resultant $r(\overline{u})$ is a homogeneous polynomial of degree $\prod_{i=1}^n \deg f_i$, and*[3]

$$r(\overline{u}) \doteq \prod_{\overline{\xi} \in V} \left(\sum_{i=0}^{n} \xi_i u_i \right)^{\mu(\overline{\xi})} . \tag{13}$$

Note that, since the $\overline{\xi}$'s are points in projective space, this polynomial is unique only up to a nonzero constant factor.

The argument sketched above can be extended to show that $r(\overline{u}) \not\equiv 0$ if and only if $V(f_1, \ldots, f_n)$ is finite. From Lemma 29 it follows that this u-resultant is a quotient of determinants $a(\overline{u})$ and $b(\overline{u})$. By the same lemma, we can construct these polynomials so that b is independent of the coefficients of one of the given polynomials. In particular, we can construct a and b from the f_i's and the polynomial $\overline{u} \cdot \overline{x}$ so that b is independent of the variables $\overline{u}$. Thus when $f_1, \ldots, f_n \in k[\overline{x}]$, b is a constant in k. Whenever $b \neq 0$, $a \not\equiv 0$ and $a(\overline{u}) \doteq r(\overline{u})$, so that a is itself a u-resultant polynomial. In this case, the u-resultant can thus be constructed using a single determinant computation over $k[\overline{u}]$.

4.4 Generalized Characteristic Polynomials

The algorithms developed in §4.2 and §4.3 for computing resultants and u-resultants took advantage of the fact that the quotient of the characterisitic polynomials of the matrices A and B has two important features: it never vanishes identically, and it is the resultant (u-resultant) of a perturbation of the original system of equations by a new indeterminate. This observation was used to produce an efficient algebraic algorithm for computing the resultant of n polynomials in n variables directly from their coefficients, using only the operations of the coefficient ring. This quotient was given the name *generalized characteristic polynomial* by Canny [7] because it specializes in the case of linear equations to the notion of the characteristic polynomial of the matrix of the linear system.

Recently such polynomials have found wide use in extending resultants and u-resultants to the case of affine (inhomogeneous) sets. Let us begin by considering how the algorithms of §4.3 might be adapted to handle inhomogeneous polynomials. Recall the standard embedding of n-dimensional affine space into n-dimensional projective space:

$$(\alpha_1, \ldots, \alpha_n) \mapsto (1 : \alpha_1 : \cdots : \alpha_n) .$$

We can exploit this correspondence in the following manner. Let $f \in R$ be a (possibly inhomogeneous) polynomial in the n variables $x_1, \ldots, x_n$ and let x_0 be a new variable. We define the

[3] We write $f \doteq g$ to signify that f and g are equal up to a nonzero constant factor; *i.e.* there is a $\kappa \in k$, $\kappa \neq 0$ such that $f = \kappa g$.

homogenization of f, written f^h, to be the polynomial

$$f^h(x_0, \ldots, x_n) \;=\; x_0^{\deg f} f(\frac{x_1}{x_0}, \ldots, \frac{x_n}{x_0}) \;,$$

where $\deg f$ is the total degree of f. Operationally, this means that we multiply each term of f by a sufficiently large power of the new variable x_0 to bring it up to degree $\deg f$.

Although homogenization gives an operational way of obtaining a related homogeneous polynomial from an inhomogeneous one, we still have not shown that there is a relation between the roots these polynomials which can be exploited in resultant-based algorithms. On the one hand, the reader can easily verify that $(\alpha_1, \ldots, \alpha_n) \in \mathbf{A}^n$ is a zero of f exactly when $(1 : \alpha_1 : \cdots : \alpha_n)$ is a zero of f^h. However, the polynomial f^h may have zeros which do not correspond to zeros of the original polynomial when we set $x_0 = 0$. Such solutions lie on the hyperplane at infinity and are called *improper* or *infinite.*

Let us consider the case of the u-resultant. We know that a u-resultant polynomial can be computed for n homogeneous polynomials in S whenever they have only finitely many projective zeros. Now suppose we are given n possibly inhomogeneous polynomials $f_1, \ldots, f_n \in R$ and wish to compute their u-resultant. We homogenize them and compute the u-resultant of $f_1^h, \ldots, f_n^h$. This u-resultant should be a constant multiple of the polynomial

$$\left(\prod_{\overline{\xi}} (u_0 + \sum_{i=1}^{n} \xi_i u_i)\right) \cdot \left(\prod_{\overline{\xi}'} \overline{\xi}' \cdot \overline{u}\right) \;,$$

where $\overline{\xi}$ ranges over all common zeros of $f_1, \ldots, f_n$ and $\overline{\xi}'$ ranges over all improper solutions of $f_1^h, \ldots, f_n^h$, and from it we can still recover the common zeros of the original system. But this is only possible when the number of improper solutions is finite. For if there are infinitely many improper projective solutions, the u-resultant polynomial vanishes identically and no information about the proper solutions will be available.

What we would like is a guaranteed method of obtaining all proper solutions to the original system—perhaps together with a finite number of the improper solutions—in the manner that the u-resultant provides for the strictly homogeneous case. This is given by the following lemma, proved by Canny [7] for the case $k = \mathbf{C}$ and by Ierardi [12, Ch. 4] for fields of arbitrary characteristic. We state the lemma only in its simplest form.

Lemma 33 *Let $f_1, \ldots, f_n$ be inhomogeneous polynomials in n variables with only a finite number of common zeros. Then the u-resultant of the system*

$$f_1^h(\overline{x}) - t x_1^{\deg f_1}, \quad \ldots, \quad f_n^h(\overline{x}) - t x_n^{\deg f_n}$$

with respect to the variables $x_0, \ldots, x_n$ is a polynomial $r(\overline{u}, t) \doteq \sum_{i=c}^{e} r_i(\overline{u}) t^i$ in which the least *nonzero coefficient $r_c(\overline{u})$ factors as*

$$\left(\prod_{\overline{\xi}}(x_0 + \sum_{i=1}^{n} \xi_i u_i)\right) \cdot \left(\prod_{\overline{\xi}'} \overline{\xi}' \cdot \overline{u}\right) ,$$

where $\overline{\xi}$ ranges over all points in $V(f_1, \ldots, f_n)$.

4.5 Applications of u-Resultants

Many of the applications of resultants and u-resultants arise from the possibility of recovering the coordinates of points defined as the zero set of a number of multivariate polynomial equations. To a large extent, the efficiency and elegance of these parallel algorithms stems from their ability to recover information about these points *symbolically*, without resorting to numerical approximation.

For example, suppose that a u-resultant polynomial $r(\overline{u})$ has been constructed by one of the methods outlined in §4.4, and we wish to compute the i^{th} coordinate of the affine points represented by this form. One way of obtaining this information is by choosing new indeterminates y and t and substituting into $r(\overline{u})$ the following values for the variables $\overline{u}$:

$$u_j \quad \mapsto \quad \begin{cases} y, & \text{if } j = 0, \\ t^i - 1, & \text{if } j = i, \\ t^j, & \text{otherwise.} \end{cases}$$

A simple calculation shows that if the original polynomial r factored as

$$r(\overline{u}) \quad = \quad \prod_{\overline{\xi}} \overline{\xi} \cdot \overline{u} ,$$

then after the substitution we obtain a polynomial r' which factors as

$$r'(y,t) \quad = \quad \prod_{\overline{\xi}} \left(y + (\sum_{j=0}^{n} \frac{\xi_j}{\xi_0} t^j) - \frac{\xi_i}{\xi_0} \right)$$

and the least nonzero coefficient of r' as a polynomial in t is a polynomial r_i,

$$r_i(y) \quad = \quad \prod_{\xi_0 \neq 0} (y - \frac{\xi_i}{\xi_0}) ,$$

the roots of which are just the i^{th} coordinates of the affine solutions of the original system.

Using subresultant techniques and the notion of primitive elements, this method can be extended to provide a tool which is even more useful. As in the sketch above, the following result relies on the fact that the factors of the polynomial in which we are interested are all linear. Together with the construction of generalized characteristic polynomials presented in §4.4, it permits us to reduce the problem of finding all of the zeros of n polynomial in n variables to a univariate problem, provided that there are only a finite number of solutions altogether. More importantly, it provides a way of representing these solutions symbolically.

Algorithm 34 (Canny, Renegar) *Let $f_1, \ldots, f_n$ be polynomials with only a finite number of common zeros. We can compute a polynomial $q(t)$ and rational functions $r_1(t), \ldots, r_n(t)$ such that the points $(r_1(\theta), \ldots, r_n(\theta))$ include all of these common zeros as θ ranges over the roots of q.*

The complete construction and its proof are presented in [19] and [8], where applications to problems in real geometry are also discussed. The constructions have already proven useful in the following context: when we have found a finite set of points defined by a set of multivariate equations, this algorithm allows us to reduce the multivariate problem to a univariate problem involving just those points.

4.6 Extensions to the Affine Case

In the affine case—the case of inhomogeneous polynomials—there cannot be a purely algebraic criterion for the existence of a common solution to sets of polynomial equations. However, recent results of Kollàr [13] and Galligo, Heintz and Morgenstern [5] do yield a parallel polynomial time algebraic algorithm for deciding this question. The algorithm again depends on obtaining good degree bounds for the Nullstellensatz.

According to the Nullstellensatz, whenever $(f_1, \ldots, f_m) \subseteq R$,

$$\begin{aligned} V(f_1, \ldots, f_m) = \emptyset \quad &\Leftrightarrow \quad 1 \in (f_1, \ldots, f_m) \\ &\Leftrightarrow \quad \exists g_1, \ldots, g_m \in R \quad \sum_{i=1}^{m} g_i f_i = 1 \; . \end{aligned} \tag{14}$$

The discussion of the previous sections suggest that if we can find a degree bound for the polynomials g_i in (14), then we can reduce the problem of determining whether 1 is an element of this ideal to the problem of determining the rank of an appropriate matrix formed from the coefficients of the given polynomials. By reducing the problem to the problem of determining the rank of a matrix, one can then apply Mulmuley's algorithm to give a parallel polynomial-time solution to the problems of deciding the solvability of a set of polynomial equations, and even the problem of quantifier elimination in algebraically closed fields. The essential facts are stated in the next theorem.

Theorem 35 (Kollàr) *Let $f_1, \ldots, f_m \in R$ be polynomials of degree at most d. If $V(f_1, \ldots, f_m) = \emptyset$, then there exist polynomials $g_1, \ldots, g_m$ satisfying (14) with $\deg f_i + \deg g_i \leq d^n$ for $1 \leq i \leq m$.*

This bound now leads to the following algorithm for solving systems of equations. For each $e \geq 0$, let R_e denote the additive subgroup of R consisting of all polynomials of degree at most e. Note that R_e is a k-vector space of dimension n_e, and as a basis we can take all monic monomials of degree at most e. Let Ψ be the matrix of the linear map

$$\begin{aligned} \psi \quad : \quad \prod_{i=1}^{m} R_{d^n - \deg f_i} \quad & \rightarrow \quad R_{d^n} \\ : \quad (g_1, \ldots, g_m) \quad & \mapsto \quad \sum_{i=1}^{m} g_i f_i \end{aligned}$$

with repect to this basis, and let $\hat{1}$ denote the vector corresponding to the multiplicative identity $1 \in R_{d^n}$. Now we know that the polynomials f_i have a solution if and only if $\hat{1}$ is in the linear span of the columns of Ψ. If it is, then the rank of Ψ is the same as the rank of the matrix $(\Psi, \hat{1})$, obtained by adding $\hat{1}$ as a new column of Ψ. If not, then the ranks differ. Invoking Mulmuley's *NC* algorithm for determining the rank of a matrix, we thus obtain an efficient parallel algorithm for this problem.

5 Exercises

1. Let D be a square matrix with entries in a graded ring $S = \bigoplus_{d=0}^{\infty} S_d$. Suppose that every entry of D is nonzero and homogeneous, and every 2×2 minor of D is homogeneous. Prove that $\det D$ is homogeneous.

2. Let A be a set of distinct indeterminates, $|A| = d$, and let x be another indeterminate. Show that the coefficient of x^{d-i} in the polynomial $\prod_{a \in A}(x - a)$ is a homogeneous polynomial in $k[A]$ of degree i. (This coefficient is called the i^{th} *elementary symmetric polynomial* in A.)

3. Let f, g be multivariate polynomials such that

   - all irreducible factors of f are distinct (*i.e.*, f is squarefree);
   - f and g are homogeneous of the same degree; and
   - $V(f) = V(g)$.

   Show that $f \doteq g$.

4. Let A be the multiset of roots of a monic univariate polynomial f; thus $f = \prod_{\alpha \in A}(x - \alpha)$. Similarly, let B be the multiset of roots of a monic univariate polynomial g. Prove that the resultant of f and g is

$$\prod_{\substack{\alpha \in A \\ \beta \in B}} (\beta - \alpha) \, .$$

5. Show that the multivariate resultant of two homogeneous polynomials in two variables is essentially the same as the univariate resultant of two polynomials. Explain why this is so.

6. Let f be a univariate polynomial with rational coefficients. The resultant of f and its formal derivative f' is called the *discriminant* of f.

   (a) Calculate the discriminant of the quadratic $ax^2 + bx + c$.

   (b) Prove that f has a multiple root if and only if its discriminant vanishes.

7. Let s_m and t_m be as in Definition 10, and let p and q be arbitrary polynomials. Prove that for $0 \leq m \leq n$,

$$\gcd(p, q) \;=\; \gcd(s_m p + t_m q, s_{m+1} p + t_{m+1} q) \, .$$

8. Let $f_0, f_1, \ldots, f_n, f_{n+1}$ be the PRS of f_0 and f_1, and let s_m and t_m be as in Definition 10. Prove that for $0 \leq m \leq n$,

$$f_m \;=\; \gcd(f_0 + (-1)^m t_m f_{m+1}, f_1 - (-1)^m s_m f_{m+1}) \, .$$

References

[1] Michael Ben-Or, Dexter Kozen, and John Reif. The complexity of elementary algebra and geometry. *JCSS*, 32(2):251–264, 1986.

[2] Stuart J. Berkowitz. On computing the determinant in small parallel time using a small number of processors. *Information Processing Letters*, 18:147–150, 1984.

[3] Allan Borodin, Joachim von zur Gathen, and John Hopcroft. Fast parallel matrix and gcd computations. *Information and Control*, 52(3):241–256, 1982.

[4] W. Brown and J. F. Traub. On Euclid's algorithm and the theory of subresultants. *J. ACM*, 18:505–514, 1971.

[5] Leandro Caniglia, André Galligo, and Joos Heintz. Some new effectivity bounds in computational geometry. Preprint.

[6] John Canny. *The Complexity of Robot Motion Planning*. PhD thesis, MIT, 1987.

[7] John Canny. Generalized characteristic polynomials. Preprint, 1988.

[8] John Canny. Some algebraic and geometric problems in PSPACE. In *Proc. 20th ACM Symp. Theory of Computing*, pages 460–467. Assoc. Comput. Mach., May 1988.

[9] A. L. Chistov. Fast parallel calculation of the rank of matrices over a field of arbitrary characteristic. In *Proc. Conf. Foundations of Computation Theory, Lecture Notes in Computer Science 199*, pages 63–69. Springer-Verlag, 1985.

[10] L. Csanky. Fast parallel matrix inversion algorithms. *SIAM J. Comput.*, 5(4):618–623, 1976.

[11] Robert Hartshorne. *Algebraic Geometry*. Springer-Verlag, 1977.

[12] Doug Ierardi. *The Complexity of Quantifier Elimination in the Theory of an Algebraically Closed Field*. PhD thesis, Cornell University, 1989.

[13] J. Kollàr. Sharp effective Nullstellensatz. Preprint.

[14] Daniel Lazard. Résolution des systems d'èquations algèbriques. *Theor. Comput. Sci.*, 15:77–110, 1981.

[15] F. S. Macaulay. *The Algebraic Theory of Modular Systems*. Cambridge Tracts in Mathematics and Mathematical Physics. Cambridge U. Press, 1916.

[16] H. Matsumura. *Commutative Ring Theory*. Cambridge Studies in Advanced Mathematics 8. Cambridge U. Press, 1988.

[17] Ketan Mulmuley. A fast parallel algorithm to compute the rank of a matrix over an arbitrary field. *Combinatorica*, 7(1):101–104, 1987.

[18] James Renegar. On the worst case arithmetic complexity of approximating zeros of systems of polynomials. Technical Report 748, School of Operations Research and Industrial Engineering, Cornell U., Ithaca, New York, 1987.

[19] James Renegar. A faster PSPACE algorithm for deciding the existential theory of the reals. Technical Report 792, School of Operations Research and Industrial Engineering, Cornell U., 1988.

[20] James Renegar. On the computational complexity and geometry of the first-order theory of the reals, Parts I, II, III. Technical report, School of Operations Research and Industrial Engineering, Cornell U., Ithaca, New York, 1989.

[21] J. T. Schwartz. Fast probabilistic algorithms for verification of polynomial identities. *J. ACM*, 27(4):701–717, October 1980.

[22] B. L. van der Waerden. *Modern Algebra*, volume 2. F. Ungar Publishing Co., third edition, 1950.

[23] B. L. van der Waerden. *Modern Algebra*, volume 2. F. Ungar Publishing Co., fifth edition, 1970.

[24] Joachim von zur Gathen. Parallel linear algebra. *This volume.*

[25] Joachim von zur Gathen. Parallel algorithms for algebraic problems. *SIAM J. Comput.*, 13(4):802–824, 1984.

[26] Richard E. Zippel. Probabilistic algorithms for sparse polynomials. In Ng, editor, *Proc. EUROSAM 79, Lecture Notes in Computer Science 72*, pages 216–226. Springer-Verlag, 1979.