Symbolic and Numerical Computation for Artificial Intelligence

edited by

Bruce Randall Donald
Department of Computer Science
Cornell University, USA

Deepak Kapur
Department of Computer Science
State University of New York, USA

Joseph L. Mundy
AI Laboratory
GE Corporate R&D, Schenectady, USA

Academic Press
Harcourt Brace Jovanovich, Publishers
London San Diego New York
Boston Sydney Tokyo Toronto

ACADEMIC PRESS LIMITED
24–28 Oval Road
London NW1

US edition published by
ACADEMIC PRESS INC.
San Diego, CA 92101

This book is printed on acid-free paper

A catalogue record for this book is available from the British Library

ISBN 0-12-220535-9

Printed and Bound in Great Britain by
The University Press, Cambridge

Chapter 2

Elimination Methods: an Introduction

Deepak Kapur [†]

Institute for Programming and Logics, Department of Computer Science

State University of New York, Albany, NY 12222

kapur@cs.albany.edu

Yagiti N. Lakshman [‡]

Department of Computer and Information Sciences

University of Delaware, Newark, DE 19716

lakshman@cis.udel.edu

This is an introductory paper on elimination methods applicable to a system of nonlinear polynomials equations. We discuss three different approaches for solving a system of nonlinear polynomial equations. The first approach of resultants is based on the theory of determinants, and was developed in the late 19th century and the early 20th century. The main idea is to generate from a system of nonlinear polynomial equations, a possibly larger system of independent polynomial equations such that there are as many equations as terms in the polynomials so that each term can be used as an unknown and the theory of linear system of equations can be applied. We describe the resultants of Sylvester, Bezout, Dixon's extension of Bezout's resultant, and Macaulay's resultant, followed by some recent extensions of Macaulay's resultant.

The second approach is based on polynomial ideal theory and generates special bases of polynomial ideals, called Gröbner bases. An algorithm for computing such bases was given by Buchberger and is described here. We review some recent developments in the Gröbner basis theory, by the use of which, nonlinear polynomial equations with finitely many solutions can be efficiently solved.

The third approach is based on Ritt's characteristic set construction, motivated by the analysis and decomposition of the zero sets of systems of polynomial equations. This approach has been recently popularized by Wu Wen-tsun who has impressively demonstrated its application to automated geometry theorem proving. Using Wu's method, it is possible to automatically prove, in a matter of seconds, nontrivial theorems in plane Euclidean geometry which human experts find difficult to prove.

[†] Supported in part by a grant from United States Air Force Office of Scientific Research AFOSR-91-0361.

[‡] Supported by the Center for Advanced Technology in Automation and Robotics at Rensselaer Polytechnic Institute under a grant (Grant No. 90-39) from the New York State Science and Technology Foundation.

1. Introduction

In high school, we learn how to manipulate and solve a system of linear equations. In particular, we learn methods for determining whether a system of linear equations has a solution or not. With a little more effort, it is also possible to determine whether a system of equations has a single solution or infinitely many solutions. In the latter case, it is possible to study the structure of the solution space by classifying the variables into independent and dependent subsets and specifying the solutions in terms of independent variables.

A related problem of determining whether an equation c (or a system of equations) follows from a system $\mathcal{S}$ of equations can also be answered easily depending upon the number of solutions of $\mathcal{S}$. If $\mathcal{S}$ has no solution, i.e. $\mathcal{S}$ is inconsistent, then it is possible to take either of the two views – any equation follows from an inconsistent system of equations that has no solution, or alternatively, it is meaningless to ask whether an equation follows from an inconsistent set of equations. If $\mathcal{S}$ has a unique solution, then if the solution of $\mathcal{S}$ is also a solution of c, then c follows from $\mathcal{S}$. If $\mathcal{S}$ has infinitely many solutions, which can be represented by expressing each dependent variable in terms of the independent variables, the expression for each of the dependent variables can be substituted into the given equation c to check whether the equation holds. This is equivalent to checking whether the solution space of the given equation c includes the solution space of the system $\mathcal{S}$ of equations.

The above problems can also be studied in a slightly different setting using their formulation in terms of properties of matrices and determinants. That is what one often learns in a first course on linear algebra at a college level. What has been missing, of late, is a similar discussion for studying these problems for nonlinear polynomial equations. That is the case even though elegant generalizations were developed for solving these problems for nonlinear polynomial equations in the late 19th century and early 20th century. A number of books on the theory of equations were written which are now out of print. For an excellent discussion of the history of constructive methods in algebra, the reader is referred to a thought-provoking article by Professor Abhyankar (Abhyankar, 1976).

In this paper, we discuss these problems for nonlinear equations and discuss some of the approaches proposed in the late 19th century and early 20th century. We also discuss two additional constructive approaches – the approach proposed by Ritt and recently revived by Wu based on algebraic geometry, and another approach by Buchberger based on polynomial ideal theory.

We study two additional subproblems that turn out to be useful in their own right as well as intermediate steps while discussing the problems mentioned earlier. The first problem has to do with the equivalence of two systems of equations. The second problem is that of *projection* or *elimination*, which is to project the solution space of a system $\mathcal{S}$ of equations along certain dimensions and compute another system $\mathcal{S}'$ of equations in an appropriate subset of the variables, such that the solution space of $\mathcal{S}'$ coincides with the projection of the solution space of $\mathcal{S}$ on the chosen variables.

To summarize, here is a list of problems being considered. Given a system $\mathcal{S}$ of polynomial equations in n variables,

- does $\mathcal{S}$ have a solution?
- if $\mathcal{S}$ has a solution, what is the structure of the solution space? What is a good characterization of the structure of the solutions of $\mathcal{S}$?
- eliminate m variables from $\mathcal{S}$ to obtain a system of equations $\mathcal{S}'$ in the remaining $n-m$ variables such that the solutions of $\mathcal{S}'$ are the projections of the solutions of $\mathcal{S}$ to the coordinates corresponding to these $n-m$ variables.
- Given two systems of polynomial equations $\mathcal{S}$ and $\mathcal{S}'$, do they have the same solutions, or, is the set of solutions of $\mathcal{S}'$ properly contained in the set of solutions of $\mathcal{S}$?

1.1. Overview

We begin with the basic definitions of zero sets of polynomial systems and ideals. This is followed by an exposition on resultants. We discuss Dixon's formulation of the resultant of two polynomials in one variable as well as three polynomials in two variables, followed by a presentation of Macaulay's construction of the multivariate multi-polynomial resultant. The concept of the u-resultant for determining the common zeros of polynomials is discussed.

Section 3 is about Gröbner bases. The concepts of an ordering on power products and viewing polynomial equations as simplification rules are discussed. A Gröbner basis of an ideal is a basis with many useful properties. Among other things, Gröbner bases can be used to solve systems of equations. After an overview of the basic properties of Gröbner bases, we pay special attention to systems of polynomials that have finitely many common solutions. For computing the common zeros of such systems, we describe a recent algorithm due to Faugère *et al.* (1989).

Section 4 discusses Ritt's characteristic sets. Our presentation is based on Wu's treatment of characteristic sets and an algorithm for computing a characteristic set. A breadth-first algorithm, which is Ritt's original algorithm, as well as a depth-first algorithm for computing characteristic sets are described.

An expanded version of this article includes illustrations of the applications of different elimination techniques to curve and surface implicitization, detection of unfaithful parameterizations and geometry theorem-proving. The expanded version will be available as a technical report from the Department of Computer and Information Science of the University of Delaware or from Institute for Programming and Logics, the Department of Computer Science, State University of New York at Albany.

1.2. Preliminaries

Let $\mathcal{Q}$ denote the field of rational numbers and $\mathcal{C}$ denote the field of complex numbers. Unless specified otherwise, by a polynomial, we mean a multivariate polynomial with rational or integer coefficients. A univariate polynomial $p(x)$ is an element of the polynomial ring $\mathcal{Q}[x]$.

A multivariate can be viewed in one of two ways: as an element of the ring $\mathcal{Q}[x_1, \ldots, x_n]$ or as an element of the ring $\mathcal{Q}[x_1, \ldots, x_{n-1}][x_n]$. Under the first view, the polynomial $p(x_1, \ldots, x_n)$ is seen as a sum of products with nonzero coefficients, where each product

${x_1}^{d_1}{x_2}^{d_2}\cdots{x_n}^{d_n}$, is called a *term* or *power product*; together with its coefficient, it is called a *monomial*; the degree of a term ${x_1}^{d_1}{x_2}^{d_2}\cdots{x_n}^{d_n}$ is $d_1+d_2+\cdots+d_n$.

Under the second view, the polynomial $p(x_1,\ldots,x_n)$ is seen as a univariate polynomial in x_n (also known as the *main* or *leading* variable) with coefficients that are themselves multivariate polynomials in the remaining variables. For example, the polynomial

$$q(x_1,x_2,x_3)=2x_1^2x_2x_3^2+x_2^2x_3^2-2x_1+x_2^2+x_3$$

can be viewed as a sum of monomials appearing in it, or can be viewed as the polynomial

$$\widehat{q}(x_3)=(2x_1^2x_2+x_2^2)x_3^2+x_3+(-2x_1+x_2^2)$$

in the variable x_3 (q could be also considered as a univariate polynomial with x_1 as the main variable or x_2 as the main variable).

The degree of a univariate polynomial $p(x)$ is the maximum degree, say d, of x in $p(x)$; the leading term of $p(x)$ is then x^d, and the leading coefficient (also called the *initial*) of $p(x)$ is the coefficient of x^d in $p(x)$. For a multivariate polynomial, the leading term and the leading coefficient can be determined only after an ordering on terms is chosen. If a multivariate polynomial is considered as a polynomial in one of its variables, say x_3 in the case of $\widehat{q}$ above, then its degree is the maximum degree in that variable. For $\widehat{q}$, the degree of the polynomial is 2. Its leading term is x_3^2 and the leading coefficient is the polynomial $2x_1^2x_2+x_2^2$.

A univariate polynomial $p(x)$ is said to vanish at $x=a$ if the polynomial p evaluates to zero when a is uniformly substituted for x in p. The value of p at a is denoted by $p(a)$; if $p(a)=0$, then a is called a *zero* of $p(x)$. Equivalently, the polynomial equation $p(x)=0$ is said to have a solution a. The domain from which the values are picked to be checked for zeros of p is very important. It is possible for p to have a zero in one domain and not in another domain. For instance, x^2+1 does not have a zero in the reals, but it does have a zero in the complex numbers. A univariate polynomial of degree n with complex coefficients has exactly n complex roots. They may or may not be distinct.

The above definitions extend to multivariate polynomials also. Given $p(x_1,\ldots,x_n)$, an n-tuple $(a_1,a_2,\ldots,a_n)\in\mathcal{C}^n$, the affine n-space over complex numbers, is a zero of p if p evaluates to zero when, for each $1\le i\le n$, a_i is uniformly substituted for x_i, i.e. $p(a_1,a_2,\ldots,a_n)=0$. Equivalently, $(a_1,a_2,\ldots,a_n)$ is called a solution of the multivariate polynomial equation $p(x_1,\ldots,x_n)=0$ if $p(a_1,\ldots,a_n)=0$.

Given a system $\{f_1(x_1,x_2,\ldots,x_n),\ f_2(x_1,x_2,\ldots,x_n),\ \ldots,\ f_r(x_1,x_2,\ldots,x_n)\}$ of polynomials (equivalently a system $\{f_i(x_1,x_2,\ldots,x_n)=0,\ i=1,\ldots,r\ \}$ of polynomial equations), $(a_1,a_2,\ldots,a_n)$ is a common zero (respectively, a common solution) of the system, if, for each $1\le i\le r$, $f_i(a_1,a_2,\ldots,\ldots a_n)=0$, i.e. $(a_1,a_2,\ldots,a_n)$ is a zero of every polynomial in the system. The n-tuple $(a_1,a_2,\ldots,a_n)$ is also called a common root of these polynomials. Henceforth, we will abuse the terminology; by a system, we will mean either a set of polynomials or a set of polynomial equations, with the hope that the context can resolve the ambiguity.

Given a system $\mathcal{S}$ of polynomials, the set

$$\mathsf{Zero}(\mathcal{S})=\{(a_1,a_2,\ldots,a_n)\in\mathcal{C}^n\mid\forall f\in\mathcal{S}, f(a_1,a_2,\ldots,a_n)=0\}$$

is called the *zero set* of the system $\mathcal{S}$. The zero set defined by a system of polynomials is also called an *algebraic* set. Throughout this paper, we have focused on zeros in the field of complex numbers (in general, an algebraically closed field). For details about zero sets of polynomials over the reals (these are called semi-algebraic sets), the reader may consult Tarski (1948), Arnon *et al.* (1984), Coste and Roy (1988) and Renegar (1989).

Example: Let $f(x,y) = x^3 - y^2$, $g(x,y) = 2x^2 + (y-1)^2 - 2$, $\mathcal{S} = \{f = 0, g = 0\}$,

$$\begin{aligned}
\mathsf{Zero}(\mathcal{S}) \;=\; & \{(1,1),\\
& (-2.253012931 + 1.322062452i, 3.016885573 + 2.953686458i),\\
& (-2.253012931 - 1.322062452i, 3.016885573 - 2.953686458i),\\
& (-.4604205396 + .3623712296i, -.3774386961 - .2422512995i),\\
& (-.4604205396 - .3623712296i, -.3774386961 + .2422512995i),\\
& (.4268669419, -.2788937516)\}
\end{aligned}$$

(the values shown here are floating point approximations). Geometrically, $\mathsf{Zero}(\mathcal{S})$ consists of the coordinates of all the points of intersection of the cusp $f(x,y) = 0$ and the ellipse $g(x,y) = 0$.

The zero set of a system of polynomials could be infinite. In that case, it is possible, as in linear algebra, to talk about the dimension of the zero set. If the zero set is finite, the system is called *zero-dimensional.* If the zero set is empty, then the system is said to be of dimension -1. If the zero set is infinite, then the system has positive dimension.

A zero a of a univariate polynomial $p(x)$ is said to be of multiplicity $k+1$ if $p^{(i)}(a)$ for $i = 0, \ldots, k$ all evaluate to zero, where $p^{(i)}$ is the i-th derivative of p with respect to x. This is equivalent to saying that $(x-a)^{k+1}$ divides $p(x)$.

1.2.1. Affine Zeros and Projective Zeros

The familiar complex n-space is known as the affine n-space over the complex numbers. It consists of all n-tuples of the complex numbers. In cartesian coordinates, each n-tuple $(c_1, c_2, \ldots, c_n) \in \mathcal{C}^n$ represents a point in the affine n-space. By $\mathsf{Zero}(\mathcal{S})$, we mean the set of affine zeros of $\mathcal{S}$. We now introduce the notion of projective space and projective zeros.

The projective n-space over the complex numbers, denoted by $\mathcal{P}^n$, consists of all $(n+1)$-tuples over the complex numbers except $(0, 0, \ldots, 0)$. Each $(n+1)$-tuple in $\mathcal{P}^n$ is said to represent a point in the projective n-space. However, the tuples $(a_0, a_1, \ldots, a_n)$ and $(\lambda a_0, \lambda a_1, \ldots, \lambda a_n)$ are said to represent the same point for any non-zero complex number λ and the two $(n+1)$-tuples are said to be equivalent. Thus, each point in $\mathcal{P}^n$ is represented by any member of an infinite set of proportionate (equivalent) n-tuples. We can embed the affine n-space into the projective n-space as follows. To each point $(c_1, c_2, \ldots, c_n) \in \mathcal{C}^n$, we associate the $(n+1)$-tuple $(1, c_1, c_2, \ldots, c_n)$. As mentioned earlier, any other $(n+1)$-tuple of the form $(\lambda, \lambda c_1, \ldots, \lambda c_n)$ represents the same point as long as $\lambda \neq 0$. Any $(n+1)$-tuple $B = (b_0, b_1, b_2, \ldots, b_n)$ in which $b_0 \neq 0$ is associated with the unique point in affine n-space whose coordinates are

$$\bar{B} = (b_1/b_0, b_2/b_0, \ldots, b_n/b_0).$$

Note that any other $(n+1)$-tuple equivalent to B gives rise to the same point $\bar{B}$ in the affine n-space. Those $(n+1)$-tuples which have $b_0 = 0$ are said to represent *points at infinity.* The set of all points at infinity in $\mathcal{P}^n$ is known as the *hyper plane at infinity.* To summarize what we have said so far, projective n-space consists of the affine n-space together with the hyper-plane at infinity.

A polynomial $f(x_0, x_1, \ldots, x_n)$ is said to be homogeneous of degree d if each *term* in f has degree d.

Example: $f(x_0, x_1, x_2) = x_0^2 x_1 - 2x_1 x_2^2 + x_1^3$ is a homogeneous polynomial of degree 3. $g(x_1, x_2) = x_1^3 + x_2 - x_1 x_2$ is *not* homogeneous.

If $(a_0, a_1, \ldots, a_n)$ is a zero of the homogeneous polynomial $f(x_0, x_1, \ldots, x_n)$, i.e.

$$f(a_0, a_1, \ldots, a_n) = 0,$$

then any other $(n+1)$-tuple $(\lambda a_0, \lambda a_1, \ldots, \lambda a_n)$ is also a zero of $f(x_0, x_1, \ldots, x_n)$.

Given a non-homogeneous polynomial $f(x_1, x_2, \ldots, x_n)$, of degree d, it can be *homogenized* as follows. Consider

$$^h f(x_0, x_1, x_2, \ldots, x_n) = x_0^d f(x_1/x_0, x_2/x_0, \ldots x_n/x_0)$$

where x_0 is a new variable. $^h f(x_0, x_1, \ldots, x_n)$ is a homogeneous polynomial of degree d such that

$$^h f(1, x_1, x_2, \ldots, x_n) = f(x_1, x_2, \ldots x_n).$$

Let $(a_1, a_2, \ldots, a_n) \in \mathcal{C}^n$ be a zero of f, i.e. $f(a_1, a_2, \ldots a_n) = 0$. Then $(1, a_1, a_2, \ldots, a_n) \in \mathcal{P}^n$ (or any $(n+1)$-tuple equivalent to it) is a zero of $^h f$. Conversely, if $(a_0, a_1, a_2, \ldots, a_n) \in \mathcal{P}^n$ is a zero of $^h f$, and $a_0 \neq 0$, then $(a_1/a_0, a_2/a_0, \ldots, a_n/a_0) \in \mathcal{C}^n$ is a zero of f. If $a_0 = 0$, then there is no corresponding point in the affine space that is a zero of f. Such zeros of $^h f$ are called *zeros at infinity* of f.

Most of the resultant calculations work over projective space with homogeneous polynomials.

1.2.2. Ideals

Consider a commutative ring $\mathcal{R}$. Let $\mathcal{A} \subseteq \mathcal{R}$. $\mathcal{A}$ is called an *ideal* in $\mathcal{R}$ iff:

- for all $f, g \in \mathcal{A}$, $f + g \in \mathcal{A}$, and,
- for all $f \in \mathcal{A}$, $gf \in \mathcal{A}$ for any $g \in \mathcal{R}$.

Let $f_1, f_2, \ldots, f_r \in \mathcal{R}$. Consider an ideal $\mathcal{J}$ that contains all of $f_1, \ldots, f_r$. By the above definition, the element

$$f = g_1 f_1 + g_2 f_2 + \ldots + g_r f_r \in \mathcal{J}$$

for *any* $g_1, g_2, \ldots, g_r \in \mathcal{R}$. Indeed, the set

$$\mathcal{I} = \{\sum_{i=1}^{r} g_i f_i \mid g_i \in \mathcal{R}\}$$

is an ideal in $\mathcal{R}$ and it is the *smallest* ideal in $\mathcal{R}$ containing the set $\{f_1, f_2, \ldots, f_r\}$. $\mathcal{I}$ is called the *ideal generated by* $f_1, f_2, \ldots, f_r$ and denoted by $(f_1, f_2, \ldots, f_r)$. The set $\{f_1, f_2, \ldots, f_r\}$ is called a *generating set* or a *basis* for the ideal $\mathcal{I}$.

Examples:

1. $\mathcal{R}$ = ring of integers, $\mathcal{A}$ = set of all multiples of 3.
2. $\mathcal{R}$ = ring of integers, and $\mathcal{A}$ = the set of integers of the form $9a + 30b$, where a, b are integers.
3. $\mathcal{R} = \mathcal{Q}[x, y]$, and $\mathcal{A}$ is the set of all polynomials $f(x, y) \in \mathcal{Q}[x, y]$ such that $f(a, b) = 0$ for some fixed constants $a, b \in \mathcal{Q}$.
4. $\mathcal{R} = \mathcal{Q}[x_1, x_2, \ldots, x_n]$ and $\mathcal{A} = (x_1 - a_1, x_2 - a_2, \ldots, x_n - a_n)$ where $a_1, a_2, \ldots, a_n \in \mathcal{Q}$.

An ideal can have many bases. For example, the ideal in the first example has $\{3\}$ as its basis whereas the ideal in the second example has $\{9, 30\}$ as it basis, but the two ideals are the same.

Let

$$\mathcal{I} = (f_1, f_2, \ldots, f_r) \subseteq \mathcal{Q}[x_1, x_2, \ldots, x_n].$$

Let $(a_1, a_2, \ldots, a_n) \in \mathcal{C}^n$ be a *common zero* of $f_1, f_2, \ldots, f_r$, i.e.

$$f_i(a_1, a_2, \ldots, a_n) = 0, \quad i = 1, \ldots, r.$$

Since, for any $f \in \mathcal{I}$, there exist $g_i \in \mathcal{Q}[x_1, x_2, \ldots, x_n]$ such that $f = \sum_1^r g_i f_i$, it follows that $f(a_1, a_2, \ldots, a_n) = 0$, i.e. $(a_1, a_2, \ldots, a_n)$ is a *zero of every polynomial* in the ideal. The set

$$\mathsf{Zero}(\mathcal{I}) = \{(a_1, a_2, \ldots, a_n) \in \mathcal{C}^n \mid \forall f \in \mathcal{I}, f(a_1, a_2, \ldots, a_n) = 0\}$$

is called the *zero set* of the ideal $\mathcal{I}$.

Earlier we considered a polynomial equation defining a cusp, $f(x, y) = x^3 - y^2$, and another polynomial equation defining an ellipse, $g(x, y) = 2x^2 + (y-1)^2 - 2$. Let $\mathcal{I} = (f, g)$. The location of the points of intersection of the two curves is not evident from the equations $f(x, y) = 0, g(x, y) = 0$. It will be shown later that the set

$$\begin{aligned} G \;=\; & \{g_1(x) = x^6 + 4x^5 + 4x^4 - 6x^3 - 4x^2 + 1, \\ & g_2(x, y) = y + 1/2(-x^3 - 2x^2 + 1)\} \end{aligned}$$

is another basis for the ideal $\mathcal{I}$. Notice that G has one polynomial that depends only on x, namely $g_1(x)$, and one that depends on both x and y, i.e. $g_2(x, y)$. The roots of $g_1(x)$ are the x-coordinates of the points of intersection of the cusp $f(x, y) = 0$ and the ellipse $g(x, y) = 0$. For each root α of $g_1(x)$, the y-coordinates of the corresponding intersection points are found by computing the roots of $g_2(\alpha, y)$. As in linear algebra, we say that the above two polynomials in G are in *triangular form*.

2. Resultants

Given two polynomials $f(x), g(x) \in \mathcal{Q}[x]$ of degrees m and n respectively, i.e.

$$\begin{aligned} f(x) &= f_n x^n + f_{n-1} x^{n-1} + \ldots f_1 x + f_0, \quad \text{and,} \\ g(x) &= g_m x^m + g_{m-1} x^{m-1} + \ldots + g_1 x + g_0, \end{aligned}$$

when do f and g have common roots? The question leads naturally to a condition that has to be satisfied by the coefficients of f and g. This condition was discovered by Euler and is now commonly referred to as the *vanishing of the Sylvester resultant of f and g.* The Sylvester resultant of f, g is the determinant of the following matrix:

$$R = \begin{pmatrix} f_0 & 0 & 0 & \dots & 0 & g_0 & 0 & 0 & \dots & 0 \\ f_1 & f_0 & 0 & \dots & 0 & g_1 & g_0 & 0 & \dots & 0 \\ f_2 & f_1 & f_0 & \dots & 0 & g_2 & g_1 & g_0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ f_n & f_{n-1} & f_{n-2} & \cdots & f_{n-m+1} & g_n & g_{n-1} & g_{n-2} & \cdots & g_0 \\ 0 & f_n & f_{n-1} & \cdots & f_{n-m+2} & g_{n+1} & g_n & g_{n-1} & \cdots & g_1 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & f_n & 0 & 0 & 0 & \dots & g_m \end{pmatrix}$$

Assuming that at least one of f_n, g_m is non-zero, *the vanishing of the Sylvester resultant is a necessary and sufficient condition for f and g to have common roots.*

Resultants are most commonly used for computing projections and for successive elimination of variables. The Sylvester resultant has been studied extensively in the past. We refer the interested reader to the beautiful subresultant theory developed simultaneously and independently by G.E. Collins and W.S. Brown. For an exposition of the theory, see Knuth (1980, pp. 407–408), Loos (1983), Collins (1967, 1971) and Brown and Traub (1971). Efficient implementations of algorithms for computing resultants are available in most computer algebra systems including REDUCE, MACSYMA, MAPLE and MATHEMATICA.

2.1. Dixon's Formulation

In 1779, Bezout had already developed a method for computing the resultant of two univariate polynomials. We describe Cayley's reformulation of Bezout's method. It is simple to explain and extends naturally to the bivariate case as shown by Dixon. Cayley proposed viewing the resultant of $f(x)$ and $g(x)$ as follows. Replace x by α in both $f(x)$ and $g(x)$ and we get polynomials $f(\alpha)$ and $g(\alpha)$. The determinant $\Delta(x, \alpha)$ of the matrix

$$\begin{vmatrix} f(x) & g(x) \\ f(\alpha) & g(\alpha) \end{vmatrix}$$

is a polynomial in x and α and it obviously is equal to zero if $x = \alpha$. This implies that the determinant has $(x - \alpha)$ as a factor. The polynomial

$$\delta(x, \alpha) = \frac{\Delta(x, \alpha)}{(x - \alpha)}$$

is an $n - 1$ degree polynomial in α and is symmetric in x and α. It vanishes at every common zero x_0 of $f(x)$ and $g(x)$ no matter what values α has. So, at $x = x_0$, the coefficient of every power product of α in $\delta(x, \alpha)$ is 0. This gives n equations which are polynomials in x, and the maximum degree of these polynomials is $n - 1$. Any common zero of $f(x)$ and $g(x)$ is a solution of these polynomial equations, and these polynomial equations have a common solution if the determinant of their coefficients is

0. Unlike in Sylvester's formulation, where the resultant of f and g is the determinant of an $(m+n) \times (m+n)$ matrix, in the Cayley-Dixon formulation, the resultant is obtained as the determinant of a $n \times n$ matrix.

For example, consider a generic cubic polynomial:

$$p = ax^3 + bx^2 + cx + d.$$

The discriminant of p is the resultant of p and dp/dx. The polynomial p has multiple roots if and only if its discriminant is zero. Let us compute the discriminant of p by Cayley's method. We have

$$dp/dx = 3ax^2 + 2bx + c.$$

The determinant of the matrix:

$$\begin{vmatrix} ax^3 + bx^2 + cx + d & 3ax^2 + 2bx + c \\ a\alpha^3 + b\alpha^2 + c\alpha + d & 3a\alpha^2 + 2b\alpha + c \end{vmatrix}$$

when divided by $x - \alpha$ gives the polynomial:

$$(3a^2x^2+2abx+ac)\alpha^2+(2abx^2+(2b^2-2ac)x+(bc-3ad))\alpha+(acx^2+(bc-3ad)x+(c^2-2bd).$$

We get three equations by equating the coefficients of the power products of α above to 0:

$$\begin{array}{rlllll} 3a^2 & x^2 + 2ab & x + ac & = & 0, \\ 2ab & x^2 + (2b^2 - 2ac) & x + (bc - 3ad) & = & 0, \\ ac & x^2 + (bc - 3ad) & x + (c^2 - 2bd) & = & 0. \end{array}$$

Treating x^0, x^1, x^2 as unknowns, we have three homogeneous equations in three unknowns; they have a common solution if and only if the determinant of the coefficient matrix is 0, i.e.

$$\begin{vmatrix} 3a^2 & 2ab & ac \\ 2ab & (2b^2 - 2ac) & (bc - 3ad) \\ ac & (bc - 3ad) & (c^2 - 2bd) \end{vmatrix} = -a^2(-c^2b^2 + 4ac^3 + 4b^3d - 18abdc + 27a^2d^2) = 0.$$

The reader may want to compare this determinant with the Sylvester resultant given by the determinant

$$\begin{vmatrix} a & b & c & d & 0 \\ 0 & a & b & c & d \\ 3a & 2b & c & 0 & 0 \\ 0 & 3a & 2b & c & 0 \\ 0 & 0 & 3a & 2b & c \end{vmatrix} = -a(-c^2b^2 + 4ac^3 + 4b^3d - 18abdc + 27a^2d^2).$$

The Dixon resultant has an extraneous factor of a as compared to the Sylvester resultant. This factor arises because Cayley's formulation assumes that both the polynomials are of the same degree. In general, the Bezout resultant computed using the Cayley-Dixon formulation will have an extraneous factor $I_f{}^{(degree(f)-degree(g))}$, where I_f is the initial of $f(x)$.

Dixon (1908) showed how to extend this formulation to three polynomials in two variables. Consider the following three generic bi-degree polynomials which have all the power products of the type $x^i y^j$, where $0 \leq i \leq m, 0 \leq j \leq n$, i.e.

$$f(x,y) = \sum_{i,j} a_{ij} x^i y^j, \quad g(x,y) = \textstyle\sum_{i,j} b_{ij} x^i y^j, \quad h(x,y) = \sum_{i,j} c_{ij} x^i y^j.$$

Just as in the earlier case, Dixon observed that the determinant

$$\Delta(x,y,\alpha,\beta) = \begin{vmatrix} f(x,y) & g(x,y) & h(x,y) \\ f(\alpha,y) & g(\alpha,y) & h(\alpha,y) \\ f(\alpha,\beta) & g(\alpha,\beta) & h(\alpha,\beta) \end{vmatrix}$$

vanishes when α or β are substituted for x or y, respectively, implying that $(x-\alpha)(y-\beta)$ is a factor of the above determinant. The expression

$$\delta(x,y,\alpha,\beta) = \frac{\Delta(x,y,\alpha,\beta)}{(x-\alpha)(y-\beta)}$$

is a polynomial of degree $2m-1$ in α, $n-1$ in β, $m-1$ in x and $2n-1$ in y. Since the above determinant vanishes when we substitute $x = x_0, y = y_0$ where (x_0, y_0) is a common zero of $f(x,y), g(x,y), h(x,y)$, into the above matrix, $\delta(x_0, y_0, \alpha, \beta)$ must vanish no matter what α and β are. The coefficients of each power product $\alpha^i \beta^j, 0 \leq i \leq 2m-1, 0 \leq j \leq n-1$, have common zeros which include the common zeros of $f(x,y), g(x,y), h(x,y)$. This gives $2mn$ equations in power products of x, y, and the number of power products $x^i y^j, 0 \leq i \leq m-1, 0 \leq j \leq 2n-1$ is also $2mn$. The determinant of the coefficient matrix from these equations is a multiple of the resultant. Using a simple geometric argument, Dixon proved that in this case, the determinant is in fact the resultant up to a constant factor. For three arbitrary polynomials, Dixon developed some special methods which selected some coefficients of $\alpha^i \beta^j$ from δ, and used dialytic expansion of $f(x,y), g(x,y), h(x,y)$ to come up with a system of k linearly independent polynomial equations expressed using k power products in x, y.

As an example, consider the following two bi-quadratics and a linear form.

$$\begin{aligned} f_1 &= a_1 x_1^2 x_2^2 + a_2 x_1^2; \\ f_2 &= b_1 x_1^2 x_2^2 + b_2 x_2^2; \\ f_3 &= u_1 x_1 + u_2 x_2 + u_3 \end{aligned}$$

The polynomial $\delta(x_1, x_2, \alpha, \beta)$ can be given as:

$$\begin{pmatrix} 1 & \alpha & \beta & \alpha^2 & \alpha^2\beta & \alpha\beta & \alpha^3 & \alpha^3\beta \end{pmatrix} \quad \mathbf{D} \quad \begin{pmatrix} 1 & x_2^3 & x_1 & x_2 & x_2^2 & x_1 x_2^3 & x_1 x_2^2 & x_1 x_2 \end{pmatrix}^{\mathbf{Tr}}$$

where $\mathbf{D}$ is the 8 by 8 matrix :

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & a_1u_3b_2 & 0 & a_2u_3b_2 \\ 0 & a_1u_3b_2 & 0 & a_2u_3b_2 & 0 & a_1u_1b_2 & 0 & a_2u_1b_2 \\ 0 & 0 & a_2u_3b_2 & 0 & 0 & a_1u_2b_2 & a_1u_3b_2 & a_2u_2b_2 \\ 0 & a_1u_1b_2 & 0 & 0 & 0 & 0 & a_2u_2b_1 & a_2u_3b_1 \\ 0 & 0 & a_2u_3b_1 & 0 & a_1u_1b_2 & 0 & 0 & a_2u_2b_1 \\ a_2u_3b_2 & a_1u_2b_2 & a_2u_1b_2 & a_2u_2b_2 & a_1u_3b_2 & 0 & a_1u_1b_2 & 0 \\ 0 & 0 & 0 & a_2u_3b_1 & a_2u_2b_1 & 0 & 0 & a_2u_1b_1 \\ a_2u_3b_1 & 0 & a_2u_1b_1 & a_2u_2b_1 & 0 & 0 & 0 & 0 \end{pmatrix}$$

The determinant of $\mathbf{D}$,

$$a_1^2\, a_2^4\, b_1^2\, b_2^4\, u_3^4(a_1^2\, b_1^2\, u_3^4+2a_1^2\, b_1\, b_2\, u_1^2\, u_3^2+a_1^2\, b_2^2\, u_1^4+2a_1a_2b_1^2\, u_2^2\, u_3^2-2a_1a_2b_1b_2u_1^2\, u_2^2+a_2^2\, b_1^2\, u_2^4),$$

is the resultant up to a constant factor.

Chionh's thesis (Chionh, 1990) gives a good summary of Dixon's approach to resultants. He also discusses how Dixon's resultants for the two variable case can be used for implicitization and finding base points.

2.2. Multivariate Resultants

Macaulay constructed a resultant (henceforth referred to as the Macaulay resultant) for n homogeneous polynomials in n variables (Macaulay, 1916). It simultaneously generalizes the Sylvester resultant and the determinant of a system of linear equations (in the sense that the Macaulay resultant for two homogeneous polynomials in two variables is the same as their Sylvester resultant, and the Macaulay resultant for a system of n homogeneous linear equations in n variables is the same as the determinant of the system). Macaulay's resultant disappeared from the literature for several decades until it was used in a slightly different form by Lazard (1981) for equation-solving. More recently, Canny (1988) resurrected the Macaulay resultant and used it in his roadmap algorithm for the robot motion-planning problem.

The Macaulay resultant can be used to eliminate several variables at once from a system of polynomial equations. Macaulay used his resultant construction to actually determine the solutions of a system of homogeneous polynomial equations. If one wishes to solve non-homogeneous polynomial equations using the Macaulay resultant, one has to homogenize the polynomials first. Methods based on Macaulay's matrix give out zeros in $\mathcal{P}^n$ for the homogenized system of equations and they can include zeros at infinity, which will have to be dealt with if one is interested only in affine common zeros.

2.3. Macaulay's Matrices

In this section, we describe Macaulay's construction. The key idea is to show which power products are sufficient in the dialytic method to be used as multipliers for the polynomials, so that we get a square system of l linear equations in l power products which can be considered as the unknowns.

Let $f_1, f_2, \ldots, f_n$ be n *homogeneous* polynomials in $x_1, x_2, \ldots, x_n$. Let $d_i = \deg(f_i)$ and

$$d_M = 1 + \sum_1^n (d_i - 1).$$

Let T denote the set of all terms of degree d_M in the n variables $x_1, x_2, \ldots, x_n$, i.e.

$$T = \{x_1^{\alpha_1} x_2^{\alpha_2} \ldots x_n^{\alpha_n} | \alpha_1 + \alpha_2 + \ldots + \alpha_n = d_M\}$$

and

$$|T| = \binom{d_M + n - 1}{n - 1}.$$

The polynomials are multiplied with appropriate power products to generate $|T|$ equations in $|T|$ unknowns which are power products of degree d_M. The order in which the polynomials are considered for selecting multipliers results in different systems of linear equations.

Example: We will use a generic system of one linear and two quadratic polynomials in three variables to illustrate various aspects of Macaulay's construction. Let

$$\begin{aligned}
f_1 &= a_{1,1}x_1^2 + a_{1,2}x_1x_2 + a_{1,3}x_1x_3 + a_{2,2}x_2^2 + a_{2,3}x_2x_3 + a_{3,3}x_3^2,\ d_1 = 2,\\
f_2 &= b_{1,1}x_1^2 + b_{1,2}x_1x_2 + b_{1,3}x_1x_3 + b_{2,2}x_2^2 + b_{2,3}x_2x_3 + b_{3,3}x_3^2,\ d_2 = 2,\\
f_3 &= c_1x_1 + c_2x_2 + c_3x_3,\ d_3 = 1 \text{ and } d_M = 3.
\end{aligned}$$

The number of terms in three variables of degree 3 is 10. To determine the power products to be used as multipliers to obtain a 10 by 10 system, we will use the ordering (f_1, f_2, f_3) to illustrate the construction. The first three rows are obtained by multiplying f_1 by the power products of degree 1, the difference of d_M and the degree of f_1; these power products are: x_1, x_2, x_3. The next three rows are obtained also by multiplying f_2 by the power products of degree 1 that are not multiples of x_1^2. In this case, they are x_1, x_2, x_3. The last four rows are obtained by multiplying f_3 by the power products of degree 2, the difference of d_M and the degree of f_3, that are not multiples of either x_1^2 or x_2^2. These power products are: $x_1x_2, x_1x_3, x_2x_3, x_3^2$.

Macaulay's matrix in this case is:

$$\begin{array}{c|cccccccccc}
 & x_1^3 & x_1^2x_2 & x_1^2x_3 & x_1x_2^2 & x_1x_2x_3 & x_1x_3^2 & x_2^3 & x_2^2x_3 & x_2x_3^2 & x_3^3 \\
\hline
x_1 & a_{1,1} & a_{1,2} & a_{1,3} & a_{2,2} & a_{2,3} & a_{3,3} & 0 & 0 & 0 & 0 \\
x_2 & 0 & a_{1,1} & 0 & a_{1,2} & a_{1,3} & 0 & a_{2,2} & a_{2,3} & a_{3,3} & 0 \\
x_3 & 0 & 0 & a_{1,1} & 0 & a_{1,2} & a_{1,3} & 0 & a_{2,2} & a_{2,3} & a_{3,3} \\
x_1 & b_{1,1} & b_{1,2} & b_{1,3} & b_{2,2} & b_{2,3} & b_{3,3} & 0 & 0 & 0 & 0 \\
x_2 & 0 & b_{1,1} & 0 & b_{1,2} & b_{1,3} & 0 & b_{2,2} & b_{2,3} & b_{3,3} & 0 \\
x_3 & 0 & 0 & b_{1,1} & 0 & b_{1,2} & b_{1,3} & 0 & b_{2,2} & b_{2,3} & b_{3,3} \\
x_1x_2 & 0 & c_1 & 0 & c_2 & c_3 & 0 & 0 & 0 & 0 & 0 \\
x_1x_3 & 0 & 0 & c_1 & 0 & c_2 & c_3 & 0 & 0 & 0 & 0 \\
x_2x_3 & 0 & 0 & 0 & 0 & c_1 & 0 & 0 & c_2 & c_3 & 0 \\
x_3^2 & 0 & 0 & 0 & 0 & 0 & c_1 & 0 & 0 & c_2 & c_3
\end{array}$$

The general construction is given below. Let

$$\begin{aligned}
T^{(0)} &= \{\text{terms of degree } d_M - d_1\},\\
T^{(1)} &= \{\text{terms of degree } d_M - d_2 \text{ and not divisible by } x_1^{d_1}\},\\
T^{(2)} &= \{\text{terms of degree } d_M - d_2 \text{ and not divisible by } x_1^{d_1}, \text{ or by } x_2^{d_2}\},\\
&\ \ \vdots\\
T^{(n-1)} &= \{\text{terms of degree } d_M - d_n \text{ and not divisible by}\\
&\qquad x_1^{d_1} \text{ or } x_2^{d_2} \text{ or } \ldots \text{ or } x_{n-1}^{d_{n-1}}\}.
\end{aligned}$$

Macaulay refers to $T^{(i)}$ as the *reduced* set of terms with respect to $x_1, \ldots, x_i$ (so a term is reduced with respect to $x_1, \ldots, x_i$ if it is not divisible by any of $x_1^{d_1}$, $x_2^{d_2}$, …, $x_i^{d_i}$). Now, construct a matrix A with $|T|$ columns and $\sum_i |T^{(i)}|$ rows. The columns of A are labeled by the terms in T in some order. The first $|T^{(0)}|$ rows are labeled by the terms in $T^{(0)}$, the next $|T^{(1)}|$ rows are labeled by the terms in $T^{(1)}$ and so on. In the row labeled by the term $t \in T^{(i)}$, arrange the coefficients of tf_{i+1}, with the coefficient of a term t' in tf_{i+1} appearing under the column labeled by t' (note that tf_{i+1} has degree d_M).

We reproduce below Macaulay's argument showing that the matrix A thus constructed is square. Let a_i denote the coefficient of $x_i^{d_i}$ in f_i for $i = 1, 2, \ldots, n$ (note that this coefficient could possibly be 0). Since each row contains the coefficients of some f_i alone (shifted appropriately), every row contains exactly one a_i.

Every column contains at least one a_i. Suppose not. Let the label of a column not containing a_i be t. So, there is no term t' and no i such that $t'x_i^{d_i} = t$ which implies that t is not divisible by $x_i^{d_i}$ for any i. This implies that $\deg(t) \leq \sum_i (d_i - 1) < d_M$, which is a contradiction since $\deg(t) = d_M$.

Each column contains at most one a_i. The proof is again by contradiction. Suppose a column labeled by t has a_i, a_j with $i < j$; this means that there are terms $t_1 \in T^{(i-1)}$ and $t_2 \in T^{(j-1)}$ such that $t_1 x_i^{d_i} = t_2 x_j^{d_j} = t$. This implies that $x_i^{d_i}$ divides t_2; but $t_2 \in T^{(j-1)}$, which means that it is not divisible by $x_1^{d_1}$ or $x_2^{d_2}$, so on up to $x_{j-1}^{d_{j-1}}$; in particular, t_2 is *not divisible by* $x_i^{d_i}$ since $i < j$, which is a contradiction.

We have thus shown that each row and each column has exactly one a_i and this establishes a one-to-one correspondence between rows and columns, which implies that

$$|T| = \sum_i |T^{(i)}|$$

and A is a square matrix.

Let $\det(A)$ denote the determinant of A, which is a polynomial in the coefficients of the f_i's. It is homogeneous in the coefficients of each f_i and the degree of $\det(A)$ in the coefficients of f_n is $d_1 d_2 \ldots d_{n-1}$ (this is the number of rows of A labeled by the terms in $T^{(n)}$; these rows contain the coefficients of f_n).

The construction of A depends on how the polynomials are ordered. A different order produces a different matrix. The ideal generated by the determinants of all such matrices is the so-called *ideal of inertial forms,* a concept that goes back to Hurwitz (van der Waerden, 1950) and this ideal is known to be principle. Let $\det(A_\sigma)$ denote the determinant of a matrix constructed using a permutation σ of the polynomials $f_1, f_2, \ldots, f_n$. The greatest common divisor of the $\det(A_\sigma)$'s ($\sigma \in S_n$, the symmetric group on n letters) regarded as polynomials in the indeterminate coefficients of the f_i is defined to be the resultant (denoted R) of the system $\{f_1, f_2, \ldots, f_n\}$.

For the above example, the resultant is the greatest common divisor of the determinants of all such matrices ($3! = 6$ of them in this case).

We list some important properties of the resultant below.

1. $R = 0$ if and only if the f_i's have a non-trivial common zero.
2. R is absolutely irreducible and invariant under linear coordinate transformations. The

vanishing of R is thus a necessary condition for the system to have a common zero and it is the smallest such condition.

3. R is homogeneous in the coefficients of each f_i and has degree $(\prod_{j=1}^{n} d_j)/d_i$ in coefficients of f_i. For instance, in the above example, the resultant is a polynomial in which each term has degree 2 in $a_{j,k}$, degree 2 in $b_{j,k}$ and degree 4 in c_j.
4. If $f_n = gh$, then the resultant of $f_1, f_2, \ldots, f_n$ is a product of two resultants R_1 (of $f_1, f_2, \ldots, g$) and R_2 (of $f_1, f_2, \ldots, h$).

Macaulay also constructed the following formula relating R and $\det(A_\sigma)$:

$$R \det(B_\sigma) = \det(A_\sigma)$$

where $\det(B_\sigma)$ is the determinant of a submatrix of A_σ. The submatrix B_σ is obtained by deleting all columns labeled by terms *reduced* (in Macaulay's sense) in *any* $n-1$ *of the variables*, and, those rows which contain one of the a_is in the deleted columns.

Example: In the earlier example, the submatrix B for the Macaulay matrix is

$$\begin{array}{c} \\ x_3 \\ x_3 \end{array} \begin{array}{c} \begin{array}{cc} x_1^2x_3 & x_2^2x_3 \end{array} \\ \begin{pmatrix} a_{1,1} & a_{2,2} \\ b_{1,1} & b_{2,2} \end{pmatrix} \end{array}.$$

The matrix B was obtained by deleting columns that are *reduced* (in Macaulay's sense) in any two variables (e.g. x_1^3 is not divisible by x_2^2 or by x_3, so it is *reduced* in x_2, x_3; hence, the column labeled x_1^3 was deleted, a similar reason for deleting the other columns). The surviving columns are reduced in fewer than $n-1$ variables; for example, $x_1^2x_3$ is divisible by x_1^2, x_3 but not by x_2^2. Hence, it is reduced in x_2 only. The rows that contained an a_i (in this example, this means one of $a_{1,1}, b_{2,2}, c_3$) in the deleted columns are also deleted. For example, the first row was deleted because it contained $a_{1,1}$ in a deleted column, namely, the column labeled x_1^3.

The crucial point here is that Macaulay's formula works in "general", or when the coefficients are taken to be indeterminates. If one wants to compute R for a specific system of polynomials using this formula, one may encounter the problem of having $\det(B_\sigma) = 0$. A similar problem can arise if we try to compute R as the gcd of $\det(A_\sigma)$'s due to the vanishing of some or all of the $\det(B_\sigma)$'s. The computation of the resultant as a *generic polynomial* in the indeterminate coefficients of f_i is infeasible, as it is very large, even for low degree input polynomials.

2.4. The U-Resultant

Suppose we have a system of n polynomials in n variables, i.e.

$$f_1(x_1, x_2, \ldots, x_n), \ldots, f_n(x_1, x_2, \ldots, x_n),$$

and we wish to find their common zeros. Let d_i denote the total degree of f_i. Introduce a new homogenizing variable x_0 and let R_u denote the Macaulay resultant of the $(n+1)$ homogeneous polynomials ${}^hf_1, {}^hf_2, \ldots, {}^hf_n, f_u$ in $(n+1)$ variables $x_0, x_1, \ldots, x_n$ where f_u is the linear form

$$f_u = x_0u_0 + x_1u_1 + \ldots + x_nu_n$$

and $u_0, u_1, \ldots, u_n$ are *new unknowns*. R_u is a polynomial in $u_0, u_1, \ldots, u_n$. It is homogeneous in the u_i of degree $B = \prod_{i=1}^{n} d_i$ (these observations follow from the properties of the Macaulay resultant listed earlier). R_u is known as the u-resultant of the given system of polynomials. It can be shown that R_u factors into *linear factors* over the complex numbers, i.e.

$$R_u = \prod_{j=1}^{B}(u_0\alpha_{0,j} + u_1\alpha_{1,j} + \ldots + u_n\alpha_{n,j})$$

and if $(u_0\alpha_{0,j} + u_1\alpha_{1,j} + \ldots + u_n\alpha_{n,j})$ is a factor of R_u, then $(\alpha_{0,j}, \alpha_{1,j}, \ldots, \alpha_{n,j})$ is a common zero of of ${}^hf_1, {}^hf_2, \ldots, {}^hf_n$. The converse can also be proved, i.e. if $(\beta_{0,j}, \beta_{1,j}, \ldots, \beta_{n,j})$ is a common zero of ${}^hf_1, {}^hf_2, \ldots, {}^hf_n$ then $(u_0\beta_{0,j} + u_1\beta_{1,j} + \ldots + u_n\beta_{n,j})$ divides R_u. This gives an algorithm for finding all the common zeros of ${}^hf_1, {}^hf_2, \ldots, {}^hf_n$.

Example: Consider the unit circle $x_1^2 + x_2^2 - 1 = 0$ and the pair of straight lines $(x_1 - x_2 - 1)\ (x_1 - x_2 + 1)$. To find their intersection, we compute their u-resultant which is the Macaulay resultant of

$$\begin{aligned} f_1 &= x_1^2 + x_2^2 - x_0^2 \\ f_2 &= x_1^2 - 2x_1x_2 + x_2^2 - x_0^2 \\ f_u &= u_0x_0 + u_1x_1 + u_2x_2. \end{aligned}$$

The u-resultant is computed to be the polynomial

$$u_1^2u_2^2 - u_1^2u_0^2 - u_2^2u_0^2 + u_0^4$$

which factors as

$$(0.u_1 - 1.u_2 + 1.u_0)(0.u_1 + 1.u_2 + 1.u_0)(-1.u_1 + 0.u_2 + 1.u_0)(1.u_1 + 0.u_2 + 1.u_0).$$

We can read off the four intersection points from the linear factors as

$$(0, -1), (0, 1), (-1, 0), (1, 0).$$

Constructing the full u-resultant, however, is an almost impossible task (it is a polynomial of degree $\prod_{i=1}^{n} d_i$ in n variables). So, if one is interested in computing the common zeros of a set of polynomials, one does so by computing specializations of R_u. For example, the univariate polynomial $R_1(u_0)$ obtained by substituting $u_i = 0$ for $i = 2, \ldots, n$ and $u_1 = -1$ has as its roots the x_1 coordinates of the common zeros. $R_1(u_0)$ can be computed from the Macaulay matrices by evaluation of determinants with rational (or complex) entries and interpolation and without constructing the full u-resultant. For more details, see Canny (1988), Lakshman (1990a), Lakshman and Lazard (1991) and Manocha and Canny (1991).

This method does not always work, however. Since for each common zero $(\beta_{1,j}, \ldots, \beta_{n,j})$ of the f_i's, the linear form $(u_0 + u_1\beta_{1,j} + \ldots + u_n\beta_{n,j})$ divides the u-resultant R_u, if the given system of polynomials $f_1, f_2, \ldots, f_n$ has infinitely many common zeros, then the u-resultant R_u of the system is identically zero and one cannot compute the common zeros by this method. Therefore, we assume that the given system of polynomials $f_1, f_2, \ldots, f_n$ has only finitely many common zeros. However, even this is not sufficient since the u-resultant vanishes whenever there are infinitely many common zeros of the *homogeneous polynomials* ${}^hf_1, {}^hf_2, \ldots, {}^hf_n$. It may happen that $f_1, f_2, \ldots, f_n$ have only finitely many common zeros but ${}^hf_1, {}^hf_2, \ldots, {}^hf_n$ have infinitely many common zeros – all but a finite number of them at infinity (when this happens, the zero set is said to have an excess component at infinity).

Example: The u-resultant of

$$f_1 = x_1^2 - x_2^2 + 2x_2 + 1, \text{ and } f_2 = x_1^2 - 2x_1x_2 + x_2^2 + x_1 + 2$$

is zero because they have a common component at infinity given by $x_1 = x_2$.

Often, one has a system of non-homogeneous polynomials that are known to have only finitely many common zeros but the corresponding homogeneous system may have excess components at infinity. The u-resultant algorithm cannot be used as it is in this situation. Grigoriev and Chistov (1983) and Canny (1990) suggest a modification of the algorithm that will give all the affine zeros of the original system (as long as they are finite in number) even in the presence of excess components at infinity. We now briefly describe their approach.

Let $f_1, f_2, \ldots, f_n$ be as before. Let

$$g_i = {}^hf_i + \lambda x_i^{d_i}, \quad \text{for } i = 1, \ldots, n,$$

and

$$g_u = (u_0 + \lambda)x_0 + u_1x_1 + \ldots + u_nx_n$$

where λ is a new unknown. Let $R_u(\lambda, u_0, \ldots, u_n)$ be the Macaulay resultant of $g_1, g_2, \ldots, g_n$ and g_u, regarded as homogeneous polynomials in $x_0, x_1, \ldots, x_n$. $R_u(\lambda, u_0, \ldots, u_n)$ is called the *generalized characteristic polynomial* of $f_1, \ldots, f_n$. Now, look at $R_u(\lambda, u_0, \ldots, u_n)$ as a polynomial in λ whose coefficients are polynomials in $u_0, u_1, \ldots, u_n$, i.e.

$$R_u(\lambda, u_0, \ldots, u_n) = \lambda^\delta + R_{\delta-1}\lambda^{\delta-1} + \ldots + R_k\lambda^k$$

where $k \geq 0$ and the R_i are polynomials in $u_0, u_1, \ldots, u_n$ (if $k = 0$, R_k will be the same as the u-resultant R_u; however, if there are excess components at infinity, then $k > 0$). The trailing coefficient R_k shares a very useful property with the u-resultant, namely, it can be shown that R_k factors into *linear factors* over the complex numbers, i.e.

$$R_k = \prod_{j=1}^{B}(u_0\alpha_{0,j} + u_1\alpha_{1,j} + \ldots + u_n\alpha_{n,j})$$

and if $(u_0\alpha_{0,j} + u_1\alpha_{1,j} + \ldots + u_n\alpha_{n,j})$ is a factor of R_k, then $(\alpha_{0,j}, \alpha_{1,j}, \ldots, \alpha_{n,j})$ is a common zero of of ${}^hf_1, {}^hf_2, \ldots, {}^hf_n$. The converse can also be proved, i.e. if $(\beta_{1,j}, \ldots, \beta_{n,j})$ is an *affine common zero* of $f_1, f_2, \ldots, f_n$ then $(u_0 + u_1\beta_{1,j} + \ldots + u_n\beta_{n,j})$ divides R_k. This gives us a way to recover all the affine common zeros of $f_1, f_2, \ldots, f_n$ even in the presence of excess components at infinity. Again, in practice, one never constructs the complete generalized characteristic polynomial of a system of polynomials. As with the u-resultant, one can recover the affine common zeros by computing specializations of the generalized characteristic polynomial.

2.5. Implementations of Multivariate Resultants

Multivariate resultant algorithms (both Dixon's method for the bivariate case as well as Macaulay's method) can be easily implemented on computer algebra systems since the main calculation needed is that of determinant of a matrix. Sederberg (1983) presents an implementation of Dixon's method. Chionh (1990) reports experimenting with the implementations of Dixon's method and Macaulay resultants in MAPLE for implicitization, parameterization and surface intersection problems. Since the matrices that arise are large with polynomial entries, special techniques such as interpolation and modular methods are needed to do slightly nontrivial examples.

Manocha and Canny (1991) have recently reported impressive results in using multivariate resultant-based methods for implicitization problems using interpolation and modular methods. This is the first and very impressive illustration of multivariate resultants outperforming Gröbner basis methods and characteristic set methods for the implicitization problem.

3. Gröbner Bases Computations

We first establish some of the basic notions needed for our exposition of Gröbner basis computations. The discussion below assumes the coefficient field to be Q. However, the development of Gröbner basis theory as discussed below carries over to polynomial rings in a finite number of variables over most fields (field of complex numbers, finite fields, field of rational functions in k variables over the complex numbers, ...).

3.1. Term Orderings

As said before, a term or power product is any product of powers $x_1^{\alpha_1}x_2^{\alpha_2}\ldots x_n^{\alpha_n}$ of the variables $x_1, x_2, \ldots, x_n$ with $\alpha_j \geq 0$. For a term $t = x_1^{\alpha_1}x_2^{\alpha_2}\ldots x_n^{\alpha_n}$, $\deg(t)$ denotes the total degree of the term t, i.e. $\deg(t) = \alpha_1 + \alpha_2 + \ldots + \alpha_n$. We are interested in *total orderings* (denoted by $\prec$) on terms that satisfy the following properties.

1. *Compatibility with multiplication:* if t, t_1, t_2 are terms, then, $t_1 \prec t_2 \Longrightarrow t\,t_1 \prec t\,t_2$.
2. *Termination:* there can be no strictly decreasing infinite sequence of terms such as

$$t_1 \succ t_2 \succ t_3 \succ \ldots.$$

Such term orderings are called *admissible* orderings and they play a key role in the development of Gröbner basis theory. Commonly used term orderings are

(i) the *Lexicographic Order*, $\prec_l$, in which terms are ordered as in a dictionary i.e. for terms t_1, t_2 with $t_1 = x_1^{\alpha_1}x_2^{\alpha_2}\ldots x_n^{\alpha_n}$ and $t_2 = x_1^{\beta_1}x_2^{\beta_2}\ldots x_n^{\beta_n}$ then $t_1 \prec_l t_2$ iff $\exists i \leq n$ such that $\alpha_j = \beta_j$ for $i < j \leq n$ and $\alpha_i < \beta_i$. For example, for terms made up of two variables x_1, x_2, where $x_1 \prec x_2$, we have

$$1 \prec_l x_1 \prec_l x_1^2 \prec_l x_1^3 \ldots \prec_l x_2 \prec_l x_1x_2 \prec_l x_1^2x_2 \ldots \prec_l x_2^2 \prec_l x_1x_2^2 \prec_l x_1^2x_2^2 \ldots$$

(ii) the *Degree Order*, $\prec_d$, in which terms are compared first by their degrees, and equal degree terms are compared lexicographically i.e.

$$t_1 \prec_d t_2 \text{ iff } \deg(t_1) < \deg(t_2) \text{ or } \deg(t_1) = \deg(t_2) \text{ and } t_1 \prec_l t_2.$$

For example, in the bivariate case, assuming $x_1 \prec x_2$ we have

$$1 \prec_d x_1 \prec_d x_2 \prec_d x_1^2 \prec_d x_1x_2 \prec_d x_2^2 \prec_d x_1^3 \prec_d x_1^2x_2 \prec_d x_1x_2^2 \prec_d x_2^3 \dots$$

3.2. Head Terms and the Notion of Reduction

Given an admissible term order $\prec$, for every polynomial f in $Q[x_1, x_2, \dots, x_n]$, we call the largest term (under $\prec$) in f that has a non-zero coefficient as the head term of f, denoted by head(f). By ldcf(f), we denote the leading coefficient of f, i.e. the coefficient of head(f) in f. Clearly, for every polynomial f, we can write

$$f = \mathrm{ldcf}(f)\ \ \mathrm{head}(f) + g \quad \text{where} \qquad \mathrm{head}(g) \prec \mathrm{head}(f).$$

We write tail(f) for g. For example, if

$$f(x, y) = x^3 - y^2, \text{ then}$$

$$\mathrm{head}(f) = x^3 \text{ and } \mathrm{tail}(f) = -y^2$$

under the total degree ordering $\prec_d$, and

$$\mathrm{head}(f) = y^2, \qquad \mathrm{tail}(f) = x^3$$

under the purely lexicographic ordering with $x \prec_l y$.

Let f and g be two polynomials; suppose g has a term t with a non-zero coefficient that is a multiple of head(f), i.e.

$$g = at + \hat{g} \text{ where } a \in Q \text{ and } t = t' \ \ \mathrm{head}(f)$$

for some term t'. We say that g is *reducible* with respect to f and denote a *reduction by* f as

$$g \xrightarrow{f} h$$

where

$$h = g - \hat{a}\ \ t'\ \ f = \hat{a}\ \ t'\ \ \mathrm{tail}(f) + \hat{g}, \qquad \text{and} \qquad \hat{a}\ \mathrm{ldcf}(f) \ = \ a.$$

The polynomial g is said to be *reducible* with respect to a set (or basis) of polynomials $F = \{f_1, f_2, \dots, f_r\}$ if it is reducible with respect to one or more polynomials in F; else we say that *g is reduced* or *g is a normal form* with respect to F.

Given a polynomial g and a basis $F = \{f_1, f_2, \dots, f_r\}$, through a *finite sequence of reductions*

$$g = g_1 \xrightarrow{F} g_2 \xrightarrow{F} g_3 \dots \xrightarrow{F} g_s,$$

we can obtain a polynomial g_s that is reduced with respect to F. Two things to note:

- Clearly, any sequence of reductions has to end after a finite number of reductions; if not, we can create an infinite decreasing chain of terms from this sequence which contradicts the assumption that the term ordering being used is terminating.

- For every g_i in the above reduction sequence,

$$g_i - g \in (f_1, f_2, \ldots, f_r).$$

Let us consider an example. Let

$$f_1 = x_1^2x_2 - 2x_2x_3 + 1,\ f_2 = x_1x_2^2 - x_3^2 + 2x_1,\ f_3 = x_2^2x_3 - x_1^2 + 5,\ \text{ and, }\ g = 3x_1^2x_2^2 + x_1^3 - 1.$$

Let $F = \{f_1, f_2, f_3\}$. Under $\prec_d$,

$$\text{head}(f_1) = x_1^2x_2, \qquad \text{head}(f_2) = x_1x_2^2, \qquad \text{head}(f_3) = x_2^2x_3$$

and g is reducible with respect to F. One possible reduction sequence is:

$$g = g_1 \xrightarrow{f_1} g_2 = 6x_2^2x_3 - 3x_2 + x_1^3 - 1 \xrightarrow{f_3} g_3 = 6x_1^2 - 3x_2 + x_1^3 - 31$$

and g_3 is a normal form with respect to F. It is possible to reduce g in another way that leads to a different normal form! For example, we have

$$g = g_1 \xrightarrow{f_2} g_2' = 3x_1x_3^2 + x_1{}^3 - 6x_1^2 - 1$$

and the normal form g_2' is different from g_3. For an arbitrary set of basis polynomials, we cannot expect to avoid this phenomenon. A Gröbner basis has the following special property:

DEFINITION 3.1. *A basis $G \in Q[x_1, x_2, \ldots, x_n]$ is called a Gröbner basis for the ideal it generates if and only if every polynomial in $Q[x_1, x_2, \ldots, x_n]$ has a unique normal form with respect to G.*

Gröbner bases were introduced by Buchberger (1965, 1976). One of his fundamental contributions was to show that every ideal in $Q[x_1, x_2, \ldots, x_n]$ has a Gröbner basis. He designed an algorithm to construct a Gröbner basis for any ideal I in $Q[x_1, x_2, \ldots, x_n]$ starting from an arbitrary basis for I.

Let us consider the earlier example for a moment. We found at least two normal forms (g_3 and g_2') for the polynomial g with respect to F. The reason was that g had a monomial that was reducible by two polynomials in the basis F, i.e. head(g) was a common multiple of head(f_1) and head(f_2). The ambiguity concerning the normal form of g with respect to F can be resolved by augmenting the basis F by the polynomial $g_3 - g_2'$ (the augmented basis still generates the same ideal since $g_3 - g_2' \in (F)$). Buchberger's algorithm attempts to resolve the situation for *all terms that have more than one normal form* with respect to F. The key insight in Buchberger's algorithm is to show that we need to consider only a finite set of terms. To this end, we define an s-polynomial of two polynomials f_1, f_2. Let

$$m = \text{lcm}(\text{head}(f_1), \text{head}(f_2)) = \quad m_1 \ \ \text{head}(f_1) = \quad m_2 \ \ \text{head}(f_2)$$

where m_1, m_2 are terms. Define

$$s\text{-poly}(f_1, f_2) = m_1 \ \ \text{ldcf}(f_2)f_1 - m_2 \ \ \text{ldcf}(f_1)f_2.$$

In the following description of Buchberger's algorithm, by $\mathcal{NF}_G(f)$, we denote *any* normal form of f with respect to the *current elements* of the basis G. The basis G is augmented until $\mathcal{NF}_G(s\text{-poly}(g_i, g_j))$ is zero for the s-polynomial of every pair of polynomials in G.

Given a basis F for an ideal $\mathcal{I}$ and an admissible term ordering $\prec$, the algorithm returns a Gröbner basis for $\mathcal{I}$ for the term ordering $\prec$.

```
$G := F$;
$B := \{\langle f_i, f_j\rangle \mid f_i, f_j \in F, i < j\}$;
while $B$ is non-empty do
    $h := \mathcal{NF}_G(s\text{-poly}(f_i, f_j))$;  % for some $\langle f_i, f_j\rangle \in B$
    if $h \neq 0$ then
        $B := B \cup \{\langle h, g\rangle | g \in G\}$;
        $G := G \cup \{h\}$;
    fi;
    $B := B \setminus \{\langle f_i, f_j\rangle\}$;
od;
```

For the well-known proofs of termination and correctness of the algorithm, the reader is referred to Buchberger (1965, 1976). We now list some of the most important properties of Gröbner bases in the form of a theorem.

THEOREM 3.1. *The following properties are equivalent:*

1. *G is a Gröbner basis for the ideal I with respect to a term order $\prec$.*
2. *For every pair of polynomials $g_1, g_2 \in G$, the normal form of the s-polynomial of g_1, g_2 with respect to G is zero.*
3. *Every polynomial f in $Q[x_1, x_2, \ldots, x_n]$ has a unique normal form with respect to G.*
4. *A polynomial f is a member of the ideal I if and only if its normal form with respect to G is zero.*

Examples: Consider the ideal $\mathcal{I}$ generated by

$$f(x,y) = x^3 - y^2, \; g(x,y) = 2x^2 + (y-1)^2 - 2$$

that we saw earlier (f defines a cusp and g defines an ellipse).

$$G_1 = \{y^2 + 2x^2 - 2y - 1, x^3 + 2x^2 - 2y - 1\}$$

is a Gröbner basis for $\mathcal{I}$ under the degree ordering with $x \prec y$.

$$\begin{aligned} G_2 &= \{g_1(x) = x^6 + 4x^5 + 4x^4 - 6x^3 - 4x^2 + 1, \\ &\quad g_2(x,y) = y + 1/2(-x^3 - 2x^2 + 1)\} \end{aligned}$$

is a Gröbner basis for $\mathcal{I}$ under the lexicographic ordering with $x \prec y$.

$$\begin{aligned} G_3 &= \{h_1(y) = y^6 - 6y^5 + 17y^4 + 4y^3 - 9y^2 - 6y - 1, \\ &\quad h_2(x,y) = x + 1/4(2y^5 - 13y^4 + 40y^3 - 10y^2 - 18y - 5)\} \end{aligned}$$

is a Gröbner basis for $\mathcal{I}$ under the lexicographic ordering with $y \prec x$.

These examples illustrate the fact that, in general, an ideal has different Gröbner bases depending on the term ordering that we choose. But, can an ideal have different Gröbner bases for the same term ordering? The answer is *no,* provided we restrict our attention to the so-called reduced Gröbner bases.

DEFINITION 3.2. *A Gröbner basis is called a* reduced *Gröbner basis iff for every $g \in G$,*

$$\mathcal{NF}_{G'}(g) = g$$

where $G' = G \setminus \{g\}$, i.e. each polynomial in the basis G is reduced with respect to all the other polynomials in G.

In the above example, G_1, G_2, G_3 are all reduced Gröbner bases. The basis

$$G_0 = \{f\} \cup G_1$$

is a Gröbner basis for $\mathcal{I}$ under the degree lexicographic ordering. However, it is not a reduced basis since f is reducible with respect $G_0 \setminus \{f\}$. From here on, by a Gröbner basis, we mean a reduced Gröbner basis unless specified otherwise. The choice of the term ordering depends on what we wish to use a Gröbner basis for. Also, the time needed to compute a Gröbner basis is very sensitive to the term ordering used.

In the Gröbner basis G_2 for the above example, there is a polynomial $g_1(x)$ that depends only on x and one that depends on both x, y (in general, there can be several polynomials in a Gröbner basis that depend on x, y). We say that the *variables are separated in the basis* G_2. A basis in which variables are separated is similar to a triangular form; in a triangular form, there is at most one polynomial in $x_1, x_2, \cdots, x_i$ for each $1 \leq i \leq n$, but in a basis in which variables are separated, there can be more than one polynomials in $x_1, x_2, \cdots, x_i$ for each $1 \leq i \leq n$. Such a separation of variables can be used to compute all the common zeros of the ideal $\mathcal{I}$. We first find all the roots of the univariate polynomial $g_1(x)$. These give the x-coordinates of the common zeros of the ideal $\mathcal{I}$. Similar to the case of a linear system of equations, we can perform back substitution. For each root α of g_1, we can find the common roots of $g_2(\alpha, y), \ldots$ which give the y-coordinates of the corresponding common zeros of $\mathcal{I}$. This algorithm is mentioned only to indicate the kind of uses that one can derive from Gröbner bases.

Ideal bases in which the variables are separated are very useful in solving a variety of problems. In fact, the separation observed in Gröbner bases G_2, G_3 above is not accidental. It was observed by Trinks (and Buchberger) that such a separation of variables exists in Gröbner bases whenever the term ordering used is a lexicographic one. We now illustrate the observations of Trinks (1978) and Buchberger (1985).

Let $\mathcal{I}$ be an ideal in the polynomial ring $Q[x_1, x_2, \ldots, x_n]$ and let $j < n$. The set of all the polynomials in $\mathcal{I}$ that depend only on $x_1, \ldots, x_j$ constitute a sub-ideal of $\mathcal{I}$. More precisely, let

$$\mathcal{I}_j = \mathcal{I} \cap Q[x_1, \ldots, x_j].$$

The ideal $\mathcal{I}_j$ is called the *contraction of the ideal* $\mathcal{I}$ *to the subring* $Q[x_1, \ldots, x_j]$. Some authors refer to it as the j-th elimination ideal of $\mathcal{I}$.

THEOREM 3.2. *Let G be the reduced Gröbner basis for an ideal $\mathcal{I} \subseteq Q[x_1, x_2, \ldots, x_n]$ under the lexicographic order on terms with $x_1 \prec x_2 \prec \ldots \prec x_n$. Then,*

$$\mathcal{I} \cap Q[x_1, \ldots, x_j] = (G \cap Q[x_1, \ldots, x_j])Q[x_1, \ldots, x_j]$$

for each $i = 1, 2, \ldots, n$.

Here, by $(G \cap Q[x_1, \ldots, x_j])Q[x_1, \ldots, x_j]$, we mean the ideal generated by $(G \cap Q[x_1, \ldots, x_j])$ in the ring $Q[x_1, \ldots, x_j]$, i.e.

$$(G \cap Q[x_1, \ldots, x_j])Q[x_1, \ldots, x_j] = \{\sum g_i h_i \mid g_i \in (G \cap Q[x_1, \ldots, x_j]), h_i \in Q[x_1, \ldots, x_j]\}.$$

Example: Let $F = \{f_1, f_2, f_3\}$ where

$$\begin{aligned}
f_1 &= x_3^4 - 6\,x_1x_3^3 + 13\,x_3^2x_1^2 - 12\,x_3x_1^3 + 4\,x_1^4,\\
f_2 &= x_2^2 - 2\,x_2x_1 - 2\,x_2x_3 + x_1^2 + 2\,x_1x_3 + x_3^2,\\
f_3 &= x_1^2 + 4\,x_1 + 3,\\
f_4 &= x_1x_3^3 - 3\,x_3^2x_1^2 + 3\,x_3x_1^3 - x_1^4\\
&\quad + x_3^3 - 3\,x_3^2x_1 + 3\,x_3x_1^2 - x_1^3 + x_1x_3 - 2\,x_1^2 + 3\,x_3 - 6\,x_1
\end{aligned}$$

Let G denote the reduced Gröbner basis for (F) under the lexicographic ordering with $x_1 \prec x_2 \prec x_3$. We have

$$G = \{g_1, g_{2,1}, g_{2,2}, g_{3,1}, g_{3,2}, g_{3,3}\}$$

where

$$\begin{aligned}
g_1 &= x_1^2 + 4\,x_1 + 3,\\
g_{2,1} &= x_1x_2^2 + 3\,x_2^2 + 9\,x_1 + 27 + 6\,x_2x_1 + 18\,x_2,\\
g_{2,2} &= 2\,x_2^3 + 36\,x_2^2 - 27\,x_2x_1 + 135\,x_2 - 108\,x_1 + 108,\\
g_{3,1} &= x_1x_3 + 2\,x_1 + 3\,x_3 + 6,\\
g_{3,2} &= 4\,x_2x_3 + 24\,x_3 - 2\,x_2^2 + 4\,x_2x_1 + 15\,x_1 + 45,\\
g_{3,3} &= 2\,x_3^2 + 12\,x_3 - x_1 + 15,
\end{aligned}$$

and,

$$\mathcal{I}_1 = (g_1)\mathcal{Q}[x_1],\ \mathcal{I}_2 = (g_1, g_{2,1}, g_{2,2})\mathcal{Q}[x_1, x_2],\ \mathcal{I}_3 = (G)\mathcal{Q}[x_1, x_2, x_3] = \mathcal{I}.$$

In principle, any system of polynomial equations can be solved using a lexicographic Gröbner basis for the ideal generated by the given polynomials by the algorithm outlined above. The fact that Gröbner bases under lexicographic term orderings exhibit separation of variables is useful in many situations. In geometric modeling, this property has been used

- to compute intersections of curves and surfaces,
- to find an implicit equation for a curve or surface given parametrically, and
- to determine whether a given rational parameterization of a curve or surface is faithful.

For further details and illustrations of the above applications, we refer the reader to Manocha and Canny (1990), Hoffman (1989, 1990) and Hoffman and Vermeer (1991).

If a set of polynomials does not have a common zero, i.e. its ideal is the whole ring, then it is easy to see that a Gröbner basis of such a set of polynomials includes 1 no matter what term ordering is used. Gröbner basis computations can thus be used to check for the consistency of a system of nonlinear polynomial equations.

THEOREM 3.3. *A set of polynomials in $\mathcal{Q}[x_1, \cdots, x_n]$ has no common zero in $\mathcal{C}$ if and only if their reduced Gröbner basis with respect to any admissible term ordering is* $\{1\}$.

A refutational method based on this result is proposed for automatically proving geometry theorems in Kapur (1986, 1988). A refutational approach using Gröbner basis computations for propositional calculus as well as for first-order predicate calculus is discussed in Kapur and Narendran (1985).

3.3. Gröbner Bases in Practice

In general, Gröbner bases are hard to compute. That it is inherently so was shown by Mayr and Meyer (1982) whose result can be used to exhibit ideals for which doubly exponential degree explosions during Gröbner basis computations are inevitable. In addition, the coefficients of polynomials that get generated during Gröbner bases computations can get extremely large. Another problem can arise due to the choice of the term ordering used in a Gröbner basis computation. We have seen that lexicographic bases or those similar to lexicographic bases are the most useful. However, in practice, lexicographic bases are known to be the hardest to compute.

Despite these difficulties, highly non-trivial Gröbner bases computations have been performed. Computations with ideals in polynomial rings over the rational numbers with 8-10 variables with degrees of polynomials in the initial basis about 5 are feasible. If the coefficients belong to a finite field (typically $\mathcal{Z}_p$ where p is a word sized prime), much larger computations are possible. Macaulay (Bayer and Stillman, 1989) and CoCoA (Giovini and Niesi, 1990) are specialized computer algebra systems built for performing large computations in algebraic geometry and commutative algebra; they are quite easy to use and provide a variety of built-in functions for computing with Gröbner bases. Most general computer algebra systems (such as MAPLE, MACSYMA, MATHEMATICA, REDUCE) also provide the basic Gröbner basis functions. An implementation of Gröbner basis algorithm also exists in GeoMeter (Cyrluk *et al.*, 1988; Connolly *et al.*, 1989), a programming environment for geometric modeling and algebraic reasoning. This implementation has been used for proving nontrivial plane geometry theorems using a refutational approach discussed in Kapur (1988).

Most Gröbner basis implementations use several modifications to Buchberger's algorithm in order to speed up the computations. We now briefly describe some of the common improvements.

Recall that in Buchberger's algorithm, one computes a normal form of an s-polynomial with respect to the current basis and, if it is non-zero, augments the current basis with the normal form. The s-polynomial reductions are repeated until all s-polynomials have normal form zero. An s-polynomial reduction is said to be ***useless*** if it does not produce a new polynomial to augment the current basis with (i.e. a reduction that produces a zero normal form). It has been observed that a lot of time is spent in performing useless reductions and one would like to avoid as many useless reductions as possible. In order to facilitate this, Buchberger proposed some simple conditions for predicting useless reductions. Since then, several researchers have invented variations and extensions to Buchberger's criteria. We only present Buchberger's criteria.

- If the head terms of g_i and g_j are co-prime, then s-polynomial of g_1, g_2 can be reduced to zero by the current basis. Hence, there is no need to perform the reduction.
- If g, g_1, g_2 are such that head(g) divides lcm(head(g_1), head(g_2)) and reductions of s-poly(g, g_1) and s-poly(g, g_2) are already done, then it is not necessary to perform the reduction of s-poly(g_1, g_2).

The implementation of these criteria for predicting useless reductions along with the so-called ***normal selection strategy*** (each time, the s-polynomial of the pair g_i, g_j for which lcm(head(g_1), head(g_2)) is the smallest among all the untried pairs in the current basis) is known to improve the running time of Buchberger's algorithm significantly. For complete details, we refer the reader to Gebauer and Möller (1988) and Giovini *et al.* (1991).

3.4. Zero Dimensional Ideals and Basis Conversion

We mentioned earlier the sensitivity of Gröbner basis computations to the term ordering being used. It is observed in practice that lexicographically ordered Gröbner bases (and lexicographic-like ordered bases, namely, the block ordered bases) are much harder to compute than the total degree ordered bases. However, for a number of applications, as we have already seen, one needs to compute a Gröbner basis under a lexicographic or lexicographic-like ordering. This raises the following question:

- *Suppose we are given the reduced Gröbner basis G_1 for an ideal $\mathcal{I}$ under a degree ordering, can one compute the reduced Gröbner basis for $\mathcal{I}$ under a lexicographic ordering much faster than by a direct computation using Buchberger's algorithm on G_1?*

Faugère *et al.* (1989) provided an elegant answer to this question for a special class of ideals called *zero-dimensional ideals.*

Recall that an ideal $\mathcal{I} \in \mathcal{Q}[x_1, x_2, \ldots, x_n]$ is said to be zero-dimensional if $\mathsf{Zero}(\mathcal{I})$ is finite. In other words, there are only finitely many common zeros of the polynomials in $\mathcal{I}$. The following property of Gröbner bases, observed first by Buchberger, characterizes 0-dimensional ideals:

Theorem 3.4. *Let G be the reduced Gröbner basis for an ideal $\mathcal{I} \in \mathcal{Q}[x_1, x_2, \ldots, x_n]$ under* any *admissible term ordering. $\mathcal{I}$ is zero-dimensional iff, for each i, $1 \le i \le n$, G contains a polynomial whose head term is a pure power of x_i, i.e. of the form $x_i^{d_i}$ for some integer d_i.*

A term t is said to be reduced with respect to G if t is not divisible by the head term of any polynomial in G. The condition just mentioned amounts to saying that the number of terms reduced with respect to G is finite iff the ideal $\mathcal{I}$ is zero dimensional. The number of terms reduced with respect to G is an important invariant (we denote it by D) of the ideal $\mathcal{I}$. It is the same as the cardinality of $\mathsf{Zero}(\mathcal{I})$ or the number of common zeros, *counted with multiplicities.* The algorithm of Faugère, Gianni, Lazard and Mora does the following:

- *Given a Gröbner basis G_1 for a zero-dimensional ideal $\mathcal{I}$ under some term ordering $\prec_1$, compute a reduced Gröbner basis G_2 for $\mathcal{I}$ under a second term ordering $\prec_2$.*

The original intent of Faugère *et al.* (1989) was to find all the common zeros of $\mathcal{I}$ quickly. Their algorithm, which we refer to as the *basis conversion algorithm*, is typically used as follows for solving zero-dimensional systems of equations:

- Compute a Gröbner basis G_1 for the ideal generated by the given system of polynomials under a total degree ordering $\prec_1$.
- Use the *basis conversion algorithm* to obtain a Gröbner basis G_2 under a lexicographic ordering $\prec_2$.

- Apply the *back substitution* phase (described earlier) on G_2 to obtain the common zeros.

The authors report spectacular success for this approach on several bench-mark systems of equations including cases where a direct computation of a lexicographic basis had never been carried out (usually the machine would run out of space on these examples!). Since then, variations of this technique have been developed to solve several problems related to zero-dimensional ideals (Lakshman, 1990).

The basis conversion algorithm enumerates terms starting from 1, in the *increasing order with respect to the second (new) term ordering* $\prec_2$. As it considers a term t, it classifies t as one of the following using the Gröbner basis G_1.

- t is *reduced* with respect to G_2, or
- t is *a lead term* of some polynomial in G_2, or
- t is *a multiple* of some lead term with respect to G_2.

Note that G_2 is the *desired basis* and hence the above classification is non-trivial. In fact, this classification is what leads to the construction of G_2. The complete description of the basis conversion algorithm follows:

- Let $\mathcal{NF}(t)$ denote the normal form of the term t with respect to G_1.
- Newbasis: Gröbner basis being built.
- ReducedTerm: Set of monomials that are known to be reduced with respect to Newbasis; Initialized to $\{1\}$.
- NextTerm: Function that returns the smallest monomial (under the desired admissible term ordering) that is neither in ReducedTerm nor is a multiple of some lead term in Newbasis. Returns false if no such monomial exists.

```
ReducedTerm:= $\{1\}$;
Newbasis:= $\{\ \}$;
while ($t$ := NextTerm()) do
    If there exist $t_1, \ldots, t_s$ in ReducedTerm, and $\lambda_j \in Q$
    such that $\mathcal{NF}(t) + \sum_{j=1}^{s} \lambda_j \mathcal{NF}(t_j) = 0$, then,
        Newbasis := Newbasis $\cup$ $\{t + \sum_{j=1}^{s} \lambda_j t_j\}$
    else
        ReducedTerm := ReducedTerm $\cup$ $\{t\}$;
        Save $\mathcal{NF}(t)$;
    fi
od end;
```

Example: Consider the ideal $\mathcal{I} = (f, g)$ where $f = x^3 - y^2$, $g = (y-1)^2 + 2x^2 - 2$. The basis $G = \{f_1 = y^2 - 2y - 1 - 2x^2, f_2 = x^3 + 2x^2 - 2y - 1\}$ is the reduced Gröbner basis for $\mathcal{I}$ under the degree ordering with $x \prec y$. Note that under this ordering, x^3 is the head term of f_2 and y^2 is the head term of f_1. Indeed, $\mathcal{I}$ is a zero-dimensional ideal since G has a polynomial whose lead term is a pure power of x and one whose lead term is a pure power of y (we have already seen that there are only finitely many common zeros of $\mathcal{I}$). Suppose we wish to compute the reduced Gröbner basis G_2 for $\mathcal{I}$ under the lexicographic term order with $y \prec_l x$. We proceed thus:

We begin by looking at the term y (the smallest term under $\prec_l$ that is not yet classified). At this point, we only know that the term 1 is reduced with respect to G_2. We note that

$\mathcal{NF}(y) = y$. Our attempt to find a rational constant λ_1 such that

$$\mathcal{NF}(y) + \lambda_1(\mathcal{NF}(1)) = 0$$

ends in failure. We therefore conclude that y is *reduced* with respect to G_2 and consider y^2. We find that $\mathcal{NF}(y^2) = 2y + 1 - 2x^2$ and this time, we look for rational constants λ_1, λ_2 such that

$$\mathcal{NF}(y^2) + \lambda_2(\mathcal{NF}(y)) + \lambda_1(\mathcal{NF}(1)) = 0. \qquad (iv)$$

Since rational constants λ_1, λ_2 satisfying the above relation do not exist, we conclude that y^2 is also reduced with respect to G_2. We then consider y^3, y^4, y^5 (all of which fail to produce a linear relation of the type (iv)) and finally, y^6. At this point, we are looking for rational constants $\lambda_1, \lambda_2, \lambda_4, \lambda_4, \lambda_5, \lambda_6$ such that

$$\begin{gathered}\mathcal{NF}(y^6) + \lambda_6(\mathcal{NF}(y^5)) + \lambda_5(\mathcal{NF}(y^4)) + \lambda_4(\mathcal{NF}(y^3)) \\ +\lambda_3(\mathcal{NF}(y^2)) + \lambda_2(\mathcal{NF}(y)) + \lambda_1(\mathcal{NF}(1)) = 0.\end{gathered}$$

Substituting the appropriate normal forms, we have

$$\begin{array}{cl}(72yx^2+ & 176xy - 570y - 259 + 506x^2 + 76x) \\ + & \lambda_6(-12yx^2 + 52xy - 107y - 52 + 96x^2 + 24x) \\ + & \lambda_5(-4y - 8yx^2 - 3 + 4x^2 + 8xy + 4x) \\ + & \lambda_4(5y - 2yx^2 + 2 - 4x^2) \\ + & \lambda_3(2y + 1 - 2x^2) \\ + & \lambda_2(y) \\ + & \lambda_1(1) = 0.\end{array}$$

Equating the coefficients of like terms, we have

$$\begin{array}{ccl}-12\lambda_6 - 8\lambda_5 - 2\lambda_4 + 72 & = 0, & \{\text{coeff. of } yx^2\} \\ 52\lambda_6 + 8\lambda_5 + 176 & = 0, & \{\text{coeff. of } xy\} \\ 96\lambda_6 + 4\lambda_5 - 4\lambda_4 - 2\lambda_3 + 506 & = 0, & \{\text{coeff. of } x^2\} \\ -107\lambda_6 - 4\lambda_5 + 5\lambda_4 + 2\lambda_3 + \lambda_2 & = 0, & \{\text{coeff. of } y\} \\ 24\lambda_6 + 4\lambda_5 + 76 & = 0, & \{\text{coeff. of } x\} \\ -52\lambda_6 - 3\lambda_5 + 2\lambda_4 + \lambda_3 + \lambda_1 & = 0, & \{\text{constant term}\}\end{array}$$

This is a system of linear equations in the variables λ_i and can be solved easily. The unique solution is

$$\lambda_6 = -6,\ \lambda_5 = 17,\ \lambda_4 = 4,\ \lambda_3 = -9,\ \lambda_2 = -6,\ \lambda_1 = -1$$

(the uniqueness of the solution can be deduced from the uniqueness of the reduced Gröbner basis for $\mathcal{I}$ with respect to the term order $\prec_l$). Therefore, we classify y^6 as a head term with respect to G_2 and add the polynomial

$$y^6 - 6y^5 + 17y^4 + 4y^3 - 9y^2 - 6y - 1$$

to the basis G_2. The next term that is considered by the algorithm is x (at this point, x is the *smallest unexamined* term according to $\prec_l$ that is not a multiple of any term known to be a lead term with respect to G_2). We now look for a linear relation among $\mathcal{NF}(x), \mathcal{NF}(y^5), \mathcal{NF}(y^4), \mathcal{NF}(y^3), \mathcal{NF}(y^2), \mathcal{NF}(y)$ and 1. Such a relation exists and we find that

$$x + 1/4(2y^5 - 10y2 - 13y4 + 40y3 - 18y - 5)$$

to be the polynomial resulting from the linear relation. Therefore, we add it to the basis G_2. We now know two lead terms in G_2, namely, y^6, x. The next term that we examine must not be divisible by either y^6 or x. But we have already examined all such terms and classified them. Hence, the algorithm terminates, and we have the basis

$$G_2 = \{x + 1/4(2y^5 - 10y2 - 13y4 + 40y3 - 18y - 5), y^6 - 6y^5 + 17y^4 + 4y^3 - 9y^2 - 6y - 1\}.$$

Note that the λ_i are determined by solving a linear system of equations. For the purposes of illustration, we wrote down a complete linear system. The linear systems of equations that arise in this algorithm have a nice structure (a consequence of the way they are generated) and in practice, one takes advantage of the structure to find the λ_i efficiently. It is shown in Faugère *et al.* (1989) that the number of rational arithmetic operations performed by the basis conversion algorithm is $O(n^2D^3 + n^2D^2\log(nD))$ (recall that D is the number of reduced terms with respect to G_1; n is the number of variables). In order to achieve the above bound for the running time of the algorithm, it is necessary to save all the normal forms of the terms computed by the algorithm (in the step save $\mathcal{NF}(t)$).

The termination of the basis conversion algorithm follows from the property that zero-dimensional ideals have only finitely many terms that are in normal form with respect to its Gröbner basis constructed using an admissible ordering. The fact that the basis generated from the above construction is a Gröbner basis is a corollary of the following property of Gröbner bases.

DEFINITION 3.3. *Given a basis F, and a term order $\prec$, define $init_\prec(F)$ to be the set of all head terms of the polynomials in F. For an ideal I, define the* initial ideal *with respect to $\prec$ to be the ideal generated by the set $init_\prec(\mathcal{I})$.*

It can be shown that a basis G of an ideal $\mathcal{I}$ is a Gröbner basis with respect to $\prec$ if and only if the initial ideal of $\mathcal{I}$ with respect to $\prec$ is generated by the head terms of the polynomials in the basis G. The basis conversion algorithm has been used for computing the implicit equation of a parametrically given surface (see Hoffman, 1989).

3.5. Triangular sets

We have seen earlier how lexicographic Gröbner bases can be used to solve systems of polynomial equations. A Gröbner basis contains all the information about the zeros of the ideal it generates, including multiplicities (this fact is essential when one wants to compute the primary decomposition of an ideal; see Lakshman, 1990, for instance). However, if one is interested merely in the location of the zeros of an ideal (let us assume for the moment that we are dealing with zero dimensional ideals), then the multiplicities can slow down the computation of the coordinates of the zeros of the ideal. It is possible to obtain sets of polynomials in which all the variables are separated as in a lexicographic Gröbner basis but the sets are much *simpler* than a lexicographic Gröbner basis.

Kandri-Rody (1984) showed how to construct such a set from a lexicographic Gröbner basis and called it an *extracted characteristic* set following Ritt. Extracted characteristic sets were used by Kandri-Rody for testing the primality of an ideal as well as for computing the dimension of an ideal. Lazard (1989a, 1989b) also defined triangular sets for studying zero sets ideals. In the rest of this section, we describe Lazard's triangular sets for zero dimensional ideals and present an algorithm due to Lazard for computing triangular sets from a lexicographic Gröbner basis.

A triangular set is any set of polynomials

$$f_1, f_2, \ldots, f_n \in Q[x_1, x_2, \ldots, x_n]$$

such that

$$f_1 \in Q[x_1],\ f_2 \in Q[x_1, x_2],\ f_3 \in Q[x_1, x_2, x_3],\ \ldots, f_n \in Q[x_1, x_2, \ldots, x_n].$$

The highest variable under the ordering $x_1 \prec x_2 \prec \ldots \prec x_n$ appearing in a polynomial f is called the *main variable* of f. By $\deg(f_i)$, we mean the degree of f_i in its main variable x_i, and, by $\deg_j(f_i)$, we mean the degree of f_i in x_j. The triangular set is called *reduced* if $\deg_j(f_i) < \deg(f_j)$ for all $j < i \leq n$. In this subsection, by a triangular set, we mean a reduced triangular set in which every polynomial is monic in its main variable, i.e. its initial is 1. It can be shown that *the zero set of every zero-dimensional ideal $\mathcal{I}$ is the union of the zero-sets of finitely many distinct triangular sets.* We say that two sets of polynomials are *equivalent* if they have the same zero set.

Example: Consider the ideal $\mathcal{I}$ given by the basis G below. In fact, G is the reduced Gröbner basis for $\mathcal{I}$ under the pure lexicographic ordering with $x \prec y \prec z$.

$$\begin{aligned} G = \{g_6 &= 4z^2 + 4xz + 13x^3 - 88x^2 + 173x - 94, \\ g_5 &= yz - y - z + 1, \\ g_4 &= x^2z - x^2 - 3xz + 3x + 2z - 2, \\ g_3 &= y^2 - 1, \\ g_2 &= yx^2 - 6yx + 9y - x^2 + 6x - 9, \\ g_1 &= x^4 - 9x^3 + 29x^2 - 39x + 18\} \end{aligned}$$

Given below is one possible triangular set decomposition of the zero set of $\mathcal{I}$,

$$\{(x^2 - 3x + 2, y - 1, z^2 + xz + 1),\ (x^2 - 6x + 9, y^2 - 1, z - 1)\}$$

whose structure is simpler than that of the Gröbner basis G.

We now sketch Lazard's algorithm for computing a triangular set decomposition of a zero set informally using the above example. In a reduced lexicographic Gröbner basis of a zero dimensional ideal, there is always a single polynomial in the lowest variable; this polynomial goes into a triangular set. However, there can be many polynomials in the other variables. For the above example, there are two polynomials, g_2, g_3, in x, y and there are three polynomials, g_4, g_5, g_6, in x, y, z. We attempt to make the smallest polynomial in x, y, i.e. g_2, monic. When considered as a polynomial in y, it has $x^2 - 6x + 9$ as the leading coefficient. We attempt to compute the inverse of $x^2 - 6x + 9$ modulo the polynomial g_1, i.e. try to compute a polynomial h such that

$$(x^2 - 6x + 9)h \equiv 1 \pmod{g_1}.$$

But $g_1 = x^4 - 9x^3 + 29x^2 - 39x + 18 = (x^2 - 6x + 9)(x^2 - 3x + 2)$, implying that $x^2 - 6x + 9$ cannot be inverted modulo g_1. Therefore, we try to split the triangular set constructed thus far into two triangular sets: $T^{(1)}$ containing $x^2 - 6x + 9$ and $T^{(2)}$

containing the $x^2 - 3x + 2$. For each triangular set, we repeat the above operation of making the remaining polynomials in the Gröbner basis monic.

With respect to $T^{(1)}$, the second polynomial $g_2 = yx^2 - 6yx + 9y - x^2 + 6x - 9$ in the Gröbner basis simplifies to 0. The third polynomial g_3 is already monic, so it is added to $T^{(1)}$. Consider g_4, the smallest polynomial in x, y, z, whose leading coefficient is $x^2 - 3x + 2$. We attempt to invert it with respect to $T^{(1)}$; using the extended Euclidean algorithm, as discussed below, its inverse with respect to $x^2 - 6x + 9$ is $-3/4x + 11/4$, i.e.

$$(x^2 - 3x + 2)(-3/4x + 11/4) = 1 \ \ mod \ \ x^2 - 6x + 9.$$

When we multiply g_4 by $-3/4x + 11/4$ and simplify by $T^{(1)}$, we get $z - 1$, which is added to $T^{(1)}$; we have now completed the computation of one triangular set. It is easy to see that g_5 and g_6 simplify to 0 with respect to $T^{(1)}$.

The computation of the second triangular set $T^{(2)}$ containing $x^2 - 3x + 2$ is done in the same way. The inverse of the leading coefficient of g_2 can be computed with respect to $T^{(2)}$ using which g_2 is made monic, and $y - 1$ is added to $T^{(2)}$. Polynomial g_3 simplifies to 0 using $T^{(2)}$. Polynomials g_4, g_5 also simplify to 0 using $T^{(2)}$. Simplifying g_6 gives $z^2 + xz + 1$ which is added to $T^{(2)}$.

As the reader might have noticed, two operations are needed with respect to a triangular set: *simplification/reduction* and *inversion.* They are discussed next.

3.5.1. Reduction with respect to a Triangular Set

Given a triangular set $T = \{f_1, f_2, \ldots, f_n\}$ with x_i being the main variable in f_i and a polynomial $f \in Q[x_1, \ldots, x_n]$. Let d_i be the degree of f_i in its main variable x_i. The *remainder* of f with respect to T is defined as follows. We *divide* f by f_n with a remainder, i.e.

$$f = Q_n f_n + r_n$$

where Q_n is the quotient, and r_n is the remainder. The degree of r_n in x_n is less than the degree of f_n in x_n. We can now divide r_n by f_{n-1} treating them both as a polynomials in x_{n-1}. The coefficients of f_{n-1} are polynomials in $x_1, \ldots, x_{n-2}$ and the coefficients of r_n are polynomials in $x_1, \ldots, x_{n-2}, x_n$, i.e. we have

$$r_n = Q_{n-1} f_{n-1} + r_{n-1}$$

where the degree of the remainder r_{n-1} in $x_n < d_n$ and its degree in $x_{n-1} < d_{n-1}$. We can compute successive remainders $r_{n-2}, \ldots, r_1$ with respect to $f_{n-2}, \ldots, f_1$. The last remainder r_1 is such that $\deg_i(r_1) < d_i$ for $1 \leq i \leq n$ and it is called the *remainder* of f with respect to T. The process of obtaining r_1 from f and T is called *reducing* f by T.

Example: Let $f = z^2 + y^2 + x^2$. Its remainder with respect to the triangular set $(x^2 - 3x + 2, y - 1, z^2 + xz + 1)$ with $x \prec y \prec z$ is computed as follows:

$$\begin{aligned} \text{the remainder of } f \text{ using } (z^2 + xz + 1) &= x^2 + y^2 - xz - 1 \\ \text{the remainder of } (x^2 + y^2 - xz - 1) \text{ using } (y - 1) &= x^2 - xz \\ \text{the remainder of } (x^2 - xz) \text{ using } (x^2 - 3x + 2) &= -xz + 3x - 2 \end{aligned}$$

The polynomial $-xz + 3x - 2$ is the remainder of f with respect to the triangular set. Notice that the reduction of a polynomial f with respect to a triangular set T is the same as computing the normal form of f with respect to T (which also happens to be a reduced Gröbner basis).

3.5.2. Inversion with respect to a Triangular Set

Let $T = \{f_1, f_2, \ldots, f_n\}$ be a triangular set with x_i being the main variable of f_i. Without any loss of generality, we can assume that f is already reduced with respect to T (if f is not already reduced with respect to T, reduce it using T). By inverting a polynomial f with respect to T, we mean finding a polynomial g such that the remainder of gf with respect to T is 1. Not every polynomial has an inverse with respect to a triangular set T. In case f does not have an inverse with respect to T, the process of trying to invert f *may* lead to a splitting of T. Let x_i be the highest variable appearing in f. Treating f and f_i as polynomials with coefficients that are polynomials in $x_1, \ldots, x_{i-1}$, i.e. $f, f_i \in Q[x_1, \ldots, x_{i-1}][x_i]$, compute their greatest common divisor using the extended Euclidean algorithm (see Knuth, 1980, pp. 407–408). The crucial point here is that we proceed as though the leading coefficients of the remainders that appear in performing Euclid's algorithm on f and f_i are *invertible* in T. Let g be their gcd. We have

$$g = fp + f_i q$$

where p, q are the multipliers produced by the extended Euclidean algorithm. There are three possibilities:

- if $g = 1$, then p is the inverse of f with respect to T;
- if $g = f_i$, then f_i divides f and f is *not* invertible with respect to T.
- if $f_i \neq g \neq 1$, then f_i has factors $g, f_i/g$ and the triangular set T is now *equivalent* to the union of the two triangular sets

  $$T^{(1)} = \{f_1, f_2, \ldots, g, f_{i+1}, \ldots, f_n\}, \text{ and, } T^{(2)} = \{f_1, f_2, \ldots, f_i/g, f_{i+1}, \ldots, f_n\}.$$

  f is not invertible with respect to $T^{(1)}$ and the inverse of f with respect to $T^{(2)}$ is given by $g^{-1}p$ where g^{-1} denotes the inverse of g with respect to $T^{(2)}$.

The crucial point is that the triangular set T can split at *lower levels.* This is because the extended Euclidean algorithm used to compute the gcd of f, f_i needs to compute *inverses* of polynomials with respect to f_j, $j < i$, which is done recursively.

To summarize, the operation of inverting a polynomial f with respect to a triangular set T produces a family of triangular sets $T^{(1)}, T^{(2)}, \ldots, T^{(k)}$ and polynomials $h_1, \ldots, h_k$ such that

- T is *equivalent* to the union of the triangular sets $T^{(1)}, T^{(2)}, \ldots, T^{(k)}$.
- if $h_i \neq 0$, then h_i is the inverse of f with respect to $T^{(i)}$, i.e. the remainder of fh_i with respect to $T^{(i)}$ is 1; if $h_i = 0$, then f is not *invertible* with respect to $T^{(i)}$.

What we have sketched here is, in essence, the *D5 method* of Duval for handling algebraic numbers. We illustrate the operation of inversion with respect to a triangular set below. For complete details on the D5 technique, we refer the reader to Duval (1991).

Example: Let us invert the polynomial $f = (x-1)z + 1$ with respect to the triangular set $T = (x^2 - 3x + 2, y - 1, z^2 + xz + 1)$ of the previous example with the ordering $x \prec y \prec z$.

The first step is to compute the gcd of $f, z^2 + xz + 1$, treating them as polynomials in z. At this point, we are required to invert $x - 1$ (the leading coefficient of f) with respect to T. Since $x - 1$ is independent of y, z, we try to compute the gcd of $x^2 - 3x + 2, x - 1$ treating them as polynomials in x; we find that $x - 1$ is their *gcd.* This leads to a split of T into $T^{(1)} = (x - 1, y - 1, z^2 + xz + 1)$ and $T^{(2)} = (x - 2, y - 1, z^2 + xz + 1)$. The polynomial $x - 1$ is not invertible with respect to $T^{(1)}$; it has an inverse with respect to $T^{(2)}$, which is 1 (since $(x - 1).1 \equiv 1 \bmod (x - 2)$). $T^{(1)}, T^{(2)}$ can be reduced to give $T^{(1)} = (x-1, y-1, z^2 + z + 1)$ and $T^{(1)} = (x-1, y-1, z^2 + 2z + 1)$. Note that we may have to *reduce* the triangular sets as we propagate the split upwards.

Our task now is to invert f with respect to $T^{(1)}$ and $T^{(2)}$ separately. The polynomial f has remainder 1 with respect to $T^{(1)}$ and 1 is its own inverse; f has remainder $z + 1$ with respect to $T^{(2)}$. We find that the gcd of $z + 1, z^2 + 2z + 1$ is $z + 1$. Therefore, f is not invertible with respect to $T^{(2)}$.

3.5.3. Lazard's Algorithm for Computing a Family of Triangular Sets

The algorithm that we present is called *D5Lextriangular* in Lazard (1989a). It uses the D5 method to compute a triangular set decomposition of the zero set of a zero dimensional ideal given a lexicographic Gröbner basis for the ideal. We would like to point out that there are algorithms for computing triangular sets that do not need a lexicographic Gröbner basis for input. We refer the interested reader to Lazard (1989a) and Lakshman (1990a).

Algorithm D5Lextriangular

Input: A reduced Gröbner basis G for a zero dimensional ideal $\mathcal{I}$ under the lexicographic ordering with $x_1 \prec x_2 \prec \ldots \prec x_n$; it is assumed that G is presented as a list with the head terms of the polynomials sorted in increasing order under $\prec$.

Output: A list of triangular sets equivalent to G.

Functions used:

Reduce(f, T): reduces the polynomial modulo the triangular set T.

Inverse(f, TL): f is a polynomial and TL is a list of triangular sets $U_1, \ldots, U_l$ (the list can be empty). The function returns a list of pairs, $[(h_1, T_1), \ldots, (h_k, T_k)]$, $k \geq 1$, such that the union of the triangular sets in TL is equivalent to the union of triangular sets $T_1, \ldots, T_k$; if $h_i \neq 0$, then, h_i is the inverse of f modulo T_i, else f is not invertible modulo T_i.

Lcoeff(f, x): returns the leading coefficient of f treated as a univariate polynomial in the variable x.

```
$TL := [[first(G)]]$; % start with the univariate polynomial in $G$.
for $i$ from 2 to $n$ do
    $H$ :=sublist of polynomials in $G$ that depend on $x_i$
        but not on $x_{i+1}, \ldots, x_n$.            repeat
    $f :=$ first($H$);
    $H :=$ rest($H$);
```

```
                $g$ := Lcoeff($f, x_i$);
                $L$ := Inverse($g, TL$);
                $TL$ := [];
                for each pair $(h_j, T_j)$ in $L$
                   if ($h_j \neq 0$) then
                       add Reduce($fh_j, T_j$) to $T_j$;
                       append $T_j$ to the list $TL$;
                   fi od;
          until $H$ is empty;
      od;
      return($TL$);
```

The above algorithm can be further optimized as follows:

- In the inner for-loop, if a $T_j \in TL$ already contains a polynomial that depends on x_i (for the current i), then, it is not necessary to perform the body of the for-loop for that T_j as Reduce(fh_j, T_j) will turn out to be zero.

We illustrate the algorithm and the optimization using the previous example. For y, H includes two polynomials, g_2 and g_3; Inverse is invoked on the initial of g_2 with respect to g_1. This leads to a split of the triangular set consisting of g_1 giving two triangular sets $T^{(1)} = \{x^2 - 6x + 9\}$ and $T^{(2)} = \{x^2 - 3x + 2\}$. and $h_1 = 0$ and $h_2 = -3/4x + 11/4$. At the end of the first iteration of the second inner loop, $T^{(1)} = \{x^2 - 6x + 9\}$ and $T^{(2)} = \{x^2 - 3x + 2, y - 1\}$. Now, g_3 reduces to 0 using $T^{(2)}$. The inverse of the initial of g_3 is computed with respect to $T^{(1)}$ (since the initial is 1, the inverse is also 1) which results in the second element $y^2 - 1$ added to $T^{(1)}$. Similarly, for z there are three polynomials g_4, g_5, g_6 in ascending order. The inverse of the initial of g_4 is computed with respect to $T^{(1)}$ as well as $T^{(2)}$. Using $T^{(1)}$, the inverse is $-3/4x + 11/4$ which is multiplied with g_4 to give $z - 1$ and that is added to $T^{(1)}$. For $T^{(2)}$, the inverse of the initial of g_4 does not exist. Polynomials g_5 and g_6 reduce to 0 using $T^{(1)}$. The inverse of the initial of g_5 does not exist with respect to $T^{(2)}$ either, but the inverse of the initial of g_6 with respect to $T^{(2)}$ can be computed thus giving the third element $z^2 + xz + 1$ for $T^{(2)}$. This gives the decomposition in terms of triangular sets.

The correctness of the algorithm is based on a theorem due to Gianni and Kalkbrenner (see Lazard, 1989a; Gianni, 1987; Kalkbrenner, 1987). These ideas are extended to positive dimensional ideals in Lazard (1989b).

4. Characteristic Set Construction

In this section, we discuss Ritt's characteristic set construction. The discussion is based on Ritt's presentation in his book *Differential Algebra* (Ritt, 1950, Chapter 4) and Wu's exposition of Ritt's approach as discussed in Wu (1986a). We first give an informal explanation of Ritt's characteristic set method and then give the technical details. For many applications of the characteristic set construction, the reader can consult Wu (1984, 1986a, 1986b), Chou (1988), Chou and Gao (1990a, 1990b, 1990c, 1990d), Kapur and Mundy (1988) and Kapur and Wan (1990).

Given a system $\mathcal{S}$ of polynomial equations, the characteristic set algorithm transforms $\mathcal{S}$ into *a triangular form* $\mathcal{S}'$ so that the zero set of $\mathcal{S}$ is "roughly equivalent" to the zero set of S'. A total ordering on variables is assumed. Unlike in Gröbner basis computations, polynomials are treated as univariate polynomials in their highest variable. The primitive operation used in the transformation is that of pseudo-division of a polynomial by another polynomial. The algorithm is similar in flavor to Gaussian elimination for linear equations and computing a lexicographic Gröbner basis.

If the number of equations in S is less than n, the number of variables, then the variable set $\{x_1, \cdots, x_n\}$ is typically classified into two subsets: ***independent*** variables (also called ***parameters***) and ***dependent*** variables, and a total ordering on the variables is so chosen that the dependent variables are all higher in the ordering than the independent variables. We denote the independent variables by $u_1, \cdots, u_k$ and dependent variables by $y_1, \cdots, y_l$, and the total ordering is $u_1 \prec \ldots \prec u_k \prec y_1 \prec \ldots \prec y_l$, where $k + l = n$. Choosing independent variables and defining an ordering on dependent variables is similar to choosing an ordering on variables for defining a lexicographic term ordering for computing a Gröbner basis.

To check whether an equation $f = 0$ follows from S, f is pseudo-divided using the polynomials in a triangular form S' of S. If the remainder of pseudo-division is 0, then $f = 0$ is said to follow from S under the condition that the initials of polynomials in S' are not zero. We discuss later why this condition on the initials is needed.

4.1. Ritt-Wu's Theorem

First, we formally define concepts used above in the characteristic set construction. Then, we give the definition of a characteristic set and the main theorem about characteristic sets.

Given a polynomial p, the ***highest variable*** of p is y_i if $p \in Q[u_1, \ldots, u_k, y_1, \ldots, y_i]$ and $p \notin Q[u_1, \ldots, u_k, y_1, \ldots, y_{i-1}]$ (there is a similar definition for u_i). The *class* of p is then called i.

A polynomial p is $\geq$ another polynomial q if and only if

1. the highest variable of p, say y_i, is $\succ$ the highest variable of q, say y_j, i.e. the class of p is higher than the class of q, or
2. the class of p = the class of q, and the degree of p in y_i is $\geq$ the degree of q in y_i.

If $p \geq q$ and $q \geq p$, then p and q are said to be *equivalent,* written as $p \cong q$; this means that p and q are of the same class i, and the degree of y_i in p is the same as the degree of y_i in q.

A polynomial p is *reduced with respect to* another polynomial q if

(a) the highest variable, say y_i, of p is $\prec$ the highest variable of q, say y_j (i.e. $p \prec q$), or

(b) $y_i \succ y_j$ and the degree of the y_j in q is $>$ the degree of y_j in p.[†]

If p is not reduced with respect to q, then p *reduces* to r using q by pseudo-dividing p by q giving r as the remainder of the result of pseudo-division.

A list C of polynomials, $\langle p_1, \cdots, p_m \rangle$ is called a *chain* if either

[†] As we shall see later, this definition of reduction will be weakened in order to improve the efficiency of various algorithms for computing a characteristic set.

(i) $m = 1$ and $p_1 \neq 0$, or
(ii) $m > 1$ and the class of p_1 is > 0, and for $j > i$, p_j is of higher class than p_i and reduced with respect to p_i; we thus have $p_1 \prec p_2 \prec \ldots \prec p_m$. (Wu, 1986a).

(A chain is the same as an *ascending set* defined by Wu.)

A polynomial p reduces to p' with respect to a chain $C = \langle p_1, \cdots, p_m \rangle$ if there exist nonnegative integers $i_1, \cdots, i_m$ such that

$$I_1{}^{i_1} \cdots I_m{}^{i_m} p = q_1 p_1 + \cdots + q_m p_m + p'$$

where p' is reduced with respect to each of $p_i, 1 \leq i \leq m$. Typically pseudo-division is successively done using p_i's starting with p_m, the polynomial in the highest variable.

Given two chains $C = \langle p_1, \cdots, p_m \rangle$ and $C' = \langle p'_1, \cdots, p'_{m'} \rangle$, $C \succ C'$ if (i) there is a $j \leq m$ as well as $j \leq m'$ such that $p_i \cong p'_i$ for all $i < j$ and $p_j \succ p'_j$, or (ii) $m' > m$ and for $i \leq m$, $p_i \cong p'_i$.

As stated earlier, a set $G = \{g_1, \cdots, g_m\}$ of polynomials is said to be in *triangular form* if and only if $g_1, g_2, \cdots, g_m$ are, respectively, polynomials in $\{u_1, \cdots, u_k, y_1\}$, $\{u_1, \cdots, u_k, y_1, y_2\}$, $\cdots$, $\{u_1, \cdots, u_k, y_1, \cdots, y_m\}$. If $m = l$, then G is said to be fully triangular. It is easy to see that every chain is in triangular form.

4.1.1. Characteristic Set

Ritt was apparently interested in associating characteristic sets only with *prime* ideals. A prime ideal is an ideal with the property that if an element h of the ideal can be factored as $h = h_1 h_2$, then, either h_1 or h_2 must be in the ideal. For a prime ideal Σ, Ritt defined a subset of Σ that forms the lowest chain to be a *characteristic set* of Σ. In chapter 4 in the section *Components of Finite Systems* of his book *Differential Algebra* (Ritt, 1950, p. 95), Ritt describes an algorithm for computing characteristic sets for all the minimal prime ideals containing an ideal $\mathcal{I}$, given a basis for $\mathcal{I}$.

Wu (1986a) altered Ritt's notation somewhat to make it more useful. Wu associated a characteristic set with the zero set of an arbitrary set of polynomials. He called a characteristic set associated with a prime ideal (equivalently, an irreducible zero set) as *irreducible*. A zero set is irreducible if it cannot be expressed as a union of proper algebraic subsets.[†]

A characteristic set $\{g_1, \cdots, g_l\}$ as defined by Wu could be irreducible or reducible, whereas a characteristic set defined by Ritt is always irreducible. It is interesting to note that in his first paper on geometry theorem-proving (Wu, 1984), Wu defined a characteristic set to be a triangular set in which for $i = 1$ to l,

1 the initial of g_i is a polynomial in the parameters only, and
2 g_i is irreducible over Q_{i-1} where $Q_0 = \mathcal{Q}(u_1, \cdots, u_k)$ and $Q_j = Q_{j-1}(\alpha_j)$ is an algebraic extension of Q_{j-1} obtained by adjoining a root α_j of $g_j = 0$ to Q_{j-1}, i.e. $g_j(\alpha_j) = 0$ in Q_j for $1 \leq j < i$.

Wu called such a triangular set as a *privileged basis* associated with a prime ideal. He

[†] Recall that a zero set is algebraic if it is the zero set of a set of a polynomials.

credited this definition to Gröbner 's book on algebraic geometry (Gröbner , 1949), where it is called a *prime basis*, and to Ritt for "his intimately related concept of a characteristic set" (see Wu, 1984, p. 219). In his subsequent papers, Wu relaxed this requirement on initials on pragmatic grounds, and required a polynomial in a characteristic set to be reduced with respect to other polynomials. In the sequel, we have followed Wu's definitions as they seem to be more useful.

Wu (1986a) attributes the following definition and theorem to Ritt.

DEFINITION 4.1. *Given a finite set Σ of polynomials in $u_1, ..., u_k, y_1, ..., y_l$, a* characteristic set *Φ of Σ is defined to be either*

- *$\{p_1\}$, where p_1 is a polynomial in $u_1, ..., u_k$, or*
- *a chain $\langle p_1, ..., p_l\rangle$, where p_1 is a polynomial in $y_1, u_1, ..., u_k$ with initial I_1, p_2 is a polynomial in $y_2, y_1, u_1, ..., u_k$, with initial I_2, ..., p_l is a polynomial in $y_l, ..., y_1, u_1, ..., u_k$ with initial I_l, such that*
  - ○ *any zero of Σ is a zero of Φ, and*
  - ○ *any zero of Φ that is not a zero of any of the initials I_i, is a zero of Σ.*

THEOREM 4.1. (RITT) *Given a finite set Σ of polynomials in $y_l, ..., y_1, u_1, ..., u_k$, there is an algorithm which computes a* characteristic set *Φ of Σ.*

The algorithm discussed in the next subsection involves augmenting Σ with additional polynomials from the ideal generated by Σ obtained through pseudo-division until we have a set Δ such that

1 $\Sigma \subseteq \Delta$,
2 Σ and Δ generate the same ideal, and
3 a minimal chain Φ of Δ pseudo-divides every polynomial in Δ to 0,

then Φ is a characteristic set of Σ as well as Δ. We will call Δ to be a *saturation* of Σ.

Since every polynomial q in Δ pseudo-divides to 0 using Φ, i.e.

$$I_1{}^{i_1} \cdots I_l{}^{i_l} q = q_1 p_1 + \cdots + q_m p_l.$$

This implies:

$$\mathsf{Zero}(\Sigma) = \mathsf{Zero}(\Delta) \subseteq \mathsf{Zero}(\Phi),$$

and

$$\mathsf{Zero}(\Sigma) = (\mathsf{Zero}(\Phi) \setminus \mathsf{Zero}(\prod_{i=1}^{l} I_i)) \cup \bigcup_i \mathsf{Zero}(\Sigma \cup \{I_i\}).$$

Using Wu's notation, $\mathsf{Zero}(\Phi/I)$ to stand for $\mathsf{Zero}(\Phi)\setminus\mathsf{Zero}(I)$, the above can be rewritten as:

$$\mathsf{Zero}(\Sigma) = \mathsf{Zero}(\Phi/\prod_{i=1}^{l} I_i) \cup \bigcup_i \mathsf{Zero}(\Sigma \cup \{I_i\}).$$

THEOREM 4.2. *Given a finite set Σ of polynomials, if its characteristic set Φ includes a constant (in the case $l = n$), then Σ does not have a common zero.*

If Φ does not include a constant, it does not mean that Σ has common zeros. For example, consider $\Sigma = \{(x^2-2x+1) = 0, (x-1)z-1 = 0\}$. Under the ordering $x \prec z$, Σ is a characteristic set. The two polynomials do not have a common zero (however, under the ordering $z \prec x$, the characteristic set of Σ includes 1).

The converse of the above theorem holds only for irreducible characteristic sets. This is discussed in a later subsection.

4.2. Algorithms for Computing a Characteristic Set

A characteristic set Φ is computed from a set Σ of polynomials by successively adjoining Σ with remainder polynomials obtained by pseudo-division. Starting with $\Sigma_0 = \Sigma$, we extract a minimal chain (called a *basic set* by Wu) from Σ_i, and compute non-zero remainders of polynomials in Σ_i with respect to the minimal chain.† If this remainder set is nonempty, we adjoin it to Σ_i to obtain Σ_{i+1} and repeat the computation until we have Σ_n such that every polynomial in Σ_n pseudo-divides to 0 with respect to its minimal chain. The set Σ_n is a saturation of Σ and the minimal chain of Σ_n is a characteristic set of Σ_n as well as Σ. This algorithm is given in chapter 4 in the section *Components of Finite Systems* in Ritt's *Differential Algebra* (Ritt, 1950, p. 95). The above construction terminates since the minimal chain of Σ_i is $\succ$ the minimal chain of Σ_{i+1} and the ordering on chains is well-founded (the maximum size of a chain is l, and the degree of at least one polynomial in Σ_{i+1} is lower than the degree of the corresponding polynomial in the same variable in Σ_i).

In the process of computing a characteristic set from Σ, if an element of the coefficient field (a rational number if $l = n$ or a rational function in $Q(u_1, \cdots, u_k)$) is generated as a remainder, this implies that Σ does not have a solution, or is *inconsistent*.

Below, we give two algorithms for computing characteristic sets as implemented in GeoMeter for geometry theorem proving (Connolly *et al.*, 1989; Kapur and Wan, 1990). These algorithms differ on the order in which remainders are computed and processed. The first one is Ritt's algorithm.‡

Char-Set-Breadth-first$(H, \prec)$

Input: A set of polynomials H and a variable ordering $\prec$.
Output: A characteristic set for H under the variable ordering $\prec$.
Functions used:

Basic-set$(E, \prec)$: Described below.

pseudo-divide-reduction$(p, B, \prec)$: successively reduces (pseudo-divides) the polynomial p with respect to the polynomials in the basic set B starting with the largest polynomial with respect to $\prec$.

```
        $E := \emptyset$; $R := H$;
        while $R \neq \emptyset$ do
```

† Note that a minimal chain extracted from a set need not be unique.

‡ Recall that by the smallest polynomial in a set S, we mean a polynomial p whose highest variable, say y_i, is not greater than the highest variable of any other polynomial in S and furthermore, among the polynomials with y_i as the highest variable, the degree of p in y_i is the least.

```
        $E := E \cup R$;
        $B$ :=Basic-set($E, \prec$);
        $R := \{q \mid q =$pseudo-divide-reduction$(p, B, \prec), q \neq 0, p \in E \setminus B\}$;
    od;
    Return $B$;
```

Basic-set$(S, \prec)$

Input: A set of polynomials S and a variable ordering $\prec$.
Output: A basic set contained in S with respect to $\prec$.

```
    $B := \emptyset$; $T := S$;
    while $T \neq \emptyset$ do
        $p :=$ a smallest polynomial in $T$;% "smallest" with respect to $\prec$.
        $B := B \cup \{p\}$;
        $T := \{q \mid q \in T \setminus \{p\}, q$ is reduced with respect to $p\}$;
    od;
    Return $B$;
```

The correctness of the above algorithm is based on the fact that q, the remainder from pseudo-division of p by a basic set B, is in the ideal generated by $B \cup \{p\}$, which is a subideal of the ideal generated by H. An invariant of the while loop in the procedure Char-Set-Breadth-first is that the ideal generated by E is the same as the ideal generated by H which implies that the zero sets of H and E are the same. If B is the result of Char-Set-Breadth-first, then every polynomial in E of the last iteration (which is a saturation of Σ as defined in the last subsection) pseudo-divides to 0 using B. Consequently, the zero set of B includes the zero set of H as a subset; in particular, the zero set of B minus the zero set of the initials of polynomials in B is the zero set of H.

The breadth-first strategy is quite inefficient for computing a characteristic set for detecting inconsistency. In each stage, all possible remainders are computed first before the next basic set is computed. It is much better to use a depth-first strategy given below. In this strategy, the basic set is updated each time a new remainder is obtained; in this way, remainders of lower classes can be obtained more quickly.

Char-Set-Depth-first$(H, \prec)$

Input: A set of polynomials H and a variable ordering $\prec$.
Output: A characteristic set for H with respect to $\prec$.
Functions used:
Basic-set$(E, \prec)$: same as before.
pseudo-divide-reduction$(p, B, \prec)$: same as before.
Update-Basic-set$(p, B, \prec)$: described below.

```
    $E := H$;
    $B$ :=Basic-set($E, \prec$);
    $ER := E \setminus B$;
    while ($ER \neq \emptyset$ and $1 \notin B$) do
        $S := E \setminus B$;
        $ER := \emptyset$;
        for each $p$ (starting from the smallest polynomial) $\in S$ do
            $ER := ER \cup$ Update-Basic-set($p, B, \prec$) od;
        $E := E \cup ER$;
    od;
    Return  $B$;
```

Update-Basic-set$(p, B, \prec)$

Input: A polynomial p, a basic set B with respect to a variable ordering $\prec$.
Output: The set of remainder polynomials generated when p is added to B. As a side effect, B is a new basic set.

```
    $r$ :=pseudo-divide-reduction($p, B, \prec$);
    if $r = 0$ then Return $\emptyset$
    else if $r$ is a constant then
        $B := \{1\}$;
        Return $\emptyset$;
    fi;
    $R := \{r\}$;
    $T := \{q \mid q \in B, q$ is not reduced with respect to $r\}$;
    $B := B \cup \{r\} \setminus T$;
    for each $q$ (starting from the smallest polynomial) $\in T$ do
        $R := R \cup$ Update-basic-set($q, B, \prec$) od;
    Return $R$;
```

The breadth-first and depth-first strategies do not always produce the same result. For the breadth-first strategy, in generating remainders during each stage, one remainder obtained does not in any way affect the computation of the next remainder. For the depth-first strategy however, since one remainder may affect what the next remainder might be, it is possible that some remainders computed in the breadth-first strategy may never be computed in the depth-first strategy.

4.3. Proving Conjectures from a System of Equations

A direct way to check whether an equation $c = 0$ follows (under certain conditions) from a system $\mathcal{S}$ of equations is to compute a characteristic set $\Phi = \{p_1, \cdots, p_l\}$ from $\mathcal{S}$ and check whether c pseudo-divides to 0 with respect to Φ. If c has a zero remainder with respect to Φ, then the equation $c = 0$ follows from Φ under the conditions that none of the initials used to multiply c is 0. The algebraic relation between the conjecture c and the polynomials in Φ can be expressed as:

$$I_1{}^{i_1} \cdots I_l{}^{i_l}\, c = q_1 p_1 + \cdots + q_l p_l,$$

where I_j is the initial of $p_j, j = 1, \cdots, l$.

This approach is used by Wu and Chou for geometry theorem proving. A characteristic set is computed from the hypotheses of a geometry problem; then a conjecture is pseudo-divided by the characteristic set to check whether the remainder is 0. If the remainder is 0, the conjecture is said to be generically valid from the hypotheses.

A refutational way to check whether $c = 0$ follows from $\mathcal{S}$ is to compute a characteristic set of $\mathcal{S} \cup \{cz - 1 = 0\}$, where z is a new variable. This approach has also been used for geometry theorem proving and discussed in Kapur and Wan (1990).

4.4. Wu's Structure Theorem

If a polynomial c representing a geometric conjecture does not pseudo-divide to 0 with respect to a characteristic set Φ, it cannot always be said that $c = 0$ does not follow from Φ or, for that matter, from Σ from which Φ is computed. Some additional properties need to be checked. We need to make sure that the characteristic set is *irreducible.*

Definition 4.2. *A characteristic set* $\Phi = \{p_1, \cdots, p_l\}$ *is* irreducible *over* $Q[u_1, ..., u_k, y_1, ..., y_l]$ *iff for* $i = 1$ *to* l, p_i *is irreducible over* Q_{i-1} *where* $Q_0 = Q(u_1, \cdots, u_k)$ *and* $Q_j = Q_{j-1}(\alpha_j)$ *is an algebraic extension of* Q_{j-1}, *obtained by adjoining a root* α_j *of* $p_j = 0$ *to* Q_{j-1}, *i.e.* $p_j(\alpha_j) = 0$ *in* Q_j *for* $1 \leq i < j$.

Similarly, if Σ does not have a solution, either a characteristic set Φ of Σ includes an element $Q(u_1, \cdots, u_k)$, or Φ is reducible and each of the irreducible characteristic sets includes an element of $Q(u_1, \cdots, u_k)$.

Theorem 4.3. *Given a finite set* Σ *of polynomials, if (i) its characteristic set* Φ *is irreducible, (ii)* Φ *does not include a constant, and (iii) the initials of the polynomials in* Φ *do not have a common zero with* Φ, *then* Σ *has a common zero.*

To deal with a reducible characteristic set, Ritt and Wu advocated the use of factorization over algebraic extensions of $Q(u_1, \cdots, u_k)$, which is an expensive operation. With reducibility check by factorization over extension field, the check for consistency is complete using the characteristic set method.†

In the case that any of the polynomials in a characteristic set can be factored, there is a branch for each irreducible factor as the zeros of p_j are the union of the zeros of its irreducible factors. Suppose we compute a characteristic set $\Phi = \{p_1, \cdots, p_l\}$ from Σ such that for $i > 0$, $p_1, \cdots, p_i$ are irreducible over $Q_0, \cdots, Q_{i-1}$, respectively, but p_{i+1} can be factored over Q_i. It can be assumed that

$$g \quad p_{i+1} = p^1{}_{i+1} \cdots p^j{}_{i+1},$$

where g is in $Q[u_1, \cdots, u_k, y_1, \cdots, y_i]$, and $p^1{}_{i+1} \cdots p^j{}_{i+1} \in Q[u_1, \cdots, u_k, y_1, \cdots, y_i, y_{i+1}]$ and these polynomials are reduced with respect to $p_1, \cdots, p_i$. Wu (1986a) proved that

$$\mathsf{Zero}(\Sigma) = \mathsf{Zero}(\Sigma^{11}) \cup \cdots \cup \mathsf{Zero}(\Sigma^{1j}) \cup \mathsf{Zero}(\Sigma^{21}) \cup \cdots \cup \mathsf{Zero}(\Sigma^{2i}),$$

where $\Sigma^{1h} = \Sigma \cup \{p^h{}_{i+1}\}, 1 \leq h \leq j$, and $\Sigma^{2h} = \Sigma \cup \{I_h\}$, where I_h is the initial of $p_h, 1 \leq h \leq i$. So characteristic sets are instead computed from new polynomials sets to give a system of characteristic sets. (To make full use of the intermediate computations already performed, Σ above can be replaced by a saturation of Σ.) The final result of this decomposition is a system of irreducible characteristic sets:

$$\mathsf{Zero}(\Sigma) = \bigcup_i \mathsf{Zero}(\Phi_i / J_i),$$

† We would like to remind the reader that we are only considering zeros over the complex numbers whereas, strictly speaking, for geometry theorem-proving one must check whether real zeros of the hypotheses are contained in the real zeros of the conclusion.

where Φ_i is an irreducible characteristic set and J_i is the product of the initials of all the polynomials in Φ_i. For further details, the reader should consult Wu (1986a).

Example: Consider the polynomial set

$$\begin{aligned} H = \{&13z^3 + 12x^2 - 23xz + 16z^2 - 49x + 39z + 40,\\ &13xz^2 - 12x^2 + 10xz - 29z^2 + 49x - 26z - 53,\\ &13x^3 - 88x^2 + 4xz + 4z^2 + 173x - 94,\\ &yz - y - z + 1, x^2z - x^2 - 3xz + 3x + 2z - 2,\\ &y^2 - 1, yx^2 - 6yx + 9y - x^2 + 6x - 9\}. \end{aligned}$$

The following characteristic set can be computed from it using the ordering $x \prec y \prec z$.

$$\begin{aligned} C = \{f_1 &= (x^2 - 3x + 2)z - x^2 + 3x - 2,\\ f_2 &= (x^2 - 6x + 9)y - x^2 + 6x - 9,\\ f_3 &= x^6 - 12x^5 + 58x^4 - 144x^3 + 193x^2 - 132x + 36\}. \end{aligned}$$

The polynomial f_3 in C can be factored as $(x-1)^2(x-2)^2(x-3)^2$, giving the following characteristic sets which can be further decomposed:

$$\begin{aligned} C_1 &= \{z^2 + z + 1, y - 1, x - 1\},\\ C_2 &= \{z^2 + 2z + 1, y - 1, x - 2\},\\ C_3 &= \{z - 1, y^2 - 1, x - 3\}. \end{aligned}$$

For a system of polynomials in which no variable has degree more than 2, such as in problems arising in Euclidean plane geometry, Chou developed an algorithm to check irreducibility and perform factorization in the case of reducibility to generate irreducible characteristic sets; the reader may consult Chou (1988).

4.5. Implementation

To our knowledge, none of the computer algebra systems supports an implementation of characteristic set construction. Characteristic set algorithms can be, however, easily implemented in any computer algebra system that supports an efficient implementation of pseudo-division. Wu and Chou have reported implementations using which they got extremely impressive results for plane geometry problems. GeoMeter has an implementation of the algorithms reported earlier in this section. Considerable work is needed to further improve characteristic set algorithms.

Based on their extensive experience in using the characteristic set construction, Wu and Chou have proposed modifications in the definitions of reduction as well as a polynomial being reduced with respect to a chain. In Nguyen *et al.* (1991), a lazy approach to the evaluation of the coefficients of a polynomial represented in recursive form is discussed.

As in Gröbner basis computations, the performance of algorithms to compute a characteristic set is sensitive to the classification of variables into dependent and independent variables, as well as the order of dependent variables. Several heuristics have been proposed for deducing a good ordering from a problem formulation in the application of geometry theorem proving.

5. Conclusion

When we started writing this introductory article, we had planned to discuss many additional topics. As evident, we will need a lot more space. We had to omit many issues, in particular a detailed discussion of the applications pointing out limitations of these approaches as they are currently practiced. We are, however, quite optimistic about the potential use of these methods.

One of the areas that we feel should be investigated is the apparent close relationship between the three approaches in the zero-dimensional case. We believe a lot can be learned from the study of a relationship among various approaches. In particular, it might be possible to achieve better bounds on the worst case complexity of Gröbner basis and characteristic set computations. We may also have a better idea about why these methods work better for sparse systems. It might be possible to apply results about redundant computations in Gröbner basis theory to resultant computation. On the other hand, identification of a close relationship between these approaches might give a clue about how techniques that work well for resultant computations can be used in Gröbner basis and characteristic set methods. Most importantly, there is a need for studying techniques for efficiently implementing these algorithms for polynomials arising in various application domains.

References

S.S. Abhyankar (1976), "Historical ramblings in algebraic geometry and related algebra", *American Math. Monthly,* **83**(6), 409–448.

D.S. Arnon, G.E. Collins and S. McCallum (1984), "Cylindrical algebraic decomposition I: the basic algorithm, II: an adjacency algorithm for the plane", *SIAM J. Comput.,* **13**, 865–889.

D. Bayer and M. Stillman (1989), *Macaulay User's Manual,* Cornell University, Ithaca, NY.

J.S. Brown and J.F. Traub (1971), "On Euclid's algorithm and the theory of subresultants", *J. ACM,* **18**(4), 505–514.

B. Buchberger (1965), *Ein Algorithmus zum Auffinden der Basiselemente des Restklassenringes Nach Einem Nulldimensionalen Polynomideal,* Ph.D. Thesis (in German), Universitat Innsbruck, Austria.

B. Buchberger (1976), "A theoretical basis for the reduction of polynomials to canonical form", *ACM SIGSAM Bull.,* **10**(3), 19–29.

B. Buchberger (1983), "A note on the complexity of constructing Gröbner bases", *Proc. EUROCAL '83,* London, Lecture Notes in Comput. Sci. 162, Springer-Verlag, NY, 137–145.

B. Buchberger (1985), "Gröbner bases: an algorithmic method in polynomial ideal theory", *Multidimensional Systems Theory,* N.K. Bose, ed., D. Reidel Publishing Co., Netherlands, 184–232.

J.F. Canny (1988), *The Complexity of Robot Motion Planning,* ACM Doctoral Dissertation Series, MIT Press, Cambridge, MA.

J.F. Canny (1990), "Generalized characteristic polynomials", *J. Symbolic Computation,* **9**, 241–250.

J.F. Canny, E. Kaltofen and Y.N. Lakshman (1989), "Solving systems of nonlinear polynomial equations faster", *Proc. Int. Symp. Symbolic Algebraic Computation (ISSAC-89),* Portland, OR, 121–128.

A. Cayley (1865), "On the theory of elimination", *Cambridge and Dublin Math. J.,* III, 210–270.

E. Chionh (1990), *Base Points, Resultants, and the Implicit Representation of Rational Surfaces,* Ph. D. Thesis, Dept. of Comput. Sci., University of Waterloo, Waterloo, Canada.

S.-C. Chou (1988), *Mechanical Geometry Theorem Proving,* D. Reidel Publishing Co., Netherlands.

S.-C. Chou and X.-S. Gao (1990a), "Ritt-Wu's decomposition algorithm and geometry theorem proving", *Proc. 10th Int. Conf. Automated Deduction (CADE-10),* Kaiserslautern, Germany, Lecture Notes in Comput. Sci., Springer-Verlag, NY, 207–220.

S.-C. Chou and X.-S. Gao (1990b), *On the Parameterization of Algebraic Curves,* Technical Report 90-18, Dept. of Comput. Sci., University of Texas, Austin.

S.-C. Chou and X.-S. Gao (1990c), *On the Normal Parameterization of Curves and Surfaces,* Technical Report 90-19, Dept. of Comput. Sci., University of Texas, Austin.

S.-C. Chou and X.-S. Gao (1990d), *Independent Parameters, Inversions and Proper Parameterization,* Technical Report 90-30, Dept. of Comput. Sci., University of Texas, Austin.

G.E. Collins (1967), "Subresultants and polynomial remainder sequences", *J. ACM,* **14**(1), 128–142.

G.E. Collins (1971), "The calculation of multivariate polynomial resultants", *J. ACM,* **18**(4), 515–532.

C.I. Connolly, D. Kapur, J.L. Mundy and R. Weiss (1989), "GEOMETER: a system for modeling and algebraic manipulation", *Proc. DARPA Workshop on Image Understanding,* Palo Alto, CA, 797–804.

M. Coste and M.F. Roy (1988), "Thom's lemma, the coding of real algebraic numbers and the computation of the topology of semi-algebraic sets", *J. Symbolic Computation,* **5**, 121–129.

D. Cyrluk, R. Harris and D. Kapur (1988), "GEOMETER: a theorem prover for algebraic geometry", *Proc. 9th Int. Conf. Automated Deduction (CADE-9),* Argonne, IL, Lecture Notes in Comput. Sci., Springer-Verlag, NY, 770-771.

A.L. Dixon (1908), "The eliminant of three quantics in two independent variables", *Proc. London Math. Soc.,* **6**, 468–478.

D. Duval (1991), "Computation with algebraic numbers: the D5 method", *J. Symbolic Computation,* to appear.

J.C. Faugère, P. Gianni, D. Lazard and T. Mora (1989), *Efficient Computation of Zero-Dimensional Gröbner Bases by Change of Ordering,* Technical Report 89–52, LITP, Universite Paris, Paris.

R. Gebauer and H.M. Möller (1988), "On an installation of Buchberger's algorithm", *J. Symbolic Computation,* **6**, 275–286.

P. Gianni (1987), "Properties of Gröbner bases under specializations", *Proc. EUROCAL '87,* Leipzig, Lecture Notes in Comput. Sci. 378, Springer-Verlag, NY, 293–297.

A. Giovini, T. Mora, G. Niesi, L. Robbiano and C. Traverso (1991), "One sugar cube, please" or "Selection strategies in the Buchberger algorithm", *Proc. Int. Symp. Symbolic Algebraic Computation (ISSAC '91),* ACM Press, Bonn, 49–54.

A. Giovini and G. Niesi (1990), *CoCoA User Manual,* (obtainable by anonymous ftp from gauss.dm.unipi.it).

D.Yu. Grigoriev and A.L. Chistov (1983), "Sub-exponential time solving of systems of algebraic equations", LOMI Preprints E-9-83 and E-10-83, Leningrad, USSR.

W. Gröbner (1968), *Algebraische Geometrie I* (in German), Bibliographisches Institut Mannheim.

W. Gröbner (1970), *Algebraische Geometrie II* (in German), Bibliographisches Institut Mannheim.

C.M. Hoffman (1989), *Geometric and Solid Modeling: An Introduction,* Morgan Kaufmann, San Mateo, CA.

C.M. Hoffman (1990), "Algebraic and numerical techniques for offsets and blends", *Computation of Curves and Surfaces,* W. Dahmen *et al.*, eds., Kluwer Academic Publishers, Boston, MA, 499–528.

C.M. Hoffman and P.J. Vermeer (1991), "Eliminating extraneous solutions in curve and surface operations", *Int. J. Comput. Geom. Appl.,* **1**(1), 47–66.

M. Kalkbrenner (1987), "Solving systems of algebraic equations by using Gröbner bases", Proc. *EUROCAL '87,* Leipzig, Lecture Notes in Comput. Sci. 378, Springer-Verlag, NY, 282–292.

A. Kandri-Rody (1984), *Effective Methods in the Theory of Polynomial Ideals,* Ph.D. Thesis, Dept. of Math., Rensselaer Polytechnic Institute, Troy, NY.

D. Kapur (1986), "Geometry theorem proving using Hilbert's Nullstellensatz", *Proc. Symp. Symbolic and Algebraic Computation (SYMSAC '86),* Waterloo, Canada, 202–208.

D. Kapur (1988), "A refutational approach to geometry theorem proving", *Artif. Intell.,* **37**, 61–94.

D. Kapur and J.L. Mundy (1988), "Wu's method and its application to perspective viewing", *Artif. Intell.,* **37**, 15–36.

D. Kapur and P. Narendran (1985), "An equational approach to theorem proving in first-order predicate calculus", *Proc. 7th Int. Joint Conf. Artif. Intell. (IJCAI-85),* Los Angeles, CA, 1146–1153.

D. Kapur and H. Wan (1990), "Refutational proofs of geometry theorems via characteristic set computation", *Proc. ACM-SIGSAM 1990 Int. Symp. Symbolic and Algebraic Computation (ISSAC '90),* Japan, 277–284.

M. Kline (1972), *Mathematical Thought: From Ancient to Modern Times:* I, II, and III, Oxford University Press, Oxford.

D.E. Knuth (1980), *Seminumerical Algorithms: The Art of Computer Programming,* 2nd edn., Addison-Wesley, Reading, MA.

Y.N. Lakshman (1990a), *On the Complexity of Computing Gröbner Bases for Zero Dimensional Polynomial Ideals,* Ph.D. Thesis, Dept. of Comput. Sci., Rensselaer Polytechnic Institute, Troy, NY.

Y.N. Lakshman (1990b), "A single exponential bound on the complexity of computing Gröbner bases of zero dimensional ideals", *Effective Methods in Algebraic Geometry (MEGA '90),* Progress in Math., **94**, Birkhäuser Boston, 227–234.

Y.N. Lakshman and D. Lazard (1990), "On the complexity of zero-dimensional algebraic systems", *Effective Methods in Algebraic Geometry (MEGA '90),* Progress in Math., **94**, Birkhäuser Boston, 217–226.

D. Lazard (1981), "Résolution des systèmes d'équations algébriques", *Theoretical Comput. Sci.,* **15**, 77–110.

D. Lazard (1989a), *Solving Zero-Dimensional Algebraic Systems*, Technical Report 89-48, Universite Paris, Paris.

D. Lazard (1989b), *A New Method for Solving Algebraic Systems of Positive Dimension*, Technical Report 89–77, LITP, Universite Paris, Paris.

R. Loos (1983), "Generalized polynomial remainder sequences", *Symbolic and Algebraic Computation (Computing Supplement 4)*, B. Buchberger, G.E. Collins and R. Loos, eds., Springer-Verlag, NY, 2nd edn., 115–137.

F.S. Macaulay (1916), *The Algebraic Theory of Modular Systems,* Cambridge Tracts in Math. and Math. Phys., **19**.

D. Manocha and J.F. Canny (1990), *Implicitizing Rational Parametric Surfaces,* Technical Report UCB/CSD 90/592, University of California, Berkeley.

D. Manocha and J.F. Canny (1991), "Efficient techniques for multipolynomial resultant algorithms", *Proc. Int. Symp. Symbolic and Algebraic Computation (ISSAC '91),* ACM Press, Bonn, 86–95.

E. Mayr and A. Meyer (1982), "The complexity of the word problem for commutative semigroups and polynomial ideals", *Advances Math.,* **46**, 305–329.

V.-D. Nguyen, J.L. Mundy and D. Kapur (1991), "Modeling generic polyhedral objects by constraints", *Proc. Comput. Vision Patt. Recog.,* Lahaina Maui, Hawaii, 479–485.

J. Renegar (1989), *On the Computational Complexity and Geometry of the First Order Theory of the Reals: I, II, and III,* Technical Reports 854, 856, 858, School of Oper. Res., Cornell University, Ithaca, NY.

J.F. Ritt (1950), *Differential Algebra,* AMS Colloquium Publications, New York.

T.W. Sederberg (1983), *Implicit and Parametric Curves and Surfaces*, Ph.D. Thesis, Purdue University, West Lafayette, IN.

D. Spear (1977), "A constructive approach to commutative ring theory", *Proc. 1977 MACSYMA User's Conf.,* Berkeley, CA, 369–376.

A. Tarski (1951), *A Decision Method for Elementary Algebra and Geometry,* University of California Press, Berkeley, CA.

W. Trinks (1978), "Über B. Buchberger's verfahren, systeme algebraischer gleichungen zu lösen", *J. Number Theory,* **10**, 475–488.

B.L. van der Waerden (1970), *Algebra,* **1** and **2**, Frederick Ungar Publishing Co., New York.

W. Wu (1984), "On the decision problem and the mechanization of theorem proving in elementary geometry", *Theorem Proving: After 25 years*, Contemporary Math., **29**, W. Bledsoe and D. Loveland, eds., American Math. Soc., Providence, RI, 213–234. Also in *Scientia Sinica,* (1978), **21**, 150–172.

W. Wu (1986a), "Basic principles of mechanical theorem proving in geometries", *J. Automated Reasoning,* **2**, 221–252. Also appeared in *J. Syst. Sci. Math. Sci.,* **4**(3), (1984), 207–235.

W. Wu (1986b), "On zeros of algebraic equations - an application of Ritt's principle", *Kexue Tongbao,* **31**(1), 1–5.