

Chapter 15

Algebraic Methods for Computer Aided Geometric Design

Thomas W. Sederberg and Jianmin Zheng

The concepts and methods of algebra and algebraic geometry have found significant applications in many disciplines. This chapter presents a collection of gleanings from algebra or algebraic geometry that hold practical value for the field of computer aided geometric design. We focus on the insights, algorithm enhancements and practical capabilities that algebraic methods have contributed to CAGD. Specifically, we examine resultants and Gröbner basis, and discuss their applications in implicitization, inversion, parametrization and intersection algorithms. Other topics of CAGD research work using algebraic methods are also outlined.

15.1. INTRODUCTION

CAGD draws from several branches of mathematics and computer science, such as approximation theory, differential geometry, and numerical analysis. This chapter reviews some of the tools of algebra and algebraic geometry that have been brought to bear on problems in CAGD [11,17,27,28,33,37,45].

Most of the free-form curves and surfaces used in CAGD are given by parametric equations. Planar rational curves in CAGD are typically defined as

$$x = \frac{a(t)}{c(t)}, \quad y = \frac{b(t)}{c(t)} \tag{15.1}$$

where $a(t)$, $b(t)$, and $c(t)$ are polynomials in the Bernstein basis for rational Bézier curves or in the B-spline basis for NURBS. Algebraic methods most commonly use polynomials in the power basis: $a(t) = a_0 + a_1t + \cdots + a_nt^n$, etc. Polynomials can be converted from Bernstein basis to power basis, although some algebraic methods such as resultants can be formulated using the Bernstein basis directly [26].

Parametric surfaces in CAGD are defined

$$x = \frac{a(s, t)}{d(s, t)}, \quad y = \frac{b(s, t)}{d(s, t)}, \quad z = \frac{c(s, t)}{d(s, t)} \quad (15.2)$$

where $a(s, t)$, $b(s, t)$, $c(s, t)$ and $d(s, t)$ are polynomials.

Surfaces and plane curves can also be defined using implicit equations. One contribution that algebraic methods make to CAGD is in solving the problem of implicitization and inversion of parametric curves and surfaces. For any parametric curve given by (15.1), an implicit equation $f(x, y) = 0$ (where $f(x, y)$ is a polynomial) exists that describes exactly the same curve. Likewise, for any parametric surface given by (15.2), there exists an implicit equation $f(x, y, z) = 0$ that describes exactly the same surface. The process of finding the implicit equation of a parametric curve or surface is called *implicitization*. Implicitization is of value in CAGD because the problem of determining whether a given point lies on a curve or surface is addressed much more easily using the implicit form than the parametric form. Curve implicitization is discussed in Section 15.4.

An *inversion* formula for a parametric curve (15.1) is of the form $t = \frac{g(x, y)}{h(x, y)}$ where g and h are polynomials. If the parametrization of a curve is a generally one-to-one map between parameter values and points on the curve, the inversion formula returns the parameter value t corresponding to a point (x, y) that lies on the curve. Inversion is discussed also in Section 15.4.

The process of finding the rational parametric equations of implicitly defined algebraic curves and surfaces is called *parametrization*. Some methods for parametrizing plane algebraic curves are shown in Section 15.5.

Algebraic methods also can facilitate the design of algorithms for computing intersections between curves and surfaces. The curve intersection problem is surveyed in Section 15.6.

The problems of implicitization, parametrization and intersection for surfaces are discussed in Section 15.7. Some other important applications of algebraic methods to CAGD are listed in Section 15.8.

Many of the algebraic methods reviewed in this chapter come from classical analytic geometry [31,32,46,48]. The twentieth century witnessed a marked shift from the constructive approach to non-constructive [5]. Section 15.2 gives a brief overview of polynomial ideals, varieties and Gröbner bases, and Section 15.3 introduces three popular resultant formulations.

15.2. POLYNOMIALS, IDEALS, AND VARIETIES

This section first introduces the notation and terminology which will be used later, and then presents the fundamental concepts of ideals and varieties and suggests some ways how these topics fit into CAGD. An excellent treatment of ideals and varieties and their application to CAGD can be found in [17].

15.2.1. Notation and terminology

In general, a polynomial in n variables x_1, \dots, x_n is defined

$$f(x_1, x_2, \dots, x_n) = \sum_{i=1}^{\tau} c_i x_1^{e_{1,i}} x_2^{e_{2,i}} \cdots x_n^{e_{n,i}}. \quad (15.3)$$

Each summand $c_i x_1^{e_{1,i}} x_2^{e_{2,i}} \cdots x_n^{e_{n,i}}$ is called a *term*, $x_1^{e_{1,i}} x_2^{e_{2,i}} \cdots x_n^{e_{n,i}}$ is a *monomial*, and c_i is the *coefficient* of the monomial. By convention, any given monomial occurs in at most one term in a polynomial.

$k[x_1, \dots, x_n]$ signifies the set of all polynomials in the variables x_1, \dots, x_n whose coefficients belong to a field k . For example, $R[x, y]$ is the set of all polynomials

$$\sum c_i x^{e_{1,i}} y^{e_{2,i}} \quad (15.4)$$

where $c_i \in R$ and $e_{1,i}, e_{2,i} \in \{0, 1, 2, \dots\}$. Thus, “ $f \in R[x, y, z]$ ” means that f is a polynomial whose variables are x, y and z and whose coefficients are real numbers. All polynomials in this chapter have coefficients that are real numbers.

It is often useful to list the terms of a polynomial in decreasing order, beginning with the *leading term*. This is done using a *term order* — a way to compare any two distinct terms of a polynomial and declare which is “greater”.

For linear polynomials, term order amounts to merely declaring an order on the variables. For example, the terms of the polynomial

$$2x + 3y - 4z$$

are in proper order if we declare $x > y > z$. If we declare $y > z > x$, the proper order would be $3y - 4z + 2x$. For non-linear polynomials, we begin by declaring an order on the variables, and then we must also choose one of several schemes that decide how the exponents in a polynomial influence term order. One such scheme is called *lexicographical order* (nicknamed *lex*), defined as follows. If the variables of a polynomial are ordered $x_1 > x_2 > \dots > x_n$, then given two distinct terms $T_i = c_i x_1^{e_{1,i}} x_2^{e_{2,i}} \cdots x_n^{e_{n,i}}$ and $T_j = c_j x_1^{e_{1,j}} x_2^{e_{2,j}} \cdots x_n^{e_{n,j}}$, $T_i > T_j$ if

1. $e_{1,i} > e_{1,j}$, or if
2. $e_{1,i} = e_{1,j}$ and $e_{2,i} > e_{2,j}$, or, in general, if
3. $e_{k,i} = e_{k,j}$ for $k = 1, \dots, m-1$ and $e_{m,i} > e_{m,j}$.

For example, the polynomial

$$3x^2y^2z + 4xy^3z^2 + 5x^3z + 6y^2 + 7xz^3 + 8$$

using lex with $x > y > z$ would be written $5x^3z + 3x^2y^2z + 4xy^3z^2 + 7xz^3 + 6y^2 + 8$ and its leading term is $5x^3z$. Using lex with $z > x > y$ it would be written $7z^3x + 4z^2xy^3 + 5zx^3 + 3zx^2y^2 + 6y^2 + 8$ and the leading term would be $7z^3x$. Or using lex with $y > z > x$ it would be written $4y^3z^2x + 3y^2zx^2 + 6y^2 + 7z^3x + 5zx^3 + 8$ and the leading term would be $4y^3z^2x$.

Another choice for term order is the *degree lexicographical order* (abbreviated *deglex*). If the variables are ordered $x_1 > x_2 > \dots > x_n$, then using deglex, $T_i > T_j$ if

1. $e_{1,i} + e_{2,i} + \dots + e_{n,i} > e_{1,j} + e_{2,j} + \dots + e_{n,j}$, or
2. $e_{1,i} + e_{2,i} + \dots + e_{n,i} = e_{1,j} + e_{2,j} + \dots + e_{n,j}$ and $T_i > T_j$ with respect to lex.

Using deglex with $x > y > z$, the terms of $3x^2y^2z + 4xy^3z^2 + 5x^3z + 6y^2 + 7xz^3 + 8$ would be ordered $4xy^3z^2 + 3x^2y^2z + 5x^3z + 7xz^3 + 6y^2 + 8$.

As observed in the lex and deglex examples, term orders ignore the coefficient of a term, so a term order might more properly be called a monomial order.

Other term orders can also be defined, such as degree reverse lexicographical order. The precise requirements for any term order are discussed in reference [6], page 18.

The n -dimensional real affine space is denoted R^n and is the set of n -tuples:

$$R^n = \{(a_1, \dots, a_n) : a_1, \dots, a_n \in R\} \quad (15.5)$$

15.2.2. Ideals and varieties

The *polynomial ideal* generated by $f_1, \dots, f_s \in k[x_1, \dots, x_n]$, denoted $\langle f_1, \dots, f_s \rangle$, is defined

$$\langle f_1, \dots, f_s \rangle = \{p_1f_1 + \dots + p_sf_s : p_i \in k[x_1, \dots, x_n]\}.$$

The polynomials f_1, \dots, f_s are called the *generators* of this ideal.

Consider a set of polynomials $f_1, f_2, \dots, f_s \in k[x_1, \dots, x_n]$. Let (a_1, \dots, a_n) be a point in k^n satisfying $f_i(a_1, \dots, a_n) = 0$, $i = 1, \dots, s$. The set of all such points (a_1, \dots, a_n) is called the *variety* defined by f_1, \dots, f_s , and is denoted by $V(f_1, \dots, f_s)$:

$$V(f_1, \dots, f_s) = \{(a_1, \dots, a_n) \in k^n \mid f_i(a_1, \dots, a_n) = 0, i = 1, \dots, s\}. \quad (15.6)$$

A variety defined by a single polynomial—called a *hypersurface*—is the most familiar instance of a variety. A hypersurface in R^2 is a planar curve defined using an implicit equation, and a hypersurface in R^3 is what is normally called an implicit surface in CAGD. For example, $V(x^2 + y^2 - 1)$ is a circle defined in terms of the implicit equation $x^2 + y^2 - 1 = 0$ and $V(2x + 4y - z + 1)$ is the plane whose implicit equation is $2x + 4y - z + 1 = 0$.

A variety $V(f_1, \dots, f_s)$ defined by more than one polynomial ($s > 1$) is the intersection of the varieties $V(f_1) \dots V(f_s)$.

15.2.3. Gröbner bases

It can be very useful to devise alternative generators for an ideal. Necessary and sufficient conditions for $\langle f_1, \dots, f_n \rangle = \langle g_1, \dots, g_m \rangle$ are $f_1, \dots, f_n \in \langle g_1, \dots, g_m \rangle$ and $g_1, \dots, g_m \in \langle f_1, \dots, f_n \rangle$.

A *Gröbner basis* of an ideal I is a set of polynomials $\{g_1, \dots, g_t\}$ such that the leading term of any polynomial in I is divisible by the leading term of at least one of the polynomials g_1, \dots, g_t . This, of course, requires that a term order be fixed for determining the leading terms: different term orders produce different Gröbner bases. Several excellent books have been written on Gröbner bases that do not presuppose that the reader has advanced training in mathematics [6,10,17]

A Gröbner basis is a particularly attractive set of generators for an ideal, as illustrated by two familiar examples. If $\{f_1, \dots, f_s\}$ are polynomials in one variable, the Gröbner basis of $\langle f_1, \dots, f_n \rangle$ consists of a single polynomial: the greatest common divisor (GCD) of f_1, \dots, f_s . If $\{f_1, \dots, f_s\}$ are linear polynomials in several variables, the Gröbner basis is an uppertriangular form of a set of linear equations. The Gröbner basis of these special cases provides significant computational advantage and greater insight, and the same is true of the Gröbner basis of a more general ideal.

Gröbner bases are the fruit of Bruno Buchberger's Ph.D. thesis [12], and are named in honor of his thesis advisor. Buchberger devised an algorithm for computing Gröbner bases [13,17,18]. Also, commercial software packages such as Maple and Mathematica include capabilities for computing Gröbner bases.

15.3. RESULTANTS

Given a set of polynomials $\{f_1, \dots, f_s\}$, a resultant is a polynomial expression in the coefficients of f_1, \dots, f_s such that the vanishing of the resultant is a necessary and sufficient condition for $V(f_1, \dots, f_s)$ to be non-empty [16]. Thus, a resultant determines whether or not $V(f_1, \dots, f_s)$ is empty without explicitly computing the variety. Gröbner basis methods can also be used for this task. However, resultants are usually more efficient than Gröbner bases in practical applications.

Resultants play an important role in elimination theory—a systematic approach for finding polynomials in an ideal that do not contain as many variables as generic elements of the ideal. Various formulations for resultants were extensively studied in the late 19th century and the early 20th century [31,48]. The main idea is to identify a (possibly large) set of n linearly independent polynomials that generate the ideal and that contain n terms. Then each term can be used as an unknown and the theory of linear system of equations can be applied. In practice, the resultants for two univariate polynomials and for three bivariate polynomials are of most interest.

15.3.1. Sylvester's resultant

Consider the two polynomials

$$\begin{cases} f(t) = a_n t^n + a_{n-1} t^{n-1} + \dots + a_1 t + a_0, & a_n \neq 0 \\ g(t) = b_m t^m + b_{m-1} t^{m-1} + \dots + b_1 t + b_0, & b_m \neq 0 \end{cases}$$

Using Sylvester's dialytic method, we multiply $f(t)$ by $t^{m-1}, t^{m-2}, \dots, t, 1$ and $g(t)$ by $t^{n-1}, \dots, t, 1$, arriving at $m+n$ polynomials, which can be arranged in matrix form:

$$\begin{bmatrix} t^{m-1}f(t) \\ \vdots \\ tf(t) \\ f(t) \\ t^{n-1}g(t) \\ \vdots \\ tg(t) \\ g(t) \end{bmatrix} = \begin{bmatrix} a_n & a_{n-1} & \dots & a_1 & a_0 & & & \\ & a_n & a_{n-1} & \dots & a_1 & a_0 & & \\ & & \dots & \dots & & & & \\ & & & & a_n & \dots & \dots & a_0 \\ b_m & b_{m-1} & \dots & b_1 & b_0 & & & \\ & b_m & b_{m-1} & \dots & b_1 & b_0 & & \\ & & \dots & \dots & & & & \\ & & & & b_m & \dots & \dots & b_0 \end{bmatrix} \begin{bmatrix} t^{m+n-1} \\ t^{m+n-2} \\ \vdots \\ \vdots \\ \vdots \\ t \\ 1 \end{bmatrix}$$

The determinant of the coefficient matrix is known as Sylvester's resultant for $f(t)$ and $g(t)$. The ideal generated by these $m+n$ polynomials is just the ideal $\langle f(t), g(t) \rangle$. Thus they have the same variety. When there exists one value of t in the variety, Sylvester's resultant must vanish.

Sylvester's resultant can also be derived using a method invented by Euler. Euler introduced two polynomials $h(t)$ of degree $m-1$ and $k(t)$ of degree $n-1$ with coefficients

undetermined. Letting $h(t)f(t) - k(t)g(t) = 0$ leads to $m+n$ linear equations with $m+n$ unknowns which are the coefficients of $h(t)$ and $k(t)$. The determinant of the coefficient matrix of these $m+n$ linear equations is exactly Sylvester's resultant. Obviously the determinant vanishes if and only if there exist nonzero polynomials $h(t)$ of degree not greater than $m-1$ and $k(t)$ of degree not greater than $n-1$ such that $h(t)f(t) = k(t)g(t)$ holds. This is equivalent to the existence of the common roots of polynomials $f(t)$ and $g(t)$. Therefore Euler's method shows that $R(f, g) = 0$ is not only the necessary but also sufficient condition for $f(t)$ and $g(t)$ to have common roots.

15.3.2. Bezout's resultant

Another popular resultant formulation for two univariate polynomials is Bezout's resultant. A nice derivation of Bezout's resultant is due to Cayley. Without loss of generality, we assume the degree of the polynomial $g(t)$ is less than the degree of $f(t)$, i.e., $m \leq n$. Construct a symmetric function

$$\delta(t, s) = \begin{vmatrix} f(t) & f(s) \\ g(t) & g(s) \end{vmatrix} / (s - t) = \frac{f(t)g(s) - f(s)g(t)}{s - t}.$$

Some algebraic manipulation shows that $\delta(t, s)$ is a degree $n-1$ polynomial in s whose coefficients are polynomials of t :

$$\begin{aligned} \delta(t, s) &= \frac{f(t)(g(t) - g(s))}{(t - s)} - \frac{g(t)(f(t) - f(s))}{(t - s)} \\ &= \sum_{k=0}^{m-1} \left(f(t) \sum_{i=k+1}^m b_i t^{i-k-1} - g(t) \sum_{i=k+1}^n a_i t^{i-k-1} \right) s^k - \sum_{k=m}^{n-1} \left(g(t) \sum_{i=k+1}^n a_i t^{i-k-1} \right) s^k. \end{aligned}$$

The variety of the ideal generated by these n polynomials is the same as $V(\langle f(t), g(t) \rangle)$. Write these polynomials in matrix form:

$$\begin{bmatrix} f \sum_{i=1}^m b_i t^{i-1} - g \sum_{i=1}^n a_i t^{i-1} \\ \vdots \\ f b_m - g \sum_{i=m}^n a_i t^{i-m} \\ -g \sum_{i=m+1}^n a_i t^{i-m-1} \\ \vdots \\ -g a_n \end{bmatrix} = \begin{bmatrix} c_{00} & \cdots & c_{0,n-1} \\ \vdots & \ddots & \vdots \\ c_{n-1,0} & \cdots & c_{n-1,n-1} \end{bmatrix} \begin{bmatrix} 1 \\ t \\ \vdots \\ t^{n-1} \end{bmatrix} \quad (15.7)$$

with the entry $c_{ij} = \sum_{\substack{k < \min(i,j) \\ k+h=i+j+1}} (a_k b_h - a_h b_k)$ and the convention that $b_{m+1} = \cdots = b_n = 0$.

This coefficient matrix is called Bezout's matrix. If $V(\langle f, g \rangle)$ is non-empty, the determinant of Bezout's matrix must vanish. The converse is also true when $n = m$, the proof of which can be found in [22,26]. The determinant is therefore a resultant for f and g , known as Bezout's resultant. In general, Bezout's resultant has dimension $n \times n$ while Sylvester's resultant has dimension $(n+m) \times (n+m)$.

When $n > m$, Bezout's determinant has an extraneous factor of a_n^{n-m} . This extraneous factor can be removed by modifying Bezout's resultant as follows [19]: The first m

polynomials are the same as in (15.7); and the remaining $n - m$ polynomials are obtained from $t^{i-m}g(t)$, $i = m, \dots, n$. Thus $c_{ij} = b_{i+j-m}$ for $i \geq m, 0 \leq j \leq n - 1$. Look at the example of $f(t) = a_2t^2 + a_1t + a_0$ and $g(t) = b_1t + b_0$. The original Bezout's determinant is

$$\begin{vmatrix} a_0b_1 - a_1b_0 & -a_2b_0 \\ -a_2b_0 & -a_2b_1 \end{vmatrix} = a_2 \begin{vmatrix} a_0b_1 - a_1b_0 & -a_2b_0 \\ -b_0 & -b_1 \end{vmatrix}$$

and the variant of Bezout's resultant is $\begin{vmatrix} a_0b_1 - a_1b_0 & -a_2b_0 \\ b_0 & b_1 \end{vmatrix}$.

15.3.3. Dixon's resultant

Cayley's formulation can be extended to the case of three bivariate polynomials. Consider three polynomials:

$$f(s, t) = \sum_{i=0}^n \sum_{j=0}^m a_{ij} s^i t^j, \quad g(s, t) = \sum_{i=0}^n \sum_{j=0}^m b_{ij} s^i t^j, \quad h(s, t) = \sum_{i=0}^n \sum_{j=0}^m c_{ij} s^i t^j$$

Dixon observed that the expression

$$\delta(s, t, \alpha, \beta) = \begin{vmatrix} f(s, t) & g(s, t) & h(s, t) \\ f(s, \beta) & g(s, \beta) & h(s, \beta) \\ f(\alpha, \beta) & g(\alpha, \beta) & h(\alpha, \beta) \end{vmatrix} / (s - \alpha)(t - \beta)$$

is actually a polynomial of degree $2n - 1, m - 1, n - 1$ and $2m - 1$ in s, t, α, β respectively. Thus it can be written as $\delta(s, t, \alpha, \beta) = \sum_{ijkl} d_{ijkl} s^i t^j \alpha^k \beta^l$ where d_{ijkl} are expressions in a_{ij}, b_{ij} and c_{ij} .

For any $(s, t) \in V(\langle f, g, h \rangle)$, $\delta(s, t, \alpha, \beta)$ vanishes no matter what α and β are. Thus the coefficients of each $\alpha^k \beta^l$ must vanish at these (s, t) pairs. This gives $2mn$ polynomials, each of which has $2mn$ terms in s and t since s has degree $2n - 1$ and t is degree $m - 1$. The determinant of the coefficient matrix from these polynomials serves as a resultant for f, g and h , called Dixon's resultant [23].

Consider the example:

$$\begin{cases} f(t) &= a_{21}s^2t + a_{11}st + a_{01}t + a_{20}s^2 + a_{10}s + a_{00} \\ g(t) &= b_{21}s^2t + b_{11}st + b_{01}t + b_{20}s^2 + b_{10}s + b_{00} \\ h(t) &= c_{21}s^2t + c_{11}st + c_{01}t + c_{20}s^2 + c_{10}s + c_{00} \end{cases}$$

Dixon's method gives

$$\delta(s, t, \alpha, \beta) = [1, \alpha, \beta, \alpha\beta] \begin{bmatrix} (00, 01, 10) & (00, 01, 20) & (00, 21, 10) & (00, 21, 20) \\ & +(00, 11, 10) & +(00, 11, 20) & \\ (00, 01, 11) & (00, 01, 21) & (01, 21, 10) & (01, 21, 20) \\ & +(01, 11, 10) & +(01, 11, 20) & \\ (00, 01, 20) & (10, 01, 20) & (00, 21, 20) & (10, 21, 20) \\ & +(00, 11, 20) & +(10, 11, 20) & \\ (00, 01, 21) & (00, 11, 21) & (01, 21, 20) & (11, 21, 20) \\ & +(10, 01, 21) & +(10, 11, 21) & \end{bmatrix} \begin{bmatrix} 1 \\ s \\ s^2 \\ s^3 \end{bmatrix}$$

where (ij, kj, pq) stands for the 3×3 determinant $(ij, kj, pq) = \begin{vmatrix} a_{ij} & a_{kl} & a_{pq} \\ b_{ij} & b_{kl} & b_{pq} \\ c_{ij} & c_{kl} & c_{pq} \end{vmatrix}$.

15.4. CURVE IMPLICITIZATION AND INVERSION

The algebraic tools of Gröbner bases and resultants empower us to solve several problems of interest to CAGD. This section looks at several examples of implicitization and inversion.

It is known from classical algebraic geometry that any degree n polynomial or rational parametric curve can be represented exactly using a degree n algebraic equation. For example, a circle can be expressed using the parametric equation

$$x = (1 - t^2)/(t^2 + 1), \quad y = 2t/(t^2 + 1) \quad (15.8)$$

or using the implicit equation $x^2 + y^2 - 1 = 0$. In the following we discuss three approaches for the conversion from the parametric equation to the implicit equation.

15.4.1. Resultant-based method

We have presented the resultant tool for determining whether two polynomials have a common root. We now apply that tool to converting the parametric equation of a curve given by (15.1) into an implicit equation of the form $f(x, y) = 0$.

We proceed by forming two auxiliary polynomials:

$$g(x, t) = c(t)x - a(t), \quad h(y, t) = c(t)y - b(t)$$

View $g(x, t)$ as a polynomial in t whose coefficients are linear in x , and view $h(y, t)$ as a polynomial in t whose coefficients are linear in y . If we compute the resultant of $g(x, t)$ and $h(y, t)$, we do not arrive at a numerical value, but rather a polynomial in x and y which we call $f(x, y)$. Note that $g(x, t) = h(y, t) = 0$ only for values of x, y and t which satisfy the relationships $x = a(t)/c(t)$, $y = b(t)/c(t)$. Clearly, for these values of x, y and t , the resultant $f(x, y)$ must vanish. Conversely, any (x, y) pair for which $f(x, y) = 0$, causes the resultant of g and h to be zero. But, if the resultant is zero, then we know that there exists a value of t for which $g(x, t) = h(y, t) = 0$. In other words, all (x, y) for

which $f(x, y) = 0$ lie on the parametric curve and therefore $f(x, y) = 0$ is the implicit equation of that curve.

As discussed, the resultant is the determinant of a matrix of coefficients for a set of polynomials. Inversion—computing the parameter t for a point (x, y) known to lie on the curve—can be performed by solving such a set of polynomial equations using Gauss elimination or Cramer's rule.

We illustrate with the circle parametrized by (15.8). We have

$$g = (x + 1)t^2 + (x - 1), \quad h = yt^2 - 2t + y$$

Using Sylvester's resultant, we obtain a 4×4 determinant

$$f(x, y) = \begin{vmatrix} x+1 & 0 & x-1 & 0 \\ 0 & x+1 & 0 & x-1 \\ y & -2 & y & 0 \\ 0 & y & -2 & y \end{vmatrix} = 4(x^2 + y^2 - 1).$$

Bezout's resultant provides a 2×2 determinant

$$f(x, y) = \begin{vmatrix} -2x+2 & -2y \\ -2y & 2x+2 \end{vmatrix} = -4(x^2 + y^2 - 1).$$

We could obtain an inversion equation by solving the equations:

$$\begin{bmatrix} -2x+2 & -2y \\ -2y & 2x+2 \end{bmatrix} \begin{bmatrix} 1 \\ t \end{bmatrix} = 0,$$

from which $t = y/(x + 1)$ or $t = (1 - x)/y$.

Two remarks should be made. First, if $a(t)$, $b(t)$ and $c(t)$ in (15.1) are not relatively prime, the common factor should be removed before the resultant method is applied. Otherwise, the resultant will be identically zero, containing no information about the curve, since p and q have always common solutions for arbitrary (x, y) pair. Second, if the degrees of $p(x, t)$, $q(y, t)$ with respect to t are not the same, the variant of Bezout's resultant is used.

15.4.2. Gröbner basis technique

In order to use Gröbner basis method for implicitizing a rational parametric curve defined by (15.1) with $GCD(a(t), b(t), c(t)) = 1$ (otherwise, the common factor can be removed), we define the ideal

$$I = \langle c(t)x - a(t), c(t)y - b(t) \rangle \subset R[x, y, t]. \quad (15.9)$$

If $f(x, y) = 0$ is the implicit equation of (15.1), then $f \in I \cap R[x, y]$. To guarantee f appears in the Gröbner basis, we order the variables $t > x > y$, and then construct the Gröbner basis with the lexicographic ordering for the ideal I . The lexicographic ordering results in a Gröbner basis that has a triangular structure. Thus the Gröbner basis obtained will contain the curve's implicit form—an element which does not involve t , and an inversion—an element which is linear in t .

In the example of the circle (15.8), $I = \langle (1 + t^2)x - (1 - t^2), (1 + t^2)y - 2t \rangle$. Using the computer algebra system MAPLE, we obtain the Gröbner basis $I = \langle -y + x + tx, x - 1 + yt, y^2 + x^2 - 1 \rangle$. Therefore the polynomial $y^2 + x^2 - 1$ gives the implicit equation $x^2 + y^2 - 1 = 0$. The two other polynomials $-y + x + tx$ and $x - 1 + yt$ are linear in t and thus provide the inversion $t = y/(1 + x)$ or $t = (1 - x)/y$.

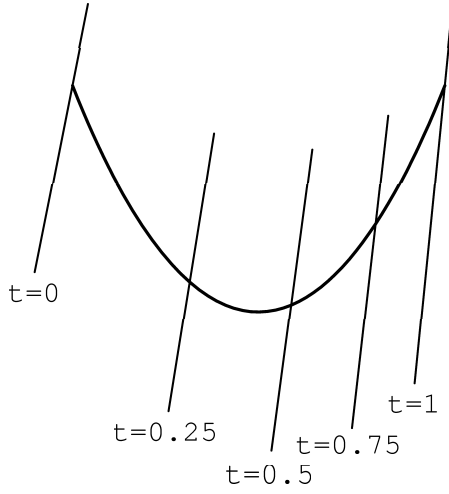


Figure 15.1. Parametric Curve and Moving Lines

15.4.3. Moving curve technique

A recent development in impliciting a planar rational parametric curve is the “moving curve” method [19,42,44]. A *moving curve* is defined as

$$C(x, y, w; t) := \sum_{i=0}^m f_i(x, y, w)t^i$$

where $f_i(x, y, w)$ is a homogeneous polynomial of degree d . Thus $C(x, y, w; t) = 0$ is a family of algebraic curves, with one curve corresponding to each t . In particular, when $d = 1$, $C(x, y, w; t) = 0$ is a family of implicitly defined lines. Therefore we call it a *moving line* of degree m . Likewise, $C(x, y, w; t) = 0$ is called a *moving conic* of degree m when $d = 2$. A moving curve $C(x, y, w; t) = 0$ is said to *follow* a planar rational curve (15.1) if $C(a(t), b(t), c(t); t)$ is identically zero, that is, if for all values of t , the point $(a(t)/c(t), b(t)/c(t))$ lies on the moving curve $C(x, y, w; t) = 0$. For example, each row of Bezout’s matrix or Sylvester’s matrix corresponds to a moving line following the curve. Figure 15.1 illustrates a parametric curve and a few moving lines.

The moving curve technique identifies $m + 1$ independent moving curves that follow a given rational curve. A square matrix can then be formed from the coefficients of these moving curves with respect to t and the determinant of the matrix gives the desired implicit equation. In general, such a collection of moving curves can be found by solving a set of linear equations [43]. For example, a degree $n - 1$ moving line

$$C(x, y, w; t) = \sum_{i=0}^{n-1} (A_i x + B_i y + C_i w) t^i = 0 \quad (15.10)$$

follows a rational curve (15.1) if

$$\sum_{i=0}^{n-1} (A_i a(t) + B_i b(t) + C_i c(t)) t^i \equiv 0$$

which can be expressed as $2n$ linear equations with $3n$ unknowns

$$\begin{bmatrix} a_0 & b_0 & c_0 & 0 & \cdots & 0 & 0 & 0 \\ a_1 & b_1 & c_1 & a_0 & \cdots & 0 & 0 & 0 \\ a_2 & b_2 & c_2 & a_1 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_n & b_n & c_n & a_{n-1} & \cdots & 0 & 0 & 0 \\ 0 & 0 & 0 & a_n & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & a_{n-1} & b_{n-1} & c_{n-1} \\ 0 & 0 & 0 & 0 & \cdots & a_n & b_n & c_n \end{bmatrix} \begin{bmatrix} A_0 \\ B_0 \\ C_0 \\ \vdots \\ \vdots \\ A_{n-1} \\ B_{n-1} \\ C_{n-1} \end{bmatrix} = 0. \quad (15.11)$$

where a_i, b_i, c_i are the coefficients of the polynomials $a(t), b(t), c(t)$. Thus solving the equations for A_i, B_i, C_i yields n linearly independent moving lines of degree $n - 1$ that follows the curve.

Recently the problem of finding an appropriate set of moving lines was illuminated by the description of the μ -basis [19]. A μ -basis for a degree- n planar rational curve consists of two moving lines $p(x, y; t)$ and $q(x, y; t)$, of degree μ and $n - \mu$ respectively, which form an ideal basis for all moving lines that follow the curve. An efficient method of computing the μ -basis is given in [52]. Once the μ -basis of a curve is known, the problem of finding $m + 1$ linearly independent moving curves is greatly simplified. For example, consider the degree four curve:

$$x = \frac{t^4 + 2t^3 + t^2 + t + 1}{-t^4 - 2t^2 - 2}, \quad y = \frac{-t^3 + t^2 + 2t}{-t^4 - 2t^2 - 2}.$$

This curve has a μ -basis of

$$p = (x + y + 1)t^2 + t + y, \quad q = (x + 1)t^2 + 2t + 2x + y + 1.$$

Thus four moving lines of degree 3 are $p, t p, q, t q$. Moreover, two moving conics of degree 1 can be obtained by taking Bezout's resultant of p and q . Each row in Bezout's matrix corresponds to a moving conic. Therefore the implicit equation can be expressed as a 2×2 determinant whose elements are quadratic in x and y . In general, for a degree n rational curve, using a variant of Bezout's resultant on the μ -basis, we can write the implicit equation of the rational curve as the determinant of an $(n - \mu) \times (n - \mu)$ matrix with μ rows whose elements are quadratic in x and y , and the remaining $n - 2\mu$ rows with elements linear in x and y , while conventional implicitization methods generate the determinant of an $n \times n$ matrix [22,34].

15.5. CURVE PARAMETRIZATION

15.5.1. Planar algebraic curves

Implicitization shows that a degree n parametric curve can be represented using a degree n implicit equation. Any implicit equation that can be obtained by implicitizing a parametric curve is said to be *rational* (in other words, a rational curve is any curve which can be parametrized using rational functions). All algebraic quadratic curves have rational

quadratic parametrizations, whereas algebraic curves of degree greater than two are not generally rational.

A degree n algebraic curve is defined by a degree n polynomial $f(x, y)$, which has $(n+1)(n+2)/2$ terms. For example, a degree two algebraic curve has six terms: $a_1x^2 + a_2xy + a_3y^2 + a_4x + a_5y + a_6 = 0$. However, any one coefficient can be specified by scaling all of the other coefficients. (For example, the coefficient of x^2 in the quadratic example can be set to 1 by dividing all the coefficients by a_1). Thus, there is an $(n+1)(n+2)/2 - 1 = n(n+3)/2$ dimensional family of degree n algebraic curves. Geometrically, this means a degree n algebraic curve can be forced to interpolate $n(n+3)/2$ points in general position.

The parametric equation of a rational degree n curve has $3(n+1)$ coefficients. However, any one of these can be specified by scaling all of the other coefficients. Also, three other coefficients can be specified by changing the parametrization by a rational linear transformation $s = \frac{at+b}{ct+d}$, and consequently there is a $3n - 1$ dimensional family of degree n rational curves.

15.5.2. Genus and rationality

The condition under which an implicit algebraic curve can be parametrized using rational polynomials is that its *genus* must be zero [49]. Basically, the genus of a curve is given by the formula $g = \frac{(n-1)(n-2)}{2} - d$ where g is the genus, n is the degree, and d is the number of double points. There are some subtleties involved in this equation if the singularities are not simple double points, but we will not concern ourselves with them.

A double point on a curve is a point for which $f(x, y) = f_x(x, y) = f_y(x, y) = 0$ where the subscripts x and y denote partial differentiations, and for a point of multiplicity k , all partials up to order $k - 1$ vanish. Geometrically, a double point means that any straight line through it intersects the curve at least twice at this point.

We see immediately that all curves of degree one and two have genus zero and thus can be parametrized using rational polynomials. A degree three algebraic curve is rational only if it has a double point.

An irreducible curve is one whose implicit equation $f(x, y) = 0$ cannot be factored. Rational curves (that can be parametrized using a single parametric equation) are irreducible, and an irreducible curve of degree n can have at most $(n-1)(n-2)/2$ double points. Thus, a rational curve has the most double points possible for a curve of its degree.

15.5.3. Parametrizing curves

One way to parametrize a degree two algebraic curve is to transform the conic into the standard form which has already a parametrization, and then to transform the standard parametrization back. The standard equations of conics are: $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ for an ellipse, $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$ for a hyperbola, and $y^2 = 2px$ for a parabola. They can be parametrized as $(a\frac{1-t^2}{1+t^2}, b\frac{2t}{1+t^2})$, $(a\frac{1+t^2}{1-t^2}, b\frac{2t}{1-t^2})$, and $(\frac{t^2}{2p}, t)$, respectively. Therefore the main step is to find a nonsingular affine coordinate transformation. This can be carried out as follows.

Suppose the conic equation is $Ax^2 + 2Bxy + Cy^2 + 2Dx + 2Ey + F = 0$. First we convert the equation into the form

$$\bar{A}\bar{x}^2 + \bar{C}\bar{y}^2 + 2\bar{D}\bar{x} + 2\bar{E}\bar{y} + \bar{F} = 0. \quad (15.12)$$

If $B = 0$, it is done. If $A = C = 0$ and $B \neq 0$, then set $x = \bar{x} + \bar{y}$ and $y = \bar{x} - \bar{y}$.

Otherwise, if $A \neq 0, B \neq 0$, then

$$A(x + \frac{B}{A}y)^2 + (C - \frac{B^2}{A})y^2 + 2Dx + 2Ey + F = 0.$$

Thus we only need to set $\bar{x} = x + \frac{B}{A}y$ and $\bar{y} = y$. A similar transformation can be derived if $C \neq 0$ and $B \neq 0$.

Second, if one of \bar{A} and \bar{C} in (15.12) is zero, the curve is a parabola. Assume that $\bar{A} \neq 0$ and $\bar{C} = 0$. Then $\bar{A}(\bar{x} + \frac{\bar{D}}{\bar{A}})^2 = -2\bar{E}\bar{y} - \bar{F} + \frac{\bar{D}^2}{\bar{A}}$. Thus setting

$$x' = \bar{x} + \frac{\bar{D}}{\bar{A}}, \quad y' = \bar{y} + \frac{F - D^2/A}{2\bar{E}}$$

arrives at the standard parabola equation. A similar process deals with the case of $\bar{A} = 0, \bar{C} \neq 0$. If both of \bar{A} and \bar{C} are nonzero, we have $\bar{A}(\bar{x} + \frac{\bar{D}}{\bar{A}})^2 + \bar{C}(\bar{y} + \frac{\bar{E}}{\bar{C}})^2 = \frac{\bar{D}^2}{\bar{A}} + \frac{\bar{E}^2}{\bar{C}} - \bar{F}$. Therefore we can take a translation of $x' = \bar{x} + \frac{\bar{D}}{\bar{A}}$ and $y' = \bar{y} + \frac{\bar{E}}{\bar{C}}$ to make the curve equation in the standard form. Composing the transformations in the above two steps gives the required coordinate transformation.

Another method to parametrize a conic is to establish a one-one correspondence between points on the curve and a family of lines through a point on the curve, which is called a *pencil-of-lines*. This pencil-of-lines method is most easily illustrated by translating the curve so that it passes through the origin, such as does the curve

$$x^2 - 2x + 4y^2 = 0$$

which is an ellipse centered at $(1, 0)$. We next make the substitution $y = tx$ and solve for x as a function of t : $x^2(1 + 4t^2) - 2x = 0$. Then

$$x = \frac{2}{1 + 4t^2}, \quad y = tx = \frac{2t}{1 + 4t^2}$$

Notice that $y = tx$ is a family of lines through the origin. The variable line $y = tx$ intersects the curve once at the origin, and at exactly one other point (because of Bezout's theorem: two algebraic curves of degree m and n intersect at either mn points or else they have common components [49]). Thus, we have established a one-one correspondence between points on the curve and values t which correspond to lines containing that point and the origin. The ellipse parametrized in this manner is shown in Figure 15.2.

To parametrize a genus zero cubic curve, one must first find its double point, which is done by solving $h(x, y) = h_x(x, y) = h_y(x, y) = 0$. Once the location of the double point is determined, one can translate the curve so that the double point lies on the origin. Then the same trick in the pencil-of-lines approach for conics can be played with this cubic curve since the curve now has an equation involving terms of degree two and three only.

Consider this example of the cubic curve

$$f(x, y) = -21 + 46x - 13x^2 + x^3 + 25y - 23xy + 3x^2y - 9y^2 + 3xy^2 + y^3 = 0$$

for which

$$f_x(x, y) = 46 - 26x + 3x^2 - 23y + 6xy + 3y^2$$

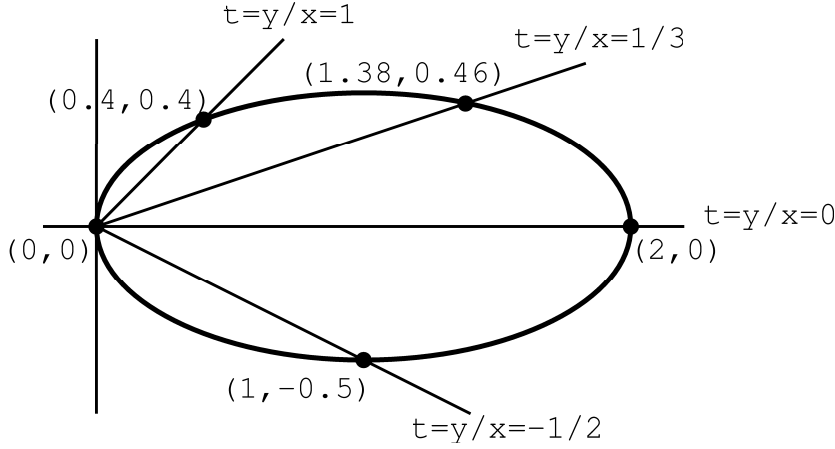


Figure 15.2. Parametrizing an Ellipse

and

$$f_y(x, y) = 25 - 23x + 3x^2 - 18y + 6xy + 3y^2$$

We compute the x coordinates of the intersections of $f_x = 0$ and $f_y = 0$ by taking the resultant of f_x and f_y with respect to y :

$$\text{Resultant}(f_x, f_y, y) = 174 - 159x + 36x^2$$

whose roots are $x = 2$ and $x = \frac{29}{12}$. Likewise the y coordinates of the intersections of $f_x = 0$ and $f_y = 0$ are found by taking the resultant of f_x and f_y with respect to x :

$$\text{Resultant}(f_x, f_y, x) = 297 - 207y + 36y^2$$

whose roots are $y = 3$ and $y = \frac{11}{4}$. From these clues, we find that the only values of (x, y) which satisfy $f(x, y) = f_x(x, y) = f_y(x, y) = 0$ are $(x, y) = (2, 3)$. This is therefore the double point.

The double point can also be found by computing the Gröbner basis of $\langle f, f_x, f_y \rangle$ using lex ordering with $x > y$, the Gröbner basis is $\{x - 2, y - 3\}$.

This curve can be parametrized by translating the implicit curve so that the double point lies at the origin. This is done by making the substitution $x = \bar{x} + 2$, $y = \bar{y} + 3$, yielding

$$2\bar{x}^2 + \bar{x}^3 + 7\bar{x}\bar{y} + 3\bar{x}^2\bar{y} + 6\bar{y}^2 + 3\bar{x}\bar{y}^2 + \bar{y}^3 = 0$$

Parametrization is then performed using the method discussed earlier in this section,

$$\bar{x} = -\frac{6t^2 + 7t + 2}{t^3 + 3t^2 + 3t + 1}; \quad \bar{y} = -\frac{6t^3 + 7t^2 + 2t}{t^3 + 3t^2 + 3t + 1}$$

and the parametrized curve is translated back so that the doubled point is again at $(2, 3)$ (see Figure 15.3):

$$x = -\frac{6t^2 + 7t + 2}{t^3 + 3t^2 + 3t + 1} + 2 = \frac{2t^3 - t}{t^3 + 3t^2 + 3t + 1};$$

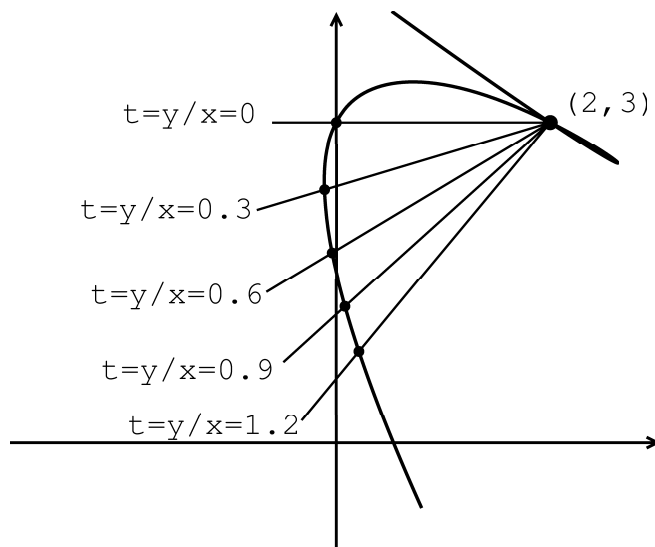


Figure 15.3. Parametrizing a Cubic Curve

$$y = -\frac{6t^3 + 7t^2 + 2t}{t^3 + 3t^2 + 3t + 1} + 3 = \frac{-3t^3 + 2t^2 + 7t + 3}{t^3 + 3t^2 + 3t + 1}$$

For a general algebraic curve, the parametrization problem involves two steps: determine whether it admits a rational parametric representation, and find one if so. The algorithms, in general, are not as simple as for conics or cubics. References [1–4,47] provide various computational techniques for parametrizing algebraic curves.

15.6. INTERSECTION COMPUTATIONS

We now consider how to compute the points at which two curves intersect. Intersection algorithms for two Bézier curves are commonly based on subdivision, or using some numerical algorithms such as a multivariate Newton method. The former takes advantage of the properties of Bézier or B-spline representations and focuses on the intersection points within the specified intervals. The latter method is not robust: it is difficult or impossible to assure that all intersection points have been found.

Algebraic methods provide a systematic way for computing intersections. As noted in Section 15.2, a variety $V(f_1, \dots, f_s)$ defined by more than one polynomial ($s > 1$) is the intersection of the varieties $V(f_1), \dots, V(f_s)$. Therefore, intersection computation is equivalent to determining a variety. Using either resultants or Gröbner bases, this finally reduces to the problem of finding the roots of a polynomial on one variable.

15.6.1. Parametric curve and implicit curve

Given one curve defined by the parametric equation $(x, y) = (x(t), y(t))$ and a second curve defined by the implicit equation $f(x, y) = 0$, we replace all occurrences of x and y in the implicit equation by $x(t)$ and $y(t)$, respectively. These substitutions create a polynomial $f(x(t), y(t)) = g(t)$ whose roots are the parameter values of the intersection points. The (x, y) coordinates of these intersection points can be easily obtained by substituting the parametric values into the equation of the parametric curve.

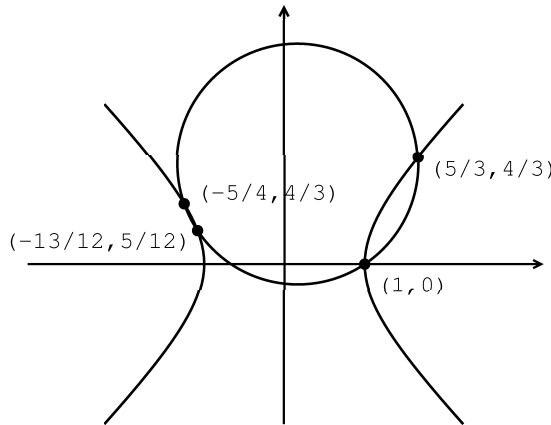


Figure 15.4. Circle and Hyperbola

15.6.2. Implicit curve and implicit curve

For two algebraic curves $f_1(x, y) = 0$ and $f_2(x, y) = 0$, the problem of computing the intersection amounts to computing the variety $V(f_1, f_2)$.

One direct method for computing the variety $V(f_1, f_2)$ is to take the resultant of f_1 and f_2 with respect to x or y . The x -resultant is computed by treating f_1 and f_2 as polynomials in x whose coefficients are polynomials in y . The x -resultant eliminates x and produces a polynomial in y whose roots are the y coordinates of the intersection points.

We illustrate with a circle $6x^2 + 6y^2 - 2x - 15y - 4 = 0$ and a hyperbola $x^2 - y^2 - 1 = 0$ (see Figure 15.4).

The x -resultant of these two implicit equations is $144y^4 - 360y^3 + 269y^2 - 60y$ whose roots are $y = 0$, $y = 4/3$, $y = 3/4$, and $y = 5/12$. These are the y -coordinates of the points of intersection of the two curves.

We can use the y -resultant to find the x -coordinates of the points of intersection. The y -resultant is $144x^4 - 48x^3 - 461x^2 + 40x + 325$ which has roots $x = 1$, $x = 5/3$, $x = -5/4$, and $x = -13/12$.

We now know the x and y components of the points of intersection, but we don't know which x goes with which y ! One way to determine that is simply to evaluate each curve equation with every x and every y to see which (x, y) pairs satisfy both curve equations simultaneously. A more clever way is to use Euclid's algorithm which computes the GCD of two polynomials. In fact, Euclid's algorithm spares us the trouble of computing both the x -resultant and the y -resultant.

Suppose we had only computed the y -resultant and we wanted to find the y -coordinate of the point of intersection whose x -coordinate is $5/3$. That is to say, we want to find a point $(\frac{5}{3}, y)$ which satisfies both curve equations. We substitute $x = 5/3$ into the circle equation to get $16/6 - y^2 = 0$ and into the hyperbola equation to get $6y^2 - 15y + 28/3 = 0$. We now simply want to find a value of y which satisfies both of these equations. Euclid's algorithm tells us that the GCD of these two is $3y - 4 = 0$, and thus one point of intersection is $(\frac{5}{3}, \frac{4}{3})$.

Gröbner bases also provide a systematic computational method with the assurance

that all intersection points have been found. The general strategy is based on the simple observation that $V(f_1, \dots, f_s) = V(\langle f_1, \dots, f_s \rangle)$. Consequently, if f_1, \dots, f_s and g_1, \dots, g_t are generators of the same ideal, then

$$V(f_1, \dots, f_s) = V(g_1, \dots, g_t).$$

Consider the previous example of a circle $V(6x^2 + 6y^2 - 2x - 15y - 4)$ and a hyperbola $V(x^2 - y^2 - 1)$. It can be verified that

$$\begin{aligned} \langle 6x^2 + 6y^2 - 2x - 15y - 4, x^2 - y^2 - 1 \rangle = \\ \langle -12x^2 + 2x + 15y + 10, 144x^4 - 48x^3 - 461x^2 + 40x + 325 \rangle \quad (\text{the Gröbner basis}) \end{aligned}$$

Since any point of intersection must be zeros of all generators of the ideal, the only possible x -coordinates for the intersection points must be roots of $144x^4 - 48x^3 - 461x^2 + 40x + 325 = 0$ (the roots are $\frac{5}{3}$, 1 , $-\frac{13}{12}$, and $-\frac{5}{4}$). The corresponding y coordinates can then be solved using $-12x^2 + 2x + 15y + 10 = 0$.

15.6.3. Parametric curve and parametric curve

If we begin with two parametric curves, we can first implicitize one of them, and then use the substitution method to compute the intersection points. We illustrate this process by intersecting the curve

$$(x, y) = \left(-\frac{1+s^2}{1-s^2}, \frac{2s}{1-s^2} \right)$$

with the curve

$$(x, y) = \left(\frac{27t^2 - 239t - 116}{6t^2 + 81t + 165}, \frac{9t^2 + 96t - 33}{6t^2 + 81t + 165} \right)$$

(see Figure 15.5). The two curves intersect four times, which is the most that two quadratic curves can intersect. We implicitize the first curve and get the implicit equation $x^2 - y^2 - 1 = 0$. Substituting the parametric equation of the second curve into this implicit equation and clearing the denominator, we arrive at the intersection equation:

$$18t^4 - 459t^3 + 991t^2 + 1031t - 437 = 0.$$

We now compute the roots of this degree four polynomial, which are 23 , $1/3$, $19/6$ and -1 . These are the parametric values on the second curve for the intersection points. From the parametric equation of the second curve, the corresponding (x, y) coordinates can be easily found: $(5/3, 4/3)$, $(-1, 0)$, $(-5/4, 3/4)$, $(5/3, -4/3)$. The parametric values on the first curve for the intersection points can be found from the inversion formulas $s = y/(1-x)$ and $s = -(1+x)/y$. They are -2 , 0 , $1/3$ and 2 , respectively.

Tests indicate that this implicitization-based intersection algorithm is several times faster than subdivision methods for quadratic and cubic curves, but subdivision methods are faster for curves of degree five and greater [38].

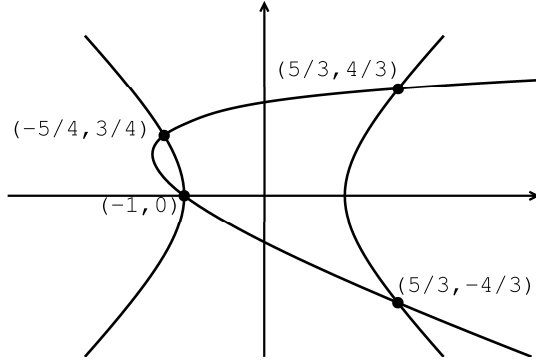


Figure 15.5. Intersections of Two Parametric Curves

Gröbner bases and resultants can also be used for finding the intersections between two parametric curves. The intersections of the curves $(x, y) = (a_1(s)/c_1(s), b_1(s)/c_1(s))$ and $(x, y) = (a_2(t)/c_2(t), b_2(t)/c_2(t))$ satisfy

$$\frac{a_1(s)}{c_1(s)} = \frac{a_2(t)}{c_2(t)}, \quad \frac{b_1(s)}{c_1(s)} = \frac{b_2(t)}{c_2(t)}.$$

This prompts us to compute the variety $V(a_1(s)c_2(t) - a_2(t)c_1(s), b_1(s)c_2(t) - b_2(t)c_1(s))$ in the (s, t) -space. The methods discussed in Section 15.6.2 can then be employed. For the previous example, the Gröbner basis using lexicographic ordering with $t > s$ is

$$\langle 144t^2 + 1944t - 4389s^5 - 6448s^4 + 32169s^3 + 11072s^2 - 27780s - 664, \\ 48s - 6t + 6s^2t + 2 + 14s^2 - 69s^3 - 16s^4 + 21s^5, -4s + 12s^2 + 5s^3 - 15s^4 - s^5 + 3s^6 \rangle.$$

Solving the third polynomial in the Gröbner basis gives six roots: $0, 1, 2, 1/3, -2, -1$. Note that the solutions of 1 and -1 are actually the roots of the denominator $1 - s^2$ of the first curve, and thus should be discarded. We then substitute the rest four roots into the second polynomial in the Gröbner basis and solve for the corresponding values of t . These values of s and t are the parametric values of the intersection points on the first and second parametric curves.

15.7. SURFACES

This section briefly overviews some applications of algebraic methods to surfaces. A rational parametric surface is usually defined by (15.2). We denote the maximum of the total degrees of the polynomials $a(s, t)$, $b(s, t)$, $c(s, t)$ and $d(s, t)$ by n and call it the *parametric degree*. The implicit equation of an algebraic surface is given by $f(x, y, z) = 0$ where $f(x, y, z)$ is a polynomial in x, y, z , and its maximum degree is denoted by m , called the *implicit degree*. Like curves, it is *always* possible to find an implicit equation of a parametric surface, but parametric equations can generally be found only for a very select class of implicit surfaces. Algorithms for implicitization and inversion also exist for surfaces [15,30], but the process for surfaces is much more complicated than the curve case. For example, a degree n plane parametric curve has an implicit equation that is also degree n . For surfaces, however, the implicit degree m can be as high as n^2 if the parametric degree is n .

15.7.1. Implicit degree of a rational parametric surface

The implicit degree can be thought of as the number of times that the surface is intersected by a generic straight line [15,48]. Define a generic straight line as the intersection of two distinct planes in general position $a_1x + a_2y + a_3z + a_4 = 0$ and $b_1x + b_2y + b_3z + b_4 = 0$. The planes intersect the parametric surface (15.2) in curves

$$a_1a(s, t) + a_2b(s, t) + a_3c(s, t) + a_4d(s, t) = 0 \quad (15.13)$$

and

$$b_1a(s, t) + b_2b(s, t) + b_3c(s, t) + b_4d(s, t) = 0 \quad (15.14)$$

These curves are each degree n in s, t . By Bezout's theorem, these two curves intersect in n^2 points, which must also be the number of times that the straight line common to the two planes intersects the surface. Thus, the degree of the surface, and of its implicit equation, is n^2 .

It seems curious that there are gaps in the sequence of the implicit degrees of parametric surfaces: 1, 4, 9, Are there no parametric surfaces whose implicit degree is 3 or 5 for example, or under what conditions will the degree decrease?

It may happen that there are values s_b and t_b satisfying $a(s_b, t_b) = b(s_b, t_b) = c(s_b, t_b) = d(s_b, t_b) = 0$. These parameter pairs (s_b, t_b) are referred to as *base points*. If a base point exists, the intersection curve of any plane with the surface will contain the base point. Thus, the above two curves will intersect at the base point and at $n^2 - 1$ other points. However, since the base point does not map to a unique point on the surface ($x = y = z = 0/0$ is undefined), this does not represent a point at which the straight line intersects the surface, and the degree of the surface is therefore $n^2 - 1$. Each additional simple base point diminishes the degree of the surface by one. Base points at infinity occur when all plane sections have a common asymptotic direction.

To understand the influence of more complicated base points on the implicit degree, consider the linear system of all curves given by (15.13), where each curve in the linear system is the intersection of the surface with the plane $a_1x + a_2y + a_3z + a_4 = 0$. A base point is any point in common with all members of the linear system. If two general curves in the linear system are tangent at a base point, they intersect twice at the base point and the degree of the surface becomes $n^2 - 2$. If two general curves in the linear system have a double point in common, they intersect four times at that base point and the degree becomes $n^2 - 4$. Thus, a general degree formula is $n^2 - \rho$ where ρ is the total number of times that two general curves in the linear system intersect at base points. This also assumes that the surface has a one-to-one parametrization.

If the surface (15.2) is a tensor product surface—one of the most popular representations for surfaces in CAGD—the parametrization is a bi-degree (n_s, n_t) parametrization in s and t . That means, n_s and n_t are the highest degrees of the parametric equations with respect to s and t . In this case, the total parametric degree is $n = n_s + n_t$. But there exist two base points at infinity, corresponding to $s = \infty, t = \infty$, counted at least n_t^2 and n_s^2 times respectively [14]. Thus the implicit degree of a bi-degree (n_s, n_t) parametrized rational surface is at most $(n_s + n_t)^2 - n_s^2 - n_t^2 = 2n_sn_t$. For example, a bicubic surface is usually of implicit degree 18.

15.7.2. Surface intersection curves

Surface/surface intersection (i.e., finding the intersection curve of two surfaces) is an important geometric operation in CAGD. The usual approach is to compute an approximation for the intersection curve. Algebraic geometry provides important information on the nature of intersections of parametric surfaces. For example, the degree of the intersection curve is easy to determine using Bezout's theorem which states that two surfaces of degree m and n respectively intersect in a curve of degree mn . Therefore if two surfaces have implicit degree n_1 and n_2 , the intersection curve has a degree $n_1 n_2$ (unless the surfaces have common components). Thus, two bicubic patches generally intersect in a curve of degree 324.

15.7.3. Implicitization

Resultants can be used to implicitize a rational parametric surface. Dixon's resultant is a good choice, because it works on three polynomials in two variables. Given a rational parametric surface (15.2), construct three auxiliary polynomials:

$$\begin{aligned} p(x, s, t) &= d(s, t)x - a(s, t), \\ q(y, s, t) &= d(s, t)y - b(s, t), \\ h(z, s, t) &= d(s, t)z - c(s, t) \end{aligned}$$

Note that $p(x, s, t) = q(y, s, t) = h(z, s, t) = 0$ only for values of x, y, z and s, t which satisfy (15.2). View $p(x, s, t)$, $q(y, s, t)$ and $h(z, s, t)$ as polynomials in s and t whose coefficients are linear in x , in y and in z , respectively. Then applying Dixon's resultant to these polynomials to eliminate s and t , we obtain a polynomial in x, y and z which we denote $f(x, y, z)$. Thus $f(x, y, z) = 0$ defines the implicit equation of the rational surface. In addition, by Cramer's rule, taking the ratio of the determinants of the submatrices from the Dixon's matrix corresponding to the terms $s^i t^j$ and $s^{i-1} t^j$, or the terms $s^i t^j$ and $s^i t^{j-1}$ yields the inversion equations of $s = s^i t^j / s^{i-1} t^j$ and $t = s^i t^j / s^i t^{j-1}$. Unfortunately, if the surface has finite basepoints, the resultant is identically zero and the algorithm fails.

The implicitization of rational parametric surfaces can also be accomplished by computing the elimination ideal. The Rational Implicitization Theorem [17] states that if $J = \langle dx - a, dy - b, dz - c, 1 - dw \rangle$, then $V(J \cap R[x, y, z])$ is the smallest variety in R^3 containing the parametric surface. The polynomial $1 - dw$ is introduced to assure that the method will work even if base points are present. Otherwise, base points would cause $J \cap R[x, y, z] = \{0\}$. In practical computation, we construct the Gröbner basis with the lexicographic ordering for the ideal J with $s > t > x > y > z$. The Gröbner basis will contain a polynomial in x, y, z . This is the implicit equation. If the parametrization of the surface is a one-to-one map, two polynomials linear in t and s are also contained in the Gröbner basis. They can produce the inversion maps.

Notwithstanding the robustness and elegance of the Gröbner basis solution to surface implicitization, it is not very computationally efficient. Recently, a promising new method, called the *moving surface method*, has been proposed for implicitizing rational surfaces [43]. Like resultants, the implicit equation is expressed as the determinant of a matrix.

We define a moving surface as

$$g(x, y, z, s, t) = \sum_{i=1}^{\sigma} f_i(x, y, z) \gamma_i(s, t) = 0$$

where the equations $f_i(x, y, z) = 0, i = 1, \dots, \sigma$ define a collection of implicit surfaces and where the $\gamma_i(s, t), i = 1, \dots, \sigma$ are a collection of polynomials in s and t . We require the $\gamma_i(s, t)$ to be linearly independent and to be relatively prime. A moving surface is said to “follow” a rational parametric surface (15.2) if

$$g\left(\frac{a(s, t)}{d(s, t)}, \frac{b(s, t)}{d(s, t)}, \frac{c(s, t)}{d(s, t)}, s, t\right) \equiv 0.$$

If we can find a set of σ moving surfaces

$$g_j(x, y, z, s, t) = \sum_{i=1}^{\sigma} f_{ji}(x, y, z) \gamma_i(s, t) = 0, \quad j = 1, \dots, \sigma$$

each of which follows a given rational surface, then

$$f(x, y, z) = \begin{vmatrix} f_{11}(x, y, z) & \cdots & f_{1\sigma}(x, y, z) \\ \vdots & \ddots & \vdots \\ f_{\sigma 1}(x, y, z) & \cdots & f_{\sigma\sigma}(x, y, z) \end{vmatrix} = 0$$

gives the implicit equation — as long as the degree of $f(x, y, z)$ is equal to the degree of the implicit equation of the rational surfaces.

In comparison, Gröbner basis method theoretically provides an elegant solution to implicitization of parametric surfaces, but involves a huge computation which limits its use in practice. The method of resultants is efficient, but fails when base points occur. Multivariate resultants can also be used for implicitization [15]. The moving surface method actually simplifies in the presence of base points. Furthermore, the method of moving surfaces provides a very compact representation for the implicit equation of a surface. For example, a bicubic patch can, in general, be written as a 9×9 determinant whose elements are all degree two in x, y, z . By contrast, Dixon’s resultant produces an 18×18 determinant. However, further study is needed on the moving surface method.

In summary, the operation of surface implicitization has not yet gained widespread use in practice, partly because the degree explosion that one encounters when moving from the parametric to the implicit form counteracts most algorithmic advantages that the implicit form might have over the parametric, and also because the computational complexity is very large, especially in the event of base points.

15.7.4. Parametrization

For an algebraic surface of arbitrary degree, Castelnuovo gave a necessary and sufficient condition for the existence of the rational parametrization [51]. Unfortunately, this criterion does not provide a constructive approach to parametrization. A systematic method for parametrizing a general rational algebraic surface is under investigation. Recently, various computational algorithms for parametrizing certain lower degree algebraic surfaces have been developed.

All degree two surfaces are rational. We can parametrize a quadric surface much like we did for conics. For example, we can use a pencil-of-lines approach by first translating the surface so that it touches the origin. Then define a line through the origin as the intersection of two pencils of planes: $y = sx$ and $z = tx$, say, and intersect the surface with that line. We demonstrate the procedure with a sphere $x^2 + y^2 + z^2 - 2z = 0$. Substituting $y = sx$ and $z = tx$, we obtain $x^2(1 + s^2 + t^2) - 2tx = 0$. This gives the x coordinates of the two points where the line intersects the sphere. Discarding the intersection at the origin gives the parametrization

$$x = \frac{2t}{1 + s^2 + t^2}, \quad y = \frac{2st}{1 + s^2 + t^2}, \quad z = \frac{2t^2}{1 + s^2 + t^2}$$

Most cubic surfaces are also rational. The only exception is the ruled cubic generated by a non-rational cubic curve. The cubic surface has a fascinating geometry. For example, the general cubic surface contains 27 straight lines, and those lines can be used in determining a parametrization for the surface. A detailed discussion can be found in [9,39,41].

15.8. OTHER ISSUES

The algebraic approaches to implicitization, parametrization and inversion illustrate how algebraic concepts and methods, such as resultants and Gröbner bases, help us analyze and solve some common problems in CAGD. Space limitations have prohibited the inclusion of several additional related topics, such as the following:

- A curve or a surface is said to be properly parametrized if to each point on the curve, except for possibly a finite number of points, there corresponds only one parameter value. It is natural to ask whether any improperly parametrized curve or surface can be reparametrized to become properly parametrized. For a rational curve, a classical theorem due to Lüroth guarantees the existence of a reparametrization [49]. However, for a rational surface, it depends on the base field where the surface is defined. Gröbner basis methods can be used for detecting and correcting improper parametrization [24].
- Geometric continuity was originally introduced as a smoothness measure for parametric curves and surfaces. This concept is also meaningful for implicit curves and surfaces. The geometric continuity conditions for implicit surfaces is studied in [25]. This consideration is important when using algebraic surfaces in geometric modeling. One application is to construct blending algebraic surfaces with a specified continuity [50]. Algebraic surfaces have been shown to have advantages over parametric surfaces when performing the blending operation.
- The fact that the most popular free-form surface patches – bicubic patches – are actually algebraic degree 18 has tempted some researchers to investigate the possibility of using lower degree algebraic surfaces for modeling purposes. Surfaces in algebraic geometry are usually global in nature while surfaces in CAGD are usually finitely defined (i.e., patches). The main reasons that parametric surface patches have been so popular in CAGD are that they can be pieced together with any desired degree of

continuity, and that there exist many elegant, intuitively meaningful techniques for controlling their shape. Therefore, to make algebraic surfaces useful in CAGD, the Bernstein-Bézier techniques have been adapted in defining algebraic surfaces [36]. Meaningful and efficient methods, such as interpolation, least-squares approximation and interactive modification, have been developed to model complicated shapes using *piecewise* implicit algebraic surfaces [7,8,20,21].

- Algebraic tools such as resultants and discriminants can help compute the intersection points between a ray and a surface (useful for performing ray tracing), and can help compute the silhouette points or curves in a scene. Some techniques based on algebraic methods have been developed to accurately render surfaces using computer graphics, such as ray-tracing [29,35] or in scan-line algorithms [40].
- The methods in algebraic geometry assume the procedure is carried out using exact (integer or rational number) arithmetic. Nevertheless, commercial computer systems dealing with CAGD use floating point arithmetic. This fact heavily hinders applying algebraic geometry methods to the practical problems of CAGD. Therefore computational theories and techniques of algebraic geometry in floating point arithmetic are of high interest. Some strategies for using algebraic methods in a floating point environment are discussed in [45].

Interest in algebraic techniques for the CAGD is growing, and it is evident that algebraic geometry is a valuable resource for computer aided geometric design.

REFERENCES

1. S. Abhyankar and C. Bajaj. Automatic parametrization of rational curves and surfaces I: Conics and conicoids. *Computer-Aided Design*, 19(1):11-14, 1987.
2. S. Abhyankar and C. Bajaj. Automatic parametrization of rational curves and surfaces II: Cubics and cubicoids. *Computer-Aided Design*, 19(9):499-502, 1987.
3. S. Abhyankar and C. Bajaj. Automatic parametrization of rational curves and surfaces III: Algebraic plane curves. *Computer Aided Geometric Design*, 5:309-321, 1988.
4. S. Abhyankar and C. Bajaj. Automatic parametrization of rational curves and surfaces IV: Algebraic space curves. *ACM Transactions on Graphics*, 8(4):324-333, 1989.
5. S.S. Abhyankar. *Algebraic Geometry for Scientists and Engineers*. American Mathematical Society, Providence, R.I., 1990.
6. W. Adams and P. Loustau. *An introduction to Gröbner bases*. American Mathematical Society, Providence, R.I., 1994.
7. C. Bajaj and I. Ihm. Smoothing of polyhedra with implicit algebraic splines. *Computer Graphics*, 26(2):79-88, SIGGRAPH 92, Chicago, Illinois, 1992.
8. C. Bajaj, J. Chenm, and G. Xu. Modeling with cubic A-patches. *ACM Transactions on Graphics*, 14(2):103-133, 1995.
9. C. Bajaj, R. Holt, and A. Netravali. Rational parametrizations of nonsingular cubic surfaces. *ACM Transactions on Graphics*, 17(1):1-31, 1998.
10. T. Becker and V. Weispfenning. *Gröbner Bases: A Computational Approach to Commutative Algebra*. Springer-Verlag, 1993.

11. W. Boehm and H. Prautzsch. *Geometric Concepts for Geometric Design*. A K Peters, Welleley, Massachusetts, 1993.
12. B. Buchberger. *Ein Algorithmus zum Auffinden der Basiselemente des Restklassenringes Nach Einem Nulldimensionalen Polynomideal*. PhD thesis (in German), Universitat Innsbruck, Austria, 1965.
13. B. Buchberger. Gröbner bases: An algorithmic method in polynomial ideal theory. In N.K. Bose, editor, *Multidimensional Systems Theory*, pages 184-232. D.Reidel Publishing Co., Netherlands, 1985.
14. E. Chionh. *Base Points, Resultants, and the Implicit Representation of Rational Surfaces*. PhD thesis, Dept. of Computer Science, University of Waterloo, Canada, 1990.
15. E. Chionh and R. Goldman. Degree, multiplicity, and inversion formulas for rational surfaces using u-resultants. *Computer Aided Geometric Design*, 9:93-108, 1992.
16. E. Chionh and R. Goldman. Elimination and resultants, Part 1: Elimination & bivariate resultants and Part 2: Multivariate resultants. *IEEE Computer Graphics and Applications*, 15(1):69-77 and 15(2):60-69, 1995.
17. D. Cox, J. Little, and D. O'Shea. *Ideals, Varieties, and Algorithms*. Springer-Verlag, 1992.
18. D. Cox, J. Little, and D. O'Shea. *Using Algebraic Geometry*. Springer-Verlag, 1998.
19. D. Cox, T. Sederberg, and F. Chen. The moving line ideal basis of planar rational curves. *Computer Aided Geometric Design*, 15:803-827, 1998.
20. W. Dahmen. Smooth piecewise quadratic surfaces. In T. Lyche and L. Schumaker, editors, *Mathematical Methods in Computer Aided Geometric Design*. Boston, Academic Press, 1989.
21. W. Dahmen and T. Thamm-Schaar. Cubicoids: modeling and visualization. *Computer Aided Geometric Design*, 10(2):89-108, 1993.
22. Y. De Montaudouin and W. Tiller. The Cayley method in computer aided geometric design. *Computer Aided Geometric Design*, 1:309-326, 1984.
23. A.L. Dixon. The eliminant of three quantics in two independent variables. *Proc. London Math. Soc.*, 6(Ser.2):468-478, 1908.
24. X. Gao and S. Chou. Implicitization of rational parametric equations. *Journal of Symbolic Computation*, 14:459-470, 1992.
25. T. Garritty and J. Warren. Geometric continuity. *Computer Aided Geometric Design*, 8(1):51-66, 1991.
26. R. Goldman, T. Sederberg, and D. Anderson. Vector elimination: A technique for the implicitization, inversion, and intersection of planar parametric rational polynomial curves. *Computer Aided Geometric Design*, 1:327-356, 1984.
27. C. Hoffmann. *Geometric and Solid Modeling: An Introduction*. Morgan Kaufmann, San Mateo, 1989.
28. C. Hoffmann. Implicit curves and surfaces in computer-aided geometric design. *IEEE Computer Graphics and Applications*, 13(1):79-88, 1993.
29. J. Kajiya. Ray tracing parametric patches. *Computer Graphics*, 16(3):245-254, 1982.
30. D. Manocha and J. Canny. Algorithm for implicitizing rational parametric surfaces. *Computer Aided Geometric Design*, 9(1):25-51, 1992.
31. G. Salmon. *Modern Higher Algebra*. G.E. Steckert & Co., New York, 1985.
32. G. Salmon. *A Treatise on the Analytic Geometry of Three Dimensions*, Vol I, II,

- R. Rogers, editor, Reprinted. Chelsea Publishing, 1914.
33. T. Sederberg. *Implicit and Parametric Curves and Surfaces*. PhD thesis, Purdue University, West Lafayette, IN, 1983.
 34. T. Sederberg, D. Anderson, and R. Goldman. Implicit representation of parametric curves and surfaces. *Computer Vision, Graphics and Image Processing*, 28:72-84, 1984.
 35. T. Sederberg and D. Anderson. Ray tracing of Steiner patches. *Computer Graphics*, 18(3):159-164, 1984.
 36. T. Sederberg. Piecewise algebraic surface patches. *Computer Aided Geometric Design*, 2:53-59, 1985.
 37. T. Sederberg and R. Goldman. Algebraic geometry for computer-aided geometric design. *IEEE Computer Graphics and Applications*, 6(6):5-59, 1986.
 38. T. Sederberg and S. Parry. A comparison of curve-curve intersection algorithms. *Computer-Aided Design*, 18:58-63, 1986.
 39. T. Sederberg and J. Snively. Parametrizing cubic algebraic surfaces. In R.R. Martin, editor, *The Mathematics of Surfaces II*, pages 299-320. Oxford University Press, Oxford UK, 1987.
 40. T. Sederberg and A. Zundel. Scan line display of algebraic surfaces. *Computer Graphics*, 23(3):147-156, 1989.
 41. T. Sederberg. Techniques for cubic algebraic surfaces, parts 1 and 2. *IEEE Computer Graphics and Applications*, July 1990, 14-25 and September 1990, 12-21.
 42. T. Sederberg, T. Saito, D. Qi, and K. Klimaszewski. Curve implicitization using moving lines. *Computer Aided Geometric Design*, 11:687-706, 1994.
 43. T. Sederberg and F. Chen. Implicitization using moving curves and surfaces. *SIGGRAPH 95, Computer Graphics Proceedings, Annual Conference Series*, pages 301-308, 1995.
 44. T. Sederberg, R. Goldman, and H. Du. Implicitizing rational curves by the method of moving algebraic curves. *Journal of Symbolic Computation*, 23:153-175, 1997.
 45. T. Sederberg. Applications to computer aided geometric design. *Proceedings of Symposia in Applied Mathematics*, 53:67-89, 1998.
 46. J. Semple and L. Roth. *Introduction to Algebraic Geometry*. Oxford University Press, Oxford, U.K., 1949.
 47. J. Sendra and F. Winkler. Symbolic parametrization of curves. *Journal of Symbolic Computation*, 6:607-632, 1991.
 48. B.L. Van der Waerden. *Modern Algebra*. Frederick Ungar, New York, 2nd edition, 1950.
 49. R. Walker. *Algebraic Curves*. Dover Publications, New York, 1949.
 50. J. Warren. Blending algebraic surfaces. *ACM Transactions on Graphics*, 8:263-278, 1989.
 51. O. Zariski. Castelnuovo's criterion of rationality $p_a = P_2 = 0$ of an algebraic surface, Ill. *J. Math.*, 2:303-315, 1958.
 52. J. Zheng and T. Sederberg. A direct approach to computing the mu-basis of planar rational curves. *Journal of Symbolic Computation*, 31(5):619-629, 2001.

