Implicitization using Moving Curves and Surfaces

Thomas W. Sederberg[1] and Falai Chen[2]
Brigham Young University

This paper presents a radically new approach to the century old problem of computing the implicit equation of a parametric surface. For surfaces without base points, the new method expresses the implicit equation in a determinant which is one fourth the size of the conventional expression based on Dixon's resultant. If base points do exist, previous implicitization methods either fail or become much more complicated, while the new method actually simplifies.

The new method is illustrated using the bicubic patches from Newell's teapot model. Dixon's method can successfully implicitize only 8 of those 32 patches, expressing the implicit equation as an 18×18 determinant. The new method successfully implicitizes *all* 32 of the patches. Four of the implicit equations can be written as 3×3 determinants, eight can be written as 4×4 determinants, and the remaining 20 implicit equations can be written using 9×9 determinants.

Categories and Subject Descriptors: I.3.5 [**Computer Graphics**]: Computational Geometry and Object Modeling.

General Terms: Algorithms

Additional Key Words and Phrases: Bézier patches, implicitization, base points.

1   INTRODUCTION

For any 2–D parametric curve $x = \frac{a(t)}{d(t)}$, $y = \frac{b(t)}{d(t)}$ where a, b, and d are polynomials, there exists an implicit equation $f(x, y) = 0$, where f is also a polynomial, which defines exactly the same curve. For example, a circle can be defined by the parametric equation $x = \frac{1-t^2}{1+t^2}$, $y = \frac{2t}{1+t^2}$ or by the implicit equation $x^2 + y^2 - 1 = 0$. The process of finding the implicit equation given the parametric equations is known as *implicitization*. Implicitization of 2–D curves leads to many practical algorithms. For example, a very fast algorithm for computing the intersection of two 2–D curves of low degree is based on implicitization [23]. Implicitization reduces the problem of curve intersection to one of finding the roots of a single polynomial.

Similarly, for any parametric surface $x = \frac{a(s,t)}{d(s,t)}$, $y = \frac{b(s,t)}{d(s,t)}$, $z = \frac{c(s,t)}{d(s,t)}$ where a, b, c, and d are polynomials in s, t, there exists a polynomial implicit equation $f(x, y, z) = 0$ which defines the same surface.

[1]368 Clyde Building, Brigham Young University, Provo UT 84602-4014. tom@byu.edu

[2]Permanent Address: Dept. of Mathematics, University of Science and Technology of China, Hefei 230026, Anhui, P. R. China

The general problem of surface implicitization has been studied for well over a century. In 1862, Salmon [18] noted that surface implicitization can be performed by eliminating the parameters from the parametric surface equations. Presumably he had in mind using Sylvester's "dialytic method" with which one could eliminate two variables from three polynomials, though the result generally needed to be "expressed as the quotient of one determinant divided by another" [17]. In 1908, Dixon published a more compact resultant for eliminating two variables from three polynomials which has become the standard method for surface implicitization, at least in the absence of base points. In 1983, Sederberg [19] resurrected Dixon's and Salmon's work in addressing the problem of how to implicitize surface patches. Other implicitization methods are surveyed in [10], and include ones based on Gröbner bases [1], numerical techniques [16], and multivariate resultants [4].

To implicitize a tensor product surface of degree $m \times n$, Dixon's resultant produces a $2mn \times 2mn$ matrix whose elements are linear in x, y, z. The determinant of that matrix is the implicit equation. For a biquadratic surface, the matrix is 8×8, and for a bicubic patch, the matrix is 18×18. (In this paper, any statement that a determinant is the implicit equation of a curve or surface should be taken as shorthand for "setting that determinant to zero gives the implicit equation".)

Surface implicitization has seen limited practical use partly because of the huge expressions involved, but also because in the event of base points (see section 4.1), things can get even more complicated. For example, if base points exist, Dixon's resultant is identically zero and hence fails to produce the implicit equation. Manocha has shown that in many cases the largest non-zero minor of Dixon's determinant is the implicit equation, but often it includes an extraneous factor [14]. Substantial further work is then needed to remove the unwanted factor. Implicitization using Gröbner bases [6] also usually fails when base points occur, since the implicit equation does not belong to the ideal generated by the parametric equations [15], although this problem can sometimes be circumvented by introducing some auxiliary polynomials into the Gröbner system [12, 10]. Nonetheless, Gröbner bases are known to be very slow in implicitizing bicubic patches.

Several other procedures have been devised to implicitize surfaces with base points [2, 9, 15, 16]. We don't review those methods here, but observe that those methods are generally more complicated than Dixon's method. Furthermore, base points are *not* a rare occurrence; most of the teapot patches have numerous base points.

This paper presents a fundamentally new procedure for implicitizing curves and surfaces in which the implicit equation can be written in much more compact form than before. Furthermore, in the presence of base points, the expressions actually *simplify*. In particular, the new method allows the implicit equation for a general bicubic patch to be written as a 9×9 determinant whose elements

are all degree two in x, y, z. If a base point exists, one row of that determinant can in general be replaced by degree one elements. Section 2 introduces the new strategy as it applies to curve implicitization. We refer to this method as the *moving curve* method. Section 3 introduces the *moving surface* method for surface implicitization. Sections 4 and 5 discuss how to implicitize general tensor product and triangular patches using moving surfaces. Section 6 reports on what happens when those methods are used to implicitize the 32 bicubic patches of the teapot. It turns out that those patches show surprising diversity in the number of base points.

2 CURVE IMPLICITIZATION

The standard method for implicitizing a 2–D curve is to use Bezout's resultant [22]. For a degree n rational curve, Bezout's resultant is the determinant of an $n \times n$ matrix whose elements are linear in x, y. For example, the implicit equation of the curve

$$x = \frac{2t^2 + 4t + 5}{t^2 + 2t + 3}; \quad y = \frac{3t^2 + t + 4}{t^2 + 2t + 3} \qquad (1)$$

can be found by taking the resultant of

$$t^2(x - 2) + t(2x - 4) + (3x - 5)$$

and

$$t^2(y - 3) + t(2y - 1) + (3y - 4).$$

Bezout's resultant for these two polynomials in t is:

$$
\begin{Vmatrix}
\begin{vmatrix} x - 2 & 2x - 4 \\ y - 3 & 2y - 1 \end{vmatrix} & \begin{vmatrix} x - 2 & 3x - 5 \\ y - 3 & 3y - 4 \end{vmatrix} \\[2ex]
\begin{vmatrix} x - 2 & 3x - 5 \\ y - 3 & 3y - 4 \end{vmatrix} & \begin{vmatrix} 2x - 4 & 3x - 5 \\ 2y - 1 & 3y - 4 \end{vmatrix}
\end{Vmatrix}
$$

$$
= \begin{vmatrix} 5x - 10 & 5x - y - 7 \\ 5x - y - 7 & -5x - 2y + 11 \end{vmatrix}. \qquad (2)
$$

Setting (2) to zero gives the implicit equation for (1).

Bezout's resultant is purely an algebraic device. By studying the following geometric interpretation of Bezout's resultant, we are led to the new implicitization algorithm.

A pencil of lines can be described by the equation

$$(a_0 x + b_0 y + c_0)(1 - t) + (a_1 x + b_1 y + c_1)t = 0 \qquad (3)$$

where the equations $a_0 x + b_0 y + c_0 = 0$ and $a_1 x + b_1 y + c_1 = 0$ define any two distinct lines.

Given two distinct pencils, $(a_{00} x + b_{00} y + c_{00})(1 - t) + (a_{10} x + b_{10} y + c_{10})t = 0$ and $(a_{01} x + b_{01} y + c_{01})(1 - t) + (a_{11} x + b_{11} y + c_{11})t = 0$, one line from each pencil corresponds to each value of t, and those two lines intersect in a point. The locus of points thus created for $-\infty \leq t \leq \infty$ is a conic section, as illustrated in Figure 1. This observation is attributed to Steiner and Chasles in the 1830's, though a roughly equivalent method for generating conic sections dates back to Newton [5].

It is easily shown that the implicit equation of this curve is

$$
\begin{vmatrix} a_{00} x + b_{00} y + c_{00} & a_{10} x + b_{10} y + c_{10} \\ a_{01} x + b_{01} y + c_{01} & a_{11} x + b_{11} y + c_{11} \end{vmatrix} = 0. \qquad (4)
$$

This same curve can actually be defined using many different pairs of pencils of lines, and it turns out that Bezout's resultant for degree two curves is nothing more than one manifestation of this
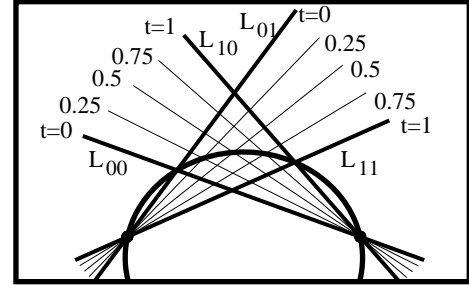

Figure 1: Intersection of Two Pencils of Lines

fact. For example, the two rows of the Bezout resultant (2) can be shown to define pencils of lines

$$(5x - 10)t + (5x - y - 7) = 0 \qquad (5)$$

and

$$(5x - y - 7)t + (-5x - 2y + 11) = 0 \qquad (6)$$

which intersect in the parametric curve (1).

The idea of defining a conic section as the intersection of two pencils of lines can be generalized to curves of any degree. One such generalization is given in [24]. This section presents some results similar to those in [24], but using a different approach. The approach taken here lays the foundation for the surface implicitization method presented in section 3.

Given a degree ν rational curve $\mathbf{Q}(t) = (x(t), y(t), w(t))$, define $g(t)$ to be the GCD of $x(t)$, $y(t)$, and $w(t)$. We can write

$$\mathbf{Q}(t) = g(t)\mathbf{P}(t) = g(t) \sum_{i=0}^{n} \mathbf{X}_i t^i, \qquad (7)$$

where $\mathbf{X}_i \equiv (x_i, y_i, w_i)$. If β is the degree of $g(t)$, $n = \nu - \beta$, we will say that $\mathbf{Q}(t)$ has β base points.

A *moving line*

$$\mathbf{X} \cdot L(t) := \mathbf{X} \cdot \sum_{j=0}^{m} L_j t^j = 0; \quad L_j = (a_j, b_j, c_j), \quad \mathbf{X} \equiv (x, y, w) \qquad (8)$$

is a parametric family of implicitly defined lines, with one line corresponding to each "time" t. For $m = 1$, the moving line is a pencil. A moving line is said to "follow" a rational curve if

$$\mathbf{Q}(t) \cdot L(t) \equiv 0, \qquad (9)$$

which means that at any time t, point $\mathbf{Q}(t)$ lies on line $L(t)$.

A set of moving lines $L_i(t)$, $i = 0, \ldots, \lambda$ is linearly independent if there do not exist constants c_i, $i = 0, \ldots, \lambda$ (not all zero) such that $\sum_{i=0}^{\lambda} c_i L_i(t) \equiv 0$.

Theorem 1. *For a degree ν curve $\mathbf{Q}(t)$ with β base points, there exist at least $2m + 2 + \beta - \nu$ linearly independent moving lines of degree m which follow the curve.*

Proof:

$$\mathbf{Q}(t) \cdot L(t) \equiv g(t)\mathbf{P}(t) \cdot L(t) \equiv 0 \qquad (10)$$

implies

$$\mathbf{P}(t) \cdot L(t) \equiv 0 \qquad (11)$$

since $g(t) \not\equiv 0$. We define

$$\mathbf{P}(t) = \sum_{i=0}^{n} (x_i, y_i, w_i) t^i, \qquad (12)$$

$$L(t) = \sum_{i=0}^{m} (a_i, b_i, c_i) t^i, \tag{13}$$

$$\mathbf{P}(t) \cdot L(t) = \sum_{k=0}^{n+m} \left\{ \sum_{i+j=k} (a_i x_j + b_i y_j + c_i w_j) \right\} t^k \tag{14}$$

The condition $\mathbf{P}(t) \cdot L(t) \equiv 0$ can be expressed $M b = 0$ where M is the $(n + m + 1) \times (3m + 3)$ matrix

$$\begin{bmatrix}
x_0 & y_0 & w_0 & 0 & \cdots & 0 & 0 & 0 \\
x_1 & y_1 & w_1 & x_0 & \cdots & 0 & 0 & 0 \\
x_2 & y_2 & w_2 & x_1 & \cdots & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
x_n & y_n & w_n & x_{n-1} & \cdots & . & . & . \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & 0 & \cdots & x_{n-1} & y_{n-1} & w_{n-1} \\
0 & 0 & 0 & 0 & \cdots & x_n & y_n & w_n
\end{bmatrix} \tag{15}$$

and

$$b = \begin{bmatrix} a_0 & b_0 & c_0 & \cdots & a_m & b_m & c_m \end{bmatrix}^T. \tag{16}$$

The dimension of the solution set is $3m + 3 - \text{rank}(M)$. But $\text{rank}(M) \leq n + m + 1$, so at least $2m + 2 - n = 2m + 2 + \beta - \nu$ linearly independent moving lines follow $\mathbf{Q}(t)$. $\blacksquare$

If $m = n - 1$, we know from theorem 1 that there exist $m + 1$ linearly independent moving lines $L_i(t) \cdot \mathbf{X} = 0$ where

$$L_i(t) = \sum_{j=0}^{m} L_{ij} t^j; \quad L_{ij} = (a_{ij}, b_{ij}, c_{ij}); \quad i = 0, \ldots m \tag{17}$$

such that

$$\mathbf{Q}(t) \cdot L_i(t) \equiv 0, \quad i = 0, \ldots, m. \tag{18}$$

Theorem 2. *Select any set of $m + 1$ linearly independent moving lines that follow $\mathbf{Q}(t)$ and define*

$$f(\mathbf{X}) = \begin{vmatrix}
L_{00} \cdot \mathbf{X} & \ldots & L_{0m} \cdot \mathbf{X} \\
. & \ldots & . \\
. & \ldots & . \\
. & \ldots & . \\
L_{m0} \cdot \mathbf{X} & \ldots & L_{mm} \cdot \mathbf{X}
\end{vmatrix}. \tag{19}$$

Then $f(\mathbf{X}) = 0$ is the implicit equation of $\mathbf{Q}(t)$. (In the case of an improperly parametrized curve [20], $f(\mathbf{X}) = 0$ will actually be some power of the implicit equation, since we generally take the implicit equation to be an irreducible polynomial).

Proof: $f(\mathbf{X}) = 0$ is the implicit equation of $\mathbf{Q}(t)$ if the following conditions are met:
1. $f(\mathbf{Q}(t)) \equiv 0$.
2. For all $\mathbf{X}$ for which $f(\mathbf{X}) = 0$, there exists a value of t such that $\mathbf{X} = \kappa \mathbf{Q}(t)$ where κ is a scalar constant. Since we are dealing with homogeneous (projective) coordinates, this simply means that $\mathbf{X}$ and $\mathbf{Q}(t)$ map to the same point in 2–D Cartesian space.

As for requirement 1, (17) and (18) can be written

$$\begin{bmatrix}
L_{00} & \ldots & L_{0m} \\
. & \ldots & . \\
. & \ldots & . \\
. & \ldots & . \\
L_{m0} & \ldots & L_{mm}
\end{bmatrix} \left\{ \begin{array}{c} 1 \\ t \\ \vdots \\ t^m \end{array} \right\} \cdot \mathbf{Q}(t) \equiv \left\{ \begin{array}{c} 0 \\ 0 \\ \vdots \\ 0 \end{array} \right\}. \tag{20}$$

Specializing $t = \tau$ and letting $\mathbf{X} = \mathbf{Q}(\tau)$, (20) can be expressed

$$\begin{bmatrix}
L_{00} \cdot \mathbf{X} & \ldots & L_{0m} \cdot \mathbf{X} \\
. & \ldots & . \\
. & \ldots & . \\
. & \ldots & . \\
L_{m0} \cdot \mathbf{X} & \ldots & L_{mm} \cdot \mathbf{X}
\end{bmatrix} \left\{ \begin{array}{c} 1 \\ \tau \\ \vdots \\ \tau^m \end{array} \right\} = \left\{ \begin{array}{c} 0 \\ 0 \\ \vdots \\ 0 \end{array} \right\}. \tag{21}$$

For (21) to be valid, either

$$\left\{ \begin{array}{c} 1 \\ \tau \\ \vdots \\ \tau^m \end{array} \right\} = \left\{ \begin{array}{c} 0 \\ 0 \\ \vdots \\ 0 \end{array} \right\} \tag{22}$$

or

$$\begin{vmatrix}
L_{00} \cdot \mathbf{X} & \ldots & L_{0m} \cdot \mathbf{X} \\
. & \ldots & . \\
. & \ldots & . \\
L_{m0} \cdot \mathbf{X} & \ldots & L_{mm} \cdot \mathbf{X}
\end{vmatrix} = 0. \tag{23}$$

Equation 22 never holds; hence (23) must be true, and condition 1 is satisfied.

Condition 2 requires that $f(\mathbf{X}) \not\equiv 0$. We can prove that this holds because the moving lines are linearly independent (see the appendix in the electronic version of this paper). Therefore $f(\mathbf{X})$ must be a non-zero polynomial of degree at most $n = m + 1$. Now consider a point $\mathbf{X}^*$ for which $f(\mathbf{X}^*) = 0$. Take an arbitrary line $ax + by + cw = 0$ that contains $\mathbf{X}^*$. There are ν roots of the equation $(a, b, c) \cdot \mathbf{Q}(t) = 0$, but β of those roots map to the undefined point $(0, 0, 0)$, leaving n values of t which map to actual intersection points between the line and $\mathbf{Q}(t)$. Each of those values of t satisfy condition 1, that is, they map to n points for which $f(\mathbf{Q}(t)) = 0$, and each of those points lies on the line $(a, b, c) \cdot \mathbf{X} = 0$.

The equation $f(\mathbf{X}) = 0$ defines an algebraic curve of degree n which intersects a general line in n points. Suppose $f(\mathbf{X}^*) = 0$, but there does not exist a value of t for which $\mathbf{Q}(t) = \kappa \mathbf{X}^*$. This means that we have found $n + 1$ points at which a degree n algebraic curve intersects a line, a violation of Bezout's theorem. $\blacksquare$

So far, we have merely examined well known curve impliciti-zation methods from a different angle, but yielding no significant computational advantage. We are now prepared to venture into profitable new territory. We define a 2–D *moving curve* as

$$C(\mathbf{X}; t) := \sum_{j=0}^{m} f_j(\mathbf{X}) t^j = 0 \tag{24}$$

where $\mathbf{X} = (x, y, w)$ and $f_j(\mathbf{X})$ is a polynomial of degree d. Thus $C(\mathbf{X}; t) = 0$ is a family of algebraic curves that vary with t. A moving curve is said to "follow" a rational curve $\mathbf{P}(t) = (x(t), y(t), w(t))$ if for all values of t, the point $\mathbf{P}(t)$ lies on the moving curve:

$$C(\mathbf{P}(t); t) = \sum_{j=0}^{m} f_j(x(t), y(t), w(t)) t^j \equiv 0. \tag{25}$$

For a degree n curve $\mathbf{P}(t)$, there are at least $\frac{d(d+3)(m+1)}{2} - nd$ linearly independent moving curves of degree d in $\mathbf{X}$ and degree m in t that follow the curve $\mathbf{P}(t)$. This can be shown as follows. Since $C(\mathbf{P}(t); t)$ is a degree $nd + m$ polynomial in t and the total number of coefficients in the polynomials $f_j(\mathbf{X})$ $(j = 0, 1, \ldots, m)$ is $\frac{(d+1)(d+2)(m+1)}{2}$, condition (25) is equivalent to a system of $nd + m + 1$ linear equations with $\frac{(d+1)(d+2)(m+1)}{2}$ unknowns.

Hence there are at least $\frac{(d+1)(d+2)(m+1)}{2} - (nd + m + 1) = \frac{d(d+3)(m+1)}{2} - nd$ linearly independent moving curves that follow $\mathbf{P}(t)$. If $(m + 1)(d^2 + 3d - 2) \geq 2nd$, there exist at least $m + 1$ linearly independent moving curves that follow $\mathbf{P}(t)$. For example, $d = 1$ and $m = n - 1$ is the case in theorems 1 and 2.

Theorem 3. *Given $m + 1$ moving curves*

$$C_i(\mathbf{X}; t) := \sum_{j=0}^{m} f_{ij}(\mathbf{X}) t^j = 0, \quad i = 0, \ldots, m \qquad (26)$$

which follow $\mathbf{P}(t)$, we define

$$f(\mathbf{X}) = \begin{vmatrix} f_{00}(\mathbf{X}) & \ldots & f_{0m}(\mathbf{X}) \\ \cdot & \ldots & \cdot \\ \cdot & \ldots & \cdot \\ \cdot & \ldots & \cdot \\ f_{m0}(\mathbf{X}) & \ldots & f_{mm}(\mathbf{X}) \end{vmatrix} \qquad (27)$$

If the degree of $f(\mathbf{X})$ is $n - \beta$ (which implies that $f(\mathbf{X}) \not\equiv 0$), then $f(\mathbf{X}) = 0$ is the implicit equation of $\mathbf{P}(t)$.

Proof: Essentially the same as the proof for theorem 2. ∎

We now explore the possible degrees for moving curves which follow $\mathbf{P}(t)$. Letting $d = 2$ and $m = \left[\frac{n-1}{2}\right]$, we find that there exist at least $m + 1$ linearly independent curves of degree 2 in $\mathbf{X}$ and degree m in t that follow the curve $\mathbf{P}(t)$. In the case where n is odd, theorem 1 assures that there will also be one moving line of degree m which follows $\mathbf{P}(t)$. Thus, from theorem 3, a rational curve with no base points, and of even degree, can *generally* be implicitized as the determinant of a $\frac{n}{2} \times \frac{n}{2}$ matrix whose elements are degree 2 in x, y, w. Likewise, a rational curve of odd degree and no base points can *generally* be implicitized as the determinant of an $\frac{n+1}{2} \times \frac{n+1}{2}$ matrix with one linear row, and the remaining rows quadratic.

We emphasize the word *generally* because theorem 3 requires $f(\mathbf{X}) \not\equiv 0$. Under certain conditions, the determinant in (27) *will* vanish, even though the rows are linearly independent. The reason is that the rows might be linearly independent, but *polynomially* dependent. The following theorem shows that high order singularities can create such a condition. A thorough description of the geometric properties of curves for which this condition occurs remains an open question, though it appears at present that singularities are not the complete answer.

Theorem 4. *The implicit equation of a quartic curve with no base points can be written as a 2×2 determinant. If the curve doesn't have a triple point, then each element of the determinant is a quadratic; otherwise one row is linear and one row is cubic.*

The rather tedious proof includes showing that if $f(\mathbf{X})$ is formed by a 2×2 determinant with quadratic elements, the following four statements are equivalent:

1. $f(\mathbf{X}) \equiv 0$.

2. There exists a degree one moving line that follows $\mathbf{P}(t)$.

3. $\mathbf{P}(t)$ has a triple point.

4.
$$\begin{vmatrix} x_0 & y_0 & w_0 & 0 & 0 & 0 \\ x_1 & y_1 & w_1 & x_0 & y_0 & w_0 \\ x_2 & y_2 & w_2 & x_1 & y_1 & w_1 \\ x_3 & y_3 & w_3 & x_2 & y_2 & w_2 \\ x_4 & y_4 & w_4 & x_3 & y_3 & w_3 \\ 0 & 0 & 0 & x_4 & y_4 & w_4 \end{vmatrix} = 0. \qquad (28)$$

The details of the proof are omitted here (see the appendix of the electronic version of the paper). However, this discussion is important because similar phenomena can occur with surfaces.

3 SURFACES

It is convenient to define a rational surface in homogeneous form:

$$\mathbf{X}(s, t) = (X(s, t), Y(s, t), Z(s, t), W(s, t)) \qquad (29)$$

where $X(s, t), Y(s, t), Z(s, t), W(s, t)$ are polynomials in s, t. The Cartesian coordinates of points on the surface are given by

$$x = \frac{X(s, t)}{W(s, t)}, \quad y = \frac{Y(s, t)}{W(s, t)}, \quad z = \frac{Z(s, t)}{W(s, t)}. \qquad (30)$$

Among the most common rational surfaces used in computer graphics are the tensor product patches for which, in power basis,

$$\mathbf{X}(s, t) = \sum_{i=0}^{d_1} \sum_{j=0}^{d_2} \mathbf{X}_{ij} s^i t^j, \quad \mathbf{X}_{ij} = (x_{ij}, y_{ij}, z_{ij}, w_{ij}). \quad (31)$$

3.1 Base Points

A base point is a value of (s, t) for which $\mathbf{X}(s, t) = (0, 0, 0, 0)$. In the absence of base points, the implicit equation of a tensor product surface can be expressed using Dixon's resultant [8], which is the determinant of a $2d_1d_2 \times 2d_1d_2$ matrix whose elements are linear in (x, y, z). For example, a biquadratic surface requires an 8×8 determinant, and a bicubic surface an 18×18 determinant. The method presented in this section expressed the implicit equation of a bicubic patch with no base points as a 9×9 determinant whose elements are degree two in (x, y, z). (It is well known that all tensor product surfaces have multiple base points at $s = \infty$ and $t = \infty$. Here we mean no *additional* base points.)

Base points are of interest for two reasons. First, each simple base point decreases the degree of the implicit equation of the rational surface by one. The full story on the relationship between base points and degree becomes more complicated when considering base points with higher multiplicity [3, 13].

Second, if base points exist, Dixon's resultant vanishes identically. To implicitize surfaces which contain base points, more complicated methods have been devised [2, 9, 15, 16] such as the method of undetermined coefficients, successive elimination, perturbations, and customized resultants. In general, these methods are much more complicated in the presence of base points.

By contrast, the implicitization approach in this paper simplifies in the case of base points. In general, for each base point on a bicubic patch, one of the rows of the 9×9 determinant can be converted from degree two to degree one in (x, y, z).

3.2 Moving Surfaces

We define a *moving surface* as

$$g(\mathbf{X}, s, t) := \sum_{i=1}^{\sigma} h_i(\mathbf{X}) \gamma_i(s, t) = 0 \qquad (32)$$

where the equations $h_i(\mathbf{X}) = 0$, $i = 1, \ldots, \sigma$ define a collection of implicit surfaces and where the $\gamma_i(s, t)$, $i = 1, \ldots, \sigma$ are a collection of polynomials in s and t. We will refer to the $\gamma_i(s, t)$ as the blending functions for the moving surface. We require the blending functions to be linearly independent and to have no non-constant factor common to all of them. A moving surface is said to follow a rational surface $\mathbf{X}(s, t)$ (29) if

$$g(\mathbf{X}(s, t), s, t) \equiv 0. \qquad (33)$$

When we make a statement such as "so many moving surfaces exist", it is implied that those moving surfaces follow the parametric surface under discussion, even though we may not explicitly say so.

Theorem 5. *Given a set of σ moving surfaces*

$$g_j(\mathbf{X}, s, t) = \sum_{i=1}^{\sigma} h_{ji}(\mathbf{X})\gamma_i(s, t) = 0, \quad j = 1, \ldots, \sigma, \quad (34)$$

each of which follows a given rational surface $\mathbf{X}(s, t)$ (29). Define

$$f(\mathbf{X}) = \begin{vmatrix} h_{11}(\mathbf{X}) & \ldots & h_{1\sigma}(\mathbf{X}) \\ \vdots & \vdots & \vdots \\ h_{\sigma 1}(\mathbf{X}) & \ldots & h_{\sigma\sigma}(\mathbf{X}) \end{vmatrix}. \quad (35)$$

If the degree d of $f(\mathbf{X})$ is equal to the degree of the implicit equation of the rational surface $\mathbf{X}(s, t)$, then $f(\mathbf{X}) = 0$ is the implicit equation of $\mathbf{X}(s, t)$.

Proof: $f(\mathbf{X}) = 0$ is the implicit equation of $\mathbf{X}(s, t)$ if the following conditions are met:
1. $f(\mathbf{X}) \not\equiv 0$.
2. $f(\mathbf{X}(s, t)) \equiv 0$.
3. For all $\mathbf{X}$ for which $f(\mathbf{X}) = 0$, there exists a parameter pair s, t such that $\mathbf{X} = \kappa \mathbf{X}(s, t)$ where κ is a scalar constant.

Condition 1 is satisfied by the requirement in the theorem that the degree of $f(\mathbf{X})$ is equal to the degree of the implicit equation.

The fact that each of the σ moving surfaces follow $\mathbf{X}(s, t)$ means that the set of equations

$$\begin{bmatrix} h_{11}(\mathbf{X}) & \ldots & h_{1\sigma}(\mathbf{X}) \\ \vdots & \vdots & \vdots \\ h_{\sigma 1}(\mathbf{X}) & \ldots & h_{\sigma\sigma}(\mathbf{X}) \end{bmatrix} \begin{Bmatrix} \gamma_1(s, t) \\ \vdots \\ \gamma_\sigma(s, t) \end{Bmatrix} = 0 \quad (36)$$

is satisfied for $\mathbf{X} = \mathbf{X}(s, t)$. If for some (s, t) $f(\mathbf{X}(s, t)) \neq 0$, then $\gamma_1(s, t) = \ldots = \gamma_\sigma(s, t) = 0$. But since $\gamma_1(s, t), \ldots, \gamma_\sigma(s, t)$ have no common factor, there are at most a finite number of (s, t) values for which $\gamma_1(s, t) = \ldots = \gamma_\sigma(s, t) = 0$. Consequently if there are any (s, t) pairs such that $f(\mathbf{X}(s, t)) \neq 0$, the number of such pairs is finite. But since $f(\mathbf{X}(s, t))$ is a polynomial in (s, t), it is therefore identically zero and condition 2 is met.

Base points map ("blow up") to entire curves on the surface known as seam curves [15]. Some authors haved argued that these curves can be interpreted as lying on the implicit surface but not on the parametric surface since there is not a parameter value for which $\mathbf{X} = \kappa \mathbf{X}(s, t)$ if $\mathbf{X}$ lies on a seam curve [3]. We avoid that debate here, and are content to prove that condition 3 holds at least for points not on seam curves. Suppose then that there exists a point $\mathbf{X}^*$ such that $f(\mathbf{X}) = 0$ but $\mathbf{X}^* \neq k\mathbf{X}(s, t)$ for any (s, t). Choose a line through $\mathbf{X}^*$ which does not intersect any seam curves and which is not tangent to the surface. Take any two planes containing that line and compute their intersection with the parametric surface $\mathbf{X}(s, t)$, yielding two curves in parameter space $g_1(s, t) = 0$ and $g_2(s, t) = 0$. Those two curves will intersect at all base points of the surface, and at d other (s, t) parameter pairs where d is the degree of the implicit equation of the surface [13]. But from condition 2, those d parameter pairs map to points for which $f(\mathbf{X}) = 0$, making a total of $d + 1$ points lying on a line for which $f(\mathbf{X}) = 0$, a contradiction of Bezout's theorem. ∎

Sections 4 and 5 will prove that it is *always* possible to find a square matrix of moving surfaces that follow any given tensor product or triangular surface patch, and they present a systematic way of finding such matrices. It is very difficult to give a rigorous proof that for any given $\mathbf{X}(s, t)$ a matrix (35) can always be found so that the degree of $f(\mathbf{X})$ is equal to the degree of the implicit equation of the rational surface $\mathbf{X}(s, t)$ (and hence the determinant is not identically zero). However, in scores of example cases, we have never failed to find such a matrix.

3.3 Examples

Theorem 5 proposes a method for implicitizing rational surfaces by finding sets of moving surfaces which follow it. We here illustrate that concept with a few simple cases. These may seem somewhat ad hoc, but they have the advantage of being concrete numerical examples which are small enough to verify by hand. No explanation is given in these examples of how to find the moving surfaces; sections 4 and 5 outline a procedure for that.

3.3.1 Explicit Surface

The simplest parametric surface to implicitize is the explicit surface

$$x = s; \quad y = t; \quad z = q(s, t) \quad (37)$$

for which the implicit equation is merely $q(x, y) - z = 0$. This case is so trivial, that it actually becomes a little more complicated to implicitize it using moving surfaces than to merely write $q(x, y) - z = 0$. However, it serves as a simple introduction to moving surfaces. In this case, we can take as blending functions $\gamma_1(s, t) = s$, $\gamma_2(s, t) = t$, $\gamma_3(s, t) = 1$. The three moving surfaces, in matrix form, are

$$\begin{bmatrix} 1 & 0 & -x \\ 0 & 1 & -y \\ a(x, y) & b(x, y) & -z + q(0, 0) \end{bmatrix} \begin{Bmatrix} s \\ t \\ 1 \end{Bmatrix} = 0 \quad (38)$$

where $a(x, y)$ and $b(x, y)$ are chosen to satisfy $a(x, y)x + b(x, y)y + q(0, 0) = q(x, y)$. These three moving surfaces clearly follow the parametric surface, and the determinant of the 3×3 matrix in (38) is clearly the implicit equation.

3.3.2 Cubic Surface

As far as the authors are aware, the closest hint in the literature to anything like moving surfaces is the observation, dating back at least to Salmon in 1862 [18], that a degree three algebraic surface can be defined as the intersection of three "bundles" of planes (the classical term for what we here would call a moving plane with blending functions 1, s, and t). Salmon began with those three bundles of planes, and computed the parametric and implicit equations of the surface from them (see [21] for a more recent presentation).

For implicitization, we work in reverse, finding the moving planes given the parametric equations. Here are a set of parametric equations for a surface that we know in advance to have a degree three implicit equation. The parametric equations have six base points.

$$x = 2 + 2t^3 + s^2 + 4t^2 + 4ts + 2t + ts^2 + 3s$$

$$y = -2t^2s - ts - 2s^2 + s - s^3 - 2t + 2$$

$$z = -3t^2s + 2t^2 - 2ts^2 - 3ts - 2t - s^3 - 3s^2 - 2s$$

$$w = -t + ts^2 + s^2 - s + s^3 + t^3 - 1 + t^2$$

Once again, the moving surface blending functions are 1, s, and t. The three moving surfaces, in matrix form, are

$$\begin{bmatrix} x & y & z \\ y + w & 2y - z & y + 2w \\ z - y & -x + 2w & x - y \end{bmatrix} \begin{Bmatrix} s \\ t \\ 1 \end{Bmatrix} = 0 \quad (39)$$

and the implicit equation is the determinant of the matrix.

3.3.3 Steiner Surface

The canonical Steiner surface is given by parametric equations

$$x = 2st; \quad y = 2t; \quad z = 2s; \quad w = s^2 + t^2 + 1.$$

This is a special case of a triangular surface patch, a general implic-itization procedure for which is given in section 5. We can again take $\gamma_1(s, t) = s$, $\gamma_2(s, t) = t$, $\gamma_3(s, t) = 1$. The three moving surfaces, in matrix form, are

$$\begin{bmatrix} y & -2z & x \\ y & -z & 0 \\ xz & xy - xz - 2zw & x^2 + yz \end{bmatrix} \left\{ \begin{array}{c} s \\ t \\ 1 \end{array} \right\} = 0 \qquad (40)$$

The determinant of the matrix, $x^2y^2 + x^2z^2 + y^2z^2 - 2xyzw$ is indeed the implicit equation of the canonical Steiner surface.

3.3.4 Surface of Revolution

When Newell reverse-engineered his teapot in 1975 [7], rational Bézier patches were not in wide use and so the surfaces of revolu-tion were approximated using polynomial patches. While that ap-proximation is well within graphical tolerance, there are advantages to reformulating the teapot using rational Bézier patches. First, the rational case can exactly represent surfaces of revolution. Second, a rational bicubic patch can model $180°$ of a surface of revolution, thus cutting in half the number of patches used to model the rim, body, lid, and bottom. Third, and most importantly for current needs, an exact surface of revolution can be implicitized in a much more compact form than can the polynomial approximation in [7].

In this section, we implicitize a patch from the lower body of the teapot which has been modified to exactly represent a surface of revolution. As discussed in section 6, without that modification, the implicit equation takes the form of a 9×9 determinant whereas the implicit equation of the modified patch can be expressed in a 2×2 determinant.

The teapot lower body is defined by rotating around the z axis the cubic polynomial Bézier curve with control points $(2, 0, .9)$, $(2, 0, .45)$, $(1.5, 0, .225)$, and $(1.5, 0, .15)$. The exact surface of revolution can be represented in rational Bézier form by

$$\mathbf{X}(s, t) = [(1-t)^3 \; 3t(1-t)^2 \; 3t^2(1-t) \; t^3] M \left\{ \begin{array}{c} (1-s)^3 \\ 3s(1-s)^2 \\ 3s^2(1-s) \\ s^3 \end{array} \right\}$$

where $M =$

$$\begin{bmatrix} (2, 0, .9, 1) & (\frac{2,4,.9,1}{3}) & (\frac{-2,4,.9,1}{3}) & (-2, 0, .9, 1) \\ (2, 0, .45, 1) & (\frac{2,4,.45,1}{3}) & (\frac{-2,4,.45,1}{3}) & (-2, 0, .45, 1) \\ (1.5, 0, .225, 1) & (\frac{1.5,3,.225,1}{3}) & (\frac{-1.5,3,.225,1}{3}) & (-1.5, 0, .225, 1) \\ (1.5, 0, .15, 1) & (\frac{1.5,3,.15,1}{3}) & (\frac{-1.5,3,.15,1}{3}) & (-1.5, 0, .15, 1) \end{bmatrix}$$

The implicit equation of this surface can be expressed in a 2×2 determinant. In this case, we can take $\gamma_1(s, t) = 1$, $\gamma_2(s, t) = s$ and the two moving surfaces are

$$\begin{bmatrix} (x+y)C - B & -2xC \\ -yC & (x+y)C + B \end{bmatrix} \left\{ \begin{array}{c} 1 \\ s \end{array} \right\} = 0 \qquad (41)$$

where

$$B = 618 \quad + \quad 6720z + 17000z^2 - 64000/9z^3$$
$$+ \quad 1305/2y^2 - 120y^2z + 1305/2x^2 - 120x^2z$$

$$C = 3x^2 + 3y^2 + 1600z^2 + 6900z + 1197$$

The determinant then gives the implicit equation of the surface of revolution — a degree 6 polynomial.

4 TENSOR PRODUCT PATCHES

We now present a systematic method for computing moving sur-faces which follow a tensor product patch (31). It is most convenient to work in power basis. As we will see, a good choice for the $\gamma_i(s, t)$ in (32) is simply the tensor product basis:

$$\gamma_i(s, t) = s^j t^k; \qquad (42)$$
$$j = 0, \ldots, b_1; \; k = 0, \ldots, b_2; \; i = k(b_1 + 1) + j + 1,$$

so $\sigma = (b_1 + 1)(b_2 + 1)$.

A degree n polynomial in three variables has $(n+1)(n+2)(n+3)/6$ coefficients. Thus, if all σ polynomials $h_i(\mathbf{X})$ are degree n, there are a total of $\sigma(n+1)(n+2)(n+3)/6$ coefficients. We can determine η, a lower bound on the number of linearly indepen-dent families of moving surfaces $g(\mathbf{X}, s, t)$ that follow $\mathbf{X}(s, t)$ by generating a set of linear equations as we did for the curve case in (15). For the surface case, the identity in (33) can be satisfied by solving a set of $(nd_1 + b_1 + 1) \times (nd_2 + b_2 + 1)$ linear homogeneous equations in $\sigma(n+1)(n+2)(n+3)/6$ unknowns. Thus,

$$\eta = \frac{(b_1 + 1)(b_2 + 1)(n + 1)(n + 2)(n + 3)}{6}$$
$$- (nd_1 + b_1 + 1)(nd_2 + b_2 + 1). \qquad (43)$$

If we can find values of n, b_1, and b_2 so that $\eta \geq \sigma$, there will be enough $h_{ji}(\mathbf{X})$ to fill a square matrix and, if the conditions in theorem 5 are met, the determinant of that matrix will be the implicit equation of $\mathbf{P}(s, t)$. Two cases turn out to exactly give $\eta = \sigma$: moving planes, and moving quadrics. Choosing $n = 1$ (moving planes), $b_1 = 2d_1 - 1$ and $b_2 = d_2 - 1$ yields $\eta = \sigma = 2d_1d_2$. With this choice, the implicit equation of a bicubic patch with no base points occurs as the determinant of an 18×18 matrix. It can be shown that Dixon's resultant is a special case of this implicitization method using moving planes.

If we choose $n = 2$ (the moving algebraic surfaces are quadrics), $b_1 = d_1 - 1$ and $b_2 = d_2 - 1$, we then have at least $\eta = \gamma = d_1d_2$ linearly independent moving surfaces. This means, for example, that a bicubic patch with no base points can generally be implicit-ized in the form of a 9×9 determinant whose elements are quadratic in x, y, z. A biquadratic patch with no base points can generally be implicitized using a 4×4 matrix with quadratic elements.

We stress the word *generally*. For arbitrarily chosen control points, experience has shown that the conditions in theorem 5 are always satisfied. As will be seen in the teapot patches, however, when control points are placed in some coherent fashion, the likeli-hood of singularities increases and the conditions in theorem 5 may no longer be met. Section 7 comments on this in more detail.

4.1 Base Points

For $n = 1$, $b_1 = d_1 - 1$ and $b_2 = d_2 - 1$, we find that $\eta = 0$. Note that η in (43) is a lower bound on the number of linearly independent moving surfaces; the actual number might be higher, depending on the rank of the $(nd_1 + b_1 + 1) \times (nd_2 + b_2 + 1)$ matrix. Here we show that in the presence of base points, the rank does indeed drop.

Theorem 6. *If $\mathbf{P}(s, t)$ has ρ distinct base points in general position (as defined in the proof found in the appendix of the electronic version), there exist at least ρ linearly independent moving planes whose blending functions are given by (42) with $b_1 = d_1 - 1$ and $b_2 = d_2 - 1$.*

Proof: This is a brief sketch of the proof. The complete proof can be found in the appendix of the electronic version of the paper.

As noted, the identity in (33) can be satisfied by solving a set of $4d_1d_2$ homogeneous linear equations in $4d_1d_2$ unknowns. The way

we generated those equations before was to simply expand (33), producing a polynomial of degree $2d_1 - 1$ in s and degree $2d_2 - 1$ in t. This polynomial has $4d_1d_2$ terms, and the $4d_1d_2$ equations are created by setting the coefficient of each of those terms equal to zero.

We could create an equally valid set of $4d_1d_2$ equations by choosing $4d_1d_2$ different parameter pairs (s_j, t_j), $j = 1, \ldots, 4d_1d_2$ and taking the equations to be

$$\sum_{i=1}^{d_1d_2} h_i(\mathbf{X}((s_j, t_j)))\gamma_i(s_j, t_j) = 0, \quad j = 1, \ldots, 4d_1d_2.$$

However, if we take ρ of those (s_j, t_j) to be base points, those ρ equations will be identically zero and the rank of the matrix will diminish by ρ. Hence, it is possible to find at least ρ moving planes. ∎

This means that for $b_1 = d_1 - 1$ and $b_2 = d_2 - 1$, we can find ρ moving planes and (at least) $d_1d_2 - \rho$ moving quadrics that follow $\mathbf{P}(s, t)$. If the determinant of the $d_1d_2 \times d_1d_2$ matrix containing those ρ moving planes and $d_1d_2 - \rho$ moving quadrics is not identically zero, it is the implicit equation.

This discussion on base points has dealt with *distinct* base points. One might be tempted to postulate that base points with multiplicity greater than one [13] would always free up moving planes equal in number to the total base point multiplicity. While this happens in many cases (such as with the teapot), it is not always so. A more detailed analysis of this question must await a future paper.

What if more than d_1d_2 base points occur, since then there are no more quadratic rows left to convert to linear rows? Again, a complete answer to this question will be left for later. However, preliminary tests suggest that the size of the matrix can continue to shrink, usually allowing blending functions with $b_1 \leq d_1 - 2$, and/or $b_2 \leq d_2 - 2$. The example in section 3.3.4 is such a case, involving 12 base points. We note that the surface of revolution can also be implicitized as a 4×4 determinant with two linear rows and two quadratic rows. In that case, $b_1 = b_2 = 1$.

5   TRIANGULAR PATCHES

By a triangular surface patch, we mean one whose parametric equations are of pure degree d:

$$\mathbf{X}(s, t) = \sum_{i+j \leq d} \mathbf{X}_{ij} s^i t^j. \tag{44}$$

In this case, the right choice for moving surface blending functions are the $\sigma = d(d+1)/2$ monomials in s, t of total degree $< d$:

$$\gamma_i(s, t) = s^j t^k, \quad j = 0, \ldots, d-1; \quad k = 0, \ldots, d-j-1; \tag{45}$$

with $i = j(d - \frac{j-1}{2}) + k + 1$.

As in the tensor product case, if all σ polynomials $h_i(\mathbf{X})$ are degree n, there are a total of $\sigma(n+1)(n+2)(n+3)/6$ coefficients. The identity in (33) can be satisfied by solving a set of $(nd+d)(nd+d+1)/2$ linear homogeneous equations in $\sigma(n+1)(n+2)(n+3)/6$ unknowns. Thus, the number of degree n moving surfaces is at least

$$\begin{aligned} \eta &= d(d+1)(n+1)(n+2)(n+3)/12 \\ &\quad - (nd+d)(nd+d+1)/2. \end{aligned} \tag{46}$$

Letting $n = 1$, we find that there are at least d moving planes. Letting $n = 2$, there are at least $(d^2 + 7d)/2$ moving quadrics. However, $4d$ of those moving quadrics can be created from the moving planes as follows. Let a moving plane be given by (32). Then

$$(c_1 x + c_2 y + c_3 z + c_4 w)\sum_{i=1}^{\sigma} h_i(\mathbf{X})\gamma_i(s, t) = 0$$

gives a moving quadric that also follows the surface. Hence, for each moving plane, there exists four moving quadrics, and there are only $(d^2 + 7d)/2 - 4d = (d^2 - d)/2$ moving quadrics which cannot be created from moving planes. Hence, the matrix in (35) can just exactly be filled with d linear rows and $(d^2 - d)/2$ quadratic rows. If the determinant does not vanish, $f(\mathbf{X})$ is degree d^2 as expected.

The Steiner surface example in section 3.3.3 is an application of this method, for $d = 2$.

The base point discussion in section 4.1 applies also to triangular patches; each simple base point will generally allow a quadratic row to be converted to a linear row.

The cubic surface example in section 3.3.2 illustrates what can happen when several (6 in this case) base points occur.

6   THE TEAPOT

The 32 bicubic patches defining Newells' teapot [7] provide a surprisingly diverse set of tests for moving surface implicitization.

The teapot patches fall into ten groups: rim, upper body, lower body, upper handle, lower handle, upper spout, lower spout, upper lid, lower lid, and bottom. All patches in a given group are simple linear transformations of one another, so their implicit equations are similar.

In every case, it is possible to implicitize these patches using 9×9 determinants with tensor product blending functions (42) for which $b_1 = b_2 = 2$. The 16 patches in the rim, lower lid, upper body and lower body all have degree nine implicit equations (all having nine base points), and their implicit equations can each be expressed as 9×9 determinants with nine linear rows. Those nine moving planes can be found by solving a set of linear equations as discussed in section 4.

The four patches in the upper lid have five base points (degree 13 implicit equations). The determinant has five moving planes and four moving quadrics. The four bottom patches have three base points (degree 15 implicit equations), and their determinant has three moving planes and six moving quadrics.

All four spout patches have no base points (degree 18 implicit equations), and their determinant has nine moving quadratics.

The four handle patches provided a surprise. These patches have no base points (degree 18 implicit equations). However, the determinant has three moving planes, three moving quadrics, and three moving cubics! This phenomenon is reminiscent of the degree four planar curve with a triple point.

It turns out that the rim's implicit equation can also be expressed as a 3×3 determinant using $\sigma_1 = 1$, $\sigma_2 = s$, $\sigma_3 = s^2$. In this case, there are two moving quartics and one moving plane. This discovery was made purely by trial and error, but the search was motivated by the fact that for this surface, z is a quadratic function of s only, so we immediately knew there exists a moving plane with $\sigma_1 = 1$, $\sigma_2 = s$, $\sigma_3 = s^2$. The two moving quartics were pure serendipity.

Similarly, the upper lid's implicit equation can be written as a 4×4 determinant with three moving quartics and one moving plane. The blending functions are $\sigma_1 = 1$, $\sigma_2 = s$, $\sigma_3 = s^2$, $\sigma_4 = s^3$. Finally, the four bottom patches can be implicitized using a 4×4 determinant with one moving plane, one moving quartic, and two moving quintics. The blending functions are $\sigma_1 = 1$, $\sigma_2 = s$, $\sigma_3 = s^2$, $\sigma_4 = s^3$.

7 DISCUSSION

This paper has proven several theorems and provided empirical support which suggests that the method of moving surfaces is a comprehensive solution to the problem of surface implicitization, and the resulting expressions for the implicit equation are much more compact than those obtained with previous methods.

This work has largely been an adventure in experimental mathematics. The basic notion of moving surfaces was arrived at using pencil and paper, as was the realization that σ moving surfaces can always be found for tensor product and triangular patches such that, if the determinant formed by them (35) does not vanish, it must be the implicit equation. While we have not yet succeeded in theoretically showing that a non-vanishing determinant always exists, empirical substantiation has been provided using computer algebra. We have tried numerous examples with randomly chosen control points (assuring that the surface has no base points or high order singularities), and the blending functions in sections 4.1 and 5 have always worked. Thus, we conjecture that for randomly chosen control points, the methods in sections 4.1 and 5 are robust.

We have also run numerous test cases in which we generated surfaces with randomly chosen simple base points, and have found no counterexample to the conjecture that in randomly chosen cases it always works to trade one quadratic row for a linear row (until there are more base points than quadratic rows), thereby reducing the degree of the surface by one.

Based on our experience, an automatic algorithm for implicitizing bicubic patches (for example) would begin by determining the degree of the implicit equation. This can be done very quickly by firing two skew rays into it [11] and checking how many unique (s, t) pairs they each intersect the surface. This is the degree. If the degree is 18, compute how many moving quadrics exist. If there are exactly nine, they will fill a 9×9 determinant which defines the implicit equation. If there are more than nine, search for moving planes and moving cubics to fill the determinant. If the degree is between 9 and 18, search for r moving planes and s moving quadrics such that $r + 2s =$ degree. If the degree is less than 9, try blending functions (42) with $b_1 = b_2 = 1$

Whether or not a non-zero determinant can always be found which satisfies the conditions in theorem 5 remains an open question. All of the hundreds of cases we have studied thus far have yielded a non-zero determinant, though at present we have no proof that this is always so. We have observed that base points with multiplicity greater than one can produce surprising results. In one example, a case was studied involving a base point that was a common double point on the curves $x(s, t) = y(s, t) = z(s, t) = w(s, t) = 0$ (thus a base point of multiplicity four [13]). The anticipation was that four of the rows on the standard 9×9 determinant could convert from quadratic to linear. In fact, only three moving planes exist! Nonetheless, we succeeded in finding a solution involving an 8×8 determinant with six moving quadrics and two moving planes.

We note that moving surfaces can also solve the inversion problem (given a point on the surface, compute the corresponding parameter values).

REFERENCES

[1] Buchberger, Bruno. Applications of Gröbner Bases in Nonlinear Computational Geometry. In D. Kapur and J. L. Mundy, editors, *Geometric Reasoning*, pages 413–446. Elsevier Science Publisher, MIT Press, 1989.

[2] Chionh, Eng Wee. *Base Points, Resultants, and the Implicit Representation of Rational Surfaces*. PhD thesis, University of Waterloo, 1990.

[3] Chionh, Eng Wee and Ronald N. Goldman. Degree, Multiplicity, and Inversion Formulas for Rational Surfaces using U-resultants. *Computer Aided Geometric Design*, 9:93–108, 1992.

[4] Chionh, Eng Wee and Ronald N. Goldman. Using Multivariate Resultants to Find the Implicit Equation of a Rational Surface. *The Visual Computer*, 8(3):171–180, 1992.

[5] Coolidge, Julian. *A History of the Conic Sections and Quadric Surfaces*. Oxford, 1945.

[6] Cox, David, John Little and Donal O'Shea. *Ideals, Varieties, and Algorithms*. Undergraduate Texts in Mathematics. Springer-Verlag, 1992.

[7] Crow, Frank. The Origins of the Teapot. *IEEE Computer Graphics and Applications*, 7(1):8–19, 1987.

[8] Dixon, A. L. The Eliminant of Three Quantics in Two Independent Variables. *Proceedings of London Mathematical Society*, 6:46–49, 473–492, 1908.

[9] Hoffmann, Christoph. *Geometric and Solid Modeling: An Introduction*. Morgan Kaufmann, 1989.

[10] Hoffmann, Christoph. Implicit Curves and Surfaces in CAGD. *IEEE Computer Graphics & Applications*, 13(1):79–88, 1993.

[11] Kajiya, James. Ray Tracing Parametric Patches. Proceedings of SIGGRAPH 82 (Boston, July 26–30, 1982). In *Computer Graphics*, 16,3 (July 1982), 245–254.

[12] Kalkbrener, Michael. *Three Contributions to Elimination Theory*. PhD thesis, Johannes Kepler Universitat, Linz, Austria, 1991.

[13] Katz, Sheldon and Thomas W. Sederberg. Genus of the Intersection Curve of Two Parametric Surface Patches. *Computer Aided Geometric Design*, 5:253–258, 1988.

[14] Krishnan, S. and Dinesh Manocha. An Efficient Surface Intersection Algorithm Based on the Lower Dimensional Formulation. Technical Report TR94-062, Department of Computer Science, University of North Carolina, 1994.

[15] Manocha, Dinesh and John F. Canny. Algorithms for Implicitizing Rational Parametric Surfaces. *Computer-Aided Geometric Design*, 9:25–50, 1992.

[16] Manocha, Dinesh and John F. Canny. The Implicit Representation of Rational Parametric Surfaces. *Journal of Symbolic Computation*, 13:485–510, 1992.

[17] Salmon, George. *Modern Higher Algebra*. Chelsea, New York, 5th edition, pp. 83–86, 1885.

[18] Salmon, George. *A Treatise on the Analytic Geometry of Three Dimensions*. Longmans, Green and Co., London, 5th edition, p. 264, 1915.

[19] Sederberg, Thomas W. *Implicit and Parametric Curves and Surfaces for Computer Aided Geometric Design*. PhD thesis, Purdue University, 1983.

[20] Sederberg, Thomas W. Improperly Parametrized Rational Curves. *Computer Aided Geometric Design*, 3:67–75, 1986.

[21] Sederberg, Thomas W. Techniques for Cubic Algebraic Surfaces. *IEEE Computer Graphics and Applications*, 10(4):14–26, 1990.

[22] Sederberg, Thomas W., David C. Anderson, and Ronald N. Goldman. Implicit Representation of Parametric Curves and Surfaces. *Computer Vision, Graphics and Image Processing*, 28:72–84, 1984.

[23] Sederberg, Thomas W. and Scott R. Parry. A Comparison of Curve-Curve Intersection Algorithms. *Computer-Aided Design*, 18:58–63, 1986.

[24] Sederberg, Thomas W., Takafumi Saito, Dongxu Qi, and Krzysztof S. Klimaszewski. Curve Implicitization using Moving Lines. *Computer Aided Geometric Design*, 11:687–706, 1994.

Appendix

This appendix contains additional details for the proofs of theorems 2, 4 and 6.

Proof of Theorem 4. *In order to show that $f(\mathbf{X})$ in (19) is not identically zero, we prove the following facts:*

1. *$rank(M) = 2n$, where M is the matrix defined in (15).*

2. *Let $L_i'(t) \cdot \mathbf{X}$, $i = 0, 1, \ldots, m$ be another set of linearly independent moving lines and define*

$$f'(\mathbf{X}) = \begin{vmatrix} L_{00}' \cdot \mathbf{X} & \ldots & L_{0m}' \cdot \mathbf{X} \\ \cdot & \ldots & \cdot \\ \cdot & \ldots & \cdot \\ \cdot & \ldots & \cdot \\ L_{0m}' \cdot \mathbf{X} & \ldots & L_{mm}' \cdot \mathbf{X} \end{vmatrix}$$

   then there exists a non-zero constant c such that

$$f'(\mathbf{X}) = c \cdot f(\mathbf{X}).$$

3. *$f(\mathbf{X}) \not\equiv 0$.*

1. We first assume $x(t)$ and $w(t)$ do not have a common factor. Denote by M_1 the matrix obtained by deleting from M columns $2, 5, \ldots, 3n - 1$. The determinant of M_1 is none other than the resultant of $x(t)$ and $w(t)$, hence is nonzero. Therefore $rank(M) = rank(M_1) = 2n$.

If $x(t)$ and $w(t)$ do have a common factor, there exist constants $\alpha (\alpha \neq 0)$ and β such that $\alpha x(t) + \beta y(t)$ and $w(t)$ are coprime, since curve $\mathbf{P}(t) = (x(t), y(t), w(t))$ doesn't have a base point. Now consider curve $\mathbf{P}'(t) = (\alpha x(t) + \beta y(t), y(t), w(t))$. In this case, the matrix M in (15) is the $2n \times 3n$ matrix $M' =$

$$\begin{pmatrix} x_0' & y_0 & w_0 & & & & & & \\ x_1' & y_1 & w_1 & x_0' & y_0 & w_0 & & & \\ & \vdots & & x_1' & y_1 & w_1 & \ddots & & \\ & & & & & & x_0' & y_0 & w_0 \\ x_n' & y_n & w_n & & \vdots & & x_1' & y_1 & w_1 \\ & & & x_n' & y_n & w_n & & & \vdots \\ & & & & & & \ddots & x_n' & y_n & w_n \end{pmatrix}$$

where $x_i' = \alpha x_i + \beta y_i$, $i = 0, 1, \ldots, n$.

Since $\alpha x(t) + \beta y(t)$ and $w(t)$ do not have common factors, $rank(M') = 2n$. By taking column reductions to matrix M' we have $rank(M) = rank(M') = 2n$.

2. Let

$$L_i' = \sum_{j=0}^{m} L_{ij}' t^j; \quad L_{ij}' = (a_{ij}', b_{ij}', c_{ij}'); \quad i = 0, 1, \ldots, m.$$

From statement 1 we know there are exactly n independent moving lines which follow $\mathbf{P}(t)$. Hence there exist constants d_{ij} such that

$$(L_{i0}', \ldots, L_{im}') = \sum_{j=0}^{m} d_{ij}(L_{j0}, \ldots, L_{jm}), \quad i = 0, 1, \ldots, m.$$

or, in matrix form,

$$L' = D \cdot L,$$

where $L' = (L_{ij}')$, $L = (L_{ij})$, $D = (d_{ij})$.

Since $L_i(t) \cdot \mathbf{X}$, $i = 0, 1, \ldots, m$ and $L_i'(t) \cdot \mathbf{X}$, $i = 0, 1, \ldots, m$ are two sets of independent moving lines, we have $rank(L') = rank(L) = m + 1$, hence $rank(D) = m + 1$, i.e. D is an invertible matrix.

On the other hand, we also have

$$f'(\mathbf{X}) = det(D) \cdot f(\mathbf{X}).$$

Since $det(D) \neq 0$, statement 2 is proved.

3. This statement follows immediately from statement 2 and the fact that Bezout's resultant does not vanish. But we can also prove it without using that result.

First we assume $x(t)$ and $w(t)$ do not have a common factor. We can solve equations $Mb = 0$ in the following way:

Take b_i, $i = 0, 1, \ldots, n - 1$ as free parameters and solve for a_i, $i = 0, 1, \ldots, n - 1$ and c_i, $i = 0, 1, \ldots, n - 1$. Each given set of parameters b_i gives one unique solution since $rank(M_1) = 2n$. If we choose parameters $(b_0, \ldots, b_{n-1})$ to be e_i (creating an n dimensional vector with all components being 0 except i^{th} element being 1), $i = 1, 2, \ldots, n$, respectively, then we get n linearly independent moving lines. Furthermore, if we form determinant (19) from these moving lines, the coefficient for monomial term y^n of $f(\mathbf{X})$ is 1, hence $f(\mathbf{X})$ is not identically zero.

Now suppose $x(t)$ and $w(t)$ do have a common factor, then there exist constants $\alpha (\alpha \neq 0)$ and β such that $\alpha x(t) + \beta y(t)$ and $w(t)$ do not have a common factor. Similarly as in the proof of statement 1 we consider curve $\mathbf{P}'(t) := (\alpha x(t) + \beta y(t), y(t), w(t))$. To find moving lines which follow this curve, we need to solve the linear equation $M' \cdot b' = 0$, where M' is defined as before and

$$b' = (a_0', b_0', c_0', \cdots, a_{n-1}', b_{n-1}', c_{n-1}')^T, \quad i = 0, 1, \ldots, n - 1.$$

Since $\alpha x(t) + \beta y(t)$ and $w(t)$ do not have common factor, the determinant corresponding to (19) formed by the moving lines of curve $\mathbf{P}'(t)$ (denote it by $g(x, y)$) doesn't vanish, thus is the implicit equation of curve $\mathbf{P}'(t)$.

On the other hand, noticing matrices M and M' have the relationship

$$M' = M \cdot B,$$

where

$$B = diag(C, C, \ldots, C),$$

and C is a 3×3 matrix:

$$C = \begin{pmatrix} \alpha & 0 & 0 \\ \beta & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

we know the solutions of $M \cdot b = 0$ and $M' \cdot b' = 0$ have the relationship

$$b = B \cdot b',$$

from which we can easily obtain the determinant formed by the moving lines of curve $\mathbf{P}(t)$ is just $g(\alpha x + \beta y, y)$, hence is nonzero.

This complete the proof of theorem 2. ∎

Proof of Theorem 4:

$1 \Longleftrightarrow 2$.

Suppose the $f_{ij}(\mathbf{X})$ $(i, j = 0, 1)$ in (27) are all quadratics. If $f(\mathbf{X}) \equiv 0$, then $f_{00}(\mathbf{X}) f_{11}(\mathbf{X}) \equiv f_{10}(\mathbf{X}) f_{01}(\mathbf{X})$. We show that $f_{ij}(\mathbf{X})$ $(i, j = 0, 1)$ all can be factored into linears. The proof is by contradiction.

Suppose for example $f_{00}(\mathbf{X})$ is prime, then there exists a constant c such that $f_{10}(\mathbf{X}) = c f_{00}(\mathbf{X})$ or $f_{01}(\mathbf{X}) = c f_{00}(\mathbf{X})$. If

$f_{10}(\mathbf{X}) = c f_{00}(\mathbf{X})$, then $f_{11}(\mathbf{X}) = c f_{01}(\mathbf{X})$, this contradicts the fact that the coefficients of two rows are linearly independent. If $f_{01}(\mathbf{X}) = c f_{00}(\mathbf{X})$, then $f_{11}(\mathbf{X}) = c f_{10}(\mathbf{X})$. But $f_{i0}(\mathbf{P}(t)) + t f_{i1}(\mathbf{P}(t)) \equiv 0$, $i = 0, 1$, so $f_{i0}(\mathbf{P}(t)) \equiv 0$, means that $\mathbf{P}(t)$ is in fact a degree two curve and this contradicts the fact that $\mathbf{P}(t)$ doesn't have base point.

Now let $f_{00}(\mathbf{X}) = l_0(\mathbf{X}) \, l_1(\mathbf{X})$ and $f_{11}(\mathbf{X}) = l_2(\mathbf{X}) \, l_3(\mathbf{X})$, so $f_{10}(\mathbf{X}) = l_0(\mathbf{X}) \, l_2(\mathbf{X})$ and $f_{01}(\mathbf{X}) = l_1(\mathbf{X}) \, l_3(\mathbf{X})$. Hence we have $l_0(\mathbf{P}(t)) + t l_3(\mathbf{P}(t)) \equiv 0$. that is, there exists a moving line which follows $\mathbf{P}(t)$.

Conversely, if there exists a moving line $L(\mathbf{X}; t)$ which follows $\mathbf{P}(t)$, then every moving quadratic which follows $\mathbf{P}(t)$ must be $L(\mathbf{X}; t)$ scaled by a linear function in x and y. This is so because if the moving line and the moving quadratic were to form a 2×2 determinant which is not identically zero, it would be the implicit equation according to theorem 3. But that expression would be cubic in x, y and the true implicit equation is a quartic because there is no base point. Hence $f(\mathbf{X}) \equiv 0$.

$2 \Longleftrightarrow 4$.

Suppose moving line $L(\mathbf{X}; t) = (a_0 x + b_0 y + c_0 w) + t(a_1 x + b_1 y + c_1 w)$ follows $\mathbf{P}(t)$, so $L(\mathbf{P}(t); t) \equiv 0$. This is equivalent to a system of homogeneous linear equations with a_0, b_0, c_0, a_1, b_1 and c_1 being unknowns. It is easy to know that the coefficient matrix is none other than the matrix in (28). Hence the equations have non-zero solutions if and only if (28) holds.

$2 \Longleftrightarrow 3$.

Suppose $\mathbf{P}(t)$ has a triple point $\mathbf{T}$. Let $L_0(\mathbf{X}) = 0$ be the line which passes through $\mathbf{T}$ and $\mathbf{P}(0)$, and $L_1(\mathbf{X}) = 0$ be the line which passes through $\mathbf{T}$ and $\mathbf{P}(1)$. Define $L(\mathbf{X}; t) = (1 - t) c L_0(\mathbf{X}) + t L_1(\mathbf{X})$. We prove that there exists a scalar c for which $L(\mathbf{X}; t)$ follows $\mathbf{P}(t)$. For this, the intersection of line $L(\mathbf{X}; t)$ with curve $\mathbf{P}(t)$ can be expressed as the degree five polynomial $L(\mathbf{P}(t); t)$. This polynomial has zeros at the three triple point parameters, as well as zeros at $t = 0$ and $t = 1$. We can choose c so that $L(\mathbf{P}(t); t) = 0$ for one other value of t, which makes it identically zero, and hence $L(\mathbf{X}; t)$ follows $\mathbf{P}(t)$.

Conversely, if there exists a pencil of lines which follows a quartic curve, the pencil axis must lie at a triple point on the curve, because as t sweeps monotonically from $-\infty$ to $+\infty$, $L(\mathbf{X}; t)$ rotates monotonically $180°$ about a fixed axis $\mathbf{A}$. However, if $L(\mathbf{X}; t)$ were to intersect $\mathbf{P}(t)$ at more than one point other than $\mathbf{A}$, then as t sweeps monotonically from $-\infty$ to $+\infty$, the line $\mathbf{A}$–$\mathbf{P}(t)$ would not rotate monotonically.

The equivalence of the above four statments shows that if the quartic curve doesn't have a triple point then the implicit equation can be formed by two quadratic rows, and if the quartic curve does have a triple point we can find a moving line following it. If we multiply the moving line by a degree two polynomial in x, y, w (that has six coefficients), we see that there is a six parameter family of moving cubics that are degree one in t which also follow the quartic curve. In the following we prove we can also find a moving cubic which follows the curve and which is not just a quadratic scale of the moving line. To do this, we only need to show there exists at least seven moving cubics.

Suppose the quartic curve $\mathbf{P}(t)$ has a triple point $\mathbf{T} = \mathbf{P}(t_1) = \mathbf{P}(t_2) = \mathbf{P}(t_3)$. Without loss of generality, we assume that $t_1 \neq t_2 \neq t_3$ (if some of them are equal the discussion will involve derivatives). To find the moving cubics $C(\mathbf{X}; t) := f_0(\mathbf{X}) + t f_1(\mathbf{X})$ which follows the curve $\mathbf{P}(t)$, we just substitute the equation of $\mathbf{P}(t)$ into $C(\mathbf{X}; t)$ and get a degree thirteen polynomial $g(t)$, then choose fourteen different parameters $s_1, s_2, \ldots, s_{14}$ and obtain fourteen equations $g(s_i) = 0$, $i = 1, 2, \ldots, 14$. Specifically we can choose three of the parameters be t_1, t_2 and t_3 and get three equations $f_0(\mathbf{T}) + t_i f_1(\mathbf{T}) = 0$, $i = 0, 1, 2$. These three equations are linearly dependent because they are equivalent to $f_0(\mathbf{T}) = f_1(\mathbf{T}) = 0$.

Hence we only have thirteen equations, but there are twenty unknowns, so there exist at least seven moving cubics which follow the quartic curve. ∎

Proof of Theorem 6: In the theorem, we impose the requirement that the ρ base points are in "general position". Here we define what that means. Given a set of points (s_i, t_i), $i = 1, \ldots, k$, define

$$v_i = (1, \ldots, s_i^{2d_1 - 1}, t_i, \ldots, t_i s_i^{2d_1 - 1}, \ldots$$
$$\ldots, t_i^{2d_2 - 1}, \ldots, t_i^{2d_2 - 1} s_i^{2d_1 - 1}), \quad i = 1, \ldots, k.$$

The points are in general position if the vectors v_i, $i = 1, \ldots, k$, are linearly independent.

We prove the following facts:

1. If (s_i, t_i), $i = 1, \ldots, k$ ($k \leq d := 4 d_1 d_2$) are in general position, then there exist additional points (s_i, t_i), $i = k + 1, \ldots, d$ such that (s_i, t_i), $i = 1, \ldots, d$ are also in general position, that is,

$$M = (v_1^T, v_2^T, \ldots, v_d^T)^T$$

   has full rank.

2. Let

$$h(s, t) := \sum_{i=1}^{d_1 d_2} h_i(\mathbf{X}(s, t)) \gamma_i(s, t),$$

   where $h_i(\mathbf{X})$ are linear functions in x, y, z and $\gamma_i(s, t)$ are blending functions defined in (42). If (s_i, t_i), $i = 1, \ldots, d$ are in general position, then $h(s, t) \equiv 0$ iff

$$h(s_i, t_i) = 0, \quad i = 1, \ldots, d.$$

3. If $\mathbf{P}(s, t)$ has ρ distinct base points and these base points are in general position, then there are at least ρ moving planes which follow $\mathbf{P}(s, t)$.

1. We only need to prove that there exists (s_{k+1}, t_{k+1}) such that (s_i, t_i), $i = 1, \ldots, k + 1$ are in general position, the general case follows from induction.

By assumption, (s_i, t_i), $i = 1, \ldots, k$ are in general position, that is $v_1, \ldots, v_k$ are linearly independent. Therefore there exists a $k \times k$ submatrix of M, say $M_1 = M \begin{pmatrix} 1 & 2 & \ldots & k \\ j_1 & j_2 & \ldots & j_k \end{pmatrix}$ such that $det(M_1) \neq 0$. Now define

$$M' = (v_1^T, \ldots, v_k^T, v^T)^T,$$

where $v = (1, \ldots, s^{2d_1 - 1}, t, \ldots, t s^{2d_1 - 1}, \ldots, t^{2d_2 - 1}, \ldots, t^{2d_2 - 1} s^{2d_1 - 1})$, and consider a $(k + 1) \times (k + 1)$ submatrix of M',

$$M_1' = M' \begin{pmatrix} 1 & 2 & \ldots & k & k+1 \\ j_1 & j_2 & \ldots & j_k & l \end{pmatrix},$$

here $l \neq j_i, i = 1, \ldots, k$. Because $det(M_1) \neq 0$, $det(M_1')$ is a non-zero polynomial in s and t (denoted by $f(s, t)$). Hence there exists some $(s, t) = (s_{k+1}, t_{k+1})$ such that $det(M_1') = f(s_{k+1}, t_{k+1}) \neq 0$. Thus M_1' has full rank, i.e., (s_i, t_i), $i = 1, \ldots, k + 1$ are in general position.

2. Noticing $h(s, t)$ is a tensor product polynomial of degree $2 d_1 - 1$ in s and degree $2 d_2 - 1$ in t, we can rewrite $h(s, t)$ as

$$h(s, t) = \sum_{i=0}^{2 d_1 - 1} \sum_{j=0}^{2 d_2 - 1} h_{ij} s^i t^j.$$

If $h(s,t) \equiv 0$, we certainly have $h(s_i, t_i) = 0$, $i = 1, \ldots, d$.

Conversely, if $h(s_i, t_i) = 0$, $i = 1, \ldots, d$, we get $M \cdot b = 0$, where M is define as before and $b = (h_{00}, \ldots, h_{2d_1-1,0}, h_{01}, \ldots, h_{2d_1-1,1}, \ldots, h_{0,2d_2-1}, \ldots, h_{2d_1-1,2d_2-1})$. Since M is an invertible matrix we must have $b = 0$, thus $h(s,t) \equiv 0$.

3. Let (s_i, t_i), $i = 1, \ldots, \rho$ be the ρ base points of $\mathbf{P}(s,t)$ and these base points are in general position. According to statement 1, we can find pairs (s_i, t_i), $i = \rho + 1, \ldots, d$ such that (s_i, t_i), $i = 1, \ldots, d$ are in general position. Therefore, $h(s,t) \equiv 0$ is equivalent to $(s_i, t_i) = 0$, $i = 1, \ldots, d$ by statement 2.

But since (s_i, t_i), $i = 1, \ldots, \rho$ are base points of $\mathbf{P}(s,t)$, equations $h(s_i, t_i) = 0$, $i = 1, \ldots, \rho$ are identities. Hence $h(s,t) \equiv 0$ has at least ρ solutions, i.e. there are at least ρ moving planes which follow $\mathbf{P}(s,t)$. ∎