Polynomial Curves and Surfaces

Chandrajit Bajaj and Andrew Gillette

September 8, 2010

Contents

1 What is an Algebraic Curve or Surface?

We review some basic terminology from algebraic geometry that we shall use in subsequent sections. These and additional facts can be found for example in [38].

The set of real and complex solutions (or *zero* set $Z(C)$) of a collection C of polynomial equations

$$
\begin{aligned}
f_1(x_1, ..., x_d) &= 0 \\
&\vdots \\
f_m(x_1, ..., x_d) &= 0
\end{aligned}
\tag{1}
$$

with coefficients over the reals $\mathbb{R}$ or complexes $\mathbb{C}$, is referred to as an *algebraic set.* The algebraic set defined by a single equation ($m = 1$) is also known as a hypersurface. A algebraic set that cannot be represented as the union of two other distinct algebraic sets, neither containing the other, is said to be *irreducible.* An irreducible algebraic set $Z(C)$ is also known as an *algebraic variety* V.

A hypersurface in $\mathbb{R}^d$, some d dimensional space, is of *dimension* $d-1$. The *dimension* of an algebraic variety V is k if its points can be put in one-to-one rational correspondence with the points of an irreducible hypersurface in $k+1$ dimensional space. In $\mathbb{R}^d$, a variety V_1 of dimension k intersects a variety V_2 of dimension h, with $h \geq d-k$, in an algebraic set $Z(S)$ of dimension at least $h+k-d$. The resulting intersection is termed *proper* if all subvarieties of $Z(S)$ are of the same minimum dimension $h+k-n$. Otherwise the intersection is termed *excess* or *improper.*

Let the *algebraic degree* of an algebraic variety V be the *maximum* degree of any defining polynomial. A degree 1 hypersurface is also called a *hyperplane* while a degree 1 algebraic variety of dimension k is also called a *k-flat.* The *geometric degree* of a variety V of dimension k in some $\mathbb{R}^d$ is the maximum number of intersections between V and a $(d-k)$-flat, counting both real and complex intersections and intersections at infinity. Hence the geometric degree of an algebraic hypersurface is the maximum number of intersections between the hypersurface and a line, counting both real and complex intersections and at infinity.

The following theorem, perhaps the oldest in algebraic geometry, summarizes the resulting geometric degree of intersections of varieties of different degrees.

Theorem 1.1 (Bezout)**.** *A variety of geometric degree p which* properly *intersects a variety of geometric degree q does so in an algebraic set of geometric degree either at most pq or infinitely often.*

The *normal* or *gradient* of a hypersurface $\mathcal{H} : f(x_1, ..., x_n) = 0$ is the vector $\nabla f = (f_{x_1}, f_{x_2}, \ldots, f_{x_n})$. A point $\mathbf{p} = (a_0, a_1, \ldots a_n)$ on a hypersurface is a *regular* point if the gradient at $\mathbf{p}$ is not null; otherwise the point is *singular.* A singular point $\mathbf{q}$ is of multiplicity e for a hypersurface $\mathcal{H}$ of degree d if any line through $\mathbf{q}$ meets $\mathcal{H}$ in at most $d-e$ additional points. Similarly a singular point $\mathbf{q}$ is of multiplicity e for a variety V in $\mathbb{R}^n$ of dimension k and degree d if any sub-space $\mathbb{R}^{n-k}$ through $\mathbf{q}$ meets V in at most $d-e$ additional points. It is important to note that even if two varieties intersect in a *proper* manner, their intersection in general may consist of sub-varieties of various multiplicities. The total degree of the intersection, however is bounded by Bezout's theorem. Finally, one notes that a hypersurface $f(x_1, ..., x_n) = 0$ of degree d has $K = \binom{n+d}{n}$ coefficients and one less than that number of independent coefficients. Hypersurfaces $f(x_1, ..., x_n) = 0$ of degree d form $K - 1$ dimensional vector spaces over the field of coefficients of the polynomials.

Finally, two hypersurfaces $f(x_1, \ldots, x_n) = 0$ and $g(x_1, \ldots, x_n) = 0$ meet with C^k-continuity along a common subvariety V if and only if there exist functions $\alpha(x_1, \ldots, x_n)$ and $\beta(x_1, \ldots, x_n)$ such that all derivatives up to order k of $\alpha f - \beta g$ equals zero at all points along V, see for instance [25].

1.1 Algebraic Curves

Two dimensional curves are defined as plane curves. They thus have a reduced representation when compared with space curves and can be parameterized (if possible) more efficiently. Algebraic plane curves are defined as $f(x, y) = 0$ and the parametric representation is $\{x = f_1(t) \text{ and } y = f_2(t)\}$. All degree two curves are rational. Degree three curves which are non-singular like ellipses are not. In general, curves with degree higher than two need not be rational. We will next give the conditional for rationality. The genus of a curve is defined as a measure of how much the curve is deficient from its maximum allowable limit of singularities. For a curve of degree d, genus g is given as $g = 1/2(d-1)(d-2) - DP$, where DP stands for Double Points, a sum of all singularities of the curve C_d which are counted.

Cayley - Riemann criterion The genus of a curve is zero if and only if the curve has a rational parametrization.

1.2 Algebraic Surfaces

Why is low degree important? The **geometric degree of an algebraic surface** is the maximum number of intersections between the surface and a line, counting complex, infinite and multiple intersections. It is a measure of the "wavi-ness" of the surface. This geometric degree is the same as the degree of the defining polynomial f of the algebraic surface in the implicit definition, but may be as high as n^2 for a parametrically defined surface with rational functions G_i of degree n. The **geometric degree of an algebraic space curve** is the maximum number of intersections between the curve and a plane, counting complex, infinite and multiple intersections. A well known theorem of algebraic geometry (**Bezout**'s theorem) states that the geometric degree of an algebraic intersection curve of two algebraic surfaces may be as large as the product of the geometric degrees of the two surfaces [38]. The use of low degree surface patches to construct models of physical objects thus results in faster computations for subsequent geometric model manipulation operations such as computer graphics display, animation, and physical object simulations, since the time complexity of these manipulations is a direct function of the degree of the involved curves and surfaces. Furthermore, the number of singularities[1] (sources of numerical ill-conditioning) of a curve of geometric degree m may be as high as m^2 [40]. Keeping the degree low of the curves and surfaces thus leads to potentially more robust numerical computations.

Theorem 1.2. *[Castelnuovo] An algebraic surface S is rational if and only if the arithmetic genus(S)= second pluri-genus (S) = 0.*

A proof can be found in [43].

[1] Points on the curve where all derivatives are zero

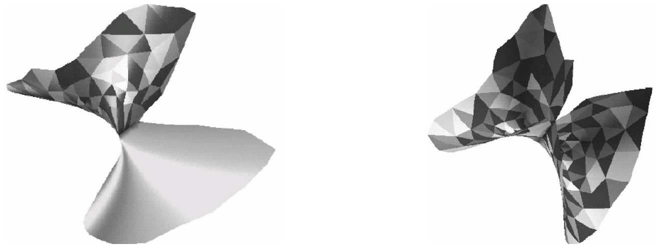
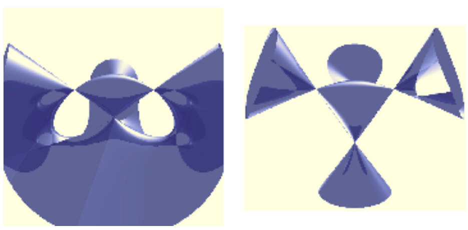
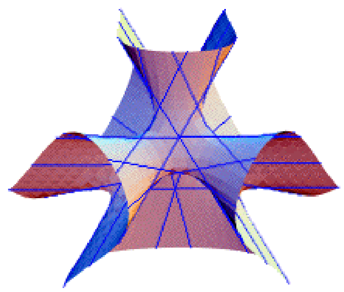
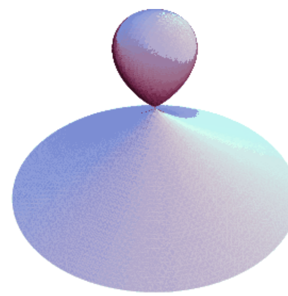


Figure 1: Cartan Surface $f = x^2 - yz^2 = 0$

2 Singularities and Extreme Points

		
The Clebsch Diagonal Cubic	The Cayley Cubic	The Ding-Dong Surface
$81x^3+81y^3+81z^3-189x^2y-189x^2z$ $-189y^2x-189y^2z-189z^2x-189z^2y$ $+54xyz+126xy+126xz+126yz-$ $9x^2,-9y^2-9z^2-9x-9y-9z+1$	$-5x^2y - 5x^2z - 5y^2x - 5y^2z - 5z^2y - 5z^2x + 2xy + 2xz + 2yz$	$x^2+y^2-(1-z)z^2$
(27 real lines with 10 triple points)	(9 real lines = 6 connecting 4 double points, and 3 in a coplanar config)	

2.1 Singularities and Genus

Consider an irreducible plane algebraic curve C_d of degree d. Lines through a point P intersects C_d (outside P) in general at $d - mult_p C_d$ points, where $mult_p C_d = e =$ multiplicity of C_d at $P =$ order at P of the polynomial equation describing C_d . The order of a polynomial equation at a point $P = (a, b)$, is the minimum $(i+j)$, when the polynomial is expressed with terms $(x-a)^i(y-b)^j$. If $e = 0$: P is not on C_d. If $e = 1$ then P is called a simple point. If $e > 1$ we say P is a *singular point* of the curve C_d with multiplicity e or an e-fold point. A 2-fold point is also called a double point and a 3-fold point a triple point.

By Bezout's theorem one may see that the maximum number of double points of C_d is $\leq \frac{(d-1)(d-2)}{2}$. Further, the number of independent conditions needed to specify C_d is $\frac{(d+2)(d+1)}{2} - 1$. One definition of the genus G of a curve C_d is a measure of how much the curve is deficient from its maximum allowable limit of singularities,

$$G = \frac{(d-1)(d-2)}{2} - DP \quad (2)$$

where DP is a 'proper' counting of the number of double points of C_d (summing over all singularities, in the projective complex plane).

Distinct singularities of a plane curve can computationally be obtained by simultaneously solving for the roots of the system of polynomial equations $f = f_x = f_y = 0$ where f_x and f_y are the x and y partial derivatives of f, respectively. One way of obtaining the common solutions is to find those roots of $Res_x(f_x, f_y) = 0$ and $Res_y(f_x, f_y) = 0$ which are also the roots of $f = 0$. Here $Res_x(f_x, f_y)$ (similarly $Res_y(f_x, f_y)$) is the Sylvester resultant of f_x and f_y treating them as polynomials in x (similarly y). For a classical treatment of the Sylvester resultant see [32]. Other methods of computing the roots of a system of polynomial equations, for example via the $U-$resultant may also be used [31]. This method yields an overall time bound of $O(d^6 + T(d^2))$ for computing all the $O(d^2)$ possible singularities of C_d, using the Sylvester resultant which for two j-variate polynomials of maximum degree d can be computed in $O(d^{2j} log^3 d)$ time [12]. Note that singularities at infinity can be obtained in a similar way after replacing the line at infinity with one of the affine coordinate axes. In particular, on homogenizing $f(x, y)$ to $F(X, Y, Z)$ we can set $Y = 1$ to obtain $\tilde{f}(x, z)$ thereby swapping the line at infinity $Z = 0$ with the line $Y = 0$. Now the above computation of roots can be applied to $\tilde{f} = \tilde{f}_x = \tilde{f}_z = 0$ to compute singularities at infinity.

Having computed the singular points one next obtains a proper count of the total number of double points DP of C_d. A proper counting was achieved by Noether using (projective) Cremona quadratic transformations, see [40] Following [2], the same can be achieved using (affine) quadratic transformations.

Affine Quadratic Transforms In a general procedure for counting double points, given an e-fold point P of a plane curve C_d, we choose our coordinates to bring P to the origin and then apply the quadratic transformation Q_1 or Q_2.

$$Q_1 \quad : \qquad x \;=\; x_1 \;\;, \qquad y \;=\; x_1 y_1 \tag{3}$$

$$Q_2 \quad : \qquad x \;=\; x_2 y_2 \;\;, \qquad y \;=\; y_2 \tag{4}$$

Affine quadratic transformations are centered on a singularity and affect the curve locally, allowing one to treat each singularity of C_d in isolation. If now $C_d : \; f(x,\; y) \;=\; 0$, then the quadratic transformation Q_1 transforms C_d into the curve $C^1 : \; f_1(x_1,\; y_1) \;=\; 0$ given by

$$f(x_1,\; x_1\, y_1) \;=\; {x_1}^e \; f_1(x_1, y_1)$$

C^1 will intersect the exceptional line $E \; : \; x_1 \;=\; 0$ in the points $P^1 \;, ..., \; P^m$, the roots of $f_1(0,\; y) \;=\; 0$. If P^i is a e_i-fold point of C^1, then we shall have $e_1 + ... + e_m \;\leq\; e$. The $P^1 \;, ..., \; P^m$ are termed the points of C_d in the first neighborhood of P. The quadratic transformations can be repeated at each of the P^ipoints of C^1 with $e_i \;>\; 1$, yielding points P^{ij} points in the second neighborhood of P and so on. The collection of these neighborhod points are termed the points *infinitely near* P and form in general a *singularity tree* at P. At each node of this tree (including the root) keep a count equal to the multiplicity of the curve (transformed curve) at that point. The desingularization theorem for algebraic plane curves, see [2, 40], states that at every node beyond a certain level, the count equals one; in other words, C has only a finite number of singular points infinitely near P. Next (using Bezout) take $\frac{e(e-1)}{2}$ double points towards DP for a count e and sum over all nodes of a singularity tree and additionally over all singularities of C_d and their corresponding singularity trees, to obtain a precise count for the total number of double points DP of C_d. This proper counting of double points then yields the genus of C_d via the above genus formula, (1).

Theorem 2.1: The Genus G for C_d of degree d can be computed in $O(d^6 \;+\; d^2 T(d^2))$ time.

Proof. The time taken to compute G is bound by the time $O(d^6 \;+\; T(d^2))$ taken to compute the $O(d^2)$ possible singular points of C_d, plus the time taken by the refinement of singularities via quadratic transformations, which we now bound. As many as $O(d^2)$ quadratic transformations may be needed for all *infinitely near* singularities of C_d where a single quadratic transformation takes $O(d^2)$ time. Then there is the $O(d^2\ T(d^2))$ time spent in computing intersections with the exceptional line accounting also for a degree blowup of $O(d^2)$ for the transformed curve in a sequence of quadratic transformations. Additionally, there is the time spent in translating the singularity to the origin which entails an algebraic simplification with an overall cost of $O(d^4)$. This results in the overall time bound of $O(d^6 + d^2\ T(d^2))$. □

There is then the concise characterization for curves having rational parametric equations

Theorem [Cayley-Riemann]: C_d has a rational parameterization *iff* $G = 0$.

In other words if the given plane curve has its maximum allowable limit of singularities, then it is rational.

2.2 Parameterizing with a Pencil of Lines

From Cayley-Riemann Theorem of the earlier section, we know that all degree d curves C_d with one distinct $d-1$ fold point, are rational. One way then of parameterizing these curves C_d is to symbolically intersect them with a pencil of lines $(y-y_0) = t(x-x_0)$ through the $d-1$ fold point (x_0, y_0) on the curve. This pencil intersects C_d in only one additional point, the coordinates of which can be expressed as rational functions of the parameter t. Alternatively, the same can be achieved by mapping the $d-1$ fold point on C_d to infinity along one of the coordinate axis. We illustrate this below.

Mapping Points to Infinity Consider $f(x,\ y)$ a polynomial of degree d in x and y representing a plane algebraic curve C_d of degree d with a *distinct* $d-1$ fold singularity. We first determine the $d-1$ fold singularity of the curve C_d and translate it to the origin. Then we can write

$$f(x,\ y) = f_d(x,\ y) + f_{d-1}(x,\ y) = 0$$

where f_i consists of the terms of degree i. Note that f_d and f_{d-1} are the only terms that will exist, since a $d-1$ fold singularity at the origin implies that $\forall (i+j) < d-1$, $\frac{\partial f^{i+j}}{\partial x^i \partial y^j} = 0$ at $(0,0)$.

On homogenizing $f(x,y)$ we obtain

$$\begin{aligned} F(X,Y,Z) = \quad & a_0Y^d + a_1Y^{d-1}X + \ldots + a_dX^d \\ & + b_0Y^{d-1}Z + b_1Y^{d-2}XZ + \ldots + b_dX^{d-1}Z = 0 \end{aligned} \tag{5}$$

Now by sending the singular point $(0,0,1)$ to infinity along the Y axis we eliminate the Y^d term. Algebraically this is achieved by a homogeneous linear transformation which maps the point $(0,0,1)$ to the point $(0,1,0)$ and is given by $X = X_1$, $Y = Z_1$, $Z = Y_1$, which yields

$$\begin{aligned} F(X_1,Y_1,Z_1) = a_0Z_1^d \quad & + \quad a_1Z_1^{d-1}X_1 + \ldots + a_dX_1^d \\ & + \quad b_0Z_1^{d-1}Y_1 + b_1Z_1^{d-2}X_1Y_1 + \ldots + b_dX_1^{d-1}Y_1 = 0 \end{aligned} \tag{6}$$

Then one easily obtains

$$Y_1 = -\frac{a_0Z_1^d + a_1Z_1^{d-1}X_1 + \ldots + a_dX_1^d}{b_0Z_1^{d-1} + b_1Z_1^{d-2}X_1 + \ldots + b_dX_1^{d-1}} \tag{7}$$

Letting $X_1 = t$ and dehomogenizing by setting $Z_1 = 1$ and using the earlier homogeneous linear transformation, we construct the original affine coordinates

$$\begin{aligned} x &= \frac{X}{Z} = \frac{X_1}{Y_1} \\ y &= \frac{Y}{Z} = \frac{Z_1}{Y_1} \end{aligned} \tag{8}$$

as rational functions of the single parameter t.

Theorem 3.1: An algebraic plane curve of degree d with a *distinct* $d-1$ fold point can be rationally parameterized in $O(d^4 log^3 d)$ time.

Proof. The time taken to determine the $d-1$-fold singularity is bound by $O(d^4 log^3 d)$ the time taken to determine a single mutliple root of a univariate polynomial of degree d is $O(d\ log^2 d$ [30]. This also yields the overall time bound, since the homogeneous linear transformation after a translation of the singularity to the origin, is bound by $O(d^4)$. □

2.3 Parameterizing with a Pencil of Curves

In the general case we consider a curve C_d with the appropriate number of *distinct* and *infinitely near* singularities which make C_d rational (*genus* 0). We pass a pencil of curves $C_{d-2}(t)$ through these singular points and $d-3$ additional simple points of C_d. This pencil intersects C_d in only one additional point, the coordinates of which can be expressed as rational functions of the parameter t.

Let $F(X, Y, Z) = 0$ and $G(X, Y, Z) = 0$ be the homogeneous equations of the curves C_d and $C_{d-2}(t)$ respectively. For a distinct singular point of multiplicity m of C_d at the point (X_i, Y_i, Z_i) we pass the curve $C_{d-2}(t)$ through it with multiplicity $m-1$. To achieve this we equate

$$G(X_i, Y_i, Z_i) \;=\; F(X_i, Y_i, Z_i) \;=\; 0 \qquad (9)$$

$$G_{X^jY^k}(X_i, Y_i, Z_i) \;=\; F_{X^jY^k}(X_i, Y_i, Z_i) \;=\; 0\ ,\quad 1 \leq j + k \leq m-2 \qquad (10)$$

where $G_{X^jY^k} = \frac{\partial G^{j+k}}{\partial X^j \partial Y^k}$. Similarly for $F_{X^jY^k}$.

For an *infinitely near* singular point of C_d we construct its associated *singularity tree* and pass the curve $C_{d-2}(t)$ with multiplicity $r-1$ through each of the points of multiplicity r in the first, second, third, ..., neighborhoods. To achieve this we apply quadratic transformations T_i to both $F(X, Y, Z)$ and $G(X, Y, Z)$ centered around the *infinitely near* singular points corresponding to the singularity tree. The appropriate multiplicity of passing is achieved by equating the transformed equations F_{T_i} and G_{T_i} and their partial derivatives as above. All the above conditions in totality lead to a square system of homogeneous linear equations where the unknowns are the coefficients of $C_{d-2}(t)$ having one variable parameter t.

A counting argument shows that this method generates the correct number of conditions which specifies $C_{d-2}(t)$ and furthermore the total intersection count between C_d and $C_{d-2}(t)$ satisfies *Bezout*. A curve C_d of *genus* $= 0$ has the equivalent of exactly $\frac{(d-1)(d-2)}{2}$ double points. To pass a curve $C_{d-2}(t)$ through these double points and $d-3$ other fixed simple points of C_d, the total number of conditions (= the total number of linear equations) is given by

$$\frac{(d-1)(d-2)}{2} + (d-3) = \frac{d\,(d-1)}{2} - 2$$

which is exactly the number of independent unknowns to determine a pencil of $C_{d-2}(t)$. Having determined the pencil of $C_{d-2}(t)$ curves we compute the resultant $Res_x(C_d,\ C_{d-2}(t))$ which yields a polynomial of degree $d(d-2)$ in y which on dividing by the common factors corresponding to the $(d-3)$ simple points and $\frac{(d-2)(d-1)}{2}$ double points, yields a polynomial in y and t which is linear in y and thereby gives y as a rational function of t. Similarly repeating with $Res_y(C_d,\ C_{d-2}(t))$ yields x as a rational function of t.

Theorem 4.1: A rational algebraic plane curve of degree d can be rationally parameterized in $O(d^6 log^3 d + d^2 T(d^2))$ time.

Proof. The time taken to compute the $O(d^2)$ point singularities with refinement for infinitely near singularitites is bound as before by the time $O(d^6 + d^2 T(d^2))$. The time taken to determine $d-3$

simple points requires at worst no more than $O(d\,T(d))$ time (most points are simple). Then there is the time taken to solve the homogeneous linear system of size $O(d^2)$. Using a technique similar to Gaussian elimination (which requires $O(d^3)$ for a linear system of size d), the time can be bound by $O(d^6)$. Finally there is the computation of the resultants of the equations for C_d and C_{d-2} involving variables x, y and t and division by univariate polynomials [7], all bound by $O(d^6 log^3 d)$ time. Hence the overall time bound above. □

2.4 Algebraic Space Curves

Consider an irreducible algebraic space curve C_d of degree d, which is implicitly defined as the intersection of two algebraic surfaces $f(x,y,z) = 0$ and $g(x,y,z) = 0$. There always exists a birational correspondence between the points of C_d and the points of an irreducible plane curve P_d of degree d, whose genus is the same as that of C_d [1]. Birational correspondence between C_d and P_d means that the points of C_d can be given by rational functions of points of P_d and vice versa (i.e there exists a 1-1 mapping between points of C_d and P_d, except for a finite number of exceptional points). Consequently, knowing how to compute the genus and rational parameterization of algebraic plane curves from sections 2, 3 and 4, yields an algorithm to compute the genus of the space curve C_d and if genus = 0 the rational parametric equations of C_d.

To determine the equation of the plane curve P_d we consider the projection of the space curve C_d along one of the coordinate axis. Projecting C_d along, say the z axis, can be achieved by treating both f and g as polynomials in z with coefficients in x and y and then computing the Sylvester resultant. The resultant yields a polynomial in the coefficients of f and g, viz., a plane curve P_d described by the polynomial in x and y. However this projected plane curve P_d in general, is not in birational correspondence with the space curve C_d. For a chosen projection direction it is quite possible that most points of P_d may correspond to more than one point of C_d (i.e. a multiple covering of P_d by C_d). However this may be rectified by choosing a valid projection direction.

Valid Projection Direction To find an appropriate axis of projection, the following general procedure may be adopted. Consider the linear transformation $x = a_1x_1 + b_1y_1 + c_1z_1$, $y = a_2x_1 + b_2y_1 + c_2z_1$ and $z = a_3x_1 + b_3y_1 + c_3z_1$. On substituting into the equations of the two surfaces defining the space curve we obtain the transformed equations $f_1(x_1, y_1, z_1) = 0$ and $g_1(x_1, y_1, z_1) = 0$. Next compute the $Res_{z_1}(f_1, g_1)$ which yields a polynomial $h(x_1, y_1)$ which is the equation of the projected plane curve. Choose the coefficients of the linear transformation, a_i, b_i and c_i such that (i) the determinant of a_i, b_i and c_i is non zero and (ii) the equation of the projected plane curve $h(x_1, y_1)$ is not a power of an irreducible polynomial. The latter can be achieved by making the discriminant $Res_{x_1}(h, h_{x_1})$ to be non zero. Such a choice of coefficients ensures that the projected irreducible plane curve given by $h(x_1, y_1)$ is in birational correspondence with the irreducible space curve and thus of the same genus. As "bad" values for a_i, b_i, c_i, $i = 1 \ldots 3$, satisfy a lower dimension hypersurface, any random choice of values will suffice with probability 1, see [34].

Constructing the Birational Map There remains the problem of constructing the birational mapping between points on P_d and C_d. Let the projected plane curve P_d be defined by the polynomial $h(x_1, y_1)$. The map one way is linear and is given trivially by $x_1 = x$ and $y_1 = y$. To construct the reverse rational map one only needs to compute $z = I(x_1, y_1)$ where I is a rational function. We now show how it is always possible to construct this rational function by use of a polynomial remainder sequence along a valid projection direction.

Let the surfaces $f(x,y,z) = 0$ and $g(x,y,z) = 0$ be of degrees m_1 and m_2 respectively. Without loss of generality let this direction be the z axis and that $m_1 \geq m_2$. Both m_1 and m_2 are bound by d, the degree of the space curve C_d. Let $F_1 = f(x,y,z)$ and $F_2 = g(x,y,z)$ be given by

$$
\begin{aligned}
F_1 &= f_0\, z^{m_1} + f_1\, z^{m_1 - 1} + \ldots + f_{m_1 - 1}\, z + f_{m_1} \\
F_2 &= g_0\, z^{m_2} + g_1\, z^{m_2 - 1} + \ldots + g_{m_2 - 1}\, z + g_{m_2}
\end{aligned} \tag{11}
$$

with f_j, $(j = 0 \ldots m_1)$ and g_k, $(k = 0 \ldots m_2)$, denoting polynomials in x, y. Then, there exist polynomials $F_{i+2}(x, y, z)$, for $i = 1 \ldots k$, such that $A_i\ F_i = Q_i\ F_{i+1} + B_i\ F_{i+2}$ where m_{i+2}, the degree of z in F_{i+2}, is less than m_{i+1}, the degree of z in F_{i+1} and certain polynomials $A_i(x, y)$, $Q_i(x, y, z)$ and $B_i(x, y)$. The polynomials F_{i+2}, $i = 1, 2, \ldots$ form, what is known as a polynomial remainder sequence (PRS) and can be computed in various different ways [28]. We choose the subresultant PRS scheme for its computational superiority and also because each $F_i = S_{m_{i-1}-1}$, $1 \geq i \geq r$, where S_k is the k^{th} subresultant of F_1 and F_2. This together with making the z axis a valid projection direction ensures that in the polynomial remainder sequence there exists a polynomial remainder which is linear in z, i.e., $F_{r-1} = z\Phi_1(x,\ y) - \Phi_2(x,\ y) = 0$. This then yields z as a rational function of x and y and the inverse rational map.

Theorem 5.1: For an irreducible algebraic space curve C_d, the equations of the birational map and the projected plane curve P_d can be computed in $O(d^6 log^3 d)$ time.

Proof. The time for computing the valid projection direction via a random choice of values and the above polynomial remainder sequence is bound by the resultant computation for the projection. □

This together with Theorems 2.1 and 4.1 yields

Corollary 5.2: The genus of an algebraic space curve of degree d and the parametric equations of a rational space curve of degree d can be computed in $O(d^6 log^3 d + d^2 T(d^2))$ time.

2.5 Faithful Parameterizations

Given a polynomial parameterization

$$
\begin{aligned}
x &= P(t) = a_m t^m + a_{m-1} t^{m-1} + \ldots + a_0 \\
y &= Q(t) = b_n t^n + b_{n-1} t^{n-1} + \ldots + b_0
\end{aligned} \tag{12}
$$

of an affine algebraic curve $f(x, y)$ we now give an algorithm to check if the parameterization is faithful, i.e., for all but a finite number of points of the curve there corresponds a single parameter value and vice versa. Both m and n are bound by the degree d of the plane curve. Take the Taylor expansion with a single shift and let

$$
C(t) = Res_\tau \begin{pmatrix} \frac{P(t+\tau) - P(t)}{\tau} \\ \frac{Q(t+\tau) - Q(t)}{\tau} \end{pmatrix} \tag{13}
$$

$$
= Res_\tau \begin{pmatrix} P^{(1)}(t) & +\frac{1}{2}P^{(2)}(t)\tau & +\ldots & +\frac{1}{m!}P^{(m)}(t)\tau^{m-1} \\ Q^{(1)}(t) & +\frac{1}{2}Q^{(2)}(t)\tau & +\ldots & +\frac{1}{n!}Q^{(n)}(t)\tau^{n-1} \end{pmatrix} \tag{14}
$$

where P^k is the k^{th} derivative of P. Similarly for Q^k.

Then $C(t) \neq 0$ if and only if the parameterization is faithful. Further, if $C(t)$ is a nonzero polynomial, its roots give the singular points with multiplicities of the affine curve. Finally, if $C(t)$ is a non-zero constant then the affine plane curve is non-singular, or equivalently, since the curve is of genus 0, the curve has a single $d - 1$ fold singularity at infinity.

For a rational parameterization

$$
\begin{aligned}
x &= \frac{P(t)}{R(t)} \\
y &= \frac{Q(t)}{R(t)}
\end{aligned} \tag{15}
$$

of $f(x,y)$, again take the Taylor expansion with a single shift and let

$$C(t) = Res_\tau \begin{pmatrix} \frac{P(t+\tau)R(t)-R(t+\tau)P(t)}{\tau} \\ \frac{Q(t+\tau)R(t)-R(t+\tau)Q(t)}{\tau} \end{pmatrix} \tag{16}$$

$$= Res_\tau \begin{pmatrix} R(t)P^{(1)}(t) - P(t)R^{(1)}(t) & +\frac{1}{2}(R(t)P^{(2)}(t) - P(t)R^{(2)}(t))\tau + ... \\ R(t)Q^{(1)}(t) - Q(t)R^{(1)}(t) & +\frac{1}{2}(R(t)Q^{(2)}(t) - Q(t)R^{(2)}(t))\tau + ... \end{pmatrix} \tag{17}$$

Then again $C(t) \neq 0$ if and only if the parameterization is faithful.

Theorem 6.1: The faithfulness of parameterizations as well as the singularities of parameterically defined algebraic curves of degree d can be computed in $O(d^4 log^3 d)$ time.

Proof. The time for the Taylor expansion is at most $O(d^2)$ and is bound by the time taken to compute the resultant.

□

3 Triangulation and Display

See Section 3.4.

4 Polynomial and Power Basis

Let $\mathbf{p}_1, \ldots \mathbf{p}_d \in \mathbb{R}^k$. Then the convex hull of $\mathbf{p}_i$ is the set

$$[\mathbf{p}_1, \ldots, \mathbf{p}_d] = \{p : p \in \mathbb{R}^k, \quad p = \sum_{i=1}^{d} \lambda_i \mathbf{p}_i, \quad \sum_{i=1}^{d} \lambda_i = 1, \quad \lambda_i \geq 0\}$$

If $d = k+1$, then $[\mathbf{p}_1, \ldots, \mathbf{p}_d]$ is called a simplex. Let $P \subset \mathbb{R}^p$, $Q \subset \mathbb{R}^q$, then $P \times Q$ is defined by

$$P \times Q = \{p \in \mathbb{R}^{p+q} : p = (x,y)^T, \ x \in P, \ y \in Q\}.$$

A. Tensor-product form

$$p(x_1, \ldots, x_d) = \sum_{i_1=0}^{n_1} \sum_{i_2=0}^{n_2} \cdots \sum_{i_d=0}^{n_d} a_{i_1 i_2 \ldots i_d} \, x_1^{i_1} x_2^{i_2} \ldots x_d^{i_d} \tag{18}$$

where $(x_1, \ldots, x_d)^T \in \mathbb{R}^d$.

B. Total degree form

$$p(x_1, \ldots, x_d) = \sum_{i_1+\ldots+i_d \leq n} a_{i_1 i_2 \ldots i_d} x_1^{i_1} x_2^{i_2} \ldots x_d^{i_d} \tag{19}$$

where $(x_1, \ldots, x_d)^T \in \mathbb{R}^d$.

C. Mixed form

Let $d = d_1 + d_2$, then the mixed form is

$$p(x_1 \ldots x_d) = \sum_{i_1+\ldots+i_{d_1} \leq m} \sum_{j_1=0}^{n_1} \cdots \sum_{j_{d_2}=0}^{n_{d_2}} a_{i_1 \ldots i_{d_1} j_1 \ldots j_{d_2}} \, x_1^{i_1} \ldots x_{d_1}^{i_{d_1}} x_{d_1+1}^{j_1} \ldots x_d^{j_{d_2}} \tag{20}$$

where $(x_1 \ldots, x_d) \in \mathbb{R}^d$. If $d_1 = 0$, p is the tensor-produce form. If $d_2 = 0$, p is the total degree form.

5 Power Series and Puiseux Expansions

5.1 Weierstrass Factorization

Consider $f(x,y)$ with degree d and $ord_y\ f(0,y) = e < \infty$. The $ord_y\ f(0,y)$ is the y-exponent of the lowest degree term in $f(0,y)$ and is equal to ∞ if $f(0,y) = 0$. The occurrence of $f(0,y) = 0$ can be rectified by a simple linear transformation (rotation) of $f(x,y)$, which avoids making the x-axis, a tangent to the curve $f(0,y) = 0$ at the origin, and hence yields a nonzero $f(0,y)$ and a finite $ord_y\ f(0,y)$. A Weierstrass power series factorization is of the form $f(x,y) = g(x,y)\underbrace{(y^e + a_{e-1}(x)y^{e-1} + \cdots + a_0(x))}_{h(x,y)}$ where $g(x,y)$ is a unit power series, i.e., $g(0,0) \neq 0$ while $h(x,y)$ is a "distinguished" polynomial in y with coefficients $a_i(x)$, $i = 0 \ldots e-1$ being non-unit power series, i.e., $a_i(0) = 0$.

The Weierstrass preparation can efficiently be achieved via Hensel Lifting. Given

$$f(x,y) = f_0(y) + f_1(y)x + \cdots + f_k(y)x^k + \cdots$$

with

$$f(0,y) = f_0(y) = \underbrace{(a_0 + a_1 y + \cdots)}_{g_0(y)}\underbrace{y^e}_{h_0(y)}, \qquad a_0 \neq 0$$

in general for $k \geq 1$, we wish to compute $h_k(y)$ and $g_k(y)$ using Hensel, yielding factors similar to (22) such that

$$f_k(y) - \sum_{i>0,\ j>0,\ i+j=k} g_i(y)h_j(y) = g_0(y)h_k(y) + y^e g_k(y) \tag{21}$$

with degree $h_k(y) < e$.

To achieve this we compute $A(y) = \frac{f_k(y) - \sum_{i>0,\ j>0,\ i+j=k} g_i(y)h_j(y)}{g_0(y)}$ and then set $h_k(y) =$ terms of $A(y)$ with degree $< e$ and $g_k(y) =$ terms of $A(y)$ with degree $\geq e$.

5.2 Hensel Lifting

Consider $f(x,y)$ of degree d. Assume it is monic in y .

$$f(x,y) = f_0(y) + f_1(y)x + \cdots + f_k(y)x^k + \cdots$$

We wish to compute real power series factors $g(x,y)$ and $h(x,y)$ where $f(x,y) = g(x,y)h(x,y)$. The technique of Hensel lifting allows one to reconstruct the power series factors

$$\begin{aligned} g(x,y) &= g_0(y) + g_1(y)x + \cdots + g_i(y)x^i + \cdots \\ h(x,y) &= h_0(y) + h_1(y)x + \cdots + h_j(y)x^j + \cdots \end{aligned} \tag{22}$$

from initial factors $f(0,y) = f_0(y) = g_0(y)h_0(y)$.

Consider the factorization of $f(0,y) = f_0(y)$ as the base case of $k = 0$. Assume $f_0(y)$ is of degree d. Choose real coprime factors $g_0(y)$ of degree p and $h_0(y)$ of degree q satisfying: $p + q = d$. Real coprimeness is achieved by ensuring that g_0 and h_0 contain distinct real roots of f_0 and that complex conjugate pairs are not split up. However, it may arise that the only coprime factors of f_0 are complex, that is, the distinct roots are complex conjugates, in which case the curve $f(x,y) = 0$ does not intersect the y-axis and there is no real Newton power series factorization. Since $GCD(g_0(y), h_0(y)) = 1$ using the fast GCD algorithm we can also compute $\alpha(y)$ and $\beta(y)$ such that $\alpha(y)g_0(y) + \beta(y)h_0(y) = 1$.

In the iterative Case of $k \geq 1$, we compute $g_k(y)$ and $h_k(y)$ of the desired factorization (22), with degree of $g_k(y) < p$ and degree of $h_k(y) < q$, as follows. We note from (22) that

$$f_k(y) = \sum_{i + j = k} g_i(y)h_j(y)$$

and additionally

$$f_k(y) - \sum_{i>0,\ j>0,\ i+j=k} g_i(y)h_j(y) = g_0(y)h_k^*(y) + h_0(y)g_k^*(y) \tag{23}$$

Hence,

$$\begin{aligned} h_k^*(y) &= \alpha(y)[f_k(y) - \sum_{i>0,\ j>0,\ i+j=k} g_i(y)h_j(y)] \\ g_k^*(y) &= \beta(y)[f_k(y) - \sum_{i>0,\ j>0,\ i+j=k} g_i(y)h_j(y)] \end{aligned}$$

If degree $h_k^*(y) \geq q$ then compute $h_k(y) = h_k^*(y) \bmod h_0(y)$ and set $g_k(y) = \gamma(y)g_0(y) + g_k^*(y)$ where $h_k^*(y) = \gamma(y)h_0(y) + h_k(y)$.

$$f_k(y) - \sum_{i>0,\ j>0,\ i+j=k} g_i(y)h_j(y) = g_0(y)h_k(y) + h_0(y)g_k(y) \tag{24}$$

Clearly degree $h_k(y)$ is $< q$. Additionally in (24) the degree of $g_k(y)$ must also be $< p$. This is so because in (24) the degree of the LHS is $< d$ and since degree $g_0(y)h_k(y)$ is $< d$ and degree $h_0(y)$ is $= q$, it must be that degree $g_k(y)$ is $< p$.

6 Derivatives, Tangents, Curvatures

6.1 Curvature Computations

6.1.1 Curvature Formulas

The aim of this section is to provide readers with a quick reference for the curvature computation formulas. The detail derivation of these formulas are given in the section that follows.

Let M be a 2-dimensional Riemannian manifold in $\mathbb{R}^k$ with a Riemannian metric defined by the scalar inner product. Let (ξ_1, ξ_2) be a local coordinate system of the 2-manifold M at the point $x \in M$. Then $x \in \mathbb{R}^k$ can be expressed as

$$x = [x_1(\xi_1, \xi_2), \cdots, x_k(\xi_1, \xi_2)]^T. \tag{25}$$

Let $t_i = \frac{\partial x}{\partial \xi_i}$, $t_{ij} = \frac{\partial^2 x}{\partial \xi_i \partial \xi_j}$, $g_{ij} = t_i^T t_j$, and

$$G = \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix}, \quad Q = I - [t_1, t_2]G^{-1}[t_1, t_2]^T \in \mathbb{R}^{k \times k},$$

where

$$G^{-1} = \frac{1}{\det(G)} \begin{bmatrix} g_{22} & -g_{12} \\ -g_{21} & g_{11} \end{bmatrix}.$$

Then we have the following formulas:

Riemannian Curvature:

$$K(x) = \frac{t_{11}^T Q t_{22} - t_{12}^T Q t_{12}}{\det(G)}. \tag{26}$$

The Riemannian curvature is a counterpart of the Gaussian curvature of the classical surface. If $k = 3$, the Riemannian curvature coincides with the Gaussian curvature for surfaces.

Mean Curvature Vector:

$$H(x) = \frac{Q(g_{22}t_{11} + g_{11}t_{22} - 2g_{12}t_{12})}{2\det(G)}. \tag{27}$$

The mean curvature vector is a vector in the normal space. If $k = 3$, the mean curvature vector is in the normal direction, and its length is the classical mean curvature of the surface.

Principal Curvatures and Principal directions:

To obtain formulas for the principal curvatures and the principal directions, we first introduce an auxiliary result: *Let $A = (a_{ij})_{i,j=1}^2 \in \mathbb{R}^{2\times 2}$ be a symmetric matrix. Then the eigenvalues of A are*

$$\lambda_\pm = \frac{a_{11} + a_{22} \pm \sqrt{(a_{11} - a_{22})^2 + 4a_{12}^2}}{2} \tag{28}$$

and the corresponding eigenvectors are $[cos\theta_\pm, sin\theta_\pm]^T$, where $\theta_\pm$ are given (modulo π) by

$$\theta_+ = \frac{1}{2}\text{arctan}\frac{2a_{12}}{a_{11} - a_{22}}, \quad \theta_- = \theta_+ + \frac{\pi}{2}. \tag{29}$$

Now we give formulas for computing the principal curvatures and the principal directions. Let $h(x) = H(x)/\|H(x)\|$,

$$A = \Lambda^{-\frac{1}{2}} K F_h K^T \Lambda^{-\frac{1}{2}} \in \mathbb{R}^{2\times 2}, \quad [u_1, u_2] = [t_1, t_2] K^T \Lambda^{-\frac{1}{2}}, \tag{30}$$

where $F_h = -\left(t_{ij}^T\ h(x)\right)_{ij=1}^2$, $K \in \mathbb{R}^{2\times 2}$ and $\Lambda \in \mathbb{R}^{2\times 2}$ are defined by

$$G = K^T \Lambda K, \quad K^T K = I, \quad \Lambda = \text{diag}(\lambda_1, \lambda_2) \tag{31}$$

and they can be computed by (28)–(29). Let A be expressed, by virtue of (28) and (29), as

$$A = P\ \text{diag}(k_1, k_2)\ P^T, \quad \text{with} \quad P^T P = I. \tag{32}$$

Then k_1 and k_2 are the principal curvatures and v_1 and v_2, defined by

$$[v_1, v_2] := [u_1, u_2]P = [t_1, t_2] K^T \Lambda^{-\frac{1}{2}} P, \tag{33}$$

are the corresponding principal directions with respect to the direction vector h.

Again, the principal curvatures and the principal directions are the counterparts of the same concepts for surfaces. If $k = 3$, they are the same.

6.1.2 Derivation

In this section we derive the curvature formulas from the field of Riemannian geometry. However, we have tried to make the paper self-contained, so that readers can understand the derivation without having to consult the Riemannian geometry literature. Readers may simply skim over this section if they merely intend to use the curvature formulas.

Notations and Terminologies **Differential Manifold**. A *differentiable manifold* of dimension n is a set M and a family of injective mappings $x_\alpha : U_\alpha \subset \mathbb{R}^n \to M$ of open sets U_α into M such that

(1). $\bigcup_\alpha x_\alpha(U_\alpha) = M$.

(2). For any pair α, β with $x_\alpha(U_\alpha) \cap x_\beta(U_\beta) = W \neq \emptyset$, the sets $x_\alpha^{-1}(W)$ and $x_\beta^{-1}(W)$ are open in $\mathbb{R}^n$ and the mappings $x_\beta^{-1} \circ x_\alpha$ are differentiable.

The mapping x_α with $x \in x_\alpha(U_\alpha)$ is called a parameterization of M at x. In our case, we use the 2-dimensional manifold ($n = 2$). Denoting the coordinate U_α as (ξ_1, ξ_2), then the tangent space T_xM at $x \in M$ is spanned by $\{\frac{\partial}{\partial \xi_1}, \frac{\partial}{\partial \xi_2}\}$. The set $TM = \{(x, v);\ x \in M,\ v \in T_xM\}$ is called a tangent bundle.

Vector Field ([23], page 25). A vector field X on a differentiable manifold M is a correspondence that associates to each point $x \in M$ a vector $X(x) \in T_xM$. The field is differentiable if the mapping $X : M \to TM$ is differentiable.

Considering a parameterization $x : U \subset \mathbb{R}^2 \to M$, there exist $a_i(x)$, such that

$$X(x) = \sum_i a_i(x) \frac{\partial}{\partial \xi_i}.$$

Let $\mathcal{D}$ be the set of differentiable functions on M, and X be a vector field on M. Then X can be regarded as a mapping $X : \mathcal{D} \to \mathcal{D}$ such that

$$(Xf)(x) = \sum_i a_i(x) \frac{\partial f}{\partial \xi_i}(x). \tag{34}$$

It is easy to check that Xf does not depend on the choice of parameterization x. Let X and Y be differentiable vector fields on a differentiable manifold M. Then there exists a unique vector field Z such that, for all $f \in \mathcal{D}$, $Zf = (XY - YX)f$. The vector field $Z := XY - YX$ is called the *bracket* of X and Y (see [23], pages 26-27), denoted by $[X, Y]$. Let

$$X(x) = \sum_i a_i(x) \frac{\partial}{\partial \xi_i}, \quad Y(x) = \sum_i b_i(x) \frac{\partial}{\partial \xi_i}.$$

Then from (34) we can derive that

$$[X, Y] = \sum_j \left(\sum_i a_i \frac{\partial b_j}{\partial \xi_i} - b_i \frac{\partial a_j}{\partial \xi_i} \right) \frac{\partial}{\partial \xi_j}.$$

Riemannian Manifold. A differentiable manifold with a given Riemannian metric is called a *Riemannian Manifold*. A Riemannian metric $\langle\ ,\ \rangle_x$ of M is a symmetric, bilinear and positive-definite form on the tangent space T_xM. Since M is a sub-manifold of Euclidean space $\mathbb{R}^k$, we use the *induced metric*:

$$\langle u, v \rangle_x = u^T v, \quad u, v \in T_xM.$$

Connection. Let us indicate by $\mathcal{X}(M)$ the set all vector fields of class C^∞ on M and by $\mathcal{D}(M)$ the ring of real-valued functions of class C^∞ defined on M. An affine connection ∇ on a differentiable manifold M is a mapping $\nabla :\ \mathcal{X}(M) \times \mathcal{X}(M) \to \mathcal{X}(M)$ which is denoted by $(X, Y) \to \nabla_X Y$ and which satisfies the following properties:

1) $\nabla_{fX+gY} Z = f\nabla_X Z + g\nabla_Y Z$.

2) $\nabla_X(Y + Z) = \nabla_X Y + \nabla_X Z$.

3) $\nabla_X(fY) = f\nabla_X Y + X(f)Y$,

in which $X, Y, Z \in \mathcal{X}(M)$ and $f, g \in \mathcal{D}(M)$.

Choose a system of coordinates (ξ_1, ξ_2),

$$X = \sum_i a_i t_i, \; Y = \sum_j b_j t_j,$$

where $t_i = \frac{\partial}{\partial \xi_i}$, then from properties 1)–3) we have

$$\nabla_X Y = \sum_k \left(\sum_{ij} a_i b_j \Gamma_{ij}^k + X(b_k) \right) t_k, \tag{35}$$

where Γ_{ij}^k is defined by

$$\nabla_{t_i} t_j = \sum_k \Gamma_{ij}^k t_k. \tag{36}$$

An affine connection ∇ on M is said to be symmetric if

$$\nabla_X Y - \nabla_Y X = [X, Y] \quad \text{for all} \quad X, Y \in \mathcal{X}(M).$$

A connection ∇ on a Riemannian manifold M is compatible with the metric $\langle \; , \; \rangle$ if and only if ([23], page 54)

$$X\langle Y, Z\rangle = \langle \nabla_X Y, Z\rangle + \langle Y, \nabla_X Z\rangle, \quad X, Y, Z \in \mathcal{X}(M).$$

Riemannian Curvature Here we start with the Levi-Civita theorem ([23] page 55): *Given a Riemannian manifold M, there exists a unique affine connection ∇ on M satisfying the conditions:*

a) ∇ is symmetric.

b) ∇ is compatible with the Riemannian metric.

The connection defined by the Levi-Civita theorem is called the Riemannian connection. For the Riemannian connection, the number Γ_{ij}^k defined by (36), which is called the *Christoffel Symbols*, is calculated by

$$\Gamma_{ij}^m = \frac{1}{2} \sum_k \left\{ \frac{\partial g_{jk}}{\partial \xi_i} + \frac{\partial g_{ki}}{\partial \xi_j} - \frac{\partial g_{ij}}{\partial \xi_k} \right\} g^{km}, \tag{37}$$

where $(g^{kl}) = (g_{ij})^{-1}$, $(g_{ij}) = G$. Note that $\Gamma_{ij}^m = \Gamma_{ji}^m$, since $g_{ij} = g_{ji}$. It is easy to recognize that if M is an Euclidean space $\Gamma_{ij}^m = 0$.

Curvature. The *curvature* ([23], page 89) of a Riemannian manifold M is a correspondence which associates to every pair of vector fields X, Y a mapping $R(X, Y)$ which maps a vector field of M to another vector field of M given by

$$R(X, Y)Z = \nabla_Y \nabla_X Z - \nabla_X \nabla_Y Z + \nabla_{[X,Y]} Z, \tag{38}$$

where ∇ is the Riemannian connection of M.

Let (U_α, x_α) be a coordinate system at point $x \in M$. Let $\frac{\partial}{\partial \xi_i} = t_i$ and put

$$R(t_i, t_j)t_k = \sum_{l=1}^{2} R_{ijk}^l t_l. \tag{39}$$

Then $R(X, Y)Z$ can be expressed as

$$R(X, Y)Z = \sum_{i,j,k,l} R_{ijk}^l a_i b_j c_k t_l,$$

where $X = \sum_i a_i t_i,\ Y = \sum_j b_j t_j,\ Z = \sum_k c_k t_k$. From (35) we can derive that R^l_{ijk} are given as

$$R^s_{ijk} = \sum_l \Gamma^l_{ik}\Gamma^s_{jl} - \sum_l \Gamma^l_{jk}\Gamma^s_{il} + \frac{\partial \Gamma^s_{ik}}{\partial \xi_j} - \frac{\partial \Gamma^s_{jk}}{\partial \xi_i}. \tag{40}$$

Riemannian curvature. The counterpart of the Gaussian curvature for a surface is the *Riemannian curvature* for a Riemannian manifold ([23, 41]). For $x \in M$, let $X, Y \in T_xM$ be two linearly independent vectors. Then the *Riemannian curvature* of the tangent space T_xM is defined by

$$K(x) = \frac{\langle R(X,Y)X, Y\rangle_x}{\|X\|^2\|Y\|^2 - \langle X, Y\rangle_x^2}.$$

The Riemannian curvature, also called *sectional curvature*, is originally defined for a two-dimensional subspace of the tangent space T_xM. However, since T_xM is assumed to be two-dimensional, the Riemannian curvature is then uniquely defined. For a regular surface in $\mathbb{R}^3$, the Riemannian curvature is the Gaussian curvature. It is not difficult to realize ([23], page 94) $K(x)$ does not depend on the choice of the vectors $X, Y \in T_xM$. Hence, we can use $X = t_1,\ Y = t_2$ and from (39)–(40) we have

$$\begin{aligned} \langle R(X,Y)X, Y\rangle_x &= \sum_{s=1}^{2} R^l_{121} g_{s2} \\ &= \sum_{s=1}^{2}\left[\sum_{l=1}^{2}\left(\Gamma^l_{11}\Gamma^s_{2l} - \Gamma^l_{21}\Gamma^s_{1l}\right)\right] g_{s2} + \sum_{s=1}^{2}\left(\frac{\partial \Gamma^s_{11}}{\partial \xi_2} - \frac{\partial \Gamma^s_{21}}{\partial \xi_1}\right) g_{s2}. \end{aligned} \tag{41}$$

It follows from (37) that

$$[\Gamma^1_{ij},\ \Gamma^2_{ij}] = t^T_{ij}[t_1,\ t_2]G^{-1}, \quad i, j = 1, 2. \tag{42}$$

Substituting these into (41) we have

$$\begin{aligned} \sum_{s=1}^{2}\left(\frac{\partial \Gamma^s_{11}}{\partial \xi_2} - \frac{\partial \Gamma^s_{21}}{\partial \xi_1}\right) g_{s2} &= (t^T_{11}t_{22} - t^T_{12}t_{12}) + t^T_{11}[t_1, t_2]G^{-1}\left\{[t_{12}, t_{22}]^T t_2 + [t_1, t_2]^T t_{22}\right\} \\ &- \ t^T_{12}[t_1, t_2]G^{-1}[t^T_{11}t_2 + t^T_1 t_{12}, t^T_{12}t_2 + -t^T_2 t_{12}]^T. \end{aligned} \tag{43}$$

Using (42), the first summation of (41) can be written as

$$\begin{aligned} &(\Gamma^2_{11}\Gamma^1_{22} - \Gamma^1_{12}\Gamma^2_{12})g_{12} \ + \ [\Gamma^2_{12}(\Gamma^1_{11} - \Gamma^2_{12}) + \Gamma^2_{11}(\Gamma^2_{22} - \Gamma^1_{21})]g_{22} \\ &\qquad = [\Gamma^1_{11}, \Gamma^2_{11}][t^T_{12}t_2, t^T_{22}t_2]^T - [\Gamma^1_{21}, \Gamma^2_{21}][t^T_{11}t_2, t^T_{12}t_2]^T. \end{aligned} \tag{44}$$

Combining (43) with (44) we arrive at formula (26). Note that (41) involves the third order partial derivatives of M, but (26) does not.

Mean Curvature For a 2-dimensional Riemannian sub-manifold M of $\mathbb{R}^k$, the mean curvature vector is defined by ([42] page 119)

$$H(x) = \frac{1}{2}[h(e_1, e_1) + h(e_2, e_2)],$$

where (e_1, e_2) is an orthonormal frame for the tangent space to M at x. $h(X, Y)$ is defined by

$$h(X, Y) = \tilde{\nabla}_X Y - \nabla_X Y,$$

where ∇ and $\tilde{\nabla}$ are the Riemannian connection in M and $\mathbb{R}^k$, respectively. Since $\nabla_X Y \in TM$, $h(X,Y) \in TM^{\perp}$, we may consider only the computation of $\tilde{\nabla}_X Y$ and then project it into the normal space to obtain $h(X,Y)$. It follows from (35) that

$$\tilde{\nabla}_{e_l} e_l = \left[\frac{\partial e_l}{\partial x_1}, \ \cdots, \ \frac{\partial e_l}{\partial x_k}\right] e_l, \tag{45}$$

where the fact $\Gamma_{ij}^l = 0$ for the Euclidean space $\mathbb{R}^k$ has been used. The orthonormal frame (e_1, e_2) can be obtained by the Gram-Schmitt process from (t_1, t_2):

$$e_1 = t_1/\sqrt{g_{11}}, \quad e_2 = (g_{11}t_2 - g_{12}t_1)/\sqrt{g_{11}\det(G)}, \tag{46}$$

Since $x_j = x_j(\xi_1, \xi_2), \ j = 1, \cdots, k$, we have

$$\varepsilon_j = t_1 \frac{\partial \xi_1}{\partial x_j} + t_2 \frac{\partial \xi_2}{\partial x_j}, \quad j = 1, \cdots, k, \tag{47}$$

where $\varepsilon_j = [0, \cdots, 0, 1, 0, \cdots, 0]^T$ is the j-th unit vector in $\mathbb{R}^k$. Performing the inner product of both sides of (47) with t_1 and t_2 and then solving the linear system derived for the unknowns $\frac{\partial \xi_1}{\partial x_j}, \frac{\partial \xi_2}{\partial x_j}$, we get

$$\left[\frac{\partial \xi_1}{\partial x_j}, \frac{\partial \xi_2}{\partial x_j}\right]^T = G^{-1}\left[t_1^T \varepsilon_j, t_2^T \varepsilon_j\right]^T.$$

Then by

$$\frac{\partial e_l}{\partial x_j} = \frac{\partial e_l}{\partial \xi_1}\frac{\partial \xi_1}{\partial x_j} + \frac{\partial e_l}{\partial \xi_2}\frac{\partial \xi_2}{\partial x_j}, \quad l = 1, 2; \ j = 1, \cdots, k,$$

we have

$$\tilde{\nabla}_{e_l} e_l = \left[\frac{\partial e_l}{\partial \xi_1}, \frac{\partial e_l}{\partial \xi_2}\right] G^{-1}[t_1, t_2]^T e_l. \tag{48}$$

Taking $l = 1, 2$ and using (46) we get

$$\tilde{\nabla}_{e_1} e_1 = \frac{\partial e_1}{\partial \xi_1}/\sqrt{g_{11}}, \quad \tilde{\nabla}_{e_2} e_2 = \left(-g_{12}\frac{\partial e_2}{\partial \xi_1} + g_{11}\frac{\partial e_2}{\partial \xi_2}\right)/\sqrt{g_{11}\det G}.$$

Since what we required is the part of $\tilde{\nabla}_{e_l} e_l$ orthogonal to the tangent space, we get

$$\left[\tilde{\nabla}_{e_1} e_1\right]^{\perp} = \frac{[t_{11}]^{\perp}}{g_{11}}, \quad \left[\tilde{\nabla}_{e_2} e_2\right]^{\perp} = \frac{[g_{12}^2 t_{11} + g_{11}^2 t_{22} - 2g_{11}g_{12}t_{12}]^{\perp}}{g_{11}\det G},$$

where $[\cdot]^{\perp}$ denotes the normal component of a vector. From this we have

$$H(x) = \frac{[g_{22}t_{11} + g_{11}t_{22} - 2g_{12}t_{12}]^{\perp}}{2(g_{11}g_{22} - g_{12}^2)}, \tag{49}$$

and thereafter (27) is derived, since Q is a projector that maps a vector to the normal space.

Principal Curvatures and Principal Directions Since $M \subset \mathbb{R}^k$, the normal space, denoted by $T_x M^{\perp}$, can be defined at each point $x \in M$:

$$T_x M^{\perp} = \{n \in \mathbb{R}^k : \ \langle t, n \rangle_x = 0, \ \ \forall t \in T_x M\}.$$

Let n be a normal vector field on M and X be a vector field tangent to M. Then we have

$$\tilde{\nabla}_X n = -\mathcal{A}_n X + \nabla_X^{\perp} n,$$

where $-\mathcal{A}_n X$ and $\nabla^{\perp}_X n$ are respectively the tangent and the normal components. Then $\mathcal{A}_n$ is a self-adjoint map from TM to TM, called *second fundamental tensor* with respect to n ([42] pages 119-121). The *principal curvatures* $k_1(x)$, $k_2(x)$ and the *principal directions* $v_1(x)$, $v_2(x)$ with respect to n are defined as the eigenvalues and the orthonormal eigenvectors of $\mathcal{A}_n$. However, the principal curvatures and the principal directions are not uniquely defined since the normal vector field is not uniquely defined for $k > 3$.

We have chosen a special normal vector field $h = H(x)/\|H(x)\|$, which is the normalized mean curvature vector field of the manifold M and is uniquely defined.

To calculate the spectrum of $\mathcal{A}_h$, we need to obtain its matrix representation. Let e_1, e_2 be the orthonormal basis of $T_x M$ defined by (46):

$$[e_1, e_2] = [t_1, t_2]W, \quad \text{with} \quad W = \begin{bmatrix} g_{11}^{-\frac{1}{2}} & -g_{12}[g_{11}\det(G)]^{-\frac{1}{2}} \\ 0 & g_{11}[g_{11}\det(G)]^{-\frac{1}{2}} \end{bmatrix}.$$

Let

$$\mathcal{A}_h e_i = a_{1i}e_1 + a_{2i}e_2, \quad i = 1, 2. \tag{50}$$

Then

$$\mathcal{A}_h[e_1, e_2] = [e_1, e_2]A_h,$$

where A_h is a 2×2 matrix which needs to be calculated in the following. It will be clear soon that A_h is symmetric. Before giving an explicit form of A_h, we first show that the eigenvalues of A_h are the eigenvalues of $\mathcal{A}_h$. Let

$$A_h = S \ \mathrm{diag}(\lambda_1, \lambda_2) \ S^T, \quad S^T S = I, \quad \text{and} \quad [v_1, v_2] = [e_1, e_2]S. \tag{51}$$

Then

$$\begin{aligned} \mathcal{A}_h[v_1, v_2] &= \mathcal{A}_h[e_1, e_2]S \\ &= [e_1, e_2]A_h S \\ &= [e_1, e_2]S\mathrm{diag}(\lambda_1, \lambda_2) \\ &= [v_1, v_2]\mathrm{diag}(\lambda_1, \lambda_2). \end{aligned}$$

Hence, v_i is the eigenvector of $\mathcal{A}_h$ with respect to the eigenvalue λ_i. Furthermore,

$$[v_1, v_2]^T[v_1, v_2] = S^T[e_1, e_2]^T[e_1, e_2]S = I.$$

That is, v_1, v_2 are orthonormal.

Now we calculate the matrix A_h. To this end, we need to calculate $\tilde{\nabla}_{e_i} h$. Paralleling to the derivation of $\tilde{\nabla}_{e_i} e_i$, we have an expression similar to (48):

$$\tilde{\nabla}_{e_l} h = \left[\frac{\partial h}{\partial \xi_1}, \frac{\partial h}{\partial \xi_2}\right] G^{-1}[t_1, t_2]^T e_l, \quad l = 1, 2. \tag{52}$$

If we project $\tilde{\nabla}_{e_l} h$ into the tangent space and express $\mathcal{A}_h e_l$ as (50), A_h can be expressed as

$$\begin{aligned} A_h &= -[e_1, e_2]\left[\frac{\partial h}{\partial \xi_1}, \frac{\partial h}{\partial \xi_2}\right] G^{-1}[t_1, t_2]^T[e_1, e_2] \\ &= -W^T[t_1, t_2]^T \left[\frac{\partial h}{\partial \xi_1}, \frac{\partial h}{\partial \xi_2}\right] W. \end{aligned} \tag{53}$$

Now we need to calculate $[t_1, t_2]^T \left[\frac{\partial h}{\partial \xi_1}, \frac{\partial h}{\partial \xi_2}\right]$. Substituting the expression h into (53) and with some additional calculations, we have

$$A_h = W^T F_h W, \tag{54}$$

where F_h is a 2×2 symmetric matrix defined by

$$F_h = \frac{g_{22}(B_{11} - A_{11}) + g_{11}(B_{22} - A_{22}) - 2g_{12}(B_{12} - A_{12})}{2\det(G)\|H(x)\|}, \quad (55)$$

$$A_{kl} = (t_{ij}^T t_{kl})_{ij=1}^2, \quad B_{kl} = (c_{ij}^T G^{-1} c_{kl})_{ij=1}^2, \quad c_{ij} = [t_1, t_2]^T t_{ij}. \quad (56)$$

Substituting (56) into (55) and using the mean curvature formula (27), we have a simple expression for F_h:

$$F_h = -\left(t_{ij}^T\, h(x)\right)_{ij=1}^2.$$

Having an explicit expression for A_h, we are able to compute the principal curvatures and the principal directions by (28), (29) and (51). However, since A_h involves W, it is not intrinsic. To obtain more elegant formulas, we rewrite $A_h = W^T F_h W$ as follows

$$A_h = (\Lambda^{\frac{1}{2}} K W)^T A (\Lambda^{\frac{1}{2}} K W) \quad \text{with} \quad A = \Lambda^{-\frac{1}{2}} K F_h K^T \Lambda^{-\frac{1}{2}}.$$

It follows from (30) and (31) that $[u_1, u_2]^T [u_1, u_2] = I$. We then have

$$\begin{aligned} I &= [e_1, e_2]^T [e_1, e_2] \\ &= W^T [t_1, t_2]^T [t_1, t_2] W \\ &= W^T K^T \Lambda^{\frac{1}{2}} [u_1, u_2]^T [u_1, u_2] \Lambda^{\frac{1}{2}} K W \\ &= (\Lambda^{\frac{1}{2}} K W)^T (\Lambda^{\frac{1}{2}} K W), \end{aligned} \quad (57)$$

that is, $\Lambda^{\frac{1}{2}} K W$ is an orthgonal matrix. Hence, the eigenvalues of A_h and A are the same, and therefore we can use A to compute the principal curvatures instead of A_h. Using relation (32), we have

$$A_h = (\Lambda^{\frac{1}{2}} K W)^T P \text{diag}(k_1, k_2) P^T (\Lambda^{\frac{1}{2}} K W).$$

Hence S could be written as

$$S = (\Lambda^{\frac{1}{2}} K W)^T P.$$

Therefore, the eigenvectors are given by

$$\begin{aligned} [v_1, v_2] &= [e_1, e_2] S \\ &= [t_1, t_2] W S \\ &= [u_1, u_2] (\Lambda^{\frac{1}{2}} K W)(\Lambda^{\frac{1}{2}} K W)^T P \\ &= [u_1, u_2] P, \end{aligned}$$

and hence (33) is derived.

Remark 1. Since h uses the second order partial derivatives of $x \in M$, $\frac{\partial h}{\partial \xi_i}$ uses the third order partials. A nice property is that all the third order partials are canceled in A_h. The final result only uses the first and the second order partials.

Remark 2. Both the matrix A for computing the principal curvatures and the formula (33) for computing the principal directions do not involve W. Furthermore, it can be proved that all the curvatures do not depend on the choice of the local coordinate system (ξ_1, ξ_2). Therefore, they are intrinsic to the manifold M. The proof of this claim is not the theme of this paper.

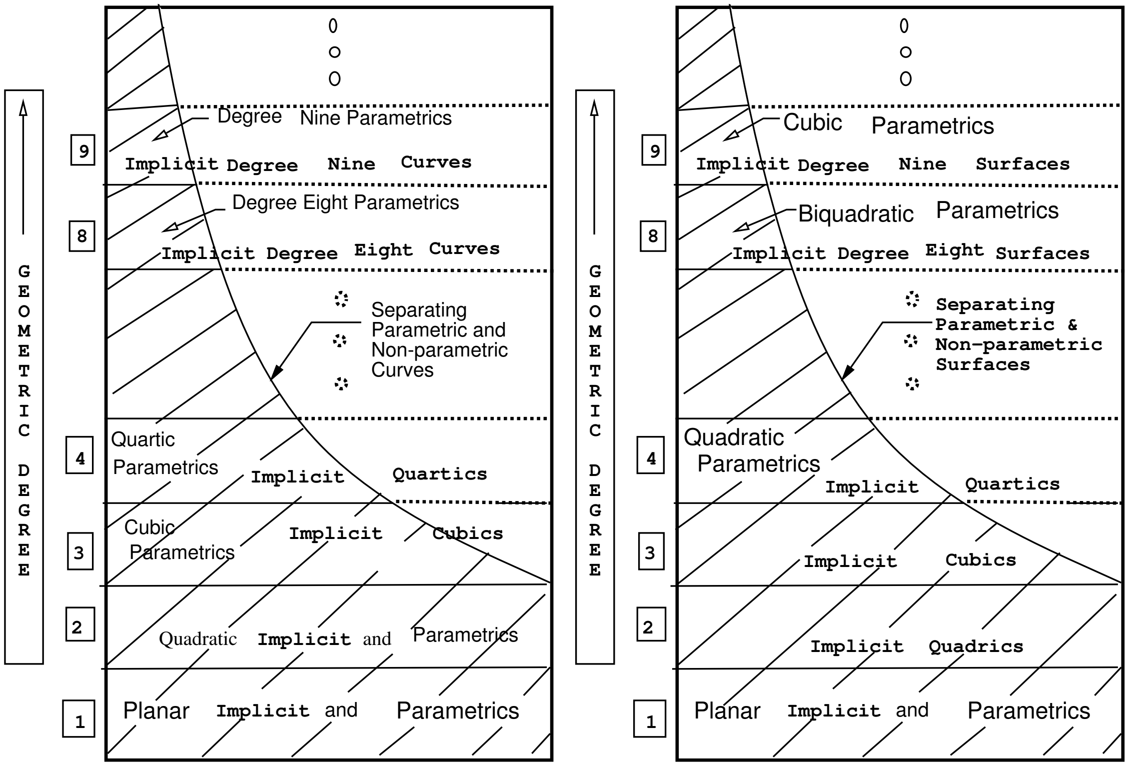


Figure 2: A classification of low degree algebraic curves (left) and surfaces (right)

7 Converting Between Implicit and Parametric Forms

A real implicit algebraic plane curve $f(x, y) = 0$ is a hypersurface of dimension 1 in $\mathbb{R}^2$, while a parametric plane curve $[f_3(s)x - f_1(s) = 0, f_3(s)y - f_2(s) = 0]$ is an algebraic variety of dimension 1 in $\mathbb{R}^3$, defined by the two independent algebraic equations in the three variables x, y, s. Similarly, a real implicit algebraic surface $f(x, y, z) = 0$ is a hypersurface of dimension two in $\mathbb{R}^3$, while a parametric surface $[f_4(s,t)x - f_1(s,t) = 0, f_4(s,t)y - f_2(s,t) = 0, f_4(s,t)z - f_3(s,t) = 0]$ is an algebraic variety of dimension 2 in $\mathbb{R}^5$, defined by three independent algebraic equations in the five variables x, y, z, s, t.

A plane parametric curve is a very special algebraic variety of dimension 1 in x, y, s space, since the curve lies in the 2-dimensional subspace defined by x, y and furthermore points on the curve can be put in $(1, 1)$ rational correspondence with points on the 1-dimensional sub-space defined by s. Parametric curves are thus a special subset of algebraic curves, and are often also called rational algebraic curves. Figure 2 depicts the relationship between the set of parametric curves and non-parametric curves at various degrees.

Example parametric (rational algebraic) curves are degree two algebraic curves (conics) and degree three algebraic curves (cubics) with a singular point. The non-singular cubics are not rational and are also known as elliptic cubics. In general, a necessary and sufficient condition for the rationality of an algebraic curve of arbitrary degree is given by the Cayley-Riemann criterion: a curve is rational if and only if $g = 0$, where g, the genus of the curve is a measure of the deficiency of the curve's singularities from its maximum allowable limit [40]. Algorithms for computing the genus of an algebraic curve and for symbolically deriving the parametric equations of genus 0 curves, are given for example in [5].

For implicit algebraic plane curves and surfaces defined by polynomials of degree d, the maximum number of intersections between the curve and a line in the plane or the surface and a line in space, is equal to the maximum number of roots of a polynomial of degree d. Hence, here the geometric degree is the same as the algebraic degree which is equal to d. For parametric curves defined by polynomials of degree d, the maximum number of intersections between the curve and a line in the plane is also equal to the maximum number of roots of a polynomial of degree d. Hence here again the geometric degree is the same as the algebraic degree.

For parametric surfaces defined by polynomials of degree d the geometric degree can be as large as d^2, the square of the algebraic degree d. This can be seen as follows. Consider the intersection of a generic line in space $[a_1x+b_1y+c_1z-d_1=0, a_2x+b_2y+c_2z-d_2=0]$ with the parametric surface. The intersection yields two implicit algebraic curves of degree d which intersect in $O(d^2)$ points (via Bezout's theorem), corresponding to the intersection points of the line and the parametric surface.

A parametric curve of algebraic degree d is an algebraic curve of genus 0 and so have $\frac{(d-1)(d-2)}{2}$ $= O(d^2)$ singular (double) points. This number is the maximum number of singular points an algebraic curve of degree d may have. From Bezout's theorem, we realize that the intersection of two implicit surfaces of algebraic degree d can be a curve of geometric degree $O(d^2)$. Furthermore the same theorem implies that the intersection of two parametric surfaces of algebraic degree d (and geometric degree $O(d^2)$) can be a curve of geometric degree $O(d^4)$. Hence, while the potential singularities of the space curve defined by the intersection of two implicit surfaces defined by polynomials of degree d can be as many as $O(d^4)$, the potential singularities of the space curve defined by the intersection of two parametric surfaces defined by polynomials of degree d can be as many as $O(d^8)$.

7.1 Parameterization of Curves

An irreducible algebraic curve C_d of degree d in the plane is one which is met by most lines in d points. Lines through a point P meet C_d (outside P) in general at $d - mult_pC_d$ points, where $mult_pC_d = e =$ multiplicity of C_d at P. If $e = 1$ then P is called a simple point. If $e = 2$ then P is called a double point. Similarly we talk about an e-ple point or an e-fold point. If $e = 0$: P is not on C_d. If $e > 1$ we say P is a *singular point* of the curve C_d with multiplicity e. This also leads to the following theorem for curves

Theorem 7.1. *[Bezout] Curves of degree d and curves of degree e, with no common components, meet at dcdote points, counting multiplicities and points at infinity. ($C_d \cdot C_e = d \cdot e$ points.)*

Consider curve C_d of degree d to be also of order e.

$$C_d : f(x,y) = \sum_{e \le i+j \le d} a_{ij}x^iy^j = f_d(x,y) + f_{d-1}(x,y) + \cdots + f_e(x,y)$$

where $f_i(x,y)$ are homogeneous polynomials of degree i together with $f_d(x,y) \neq 0$ and $f_e(x,y) \neq 0$ so that $d =$ degree and e order. Thus $f_d(x,\ y)$ is the degree form and $f_e(x,\ y)$ is the initial or order form. Again, the multiplicity of a point P on C_d is geometrically, the number of points that a line through that point P meets C_d at P. By translation, we can assume the point P to be the origin. Then the equation of a line through the origin is $y = mx$. Its intersection with the curve is given by

$$\begin{aligned} f(x,\ mx) &= f_d(x,\ mx) + f_{d-1}(x,\ mx) + \cdots + f_e(x,\ mx) \\ &= x^d f_d(1,\ m) + x^{d-1} f_{d-1}(1,\ m) + \cdots x^e f_e(1,\ mx) \\ &= x^e[f_d(1,\ m)x^{d-e} + \cdots + f_e(1,\ m) \end{aligned}$$

Lines through the origin meet the curve, outside the origin, in $d-e$ points. Hence the multiplicity of the origin $= e =$ (*order* of the curve). Thus if the curve C_d has a $d-1$ fold point (origin), then lines through that point meet F at one other point, and thereby parameterizes the curve (rational).

Here we can also note that for most values of m , $f_e(1,\ m) \neq 0$. The values of m for which it is zero correspond to the tangents $f_e(x,\ y) = \prod_{i=1}^{e}(y - m_ix)$ to the curve at the origin. (Tangents

degree of curve	1	2	3	4	5	6	$\cdots$	d
the maximum number of double points	0	0	1	3	6	10	$\cdots$	$\frac{1}{2}(d - 1)(d - 2)$
the number of independent parameters	2	5	9	14	20	27	$\cdots$	$\frac{1}{2}d(d + 3)$

Table 1: Relation of degree of curve to number of double points and independent parameters.

at P are thus those special lines which meet C_d at P at more than e points, where $e =$ multiplicity of C_d at P.)

Now note for example that the equation of a conic has five independent coefficients and if we take five 'independent' points in the plane and consider a conic passing through these points then this will give five linear homogeneous equations in the five coefficient variables. If the rank of the matrix is 5 then there is a unique conic through these points. In general, the number of independent coefficients of a plane algebraic curve C_d of degree d is $\frac{1}{2}d(d+3)$.

One can easily prove by Bezout's theorem that a curve of degree 4, for example, cannot have 4 double points. In general one may see that the number of double points, say DP, of C_d is $\leq \frac{1}{2}(d - 1)(d - 2)$. Assume $DP > \frac{1}{2}(d - 1)(d - 2)$. Then since $\frac{1}{2}(d - 2)(d + 1)$ fixed points determine a C_{d-2} curve and if we choose $\frac{1}{2}(d - 1)(d - 2) + 1$ double points of C_d then to determine C_{d-2} one needs a remaining

$$\frac{1}{2}(d-2)(d+1) - (\frac{1}{2}(d-1)(d-2)+1) = (d-2) - 1 = d-3 \quad \text{points.}$$

So take $(d - 3)$ other fixed simple points of C_d. Then we can pass a C_{d-2} curve through the above $\frac{1}{2}(d - 1)(d - 2) + 1$ double points of C_d and $(d - 3)$ other simple points of C_d. Then counting the number of points of intersection of C_d and C_{d-2} together with multiplicities yields

$$(d - 1)(d - 2) + 2 + d - 3 = d^2 - 2d + 1 = (d-2)d + 1 = C_d \cdot C_{d-2} + 1$$

which contradicts Bezout. Thus assuming Bezout we see that

$$DP \leq \frac{1}{2}(d - 1)(d - 2)$$

In general, we have Table 7.1

One definition of the genus g of a curve C_d is a measure of how much the curve is deficient from its maximum allowable limit of singularities,

$$g = \frac{1}{2}(d - 1)(d - 2)$$

where DP is a 'proper' counting of the number of double points of C_d (summing over all singularities). From the earlier discussion and Bezout, we can see that in counting the number of double points DP of C_d an e-ple point of C is to be counted as $\frac{1}{2}e(e - 1)$ double points.

However this counting is not very precise as such is the case only for the so called *distinct* multiple points of C. For a multiple point, that is not *distinct*, one has also to consider *infinitely near* singularities. In general a double point is roughly either a *node* or a *cusp*. If a cusp is given by $y^2 - x^3$ we call it a *distinct* cusp and is counted as a single double point. Cusps other than distinct look like $y^2 - x^{2m+1}$ (an $m-fold$ cusp). Though the multiplicity of the origin is two (= *order* of the curve) the origin accounts for m double points when counted properly. The proper counting was achieved by Noether using homogeneous "Cremona quadratic transformations", see also [40]. Following [2] we can achieve the same thing by using "affine quadratic transformations".

Consider for example, the cusp $y^2 - x^3 = 0$ which has a double point at the origin. The quadratic transformation $\overline{q}$ given by

$$x = \bar{x} \quad \text{and} \quad y = \bar{x}\bar{y} \tag{58}$$

yields

$$0 = y^2 - x^3 = \bar{x}^2\bar{y}^2 - \bar{x}^3 = \bar{x}^2(\bar{y}^2 - \bar{x}),$$

and canceling out the extraneous factor $\bar{x}^2$ we get the nonsingular parabola $\bar{y}^2 - \bar{x} = 0$. So the origin in this case was a *distinct* singular point and counted as a single double point.

To desingularize the $m - fold$ cusp one has to make a succession of m transformations of the type (58). Only the m^th successive application of (1) changes the multiplicity of the origin from two to one. Hence in this case, counting properly, we say that the cusp has one *distinct* double point and $(m - 1)$ *infinitely near* double points, giving a total DP count of m.

In a general procedure for counting double points, given an e-fold point P of a plane curve C, we choose our coordinates to bring P to the origin and then apply (58). If now $C : f(x, y) = 0$, then the substitution (58) transforms C into the curve $\overline{C} : \overline{f} : (\overline{x}, \overline{y}) = 0$ given by

$$f(\overline{x}, \overline{xy}) = \overline{x}^e\overline{f}(\overline{x}, \overline{y}).$$

$Cbar$ will meet the line $E : \bar{x} = 0$ in the points $P^1, \ldots, P^m$, the roots of $\bar{f}(0, y) = 0$ which corresponds to the tangents to C at P. If P^i is a e_i-fold point of $Cbar$, then we shall have $e_1 + ... + e_m \leq e$. We say that $P^1, \ldots, P^m$ are the points of $\bar{C}$ in the first neighborhood of P, and the multiplicity of $\bar{C}$ at P^i is e_i. Now iterate this procedure. The points of C *infinitely near* P can be diagrammed by the *singularity tree* of C at P.

At every node of this tree (including the root) we keep a count equal to the multiplicity of C at that point which will then be $\geq$ the number of branches arising at that node. It follows that every node higher than a certain level will be unforked, that is have a single branch. The desingularization theorem for algebraic plane curves, see [2] or [40] says that at every node higher than a certain level, the count equals one; in other words, C has only a finite number of singularities infinitely near P. Thus, since C has only finitely many *distinct* singularities, it follows that C has only a finite number of singular points, *distinct* as well as *infinitely near*.

Thus, by summing the counts of each node and counting $\frac{1}{2}e(e - 1)$ double points for a count e and additionally summing over all singularities of C and their corresponding singularity trees, we obtain a precise count of the total number of double points DP of C. This proper counting of double points yields the following theorem.

Theorem 7.2. *[Cayley-Reimann] $g = 0$ if and only if C has a rational parametrization.*

In other words the given plane curve has its maximum allowable limit of singularities if and only if it is rational.

Note also that in counting singularities we consider all the singularities of the projective curve. That is we consider the singularities at both finite distance as well as at infinity. The process of considering singularities at infinity is no different than that at finite distance. With regard to homogeneous coordinates let us consider $Z = 0$ to be the line at infinity. By swapping one of the axis lines $x = 0$ or $y = 0$ with the line at infinity we can bring the points at infinity to the affine plane. We illustrate this as well as *Theorem 2* by means of an example. Consider again the $m - fold$ cusp $y^2 - x^{2m+1}$. We have seen earlier that the origin accounts for m double points when counted properly. Now consider the singularity at infinity. We swap the $Z = 0$ line with the $Y = 0$ line by homogenizing and then setting $Y = 1$ to dehemogenize.

$$Y^2Z^{2m-1} - X^{2m+1} \Rightarrow z^{2m-1} - x^{2m+1}$$

The singularity at infinity is again at the origin and of multiplicity $2m - 1$ accounting for $\frac{1}{2}(2m - 1)(2m - 2)$ double points. On applying an appropriate quadratic transformation $x = \overline{x}$ and $z = \overline{xz}$, the multiplicity is reduced to 2:

$$\overline{z}^{2m - 1} - \overline{x}^2.$$

After a sequence of $m-1$ additional quadratic transformations the multiplicity at the origin finally reduces to one. These *infinitely near* singularities then account for totally $m-1$ additional double points, resulting in a total DP count for the curve to be equal to

$$m + \frac{1}{2}(2m - 1)(2m - 2) + m - 1 = \frac{1}{2}(2m)(2m-1)$$

which is exactly the maximum number of allowable double points for a curve of degree $2m + 1$. Hence the m-fold cusp has genus 0 and is rational with a parametrization given by

$$x = t^2, \qquad y = t^{2m + 1}.$$

7.1.1 Parameterizing with lines

The geometric idea of parameterizing a circle or a conic is to fix a point and take lines through that point which meet the conic at one additional point. Hence conics always have a rational parametrization, with the slope of the line being the single parameter. Next, consider a cubic curve, C_3. A cubic curve is a curve to which most lines intersect in three points. If we consider a *singular* cubic curve then lines through the singular (double) point meet the curve at one additional point and hence rationally parameterize the cubic curve. If C_3 has no singular points, then C_3 cannot be parameterized by rational functions. Now intersecting a curve C with a pencil of lines through a fixed point P on it, can be achieved by sending the point P on C to infinity. To understand this, let us first consider an irreducible conic which is represented by the equation

$$g(x, y) = ax^2 + by^2 + cxy + dx + ey + f = 0$$

Bezout confirms that the irreducible conic cannot contain a double point for otherwise the conic consists of two lines. We observe that the trivial parameterizable cases are the parabola $y^2 = x$ which has no term in x^2; the parabola $x^2 = y$ which has no term in y^2; and the hyperbola $xy = 1$ which has no terms in x^2 and y^2. The non-trivial case arises when a and b are both non-zero, e.g. the ellipse. This then suggests that to obtain a rational parametrization all we need to do is to kill the y^2 term. This can always be achieved by a suitable linear transformation resulting in the equation

$$(rx + s)y + (ux^2 + vx + w) = 0$$

from which one can easily obtain a rational parametrization

$$x = t, \quad y = \frac{-(ut^2 + vt + w)}{(rt + s)}$$

The elimination of the x^2 or the y^2 term through a coordinate transformation is said to make the conic irregular in x or y respectively. Geometrically speaking, a conic being irregular in x or y means that most lines parallel to the x or y axis respectively, intersect the conic in one point. Note that most lines through a fixed point on the conic meet the conic in one additional varying point. By sending the fixed point to infinity we make all these lines parallel to some axis and the curve irregular in one of the variables (x, or y) and hence amenable to parametrization. The coordinate transformation we select is thus one which sends any point on the conic to infinity along either of the coordinate axis x or y.

As an example consider the unit circle and fix a simple point $P(-1,\ 0)$ on it: (x, y) affine coordinates $(-1,\ 0)$ and $(X,\ Y,\ Z)$ homogeneous coordinates $(-1,\ 0,\ 1)$. Send P to a point at infinity along the y-axis. That is, send $(-1,\ 0,\ 1)$ to $(0,\ 1,\ 0)$. (Explanation: A point on y-axis is like $(0,\ p,\ 1)$ divide by p and let $p \to \infty$). Thus we achieve by a homogeneous linear transformation which transforms $(-1,\ 0,\ 1)$ to $(0,\ 1,\ 0)$

$$\begin{aligned} X &\to \alpha\overline{X} + \beta\overline{Y} + \gamma\overline{Z} \\ Y &\to \hat{\alpha}\overline{X} + \hat{\beta}\overline{Y} + \hat{\gamma}\overline{Z} \\ Z &\to \alpha^*\overline{X} + \beta^*\overline{Y} + \gamma^*\overline{Z} \end{aligned}$$

The chosen point on the circle $(-1,\ 0,\ 1)$ determines

$$-1 = \beta, \quad 0 = \hat{\beta}, \quad 1 = \beta^*$$

and the α's and γ's are chosen such that the *det* (α's , β's, γs) $! = 0$, yielding a well defined invertible transformation. So let us take as our homogeneous linear transformation

$$\begin{aligned} X &\to -\overline{Y}, \\ Y &\to \overline{Z}, \\ Z &\to \overline{X} + \overline{Y}. \end{aligned}$$

We first homogenize the circle $x^2 + y^2 - 1 = 0$ to $X^2 + Y^2 - Z^2 = 0$. On applying the above linear transformation we eliminate the $\overline{Y}^2$ term

$$\overline{Y}^2 + \overline{Z}^2 - (\overline{X} + \overline{Y})^2 = 0$$

$$\Rightarrow \quad -2\overline{XY} = \overline{X}^2 - \overline{Z}^2 \Rightarrow \overline{Y} = \frac{\overline{Z}^2 - \overline{X}^2}{2\overline{X}}.$$

Then on dehomogenizing by setting $\overline{Z} = 1$ and using the linear transformation to obtain the original affine coordinates and setting $\overline{X} = t$, we obtain the rational parametrization of the circle:

$$x = \frac{X}{Z} = \frac{-\overline{Y}}{\overline{X} + \overline{Y}}, \quad y = \frac{Y}{Z} = \frac{1}{\overline{X} + \overline{Y}}$$

$$\left\{ \begin{array}{lcl} \overline{X} &=& t \\ \overline{Y} &=& \frac{1 - t^2}{2t} \end{array} \right. \Rightarrow \left\{ \begin{array}{lclcl} x &=& \frac{-(1 - t^2)/2t}{t + (1 - t^2)/2t} &=& -\frac{1 - t^2}{1 + t^2} \\ y &=& \frac{1}{t + (1 - t^2)/2t} &=& \frac{2t}{1 + t^2} \end{array} \right.$$

In general, curves of degree d with a *distinct* $d-1$ fold point can be rationally parameterized by sending the $d-1$ fold point to infinity. Consider $f(x,\ y)$ a polynomial of degree d in x and y representing a plane algebraic curve C_d of degree d with a *distinct* $d-1$ fold singularity.

Note that singularities of a plane curve can be computed by simultaneously solving the equations $f = f_x = f_y = 0$ where f_x and f_y are the x and y partial derivatives of f, respectively. One way of obtaining the common solutions is to find those roots of $Res_x(f_x, f_y) = 0$ which are also the roots of $f = 0$. Here $Res_x(f_x, f_y)$ is the resultant of f_x and f_y treating them as polynomials in x. Note singularities at infinity can be obtained the same way after replacing the line at infinity with one of the coordinate axes. In particular on homogenizing a plane curve $f(x, y)$ to $F(X, Y, Z)$ we can set $Y = 1$ to obtain $fbar(x, z)$ thereby swapping the line at infinity $Z = 0$ with the line $Y = 0$. Now the above procedure can be applied to $fbar(x, z)$ to find the singularities at infinity.

Let us then compute the $d-1$ fold singularity of the curve C_d and translate it to the origin by a simple linear transformation. Then the polynomial describing the curve will be of the form

$$f(x,\ y) = f_d(x,\ y) + f_{d-1}(x, y)$$

where f_d, (degree form), consists of the terms of degree d and f_{d-1} consists of terms of degree $d-1$. On homogenizing this curve we obtain

$$F(X,Y,Z) = a_0Y^d + a_1Y^{d-1}X + \ldots + a_dX^d + b_0Y^{d-1}Z + b_1Y^{d-2}XZ + \ldots + b_dX^{d-1}Z$$

Now by sending the singular point $(0,0,1)$ to infinity along the Y axis we can eliminate the Y^d term. This as before by a homogeneous linear transformation which maps the point $(0,0,1)$ to the point $(0,1,0)$ and given by

$$X = \overline{X} \quad Y = \overline{Z}, \quad Z = \overline{Y}$$

which yields

$$F(\overline{X},\overline{Y},\overline{Z}) = a_0\overline{Z}^d + a_1\overline{Z}^{d-1}\overline{X} + \cdots + a_d\overline{X}^d + b_0\overline{Z}^{d-1}\overline{Y} + b_1\overline{Z}^{d-2}\overline{XY} + \cdots + b_d\overline{X}^{d-1}\overline{Y},$$

$$\overline{Y} = -\frac{a_0\overline{Z}^d + a_1\overline{Z}^{d-1}\overline{X} + \cdots + a_d\overline{X}^d}{b_0\overline{Z}^{d-1} + b_1\overline{Z}^{d-2}\overline{X} + \ldots + b_d\overline{X}^{d-1}}.$$

Then dehomogenizing, by setting $\overline{Z} = 1$ and using the linear transformation to obtain the original affine coordinates

$$x = \frac{X}{Z} = \frac{\overline{X}}{\overline{Y}}, \quad y = \frac{Y}{Z} = \frac{\overline{Z}}{\overline{Y}}$$

and setting $\overline{X} = t$ we obtain the rational parametrization of the curve.

Alternatively we could have symbolically intersected a single parameter family (*pencil*) of lines through the $d-1$ fold singularity with C_d and obtained a rational parameterization with respect to this parameter. This concept of passing a pencil of curves through singularities is generalized in the next section.

7.1.2 Parameterizing with Higher Degree Curves

From the genus formula and Bezout's theorem we note that an irreducible rational quartic curve in the plane has either a *distinct* triple point or three *distinct* double points. The rational parameterization of the quartic with a *distinct* triple point is handled by the method of the previous section. Let us then consider an irreducible quartic curve C_4 with three *distinct* double points. From Table 7.1 we know that through 5-points a conic can be passed. Choose three double points and a simple point on the curve C_4, yielding a one parameter family (pencil) of conics, $C_2(t)$. Now $C_4 \cdot C_2(t) = 8$ points. Since the fixed points (3 double points and a simple point) account for $2 + 2 + 2 + 1 = 7$ points, the remaining point on C_4 is the variable point, giving us a rational parametrization of C_4, in terms of parameter t.

Computationally we proceed as follows. Consider first C_4 with three distinct double points. We first obtain the three double point singularities of the homogeneous quartic $F(X,Y,Z)$ as well as a simple point on it. Let them be given by (X_1,Y_1,Z_1), (X_2,Y_2,Z_2), (X_3,Y_3,Z_3) and (X_4,Y_4,Z_4) respectively. Consider next the general equation of a homogeneous conic C_2 given by

$$G(X,Y,Z) = aX^2 \ + \ bY^2 + cXY + dXZ + eYZ + fZ^2 = 0$$

which has six coefficients however five independent unknowns as we can always divide out by one of the nonzero coefficients. We now try to determine these unknowns to yield a one parameter family of curves, $C_2(t)$. We pass C_2 simply through the singular double points and the simple point of C_4. (In general we shall pass a curve through an m-fold singularity with multiplicity $m-1$). In other words we equate for $i = 1, ..., 4$,

$$F(X_i,Y_i,Z_i) \ = \ G(X_i,Y_i,Z_i) \ = \ 0$$

This yields a linear system of 4 equations in five unknowns. Set one of the unknowns to be t and solve for the remaining unknowns in terms of t.

Next compute the intersection of C_4 and $C_2(t)$, by computing $Res_Y(F, G)$ which is a polynomial in X, Z and t. On dehomegenizing this polynomial by setting $Z = 1$, (since resultants of homogeneous polynomials are homogeneous) and dividing by the common factors $(x - x_i)^2$ for $i = 1..3$ and $(x - x_4)$ we obtain a polynomial linear in x which yields the rational parameterization. The process when repeated for y by taking the $Res_X(F, G)$ and dividing by the common factors $(y - y_i)^2$ for $i = 1..3$ and $(y - y_4)$ yields a polynomial in y and t and linear in y which yields the rational parameterization.

Next consider an example of a quintic curve with *infinitely near* singularities. In particular, the homogenized quintic cusp $C_5 : F(X, Y, Z) = Y^2Z^3 - X^5$ has a *distinct* double point and an *infinitely near* double point (in the first neighborhood) at (0 , 0 , 1), and a *distinct* triple point and an *infinitely near* double point at (0 , 1 , 0). Counting all the double points, properly, we see that $Csub5$ has 6 double points and hence is of genus 0 and rational. To obtain the parameterization we pass a one parameter family of cubics $C_3(t)$ given by

$$G(X, Y, Z) = aX^3 + bY^3 + cX^2Y + dXY^2 + eX^2Z + fY^2Z + gXYZ + hXZ^2 + iYZ^2 + jZ^3$$

through the singularities of C_5. Passing $C_3(t)$ through the *distinct* double point (with multiplicity $2 - 1 = 1$) is obtained as before by equating

$$F(0,\ 0,\ 1) = G(0,\ 0,\ 1) = 0 \tag{59}$$

and the *distinct* triple point, (with multiplicity $3 - 1 = 2$) by equating

$$F(0,\ 1,\ 0) = G(0,\ 1,\ 0) = 0 \tag{60}$$

$$F_X(0,\ 1,\ 0) = G_X(0,\ 1,\ 0) = 0 \tag{61}$$

$$F_Z(0,\ 1,\ 0) = G_Z(0,\ 1,\ 0) = 0 \tag{62}$$

These conditions for our example curve C_5 makes $j = 0$, $b = 0$, $d = 0$ and $f = 0$ in $C_3(t)$ yielding the curve

$$\overline{G}(X,\ Y,\ Z) = aX^3 + cX^2Y + eX^2Z + gXYZ + hXZ^2 + iYZ^2$$

We now wish to pass $C_3(t)$ through the *infinitely near* double point in the first neighborhood of the singularity at (0, 0, 1) of C_5. To achieve this we apply the quadratic transformation $X = \overline{X}$, $Y = \overline{XY}$, $Z = \overline{Z}$ centered at (0, 0, 1) to both $F(X,\ Y,\ Z)$ and $\overline{G}(X,\ Y,\ Z)$. The transformed equation $F_T = \overline{Y}^2\overline{Z}^3 - \overline{X}^3$ has a double point at (0, 0, 1) and we pass the curve of the transformed equation $\overline{G}_T = a\overline{X}sup2 + c\overline{X}^2Y + e\overline{XZ} + g\overline{XYZ} + h\overline{Z}^2 + i\overline{YZ}^2$ through the double point as before by equating

$$F_T(0,\ 0,\ 1) = \overline{G}_T(0,\ 0,\ 1) = 0... \tag{63}$$

This condition makes $h = 0$ in $C_3(t)$ yielding

$$\hat{G}(x,\ y,\ z) = aX^3 + cX^2Y + eX^2Z + gXYZ + iYZ^2.$$

Similarly we pass C_3 through the *infinitely near* double point in the first neighborhood of the singularity at (0, 1, 0) of C_5. To achieve this we apply the quadratic transformation $X = \hat{X}$, $Y = \hat{Y}$, $Z = \hat{X}\hat{Z}$ centered at (0, 1, 0) to both $F(X,\ Y,\ Z)$ and $Ghat(X,\ Y,\ Z)$. The transformed equation $F_T = \hat{Y}^2\hat{Z}^3 - Xhat^2$ has a double point at (0, 1, 0) and we pass the curve of the transformed equation $Ghat_T = a\hat{X} + c\hat{Y} + e\hat{X}\hat{Z} + g\hat{Y}\hat{Z} + i\hat{Y}\hat{Z}^2$ through the double point as before by equating

$$F_T(0,\ 1,\ 0) = \hat{G}_T(0,\ 1,\ 0) = 0 \tag{64}$$

This condition makes $c = 0$ in C_3 yielding

$$\tilde{G}(x,\ y,\ z) = aX^3 + eX^2Z + gXYZ + iYZ^2.$$

Our final condition to determine pencil of cubics $C_3(t)$ is to choose two simple points on C_5, say $(1,\ 1,\ 1)$ and $(1,\ -1,\ 1)$ and pass C_3 through it by equating.

$$F(1,\ 1,\ 1) = \tilde{G}(1,\ 1,\ 1) = 0 \tag{65}$$
$$F(1,\ -1,\ 1) = \tilde{G}(1,\ -1,\ 1) = 0. \tag{66}$$

Note that in total we applied eight conditions to determine the pencil, since nine conditions completely determine the cubic. The last two conditions yield the equations

$$a + e + g + i = 0$$
$$a + e - g - i = 0$$

In choosing the pencil $C_3(t)$ we allow one of the coefficients to be t and we may divide out by another coefficient (or choose it to be 1). The above equations yield $a + e = 0$ and $g + i = 0$ and on choosing $a = t$ and $g = 1$ we obtain $e = -t$ and $i = -1$. Hence our homogeneous cubic pencil is given by

$$G_3(X,\ Y,\ Z,\ t) = tX^3 - tX^2Z + XYZ - YZ^2$$

or the dehomogenized pencil $G_3(x,\ y,\ t) = tx^3 - tx^2 + xy - y = 0$. This yields $y = -tx^2$. Intersecting it with the dehomogenized quintic $C_5 : y_2 - x^5$ yields $t^2x^4 - xsup5 = 0$ or $x = t^2$ on dividing out by the common factor x^4. Finally the parametric equations of the rational quintic C_5 are given by $x = t^2$ and $y = -t^5$.

In the general case we consider an irreducible curve C_d with the appropriate number of *distinct* and *infinitely near* singularities which make C_d rational (*genus* 0). We pass a curve $C_d - 2$ through these singular points and $d-3$ additional simple points of C_d. Consider again $F(X,Y,Z)$ and $G(X,Y,Z)$ as the homogeneous equations of curves C_d and $C_d - 2$ respectively. For a distinct singular point of multiplicity m of C_d at the point (X_i, Y_i, Z_i) we pass the curve $C_d - 2$ through it with a multiplicity of $m-1$. To achieve this we equate

$$\begin{aligned}
F(X_i,Y_i,Z_i) &= G(X_i,Y_i,Z_i)\\
F_X(X_i,Y_i,Z_i) &= G_X(X_i,Y_i,Z_i)\\
F_Y(X_i,Y_i,Z_i) &= G_Y(X_i,Y_i,Z_i)\\
F_{XX}(X_i,Y_i,Z_i) &= G_{XX}(X_i,Y_i,Z_i)\\
F_{XY}(X_i,Y_i,Z_i) &= G_{XY}(X_i,Y_i,Z_i)\\
F_{YY}(X_i,Y_i,Z_i) &= G_{YY}(X_i,Y_i,Z_i)\\
&\vdots\\
F_{X^jY^k}(X_i,Y_i,Z_i) &= G_{X^jY^k}(X_i,Y_i,Z_i), \quad 0 \le j+k \le m-2.
\end{aligned}$$

For an *infinitely near* singular point of C_d with its associated singularity tree we pass the curve $C_d - 2$ with multiplicity $r-1$ through each of the points of multiplicity r in the first, second, third, ..., neighborhoods. To achieve this we apply quadratic transformations T_i to both $F(X,Y,Z)$ and $G(X,Y,Z)$ centered around the *infinitely near* singular points corresponding to the singularity tree. The appropriate multiplicity of passing is achieved by equating the transformed equations F_{T_i} and G_{T_i} and their partial derivatives as above.

A simple counting argument now shows us that this method generates the correct number of conditions which specifies $C_d - 2$ and furthermore the total intersection count between C_d and

$C_d - 2$ satisfies *Bezout*. A curve C_d of *genus* $= 0$ has the equivalent of exactly $\frac{1}{2}(d-1)(d \ - \ 2)$ double points. Then to pass a curve $C_d - 2$ through these double points and $d-3$ other fixed simple points of C_d and one variable point specified by t, the total number of conditions (= to the total number of linear equations) is given by

$$\frac{1}{2}(d \ - \ 1)(d \ - \ 2) + (d \ - \ 3) \ + \ 1 \ = \frac{1}{2}(d \ - \ 2)(d \ + \ 1)$$

which is exactly the number of independent unknowns to determine $C_d - 2$ (see table 7.1). Next, counting the number of points of intersection of C_d and $C_{d \ - \ 2}$

$$(d \ - \ 1)(d \ - \ 2) \ + \ d \ - \ 3 \ + \ 1 = (d \ - \ 2)d = C_{d-2} \cdot \ C_d$$

satisfying Bezout.

For further details of the applicability of Bezout's theorem with respect to *infinitely near* singularities, see Abhyankar (1973). Then computing the $Res_x(C_d, \ C_d - 2)$ which yields a polynomial of degree $d(d-2)$ in y and dividing by the common factors corresponding to the $(d-3)$ simple points (a polynomial of degree $(d-3)$ in y) and $\frac{1}{2}(d-2)(d-1)$ double points (a polynomial of degree $(d-2)(d-1)$ in y) yields a polynomial in y and t which is linear in y, (for the single variable point) and thus gives a rational parameterization of y in terms of t. Similarly repeating with Res""$_y(C_d, \ C_d - 2)$ yields a rational parameterization of x in terms of t.

As a final example consider the m-fold cusp $y^2 \ - \ x^2m + 1$ once again (for the last time). We know from the above that it is a rational curve with genus 0 and with a distinct double point and $m-1$ infinitely near double points at the origin (0,0,1) and a distinct $(2m-1)$-fold singularity and $m-1$ infinitely near double points at infinity (0,1,0). Now we pass a pencil of curve C_{2m-1} of degree $2m-1$ appropriately (as explained above) through these singularities and also through $2m+1 \ - \ 3 \ = \ 2m-2$ simple points of the m-fold cusp C_{2m+1}.

In the following let $F(X, \ Y, \ Z) \ = \ 0$ be the equation of C_{2m+1} and $G(X, \ Y, \ Z)$ the equation of C_{2m-1}. Now the conditions available to specify a pencil of curves C_{2m-1} is given as follows. A total of $2m-2$ conditions are given by equating F and G at the $2m-2$ simple points of C_{2m+1}. Further by equating F and G and the corresponding transformed F_{T_i} and G_{T_i} (transformed by a sequence of quadratic transformations) at the *distinct* and *infinitely near* double points of the origin (0,0,1) and *infinitely near* double points of infinity (0,1,0). This totally accounts for $m \ + \ m-1 \ = \ 2m-1$ additional conditions. Finally through the $(2m-1)$ fold singularity at infinity of C_{2m+1} the pencil C_{2m-1} is passed with multiplicity $2m-2$ which is obtained by equating the equations and the partial derivatives $F_{X^iY^k} \ = \ G_{X^iY^k}$ for all $0 \ <= \ j \ + \ k \ < \ 2m-2$ which yields $\frac{1}{2}(2m-2)(2m-1)$ conditions. One final condition is achieved by equating one of the coefficients of C_{2m-1} to 't'. Hence totally the conditions available to specify the pencil of curves C_{2m-1} is given by

$$1 \ + \ 2m-2 \ + \ 2m-1 + \frac{1}{2}(2m-2)(2m-1) = \frac{1}{2}(2m-1)(2m+2)$$

which is exactly the number of conditions required to specify a pencil of curve C_{2m-1} as given by Table 7.1 This then yields a linear system of $(2m-1)(m+1)$ equations in the same number of unknowns and can be easily solved.

Finally, note that the total number of intersections (counting multiplicities) between C_{2m-1} are given by 1 single variable point + $(2m-2)$ fixed simple points + $2(2m-1)$ double points + $(2m-1)(2m-2)$ $2m-2$ multiplicity of C_{2m-1} at the $(2m-1)$-fold singularity of $C_{2m+1} = (2m-1)(2m+1)$ satisfying Bezout. Hence on computing the $Res_x(C_{2m+1}, \ C_{2m-1}$ and dividing by the common factors corresponding to the $(2m-2)$ simple points, $(2m-1)$ double points and the $2m-2$ multiplicity of C_{2m-1} at the $(2m-1)$-fold singularity of C_{2m+1} yields a polynomial in y and t which is linear in y, (for the single variable point) and thus gives a rational parameterization of y in terms of t. Similarly repeating with $Res_y(C_{2m+1}, \ C_{2m-1})$ yields a rational parameterization of x in terms of t.

7.1.3 Parameterization of conic, cubic plane curves

Conics A general degree two algebraic plane curve is given as $f(x,y) = ax^2 + by^2 + cxy + dx + ey + f = 0$. If either a or b is 0, then set the the solution for the corresponding variable gives us a parameterization. Otherwise, a point on the curve is sent to infinity, thus making the curve linear in one of the variables. We obtain a rational parameterization with polynomials of maximum degree two.

Cubics A general degree two algebraic plane curve is given as $f(x,y,z) = ax^3 + by^3 + cx^2y + dxy^2 + ex^2 + fy^2 + gxy + hx + iy + j = 0$. Since a cubic equation has a real root, a real point at infinity, remove say the y^3 term with a linear transformation of x and y. Rearrange the terms to obtain a quadratic in transformed y. Taking the root of this equation, we will obtain a function for the transformed y. Hence if there exists a parameterization for this function, we have obtained the parameterization of the cubic curve.

7.2 Parameterization of Algebraic Space Curves

Consider an irreducible algebraic space curve C which is implicitly defined as the intersection of two algebraic surfaces $f(x,y,z) = 0$ and $g(x,y,z) = 0$. We concern ourselves with space curves defined by two surfaces since they are of direct interest to applications in computer-aided design and computer graphics, see Boehm, et. al [17]. Irreducible space curves in general, defined by more than two surfaces are difficult to handle equationally and one needs to resort to computationally less efficient ideal-theoretic methods, Buchberger [19]. However general space curves is a topic with various unresolved issues of mathematical and computational interest and an area of important future research, Abhyankar [1].

Now for an irreducible algebraic space curve C as above, there always exists a birational correspondence between the points of C and the points of an irreducible plane curve P whose genus is the same as that of C, see Walker [40]. Birational correspondence between C and P means that the points of C can be given by rational functions of points of P and vice versa (i.e a 1-1 mapping, except for a finite number of exceptional points, between points of C and P). Together, (i) the method of computing the genus and rational parameterization of algebraic plane curves, Abhyankar and Bajaj [5], and (ii) the method of this paper of constructing a plane curve P along with a birational mapping between the points of P and the given space curve C, gives an algorithm to compute the genus of C and if genus = 0 the rational parametric equations of C.

We now show how, given an irreducible space curve C, defined implicitly as the intersection of two algebraic surfaces, one is able to construct the equation of a plane curve P and a birational mapping between the points of P and C. As a first attempt in constructing P, we may consider the projection of the space curve C along any of the coordinate axis yielding a plane curve whose points are in correspondence with the points of C. Projecting C along, say the z axis, can be achieved by computing the Sylvester resultant of f and g, treating them as polynomials in z, yielding a single polynomial in x and y the coefficients of f and g. The Sylvester resultant eliminates one variable, in this case z, from two equations, see Salmon [32]. Efficient methods are known for computing this resultant for polynomials in any number of variables, see Collins [22], Bajaj and Royappa [12]. The Sylvester resultant of f and g thus defines a plane algebraic curve P. However this projected plane curve P in general, is not in birational correspondence with the space curve C. For a chosen projection direction it is quite possible that most points of P may correspond to more than one point of C (i.e. a multiple covering of P by C) and hence the two curves are not birationally related. However this approach may be rectified by choosing a valid projection direction which yields a birationally related, projected plane curve P.

There remains the problem of constructing the birational mapping between points on P and C. Let the projected plane curve P be defined by the polynomial $h(\tilde{x}, \tilde{y})$. The map one way is

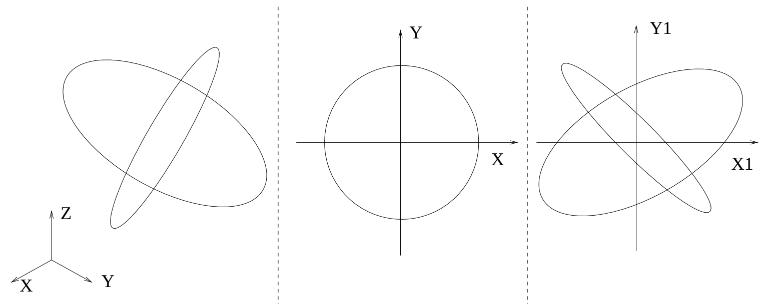


Figure 3: Two intersecting space curves. A valid projection direction will yield two planar curves intersecting transversally.

linear and is given trivially by $\tilde{x} = x$ and $\tilde{y} = y$. To construct the reverse rational map one only needs to compute $z = I(\tilde{x}, \tilde{y})$ where I is a rational function. We show how it is always possible to construct this rational function by use of a polynomial remainder sequence along a valid direction. In fact the resultant is no more than the end result of a polynomial remainder sequence, see Bocher [16], Collins [21].

Note additionally that the reverse rational map, $z = I(\tilde{x}, \tilde{y})$ where I is a rational function is also the rational parametric equation of a rational surface containing the space curve C. Hence constructing a birational mapping between space and plane curves which always exists, also yields an explicit rational surface containing the space curve. By an explicit rational surface we mean one with a known or trivially derivable rational parameterization. For irreducible space curves C, a method of obtaining an explicit rational surface containing C, is given (without proof) in Snyder and Sisam [39]. The technique presented here is similar, but uses a subresultant polynomial remainder sequence, which for an appropriately chosen coordinate direction, provides an efficient way of obtaining the reverse rational map as well as an explicit rational surface containing C.

It is important to note that conversely knowing the rational parametric equations of a rational surface containing a space curve, yields a birational mapping between points on the space curve and a plane curve. Namely, if one of the two surfaces f or g defining the space curve C, or actually any known surface in $I(f, g)$, the Ideal[2] of the curve generated by f and g is rational with a known rational parameterization, then points on C are easily mapped to a single polynomial equation $h(s, t) = 0$ describing a plane curve P in the parametric plane $s - t$ of the rational surface. This mapping between the (x, y, z) points of C and the (s, t) points of P is birational with the reverse rational map, from the points on P to points on C being given by the parametric equations of the rational surface. For space curves C which have a quadric or a rational cubic surface in its Ideal, the plane curve P and the rational mapping from the points on P to C are easily constructed by using known techniques for parameterizing these rational surface, see Abhyankar and Bajaj [3, 4], Sederberg and Snively [36].

The rest of this paper is structured as follows. Section 2 describes a method of choosing a valid direction of projection for the space curve C. This then also yields a projected plane curve P in birational correspondence to C. Using these results, Section 3 describes a method of constructing the reverse rational map between points on the plane curve P and points on C.

Valid Projection Direction To find an appropriate axis of projection, the following general procedure may be adopted. Consider the linear transformation $x = a_1x_1 + b_1y_1 + c_1z_1$, $y = a_2x_1 + b_2y_1 + c_2z_1$ and $z = a_3x_1 + b_3y_1 + c_3z_1$. On substituting into the equations of the two surfaces defining the space curve we obtain the transformed equations $f_1(x_1, y_1, z_1) = 0$ and

[2] $I(f, g) = \{h(x, y, z) \mid h = \alpha f + \beta g$ for any polynomials $\alpha(x, y, z)$ and $\beta(x, y, z)\}$.

$g_1(x_1, y_1, z_1) = 0$. Next compute the $Res_{z_1}(f_1, g_1)$ which is a polynomial $h(x_1, y_1)$ describing the projection along the Z axis of the space curve C onto the $z = 0$ plane.

Since C is irreducible and f and g are not tangent along C, the order of $h(x_1, y_1)$ is exactly equal to the projection degree, see [1]. By order of $h(x_1, y_1)$ we mean k, if $h(x_1, y_1) = (g(x_1, y_1))^k$. For a birational mapping we desire a projection degree equal to one. Hence, we choose the coefficients of the linear transformation, a_i, b_i and c_i such that (i) the determinant of a_i, b_i and c_i is non zero and (ii) the equation of the projected plane curve $h(x_1, y_1)$ is not a power of an irreducible polynomial. The latter can be achieved by making the discriminant $Res_{x_1}(h, h_{x_1})$ to be non zero. Note, a random choice of coefficients would also work with probability 1, since the set of coefficients which make the determinant and $Res_{x_1}(h, h_{x_1})$ equal to zero, are restricted to the points of a lower dimensional hypersurface. See [34] where the notion of randomized computations with algebraic varieties is made precise. A suitable choice of coefficients thus ensures that the projected irreducible plane curve given by $h(x_1, y_1)$ is in birational correspondence with the irreducible space curve and thus of the same genus. The parameterization methods of Abhyankar and Bajaj [4] for algebraic plane curves are now applicable and thereby yield a genus computation as well as an algorithm for rationally parameterizing the space curve.

Constructing the Birational Map We choose a valid projection direction by the method described in the earlier section. Without loss of generality let this direction be the Z axis. Let the surfaces $f(x, y, z) = 0$ and $g(x, y, z) = 0$ be of degrees m_1 and m_2 in z, respectively. Again, without loss of generality, assume $m_1 \geq m_2$. Let $F_1 = f(x, y, z)$ and $F_2 = g(x, y, z)$ be given by

$$\begin{aligned} F_1 &= f_0\, z^{m_1} + f_1\, z^{m_1 - 1} + \ldots + f_{m_1 - 1}\, z + f_{m_1} \\ F_2 &= g_0\, z^{m_2} + g_1\, z^{m_2 - 1} + \ldots + g_{m_2 - 1}\, z + g_{m_2} \end{aligned} \tag{67}$$

with f_j, $(j = 0 \ldots m_1)$ and g_k, $(k = 0 \ldots m_2)$, denoting polynomials in x, y. Then, there exist polynomials $F_{i+2}(x, y, z)$, for $i = 1 \ldots k$, such that $A_i\, F_i = Q_i\, F_{i+1} + B_i\, F_{i+2}$ with m_{i+2}, the degree of z in F_{i+2}, less than m_{i+1}, the degree of z in F_{i+1} and certain polynomials $A_i(x, y)$, $Q_i(x, y, z)$ and $B_i(x, y)$. The polynomials F_{i+2}, $i = 1, 2, \ldots$ form, what is known as a polynomial remainder sequence and can be computed in various different ways, as we now describe.

Let $lc(A)$ denote the leading coefficient of polynomial A, viewed as a polynomial in z, (i.e. coefficient of term with highest z degree). Further let c_i denote $lc(F_i)$. To compute F_{i+2} from F_i and F_{i+1} we first begin with $R_i^0 = F_i$ and then,

$$\begin{array}{llll} for & k & = 1, \ldots, m_i - m_i + 1 + 1 & \\ & if & lc(R_i^k - 1) = 0 & \\ & then & R_i^k = R_i^k - 1 & \\ & else & R_i^k = c_{i+1}\, R_i^{k-1} - z^{m_i - m_{i+1} + 1 - k}\, lc(R_i^{k-1})\, F_{i+1} & \end{array} \tag{68}$$

The polynomial $R_i^{m_i - m_{i+1} + 1}$ is known as the psuedo-remainder of F_i and F_{i+1}. Using Collin's reduced PRS method [21], one constructs the polynomial $F_{i+2} = \frac{R_i^{m_i - m_{i+1} + 1}}{d_{i-1}}$ where $d_0 = 1$ and $d_i = c_{i+1}^{m_i - m_{i+1} + 1}$. Using Brown's subresultant PRS scheme [18], one constructs the polynomial $F_{i+2} = (-1)^{m_i - m_{i+1} + 1} \frac{R_i^{m_i - m_{i+1} + 1}}{c_i\, E_{m_i}}$ where $E_{m_1} = 1$ and $E_{m_{i+1}} = \frac{c_{i+1}^{m_i - m_{i+1}}}{E_{m_i}^{m_i - m_{i+1} - 1}}$. As shown by Loos [28], both the above methods, as well as others, follow naturally from the subresultant theorem of Habicht.

Thus starting with polynomials F_1 and F_2 one constructs the polynomial remainder sequence, $F_1, F_2, F_3, \ldots F_i, \ldots F_r$ with m_i, the z degree of F_i less than m_{i-1}, the z degree of F_{i-1} and $m_r = 0$ (i.e. F_r being independent of z). We choose the subresultant PRS scheme for its computational superiority and also because each $F_i = S_{m_{i-1}-1}$, $1 \le i \le r$, where S_k is the k^{th} subresultant of F_1 and F_2, see [16, 18, 21].

Now for any i, if F_{i-1} and F_i are of degree greater than two and F_{i+1} is independent of z then the Z axis is not a valid projection direction. This may be seen as follows. Since the Z axis was chosen as a valid projection direction, the $Res_z[f(x,y,z), g(x,y,z)] = Res_z[F_1, F_2]$ is non-zero and not a multiple of some irreducible polynomial. This holds for any two surfaces $f = F_{i-1}$ and F_i in the polynomial remainder sequence where each of the subresultants is also not a multiple of some irreducible polynomial. To complete the argument, it remains to see that by induction if F_{i-1} and F_i are of say degree three and two respectively and F_{i+1} is independent of z then the $Res_z(F_{i-1}, F_i)$ is equal to some $h^3(x,y)$, which is impossible.

Hence in the polynomial remainder sequence there exists a polynomial remainder which is linear in z, i.e., $F_{r-1} = z\Phi_1(x, y) - \Phi_2(x, y) = 0$. Thus on computing the polynomial remainder sequence and obtaining F_{r-1}, one is able to construct the required inverse map, $z = \frac{\Phi_2(x,y)}{\Phi_1(x,y)}$, which also is a rational surface containing the space curve. The rational parameterization of this rational surface is trivially given by $x = s$, $y = t$ and $z = \frac{\Phi_2(s, t)}{\Phi_1(s, t)}$.

The method of the earlier sections of constructing the inverse rational map as well as a rational surface containing the space curve can be applied for reducible as well as irreducible curves, defined implicitly as the intersection of two surfaces. The one limitation however is the assumption of non-tangency of the surfaces meeting along the space curve. It remains open to construct a birational map as well as a rational surface containing a space curve when the two surfaces defining the space curve are also tangent along the entire curve.

7.3 Automatic Parametrization of Degree 2 Curves and Surfaces

General curves and surfaces can be represented by implicit or parametric equations. A general (degree two) conic implicit equation is given by $C(x,y) = ax^2 + by^2 + cxy + dx + ey + f = 0$, and rational parametric equations given by $x = u(t)/w(t)$ and $y = v(t)/w(t)$, where u, v and w are no more than quadratic polynomials. Further a general (degree two) conicoid implicit equation is given by $C(x,y,z) = ax^2 + by^2 + cz^2 + dxy + exz + fyz + gx + hy + iz + j = 0$, with corresponding rational parametric equations $x = u(s,t)/q(s,t)$, $y = v(s,t)/q(s,t)$, and $z = w(s,t)/q(s,t)$, where again u, v, w and q are no more than quadratic polynomials. The rational parametric form of representing a surface allows greater ease for transformation and shape control, Tiller (1983), Mortenson (1985). The implicit form is preferred for testing whether a point is above, on, or below the surface, where above and below is determined relative to the direction of the surface normal. As both forms have their inherent advantages it becomes crucial to be able to go efficiently from one form to the other, especially when surfaces of an object are automatically generated in one of the two representations.

Both conics and conicoids always have a rational parameterization. We describe algorithms to obtain rational parametric equations for the conics and conicoids, given the implicit equations. Polynomial parameterizations are also obtained whenever they exist for the conics and conicoids. These parameterizations are at most degree 2 and are over the field of Reals, or the field of Complex numbers when real solutions do not exist. We consider obtaining rational parameterizations over Q, the fields of rationals. Computations over Q are exact and hence give rise to stable computational algorithms as opposed to finite precision calculations with real numbers. Additionally, considering rational parameterizations over Q proves to be an interesting mathematical question in its own right.

Cubicoids (degree 3) surfaces also always have a rational parameterization. On the other hand cubics (degree 3) plane curves do not always have a rational parameterization. However they always have a parameterization of the type which allows a single square root of rational functions. In a companion paper, Abhyankar and Bajaj (1986), show how to obtain the rational and special parametric equations for cubics and cubicoids. Higher degree curves and surfaces in general are not rational. The reverse problem of converting from parametric to implicit equations, called implicitization has been considered computationally by various authors in the past, see Collins (1971) and Sederberg et. al., (1985). However as yet no correct closed form solution is known for implicitizing rational surfaces or in general, implicitizing parametric algebraic varieties.

7.3.1 Conics

The general conic implicit equation is given by $C(x,y) = ax^2 + by^2 + cxy + dx + ey + f = 0$. The non-trivial case in converting this to a rational parameterization arises when a and b are both non-zero. Otherwise one already has one variable (x, or y) in linear form and expressible as a rational polynomial expression of the other, and hence a rational parameterization. This then suggests that to obtain a rational parameterization all we need to do is to make $C(x,y)$ non-regular in x or y. That is, eliminate the x^2 or the y^2 term through a coordinate transformation. For then one of the variables is again in linear form and is expressible as a rational polynomial expression of the other. We choose to eliminate the y^2 term, by an appropriate coordinate transformation applied to $C(x,y)$. This is always possible and the algorithm is now described below. (The entire algorithm which also handles all trivial and degenerate cases of the conic is implemented on a VAX-780 using VAXIMA.)

Geometrically speaking, a conic being irregular in x or y means that most lines parallel to the x or y axis respectively, intersect the curve in one point. Also, most lines through a point (b_1, b_2) on the conic meet the conic in one additional point. By sending this point (b_1, b_2) to infinity we make all these lines parallel to some axis and the curve irregular in one of the variables (x, or y) and hence amenable to parameterization. The coordinate transformation we select is thus one which sends the point (b_1, b_2) on the conic to infinity. The rational parameterization we obtain is global, of degree at most 2 and with parameter t corresponding to the slopes of the lines through the point (b_1, b_2) on the conic. Further t ranges from $(-\infty, \infty)$ and covers the entire curve. The selection of the point (b_1, b_2) on the conic becomes important and may be made appropriately, when the parameterization is desired only for a specific piece of the conic.

If $C(x,y)$ has a real root at infinity, a *linear* transformation of the type $x' = a_1x + b_1y + c_1$ and $y' = a_2x + b_2y + c_2$ will suffice. If $C(x,y)$ has no real root at infinity, we must use a *linear transformation* of the type $x' = (a_1x + b_1y + c_1)/(a_3x + b_3y + c_3)$ and $y' = a_2x + b_2y + c_2)/(a_3x + b_3y + c_3)$. This is equivalent to a *homogeneous linear transformation* of the type $X' = a_1X + b_1Y + c_1H$, $Y' = a_2X + b_2Y + c_2H$ and $H' = a_3X + b_3Y + c_3H$ applied to the homogeneous conic $C(X,Y,H) = aX^2 + bY^2 + cXY + dXH + eYH + fH^2 = 0$.

Step (2) Points at infinity for $C(x,y)$ are given by the linear factors of the *degree form* (highest degree terms) of I. For the conic this corresponds to a real root at infinity if $c^2 >= 4ab$, (e.g. parabolas and hyperbolas). For otherwise both roots at infinity are complex , (complex roots arise in conjugate pairs). Further for $c^2 = 4ab$, (e.g. parabolas), the degree form is a perfect square and this gives a polynomial parameterization for the curve.

Step (3) Applying a linear transformation for $c^2 >= 4ab$, gives rise to $C(x', y') = I(a_1x + b_1y + c_1, a_2x + b_2y + c_2)$. To eliminate the y^2 term we need to choose b_1 and b_2 such that $ab_1^2 + cb_1b_2 + bb_2^2 = 0$. Here both the values of b_1 and b_2 can always be chosen to be real.

Step (4) Applying a homogeneous linear transformation for $c^2 < 4ab$, gives rise to $C(X', Y', H') = C(a_1X + b_1Y + c_1H, a_2X + b_2Y + c_2H, a_3X + b_3Y + c_3H)$. To eliminate the Y^2 term we need to choose b_1, b_2 and b_3 such that $ab_1^2 + bb_2^2 + cb_1b_2 + db_1b_3 + eb_2b_3 + fb_3^2 = 0$. This is equivalent to finding a point (b_1, b_2, b_3) on the homogeneous conic. The values of b_1 and b_2

are both real if $(cd - 2ae)$ is not less than the *geometric mean* of $4af - d^2$ and $4ab - c^2$, or alternatively $(ce - 2bd)$ is not less than the *geometric mean* of $4bf - e^2$ and $4ab - c^2$.

Step (5) Finally choose the remaining coefficients a_i' s, c_i' s, ensuring that the appropriate transformation is well defined. In the case of a linear transformation, this corresponds to ensuring that the matrix

$$\begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix}$$

is non-singular. Hence c_i' s can be chosen to be 0 and $a_1 = 1$, $a_2 = 0$. In the case of a homogeneous linear transformation, one needs to ensure that the matrix

$$\begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}$$

is non-singular. Here $a_1 = 1$, $c_2 = 1$ and the rest set to 0 suffices. These remaining coefficients provide a measure of local control for the curve and may also be chosen in a way that gives specific local parameterizations for pieces of the curve, appropriate for particular applications.

Conicoids

The case of the conicoid is a generalization of the method of the conic. The general conicoid implicit equation is given by $C(x, y, z) = ax^2 + by^2 + cz^2 + dxy + exz + fyz + gx + hy + iz + j = 0$. Again the main case of concern is when a, b and c are all non-zero. Otherwise one already has one of the variables (x, y, or z) in linear form and expressible as a rational polynomial expression of the other two. This then suggests that to obtain the rational parameterization all we need to do again is to make $C(x, y, z)$ *non-regular* in say, y. That is, eliminate the y^2 term through a coordinate transformation. For then y is in linear form and is expressible as a rational polynomial expression of the other two. We eliminate the y^2 term by an appropriate coordinate transformation applied to $C(x, y, z)$. This is always possible and the algorithm is now described below. (The entire algorithm which also handles all trivial and degenerate cases of the conicoid is implemented on a VAX-780 using VAXIMA.)

Geometrically speaking, a conicoid being irregular in x, y or z means that most lines parallel to the x, y or z axis respectively, intersect the surface in one point. Also, most lines through a point (b_1, b_2, b_3) on the conicoid meet the conicoid in one additional point. By sending this point (b_1, b_2, b_3) to infinity we make all these lines parallel to some axis and the surface irregular in one of the variables (x, or y) and hence amenable to parameterization. The coordinate transformation we select is thus one which sends the point (b_1, b_2, b_3) on the conicoid to infinity. The rational parameterization we obtain is global, of degree at most 2 and with parameters s and t corresponding to the ratio of the direction cosines of the lines through the point (b_1, b_2, b_3) on the conicoid. Further s and t both range from $(-\infty, \infty)$ and cover the entire surface. The selection of the point (b_1, b_2, b_3) on the conicoid becomes important and may be made appropriately, when the parameterization is desired only for a specific patch of the conicoid.

Step (1) If $C(x, y, z)$ has a real root at infinity, a*linear transformation* of the type $x' = a_1x + b_1y + c_1z + d_1$, $y' = a_2x + b_2y + c_2z + d_2$ and $z' = a_3x + b_3y + c_3z + d_3$ will suffice. If $C(x, y, z)$ has no real root at infinity, we must use a *linear transformation* of the type $x' = (a_1x + b_1y + c_1z + d_1)/(a_4x + b_4y + c_4z + d_4)$, $y' = a_2x + b_2y + c_2z + d_2)/(a_4x + b_4y + c_4z + d_4)$. and $z' = (a_3x + b_3y + c_3z + d_3)/(a_4x + b_4y + c_4z + d_4)$. This is equivalent to a *homogeneous linear transformation* of the type $X' = a_1X + b_1Y + c_1Z + d_1H$, $Y' = a_2X + b_2Y + c_2Z + d_2H$, $Z' = a_3X + b_3Y + c_3Z + d_3H$ and $H' = a_4X + b_4Y + c_4Z + d_4H$ applied to the homogeneous conicoid $C(X, Y, Z, H) = aX^2 + bY^2 + cZ^2 + dXY + eXZ + fYZ + gXH + hYH + iZH + jH^2 = 0$,

Step (2) Points at infinity for $C(x, y)$ are given by the linear factors of the *degree form* (highest degree terms) of I. For the conicoid this corresponds to the roots of the homogeneous conic equation $C(x, y, z) = ax^2 + by^2 + dxy + exz + fyz + cz^2 = 0$. Also, here the simultaneous truth of $d^2 = 4ab$, $e^2 = 4ac$ and $f^2 = 4bc$ corresponds to the existence of a polynomial parameterization for the conicoid, as then the degree form is a perfect square.

Step (3) Apply a linear transformation if a real root (r_x, r_y, r_z) exists for the homogeneous conic $C(x, y, z)$ of (2). This gives rise to $C(x', y', z') = I(\ a_1x + b_1y + c_1z + d_1, a_2x + b_2y + c_2z + d_2, a_3x + b_3y + c_3z + d_3)$. To eliminate the y^2 term we can take $(\ b_1, b_2, b_3) = (\ r_x, r_y, r_z)$, the real point on $C(x, y, z)$.

Step (4) Apply a homogeneous linear transformation if only complex roots exist for the homogeneous conic $C(x, y, z)$ of (2). This gives rise to $C(X', Y', Z', H') = I(\ a_1X + b_1Y + c_1Z + d_1H, a_2X + b_2Y + c_2Z + d_2H, a_3X + b_3Y + c_3Z + d_3H, a_4X + b_4Y + c_4Z + d_4H)$. To eliminate the Y^2 term we choose $b_4 = 1$ and $(\ b_1, b_2)$ to be a point on either the conic $ax^2 + by^2 + dxy + gxz + hyz + jz^2 = 0$ with $b_3 = 0$ or a point on the conic $ax^2 + by^2 + dxy + (e + g)xz + (f + h)yz + (c + i + j)z^2 = 0$ with $b_3 = 1$. Real values exist for b_1 and b_2 if there exists a real point on either of the above conics.

Step (5) Finally choose the remaining coefficients a_i' s, c_i' s, and d_i' s, ensuring that the appropriate transformation is well defined. In the case of a linear transformation, this corresponds to ensuring that the matrix

$$\begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}$$

is non-singular. Here the d_i' s can be chosen to be 0. Further $a_2 = 1$, $c_3 = 1$ if b_1 is non-zero or else $a_1 = 1$, $c_3 = 1$ if b_2 is non-zero or else $a_1 = 1$, $c_2 = 1$, with the rest set to 0. In the case of a homogeneous linear transformation one needs to ensure that the matrix

$$\begin{pmatrix} a_1 & b_1 & c_1 & d_1 \\ a_2 & b_2 & c_2 & d_2 \\ a_3 & b_3 & c_3 & d_3 \\ a_4 & b_4 & c_4 & d_4 \end{pmatrix}$$

is non-singular. Here $a_1 = 1$, $c_3 = 1$, $d_2 = 1$ with the rest set to 0 suffices. These remaining coefficients provide a measure of local control for the surface and may also be chosen in a way that gives specific local parameterizations for pieces of the surface, appropriate for particular applications.

7.3.2 Rational Fields

As seen from the previous sections one obtains parameterizations over the reals or the complex numbers if the corresponding coefficients of the appropriate transformations are over the fields of reals or complex numbers respectively. The coefficients themselves correspond to finding real or complex points on various conic equations. Thus the question of whether the rational parameterization for conics and conicoids is possible over Q, the field of rationals, reduces to the question of whether there exists a rational root of a certain conic equation with integral coefficients, (or an integral root of the homogenized conic equation). The answer to the latter question is given by a succint criterion of Smith (1864) involving the equality of Legendre quadratic symbols. When such an integral root exists one can obtain it by solving an appropriate diophantine equation of the type x^2 - $D * y^2 = N$, for integer D and N, as we illustrate below. (Such a diophantine equation has come to be known as Pell's equation, though could also be accredited to Bhaskaracharya, the Indian mathematician of 1150 A.D).

To compute an integral point on a homogeneous conic, $C(X,Y,Z) = aX^2 + bY^2 + cXY + dXZ + eYZ + fZ^2 = 0$, one could find the point at infinity ($Z = 0$), or at finite distances ($Z = 1$). Such an integral point exists at infinity if $c^2 - 4ab$ is a perfect square. Finding integral points at finite distances is equivalent to finding rational points of the dehomogenized conic $C(X,Y,1)$. This corresponds to finding a rational solution (b_1, b_2) of the equation, $ab_1^2 + (cb_2 + d)b_1 + (bb_2^2 + eb_2 + f) = 0$. A solution exists when the discriminant of the equation is a perfect square, (equal to y^2 for integer y). This reduces to finding a rational solution of the equation, $(c^2 - 4ab)b_2^2 + 2(cd - 2ae)b_2 + (d^2 - 4af - y^2) = 0$, where such a rational solution again exists when its discriminant is a perfect square, (equal to x^2 for integer x). Hence we need to solve the equation x^2 - $Dy^2 = N$, for diophantine solutions x and y, with $D = c^2 - 4ab$ and $N = (cd - 2ae)^2 - (c^2 - 4ab)(d^2 - 4af)$. If D is negative or a perfect square there are only a finite number of solutions to this equation. If D is positive, solutions can be obtained by simple continued fractions, Niven and Zuckerman (1972).

To compute an integral point on a homogeneous conicoid, $C(X,Y,Z,W) = aX^2 + bY^2 + cZ^2 + dXY + eXZ + fYZ + gXW + hYW + iZW + jW^2 = 0$, one could again find the point at infinity ($Z = 0$), or at finite distances ($Z = 1$). Finding an integral point at infinity reduces to the above case of finding an integral root of a homogeneous conic $aX^2 + bY^2 + dXY + eXZ + fYZ + cZ^2 = 0$. Finding integral points at finite distances also reduces to the earlier case of solving for a rational point of a conic $ax^2 + by^2 + dxy + gx + hy + j^2 = 0$, or a rational point of the conic $ax^2 + by^2 + dxy + (e+g)x + (f+h)y + (c+i+j)^2 = 0$.

Conclusion Both implicit and parametric representations for curves and surfaces have their inherent advantages. It thus becomes crucial to be able to go efficiently from one form to the other, especially when surfaces of an object are automatically generated in one of the two representations, see Hoffmann and Hopcroft (1985), Bajaj and Kim (1986). Also simpler algorithms are at times possible when both representations are available. For example a straightforward method for computing surface - surface intersections exists when one of the surfaces is in its implicit form and the other in its parametric form.

For surfaces of degree higher than three no rational parametric forms exist in general, although parameterizable subclasses can be identified. For low degree curves and surfaces, in this paper and in Abhyankar and Bajaj (1986a) procedures have been developed and implemented for parameterizing implicit forms. The approach has been extended to parameterize planar curves of higher degree and special space curves, Abhyankar and Bajaj (1986b). These methods can be specialized to work over rational or real fields. Currently efforts are being made to obtain explicit parameterizations of special families of quartic surfaces and surfaces of higher degree which would prove useful for representing blending surfaces.

7.4 Automatic Parametrization of Degree 3 Curves and Surfaces

Rational algebraic curves and surfaces can be represented by implicit or parametric equations. A general algebraic curve of degree three (cubics) is represented implicitly by

$$C(x,y) = ax^3 + by^3 + cx^2y + dxy^2 + ex^2 + fy^2 + gxy + hx + iy + j = 0,$$

All singular cubics are rational, or stated equivalently, they can also be represented by a pair of rational parametric equations $x = u(t)/w(t)$ and $y = v(t)/w(t)$, where u, v and w are no more than cubic polynomials. Non-singular cubics are not rational and the best one can achieve is a parameteric representation with a single square root of rational functions. Next, a general algebraic surface of degree three (cubicoids) has an implicit equation given by

$$C(x,y,z) = ax^3 + by^3 + cz^3 + dx^2y + ex^2z + fxy^2 + gy^2z + hxz^2 + iyz^2 + jxyz \\ + kx^2 + ly^2 + mz^2 + nxy + oxz + pyz + qx + ry + sz + t = 0.$$

All singular cubicoids are rational as are all non-singular cubicoids except cones and cylinders with non-singular cubic generating curves. Rational cubicoids can also be represented by a triad of rational parametric equations $x = u(s,t)/q(s,t)$, $y = v(s,t)/q(s,t)$, and $z = w(s,t)/q(s,t)$ with again u, v, w and q being no more than cubic polynomials. Also, non-rational cubicoids can be represented by a parameteric representation with a single square root of rational functions.

Both implicit and parametric representations for curves and surfaces have their inherent advantages. The rational parametric form of representing a surface allows greater ease for transformation and shape control, Tiller (1983), Mortenson (1985). The implicit form is preferred for testing whether a point is on the surface and is conducive to the direct application of algebraic techniques. It thus becomes crucial to be able to go efficiently from one form to the other, especially when surfaces of an object are automatically generated in one of the two representations, see Hoffmann and Hopcroft (1985), Bajaj and Kim (1987a,b). Also simpler algorithms are at times possible when both representations are available. For example a straightforward method for computing surface - surface intersections exists when one of the surfaces is in its implicit form and the other in its parametric form.

While rational curves are of the same parametric and implicit degree, rational surfaces of parametric degree n in general may have a corresponding implicit degree of n^2, Sederberg, et. al. (1985). The process of conversion from implicit representations to parametric, also known as parameterization, becomes then of added importance for rational surfaces as it yields in general lower degree representations. Computational methods for parameterizations have been given for degree two surfaces by Levin (1976), intersection curves of two degree two surfaces by Ocken, et. al. (1983), degree two curves and surfaces by Abhyankar, Bajaj (1987a), degree three surfaces by Sederberg, Snively (1987) and a general method for rational algebraic curves of arbitrary degree by Abhyankar, Bajaj (1987b).

We describe specific algorithms to obtain rational and special parametric equations for the cubics and cubicoids, given their implicit equations. The algorithm for parameterizing cubics is based on the simple idea of mapping a point on a curve to infinity via a linear transformation and consequently is very efficient. Parameterizing non-singular cubicoids relies on being able to generate rational curves (straight lines, conics or singular cubics) on the cubicoid surface and uses a novel and efficient method of intersecting the surface with tangent planes. The parameterization method for cubicoids of Sederberg, Snively (1987) is restricted to generating straight lines on the cubicoid surface which are computed by intersecting a parametric line with the cubicoid and then simultaneously solving a nonlinear system of four equations in four unknowns.

7.4.1 Cubics

Geometric Viewpoint The idea of parametrizing a conic was to fix a point on the conic and take lines through that point, which intersects the conic in only one additional point, Abhyankar and Bajaj (1987a). The conic was thus rationally parametrized by a pencil of lines with parameter t corresponding to the slope of the lines. A cubic is a curve which intersect most lines in three points. However if we consider a singular cubic then lines through the singular point, (a double point), give a rational parameterization for the curve as again these lines of slope t intersect the cubic in only one additional point. Such is not the case for non-singular cubic curves and they correspond to the cubics which do not have a rational parameterization. Intersecting a curve by a pencil of lines through a point P on the curve can algebraically also be achieved by mapping the point P to infinity along one of the coordinate axis. The pencil of lines then become parallel lines to that coordinate axis all passing through the point P at infinity.

Algebraic Method A plane cubic curve is given by

$$C(x,\ y)\ =\ ax^3\ +\ bx^2y\ +\ cxy^2\ +\ dy^3\ +\ ex^2\ +\ fy^2\ +\ gxy\ +\ hx\ +\ iy\ +\ j$$

Make it nonregular in y (by eliminating the y^3 term through a coordinate transformation). If there exists a real point at infinity then a linear transformation suffices. Recall that, points at infinity are given by the degree form of $C(x, y)$, (terms of highest degree). For conics, we have real points at infinity only in the case of parabolas and hyperbolas. However all cubics have a real point at infinity. The reason being: the degree form always has a real root as it is of degree 3 and complex roots occur in conjugate pairs. The degree form on dehomogenizing ($y\ =\ 1$), gives $f(x)\ =\ ax^3\ +\ bx^2\ +\ cx\ +\ d$ which always has a real root (if $a\ !=\ 0$). When a is zero, the cubic $C(x,\ y)$ is already nonregular in x. Thus to make $C(x, y)$ nonregular in y we may use a linear transformation given by

$$\begin{array}{rcl} x & \mapsto & \alpha\bar{x} + \beta\bar{y} \\ y & \mapsto & \gamma\bar{x} + \delta\bar{y} \end{array} \qquad (69)$$

To make the $\bar{y}^3$ term to be zero, we set its coefficient $(a\beta^3\ +\ b\beta^2\ +\ c\beta\ +\ d)\ =\ 0$ by taking $\delta\ =\ 1$ and β to be the real point at infinity. The remaining parameters $alpha$ and γ are then chosen to ensure $alpha\delta\ -\ \beta\gamma\ !=\ 0$, (for a well formed linear transformation). Additionally the parameters may also be chosen appropriately to obtain suitable desired parameterizations for specific pieces of the cubic.

Now the transformed cubic, in a somewhat rearranged fashion, is given by

$$C(\bar{x},\bar{y})\ =\ (u\bar{x}\ +\ v)\bar{y}^2\ +\ (p\bar{x}^2\ +\ q\bar{x}\ +\ r)\bar{y}\ +\ (k\bar{x}^3\ +\ l\bar{x}^2\ +\ m\bar{x}\ +\ n)$$

which is the usual quadratic equation. Using the old Indian method of Shreedharacharya (5th century), of solving the quadratic equation, ("multiply by 4 times the coefficient of the square term and add the square of the coefficient of the unknown, and the rest follows"), we obtain

$$4(u\bar{x}\ +\ v)^2\bar{y}^2\ +\ 4(u\bar{x}\ +\ v)(p\bar{x}^2\ +\ q\bar{x}\ +\ r)\bar{y}\ +\ 4(u\bar{x}\ +\ v)(k\bar{x}^3\ +\ l\bar{x}^2\ +\ m\bar{x}\ + n)\ =\ 0$$

which on completing the square becomes

$$[2(u\bar{x}\ +\ v)\bar{y}\ +\ (p\bar{x}^2\ +\ q\bar{x}\ +\ s)]^2\ =\ (p\bar{x}^2\ +\ q\bar{x}\ +\ s)^2\ -\ 4(u\bar{x}\ +\ v)(k\bar{x}^3\ +\ l\bar{x}^2\ +\ m\bar{x}\ + n)$$

If we let

$$y^*\ =\ [2(u\bar{x}\ +\ v)\bar{y}\ +\ (p\bar{x}^2\ +\ q\bar{x}\ +\ s)] \qquad (70)$$

then equation (70) becomes of the type

$$y^{*2}\ =\ g(\bar{x}), \qquad deg.\ g(\bar{x})\ \le\ 4 \qquad (71)$$

We only need to analyze (71) and see if we can obtain a parametrization for $\bar{x}$ and y^* for then using transformations (69) and (70) we obtain directly the parameterization for x and y. To do this we consider several cases as follows: $g(\bar{x})\ =\ 0$ has only one distinct root, $g(\bar{x})$ has two distinct roots, ... etc., where both real and imaginary roots of $g(\bar{x})\ =\ 0$ are considered. In the case of multiple roots, we may use the following general method to get rid of them.

Suppose

$$y^{*2}\ =\ [\prod_{i\ =\ 1}^{d}(\bar{x}\ -\mu_i)^2]\Omega(\bar{x}) \qquad d\ =\ 1\ \ or\ \ 2$$

so each root μ_i occurs an even number of times and $\Omega(x)$ has no multiple roots. Then if we let

$$y^{**}\ =\ \left[\frac{y^*}{\prod_{i\ =\ 1}^{d}(\bar{x}\ -\mu_i)}\right] \qquad (72)$$

then equation (71) reduces to

$$y^{**2} \;=\; \Omega(\bar{x}) \tag{73}$$

If $deg.\Omega(\bar{x}) \;<=\; 2$, then the above equation (73) is a conic and a rational parametrization is always possible, Abhyankar and Bajaj (1987a). This then, together with transformations (69), (70) and (72), gives a parameterization for x and y of the original curve. Otherwise, $g(\bar{x})$ has either 3 or 4 distinct roots, and a rational parametrization is not possible. This case arises for the non-singular cubics, also known as elliptic curves or curves of genus 1 which do not have a rational parametrization. Genus 0 is both a necessary and sufficient condition for rational curves, see Abhyankar and Bajaj (1987b). However, by solving the above equation (73), quadratic in y^{**} when $g(\bar{x})$ has 3 or 4 distinct roots, a parameterization for the non-singular cubic is obtained and is of the type that includes a single square root of rational functions. The above parameterizations obtained are global, and of degree at most 3 with the parameter t ranging from $(-\infty, \infty)$ and spanning the entire curve.

7.4.2 Cubicoids

Geometric Viewpoint If we intersect a cubicoid with a plane we get a cubic curve in general. However if we intersect it with a tangent plane (for a point on the cubicoid) then something special happens, namely, we get a singular cubic curve or a reducible curve (either a straight line and a conic, or three straight lines). In general we obtain a singular cubic curve as there are only a finite number of real straight lines on a cubic surface, see Henderson (1911). In either of the three cases, a straight line, a conic or a singular cubic, the intersection curve is always rational and can be parameterized by a single parameter.

To obtain a rational parameterization of the cubicoid we need to generate two rational curves on its surface. Let t and τ correspond to independent parameterizations of the two chosen rational curves. Then the *net* of lines defined by two varying points t and τ (a variable point t on one rational curve and a variable point τ on the other), intersect the cubic surface in one additional point giving a rational parameterization of the cubicoid. For two non-intersecting rational curves on the cubicoid with independent parameterization parameters t and τ, a point (x, y, z) on the rational cubic surface can be seen to correspond to a single pair (t, τ) yielding what is known as a *1-fold* parameterization or a *1-fold* covering of the plane. Higher fold parameterizations are obtained for arbitrary choices of rational curves on the cubicoid.

One algorithm for obtaining two different rational curves on the cubicoid is to repeat the tangent plane intersection method for two different simple (non-singular) points on the cubicoid. Alternatively, one can generate two non-intersecting straight lines from the twenty seven lines on a cubicoid, (first found by Cayley and Salmon in 1849). All possible configurations as well as the number of real and imaginary straight lines on cubicoids have been accurately classified by various authors in the past, see Blythe (1905), Henderson (1911) and Segre (1942). Straight lines on the cubicoid can be computed again by the method of tangent plane intersections. Here points are carefully chosen such that the tangent planes to the surface at these point yield reducible intersections with the cubicoid. Each point on the cubicoid can yield one or three straight lines lying on the same tangent plane. However two specific points are chosen to yield two non-intersecting straight lines of the cubicoid lying on different tangent planes.

Algebraic Method A general degree three surface has an implicit equation given by

$$C(x,y,z) \;=\; ax^3 + by^3 + cz^3 + dx^2y + ex^2z + fxy^2 + gy^2z + hxz^2 + iyz^2 + jxyz$$
$$+ kx^2 + ly^2 + mz^2 + nxy + oxz + pyz + qx + ry + sz + t = 0$$

Take a simple (non-singular) point (x_0, y_0, z_0) on it. Most points on the cubicoid are simple, so this is not a problem. Bring the simple point to the origin by a simple translation $x = x' + x_0$, $y = y' + y_0$ and $z = z' + z_0$.

$$C(x',\ y',\ z') = a'x' + b'y' + c'z' + \ldots + \text{terms of higher degree.}$$

Next rotate the tangent plane to the surface at the origin, given by the order form (terms of lowest degree), to the $z = 0$ plane. This by using a simple rotation, $x' = x$, $y' = y$ and $z' = 1overc'z - a'overc'x - b'overc'y$ which gives

$$C(x,\ y,\ z) = z + [f_2(x,\ y) + f_1(x,\ y)z + f_0 z^2]$$

$$+ [g_3(x,\ y) + g_2(x,\ y)z + g_1(x,\ y)z^2 + g_0 z^3]$$

where $f_i(x, y)$ and $g_i(x, y)$ are appropriate terms of degree i. Its intersection with the tangent plane $z = 0$ is simply,

$$f_2(x,\ y) + g_3(x,\ y) = 0 \tag{74}$$

which is either a reducible curve or a cubic curve with a double point at the origin (lowest degree terms in (74) are ≥ 2). In all cases the curve (74) can be rationally parameterized with a single independent parameter t and rational functions K and L

$$\begin{aligned} x &= K(t) \\ y &= L(t) \\ z &= 0 \end{aligned}$$

which can also be expressed in terms of the original x, y, z coordinates by using the above linear transformations. At this stage the above procedure can be repeated for a second simple point $(x_1,\ y_1,\ z_1)$ on the cubicoid. Alternatively to obtain non-intersecting straight lines on the cubicoid we need to choose the second point carefully. To do this we bring a general point specified by parameter t on this parameterized curve to the origin again by a simple translation

$$\begin{aligned} x &= \bar{x} + K(t) \\ y &= \bar{y} + L(t) \\ z &= \bar{z} \end{aligned} \tag{75}$$

Since this point also lies on the cubic surface $C(x, y, z)$, the surface equation has a zero constant term and is given by

$$C(\bar{x},\ \bar{y},\ \bar{z}) = \bar{a}(t)\bar{x} + \bar{b}(t)\bar{y} + \bar{c}(t)\bar{z} + \ldots + \text{terms of higher degree.}$$

Next a simple rotation

$$\begin{aligned} \bar{x} &= \hat{x} \\ \bar{y} &= \hat{y} \\ \bar{z} &= \frac{1}{\bar{c}(t)}\hat{z} - \frac{\bar{a}(t)}{\bar{c}(t)}\hat{x} - \frac{\bar{b}(t)}{\bar{c}(t)}\hat{y} \end{aligned} \tag{76}$$

makes the tangent plane to the surface at the origin to be the $\hat{z} = 0$ plane, resulting again in

$$\begin{aligned} C(\hat{x}, \hat{y}, \hat{z}) &= \hat{z} + [\hat{f}_2(\hat{x},\ \hat{y}) + \hat{f}_1(\hat{x},\ \hat{y})\hat{z} + \hat{f}_0 \hat{z}^2] \\ &+ \ [\hat{g}_3(\hat{x},\ \hat{y}) + \hat{g}_2(\hat{x},\ \hat{y})\hat{z} + \hat{g}_1(\hat{x},\ \hat{y})\hat{z}^2 \\ &+ \ \hat{g}_0(\hat{x},\ \hat{y})\hat{z}^3] \end{aligned}$$

Its intersection with $\hat{z} = 0$ plane will give

$$\hat{f}_2(\hat{x},\ \hat{y}) + \hat{g}_3(\hat{x},\ \hat{y}) = 0 \tag{77}$$

which is a plane curve with coefficients involving t. For certain values of t, the plane curve is reducible, which then gives the lines on the cubic surface. Specifically (77) is reducible for those values of t for which the two polynomials $\hat{f}_2(\hat{x},\ \hat{y})$ and $\hat{g}_3(\hat{x},\ \hat{y})$ have a linear or quadratic common factor. One way of obtaining these t values is as follows. Consider $\hat{f}_2(\hat{x},\ \hat{y}) = 0$ the homogeneous equation of degree 2 with coefficients involving t. It has two linear factors $\hat{y} = m_1(t)\hat{x}$ and $\hat{y} = m_2(t)\hat{x}$. Substituting either of these into the homogeneous equation $g_3(\hat{x},\ \hat{y}) = 0$ yields a cubic equation of the form $p(t)\hat{x}^3 = 0$ where $p(t)$ is a function of t. Specific solutions t of the equation $p(t) = 0$ can easily be obtained by using known methods for obtaining roots of univariate polynomial equations, by either numerical methods, see Jenkins and Traub (1972), or by symbolic methods, see Collins and Loos (1982). With (75) and (76) and for two appropriate choices of t one obtains the equations of two distinct lines on the cubicoid.

Having obtained two rational curves on the cubicoid, say parameterized respectively by $x_1 = f_1(t)$, $y_1 = f_2(t)$, $z_1 = f_3(t)$ and $x_2 = g_1(\tau)$, $y_2 = g_2(\tau)$, $z_2 = g_3(\tau)$, consider next the *net* of straight lines passing through a point on each of the two curves. This in space is given by two equations

$$\frac{z - z_1}{x - x_1} = \frac{z_2 - z_1}{x_2 - x_1} \tag{78}$$

$$\frac{y - y_1}{x - x_1} = \frac{y_2 - y_1}{x_2 - x_1} \tag{79}$$

and defines a two parameter family (a *net*) of straight lines for varying t and τ. Substituting for y and z in terms of x in the equation of the cubic surface $C(x, y, z) = 0$ yields a cubic equation in x with coefficients in t and τ, viz., $G(x, t, \tau) = 0$. However both $x = x_1$ and $x = x_2$ satisfy this equation and thus

$$\frac{G(x, t, \tau)}{(x - f_1(t))(x - g_1(\tau))} = 0$$

is linear in x, yielding x as a rational function of t and τ. Together with (78) and (79) this yields a rational parameterization of the cubic surface in terms of the independent parameters t and τ. The rational parameterization obtained are global and of degrees at most 3, with parameters t and τ both ranging from $(-\infty, \infty)$ and spanning the entire cubicoid (except the two rational curves on the surface).

It suffices to mention that non-rational cubicoids can be parameterized by single square roots of rational functions in a way exactly similar to non-rational cubics. By mapping any point on the cubicoid to infinity along the Z axis by a real linear transformation, the cubicoid equation can be made non-regular in z (equation with no z^3 term). What remains then is the transformed cubicoid equation which is quadratic in z and from which z can be easily expressed in terms of square roots of rational functions of x and y using the quadratic equation formula.

Various computational issues in extending the above algorithmic methods approach to parameterize planar algebraic curves of arbitrary degree are discussed in Abhyankar and Bajaj (1987b). For surfaces of degree higher than three no rational parametric forms exist in general, although parameterizable subclasses can be identified. For example degree four surfaces with a triple point such as the Steiner surfaces or degree four surfaces with a double curve such as the such as the Plucker surfaces are rational, Jessop (1916). Currently efforts are being made to obtain explicit parameterizations of special families of quartic and higher degree surfaces which prove useful for representing blending surfaces.

7.5 Parameterizations of Real Cubic Surfaces

Low degree real algebraic surfaces (quadrics, cubics and quartics) play a significant role in constructing accurate computer models of physical objects and environments for purposes of simulation and prototyping[10]. While quadrics such as spheres, cones, hyperboloids and paraboloids prove sufficient for constructing restricted classes of models, cubic algebraic surface patches are sufficient to model the boundary of objects with arbitrary topology in a C^1 piecewise smooth manner [11].

Real cubic algebraic surfaces are the real zeros of a *polynomial* equation $f(x, y, z) = 0$ of degree three. In this representation the cubic surface is said to be in *implicit* form. The irreducible cubic surface which is not a cylinder of a nonsingular cubic curve, can alternatively be described explicitly by rational functions of parameters u and v:

$$x = \frac{f_1(u,v)}{f_4(u,v)},\ y = \frac{f_2(u,v)}{f_4(u,v)},\ z = \frac{f_3(u,v)}{f_4(u,v)}, \tag{80}$$

where f_i, $i = 1 \ldots 4$ are polynomials. In this case the cubic surface is said to be in *rational parametric* form.

Real cubic algebraic surfaces thus possess dual implicit-parametric representations and this property proves important for the efficiency of a number of geometric modeling and computer graphics display operations [10, 27]. For example, with dual available representations the intersection of two surfaces or surface patches reduces simply to the sampling of an algebraic curve in the planar parameter domain [8]. Similarly, point-surface patch incidence classification, a prerequisite for boolean set operations and ray casting for graphics display, is greatly simplified in the case when both the implicit and parametric representations are available [8]. Additional examples in the computer graphics domain which benefit from dual implicit-parametric representations are the rapid triangulation for curved surface display and image texture mapping on curved surface patches.

Deriving the rational parametric form from the implicit representation of algebraic surfaces, is a process known as rational parameterization. Algorithms for the rational parameterization of cubic algebraic surfaces have been given in [4, 37], based on the classical theory of skew straight lines and rational curves on the cubic surface [15, 26]. One of the main results of our current paper is to constructively address the parameterization of cubic surfaces based on the reality of the straight lines on the real cubic surface. In doing so we provide an algorithm to construct all twenty-seven straight lines (real and complex) on the real nonsingular cubic surface. We prove that the parameterizations of the real cubic surface components are constructed using a pair of real skew lines for those three families which have them, and remarkably using a complex conjugate pair of skew lines, in a fourth family. There does not appear to be a similar rational parameterization for the fifth family that covers all or almost all of the surface, so instead we use two disjoint parameterizations which involve one square root each. A rational parameterization that covers part of the surface is described in [37]. In that scheme points which lie on tangent planes through points on a real line are covered, but these points do not necessarily comprise most of the surface, and the covering is in general two-to-one instead of one-to-one. All of the parameterizations described are one-to-one, meaning that for any point on the cubic surface there can be just one set of values (u, v) which give rise to that point.

We also analyze the image of the derived rational parameterization for both real and complex parameter values, together with "base" points where the parameterizations are ill-defined. These base points cause a finite number (at most five) of lines and points, and possibly two conic sections lying on the surface, to be missed by the parameterizations. One of these conics can be attained by letting $u \to \pm\infty$ and the other with $v \to \pm\infty$ separately, or by using projective coordinates $\{u, u^*\}$ and $\{v, v^*\}$ instead of (u, v) and setting $v = 0$ and $u = 0$, respectively.

One of the gems of classical algebraic geometry has been the theorem that twenty-seven distinct straight lines lie completely on a nonsingular cubic surface [33]. Schläfi's double-six notation elegantly captures the complicated and many-fold symmetry of the configurations of the twenty-seven lines. He also partitions all nonsingular cubic surfaces $f(x, y, z) = 0$ into five families $F_1, \ldots, F_5$ based on the reality of the twenty-seven lines. Family F_1 contains 27 real straight lines, family F_2

contains 15 real lines, and family F_3 contains 7 real lines while families F_4 and F_5 contain 3 real lines each. What distinguishes F_4 from F_5 is that while 6 of the 12 conjugate complex line pairs of F_4 are skew (and 6 pairs are coplanar), each of the 12 conjugate pairs of complex line pairs of F_5 is coplanar. When a nonsingular cubic surface F tends to a singular cubic surface G (with an isolated double point or a double line) 12 of F's straight lines (constituting a double six) tend to 6 lines through a double point of G. Hence singular cubic surfaces have only twenty-one distinct straight lines.

Alternatively a classification of cubic surfaces can be obtained from computing all 'base' points of its parametric representation,

$$x = \frac{f_1(u,v)}{f_4(u,v)},\ y = \frac{f_2(u,v)}{f_4(u,v)},\ z = \frac{f_3(u,v)}{f_4(u,v)},$$

Base points of a surface parameterization are those isolated parameter values which simultaneously satisfy $f_1 = f_2 = f_3 = f_4 = 0$. It is known that any nonsingular cubic surface can be expressed as a rational parametric cubic with six base points. The classification of nonsingular real cubic surfaces then follows from:

1. If all six base points are real, then all 27 lines are real, i.e. the F_1 case.
2. If two of the base points are a complex conjugate pair then 15 of the straight lines are real, i.e. the F_2 case.
3. If four of the base points are two complex conjugate pairs then 7 of the straight lines are real, i.e. the F_3 case.
4. If all base points are complex then three of the straight lines are real. In this case the three real lines are all coplanar, i.e. the F_4 and F_5 cases.

7.5.1 Real and Rational Points on Cubic Surfaces

We first begin by computing a simple real point (with a predefined bit precision) on a given real cubic surface $f(x,y,z) = 0$. For obvious reasons of exact calculations with bounded precision it is very desirable to choose the simple point to have rational coordinates. Mordell in his 1969 book [29] mentions that no method is known for determining whether rational points exist on a general cubic surface $f(x,y,z) = 0$, or finding all of them if any exist. We are unaware if a general criterion or method now exists or whether Mordell's conjecture below has been resolved.

The following theorems and conjecture exhibit the difficulty of this problem, and are repeated here for information.
Theorem[[29],chap 11]: All rational points on a cubic surface can be found if it contains two lines whose equations are defined by conjugate numbers of a quadratic field and in particular by rational numbers.
Theorem[[29],chap 11]: The general cubic equation (irreducible cubic and not a function of two independent variables nor a homogeneous polynomial in linear functions of its variables) has either none or an infinity of rational solutions.
Mordell Conjecture[[29],chap 11]: The cubic equation $F(X,Y,Z,W) = 0$ is solvable if and only if the congruence $F(X,Y,Z,W) \equiv 0 \pmod{p^r}$ is solvable for all primes p and integers $r > 0$ with $(X,Y,Z,W,p) = 1$.

We present a straightforward search procedure to determine a real point on $f(x,y,z) = 0$, and if lucky one with rational coordinates.

Collect the highest degree terms of $f(x,y,z)$ and call this homogeneous form $F_3(x,y,z)$. Recursively determine if $F_3(x,y,z) = 0$ has a rational point. Being homogeneous, one only needs to check for $F_3(x,y,1) = 0$ and $F_3(x,y,0) = 0$, which are both polynomials in one less variable, and hence the recursion is in dimension. Now for a univariate polynomial equation $g(x) = 0$ we use the technique of [28] to determine the existence and coordinates of a rational root. If not, one computes a real root having the desired bit precision as explained below.

Additionally, if the highest degree terms of $f(x, y, z)$ do not yield a rational point, we compute the resultant and linear subresultants of f and f_x, eliminating x to yield new polynomials $f_1(y, z)$ and $xf_2(y, z) + f_3(x, y, z)$ (see [9] for details of this computation). Recursively compute the rational points of $f_1(y, z) = 0$, using the equation $xf_2(y, z) + f_3(x, y, z) = 0$ to determine the rational x coordinate given rational y and z coordinates of the point.

In the general case, therefore, we are forced to take a real simple point on the cubic surface. We can bound the required precision of this real simple point so that the translations and resultant computations in the straight line extraction and cubic surface parameterization algorithm of the next section, are performed correctly. The lower bound of this value can be estimated as in [20] by use of the following gap theorem:

Gap Theorem ([20],p70). *Let $\mathcal{P}(d, c)$ be the class of integral polynomials of degree d and maximum coefficient magnitude c. Let $f_i(x_1, \cdots, x_n) \in \mathcal{P}(d, c)$, $i = 1, \cdots, n$ be a collection of n polynomials in n variables which has only finitely many solutions when homogenized. If $(\alpha_1, \cdots, \alpha_n)$ is a solution of the system, then for any j either $\alpha_j = 0$, or $|\alpha_j| > (3dc)^{-nd^n}$.*

7.5.2 Algebraic Reduction

Given two skew lines $\mathbf{l}_1(u) = \begin{bmatrix} x_1(u) \\ y_1(u) \\ z_1(u) \end{bmatrix}$ and $\mathbf{l}_2(v) = \begin{bmatrix} x_2(v) \\ y_2(v) \\ z_2(v) \end{bmatrix}$ on the cubic surface $f(x, y, z) = 0$, the cubic parameterization formula for a point $\mathbf{p}(u, v)$ on the surface is :

$$\mathbf{p}(u, v) = \begin{bmatrix} x(u, v) \\ y(u, v) \\ z(u, v) \end{bmatrix} = \frac{a\mathbf{l}_1 + b\mathbf{l}_2}{a + b} = \frac{a(u, v)\mathbf{l}_1(u) + b(u, v)\mathbf{l}_2(v)}{a(u, v) + b(u, v)} \tag{81}$$

where

$$\begin{aligned} a = a(u, v) = \nabla f(\mathbf{l}_2(v)) \cdot [\mathbf{l}_1(u) - \mathbf{l}_2(v)] \\ b = b(u, v) = \nabla f(\mathbf{l}_1(u)) \cdot [\mathbf{l}_1(u) - \mathbf{l}_2(v)] \end{aligned}$$

The total degree of the numerator of the parameterization formula in $\{u, v\}$ is 4 while the denominator total degree is 3. Note that if the lines are coplanar, formula (81) can only produce points on the plane of the lines, hence the search for skew lines on the cubic surface.

Following the notation of [4], a real cubic surface has an implicit representation of the form

$$\begin{aligned} f(x, y, z) = Ax^3 + By^3 + Cz^3 + Dx^2y + Ex^2z + Fxy^2 + Gy^2z + Hxz^2 + Iyz^2 + Jxyz \\ + Kx^2 + Ly^2 + Mz^2 + Nxy + Oxz + Pyz + Qx + Ry + Sz + T = 0 \, . \end{aligned}$$

Compute a simple (nonsingular) point (x_0, y_0, z_0) on the surface. We can move the simple point to the origin by a translation $x = x' + x_0$, $y = y' + y_0$, $z = z' + z_0$, producing

$$f'(x', y', z') = Q'x' + R'y' + S'z' + \ldots \text{ terms of higher degree.}$$

Next, we wish to rotate the tangent plane to $f(x', y', z')$ at the origin to the plane $z'' = 0$. This can be done by the transformation

$$\begin{array}{ll} x' = x'',\ y' = y'',\ z' = (z'' - Q'x'' - R'y'')/S' & \text{if } S' \neq 0 \\ x' = x'',\ y' = (z'' - Q'x'')/R',\ z' = y'' & \text{if } S' = 0 \text{ and } R \neq 0 \\ x' = z''/Q',\ y' = x'',\ z' = y'' & \text{if } S' = 0,\ R' = 0, \text{ and } Q' \neq 0 \ . \end{array}$$

Fortunately Q', R', and S' cannot all be zero, because then the selected point (x_0, y_0, z_0) would be a singular point on the cubic surface.

The transformed surface can be put in the form

$$f''(x'',y'',z'') = z''+[f_2(x'',y'')+f_1(x'',y'')z''+f_0z''^2]+[g_3(x'',y'')+g_2(x'',y'')z''+g_1(x'',y'')z''^2+g_0z''^3] ,$$

where $f_j(x'', y'')$ and $g_j(x'', y'')$ are terms of degree j in x'' and y''. In general, this surface intersects the tangent plane $z'' = 0$ in a cubic curve with a double point at the origin (as its lowest degree terms are quadratic). This curve can be rationally parametrized as

$$\begin{aligned} x'' &= K(t) = -\frac{L''t^2 + N''t + K''}{B''t^3 + F''t^2 + D''t + A''} \\ y'' &= L(t) = tK(t) = -\frac{L''t^3 + N''t^2 + K''t}{B''t^3 + F''t^2 + D''t + A''} \\ z'' &= 0 , \end{aligned} \tag{82}$$

where $A'', B'', \ldots$ are the coefficients in $f''(x'', y'', z'')$ that are analogous to $A, B, \ldots$ in $f(x, y, z)$. In the special case that the singular cubic curve is reducible (a conic and a line or three lines), a parameterization of the conic is taken instead.

We transform the surface again to bring a general point on the parametric curve specified by t to the origin by the translation

$$\begin{aligned} x'' &= \overline{x} + K(t) \\ y'' &= \overline{y} + L(t) \\ z'' &= \overline{z} . \end{aligned}$$

The cubic surface can now be expressed by

$$\bar{f}(\overline{x}, \overline{y}, \overline{z}) = \overline{Q}(t)\overline{x} + \overline{R}(t)\overline{y} + \overline{S}(t)\overline{z} + \cdots \text{ terms of higher degree} .$$

We make the tangent plane of the surface at the origin coincide with the plane $\hat{z} = 0$ by applying the transformation

$$\begin{aligned} \overline{x} &= \hat{x} \\ \overline{y} &= \hat{y} \\ \overline{z} &= -\frac{\overline{Q}(t)}{\overline{S}(t)}\hat{x} - \frac{\overline{R}(t)}{\overline{S}(t)}\hat{y} + \frac{1}{\overline{S}(t)}\hat{z} . \end{aligned}$$

The equation of the surface now has the form

$$f(\hat{x}, \hat{y}, \hat{z}) = \hat{z} + [\hat{f}_2(\hat{x}, \hat{y}) + \hat{f}_1(\hat{x}, \hat{y})\hat{z} + \hat{f}_0\hat{z}^2] + [\hat{g}_3(\hat{x}, \hat{y}) + \hat{g}_2(\hat{x}, \hat{y})\hat{z} + \hat{g}_1(\hat{x}, \hat{y})\hat{z}^2 + \hat{g}_0\hat{z}^3] .$$

The intersection of this surface with $\hat{z} = 0$ gives

$$\hat{f}_2(\hat{x}, \hat{y}) + \hat{g}_3(\hat{x}, \hat{y}) = 0 . \tag{83}$$

Recall that $\hat{x}$ and $\hat{y}$, and hence $\hat{f}_2$ and $\hat{g}_3$, are functions of t. As shown in [4], equation (83) is reducible, and hence contains a linear factor, for those values of t for which $\hat{f}_2(\hat{x}, \hat{y})$ and $\hat{g}_3(\hat{x}, \hat{y})$ have a linear or quadratic factor in common. These factors correspond to lines on the cubic surface, and our goal is to find the values of t which produce these lines.

The values of t may be obtained by taking the resultant of $\hat{f}_2(\hat{x}, \hat{y}, t)$ and $\hat{g}_3(\hat{x}, \hat{y}, t)$ by eliminating either $\hat{x}$ or $\hat{y}$. Since $\hat{f}_2$ and $\hat{g}_3$ are homogeneous in $\{\hat{x}, \hat{y}\}$ it does not matter with respect to which variable the resultant is taken[40]; the result will have the other variable raised to the sixth power as a factor. Apart from the factor of $\hat{x}^6$ or $\hat{y}^6$, the resultant consists of an 81st degree polynomial $P_{81}(t)$ in t. At first glance it would appear that there could be 81 values of t for which a line on the cubic surface is produced, but this is not the case:

Theorem 1: The polynomial $P_{81}(t)$ obtained by taking the resultant of $\hat{f}_2$ and $\hat{g}_3$ factors as $P_{81}(t) = P_{27}(t)[P_3(t)]^6[P_6(t)]^6$, where $P_3(t) = B''t^3 + F''t^2 + D''t + A''$, the denominator of $K(t)$ and $L(t)$, and $P_6(t)$ is the numerator of $\overline{S}(t) \quad (P_6(t) = \overline{S}(t)[P_3(t)^2])$.

Sketch of proof: This proof was performed through the use of the symbolic manipulation program Maple. When expanded out in full, $P_{81}(t)$ contains hundreds of thousands of terms, so a direct approach was not possible. Instead, $P_{81}(t)$ was shown to be divisible by both $[P_3(t)]^6$ and $[P_6(t)]^6$.

When $\hat{f}_2$ and $\hat{g}_3$ were expressed in terms of the numerators of $\overline{Q}(t)$, $\overline{R}(t)$, and $\overline{S}(t)$, it was possible to take the resultant without overflowing the memory capabilities of the machine. The resultant could be factored, and $[P_6(t)]^6$ was found to be one of the factors.

The factor $[P_3(t)]^6$ proved to be more difficult to obtain. After the factor $[P_6(t)]^6$ was removed, the remaining factor was split into several pieces, according to which powers of $\overline{Q}(t)$, $\overline{R}(t)$, and $\overline{S}(t)$ they contained. These pieces were each divided by $[P_3(t)]^6$, and the remainders taken. The remainders were expressed as certain polynomials times various powers of $P_3(t)$, as in $a_0(t) + a_1(t)P_3(t) + a_2(t)[P_3(t)]^2 + a_3(t)[P_3(t)]^3 + a_4(t)[P_3(t)]^4 + a_5(t)[P_3(t)]^5$. We were able to show that $a_0(t)$ is in fact divisible by $P_3(t)$. Then we could show that $a_0(t)/P_3(t) + a_1(t)$ is also divisible by $P_3(t)$, and so on up the line until we could show the whole remaining factor is divisible by $[P_3(t)]^6$.

The solutions of $P_{27}(t) = 0$ correspond to the 27 lines on the cubic surface. A method of partial classification is suggested by considering the number of real roots of $P_{27}(t)$: if it has 27, 15, or 7 real roots the cubic surface is F_1, F_2, or F_3, respectively, and if $P_{27}(t) = 0$ has three real roots the surface can be either F_4 or F_5. However, this is not quite accurate. In exceptional cases, $P_{27}(t)$ may have a double root at $t = t_0$, which corresponds to $\hat{f}_2$ and $\hat{g}_3$ sharing a quadratic factor. If this quadratic factor is reducible over the reals, the double root corresponds to two (coplanar) real lines; if the quadratic factor has no real roots it corresponds to two coplanar complex conjugate lines.

Theorem 2: *Simple real roots of $P_{27}(t) = 0$ correspond to real lines on the surface.*

Proof: Let t_0 be a simple real root of $P_{27}(t) = 0$. Since $P_{27}(t)$ is a factor of the resultant of $\hat{f}_2$ and $\hat{g}_3$ obtained by eliminating $\hat{x}$ or $\hat{y}$, $\hat{f}_2(\hat{x}, \hat{y}, t_0)$ and $\hat{g}_3(\hat{x}, \hat{y}, t_0)$ must have a linear or quadratic factor in common. If $\hat{f}_2(\hat{x}, \hat{y}, t_0)$ and $\hat{g}_3(\hat{x}, \hat{y}, t_0)$ have just a linear factor in common, then that factor is of the form $c_1\hat{x} + c_2\hat{y}$ where c_1 and c_2 are real constants since all the coefficients of $\hat{f}_2(\hat{x}, \hat{y}, t_0)$ and $\hat{g}_3(\hat{x}, \hat{y}, t_0)$ are real and $\hat{f}_2(\hat{x}, \hat{y}, t_0)$ and $\hat{g}_3(\hat{x}, \hat{y}, t_0)$ are homogeneous in $\hat{x}$ and $\hat{y}$. In this case the real line $c_1\hat{x} + c_2\hat{y} = 0$ lies on the surface.

If $\hat{f}_2(\hat{x}, \hat{y}, t_0)$ and $\hat{g}_3(\hat{x}, \hat{y}, t_0)$ have a quadratic factor in common, then that factor is of the form $c_1\hat{x}^2 + c_2\hat{x}\hat{y} + c_3\hat{y}^2$. We will show that if this is the case, then $P_{27}(t)$ has at least a double root at $t = t_0$. This will be sufficient to prove that simple roots of $P_{27}(t)$ can only correspond to common linear factors of $\hat{f}_2(\hat{x}, \hat{y}, t_0)$ and $\hat{g}_3(\hat{x}, \hat{y}, t_0)$, and hence real lines on the cubic surface.

If we write $\hat{f}_2(\hat{x}, \hat{y}, t) = Q_1(t)\hat{x}^2 + Q_2(t)\hat{x}\hat{y} + Q_3(t)\hat{y}^2$ and $\hat{g}_3(\hat{x}, \hat{y}, t) = Q_4(t)\hat{x}^3 + Q_5(t)\hat{x}^2\hat{y} + Q_6(t)\hat{x}\hat{y}^2 + Q_7(t)\hat{y}^3$, then the resultant of $\hat{f}_2(\hat{x}, \hat{y}, t)$ and $\hat{g}_3(\hat{x}, \hat{y}, t)$ obtained by eliminating $\hat{x}$ is

$$R(\hat{f}_2, \hat{g}_3) = \begin{vmatrix} Q_1(t) & Q_2(t) & Q_3(t) & 0 & 0 \\ 0 & Q_1(t) & Q_2(t) & Q_3(t) & 0 \\ 0 & 0 & Q_1(t) & Q_2(t) & Q_3(t) \\ Q_4(t) & Q_5(t) & Q_6(t) & Q_7(t) & 0 \\ 0 & Q_4(t) & Q_5(t) & Q_6(t) & Q_7(t) \end{vmatrix} \hat{y}^6 \ . \tag{84}$$

We need to show that if $\hat{f}_2(\hat{x}, \hat{y}, t)$ and $\hat{g}_3(\hat{x}, \hat{y}, t)$ have a quadratic factor in common when $t = t_0$, then $R(\hat{f}_2, \hat{g}_3)/\hat{y}^6$ has a double root at $t = t_0$. This is equivalent to showing that

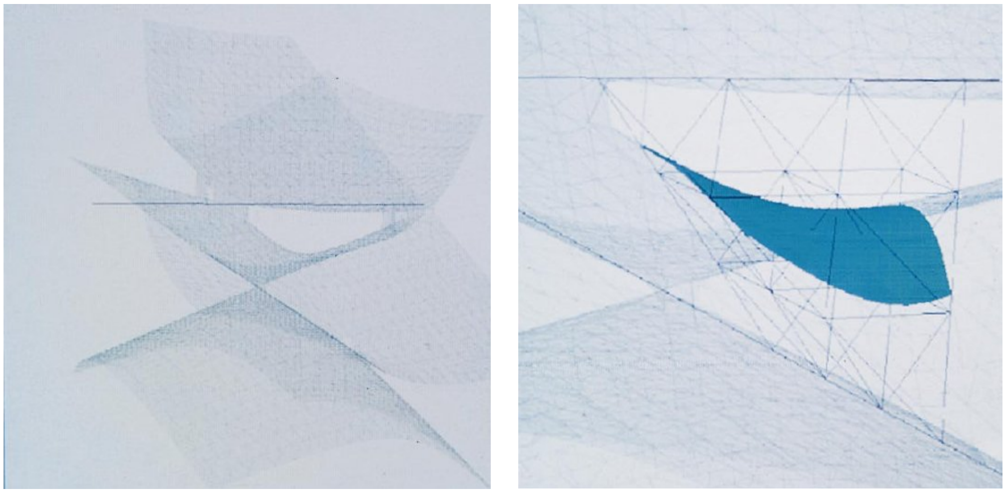


Figure 4: An F_2 cubic surface with two skew lines out of its 15 real straight lines (left), and a zoom-in on that surface showing the Bézier patch with its bounding tetrahedron, determined by the two skew lines as opposite edges (right).

$R(\hat{f}_2(t_0), \hat{g}_3(t_0)) = 0$ and $(d/dt)[R(\hat{f}_2(t_0), \hat{g}_3(t_0))] = 0$. If $\hat{f}_2(\hat{x}, \hat{y}, t_0)$ and $\hat{g}_3(\hat{x}, \hat{y}, t_0)$ have a quadratic factor in common, then $\hat{g}_3(t_0) = k(c_1\hat{x} - c_2\hat{y})\hat{f}_2(t_0)$ for some real constants k, c_1, and c_2. Thus $Q_4(t_0) = kc_1Q_1(t_0)$, $Q_5(t_0) = k[c_1Q_2(t_0) - c_2Q_1(t_0)]$, $Q_6(t_0) = k[c_1Q_3(t_0) - c_2Q_2(t_0)]$, and $Q_7(t_0) = -kc_2Q_3(t)$. Making these substitutions in (84), we find that indeed both $R(\hat{f}_2(t_0), \hat{g}_3(t_0)) = 0$ and $(d/dt)[R(\hat{f}_2(t_0), \hat{g}_3(t_0))] = 0$. ■

To summarize, the simple real roots of $P_{27}(t) = 0$ correspond to real lines on the cubic surface. Double real roots may correspond to either real or complex lines, depending on whether the quadratic factor $\hat{f}_2(\hat{x}, \hat{y}, t)$ and $\hat{g}_3(\hat{x}, \hat{y}, t)$ have in common is reducible or not over the reals. Higher order roots indicate some type of singularity. Complex roots can only correspond to complex lines in nonsingular cases. If t_0, a complex root of $P_{27}(t) = 0$, corresponded to a real line $c_1\hat{x} - c_2\hat{y}$ on the surface, then $\overline{t_0}$ would correspond to the same line, as a real line is its own complex conjugate. Thus one real line would be leading to two distinct values for t_0.

When the cubic surface is of class F_1, F_2, or F_3, it contains at least two real skew lines, and the parameterization in [4] is used. Having obtained skew lines $\mathbf{l}_1(u) = [x_1(u) \;\; y_1(u) \;\; z_1(u)]$ and $\mathbf{l}_2(v) = [x_1(v) \;\; y_1(v) \;\; z_1(v)]$, we consider the net of lines passing through a point on each. This is given by

$$\frac{z - z_1}{x - x_1} = \frac{z_2 - z_1}{x_2 - x_1} \qquad \frac{y - y_1}{x - x_1} = \frac{y_2 - y_1}{x_2 - x_1}\ .$$

Solving these for y and z in terms of x, and substituting into the cubic surface $f(x, y, z) = 0$ gives a cubic equation in x with coefficients in u and v, say $G(x, u, v) = 0$. Since $x = x_1$ and $x = x_2$ satisfy this equation, $G(x, u, v)$ is divisible by $x - x_1$ and $x - x_2$, and we have that

$$H(u, v, x) = \frac{G(x, u, v)}{[x - x_1(u)][x - x_2(v)]} \tag{85}$$

is a linear polynomial in x. This is solved for x as a rational function of u and v. Rational functions for y and z are obtained analogously.

The parameterization (80) is then computed as in (81):

$$(x,\ y,\ z) = (x(u, v),\ y(u, v),\ z(u, v)) = (f_1(u, v)/f_4(u, v),\ \ f_2(u, v)/f_4(u, v),\ f_3(u, v)/f_4(u, v))$$

where

$$
\begin{aligned}
f_1(u,v) &= a(u,v)x_1(u) + b(u,v)x_2(v) \\
f_2(u,v) &= a(u,v)y_1(u) + b(u,v)y_2(v) \\
f_3(u,v) &= a(u,v)z_1(u) + b(u,v)z_2(v) \\
f_4(u,v) &= a(u,v) + b(u,v) \ ,
\end{aligned}
\tag{86}
$$

with

$$
a(u,v) = \nabla f(\mathbf{l}_2(v)) \cdot [\mathbf{l}_1(u) - \mathbf{l}_2(v)] \ , \qquad b(u,v) = \nabla f(\mathbf{l}_1(u)) \cdot [\mathbf{l}_1(u) - \mathbf{l}_2(v)]
$$

In this notation $-f_1(u,v)$ and $f_4(u,v)$ are the coefficients of x^0 and x^1, respectively, in $H(u,v,x)$. The symbolic manipulation program Maple was used to verify that the expressions $f_1(u,v)/f_4(u,v)$, $f_2(u,v)/f_4(u,v)$, and $f_3(u,v)/f_4(u,v)$ do simplify to x, y, and z respectively.

Using floating-point arithmetic, it may be the case that some terms with very small coefficients appear in $f_1(u,v)$, $f_2(u,v)$, $f_3(u,v)$, and $f_4(u,v)$ when the coefficients should in fact be zero. Specifically, these are the terms containing u^3, v^3, u^4, v^4, u^3v and uv^3 in f_1, f_2, and f_3, and terms containing u^3 and v^3 in f_4. These coefficients were shown to be zero using Maple, so in the algorithm they are subtracted off in case they appear in the construction of f_1, f_2, f_3, and f_4.

7.5.3 Parameterizations without Real Skew Lines

When the cubic surface is of class F_4 or F_5 it does not contain any pair of real skew lines. In the F_4 case we derive a parameterization using complex conjugate skew lines, and in the F_5 case we obtain a parameterization by parametrizing conic sections which are the further intersections of the cubic surface with planes through a real line on the surface.

The F_4 Case In this case there are 12 pairs of complex conjugate lines. For 6 of these pairs, the two lines intersect (at a real point). In the other 6 pairs, the two lines are skew. Let one pair of complex conjugate skew lines be given by $(x_1(u+vi),\ y_1(u+vi),\ z_1(u+vi))$ and $(x_1(u-vi),\ y_1(u-vi),\ z_1(u-vi))$. Here x_1, y_1, and z_1 are (linear) complex functions of a complex variable, and x_2, y_2, z_2 may be considered to be the complex conjugates of x_1, y_1, z_1. Also the real parameters u and v are unrestricted. Then the parameterization is again given by (86). Even though the quantities x_i, y_i, and z_i are complex, the expressions for $x(u,v)$, $y(u,v)$, and $z(u,v)$ turn out to be real when x_2, y_2, and z_2 are the complex conjugates of x_1, y_1, and z_1. The symbolic manipulation program Maple was used to verify that the quantities $f_1(u,v)/i$, $f_2(u,v)/i$, $f_3(u,v)/i$ and $f_4(u,v)/i$ are all real when (x_1,y_1,z_1) and (x_2,y_2,z_2) are complex conjugates.

Using floating-point arithmetic, it may be the case that some terms with very small coefficients appear in $f_1(u,v)$, $f_2(u,v)$, and $f_3(u,v)$ when the coefficients should in fact be zero. Specifically, these are the terms containing u^3v and uv^3. These coefficients were shown to be zero using Maple, so in the algorithm they are subtracted off in case they appear in the construction of f_1, f_2, and f_3.

Theorem 3: The algorithm provides a valid parameterization of an F_4 cubic surface when u and v are related as follows: u is unrestricted (both real and imaginary parts), and v is the complex conjugate of u. Each real point on the F_4 surface, except for those corresponding to base points of the parameterization, is obtained for exactly one complex value of u.

Lemma: Given two skew complex conjugate lines $\mathbf{l}_1(u) = (A+Bi,\ C+Di,\ E+Fi) + (G+Hi,\ I+Ji,\ K+Li)u$ and $\overline{\mathbf{l}_1}(v) = (A-Bi,\ C-Di,\ E-Fi) + (G-Hi,\ I-Ji,\ K-Li)v$, then for an arbitrary real point $\mathbf{p} = (x,y,z)$, there exists a unique complex value u_0 such that the points $\mathbf{p}$, $\mathbf{l}_1(u_0)$, and $\overline{\mathbf{l}_1(u_0)}$ are collinear.

Proof of Lemma: The points $\mathbf{p}$, $\mathbf{l}_1(u_0)$, and $\overline{\mathbf{l}_1(u_0)}$ will be collinear if and only if the vectors $\mathbf{p} - \mathbf{l}_1(u_0)$ and $\mathbf{p} - \overline{\mathbf{l}_1(u_0)}$ are parallel. Setting the cross product of these two vectors equal to zero and splitting u_0 into real and imaginary parts as $a_0 + b_0 i$, we find that there is a solution when

$$a_0 = \frac{M_1 M_3 + M_2 M_4}{M_1^2 + M_2^2} \qquad \text{and} \qquad b_0 = \frac{M_2 M_3 - M_1 M_4}{M_1^2 + M_2^2} \quad , \tag{87}$$

and

$$M_1 = \begin{vmatrix} B & G & H \\ D & I & J \\ F & K & L \end{vmatrix} \qquad M_2 = \begin{vmatrix} x-A & G & H \\ y-C & I & J \\ z-E & K & L \end{vmatrix}$$

$$M_3 = \begin{vmatrix} B & x-A & H \\ D & y-C & J \\ F & z-E & L \end{vmatrix} \qquad M_4 = \begin{vmatrix} B & G & x-A \\ D & I & y-C \\ F & K & z-E \end{vmatrix}$$

The denominators $M_1^2 + M_2^2$ are positive because M_1 is nonzero exactly when $\mathbf{l}_1$ does not contain a real point. $\mathbf{l}_1$ contains a real point if and only if the vectors $[B \;\; D \;\; F]^T$, $[G \;\; I \;\; K]^T$, and $[H \;\; J \;\; L]^T$ are linearly dependent, and this is equivalent to $M_1 = 0$.

There certainly cannot be two distinct complex values u_1 and u_2 such that $\mathbf{p}$, $\mathbf{l}_1(u_1)$, and $\overline{\mathbf{l}_1(u_1)}$ are collinear and also $\mathbf{p}$, $\mathbf{l}_1(u_2)$, and $\overline{\mathbf{l}_1(u_2)}$ are collinear as that would imply $\mathbf{l}_1(u_1)$, $\mathbf{l}_1(u_2)$, $\overline{\mathbf{l}_1(u_1)}$ and $\overline{\mathbf{l}_1(u_2)}$ are coplanar, which is impossible as $\mathbf{l}_1$ and $\overline{\mathbf{l}_1}$ are skew. ■

Proof of Theorem 3: Given an arbitrary real point (x_0, y_0, z_0) on the cubic surface, Equation (87) can be used to obtain a specific parameter value $u_0 = (a_0, b_0)$. This value of (a_0, b_0), when inserted into the parameterization (86), gives back (x_0, y_0, z_0), unless (a_0, b_0) happens to make the fractions in (86) 0/0, which means that (a_0, b_0) is a base point of the parameter map. ■

As will be shown in Section 6, there are five base points in this F_4 parameterization, with one of them being real. The points on the cubic surface which may be missed include one real line, which corresponds to the real base point. The other base points correspond to two pair of complex conjugates lines. For each pair, if the two lines are coplanar, and thus have a real point in common, that point is also missed in the parameterization. Skew complex lines corresponding to base points result in no missed real surface points.

It may seem odd that a real line may be missed by this parameterization, but in fact the real line does intersect the two skew complex conjugate lines. Here an extended notion of a real line is used: a line may be of the form $\mathbf{p} = \mathbf{d}u$ where $\mathbf{p}$ is a real 3D point and $\mathbf{d}$ is a real 3D vector, but in the context here we have to allow u to take on all complex values. With this understanding it is possible for an apparent real line to intersect both complex conjugate skew lines in complex points, and when it does, the points of intersection are complex conjugates. All points on this real line map into the same (a_0, b_0).

The F_5 Case When the cubic surface is of class F_5 it does not have any complex conjugate skew lines. One could attempt to use one real line and one complex line, or two non-conjugate complex skew lines, and proceed as before. However, there is no simple way to describe the values the parameters u and v may take on. In the F_1, F_2, and F_3 cases, u and v were unrestricted real parameters. In the F_4 case, when we let $u = \Re(u) + \Im(u)i$ and $v = \Re(v) + \Im(v)i$, we obtained a parameterization in which $\Re(u)$ and $\Im(u)$ are unrestricted, and then $\Re(v) = \Re(u)$ and $\Im(v) = -\Im(u)$. If we try the same idea with one real and one complex line, or two complex lines which are not conjugates, and let $\Re(u)$ and $\Im(u)$ be unrestricted, then $\Re(v)$ and $\Im(v)$ are complicated functions of $\Re(u)$ and $\Im(u)$, typically seventh degree polynomials.

In [37], a rational parameterization based on tangent planes at points lying on a real line is given. However, in general this only parametrizes part of the cubic surface. Points on the surface which do not lie on any tangent plane through a point on the chosen real line are missed, and these may account for a substantial portion of the surface. Since our goal is to parametrize the entire surface we instead parametrize the surface by parametrizing planes through one of the real lines on the surface, and then by parametrizing the conic sections which are the further intersections of these planes with the cubic surface. The parameterization of the conics will be that of [3]. One cost of parametrizing the whole surface is that we now have to use a square root in the parameterization. Another drawback of this parameterization is that there are typically two values of (u, v) corresponding to points on the cubic surface, instead of the one-to-one map resulting when both curves used in the parameterization are line, as in the F_1 through F_4 cases. Also, we have to use two distinct parameterizations; one which works when the conics are ellipses and the other for hyperbolas.

The procedure for finding the parameterization starts out like the ones for the F_1 through F_4 cases. In this case three coplanar real lines and 24 complex lines are determined, and the complex lines are found to come in 12 coplanar conjugate pairs. Since the methods of the other cases involving skew lines do not work here, one of the real lines is chosen to be mapped into the x-axis and the plane of the three real lines is mapped into the xy-plane. Specifically, suppose a real line $\mathbf{l}$ is given by $\mathbf{l}(u) = (A + Bu, C + Du, E + Fu)$ and that the normal to the plane is given by $\mathbf{N} = (N_1, N_2, N_3)$. $\mathbf{N}$ is obtained by taking the cross product of the (unit) direction vectors of two of the real lines, or by taking any unit vector perpendicular to the real lines if they are all parallel. Next, let $\mathbf{B} = (B_1, B_2, B_3)$ be the cross product of the direction vector of $\mathbf{l}$ with $\mathbf{N}$. We move a point on $\mathbf{l}$ to the origin by the translation $x = x' + A$, $y = y' + C$, $z = z' + E$, and then apply the transformation

$$\begin{aligned} x' &= (B_2N_3 - B_3N_2)x'' + (FN_2 - DN_3)y'' + (DB_3 - FB_2)z'' \\ y' &= (B_3N_1 - B_1N_3)x'' + (BN_3 - FN_1)y'' + (FB_1 - BB_3)z'' \\ z' &= (B_1N_2 - B_2N_1)x'' + (DN_1 - BN_2)y'' + (BB_2 - DB_1)z'' \ . \end{aligned} \tag{88}$$

This brings $\mathbf{l}$ to the x'' axis and the plane of the real lines to $z'' = 0$.

Planes through the x''-axis can be parametrized by $z'' = uy''$ for real values of u. All planes through the x''-axis are obtained except for $z'' = 0$, the plane containing the three real lines already found. The cubic surface now has an equation of the form $f''(x'', y'', z'') = 0$, and satisfies $f''(x'', 0, 0) = 0$. If we now make the substitution $z'' = uy''$ into $f''(x'', y'', z'')$, we obtain and equation that factors as $y''g''(x'', y'') = 0$, where $g''(x'', y'')$ is a quadratic in x'' and y''. The factor of y'' indicates that the line $y'' = 0$ is in the intersection of the cubic surface and the plane $z'' = uy''$ for any real u. The conic section $g(x'', y'') = 0$ is parametrized as in [3]: Let $g(x'', y'') = ax''^2 + by''^2 + cx''y'' + dx'' + ey'' + f$, and the discriminant $k = c^2 - 4ab$. The quantities a through f are polynomials in u.

If $k < 0$, the conic is an ellipse, and is parametrized by

$$\begin{aligned} x'' &= \frac{[af(ce - 2bd) - d(t_2 + t_3)]v^2 + [df(ce - 2bd) - 2ft_3]v + f^2(ce - 2bd)}{a(t_1 + t_3)v^2 - df(c^2 - 4ab)v + f(t_1 - t_3)} \\ y'' &= \frac{f(c^2 - 4ab)(av^2 + dv + f)}{a(t_1 + t_3)v^2 - df(c^2 - 4ab)v + f(t_1 - t_3)} \ , \end{aligned}$$

where

$$t_1 = ae^2 + bd^2 - cde \ , \quad t_2 = t_1 + f(c^2 - 4ab) \ , \quad t_3 = \sqrt{t_1 t_2} \ .$$

This gives real points only when the terms t_1 and t_2 have the same sign or are zero. If t_1 and t_2 have opposite sign, $g(x'', y'') = 0$ has no real points, and geometrically this means that the plane $z'' = uy''$ intersects the cubic surface only in the x''-axis. Thus values of u should be restricted to those that give non-negative values for t_1t_2. Upon back substitution using $z'' = uy''$ and (88), in the final parameterization x, y, and z are given by quotients of functions of the form $Q_1(u, v) + Q_2(u, v)\sqrt{Q_9(u)}$, where $Q_1(u, v)$ is of degree six in u and two in v, $Q_2(u, v)$ is of degree one in u and two in v, and $Q_3(u)$ is of degree nine in u alone. Due to the use of floating-point arithmetic, a nonzero coefficient for u^{10} may appear in $Q_3(u)$, and this is subtracted off in case it does show up.

If $k \geq 0$, the conic is a hyperbola or parabola, and is parametrized by

$$\begin{aligned} x'' &= \frac{a(c+\sqrt{c^2-4ab})v^2 + 2aev + f(c-\sqrt{c^2-4ab})}{2a\sqrt{c^2-4ab}\,v + 2ae - cd + d\sqrt{c^2-4ab}} \\ y'' &= \frac{-2a(av^2+dv+f)}{2a\sqrt{c^2-4ab}\,v + 2ae - cd + d\sqrt{c^2-4ab}} \,. \end{aligned}$$

Here real values are given for all u and v for which the denominators are nonzero. In the final parameterization x, y, and z are given by quotients of functions of the form $[Q_1(u,v) + Q_2(u,v)\sqrt{Q_3(u)}]$ $/[Q_4(u,v) + Q_5(u,v)\sqrt{Q_3(u)}]$, where $Q_1(u,v)$ is of degree three in u and two in v, $Q_2(u,v)$ is of degree one in u and two in v, $Q_3(u)$ is of degree four in u alone, $Q_4(u,v)$ is of degree three in u and one in v, and $Q_5(v)$ is of degree one in each of u and v.

7.5.4 Classification and Straight Lines from Parametric Equations

We also consider the question of deriving a classification and generating the straight lines of the cubic surface given its rational parametric equations (equation (80) above):

$$x = \frac{f_1(u,v)}{f_4(u,v)},\ y = \frac{f_2(u,v)}{f_4(u,v)},\ z = \frac{f_3(u,v)}{f_4(u,v)},$$

Note that given an arbitrary parameterization, the fact that it belongs to a cubic surface can be computed by determining the parameterization base points and multiplicities.

The computation of real base points which are the simultaneous zeros of $f_1 = f_2 = f_3 = f_4 = 0$, are obtained by first computing the real zeros of $f_1 = f_2 = 0$ using resultants and subresultants, via the method of birational maps [9] and then keeping those zeros which also satisfy $f_3 = f_4 = 0$. The classification follows from the reality of the base points, as detailed in the preliminaries section.

Having determined the base points, the straight lines on the cubic surface are then determined by the image of these points and combinations of them. In general there can be six real base points for cubic surfaces. The image of each of the six base points under the parameterization map yields a straight line on the surface. Next the fifteen pairs of base points define lines in the u, v parameter space, whose images under the parameterization map also yield straight lines. Finally the six different conics in the u, v parameter space which pass through distinct sets of five base points, also yield straight line images under the parameterization map. See Bajaj and Royappa for techniques to find parametric representations of the straight lines which are images of these base points. The question of determining parametric representations of the straight lines which are the images of parameter lines or parameter conics is for now, open.

Normally a cubic surface parameterization has six base points, but in the case of our parameterizations for the F_1, F_2, F_3, and F_4 surfaces, the number of base points is reduced to five. This happens because the degree of the parameterization is sufficiently small: neither u nor v appears to a power higher than the second. Consider the intersection of the parametrized surface with a line in 3-space. Let the line be given as the intersection of two planes $a_ix + b_iy + c_iz + d_i = 0$ for $i = 1, 2$. Then when the substitutions $x = f_1(u,v)/f_4(u,v)$, $y = f_2(u,v)/f_4(u,v)$, $z = f_3(u,v)/f_4(u,v)$ are made into the equations of the lines, we obtain polynomials of degree two in each of u and v. When resultants of these polynomials are taken to eliminate either u or v, univariate polynomials of degree eight are obtained. This indicates that there could be as many as eight intersection points of the line with the surface. However, a cubic surface will intersect the line in only three (possibly complex) points, counting multiplicity and solutions at infinity. The difference between these two results (eight and three) is the number of base points. A cubic parameterization would have led to nine possible intersection points when considering the algebraic equations, and hence six, the difference of nine and three, is the number of base points for such a parameterization.

Let $\mathbf{l}_1$ and $\mathbf{l}_2$ be the two skew lines used in the parameterization, whether they be real or complex. The base points (u, v) correspond to lines on the cubic surface which intersect both $\mathbf{l}_1$ and $\mathbf{l}_2$. Real base points correspond to real lines and complex base points correspond to complex lines. One of the many useful results on nonsingular cubic surfaces is that given any two (real or complex) skew lines on the surface, there are exactly five lines that intersect both . For an F_1 surface, the five transversal lines, and the base points, are all real. Thus those five real lines are missed by the parameterization (80). For an F_2 surface, three of the base points are real and the other two form a complex conjugate pair. The parameterization (80) consequently misses the three real lines incident to both $\mathbf{l}_1$ and $\mathbf{l}_2$. In addition, if the two transversal complex conjugate lines are coplanar and have a real intersection point, that point is also missed. For both F_3 and F_4 surfaces, one base point is real and the other four form two conjugate pairs. In each of these cases there is one real line through both $\mathbf{l}_1$ and $\mathbf{l}_2$, and that line is missed. Again, if a pair of transversal complex conjugate lines is coplanar, their real intersection point is missed, so there may be two such isolated points for F_3 and F_4. As will be demonstrated in the example below, the missing points on the surface can be approached as (u, v) approaches the corresponding base point in an appropriate manner.

In addition to the transversal lines, two conic sections are also missed in the parameterization of the F_1, F_2, and F_3 surfaces. One conic is obtained as follows: take the intersection of the plane containing $\mathbf{l}_1(u)$ and perpendicular to $\mathbf{l}_2(v)$ with the cubic surface. This intersection consists of $\mathbf{l}_1$ plus a conic. It turns out that the value of v at which $\mathbf{l}_2$ intersects this plane tends to $\pm\infty$. Thus points on the conic are not obtained for finite values of v, even though the line $\mathbf{l}_1$ does turn out to be reachable. The other missing conic is found by interchanging the roles of $\mathbf{l}_1$ and $\mathbf{l}_2$. These two conics lie on parallel planes, and may be obtained if the parameterization uses projective coordinates as indicated in the example below.

We have presented a method of extracting real straight lines and from there a rational parameterization of each of four families of nonsingular cubic surfaces. The parameterizations of the real cubic surface components are constructed using a pair of real skew lines for those three families which have them, and remarkably using a complex conjugate pair of skew lines, in a fourth family. In each of these, the entire real surface is covered except for one, three, or five lines which intersect both skew lines, one or two isolated points, and two conic sections. The missing conics can be recovered through the use of projective instead of real coordinates. For the last family, in which two real skew lines do not exist, in order to cover the whole surface we had to use two separate parameterizations, each involving a square root. Fortunately many graphics applications, such as the triangulation of a real surface, will involve only the classes of cubics which do contain real skew lines. These real skew lines will correspond to non-intersecting edges of the tetrahedra. Open problems remain in computing the images of curves containing the real base points in the parameter plane.

An additional associated line of future research is in computing invariants for cubic surfaces based on its straight lines. In Computer Vision, as pointed out by Holt-Netravali and Mundy-Zisserman , it is essential to derive properties of curves and surfaces which are invariant to perspective projection and to be able to compute these invariants reliably from perspective image intensity data. In connection with FFT (= First Fundamental Theorem of Invariant Theory), referring to Abhyankar and Mundy-Zisserman for details, we attempt to calculate complete systems of symbolic invariants of cubic surfaces. In doing these calculations, it is important to know all the relations between a set of invariants which is the content of SFT (= Second Fundamental Theorem of Invariant Theory).

Turning to our specific situation, we may derive projective invariants of a cubic surface from simultaneous invariants of the 27 lines on it. Namely, by taking the coefficients of two planes through a line in 3-space we get a 2×4 matrix whose 2×2 minors are the six Grassmann coordinates of the line. Thus we get a 27×6 matrix; its 6×6 minors are invariants and pure covariants as well as dot products between them. This is the FFT of vector invariants. Since we have derived an effective classification of cubic surfaces based on the line configurations, we can now derive these invariants (symbolically). Details of this procedure are left to a subsequent paper.

7.5.5 Parameterization of general algebraic plane curves by A-splines

In general, a degree d curve can be parameterized if it satisfies the Cayley - Riemann criterion. Consider a curve C_d, with a $d-1$ singular point. By sending that point to infinity, we can draw lines which intersect the curve at one point each. The slopes of these pencil of lines obtains the parameterization of the curve.

An A-spline of degree n over the triangle $[\mathbf{p}_1\mathbf{p}_2\mathbf{p}_3]$ is defined by

$$G_n(x, y) := F_n(\boldsymbol{\alpha}) = F_n(\alpha_1, \alpha_2, \alpha_3) = 0, \tag{89}$$

where

$$F_n(\alpha_1, \alpha_2, \alpha_3) = \sum_{i+j+k=n} b_{ijk} B^n_{ijk}(\alpha_1, \alpha_2, \alpha_3), \quad B^n_{ijk}(\alpha_1, \alpha_2, \alpha_3) = \frac{n!}{i!j!k!}\alpha_1^i \alpha_2^j \alpha_3^k,$$

and $(x, y)^T$ and $(\alpha_1, \alpha_2, \alpha_3)^T$ are related by

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} \mathbf{p}_1 & \mathbf{p}_2 & \mathbf{p}_3 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix}. \tag{90}$$

Here the objective is to get an A-spline parameterization of the following form:

$$\mathfrak{X}(t) = \sum_{i=0}^{n} w_i B_i^n(t)\mathbf{b}_i \Big/ \sum_{i=0}^{n} w_i B_i^n(t), \qquad t \in [0, 1], \tag{91}$$

where $\mathbf{b}_i \in \mathbb{R}^3$, $w_i \in \mathbb{R}$, and $B_i^n(t) = \{n!/[i!(n-i)!]\}t^i(1-t)^{n-i}$. Without loss of generality, we may assume that $w_0 = 1$ (otherwise we could divide through by t and have a parameterization of one lower degree). Next, under the transformation

$$t = \frac{t' + at'}{1 + at'}, \quad a > -1, \quad t' \in [0, 1], \tag{92}$$

the curve (91) will preserve its form, that is

$$\mathfrak{X}(t) = \sum_{i=0}^{n}(1+a)^i w_i B_i^n(t')\mathbf{b}_i \Big/ \sum_{i=0}^{n}(1+a)^i w_i B_i^n(t'), \qquad t' \in [0, 1].$$

Therefore, we may assume further that $w_n = 1$ by setting $a = w_n^{-1/n} - 1$, which makes $(1+a)^n w_n = 1$, in the transformation (92).

We consider first convex C^1 continuous A-splines (see Figure 5(a)). An A-spline being C^1 implies that $b_{n00} = b_{0n0} = b_{n-1,01} = b_{0,n-1,1} = 0$, as shown in [14]. The C^0 continuous A-splines on the triangle $[\mathbf{p}_1\mathbf{p}_2\mathbf{p}_3]$ can be made into C^1 continuous A-splines on the triangle $[\mathbf{p}_1\mathbf{p}_2\mathbf{p}'_3]$ (see Figure 5(c)) through the use of the subdivision formula (see [24]). In our applications in the parameterization of cubic ($n = 3$) A-patches, the coefficients $a,\ b,\ c$ are fixed and $d,\ e,\ f$ are parameters to be determined, where

$$a = b_{210}, \quad b = b_{120}, \quad c = b_{111}, \quad d = b_{102}, \quad e = b_{012}, \quad f = b_{003}.$$

The non-convex case (see Figure 5(b)) can be converted to the convex case by first computing the intersection point $\mathbf{p}'_2$, which leads to a linear equation for $n = 3$, and then computing the tangent of the curve at $\mathbf{p}'_2$. Note that this tangent does not depend upon the coefficients $d,\ e,\ f$.

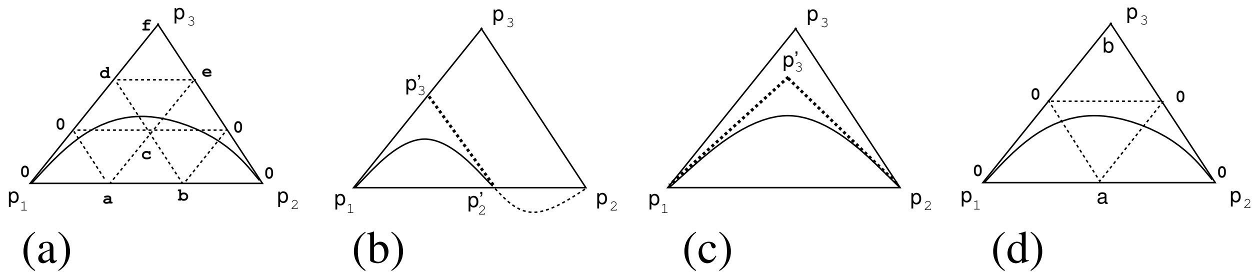


Figure 5: (a): Convex case; (b) Non-convex case; (c) C^0 A-spline; (d) Quadratic A-spline.

Quadratic A-splines It is not difficult to see that the parametric form of a C^1-continuous quadratic A-spline should have the following form (see Figure 5(d)) since it interpolates the points $\mathbf{p}_1$ and $\mathbf{p}_2$ and is tangent to the lines $[\mathbf{p}_1\mathbf{p}_3]$ and $[\mathbf{p}_2\mathbf{p}_3]$ at the points $\mathbf{p}_1$ and $\mathbf{p}_2$, respectively.

$$\mathfrak{X}(t) = \frac{\mathbf{p}_1 B_0^2(t) + w_1\mathbf{p}_3 B_1^2(t) + \mathbf{p}_2 B_2^2(t)}{B_0^2(t) + w_1 B_1^2(t) + B_2^2(t)}, \qquad t \in [0,1], \tag{93}$$

where w_1 is a parameter to be determined. This is called a (2/2) rational parameterization because the of the numerator and denominator are each of degree 2 in t. In order for the quadratic A-spline to be rationally parameterizable, we must have

$$w_1 = \sqrt{-\frac{b_{110}}{2b_{002}}} \geq 0. \tag{94}$$

Cubic A-splines We first show that an irreducible C^1-continuous cubic A-spline never has a (2/2) rational parameterization. If we substitute the αs defined by (90) into $F_3(\boldsymbol{\alpha}) = 0$, we have $\sum_{i=0}^{6} c_i B_i^6(t) \equiv 0$, where

$$c_0 = b_{300}, \quad c_1 = b_{201}w_1, \quad c_2 = \frac{1}{5}b_{210} + \frac{4}{5}b_{102}w_1^2, \quad c_3 = \frac{3}{5}b_{111}w_1 + \frac{2}{5}b_{003}w_1^3,$$
$$c_6 = b_{030}, \quad c_5 = b_{021}w_1, \quad c_4 = \frac{1}{5}b_{120} + \frac{4}{5}b_{012}w_1^2.$$

Since $B_i^6(t),\ i = 0,\ \ldots,\ 6$, are linearly independent, we have $c_i = 0, \quad i = 0,\ \ldots,\ 6$. It then follows that

$$a + 4dw_1^2 = 0, \quad 3cw_1 + 2fw_1^3 = 0, \quad b + 4ew_1^2 = 0$$

and hence $w_1 = \sqrt{-a/4d}$. The coefficients of the A-spline must satisfy

$$\frac{d}{a} = \frac{f}{6c} = \frac{e}{b}, \tag{95}$$

where

$$a = b_{210}, \quad b = b_{120}, \quad c = b_{111}, \quad d = b_{102}, \quad e = b_{012}, \quad f = b_{003}.$$

However, the substitutions (95) turn the A-spline $F_3(\boldsymbol{\alpha}) = 3a\alpha_1^2\alpha_2 + 3b\alpha_1\alpha_2^2 + 6c\alpha_1\alpha_2\alpha_3 + 3d\alpha_1\alpha_3^2 + 3e\alpha_2\alpha_3^2 + f\alpha_3^3 = 0$ into $F_3(\boldsymbol{\alpha}) = (\alpha_1\alpha_2 + d/a\ \alpha_3^2)(a\alpha_1 + b\alpha_2 + 2c\alpha_3) = 0$, which is the product of a line and an ellipse. The parameterization covers the ellipse, and is essentially the same as the (2/2) parameterization of a quadratic A-spline.

The (3/3) rational parametric form of a C^1-continuous cubic A-spline should have the following form in order to interpolate the points $\mathbf{p}_1$ and $\mathbf{p}_2$ and be tangent to the lines $[\mathbf{p}_1\mathbf{p}_3]$ and $[\mathbf{p}_2\mathbf{p}_3]$ at $\mathbf{p}_1$ and $\mathbf{p}_2$, respectively:

$$\mathfrak{X}(t) = \frac{\mathbf{p}_1 B_0^3(t) + w_1[\mathbf{p}_1 + \alpha(\mathbf{p}_3 - \mathbf{p}_1)]B_1^3(t) + w_2[\mathbf{p}_2 + \beta(\mathbf{p}_3 - \mathbf{p}_2)]B_2^3(t) + \mathbf{p}_2 B_3^3(t)}{B_0^3(t) + w_1 B_1^3(t) + w_2 B_2^3(t) + B_3^3(t)}, \tag{96}$$

where α, β, w_1, w_2 are parameters to be determined.

We will show that the equation

$$\begin{aligned} G_{[\mathbf{p}_1\mathbf{p}_2\mathbf{p}_3]}(a,b,c,d,e,f) = 48a^3e^3f^2 - 9a^2b^2f^4 + 72a^2bcef^3 - 72a^2bde^2f^2 - 96a^2c^2e^2f^2 \\ - 288a^2cde^3f + 432a^2d^2e^4 + 72ab^2cdf^3 - 72ab^2d^2ef^2 - 8abc^3f^3 \\ - 552abc^2def^2 + 1152abcd^2e^2f - 864abd^3e^3 + 48ac^4ef^2 + 576ac^3de^2f \\ - 864ac^2d^2e^3 + 48b^3d^3f^2 - 96b^2c^2d^2f^2 - 288b^2cd^3ef + 432b^2d^4e^2 \\ + 48bc^4df^2 + 576bc^3d^2ef - 864bc^2d^3e^2 - 288c^5def + 432c^4d^2e^2 = 0 \end{aligned} \tag{97}$$

gives a condition on the A-spline coefficients that guarantee the A-spline has a rational parameterization. The proof of this is rather technical.

We will wish to construct rationally parameterizable cubic A-splines defined on a triangle $[\mathbf{p}_1\mathbf{p}_2\mathbf{p}_3]$ and passing through $\mathbf{p}_1$ and $\mathbf{p}_2$, that are not necessarily tangent to the edges $[\mathbf{p}_1\mathbf{p}_3]$ and $[\mathbf{p}_2\mathbf{p}_3]$ at $\mathbf{p}_1$ and $\mathbf{p}_2$. This situation is illustrated in Figure 5(c), where the tangent lines at $\mathbf{p}_1$ and $\mathbf{p}_2$ intersect at some other point $\mathbf{p}_3'$. These cubic A-splines will have one degree of freedom, the weight b_{003}, which we will use to satisfy (97). In order to accomplish this we define a coordinate system $\alpha_1'\alpha_2'\alpha_3'$ (where $\alpha_1' + \alpha_2' + \alpha_3' = 1$) that has its origin $(0,0,1)$ at $\mathbf{p}_3'$ instead of $\mathbf{p}_3$, while keeping the points $(1,0,0)$ and $(0,1,0)$ fixed.

The general cubic curve passing through $\mathbf{p}_1$ and $\mathbf{p}_2$ is

$$\begin{aligned} 3b_{210}\alpha_1^2\alpha_2 + 3b_{201}\alpha_1^2\alpha_3 + 3b_{120}\alpha_1\alpha_2^2 + 6b_{111}\alpha_1\alpha_2\alpha_3 \\ + 3b_{102}\alpha_1\alpha_3^2 + 3b_{021}\alpha_2^2\alpha_3 + 3b_{012}\alpha_2\alpha_3^2 + b_{003}\alpha_3^3 = 0 \ . \end{aligned} \tag{98}$$

The tangent lines to this curve at $\mathbf{p}_1$ and $\mathbf{p}_2$ are

$$\begin{aligned} b_{210}\alpha_2 + b_{201}\alpha_3 = 0 \\ b_{120}\alpha_1 + b_{021}\alpha_3 = 0 \ , \end{aligned}$$

and these intersect at the point

$$(\alpha_1, \alpha_2, \alpha_3) = \frac{(b_{210}b_{021}, b_{201}b_{120}, -b_{210}b_{120})}{b_{210}b_{021} + b_{201}b_{120} - b_{210}b_{120}} \ .$$

The linear transformation that maps $(\alpha_1, \alpha_2, \alpha_3) = (1,0,0), (0,1,0), (b_{210}b_{021}, b_{201}b_{120}, -b_{210}b_{120})/(b_{210}b_{021}+b_{201}b_{120}-b_{210}b_{120})$ into $(\alpha_1', \alpha_2', \alpha_3') = (1,0,0), (0,1,0), (0,0,1)$, respectively is

$$\begin{aligned} \alpha_1 &= \alpha_1' + \frac{b_{210}b_{021}}{b_{210}b_{021} + b_{201}b_{120} - b_{210}b_{120}}\,\alpha_3' \\ \alpha_2 &= \alpha_2' + \frac{b_{201}b_{120}}{b_{210}b_{021} + b_{201}b_{120} - b_{210}b_{120}}\,\alpha_3' \\ \alpha_3 &= -\frac{b_{210}b_{120}}{b_{210}b_{021} + b_{201}b_{120} - b_{210}b_{120}}\,\alpha_3' \end{aligned} \tag{99}$$

with the inverse

$$\begin{aligned} \alpha_1' &= \alpha_1 - \frac{b_{021}}{b_{120}}\,\alpha_3 \\ \alpha_2' &= \alpha_2 - \frac{b_{201}}{b_{210}}\,\alpha_3 \\ \alpha_3' &= -\frac{b_{210}b_{021} + b_{201}b_{120} - b_{210}b_{120}}{b_{210}b_{120}}\,\alpha_3 \end{aligned} \tag{100}$$

Thus the transformation (99) maps (98) into an equation of the form

$$a'{\alpha_1'}^2\alpha_2' + 3b'\alpha_1'{\alpha_2'}^2 + 6c'\alpha_1'\alpha_2'\alpha_3' + 3d'\alpha_1'{\alpha_3'}^2 + 3e'\alpha_2'{\alpha_3'}^2 + f'{\alpha_3'}^3 = 0 \ .$$

7.5.6 Parameterization of algebraic space curves

Space curves are projected to planes to obtain plane curves, which are then parameterized by the method described in the previous sections. It is shown that such a projection always exists. We need to obtain a birational map, which maps the space curve to a plane curve which has the same genus.

7.5.7 Inversion of Parameterizations

Parameterizations tell how to map a point in parameter space to a curve. The inversion of this map, called a *chart* in differential topology, tells how to map a point on a curve to its parameter value. Certain assumptions must be made on the parameterization in order for it to have a computable inverse.

Similar to the case of curves, a parametric surface is a very special algebraic variety of dimension 2 in x, y, z, s, t space, since the surface lies in the 3-dimensional subspace defined by x, y, z and furthermore points on the surface can be put in 1-to-1 rational correspondence with points on the 2-dimensional sub-space defined by s, t. Figure 2 depicts the relationship between parametric and non-parametric surfaces.

Example parametric (rational algebraic) surfaces are degree two algebraic surfaces (quadrics) and most degree three algebraic surfaces (cubic surfaces). The cylinders of nonsingular cubic curves and the cubic surface cone are of not rational.

Other examples of rational algebraic surfaces are Steiner surfaces which are degree four surfaces with a triple point, and Plücker surfaces which are degree four surfaces with a double curve. In general, a necessary and sufficient condition for the rationality of an algebraic surface of arbitrary degree is given by Castelnuovo's criterion: $P_a = P_2 = 0$, where P_a is the arithmetic genus and P_2 is the second plurigenus [43]. Algorithms for symbolically deriving the parametric equations of degree two and three rational surfaces are given in [5, 6, 35].

Both the parametric and the implicit representation of algebraic curve segments and algebraic surface patches can be represented in either Bernstein-Bézier or B-spline bases.

Rational parametric algebraic surface The canonical representation of a rational parametric algebraic surface patch in x, y, z space are given by

$$\begin{cases} X = P_1(s,t), \\ Y = P_2(s,t), \\ Z = P_3(s,t), \\ W = P_4(s,t). \end{cases}$$

or

$$\begin{cases} x = X/W, \\ y = Y/W, \\ z = Z/W, \end{cases}$$

where the P_i are polynomials in any of the above appropriate bases and the variables/parameters s, and t range over a finite interval (or canonically the unit interval [0,1], see [13]).

The domain of the mapping for rational algebraic parametric surfaces is usually one of the following two kinds:

- Tensor domain: The parameters s, t are defined over the interval [0,1]. ($s \in [0,1], t \in [0,1]$)
- Barycentric domain: This is a triangular domain, with the parameters ranging over a finite interval and satisfying the condition: $0 \leq s, t \leq 1$.

Multi-sided patches:

Base points are isolated pairs of parameter values which satisfy $P_1 = P_2 = P_3 = P_4 = 0$ and hence cause the parametric map to be ill-defined (0/0).

For example, in the hyperboloid of 1 sheet, we see a pair of lines being absent due to the ill-formed mapping.

The image in x, y, z of the base points in the parameter domain, are in general curves, yielding multi-sided patches.

Implicit Algebraic Surface Patches An implicit algebraic surface patch can be defined in x, y, z space by :

$$w = P(x, y, z) \wedge w = 0$$

where the P is a polynomial in any of the above appropriate basis and the variables x, y, z range over the unit interval [0,1]. Alternatively, the surface patch can be defined by a closed cycle of trimming curves which may be defined with rational parametric equations or implicitly or both. In section 3 the surfaces patches are defined implicitly with a closed triple (triangle) of rational trimming curves.

As in the *Rational parametric representation*, we have many elements over which we can define the patches. In general, the simplest polyhedron we consider are

- *Cube (Tensor domain)* The parameters x, y, z are defined over the interval [0,1]. ($x \in [0, 1], y \in [0, 1], z \in [0, 1]$) This yield a tensor product Bernstein-Bézier coordinate system for trivariate polynomials.

- *Tetrahedron* The parameters x, y, z are defined with the condition: $0 \leq x + y + z \leq 1$. This yields a barycentric coordinate system for trivariate polynomials.

- *Triangular prism* The parameters x, y, z are now defined as follows: z is defined over the interval [0,1]. ($z \in [0, 1]$) and x, y range over $0 \leq x + y \leq 1$

- *Square pyramid* The parameters x and y satisfy $x \in [0, 1]$ and $y \in [0, 1]$, while z satisfies the condition $0 \leq x + y + z \leq 1$

References

[1] S. Abhyankar. *Algebraic space curves.* Gaetan Morin Editeur Ltee, 1971.

[2] S. Abhyankar. Desingularization of plane curves. *Singularities*, 40(Part 1):1, 1983.

[3] S. Abhyankar and C. Bajaj. Automatic Parameterization of Rational Curves and Surfaces I: Conics and Conicoids. *Computer Aided Design*, 19,1:11–14, 1987.

[4] S. Abhyankar and C. Bajaj. Automatic Parameterization of Rational Curves and Surfaces II: Cubics and Conicoids. *Computer Aided Design*, 19,9:499–502, 1987.

[5] S. S. Abhyankar and C. L. Bajaj. Automatic parameterization of rational curves and surfaces III: algebraic plane curves. *Comput. Aided Geom. Des.*, 5(4):309–321, 1988.

[6] S. S. Abhyankar and C. L. Bajaj. Automatic parameterization of rational curves and surfaces IV: algebraic space curves. *ACM Trans. Graph.*, 8(4):325–334, 1989.

[7] A. Aho and J. Hopcroft. *Design & Analysis of Computer Algorithms.* Pearson Education India, 1974.

[8] C. Bajaj. Geometric modeling with algebraic surfaces. In D. Handscomb, editor, *The Mathematics of Surfaces III*, pages 3–48. Oxford Univ. Press, 1988.

[9] C. Bajaj. Geometric computations with algebraic varieties of bounded degree. In *Proc. of the Sixth ACM Symposium on Computational Geometry*, pages 148–156, Berkeley, California, 1990.

[10] C. Bajaj. The Emergence of Algebraic Curves and Surfaces in Geometric Design. In R. Martin, editor, *Directions in Geometric Computing*, pages 1–29. Information Geometers Press, 1993.

[11] C. Bajaj, J. Chen, and G. Xu. *Modeling with Cubic A-Patches. ACM Transactions on Graphics*, 14, 2:103–133, 1995.

[12] C. Bajaj and A. Royappa. A note on an efficient implementation of the Sylvester resultant for multivariate polynomials. *Computer Science Technical Report CSD-TR-718, Purdue University*, 1988.

[13] C. Bajaj and A. Royappa. Finite Representations of Real Parametric Curves and Surfaces. *International Journal of Computational Geometry and Applications*, 5(3):313–326, 1995.

[14] C. Bajaj and G. Xu. Data fitting with cubic A-splines. In M. Gigante and T. Kunii, editors, *Proceedings of Computer Graphics International, CGI'94*, pages 329–340, Melbourne, Australia, 1996. World Scientific Publishing.

[15] W. Blythe. *On Models of Cubic Surfaces*, 1905.

[16] M. Bôcher and E. Duval. *Introduction to higher algebra.* Macmillan, 1922.

[17] W. Böhm, G. Farin, and J. Kahmann. A survey of curve and surface methods in CAGD. *Computer Aided Geometric Design*, 1:1–60, 1984.

[18] W. Brown. The subresultant PRS algorithm. *ACM Transactions on Mathematical Software (TOMS)*, 4(3):237–249, 1978.

[19] B. Buchberger. Gröbner-bases: An algorithmic method in polynomial ideal theory, 1985.

[20] J. Canny. The Complexity of Robot Motion Planning, 1988.

[21] G. Collins. Subresultants and reduced polynomial remainder sequences. *Journal of the ACM (JACM)*, 14(1):128–142, 1967.

[22] G. Collins. The calculation of multivariate polynomial resultants. In *Proceedings of the second ACM symposium on Symbolic and algebraic manipulation*, page 222. ACM, 1971.

[23] M. do Carmo. *Riemannian Geometry.* Boston, 1992.

[24] G. Farin. *Curves and Surfaces for Computer Aided Geometric Design: A Practical Guide, Second Edition.* Academic Press Inc., 1990.

[25] T. Garrity and J. Warren. Geometric Continuity. *CAGD*, 8:51 – 65, 1991.

[26] A. Henderson. *The Twenty Seven Lines upon the Cubic Surface*, 1911.

[27] S. Lodha and J. Warren. Bézier Representaion for Cubic Surface Patches. *Computer Aided Design*, 24(12):643 – 650, 1992.

[28] R. Loos. Generalized Polynomial Remainder Sequences, 1983.

[29] L. J. Mordell. *Diophantine equations.* Pure and Applied Mathematics, Vol. 30. Academic Press, London, 1969.

[30] V. Pan. Fast and efficient algorithms for sequential and parallel evaluation of polynomial zeros and of matrix polynomials. In *26th Annual Symposium on Foundations of Computer Science, 1984.*, pages 522–531, 1985.

[31] J. Renegar. On the worst-case arithmetic complexity of approximating zeros of systems of polynomials. *SIAM Journal on Computing*, 18:350, 1989.

[32] G. Salmon. *Lessons introductory to the modern higher algebra.* Hodges, Figgis, and co., 1885.

[33] G. Salmon. *A Treatise on the Analytic Geometry of Three Dimensions, vol. I and II*, 1914.

[34] J. Schwartz. Fast probabilistic algorithms for verification of polynomial identities. *Journal of the ACM (JACM)*, 27(4):717, 1980.

[35] T. Sederberg and J. Snively. Parameterization of Cubic Algebraic Surfaces. In R. Martin, editor, *The Mathematics of Surfaces II*, pages 299 – 319, 1987.

[36] T. Sederberg and J. Snively. Parametrization of cubic algebraic surfaces. In *Proceedings on Mathematics of surfaces II*, pages 299–319. Clarendon Press, 1987.

[37] T. Sederberg and J. Snively. Parametrization of cubic algebraic surfaces. In R. R. Martin, editor, *Mathematics of Surfaces II*, pages 299–310. Clarendon Press, Oxford, 1987.

[38] J. Semple and L. Roth. *Introduction to Algebraic Geometry.* Oxford University Press, Oxford, U.K., 1949.

[39] V. Snyder and C. Sisam. Analytic geometry of space. *Bull. Amer. Math. Soc. 22 (1916), 350-354. DOI: 10.1090/S0002-9904-1916-02797-2 PII: S*, 2(9904):02797–2, 1916.

[40] R. Walker. *Algebraic Curves.* Springer Verlag, New York, 1978.

[41] T. J. Willmore. *Total Curvature in Riemannian Geometry.* Ellis Horwood Limited, 1982.

[42] T. J. Willmore. *Riemannian Geometry.* Clarendon Press, 1993.

[43] Zariski, O. *Algebraic Surfaces.* Ergebnisse der Mathematik und ihre Grenzgebiete 4, 1935.