1992

R.C. Veltkamp

Survey of continuities of curves and surfaces

Computer Science/Department of Interactive Systems Report CS-R9202 January

CWI is the research institute of the Stichting Mathematisch Centrum, which was founded on February 11, 1946, as a non-profit institution aiming at the promotion of mathematics, computer science, and their applications. It is sponsored by the Dutch Government through the Netherlands organization for scientific research (NWO).

Survey of Continuities of Curves and Surfaces

Remco C. Veltkamp
CWI, Department of Interactive Systems
Kruislaan 413, 1098 SJ Amsterdam, the Netherlands
email: remco@cwi.nl

Abstract

This survey presents an overview to various types of continuity of curves and surfaces, in particular parametric (C^n), visual or geometric (V^n, G^n), Frenet frame (F^n), and tangent surface continuity (T^n), and discusses the relation with curve and surface modeling, visibility of (dis)continuities, and graphics rendering algorithms. It is the purpose of this paper to provide an overview of types of continuity, and to put many terms and definitions on a common footing in order to give an understanding of the subject.

1991 Mathematics Subject Classification: 41A15, 65D07, 68U05
1991 Computing Reviews Classification:
I.3.5 [Computer Graphics] Computational geometry and object modeling.
I.3.7 [Computer Graphics] Three-Dimensional Graphics and Realism.

Key Words and Phrases: Curves, Surfaces, Continuity, Shading, Modeling.

1 Introduction

Many researchers in the wide field of graphics occasionally encounter applications where aspects of continuity of curves and surface are important. However, many papers on these topics use different terms and techniques, or are focussed on a single aspect. This holds even for general textbooks about curves and surfaces for computer graphics and computer aided geometric design [8, 32], and for survey and tutorial type articles [6, 7, 14, 40, 46]. This confusion of terminology, obscures the relation between the different methods that describe continuity.

In this paper we explain many notions of continuity (parametric, geometric, visual, β-, Frenet frame, tangent surface), and much of the terminology and fundamental notions involved (for example n-jet, generalised curvature, Dupin indicatrix). I briefly discuss the relation between types of continuity and types of splines, such as β-, γ-, ν-, τ-splines, as well as Catmull-Rom spline curve and tensor product surface constructions. I further relate the notions of continuity to illumination models and shading algorithms for computer graphics rendering, with visual aspects of continuity.

Although the emphasis of this paper is on breadth rather than depth, many specialized details are given when that is necessary to develop definitions of continuity etc. However, all proofs of theorems are omitted; the reader is referred to the appropriate research papers instead.

The rest of this paper has the following structure: Section 2 is about continuity of curves and Section 3 about surfaces. Section 4 discusses applications in the field of curve and surface modeling and Section 5 mentions the implications of continuity for shaded rendering.

Remark 1 The remarks can be skipped without affecting the understanding of the material. All following remarks only provide additional terminology related to concepts introduced in the rest of the text.

Report CS-R9202
ISSN 0169-118X
CWI
P.O. Box 4079, 1009 AB Amsterdam, The Netherlands

2 Continuity of curves

2.1 Fundamentals

An arbitrary curve c is a topologically one-dimensional object, in an embedding space of arbitrary dimension larger than one.

A *functional* curve is a scalar function $c : \mathbb{R} \rightarrow \mathbb{R}$. A *parametric* curve is a vector-valued function $c : \mathbb{R} \rightarrow \mathbb{R}^d$, each coordinate component being a function of the parameter: $c(t) = (c^1(t), \ldots, c^d(t))^T$, $t_{min} \leq t \leq t_{max}$, for a d-dimensional curve. Parametric curves are often used because of their independence of a particular coordinate frame. In the rest of the paper we deal only with parametric curves.

A *rational* curve $c : \mathbb{R} \rightarrow \mathbb{R}^d$ has rational components: $c(t) = (f^1(t)/g(t), \ldots, f^d(t)/g(t))^T$. Note that arbitrary rational components can always be rewritten so as to have a common denominator, as in the given formulation. The *homogeneous* curve $C : \mathbb{R} \rightarrow \mathbb{R}^{d+1}$ associated with the rational, or projected, curve is $C(t) = (f^1(t), \ldots, f^d(t), g(t))^T$. In the following, all curves can also be homogeneous curves, but a rational curve is always denoted explicitly.

One particular type of parametric curve is the *polynomial* curve:

$$c(t) = a_0 + a_1 t + a_2 t^2 + \ldots + a_k t^k \tag{1}$$

for some integer $k > 0$, the degree of the polynomial curve. The coefficients a_i are vectors whose components are the coefficients for each coordinate. The number of coefficient vectors, $k + 1$, is the *order* of the polynomial. Polynomial curves are often used because they are easy to handle, for example for evaluation and determination of derivatives.

DEFINITION 1 (SUPPORT) *The support of a (basis) function is the closure of the set of parameter values for which it is non-zero.*

A function has local support if its support is finite.

In Equation (1) the curve is represented as a linear combination of the so called power basis functions $1, t^1, \ldots, t^k$. We can also use other basis functions $B_i^k(t)$:

$$p(t) = \sum_{i=0}^{k} a_i B_i^k(t). \tag{2}$$

The B_i^k are often called blending functions, and the a_i weights or control points. A well known example of such blending functions are the Bernstein polynomials [9]:

$$B_i^k(t) = \binom{k}{i} t^i (1-t)^{k-i}. \tag{3}$$

The derivative of a curve is a vector, the derivatives are taken component-wise. The n-th derivative of a curve $c(t)$ is denoted $c^{(n)}(t)$. To avoid potential problems with the parameterization of the curve, we assume in the rest of the paper that the (first) derivative vector of all curves is not equal to the null-vector: $c^{(1)}(t) \neq 0$. Such a curve and its parameterization are called *regular*.

A *piecewise* polynomial curve is defined segment by segment. The parameter range is then partitioned into subranges: $t_{min} = t_0 \leq t_1 \leq \ldots t_m = t_{max}$, where t_i are fixed parameter values. The curve is defined for each subrange: $s(t) = s_i(t)$, for $t_{i-1} \leq t < t_i$. We are usually interested in positionally continuous piecewise curves:

$$\lim_{t \uparrow t_i} s_i(t) = s_{i+1}(t_i),$$

as in Figure 1, but that is not always necessary.

DEFINITION 2 (KNOT, BREAKPOINT, JOINT) *The t_i are called knots, $(t_0, \ldots, t_m)$ is a knot vector or knot sequence. Multiple knots are consecutive equal knots. Breakpoints are* distinct *knots. $s(t_i)$ is a junction point or joint, if s is positional continuous at $s(t_i)$.*

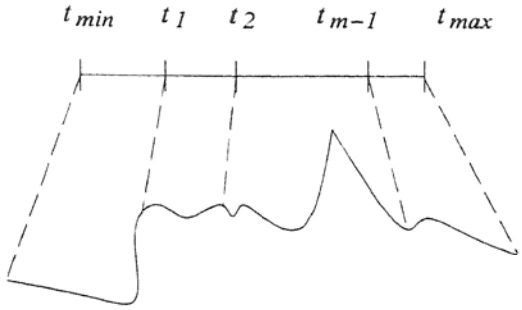

Figure 1: Piecewise curve.

Multiple knots can be used to influence the shape and the parameterization of the curve.

The parameter subranges can be transformed so as to provide a local parameter v, $v_{min} \leq v \leq v_{max}$, for example to normalize into $[0, 1]$: $v = (t - t_i)/(t_{i+1} - t_i)$, $t_{i+1} \neq t_i$. The curve $s_i(t)$ can then be *reparameterized* into $\tilde{s}_i(v) = s_i(t(v))$, $v_{min} \leq v \leq v_{max}$, in the previous example $t(v) = t_i + v(t_{i+1} - t_i)$. In the general case the local parameter ranges are independent of each other, so the ranges may be disconnected or overlap.

A given curve can have many parameterizations. One important parameterization is the arc-length parameterization:

DEFINITION 3 (ARC-LENGTH PARAMETERIZATION) *For a curve $s(v)$, $v_{min} \leq v \leq v_{max}$, the arc-length w as a function of v, is defined as*

$$w(v_0) = \int_{v_{min}}^{v_0} \|s^{(1)}(v)\| dv,$$

which is the length of s from $s(v_{min})$ to $s(v_0)$. The curve $\tilde{s}(w) = s(v^{-1}(w))$ is the corresponding arc-length parameterized curve.

Here $\|\ \|$ denotes the Euclidean norm or length. Although the arc-length is an important concept it is used primarily for theoretical purposes, e.g. to develop continuity conditions.

The word *spline* is an East Anglian dialect word, denoting a metal or wooden strip, bended around pins to form a pleasing shape. It was observed that under gentle bending the shape corresponds to a piecewise *cubic* polynomial function having continuous first and second derivatives. In the context of mathematical curves a spline can be of any degree. The theory of splines originates from approximation theory. Spline approximation in its present form first appeared in a paper by Schoenberg [72], who developed methods for the smooth approximation of empirical tables. In approximation theory a spline of order $n+1$ is generally defined as a piecewise polynomial of degree n that is everywhere C^{n-1}-continuous (see Section 2.2 for C^n-continuity). In geometric modeling it is sometimes desired to model discontinuities on purpose, so that the continuity requirement in the definition is left out:

DEFINITION 4 (SPLINE) *A spline function is a piecewise polynomial function, a spline curve a curve whose components are spline functions.*

One particular spline function is the B-spline (basic spline), so called because translates of this function form a basis for the space of spline functions. B-splines were probably already known to Hermite, and certainly to Peano early this century [73] but were introduced in geometric modeling by Riesenfeld [68] and have become a popular device for curve and surface design. An easy introduction to B-spline curves and surfaces is given by Bartels et al. [8]

We are interested in the continuity of spline curves segments $r(u)$, $u_{min} \leq u \leq u_{max}$ and $s(v)$, $v_{min} \leq v \leq v_{max}$, at points $s(v_0)$ and $r(u_0)$ on the curves, in particular at the endpoints $r(u_{max})$

and $s(v_{min})$ where they are supposed to connect. An important device for the algebraic aspects of continuity is the connection matrix [37] which describes how two segments are connected. I define the connection matrix in terms of the n-jet:

DEFINITION 5 (n-JET) *For a function $f(v) : \mathbb{R} \to \mathbb{R}^d$ and a fixed v_0 the* n-jet *$D_n f(v_0)$ is defined as $(f(v_0), f^{(1)}(v_0), \ldots, f^{(n)}(v_0))^T$.*

One can also specify an n-jet $(\alpha_0, \ldots, \alpha_n)$ without giving the original function.

DEFINITION 6 (CONNECTION MATRIX) *The matrix $M = (M_{ij})$, $i, j = 0, \ldots, n$ is a* connection matrix *for $r(u)$ and $s(v)$ and fixed parameter values u_0 and v_0, if and only if $D_n(s)(v_0) = M D_n(r)(u_0)$.*

Note that M applies to all the components of the curve.

2.2 Parametric continuity

Parametric continuity is the classical notion of continuity in analysis: if a function is n times continuously differentiable, or more exactly, the derivatives exist and are continuous, then the function is n-th order parametric continuous. In our context we get:

DEFINITION 7 (PARAMETRIC CONTINUITY) *Curves $r(u)$ and $s(v)$ are n-th order parametric continuous, $n \geq 0$, at u_0 and v_0 if and only if $D_n(s)(v_0) = D_n(r)(u_0)$.*

Parametric continuity of order n is denoted C^n. Positional *dis*continuity is denoted C^{-1}; C^0-continuity amounts to positional continuity. A one-segment polynomial curve of arbitrary degree, see Equation (1), is C^∞.

If two homogeneous curves in $\mathbb{R}^{d+1}$ are C^n, then so are the projected rational curves in $\mathbb{R}^d$, provided that the denominaters are non-zero. Therefore, C^n-continuity is said to be projectively invariant. However, the C^n-continuity of the homogeneous curves is not necessary for the rational curves to be C^n. The necessary and sufficient conditions are as follows [48].

THEOREM 1 (RATIONAL PARAMETRIC CONTINUITY) *Let $r(u)$ and $s(v)$ be two homogeneous curves in $\mathbb{R}^{d+1}$. The associated rational curves $R(u)$ and $S(v)$ in $\mathbb{R}^d$ are C^n at u_0 and v_0 if and only if the denominators of the rationals are non-zero and there exists an n-jet $(\alpha_0, \ldots, \alpha_n)$ such that $D_n(s)(v_0) = A_n D_n(r)(u_0)$, where matrix $A_n = (a_{ij})$ has entries defined as $a_{ij} = \binom{i}{j}\alpha_{i-j}$ if $i \geq j$ and zero otherwise.*

So, matrix A_n looks like

$$(4) \qquad A_n = \begin{pmatrix} \alpha_0 & 0 & 0 & \cdots & 0 \\ \alpha_1 & \alpha_0 & 0 & \cdots & 0 \\ \alpha_2 & 2\alpha_1 & \alpha_0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ \binom{n-1}{0}\alpha_{n-1} & \binom{n-1}{1}\alpha_{n-2} & \binom{n-1}{2}\alpha_{n-3} & \cdots & 0 \\ \binom{n}{0}\alpha_n & \binom{n}{1}\alpha_{n-1} & \binom{n}{2}\alpha_{n-2} & \cdots & \binom{n}{n}\alpha_0 \end{pmatrix}.$$

Two curve segments need not have the same derivative vector at their joint in order to have the same tangent line, as illustrated in Figure 2. Similarly, they need not be C^2 in order to have the same normal curvature (defined in Section 2.4). A crucial observation here is that derivatives depend on the parameterization while the tangent line and curvature depend on the shape of the spline and are independent of parameterization, i.e. they are *intrinsic*.

In order to base the notion of continuity on intrinsic aspects of the curve, we can follow two approaches: take a closer look at the effects of parameterizations (algebraic approach, next subsection) or take the intrinsic notions like tangent and curvature as a starting point (differential geometry approach, Section 2.4).

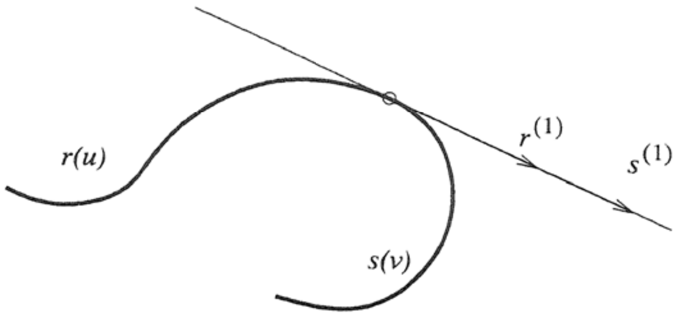

Figure 2: Two curve segments joining with derivative of same direction but different magnitude.

2.3 Geometric continuity

Since parametric continuity depends on the parameterization, one possibility for an intrinsic notion of continuity is to avoid any dependency on a specific parameterization. This leads to a definition of continuity called geometric, or visual continuity:

DEFINITION 8 (GEOMETRIC CONTINUITY) *Curves $r(u)$ and $s(v)$ are n-th order geometric continuous at u_0 and v_0, if and only if there exists a reparameterization $u = u(\tilde{u})$ such that $\tilde{r}(\tilde{u}) = r(u(\tilde{u}))$ and $s(v)$ are C^n at $s(v_0)$.*

Geometric continuity of order n is denoted G^n or GC^n, visual continuity VC^n or V^n.

The term geometric continuity was first used by Barsky [5], the term visual continuity by Farin [28]. However, the concepts of geometric continuity were already exploited in for example [34].

Definition 8 is impractical for applications, since it is non-constructive. By contrast, the chain and product rule of differentiation show that $\tilde{r}^{(k)}(\tilde{u}) = d^k\tilde{r}(\tilde{u})/d\tilde{u}^k$ can be rewritten in terms of $d^i r(u)/du^i$ and $d^i u(\tilde{u})/d\tilde{u}^i$, $i = 1, \ldots, k$. For example

$$\frac{d^2\tilde{r}}{d\tilde{u}^2} = \frac{d^2 r}{du^2}\left(\frac{du}{d\tilde{u}}\right)^2 + \frac{dr}{du}\frac{d^2u}{d\tilde{u}^2}.$$

Letting β_i indicate $u^{(i)}(\tilde{u}_0)$, we get the following equations:

(5a) $$s^{(0)}(v_0) = r^{(0)}(u_0)$$

(5b) $$s^{(1)}(v_0) = \beta_1 r^{(1)}(u_0)$$

(5c) $$s^{(2)}(v_0) = \beta_1^2 r^{(2)}(u_0) + \beta_2 r^{(1)}(u_0)$$

(5d) $$s^{(3)}(v_0) = \beta_1^3 r^{(3)}(u_0) + 3\beta_1\beta_2 r^{(2)}(u_0) + \beta_3 r^{(1)}(u_0)$$

(5e) $$s^{(4)}(v_0) = \beta_1^4 r^{(4)}(u_0) + 6\beta_1^2\beta_2 r^{(3)}(u_0) + (4\beta_1\beta_3 + 3\beta_2^2) r^{(2)}(u_0) + \beta_4 r^{(1)}(u_0)$$

$$\vdots$$

where $\beta_1 > 0$ is required to let the tangent vectors have the same direction, in which case the curves are said to have the same orientation. Of course $\beta_1 = 1$ and $\beta_i = 0$, $i > 1$, amounts to C^n-continuity.

Equation (5a) amounts to positional continuity, Equation (5b) means that the derivative vectors differ only a scalar factor, and (5c) prescribes a dependency as depicted in Figure 3.

The parameters β_i are called shape-handles because they can be used to model the shape of the curve. In particular, β_1 is called bias and β_2 tension, because of their specific shape changing effect. One particular spline that satisfies the β-constraints is the β-spline [3, 4], a geometrically continuous analogue of the B-spline, see also Section 4.1.

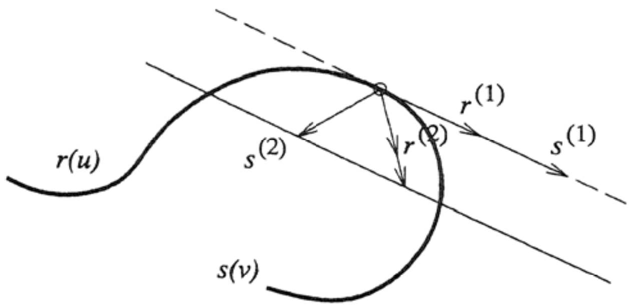

Figure 3: Relation between derivative vectors for geometric continuity.

Equations (5) can be put in a matrix notation:

$$D_n(s)(v_0) = M_n(\beta)D_n(r)(u_0), \tag{6}$$

where the connection matrix $M_n(\beta) = (m_{ij}(\beta))$ result from the chain rule. The entries m_{ij}, $i \geq j$, are polynomials in the β_i-s:

$$m_{ij}(\beta) = \sum_{\substack{k_1+k_2+\cdots+k_i=j \\ k_1+2k_2+\cdots+ik_i=i \\ k_1,k_2,\ldots,k_i\geq 0}} c(i,j,k_1,\ldots,k_i)\beta_1^{k_1}\cdots\beta_i^{k_i},$$

where the coefficients are defined recursively by

$$\begin{aligned} c(i+1,j,k_1,\ldots,k_{i+1}) = {} & c(i,j-1,k_1-1,\ldots,k_i) \\ & + (k_1+1)\,c(i,j,k_1+1,k_2-1,k_3,\ldots,k_i) \\ & + (k_2+1)\,c(i,j,k_1,k_2+1,k_3-1,\ldots,k_i) \\ & \vdots \\ & + (k_{i-1}+1)\,c(i,j,k_1,k_2,\ldots,k_{i-1}+1,k_i-1), \end{aligned}$$

where $c(i,j,\ldots) = 0$ for $j > i$ and $c(1,1,1) = 1$. An explicit expression for $c(i,j,k_1,\ldots,k_i)$ is part of Faa di Bruno's formula [52]:

$$c(i,j,k_1,\ldots,k_i) = \frac{i!}{k_1!(1!)^{k_1}\cdots k_i!(i!)^{k_i}}.$$

The connection matrix $M_n(\beta)$ is called a reparameterization matrix. As we can see from the formula, $M_n(\beta)$ is a lower triangular matrix with $m_{ii}(\beta) = \beta_1^i$, $m_{i0}(\beta) = \delta_{i0}$, $m_{i1}(\beta) = \beta_i$, $i \geq 1$, and the remaining entries in the subdiagonals are polynomials in the β_i-s such that β_i does not appear in $m_{ij}(\beta)$ for $j > i$:

$$M_n(\beta) = \begin{pmatrix} 1 & 0 & 0 & & \cdots & & 0 \\ 0 & \beta_1 & 0 & & \cdots & & 0 \\ 0 & \beta_2 & \beta_1^2 & & & & \\ 0 & \beta_3 & 3\beta_1\beta_2 & \beta_1^3 & & & \vdots \\ \vdots & \vdots & \vdots & \vdots & \ddots & & \\ 0 & \beta_{n-1} & & \cdots & & \beta_1^{n-1} & 0 \\ 0 & \beta_n & & \cdots & & & \beta_1^n \end{pmatrix}. \tag{7}$$

So, an alternative definition of geometric continuity is the following:

THEOREM 2 (GEOMETRIC CONTINUITY) *Curves $r(u)$ and $s(v)$ are nth order geometric continuous at u_0 and v_0, if and only if there exists a reparameterization matrix $M_n(\beta)$ such that*

$$D_n(s)(v_0) = M_n(\beta)D_n(r)(u_0).$$

The conditions in Definition 8 and Theorem 2 are equivalent. It can be shown [6] that the following theorem provides another equivalent formulation:

THEOREM 3 (GEOMETRIC CONTINUITY) *Curves $r(u)$ and $s(v)$ are nth order geometric continuous at u_0 and v_0, if and only if the corresponding arc-length parameterized curves $\tilde{r}(t)$ and $\tilde{s}(w)$ are C^n at $\tilde{s}(w(v_0))$.*

G^n-continuity is projectively invariant. Moreover, the reparameterization matrix is also projectively invariant. So, if two homogeneous curves in $\mathbb{R}^{d+1}$ are G^n, the projected rational curves in $\mathbb{R}^d$ are G^n with the same reparameterization matrix [37]. However, the G^n-continuity of the homogeneous curves is not necessary for the rational curves to be G^n. The necessary and sufficient conditions are as follows [48].

THEOREM 4 (RATIONAL GEOMETRIC CONTINUITY) *Let $r(u)$ and $s(v)$ be two homogeneous curves in $\mathbb{R}^{d+1}$. The associated rational curves $R(u)$ and $S(v)$ in $\mathbb{R}^d$ are G^n at u_0 and v_0 if and only if the denominator of the rationals are non-zero and there exist a reparameterization matrix $M_n(\beta)$ and an n-jet $(\alpha_0, \ldots, \alpha_n)$ such that*

$$D_n(s)(v_0) = A_n M_n(\beta) D_n(r)(u_0),$$

where matrix A_n is defined in terms of α_i by Equation (4).

2.4 Frenet frame continuity

This subsection takes a differential geometry approach to continuity. All concepts from differential geometry that are used here can be found in the text book on this subject by do Carmo [22].

The direction of the tangent line is determined by the tangent vector, the normalized derivative vector $s^{(1)}(v)/\|s^{(1)}(v)\|$ which has unit length. Two curves need not have the same derivative vector in order to have the same tangent vector, as was illustrated in Figure 2. There is a particular parameterization that gives a unit length derivative vector at every point, so that the derivative and tangent vector are equal: the arc-length parameterization.

Let $s(w)$ be an arc-length parameterized curve. The tangent vector t_1 is now

$$t_1(w) = s^{(1)}(w), \qquad \|t_1\| = 1,$$

The normal curvature vector is defined as

$$t_2(w) = t_1^{(1)}/\kappa_1(w)$$

where $\kappa_1(w)$ is a scalar such that $\|t_2(w)\| = 1$; $\kappa_1(w)$ is called curvature.

Remark 2 The normal curvature vector is often called main or principal normal vector.

The *bi*normal curvature vector, normal to t_1 and t_2, is defined as

$$t_3(w) = \frac{t_2^{(1)}(w) + \kappa_1(w)t_1(w)}{\kappa_2(w)}$$

where $\kappa_2(w)$ is a scalar such that $\|t_3(w)\| = 1$; $\kappa_2(w)$ is called torsion. Planar curves have zero torsion.

Remark 3 The torsion is often indicated by $\tau(w)$.

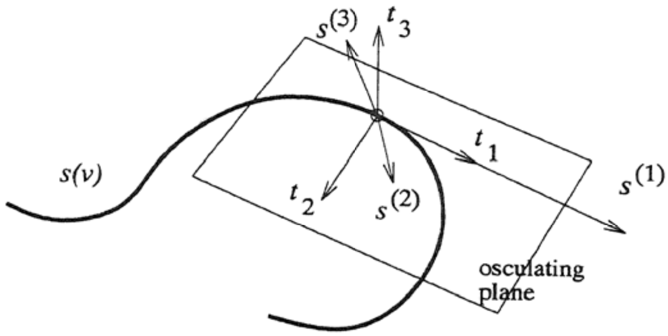

Figure 4: Frenet frame.

Curvature and torsion have an intuitive geometrical meaning: $\kappa_1(w)$ and $-\kappa_2(w)$ are the angular velocities of $t_1(w)$ and $t_2(w)$. The plane spanned by the tangent and the normal curvature vector is called the osculating plane, see Figure 4.

The notions of tangent, normal curvature and binormal curvature can be generalized to so called generalized curvatures [37, 48]:

$$\begin{aligned} t_1(w) &= s^{(1)}(w) \\ \kappa_0(w) &= 0 \\ t_{i+1}(w) &= \frac{t_i^{(1)}(w) + \kappa_{i-1}(w)t_{i-1}(w)}{\kappa_i(w)}, \\ &\quad i = 1, \ldots, n-1, \end{aligned} \tag{8}$$

where $\kappa_i(w)$ is such that $\|t_{i+1}(w)\| = 1$.

Remark 4 The generalized curvature vectors t_i and scalars κ_i can also be derived from $s^{(i)}$ by the Gram-Schmidt orthogonalization process [40].

Analogous to the osculating plane, the linear space spanned by the first $d-1$ curvature vectors is called an osculating linear space.

Definition 9 (Frenet frame) *In* $\mathbb{R}^d$, *the* Frenet frame *is defined as the first d curvature vectors* $(t_1(w), \ldots, t_d(w))$.

Remark 5 The Frenet frame is also called Frenet-Serret or Serret-Frenet frame, and in 3D also Frenet trihedron.

The moving Frenet frame as a function of w uniquely defines the shape of the curve. So when the tangent, curvature and torsion vectors of a curve in 3D are given, we are able to draw the curve, see also Koenderink [53].

Note that the curve must be arc-length parameterized in the above definition but the generalized curvature vectors and curvature scalars are intrinsic properties of the curve. That is, independent of the actual parameterization any curve possesses an arc length parameterization and corresponding generalized curvatures, so that we can speak of a Frenet frame of an arbitrarily parameterized curve.

Frenet frame continuity of a curve in $\mathbb{R}^d$ can be defined as the continuity of the Frenet frame $(t_1(w), \ldots, t_d(w))$ as a whole [40]. A more general definition is the following [37, 48]:

Definition 10 (Frenet frame continuity) *Two curves* $r(u)$ *and* $s(v)$ *are* n*-th order Frenet frame continuous,* $n > 0$, *at* u_0 *and* v_0, *if and only if the first* n *curvature vectors and scalar curvatures coincide.*

Strictly speaking, in 'Frenet frame continuity' it is not the d-dimensional Frenet frame, but the n generalized curvatures that must be continuous. Frenet frame continuity of order n, $n > 0$is denoted F^n. F^0 and F^{-1} can be used to indicate positional continuity and discontinuity.

A curve in $\mathbb{R}^d$ that is F^d, is also F^n for $n > d$. So, Frenet frame continuity distinguishes only between $d+2$ classes of curves: F^{-1}, F^0, ..., F^d. In particular, a curve in 3-space that is F^3 continuous, is also F^n, $n > 3$, for example a F^2 planar curve in 3-space (it has a zero torsion).

The Frenet frame at $s(v_0)$ can also be defined as the unique set of orthonormal vectors $t_1, \ldots, t_n$ satisfying

$$D_n(s)(v_0) = M\ (s(v_0), t_1, \ldots, t_n)^T$$

where M is a lower triangular matrix with $m_{i0} = \delta_{i0}$, and a positive diagonal [26].

Frenet frame continuity can also be defined in matrix form [26] :

THEOREM 5 (FRENET FRAME CONTINUITY) *Two curves $r(u)$ and $s(v)$ are n-th order Frenet frame continuous, $n > 0$, at u_0 and v_0, if and only if*

$$D_n(s)(v_{max}) = N_n D_n(r)(u_{min})$$

where $N_n = (n_{ij})$ is a lower triangular matrix with $n_{i0} = \delta_{i0}$, $n_{11} > 0$, and $n_{ii} = n_{11}^i$.

Thus the diagonal of N_n is the same as that for $M_n(\beta)$, see Equation (7), but the $n(n-1)/2$ non-zero subdiagonal entries are entirely arbitrary. N is called the Frenet frame connection matrix. It is immediately clear that Frenet frame continuity is a weaker restriction than geometric continuity.

If we relax the condition $n_{ii} = n_{11}^i$ and permit arbitrary non-zero values along the diagonal, we get continuity of the first n osculating linear spaces. By relaxing the condition that N_n is lower triangular, we get even weaker forms of continuity [26].

F^n-continuity is projectively invariant. However, the Frenet frame connection matrix is not projectively invariant. So, if two homogeneous curves in $\mathbb{R}^{d+1}$ are F^n, the projected rational curves in $\mathbb{R}^d$ are F^n but generally with another Frenet frame connection matrix [13, 37]. However, the F^n-continuity of the homogeneous curves is not necessary for the rational curves to be F^n. The necessary and sufficient conditions are as follows [48].

THEOREM 6 (RATIONAL FRENET FRAME CONTINUITY) *Let $r(u)$ and $s(v)$ be two homogeneous curves in $\mathbb{R}^{d+1}$. The associated rational curves $R(u)$ and $S(v)$ in $\mathbb{R}^d$ are F^n at u_0 and v_0 if and only if the denominator of the rationals are non-zero and there exist a Frenet frame connection matrix N and an n-jet $(\alpha_0, \ldots, \alpha_n)$ such that $D_n(s)(v_0) = A_n N_n D_n(r)(u_0)$, where matrix A_n is defined in terms of α_i by Equation (4).*

Continuity of tangent line and curvature were already used at an early stage [33, 69, 10, 58, 59], but the generalization has only been carried out recently [26, 48].

2.5 Comparing G^n and F^n

G^n and F^n continuity agree for $n = 1$ and $n = 2$, but do not agree for $n > 2$. For $n \geq 3$, F^n continuity does not insure C^n-continuity of the corresponding arc-length parameterized curves. F^n is more general than G^n continuity, that is, there are curves that are Frenet frame continuous, but do not satisfy the β-constraints, and so do not possess a regular C^n parameterization.

Frenet frame continuity discriminates only $d+2$ classes but geometric continuity distinguishes between an infinite number of classes: G^n, $n = -1, 0, 1, \ldots$. Two planar curves in 3-space (which have zero torsion), can meet with F^3-continuity, while their derivatives of normal curvature $\kappa_1'(w)$ are not equal.

Because G^n and F^n agree for $n = 1$ and $n = 2$, the terms geometric, visual and Frenet frame continuity are not used in a consistent way in the literature. The way we have used them until now is mostly used in the computer graphics community. However, sometimes geometric continuity refers to Frenet frame continuity, and visual continuity to the β-constraints [64]. Sometimes geometric continuity is used in a broad sense to mean both 'arc length GC' and 'Frenet frame GC' [40]. An

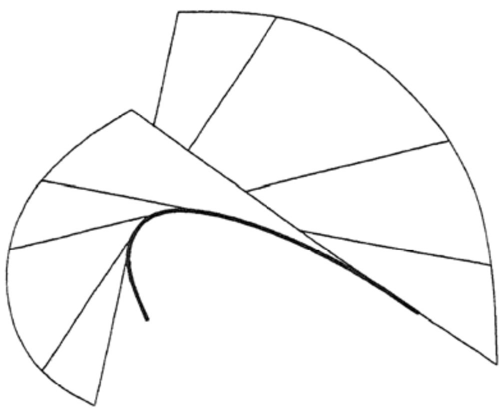

Figure 5: The 1-tangent surface $t^1(u,v)(s)$ of a non-planar curve.

even more general notion of geometric geometry is the existence of an arbitrary connection matrix [37], while phrases like 'GC in the sense of reparameterization' and 'Frenet frame GC' are used to be specific. Adding to the confusion, the term 'geometric spline' is used for a G^2 spline specifically. The meaning of 'β-continuity' [36] and 'algebraic continuity' [12] should be clear; it is also called 'contact of order n' especially by German authors [34, 14, 64].

2.6 Tangent surface continuity

This subsection is about the classification of Frenet frame continuous curves into projectively invariant classes, developed by Pottmann [64]. It is based on geometric continuity of so-called ruled tangent surfaces, so that notions of continuity of surfaces, treated in Section 3, are used.

Ruled tangent surfaces are special ruled surfaces. A ruled surface in 3-space is defined as

(9) $$r(v,w) = (1-w)s_1(v) + ws_2(v).$$

The curves s_1 and s_2 are joined by a family of straight lines called rulings performing a linear interpolation between s_1 and s_2.

Remark 6 Rulings are also called generators, s_1 and s_2 are called directrices.

We can write Equation (9) as $s(v,w) = s_1(v) - w(s_1(v) - s_2(v))$ where $s_1(v) - s_2(v)$ are direction vectors of the rulings. This can be generalized to a $(k+1)$-dimensional ruled surface in d-space as follows (w_i are separate variables, not fixed values of one variable):

(10) $$r^k(v, w_1, \dots, w_k) = s(v) + \sum_{i=1}^{k} w_i g_i(v),$$

where $g_1(v), \dots, g_k(v)$ are linearly independent, forming a basis of the k-dimensional ruler at each value of v.

We obtain a special ruled surface if we set $g_i(v) = s^{(i)}(v)$. If $s^{(1)}(v), \dots, s^{(k)}(v)$ are linearly independent, they span k-dimensional osculating spaces. The resulting ruled surface is called the t-tangent surface of s and denoted $t^k(s)$:

(11) $$t^k(s)(v, w_1, \dots, w_k) = s(v) + \sum_{i=1}^{k} w_i s^{(i)}(v).$$

See Figure 5 for a 1-tangent surface.

THEOREM 7 (FRENET FRAME CONTINUITY) *Two curves $r(u)$ and $s(v)$ are n-th order Frenet frame continuous, $n > 0$, at u_0 and v_0, if and only if their k-tangent surfaces are all G^2 at u_0 and v_0 for $k = 0, \ldots, n-2$.*

Projective transformations map the k-tangent surface of a homogeneous curve $s(v)$ onto the k-tangent surface of the corresponding rational curve $S(v)$. Since G^2 is projectively invariant, it is clear from Theorem 7 that F^n is also projectively invariant, something that we already know from Section 2.4.

By Theorem 7, F^n-continuity is a classification based on k-tangent surfaces and G^2-continuity. This classification can be generalized as follows:

DEFINITION 11 (TANGENT SURFACE CONTINUITY CHARACTERISTIC) *Curves r and s possess tangent surface continuity characteristic $[(0, n_0)(1, n_1)\ldots(k, n_k)]$ at u_0 and v_0 if and only if the i-tangent surfaces $t^i(r)$ and $t^i(s)$ are G^{n_i} for $i = 0, \ldots, k$, at u_0 and v_0.*

The continuity constraints on a k-tangent surfaces only make sense if $d \geq k+2$. In particular in $\mathbb{R}^3$ we have only characteristics of the form $[(0, n_0)(1, n_1)]$, where the highest dimensional tangent surface involved is the 1-tangent surface, illustrated in Figure 5.

The tangent surface continuity characteristic is projectively invariant for the same reason as F^n is projectively invariant. By Theorem 7, F^n-continuity corresponds to the characteristic $[(0,2)(1,2)\ldots(n-2,2)]$.

DEFINITION 12 (TANGENT SURFACE CONTINUITY) *A curve with characteristic $[(0, n-2)(1, n-1)(2, n-2)\ldots(n, 0)]$ is called n-th order tangent surface continuous.*

Tangent surface continuity is denoted T^n or TC^n.

The tangent surface continuity characteristic of a curve is not unique. By definition of the tangent surface, G^0-continuity of the $(k+1)$-tangent surface implies G^1-continuity of the k-tangent surface. So, T^n requires characteristic $[(0, n-2)(1, n-1)(2, n-2)\ldots(n, 0)]$ and this is equivalent to $[(0, n-2)(1, n-1)(2, n-2)\ldots(n-1, 1)]$.

Like G^n and F^n, T^n-continuity can be expressed with a connection matrix. The following theorem provides the parameters that appear in the matrix:

THEOREM 8 (TANGENT SURFACE CONTINUITY) *Two curves $r(u)$ and $s(v)$ that are C^{n-2} at u_0 and v_0 are T^n there, if and only if*

$$\begin{aligned} s^{(n-1)} &= r^{(n-1)} + \nu_1 r^{(2)} + \nu_2 r^{(1)} \\ s^{(n)} &= r^{(n)} + (n-1)\nu_1 r^{(3)} + \nu_3 r^{(2)} + \nu_4 r^{(1)} \end{aligned}$$

at u_0 and v_0 for $\nu_i \in \mathbb{R}$.

For $n \geq 4$ the connection matrix $M(T^n)$ is given by

$$M(T^n) = M_n(\beta) \begin{pmatrix} 1 & 0 & 0 & & & \cdots & & 0 \\ 0 & 1 & 0 & & & \cdots & & 0 \\ 0 & 0 & 1 & & & & & \\ \vdots & \vdots & \ddots & & \ddots & & & \vdots \\ 0 & 0 & 0 & & & & & \\ 0 & \nu_2 & \nu_1 & 0 & \ddots & & \ddots & 0 \\ 0 & \nu_4 & \nu_3 & \nu_1(n-1) & 0 & \cdots & 0 & 1 \end{pmatrix}, \tag{12}$$

where $\nu_1, \ldots, \nu_4$ are given by Theorem 8. The connection matrices of T^3 and T^2 continuity do not fit in this scheme but are as follows:

$$M(T^3) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \alpha_1 & 0 & 0 \\ 0 & \alpha_2 & \alpha_1\alpha_3 & 0 \\ 0 & \alpha_4 & \alpha_5 & \alpha_1\alpha_3^2 \end{pmatrix}, \tag{13}$$

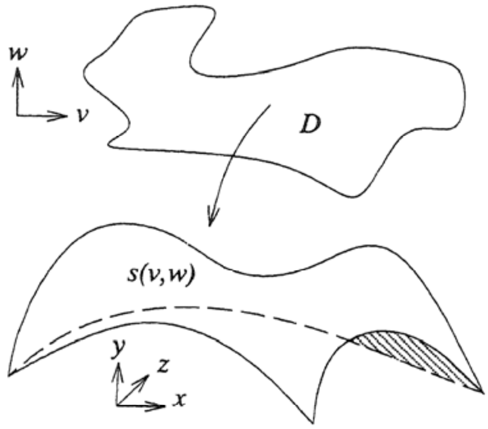

Figure 6: Mapping from a parameter domain to a surface.

$$M(T^2) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \alpha_1 & 0 \\ 0 & \alpha_2 & \alpha_3 \end{pmatrix}. \tag{14}$$

Since the tangent surface continuity characteristic is projectively invariant, T^n-continuity is also projectively invariant. Therefore, if two homogeneous curves are T^n-continuous, the corresponding rational curves are also T^n-continuous. However, a result analogous to Theorem 6 that would state the exact conditions on two homogeneous curves to let the rational curves be T^n has not been derived.

3 Continuity of surfaces

3.1 Fundamentals

In this section we consider surfaces that are topologically two-dimensional objects, in an ambient space $\mathbb{R}^d$ of arbitrary dimension d larger than two, except where explicitly stated otherwise.

A *functional* surface is a scalar function $s : \mathbb{R}^2 \to \mathbb{R}$. A *parametric* surface s is defined componentwise, each coordinate component being a function of two parameters (bivariate functions). The parameters are allowed to range over some arbitrarily shaped region $D \in \mathbb{R}^2$: $s(v, w) : D \to \mathbb{R}^d$. Thus in particular, the parameter domain need not be rectangular, see Figure 6. In its most general form, the parameter domain may have an arbitrary topology with disconnected pieces and holes. Such a trimmed domain corresponds to a *trimmed surface*, frequently resulting from intersections with other curved surfaces, for example in CSG-modeling. The domain over which the surface is defined is then modified, while the coordinate component functions are left unchanged [16].

A *rational* surface has rational components: $s(v, w) = (f^1(v, w)/g(v, w), \ldots , f^d(v, w)/g(v, w))^T$. The homogeneous surface $S(v, w)$ associated with the projected surface $s(v, w)$ is $S(v, w) = (f^1(v, w), \ldots , f^d(v, w), g(v, w))^T$. In the following, all surfaces can also be homogeneous surfaces, but a rational surface is always denoted explicitly.

A *polynomial* surface of *total degree* n has the following form:

$$s(v, w) = \sum_{i+j\leq n} a_{ij} v^i w^j.$$

A polynomial surface of *coordinate degree* (m, n) is of the form

$$s(v, w) = \sum_{i=0}^{m} \sum_{j=0}^{n} a_{ij} v^i w^j.$$

A *piecewise* polynomial surface $s(v, w)$ is defined patch by patch. The parameter domain is then partitioned into sub-domains, which can be transformed so as to provide local parameters for each patch. A *spline* surface is a piecewise polynomial surface.

The ith partial derivative with respect to v and the jth with respect to w is denoted $s^{(i,j)}(v, w)$:

$$s^{(i,j)}(v, w) = \frac{\partial^{i+j} s}{\partial v^i \partial w^j} (v, w)$$

Note that $s^{(i,j)}(v, w)$ is a vector. To avoid potential problems with the parameterization of a surface $s(v, w)$, we assume in the rest of the paper that the first partial derivatives $s^{(1,0)}(v, w)$ and $s^{(0,1)}(v, w)$ exist, and are linearly independent. The surface and its parameterization are then said to be *regular*.

A partial derivative is a special case of a *directional* derivative.

DEFINITION 13 (DIRECTIONAL DERIVATIVE) *The directional derivative at a surface point $s(v_0, w_0)$ in the direction $d = (d_v, d_w)$ in the parameter space, is*

$$\nabla_d s(v_0, w_0) = \lim_{h \to 0} \frac{s(v_0 + hd_v, w_0 + hd_w) - s(v_0, w_0)}{h}.$$

The partial derivative are obtained in the direction of the parameter space axes: $\nabla_{(1,0)} s(v_0, w_0) = s^{(1,0)}(v_0, w_0)$, and $\nabla_{(0,1)} s(v_0, w_0) = s^{(0,1)}(v_0, w_0)$.

Remark 7 The symbol "∇" is called nabla. The directional derivative $\nabla_d s$ is also called the covariant derivative of s by d.

Remark 8 Surface s is defined to be *weakly* or *Gateaux* differentiable at (v_0, w_0), if and only if the mapping $d \mapsto \nabla_d s(v_0, w_0)$ is linear and continuous. s is defined to be *strongly* or *Fréchet* differentiable, if and only if s is Gateaux differentiable, and the mapping $(v_0, w_0) \mapsto \nabla s(v_0, w_0)$ is continuous.

We are interested in the continuity of spline surfaces $r(t, u)$ and $s(v, w)$, in particular along a common curve or edge.

3.2 Parametric continuity

DEFINITION 14 (PARAMETRIC CONTINUITY) *Two surfaces $r(t, u)$ and $s(v, w)$ are C^n-continuous at (t_0, u_0) and (v_0, w_0), if and only if $r^{(i,j)}(t_0, u_0) = s^{(i,j)}(v_0, w_0)$, $i + j = 0, \ldots, n$.*

The surfaces are C^n along a common curve if they are C^n at each point on that curve. C^{-1} denotes positional discontinuity.

Note that two surfaces need not have the same first order partial derivatives in order to have the same tangent plane. As with curves the derivatives depend on the parameterization while the tangent plane does not. Moreover, on closed surfaces singularities occur where the derivative of the surface is not defined. As an example see Figure 7, showing a closed piecewise triangular surface s, together with a global parameterization and a partial local parameterization. Corresponding points in the parameter domain and the surface are indicated; the global domain is 'folded' so as to join the D_i to point D on the surface. Taking the derivative at $s(A)$ in the directions $D_1 - A$ and $D_2 - A$ in the global parameterization, it follows from Definition 13 that $\nabla_{D_1 - A} s(A) \neq \nabla_{D_2 - A} s(A)$ (in fact if $D_1 - A = A - D_2$ then $\nabla_{D_1 - A} s(A) = -\nabla_{D_2 - A} s(A)$, see also Herron [45]). By contrast, the derivative at the closed surface requires $\nabla_{D_1 - A} s(A) = \nabla_{D_2 - A} s(A)$. One might think that a local parameterization solves the problem, but a parameterization of patch (A, B, C) as shown in the Figure 7, determine $\nabla_{B-A} s(A)$ and $\nabla_{C-A} s(A)$ which imply the parameterization for patch (A, B, D) as shown, which in turn determines $\nabla_{D-A} s(A)$. But the parameterization of (A, B, C) also implies the parameterization for patch (A, C, D) as shown. The resulting $\nabla_{D-A} s(A)$ conflicts with the previous one. Thus, in both the global and the local parameterization the derivative is

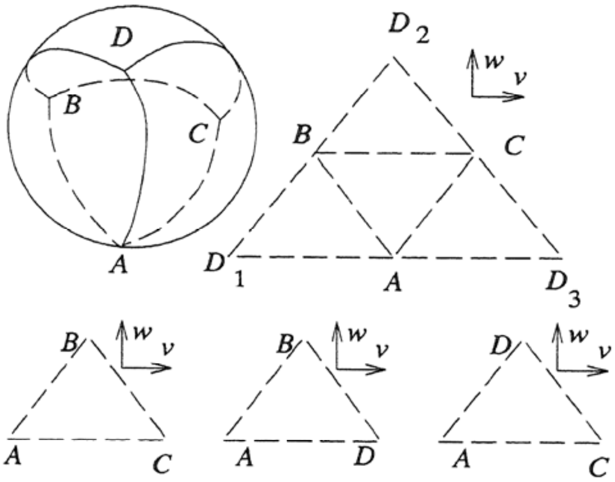

Figure 7: Closed surface (upper left) with a global parameterization (upper right), and a partial local parameterization (lower row).

not properly defined, while the tangent plane may still be coincident at common patch boundaries.

As for curves, C^n-continuity of surfaces is projectively invariant, but C^n-continuity of the homogeneous surface is not necessary for C^n-continuity of the corresponding rational surface. However, a result analogous to Theorem 1 has not been derived for surfaces.

Again we can take an algebraic and a differential geometry approach in order to base the notion of continuity on intrinsic aspects of the surface.

3.3 Geometric continuity

Analogous to curves, geometric continuity can be defined as follows [20]:

DEFINITION 15 (GEOMETRIC CONTINUITY) *Two surfaces $r(t,u)$ and $s(v,w)$ are G^n-continuous at (t_0,u_0) and (v_0,w_0) if and only if there exist a reparameterization $t = t(\tilde{t},\tilde{u})$, and $u = u(\tilde{t},\tilde{u})$, such that $\tilde{r}(\tilde{t},\tilde{u}) = r(t(\tilde{t},\tilde{u}),u(\tilde{t},\tilde{u}))$ and $s(v,w)$ are C^n-continuous at $r(t_0,u_0)$ and $s(v_0,w_0)$.*

Note that the reparameterization may be different at another point $s(v_1,w_1)$, otherwise we would again run into trouble with closed surfaces.

Since this definition is non-constructive, we apply the bivariate chain and product rule to the reparameterization, and observe that $\tilde{r}^{(k,l)}(\tilde{t},\tilde{u})$ can be written in terms of $\partial^{i+j}r(t,u)/\partial t^i\partial u^j$, $\partial^{i+j}t(\tilde{t},\tilde{u})/\partial\tilde{t}^i\partial\tilde{u}^j$, and $\partial^{i+j}u(\tilde{t},\tilde{u})/\partial\tilde{t}^i\partial\tilde{u}^j$, $i+j = k+l$, and $i+j = 1,\ldots,n$. For example

$$\frac{\partial\tilde{r}}{\partial\tilde{t}} = \frac{\partial r}{\partial t}\frac{\partial t}{\partial\tilde{t}} + \frac{\partial r}{\partial u}\frac{\partial u}{\partial\tilde{t}}.$$

Letting $\beta^t_{i,j}$ denote $\partial^{i+j}t/\partial^i\tilde{t}\partial^j\tilde{u}$ at $(\tilde{t}_0,\tilde{u}_0)$, and $\beta^u_{i,j}$ the analogue for $u(\tilde{t},\tilde{u})$, we get the β-constraints for surfaces:

$$
\begin{aligned}
&(15a) && s^{(0,0)} = r^{(0,0)} \\
&(15b) && s^{(1,0)} = \beta^t_{1,0} r^{(1,0)} + \beta^u_{1,0} r^{(0,1)} \\
&(15c) && s^{(0,1)} = \beta^t_{0,1} r^{(1,0)} + \beta^u_{0,1} r^{(0,1)} \\
&(15d) && s^{(2,0)} = (\beta^t_{1,0})^2 r^{(2,0)} + 2\beta^t_{1,0}\beta^u_{1,0} r^{(1,1)} + (\beta^u_{1,0})^2 r^{(0,2)} + \beta^t_{2,0} r^{(1,0)} + \beta^u_{2,0} r^{(0,1)} \\
&(15e) && s^{(1,1)} = \beta^t_{1,0}\beta^t_{0,1} r^{(2,0)} + (\beta^t_{1,0}\beta^u_{0,1} + \beta^t_{0,1}\beta^u_{1,0}) r^{(1,1)} + \beta^u_{1,0}\beta^u_{0,1} r^{0,2} + \beta^t_{1,1} r^{(1,0)} + \beta^u_{1,1} r^{(0,1)} \\
&(15f) && s^{(0,2)} = (\beta^t_{0,1})^2 r^{(2,0)} + 2\beta^t_{0,1}\beta^u_{0,1} r^{(1,1)} + (\beta^u_{0,1})^2 r^{(0,2)} + \beta^t_{0,2} r^{(1,0)} + \beta^u_{0,2} r^{(0,1)} \\
& && \vdots
\end{aligned}
$$

where it is understood that the derivatives of r and s are taken at $(t(\tilde{t}_0, \tilde{u}_0), u(\tilde{t}_0, \tilde{u}_0))$ and $s(v_0, w_0)$ respectively. Note that for continuity along a whole curve these equations must hold for all points along the curve but that the β-s need not be constant.

This set of equations is not put into matrix form in a straightforward manner. The β-s provide $n(n+3)$ shape handles for G^n-continuity. Of course if $\beta^t_{1,0} = \beta^u_{0,1} = 1$ and $\beta^t_{i,j} = \beta^u_{j,i} = 0$, $i \neq 1$ and $j \neq 0$ amounts to parametric continuity.

The reparameterization approach is easily extended to k-variate surfaces in $\mathbb{R}^d$, $d > k$. Geometric continuity of k-dimensional surfaces is applied in Section 2.6 to k-tangent surfaces.

As for curves, G^n-continuity of surfaces is preserved under arbitrary C^n diffeomorphisms [64] (C^n mappings that have a C^n inverse), particularly under projective transformations. However, a result analogous to Theorem 4 that would state the exact conditions on a homogeneous surface so as to let the rational surface be G^n has not been derived.

3.4 Differential geometry approach

Just as we considered the tangent line of a curve we now consider the tangent plane of a surface.

The tangent plane at $s(v_0, w_0)$ is spanned by the vectors $s^{(1,0)}(v_0, w_0)$ and $s^{(0,1)}(v_0, w_0)$. Equivalently, the tangent plane is normal to the surface normal vector

$$
N(v_0, w_0) = \frac{s^{(1,0)}(v_0, w_0) \times s^{(0,1)}(v_0, w_0)}{\|s^{(1,0)}(v_0, w_0) \times s^{(0,1)}(v_0, w_0)\|}, \tag{16}
$$

where '$\times$' denotes the vector or cross product. The tangent planes at $r(t_0, u_0)$ and $s(v_0, w_0)$ coincide if and only if $r^{(1,0)}(t_0, u_0)$, $r^{(0,1)}(t_0, u_0)$, $s^{(0,1)}(v_0, w_0)$, and $s^{(0,1)}(v_0, w_0)$ are coplanar, that is, if Equations (15b) and (15c) hold.

Theorem 9 (Geometric continuity) *Two surfaces are first order geometric continuous at a common point if and only if their normal vectors coincide.*

Second order geometric continuity is based on curvature. For any direction d in the tangent plane at $s(v_0, w_0)$, the plane through d and $N(v_0, w_0)$ intersects $s(v, w)$ in a curve. The normal curvature of this curve is the normal curvature of the surface in the direction of d: $\kappa_d(v_0, w_0)$. Unless $\kappa_d(v_0, w_0)$ is the same in all directions, there are two directions d_1 and d_2 in which $\kappa_d(v_0, w_0)$ takes the maximum and minimum values: the principal curvatures $\kappa_1(v_0, w_0)$ and $\kappa_2(v_0, w_0)$ respectively.

Remark 9 The *first* curvature at $s(v_0, w_0)$ is defined as $\kappa_1(v_0, w_0)+\kappa_2(v_0, w_0)$. The *Gauss* or *second* curvature is $\kappa_1(v_0, w_0)\kappa_2(v_0, w_0)$. The *mean* normal curvature is $(\kappa_1(v_0, w_0) + \kappa_2(v_0, w_0))/2$. The *amplitude* of the normal curvature is $(\kappa_1(v_0, w_0) - \kappa_2(v_0, w_0))/2$.

Second order geometric continuity can be defined in terms of the so called Dupin indicatrix based on the principal curvatures.

DEFINITION 16 (DUPIN INDICATRIX) *Taking the principal directions (d_1, d_2) as coordinate axes, the Dupin indicatrix is the set of points (x_1, x_2) defined by the conic section:*

$$
\begin{aligned}
\kappa_1 x_1^2 + \kappa_2 x_2^2 &= 1 && \text{if } \kappa_1\kappa_2 > 0, \\
\kappa_1 x_1^2 + \kappa_2 x_2^2 &= \pm 1 && \text{if } \kappa_1\kappa_2 < 0, \\
\kappa_1 x_1^2 &= 1 && \text{if } \kappa_2 = 0.
\end{aligned} \tag{17}
$$

So, the Dupin indicatrix at a point $s(v_0, w_0)$ is a conic lying in the tangent plane at that point. If κ_1 and κ_2 have the same sign, the Dupin indicatrix is an ellipse. If κ_1 and κ_2 have opposite signs, it consists of two hyperbolas, if $\kappa_1 \neq 0$ and $\kappa_2 = 0$, it degenerates into a pair of parallel lines, and if $\kappa_1 = \kappa_2 = 0$ it is void. The Dupin indicatrix fully characterizes the curvature at $s(v_0, w_0)$.

Remark 10 Point $s(v_0, w_0)$ is called a synclastic or elliptic point if $\kappa_1\kappa_2 > 0$, an anticlastic or hyperbolic point if $\kappa_1\kappa_2 < 0$, a parabolic point if $\kappa_1 \neq 0$ and $\kappa_2 = 0$, and a planar point if $\kappa_1 = \kappa_2 = 0$.

THEOREM 10 (GEOMETRIC CONTINUITY) *Two surfaces are second order geometric continuous if and only if their Dupin indicatrices coincide.*

Arc length doesn't apply to a surface as a whole, but of course it *does* apply to curves on the surface. A normalized parameterization similar to arc length parametrization for curves, is based on the notion of a line of curvature:

DEFINITION 17 (LINE OF CURVATURE) *A line of curvature on a surface is a regular curve such that at all points on the curve the tangent vector has a principal direction.*

The following surface parameterization is of interest: lines of constant parameter value are lines of curvature and are arc length parameterized. If two surfaces are parameterized in that way, and they are C^n $(n = 1, 2)$, then they are also G^n. The reverse is only conjectured [76]:

CONJECTURE 1 *Two surfaces $r(t, u)$ and $s(v, w)$ are G^n-continuous, $n = 1, 2$, at (t_0, u_0) and (v_0, w_0) if the corresponding surfaces $\tilde{r}(\tilde{t}, \tilde{u})$ and $\tilde{s}(\tilde{v}, \tilde{w})$ whose lines of constant parameter value are lines of curvature that are arc length parameterized, satisfy $\partial^n \tilde{r}/\partial \tilde{t}^n = \partial^n \tilde{s}/\partial \tilde{v}^n$ and $\partial^n \tilde{r}/\partial \tilde{u}^n = \partial^n \tilde{s}/\partial \tilde{w}^n$, $n = 1, 2$, at $(\tilde{t}_0, \tilde{u}_0)$ and $(\tilde{v}_0, \tilde{w}_0)$.*

Note that for $n = 2$ the condition on the derivatives does not amount to C^2-continuity, since the mixed partial derivatives $\partial^2 \tilde{r}/\partial \tilde{t} \partial \tilde{u}$ and $\partial^2 \tilde{s}/\partial \tilde{v} \partial \tilde{w}$ need not agree.

For order higher than two there seems to be no notion of continuity based of differential geometry.

A particular but frequently used parameterization is that two patches r and s agree along a *constant* value of one parameter, for example $r(t_0, u) = s(v, w_0)$. The patches have coincident tangent planes at every point of their common curve if the three vectors $s^{(0,1)}(v_0, w_0)$, $r^{(1,0)}(t_0, u_0)$, and $r^{(0,1)}(t_0, u_0)$ are coplanar, see Figure 8. That is the case, if

$$
s^{(0,1)}(v_0, w_0) = \alpha(u) r^{(1,0)}(t_0, u_0) + \gamma(u) r^{(0,1)}(t_0, u_0), \tag{18}
$$

where of course u_0 and w_0 are such that $r(t_0, u_0) = s(v_0, w_0)$, and $\alpha(u)$ does not vanish. Indeed, along the common curve $r(t_0, u) = s(v, w)$, the β-s in Equation (15b) depend only on u, so that Equation (15b) reduces to Equation (18). For a smooth transition from one surface into the other, it is required that $\alpha(u) < 0$, otherwise a sharp edge results. An alternative formulation to Equation (18) is the following:

$$
\delta(v) s^{(1,0)}(v_0, w_0) = \alpha(u) r^{(1,0)}(t_0, u_0) + \gamma(u) s^{(0,1)}(t_0, u_0),
$$

where $\alpha(u)$ and $\delta(v)$ do not vanish. Special choices of $\alpha(u)$, $\gamma(u)$, and $\delta(v)$ give special constructions of tangent plane continuity; see Du & Schmitt [25] for a review.

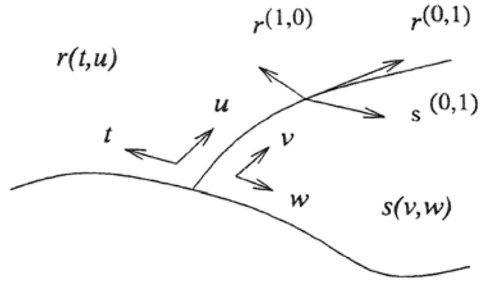

Figure 8: A particular parameterization such that $r(t_0, u) = s(v, w_0)$.

Kahmann [50] further shows that the patches have coincident Dupin indicatrices if

$$s^{(0,2)} = \alpha^2 r^{(0,2)} + 2\alpha\gamma r^{(1,1)} + \gamma^2 r^{(2,0)} + \varepsilon r^{(1,0)} + \zeta r^{(0,1)}, \tag{19}$$

where ε and ζ are, like α and γ, functions of u. Compare Equations (19) and (15f).

Continuity of tangent plane and Dupin indicatrix were discussed in [70, 74, 50].

For general rational surfaces no continuity conditions have been derived, but for Bézier surfaces G^1 conditions have been elaborated [56, 21].

3.5 Ruled tangent surface continuity

For curves we used k-dimensional osculating linear spaces as the rulings that are swept along a curve so as to generate a $k+1$-dimensional ruled tangent surface. For surfaces, we can use k-dimensional osculating linear spaces as rulings that are swept along the surface so as to form a $k+2$-dimensional ruled tangent surface which we can subject to continuity constraints. However, in 3-space and 4-space any $k+2$-tangent surface is trivially continuous. This is similar to planar curves, whose $k+1$-tangent surfaces are also trivially continuous.

Geometric continuity, which is based on reparameterization, is properly defined in all these cases.

4 Modeling

Curve and surface design is typically done in a piecewise way. After all the segments have been defined, they must be properly put together to construct a 'smooth' curve or surface.

If a curve is C^k then it is also G^k, but the reverse is not true. However, G^k curves are not less smooth than C^k ones. A G^k curve can be reparameterized without altering its shape so that it will be C^k. Thus geometric continuity is a relaxation of parameterization but not of smoothness. This relaxation results in several degrees of freedom, for example the β-parameters. Such parameters are called shape parameters or shape handles. They can be used to modify the shape without destroying the imposed continuity conditions. These shape handles have proved to be useful in interactive design of curves and surfaces.

4.1 Curve modeling

A particular G^2 (curvature continuous) cubic spline is the ν-spline, introduced in 1974 by Nielson [59]. The ν-spline is expressed in a global parameter $v_{min} \leq v \leq v_{max}$, with respect to which it is C^1, and interpolates given data points at the knots. Considering two segments $r(v)$, $v_{i-1} \leq v \leq v_i$ and $s(v)$, $v_i \leq v \leq v_{i+1}$, the ν-spline satisfies

$$s(v_i) = r(v_i), \tag{20a}$$

$$s^{(1)}(v_i) = r^{(1)}(v_i), \tag{20b}$$

$$s^{(2)}(v_i) = r^{(2)}(v_i) + \nu_i r^{(1)}(v_i). \tag{20c}$$

The breakpoints (20a) are prescribed. The tangent vectors (20b) are uniquely determined by the interpolation requirement. The ν_i are constant shape parameters, affecting how tight the spline runs through the breakpoints; they are therefore called tension parameters. These shape parameters can be used to alter the shape of the curve.

The γ-spline is a G^2 (curvature continuous) cubic spline in a special formulation involving control points of both the B-spline and the Bézier curve representation at the same time, developed by Böhm [11]. Like the ν-spline, the γ-spline is expressed in a global parameter $v_{min} \le v \le v_{max}$, with respect to which it is C^1. At the knots v_i there are weight factors γ_i defined in terms of ratios of distances between control points. The relation between γ_i and ν_i is given by

$$\nu_i = 2\left(\frac{1}{v_i - v_{i-1}} + \frac{1}{v_{i+1} - v_i}\right)\left(\frac{1}{\gamma_i + 1}\right)$$

If the knot sequence is rescaled such that each interval has unit length, the curve becomes G^1 instead of C^1, with the β_1 from Equation (5) defined at knot v_i as $\beta_{1,i} = (v_{i+2} - v_{i+1})/(v_{i+1} - v_i)$. The resulting spline is the β-spline. Note that the β-spline is only one particular spline among all splines satisfying Equation (5). In the formulation of the β-spline, the $\beta_{2,i}$ are equivalent to ν_i. Setting all $\beta_{1,i} = \beta_1$ and $\beta_{2,i} = \beta_2$ yields the β-spline as originally introduced by Barsky [3, 4].

The τ-spline is a quintic F^3 (torsion) continuous spline, introduced by Hagen [41]. The τ-spline is F^3 but generally not G^3, while the β-, γ-, and ν-spline are both G^2 and F^2 since G^2 is equivalent to F^2. We do not need a curve of degree five for torsion continuity: a quartic F^3 spline is given by Böhm [12].

A Catmull-Rom spline [17] is defined as

$$s(v) = \sum_{i}^{k} B_i(v) r_i(v), \tag{21}$$

where $r_i(v)$ are curves with the same control points; they have in general different local supports and interpolate the control points that correspond to their supports. The blending functions B_i do not only weight the control points but the entire curves.

When r_i and B_i are parametrically continuous, s is also parametrically continuous. Moreover, when r_i and B_i are all G^n for the same set of β-values $(\beta_0, \beta_1, \dots, \beta_n)$, s is also G^n with these β-values. However, if r_i and B_i are all F^n for the same connection matrix N from Equation (5), the Catmul-Rom spline need not be F^n for the same connection matrix. Whether s is Frenet frame continuous for another connection matrix is still an open problem [36].

4.2 Tensor product patches

A function f on $\mathbb{R}^2$ defined as $f(v,w) = g(v)h(w)$ for all $v, w \in \mathbb{R}$ whith g and h univariate functions, is called a tensor product of g and h [18].

Remark 11 The tensor product of g and h is often denoted $g \otimes h$.

Analogous to Equation (1), we can form a linear combination of tensor product functions:

$$f(v,w) = \sum c_i g_i(v) h_i(w).$$

A frequently used form of a tensor product surface is [18, 15]:

$$s(v,w) = \sum_{i=0}^{k} \sum_{j=0}^{l} a_{i,j} B_i^k(v) B_j^l(w), \tag{22}$$

compare to Equation (2). For a constant v_0, $s(v_0, w)$ is a curve on the surface depending on w only; a constant value for w gives a curve in v. Tensor product surfaces are frequently used because of

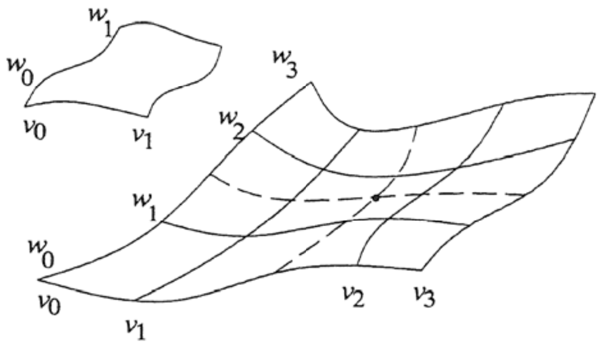

Figure 9: Tensor product patch, and rectangular grid of patches.

their simplicity: many operations such as subdivision or derivation result from the corresponding univariate operations in a straightforward manner.

The projection of a tensor product surface is often written in the following form:

$$S(v,w) = \frac{\sum_{i=0}^{k}\sum_{j=0}^{l} a_{i,j} B_i^k(v) B_j^l(w)}{\sum_{i=0}^{k}\sum_{j=0}^{l} w_{i,j} B_i^k(v) B_j^l(w)}, \tag{23}$$

where each $a_{i,j}$ is still a vector but the $w_{i,j}$ are scalar weight factors. It is a common misconception to call the surface of Equation (23) itself a tensor product [31]. In the following we discuss only tensor product surfaces (Equation (22)), and not the projection thereof.

If $B_i^k(v)$ and $B_j^l(w)$ are parametrically continuous, then so is $s(v,w)$. But, if the blending functions are only geometrically continuous, the tensor product surface need not be geometrically continuous. However, a sufficient condition is the following: if all the $B_i^k(v)$ are G^n at v_0 for the same set of β-values, and all $B_j^l(w)$ for another set of β-values at w_0, then $s(v_0, w_0)$ is also G^n-continuous [36].

When $s_1(v)$ and $s_2(w)$ are splines of multiple segments, we get a rectangular grid or mesh of patches, see Figure 9. When the curves are single segments, the tensor product is a single patch. The patches can be independently parameterized, and can placed in an arbitrary mosaic. Then the number of patches meeting at the corners of each patch is not restricted to four, as in the case of a rectangular mesh.

A rectangular mesh is not suitable for modeling all shapes. For example a sphere-like object can be covered by six rectangular patches but cannot be covered with a rectangular mesh of patches without degenerate patches. Degenerate rectangular patches have an edge of zero length, collapsing to a point. At that point the surface normal is not defined which causes problems to many standard intersection and shading algorithms. In shapes where n-sided patches naturally arise ($n \neq 4$) they can be modeled by n rectangular patches meeting at a common point, see Figure 10 for an example with a triangular patch, the 'suitcase corner', and a pentagonal patch.

So, for a modeling system based on rectangular patches to be general it must be able to handle an arbitrary mosaic of patches, allowing n patches the meet at a corner. For G^1-continuity around such a corner, all adjacent patches must satisfy Equation (18). For particular types of surfaces this poses constraints on its definition, for example on the control points in Equation (22).

In order to find constraints on certain control points we can differentiate Equation (18) [71, 75, 25]:

$$\frac{\partial}{\partial v} s^{(0,1)}(v_0, w_0) = \frac{\partial}{\partial u}\Big(\alpha(u) r^{(1,0)}(t_0, u_0) + \gamma(u) r^{(0,1)}(t_0, u_0)\Big). \tag{24}$$

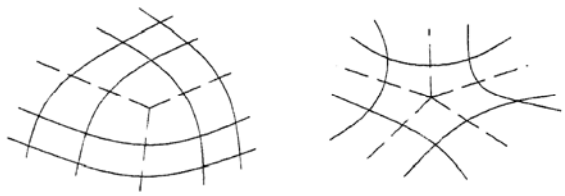

Figure 10: Non-rectangular meshes with 3- and 5-sided patch, subdivided into 3 and 5 rectangular patches meeting at a common corner.

If we have n patches at a corner numbered as in Figure 11, and evaluate Equation (24) at the corner we get the following set of equations:

$$s_{i+1}^{(1,1)} = \alpha_i s_i^{(1,1)} + \alpha_i^{(1)} s_i^{(1,0)} + \gamma_i s_i^{(0,2)} + \gamma_i^{(1)} s_i^{(0,1)} \tag{25}$$

or equivalently

$$\begin{pmatrix} -\alpha_1 & 1 & 0 & \cdots & 0 \\ 0 & -\alpha_2 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & & \ddots & -\alpha_{n-1} & 1 \\ 1 & 0 & \cdots & 0 & -\alpha_n \end{pmatrix} \begin{pmatrix} s_1^{(1,1)} \\ s_2^{(1,1)} \\ \vdots \\ s_{n-1}^{(1,1)} \\ s_n^{(1,1)} \end{pmatrix} = \begin{pmatrix} \alpha_1^{(1)} s_1^{(1,0)} + \gamma_1 s_1^{(0,2)} + \gamma_1^{(1)} s_1^{(0,1)} \\ \vdots \\ \alpha_n^{(1)} s_n^{(1,0)} + \gamma_n s_n^{(0,2)} + \gamma_n^{(1)} s_n^{(0,1)} \end{pmatrix}. \tag{26}$$

The determinant of the first matrix is given by

$$\prod_{i=1}^{n} -\alpha_i - (-1)^n.$$

Recall that the α_i are negative, so that the determinant is positive if n is odd. The matrix is singular if n is even and $\prod -\alpha_i = 1$. In that case Equation (25) cannot be used to derive constraints on control points or other terms in the patch definition. It does not mean that Equation (25) is necessarily inconsistent, since the equations can as well be dependent, allowing infinitely many solutions.

Triangular patches provide an alternative to the ubiquitous tensor product schemes. Especially Bézier triangles[29], interpolating prescribed vertices, have gained much attention. Transfinite triangular schemes, interpolating prescribed boundaries, have also been developed for G^1-continuity[60] and G^2-continuity[42]. General n-sided patches are typically subdivided into rectangular patches[49], see also Section 4.2.

Both rectangular and triangular Bézier schemes have gained much attention for developing continuity conditions because their control points are geometrically meaningful. G^1-continuity conditions on the Bézier control points have been derived[27, 63, 55, 23, 24, 21, 61] but not all of these apply to n patches around a common vertex. G^2-continuity conditions have been elaborated too[50, 19].

Also for rational Bézier surfaces G^1 conditions are derived [56, 21].

5 Visual aspects of continuity

Rendering of curves and surfaces is usually limited to curves in the plane or 3-space and surfaces in 3-space, although higher dimensional applications such as scientific visualization [44] and multivariate analysis [1] for which higher order continuity apply are becoming important.

Manning [58] described geometric continuity conditions for curves with application to insole shape design. G^2-continuity seems to be sufficient for a smooth appearance of the curve. A

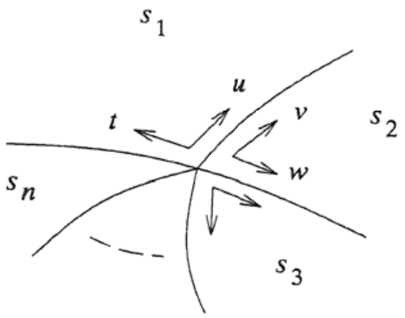

Figure 11: n patches meeting at a corner.

discontinuity in the curvature can be detected by a practised eye, but it seems that a discontinuity in the torsion is not visible.

The visual aspects of surfaces are much more complicated, since these depend on the surface contour, texture, contrast, and illumination, see also Koenderink[53]. In this paper I survey some aspects of illumination. First, the two following subsections treat simple concepts of illumination models and shading algorithms. Then they are related to surface continuity in Section 5.3. Much more about illumination models and shading algorithms can be found in the text books by Glassner[35] and Hall[43].

5.1 Illumination models

Many illumination models that are used in computer graphics for the rendering of surfaces depend on the surface normal $N(v_0, w_0)$ defined by Equation (16). Local illumination models describe how an imaginary ray of light from a point light source in the direction L with intensity I_{light} is scattered from an idealized surface. Diffuse reflection scatters light in all directions with equal intensity, which is proportional to the cosine of the angle θ between N and L, see Figure 12. The Lambert or diffuse illumination equation, based on Lambert's cosine law is

$$I_{pixel} = I_{light} k_d \cos\theta = I_{light} k_d (N \cdot L), \tag{27}$$

where k_d is the fraction of diffuse light reflected which depends on the surface material and is mostly kept constant, and the vectors are normalized. Note that N varies along the surface; L also varies, and is only constant when the light source is at infinity.

A mirror reflects light from direction L only in the direction R, see Figure 12. Glossy surfaces exhibit specular reflection by scattering light unequally in different directions. The empirical

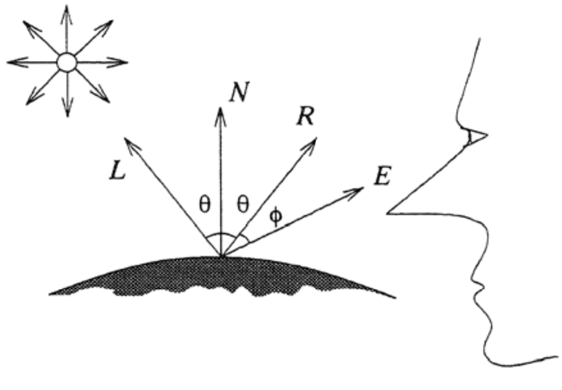

Figure 12: Point light source reflection geometry.

model for the specular reflection of glossy surfaces described by Phong [62] assumes that maximum specular reflection occurs when the angle ϕ between R and the eye vector E is zero, and falls off sharply as ϕ increases. This falloff is modeled by $\cos^n \phi$ where n is the specular reflection exponent depending on the surface material. Phong combines the diffuse and specular reflection giving

$$\begin{aligned} I_{pixel} &= I_{light}(k_d \cos\theta + k_s \cos^n \phi) \\ &= I_{light}(k_d N \cdot L + k_s (R \cdot E)^n) \end{aligned} \tag{28}$$

where k_s is the fraction of specular light reflected which depends on the surface material and is mostly kept constant, and the vectors are normalized. Note that R varies along the surface. E also varies, except when the eye is at infinity.

The empirical model described by Phong has been improved and enhanced in several ways, so that local reflection models produce realistic looking objects. However, inter-object illumination effects such as one object casting a shadow on another or reflecting light towards another cannot be captured by only a ray of light and a single surface. These global effects require other algorithms. In 1980, Whitted [77] introduced an improved illumination model that captures diffuse and specular reflection as well as transmission through transparent objects, and inter-object illumination by *specular* reflection. Another new algorithm, radiosity, was introduced in computer graphics in 1984 by Goral et al. [38], to model the energy balance for perfectly diffuse surfaces. Radiosity gives softer shadowing and inter-object illumination by *diffuse* reflection.

However, in relation to surface continuity the prime test case is the reflection of a linear source of light. Specular reflection of a linear light source gives a reflection curve on the surface. Reflection curves are used in for example car industry to test surface smoothness[51]. In Section 5.3 is shown why reflection curves are interesting.

Linear light sources can be modeled as follows:

$$\begin{aligned} I_{pixel} &= \int I_{light}(k_d N \cdot L + k_s (R \cdot E)^n) dl \\ &= I_{light} k_d \int N \cdot L \, dl \; + \; I_{light} k_s \int (R \cdot E)^n \, dl, \end{aligned} \tag{29}$$

where the integral is taken over the length of the linear light, and the intensity I_{light} is assumed to be constant along the line.

5.2 Shading algorithms

A polygonal C^0 surface has a constant normal at each facet so that diffuse reflection is constant if the light source is at infinity. Rendering a polygonal surface in this way is called flat or constant shading.

Any G^k, $k \geq 1$ surface can be rendered by calculating the normal at every point on the surface for use in any illumination model, but this is computationally expensive. The surface is often approximated by a polygonal C^0 surface obtained by adaptive subdivision of patches into several patches that together have the same shape as the parent patch, until the polygonal approximation of the patch is within a certain error bound from the real surface. More general shading algorithms can handle patches bounded by curved segments [54]. The approximating surface can be displayed by flat shading but this gives clear intensity discontinuities.

The surface can also be rendered by interpolation algorithms. The Gouraud shading algorithm requires that the normal is known at each vertex of the polygonal mesh. Vertex intensities can be found by any desired illumination model, but originally Gouraud [39] adopted a Lambertian diffuse illumination model like Equation (27). The Gouraud shading algorithm is a scan line method: each polygon is shaded by linear interpolation of vertex intensities along each edge and then between edges along each scan line. Gouraud shading is implemented in hardware or firmware in many graphics workstations.

The Phong shading algorithm is also a scan line method. It interpolates the surface normal, eye, light, and reflection vectors, instead of the light intensity. Again the vectors at the vertices must be

Figure 13: Cause of Mach band effect: dashed line is true intensity with dicontinuous derivative, solid line is the perceived intensity.

known. These vectors are interpolated along the edges and then between edges along each scan line. At each pixel along the scan line the light intensity is calculated using the corresponding vectors and any desired illumination model, but originally Phong [62] adopted a specular illumination model similar to Equation (28). Phong shading is implemented in many graphics workstation system software.

The illumination model of Whitted [77] asked for a new shading algorithm, called ray tracing. The algorithm traces a ray backwards form the eye through a pixel on the imaginary screen into the object space and towards the light sources. Ray tracing gives better results than Phong shading, but a linear light source is usually modeled by a series of collinear point sources. Many points lights are needed for a good approximation. Alternatively, a linear light can be modeled as an area in order to apply the radiosity approach. However, radiosity models perfectly diffuse surfaces. Specular reflection (what we are interested in) is thus not considered.

Using a linear light source model (Equation (29)) in a scan line method requires the evaluation of the integral at each pixel along a scan line. Recently Poulin and Amanatides [65] have shown that the diffuse integral term of Equation (29) can be solved analytically, and they approximate the specular integral term by Chebyshev polynomials.

5.3 Visibility of surface discontinuity

Flat shading of a C^0 polygonal surface gives clear color intensitiy discontinuities, which is correct.

Gouraud and Phong shading use a polygonal surface as an approximation of a smooth surface, and try to give and impression of smoothness by a continuously varying color intensity. However, when the intensity changes continuously between two surface patches, but the derivative of the intensity does not, one perceives the well known Mach bands, discovered by Mach in 1865 [66, 67]. The Mach band effect is the perception of a darker band on the dark patch and a lighter band on the light patch along the border between the patches. The perception of Mach bands is caused by our visual system that exaggerates intensity changes where the actual intensity or intensity derivative exhibits a discontinuity, see Figure 13.

So, the Gouraud shading algorithm creates visual artifacts due to the approximation of curved objects by planar facets. Although the intensity changes continuously, Mach bands appear because of the discontinuous derivative of the intensity across the edges, since the interpolation is only linear.

Also when interpolating vectors instead of intensities, as in Phong shading, the computed intensity need not be C^1-continuous. So Phong shading also suffers from Mach banding, be it much less than Gouraud shading. Both algorithms make another error however: interpolation is performed in image space instead of object space. This is more serious for Phong shading since the specular highlights are displaced.

Let us see how continuity between surface patches affects the specular reflection of a linear light source. Given a linear light source $LL(t)$, the shape of the reflection curve $LL^*(t)$ depends on the shape of the surface. If the surface is G^1-continuous the tangent vector of $LL^*(t)$ at both sides of the patch boundary lie in the surface tangent plane, but need not be collinear, see Figure 14. So, G^1 surfaces generally show a reflection line that is not G^1-continuous.

Conversely, let us assume that $LL^*(t)$ *is* G^1-continuous, as well as the surface itself. Let V be the view point (position of the eye) such that the angle between E and R is zero, then

$$E(t) = \frac{V - LL^*(t)}{\|V - LL^*(t)\|}$$

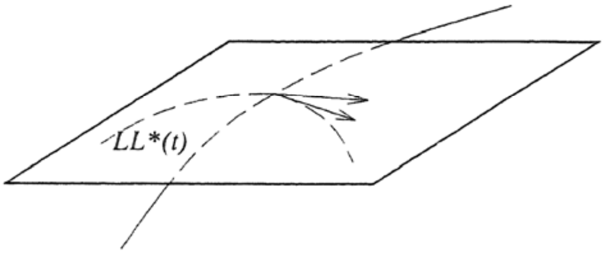

Figure 14: G^1 surface with a reflection curve $LL^*(t)$ that is *not* G^1.

and

$$L(t) = \frac{LL(t) - LL^*(t)}{||LL(t) - LL^*(t)||}.$$

A necessary and sufficient condition for a point $LL^*(t)$ to be a reflection point is (see Figure 15):

$$N(t) \cdot (E(t) - L(t)) = 0, \tag{30}$$

where $N(t)$ is the *surface* normal along the reflection curve, and '$\cdot$' the dot or scalar product of two vectors. Differentiating of Equation (30) with respect to t gives

$$N^{(1)}(t) \cdot (E(t) - L(t)) + N(t) \cdot (E^{(1)}(t) - L^{(1)}(t)) = 0. \tag{31}$$

Since the surface is G^1-continuous, $N(t)$ is continuous, and since we assume that $LL^*(t)$ is G^1-continuous, $E(t)$, $E^{(1)}(t)$, $L(t)$, and $L^{(1)}(t)$ are also continuous. Therefore, Equation (31) only holds when $N^{(1)}(t)$ is continuous. This implies that the surface curvature is continuous. So, a G^1 reflection curve implies a G^2-continuous surface.

Note that the surface can thus be inspected visually without computing explicitly curves of constant curvature along the surface, or other surface features, see Higashi et al. [47] Incorporating a linear light source into the illumination model gives surface continuity information for free. Bear in mind however, that algorithms interpolating the surface normal vector, like Phong shading, may introduce artifacts again.

6 Conclusions

This paper has given an overview of notions of continuity for curves and surfaces: the classical parametric continuity (C^n) and alternative notions of continuity based on reparameterization (G^n) and the differential geometry approaches based on Frenet frames (F^n) and ruled tangent surfaces (T^n) have been treated. In addition to other surveys, overviews, and tutorials, this paper has given explicit expressions for the β-constraints for surfaces in addition to those for curves.

I have shown that derivatives at closed surfaces exhibit singularities for piecewise parameterization in addition to the case for global parameterization, so that parametric continuity is not even properly defined More important however, other notions of continuity give more freedom in modeling curves and surfaces, in some cases giving intuitive shape parameters.

Finally I have related geometric continuity to illumination models and shading algorithms, and have shown in particular that a reflection curve is G^1 if the surface is G^2, but generally not G^1 if the surface is only G^1-continuous. Perception of (dis)continuity of a surface depends not only on the surface itself but also on the illumination, in a computer model as well as in real. The term 'visual continuity' for what has been defined as G^n-continuity is therefore inappropriate.

Acknowledgement

I like to thank Miente Bakker and Edwin Blake for discussions and careful reading of this manuscript.

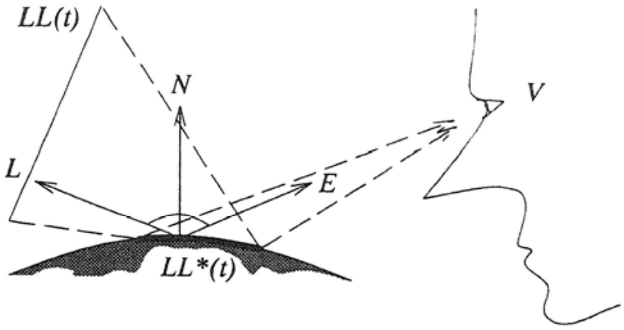

Figure 15: Linear light source reflection geometry.

References

[1] Peter Alfeld. Scattered data interpolation in three or more variables. In Lyche and Schumaker [57], pages 1 – 33.

[2] Robert E. Barnhill and Richard F. Riesenfeld, editors. *Computer Aided Geometric Design.* Academic Press, 1974.

[3] Brian A. Barsky. *The β-Spline: A Local Representation Based on Shape Parameters and Fundamental Geometric Measures.* PhD thesis, University of Utah, Salt Lake City, 1981.

[4] Brian A. Barsky. *Computer Graphics and Geometric Modeling Using β-Splines.* Springer-Verlag, 1988.

[5] Brian A. Barsky and John C. Beatty. Local control of bias and tension in beta-splines. *ACM Transactions on Graphics*, 2(2):109 – 134, April 1983.

[6] Brian A. Barsky and Tony D. DeRose. Geometric continuity of parametric curves: Three equivalent characterizations. *Computer Graphics & Applications*, 9(6):60 – 68, November 1989.

[7] Brian A. Barsky and Tony D. DeRose. Geometric continuity of parametric curves: Constructions of geometrically continuous splines. *Computer Graphics & Applications*, 10(1):60 – 68, January 1990.

[8] Richard H. Bartels, John C. Beatty, and Brian A. Barksy. *An Introduction to Splines For Use in Computer Graphics and Geometric Modeling.* Morgan Kaufmann Publishers Inc., 1987.

[9] S. Bernstein. Démonstration de théorème de Weierstrass fondeé sur le calcul des probabilités. *Harkov Soobs. Matem ob-va*, 13(1–2), 1912.

[10] Pierre E. Bézier. *Numerical Control — Mathematics and Applications, translation by A. Robin Forrest and Anne F. Pankhurst of Emploi des machines á commande numérique, Masson et Cie, Paris, 1970.* John Wiley & Sons, 1972.

[11] Wolfgang Böhm. Curvature continuous curves and surfaces. *Computer Aided Geometric Design*, 2:313 – 323, 1985.

[12] Wolfgang Böhm. Smooth curves and surfaces. In Farin [30], pages 175 – 184.

[13] Wolfgang Böhm. On the definition of geometric continuity. *Computer-Aided Design*, 20(7):370 – 372, 1988.

[14] Wolfgang Böhm. Visual continuity. *Computer-Aided Design*, 20(6):307 – 311, 1988.

[15] Wolfgang Böhm, Gerald Farin, and Jürgen Kahmann. A survey of curve and surface methods in CAGD. *Computer Aided Geometruic Design*, 1(1):1 – 60, 1984.

[16] Malcolm S. Casale. Free-form surface modelling with trimmed surface patches. *IEEE Computer Graphics & Applications*, 7(1):33 – 43, January 1987.

[17] Edwin Catmull and Raphael Rom. A class of local interpolating splines. In Barnhill and Riesenfeld [2], pages 317 – 326.

[18] Carl de Boor. *A Practicle Guide to Splines.* Applied Mathematical Sciences 27. Springer-

Verlag, 1978.
[19] Wendelin L. F. Degen. Explicit continuity conditions for adjacent Bézier surface patches. *Computer Aided Geometric Design*, 7:181 – 189, 1990.
[20] Tony DeRose and Brian A. Barsky. An intuitive approach to geometric continuity for parametric curves and surfaces. In Nadia Magnenat-Thalmann and Daniel Thalmann, editors, *Computer-Generated Images – The State of the Art*, pages 159 – 175. Springer-Verlag, 1985.
[21] Tony D. DeRose. Necessary and sufficient conditions for tangent plane continuity of Bézier surfaces. *Computer Aided Geometric Design*, 7:165 – 179, 1990.
[22] Manfredo P. do Carmo. *Differential Geometry of Curves and Surfaces*. Prentice-Hall, 1976.
[23] Wen-Hui Du and Francis J. M. Schmitt. On the G^1 smooth connection between triangular Bernstein-Bézier patches, part I: Around an n-patch corner. Technical Report 89 D 012, Telecom Paris, France, October 1989.
[24] Wen-Hui Du and Francis J. M. Schmitt. On the G^1 smooth connection between triangular Bernstein-Bézier patches, part II: In a piecewise representation. Technical Report 89 D 013, Telecom Paris, France, October 1989.
[25] Wen-Hui Du and Francis J. M. Schmitt. On the G^1 continuity of piecewise bézier surfaces: a review with new results. *Computer Aided Design*, 22(9):556 – 573, November 1990.
[26] Nira Dyn and Charles A. Michelli. Piecewise polynomial spaces and geometric continuity of curves. *Numerische Mathematik*, 54:319 – 337, 1988.
[27] Gerald Farin. A construction for visual C^1 continuity of polynomial surface patches. *Computer Graphics and Image Processing*, 20(3):272 – 282, November 1982.
[28] Gerald Farin. Visually C^2 cubic splines. *Computer Aided Design*, 14(3):137 – 139, May 1982.
[29] Gerald Farin. Triangular Bernstein-Bézier patches. *Computer Aided Geometric Design*, 3(2):83 – 127, 1986.
[30] Gerald Farin, editor. *Geometric Modeling: Algorithms and New Trends*. SIAM, 1987.
[31] Gerald Farin. Rational curves and surfaces. In Lyche and Schumaker [57], pages 215 – 238.
[32] Gerald Farin. *Curves and Surfaces for Computer Aided Geometric Design*. Academic Press, 2 edition, 1990.
[33] A. H. Fowler and C. W. Wilson. Cubic spline, a curve fitting routine. Technical Report Corporate Report Y-1400, Union Carbide, 1966.
[34] G. Geise. Über berührende Kegelschnitte einer ebenen Kurve. *Zeitschrift für Angewandte Mathematik und Mechanik*, 42(7/8):297 – 304, July/August 1962.
[35] Andrew S. Glassner, editor. *An Introduction to Ray-Tracing*. Academic Press, 1989.
[36] R. N. Goldman and B. A. Barsky. On β-continuous functions and their application to the construction of geometrically continuous curves and surfaces. In Lyche and Schumaker [57], pages 299 – 311.
[37] R. N. Goldman and C. A. Michelli. Algebraic aspects of geometric continuity. In Lyche and Schumaker [57], pages 313 – 332.
[38] C. M. Goral, K. E. Torrance, and D. P. Greenburg. Modeling the interaction of light between diffuse surfaces. *Proceedings SIGGRAPH '84, Computer Graphics*, 18(3):213 – 222, 1984.
[39] Henri Gouraud. Continuous shading of curved surfaces. *IEEE Transactions on Computers*, C-20(6):623 – 628, 1971.
[40] John A. Gregory. Geometric continuity. In Lyche and Schumaker [57], pages 353 – 371.
[41] Hans Hagen. Geometric spline curves. *Computer Aided Geometric Design*, 2(1 – 3):223 – 227, 1985.
[42] Hans Hagen and Helmut Pottmann. Curvature continuous triangular interpolants. In Lyche and Schumaker [57], pages 373 – 384.
[43] R. Hall. *Illumination and Color in Computeter Generated Imagery*. Springer-Verlag, 1989.
[44] Andrew J. Hanson and Pheng A. Heng. Visualizing the fourth dimension using geometry and light. In *Proceedings of Visualization '91, San Diego, California*, pages 321 – 328. IEEE Computer Society Press, October 1991.
[45] G. Herron. Smooth closed surfaces with discrete triangular interpolants. *Computer Aided Geometric Design*, 2(4):297 – 306, 1985.
[46] Gary Herron. Techniques for visual continuity. In Farin [30], pages 163 – 174.

[47] Masatake Higashi, Takuya Kushimoto, and Mamoru Hosaka. On formulation and display for visualizing features and evaluating quality of free-form surfaces. In C. E. Vandoni and D. A. Duce, editors, *Eurographics '90*, pages 299 – 309. North-Holland, 1990.

[48] M. E. Hohmeyer and B. A. Barsky. Rational continuity: Parametric, geometric, and Frenet frame continuity of rational curves. *ACM Transactions on Graphics*, 8(4):335 – 359, October 1989.

[49] A. K. Jones. Nonrectangular surface patches with curvature continuity. *Computer Aided Design*, 20(6):325 – 335, 1988.

[50] Jürgen Kahmann. Continuity of curvature between adjacent Bézier patches. In Robert E. Barnhill and Wolfgang Böhm, editors, *Surfaces in Computer Aided Geometric Design*, pages 65 – 75. North-Holland, 1983.

[51] Reinhold Klass. Correction of local surface irregularities using reflection lines. *Computer-Aided Design*, 12(2):73 – 77, March 1980.

[52] Donald E. Knuth. *The Art of Computer Programming Volume 1: Fundamental Algorithms*. Addison-Wesley, 2nd edition, 1973.

[53] Jan J. Koenderink. *Solid Shape*. MIT Press, 1990.

[54] Jeffrey M. Lane, Loren C. Carpenter, Turner Whitted, and James F. Blinn. Scan line methods for displaying parametrically defined surfaces. *Communications of the ACM*, 23(1):23 – 34, January 1980.

[55] D. Liu and J. Hoschek. CG^1 continuity conditions between adjacent rectangular and triangular Bézier surface patches. *Computer Aided Design*, 21(4):194–200, May 1989.

[56] Dingyuan Liu. GC^1 continuity conditions between two adjacent rational Bézier surface patches. *Computer Aided Geometric Design*, 7:151 – 163, 1990.

[57] Tom Lyche and Larry L. Schumaker, editors. *Mathematical Methods in Computer Aided Geometric Design (conference held in Olso, Norway, 1988)*. Academic Press, 1989.

[58] J. R. Manning. Continuity conditions for spline curves. *The Computer Journal*, 17(2):181 – 186, May 1974.

[59] Gregory M. Nielson. Some piecewise polynomial alternatives to splines under tension. In Barnhill and Riesenfeld [2], pages 209 – 235.

[60] Gregory M. Nielson. A transfinite, visually continuous, triangular interpolant. In Farin [30], pages 235 – 246.

[61] Pia Pflüger and Marian Neamtu. Geometrically smooth interpolation by triangular Bernstein-Bézier patches with coalescent control points. In P. J. Laurent, A. Le Méhauté, and L. L. Schumaker, editors, *Curves and Surfaces*, pages 363 – 366. Academic Press, 1991.

[62] Bui Tuong Phong. Illumination for computer-generated pictures. *Communications of the ACM*, 18(6):311 – 317, June 1975.

[63] Bruce R. Piper. Visually smooth interpolation with triangular Bézier patches. In Farin [30], pages 221 – 233.

[64] Helmut Pottmann. Projectively invariant classes of geometric continuity for CAGD. *Computer Aided Geometric Design*, 6:307 – 321, 1989.

[65] Pierre Poulin and John Amanatides. Shading and shadowing with linear light sources. In C. E. Vandoni and D. A. Duce, editors, *Eurographics '90*, pages 377 – 386. North-Holland, 1990.

[66] Floyd Ratliff. *Mach Bands: Quantitative Studies on Neural Networks in the Retina*. Holden Day, San Francisco, 1965.

[67] Floyd Ratliff. Contour and contrast. *Scientific American*, 226(6):90 – 101, June 1972.

[68] Richard F. Riesenfeld. *Applications of B-spline Approximation to Geometric Problems of Computer-Aided Design*. PhD thesis, Department of Systems and Information Sciences, Syracuse University, Syracuse, New York, 1973.

[69] Malcolm A. Sabin. Parametric splines in tension. Technical Report Technical Report VTO/MS/160, British Aircraft Corporation, 1970.

[70] Malcolm A. Sabin. *The Use of Piecewise Forms for the Numerical Representation of Shape*. PhD thesis, Hungarian Academy of Sciences, Budapest, Hungary, 1976.

[71] Ramon F. Sarraga. G^1 interpolation of generally unrestricted cubic Bézier curves. *Computer*

Aided Geometric Design, 4(1):23 – 39 (Errata: 6(2) 1989, 167 –171), 1987.

[72] I. J. Schoenberg. Contributions to the problem of approximation of equidistant data by analytic functions. *Quaterly of Applied Mathematics*, 4:Part A 45 – 99, Part B 112 – 141, 1946.

[73] I. J. Schoenberg. *Cardinal Spline Interpolation.* SIAM, 1973.

[74] M. Veron, G. Ris, and J.-P. Musse. Continuity of biparametric surface patches. *Computer-Aided Design*, 8(4):267 – 273, October 1976.

[75] M. A. Watkins. Problems in geometric continuity. *Computer Aided Design*, 20(8), October 1988.

[76] Jörg Weber. *Methoden zur Konstruktion krümmungsstetiger Freiformflächen.* PhD thesis, Technischen Universität Berlin, 1990.

[77] Turner Whitted. An improved illumination model for shaded display. *Communications of the ACM*, 23(6):343 – 349, January 1980.