A variational approach to spline curves on surfaces

Helmut Pottmann and Michael Hofer*

Abstract

Given an m-dimensional surface Φ in $\mathbb{R}^n$, we characterize parametric curves in Φ, which interpolate or approximate a sequence of given points $\mathbf{p}_i \in \Phi$ and minimize a given energy functional. As energy functionals we study familiar functionals from spline theory, which are linear combinations of L^2 norms of certain derivatives. The characterization of the solution curves is similar to the well-known unrestricted case. The counterparts to cubic splines on a given surface, defined as interpolating curves minimizing the L^2 norm of the second derivative, are C^2; their segments possess fourth derivative vectors, which are orthogonal to Φ; at an end point, the second derivative is orthogonal to Φ. Analogously, we characterize counterparts to splines in tension, quintic C^4 splines and smoothing splines. On very special surfaces, some spline segments can be determined explicitly. In general, the computation has to be based on numerical optimization.

1 Introduction

Curve design using splines is one of the most fundamental topics in CAGD. B-spline curves possess a beautiful shape preserving connection to their control polygon. They allow us the formulation of efficient algorithms for processing, especially subdivision algorithms. Moreover, at least the curves of odd degree and maximal smoothness also arise as solutions of variational problems. There is a huge body of literature on these curves, and they received many generalizations [9]. In extension of the standard spline methods, variational curve design has been investigated in a large number of contributions (see [5, 13] and the references therein).

It is quite surprising that there seem to be relatively few contributions on the design of spline curves which are restricted to surfaces or more general surfaces (manifolds) in arbitrary dimensions. Of course, these curves can no longer be expressed as linear combinations of control points with suitable basis functions.

*Research Group 'Geometric Modeling and Industrial Geometry', Vienna University of Technology, Wiedner Hauptstr. 8-10, A-1040 Wien, Austria. E-mail: pottmann@geometrie.tuwien.ac.at, hofer@geometrie.tuwien.ac.at

Moreover, shape preserving schemes on manifolds, e.g. based on subdivision, are different from the solutions of variational problems. In the present paper, we will investigate the latter topic. We will study the functionals which are minimized by splines in affine spaces, but restrict the candidate curves to a given surface. The solution is not as simple and explicit as in the unrestricted case. However, we will find remarkable counterparts to known results.

Previous work

Let us briefly sketch the literature dealing with splines on surfaces. We are focussing here only on curves which have some shape design handles or are generated as solutions of variational problems.

In 1985, Shoemake [28] introduced spherical counterparts of Bézier curves. He extended de Casteljau's algorithm to the sphere, replacing straight line segments by geodesic arcs and replacing ratios of Euclidean distances by ratios of geodesic distances. Using this on the 3-sphere in $\mathbb{R}^4$, he generated the spherical component of rigid body motions and applied it to Computer Animation. Soon, it became apparent that the transition to the sphere has the following problem: One can use the spherical de Casteljau algorithm for evaluation of the curve, but one looses the powerful and important subdivision property. Thus, alternative constructions on the sphere have been proposed (see e.g. [11, 10, 21]. Again motivated by the application of motion design, Park and Ravani defined Bézier curves on Riemannian manifolds. Moreover, Sprott and Ravani [29] extended B-spline algorithms to Lie groups and applied them to motion design. However, smoothness properties have not been proved.

An algebraic approach to NURBS curves on quadrics, which is also rooted in line geometry and kinematics, has been proposed by Dietz, Hoschek and Jüttler [7]. It has been shown that this method delivers a useful control structure and is suitable for interactive design [20]. The algebraic approach is extendable to low degree algebraic surfaces. In particular, it can be used for motion design (see e.g. [10, 24]), but it is not suitable for curve design on general surfaces and manifolds.

There is only very little work dealing with subdivision schemes on manifolds. Until very recently, the most general analysis of a nonlinear subdivision scheme has been the analysis of the counterpart to the Lane-Riesenfeld algorithm for cubic B-spline curves on Riemannian manifolds by L. Noakes [15]. Noakes proved that the limit curve is differentiable and its derivative Lipschitz. In ongoing research, Wallner and Dyn found a technique to analyze nonlinear schemes based on certain 'proximity' conditions to known linear schemes [34]. We will not pursue subdivision in this paper, but concentrate on variational design. The 'intersecting topic', variational subdivision curves on surfaces, has been addressed, but not fully analyzed in [8].

A number of papers dealing with variational design on surfaces is considering the sphere [3, 4], partially in view of the application to motion design [19, 21].

An early contribution to variational curve design on more general surfaces is due to Noakes et al. [14]. The authors characterize the minimizers of an intrinsic geometric counterpart to the L^2 norm of the second derivative. This is the integral of the squared covariant derivative of the first derivative with respect to arc length. We will not pursue this intrinsic formulation in the present paper.

The most closely related to our work is the PhD thesis of H. Bohl [2]. It deals with the usual energy functional used in CAGD, but restricts the curves to a given surface in $\mathbb{R}^3$. Bohl proves the existence of a solution and gives computed examples based on a quasi Newton optimization algorithm. A large part of his work deals with patch boundaries.

If we consider on a surface the minimizers of the curve length, we obtain the well-studied *geodesics*. Their geometric properties have been investigated in classical differential geometry. Geodesics in a scaled arc length parameterizations also arise as minimizers of the L^2 norm of the first derivative. A variety of applications of geodesics has been described within Computer Vision and Image Processing (see e.g. [17, 26]).

The present work has also been inspired by research on active contours [1], especially by work on active curves and geometric flows of curves on surfaces (see, e.g. [6, 12, 17]).

Contributions and outline of the present paper

We will study the minimizers of standard energy functionals of spline theory, which interpolate or approximate given data points, and are restricted on a given m-dimensional surface $\Phi \subset \mathbb{R}^n$. The restriction to Φ is the new aspect, and it is also the source of arising complications. However, we will find nice characterizations of the minimizers.

In Section 2, we characterize the counterparts on surfaces to C^2 cubic splines, to splines in tension and to C^4 quintic splines. It will turn out that differentiability at the interpolation points is the familiar one, but the characterization of spline segments is slightly different. We give an example. Whereas the minimizers of the L^2 norm of the second derivative have cubic segments (vanishing fourth derivative) in the unrestricted case, the corresponding splines on surfaces have segments with vanishing tangential component of the fourth derivative; we call such segments 'tangentially cubic'. Hence, the differential equation $\mathbf{c}^{(4)} = 0$ changes to $tpr\,\mathbf{c}^{(4)} = 0$, with tpr denoting the tangential projection (orthogonal projection into the corresponding tangent space of Φ). Analogous changes are found in the characterizing differential equations for the other spline schemes. Section 3 is devoted to approximating curves, in particular to the counterparts of smoothing splines. In view of the importance of tangentially cubic curves, we give in Section 4 some explicit representations of such curves on special surfaces, namely certain cylinder surfaces. In particular, we address special tangentially cubic parametrizations of a curve; these are parametrizations whose fourth derivative is orthogonal to the curve.

2 Interpolating spline curves on surfaces

In view of the applications we have in mind, it is necessary to work on surfaces of arbitrary dimension and codimension. Thus, we consider an m-dimensional surface Φ in Euclidean $\mathbb{R}^n$, $m < n$. Moreover, a sequence of points $\mathbf{p}_i \in \Phi,\ i = 1, \ldots, N$ and real numbers $u_1 < \cdots < u_N$ shall be given. We are seeking interpolating splines on the surface Φ. Sometimes we will also call Φ a manifold; if not stated otherwise, we work with a manifold *without boundary.* Thus, we have closed or unbounded surfaces.

2.1 The counterparts to cubic splines

Let us recall the situation, where we are not confined to a manifold: Among all curves $\mathbf{x}(u) \subset \mathbb{R}^n$, whose first and second derivative satisfy $\dot{\mathbf{x}} \in AC(I), \ddot{\mathbf{x}} \in L^2(I)$ on $I = [u_1, u_N]$, and which interpolate the given data, $\mathbf{x}(u_i) = \mathbf{p}_i$, the unique minimizer of

$$E_2(\mathbf{x}) = \int_{u_1}^{u_N} \|\ddot{\mathbf{x}}(u)\|^2 du, \tag{1}$$

is the interpolating C^2 *cubic spline* $\mathbf{c}(u)$.

In the following we would like to extend this well-known result to the case where the admissible curves $\mathbf{x}(u)$ are restricted to the given manifold Φ. We are not changing the functional E_2, whose interpretation requires an embedding space. We are considering the restriction to the manifold as a constraint, rather than formulating the problem in terms of the intrinsic geometry of the manifold. As we will see later, this is a suitable formulation for the problems we would like to solve.

Theorem 1 *Consider real numbers $u_1 < \ldots < u_N$ and points $\mathbf{p}_1, \ldots, \mathbf{p}_N$ on an m-dimensional C^4 surface Φ in Euclidean $\mathbb{R}^n$. We let $I = [u_1, u_N]$. Then among all C^1 curves $\mathbf{x} : I \to \Phi \subset \mathbb{R}^n$, which interpolate the given data, i.e. $\mathbf{x}(u_i) = \mathbf{p}_i$, $i = 1, \ldots, N-1$, and whose restrictions to the intervals $[u_i, u_{i+1}], i = 1, \ldots, N-1$ are C^4, a curve $\mathbf{c}$ which minimizes the functional E_2 of Equ. (1) is C^2 and possesses segments $\mathbf{c}|[u_i, u_{i+1}]$, whose fourth derivative vectors are orthogonal to Φ. Moreover, at the end points $\mathbf{p}_1 = \mathbf{c}(u_1)$ and $\mathbf{p}_N = \mathbf{c}(u_N)$ of the solution curve, the second derivative vector is orthogonal to Φ.*

Proof. If a solution curve $\mathbf{c}$ exists, the first variation of the energy functional must vanish there. To express this condition, we consider neighboring curves $\mathbf{x}(u) \subset \Phi$, written as

$$\mathbf{x}(u, \epsilon) = \mathbf{c}(u) + \mathbf{h}(u, \epsilon). \tag{2}$$

For any fixed $\tilde{u} \in I$, the curve $\mathbf{h}(\tilde{u}, \epsilon)$ is a curve in Φ with $\mathbf{h}(\tilde{u}, 0) = 0$. Its Taylor expansion at $\epsilon = 0$ reads

$$\mathbf{h}(\tilde{u}, \epsilon) = \epsilon \mathbf{h}_\epsilon(\tilde{u}, 0) + \frac{\epsilon^2}{2} \mathbf{h}_{\epsilon\epsilon}(\tilde{u}, 0) + \ldots,$$

where the subscripts indicate differentiation. Note that $\mathbf{h}_\epsilon(\tilde{u}, 0) =: \mathbf{t}(\tilde{u})$ is a tangent vector of Φ at $\mathbf{c}(\tilde{u})$. The displacement curves $\mathbf{h}$ to the given interpolation points vanish, $\mathbf{h}(u_i, \epsilon) = 0$, for all ϵ. In particular, we have $\mathbf{h}_\epsilon(u_i, 0) = \mathbf{t}(u_i) = 0$. For the following, it is important to see that the mixed partial derivative vector $\mathbf{h}_{\epsilon u}(u_i, 0) = \mathbf{t}_u(u_i)$ is a tangent vector of Φ at $\mathbf{c}(u_i) = \mathbf{p}_i$. Geometrically, this follows from the fact that the curve $\mathbf{c}(u) + \mathbf{t}(u)$ is a curve on the ruled surface $\mathbf{r}(u, v) = \mathbf{c}(u) + v\mathbf{t}(u)$, which touches Φ along $\mathbf{c}$. This curve passes through each interpolation point $\mathbf{p}_i = \mathbf{c}(u_i) + \mathbf{t}(u_i)$, and thus its derivative vector $\mathbf{c}_u(u_i) + \mathbf{t}_u(u_i)$ is a tangent vector of Φ. Together with the tangency of $\mathbf{c}_u(u_i)$ this implies tangency of $\mathbf{t}_u(u_i)$.

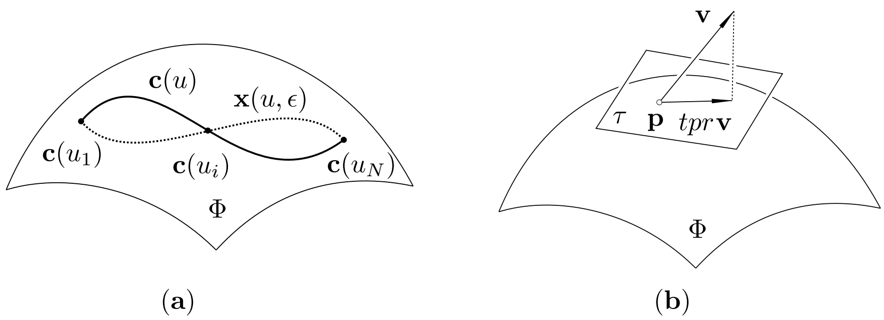

Figure 1: (**a**) Neighboring curves $\mathbf{x}(u, \epsilon) \subset \Phi$ to the solution curve $\mathbf{c}(u)$. (**b**) The orthogonal projection $tpr\mathbf{v}$ of a vector $\mathbf{v}$ at a point $\mathbf{p} \in \Phi$ onto the tangent space τ of Φ at $\mathbf{p}$.

Whatever field of displacement curves $\mathbf{h}(u, \epsilon)$ we have chosen, the energy functional must assume a stationary value at $\epsilon = 0$,

$$\frac{d}{d\epsilon} E_2(\mathbf{x}(u, \epsilon))\Big|_{\epsilon=0} = 0. \tag{3}$$

In view of

$$E_2(\mathbf{x}(u, \epsilon)) = \int_I [\ddot{\mathbf{c}}(u) + \mathbf{h}_{uu}(u, \epsilon)]^2 du = \int_I [\ddot{\mathbf{c}} + \epsilon \mathbf{t}_{uu} + \epsilon^2(\cdots)]^2 du,$$

this is equivalent to

$$\int_I \ddot{\mathbf{c}} \cdot \mathbf{t}_{uu} \; du = 0. \tag{4}$$

Integration by parts on each segment yields

$$\int_{u_i}^{u_{i+1}} \ddot{\mathbf{c}} \cdot \mathbf{t}_{uu} \; du = \ddot{\mathbf{c}} \cdot \mathbf{t}_u \Big|_{u_i}^{u_{i+1}} - \mathbf{c}^{(3)} \cdot \mathbf{t} \Big|_{u_i}^{u_{i+1}} + \int_{u_i}^{u_{i+1}} \mathbf{c}^{(4)} \cdot \mathbf{t} \; du.$$

We first note that the middle term vanishes, since $\mathbf{t}(u_i) = 0$. Summing up over all intervals and denoting left and right derivatives at the knots u_i by a subscript $-$ and $+$, respectively, the condition for a solution $\mathbf{c}$ reads,

$$
\begin{aligned}
0 = -\ddot{\mathbf{c}}_+(u_1) \cdot \mathbf{t}_u(u_1) \quad &+ \quad \sum_{i=2}^{N-1} [(\ddot{\mathbf{c}}_-(u_i) - \ddot{\mathbf{c}}_+(u_i)) \cdot \mathbf{t}_u(u_i)] \\
&+ \quad \ddot{\mathbf{c}}_-(u_N) \cdot \mathbf{t}_u(u_N) + \int_I \mathbf{c}^{(4)} \cdot \mathbf{t} \; du. \qquad (5)
\end{aligned}
$$

Since $\mathbf{t}(u)$ is an arbitrary tangent vector field along $\mathbf{c}$, and $\mathbf{t}_u(u_i)$ are arbitrary tangent vectors at the given interpolation points, a solution curve $\mathbf{c}(u)$ must satisfy the following orthogonality conditions to the manifold Φ. To formulate them, we denote by *tpr* the orthogonal projection of a vector at a point $\mathbf{p} \in \Phi$ onto the tangent space τ of Φ at $\mathbf{p}$ (see Fig. 1),

$$
\begin{aligned}
& tpr \;\ddot{\mathbf{c}}_+(u_1) = tpr \;\ddot{\mathbf{c}}_-(u_N) = 0, && (6)\\
& tpr \;(\ddot{\mathbf{c}}_-(u_i) - \ddot{\mathbf{c}}_+(u_i)) = 0, \quad i = 2, \dots, N-1, && (7)\\
& tpr \;\mathbf{c}^{(4)}(u) = 0 \quad \text{on each interval } [u_i, u_{i+1}]. && (8)
\end{aligned}
$$

Exactly these conditions are stated in Theorem 1 (for an illustration see Fig. 2). Equation (6) means that the second derivative vector at the end points is orthogonal to the surface; this implies (but is not equivalent to) vanishing geodesic curvature at the end points. Equation (7) shows that the tangential component of the second derivative is continuous along $\mathbf{c}$; we call this behaviour *tangentially* C^2 or briefly TC^2. It will be shown later that in view of the C^4 manifold Φ this implies C^2. Finally, by equation (8), the fourth derivative vectors of the curve (which may be discontinuous at the interpolation points) are orthogonal to the manifold Φ. Vanishing fourth derivative characterizes a cubic curve; since here only the tangential component vanishes, we speak of a *tangentially cubic curve* in the following.

To complete the proof, we show the following result:

Lemma 1 *On an m-dimensional C^k manifold $\Phi \subset \mathbb{R}^n$, $m < n$, we consider a curve $\mathbf{c}(u)$, which is C^k ($k \geq 2$) everywhere except at a point $\mathbf{c}(u_0)$. At $\mathbf{c}(u_0)$, all derivatives up to order k have a continuous tangential component, i.e.*

$$
tpr(\mathbf{c}_-^{(i)}(u_0) - \mathbf{c}_+^{(i)}(u_0)) = 0, \;\; i = 1, \dots, k. \qquad (9)
$$

Then, the curve is also C^k at $\mathbf{c}(u_0)$.

Proof. We may represent the manifold at least in a neighborhood of $\mathbf{c}(u_0)$ as an intersection of $n - m$ hypersurfaces, whose implicit representations shall be

$$
F_j(\mathbf{x}) = 0, \quad j = 1, \dots, n - m.
$$

Since the curve lies on all these hypersurfaces, we have the identity $F_j(\mathbf{c}(u)) = 0$ in u. By differentiation, we obtain $\nabla F_j(\mathbf{c}) \cdot \dot{\mathbf{c}} = 0$, and for the i-th derivative we find an expression of the form

$$
G_{i-1}(\mathbf{c}, \dots, \mathbf{c}^{(i-1)}) + \nabla F_j(\mathbf{c}) \cdot \mathbf{c}^{(i)} = 0. \qquad (10)
$$

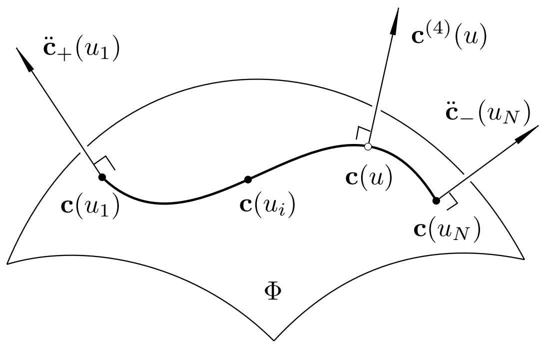

Figure 2: Conditions of Theorem 1.

Here, $G_{i-1}(\mathbf{c}, \ldots, \mathbf{c}^{(i-1)})$ abbreviates a function, which depends on F_j, up to differentiation order i, and on $\mathbf{c}$, but only up to differentiation order $i-1$. At $\mathbf{c}(u_0)$ this holds for the left and right derivatives. The first derivative is tangential, and thus $\mathbf{c}$ is C^1 at u_0. This is the basis of an induction on the differentiation order i. Assuming continuity of the derivative $\mathbf{c}^{(i-1)}$ at u_0, equation (10) expresses that the components $\nabla F_j(\mathbf{c}) \cdot \mathbf{c}^{(i)}$ of the i-th derivative are continuous at u_0, since they equal $G_{i-1}(\mathbf{c}, \ldots, \mathbf{c}^{(i-1)})$, which is continuous by the induction hypothesis. Since the vectors $\nabla F_j(\mathbf{c}(u_0))$, $j = 1, \ldots, n-m$, span the normal space of Φ at $\mathbf{c}(u_0)$, we have shown that the normal component of the i-th derivative is continuous. The tangential component is continuous by (9) and thus we have shown continuity of $\mathbf{c}^{(i)}$. This completes the proof of Lemma 1 and thus also the proof of Theorem 1. $\square$

The minimizers possess a characterization which is very similar to the familiar cubic C^2 splines. They are C^2 and tangentially cubic. However, the problem is nonlinear. Tangentially cubic curves can in general not be computed explicitly, but only by a numerical algorithm.

Existence of the solution follows from the work of Bohl [2] and Wallner [33]. Uniqueness cannot be expected, as will be clear from the following considerations.

2.2 Geodesics

Let us replace the energy in (1) by the L^2 norm of the first derivative,

$$E_1(\mathbf{x}) = \int_{u_1}^{u_N} \|\dot{\mathbf{x}}(u)\|^2 du. \tag{11}$$

Moreover, we just prescribe the two end points. Then we find with the same approach as in the proof of Theorem 1 that the curve's second derivative vectors

$\ddot{\mathbf{c}}$ are orthogonal to Φ; we may say the curve is *tangentially linear.* In particular, $\dot{\mathbf{c}}$ and $\ddot{\mathbf{c}}$ are orthogonal, and hence $\|\dot{\mathbf{c}}\|^2 = const$. This proves, that the curve is parameterized by a constant multiple of its arc length. Moreover, it shows that $\ddot{\mathbf{c}}$ represents the principal normal, and orthogonality of the principal normal to Φ characterizes a geodesic. Thus we find the known result that *the minimizers of E_1 are geodesics in a scaled arc length parameterization.* Note that the functionals we are considering are also optimizing the parameterization. In case of E_1, we get a very nice one. But also for E_2, the obtained parameterization may be very desirable, e.g. for applications in robotics. We will later show how to deal with purely geometric functionals.

Since even for geodesics we do not have simple results on uniqueness, the more involved case of splines will hardly allow us a characterization of situations with a unique solution.

2.3 Splines in tension on surfaces

A linear combination of energies (11) and (1) leads in the unrestricted case to the well known *splines in tension* as minimizers [27]. The counterpart on manifolds is characterized in the following theorem, whose proof can be omitted since it is completely analogous to that of Theorem 1.

Theorem 2 *Consider a sequence of data points, parameter values and admissible curves on a surface Φ as in Theorem 1. Then, a minimizer of the functional*

$$E_t(\mathbf{x}) = \int_{u_1}^{u_N} (\ddot{\mathbf{x}}^2 + w\dot{\mathbf{x}}^2)du, \quad w = const > 0, \tag{12}$$

is a C^2 curve which satisfies

$$tpr\ (\mathbf{c}^{(4)}(u) - w\ddot{\mathbf{c}}(u)) = 0 \tag{13}$$

on all segments. The end conditions are $tpr\,\ddot{\mathbf{c}}_+(u_1) = tpr\,\ddot{\mathbf{c}}_-(u_N) = 0$.

Increasing w brings the solution closer to the curve which minimizes E_1, i.e., the curve composed by geodesic segments between the interpolation points. Hence, the parameter w controls the tension of the curve.

2.4 Counterparts to higher order splines on surfaces

It is known from spline theory, that the minimizers of the L^2 norm of the third derivative,

$$E_3(\mathbf{x}) = \int_{u_1}^{u_N} \|\mathbf{x}^{(3)}(u)\|^2 du, \tag{14}$$

under interpolation conditions are *quintic C^4 splines.* It is not difficult to extend this result to manifolds.

Theorem 3 *Consider real numbers $u_1 < \ldots < u_N$ and points $\mathbf{p}_1, \ldots, \mathbf{p}_N$ on an m-dimensional C^6 manifold Φ in Euclidean $\mathbb{R}^n$. Then among all C^2 curves $\mathbf{x} : [u_1, u_N] \to \Phi \subset \mathbb{R}^n$, which interpolate the given data and whose restrictions to the intervals $[u_i, u_{i+1}], i = 1, \ldots, N-1$, are C^6, a curve $\mathbf{c}$ which minimizes the L^2 norm E_3 of the third derivative is C^4 and possesses segments $\mathbf{c}|[u_i, u_{i+1}]$, whose sixth derivative vectors are orthogonal to Φ. Moreover, at the end points $\mathbf{p}_1 = \mathbf{c}(u_1)$ and $\mathbf{p}_N = \mathbf{c}(u_N)$ of the solution curve, the third derivative vanishes and the fourth derivative vector is orthogonal to Φ.*

Proof. The proof resembles the one of Theorem 1. We use the same notation and just mention the few essential differences. A solution $\mathbf{c}(u)$ must satisfy

$$\int_I \mathbf{c}^{(3)} \cdot \mathbf{t}_{uuu}\, du = 0. \tag{15}$$

Integration by parts on each segment yields

$$\int_{u_i}^{u_{i+1}} \mathbf{c}^{(3)} \cdot \mathbf{t}_{uuu}\, du = \mathbf{c}^{(3)} \cdot \mathbf{t}_{uu}\Big|_{u_i}^{u_{i+1}} - \mathbf{c}^{(4)} \cdot \mathbf{t}_u\Big|_{u_i}^{u_{i+1}} + \mathbf{c}^{(5)} \cdot \mathbf{t}\Big|_{u_i}^{u_{i+1}} - \int_{u_i}^{u_{i+1}} \mathbf{c}^{(6)} \cdot \mathbf{t}\, du.$$

¿From the four terms on the right hand side, the third one vanishes because of the interpolation conditions, expressed by $\mathbf{t}(u_i) = 0$. Since $\mathbf{t}$ is an arbitrary tangent vector field along $\mathbf{c}$, the last term vanishes iff $tpr\,\mathbf{c}^{(6)} = 0$. Now we can sum up over all intervals and collect the contributions from the first two terms $\mathbf{c}^{(3)} \cdot \mathbf{t}_{uu}$ and $\mathbf{c}^{(4)} \cdot \mathbf{t}_u$. Since $\mathbf{t}_u(u_i)$ is tangential, we find a TC^4 condition at the inner knots and orthogonality of the fourth derivative to the manifold at the end points. It is not hard to see that $\mathbf{t}_{uu}(u_i)$ can be an arbitrary vector and this shows that a solution must have vanishing third derivative at the ends and it must be C^3 at the inner knots. Application of Lemma 1 shows that the curve is even C^4 at the inner knots, which completes the proof. □

Remark 1 *So far we always considered an open curve and natural end conditions. It is clear from the derivation of our results how to handle other cases. Let us discuss this at hand of the minimization of E_3. The segments are of course always tangentially quintic, i.e., $tpr\,\mathbf{c}^{(6)} = 0$. A closed spline is C^4 at all knots. Moreover, instead of working for an open curve with natural end conditions, we could prescribe first and second derivative vectors at the end points.*

The extension of the results to the minimization of the L^2 norm of even higher derivatives, or combinations of L^2 norms of different derivatives, is straightforward and thus it can be omitted.

3 Approximating spline curves on surfaces

Reinsch [22] relaxed the interpolation conditions $\mathbf{x}(u_i) - \mathbf{p}_i = 0$ by adding the sum of squared interpolation errors as a penalty term to E_2,

$$E_2^s(\mathbf{x}) = \lambda \int_{u_1}^{u_N} \|\ddot{\mathbf{x}}(u)\|^2 du + \mu \sum_{i=1}^{N} \|\mathbf{x}(u_i) - \mathbf{p}_i\|^2, \quad \lambda, \mu > 0. \tag{16}$$

The minimizers, called *smoothing splines*, found a variety of applications in the analysis of observational data [32]. The choice of the smoothing parameter $\lambda : \mu$ in this method for data approximation is not a simple problem. Often one uses statistical methods for this critical task [32]. We should also mention that an analogous functional is used for freeform surface fitting in reverse engineering applications [31].

We can use the same method as in the proof of Theorem 1 to derive the following result on the counterparts to cubic smoothing splines on surfaces.

Theorem 4 *Consider a sequence of data points, parameter values and differentiability conditions on an admissible curve on a surface Φ as in Theorem 1. Then, a minimizer of the smoothing spline functional E_2^s has all the properties of the minimizer of E_2 in Theorem 1. However, the interpolation conditions $\mathbf{c}(u_i) = \mathbf{p}_i$ are replaced by the following transition conditions between segments,*

$$tpr\ [\lambda(\mathbf{c}_+^{(3)}(u_i) - \mathbf{c}_-^{(3)}(u_i)) + \mu(\mathbf{c}(u_i) - \mathbf{p}_i)] = 0. \tag{17}$$

In an analogous way, one can investigate the minimizers of smoothing spline counterparts $E_t^s, E_3^s, \ldots,$ to other energy functionals. For example, the case of E_3^s yields in modification of Theorem 3 curves, where the interpolation conditions are replaced by the smoothness conditions

$$tpr\ [\lambda(\mathbf{c}_-^{(5)}(u_i) - \mathbf{c}_+^{(5)}(u_i)) + \mu(\mathbf{c}(u_i) - \mathbf{p}_i)] = 0. \tag{18}$$

4 Examples of tangentially cubic curves

4.1 Tangentially cubic curves on cylinder surfaces in $\mathbb{R}^3$

Let us consider a general cylinder surface $\Phi \subset \mathbb{R}^3$. We may choose a directrix curve $\mathbf{l}(u) = (x_1(u), x_2(u), 0)$ and rulings parallel to $(0, 0, 1)$, i.e. a parameterization of Φ in the form

$$\Phi: \ \mathbf{x}(u, v) = (x_1(u), x_2(u), v).$$

For a tangentially cubic spline curve $\mathbf{c}(t) = (c_1(t), c_2(t), c_3(t)) \subset \Phi$, the fourth derivative $\mathbf{c}^{(4)}$ must be orthogonal to Φ, and thus also orthogonal to the corresponding ruling of Φ. This requires a vanishing third component $c_3^{(4)}(t) = 0$ and shows that the function c_3 is *cubic*,

$$c_3(t) = a_0 + a_1 t + \ldots + a_3 t^3.$$

The other two coordinate functions $(c_1(t), c_2(t))$ have its fourth derivative orthogonal to a parameterization $(c_1(t), c_2(t), 0)$ of the (planar) orthogonal cross section $\mathbf{l}$ of Φ. This is therefore a *tangentially cubic parameterization of the cross section curve.*

Choosing the data points for a spline on a cross section, we find a result on energy minimizing parameterizations of a given curve.

Corollary 1 *Given a C^4 curve $\mathbf{l}$ in $\mathbb{R}^n$, and a sequence of points $\mathbf{p}_i$ with parameters u_i, $i = 1, \ldots, N$, on it. Then, among all C^1, piecewise C^4 parameterizations of the curve which satisfy the interpolation conditions $\mathbf{c}(u_i) = \mathbf{p}_i$, a minimizer of the L^2 norm of the second derivative is tangentially cubic, i.e., a parameterization whose fourth derivative is orthogonal to $\mathbf{l}$. Moreover, $\mathbf{c}$ is C^2 and the second derivative at the end points is orthogonal to $\mathbf{l}$.*

The derivation in $\mathbb{R}^n$ can be done with a cylinder surface in $\mathbb{R}^{n+1}$. We may also note that the case of a manifold of dimension 1 is actually included in the formulation and in the proof of Theorem 1.

Since the study of splines on cylinders is reduced to tangentially cubic parameterizations of curves, we will study those in the following in more detail.

4.2 Tangentially cubic arc length parameterizations

In view of the important role of arc length parameterizations for curves, we answer the question on which curves in $\mathbb{R}^n$ the arc length parameterization $\mathbf{c}(s)$ (or a scaled version of it), is tangentially cubic.

Denoting the Frenet frame of $\mathbf{c}$ by $\mathbf{e}_1(s), \ldots, \mathbf{e}_n(s)$, and the curvatures by $\kappa_1, \ldots, \kappa_{n-1}$, we have

$$\mathbf{c}' = \mathbf{e}_1, \ \mathbf{e}_1' = \kappa_1 \mathbf{e}_2, \ \mathbf{e}_1'' = \kappa_1' \mathbf{e}_2 - \kappa^2 \mathbf{e}_1 + \kappa_1 \kappa_2 \mathbf{e}_3.$$

Hence, the condition for a tangentially cubic parameterization becomes

$$0 = \mathbf{c}^{(4)} \cdot \mathbf{c}' = \mathbf{e}_1^{(3)} \cdot \mathbf{e}_1 = -3\kappa_1 \kappa_1'.$$

Proposition 1 *The arc length parameterization $\mathbf{c}(s)$ of a curve in $\mathbb{R}^n$ is tangentially cubic iff the curve possesses constant curvature κ_1.*

In the plane, we get only circles and straight lines. However, already in $\mathbb{R}^3$ this family is relatively rich, since the torsion $\tau = \kappa_2$ can be arbitrary. Parametric representations of special space curves with constant curvature have been given by E. Salkowski [25], whose formulae for $n \neq 1/2, m \in \mathbb{R} \backslash \{0\}$ we include for completeness:

$$\begin{aligned}
x(v) &= \frac{-1}{\sqrt{1+m^2}} \left(\frac{1-n}{4(1+2n)} \sin(1+2n)v + \frac{1+n}{4(1-2n)} \sin(1-2n)v + \frac{1}{2} \sin v \right) \\
y(v) &= \frac{1}{\sqrt{1+m^2}} \left(\frac{1-n}{4(1+2n)} \cos(1+2n)v + \frac{1+n}{4(1-2n)} \cos(1-2n)v + \frac{1}{2} \cos v \right) \\
z(v) &= \frac{1}{4m\sqrt{1+m^2}} \cos 2nv
\end{aligned}$$

Fig. 3 shows three examples with $m = \sqrt{1-n^2}$, $v \in [0, 6\pi)$ and $n = 1/6$ (left), $n = 1/4$ (center), $n = 1/3$ (right).

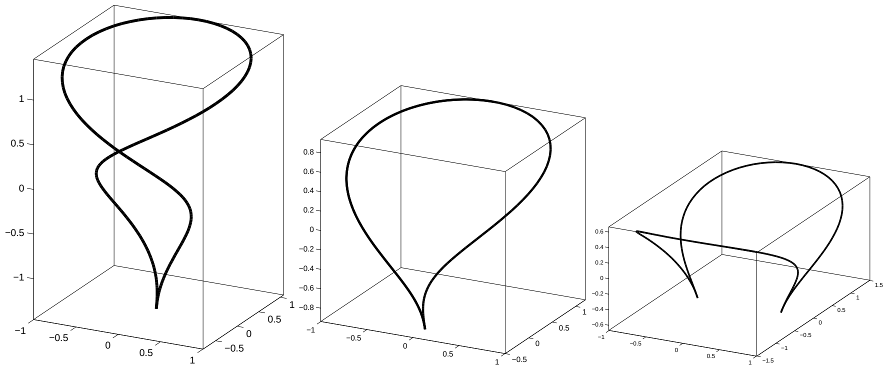

Figure 3: Examples of space curves with constant curvature.

4.3 Tangentially cubic parameterizations of the circle; spline curves on a right circular cylinder

Consider the unit circle $\mathbf{x}^2 = 1$ and its parameterization

$$\mathbf{c}(t) = (\cos\phi(t), \sin\phi(t)).$$

We want to determine $\phi(t)$ such that $\mathbf{c}(t)$ is a tangentially cubic parameterization of the circle. A simple computation shows that orthogonality of first and fourth derivative is equivalent to

$$\phi^{(4)} - 6\dot{\phi}^2\ddot{\phi} = 0. \tag{19}$$

We set $\psi(t) := \dot{\phi}(t)$ and thus have to solve the differential equation

$$\psi^{(3)} = 6\psi^2\dot{\psi}.$$

We can immediately perform a first integration,

$$\ddot{\psi} = 2\psi^3 + A_0, \tag{20}$$

with a constant $A_0 \in \mathbb{R}$. Since the particular solution $\dot{\psi} = 0$, i.e. the scaled arc length parameterization $\phi = at + b$ is already known from the previous considerations, we may assume $\dot{\psi} \neq 0$ from now on. Multiplying equation (20) by $2\dot{\psi}$ and integrating again, we arrive with $A := 2A_0$ at

$$\dot{\psi}^2 = \psi^4 + A\psi + B.$$

Separation of the variables and integration leads to

$$t + C = \pm\int \frac{d\psi}{\sqrt{\psi^4 + A\psi + B}}. \tag{21}$$

The integral on the right hand side requires Legendre normal integrals. We omit its discussion in the general case and just look at the special case $A = B = 0$. There, we obtain $\psi = \pm 1/(t + C)$ and

$$\phi(t) = \ln |C \pm t| + D. \tag{22}$$

Thus, even in the very simple case of a right circular cylinder, just two very special families of tangentially cubic curves can be written down explicitly. On the unit cylinder $x_1^2 + x_2^2 = 1$, these are the curves derived from a scaled arc length parameterization

$$\mathbf{c}(t) = (\cos(at + b), \sin(at + b), a_0 + a_1 t + \ldots + a_3 t^3), \tag{23}$$

and the curves to the circle parameterization (22),

$$\mathbf{c}(t) = (\cos(\ln |C \pm t| + D), \sin(\ln |C \pm t| + D), a_0 + a_1 t + \ldots + a_3 t^3). \tag{24}$$

A few examples of curves (23) and (24) are depicted in Fig. 4.

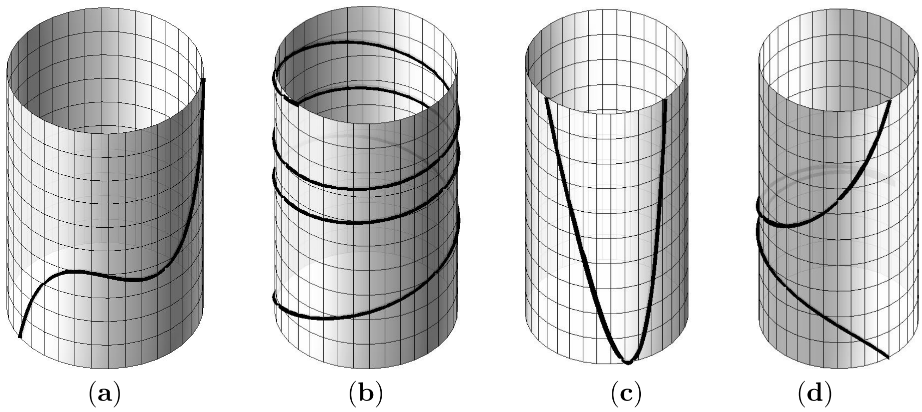

Figure 4: Examples of tangentially cubic curves derived from (**a**) and (**b**) a *scaled arc length parameterization* and (**c**) and (**d**) a *circle parameterization.*

4.4 Tangentially cubic parameterization by the direction angle of the curve tangent

It turns out that nice explicit representations of tangentially cubic curves on cylinder surfaces arise as follows. We would like to determine those planar curves, whose parameterization $\mathbf{c}(\phi)$ with respect to the direction angle ϕ of the curve tangent is tangentially cubic. Therefore, we write the family of tangent lines of such a curve in the form

$$-x_1 \sin \phi + x_2 \cos \phi + h(\phi) = 0. \tag{25}$$

Here, ϕ is the angle between the tangent, oriented by the vector $(\cos\phi, \sin\phi)$, and the x_1-axis; h denotes the signed distance of the origin to the oriented tangent. The function $h(\phi)$ is known as support function. The curve parameterization $\mathbf{c}(\phi)$ is computed as envelope of the family of tangents, i.e., from equation (25) and its derivative with respect to ϕ. We find

$$\mathbf{c}(\phi) = (h\sin\phi + \dot{h}\cos\phi, -h\cos\phi + \dot{h}\sin\phi). \tag{26}$$

A tangentially cubic parameterization requires $\dot{\mathbf{c}} \cdot \mathbf{c}^{(4)} = 0$, which is equivalent to

$$h^{(5)} - 2h^{(3)} - 3\dot{h} = 0. \tag{27}$$

The solution of this differential equation is, with real constants b_i,

$$h(\phi) = b_0 + b_1\cos\phi + b_2\sin\phi + b_3\cosh(\sqrt{3}\phi + b_4). \tag{28}$$

To understand these curves, we note that an appropriate translation of the origin can eliminate the term $b_1\cos\phi + b_2\sin\phi$, and b_4 just accounts for a rotation. Thus it is sufficient to study the case $b_1 = b_2 = b_4 = 0$. Addition of the constant b_0 to the support function means construction of the offset at signed distance b_0.

The curve with support function

$$h(\phi) = \cosh\sqrt{3}\phi, \tag{29}$$

is a special instance of a so-called *hypercycloid*. Hypercycloids are a certain counterpart to the familiar epicycloids or hypocycloids, which are obtained as locus of a point on a circle, which rolls on another circle. A hypercycloid possesses such a kinematic generation, but the generating circles are not real: the moving circle has a complex radius and the fixed circle a purely imaginary radius. Hypercycloids arise for example as orthogonal projections of geodesics on a paraboloid of revolution, when projecting parallel to the rotational axis (see [30], pp. 227–232).

All other nontrivial solutions of our problem arise from the special hypercycloid with support function (29) by offsetting and a similarity transformation. The trivial solutions belong to $b_3 = 0$ and are circles.

On a cylinder surface with the hypercycloid to (29) as cross section, the curves

$$\mathbf{c}(\phi) = \begin{pmatrix} \cosh\sqrt{3}\phi\ \sin\phi + \sqrt{3}\sinh\sqrt{3}\phi\ \cos\phi \\ -\cosh\sqrt{3}\phi\ \cos\phi + \sqrt{3}\sinh\sqrt{3}\phi\ \sin\phi \\ a_0 + a_1\phi + a_2\phi^2 + a_3\phi^3 \end{pmatrix} \tag{30}$$

are tangentially cubic. Figure 5 shows for $\phi \in [-2.5, 2.5]$ tangentially cubic curves (30) on a cylinder surface with the hypercycloid (29) as cross section, with (**a**) $a_0 = a_1 = a_3 = 0, a_2 = 5$ and (**b**) $a_0 = 0, a_1 = 1, a_2 = 2, a_3 = 1$.

The main conclusions of this section are:

- Only for very special surfaces, namely cylinders with special cross sections, we have been able to find explicit representations of spline segments.

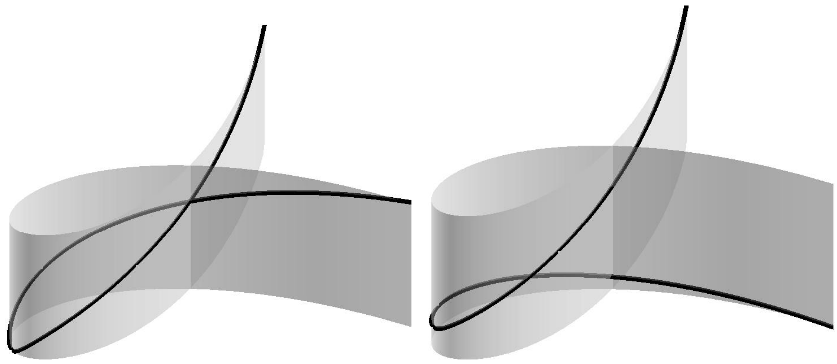

Figure 5: Tangentially cubic curves (30) on a cylinder surface with the hypercycloid (29) as cross section.

- There is no hope to find explicit representations of spline segments (tangentially cubic curves) on a sufficiently large class of surfaces and therefore further work must concentrate on efficient algorithms for the numerical computation.

Acknowledgments

This research has been supported by the Austrian Science Fund (FWF) under grant P16002-N05.

References

[1] Blake, A., Isard, M., 1998. Active Contours, Springer.

[2] Bohl, H., 1999. Kurven minimaler Energie auf getrimmten Flächen, PhD thesis, Univ. Stuttgart, Germany.

[3] Brunnett, G., Crouch, P. E., 1994. Elastic curves on the sphere, Advances in Comp. Math. 2, 23–40.

[4] Brunnett, G., Crouch, P. E., Silva Leite, F., 1995. Spline elements on spheres, In: Daehlen, M., Lyche, T., Schumaker, L., eds., Mathematical Methods for Curves and Surfaces, Vanderbilt Univ. Press, pp. 49–54.

[5] Brunnett, G., Hagen, H., Santarelli, P., 1993. Variational design of curves and surfaces, Surv. Math. Ind. 3, 1–27.

[6] Cheng, B.-T., Burchard, P., Merriman, B., Osher, S., 2000. Motion of curves constrained on surfaces using a level set approach, UCLA CAM Report 00-32.

[7] Dietz, R., Hoschek, J., Jüttler, B., 1993. An algebraic approach to curves and surfaces on the sphere and on other quadrics, Comp. Aided Geom. Design 10, 211–229.

[8] Hofer, M., Pottmann, H., Ravani, B., 2003. Geometric design of motions constrained by a contacting surface pair, Comp. Aided Geom. Design 20, 523–547.

[9] Hoschek, J., Lasser, D., 1993. Fundamentals of Computer Aided Geometric Design, A. K. Peters, Wellesley, MA.

[10] Jüttler, B., Wagner, M., 2002. Kinematics and Animation, in G. Farin, M.-S. Kim and J. Hoschek, eds., Handbook of Computer Aided Geometric Design, North Holland, Amsterdam, pp. 723–748.

[11] Kim, M.-J., Kim, M.-S., Shin, S., 1995. A general construction scheme for unit quaternion curves with simple high order derivatives, Computer Graphics (SIGGRAPH '95) 29, 369–376.

[12] Memoli, F., Sapiro, G., 2001. Fast computation of weighted distance functions and geodesics on implicit hyper-surfaces, J. Comp. Phys. 173, 730–764.

[13] Moreton, H., Sequin, C., 1993, Scale-invariant minimum-cost curves: fair and robust design implements, Computer Graphics Forum 12, 473–484.

[14] Noakes, L., Heinzinger, G., Paden, B., 1989. Cubic splines on curved spaces, IMA J. Math. Control & Information 6, 465–473.

[15] Noakes, L., 1998. Non-linear corner cutting, Adv. Comp. Math. 8, 165–177.

[16] Osher, S., Fedkiw, R., 2002. The Level Set Method and Dynamic Implicit Surfaces, Springer.

[17] Osher, S., Paragios, N., eds., 2003. Geometric Level Set Methods, Springer, New York.

[18] Park, F. C., Ravani, B., 1995. Bézier curves on Riemannian manifolds and Lie groups with kinematic applications, ASME J. of Mechanical Design 115, 36–40.

[19] Park, F. C., Ravani, B., 1997. Smooth invariant interpolation of rotations, ACM Trans. Graphics 16, 277–295.

[20] Pottmann, H., Wallner, J., 2001. Computational Line Geometry, Springer, Heidelberg.

[21] Ramamoorthi, R., Barr, A. 1997. Fast construction of accurate quaternion splines, Computer Graphics (SIGGRAPH '97) 31, 287–292.

[22] Reinsch, C. H., 1967. Smoothing by spline functions, Num. Math. 10, 177–183.

[23] Renner, G., Weiss, V., 2004. Exact and approximate computation of B-spline curves on surfaces, Computer-Aided Design 36, 351–362.

[24] Röschel, O., 1998. Rational motion design - a survey. Computer-Aided Design 30, 169–178.

[25] Salkowski, E., 1909. Zur Transformation von Raumkurven, Math. Ann. 66, 517–557.

[26] Sapiro, G., 2001. Geometric Partial Differential Equations and Image Analysis, Cambridge Univ. Press.

[27] Schweikert, D., 1966. An interpolation curve using a spline in tension, J. Math. and Phys. 45, 312–317.

[28] Shoemake, K., 1985. Animating rotations with quaternion curves, Computer Graphics (SIGGRAPH '85) 19, 245–254.

[29] Sprott, K., Ravani, B., Ruled surfaces, Lie groups and mesh generation, Proc. ASME Design Eng. Techn. Conf. 1997, No DETC97/DAC-3966.

[30] Strubecker, K., 1969. Differentialgeometrie II: Theorie der Flächenmetrik, Walter de Gruyter, Berlin.

[31] Várady, T., Martin, R., 2002. Reverse Engineering, In: Farin, G., Hoschek, J., Kim, M.-S., eds., Handbook of Computer Aided Geometric Design, Elsevier, pp. 651–681.

[32] Wahba, G., 1990. Spline Models for Observational Data, SIAM, Philadelphia.

[33] Wallner, J., 2003. Existence of set-interpolating and energy-minimizing curves, Techn. Report 111, Institute of Geometry, Vienna Univ. of Technology.

[34] Wallner, J., Dyn, N., 2003. Smoothness of subdivision schemes by proximity, Preprint, Institute of Geometry, Vienna Univ. of Technology.