

SPLINE SMOOTHING ON SURFACES

TOM DUCHAMP AND WERNER STUETZLE

ABSTRACT. We present a method for estimating functions on topologically and/or geometrically complex surfaces from possibly noisy observations. Our approach is an extension of spline smoothing, using a finite element method. The paper has a substantial tutorial component: we start by reviewing smoothness measures for functions defined on surfaces, simplicial surfaces and differentiable structures on such surfaces, subdivision functions, and subdivision surfaces. After describing our method, we show results of an experiment comparing finite element approximations to exact smoothing splines on the sphere, and we give examples suggesting that generalized cross-validation is an effective way of determining the optimal degree of smoothing for function estimation on surfaces.

1. INTRODUCTION AND MOTIVATION

We present a method for estimating functions on topologically and/or geometrically complex surfaces, like a (deformed) sphere, a three hole torus (Figure 1(a)), or the model of an elephant (Figure 1(b)). Let M be a surface. We assume that we are given observations (x_i, y_i) , with $x_i \in M$ and $y_i \in \mathbb{R}$. The y_i are, possibly noisy, observations of some unknown function f_{true} at the locations x_i , and our goal is to estimate f_{true} .

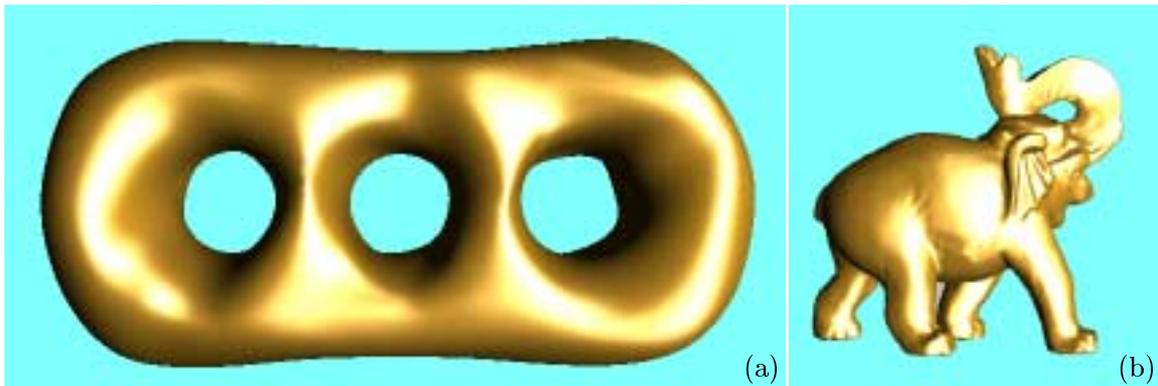


FIGURE 1. Panel (a) Three hole torus; Panel (b): Model of an elephant.

Estimation of functions defined on Euclidean domains is widely recognized as an important problem that has been addressed by researchers in areas as diverse as the earth sciences, meteorology, computer graphics, computer vision and image processing, machine learning, and statistics. The

Date: September 23, 2001, Revised: June 13, 2002.

Key words and phrases. Smoothing, splines, finite element method, subdivision surface, Laplace-Beltrami operator.

Supported by NSF grant DMS-9803226. Part of this work was performed at the Institute for Mathematics and its Applications, Minneapolis.

conceptually simple problem of estimating functions on the line alone has generated many hundreds of publications proposing a wide variety of algorithms and establishing their theoretical properties.

Among non-euclidean domains, the 2-sphere has attracted the most attention, driven by obvious applications in the earth sciences and meteorology. While there has long been some interest in function estimation on more complex domains, motivated by examples like estimating the pressure over the wing of an aircraft, research in this direction has recently received renewed impetus by the advent of 3D photography. The goal of 3D photography is to generate computer models of physical objects, reflecting their shape and other aspects of their appearance, such as color or reflectance. It is natural to think of color and reflectance as (not necessarily real valued) functions defined on the surface, which have to be estimated from physical measurements.

Our approach to function estimation on surfaces is a generalization of spline smoothing on the line (Wahba 1990). The *smoothing spline* with smoothing parameter λ is the function f_λ minimizing the *spline functional*

$$(1.1) \quad E[f] = \frac{1}{n} \sum (y_i - f(x_i))^2 + \lambda \int f''(x)^2 dx$$

in the space of functions with square integrable second derivative.

Using techniques from the calculus of variations it can be shown that for any $\lambda > 0$, f_λ is a cubic spline function with knots at the predictor values x_i . The parameter λ controls the tradeoff between smoothness of the estimate and fidelity to the data or, in statistical terms, the tradeoff between bias and variance. In the limit $\lambda \rightarrow 0$, the smoothing spline f_λ interpolates the data and therefore estimates f_{true} with small bias but (possibly) large variance. For $\lambda = \infty$, on the other hand, f_λ is the least squares straight line, which has low variance but may have a large bias. In practice, the optimal value of λ has to be estimated from the data, typically by (generalized) cross-validation (Wahba 1990).

Spline smoothing has been previously generalized to the plane, the flat torus, and the standard sphere, by replacing the single integral in Equation 1.1 with a so-called “thin-plate” energy term, which is a measure of roughness (Duchon 1977; Meinguet 1979; Wahba 1981; Wahba 1990). The theory of spline smoothing has been extended to arbitrary Riemannian manifolds (Narcowich 1995; Dyn *et al.* 1997; Kim 2001). However, these authors have not presented operational methods for actually computing such splines.

We extend the domain of spline smoothing to the class of subdivision surfaces. Subdivision surfaces can model complex shapes of arbitrary topology in a conceptually simple and parsimonious way; they are easy to represent and manipulate in the computer; and they support the notion of a “Sobolev space $W_2(\mathbb{R})$ of functions with square integrable second derivative” (Arden 2001).

In contrast to simple domains like the line or the sphere there are no closed form expressions for smoothing splines on subdivision surfaces. We therefore use the *finite element method* to approximate such splines. Similar ideas have recently been used to solve partial differential equations on subdivision surfaces (Cirak *et al.* 1999; Bajaj and Xu 2001).

Besides presenting a novel method for function estimation on surfaces, our paper has a substantial tutorial component, summarizing definitions, ideas, and results that are scattered through the literature. In Section 2 we use the simple setting of spline smoothing on the line to introduce multiresolution spaces of univariate spline functions and the finite element method for approximating smoothing splines. We also explain the idea of computing piecewise linear approximations to splines through subdivision. In Section 3 we discuss roughness measures for bivariate functions and introduce multiresolution spaces of polynomial splines in the plane and their subdivision rules. In Section 4 we consider roughness measures for functions on arbitrary compact surfaces. In Section 5 we generalize spline smoothing to surfaces parameterized over a planar region. In Section 6, we introduce subdivision functions, which are a generalization of splines, and subdivision surfaces. In Section 7 we give a precise definition of smoothness for functions on subdivision surfaces. Our principal contribution, a finite element method for spline smoothing on subdivision surfaces is described in Section 8. In Section 9 we present experimental results. Section 10, containing a review and discussion of related work and ideas for future research, concludes the paper.

2. UNIVARIATE SPLINE SMOOTHING

Suppose that we are given data $(x_1, y_1), \dots, (x_n, y_n) \in \mathbb{R}^2$, with $a = x_1 < x_2 < \dots < x_n = b$. The smoothing spline for smoothing parameter λ is the function f_λ minimizing the spline functional

$$(2.2) \quad E[f] = \frac{1}{n} \sum (y_i - f(x_i))^2 + \lambda \int_a^b f''(x)^2 dx$$

in the Sobolev space $W_2([a, b])$ of functions with square integrable second derivative. As pointed out in Section 1, for any $\lambda > 0$ there is a closed form solution to this optimization problem: f_λ is a cubic spline with knots at the x_i , and it can be found by linear algebra. We will throughout the paper assume that $\lambda > 0$.

We are interested in solving more general versions of the spline smoothing problem where the domain is not an interval, but an arbitrary, compact surface in \mathbb{R}^3 . In this situation it is no longer possible to exactly minimize (a generalized version of) the spline functional 2.2, and we will have to be content with approximations. We introduce some of the pertinent ideas here, in the context of smoothing on the line.

The basic idea is simple. Suppose that we have a *multiresolution sequence* of function spaces in $W_2([a, b])$, *i.e.* a nested sequence of finite dimensional subspaces

$$V^0 \subset V^1 \subset V^2 \subset \dots \subset W_2([a, b])$$

whose union is dense in $W_2([a, b])$. We can then compute approximations to the smoothing spline f_λ by choosing a resolution level J and restricting the functional E to V^J . These approximations converge to f_λ as J approaches infinity. Minimizing the spline functional over V^J is a finite dimensional optimization problem that can be solved by linear algebra.

2.1. Nested spaces of univariate splines. There are many ways of constructing such a multiresolution sequence. We will now discuss a particular choice that can be generalized to surfaces.

Let V^0 be the restriction of the space of cubic splines with knots at the integers to the interval $[a, b]$. This space is spanned by the integer translates of a single basis function $\phi_0^0(x)$, shown in Figure 2(d). The basis function $\phi_0^0(x)$ is the unique C^2 , piecewise cubic polynomial supported on the interval $-2 \leq x \leq 2$, such that it and its integer translates $\phi_\alpha^0(x) = \phi_0^0(x - \alpha)$ form a partition of unity:

$$\sum_{\alpha} \phi_\alpha^0(x) = 1.$$

The space V^J at resolution J is spanned by scaled and translated versions of ϕ_0^0 :

$$\phi_\alpha^J(x) = \phi_0^0(2^J x - \alpha).$$

The union of the V^J is dense in $W_2([a, b])$.

2.2. Computing finite element approximations of smoothing splines. To find an approximate minimum for $E[f]$, we choose a resolution level J and express $f(x)$ as a finite sum

$$f(x) = \sum_{\alpha=p_0}^{p_1} f^\alpha \phi_\alpha^J(x)$$

where the index ranges over the basis functions whose support intersects the interval $[a, b]$. Substituting into the formula for the spline functional $E[f]$ yields the identity

$$E[f] = \frac{1}{n} \sum_i \left(y_i - \sum_{\alpha} f^\alpha \phi_\alpha^J(x_i) \right)^2 + \lambda \sum_{\alpha, \beta} f^\alpha f^\beta B_{\alpha, \beta},$$

where the entries of the *penalty matrix* B are $B_{\alpha, \beta} = \int_a^b (\phi_\alpha^J)'' (\phi_\beta^J)'' dx$. This shows that the spline functional $E[f]$ restricted to V^J is a quadratic function of the coefficients f^α , and therefore the minimum can be found by solving the linear system

$$(2.3) \quad (X^t X + n \lambda B) \cdot \mathbf{f} = X^t \cdot \mathbf{y},$$

where

$$X = \begin{pmatrix} \phi_{p_0}^J(x_1) & \phi_{p_0+1}^J(x_1) & \dots & \phi_{p_1}^J(x_1) \\ \phi_{p_0}^J(x_2) & \phi_{p_0+1}^J(x_2) & \dots & \phi_{p_1}^J(x_2) \\ \vdots & \vdots & & \vdots \\ \phi_{p_0}^J(x_n) & \phi_{p_0+1}^J(x_n) & \dots & \phi_{p_1}^J(x_n) \end{pmatrix},$$

$\mathbf{f} = (f^{p_0}, \dots, f^{p_1})$, and $\mathbf{y} = (y_1, \dots, y_n)$.

2.3. Computing univariate splines by subdivision. We now describe a method for constructing piecewise linear approximations to functions in V^J . This method, called subdivision, dates back to the early days of computer graphics when computers were slow and had only integer arithmetic. While subdivision is no longer of great practical relevance for computing univariate splines, we introduce it here because we will later use it to generalize splines to more complex domains.

Note first of all that functions in V^J are scaled versions of functions in V^0 . We may therefore without loss of generality assume that $J = 0$.

Let $\hat{\phi}_0^0$ be the standard *hat function* shown in Figure 2(a). We denote the scaled and shifted versions of the standard hat function by $\hat{\phi}_\alpha^J$. Let $f = \sum_{\alpha} f^{0, \alpha} \hat{\phi}_\alpha^0(x)$ be a function in V^0 . The piecewise

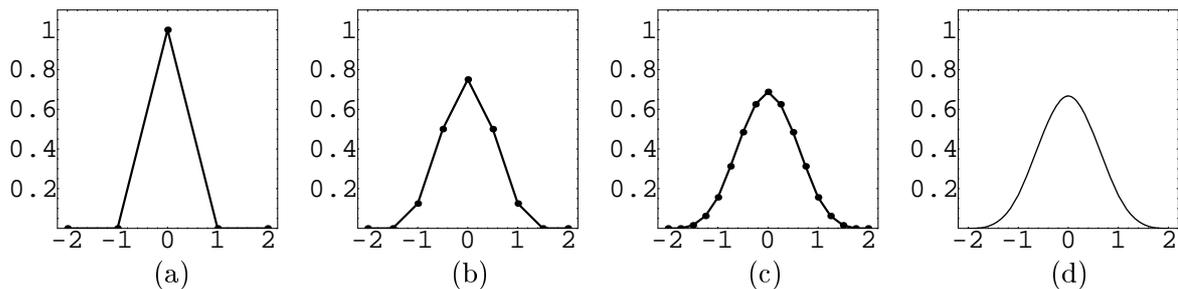


FIGURE 2. The subdivision process applied to the hat function $\hat{\phi}_0^0$ (a) yields a sequence of piecewise linear functions (b, c) that converges to the cubic spline basis function $\phi_0^0(x)$ (d).

linear function

$$f_{PL}^0(x) = \sum_{\alpha} f^{0,\alpha} \hat{\phi}_{\alpha}^0(x)$$

with vertices at $(\alpha, f^{0,\alpha})$ is called the *control net* of f at resolution level $J = 0$. As V^0 is a subset of V^1 , f can also be expressed in terms of the basis $\phi_{\alpha}^1(x)$:

$$f(x) = \sum_{\alpha} f^{1,\alpha} \phi_{\alpha}^1(x).$$

The corresponding control net is

$$f_{PL}^1(x) = \sum_{\alpha} f^{1,\alpha} \hat{\phi}_{\alpha}^1(x).$$

We can repeat this process to obtain a sequence of control nets $f_{PL}^J(x)$.

Because the basis functions of V^J can be expressed in terms of the basis functions of V^{J+1} , there are formulas, called *subdivision rules*, relating level J coefficients to level $(J + 1)$ coefficients or, equivalently, the control nets f_{PL}^J and f_{PL}^{J+1} :

$$f^{J+1,2\alpha+1} = \frac{1}{2} (f^{J,\alpha} + f^{J,\alpha+1}) \quad \text{and} \quad f^{J+1,2\alpha} = \frac{1}{8} f^{J,\alpha-1} + \frac{6}{8} f^{J,\alpha} + \frac{1}{8} f^{J,\alpha+1}.$$

As $J \rightarrow \infty$, the control nets f_{PL}^J converge to the spline f (Lane and Riesenfeld 1980). This is illustrated in Figure 2. Figure 2(a) shows the standard hat function $\hat{\phi}_0^0$, which by definition is the level $J = 0$ control net of the spline basis function ϕ_0^0 in Figure 2(d). Figures 2(b) and 2(c) show the level $J = 1$ and level $J = 2$ control nets of ϕ_0^0 , indicating that these control nets indeed converge to ϕ_0^0 .

Repeated subdivision of control nets is a computationally efficient way to approximate f . There is an alternative way of describing the subdivision process: Resample the level J control net at the grid points of the level $J + 1$ grid. Then smooth the resulting function on the level $J + 1$ grid by a moving average with weights $(1/4, 1/2, 1/4)$.

3. SPLINE SMOOTHING IN THE PLANE

To generalize spline smoothing from the line to the plane we need a measure of roughness to replace $(f'')^2$. To approximate smoothing splines with the finite element method we also need a nested sequence of function spaces whose union is dense in an appropriate Sobolev space playing the role of $W_2([a, b])$.

3.1. Measuring roughness of functions in the plane. For a bivariate function $f(x^1, x^2)$, the analogue of f'' is the Hessian

$$H_f = \begin{pmatrix} f_{x^1, x^1} & f_{x^1, x^2} \\ f_{x^1, x^2} & f_{x^2, x^2} \end{pmatrix}.$$

We want the roughness measure to be rotationally invariant. This limits the choice to the trace of powers of the Hessian $\text{tr}(H_f^k)$ and functions thereof (Weyl 1946). The trace, itself, is not useful, because the integral of the Laplacian $\Delta f = f_{x^1 x^1} + f_{x^2 x^2} = \text{tr}(H_f)$ vanishes for any f with compact support.

The next simplest functions are the square of the Laplacian of f :

$$(\text{tr}(H_f))^2 = (\Delta f)^2 = (f_{x^1 x^1} + f_{x^2 x^2})^2$$

and the standard “thin-plate” energy:

$$\text{tr}(H_f^2) = (f_{x^1 x^1})^2 + 2(f_{x^1 x^2})^2 + (f_{x^2 x^2})^2.$$

Either one could be used as a roughness penalty. In our exposition we choose the square of the Laplacian, although the techniques developed in the rest of the paper are equally applicable to the thin plate energy.

With this choice of penalty, the univariate spline smoothing functional generalizes to

$$(3.4) \quad E[f] = \frac{1}{n} \sum_i (y_i - f(x_i))^2 + \lambda \int_{\mathbb{R}^2} (\Delta f)^2 dx.$$

Note that, in contrast to the univariate case, the smoothing spline f_λ minimizing $E[f]$ is *not* a piecewise polynomial function.

3.2. Nested spaces of quartic triangular B-splines. There are many ways of generalizing the multiresolution sequence $V^0 \subset V^1 \subset \dots \subset W_2([a, b])$ of univariate spline spaces to the plane. For instance, we could use tensor products of univariate splines. For our purposes, it is more convenient to use the *quartic triangular* B-splines of Sabin (1976), which we now describe. See Loop (1992) or Loop and DeRose (1990) for a more detailed discussion.

Consider the hexagonal lattice of triangles shown in Figure 3(a). As in the univariate case, the space V^0 is spanned by translates of a basis function $\phi_0^0(x^1, x^2)$ to the vertices of the lattice. The basis function, shown in Figure 4, is the unique function satisfying the following conditions: (i) it is quartic polynomial on each triangle; (ii) it is C^2 ; (iii) it has minimum support; and (iv) its translates form a partition of unity.

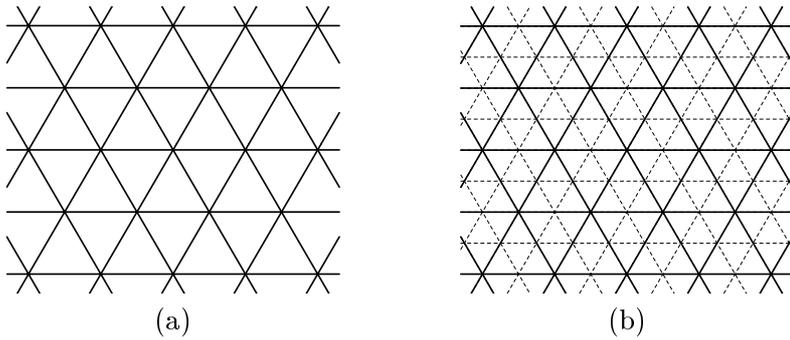


FIGURE 3. Panel (a): Hexagonal lattice in the plane; Panel (b): Lattice after one 4-1 split.

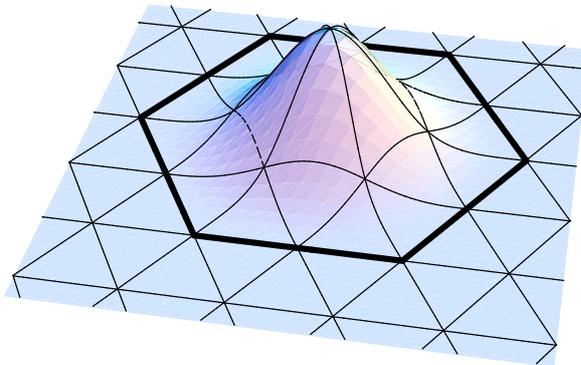


FIGURE 4. Hexagonal lattice with quartic triangular B-spline basis function ϕ_0^0 . The support of the basis function consists to the 24 triangles enclosed by the solid hexagon.

Again as in the univariate case, the space V^J at resolution J is spanned by the functions ϕ_α^J obtained by scaled and translated vesions of ϕ_0^0 :

$$\phi_\alpha^J(x^1, x^2) = \phi_0^0(2^J x^1 - \alpha^1, 2^J x^2 - \alpha^2).$$

3.3. Constructing quartic triangular B-splines by subdivision. As in the univariate case, we can construct a sequence of piecewise linear approximations to any function in V^0 (and therefore in V^J) using subdivision. We now review this process (for details see (Stollnitz *et al.* 1996 or Loop 1992). Let $\sum_\alpha f^{0,\alpha} \phi_\alpha^0$ be a function in V^0 , and let $f_{PL}^0 = \sum_\alpha f^{0,\alpha} \hat{\phi}_\alpha^0$ be the corresponding level $J = 0$ control net. The level $J = 0$ control net f_{PL}^0 is a piecewise linear function on the level $J = 0$ hexagonal lattice shown in Figure 3(a).

Because f can also be expressed in terms of level $J = 1$ basis functions: $f = \sum_\alpha f^{1,\alpha} \phi_\alpha^1$, the level 1 coefficients $f^{1,\alpha}$ can be expressed in terms of the level 0 coefficients, and therefore the level 1 control net $f_{PL}^1 = \sum_\alpha f^{1,\alpha} \hat{\phi}_\alpha^1$ can be expressed in terms of the level 0 control net.

In close analogy to the univariate case, we can obtain f_{PL}^1 from f_{PL}^0 by an interpolation step followed by a smoothing or averaging step. We first interpolate the values of f_{PL}^0 to the new (edge) vertices of the level 1 lattice shown in Figure 3(b). We then replace the value at each vertex of the level 1 lattice by a weighted average of its neighbors. Let v_0 be a vertex of the lattice, and let v_1, \dots, v_n be its neighbors. Then

$$(3.5) \quad f_{PL}^{J+1}(v_0) = \frac{1}{4}f_{PL}^J(v_0) + \frac{1}{8} \sum_i f_{PL}^J(v_i).$$

The same rules can be used to obtain the level $(J + 1)$ control net from the J control net. The sequence of control nets converges to f , so repeated subdivision of controls nets is a computationally efficient way to approximate f .

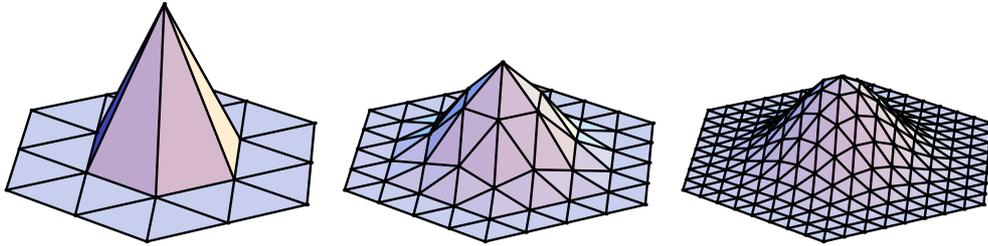


FIGURE 5. The subdivision process applies to the hat function $\hat{\phi}_0^0$ (left) yields a sequence of piecewise linear functions that converges to the quartic triangular B-spline basis function ϕ_0^0 (right).

4. MEASURING ROUGHNESS OF FUNCTIONS DEFINED ON SMOOTH SURFACES

Let $M \subset \mathbb{R}^3$ be a smooth surface. A standard way of defining a roughness measure for functions on M is to extend the concepts of the Hessian and the Laplace operator.

Consider a point $p \in M$. Without loss of generality, we may assume that p is the origin and the tangent plane at p is the (u^1, u^2) -plane. Then near p the surface has the local parameterization

$$M = \{x : x = \mathbf{F}(u^1, u^2) = (u^1, u^2, F(u^1, u^2))\}$$

where $F(u^1, u^2)$ is a scalar function with vanishing value and gradient at the origin, and any function on M can be regarded as a function $f(u^1, u^2)$.

The *Hessian* of f at p is just the matrix of second derivatives of f with respect to u^1 and u^2 at the origin, and the value of the *Laplace-Beltrami* operator Δ_M at p is the trace of the Hessian.

4.1. Computing the Laplace-Beltrami operator for a parametrized surface. Although conceptually simple, the definition of the Hessian and Laplace-Beltrami operator given above is not

useful operationally: we need explicit formulas expressing them in terms of a general parameterization

$$\mathbf{F} : \mathbb{R}^2 \rightarrow M \subset \mathbb{R}^3 : (u^1, u^2) \mapsto \mathbf{F}(u^1, u^2).$$

To state such formulas, we introduce some standard notation from differential geometry (Do Carmo 1976). Let $T_p M \subset \mathbb{R}^3$ be the tangent space to M at p . The partial derivatives $\mathbf{F}_{u^a}(u^1, u^2)$, $a = 1, 2$ form a basis of the tangent space. The *first fundamental form* is the symmetric matrix of inner products

$$(4.6) \quad I = \begin{pmatrix} g_{1,1} & g_{1,2} \\ g_{2,1} & g_{2,2} \end{pmatrix},$$

where

$$(4.7) \quad g_{a,b} = \langle \mathbf{F}_{u^a}, \mathbf{F}_{u^b} \rangle.$$

The quantities $g^{a,b}$ denote the entries in the inverse matrix of I^{-1} . The *Riemann-Christoffel symbols* are the quantities

$$\Gamma_{ab}^c = \sum_{d=1}^2 g^{c,d} \Gamma_{ab,d},$$

where

$$\Gamma_{ab,c} = \langle \mathbf{F}_{u^a, u^b}, \mathbf{F}_{u^c} \rangle.$$

The Hessian of f is the quadratic form

$$H_f : T_p M \times T_p M \rightarrow \mathbb{R},$$

defined by the formula

$$(4.8) \quad H_f(X, Y) = \begin{pmatrix} X^1 & X^2 \end{pmatrix} \cdot \begin{pmatrix} f_{|1,1} & f_{|1,2} \\ f_{|1,2} & f_{|2,2} \end{pmatrix} \cdot \begin{pmatrix} Y^1 \\ Y^2 \end{pmatrix},$$

where $f_{|a,b}$ is the so-called *covariant derivative*

$$f_{|a,b} = f_{u^a, u^b} - \sum_{c=1}^2 \Gamma_{a,b}^c f_{u^c}.$$

In these formulas we have abused notation and replaced $f \circ \mathbf{F}$ by f .

As in the Euclidean case, the invariants of H_f can be expressed in terms of the trace (Weyl 1946). But because the vectors \mathbf{F}_{u^a} do not form an orthonormal basis of the tangent space, the trace of any quadratic form $B : T_p M \times T_p M \rightarrow \mathbb{R}$ is given by the formula

$$\text{tr}(B) := \sum_{a,b} g^{a,b} B_{a,b}.$$

The *Laplace-Beltrami* operator is the trace of the Hessian and is given by

$$\Delta_M f = \sum_{a,b} g^{a,b} (f_{u^a, u^b} - \Gamma_{a,b}^c f_{u^c}).$$

In the Euclidean case, where $\mathbf{F}(u^1, u^2) = (u^1, u^2, 0)$, the matrix I is the identity matrix and the Riemann-Christoffel symbols vanish, and therefore the Laplace-Beltrami operator reduces to the Laplace operator.

4.2. Integration over a surface. Having generalized the measure of roughness to functions on surfaces, we now need to generalize the notation of integration. The metric on \mathbb{R}^3 induces an area measure on M , which we denote by dA . In local coordinates dA can be expressed in two ways:

$$dA = \sqrt{\det(I)} du^1 du^2 = |\mathbf{F}_{u^1} \times \mathbf{F}_{u^2}| du^1 du^2.$$

5. SPLINE SMOOTHING ON SURFACES PARAMETERIZED OVER A PLANAR REGION

We now consider spline smoothing in the case where the topology of the surface is simple but the geometry is non-Euclidean (Greiner 1994). Let $\Omega \subset \mathbb{R}^2$ be a region in the plane, and suppose that M is given by an embedding $\mathbf{F} : \Omega \rightarrow \mathbb{R}^3$ of this region into \mathbb{R}^3 . The data is a collection $\{(x_i, y_i) \in M \times \mathbb{R}\}$, where $x_i = \mathbf{F}(u_i^1, u_i^2)$, $1 \leq i \leq n$.

We want to find the function on M that minimizes the functional

$$E[f] = \frac{1}{n} \sum_i (y_i - f(x_i))^2 + \lambda \int_M (\Delta_M f)^2 dA,$$

where Δ_M is the Laplace-Beltrami operator defined in Section 4. Since \mathbf{F} is a diffeomorphism, f can be regarded as a function of (u^1, u^2) , so $E[f]$ assumes the form

$$E[f] = \frac{1}{n} \sum_i (y_i - f(u_i^1, u_i^2))^2 + \lambda \int_{\Omega} (\Delta_M f(u^1, u^2))^2 dA.$$

We call the function f_λ minimizing $E[f]$ a smoothing spline on M .

As in Section 2 we can use the finite element method to approximate f_λ . We choose a resolution level J and minimize $E[f]$ over functions in V^J . Kowalski and Krzysztof (1990) show that the union of the spline spaces V^J defined in Section 3.2 is dense in $W_2(\mathbb{R}^2)$. They also give results on approximation order. Therefore, for sufficiently large J our solution $f_{\lambda,J}$ will be close to $f_\lambda = \operatorname{argmin}_{f \in W_2(\mathbb{R}^2)} E[f]$.

We express f as a finite sum

$$(5.9) \quad f(u^1, u^2) = \sum_{\alpha} f^{\alpha} \phi_{\alpha}^J(u^1, u^2)$$

where the indices range over all basis functions whose support intersects Ω . Substituting into the formula for $E[f]$ yields the identity

$$(5.10) \quad E[f] = \frac{1}{n} \sum_i \left(y_i - \sum_{\alpha} f^{\alpha} \phi_{\alpha}^J(u_i^1, u_i^2) \right)^2 + \lambda \sum_{\alpha, \beta} f^{\alpha} f^{\beta} B_{\alpha, \beta}$$

where

$$(5.11) \quad B_{\alpha, \beta} = \int_{\mathbb{R}^2} \Delta_M(\phi_{\alpha}^J) \Delta_M(\phi_{\beta}^J) dA.$$

As in the univariate case, $E[f]$ is a quadratic function in the coefficients of f .

To see what is involved in computing $B_{\alpha, \beta}$, recall the definitions of Δ_M and dA :

$$\Delta_M f = \sum_{a, b} g^{a, b} (f_{u^a, u^b} - \Gamma_{a, b}^c f_{u^c}) \quad \text{and} \quad dA = |\mathbf{F}_{u^1} \times \mathbf{F}_{u^2}| du^1 du^2.$$

Consequently the integral involves derivatives of both the basis functions and the function \mathbf{F} defining the surface. Even in the simple case where \mathbf{F} is polynomial, the integrand is a complicated algebraic expression that cannot be integrated analytically, so we have to resort to numerical quadrature.

6. SUBDIVISION FUNCTIONS: A GENERALIZATION OF SPLINES TO TOPOLOGICALLY COMPLEX DOMAINS

In the previous section we considered estimation of functions defined on a surface which is an embedding of a planar region. We now want to move to more general domains, while preserving as much as possible the flavor of the previous section.

First observe that to represent a surface by an embedding of a parameter domain into \mathbb{R}^3 , the parameter domain has to be of the same topological type as the surface. We choose parameter domains, called *simplicial surfaces*, that are generalizations of the triangular lattice in the plane. This allows us to construct generalizations of quartic triangular B-splines, called *subdivision functions*, and immediately leads to a multiresolution sequence of function spaces.

6.1. Simplicial surfaces. Roughly speaking, a simplicial surface is a union of triangles glued together along their edges. More formally, a *simplicial complex* K consists of a (finite) set $\text{Vert}(K)$, called the *vertices* of K , and a collection of non-empty subsets of $\text{Vert}(K)$, called *simplices*, such that (i) each of the one point sets $\{v\}$, for $v \in \text{Vert}(K)$, is a simplex and (ii) if σ is a simplex of K and σ' is a non-empty subset of σ then σ' is also a simplex of K . A simplex containing $q + 1$ vertices is called a *q-simplex*, and it is customary to identify the 0-simplex $\{v\}$ with the vertex v , itself (i.e. $\{v\} = v$).

The *topological realization* of a simplicial complex K with m vertices is the topological subspace $|K| \subset \mathbb{R}^m$ defined as follows: Identify the vertices of K with the canonical basis vectors of \mathbb{R}^m . If $\sigma = \{v_0, \dots, v_q\}$ is a q -simplex of K let $|\sigma| \subset \mathbb{R}^m$ denote the convex hull of its vertices. Then

$$|K| = \cup_{\sigma \in K} |\sigma| .$$

If σ is a 1-simplex, we will sometimes call $|\sigma|$ an *edge* of $|K|$. And if σ is a 2-simplex, we will call $|\sigma|$ a *triangle* of $|K|$. A *simplicial surface* is a simplicial complex whose topological realization is homeomorphic to a compact, two dimensional manifold. For simplicity, we only consider manifolds without boundary.

6.2. Subdivision functions on simplicial surfaces. Notice that, just as we did in the plane, we can subdivide a simplicial surface by performing four-to-one splits on each of its faces. This suggests that we can formally extend the subdivision construction for splines on the plane to any simplicial surface. We now describe the basic idea. For a more detailed exposition see Reif (1995), Stollnitz *et al.* (1996), and Zorin (1998). Suppose K is a simplicial complex, and let K^J be the J times subdivided complex, with $K^0 = K$. We can identify the topological realizations $|K^J|$ with $|K|$ in the obvious way. From now on we will also abuse notation and no longer distinguish between a simplicial complex and its topological realization; it will be clear from context which one is meant.

In parallel to the construction of spline functions on the plane, we construct a subdivision function as the limit of a sequence of piecewise linear functions or *control nets*

$$f_{PL}^J : K \rightarrow \mathbb{R},$$

where f_{PL}^J is a piecewise linear function on the subdivided complex K^J .

We now describe the *subdivision rules* that define f_{PL}^{J+1} in terms of f_{PL}^J . Because f_{PL}^{J+1} is piecewise linear with respect to K^{J+1} , we only have to specify its values on the vertices of K^{J+1} . The vertices of K^{J+1} consist of the vertices of K^J and the new *edge vertices* introduced by subdivision.

We proceed in close analogy to the planar case. We first interpolate the values of f_{PL}^J to the new (edge) vertices of K^{J+1} . We then replace the value at each vertex of K^{J+1} by a weighted average of its neighbors. For vertices of valence 6, the weights are exactly the same as in the planar case. Note, however, that an argument using Euler's formula shows that any triangulation of a surface other than the torus must have vertices of valence other than 6; such vertices are called *extraordinary vertices*.

Specifying the averaging rules at extraordinary vertices of K^{J+1} is a delicate matter. We use the rules invented by C. Loop (Loop 1987, Loop 1992, Loop-DeRose 1990). Let v_0 be a vertex of valence n , and let v_1, \dots, v_n be its set of neighboring vertices, which are edge vertices introduced by subdivision. Then

$$(6.12) \quad f_{PL}^{J+1}(v_0) = \omega(n)f_{PL}^J(v_0) + \frac{(1 - \omega(n))}{n} \sum_i f_{PL}^J(v_i),$$

where

$$\omega(n) = \frac{(3 + 2 \cos \frac{2\pi}{n})^2 - 8}{32}$$

Notice that for vertices of valence $n = 6$, the formula assigns weight $1/4$ to the central vertex and weight $1/8$ to the surrounding vertices, as in the planar case.

Loop (1987) shows that for any choice of the initial control net f_{PL}^0 , the sequence of control nets f_{PL}^J converges to a continuous function on K . See also Reif (1995) and Zorin (1998). We denote the linear space of all such limits by V^0 . More generally, we can specify an initial control net at any level subdivision level J and compute the limit of the sequence, f_{PL}^{J+k} . Again the sequence converges to a continuous function and the set of all such functions forms a vector space V^J , which we call the space of *level J subdivision functions*. By construction, these spaces are nested:

$$V^0 \subset V^1 \subset \dots \subset V^J \subset \dots$$

In parallel with the planar case, each of the spaces V^J is spanned by a set of basis functions ϕ_α^J , where the index α ranges over the vertices of K^J . The basis function ϕ_α^J is the limit of the sequence of control nets generated by the hat function $\hat{\phi}_\alpha^J$ centered at the vertex α .

The basis functions have the following easily verified properties:

- They are non-negative: $\phi_\alpha^J \geq 0$.
- They form a partition of unity: $\sum_\alpha \phi_\alpha^J = 1$.

- The support of ϕ_α^J is contained in the “two-ring” of triangles of K^J centered at α .

Moreover, away from the extraordinary vertices, subdivision functions at any level are locally quartic triangular B-splines. To see this, consider a subdivision function $f \in V^J$. Recall that f is the limit of a sequence of control nets:

$$f = \lim_{k \rightarrow \infty} f_{PL}^{J+k},$$

and that f_{PL}^{J+k+1} depends only on f_{PL}^{J+k} for all k . Notice also that the values of f on any triangle of K^{J+k} depend only on the values of f_{PL}^{J+k} at the vertices of the one-ring of K^{J+k} surrounding the triangle. Choose a point p of K that is not an extraordinary vertex (Figure 6(a)). After a sufficiently large number of subdivisions, the one-ring of triangles surrounding p will only contain vertices of valence 6 (Figure 6(b)) and it therefore can be mapped (in an essentially unique way) into the regular hexagonal lattice of equilateral triangles in the plane (Figure 3). From then on, the subdivision process on the triangle containing the point coincides with the subdivision process in the plane, and the latter converges to a quartic triangular B-spline.

6.3. Representing submanifolds of \mathbb{R}^3 by subdivision surfaces. A *subdivision surface* M is the image of an embedding

$$\mathbf{F} : K \rightarrow \mathbb{R}^3$$

where the components of \mathbf{F} are subdivision functions.

Rather than treat each component of \mathbf{F} separately, we view \mathbf{F} as the limit of a sequence of vector-valued control nets

$$\mathbf{F}_{PL}^J : K \rightarrow \mathbb{R}^3.$$

Because \mathbf{F}_{PL}^J is piecewise linear on the subdivided complex K^J , it is determined by its values on the vertices of K^J . The image of \mathbf{F}_{PL}^J is a polyhedron $M^J \subset \mathbb{R}^3$ with triangular faces, which we call the *control mesh* of \mathbf{F} at level J . (See Figure 6.)

We can picture the subdivision process as operating on the initial control mesh M^0 . At each subdivision step, we first perform four-to-one splits on each of the triangles of the current control mesh; we then reposition the vertices of the subdivided mesh according to the subdivision rules described in Section 6.2.

7. SMOOTHNESS PROPERTIES OF SUBDIVISION FUNCTIONS

Subdivision functions are defined on simplicial surfaces, which are 2-dimensional piecewise linear manifolds. So it is not obvious what it means for a function on a simplicial surface to be differentiable. In our context, however, we clearly need a notion of smoothness to make sense of expressions like $\Delta_M f$.

7.1. Definition of smoothness for functions on topological surfaces. To address this issue, we first review how smoothness of functions on a topological manifold M is defined. Let $U \subset M$ be an open set. A *chart* (or *local coordinate system*) is a homeomorphism

$$\varphi : U \rightarrow \mathbb{R}^2$$

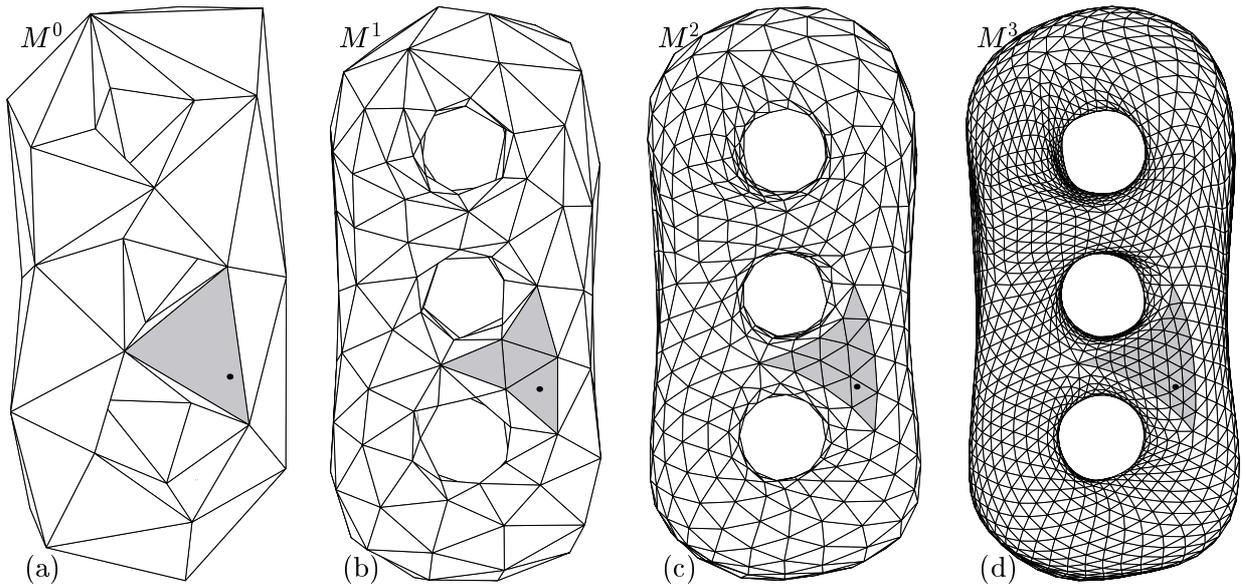


FIGURE 6. The subdivision surface $M = \mathbf{F}(K)$ is the limit of the sequence of control meshes $M^J = \mathbf{F}_{PL}^J(K)$

onto an open subset of \mathbb{R}^2 . The component functions $\varphi^a : U \rightarrow \mathbb{R}$, $a = 1, 2$ are called *coordinate functions* or *local coordinates*. A finite *atlas* for M is a finite collection

$$\{\varphi_\alpha : U_\alpha \rightarrow \mathbb{R}^2 : \alpha = 1, \dots, m\},$$

of charts such that $\{U_\alpha\}$ is an open cover of M . The maps

$$\varphi_{\alpha,\beta} = \varphi_\beta \circ \varphi_\alpha^{-1} : \varphi_\alpha(U_\alpha \cap U_\beta) \rightarrow \varphi_\beta(U_\alpha \cap U_\beta)$$

are called *transition functions*. We say that the atlas is of class C^k if the transition functions are all of class C^k . A topological surface with a C^k atlas is called a C^k -surface, and the atlas itself is sometimes referred to as a C^k structure on M . Similar definitions apply to any smoothness class of functions defined on regions in the plane. For instance, a $C_{loc}^{k,1}$ -structure is an atlas whose transition functions lie in the space $C_{loc}^{k,1}$ of C^k functions whose k -th order derivatives are locally Lipschitz.

Let M be a C^k surface. A C^r -function $f : M \rightarrow \mathbb{R}^q$, $r \leq k$, is a function such that $f \circ \varphi_\alpha^{-1} : \varphi_\alpha(U_\alpha) \rightarrow \mathbb{R}^q$ is C^r for all α . Let $f : M \rightarrow \mathbb{R}^q$ be a C^r -function. The *rank* of f at a point $p \in M$ is defined as follows. Let $\varphi : U \rightarrow \mathbb{R}^2$ be a C^k coordinate chart with $p \in U$, and let $\bar{f} = f \circ \varphi^{-1}$. The rank of f at p is defined to be the rank of the derivative $d\bar{f} : \mathbb{R}^2 \rightarrow \mathbb{R}^q$ at $\varphi(p)$. A C^r , $r > 0$ function $f : M \rightarrow \mathbb{R}^q$ is called an *immersion* if f has rank 2 at all points. An injective immersion is called an *embedding*.

7.2. Definition of smoothness for functions on simplicial surfaces. To define the notation of smoothness on a simplicial surface K we need to define a suitable atlas on K . Our choice of atlas has to produce a definition of smoothness that agrees with the obvious one on the interiors of triangles. We can accomplish this by including in the atlas the affine maps from the interior of each triangle to the interior of an equilateral triangle in the plane. Notice that because each edge of K is adjacent to exactly two triangles, we can define charts on the interior of the union of any two

adjacent triangles by mapping it in the obvious way to the interior of the diamond shaped region formed by two adjacent equilateral triangles.

The charts that we have defined so far cover all of K , except the vertices. The neighborhood of each valence 6 vertex naturally maps to the regular hexagon made of six equilateral triangles. This chart is obviously compatible with the charts defined previously. Moreover, because the subdivision functions are piecewise polynomial away from extraordinary vertices, they are C^2 functions with respect to the charts so far defined.

To complete our construction to a C^2 -atlas, we have to define charts in the neighborhood of the extraordinary vertices. The simple recipe of mapping the neighborhood of a valence n extraordinary vertex to a regular n -gon in the plane is incompatible with the previously defined charts. A suitable completion of the atlas was given by Schweitzer (1996) and Zorin (1996), using results of Reif (1995). See also Arden (2001).

For each vertex v the *characteristic chart* is a map $\varphi_v : U_v \rightarrow \mathbb{R}^2$, where U_v is the interior of the union of the triangles of K containing v . Characteristic charts are defined in terms of subdivision functions; for details see Arden (2001) and Zorin (1996). Figure 7 shows a characteristic chart for the base complex modeling the statuette of the elephant.

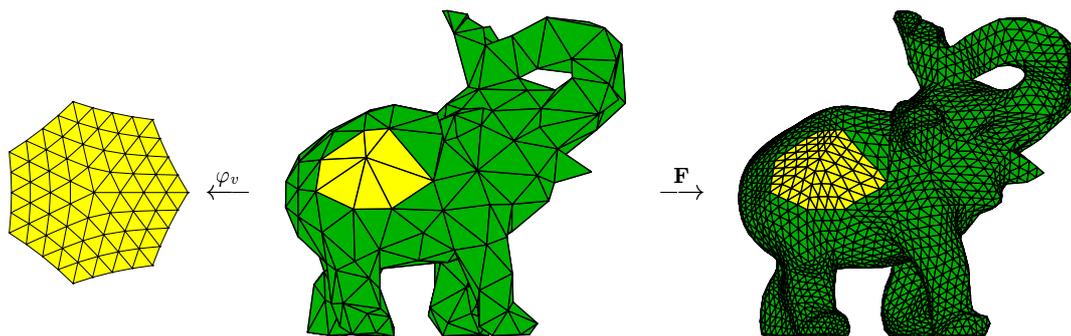


FIGURE 7. The figure above illustrates the subdivision surface that models the statuette of the elephant. The domain U_v about a vertex v of valence 7 is the interior of the union of triangular faces adjacent to v (grey region in center). The image of the characteristic chart $\varphi_v : U_v \rightarrow \mathbb{R}^2$ is a region in the plane (left). The surface, itself, is the image of the map $\mathbf{F} : K \rightarrow \mathbb{R}^3$ (right). A function $f : K \rightarrow \mathbb{R}^n$ is defined to be a C^r function if $f \circ \varphi_v^{-1}$ is smooth for each vertex v . The function \mathbf{F} is C^1 everywhere and C^2 away from extraordinary vertices.

This atlas has several useful properties: (i) Subdivision functions are of class C^1 (and in a sense that can be made precise) almost C^2 ; (ii) Subdivision functions have square integrable second derivatives; (iii) For generic control values, subdivision maps from K into \mathbb{R}^3 have maximal rank, making them suitable for modeling smooth surfaces. Because the atlas is of class C^2 , the notion of the Sobolev space $W_2(K)$ of functions on K with square integrable derivatives up to order 2 is well-defined (Arden 2001). Arden also shows that V^J is a subset of $W_2(M)$ for all J and that $\cup_J V^J$ is dense in $W_2(M)$.

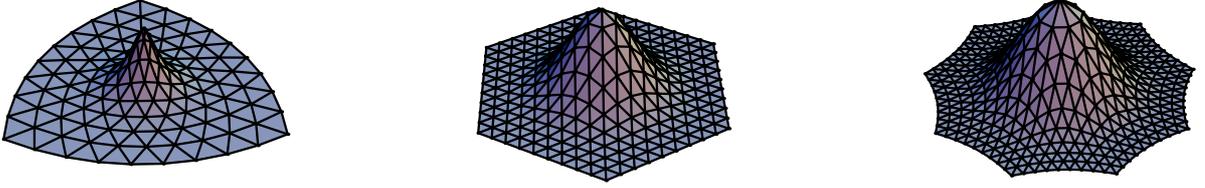


FIGURE 8. The graph of the function $\phi_v^1 \circ \varphi_v^{-1}$ centered at vertices of valence 3, 6 and 9, respectively. Notice the particularly high curvature of the graph of the basis function centered at a vertex of valence 3.

8. SPLINE SMOOTHING ON SUBDIVISION SURFACES

Suppose we are given a subdivision surface $M \subset \mathbb{R}^3$, which is the image of an embedding $\mathbf{F} : K \rightarrow \mathbb{R}^3$, together with a collection of n observations $(x_i, y_i) \in M \times \mathbb{R}$. Each point x_i is the image of a point $u_i \in K$.

As in Section 3, a *smoothing spline* is a function on K that minimizes the functional

$$(8.13) \quad E[f] = \frac{1}{n} \sum_i (y_i - f(u_i))^2 + \lambda \int_K (\Delta_M f)^2 dA.$$

One can show that there is a unique smoothing spline f_λ for each choice of $\lambda > 0$. An *interpolating spline* is the limit $f_0 = \lim_{\lambda \rightarrow 0} f_\lambda$ (Dyn *et al.* 1997; Kim 2001).

We can use the finite element method to compute approximate smoothing splines exactly as described in Section 2.2. We have a nested sequence of spaces of subdivision functions

$$V^0 \subset V^1 \subset \dots \subset V^J \subset \dots \subset W_2(K)$$

whose union is dense in $W_2(K)$. We choose a resolution level J and restrict E to V^J . Substituting the expansion

$$f = \sum_{\alpha=1}^p f^\alpha \phi_\alpha^J$$

into the formula for $E[f]$ and simplifying yields the expression

$$(8.14) \quad E[f] = \frac{1}{n} \sum_i \left(y_i - \sum_{\alpha=1}^p f^\alpha \phi_\alpha^J(u_i) \right)^2 + \lambda \sum_{\alpha, \beta} f^\alpha f^\beta B_{\alpha, \beta},$$

where

$$(8.15) \quad B_{\alpha, \beta} = \int_K \Delta_{\mathbf{F}}(\phi_\alpha^J) \Delta_{\mathbf{F}}(\phi_\beta^J) dA.$$

As before this is a quadratic function in the coefficients f^α .

8.1. **Evaluating the entries of the penalty matrix B .** We want to evaluate integrals of the form

$$B_{\alpha,\beta} = \int_K \Delta_M(\phi_\alpha^J) \Delta_M(\phi_\beta^J) dA.$$

As already pointed out in Section 5, this has to be done by numerical quadrature. We subdivide the base complex K a fixed number of times beyond the subdivision level J of the finite element approximation. Let J_{quad} be the subdivision level. The integral is then the sum of integrals over triangles of the subdivision:

$$B_{\alpha,\beta} = \sum_{T \in K^{J_{quad}}} \int_T \Delta_M(\phi_\alpha^J) \Delta_M(\phi_\beta^J) dA.$$

On each triangle T , we approximate the Laplacians $\Delta_M \phi_\alpha^J$ and $\Delta_M \phi_\beta^J$ by linearly interpolating their values at the vertices, and we assume that the area density $dA = |\mathbf{F}_{u^1} \times \mathbf{F}_{u^2}| du^1 du^2$ is constant on T . We then integrate the approximation.

There is a slight complication when one of the vertices of T is an extraordinary vertex: the map \mathbf{F} is singular at such a vertex, and therefore the Hessian and the Laplace-Beltrami operator are not defined. This is not a fundamental problem because we know that the Laplacian of every subdivision function is square integrable (Arden 2001) and therefore the integral is well defined. Rather than ignoring such triangles altogether, we evaluate the Laplacian at the barycenter of T and linearly extend to the extraordinary vertex. Note that the only *extraordinary vertices* of K^J are the extraordinary vertices of the original complex K , because every vertex introduced by subdivision has valence 6.

This process is not as computationally intensive as it might seem at first glance because the support the basis functions, on average, is contained in only 24 triangles and the number of triangles of K^J is 4^J times the number of triangles of K .

9. EXPERIMENTAL RESULTS

We present the results of two experiments that were conducted with different goals in mind. In the first experiment we applied the finite element method to a smoothing problem on the unit sphere where the smoothing spline defined by Equation (8.13) can be found exactly (Wahba 1990; Wendelberger 1982). The goals of this experiment were to validate our code and to understand how the choice of subdivision level influences the accuracy of the finite element approximation.

The goal of the second experiment was to demonstrate the use of spline smoothing for estimating functions defined on the elephant shown in Figure 1(b) and to provide some evidence that generalized cross-validation (Wahba 1990) is an effective way for choosing the optimal smoothing parameter λ .

Computing finite element approximations to smoothing splines requires solving large systems of linear equations. For our examples we used a preconditioned conjugate gradient method (Golub and Van Loan 1996, Section 10.3). However, the subdivision framework has an intrinsic multi-resolution structure, making it amenable to adaptive methods (Grinspun *et al.* 2002).

9.1. Comparison of exact smoothing splines and finite element approximations on the sphere. Our finite element approach to spline smoothing assumes that the estimation domain is a subdivision surface. We therefore approximated the unit sphere S^2 by a subdivision surface M . The control mesh of M has the connectivity of a five times subdivided tetrahedron. Using a procedure similar to the one described in Hoppe *et al.* (1994) we positioned the vertices of the control mesh to best approximate the sphere. The approximation is extremely close: the maximum radial deviation is only about 0.05%.

We then generated 100 random points x_1, \dots, x_{100} uniformly distributed over the sphere and 100 random y -values y_1, \dots, y_{100} from a standard Gaussian distribution, and computed the exact interpolating spline f_{true} .

Next we generated test samples $(x_1, y_1), \dots, (x_n, y_n)$ for $n = 100, 200, 400$. The x_i were again uniformly distributed over the sphere, and $y_i = f_{true}(x_i)$. By radially projecting the x_i onto M we obtained corresponding data sets on the subdivision surface M .

For each of the three test samples we found the exact smoothing splines f_λ for a set of values of λ spanning essentially the entire range between interpolation and averaging. We also computed the finite element approximations $f_{\lambda, J}$ for subdivision levels $J = 3, 4, 5, 6$. In all cases we used subdivision level $J_{quad} = 7$ for computing the entries of the penalty matrix B . (Level $J_{quad} = 7$ was the largest subdivision level that led to computational problems tractable on the hardware available to us.)

Our goal was to understand the dependence of the approximation error on the subdivision level and on the amount of smoothing. The latter is controlled by the smoothing parameter λ . However, λ is not a very “intuitive” parameter; for example, when presented with a dataset and a smoothing spline, one would be hard pressed to guess even the order of magnitude of λ . A more intuitive measure of the amount of smoothing is the *effective number of parameters* or the *number of degrees of freedom* of the spline. Spline smoothing is a linear operation, meaning the vector $\hat{\mathbf{y}}$ of fitted values depends linearly on the observed response vector \mathbf{y} : $\hat{\mathbf{y}} = H \mathbf{y}$ for some matrix H . The number of degrees of freedom is defined as the trace of H (see Hastie and Tibshirani 1990, Chapter 3.5). This definition of “degrees of freedom” agrees with the usual one in the case of linear least squares, where H is a projection matrix.

Figure 9 shows relative approximation error

$$\frac{\|f_\lambda - f_{\lambda, J}\|_2}{\|f_\lambda\|_2} = \frac{(\int_M |f_\lambda(x) - f_{\lambda, J}(x)|^2 dA)^{1/2}}{(\int_M |f_\lambda(x)|^2 dA)^{1/2}}$$

as a function of the number of degrees of freedom used by the exact spline, for test sample size $n = 200$.

For fixed subdivision level, the error increases as the number of degrees of freedom increases and we get closer to interpolation. This is not surprising; after all, the approximations $f_{\lambda, J}$ lie in a finite dimensional function space and therefore cannot be expected to reproduce the “high frequency” components of f_λ . Those high frequency components get bigger as we get closer to interpolation, leading to an increase in error.

For fixed smoothing parameter, on the other hand, we see that the error decreases exponentially with the subdivision level. This is in agreement with Theorem 2 of Arden (2001). There are two unusual points in the plot, corresponding to subdivision level 6, and 10 and 25 degrees of freedom, where the error is larger than one would expect given the pattern in the rest of the plot. We conjecture that this is due to insufficient accuracy in evaluating the entries of the penalty matrix, and that the abnormality would disappear if we switched to subdivision level $J_{quad} = 8$ for computing the entries of B .

The results for test sample sizes $n = 100$ and $n = 400$ look very similar to Figure 9; we do not present them here.

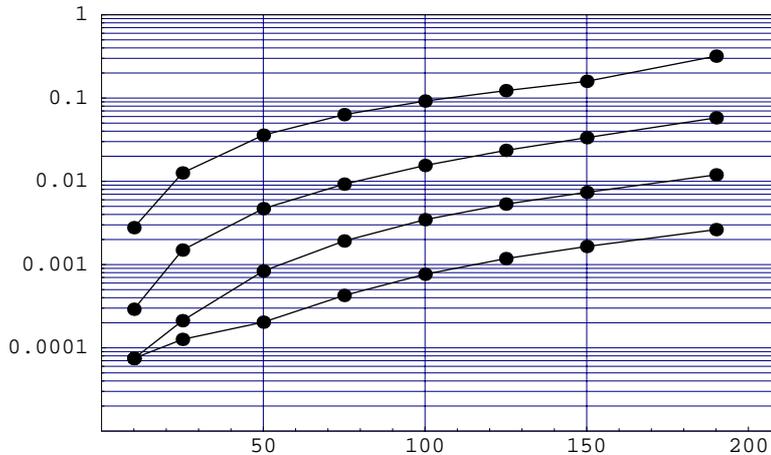


FIGURE 9. Relative error of approximate smoothing splines $f_{\lambda, J}$ for $J = 3, 4, 5, 6$ (top to bottom), as a function of the number of degrees of freedom used by the exact spline f_{λ} .

9.2. Spline smoothing on the elephant. We tested our procedure on a free form surface constructed from laser scans of a ceramic statuette of an elephant (Figure 1(b)). The elephant's trunk loops around and touches the top of the head; the surface therefore has the topology of a torus. Starting from an initial mesh with approximately 300,000 faces, we used a MAPS-like simplification procedure (Wood *et al.* 2000; Lee *et al.* 1998) to construct a simplicial complex K with 400 faces. We then approximated the initial mesh by the image $M = \mathbf{F}(K)$ of a level-3 subdivision function $\mathbf{F} : K \rightarrow \mathbb{R}^3$.

9.2.1. Construction of random functions on the elephant. We now describe how we generated random functions on the elephant with varying degrees of smoothness. The basic idea is to first determine the analogue of a Fourier basis for functions on M , and then form a Fourier series with random coefficients. Recall that the Laplace-Beltrami operator Δ_M is a self-adjoint elliptic operator. It follows from the general theory of such operators (Warner 1983) that Δ_M has a complete orthonormal set of eigenfunctions:

$$\begin{aligned} \Delta_M \psi_k + \lambda_k \psi_k &= 0 \\ \lambda_0 = 0 &\leq \lambda_1 \leq \lambda_2 \leq \dots \end{aligned}$$

with $\psi_0 = 1$, where we have normalized the area element so that M has unit area. The eigenfunctions ψ_k are the generalization of spherical harmonics to arbitrary compact surfaces. Notice that

$$\int_K |\Delta_M \psi_k|^2 dA = \lambda_k^2;$$

hence, the roughness of ψ_k increases with increasing k .

We approximate the eigenfunctions ψ_k by subdivision functions. Let p_J denote the dimension of the space V^J spanned by the level J basis functions ϕ_α^J , and let A and C be the $p_J \times p_J$ matrices defined by

$$A_{\alpha\beta} = \int_K \phi_\alpha^J \phi_\beta^J dA$$

and

$$C_{\alpha\beta} = \int_K d\phi_\alpha^J d\phi_\beta^J dA = - \int_K \phi_\alpha^J \Delta \phi_\beta^J dA,$$

respectively. By construction, A and C are symmetric, A is positive definite, and C is positive. By virtue of the results of (Arden 2001), the subdivision functions which are solutions of the generalized eigenvalue problem

$$Cv = \lambda Av$$

converge to the eigenfunctions of the Laplace-Beltrami operator on M . Let $\psi_k^J : K \rightarrow \mathbb{R}$, $k = 0, \dots, p_J$ denote the k -th (approximate) eigenfunction. We computed the eigenfunctions ψ_k^1 .

We then constructed four random functions f^1, \dots, f^4 on M by forming random linear combinations of the eigenfunctions:

$$f^s = \sum_{k=0}^{p_J} A^k(s) \psi_k^1,$$

with $A^k(s) \sim N(0, \sigma_k^2(s))$, and

$$\sigma_k^2(s) = \tanh^2(s/2).$$

We normalized the test functions f^s to have mean 0 and variance 1, and scaled the surface to have unit area. The quantities

$$\|f^s\|_1^2 = \int_K |\text{grad } f^s|^2 dA \text{ and } \|f^s\|_2^2 = \int_K |\Delta_M f^s|^2 dA$$

shown in Table 1 are then reasonable measures of the complexity of f^s . As expected, the complexity increases with s . Figure 10 displays the normalized random functions on the elephant.

s	$\ f^s\ _1^2$	$\ f^s\ _2^2$	$\min f^s$	$\max f^s$
1	481	1.72e+7	-1.97	1.23
2	708	2.17e+7	-2.83	1.51
3	1.85e+3	1.02e+8	-3.01	1.98
4	7.19e+3	3.78e+8	-2.59	2.55

TABLE 1. Statistics for the normalized random functions on the elephant.

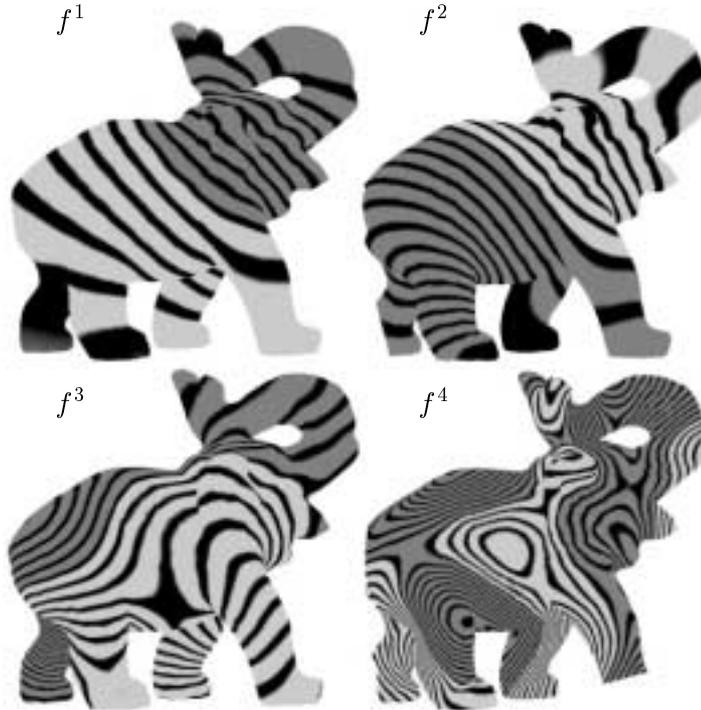


FIGURE 10. Level 1 random functions f^1, \dots, f^4 on the surface of the elephant. Light areas indicate positive function value. Dark lines are isopleths.

9.3. Test data and results. For each of the test functions f^s , $s = 1, 2, 3, 4$, we generated nine test data sets $\mathcal{S}_{n,b}^s$ with sample sizes $n = 100, 400, 1600$ and signal-to-noise ratios $b = 1, 10, 100$. The predictor values x_1, \dots, x_n in $\mathcal{S}_{n,b}^s$ are uniformly distributed over the surface, and the response values are given by $y_i = f^s(x_i) + \epsilon_i$, where the ϵ_i are Gaussian random errors with mean 0. The variance of the ϵ_i for test data set $\mathcal{S}_{n,b}^s$ is given by

$$\sigma^2 = \frac{\text{Var}(f^s)}{b}.$$

Here, $\text{Var}(f^s)$ is the variance of f^s with respect to the uniform distribution on M .

We then applied our spline smoothing procedure, using level-0 subdivision functions in the finite element method. Notice that the test functions f^s are level 1 subdivision functions and are therefore not contained in the space of the finite element approximations. For each triple (s, n, b) , we used *generalized cross validation* (GCV) to estimate the optimal smoothing parameter λ and the corresponding number of degrees of freedom (df).

We measured the estimation error between the approximate smoothing spline $f_{\lambda,0}$ and the test function f^s by

$$\text{error} = \frac{\|f_{\lambda,0} - f^s\|^2}{\text{Var}(f^s)}$$

Our results are summarized in Figure 11. There are 12 panels corresponding to the different combinations of s (determining the smoothness of the underlying function f^s), and sample size n . Each panel has three curves corresponding, from top to bottom, to signal-to-noise ratios $b = 1, 10, 100$.

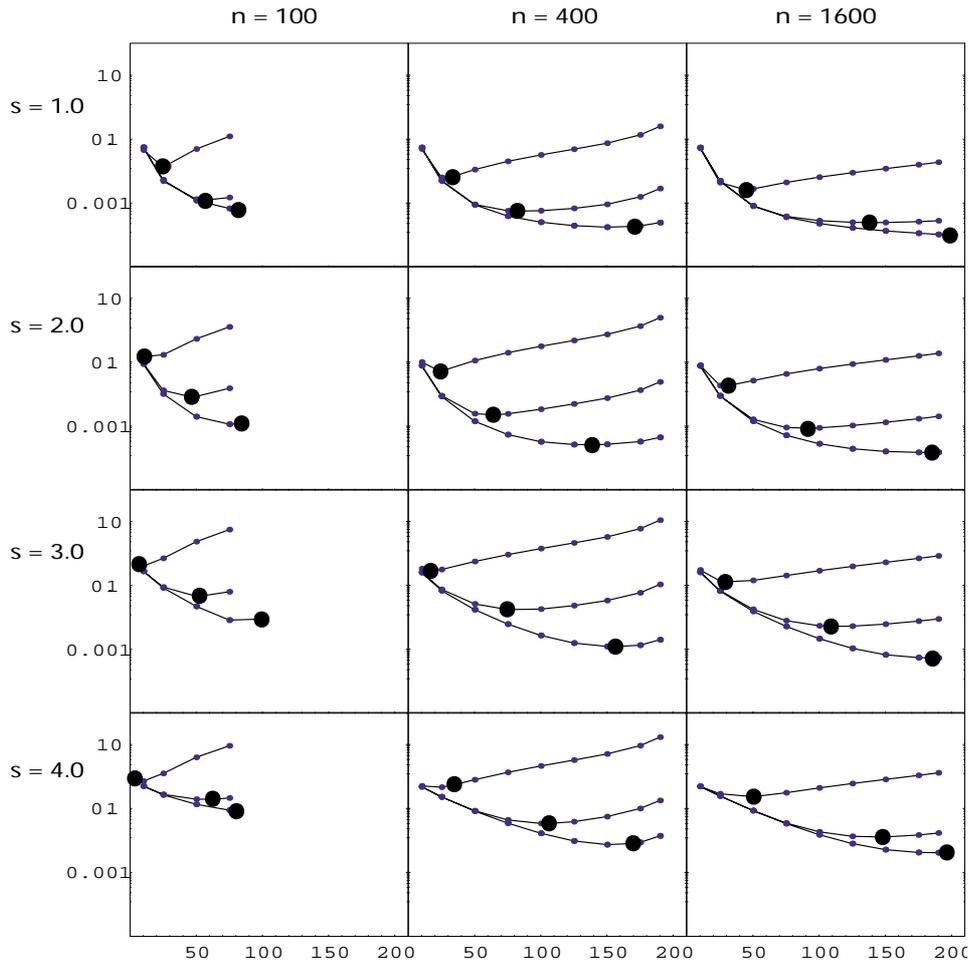


FIGURE 11. Relative estimation error when estimating f^s from a sample of size n by the approximate smoothing spline $f_{\lambda,0}$, as a function of the number of degrees of freedom used by the spline. Filled circles indicate estimates chosen by generalized cross-validation.

Each curve shows the estimation error as a function of the number of degrees of freedom used by the smooth. The filled circle on each curve indicates the number of degrees of freedom chosen by generalized cross-validation. Qualitatively, the results are what one would expect. Error is increasing in s — the more complex the true underlying function, the harder it is to estimate. It is decreasing in n , and also decreasing in the signal-to-noise ratio b . Generalized cross-validation does an excellent job in estimating the optimal degree of smoothing, at least for the 36 test data sets we considered.

10. DISCUSSION AND FUTURE WORK

We have demonstrated a finite element method for computing smoothing splines on (subdivision) surfaces with complex topology and geometry. We now review related work, discuss its connections to our approach, and point out some directions for future research.

There is a considerable literature on function estimation on the sphere; see Fasshauer and Schumaker (1998) for a review. Some of these approaches have been generalized to star-shaped surfaces (radially distorted spheres) (Alfeld *et al.* 1996).

Probably the earliest paper discussing function estimation on complex surfaces of arbitrary topology is Mallet (1992). Mallet assumes that the domain is given as a triangular mesh M , and that the observation sites x_i are the vertices of the mesh. The estimate f is a piecewise linear function on M , defined by its values f_i at the vertices. Roughly speaking, Mallet proposes to measure smoothness of f by

$$\phi(f) = \sum_i \left(f_i - \frac{1}{|N_i^k|} \sum_{x_j \in N_i^k} f_j \right)^2,$$

where N_i^k is the set of vertices that can be reached from x_i by traversing at most k edges. He then finds the function that minimizes the loss function

$$E[f] = \frac{1}{n} \sum (y_i - f_i)^2 + \lambda \phi(f).$$

The parameter λ controls the amount of smoothing. A closely related approach was more recently proposed by Joshi (1998, Section 3.2).

Mallet's approach is similar in spirit to spline smoothing, in that there is a loss function allowing for a tradeoff between smoothness and fidelity to the data, and the estimate minimizing the loss function is found by solving a penalized least squares problem. The differences are that in Mallet's approach the solution is not smooth, and the geometry of the domain enters only in a fairly crude way through the connectivity of the domain mesh.

More recent work can be characterized as variations on the idea of "smoothing by diffusion" (Perona and Malik 1990). The idea is easiest to describe on the line. Suppose we have an initial estimate $f(x, 0)$ interpolating or closely approximating the data $(x_1, y_1), \dots, (x_n, y_n)$. We can then generate a family $f(x, T)$ of smoothed versions of $f(x, 0)$ by solving the heat equation

$$\frac{\partial f(x, t)}{\partial t} - \frac{\partial^2 f(x, t)}{\partial x^2} = 0$$

with initial value $f(x, 0)$. The amount of smoothing is controlled by T — the larger T , the smoother $f(x, T)$. The approach generalizes to functions on the plane or on general surfaces by replacing the second derivative with the Laplace-(Beltrami) operator. Note that on the line or the plane $f(x, T)$ can be obtained by convolving $f(x, 0)$ with a Gaussian kernel whose width depends on T , and therefore on those domains there is not much difference between smoothing by diffusion and kernel smoothing.

More interesting smoothers are obtained if the Laplace operator is replaced by a more general non-linear operator D , leading to non-linear diffusion. For example, in the context of image denoising, Perona and Malik (1990) propose to solve the PDE

$$\frac{\partial f(x, t)}{\partial t} - \nabla_x \cdot (g(\|\nabla_x f\|) \nabla_x f) = 0$$

with

$$g(x) = \frac{1}{1 + \frac{x^2}{\alpha}}$$

The effect of this choice of operator is to change $f(x, t)$ slowly where $\nabla_x f$ is large, i.e. along sharp edges. This prevents sharp features from being smeared out.

Smoothing by nonlinear diffusion is also the basic idea in the paper by Bajaj and Xu (2001). Like (Mallet 1992) they assume that the surface is given as a triangular mesh, and that the observation sites x_i are the vertices of the mesh. They treat the mesh as the base complex K of a subdivision surface, and they consider the y_i as the control values of a subdivision function $f(x, 0)$ over the base complex. They then smooth $f(x, 0)$ by nonlinear diffusion, solving the corresponding PDE using a finite element method. The connection between the approach of Bajaj and Xu and our implementation of spline smoothing is the use of subdivision functions in a finite element method. However there are basic differences. Spline smoothing minimizes a well defined and motivated functional, and at least on simple domains the statistical performance of spline smoothing has been thoroughly analyzed. Smoothing by anisotropic diffusion is more *ad hoc* and performance of such procedures is less well understood.

On the other hand, spline smoothing is a linear method, whereas the diffusion method of Bajaj and Xu (2001) is nonlinear. The distinction between linear and non-linear smoothing methods is a fundamental one. For linear methods, the predicted values \hat{y}_i at the observation sites x_i are a linear function of the observed responses y_i :

$$\hat{\mathbf{y}} = W \mathbf{y}$$

where the matrix W does not depend on \mathbf{y} . Mallet's method, spline smoothing, and diffusion using the heat equation all are linear methods. On a fundamental level, all linear smoothers are alike. Consider, for example, function estimation on the circle, with regularly spaced x_i . In this situation, weight matrices for linear smoothers are circulant, for symmetry reasons. So these methods are smoothing by convolution, and the difference between them merely lies in the properties of the convolution kernel.

In contrast to linear smoothers, non-linear smoothers can be designed to locally adapt to discontinuities and other special features of the underlying function f_{true} by looking at the observed response values y_i . This can result in superior performance. For example, Donoho and Johnstone (1995) have shown that their non-linear, wavelet based "SureShrink" method for function estimation on the line achieves optimal convergence rates in a variety of function spaces containing non-smooth functions.

Applications of nonlinear diffusion have been mostly in problem domains where the signal-to-noise ratio is high. In this case it is possible to infer properties of the true underlying function f_{true} at some point x from the observed response values in a small neighborhood of x . It seems reasonable to conjecture that this approach will be less successful when the signal-to-noise ratio is low. Porting the basic ideas of SureShrink to function estimation on surfaces is an interesting open problem.

REFERENCES

- [1] P. Alfeld, M. Neamtu, and L.L. Schumaker, *Fitting scattered data on sphere-like surfaces using spherical splines*, J. Comput. Appl. Math. **73** (1996), no. 1-2, 5–43.
- [2] G. Arden, *Approximation properties of subdivision surfaces*, Ph.D. thesis, University of Washington, Department of Mathematics, 2001, available at <http://grail.cs.washington.edu/projects/scanning/>.

- [3] C. Bajaj and G. Xu, *Anisotropic diffusion of noisy surfaces and noisy functions on surfaces*, Proc. of Information Visualization 2001— Computer Aided Geometric Design, London, UK, 2001, 2001, (to appear).
- [4] F. Cirak, M. Ortiz, and P. Schröder, *Subdivision surfaces: a new paradigm for thin-shell finite-element analysis*, International Journal for Numerical Methods in Engineering **47** (2000), 2039–2072.
- [5] M.P. do Carmo, *Differential geometry of curves and surfaces*, Prentice-Hall, Englewood Cliffs, N.J., 1976.
- [6] D.L. Donoho and I.M. Johnstone, *Adapting to unknown smoothness via wavelet shrinkage*, J. Amer. Statist. Assoc. **90** (1995), no. 432, 1200–1224.
- [7] C.E. Duchon, *Splines minimizing rotation-invariant semi-norms in Sobolev spaces*, Constructive Theory of Functions of Several Variables (Berlin), Springer-Verlag, 1977, pp. 85–100.
- [8] N. Dyn, F. J. Narcowich, and J. D. Ward, *Variational principles and Sobolev-type estimates for generalized interpolation on a Riemannian manifold*, Constr. Approx. **15** (1997), 175–208.
- [9] G.E. Fasshauer and L.L. Schumaker, *Scattered data fitting on the sphere*, Mathematical methods for curves and surfaces, II (Lillehammer, 1997), Vanderbilt Univ. Press, Nashville, TN, 1998, pp. 117–166.
- [10] G.H. Golub and C.F. Van Loan, *Matrix computations*, third ed., The Johns Hopkins University Press, Baltimore, 1996.
- [11] G. Greiner, *Variational design and fairing of spline surfaces*, Computer Graphics Forum **13** (1994), 143–154.
- [12] E. Grinspun, P. Krysl, and P. Schröder, *Charms: A simple framework for adaptive simulation*, SIGGRAPH 2002, Computer Graphics Annual Conference Series, 2002, to appear.
- [13] T.J. Hastie and R.J. Tibshirani, *Generalized additive models*, Chapman and Hall, 1990.
- [14] H. Hoppe, T. DeRose, T. Duchamp, M. Halstead, H. Jin, J.A. McDonald, and W. Stuetzle, *Piecewise smooth surface reconstruction*, Computer Graphics (SIGGRAPH '94 Proceedings) **28** (1994), 295–302.
- [15] Sarang C. Joshi, *Large deformation diffeomorphisms and gauss random fields for statistical characterization of brain sub-manifolds*, Ph.D. thesis, Washington University, Sever Institute of Technology, Department of Electrical Engineering, 1998.
- [16] P. Kim, *Splines on Riemannian manifolds and a proof of a conjecture by Wahba*, preprint (<http://www.uoguelph.ca/~pkim/research.html>), 2001.
- [17] Jan Krzysztof Kowalski, *Application of box splines to the approximation of Sobolev spaces*, J. Approx. Theory **61** (1990), no. 1, 53–73. MR **91b**:46033
- [18] J. Lane and R. Riesenfeld, *A theoretical development for the computer generation and display of piecewise polynomial surfaces.*, IEEE Transactions on Pattern Analysis and Machine Intelligence **2** (1980), no. 1, 35–46.
- [19] A. Lee, W. Sweldens, P. Schröder, L. Cowsar, and D. Dobkin, *Maps: multiresolution adaptive parameterization*, Proceedings of the 25th annual conference on Computer Graphics, 1998, Computer Graphics Annual Conference Series, ACM SIGGRAPH, Addison Wesley, 1998, pp. 95–104.
- [20] C. Loop, *Smooth subdivision surfaces based on triangles*, Master's thesis, Department of Mathematics, University of Utah, 1987.
- [21] ———, *Generalized b-spline surfaces of arbitrary topological type*, Ph.D. thesis, Department of Computer Science and Engineering, University of Washington, 1992.
- [22] C. Loop and T. DeRose, *Generalized b-spline surfaces of arbitrary topology*, Computer Graphics (SIGGRAPH '90 Proceedings) **24** (1990), no. 4, 347–356.
- [23] J.L. Mallet, *Discrete smooth interpolation in geometric modeling*, CAD **24** (1992), no. 4, 178–191.
- [24] J. Meinguet, *Multivariate interpolation at arbitrary points made simple*, J. Appl. Math. Phys. **30** (1979), 292–304.
- [25] F. J. Narcowich, *Generalized Hermite interpolation and positive definite kernels on a Riemannian manifold*, J. Math. Anal. Appl. **190** (1995), 165–193.
- [26] P. Perona and J. Malik, *Scale-space and edge detection using anisotropic diffusion*, IEEE Trans. Pattern Anal. and Machine Intelligence **12** (1990), no. 7, 629–639.
- [27] U. Reif, *A unified approach to subdivision algorithms near extraordinary vertices*, Computer-Aided Geometric Design **12** (1995), no. 2, 153–174.
- [28] M. Sabin, *The use of piecewise forms for the numerical representation of shape*, Ph.D. thesis, Hungarian Academy of Sciences, Budapest, Hungary, 1976.
- [29] J.E. Schweitzer, *Analysis and application of subdivision surfaces*, Ph.D. thesis, University of Washington, Department of Computer Science and Engineering, August 1996.
- [30] E. Stollnitz, T. DeRose, and D. Salesin, *Wavelets for computer graphics: Theory and applications*, Morgan Kaufmann Publishers, Inc, San Francisco, 1996.
- [31] G. Wahba, *Data-based optimal smoothing of orthogonal series density estimates*, Annals of Statistics **9** (1981), 146–156.
- [32] ———, *Spline interpolation and smoothing on the sphere (corr: V3 p385-386)*, SIAM J. Sci. Stat. Computing **2** (1981), 5–16.
- [33] ———, *Spline functions for observational data*, SIAM, Philadelphia, 1990.

- [34] F.W. Warner, *Foundations of differentiable manifolds and Lie groups*, Springer-Verlag, New York, 1983, Corrected reprint of the 1971 edition.
- [35] J.G. Wendelberger, *Smoothing noisy data with multidimensional splines and generalized cross-validation*, Ph.D. thesis, University of Wisconsin–Madison, Department of Statistics, 1982.
- [36] H. Weyl, *The classical groups, their invariants and representations*, second ed., Princeton University Press, Princeton, N.J., 1946.
- [37] D. Wood, D. Azuma, K. Aldinger, B. Curless, T. Duchamp, D. Salesin, and W. Stuetzle, *Surface light fields for 3d photography*, SIGGRAPH '2000, Computer Graphics Annual Conference Series, July 2000, pp. 287–296.
- [38] D. Zorin, P. Schröder, and W. Sweldens, *Interpolating subdivision for meshes with arbitrary topology*, Computer Graphics (SIGGRAPH '96 Proceedings) **30** (1996), 189–192.
- [39] D.N. Zorin, *Stationary subdivision and multiresolution surface representations*, Ph.D. thesis, Caltech, Department of Computer Science, 1998.

TOM DUCHAMP, DEPARTMENT OF MATHEMATICS, BOX 354350, UNIVERSITY OF WASHINGTON, SEATTLE, WA. 98195

E-mail address: `duchamp@math.washington.edu`

WERNER STUETZLE, DEPARTMENT OF STATISTICS, BOX 354322, UNIVERSITY OF WASHINGTON, SEATTLE, WA. 98195

E-mail address: `wxs@stat.washington.edu`