Shape-preserving, Multiscale Fitting of Bivariate Data by Cubic L_1 Smoothing Splines

David E. Gilsinn
John E. Lavery

Abstract. Bivariate cubic L_1 smoothing splines are introduced. The coefficients of a cubic L_1 smoothing spline are calculated by minimizing the weighted sum of the L_1 norms of second derivatives of the spline and the ℓ_1 norm of the residuals of the data-fitting equations. Cubic L_1 smoothing splines are compared with conventional cubic smoothing splines based on the L_2 and ℓ_2 norms. Computational results for fitting a challenging data set consisting of discontinuously connected flat and quadratic areas by C^1-smooth Sibson-element splines on a tensor-product grid are presented. In these computational results, the cubic L_1 smoothing splines preserve the shape of the data while cubic L_2 smoothing splines do not.

§1. Introduction

Among the current options for approximating bivariate data are tensor-product, polynomial and thin-plate smoothing splines [1,4,7,8,13,16,17,18], multiquadrics [2] and wavelets [3,5]. Tensor-product, polynomial and thin-plate smoothing splines often have extraneous, "nonphysical" oscillation, especially near multiscale phenomena, that is, near regions where the magnitude of the data or the sizes of the cells in the grid change abruptly. The oscillation in these smoothing splines can be mitigated or eliminated by shifting the positions of nodes, adjusting the number of nodes, adding various constraints or penalties and a posteriori filtering, often with significant amounts of human interaction. At additional computational expense, multiquadrics and wavelets can avoid nonphysical oscillation. Development of computationally inexpensive smoothing splines that preserve shape without requiring human interaction would be of great benefit in modeling objects with multiscale phenomena such as urban and natural terrain, mechanical objects and images.

In [10,11,12], new classes of univariate L_1 interpolating splines, univariate L_1 smoothing splines and bi- and multivariate L_1 interpolating splines were

Approximation Theory X
Charles K. Chui, Larry L. Schumaker, and Joachim Stoeckler (eds.), pp. 1–10.

ISBN 0-8265-xxxx-x.

proposed. These splines preserve shape for smooth data as well as for data with abrupt changes in magnitude and spacing and for smooth sets of spline nodes as well as for those with abrupt changes in spacing. In the present paper, we extend the results of [11,12] by creating a new class of bivariate cubic L_1 smoothing splines. Our focus here is mainly on bivariate C^1-smooth cubic L_1 smoothing splines on tensor-product grids.

The objective of this paper is to present two case studies of approximation of simulated urban structures by bivariate L_1 smoothing splines. The urban structures were simulated so that data sets would be devoid of signal noise and image contamination due to preprocessing. Although it is crucial to be able to deal with signal noise and image contamination, the primary focus here is to study the performance of L_1 smoothing splines on "clean" data sets.

§2. Bivariate Data, Grids and Sibson Shape Elements

We consider fitting the data $(\hat{x}_m, \hat{y}_m, \hat{z}_m)$, $m = 1, 2, \ldots, M$. The weight of the mth data point is a positive real number $\hat{w}_m$. We will create bivariate cubic L_1 smoothing splines on tensor-product grids with nodes x_i, $i = 0, 1, \ldots, I$ and y_j, $j = 0, 1, \ldots, J$ that are strictly monotonic partitions of the finite real intervals $[x_0, x_I]$ and $[y_0, y_J]$, respectively. The domain of the spline will be $D = [x_0, x_I] \times [y_0, y_J]$.

Sibson elements [6,9,12] will be used for the computational results in the present paper. To create a Sibson element on a rectangle $(x_i, x_{i+1}) \times (y_j, y_{j+1})$, one first divides the rectangle into four triangles by drawing the two diagonals of the rectangle. The Sibson element is a shape function $z(x, y)$ that is cubic in each triangle, is C^1 at the lines separating the four triangles, is C^1 with the Sibson elements in the adjacent rectangles, has derivative $\partial z/\partial x$ that is linear along the edges $x = x_i$ and $x = x_{i+1}$ of the rectangle and has derivative $\partial z/\partial y$ that is linear along the edges $y = y_j$ and $y = y_{j+1}$. The Sibson element z in a given rectangle depends only on the values of z, $\partial z/\partial x$ and $\partial z/\partial y$ at the corners of that rectangle (12 parameters per rectangle) as described in [6] and in Sec. 2 of [12]. The values of z, $\partial z/\partial x$ and $\partial z/\partial y$ at node (x_i, y_j) will be denoted by z_{ij}, z^x_{ij} and z^y_{ij}, respectively. The vectors of the values of the z_{ij}, z^x_{ij} and z^y_{ij}, $i = 0, 1, \ldots, I$, $j = 0, 1, \ldots, J$, will be denoted by z, z^x and z^y, respectively.

§3. Minimization Principle for Cubic L_1 Smoothing Splines

A cubic L_1 smoothing spline is a function that, for a given α, $0 < \alpha < 1$, minimizes

$$\begin{aligned} E = {} & \alpha \sum_{m=1}^{M} \hat{w}_m \left| z(\hat{x}_m, \hat{y}_m) - \hat{z}_m \right| \\ & + (1-\alpha) \iint_D \left[\left| \frac{\partial^2 z}{\partial x^2} \right| + 2 \left| \frac{\partial^2 z}{\partial x \partial y} \right| + \left| \frac{\partial^2 z}{\partial y^2} \right| \right] dx\, dy \end{aligned} \tag{1}$$

over all surfaces z of a given class. The balance parameter α determines the trade-off between the closeness with which the data are fitted, represented by the sum in (1), and the tendency of the spline to be close to a piecewise planar surface, represented by the double integral in (1). The double integral in (1) is the functional that defines a cubic L_1 interpolating spline [12]. This double integral could be replaced by a double integral with different weights on the three terms in the integrand, as in expression (4) of [12], or by a double integral with different terms, as in expression (6) of [12].

Cubic L_1 smoothing splines based on Sibson elements exist because, as a function of the coefficients z, z^x and z^y, functional (1) is continuous, bounded below by 0 and convex and tends to ∞ uniformly as the Euclidean norm of the coefficients tends to ∞ (cf. Theorem 1 of [12], which states this result for interpolating splines). However, cubic L_1 smoothing splines under this definition need not be unique because functional (1) is not necessarily strictly convex. When there are several candidates for an L_1 smoothing spline, the candidate with (in some metric) the smallest absolute values of the first derivatives, that is, the flattest surface, is typically the choice of most users. For this reason, we add to E a "regularization" term:

$$E + \sum_{i=0}^{I}\sum_{j=0}^{J}\left[\varepsilon_{1ij}|z_{ij}^{x}| + \varepsilon_{2ij}|z_{ij}^{y}|\right] \tag{2}$$

where the regularization parameters ε_{1ij} and ε_{1ij} are small nonnegative numbers. Functional (2) is still not necessarily strictly convex and can therefore achieve its minimum for more than one set of coefficients z, z^x, z^y. However, standard interior-point algorithms, including the primal affine method used for the computational results presented in this paper, yield a unique set of coefficients that minimize (a discretization of) functional (2).

For comparison with cubic L_1 smoothing splines, we will calculate "cubic L_2 smoothing splines" by minimizing the functional

$$\begin{aligned}
&\alpha^2 \sum_{m=1}^{M}\left[\hat{w}_m\left(z(\hat{x}_m,\hat{y}_m)-\hat{z}_m\right)\right]^2 \\
&\quad + (1-\alpha)^2 \iint_D \left[\left(\frac{\partial^2 z}{\partial x^2}\right)^2 + 4\left(\frac{\partial^2 z}{\partial x \partial y}\right)^2 + \left(\frac{\partial^2 z}{\partial y^2}\right)^2\right] dx\,dy \\
&\quad + \sum_{i=0}^{I}\sum_{j=0}^{J}\left[\left(\varepsilon_{1ij} z_{ij}^{x}\right)^2 + \left(\varepsilon_{2ij} z_{ij}^{y}\right)^2\right]
\end{aligned} \tag{3}$$

Functional (3) is the same as the functional (2) except that the squares of the ℓ_2 and L_2 norms have replaced the ℓ_1 and L_1 norms. The integral in the functional for cubic L_2 smoothing splines is not the same as the integral for thin-plate splines because the coefficient of the middle term in the integrand is 4, not 2. L_2 smoothing splines could, of course, be calculated with a thin-plate-spline functional replacing the double integral in (3). However, L_2 smoothing splines based on minimization principle (3) were chosen for comparison with L_1 smoothing splines because the main goal of this paper is to demonstrate that the fundamental solution of the shape-preservation problem is a proper choice of the function spaces. The most relevant comparisons are therefore those in which only the function spaces (and not, for example, the coefficients in the integrals) differ. Comparison of L_1 smoothing splines of many different types (including types A_1 and B described in [12]) with L_2 smoothing splines of many different types (including thin-plate smoothing splines and smoothing splines of types A_1 and B) is an important issue for future investigation.

§4. Algorithm and Computational Results

We calculate the coefficients of a Sibson-element L_1 smoothing spline by minimizing the functional (2) in which the integral is discretized by the scheme used in [12], which can be summarized as follows. Let N be an integer ≥ 2. Divide each rectangle $(x_i, x_{i+1}) \times (y_j, y_{j+1})$ into N^2 subrectangles. Approximate the double integral over the rectangle by $1/[2N(N-1)]$ times the sum of the $2N(N-1)$ values of the integrand at the midpoints of the sides of the subrectangles that are in the interior of the rectangle. This discretization was chosen because it uses values of the integrand only in the interiors (and not on the boundaries) of the 4 triangles that make up each rectangle.

Minimization of (2) with the integral discretized in this manner was carried out by the primal affine method of Vanderbei, Meketon and Freedman [12,14,15] coded by the authors of this paper. The primal affine algorithm is known to converge globally to a unique solution no matter whether functional (2) is strictly convex or only (non-strictly) convex. Further information on the convergence of the primal affine method can be found in the second paragraph of Sec. 3 of [11] and the second-to-last paragraph of Sec. 4 of [12], which cite the original results in [14,15] and elsewhere.

For comparison with cubic L_1 smoothing splines, cubic L_2 smoothing splines with the integral in (3) discretized in the same manner as the integral in (2) were calculated. For all of the computational results presented below, $N = 5$. The authors have computational experience with the data sets of Figs. 1 and with other data sets not only with $N = 5$ but also with $N = 3$. No significant differences have been noted. Investigation of how L_1 and L_2 smoothing splines vary with N may be of interest. However, before commencing such an investigation, it may be useful to determine whether functional (2) can be minimized without discretizing the integral. Research by the optimization community may answer this question within the next few years.

The weights $\hat{w}_m$ were chosen to be 1 for all m. The regularization parameters ε_{1ij} and ε_{2ij} were chosen to be 10^{-4} for all i and j.

Cubic L_1 and L_2 smoothing splines were computed for two data sets. Data set 1, shown on the left in Fig. 1, is a set of discontinuously connected flat areas that represents a high-rise hotel complex. This data set consists of 201×201 data points at equal, 1-unit spacing. Let $\hat{\imath} = 0, 1, \ldots, 200$ and

$\hat{\jmath} = 0, 1, \ldots, 200$ be the coordinates of the data locations. The data points $\hat{z}$ of data set 1 are given in Table 1.

Tab. 1 Data Set 1 (Simulated High-rise Hotel Complex)				
$\hat{z}$	Begin $\hat{\imath}$	End $\hat{\imath}$	Begin $\hat{\jmath}$	End $\hat{\jmath}$
150	71	135	16	35
25	101	125	51	75
10	16	135	100	120
10	16	35	121	155
100	36	115	121	155
10	116	135	121	155
10	16	135	156	175
0	Otherwise	Otherwise	Otherwise	Otherwise

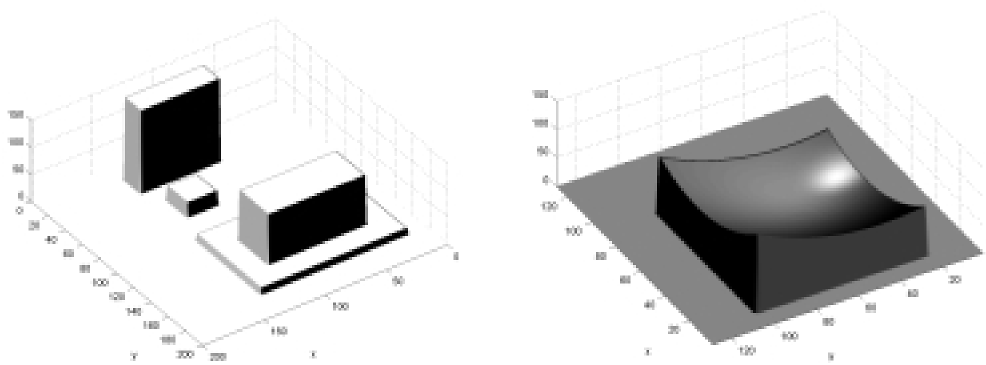

Fig. 1. Simulated high-rise hotel complex (left, data set 1) and sports stadium (right, data set 2).

Data set 2, shown on the right in Fig. 1, is a simulated sports stadium consisting of a quadratic surface discontinuously embedded in a flat area. This data set consists of 128×128 data points at equal, 1-unit spacing. Letting $\hat{\imath} = 0, 1, \ldots, 128$ and $\hat{\jmath} = 0, 1, \ldots, 128$ be the coordinates of the data locations, the data points $\hat{z}$ of data set 2 are given by

$$\hat{z}(\hat{\imath}, \hat{\jmath}) = \begin{cases} 20 + 0.024 * (\hat{\imath} - 62)^2 + 0.016 * (\hat{\jmath} - 42)^2 & \text{if } \begin{cases} \hat{\imath} = 12 : 92 \\ \hat{\jmath} = 22 : 102 \end{cases} \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

For data set 1, cubic L_1 and L_2 smoothing splines were computed on spline grids consisting of 100×100 equal cells, each of size 2 units by 2 units, with $\alpha = 0.8$. For these smoothing splines, the "raw compression ratio," that is, the number of floating-point storage locations for the original data $\hat{z}(\hat{\imath}, \hat{\jmath})$ divided by the number of floating-point storage locations for the smoothing spline parameters $z_{i,j}$, $z^x_{i,j}$, $z^y_{i,j}$, $i = 0, 1, \ldots, 100$, $j = 0, 1, \ldots, 100$ is $201^2/(3 *$

$101^2) = 1.32$. This case was chosen because it shows the different capabilities of L_1 and L_2 splines at a low compression ratio. The L_1 and L_2 smoothing spline approximations are shown in Fig. 2. Note the sharp and accurate approximations of edges and corners in the L_1 smoothing spline. In contrast, the L_2 spline has over- and undershoot at the edges of the buildings and has oscillation near the edges.

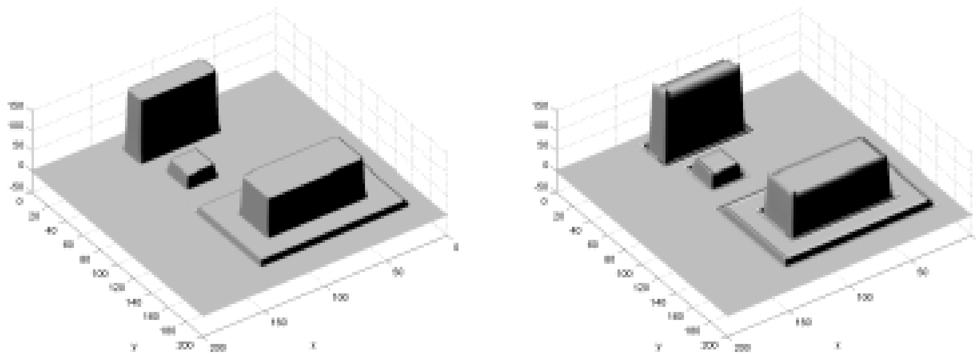

Fig. 2. L_1 smoothing spline (left) and L_2 smoothing spline (right) with $\alpha = 0.10$ for data set 1.

For data set 2, cubic L_1 and L_2 smoothing splines were computed on spline grids consisting of 16×16 equal cells, each of size 8 units by 8 units, with $\alpha = 0.99, 0.9, 0.85, 0.8, 0.75, 0.65, 0.6, 0.5, 0.4$ and 0.3. For these smoothing splines, the raw compression ratio is $129^2/(3 * 17^2) = 19.19$. This case was chosen because it shows the capabilities of L_1 and L_2 splines at a medium compression ratio. The L_1 and L_2 splines with $\alpha = 0.8$ and 0.75 are shown in Figs. 3 and 4, resp. In both of these figures, the L_1 smoothing splines fit the data well and have very little over/undershoot and extraneous oscillation. In contrast, the L_2 smoothing splines have considerable "nonphysical" oscillation.

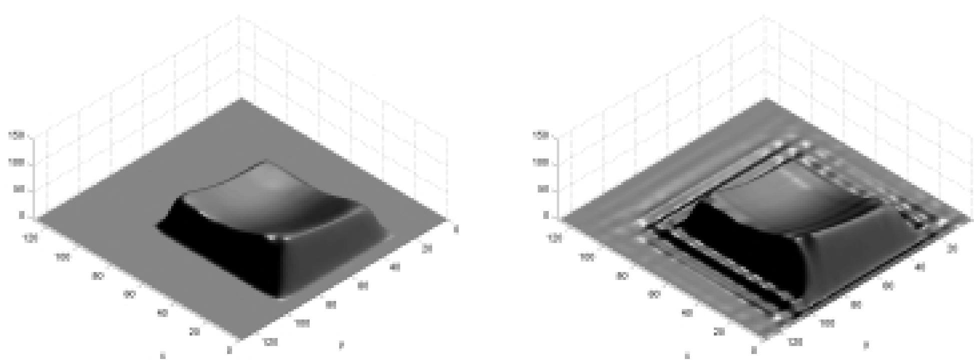

Fig. 3. L_1 smoothing spline (left) and L_2 smoothing spline (right) with $\alpha = 0.8$ for data set 2.

In the above paragraph, the measure for the performance of the smoothing splines is visual inspection because it is not yet known how to measure shape preservation quantitatively. Nevertheless, it is appropriate to give some

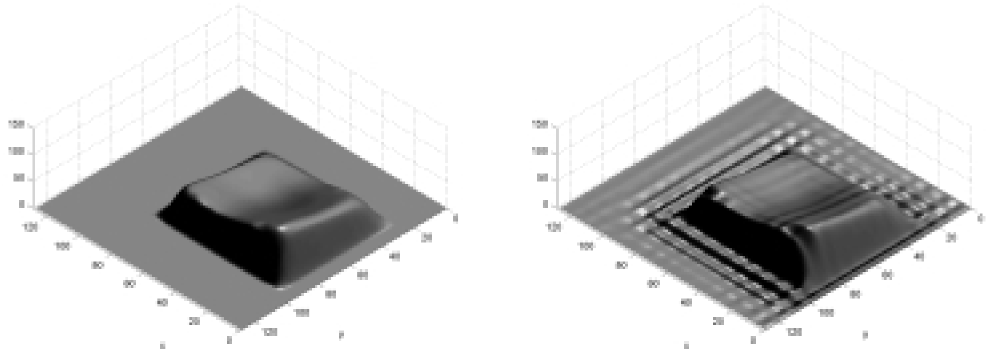

Fig. 4. L_1 smoothing spline (left) and L_2 smoothing spline (right) with $\alpha = 0.75$ for data set 2.

quantitative information about the performance of the smoothing splines. To do so, we will use the following norms: 1) the (normalized) ℓ_1 norm $||\cdot||_{\ell_1}$ (sum of the absolute values of the 129^2 values divided by 129^2), 2) the (normalized) ℓ_2 norm $||\cdot||_{\ell_2}$, also known as the RMS or root-mean-square norm (square root of the quotient that consists of the sum of the squares of the 129^2 values divided by 129^2) and 3) the ℓ_∞ norm $||\cdot||_{\ell_\infty}$ (maximum of the 129^2 absolute values). In Table 2, we present the ℓ_1, ℓ_2 and ℓ_∞ norms of the error between the values z of the L_1 and L_2 smoothing splines and the $\hat{z}$ of data set 2. In this table, we denote L_1 smoothing splines by $z_{[L_1,\alpha]}$ and L_2 smoothing splines by $z_{[L_2,\alpha]}$.

Tab. 2 Norms of Errors of Smoothing Splines and Data Set 2

$i =$	1		2	
weight $\alpha =$	0.8	0.75	0.8	0.75
$\lvert\lvert z_{[L_i,\alpha]} - \hat{z}\rvert\rvert_{\ell_1} =$	8.373	9.517	8.637	9.729
$\lvert\lvert z_{[L_i,\alpha]} - \hat{z}\rvert\rvert_{\ell_2} =$	22.80	23.81	22.18	23.23
$\lvert\lvert z_{[L_i,\alpha]} - \hat{z}\rvert\rvert_{\ell_\infty} =$	137.6	137.6	130.7	132.4

By the information in Table 2, one could not determine whether L_1 smoothing splines are better or worse than L_2 smoothing splines. Since L_1 and L_2 smoothing splines fit data in the spaces ℓ_1 and ℓ_2, respectively, it is perhaps not surprising that L_1 smoothing splines perform better in the ℓ_1 norm and that L_2 smoothing splines perform better in the ℓ_2 (and ℓ_∞) norm. However, most observers interested in geometric modeling agree that the L_1 smoothing splines represent the original data better than the L_2 smoothing splines. The error norms in Table 2 confirm what is well known in the image processing community, namely, that the ℓ_1, ℓ_2 and ℓ_∞ norms are not good measures of shape preservation.

§5. Convergence Issues

The primal affine method, which has so far been the method of choice for calculating L_1 smoothing splines, typically performs well for small data sets. However, for the large data sets of interest in this paper, the primal affine method converged slowly (100-600 iterations) for some α, converged incompletely (difference between iterations decreased until a certain point and then increased) for other α and diverged for yet other α. The computational results shown above were, of course, for cases of complete convergence. Alternative strategies for calculating L_1 smoothing splines are under investigation. One of these strategies is domain decomposition, in which the global domain is broken up into many slightly overlapping subdomains and the global L_1 smoothing spline is patched together from the local L_1 smoothing splines. Also, linear and nonlinear programming algorithms to replace the primal affine method are under investigation. The primal affine method in its current global implementation has been sufficient to "prove the principle" (in the present paper and in [10,11,12]) that L_1 interpolating and smoothing splines are able to preserve the shape of data much better than conventional L_2 splines. In the future, it is likely that other algorithms for calculating L_1 splines will be of great interest and use.

§6. Conclusion

In this paper, we have focused on providing evidence that the choice of the function spaces in smoothing spline minimization principles has far greater influence than previously expected. The seemingly minor change of the function spaces from the conventional choices ℓ_2 and L_2 to the unconventional choices ℓ_1 and L_1 results in a vast improvement in the shape-preserving, multiscale capabilities of cubic smoothing splines. The contribution of this article is not to prove that L_1 smoothing splines preserve shape better than L_2 smoothing splines but merely to observe that that is so for a limited set of test cases. A full proof that L_1 smoothing splines preserve shape better than L_2 smoothing splines requires quantitative understanding of shape preservation, something that does not yet exist.

The smoothing splines presented here were calculated with Sibson elements and with no adaptivity in the spline grids. Investigation of bivariate L_1 smoothing splines using various Sibson and non-Sibson elements on quadrangulations and triangulations in nonadaptive and adaptive settings would be of large interest. Comparison with other widely used methods for modeling irregular, multiscale data (TINs, wavelets, JPEG, etc.) needs be carried out.

One could, of course, choose to fit the data of Figs. 1 and 2 by L_2 smoothing splines on subdomains that do not cross the lines of discontinuity and therefore have much less extraneous oscillation. However, if one does so, one must identify the lines of discontinuity, introduce topology into the fitting procedure and handle issues of matching of splines at the boundaries of the subdomains. The cost of doing this has to be balanced against the advantages of having one "terrain skin," the L_1 smoothing spline, that requires none of this. Different users will make different choices that fit their needs. L_1 smoothing splines do not replace other options but do add a new, attractive option to the set of options from which the user can choose.

Cubic L_1 smoothing splines are a promising new technique for geometric modeling, especially modeling of objects with multiscale phenomena such as urban and natural terrain, mechanical objects and images. The preliminary results in the present paper indicate that further investigation of L_1 smoothing splines may have high payoff.

References

1. Anthony, G.T., and M.G. Cox, The fitting of extremely large data sets by bivariate splines, in Mason, J.C., and M.G. Cox, eds., *Algorithms for Approximation*, Clarendon Press, Oxford, 1987, 5–20.
2. Carlson, R.E., and B.K. Natarajan, Sparse approximate multiquadric interpolation, Comput. Math. Appl.**27** (1994), 99–108.
3. Cohen, A., Nonlinear wavelet approximation and image compression, in Chui, C.K., and L.L. Schumaker, eds., *Wavelets and Multilevel Approximation, Approximation Theory VIII*, vol. 2, World Scientific, Singapore, 1995, 17–38.
4. Dierckx, P., *Curve and Surface Fitting with Splines*, Clarendon Press, New York, 1993.
5. Dubuc, S., and G. Deslauriers, eds., *Spline Functions and the Theory of Wavelets*, Proceedings and Lectures Notes, vol. 18, American Mathematical Society, Providence, Rhode Island, 1999.
6. Gilsinn, D.E., Constructing Sibson Elements for a Rectangular Mesh, NISTIR 6718, National Institute of Standards and Technology, Gaithersburg, Maryland, 20899.
7. Gmelig Meyling, R.H.J., Approximation by cubic C^1 splines on arbitrary triangulations, Numer. Math. **51** (1987), 65–85.
8. Golitschek, M. von, and L.L. Schumaker, Data fitting by penalized least squares, in Mason, J.C., and M.G. Cox, eds., *Algorithms for Approximation II*, Chapman & Hall, London, 1990, 210–227.
9. Han, L., and L.L. Schumaker, Fitting monotone surfaces to scattered data using C^1 piecewise cubics, SIAM J. Numer. Anal. **34** (1997), 569–585.
10. Lavery, J.E., Univariate cubic L_p splines and shape-preserving, multiscale interpolation by univariate cubic L_1 splines, Comput. Aided Geom. Design **17** (2000), 319–336.
11. Lavery, J.E., Shape-preserving, multiscale fitting of univariate data by cubic L_1 smoothing splines, Comput. Aided Geom. Design **17** (2000), 715–727.
12. Lavery, J.E., Shape-preserving, multiscale interpolation by bi- and multivariate cubic L_1 splines, Comput. Aided Geom. Design **18** (2001), 321–343.
13. McMahon, J., R. John, and R. Franke, Knot selection for least squares thin plate splines, SIAM J. Sci. Stat. Comp. **13** (1992), 484–498.

14. Vanderbei, R.J., Affine-scaling for linear programs with free variables, Mathematical Programming **43** (1989), 31–44.

15. Vanderbei, R.J., M.J. Meketon, and B.A. Freedman, A modification of Karmarkar's linear programming algorithm, Algorithmica **1** (1986), 395–407.

16. Wahba, G., Multivariate thin plate spline smoothing with positivity and other linear inequality constraints, in Wegman, E.J., and D.J. dePriest, eds., *Statistical Image Processing and Graphics*, Marcel Dekker, New York, 1985, 275–290.

17. Wahba, G., *Spline Models for Observational Data*, Society for Industrial and Applied Mathematics, Philadelphia, Pennsylvania, 1990.

18. Willemans, K., and P. Dierckx, Surface fitting using convex Powell-Sabin splines, J. Comput. Appl. Math. **56** (1994), 263–282.

David E. Gilsinn
National Institute of Standards and Technology
100 Bureau Drive, Stop 8910
Gaithersburg, MD 20899-8910
`dgilsinn@nist.gov`

John E. Lavery
Computing and Information Sciences Division
Army Research Office
Army Research Laboratory
P.O. Box 12211
Research Triangle Park, NC 27709-2211
`lavery@arl.aro.army.mil`