ELSEVIER

Computer Aided Geometric Design 16 (1999) 39–59

COMPUTER AIDED GEOMETRIC DESIGN

Energy formulations of A-splines ☆

Chandrajit L. Bajaj [a,*], Jindong Chen [b], Robert J. Holt [c], Arun N. Netravali [c]

[a] *Department of Computer Sciences and TICAM, University of Texas, Austin, TX 78712, USA*
[b] *Digital Imaging Technology Center, CR&T, Xerox Corporation, Palo Alto, CA 94304, USA*
[c] *Bell Laboratories, Lucent Technologies, Murray Hill, NJ 07974, USA*

Received January 1997; revised March 1998

Abstract

A-splines are implicit real algebraic curves in Bernstein–Bézier (BB) form that are smooth. We develop A-spline curve models using various energy formulations, incorporating bending and stretching energy, based on the theory of elasticity. The attempt to find true energy minimizing curves usually leads to complicated integrals which can only be solved numerically, we introduce a simplified energy formulation which is much faster to compute yet still provides reasonably accurate results. Several examples for C^1-continuous quadratic A-splines using the true and simplified energy models are then presented.

1. Introduction

An A-spline is a smooth zero contour curve of a bivariate polynomial in Bernstein–Bézier (BB) form defined within a triangle (Bajaj and Xu, 1992), where the "A" stands for algebraic. Solutions to the problem of constructing a C^1 chain of implicit algebraic splines based on a polygon $\mathcal{P}$ have been given by Bajaj and Xu (1992) and for dense image or sparse scattered data by Bajaj and Xu (1996).

Several applications in image processing and computer graphics have been shown to be enhanced by using active contour models (Cohen and Cohen, 1993; Kass et al., 1988; Ronfard, 1994; Williams and Shah, 1992) and physically based modeling (see (Terzopoulos et al., 1987) and several others). In this paper we develop a spectrum of

☆ Supported in part by NSF grant CCR 92-22467, AFOSR grant F49620-97-1-0278, and ONR grant N00014-97-1-0398.

* Corresponding author. E-mail: bajaj@cs.utexas.edu.

PII: S0167-8396(98)00029-6

physical A-spline curve models using various energy formulations based on the theory of elasticity (Landau and Lifshitz, 1959). There are several advantages of this representation:

- A-splines can model arbitrary closed curves with degree bounded polynomial curve segments and various inter-segment continuity
- The class of objects represented by A-splines contains the class of objects representable by polynomial or rational B-splines

We consider an exact model for the strain energy, and conclude that simplifications are necessary in order to develop a practical algorithm for its approximation. Possible methods of approximating the strain energy include C^1-continuous quadratic and C^3-continuous cubic A-splines. We then explore in detail a simplified energy formulation for C^1-continuous quadratic A-splines.

The search for minimal energy curves has a history stretching back to Euler (curves he termed as "elastica") and more recently by pure and applied mathematicians (Birkhoff and De Boor, 1965; Bryant and Griffiths, 1986; Golomb and Jerome, 1982; Kallay, 1986; Lee and Forsythe, 1975; Malcom, 1977), computer vision experts (Brotman and Netravali, 1988; Bruckstein and Netravali, 1990; Horn, 1983; Mumford, 1994), and geometric designers (Mehlum, 1974; Jou and Han, 1990b). A characterization of plane elastica is given by Brunnett (1992), and closed-from expressions of bending energies for polynomial Pythagorean-Hodograph curves in Farouki (1996). We note that since quadratic A-splines are at most C^1-continuous, we find that energy-minimizing curves seem to be better approximated by cubic A-splines, which can achieve up to C^3 continuity while still maintaining a degree of freedom for shape control.

The rest of this paper is as follows. Section 2 gives preliminary information about A-splines as well as the theory of elasticity. In Section 3 we develop the elastic strain energy model for A-splines, taking into account both stretching and bending energy. Furthermore, we describe a simplified energy model for computational efficiency. Section 4 describes algorithms that minimize the different energy formulations of the previous section. Several example case studies detailing the use of C^1-continuous quadratic A-splines are provided. Section 5 concludes the paper.

2. Notation and preliminary details

Let $\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3 \in \mathbb{R}^2$ be non-collinear. Then the triangle (or two-dimensional simplex) with vertices $\mathbf{p}_1, \mathbf{p}_2$, and $\mathbf{p}_3$, is $T = [\mathbf{p}_1\mathbf{p}_2\mathbf{p}_3]$. Let $\mathbf{p} = (x, y)^{\mathrm{T}}$, $\mathbf{p}_i = (x_i, y_i)^{\mathrm{T}}$. Then for any $\mathbf{p} = \sum_{i=1}^{3} \alpha_i \mathbf{p}_i \in \mathbb{R}^2$ with $\sum_{i=1}^{3} \alpha_i = 1$ $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \alpha_3)^{\mathrm{T}}$ is the *barycentric coordinate* of $\mathbf{p}$, where the Cartesian and barycentric coordinates are related by

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix}. \tag{2.1}$$

The noncollinearity of $\mathbf{p}_1, \mathbf{p}_2$, and $\mathbf{p}_3$ guarantees that the 3×3 matrix in (2.1) is nonsingular and that the barycentric coordinates are well defined.

Any polynomial $F(\mathbf{p})$ of degree d can be expressed in BB form over T as $F(\mathbf{p}) = \sum_{|\boldsymbol{\lambda}|=d} b_{\boldsymbol{\lambda}} B_{\boldsymbol{\lambda}}^d(\boldsymbol{\alpha})$, $\boldsymbol{\lambda} \in \mathbb{Z}_+^3$, where $B_{\boldsymbol{\lambda}}^d(\boldsymbol{\alpha}) = (d!/\lambda_1!\lambda_2!\lambda_3!)\alpha_1^{\lambda_1}\alpha_2^{\lambda_2}\alpha_3^{\lambda_3}$ are the bivariate

Bernstein polynomials for $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \lambda_3)^{\mathrm{T}}$, and $|\boldsymbol{\lambda}|$ is defined to be $\sum_{i=1}^{3} \lambda_i$. Also, $b_{\boldsymbol{\lambda}} = b_{\lambda_1\lambda_2\lambda_3}$ (as a subscript, we simply write $\boldsymbol{\lambda}$ as $\lambda_1\lambda_2\lambda_3$) are called control points, and $\mathbb{Z}_+^3$ stands for the set of all three-dimensional vectors with nonnegative integer components. Let

$$S(\boldsymbol{\alpha}) = \sum_{|\boldsymbol{\lambda}|=d} b_{\boldsymbol{\lambda}} B_{\boldsymbol{\lambda}}^d(\boldsymbol{\alpha}), \quad |\boldsymbol{\alpha}| = 1, \tag{2.2}$$

be a given polynomial of degree d on the triangle $T = \{(\alpha_1, \alpha_2, \alpha_3)^{\mathrm{T}} \in \mathbb{R}^3 \colon \sum_{i=1}^{3} \alpha_i = 1, \alpha_i \geqslant 0\}$. The curve segment within the triangle is defined by $S(\alpha_1, \alpha_2, \alpha_3) = 0$.

Collecting the base functions $B_{\boldsymbol{\lambda}}^d$ into a vector $\mathbf{B}^d$ and the coefficients $b_{\boldsymbol{\lambda}}$ into a vector $\mathbf{b}$, Eq. (2.2) is rewritten as

$$S(\boldsymbol{\alpha}) = \mathbf{b}^{\mathrm{T}}\mathbf{B}^d.$$

Eq. (2.1) may be rewritten without the use of α_3 as

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x_1 - x_3 & x_2 - x_3 \\ y_1 - y_3 & y_2 - y_3 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} + \begin{bmatrix} x_3 \\ y_3 \end{bmatrix} = \mathbf{J}^{-1} \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} + \begin{bmatrix} x_3 \\ y_3 \end{bmatrix} \tag{2.3}$$

$\mathbf{J}$ is the Jacobian of $\boldsymbol{\alpha}$ in terms of x and y:

$$\mathbf{J} = \frac{\partial(\alpha_1, \alpha_2)}{\partial(x, y)} = \frac{1}{\Delta} \begin{bmatrix} y_2 - y_3 & x_3 - x_2 \\ y_3 - y_1 & x_1 - x_3 \end{bmatrix}, \tag{2.4}$$

where

$$\Delta = \begin{vmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ 1 & 1 & 1 \end{vmatrix}. \tag{2.5}$$

When we speak of functions of $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \alpha_3)T$, it should be noted that these are really functions of two independent variables since $\alpha_3 = 1 - \alpha_1 - \alpha_2$. The curve segment within the triangle T may thus be expressed as

$$S(\alpha_1, \alpha_2) = F\big(x(\boldsymbol{\alpha}), y(\boldsymbol{\alpha})\big).$$

We also have these relationships involving the two sets of coordinates for the gradient

$$\nabla F(x, y) = \mathbf{J}^{\mathrm{T}} \nabla S(\alpha_1, \alpha_2), \tag{2.6}$$

and the Hessian

$$\begin{aligned} \nabla^2 F(x, y) = \mathbf{H}(\boldsymbol{\alpha}) &= \mathbf{J}^{\mathrm{T}} \nabla^2 S(\alpha_1, \alpha_2) \mathbf{J} \\ &= \begin{bmatrix} \mathbf{J}_1{}^{\mathrm{T}} \nabla^2 S \mathbf{J}_1 & \mathbf{J}_1{}^{\mathrm{T}} \nabla^2 S \mathbf{J}_2 \\ \mathbf{J}_2{}^{\mathrm{T}} \nabla^2 S \mathbf{J}_1 & \mathbf{J}_2{}^{\mathrm{T}} \nabla^2 S \mathbf{J}_2 \end{bmatrix}, \end{aligned} \tag{2.7}$$

where $\mathbf{J}_i$ is the ith column of $\mathbf{J}$.

2.1. *Differential geometry of implicit functions*

2.1.1. *Arc length*

Given an implicit function $F(x, y) = 0$, assume that x and y, and thereby the arc length s, are all smooth functions of an independent variable u. Then by the arc length relation $(\mathrm{d}s/\mathrm{d}x)^2 = 1 + (\mathrm{d}y/\mathrm{d}x)^2$ and by differentiating (2.3) with respect to u to obtain $\mathrm{d}\mathbf{p}/\mathrm{d}u = \mathbf{J}^{-1}\,\mathrm{d}\boldsymbol{\alpha}/\mathrm{d}u$, we also have

$$\begin{aligned}\left(\frac{\mathrm{d}s}{\mathrm{d}u}\right)^2 &= \left(\frac{\mathrm{d}x}{\mathrm{d}u}\right)^2 + \left(\frac{\mathrm{d}y}{\mathrm{d}u}\right)^2 = \left(\frac{\mathrm{d}\mathbf{p}^{\mathrm{T}}}{\mathrm{d}u}\right)\left(\frac{\mathrm{d}\mathbf{p}}{\mathrm{d}u}\right)\\ &= \left(\frac{\mathrm{d}\boldsymbol{\alpha}^{\mathrm{T}}}{\mathrm{d}u}\right)\mathbf{J}^{-T}\mathbf{J}^{-1}\left(\frac{\mathrm{d}\boldsymbol{\alpha}}{\mathrm{d}u}\right)\\ &= \mathrm{d}_{13}^2\left(\frac{\mathrm{d}\alpha_1}{\mathrm{d}u}\right)^2 + 2c_{12}\left(\frac{\mathrm{d}\alpha_1}{\mathrm{d}u}\right)\left(\frac{\mathrm{d}\alpha_2}{\mathrm{d}u}\right) + \mathrm{d}_{23}^2\left(\frac{\mathrm{d}\alpha_2}{\mathrm{d}u}\right)^2,\end{aligned} \tag{2.8}$$

where

$$d_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2},$$
$$c_{12} = (x_1 - x_3)(x_2 - x_3) + (y_1 - y_3)(y_2 - y_3)\ .$$

2.1.2. Curvature

Given an implicit function $F(x, y) = 0$, we have $\mathrm{d}F = 0$, or

$$F_x\,\mathrm{d}x + F_y\,\mathrm{d}y = 0 \tag{2.9}$$

Assuming that in a small neighborhood of point $(x, y)T$, y is a function of the independent variable x, from (2.9) we have

$$y_x = -\frac{F_x}{F_y}. \tag{2.10}$$

By differentiating (2.10) with respect to x, we obtain

$$y_{xx} = -\frac{\partial}{\partial x}\left(\frac{F_x}{F_y}\right) = -\frac{F_{yy}F_x^2 + F_{xx}F_y^2 - 2F_{xy}F_xF_y}{F_y^3}.$$

This allows us to express the curvature of the function $F(x, y) = 0$, or $S(\alpha_1, \alpha_2) = 0$ in BB form, as

$$\begin{aligned}\kappa &= \frac{|y_{xx}|}{(1 + y_x^2)^{3/2}} = \frac{|\nabla F^{\mathrm{T}}\mathbf{P}^{\mathrm{T}}\nabla^2 F\mathbf{P}\nabla F|}{(\nabla F^{\mathrm{T}}\nabla F)^{3/2}}\\ &= \frac{|\nabla S^{\mathrm{T}}\mathbf{J}\mathbf{P}^{\mathrm{T}}\mathbf{J}^{\mathrm{T}}\nabla^2 S\mathbf{J}\mathbf{P}\mathbf{J}^{\mathrm{T}}\nabla S|}{(\nabla S^{\mathrm{T}}\mathbf{J}\mathbf{J}^{\mathrm{T}}\nabla S)^{3/2}},\end{aligned} \tag{2.11}$$

where the permutation matrix $\mathbf{P}$ is defined as

$$\mathbf{P} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}.$$

2.2. A-splines

An A-spline is a smooth zero contour curve of a bivariate polynomial in BB form defined within a triangle (Bajaj and Xu, 1992). Fig. 1 shows examples of A-splines that are defined in (Bajaj and Xu, 1992).

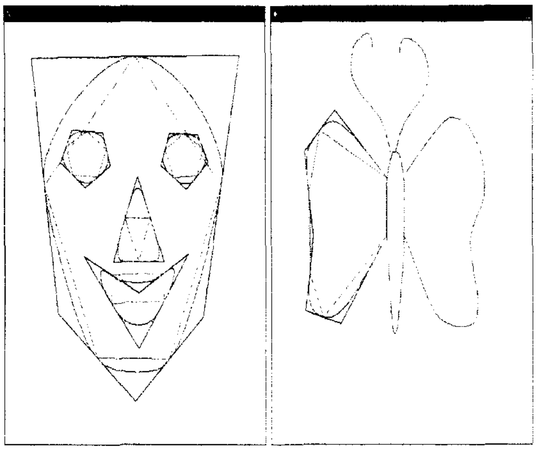


Fig. 1. Interactively designed contours with cubic A-splines.

The papers (Bajaj and Xu, 1992; 1996) explore the possibilities of building piecewise smooth curves that interpolate or approximate given polygonal data sets.

For further formulation of elastic models we briefly describe the A-spline in the following way. The formulation includes the A-spline (Bajaj and Xu, 1992), the 2D counterpart of the A-patch, which treats piecewise implicit surfaces (Bajaj et al., 1995). A piecewise A-spline curve consists of the zero contour of some piecewise smooth BB polynomials defined over a simplicial hull Σ, or a triangulation of a connected region of the space. In particular,

$$F^{(i)}(\boldsymbol{\alpha}) = \sum_{|\boldsymbol{\lambda}|=d} b_{\boldsymbol{\lambda}}^{(i)} B_{\boldsymbol{\lambda}}^{d}(\boldsymbol{\alpha}) = \mathbf{b}^{(i)\mathrm{T}} \mathbf{B}^{d} = 0$$

is the zero contour of a BB polynomial defined within the ith simplex in Σ. The base functions $B_{\boldsymbol{\lambda}}^{d}$ are grouped into a vector $\mathbf{B}$ and the coefficients $b_{\boldsymbol{\lambda}}$ into $\mathbf{b}$. Smoothness of certain degrees and local interpolation of certain degrees are enforced by some linear equality constraints

$$\mathbf{b}^{\mathrm{T}}\mathbf{C}(\mathbf{p}) = \mathbf{0}, \tag{2.12}$$

and connectedness of the curve is enforced by additional linear sign inequalities

$$\mathbf{b}^{\mathrm{T}}\mathbf{S} > \mathbf{0}. \tag{2.13}$$

See Fig. 2 for examples of sign equalities and inequalities of C^1 and C^3 cubic A-spline segments.

Vector $\mathbf{b}$ is a global collection of the coefficient vector $\mathbf{b}^{(i)}$ of all simplexes in Σ, and $\mathbf{S}$ and $\mathbf{C}(\mathbf{p})$ are defined explicitly for A-splines with C^k continuity in Bajaj and Xu (1992).

The partial derivatives of a BB polynomial $F(\boldsymbol{\alpha}) = \mathbf{b}^{\mathrm{T}}\mathbf{B}$ are given by

$$F_{\alpha_j} = \mathbf{b}^{\mathrm{T}}\mathbf{B}_{\alpha_j}^{d}(\boldsymbol{\alpha}) = \mathbf{b}^{\mathrm{T}}\mathcal{D}_j^{d}\mathbf{B}^{d-1}(\boldsymbol{\alpha}),$$

$$F_{\alpha_j\alpha_k} = \mathbf{b}^{\mathrm{T}}\mathbf{B}_{\alpha_j\alpha_k}(\boldsymbol{\alpha}) = \mathbf{b}^{\mathrm{T}}\mathcal{D}_k^{d}\mathcal{D}_j^{d-1}\mathbf{B}^{d-2}(\boldsymbol{\alpha}).$$

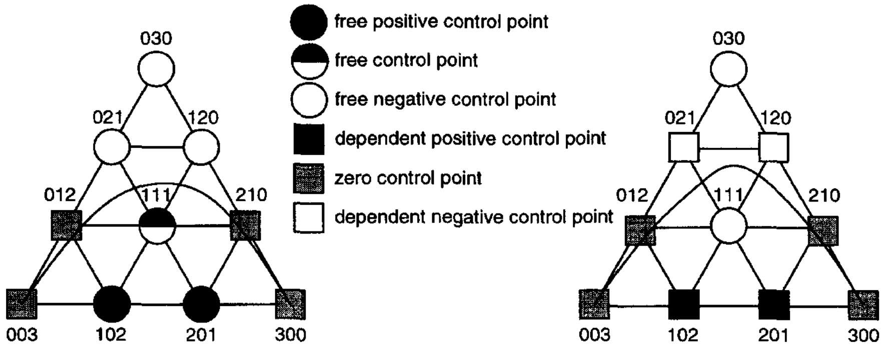


Fig. 2. Sign inequalities for the coefficients of (left) C^1 and (right) C^3 cubic A-splines. For $i+j+k=3$, the symbol labeled by ijk indicates the sign restriction, if any, for the coefficient b_{ijk}.

The $\mathcal{D}_j^d$ are matrices that relate the derivatives of the basis elements of $\mathbf{B}^d$ to the basis elements of $\mathbf{B}^{d-1}$. The $\mathcal{D}_j^d$ are independent of $\mathbf{b}$ and have dimensions $\binom{d+2}{2} \times \binom{d+1}{2}$. For example, with $d=2$, we have

$$
\begin{aligned}
\mathbf{B}^d &= \left[\alpha_1^2, 2\alpha_1\alpha_2, \alpha_2^2, 2\alpha_1(1-\alpha_1-\alpha_2), 2\alpha_2(1-\alpha_1-\alpha_2), (1-\alpha_1-\alpha_2)^2\right]^{\mathrm{T}},\\
\mathbf{B}^{d-1} &= \left[\alpha_1, \alpha_2, 1-\alpha_1-\alpha_2\right]^{\mathrm{T}},\\
\mathbf{B}_{\alpha_1}^d &= \left[2\alpha_1, 2\alpha_2, 0, -4\alpha_1-2\alpha_2+2, -2\alpha_2, 2\alpha_1+2\alpha_2-2\right]^{\mathrm{T}},\\
\mathbf{B}_{\alpha_2}^d &= \left[0, 2\alpha_1, 2\alpha_2, -2\alpha_1, -2\alpha_1-4\alpha_2+2, 2\alpha_1+2\alpha_2-2\right]^{\mathrm{T}},
\end{aligned}
$$

and

$$
\mathcal{D}_1^2 = 2\begin{bmatrix} 1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 0\\ -1 & 0 & 1\\ 0 & -1 & 0\\ 0 & 0 & -1 \end{bmatrix}, \qquad
\mathcal{D}_2^2 = 2\begin{bmatrix} 0 & 0 & 0\\ 1 & 0 & 0\\ 0 & 1 & 0\\ -1 & 0 & 0\\ 0 & -1 & 1\\ 0 & 0 & -1 \end{bmatrix}.
$$

3. Energy models for A-splines

3.1. Elastic curves

Let a be the material coordinate of a point on a plane curve C with parameterization $\mathbf{w}(a)$. For a parametric representation $\mathbf{w}(a)$, the elastic potential energy of a deformable curve is as follows:

$$
E = \int_C \left[\beta(a)|\mathbf{w}_a|^2 + \gamma(a)|\mathbf{w}_{aa}|^2\right] \mathrm{d}a \tag{3.1}
$$

where $\beta(a)$ and $\gamma(a)$ are material functions.

Most of the elastic curves that are built on parameterizations other than the intrinsic, are referred to as "elastica" (Jou and Han, 1990a; Mumford, 1994).

The first step of defining the elastic energy of a geometric entity is actually the mapping between the material coordinates and some parameterization of the geometric entity. For a piecewise representation, the joining points of the pieces could be identified as such a parameterization. Namely, each joining point is associated with fixed material coordinates during a deformation process. However, such a parameterization is too coarse. A second level parameterization is needed to describe detail changes, especially when there is substantial freedom within each entity. In the following energy curves, we assume that the material is uniformly distributed along the curve. Thus in place of the material coordinate a we will use the arc length coordinate s, and the elastic potential energy in (3.1) becomes (Mumford, 1994):

$$E = \int_C (\beta + \gamma\kappa^2)\,\mathrm{d}s. \tag{3.2}$$

The terms β and $\gamma\kappa^2$ represent the stretching and bending energy, respectively.

3.2. Elastic strain energy model of A-splines

Let $S(\alpha_1, \alpha_2) = 0$ be an A-spline defined within a triangle $\Delta\mathbf{p}_1\mathbf{p}_2\mathbf{p}_3$ (see Fig. 3). The curved piece interpolates $\mathbf{p}_1$ and $\mathbf{p}_3$ and is tangent to $\overline{\mathbf{p}_1\mathbf{p}_2}$ and $\overline{\mathbf{p}_3\mathbf{p}_2}$ at $\mathbf{p}_1$ and $\mathbf{p}_3$ respectively. Let $F(x, y) = 0$ be the representation of the spline in Cartesian coordinates. The Cartesian coordinates of $\mathbf{p}_1$, $\mathbf{p}_2$ and $\mathbf{p}_3$ are $(x_1, y_1)T$, $(x_2, y_2)T$ and $(x_3, y_3)T$, respectively, and their local barycentric coordinates, suppressing the third coordinate $\alpha_3 = 1 - \alpha_1 - \alpha_2$, are $(1, 0)T$, $(0, 1)T$, and $(0, 0)T$, respectively.

For several purposes, such as computing the energy (3.2) below, we need to express the spline coordinates as functions of a single parameter, say u. One effective way of doing this is to parametrize the A-spline within a control triangle by finding its point of intersection

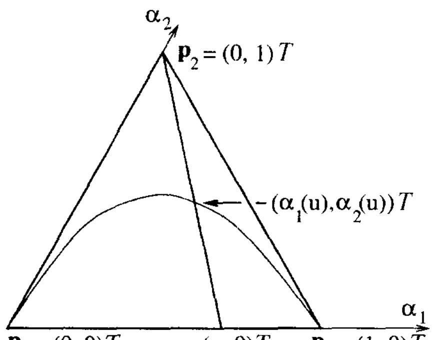


Fig. 3. Representation of points in local BB coordinates. Point $\mathbf{p}_i$ has Cartesian coordinates $(x_i, y_i)T$, $i = 1, 2, 3$. A point $(\alpha_1(u), \alpha_2(u))T$ is given by the intersection of the spline with the line connecting the points $(u, 0)T$ and $(0, 1)T$ in BB coordinates ($\mathbf{p}_2$).

with a line segment connecting the apex point ($\mathbf{p}_2$) to a point on the base ($\overline{\mathbf{p}_3\mathbf{p}_1}$). It was shown in (Bajaj and Xu, 1992) that such a line segment always intersects the A-spline exactly once when the constraints (2.12) and (2.13) are satisfied, and there this technique was used to obtain parametrizations of various quadratic and cubic A-splines. We simply let u parametrize the line segment $\overline{\mathbf{p}_3\mathbf{p}_1}$, so that in BB coordinates, $\overline{\mathbf{p}_3\mathbf{p}_1}$ is given by $(u, 0)T$, $0 \leqslant u \leqslant 1$.

Now using (3.2), (2.11), and (2.8) we can write the total energy as

$$E_{\text{total}} = \int_C (\beta + \gamma\kappa^2)\,\mathrm{d}s = \int_0^1 \left[\left(\frac{\mathrm{d}\boldsymbol{\alpha}^{\mathrm{T}}}{\mathrm{d}u}\right)\mathbf{J}^{-T}\mathbf{J}^{-1}\left(\frac{\mathrm{d}\boldsymbol{\alpha}}{\mathrm{d}u}\right)\right]^{1/2}(\beta + \gamma\kappa^2)\,\mathrm{d}u$$
$$= \int_0^1 \left[\left(\frac{\mathrm{d}\boldsymbol{\alpha}^{\mathrm{T}}}{\mathrm{d}u}\right)\mathbf{J}^{-T}\mathbf{J}^{-1}\left(\frac{d\alpha}{\mathrm{d}u}\right)\right]^{1/2}\left[\beta + \frac{\gamma(\nabla S^{\mathrm{T}}\mathbf{J}\mathbf{P}^{\mathrm{T}}\mathbf{J}^{\mathrm{T}}\nabla^2 S\mathbf{J}\mathbf{P}\mathbf{J}^{\mathrm{T}}\nabla S)^2}{(\nabla S^{\mathrm{T}}\mathbf{J}\mathbf{J}^{\mathrm{T}}\nabla S)^3}\right]\mathrm{d}u. \tag{3.3}$$

3.3. *Simplified elastic strain energy model of A-splines*

Since the integrand, call it $g(u)$, in (3.3) cannot be integrated symbolically in general, we wish to find an approximation to the total energy that can be computed quickly compared to the time-consuming numerical integration involved with (3.3). An ideal candidate is apply Simpson's rule. Using just three points in the interest of computational speed, this gives the approximation $E_{\text{total}} = [g(0) + 4g(1/2) + g(1)]/6$. However, for many common parametrizations $g(u) \to \infty$ as $u \to 0$ or 1. In this case we make a change of variable to eliminate the singularity, and use a Simpson's rule approximation to the result.

4. Energy optimization

We now consider how to minimize the different energy functions of the earlier section over the constrained degrees of freedom (domain vertices and control weights) of the A-spline curve.

4.1. *Local minimization of total energy*

Here we take into account both the bending and stretching energy as defined in Eq. (3.2) and expressed by Eq. (3.3). The minimization is obtained locally by varying the free weights of each individual A-spline curve within its triangle.

An energy-minimized setting is a solution to

$$\nabla_{\mathbf{b}} E_{\text{total}} = \mathbf{0} \tag{4.1}$$

or

$$\int_0^1 \nabla_{\mathbf{b}}\left\{\left[\left(\frac{\mathrm{d}\boldsymbol{\alpha}^{\mathrm{T}}}{\mathrm{d}u}\right)\mathbf{J}^{-T}\mathbf{J}^{-1}\left(\frac{\mathrm{d}\boldsymbol{\alpha}}{\mathrm{d}u}\right)\right]^{1/2}\left[\beta + \frac{\gamma(\nabla S^{\mathrm{T}}\mathbf{J}\mathbf{P}^{\mathrm{T}}\mathbf{J}^{\mathrm{T}}\nabla^2 S\mathbf{J}\mathbf{P}\mathbf{J}^{\mathrm{T}}\nabla S)^2}{(\nabla S^{\mathrm{T}}\mathbf{J}\mathbf{J}^{\mathrm{T}}\nabla S)^3}\right]\right\}\mathrm{d}u = 0.$$

4.1.1. Exact solutions

System (4.1) is in general a nonlinear system of $\mathbf{b}$. A nonlinear system is not guaranteed to be solvable. However, by restricting the freedom to one variable, we reduce to the system to a univariate nonlinear system, which is easy to solve.

Let $\mathbf{b}$ be a vector function of some parameter t. For an A-spline curve, system (4.1) is reduced to

$$\frac{\mathrm{d}E(t)}{\mathrm{d}t} = 0.$$

Let

$$f(t) = \frac{\mathrm{d}E(t)}{\mathrm{d}t}$$
$$= \int_0^1 \frac{\mathrm{d}}{\mathrm{d}t}\left\{\left[\left(\frac{\mathrm{d}\boldsymbol{\alpha}^{\mathrm{T}}}{\mathrm{d}u}\right)\mathbf{J}^{-T}\mathbf{J}^{-1}\left(\frac{\mathrm{d}\boldsymbol{\alpha}}{\mathrm{d}u}\right)\right]^{1/2}\left[\beta + \frac{\gamma(\nabla S^{\mathrm{T}}\mathbf{J}\mathbf{P}^{\mathrm{T}}\mathbf{J}^{\mathrm{T}}\nabla^2 S\mathbf{J}\mathbf{P}\mathbf{J}^{\mathrm{T}}\nabla S)^2}{(\nabla S^{\mathrm{T}}\mathbf{J}\mathbf{J}^{\mathrm{T}}\nabla S)^3}\right]\right\}\mathrm{d}u$$

so that

$$f'(t) = \int_0^1 \frac{\mathrm{d}^2}{\mathrm{d}t^2}\left\{\left[\left(\frac{\mathrm{d}\boldsymbol{\alpha}^{\mathrm{T}}}{\mathrm{d}u}\right)\mathbf{J}^{-T}\mathbf{J}^{-1}\left(\frac{\mathrm{d}\boldsymbol{\alpha}}{\mathrm{d}u}\right)\right]^{1/2}\right.$$
$$\left.\times\left[\beta + \frac{\gamma(\nabla S^{\mathrm{T}}\mathbf{J}\mathbf{P}^{\mathrm{T}}\mathbf{J}^{\mathrm{T}}\nabla^2 S\mathbf{J}\mathbf{P}\mathbf{J}^{\mathrm{T}}\nabla S)^2}{(\nabla S^{\mathrm{T}}\mathbf{J}\mathbf{J}^{\mathrm{T}}\nabla S)^3}\right]\right\}\mathrm{d}u.$$

$f(t)$ and $f'(t)$ are both continuous.

In order to solve these equations, we need a parametrization of the A-spline.

We can use standard methods, such as Newton's method, to solve for the roots. Note that the evaluations of $f(t)$ and $f'(t)$ would involve numerical integrations.

4.1.2. Low degree A-splines

We now proceed to derive parametrizations of linear, quadratic, and cubic A-splines. Doing the linear case first, we observe that a parametrization of a general line in BB form

$$S(\alpha_1, \alpha_2) = b_{100}\alpha_1 + b_{010}\alpha_2 + b_{001}(1 - \alpha_1 - \alpha_2) = 0 \tag{4.2}$$

is given by

$$\alpha_1 = u, \qquad \alpha_2 = 0, \qquad (\alpha_3 = 1 - u), \quad 0 \leqslant u \leqslant 1. \tag{4.3}$$

Then $\nabla S = [0\ 1]^{\mathrm{T}}$ and $\nabla^2 S$ is the 2×2 zero matrix, so $\kappa = 0$ in (2.11) and (3.3), which must be the case for a straight line segment. Thus only the stretching energy is present in this formulation. Eq. (3.3) then reduces to

$$E_{\text{total}} = \beta d_{13}, \tag{4.4}$$

a multiple of the arc length, here the distance from $\mathbf{p}_1$ to $\mathbf{p}_3$, as expected.

Now we present an example of an energy-minimizing quadratic A-spline with C^1 continuity. First we derive the general equation for such a curve in BB form. The general quadratic spline is $S(\alpha_1, \alpha_2) = b_{200}\alpha_1^2 + 2b_{110}\alpha_1\alpha_2 + 2b_{101}\alpha_1(1 - \alpha_1 - \alpha_2) + b_{020}\alpha_2^2 +$

$2b_{011}\alpha_2(1-\alpha_1-\alpha_2)+b_{002}(1-\alpha_1-\alpha_2)^2=0$. However, $S(1,0)=0 \Rightarrow b_{200}=0$ and $S(0,0)=0 \Rightarrow b_{002}=0$. Furthermore, the tangent to $S(\alpha_1,\alpha_2)$ at $(1,0)T$, is parallel to the line $\alpha_1+\alpha_2=1$. Therefore $\mathrm{d}\alpha_2/\mathrm{d}\alpha_1=-(\mathrm{d}S/\mathrm{d}\alpha_1)/(\mathrm{d}S/\alpha_2)=-1$, or $\mathrm{d}S/\mathrm{d}\alpha_1=\mathrm{d}S/\mathrm{d}\alpha_2$ at that point. This now implies that $b_{110}=0$. Also, the tangent to $S(\alpha_1,\alpha_2)$ at $(0,0)T$, is parallel to the α_2-axis. Therefore $\mathrm{d}S/\mathrm{d}\alpha_2=0$ there, and this implies that $b_{011}=0$. We now have

$$S(\alpha_1,\alpha_2)=2b_{101}\alpha_1(1-\alpha_1-\alpha_2)+b_{020}\alpha_2^2=0, \tag{4.5}$$

in accordance with Bajaj and Xu (1996).

We now parametrize the quadratic spline as described in Section 3.2. Intersecting the curve (4.5) with the line $\alpha_1=u(1-\alpha_2)$ yields (Bajaj and Xu, 1992):

$$\alpha_1(u;\mathbf{b})=\frac{u}{1+\sqrt{2b_{101}u(1-u)}},\qquad \alpha_2(u;\mathbf{b})=\frac{\sqrt{2b_{101}u(1-u)}}{1+\sqrt{2b_{101}u(1-u)}},$$
$$0\leqslant u\leqslant 1. \tag{4.6}$$

Here $\mathbf{b}$ denotes the column vector of the coefficients b_{ijk}, and in this case $\mathbf{b}$ consists of the single element b_{101}.

We find that

$$\mathbf{J}^{-T}\mathbf{J}^{-1}=\begin{bmatrix} d_{13}^2 & c_{12}\\ c_{12} & d_{23}^2\end{bmatrix},\qquad \mathbf{J}\mathbf{P}^{\mathrm{T}}\mathbf{J}^{\mathrm{T}}=\begin{bmatrix} 0 & -\frac{1}{\Delta}\\ \frac{1}{\Delta} & 0\end{bmatrix},$$

$$\mathbf{J}\mathbf{P}\mathbf{J}^{\mathrm{T}}=\begin{bmatrix} 0 & \frac{1}{\Delta}\\ -\frac{1}{\Delta} & 0\end{bmatrix},\qquad \mathbf{J}\mathbf{J}^{\mathrm{T}}=\begin{bmatrix} \frac{d_{23}^2}{\Delta^2} & -\frac{c_{12}}{\Delta^2}\\ -\frac{c_{12}}{\Delta^2} & \frac{d_{13}^2}{\Delta^2}\end{bmatrix},$$

$$\begin{aligned}&\left[\left(\frac{\mathrm{d}\boldsymbol{\alpha}^{\mathrm{T}}}{\mathrm{d}u}\right)\mathbf{J}^{-T}\mathbf{J}^{-1}\left(\frac{\mathrm{d}\boldsymbol{\alpha}}{\mathrm{d}u}\right)\right]^{1/2}\\
&\quad=\big\{4u(1-u)d_{13}^2+2b\sqrt{u(1-u)}\big[(1-2u)\big(d_{23}^2-d_{12}^2\big)+d_{13}^2\big]\\
&\qquad+b^2\big[-u(1-2u)d_{12}^2+u(1-u)d_{13}^2\\
&\qquad+(1-u)(1-2u)d_{23}^2\big]\big\}^{1/2}\big/\big\{2\sqrt{u(1-u)}\big[1+b\sqrt{u(1-u)}\big]^2\big\}\end{aligned}$$

and that the curvature satisfies

$$\begin{aligned}\kappa^2&=\frac{(\nabla S^{\mathrm{T}}\mathbf{J}\mathbf{P}^{\mathrm{T}}\mathbf{J}^{\mathrm{T}}\nabla^2 S\mathbf{J}\mathbf{P}\mathbf{J}^{\mathrm{T}}\nabla S)^2}{(\nabla S^{\mathrm{T}}\mathbf{J}\mathbf{J}^{\mathrm{T}}\nabla S)^3}\\
&=4b^2\Delta^2\big[1+b\sqrt{u(1-u)}\big]^6\Big/\Big\{4u(1-u)d_{13}^2\\
&\quad+2b\sqrt{u(1-u)}\big[(1-2u)\big(d_{23}^2-d_{12}^2\big)+d_{13}^2\big]\\
&\quad+b^2\big[-u(1-2u)d_{12}^2+u(1-u)d_{13}^2+(1-u)(1-2u)d_{23}^2\big]\Big\}^3.\end{aligned}$$

Consequently the total energy as given by (3.3) equals

$$E_{\text{total}} = \int_0^1 \Big\{ \beta \Big(\big\{ 4u(1-u)d_{13}^2 + 2b\sqrt{u(1-u)}\big[(1-2u)\big(d_{23}^2 - d_{12}^2\big) + d_{13}^2\big] + b^2\big[-u(1-2u)d_{12}^2 + u(1-u)d_{13}^2 + (1-u)(1-2u)d_{23}^2\big]\big\}^{1/2} \Big/ \Big\{2\sqrt{u(1-u)}\big[1 + b\sqrt{u(1-u)}\big]^2\Big\}\Big) + \gamma\Big[2b^2\Delta^2\big[1 + b\sqrt{u(1-u)}\big]^4 \Big/ \Big(\sqrt{u(1-u)}\Big\{4u(1-u)d_{13}^2 + 2b\sqrt{u(1-u)}\big[(1-2u)\big(d_{23}^2 - d_{12}^2\big) + d_{13}^2\big] + b^2\big[-u(1-2u)d_{12}^2 + u(1-u)d_{13}^2 + (1-u)(1-2u)d_{23}^2\big]\Big\}^{5/2}\Big)\Big]\Big\}\, du. \tag{4.7}$$

Note that as $b \to 0$, the conic spline approaches the line segment from p_1 to p_3, and $E_{\text{total}} \to \beta d_{13}$, in accord with (4.4).

Numerical integration of (4.7) can be tricky because the integrand, call it $g(b, u)$, goes to infinity as u approaches 0 or 1. It is true that $g(b, u)$ is integrable over $[0, 1]$ since infinity is approached as $u^{-1/2}$ as $u \to 0$ and as $(1-u)^{-1/2}$ as $u \to 1$. Numerical evaluation of this integral is made easier with the following changes of variable: Let $u = v^2$ for $u \in [0, 1/2]$ and let $u = 1 - w^2$ for $u \in [1/2, 1]$. These substitutions eliminate the singularities at the endpoints, and we have this equivalent expression for the total energy:

$$E_{\text{total}} = \int_0^{1/\sqrt{2}} \Big[\beta\Big(\Big\{4v^2(1-v^2)d_{13}^2 + 2bv\sqrt{1-v^2}\big[(1-2v^2)\big(d_{23}^2 - d_{12}^2\big) + d_{13}^2\big] + b^2\big[-v^2(1-2v^2)d_{12}^2 + v^2(1-v^2)d_{13}^2 + (1-v^2)(1-2v^2)d_{23}^2\big]\Big\}^{1/2} \Big/ \Big[\sqrt{1-v^2}\big(1 + bv\sqrt{1-v^2}\big)^2\Big]\Big) + 4\gamma\Delta^2 b^2\Big(\big(1 + bv\sqrt{1-v^2}\big)^4 \Big/ \sqrt{1-v^2} \times \Big\{4v^2(1-v^2)d_{13}^2 + 2bv\sqrt{1-v^2}\big[(1-2v^2)\big(d_{23}^2 - d_{12}^2\big) + d_{13}^2\big] + b^2\big[-v^2(1-2v^2)d_{12}^2 + v^2(1-v^2)d_{13}^2 + (1-v^2)(1-2v^2)d_{23}^2\big]\Big\}^{5/2}\Big)\Big]\, dv$$

$$+ \int_0^{1/\sqrt{2}} \Big[\beta\Big(\Big\{4w^2(1-w^2)d_{13}^2 + 2bw\sqrt{1-w^2} \times \big[(1-2w^2)\big(d_{12}^2 - d_{23}^2\big) + d_{13}^2\big] + b^2\big[(1-w^2)(1-2w^2)d_{12}^2 + w^2(1-w^2)d_{13}^2 - w^2(1-2w^2)d_{23}^2\big]\Big\}^{1/2} \Big/ \big[\sqrt{1-w^2}\big(1 + bw\sqrt{1-w^2}\big)^2\big]\Big)$$

$$+ 4\gamma\Delta^2 b^2\Big(\big(1 + bw\sqrt{1-w^2}\big)^4 \Big/ \sqrt{1-w^2}\Big\{4w^2(1-w^2)d_{13}^2 + 2bw\sqrt{1-w^2} \times \big[(1-2w^2)(d_{12}^2 - d_{23}^2) + d_{13}^2\big] + b^2\big[(1-w^2)(1-2w^2)d_{12}^2 + w^2(1-w^2)d_{13}^2 - w^2(1-2w^2)d_{23}^2\big]\Big\}^{5/2}\Big)\Big]\,\mathrm{d}w. \tag{4.8}$$

In a similar manner we can construct cubic A-splines with C^3 continuity. The equation of the general C^3 cubic is

$$S(\alpha_1,\alpha_2) = 3b_{201}\alpha_1^2(1-\alpha_1-\alpha_2) + 3b_{120}\alpha_1\alpha_2^2 + 6b_{111}\alpha_1\alpha_2(1-\alpha_1-\alpha_2) + 3b_{102}\alpha_1(1-\alpha_1-\alpha_2)^2 + 3b_{021}\alpha_2^2(1-\alpha_1-\alpha_2) - \alpha_2^3 = 0, \tag{4.9}$$

in accordance with Bajaj and Xu (1992, 1996). When one specifies the second and third order derivatives at $\mathbf{p}_1$ and $\mathbf{p}_3$, it is possible to express the four coefficients b_{201}, b_{120}, b_{102}, and b_{021} as linear functions of the remaining coefficient, b_{111}. When we intersect $S(\alpha_1,\alpha_2) = 0$ with the line $\alpha_1 = u(1-\alpha_2)$, we obtain a parametrization in which $\alpha_1 = u(1-\alpha_2)$, and α_2 is a root of the cubic equation in α_2. If we make the substitution $\alpha_2 = 1/(1+t)$, we obtain this cubic polynomial in t:

$$At^3 + Bt^2 + Ct - 1 = 0, \tag{4.10}$$

where

$$A = 3u(1-u)[b_{201}u + b_{102}(1-u)],$$
$$B = 6u(1-u)b_{111},$$
$$C = 3[b_{120}u + b_{021}(1-u)].$$

The inequality constraints for cubic A-splines are

$$b_{201} > 0, \qquad b_{102} > 0, \qquad b_{021} \leqslant 0, \qquad b_{120} \leqslant 0, \tag{4.11}$$

so we have $A \geqslant 0$ and $C \leqslant 0$ for all $u \in [0,1]$.

Now let

$$p = \frac{3AC - B^2}{3A^2}, \qquad q = \frac{2B^3 - 9ABC - 27A^2}{27A^3}.$$

If $(q/2)^2 + (p/3)^3 \geqslant 0$, then since $t(u) = [1-\alpha_2(u)]/\alpha_2(u)$ is a real root of (4.10), we must have

$$t(u) = \sqrt[3]{-\frac{q}{2} + \sqrt{\left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3}} + \sqrt[3]{-\frac{q}{2} - \sqrt{\left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3}} - \frac{B}{3A}. \tag{4.12}$$

If $(q/2)^2 + (p/3)^3 < 0$, then p must be negative, and (4.10) has three real roots. Define

$$r = \frac{(-p)^{3/2}}{3\sqrt{3}}, \qquad \theta = \frac{1}{3}\cos^{-1}\left(-\frac{q}{2r}\right).$$

Then the root we want is

$$t(u) = 2r^{1/3}\cos\theta - \frac{B}{3A}. \tag{4.13}$$

The final cubic parametrization is then

$$\alpha_1(u;\mathbf{b}) = u\big[1-\alpha_2(u;\mathbf{b})\big], \qquad \alpha_2(u;\mathbf{b}) = \frac{1}{1+t(u;\mathbf{b})}, \qquad 0 \leqslant u \leqslant 1, \tag{4.14}$$

where $t(u;\mathbf{b})$ is given by (4.12) or (4.13).

Expressions for the total energy of the C^3-continuous cubic A-spline analogous to (4.7) and (4.8) for the C^1-continuous quadratic case can be derived, but will be omitted here due to their great length and complexity.

Case study

Suppose a quadratic curve $F(x,y)=0$ passes through the points $\mathbf{p}_3 = (0,0)T$ and $\mathbf{p}_1 = (1,0)T$, and that we are given that the tangent lines at these points have slopes 2 and -3, respectively. Then the intersection of the tangent lines is $\mathbf{p}_2 = (3/5, 6/5)T$. In this case (2.3) gives $(\alpha_1,\alpha_2)T = ((2x-y)/2,\ 5y/6)T$. We set $b_{020} = -1$, and take $\beta = 1$ and $\gamma = 1$. Then $d_{12} = 2\sqrt{10}/5$, $d_{13} = 1$, $d_{23} = 3\sqrt{5}/5$, $\Delta = 6/5$, and the integral we wish to minimize is this specialization of (4.8):

$$\begin{aligned}
E_{\text{total}} = \int_0^{1/\sqrt{2}} &\Big[\Big(\Big\{4v^2(1-v^2) + 4bv\sqrt{1-v^2}(3-v^2)/5 \\
&\quad + b^2\big[(29v^4-30v^2+9)/5\big]\Big\}^{1/2} \Big/ \Big[\sqrt{1-v^2}\big(1+bv\sqrt{1-v^2}\big)^2\Big]\Big) \\
&\quad + (144/25)b^2\Big(\big(1+bv\sqrt{1-v^2}\big)^4 \Big/ \sqrt{1-v^2}\Big\{4v^2(1-v^2) \\
&\quad + 4bv\sqrt{1-v^2}(3-v^2)/5 + b^2\big[(29v^4-30v^2+9)/5\big]\Big\}^{5/2}\Big)\Big]\,\mathrm{d}v \\
+ \int_0^{1/\sqrt{2}} &\Big[\Big(\Big\{4w^2(1-w^2) + 4bw\sqrt{1-w^2}(2+w^2)/5 \\
&\quad + b^2\big[(29w^4-28w^2+8)/5\big]\Big\}^{1/2} \Big/ \\
&\quad \Big[\sqrt{1-w^2}\big(1+bw\sqrt{1-w^2}\big)^2\Big]\Big) \\
&\quad + (144/25)b^2\Big(\big(1+bw\sqrt{1-w^2}\big)^4 \Big/ \\
&\quad \sqrt{1-w^2}\Big\{4w^2(1-w^2) + 4bw\sqrt{1-w^2}(2+w^2)/5 \\
&\quad + b^2\big[(29w^4-28w^2+8)/5\big]\Big\}^{5/2}\Big)\Big]\,\mathrm{d}w.
\end{aligned} \tag{4.15}$$

The function in (4.15) could not be integrated symbolically, so it and its derivative with respect to b were integrated numerically for several different values of $b = \sqrt{b_{101}}$. The

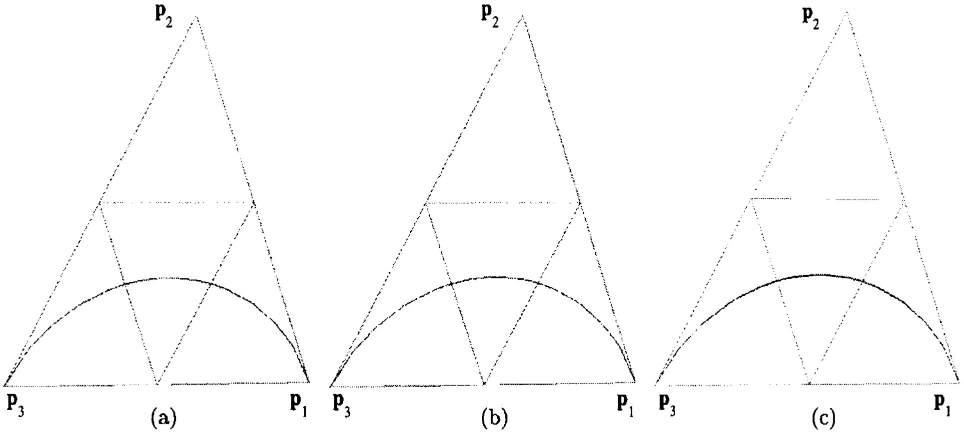


Fig. 4. In these figures $\mathbf{p}_1 = (1,0)T$, $\mathbf{p}_2 = (3/5, 6/5)T$, and $\mathbf{p}_3 = (0,0)T$ in Cartesian coordinates. (a): The exact energy-minimizing quadratic A-spline (4.16), with C^1 continuity at the endpoints, obtained when $b_{101} = 0.356$. (b) The simplified energy-minimizing A-spline (4.19), obtained when $b_{101} = 0.366$. (c) Superposition of the curves in (a) and (b).

integral attained its minimum value of 5.63763 at $b = 0.844$, or $b_{101} = 0.356$. This gives an ellipse whose equation in Cartesian coordinates is

$$6.408x^2 - 1.068xy + 5.182y^2 - 6.408x + 3.204y = 0\,. \tag{4.16}$$

This elliptic arc is shown in Fig. 4(a).

4.2. Local minimization of simplified energy

Since the expression (4.7) for the strain energy is quite complicated, a simplified form of the energy may be desired as indicated in Section 3.3. We would like a simple approximation to the integral in (4.7). This can be tricky because the integrand, call it $g(b,u)$, goes to infinity as u approaches 0 or 1. It is true that $g(b,u)$ is integrable over $[0,1]$ since infinity is approached as $u^{-1/2}$ as $u \to 0$ and as $(1-u)^{-1/2}$ as $u \to 1$. Numerical evaluation of this integral is made easier with the following changes of variable: Let $u = v^2$ for $u \in [0, 1/2]$ and let $u = (1-w)^2$ for $u \in [1/2, 1]$. These substitutions eliminate the singularities at the endpoints, and we have this equivalent expression for the total energy:

We now apply Simpson's rule using the three points at v or $w = 0$, $1/2\sqrt{2}$, and $1/\sqrt{2}$. This approximation is best for moderate values of b, say $0.8 \leqslant b \leqslant 4 \leftrightarrow 0.32 \leqslant b_{101} \leqslant 8$, and naturally more accurate approximations may be obtained using more points. We select this approximation in the interest of computational speed. The result is

$$\begin{aligned}E_{\text{simp}} = \beta\biggl(&\frac{b(d_{12}+d_{23})}{6\sqrt{2}} + \frac{2d_{13}}{3(2+b)} + \frac{32}{3\sqrt{7}(8+\sqrt{7}\,b)^2}\Bigl\{\bigl[28d_{13}^2 \\ &+ 4b\sqrt{7}\bigl(-3d_{12}^2 + 4d_{13}^2 + 3d_{23}^2\bigr) + b^2\bigl(-6d_{12}^2 + 7d_{13}^2 + 42d_{23}^2\bigr)\bigr]^{1/2} \\ &+ \bigl[28d_{13}^2 + 4b\sqrt{7}\bigl(3d_{12}^2 + 4d_{13}^2 - 3d_{23}^2\bigr) + b^2\bigl(42d_{12}^2 + 7d_{13}^2 - 6d_{23}^2\bigr)\bigr]^{1/2}\Bigr\}\end{aligned}$$

$$+\gamma\Delta^2\left(\frac{\sqrt{2}}{3b^3}\left(\frac{1}{d_{12}^5}+\frac{1}{d_{23}^5}\right)+\frac{8b^2}{3(2+b)d_{13}^5}\right.$$
$$+\frac{128b^2(8+\sqrt{7}b)^4}{3\sqrt{7}}\Big\{\big[28d_{13}^2+4b\sqrt{7}(-3d_{12}^2+4d_{13}^2+3d_{23}^2)$$
$$+b^2(-6d_{12}^2+7d_{13}^2+42d_{23}^2)\big]^{-5/2}$$
$$+\big[28d_{13}^2+4b\sqrt{7}(3d_{12}^2+4d_{13}^2-3d_{23}^2)$$
$$\left.+b^2(42d_{12}^2+7d_{13}^2-6d_{23}^2)\big]^{-5/2}\Big\}\right). \tag{4.17}$$

To illustrate the accuracy of the simplified energy, we present a couple examples, one with an equilateral control triangle and one where angle $\angle \mathbf{p}_1\mathbf{p}_2\mathbf{p}_3$ is obtuse. The first example has $d_{12}=d_{13}=d_{23}=1$, $\Delta=\sqrt{3}/2$. We obtain Table 1.

The simplified stretching and bending energies are within approximately 10% of the true energies for $0.6 < b < 4.0$, or $0.18 < b_{101} < 8.00$. This covers most of the splines occurring in actual practice.

In the second example, $d_{12}=d_{23}=1$, $d_{13}=\sqrt{3}$, $\Delta=\sqrt{3}/2$. Here we get Table 2.

Table 1

		True	Simplified	True	Simplified
b	b_{101}	Stretching energy		Bending energy	
0.6	0.180	1.120	1.126	5.000	5.419
0.8	0.320	1.166	1.172	3.941	3.984
1.0	0.500	1.209	1.215	3.628	3.646
1.5	1.125	1.304	1.316	3.841	3.792
2.0	2.000	1.380	1.408	4.500	4.297
3.0	4.500	1.494	1.580	6.158	5.680
4.0	8.000	1.573	1.749	7.958	7.323

Table 2

		True	Simplified	True	Simplified
b	b_{101}	Stretching energy		Bending energy	
0.8	0.320	1.770	1.775	1.043	1.613
1.0	0.500	1.780	1.785	0.804	1.014
1.5	1.125	1.804	1.813	0.619	0.635
2.0	2.000	1.824	1.846	0.611	0.613
3.0	4.500	1.854	1.931	0.725	0.727
4.0	8.000	1.876	2.041	0.890	0.879

In this example the simplified energies are accurate to within 20% for $1 \leqslant b \leqslant 4$, or $0.5 \leqslant b_{101} \leqslant 8$.

Case study

Using the same setup as in the example in Section 4.1.2, we find that we need to minimize this specialization of (4.17):

$$\begin{aligned}&\frac{b(4+3\sqrt{2})}{12\sqrt{5}}+\frac{2}{3(2+b)}\\&+\frac{32[(140+92\sqrt{7}b+365b^2)^{1/2}+(140+68\sqrt{7}b+317b^2)^{1/2}]}{3\sqrt{35}(8+\sqrt{7}b)^2}\\&+\frac{12}{25}\left\{\frac{25\sqrt{5}}{b^3}\left(\frac{1}{128}+\frac{\sqrt{2}}{243}\right)+\frac{8b^2}{2+b}\right.\\&+\frac{128b^2(8+\sqrt{7}b)^4}{\sqrt{35}}\left[\frac{1}{(140+92\sqrt{7}b+365b^2)^{5/2}}\right.\\&\left.\left.+\frac{1}{(140+68\sqrt{7}b+317b^2)^{5/2}}\right]\right\}.\end{aligned}\tag{4.18}$$

This function is minimized when $b = 0.856$, or $b_{101} = 0.366$, and the minimum value of the simplified energy is 5.66611. This gives an ellipse whose equation in Cartesian coordinates is

$$6.588x^2 - 1.098xy + 5.152y^2 - 6.588x + 3.294y = 0. \tag{4.19}$$

This elliptic arc is shown with in Fig. 4(b), and is quite close to the arc obtained by using the exact representation, as is evidenced by the superposition of the two curves in Fig. 4(c).

4.3. *Global minimization of simplified energy*

Here we consider the minimization of the simplified energy by varying the domain endpoint vertices of a chain of C^1 quadratic A-spline curves. With the chain we are minimizing a sum of expressions (4.17) instead of just a single one. Suppose we have $n+1$ junction points on a curve, and we wish to pass a spline through all these points. Label them $\mathbf{p}_1$, $\mathbf{p}_3 = \mathbf{p}'_1$, $\mathbf{p}'_3 = \mathbf{p}''_1, \ldots, \mathbf{p}_3^{(n-1)} = \mathbf{p}_1^{(n)}$, $\mathbf{p}_3^{(n)}$, where superscript ^{i} denotes i primes and $\mathbf{p}_3^{(n)} = \mathbf{p}_1$ if and only if the curve is to be a closed contour. The apex points $\mathbf{p}_2^{(i)}$ will be the intersections of the tangent lines through $\mathbf{p}_1^{(i)}$ and $\mathbf{p}_3^{(i)}$. Recognizing that the d_{jk} and Δ are functions of the coordinates of $\mathbf{p}_1$, $\mathbf{p}_2$, and $\mathbf{p}_3$, we can express the simplified energy in (4.17) as

$$E_{\text{simp}}^{(i)} = E_{\text{simp}}\big(b_i, \mathbf{p}_1^{(i)}, \mathbf{p}_2^{(i)}, \mathbf{p}_3^{(i)}; \beta, \gamma\big). \tag{4.20}$$

Thus our objective is to minimize

$$E_{\text{simp}} = \sum_{i=0}^{n-1} E_{\text{simp}}^{(i)}$$

for an open contour and

$$E_{\text{simp}} = \sum_{i=0}^{n} E_{\text{simp}}^{(i)}$$

for a closed one over all possible locations of the $\mathbf{p}_j^{(i)}$ and values of b_i.

Case study

As a simple example, suppose we have $(1,0)T$ and $(-1,0)T$ as two fixed points on the unit circle, and we wish to find the point $(x_0, y_0)T$ on the unit upper semicircle such that the simplified energy is minimized. This will require the sum of two components, the first of which has $\mathbf{p}_1 = (x_1, y_1)T = (1,0)T$ and $\mathbf{p}_3 = (x_0, y_0)T$, and the second with $\mathbf{p}_1' = (x_1', y_1')T = (x_0, y_0)T$ and $\mathbf{p}_3' = (-1,0)T$ (see Fig. 5). The points $\mathbf{p}_2 = (x_2, y_2)T$ and $\mathbf{p}_2' = (x_2', y_2')T$ are the intersections of the tangent lines to the circle through $\mathbf{p}_1$ and $\mathbf{p}_3$ and through $\mathbf{p}_1'$ and $\mathbf{p}_3'$, respectively; these will be $(1, (1-x_0)/y_0)T$ and $(-1, (1+x_0)/y_0)T$ in the two cases. Since all the points can be expressed in terms of x_0 via $y_0 = (1-x_0^2)^{1/2}$, this problem is a minimization over the three remaining unknowns x_0, b, and b', with the restrictions $|x_0| < 1$ and $b, b' > 0$.

We illustrate the results for $\beta = 1/12$, $\gamma = 1$ and for $\beta = 1$, $\gamma = 1/12$. In the first case, the global minimum occurs when $x_0 = 0$ and $b = b' = 1.501$ ($b_{101} = 1.126$), yielding a total energy of 3.414. For $\Delta\mathbf{p}_1\mathbf{p}_2\mathbf{p}_3$, we have $\alpha_1 = 1 - y$ and $\alpha_2 = x + y - 1$ (from (2.3)), and the Cartesian equation of this portion of the arc, that of an ellipse, is given by

$$x^2 - 0.252xy + y^2 + 0.252x + 0.252y - 1.252 = 0.$$

The second part of the spline, the piece within $\Delta\mathbf{p}_1\mathbf{p}_2\mathbf{p}_3$, is the reflection of the first piece across the y-axis. In the second case, the global minimum is at $x_0 = 0$ and $b = b' = 1.077$

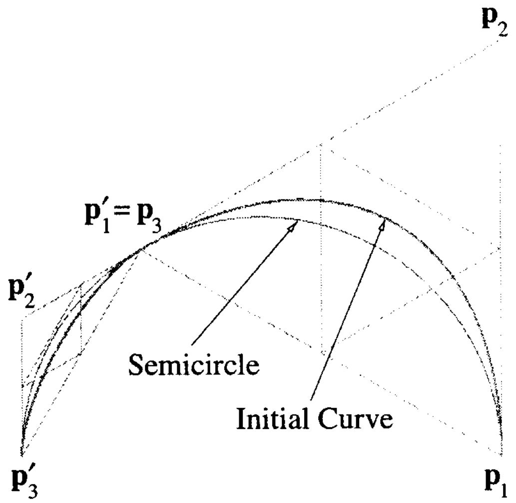
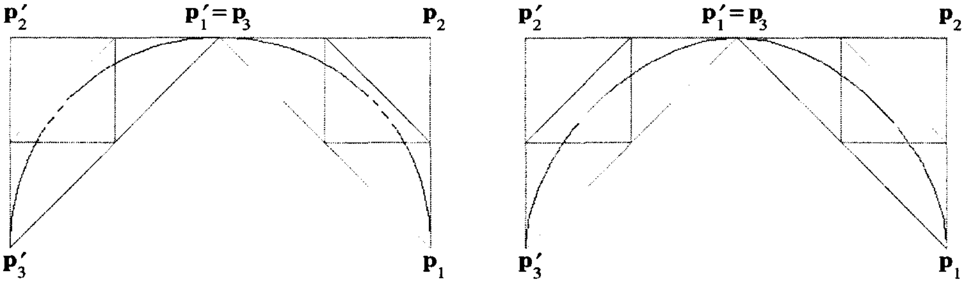


Fig. 5. Initial configuration. The point $\mathbf{p}_3 = \mathbf{p}_1'$ is allowed to slide along the semicircle. The spline labeled "initial curve" is a typical minimal-energy spline if $\mathbf{p}_3 = \mathbf{p}_1'$ is at the location indicated.



(a) (b)

Fig. 6. Optimal configurations for (a): $\beta = 1/12, \gamma = 1$; (b): $\beta = 1, \gamma = 1/12$.

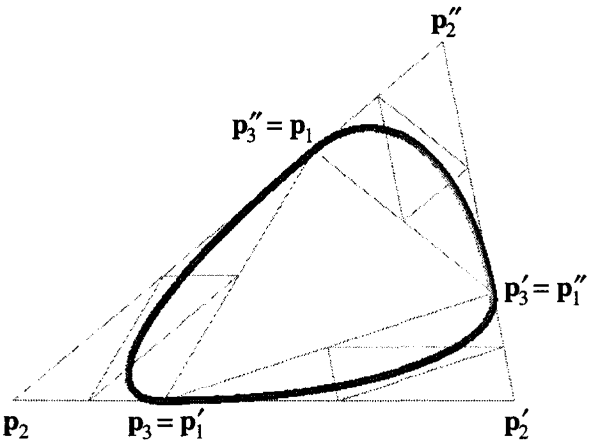


Fig. 7. Initial configuration. The points $\mathbf{p}_1$, $\mathbf{p}_1'$, and $\mathbf{p}_1''$ slide along the edges of $\triangle\mathbf{p}_2\mathbf{p}_2'\mathbf{p}_2''$ in such a way that $\overline{\mathbf{p}_2\mathbf{p}_1'}/\overline{\mathbf{p}_2\mathbf{p}_2'} = \overline{\mathbf{p}_2'\mathbf{p}_1''}/\overline{\mathbf{p}_2'\mathbf{p}_2''} = \overline{\mathbf{p}_2''\mathbf{p}_1}/\overline{\mathbf{p}_2''\mathbf{p}_2} = u$ for some u in [0, 1].

($b_{101} = 0.580$), yielding a total energy of 3.382. The Cartesian equation of the part of the arc in the first quadrant is

$$x^2 + 0.840xy + y^2 - 0.840x - 0.840y - 0.160 = 0.$$

These two cases are shown in Fig. 6. Note that as β increases with respect to γ, the stretching component of the total energy becomes more important than the bending component, and as a result the length of the simplified energy minimizing spline in 6(b) is less than that in 6(a).

Case study

For another example, we consider the problem of minimizing the total simplified energy of a closed contour with one point on each of the sides of the triangle $\triangle\mathbf{p}_2\mathbf{p}_2'\mathbf{p}_2''$, where $\mathbf{p}_2 = (0, 0)T$, $\mathbf{p}_2' = (7, 0)T$, and $\mathbf{p}_2'' = (6, 5)T$. The three points on the sides of $\triangle\mathbf{p}_2\mathbf{p}_2'\mathbf{p}_2''$ will be denoted by $\mathbf{p}_3 = \mathbf{p}_1'$, $\mathbf{p}_3' = \mathbf{p}_1''$, and $\mathbf{p}_3'' = \mathbf{p}_1$, as in Fig. 7. We will also impose the

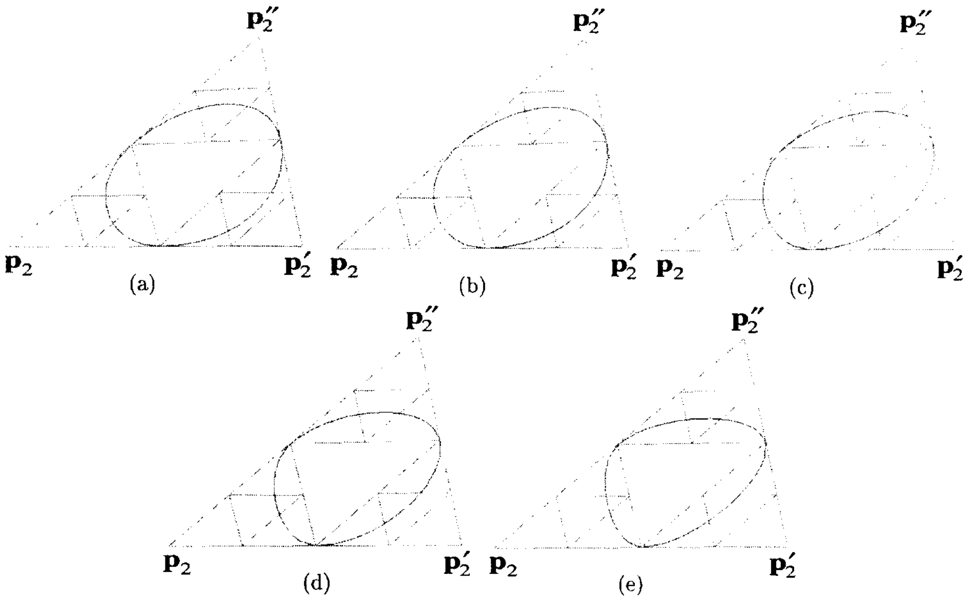


Fig. 8. Optimal configurations for (a): $\beta = \gamma/10$; (b): $\beta = \gamma/3$; (c): $\beta = \gamma$; (d): $\beta = 3\gamma$; (e): $\beta = 10\gamma$.

condition that each of $\mathbf{p}_1$, $\mathbf{p}_1'$, and $\mathbf{p}_1''$ is the same fraction along the way of their respective edges. That is,

$$\overline{\mathbf{p}_2\mathbf{p}_1'}/\overline{\mathbf{p}_2\mathbf{p}_2'} = \overline{\mathbf{p}_2'\mathbf{p}_1''}/\overline{\mathbf{p}_2'\mathbf{p}_2''} = \overline{\mathbf{p}_2''\mathbf{p}_1}/\overline{\mathbf{p}_2''\mathbf{p}_2} = u$$

for some u in $[0, 1]$. With this condition the points on the edges of $\Delta\mathbf{p}_2\mathbf{p}_2'\mathbf{p}_2''$ are found to be $\mathbf{p}_1 = (6 - 6u, 5 - 5u)T$, $\mathbf{p}_1' = (7u, 0)T$, and $\mathbf{p}_1'' = (7 - u, 5u)T$. With all points expressed in terms of u, minimizing the total simplified energy amounts to minimizing (4.20) over the four unknowns u, b, b', and b'', where $0 < u < 1$ and the $b^{(i)} > 0$.

We illustrate the results for various values of β/γ:

β/γ	1/10	1/3	1	3	10
u	0.5111	0.5109	0.5092	0.5057	0.5030
b_{101}	0.318	0.264	0.189	0.123	0.074
b'_{101}	0.916	0.789	0.604	0.409	0.243
b''_{101}	0.627	0.564	0.456	0.322	0.196

Fig. 8 shows the resulting closed contour for these five cases. Here we have the same phenomenon which occurred in the previous example: as β/γ increases, the stretching energy component becomes more important as compared to the bending component, and the length of the closed contour for the simplified energy minimizing spline decreases. Also, for each value of β/γ we have $b_{101} < b'_{101} < b''_{101}$, so that the ordering of these coefficients is the same as the order of the size of the angle of the apexes of their corresponding triangles ($\angle \mathbf{p}_1^{(i)} \mathbf{p}_2^{(i)} \mathbf{p}_3^{(i)}$).

5. Conclusion

Several elastic models using A-splines have been proposed, each of which has its own advantages and shortcomings. Besides the traditional energy model adapted from theory of elasticity, we give several different simplified models that take advantage of the A-spline formulation and also yield efficient computation of the minimum energy solution. A subsequent paper will report the use of these energy splines in image processing applications.

One problem of great interest is the efficient computation of the minimum energy for the exact model presented in Section 4.1.1. If such a problem can be restricted to having just one control weight being free, then the energy-minimizing problem reduces to a nonlinear univariate equation. In the much more common situation where this is not possible, one possibility is to generate a piecewise linear approximation of the A-spline such that the denominator in the expression (3.3), $(\nabla S^{\mathrm{T}} \mathbf{J}\mathbf{J}^{\mathrm{T}} \nabla S)^3$, over each piece is nearly constant. While the resulting system is sparse and may be solved iteratively, generally a large number of pieces will be required. For example, in the case study in Section 4.1.2, the denominator ranges from 0.254 when $\alpha_1 = 0$ down to 0.110 at $\alpha_1 = 0.535$ and back up to 0.719 when $\alpha_1 = 1$. Thus if we wanted the denominator to vary by at most 1 percent over each subinterval and were able to divide the interval [0, 1] of the α_1-axis at precisely the right points, we would need 133 subintervals. If we were willing to relax the condition to a 5 percent variance, we would still need 27 subintervals. Furthermore, finding the values of α_1 at which the break points should be located is in itself a significant problem. In practice these locations will not be known beforehand, and one may have to make a conservative subdivision of the interval $\alpha_1 \in [0, 1]$ to ensure the desired accuracy.

References

Bajaj, C.L. and Xu, G. (1992), A-splines: Local interpolation and approximation using G^k continuous piecewise real algebraic curves, Purdue University Technical Report CAPO 92-44, to appear in Computer Aided Geometric Design.

Bajaj, C.L. and Xu, G. (1996), Data fitting with cubic A-Splines, in: Gigante, M. and Kunii, T., eds., *Proceedings of Computer Graphics International, CGI'94*, Melbourne, Australia, World Scientific Publishing, 329–340.

Bajaj, C.L., Chen, J. and Xu, A. (1995), Modeling with Cubic A-patches, ACM Transaction on Graphics 14, 2, 103–133.

Birkhoff, G. and De Boor, C. (1965), Piecewise polynomial interpolation and approximation, in: Garabedian, H., ed., *Approximation of Functions,* Elsevier.

Brotman, L.S. and Netravali, A.N. (1988), Motion interpolation by optimal control, SIGGRAPH'88, Computer Graphic 22, 309–315.

Bruckstein, A.M. and Netravali, A.N. (1990), On minimal energy trajectories, Computer Vision, Graphics, and Image Processing 49, 283–296.

Brunnett, G. (1992), A new characterization of plane elastica, in: Lyche, T. and Schumaker, L.L., eds., *Mathematical Methods in Computer Aided Geometric Design II*, Academic Press, Boston, 43–56.

Bryant, R. and Griffiths, P. (1986), Reduction for constrained variational problems and $\int \frac{\kappa^2}{2}\, ds$, American Journal of Math. 108, 525–570.

Cohen, L.D. and Cohen, I. (1993), Finite-element methods for active contour modes and balloons for 2-d and 3-d images, PAMI 15, 1131–1147.

Farouki, R. (1996), The elastic bending energy of pythagorean-hodograph curves, Computer Aided Geometric Design 13, 227–241.

Golomb, M. and Jerome, J. (1982), Equilibria of the curvature functional and manifolds of non-interpolating spline curves, SIAM Journal of Math. Anal. 13, 421–458.

Horn, B.K.P. (1983), The curve of least energy, ACM Transactions on Mathematical Software 9, 259–268.

Jou, E. and Han, W. (1990a), Elastica and minimal-energy splines, in: Laurent, P., Le Méhauté, A. and Schumaker, L., eds., *Curves and Surfaces*, Academic Press, International Conference on Curves and Surfaces, Chamonix-Mont-Blanc, France, 1–4.

Jou, E. and Han, W. (1990b), Minimal-energy splines: I. Plane curves with angle constraints, Mathematical Methods in the Applied Sciences 13, 351–370.

Kallay, M. (1986), Plane curves of minimal energy, ACM Transactions on Mathematical Software 12, 219–222.

Kass, M., Witkin, A. and Terzopoulos, D. (1988), Snakes: Active contour models, International Journal of Computer Vision, 321–331.

Landau, L. and Lifshitz, E. (1959), *Theory of Elasticity*, Pergammon Press, London, U.K.

Lee, E. and Forsythe, G. (1975), Variational study of nonlinear spline curve, SIAM Review 15, 120–133.

Malcom, M. (1977), On the computation of nonlinear spline functions, SIAM Journal of Numerical Analysis 14, 254–282.

Mehlum, E. (1974), Nonlinear splines, in: Barnhill, R. and Riesenfeld, R., eds., *Computer Aided Geometric Design*, Academic Press, New York, 173–207.

Mumford, D. (1994), Elastica and computer vision, in: Bajaj C., ed., *Algebraic Geometry and its Applications*, Springer, Berlin, 491–506.

Ronfard, R. (1994), Region-based strategies for active contour models, International Journal of Computer Vision 13, 229–251.

Terzopoulos, D., Platt, J., Barr, A. and Fleischer, K. (1987), Elasticity deformable models, Computer Graphics (SIGGRAPH '87) 21, 205–214.

Williams, D.J. and Shah, M. (1992), A fast algorithm for active contours and curvature estimation, CVGIP:IU 55, 14–26.