A CONSTRUCTIVE FRAMEWORK FOR MINIMAL ENERGY PLANAR CURVES

Michael J. Johnson[1] & Hakim S. Johnson[2]

Dept. Mathematics, Faculty of Science, Kuwait University[1−corr. author]
Kuwait English School[2]

phone: +965 24985329, email: yohnson1963@hotmail.com

Abstract. Given points $P_1, P_2, \dots, P_n$ in the plane, we are concerned with the problem of finding a fair curve which interpolates the points. We assume that we have a method in hand, called a basic curve method, for solving the geometric Hermite interpolation problem of fitting a regular C^∞ curve between two points with prescribed tangent directions at the endpoints. We also assume that we have an energy functional which defines the energy of any basic curve. Using this basic curve method repeatedly, we can then construct G^1 curves which interpolate the given points $P_1, P_2, \dots, P_n$. The tangent directions at the interpolation points are variable and the idea is to choose them so that the energy of the resulting curve (i.e., the sum of the energies of its pieces) is minimal. We give sufficient conditions on the basic curve method, the energy functional, and the interpolation points for (a) existence, (b) G^2 regularity, and (c) uniqueness of minimal energy interpolating curves. We also identify a one-parameter family of basic curve methods, based on parametric cubics, whose minimal energy interpolating curves are unique and G^2 under suitable conditions. One member of this family looks very promising and we suggest its use in place of conventional C^2 parametric cubic splines.

1991 *Mathematics Subject Classification.* 65D05; 65D17, 65D07.

Key words and phrases. geometric interpolation, Hermite interpolation, parametric cubic spline, non-linear spline.

This work was supported and funded by Kuwait University, Research Project No. SM 01/14

Typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$-TEX

1. Introduction

Let $P_1, P_2, \ldots, P_n$ be a sequence of points in the complex plane $\mathbb{C}$ satisfying $P_j \neq P_{j+1}$, and consider the problem of finding a 'fair' curve which passes sequentially (i.e., interpolates) the points. Whereas there do not exist interpolating curves with minimal bending energy (see [3] and also [10]), except when the points lie sequentially along a line, it was shown recently [2] that they do exist if one imposes the additional constraint that each piece of the interpolating curve be an s-curve (a curve which turns monotonically at most $180°$ in one direction and then turns monotonically at most $180°$ in the opposite direction). Such interpolating curves with minimal bending energy are called *elastic splines*. While work on [2] was in progress, the authors of the present article took up the numerical challenge of computing elastic splines. As in [7], the problem was formulated as an optimization problem where the interpolation points $P_1, P_2, \ldots, P_n$ are given but corresponding tangent directions $d_1, d_2, \ldots, d_n$ are variable. An important sub-problem, which was extensively addressed in [2], is that of finding an s-curve with minimal bending energy which solves the first order geometric Hermite interpolation problem of constructing a curve which begins at P_j with direction d_j and ends at P_{j+1} with direction d_{j+1}. The s-curve condition places feasibility restrictions on the directions d_j and d_{j+1}. In case $P_j = 0$ and P_{j+1} lies on the positive real axis (which can be obtained by a translation and rotation) and writing $d_j = e^{i\alpha}$ and $d_{j+1} = e^{i\beta}$, these feasibility restrictions reduce to the inequalities $|\alpha|, |\beta| < \pi$ and $|\alpha - \beta| \leq \pi$. The presence of these coupled restrictions on the directions $\{d_j\}$ gives rise to a rather complicated feasible region in $\mathbb{C}^n$ and this in turn complicates the optimization algorithm (see [9] for an alternative feasible region). After completing the numerics, it was observed that the elastic splines were often fair, but could also be unsightly, particularly when the interpolation points force abrupt changes in direction (see Fig. 1a).

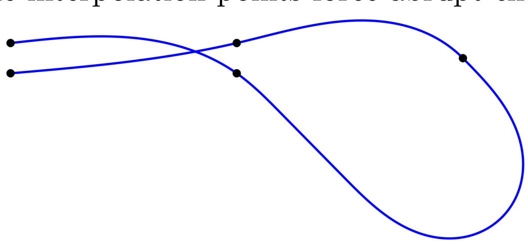

Fig. 1a

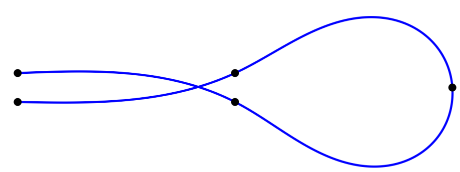

Fig. 1b

After much experimentation, we decided that the best way to eliminate the unsightly elastic splines is to replace the s-curve feasibility inequalities $|\alpha|, |\beta| < \pi$ and $|\alpha - \beta| \leq \pi$ with the simple restriction $|\alpha|, |\beta| \leq \pi/2$ (see Fig. 1b). With these simple uncoupled restrictions on $\{d_j\}$, the feasible region in $\mathbb{C}^n$ reduces to a Cartesian product and the optimization algorithm simplifies to that of optimizing each direction individually while cycling through the points. In addition to eliminating the unsightly elastic splines, the simplified restriction $|\alpha|, |\beta| \leq \pi/2$ also makes the theoretic study of elastic splines much more tractable. Moreover, the basic theory can be developed in a general context where potentially any method for solving the above-mentioned first order geometric Hermite interpolation problem can be used in place of that for elastic splines.

We now describe the basic setup. A **unit tangent vector** $u = (P, d)$ is an ordered pair of complex numbers with $|d| = 1$ and can be visualized as a directed line segment with **base-point** P and **direction** d. A C^∞ regular **curve** is a C^∞ function $f : [a, b] \to \mathbb{C}$ whose first derivative f' is non-vanishing. We say that f **connects** $u_1 = (P_1, d_1)$ to

$u_2 = (P_2, d_2)$ if $f(a) = P_1$, $f'(a) = |f'(a)|d_1$, $f(b) = P_2$ and $f'(b) = |f'(b)|d_2$. We use the term basic curve method to refer to a method for solving the first order geometric Hermite interpolation problem mentioned above. Precisely, a **basic curve method** is a mapping $(\alpha, \beta, L) \mapsto c_L(\alpha, \beta)$, which is defined for angles $\alpha, \beta \in [-\Omega, \Omega]$ and lengths $L > 0$ ($\Omega \in (0, \pi)$ is a given constant), whose image $c_L(\alpha, \beta)$ is a C^∞ regular curve which connects $u = (0, e^{i\alpha})$ to $v = (L, e^{i\beta})$ (see Fig. 2a). Associated with the basic curve method is a functional E_L, whereby the 'energy' of the curve $c_L(\alpha, \beta)$ equals $E_L(\alpha, \beta)$. In practice, the energy of $c_L(\alpha, \beta)$ is often its bending energy, defined by $\frac{1}{2}\int_a^b [\kappa(s)]^2 \frac{ds}{dt}\, dt$ with κ denoting signed curvature and s arclength, or an approximation of bending energy, such as $\frac{1}{2}\int_a^b |f''(t)|^2\, dt$, but there is no theoretical requirement that energy have a physical interpretation. It could just as well be the *cosmic energy* of the curve.

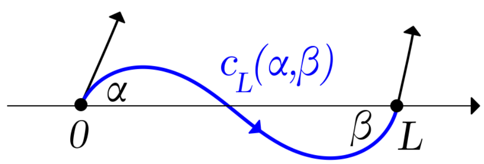

Fig. 2a

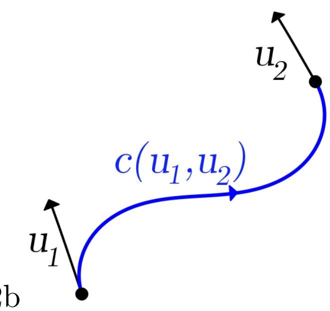

Fig. 2b

The basic curve method is extended to other pairs of unit tangent vectors by the use of translation and rotation (see Fig. 2b). Specifically, let $u_1 = (P_1, d_1)$ and $u_2 = (P_2, d_2)$ be two unit tangent vectors with distinct base points, and set $\alpha = \arg \frac{d_1}{P_2 - P_1}$, $\beta = \arg \frac{d_2}{P_2 - P_1}$ and $L = |P_2 - P_1|$. Here arg is defined, as usual, by $\arg re^{i\theta} = \theta$ when $r > 0$ and $\theta \in (-\pi, \pi]$. If the angles α and β belong to $[-\Omega, \Omega]$, then the basic curve connecting u_1 to u_2 is defined by $c(u_1, u_2) := T \circ c_L(\alpha, \beta)$, where the transformation $T(z) = a_1 z + a_2$ is determined by the requirements $T(0) = P_1$ and $T(L) = P_2$ (i.e., $a_1 = (P_2 - P_1)/L$ and $a_2 = P_1$). The energy of $c(u_1, u_2)$ is defined by $\text{Energy}(c(u_1, u_2)) := E_L(\alpha, \beta)$. As a consequence of these definitions, the extended basic curve method and its energy functional are invariant under translations and rotations.

Examples of basic curve methods pertaining to parametric cubics are given in [12] and [7]; we'll have more to say about these in Section 4. A basic curve method employing A-splines is given in [1], and we mention that second order basic curve methods (where curvature data is also interpolated) can be found in [5] and [11] and the references therein.

With a basic curve method in hand, one can construct G^1 curves which interpolate the points $P_1, P_2, \ldots, P_n$ (following [7]) by assigning 'feasible' directions $d_1, d_2, \ldots, d_n$, and then use the resultant unit tangent vectors $u_j := (P_j, d_j)$ to obtain an interpolating curve $c(u_1, u_2) \sqcup c(u_2, u_3) \sqcup \cdots \sqcup c(u_{n-1}, u_n)$, called an **admissible curve**, whose energy is defined to be the sum of energies of its constituent pieces. Here, the directions $d_1, d_2, \ldots, d_n$ are deemed **feasible** if all of the basic curves $c(u_j, u_{j+1})$ are defined. The goal is then to choose the feasible directions $\{d_j\}$ so that the energy of the corresponding admissible curve is minimized.

Our primary purpose is to prove sufficient conditions (and occasionally necessary conditions) for (a) existence, (b) G^2 regularity, and (c) uniqueness of minimal energy admissible curves. These sufficient conditions depend on the constant Ω, which limits the size of the chord angles α, β, and also on the size of the *stencil angles* $\{\psi_i\}$, where $\psi_i := \arg \frac{P_{i+1}-P_i}{P_i-P_{i-1}}$ is the angle between $P_{i+1} - P_i$ and $P_i - P_{i-1}$.

It is interesting to compare our setup with that of Brunnett and Keifer [4]. They introduce a basic curve method whose basic curves approximate pieces of rectangular elastica (specifically, the signed curvature $\kappa(s)$ of a basic curve is a polynomial function of arclength s). Once feasible tangent directions at each node are specified, an admisssible curve is constructed and then any jump discontinuities in the signed curvature across interpolation nodes are tallied. The objective function is defined as the sum of squares of these jumps in signed curvature and then their goal is to choose the tangent directions to minimize the objective function. Their objective function does not fit into our framework because it is not additive–one cannot express the objective function as the sum of energies of its constituent basic curves (basic curves have continuous signed curvature so presumably they have 0 energy). A more significant distinction, however, is that they attempt to directly minimize G^2 irregularity by defining the objective function in terms of the jump discontinuities in signed curvature. In contrast, in our approach we hope to obtain G^2 regularity as a by-product of minimized energy.

An outline of the remainder of the paper is as follows. Existence is quickly settled in Section 2 by showing that a minimal energy admissible curve exists if and only if the stencil angles satisfy $|\psi_i| \leq 2\Omega$ (note that if $\Omega = \pi/2$ then this condition is always satisfied). In Section 3, we develop a preliminary notion called 'conditional G^2 regularity' and then in Section 4 we present a family, parameterized by $\lambda \in (0, 3)$, of cubic basic curve methods which have conditional G^2 regularity. The choice $\lambda = 1$ is distinguished in that this basic curve method best emulates Elastic Splines as $(\alpha, \beta) \to (0, 0)$ (see Remark 4.5 for details and also Fig. 7 at the end of Section 7). We believe that this method, defined below, is a significant find, and we recommend it as a substitute for conventional C^2 parametric cubic splines (see [6] and [8]).

Example 1.1. For $L > 0$ and $|\alpha|, |\beta| \leq \frac{\pi}{2}$, the basic curve $c_L(\alpha, \beta)$ for the *Quasi-Elastic Cubic* is given parametrically by
$x(t) = L[t(1-t)^2 \cos\alpha + t^2(3 - 2t + (t-1)\cos\beta)]$, $y(t) = L[t(1-t)^2 \sin\alpha + t^2(t-1)\sin\beta]$, $0 \leq t \leq 1$, with energy $E_L(\alpha, \beta) = \frac{1}{L}[5 + \cos(\alpha - \beta) - 3(\cos\alpha + \cos\beta)]$.

In Section 3 we also give necessary conditions for conditional G^2 regularity (Corollary 3.6) and use these in Section 4 to argue that the basic curve methods of [12] and [7] do not have this property. In Section 5 we obtain sufficient conditions for G^2-regularity; when applied to Example 1.1, we find that if $\Omega = \pi/2$ and the stencil angles satisfy $|\psi_i| \leq \tan^{-1}\sqrt{8} \approx 70.5°$ for all i, then minimal energy admissible curves have G^2 regularity. Sufficient conditions for uniqueness are obtained in Section 6. When applied to Example 1.1, we find that if $\Omega = \cos^{-1}\frac{2}{3} \approx 48.2°$ and $|\psi_i| \leq 2\Omega$ for all i, then the minimal energy admissible curve is unique; moreover, if $\Omega = \cos^{-1}\frac{2}{3}$ and $|\psi_i| \leq \Omega - \tan^{-1}\frac{\sqrt{5}}{7} \approx 30.5°$, then the unique minimal energy admissible curve has G^2 regularity. Lastly, in Section 7 we describe a free open source `C++` program, called *Curve Ensemble*, which enables the user to construct minimal energy admissible curves using a variety of basic curve methods.

2. Existence of Minimal Energy Interpolating Curves

Given points $P_1, P_2, \dots, P_n$ in $\mathbb{C}$, with $P_i \neq P_{i+1}$, a G^1 interpolating curve is deemed **admissible** if each piece is a basic curve produced by the basic curve method. The **energy** of an admissible curve is defined to be the sum of the energies of its constituent basic curves. Whereas the interpolation nodes $P_1, P_2, \dots, P_n$ are given, the associated directions $d_1, d_2, \dots, d_n$ are variable and the goal is to choose these directions so that the energy of the resulting admissible curve is minimal (see Fig. 3).

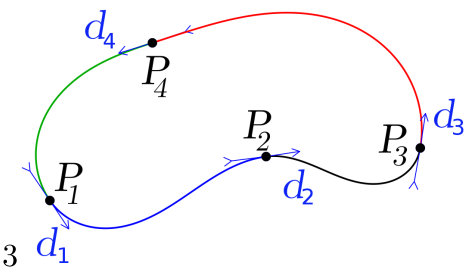

Fig. 3

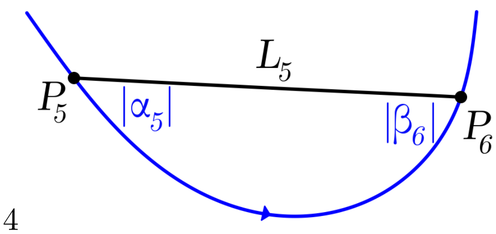

Fig. 4

Remark 2.1. As is customary, there are five different cases regarding conditions at the endpoints of the interpolating curve:
(1) The curve is *free at both* P_1 *and* P_n; i.e., both d_1 and d_n are variable (see Fig. 1b).
(2) The curve is *clamped at* P_1 *and free at* P_n; i.e., d_1 is prescribed, while d_n is variable.
(3) The curve is *free at* P_1 *and clamped at* P_n; i.e., d_1 is variable, while d_n is prescribed.
(4) The curve is *clamped at both* P_1 *and* P_n; i.e., both d_1 and d_n are prescribed.
(5) The curve is *periodic* (see Fig. 3).
We will explicitly address the last case (periodic curve), where we also assume that $P_n \neq P_1$ and, for notational convenience, we extend the interpolation points and directions periodically by the rules $P_{k+n} = P_k$ and $d_{k+n} = d_k$, $k \in \mathbb{Z}$. Results for the first four cases will be explained by remarks following theorems.

The polygon obtained when the interpolation nodes are connected by directed line segments $[P_i, P_{i+1}]$ is called the **stencil**. In order to address the matter of existence, we define the **stencil angles** $\{\psi_i\}$ as the angular change in direction from one segment to the next (see Fig. 5); precisely,

$$\psi_i := \arg \frac{P_{i+1} - P_i}{P_i - P_{i-1}}.$$

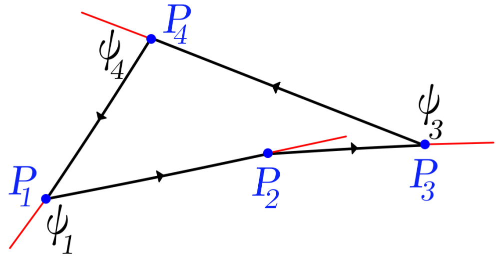

Fig. 5

Recall that each direction d_i is represented as a complex unit (i.e., $d_i \in \mathbb{C}$ with $|d_i| = 1$) and thus $u_i := (P_i, d_i)$ is a unit tangent vector. A list of directions $d_1, d_2, \dots, d_n$ is called **feasible** if the basic curve $c(u_i, u_{i+1})$ is defined for all i. We note that any curve connecting u_i to u_{i+1} has chord length $L_i := |P_{i+1} - P_i|$

and chord angles (see Fig. 4)

$$\alpha_i := \arg \frac{d_i}{P_{i+1} - P_i}, \qquad \beta_{i+1} := \arg \frac{d_{i+1}}{P_{i+1} - P_i}, \tag{2.1}$$

and therefore the list $d_1, d_2, \ldots, d_n$ is feasible if and only if $|\alpha_i|, |\beta_{i+1}| \leq \Omega$ for all i. The **feasible range** of direction d_i, denoted $\mathcal{D}_i$, is the set of all directions d_i such that $|\alpha_i|, |\beta_i| \leq \Omega$ and can be written explicitly as the intersection $\mathcal{D}_i = \mathcal{D}_i^\beta \cap \mathcal{D}_i^\alpha$, where $\mathcal{D}_i^\beta$ is the closed arc of arclength 2Ω and midpoint $\frac{P_i - P_{i-1}}{L_{i-1}}$ and $\mathcal{D}_i^\alpha$ is the same except with midpoint $\frac{P_{i+1} - P_i}{L_i}$. Being the intersection of two compact subsets of $\mathbb{C}$, $\mathcal{D}_i$ is compact, but possibly empty. We remind the reader that $0 < \Omega < \pi$. If $\Omega \geq \frac{\pi}{2}$, then $\mathcal{D}_i$ is obviously non-empty; while if $\Omega < \frac{\pi}{2}$, then $\mathcal{D}_i$ is non-empty if and only if the (shortest) arc-distance between the two midpoints is less or equal to 2Ω. This latter condition is easily seen to be equivalent to $|\psi_i| \leq 2\Omega$ and we thus conclude the following.

Lemma 2.2. *The feasible range $\mathcal{D}_i$ is non-empty if and only if the stencil angle ψ_i satisfies $|\psi_i| \leq 2\Omega$.*

Assumption 1. For all $L > 0$, the energy functional E_L is continuous on $[-\Omega, \Omega]^2$.

Existence Theorem 2.3. *Under Assumption 1, the following are equivalent.*
(i) There exists a feasible list of directions $d_1, d_2, \ldots, d_n$.
(ii) There exists a feasible list of directions $d_1, d_2, \ldots, d_n$ such that the energy of the corresponding admissible curve is minimal.
(iii) The stencil angles satisfy $|\psi_i| \leq 2\Omega$ for $i = 1, 2, \ldots, n$.

Proof. With Lemma 2.2 in view, it is clear that (i) and (iii) are equivalent since there exists a feasible list of directions if and only if all the feasible ranges are non-empty. And since (ii) implies (i), it suffices to show that (iii) implies (ii). So assume (iii). Note that the list $d_1, d_2, \ldots, d_n$ is feasible if and only if $(d_1, d_2, \ldots, d_n) \in \mathcal{D} := \mathcal{D}_1 \times \mathcal{D}_2 \times \cdots \times \mathcal{D}_n$ and in this case the energy of the corresponding admissible curve is given by

$$\text{Energy}(d_1, d_2, \ldots, d_n) = \sum_{i=1}^{n} E_{L_i}(\alpha_i, \beta_{i+1}), \tag{2.2}$$

which defines Energy as a function from $\mathcal{D}$ into $\mathbb{R}$. Under Assumption 1, and since α_i and β_i depend continuously on d_i, it follows that Energy is a continuous function and we obtain (ii) since $\mathcal{D}$ is compact. □

Remark 2.4. Although written specifically for case (5) of Remark 2.1, Existence Theorem 2.3 remains valid for cases (1)–(4) with the following modifications:
(a) Set $\psi_1 = \psi_n = 0$.
(b) If the curve is clamped at P_1, then assume additionally that the prescribed direction d_1 yields $|\alpha_1| \leq \Omega$.
(c) If the curve is clamped at P_n, then assume additionally that the prescribed direction d_n yields $|\beta_n| \leq \Omega$.

3. Conditional G^2 regularity

It is often the case that an optimal element has better qualities than those commonly held by the set of candidates. For example, with cubic splines one minimizes $\int_a^b [s''(x)]^2\,dx$ over all functions s which interpolate the given data and are globally C^1 and piecewise C^2. Whereas the candidates are only piecewise C^2, the optimal solution (the cubic spline), turns out to be globally C^2. In the present context, the candidates are the admissible curves and these have G^1 regularity by construction. One naturally hopes that minimal energy admissible curves would possess G^2 regularity, but experience has shown that one should not expect this improvement at nodes P_i where the restriction $|\alpha_i|, |\beta_i| \leq \Omega$ is active (i.e., where $|\alpha_i|$ or $|\beta_i|$ equals Ω). With this caveat, we make the following definition.

Definition 3.1. A minimal energy admissible curve has **conditional G^2 regularity** if it is G^2 across node P_i whenever the direction d_i lies in the interior of its feasible range $\mathcal{D}_i$.

Let us assume for the remainder of this section that condition (iii) of Existence Theorem 2.3 holds, and it follows that $\mathcal{D}_i$ (the feasible range of direction d_i) is non-empty for all i. It follows from the observation $\mathcal{D}_i = \mathcal{D}_i^\beta \cap \mathcal{D}_i^\alpha$, made above Lemma 2.2, that D_i is either a singleton, a closed arc, a doubleton, or the union of two disjoint closed arcs, where the latter two are only possible when $\Omega \geq \frac{\pi}{2}$. Let us assume that $d_1, d_2, \ldots, d_n$ is a feasible list of directions for which the energy of the corresponding admissible curve is minimal. Although this curve has G^1 regularity by construction, G^2 regularity may fail at interpolation nodes since the signed curvature may have jump discontinuities across the nodes. Our sufficient conditions for conditional G^2 regularity assume the following, which is stronger than Assumption 1.

Assumption 2. For all $L > 0$, the energy functional E_L is continuous on $[-\Omega, \Omega]^2$ and the partial derivatives $\frac{\partial}{\partial\alpha} E_L(\alpha, \beta)$ and $\frac{\partial}{\partial\beta} E_L(\alpha, \beta)$ exist for all $(\alpha, \beta) \in [-\Omega, \Omega]^2$ (with the understanding that when (α, β) lies on the boundary of $[-\Omega, \Omega]^2$, these partial derivatives are one-sided, as appropriate).

Definition 3.2. Let $\kappa_a(c(u, v))$ and $\kappa_b(c(u, v))$ denote the initial and terminal signed curvature of the basic curve $c(u, v)$, respectively. We say that the energy functional $E_L(\alpha, \beta)$ is **consistent with end curvatures** if there exists an odd function $\rho : \mathbb{R} \to \mathbb{R}$ such that $-\kappa_a(c_L(\alpha, \beta)) = \rho(\frac{\partial}{\partial\alpha} E_L(\alpha, \beta))$ and $\kappa_b(c_L(\alpha, \beta)) = \rho(\frac{\partial}{\partial\beta} E_L(\alpha, \beta))$ for all $L > 0$ and $(\alpha, \beta) \in [-\Omega, \Omega]^2$.

Theorem 3.3. *Under Assumption 2, if the energy functional $E_L(\alpha, \beta)$ is consistent with end curvatures, then minimal energy admissible curves have conditional G^2 regularity.*

Proof. Let $d_1, d_2, \ldots, d_n$ be a feasible list of directions for which the energy of the corresponding admissible curve is minimal, and let the chord lengths $\{L_i\}$ and chord angles $\{\alpha_i\}$, $\{\beta_{i+1}\}$ be as defined in (2.1). Fix i and assume that d_i belongs to the interior of $\mathcal{D}_i$. We will show that the curve is G^2 at node P_i. By applying a translation and rotation if necessary we can assume, without loss of generality, that $P_{i-1} = -L_{i-1}$, $P_i = 0$, and $P_{i+1} = L_i e^{i\psi_i}$. Note that $d_i = e^{i\beta_i}$. Since d_i lies in the interior of its feasible range, there exists $\varepsilon > 0$ such that the list $d_1, d_2, \ldots, d_n$ remains feasible when d_i is replaced by $d_i e^{i\delta}$,

$|\delta| < \varepsilon$. With this replacement, the energy of the resultant curve can be written as

$$\text{Energy}(\delta) := C + E_{L_{i-1}}(\alpha_{i-1}, \beta_i + \delta) + E_{L_i}(\alpha_i + \delta, \beta_{i+1}),$$

where C denotes the sum of energies of the other $n-2$ basic curves which are unaffected by δ. It follows from Assumption 2 that Energy(δ) is differentiable, and since it is minimized at $\delta = 0$, it follows that $\frac{\partial}{\partial\delta}\text{Energy}(\delta)$ vanishes at $\delta = 0$. Hence,

$$\frac{\partial}{\partial\beta} E_{L_{i-1}}(\alpha_{i-1}, \beta_i) + \frac{\partial}{\partial\alpha} E_{L_i}(\alpha_i, \beta_{i+1}) = 0.$$

Since $E_L(\alpha, \beta)$ is consistent with end curvatures (and ρ is odd), it follows that

$$\kappa_b(c_{L_{i-1}}(\alpha_{i-1}, \beta_i)) = \rho\left(\frac{\partial}{\partial\beta} E_{L_{i-1}}(\alpha_{i-1}, \beta_i)\right) = -\rho\left(\frac{\partial}{\partial\alpha} E_{L_i}(\alpha_i, \beta_{i+1})\right) = \kappa_a(c_{L_i}(\alpha_i, \beta_{i+1})),$$

and therefore the minimal energy admissible curve is G^2 at node P_i. □

Remark. Although we are specifically addressing case (5) of Remark 2.1, Theorem 3.3 also holds in cases (1)–(4). Note that in these cases, admissible curves are always G^2 at the endpoints P_1 and P_n.

Theorem 3.3 can be made more specific if one makes precise assumptions on how the basic curve $c_L(\alpha, \beta)$ and its energy $E_L(\alpha, \beta)$ depend on the parameter L. The following definitions are tailored to the case when energy equals or emulates bending energy.

Definition 3.4. The basic curve method $c_L(\alpha, \beta)$ is **scale invariant** if $c_L(\alpha, \beta) = Lc_1(\alpha, \beta)$ for all $L > 0$ and $(\alpha, \beta) \in [-\Omega, \Omega]^2$. The energy functional $E_L(\alpha, \beta)$ is **inversely proportional to scale** if $E_L(\alpha, \beta) = \frac{1}{L} E_1(\alpha, \beta)$ for all $L > 0$ and $(\alpha, \beta) \in [-\Omega, \Omega]^2$.

Theorem 3.5. *Assume that $c_L(\alpha, \beta)$ is scale invariant and $E_L(\alpha, \beta)$ is inversely proportional to scale. Then, under Assumption 2, $E_L(\alpha, \beta)$ is consistent with end curvatures if and only if there exists $\mu \in \mathbb{R}$ such that for all $(\alpha, \beta) \in [-\Omega, \Omega]^2$,*

$$-\kappa_a(c_1(\alpha, \beta)) = \mu \frac{\partial}{\partial\alpha} E_1(\alpha, \beta) \quad \text{and} \quad \kappa_b(c_1(\alpha, \beta)) = \mu \frac{\partial}{\partial\beta} E_1(\alpha, \beta). \tag{3.1}$$

Proof. Since $c_L(\alpha, \beta)$ is scale invariant, we have $\kappa_a(c_L(\alpha, \beta)) = \frac{1}{L}\kappa_a(c_1(\alpha, \beta))$ and $\kappa_b(c_L(\alpha, \beta)) = \frac{1}{L}\kappa_b(c_1(\alpha, \beta))$. It is now clear that if (3.1) holds, then $E_L(\alpha, \beta)$ satisfies Definition 3.2 with $\rho(t) = \mu t$. For the converse, assume that $E_L(\alpha, \beta)$ is consistent with end curvatures. If $\kappa_a(c_1(\alpha, \beta)) = \kappa_b(c_1(\alpha, \beta)) = 0$ for all (α, β), then (3.1) holds with $\mu = 0$; so assume there exists (α_0, β_0) such that at least one of $\kappa_a(c_1(\alpha_0, \beta_0))$, $\kappa_b(c_1(\alpha_0, \beta_0))$ is non-zero. We will address only the case when $\rho_0 := \kappa_b(c_1(\alpha_0, \beta_0)) \neq 0$, since the latter case is similar. Set $t_0 := \frac{\partial}{\partial\beta} E_1(\alpha_0, \beta_0)$ and note that $\rho_0 = \rho(t_0)$ since $E_L(\alpha, \beta)$ is consistent with end curvatures. Since $\rho_0 \neq 0$ and ρ is odd, it must be the case that $t_0 \neq 0$. Now, for $L > 0$, we have

$$\rho\left(\frac{1}{L} t_0\right) = \rho\left(\frac{\partial}{\partial\beta} E_L(\alpha_0, \beta_0)\right) = \kappa_b(c_L(\alpha_0, \beta_0)) = \frac{1}{L}\kappa_b(c_1(\alpha_0, \beta_0)) = \frac{1}{L}\rho(t_0),$$

and it now follows (since ρ is odd) that $\rho(t) = \mu t$, $t \in \mathbb{R}$, for some $\mu \in \mathbb{R}$; hence (3.1). $\square$

Remark 3.6. Theorem 3.5 demonstrates that the assumptions that the basic curve method is scale invariant and the energy functional is inversely proportional to scale have greatly simplified the system of equations in Definition 3.2. Firstly, we only have $L = 1$ in Theorem 3.5, so the variable L has been eliminated. And secondly, if one strengthens Assumption 2 to assuming that partial derivatives are continuous, then the energy functional has been eliminated from the equations because condition (3.1) is equivalent to the vector field $\mathbf{F}(\alpha, \beta) := -\kappa_a(c_1(\alpha, \beta))\mathbf{i} + \kappa_b(c_1(\alpha, \beta))\mathbf{j}$ being conservative, a condition involving only the basic curve method. Consequently, if one has a scale invariant basic curve method in hand and is seeking an energy functional which is both inversely proportional to scale and consistent with end curvatures, then it suffices to check whether $\mathbf{F}$ is conservative. If it is, then $E_1(\alpha, \beta)$ should be chosen so that $\mu E_1(\alpha, \beta)$ is a potential function for $\mathbf{F}$; if it is not, then no such energy functional exists.

The following Corollary is a consequence of the above remark and the well-known necessary condition for a vector field to be conservative.

Corollary 3.7. *Let $c_L(\alpha, \beta)$ be a basic curve method which is scale invariant, but suppose that an energy functional has yet to be defined. Assume that $\kappa_a(c_1(\alpha, \beta))$, $\kappa_b(c_1(\alpha, \beta))$, $\frac{\partial}{\partial \beta}\kappa_a(c_1(\alpha, \beta))$ and $\frac{\partial}{\partial \alpha}\kappa_b(c_1(\alpha, \beta))$ are continuous for $(\alpha, \beta) \in (-\Omega, \Omega)^2$. If there exists an energy functional $E_L(\alpha, \beta)$ such that the following hold:*
(i) $E_L(\alpha, \beta)$ is inversely proportional to scale,
(ii) $E_L(\alpha, \beta)$ satisfies Assumption 2, and
(iii) $E_L(\alpha, \beta)$ is consistent with end curvatures,
then

$$Q(\alpha, \beta) := \frac{\partial}{\partial \beta}\kappa_a(c_1(\alpha, \beta)) + \frac{\partial}{\partial \alpha}\kappa_b(c_1(\alpha, \beta)) = 0, \quad \text{for all } (\alpha, \beta) \in (-\Omega, \Omega)^2. \tag{3.2}$$

Proof. If (i),(ii) and (iii) hold, then (3.1) holds and we have

$$\frac{\partial}{\partial \beta}\kappa_a(c_1(\alpha, \beta)) = -\mu \frac{\partial^2}{\partial \beta \partial \alpha} E_1(\alpha, \beta) = -\mu \frac{\partial^2}{\partial \alpha \partial \beta} E_1(\alpha, \beta) = -\frac{\partial}{\partial \alpha}\kappa_b(c_1(\alpha, \beta)).$$

$\square$

4. Parametric Cubic Basic Curve Methods

In this section, we seek parametric cubic scale invariant basic curve methods and energy functionals which satisfy conditions (i), (ii) and (iii) of Corollary 3.7.

Following [12], we first write $c_1(\alpha, \beta)$ as the parametric cubic curve

$$\begin{aligned} x(t) &= t^2(3 - 2t) + s_0(t^3 - 2t^2 + t)\cos\alpha + s_1(t^3 - t^2)\cos\beta, \\ y(t) &= s_0(t^3 - 2t^2 + t)\sin\alpha + s_1(t^3 - t^2)\sin\beta, \quad 0 \le t \le 1, \end{aligned} \tag{4.1}$$

where $s_0 = s_0(\alpha,\beta) > 0$ and $s_1 = s_1(\alpha,\beta) > 0$ denote, respectively, the initial and terminal speeds of the parametric curve. The signed curvature at time t is given by $\kappa = (x'(t)y''(t) - y'(t)x''(t))/[(x'(t))^2 + (y'(t))^2]^{3/2}$ from which we arrive at

$$\kappa_a(c_1(\alpha,\beta)) = \frac{2s_1\sin(\alpha-\beta) - 6\sin\alpha}{s_0^2}, \quad \kappa_b(c_1(\alpha,\beta)) = \frac{2s_0\sin(\alpha-\beta) + 6\sin\beta}{s_1^2}.$$

Example 4.1. Yong and Cheng [12] propose choosing the speeds s_0 and s_1 in order to minimize $\int_0^1 (x''(t)^2 + y''(t)^2)\,dt$. They have shown that if $3\cos\alpha > \cos(\alpha - 2\beta)$ and $3\cos\beta > \cos(\beta - 2\alpha)$, then the choice $s_0(\alpha,\beta) := \frac{6\cos\alpha - 3\cos\beta\cos(\alpha-\beta)}{4-\cos^2(\alpha-\beta)}$, $s_1(\alpha,\beta) := s_0(\beta,\alpha)$ achieves this minimum, and furthermore, if $\cos\alpha, \cos\beta > 1/3$, then the resulting curve $c_1(\alpha,\beta)$ is regular. Scott Kersey (private communication) has shown that all the above inequalities hold if $|\alpha|, |\beta| \le \pi/3$ and therefore suggests using $\Omega = \pi/3$. It is easy to verify numerically that $Q(\alpha, 0) < 0$ whenever $0 < \alpha \le \pi/3$ and we therefore conclude that there does not exist an energy functional $E_L(\alpha,\beta)$ satisfying conditions (i), (ii) and (iii) of Corollary 3.7. With no better alternative in sight, we define $E_L(\alpha,\beta) := \frac{1}{L}\int_0^1 (x''(t)^2 + y''(t)^2)\,dt$.

Example 4.2. Instead of minimizing $\int_0^1 (x''(t)^2 + y''(t)^2)\,dt$, Jaklič and Žagar [7] propose that s_0 and s_1 be chosen to minimize the functional $(1 - s_0\cos\alpha)^2 + (s_0\sin\alpha)^2 + (s_1\cos\beta - 1)^2 + (s_1\sin\beta)^2$ (an approximation to the bending energy of $c_1(\alpha,\beta)$). They have shown that if $|\alpha|, |\beta| < \pi/2$, then the choice $s_0 = \cos\alpha$, $s_1 = \cos\beta$ achieves the minimum value $\sin^2\alpha + \sin^2\beta$, and furthermore, the resulting curve $c_1(\alpha,\beta)$ is regular, loop-, cusp- and fold-free. Again, it is easy to verify numerically that $Q(\alpha, 0) < 0$ whenever $0 < \alpha < \pi/2$ and we therefore conclude that there does not exist an energy functional $E_L(\alpha,\beta)$ satisfying conditions (i), (ii) and (iii) of Corollary 3.7. With no better alternative in sight, we define $E_L(\alpha,\beta) := \frac{1}{L}(\sin^2\alpha + \sin^2\beta)$.

The only choice of $s_0(\alpha,\beta)$ and $s_1(\alpha,\beta)$, known to the authors, which yields (3.2) is $s_0(\alpha,\beta) := s_1(\alpha,\beta) := \lambda$, where $\lambda > 0$ is a constant. It follows (see Remark 3.6) that the vector field $\mathbf{F}(\alpha,\beta) = (-\frac{2}{\lambda}\sin(\alpha-\beta) + \frac{6}{\lambda^2}\sin\alpha)\mathbf{i} + (\frac{2}{\lambda}\sin(\alpha-\beta) + \frac{6}{\lambda^2}\sin\alpha)\mathbf{j}$ is conservative, and one easily verifies that $p(\alpha,\beta) := \frac{2}{\lambda}\cos(\alpha-\beta) - \frac{6}{\lambda^2}(\cos\alpha + \cos\beta)$ is a potential function for $\mathbf{F}$. In order that $\mu E_1(\alpha,\beta)$ also be a potential function, we must choose E_1 in the form $E_1(\alpha,\beta) = ap(\alpha,\beta) + b$ for some constants $a \ne 0$ and b. It so happens that $\frac{1}{2}\int_0^1 (x''(t)^2 + y''(t)^2)\,dt = \lambda^3 p(\alpha,\beta) + 4\lambda^2 + 6$, which motivates the choice $a = \lambda^3$ and $b = 4\lambda^2 + 6$. We therefore define

$$E_1(\alpha,\beta) := \frac{1}{2}\int_0^1 (x''(t)^2 + y''(t)^2)\,dt = 2\lambda^2\cos(\alpha-\beta) - 6\lambda(\cos\alpha + \cos\beta) + 4\lambda^2 + 6$$

and note that (3.1) holds with $\mu = \lambda^{-3}$. By Theorem 3.5, $E_L(\alpha,\beta)$ is consistent with end curvatures.

Lemma 4.3. *Let $c_1(\alpha,\beta)$ be as defined in (4.1), with $s_0 := s_1 := \lambda$, and assume $0 < \lambda < 3$ and $|\alpha|, |\beta| \le \frac{\pi}{2}$. Then $x'(t) > 0$ for all $0 < t < 1$.*

Proof. Note that $q := x'(0) = \lambda\cos\alpha \ge 0$, $r := x'(1) = \lambda\cos\beta \ge 0$ and $x'(t) = At^2 + Bt + q$, where $A = -6 + 3q + 3r$, $B = 6 - 4q - 2r$. Since x' is a parabola with $x'(0), x'(1) \ge 0$

and $\int_0^1 x'(t)\,dt = 1$, it is easy to see that the lemma's conclusion holds in all cases except possibly in the case when $x''(0) < 0$ and $x''(1) > 0$; that is, when $B < 0$ and $2A + B > 0$. In this exceptional case, the minimum of x' occurs at $t^* = -A/(2B)$, where $x'(t^*) = r - B^2/(4A)$. So, in order to complete the proof, it suffices to show that if $2A > -B > 0$, then $Ar - B^2/4 > 0$; that is, if $2q+4r > 6$ and $4q+2r > 6$, then $6q+6r-q^2-qr-r^2-9 > 0$. In order to pursue this, we define the set $F := \{(q,r) \in [0,3]^2 : 2q + 4r \geq 6,\ 4q + 2r \geq 6\}$ and the function $f(q,r) := 6q + 6r - q^2 - qr - r^2 - 9$, $(q,r) \in F$. It is easy to verify that F is the convex hull of the four points $(1,1)$, $(3,0)$, $(0,3)$, and $(3,3)$. Since f is strictly concave and vanishes at the four corners of F, it follows that f is positive at all other points of F. Now, suppose $2q + 4r > 6$ and $4q + 2r > 6$. Then (q,r) belongs to F and it is easy to verify that (q,r) cannot equal any of the four corners of F; therefore, $6q + 6r - q^2 - qr - r^2 - 9 = f(q,r) > 0$, which completes the proof. □

Theorem 4.4. *Let $s_0 := s_1 := \lambda \in (0,3)$, and let the basic curve $c_L(\alpha,\beta)$ be defined, for $|\alpha|, |\beta| \leq \frac{\pi}{2}$ and $L > 0$, by $c_L(\alpha,\beta) := Lc_1(\alpha,\beta)$ (scale invariant), where $c_1(\alpha,\beta)$ is defined parametrically by (4.1). Define $E_L(\alpha,\beta) := \frac{1}{L}E_1(\alpha,\beta)$ (inversely proportional to scale), where E_1 is as defined above Lemma 4.3. Then $c_L(\alpha,\beta)$ is a regular C^∞ curve and conditions (i), (ii) and (iii) of Corollary 3.7 hold.*

Proof. Conditions (i) and (ii) of Corollary 3.7 clearly hold, while condition (iii) is proved just above Lemma 4.3. It is also clear that $c_L(\alpha,\beta)$ is a C^∞ curve. Since the initial and terminal speeds of $c_L(\alpha,\beta)$ are both $\lambda L > 0$ it follows from Lemma 4.3 that $c_L(\alpha,\beta)$ is regular. □

Remark 4.5. We refer to the basic curve method described in Theorem 4.4 as the *Conditionally G^2 Cubic* with shape parameter $\lambda \in (0,3)$. As mentioned in the introduction, we claim that the choice $\lambda = 1$ (the Quasi-Elastic Cubic of Example 1.1) is distinguished by its affinity with Elastic Splines as $(\alpha,\beta) \to (0,0)$. To see this, let $E_1(\alpha,\beta)$ denote the energy functional associated with Conditionally G^2 Cubics and let $\mathcal{E}_1(\alpha,\beta)$ be the bending energy of Elastic Splines. It is known that $\mathcal{E}_1(0,0) = 0$ with gradient $\nabla\mathcal{E}_1(0,0) = (0,0)$ (this follows from [2, Prop. 7.6] and its proof). Although we do not yet have a complete proof, we have observed numerically that the Hessian of $\mathcal{E}_1$ at $(0,0)$ is $\mathcal{H}_1(0,0) = 2\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$. Turning now to E_1, we see that $E_1(0,0) = 6(\lambda-1)^2$ and $\nabla E_1(0,0) = (0,0)$. That $E_1(0,0) \neq 0$ when $\lambda \neq 1$ is a minor discrepancy because one can always subtract a constant from the energy functional without losing the property of being conditionally G^2. The Hessian of E_1 is

$$H_1(\alpha,\beta) = 2\lambda \begin{bmatrix} -\lambda\cos(\alpha-\beta) + 3\cos\alpha & \lambda\cos(\alpha-\beta) \\ \lambda\cos(\alpha-\beta) & -\lambda\cos(\alpha-\beta) + 3\cos\beta \end{bmatrix},$$

so we have $H_1(0,0) = 2\lambda \begin{bmatrix} 3-\lambda & \lambda \\ \lambda & 3-\lambda \end{bmatrix}$. Note that if $\lambda = 1$, then $H_1(0,0) = \mathcal{H}_1(0,0)$ while if $\lambda \neq 1$, then $H_1(0,0)$ is not even parallel to $\mathcal{H}_1(0,0)$ (i.e., $\mathcal{H}_1(0,0) \neq \nu H_1(0,0)$ for all $\nu \in \mathbb{R}$).

5. G^2 regularity

Let $c_L(\alpha, \beta)$ be a basic curve method, defined for $|\alpha|, |\beta| \leq \Omega$, and assume that Assumption 2 holds and that the energy functional $E_L(\alpha, \beta)$ is consistent with end curvatures. We know from previous sections that if the stencil angles satisfy $|\psi_i| \leq 2\Omega$, then there exists an admissible curve with minimal energy and all such curves are conditionally G^2. Being conditionally G^2, the curve will be G^2 at points P_i where the direction d_i lies in the interior of its feasible range $\mathcal{D}_i$. It is natural then to seek sufficient conditions which will ensure that d_i lies in the interior of $\mathcal{D}_i$, and in this section we pursue this in terms of the stencil angle ψ_i. Specifically, we seek an angle $\Psi > 0$ such that d_i lies in the interior of $\mathcal{D}_i$ whenever $|\psi_i| \leq \Psi$.

Theorem 5.1. *Suppose there exists $\Psi \in (0, \Omega)$ such that:*
(i) For all $\alpha \in [-\Omega, \Omega]$, there exists β^*_α, with $|\beta^*_\alpha| \leq \Omega - \Psi$, such that
$sign(\frac{\partial}{\partial\beta} E_L(\alpha, \beta)) = sign(\beta - \beta^*_\alpha)$, for all $\beta \in [-\Omega, \Omega]$, and
(ii) For all $\beta \in [-\Omega, \Omega]$, there exists α^*_β, with $|\alpha^*_\beta| \leq \Omega - \Psi$, such that
$sign(\frac{\partial}{\partial\alpha} E_L(\alpha, \beta)) = sign(\alpha - \alpha^*_\beta)$, for all $\alpha \in [-\Omega, \Omega]$.
Then minimal energy admissible curves have G^2-regularity at all points P_i where $|\psi_i| \leq \Psi$.

Proof. Let P_i be a point where the stencil angle satisfies $|\psi_i| \leq \Psi$. We will show that d_i lies in the interior of $\mathcal{D}_i$. Without loss of generality, we can assume that $i = 2$, $P_1 = -L_1$, $P_2 = 0$ and $P_3 = L_2 e^{i\psi_2}$. Note that $\beta_2 = \arg d_2$. We will address only the case $\psi_2 \geq 0$, since the other case $\psi_2 < 0$ is similar. This case is shown in Fig. 6, where the label (α) indicates that the labeled direction is $e^{i\alpha}$.

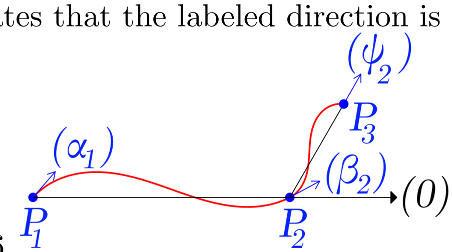

Fig. 6

Since $\pi > \Omega > \Psi \geq \psi_2 \geq 0$, it is easy to verify that $\arg \mathcal{D}_2 = [\psi_2 - \Omega, \Omega]$ and d_2 lies in the interior of $\mathcal{D}_2$ if and only if $\psi_2 - \Omega < \beta_2 < \Omega$. Since our admissible curve has minimal energy, and noting that $\alpha_2 = \beta_2 - \psi_2$, it follows that

$$f(\beta_2) \leq f(\beta) \text{ for all } \beta \in [\psi_2 - \Omega, \Omega], \tag{5.1}$$

where $f(\beta) := E_{L_1}(\alpha_1, \beta) + E_{L_2}(\beta - \psi_2, \beta_3)$. It follows from (i) that $\frac{\partial}{\partial\beta} E_{L_1}(\alpha, \beta) > 0$ at $(\alpha, \beta) = (\alpha_1, \Omega)$ and from (ii) that $\frac{\partial}{\partial\alpha} E_{L_2}(\alpha, \beta) \geq 0$ at $(\alpha, \beta) = (\Omega - \psi_2, \beta_3)$; hence $f'(\Omega) > 0$. We therefore have $\beta_2 < \Omega$, since equality would contradict (5.1). In a similar manner, we see that $\frac{\partial}{\partial\beta} E_{L_1}(\alpha, \beta) \leq 0$ at $(\alpha, \beta) = (\alpha_1, \psi_2 - \Omega)$ and $\frac{\partial}{\partial\alpha} E_{L_2}(\alpha, \beta) < 0$ at $(\alpha, \beta) = (-\Omega, \beta_3)$; hence $f'(\psi_2 - \Omega) < 0$. We therefore have $\beta_2 > \psi_2 - \Omega$, since again equality would contradict (5.1). Having established $\psi_2 - \Omega < \beta_2 < \Omega$, it follows that d_i lies in the interior of $\mathcal{D}_i$. $\square$

Remark. Although we are specifically addressing case (5) of Remark 2.1, Theorem 5.1 also holds in cases (1)–(4). We again note that in these cases, admissible curves are always G^2 at the endpoints P_1 and P_n.

Example 5.2. Consider the Conditionally G^2 Cubic basic curve method described in Theorem 4.4 with $0 < \lambda < 3$ and $\Omega = \pi/2$. We will show that conditions (i) and (ii) of Theorem 5.1 hold when $\Psi = \tan^{-1} \frac{\sqrt{9-\lambda^2}}{\lambda}$.
Since the energy functional is inversely proportional to scale and $E_1(\alpha, \beta) = 2\lambda^2 \cos(\alpha - \beta) - 6\lambda(\cos\alpha + \cos\beta) + 4\lambda^2 + 6 = E_1(\beta, \alpha)$, it suffices to establish condition (i) when $L = 1$. Differentiating yields

$$\frac{\partial}{\partial \beta} E_1(\alpha, \beta) = 2\lambda^2 \sin(\alpha - \beta) + 6\lambda \sin\beta = 2\lambda^2[\sin\alpha \cos\beta - \sin\beta(\cos\alpha - 3/\lambda)].$$

Note that $\frac{\partial}{\partial\beta} E_1(\alpha, \pi/2) = 2\lambda^2(\frac{3}{\lambda} - \cos\alpha) > 0$ and $\frac{\partial}{\partial\beta} E_1(\alpha, -\pi/2) = -2\lambda^2(\frac{3}{\lambda} - \cos\alpha) < 0$, since $0 < \lambda < 3$. Moreover, $\frac{\partial}{\partial\beta} E_1(\alpha, \beta) = 0$ if and only if $\sin\alpha\cos\beta = \sin\beta(\cos\alpha - 3/\lambda)$; that is, if and only if $\beta = \beta^*_\alpha := -\tan^{-1} \frac{\sin\alpha}{3/\lambda - \cos\alpha}$. Therefore $\operatorname{sign}(\frac{\partial}{\partial\beta} E_1(\alpha, \beta)) = \operatorname{sign}(\beta - \beta^*_\alpha)$ for all $|\alpha|, |\beta| \le \pi/2$. In order to obtain condition (i), let Ψ be defined by $\frac{\pi}{2} - \Psi = \max_\alpha |\beta^*_\alpha|$. As an exercise in Differential Calculus, we leave it to the reader to verify that the odd function $f(\alpha) := \frac{\sin\alpha}{3/\lambda - \cos\alpha}$, $|\alpha| \le \pi/2$, is uniquely maximized at $\alpha = \cos^{-1} \frac{\lambda}{3}$, where $f(\cos^{-1} \frac{\lambda}{3}) = \frac{\lambda}{\sqrt{9-\lambda^2}}$. Therefore, $\Psi = \frac{\pi}{2} - \tan^{-1} \frac{\lambda}{\sqrt{9-\lambda^2}} = \tan^{-1} \frac{\sqrt{9-\lambda^2}}{\lambda}$.

Remark. For the Quasi-Elastic Cubic (i.e., $\lambda = 1$) discussed in Remark 4.5, the angle Ψ is $\Psi = \tan^{-1} \sqrt{8} \approx 70.5°$.

6. Uniqueness

Let $c_L(\alpha, \beta)$ be a basic curve method, defined for $L > 0$ and $|\alpha|, |\beta| \le \Omega$, with energy functional $E_L(\alpha, \beta)$, and assume $\Omega < \frac{\pi}{2}$.

Assumption 3. For all $L > 0$, the energy functional E_L is continuous and strictly convex on $[-\Omega, \Omega]^2$.

Theorem 6.1. *Under Assumption 3, if the stencil angles satisfy $|\psi_i| \le 2\Omega$ for all i, then there exists a unique minimal energy admissible curve.*

Proof. Assume that the stencil angles satisfy $|\psi_i| \le 2\Omega$ for all i, and let the chord lengths $\{L_i\}$ be as defined in Section 3. Define $f : [-\Omega, \Omega]^{2n} \to \mathbb{R}$ by

$$f(a_1, b_2, a_2, b_3, \ldots, a_n, b_{n+1}) := \sum_{i=1}^{n} E_{L_i}(a_i, b_{i+1}),$$

and note that it follows from Assumption 3 that f is strictly convex. We recall from Section 3 that the set of admissible curves is parametrized by $\mathcal{D} := \mathcal{D}_1 \times \mathcal{D}_2 \times \cdots \times \mathcal{D}_n$ via the correspondence between a feasible list $(d_1, d_2, \cdots, d_n) \in \mathcal{D}$ and the corresponding admissible curve whose energy (see (2.2)) equals $f(\alpha_1, \beta_2, \alpha_2, \beta_3, \ldots, \alpha_n, \beta_{n+1})$, where the chord angles $\{\alpha_i\}$ and $\{\beta_{i+1}\}$ are defined (periodically) in (2.1). The assumption $0 < \Omega < \frac{\pi}{2}$ ensures that each feasible range $\mathcal{D}_i$ is connected (a closed arc or a singleton) and also allows us to write $\alpha_{i+1} = \beta_i - \psi_{i+1}$ for all i. Now, let $\mathcal{L}$ denote the set of all lists $(a_1, b_1, a_2, b_2, \ldots, a_n, b_n)$ in $[-\Omega, \Omega]^{2n}$ such that $a_{i+1} = b_i - \psi_{i+1}$ for $i = 1, 2, \ldots, n$, where $a_{n+1} := a_1$. It is easy to verify that $\mathcal{L}$ is a compact, convex subset of $[-\Omega, \Omega]^{2n}$. Note that if $(d_1, d_2, \ldots, d_n)$ belongs to $\mathcal{D}$, then the associated list of chord angles $(\alpha_1, \beta_1, \alpha_2, \beta_2, \ldots, \alpha_n, \beta_n)$ belongs to $\mathcal{L}$; in particular $\mathcal{L}$ is nonempty. On the other hand, if $(a_1, b_1, a_2, b_2, \ldots, a_n, b_n) \in \mathcal{L}$, then there exists a unique list $(d_1, d_2, \ldots, d_n) \in \mathcal{D}$, given by $d_i = \arg(P_{i+1} - P_i)e^{ia_i}$, whose associated list of chord angles equals $(a_1, b_1, a_2, b_2, \ldots, a_n, b_n)$. We have thus described a one-to-one correspondence between $\mathcal{D}$ and $\mathcal{L}$ and we further note that the energy of the admissible curve determined by a list $(d_1, d_2, \ldots, d_n) \in \mathcal{D}$ equals $f(\alpha_1, \beta_1, \alpha_2, \beta_2, \ldots, \alpha_n, \beta_n)$. Consequently, in order prove the theorem, it suffices to show there exists a unique list in $\mathcal{L}$ which minimizes the restriction of f to $\mathcal{L}$. But this is a simple consequence of the fact that f is continuous and strictly convex on the compact, convex set $\mathcal{L}$. $\square$

Remark. Although written specifically for case (5) of Remark 2.1, Theorem 6.1 remains valid for cases (1)–(4) with the same modifications (a),(b),(c) described in Remark 2.4.

We will now determine to what extent Theorem 6.1 applies to the Conditionally G^2 Cubic basic curve method defined in Theorem 4.4. Recall from Remark 4.5 that the Hessian of $E_1(\alpha, \beta)$ is given by $H_1(\alpha, \beta) = 2\lambda \begin{bmatrix} -\lambda\cos(\alpha-\beta) + 3\cos\alpha & \lambda\cos(\alpha-\beta) \\ \lambda\cos(\alpha-\beta) & -\lambda\cos(\alpha-\beta) + 3\cos\beta \end{bmatrix}$ and in particular $H_1(0,0) = 2\lambda \begin{bmatrix} 3-\lambda & 1 \\ 1 & 3-\lambda \end{bmatrix}$. The eigenvalues of $H_1(0,0)$ are $\mu_1 = 6 - 4\lambda$ and $\mu_2 = 6$, and therefore $H_1(0,0)$ is positive definite if and only if $0 < \lambda < \frac{3}{2}$. Assuming $0 < \lambda < \frac{3}{2}$, it follows that there exists $\Omega_{conv} > 0$ such that E_1 is strictly convex on $[-\Omega_{conv}, \Omega_{conv}]^2$. Fortunately, the largest value of Ω_{conv} can be determined.

Proposition 6.2. *Let $E_1(\alpha, \beta)$ be the energy functional associated with the Conditionally G^2 Cubic basic curve method described in Theorem 4.4, and assume $0 < \lambda < \frac{3}{2}$. Set $\Omega_{conv} := \cos^{-1}\frac{2\lambda}{3}$. Then E_1 is strictly convex on $[-\Omega_{conv}, \Omega_{conv}]^2$.*

Proof. We have seen above that the Hessian $H_1(0,0)$ is positive definite. In order to complete the proof, it suffices to show that the determinant of $H_1(\alpha, \beta)$ is positive for all non-corner points (α, β) of the square $[-\Omega_{conv}, \Omega_{conv}]^2$.
Define $f(q,r) := qr - \lambda(q+r)$ for $(q,r) \in [2\lambda, 3]^2$. Since $\frac{\partial}{\partial r} f(q,r) = q - \lambda > 0$, it follows that $f(q,r) > f(q, 2\lambda)$ whenever $r > 2\lambda$. Now, $f(q, 2\lambda) = \lambda q - 2\lambda^2 > 0$ whenever $q > 2\lambda$. It therefore follows that $f(q,r) > 0$ for all $(q,r) \in [2\lambda, 3]^2 \backslash \{(2\lambda, 2\lambda)\}$. Now let (α, β) be a non-corner point of the square $[-\Omega_{conv}, \Omega_{conv}]^2$ and set $q = 3\cos\alpha$ and $r = 3\cos\beta$. Note that $(q,r) \in [2\lambda, 3]^2 \backslash \{(2\lambda, 2\lambda)\}$. One easily verifies that $\det H_1(\alpha, \beta) = 4\lambda^2[9\cos\alpha\cos\beta - 3\lambda\cos(\alpha-\beta)(\cos\alpha + \cos\beta)]$, and therefore $\det H_1(\alpha, \beta) \geq 4\lambda^2[9\cos\alpha\cos\beta - 3\lambda(\cos\alpha + \cos\beta)] = 4\lambda^2 f(q,r) > 0$. $\square$

Remark. For the Quasi-Elastic Cubic (i.e., $\lambda = 1$) discussed in Remark 4.5, the critical angle is $\Omega_{conv} = \cos^{-1}\frac{2}{3} \approx 48.2°$.

We playfully refer to choices of Ω and Ψ which ensure existence, uniqueness and G^2 regularity of the minimal energy admissible curve as the *Green Zone.* For the Conditionally G^2 Cubic basic curve method, we have the following:

Proposition 6.3. *Consider the Conditionally G^2 Cubic basic curve method described in Theorem 4.4 with $0 < \lambda < \frac{3}{2}$, and set $\Omega = \cos^{-1}\frac{2\lambda}{3}$ and $\Psi = \Omega - \tan^{-1}\frac{\sqrt{9-4\lambda^2}}{9-2\lambda}$. If the stencil angles satisfy $|\psi_i| \le \Psi$, then the unique minimal energy admissible curve has G^2 regularity.*

Proof. Assume $|\psi_i| \le \Psi$ for all i. Since $\Psi \le 2\Omega$, it follows from Proposition 6.2 and Theorem 6.1 that there exists a unique minimal energy admissible curve. In order to show that this optimal curve has G^2 regularity, it suffices to show that conditions (i) and (ii) of Theorem 5.1 hold. For this, as explained in Example 5.2, it suffices to establish (i) for the case $L = 1$. Recall from Example 5.2 that $\mathrm{sign}(\frac{\partial}{\partial\beta}E_1(\alpha,\beta)) = \mathrm{sign}(\beta - \beta_\alpha)$ holds for $|\beta| \le \frac{\pi}{2}$, where $\beta_\alpha = -\tan^{-1} f(\alpha)$ and $f(\alpha) := \frac{\sin\alpha}{3/\lambda - \cos\alpha}$. We also recall that the odd function f is increasing on the interval $[-\cos^{-1}\frac{\lambda}{3}, \cos^{-1}\frac{\lambda}{3}]$. Since $\Omega = \cos^{-1}\frac{2\lambda}{3}$ belongs to this interval, it follows that $|\beta_\alpha| \le \tan^{-1} f(\Omega)$ for all $\alpha \in [-\Omega, \Omega]$, and we therefore obtain condition (i) of Theorem 5.1 with $\Psi = \Omega - \tan^{-1} f(\Omega)$. $\square$

7. Curve Ensemble

Here, we briefly mention a free open source `C++` computer program, called *Curve Ensemble*, that we are developing for the purpose of testing and comparing a variety of basic curve methods. *Curve Ensemble* began as a program written specifically for elastic splines in tandem with [2], but it was gradually realized that a great deal of the code (eg. optimization, plotting, inputting/editing nodes, painting, saving and exporting) could be used regardless of the particular basic curve method employed. At present, Curve Ensemble 'brings to life' six basic curve methods and is available at

`http://sourceforge.net/projects/curve-ensemble` .

As an illustration, Fig. 7 shows five different parametric cubic splines interpolating the same points. The first three are conventional C^2 parametric cubic splines with parameter intervals chosen as (a) uniform, (b) centripetal, and (c) chordal (see [6]). The last two are minimal energy curves where the basic curve method is (d) Jaklič-Žagar Cubic (Example 4.2) with $\Omega = 80°$ and (e) Quasi-Elastic Cubic (Example 1.1) with $\Omega = 90°$.

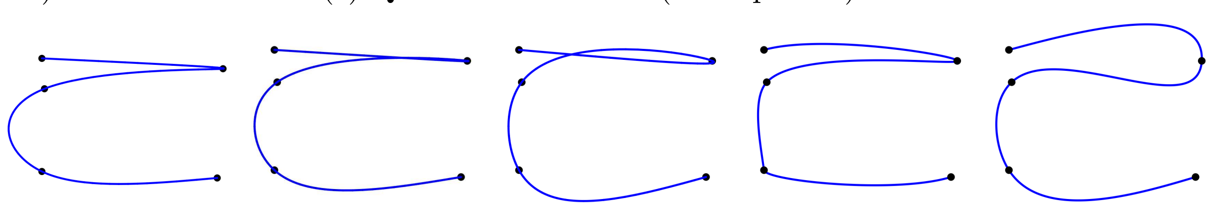

Fig. 7 (a) (b) (c) (d) (e)

Acknowledgements. We are happy to thank Scott Kersey for contributing, at an early stage, the basic curve method (from [12]) described in Example 4.1; this motivated our subsequent search for a conditionally G^2 parametric cubic method. We also heartily thank Albert Borbély for numerous discussions pertaining to sufficient conditions for G^2 regularity and uniqueness.

References

1. C.L. Bajaj, J. Chen, R.J. Holt, A.N. Netravali, *Energy formulations of A-splines*, Comput. Aided Geom. Design **16** (1999), 39–59.
2. A. Borbély & M.J. Johnson, *Elastic Splines I: Existence*, Constr. Approx. **40** (2014), 189–218.
3. G. Birkhoff & C.R. de Boor, *Piecewise polynomial interpolation and approximation*, Approximation of Functions, Proc. General Motors Symposium of 1964, H.L. Garabedian ed., Elsevier, New York and Amsterdam, 1965, pp. 164-190.
4. G. Brunnett & J. Keifer, *Interpolation with minimal energy splines*, Computer-Aided Design **26** (1994), 137–144.
5. C. de Boor, K. Höllig & M. Sabin, *High accuracy geometric Hermite interpolation*, Comput. Aided Geom. Design **4** (1987), 269–278.
6. M.S. Floater, *On the deviation of a parametric cubic spline interpolant from its data polygon*, Comput. Aided Geom. Design **25** (2008), 148–156.
7. G. Jaklič & E. Žagar, *Planar cubic G^1 interpolatory splines with small strain energy*, J. Comput. Appl. Math. **235** (2011), 2758–2765.
8. S.N. Kersey, *Best near-interpolation by curves: existence*, SIAM J. Numer. Anal. **38** (2000), 1666–1675.
9. J. Kozak & M. Krajnc, *Geometric interpolation by planar cubic G^1 splines*, BIT **47** (2007), 547–563.
10. A. Linnér, *Unified representations of nonlinear splines*, J. Approx. Theory **84** (1996), 315–350.
11. L. Lu, *Planar quintic G^2 Hermite interpolation with minimum strain energy*, J. Comput. Appl. Math. **274** (2015), 109–117.
12. J. Yong & F. Cheng, *Geometric Hermite curves with minimum strain energy*, Comput. Aided Geom. Design **21** (2004), 281–301.

Department of Mathematics, Kuwait University, P.O. Box 5969, Safat 13060, Kuwait
E-mail address: yohnson1963@hotmail.com