Available online at www.sciencedirect.com
SCIENCE DIRECT®

Journal of Approximation Theory 130 (2004) 148–161

Journal of Approximation Theory

www.elsevier.com/locate/jat

An abstract formulation of variational refinement

Scott N. Kersey

Department of Mathematical Sciences, Georgia Southern University, Statesboro, GA 30460-8093, USA

Received 23 July 2003; accepted in revised form 27 July 2004

Communicated by Ulrich Reif
Available online 11 September 2004

Abstract

In this paper, the theory of abstract splines is applied to the variational refinement of (periodic) curves that meet data to within convex sets in $\mathbb{R}^d$. The analysis is relevant to each level of refinement (the limit curves are not considered here). The curves are characterized by an application of a separation theorem for multiple convex sets, and represented as the solution of an equation involving the dual of certain maps on an inner product space. Namely,

$$T^*Tf + \tilde{\Lambda}^* w\, \Gamma(\Lambda f) = 0.$$

Existence and uniqueness are established under certain conditions. The problem here is a generalization of that studied in (Kersey, Near-interpolatory subdivided curves, author's home page, 2003) to include arbitrary quadratic minimizing functionals, placed in the setting of abstract spline theory. The theory is specialized to the discretized thin beam and interval tension problems.

MSC: 41A05; 41A15; 41A29

Keywords: Variational refinement; Subdivision; Abstract splines

1. Introduction

The variational theory of polynomial splines began in the late 1950s with the problem of best interpolation [13]. Extensions to this problem were developed in the 1960s and 1970s; most notably, the problems of smoothing splines, least-squares splines and the ν-spline. Another problem studied in this time period was the computation of best spline fits

E-mail address: skersey@georgiasouthern.edu (S. N. Kersey).

0021-9045/$ - see front matter
doi:10.1016/j.jat.2004.07.006

constrained to interpolate within *intervals*, $l_i \leqslant f(t_i) \leqslant u_i$, as studied in [2,3,6,22,23]. During about the same time frame, an abstract theory of splines was being developed, perhaps first by Atteia (see [4]), followed up by others in [1,15], and using a different approach in [5] (based on earlier papers). The interval-constrained splines were generalized to curves in the problem of best near-interpolation (see [16,17]), with constraints given by "balls" in $\mathbb{R}^d$, and this was extended to arbitrary convex sets in [16,18], while also being studied independently in [24,25]. Along different lines, the problem of variational spline interpolation was extended to subdivided curves in [21], and generalized to near-interpolatory refinement to convex sets in [19] using a discretization of the linearized thin beam functional.

It is the goal here to develop an abstract theory of variational refinement to convex sets by combining the abstract spline theory with near-interpolatory refinement techniques. In this paper, we will be viewing the curves at each level of refinement as piecewise linear splines with fixed knots (the free-knot problem is studied to some extent in [19,20]). To simplify the presentation slightly, we will be assuming that the curves are closed periodic. To characterize the minimizers at each level of refinement, we apply an elegant separation theorem for multiple convex sets that was developed in [7,8] and generalized in [12] (see [11] for an exposition of [7,8], or see [14,26]). Following this, existence and uniqueness is established, and the theory is then applied to the discretized thin beam and interval weighted splines. We do not consider the limits of the refinement scheme here, which, in particular, would depend on parametrization, and is best left for future work. As a final introductory comment, the original title of this paper was *An abstract formulation of constrained subdivision*. However, as one of the reviewers pointed out, this paper is perhaps more about an abstract formulation of constrained "minimization" (applied to (near-)interpolatory "refinement") than what is generally studied in subdivision theory. The current title reflects this point of view.

2. Constrained refinement

Let X be the linear space of all closed-periodic piecewise linear (B-spline) curves $f(t) = \sum_{i=1}^{n+1} p_i N_{i,1}(t)$ with fixed knots $t_0, \ldots, t_{n+2}, t_i < t_{i+1}$, and coefficients $p_i = (p_i^1, \ldots, p_i^d)$ in $\mathbb{R}^d$. In particular, $p_i = f(t_i)$. Let $h_i := t_{i+1} - t_i$. Since f is closed, $p_{n+1} = p_1$; since it is periodic, $h_{n+1} = h_1$ and $h_0 = h_n$. Let $\lambda_i : X \to \mathbb{R}^d$ be the "vector-valued functionals" (linear maps) defined by the action $\lambda_i f := f(t_i) = p_i$, and let $\Lambda f := (\lambda_i f : i{=}1{:}n)$. Let

$$T_i(f) := \sum_{j=1}^{n} a_{ij} \lambda_j f = \sum_{j=1}^{n} a_{ij} p_j \in \mathbb{R}^d$$

for some coefficients $a_{ij} \in \mathbb{R}$. Typically, the sequences $a_{i,:}$ have small support for each i; as, for example, when T_i is the second divided difference operator $[t_{i-1}, t_i, t_{i+1}]$, as considered later in this paper. Since the knots may be non-uniform, the sequences $a_{i,:}$ are typically different (not simply a shift of one another) for each i, and the refinement schemes non-uniform. Let T be the map

$$T : X \to Y : f \longmapsto (T_i(f) : i{=}1{:}n)$$

with $Y := \mathbb{R}^{n\times d}$. We define the energy in the curve as

$$E(f) := \langle Tf, Tf\rangle_Y := \sum_{i=1}^{n} |T_i(f)|^2 := \sum_{i=1}^{n} T_i(f)\cdot T_i(f)$$

with "·" denoting the usual *dot product* in $\mathbb{R}^d$. In particular, $E(f)$ is quadratic and positive semi-definite, and can be written

$$E(f) = p^T H p = (\Lambda f)^T H(\Lambda f) \tag{1}$$

with H a symmetric positive semi-definite matrix determined by the coefficients a_{ij}.

Since each curve $f \in X$ is identified uniquely by its coefficient sequence $\Lambda f \in Y$, the usual inner product in Y induces an inner product on X; i.e.,

$$\langle f, g\rangle_X := \langle \Lambda f, \Lambda g\rangle_Y = \sum_i \lambda_i f \cdot \lambda_i g.$$

We split the inner product by the sum

$$\langle f, g\rangle_X \;:=\; \langle Sf, Sg\rangle_{\mathrm{Ker\,T}} + \langle T\,f, T\,g\rangle_Y, \tag{2}$$

with $S : X \to \ker\, T$ defined by orthogonal projection with respect to $\langle\cdot,\cdot\rangle_X$, and $\langle\cdot,\cdot\rangle_{\mathrm{Ker\,T}}$ an inner product on $\ker\, T$. On passing to the adjoint map $T^* : Y^* \to X^*$ of T, we have

$$E(f) = \langle T\,f, T\,f\rangle_Y = \langle T^*T\,f, f\rangle_X.$$

Note that in the last inner product we have associated the functional $T^*T\,f$ with its representer in X, by the Riesz Representation theorem. We will make similar associations throughout this paper, i.e., we will interchange spaces and maps with duals and representers as needed. For example, we associate X with X^* and Y with Y^*, and so $T^* : Y \to X$, as used above.

Let $\{I_1, I_2, I_3\}$ be a partition of $1{:}n$. For each index $i \in I_1$, we associate a point q_i to be interpolated, i.e., $p_i = \lambda_i f = q_i$; for each index $i \in I_2$ we associate a convex set K_i to be near-interpolated, i.e., $p_i \in K_i$; the remaining indices $i \in I_3$ correspond to points p_i that are free to vary. One can assume that the interpolated points (indices in I_1) are fixed from previous levels of refinement, the near-interpolated points (I_2) are constrained by the sets K_i, and the remaining points (I_3) are the new points added at the next level of refinement. As we assumed at the beginning of this section, the knots t_i of our spline curves f are prescribed, and so we only need to choose the coefficients p_i.

The constraint sets K_i are defined as

$$K_i := \bigcap_{j=1}^{m} K_{ij} \quad \text{with} \quad K_{ij} := \{x \in \mathbb{R}^d : g_{ij}(x) \leqslant 0\}$$

for some functions $g_{ij} : \mathbb{R}^d \to \mathbb{R}$, and some m. We assume that the functions g_{ij} are smooth, with non-vanishing gradient ∇g_{ij} on ∂K_i, the boundary of K_i. We assume, moreover, that the sets K_i are convex with non-empty interior K_i^o. Note that we include the possibility that $K_{ij} = \mathbb{R}^d$ for some ij by allowing $g_{ij} \equiv 0$ on $\mathbb{R}^d$. In this way, each set K_i is defined only by the non-trivial functions g_{ij}; at most m for each i.

In correspondence to the index sets defined above, let

$$\pi_j : Y \to \mathbb{R}^{\#I_j \times d} : x \to (x_i : i \in I_j), \quad j = 1, 2, 3.$$

Let $q = (q_i : i \in I_1)$ and $K := \times_{i \in I_2} K_i$. Our goal is to solve the minimization problem:

$$\underset{p}{\text{minimize}} \{E(f) : \pi_1 p = q,\ \pi_2 p \in K\}.$$

Or, with $\Omega := \{f \in X : \pi_1 \Lambda f = q,\ \pi_2 \Lambda f \in K\}$,

$$\underset{f \in \Omega}{\text{minimize}}\ E(f). \tag{3}$$

3. Additional notation

Throughout this paper we will be using various sequences. To unify the presentation, the following indexing and notation will be followed:

$$\begin{array}{ll} x \in \mathbb{R}^d, & x^k \in \mathbb{R} \text{ for } k{=}1{:}d, \\ \beta \in Y = \mathbb{R}^{n \times d}, & \beta_i^k \in \mathbb{R} \text{ for } i{=}1{:}n,\ j{=}1{:}m, \\ \alpha \in \mathbb{R}^{n \times d \times m}, & \alpha_{ij}^k \in \mathbb{R} \text{ for } i{=}1{:}n,\ j{=}1{:}m,\ k{=}1{:}d, \\ w \in \mathbb{R}_{\geqslant 0}^{\#I_2 \times m}, & w_{ij} \geqslant 0. \end{array}$$

It follows, for example, that $\beta_i \in \mathbb{R}^d$, $\beta^k \in \mathbb{R}^n$ and $\alpha_{ij} \in \mathbb{R}^d$. A similar convention will be used for other variables; e.g., $p = (p_i^k : i{=}1{:}n, k{=}1{:}d)$, as before. To characterize solutions to (3), certain maps and their duals will be defined. Since our functionals (maps) λ_i are "vector-valued", the notation is perhaps non-standard. In particular, vector-vector multiplication is defined pointwise. Let:

$$\begin{array}{ll} \lambda_i^k & : f \longmapsto \lambda_i^k(f) = p_i^k, \quad \text{(the } k\text{-th coordinate of } \lambda_i f) \\ \lambda_i & : X \longrightarrow \mathbb{R}^d : f \longmapsto \lambda_i f = (\lambda_i^1 f, \ldots, \lambda_i^d f), \\ \lambda_i^* & : (\mathbb{R}^d)^* \longrightarrow X^* : x \longmapsto x\lambda_i = (x^1 \lambda_i^1, \ldots, x^d \lambda_i^d), \\ \Lambda & : X \longrightarrow Y : f \longmapsto (\lambda_1 f, \ldots, \lambda_n f), \\ \Lambda^* & : Y^* \longrightarrow X^* : \beta \longmapsto \sum_{i=1}^n \beta_i \lambda_i = \sum_{i=1}^n (\beta_i^1 \lambda_i^1, \ldots, \beta_i^d \lambda_i^d), \\ \tilde{\Lambda} & : X \longrightarrow \mathbb{R}^{\#I_2 \times d \times m} : f \longmapsto (\underbrace{\lambda_i f, \ldots, \lambda_i f}_{m \text{ times}} : i \in I_2), \\ \tilde{\Lambda}^* & : \mathbb{R}^{n \times d \times m} \longrightarrow X^* : \alpha \longmapsto \sum_{i \in I_2} \sum_{j=1}^m \alpha_{ij} \lambda_i = \sum_{i \in I_2} \sum_{j=1}^m (\alpha_{ij}^1 \lambda_i^1, \ldots, \alpha_{ij}^d \lambda_i^d), \\ \Gamma & : Y \longmapsto \mathbb{R}^{\#I_2 \times d \times m} : y \longmapsto (\nabla g_{i1}(y_i), \ldots, \nabla g_{i,m}(y_i) :\ i \in I_2). \end{array}$$

4. Characterization

In this section, solutions to (3) are characterized in Theorem 2. To do so, we apply a separation theorem by Dubovitskii and Milyutin for convex cones, generalized to arbitrary convex sets by Halkin, stated next.

Theorem 1 (see Dubovitskii and Milyutin [8, Theorem 2.1], and Halkin [12, Lemma 4.2]). *Let $C_0, \ldots, C_l$ be convex sets in a normed linear space X with C_i open for $i > 0$ and $0 \in \overline{C_i}$ (the closure of C_i) for all i. Then $\cap_i C_i = \emptyset$ iff $(C_i : i{=}0{:}l)$ is separated at 0 in the sense that there exists a sequence of functionals $(\mu_0, \ldots, \mu_l)$ in X^*, not all zero, with $\sum_i \mu_i = 0$ and $\inf \mu_i C_i \geqslant 0$, all i.*

Theorem 2. *$f \in \Omega$ solves (3) iff*

$$T^*Tf + \sum_{i\in I_2}\sum_{j=1}^{m} w_{ij}\lambda_i^*(\nabla g_{ij}(\lambda_i f)) = 0 \tag{4}$$

for some nonnegative multipliers w_{ij} with $w_{ij} = 0$ when $g_{ij}(\lambda_i f) < 0$. Equivalently,

$$T^*Tf + \tilde{\Lambda}^* w\,\Gamma(\Lambda f) = 0. \tag{5}$$

Moreover, there are only global minimizers (i.e., local minimizers are global minimizers).

Proof. To preserve the interpolation condition $\pi_1 \Lambda f = q$, (linear) *variations* $f + v$ of f must satisfy $\pi_1\Lambda(f + v) = \pi_1\Lambda f$, and so $\pi_1\Lambda v = 0$. This is easily accomplished by restricting v to $\tilde{X} := X \cap \ker \pi_1\Lambda$. And so, we will be applying Theorem 1 on $\tilde{X}$ rather than X, with the same inner product, but under the relative topology.

Let C_0 be the set

$$\begin{aligned} C_0 &:= \{v \in \tilde{X} : \langle Tf, Tv\rangle_Y < 0\} \\ &= \{v \in \tilde{X} : \langle T^*Tf, v\rangle_X < 0\} \end{aligned}$$

of directions v along which E is strictly decreasing, and let C_{ij} be the sets

$$\begin{aligned} C_{ij} &:= \{v \in \tilde{X} : \nabla g_{ij}(\lambda_i f)\cdot \lambda_i v < 0 \text{ if } g_{ij}(\lambda_i(f)) = 0\} \\ &= \{v \in \tilde{X} : \langle \lambda_i^*(\nabla g_{ij}(\lambda_i f)),\ v\rangle_X < 0 \text{ if } g_{ij}(\lambda_i(f)) = 0\} \end{aligned}$$

of feasible directions strictly into the sets K_i, for $i \in I_2$, $j{=}1{:}m$. Since f is a (local) minimizer exactly when $E(\cdot)$ is not decreasing along feasible directions into (including the boundary of) Ω, it is a minimizer iff

$$C_0 \cap \bigcap_{ij} \overline{C}_{ij} = \emptyset.$$

Moreover, for the setup here $C_0 \cap \bigcap_{ij} C_{ij} = \emptyset \implies C_0 \cap \bigcap_{ij} \overline{C}_{ij} = \emptyset$, for if $v \in C_0 \cap \bigcap_{ij} \overline{C}_{ij} \neq \emptyset$, then $v + \varepsilon(w - v) \in C_0 \cap \bigcap_{ij} C_{ij}$ for any $w \in \bigcap_{ij} C_{ij}$ and $\varepsilon > 0$ small enough, implying $C_0 \cap \bigcap_{ij} C_{ij} \neq \emptyset$. Therefore, f is a (local) minimizer iff

$$C_0 \cap \bigcap_{ij} C_{ij} = \emptyset. \tag{6}$$

The sets C_0 and C_{ij} are (relatively) open in $\tilde{X}$ and contain 0 in their closure. Indeed, as $c \downarrow 0$ in $\mathbb{R}_+$, $-cf \in C_0$ and $-cv \in C_{ij}$ for any $v \in \tilde{X}$ such that $\lambda_i v = \nabla g_{ij}(\lambda_i f)$. Therefore, by

(6) and Theorem 1 (with the sets C_{ij} in place of C_i for $i > 0$), f is a (local) minimizer iff there exist linear functionals μ_0 and μ_{ij} on $\tilde{X}$, not all zero, such that $\mu_0 + \sum_{i \in I_2} \sum_{j=1}^{m} \mu_{ij} = 0$, with $\inf \mu_0 C_0 \geqslant 0$ and $\inf \mu_{ij} C_{ij} \geqslant 0$. Moreover, since one can always choose v such that $\lambda_i v$ is directed into the convex open set K_i from $\lambda_i f$, and since Λ is an onto map, it follows that $\bigcap_{ij} C_{ij} \neq \emptyset$. As a consequence, it follows by Theorem 1 applied to the sets C_{ij} (not including C_0), that $\sum \mu_{ij} \neq 0$. Therefore, in the context above, $\mu_0 \neq 0$ when $\mu_0 + \sum_{ij} \mu_{ij} = 0$.

Since $\mu_0 C_0 \geqslant 0$, it follows that

$$\mu_0 = -w_0 \langle T^* T f, \cdot \rangle_X$$

for some $w_0 \geqslant 0$, and since $\mu_{ij} C_{ij} \geqslant 0$,

$$\mu_{ij} = -w_{ij} \langle \lambda_i^* (\nabla g_{ij}(\lambda_i f)), \cdot \rangle_X$$

for some $w_{ij} \geqslant 0$ when $g_{ij}(\lambda_i f) = 0$. On the other hand, $C_{ij} = \tilde{X}$ when $g_{ij}(\lambda_i f) < 0$, in which case $\mu_{ij} = 0$ and $w_{ij} = 0$.

Therefore, f is a local minimizer of E from Ω iff

$$-w_0 \langle T^* T f, \cdot \rangle_X + \sum_{i \in I_2} \sum_{j=1}^{m} -w_{ij} \langle \lambda_i^* (\nabla g_{ij}(\lambda_i f)), \cdot \rangle_X = 0$$

on X for some $w_0 \geqslant 0$ and $w_{ij} \geqslant 0$, with $w_{ij} = 0$ when $g_{ij}(\lambda_i f) < 0$. Moreover, since $\mu_0 \neq 0$, it follows that $w_0 > 0$. Without loss of generality, we may assume that $w_0 = 1$, and so

$$\langle T^* T f + \sum_{i \in I_2} \sum_{j=1}^{m} w_{ij} \lambda_i^* (\nabla g_{ij}(\lambda_i f)), \cdot \rangle_X = 0$$

implying, moreover, that the representer of this functional vanishes. That is,

$$T^* T f + \sum_{i \in I_2} \sum_{j=1}^{m} w_{ij} \lambda_i^* (\nabla g_{ij}(\lambda_i f)) = 0.$$

Equivalently,

$$T^* T f + \tilde{\Lambda}^* w \, \Gamma(\Lambda f) = 0.$$

Finally, local minimizers are global minimizers, since, by the convexity of Ω, $f + s(\hat{f} - f)$ is in Ω for all $s \in [0, 1]$ when $\hat{f}$ is in Ω, and, by the convexity of E,

$$E(f) \leqslant E(f + s(\hat{f} - f)) \leqslant E(f) + s(E(\hat{f}) - E(f))$$

for all s small enough (say $s \in [0, \varepsilon]$ for some small $\varepsilon > 0$) when f is a local minimizer, thereby implying that $E(f) \leqslant E(\hat{f})$ for all $\hat{f} \in \Omega$. $\square$

5. Existence and uniqueness

Definition 3. We say that E is coercive on X if $E(f) \to \infty$ as $\|f\|_X \to \infty$.

Definition 4. We say that (f_l) is a minimizing sequence for (3) if $f_l \in \Omega$ for each l and

$$\lim\ E(f_l) = \inf\{E(f) : f \in \Omega\}.$$

Theorem 5. *Assume, as above, that K is closed, nonempty and convex, and that $\Omega \neq \emptyset$. Then, solutions to* (3) *exist when either $\Omega \cap \ker\ T \neq \emptyset$, or when $\Omega \cap \ker\ T = \emptyset$ and E is coercive on X.*

Proof. Existence is trivial to establish in the case that $\Omega \cap \ker\ T \neq \emptyset$ since $E(f) = 0$ for any $f \in \Omega \cap \ker\ T$. And so, we will henceforth assume that $\Omega \cap \ker\ T = \emptyset$. Let (f_l) be a minimizing sequence for E in Ω. By the coercivity assumption, (f_l) is bounded with respect to $\|\cdot\|_X$. Since X is a finite dimensional space, all norms on it are equivalent. In particular, recalling that $p_i := \lambda_i f$ are the spline coefficients for f, the Euclidean norm of $p = \Lambda f$ in $\mathbb{R}^{n\times d}$ is a norm for f in X. Therefore, since (f_l) is bounded in X, it follows that (Λf_l) is bounded in $\mathbb{R}^{n\times d}$, and so (Λf_l) has convergent subsequences. On passing to a subsequence, we may assume that $\Lambda f_l \to p \in \mathbb{R}^{n\times d}$. Since K is closed and $\pi_2 \Lambda f_l \in K$ for each l, it follows that $\pi_2 p \in K$; since $\pi_1 \Lambda f_l = q$ for all l, $\pi_1 p = q$. This p is the coefficient sequence for some $f \in \Omega$. Since f_l is a minimizing sequence for $E(\cdot)$, it follows that $E(f) \leqslant E(f_l)$ for all l. Hence, f solves (3). $\square$

This existence result will be applied to the setup given in the next section. In particular, coercivity is established for a specific objective functional $E(\cdot)$ of practical interest. Our next goal is to establish uniqueness under certain conditions. For this we need the following result:

Lemma 6. *Suppose that f_1 and f_2 both solve* (3). *Then,* $f_1 - f_2 \in \ker\ T$.

Proof. Since f_1 and f_2 both minimize $E(\cdot) = \|T\cdot\|_Y^2$ over Ω, it follows that Tf_1 and Tf_2 are minimal norm elements in $T\Omega \subset Y$. Moreover, as the image of a convex set under a linear map, $T\Omega$ is convex in Y, and so there can be only one minimal norm element in $T\Omega$. Therefore $Tf_1 = Tf_2$, and so $f_1 - f_2 \in \ker\ T$. $\square$

The following condition is used to establish uniqueness. The terminology is borrowed from [9], but in a different context.

Definition 7. We say that the setup is well-posed if $\ker(\beta\Lambda) \cap \ker\ T = \{0\}$ whenever $\beta \in Y$ is chosen such that $0 \neq \sum_{i\in I_2} \beta_i^k \lambda_i^k \in (\ker\ T)^\perp$ for $k{=}1{:}d$.

Uniqueness is established in the next theorem. For this, we require $\ker\ T^k \cap \Omega = \emptyset$ with

$$T^k : X \longrightarrow \mathbb{R}^n : f \longmapsto (Tf)^k, \quad \text{(the } k\text{-th coordinate of } Tf)$$

k=1:d. Note that this condition is more restrictive than $\ker T \cap \Omega = \emptyset$. In the following proof, we make the association $\operatorname{ran} T^* = (\ker T)^\perp$. This follows because the subspace $\operatorname{ran} T$ is closed in Y (see [10, Theorem 4.13.6]).

Theorem 8. *Suppose that* $\ker T^k \cap \Omega = \emptyset$ *for* k=1:d, *that the sets* K_i *are strictly convex, and that the setup is well-posed. Then, there is at most one solution to* (3).

Proof. Suppose that f_1 and f_2 both solve (3), each necessarily achieving the minimum value $e := \inf\{E(f) = \|Tf\|_Y^2 : f \in \Omega\}$ of E over Ω. Let $f := (f_1 + f_2)/2$. Due to the convexity of Ω, $f \in \Omega$. Moreover, f is also a solution to (3) with value e, as follows from the inequality

$$\sqrt{e} \leqslant \sqrt{E(f)} = \left\| T\frac{(f_1+f_2)}{2} \right\| \leqslant \frac{1}{2}(\|Tf_1\|_Y + \|Tf_2\|_Y) = \sqrt{e}.$$

Since f is a solution to (3), it follows by (5) that

$$T^*Tf = -\tilde{\Lambda}^* w\,\Gamma(\Lambda f)$$

for some nonnegative multipliers w_{ij}. Let

$$\mu := -\tilde{\Lambda}^* w\,\Gamma(\Lambda f) = \sum_{i\in I_2} \beta_i \lambda_i$$

with

$$\beta_i := -\sum_{j=1}^{m} w_{ij} \nabla g_{ij}(\lambda_i f)$$

and let $\mu^k : v \mapsto (\mu v)^k$, the k-th coordinate-map of μ, for k=1:d. Since $T^*Tf = \mu$, it follows that μ is in $\operatorname{ran} T^* = (\ker T)^\perp$, and so $\mu^k \in (\ker T)^\perp$ as well. Moreover, $\mu^k \neq 0$, for otherwise $T^k f = 0$, violating the assumption $\ker T^k \cap \Omega = \emptyset$. (To see this, note that $\mu^k = 0$ implies

$$0 = \langle \mu^k, f\rangle_X = \langle (T^*Tf)^k, f\rangle_X = \langle (T^k)^* T^k f, f\rangle_X = \langle T^k f, T^k f\rangle_X.)$$

Since

$$0 \neq \mu^k = \sum_{i\in I_2} \beta_i^k \lambda_i^k \in (\ker T)^\perp, \quad k=1:d$$

and the system is well-posed, it follows that $\ker(\beta\Lambda) \cap \ker T = \{0\}$. By Lemma 6, $f_1 - f_2 \in \ker T$, and so, to prove uniqueness, it remains to show that $f_1 - f_2 \in \ker(\beta\Lambda)$.

Given a solution σ to (3), let A_σ denote the set of indices ij, restricted to $i \in I_2$, such that the ij-th constraint is active, meaning that $g_{ij}(\lambda_i \sigma) = 0$ and $w_{ij} > 0$ (i.e., when $w_{ij} \neq 0$). Now, let $f := (f_1 + f_2)/2$, as above, with multipliers w_{ij}. By convexity of the sets K_i, the ij-th constraint is active for f iff it is active for both f_1 and f_2, and so $A_f \subset A_{f_1} \cap A_{f_2}$. Moreover, by strict convexity of the sets K_i, $\lambda_i f = \lambda_i f_1 = \lambda_i f_2$ when

$ij \in A_f$. Equivalently, $\lambda_i(f_1 - f_2) = 0$ when $w_{ij} \neq 0$. Therefore, $w_{ij}\lambda_i(f_1 - f_2) = 0$ for all *ij*. By the definition of β_i given above, it follows that

$$\beta_i \lambda_i (f_1 - f_2) = -\sum_{j=1}^{m} \nabla g_{ij}(\lambda_i f) w_{ij} \lambda_i (f_1 - f_2) = 0$$

for all *i*, and so $f_1 - f_2 \in \ker(\beta\Lambda)$.

To conclude, we have shown that $f_1 - f_2 \in \ker(\beta\Lambda) \cap \ker T = \{0\}$, and so $f_1 = f_2$. Therefore, there can be at most one solution. □

To see what can go wrong when we do not have the well-posed assumption, consider the following example:

Example 9. Let *X* be the set of closed-periodic piecewise linear spline curves with knots $t_1 = 1$, $t_2 = 2$ and $t_3 = 3$ and coefficients p_1, p_2 and p_3. Let K_i be the closed balls $K_1 = B_\varepsilon(1, 0)$, $K_2 = B_\varepsilon(0, 1)$ and $K_3 = B_\varepsilon(-1, 0)$ for some “small” radius ε. Define T_i by their action: $T_1 f = f(t_1) + f(t_3) = p_1 + p_3$, $T_2 \equiv 0$ and $T_3 \equiv 0$.

Proposition 10. *The setup in Example 9 is not well-posed. Moreover, solutions to (3) exist, but are not unique.*

Proof. Solutions exist since $E(f) = 0$ for $p_1 = (1, 0)$ and $p_3 = (-1, 0)$, however, they are not unique since p_2 can be any point in K_2. To see that the setup is not well-posed, let $\beta_1 = (1, 1)$, $\beta_2 = (0, 0)$ and $\beta_3 = (1, 1)$ in Definition 7. Then, since $f \in \ker T$ iff $p_1 = -p_3$, it follows that

$$\mu f = \beta_1 p_1 + \beta_2 p_2 + \beta_3 p_3 = p_1 + p_3 = 0$$

and so $\mu \in (\ker T)^\perp$. Moreover, $\mu^k \neq 0$ for $k = 1, 2$ since $\beta_1^k \neq 0$. We have established that $0 \neq \mu^k \in (\ker T)^\perp$, while $\ker(\beta\Lambda) \cap \ker T \neq \{0\}$ because p_2 is arbitrary. Therefore, the setup is not well-posed. □

In the next section, Theorem 8 is applied to setup where $E(\cdot)$ the discretized thin beam functional, which is well-posed. Hence, we can verify uniqueness when the sets K_i are strictly convex. We also show that, although strict convexity is needed in certain cases, it is not a necessary condition in Theorem 8.

6. The discretized thin beam

In [19], the following discretization of the linearized thin beam energy functional $\frac{1}{2}\int_{t_1}^{t_{n+1}} |D^2 f(t)|^2 \,\mathrm{d}t$ was studied:

$$E(f) := \frac{1}{2}\sum_{i=1}^{n} \int_{\frac{t_i + t_{i-1}}{2}}^{\frac{t_i + t_{i+1}}{2}} |2\,\Delta_{i-1,2} f|^2 \,\mathrm{d}t = \sum_{i=1}^{n} |\Delta_{i-1,2} f|^2 \, h_{i-1,2}. \tag{7}$$

Here, $\Delta_{i-1,2} := [t_{i-1}, t_i, t_{i+1}]$ is the second divided difference operator. In the context of this paper $T_i := \sqrt{h_{i-1,2}}\,\Delta_{i-1,2}$. In this section, we establish conditions for the existence and uniqueness of solutions to (3) with $E(\cdot)$ as in (7). We also comment on the representation of linear functionals in our reproducing Hilbert space.

We may first observe that if $E(f) = 0$ for this functional, then all second divided differences vanish. Therefore, ker T is contained in the space of linear curves. But the only linear curves that are also closed are "constant curves". That is, ker T consists of curves f with coefficients $p = \Lambda f$ on the diagonal $(x, x, \ldots, x)$ in $Y = \mathbb{R}^{n\times d}$. In particular, $\dim(\ker T) = d$.

Theorem 11. *Let $E(\cdot)$ be as in (7). Solutions to (3) exist for the energy functional (7) when either $\Omega \cap \ker T \neq \emptyset$, or when $\Omega \cap \ker T = \emptyset$ and at least one of the sets K_i is bounded. Solutions are unique when the sets K_i are strictly convex and $\Omega \cap \ker T^k = \emptyset$ for k=1:d.*

Proof. Existence follows directly from Theorem 5 when $\Omega \cap \ker T \neq \emptyset$, and can be established when $\Omega \cap \ker T = \emptyset$ if we can satisfy the coercivity condition in Definition 3. To this end, recall the inner product

$$\langle f, g\rangle_X = \langle Sf, Sg\rangle_{\mathrm{Ker\,T}} + \langle T\,f, T\,g\rangle_Y$$

given in (2), with $\langle\cdot,\cdot\rangle_{\mathrm{Ker\,T}}$ some inner product on ker T. As stated above, ker T is comprised of the constant functions when $E(\cdot)$ is given by (7). In particular, we can choose

$$\langle Sf, Sg\rangle_{\mathrm{Ker\,T}} := \lambda_i f \cdot \lambda_i g$$

for any i. Here, we choose i to correspond to a bounded set K_i, as hypothesized in the theorem. By (2) and (1),

$$\begin{aligned}\|f\|_X^2 &= \langle Sf, Sf\rangle_{\mathrm{Ker\,T}} + \langle T\,f, T\,f\rangle_Y \\ &= p_i \cdot p_i + p^T H p\end{aligned}$$

is a norm (-squared) on X, with $p = \Lambda f$. Now, since K_i is bounded (for this particular i), then p_i is bounded, and so $E(f) = p^T H p$ and $\|f\|_X$ go to infinity together. In particular, $E(f) \to \infty$ when $\|f\|_X \to \infty$. This establishes coercivity, and existence when $\Omega \cap \ker T = \emptyset$.

It remains to establish uniqueness. By Theorem 8, we need to show that the setup is well-posed. To do so, suppose that $\mu^k f = 0$ for k=1:d with $\mu = \sum_i \beta_i \lambda_i$ for some coefficients β_i in $\mathbb{R}^d$. Since we are also assuming that $\mu^k \neq 0$, it follows that β_i^k is nonzero for some i, for each k (actually, there are at least two nonzero β_i^k for each k in the setup here). Then, if $f \in \ker \beta\Lambda$, it follows that $p_i^k = 0$ for at least one i, and each k. But if f is in ker T, then it is a constant curve, and so $p_1^k = p_2^k = \cdots = 0$. Therefore, $p = 0$ in Y, and so $f \equiv 0$. This establishes the well-posedness condition in Definition 7. By Theorem 8, solutions are unique. □

To see that boundedness (or perhaps some other condition) is needed to establish existence in Theorem 11, consider the following example:

Example 12. Let $K := K_1 \times K_2 \times K_3$ with

$$K_1 := \{(x, y) \in \mathbb{R}^2 : x\,y \geqslant 1,\ \ x, y > 0\},$$
$$K_2 := \mathbb{R}^2,$$
$$K_3 := \{(x, y) \in \mathbb{R}^2 : x\,y \geqslant 1,\ \ x, y < 0\}.$$

Let X be the set of closed-periodic piecewise linear spline curves with knots $t_1 = 1$, $t_2 = 2$ and $t_3 = 3$, and coefficients p_1, p_2 and p_3 in $\mathbb{R}^2$.

Proposition 13. *For the setup in Example 12, solutions to (3) do not exist.*

Proof. In Example 12, each set K_i is closed and convex in $\mathbb{R}^2$. Moreover, $\Omega \cap \ker\, T = \emptyset$ since $\ker\, T$ contains only constant curves and $\cap K_i = \emptyset$. Therefore, $E(f) > 0$ for any $f \in \Omega$. However, $\inf E(\cdot) = 0$ over Ω. To see this, let (f^k) be a sequence of spline curves with coefficients $p_1^k = \left(\frac{1}{k}, k\right)$, $p_2^k = \left(0, k + \frac{\sqrt{3}}{k}\right)$ and $p_3^k = \left(-\frac{1}{k}, k\right)$. Each of the curves f^k is an equilateral triangle in Ω. As $k \to \infty$ these triangles shrink to a point and are pushed up to ∞ in the y coordinate. Moreover, $E(f^k) = \frac{18}{k^2}$, and so $E(f^k) \to 0$ as $k \to \infty$. In particular, (f^k) is a minimizing sequence. But since $E(\cdot) = 0$ is not achieved in Ω, solutions do not exist. □

One important special case is when the sets K_i are Euclidean balls in $\mathbb{R}^d$. In this case the sets K_i are strictly convex, and so we have uniqueness by Theorem 8. To see what can go wrong when the sets K_i are not strictly convex, consider the following example:

Example 14. Let $K := K_1 \times K_2 \times K_3$ with

$$K_1 := \{(x, y) \in \mathbb{R}^2 : x \geqslant \varepsilon,\ \ y \geqslant 1\},$$
$$K_2 := \{(x, y) \in \mathbb{R}^2 : x \leqslant -\varepsilon,\ \ y \geqslant 1\},$$
$$K_3 := \{(x, y) \in \mathbb{R}^2 : y \leqslant 0\}.$$

Here, ε is some small positive number (as small as we need it to be). Let the knots be uniform.

Proposition 15. *For the setup in Example 14, solutions to (3) are not unique.*

Proof. Suppose that f solve (3) for this configuration, with coefficients p_1, p_2 and p_3. Then, either $p_1 \neq (\varepsilon, \cdot)$ or $p_2 \neq (-\varepsilon, \cdot)$, for otherwise $E(f)$ would be large due to a very small angle at p_3. But this means that f can be shifted either to the left or right without violating the constraints $f(t_i) \in K_i$, and without increasing the energy. That is, the curve $\tilde{f}$ with coefficients $\tilde{p} := p + (\delta, 0)$ is also a solution for some δ, and so, solutions are not unique. □

Unfortunately, we are often interested in configurations where the K_i are not strictly convex, and it seems tricky to have a blanket uniqueness condition in this case. However,

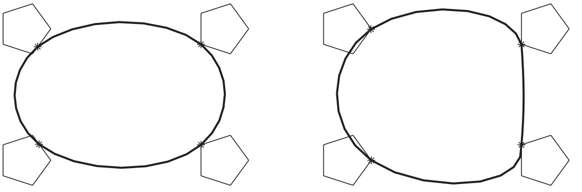

Fig. 1. Discretized thin beam (left); interval tension (right).

once we have computed a solution for a given configuration, we can often verify uniqueness by inspection. For example, consider the left image in Fig. 1. Here, we have computed a minimizer f for (3) for the minimizing functional $E(\cdot)$ given in (7). (The curve in the image was actually generated after a couple levels of refinement.) We know by Theorem 2 that this is necessarily a global minimizer. Moreover, we know by Lemma 6 that if there is another solution, $\tilde{f}$, then $f - \tilde{f} \in \ker\, T$. That is, f differs from $\tilde{f}$ by a constant function. However, by inspection of the curve in the figure, f cannot be shifted by a constant (a linear translation) without violating the constraints. Therefore, the curve computed must be the unique global minimizer to (3) (to within computational tolerances).

We conclude this section with another feature of reproducing kernel Hilbert spaces. Namely, that the representers $\lambda_i^* = \phi_i(t)$ of linear functionals λ_i on X are themselves elements of X. For the setup here, the representers of the (vector-) functionals λ_i are necessarily piecewise linear, periodic splines with knots t_i. They act by the inner product as follows:

$$\lambda_i f = \langle f, \phi_i \rangle_X = f(t_1) \cdot \phi_i(t_1) + (\Lambda f)^T H\, \Lambda \phi_i.$$

Moreover, as shown in [19],

$$E(f) = \sum_{i=1}^{n} p_i \cdot \mathrm{jmp}_{t_i}(D^3 f) = (\Lambda f)^T\, \mathrm{jmp}_t(D^3 f)$$

with $\mathrm{jmp}_t(D^3 f) := (\mathrm{jmp}_{t_i}(D^3 f) : i{=}1{:}n)$ for the “jump maps”

$$\mathrm{jmp}_{t_i}(D^3\, f) := \frac{h_{i-1,3}}{h_i}\, \Delta_{i-1,3}\, f - \frac{h_{i-2,3}}{h_{i-1}}\, \Delta_{i-2,3}\, f.$$

Therefore, the representers take on the action

$$\lambda_i f = \langle f, \phi_i \rangle_X = f(t_1) \cdot \phi_i(t_1) + (\Lambda f)^T\, \mathrm{jmp}_t(D^3 \phi_i).$$

Now, $f = \sum_j p_j N_j(t)$ in our piecewise linear spline basis, and in this basis $\lambda_i N_j = \langle N_j, \phi_i \rangle_X = \delta_{ij}$, with δ_{ij} the Kroenecker-delta function. Therefore, we can determine the representers $\phi_i(t)$ by solving linear systems whose rows are determined from

$$N_j(t_1) \cdot \phi_i(t_1) + \mathrm{jmp}_{t_j}(D^3 \phi_i) = \delta_{ij}$$

for $j{=}1{:}n$, to determine the coefficients α_k in the expansion $\phi_i(t) := \sum_{k=1}^{n} \alpha_k N_k(t)$. The linear system is almost banded, with bandwidth 5 on the banded part, just as in periodic cubic spline interpolation.

These representers can be used in computation. Assuming that we are given the multipliers w_{ij}, we have, by (5), the following:

$$\begin{aligned} T^*Tf &= -\tilde{\Lambda}^* w\,\Gamma(\Lambda f) \\ &= -\sum_{i\in I_2}\sum_{j=1}^m w_{ij}\lambda_i^*(\nabla g_{ij}(\lambda_i f)) \\ &= -\sum_{i\in I_2}\left(\sum_{j=1}^m w_{ij}\nabla g_{ij}(\lambda_i f)\right)\phi_i(\cdot). \end{aligned}$$

From this, we can recover *f*.

7. Interval tension

It is relatively straight forward to experiment with different refinement functionals. For example, to achieve interval tension, we can modify the discretized thin beam functional as follows:

$$E(f) := \frac{1}{2}\sum_{i=1}^n \int_{\frac{t_i+t_{i-1}}{2}}^{\frac{t_i+t_{i+1}}{2}} \beta_i\,|2\,\Delta_{i-1,2}f|^2\,\mathrm{d}t = \sum_{i=1}^n \beta_i|\Delta_{i-1,2}f|^2\,h_{i-1,2}. \tag{8}$$

Here, β_i are interval tension parameters, assumed to be positive. The curves in Fig. 1 were computed after a couple levels of refinement, using the energy functional (7) for the left image, and (8) for the right. The effect of interval tension is quite apparent in the right image.

8. Conclusion

In this paper, we present an abstract approach to variational refinement for curves that meet arbitrary convex constraints. We investigate the characterization, existence and uniqueness of solutions in a general, abstract framework. But the analysis is based only on one level of refinement. In particular, the smoothness of the limiting curves is not considered. This is an interesting and open problem, complicated by the non-uniformity of the knots. That is, the smoothness depends on the parametrization of the curves. The author is currently investigating this problem when $E(\cdot)$ is the energy functional in Section 6. Another useful generalization may be to allow "functionals" λ_i other than point evaluation, as is typically the case in generalized spline theory.

Acknowledgments

I would like to thank the reviewers for their thoughtful comments on the submitted manuscript.

References

[1] P.M. Anselone, P.J. Laurent, A general method for the construction of interpolating or smoothing spline-functions, Numer. Math. 12 (1968) 66–82.
[2] M. Atteia, Fonctions spline avec contraintes linéaires de type inégalite, Congrès de l'AFIRO, Nancy, Mai, 1967.
[3] M. Atteia, Fonctions spline définiés sur un ensemble convexe, Numer. Math. 12 (1968) 192–210.
[4] M. Atteia, Hilbertian kernels and Spline Functions, Elsevier Science Publishers, Amsterdam, 1992.
[5] A.Yu. Bezhaev, V.A. Vasilenko, Variational Spline Theory, Kluwer Academic/Plenum Publishers, New York, 2001.
[6] P. Copley, L.L. Schumaker, On pLg-splines, J. Approx. Theory 23 (1978) 1–28.
[7] A.Ya. Dubovitskii, A.A. Milyutin, Extremum problems with constraints, Soviet Math. 4 (1963) 452–455.
[8] A.Ya. Dubovitskii, A.A. Milyutin, Extremum problems in the presence of restrictions, U.S.S.R. Comput. Math. Math. Phys. 5 (3) (1965) 1–80.
[9] D. Ferguson, The question of uniqueness for G.D. Birkhoff interpolation problems, J. Approx. Theory 2 (1969) 1–28.
[10] A. Friedman, Foundations of Modern Analysis, Dover, New York, 1982.
[11] I.V. Girsanov, Lectures on the Mathematical Theory of Extremal Problems, Springer, New York, 1972.
[12] H. Halkin, A satisfactory treatment of equality and operator constraints in the Dubovitskii-Milyutin optimization formalism, J. Optim. Theory Appl. 6 (2) (1970) 138–149.
[13] J.C. Holladay, A smoothest curve approximation, Math. Tables Aids Comput. 11 (1957) 233–243.
[14] R.B. Holmes, A Course on Optimization and Best Approximation, Springer, 1970.
[15] J. Jerome, L.L. Schumaker, A note on obtaining natural spline functions by the abstract approach of Atteia and Laurent, SIAM J. Numer. Anal. 5 (1968) 657–663.
[16] S. Kersey, Best near-interpolation by curves: existence, SIAM J. Numer. Anal. 38 (2000) 1666–1675.
[17] S. Kersey, Near interpolation, Numer. Math. 94 (3) (2003) 523–540.
[18] S. Kersey, Near-interpolation to arbitrary constraints, in: L. Schumaker, T. Lyche, M. Mazure (Eds.), Curve and Surfaces Design: Saint Malo 2002, Nashville Press, 2003, pp. 235–244.
[19] S. Kersey, Near-interpolatory subdivided curves, manuscript can be downloaded from the author's home page, 2003.
[20] S. Kersey, The effect of parametrization on subdivided curves, manuscript can be downloaded from the author's home page, 2003.
[21] L. Kobbelt, A variational approach to subdivision, Comput. Aided Geom. Design 13 (1996).
[22] P.J. Laurent, Construction of spline functions in a convex set, in: I. Schoenberg (Ed.), Approximation with Special Emphasis on Spline Functions, Madison WI, 1969, pp. 415–446.
[23] O.L. Mangasarian, L.L. Schumaker, Splines via optimal control, in: I. Schoenberg (Ed.), Approximation with Special Emphasis on Spline Functions, Madison WI, 1969, pp. 119–155.
[24] J. Wallner, H. Pottman, Variational interpolation, Technical Report No. 84, Institut für Geometrie, Technische Universität Wien, 2001. http://www.geometrie.tuwien.ac.at/wallner/publ.html.
[25] J. Wallner, H. Pottman, Variational interpolation of subsets, Constr. Approx. 20 (2004) 233–248.
[26] E. Zeidler, Nonlinear Functional Analysis and its Applications III, Springer, New York, 1985.