BEST NEAR-INTERPOLATION BY CURVES: EXISTENCE AND CONVERGENCE

Scott N. Kersey
The Center for the Mathematical Sciences
The University of Wisconsin-Madison
kersey@math.wisc.edu

Abstract. In this paper, the conditions derived in [10] for the existence of minimizers to the nonlinear problem of best "interpolation" by curves are extended to the problem of best "near-interpolation" by curves that meet arbitrary sets, such as closed balls (as in [6]). The minimizers are spline curves with breakpoints at the data sites at which the curves meet the sets, and the nonlinearities arise as these data sites vary from curve to curve. The results here apply to Hermite-type interpolation conditions, with the possibility of repeated data sites. Following the proof of existence, we show that certain sequences of spline curves related to the minimization problem converge in norm, even when some of their breakpoints (hence, knots) coalesce.

Key words. splines, interpolation, near-interpolation, parametric curves, approximation

AMS(MOS) subject classifications. 41*A*05, 41*A*15, 41*A*29

1. Introduction. In this paper, necessary and sufficient conditions are derived for the existence of solutions to the problem

$$\underset{f,\,t}{\text{minimize}}\ \Big\{\int_0^1 |f^{(m)}(s)|^2\,\mathrm{d}s : f^{(j-1)}(t_i)\in K_{ij},\ i{=}1{:}n,\ j{=}1{:}m\Big\},$$

over "smooth" parametrized curves $f:[0\,.\,.1]\longrightarrow \mathbb{R}^d$ that meet arbitrary sets K_{ij} in $\mathbb{R}^d$ at some **data sites** t_i in $[0\,.\,.1]$. We are motivated by the special case that $K_{ij}=B^-_{\varepsilon_{ij}}(z_{ij})$ is a closed ball of radius ε_{ij} and center z_{ij} in $\mathbb{R}^d$, as in [6]. Hence, we are extending the definition of near-interpolation in that paper to include the arbitrary sets here. In the last section of this paper, we show that certain sequences $((f^l,t^l))$ related to the minimization problem converge, even when some of the data sites coalesce (i.e., when $|t_i^l-t_{i+1}^l|\longrightarrow 0$ as $l\longrightarrow\infty$ for some i).

The setup is as follows. The curves are in the Sobolev space $X:=L_2^{(m)}([0\,.\,.1]\longrightarrow\mathbb{R}^d)$ of vector-valued functions $f:[0\,.\,.1]\longrightarrow\mathbb{R}^d$ for which $f^{(m)}$ is in the Lebesgue space $Y:=L_2([0\,.\,.1]\longrightarrow\mathbb{R}^d)$, $m\geq 1$. Since X embeds into $C^{m-1}([0\,.\,.1]\longrightarrow\mathbb{R}^d)$, then the data map

$$\Lambda_t : X\longrightarrow Z:=(\mathbb{R}^{m\times n})^d : f\longmapsto (\lambda_{ij}f:=f^{(j-1)}(t_i):\ i{=}1{:}n,\ j{=}1{:}m) \tag{1.1}$$

is continuous on X for fixed data sites

$$t\in\Delta_n^- := \{t\in\mathbb{R}^n : 0{=}t_1{\leq}t_2{\leq}\cdots{\leq}t_n{=}1\},$$

and the map

$$\Lambda : X\times\Delta_n^-\longrightarrow Z : (f,t)\longmapsto \Lambda_t f \tag{1.2}$$

is continuous on $X\times\Delta_n^-$. Note that the monotonicity of the data sites forces the curves to meet the sets K_{ij} in a particular order. With

$$K:=\mathop{\times}_{ij} K_{ij}\subset Z, \tag{1.3}$$

we say that (f,t) is "feasible" if $(f,t) \in \Lambda^{-1}K$, and that f is feasible for fixed t if $f \in \Lambda_t^{-1}K$. In particular, the feasibility of a particular curve f depends on the corresponding data sites t. With

$$J : X \longrightarrow \mathrm{IR} : f \longmapsto \int_0^1 |f^{(m)}(s)|^2 \,\mathrm{d}s,$$

we say that f is a **best near-interpolant** for *fixed* data sites t if it solves the minimization problem

$$(A) \qquad \underset{f\in X}{\text{minimize}} \ \{J(f) : \ \Lambda_t f \in K\},$$

and for *free* data sites if (f,t) solves

$$(B) \qquad \underset{f\in X,\ t\in\Delta_n^-}{\text{minimize}} \ \{J(f) : \ \Lambda_t f \in K\}.$$

In [6], optimality conditions are derived for the solutions to problems (A) and (B) when the sets are closed balls, and existence of solutions to problem (A) was verified as well. In particular, the solutions are polynomial splines, which is true even when the sets are not balls. The goal here is to derive conditions for the existence of solutions to problem (B).

Problem (B) is an extension of the problem of best "interpolation" by curves, for which existence and uniqueness are verified in [7] when $m = 2$ and $d = 1$, and existence are verified in [10] when $m \geq 1$ and $d \geq 1$. The proof of existence in [10] is similar to that given here for problem (B), with their condition of "asymptotically polynomial" data replaced by "near m-order" here. In [10], the data maps are of the form $f \longmapsto f(t_i)$, while we allow Hermite/Birkhoff-type interpolation $f \longmapsto f^{(j-1)}(t_i)$. As a special case of the setup here, our results apply to the problem of best interpolation in [10] when $K_{ij} = \mathrm{IR}^d$ for $j > 1$ and K_{i1} consists of a single point in IR^d.

In problem (B), it is possible to have repeated data sites for some i (i.e., $t_i{=}t_{i+1}$) when $K_{ij} \cap K_{i+1,j} \neq \emptyset$ for $j{=}1{:}m$. This is also possible in best interpolation when $z_{ij}{=}z_{i+1,j}$ for some i, a case that was excluded in [10]. When repeated data sites are allowed, the limits of sequences of curves in X may not be of the desired kind as data sites coalesce. In particular, spline curves tend to lose differentiability as data sites (and hence knots) coalesce, and moreover, the functional J may not be bounded on such sequences, as is shown in Section 5. To verify existence, we get around this difficulty by requiring only weak convergence and compactness, etc., of certain "minimizing sequences" in the infinite-dimensional Sobolev space X. However, in Section 5, we are able to show that certain sequences of spline curves do converge in norm, even as their data sites coalesce.

The organization of this paper is as follows. In Section 2, additional notation is given; in Section 3, the sets K_{ij} are briefly discussed; in Section 4, the existence theory is developed, leading to the main result in Theorem 4.11; and, in Section 5, certain sequences of solutions to problem (A) are shown to converge as their corresponding data sites converge, as stated in Corollary 5.4.

2. Additional notation. The following notation will be used in the remainder of this paper. The derivative operator on functions in X is denoted D, and in particular, D^m is m-fold differentiation; hence $D^m f = f^{(m)}$ and $D^m X = Y$. The kernel of D^m, denoted $\ker D^m$, is the linear space of those functions in X whose restriction to $[0\,..1]$ is a polynomial (curve) of order m, i.e., of degree $< m$.

The inner products on X, $\ker D^m$ and Y are

$$\langle f,g\rangle_X := \underbrace{\sum_{j=1}^{m} f^{(j-1)}(0)\cdot g^{(j-1)}(0)}_{\langle f,g\rangle_{\ker D^m}} \;+\; \underbrace{\int_0^1 f^{(m)}(s)\cdot g^{(m)}(s)\,\mathrm{d}s}_{\langle\ f^{(m)},g^{(m)}\rangle_Y},$$

respectively, with $u\cdot v$ denoting the standard dot product on $\mathbb{R}^d$, and

$$\langle (f,t),(g,\xi)\rangle_{X\times\mathbb{R}^n} := \langle f,g\rangle_X + \langle t,\xi\rangle_{\mathbb{R}^n} \tag{2.1}$$

is an inner product on $X\times\mathbb{R}^n$ with

$$\langle t,\xi\rangle_{\mathbb{R}^n} := \sum_{i=1}^{n} t_i\xi_i.$$

The inner product on Z is

$$\langle \alpha,\beta\rangle_Z := \sum_{i=1}^{n}\sum_{j=1}^{m} \alpha_{ij}\cdot\beta_{ij}.$$

These inner products induce norms in the usual way:

$$\|f\| := \langle f,f\rangle^{1/2}.$$

In particular, under their respective norms, X and $X\times\mathbb{R}^n$ are complete inner product spaces (i.e., Hilbert spaces). Note that

$$J(f) \;=\; \|f^{(m)}\|_Y^2.$$

Hence, J is the square of a seminorm on X.

Some of the results depend on certain "weights" $w\in\mathbb{R}^{m\times n}$. These act on sequences $\alpha=(\alpha_{ij})$ in Z by

$$w\,\alpha := (w_{ij}\alpha_{ij}\,:\;\; i{=}1{:}n,\;\; j{=}1{:}m),$$

with $\alpha_{ij}\in\mathbb{R}^d$ for each i and j, and they act on the maps Λ_t by

$$w\,\Lambda_t : X \longrightarrow Z : f \longmapsto w\,(\Lambda_t f).$$

Let

$$\mathrm{dist}(A,B) := \inf\{\|a-b\| : (a,b)\in A\times B\},$$

the distance between sets A and B in a normed linear space. In particular, $\mathrm{dist}(A,B)=0$ when $A\cap B\neq\emptyset$, and, for $a\in A$,

$$\mathrm{dist}(\{a\},B) = \inf\{\|a-b\| : b\in B\}.$$

Finally, let

$$\mathrm{diam}(A) := \sup\{|a-b| : a,b\in A\}$$

denote the **diameter** of A.

3. The sets K_{ij}. The primary motivation for this paper was to verify the existence of solutions to (B) when the sets K_{ij} are closed balls in $\mathbb{R}^d$, as in [6], however the results in this paper apply to arbitrary subsets of $\mathbb{R}^d$. Indeed, the only relevant issues in verifying existence are whether or not the sets K_{ij} are closed and bounded; the "shape" of the sets is not an issue. Hence, in this paper, the sets K_{ij} are arbitrary subsets of $\mathbb{R}^d$, which are bounded iff $\varepsilon_{ij} := \operatorname{diam}(K_{ij}) < \infty$. In [6], the "tolerances" ε_{ij} were radii of the closed balls, rather than diameters, as here. Note that, if K_{ij} consists of a single point, then $\varepsilon_{ij} = 0$, and the corresponding constraint forces interpolation. Note also that the set K as defined in (1.3) is closed in Z iff all of the sets K_{ij} are closed in $\mathbb{R}^d$.

It is expected that the sets K_{ij} of particular interest are closed and convex, such as balls or cones, or perhaps portions of lower dimensional manifolds. For example, to constrain the direction but not length of a tangent vector, we may require that

$$f'(t_i) = z_{i2} + \sum_{k=1}^{o} c_{i2k}\, v_{i2k}$$

with $c_{i2k} \geq 0$ for k=1:o. In this case, the corresponding set K_{ij} is a polyhedral cone with vertex z_{i2}, constrained by the directions $v_{i2k} \in \mathbb{R}^d$. If $o = 1$, then K_{ij} would be a half-line. Or, to approximate a smooth function g, the constraints may be of the form $\lambda_{ij} f = \lambda_{ij} g$.

In the case that the sets K_{ij} are closed balls, it was shown in [6] that the curves η that solve problems (A) and (B) are polynomial splines, when they exist. The same is true for arbitrary sets K_{ij}. Indeed, a solution (η, t) to (B) necessarily solves the problem of best *interpolation* to the data $z := \Lambda_t \eta$, the solution of which is a polynomial spline in the space $\$_{2m,t}$, as defined in [6] and in Section 5 of this paper.

4. Existence of the solutions to problem (B). A useful property of reflexive Banach spaces (such as Hilbert spaces) is that, while bounded sequences do not necessarily have (norm-)convergent subsequences, they do have weakly-convergent subsequences. Hence, a reasonable strategy towards verifying the existence of solutions to variational problems set in Hilbert spaces is to show that sequences in the feasible set on which the objective functional converges to its infimum are norm-bounded, and that the infimum is attained at the weak limits of such sequences. Here, we are interested in sequences $((f^l, t^l))$ in $\Lambda^{-1}K$ on which $J(f^l)$ converges to $\inf J(\Lambda^{-1}K)$, with K as defined in (1.3).

The convergence that we will establish is weak, sequential convergence. A sequence (f^l) in X **converges weakly** to $f \in X$ iff $\mu f^l \longrightarrow \mu f$ for all $\mu \in X^*$, written

$$f^l \xrightarrow{w} f.$$

Since X is a Hilbert space, if follows by the Riesz Representation Theorem that $f^l \xrightarrow{w} f$ in X iff $\langle f^l, g\rangle_X \longrightarrow \langle f, g\rangle_X$ for all $g \in X$. The result stated next concerns weak sequential convergence in the Hilbert space $X \times \mathbb{R}^n$ under the inner product (2.1).

Lemma 4.1. *Suppose that $(f^l, t^l) \xrightarrow{w} (f, t)$ in $X \times \mathbb{R}^n$ with (t^l) in Δ_n^-. Then $\Lambda_{t^l} f^l \longrightarrow \Lambda_t f$ in Z.*

Proof: First, note that $(f^l, t^l) \xrightarrow{w} (f, t)$ in $X \times \mathbb{R}^n$ iff $f^l \xrightarrow{w} f$ in X and $t^l \longrightarrow t$. In particular, $t^l \longrightarrow t \in \Delta_n^-$ since Δ_n^- is compact.

Assume that $d = 1$. By the Sobolev embedding theorem, X compactly embeds into the space $C^{(m-1)}([0..1] \longrightarrow \mathbb{R})$. Therefore, the sequences $(D^{j-1} f^l)$ converge uniformly to $D^{j-1} f$ on $[0..1]$ for $j{=}1{:}m$, implying by [9: Theorem 7.24] that these sequences are equicontinuous. Hence, $\lambda_{ij}^l f^l \longrightarrow \lambda_{ij} f$ for $i{=}1{:}n$ and $j{=}1{:}m$, and so $\Lambda_{t^l} f^l \longrightarrow \Lambda_t f$ in Z.

When $d{>}1$, it follows as above that $\Lambda_{t^l} f_r^l \longrightarrow \Lambda_t f_r$ in Z for $r{=}1{:}d$, with $f_1^l, \ldots, f_d^l$ the components of the functions $f \in X$. Hence, again $\Lambda_{t^l} f^l \longrightarrow \Lambda_t f \in Z$. □

In the next lemma it is shown that the feasible set $\Lambda^{-1} K$ is **weakly sequentially closed** in $X \times \mathbb{R}^n$, meaning that if $((f^l, t^l))$ is a sequence in $\Lambda^{-1} K$ such that $(f^l, t^l) \xrightarrow{w} (f, t)$ for some $(f, t) \in X \times \mathbb{R}^n$, then $(f, t) \in \Lambda^{-1} K$.

Lemma 4.2. *The set $\Lambda^{-1} K$ is weakly sequentially closed in $X \times \mathbb{R}^n$ when K is closed in Z.*

Proof: In the insignificant case that $\Lambda^{-1} K = \emptyset$, the result follows trivially. Hence, assume that $\Lambda^{-1} K \neq \emptyset$. Let $((f^l, t^l))$ be a sequence in $\Lambda^{-1} K$ that converges weakly to some (f, t) in $X \times \mathbb{R}^n$. In particular, $t^l \longrightarrow t \in \Delta_n^-$ since Δ_n^- is compact. Moreover, $\Lambda_{t^l} f^l$ is in K for each l, and $\Lambda_{t^l} f^l \longrightarrow \Lambda_t f$ by Lemma 4.1, and so

$$\operatorname{dist}(\{\Lambda_t f\}, K) \leq \|\Lambda_t f - \Lambda_{t^l} f^l\|_Z \longrightarrow 0$$

as $l \longrightarrow \infty$. Therefore, $\Lambda_t f$ is in K since K is closed, and so $(f, t) \in \Lambda^{-1} K$. □

Our next goal is determine conditions for which sequences $((f^l, t^l))$ in $\Lambda^{-1} K$ are bounded. Since (t^l) is in the compact set Δ_n^-, it is trivially bounded in $\mathbb{R}^n$, hence it remains to have (f^l) bounded in X. For this, it is convenient to represent elements f of X as

$$f = Pf + Rf := \sum_{j=1}^{m} f^{(j-1)}(0)\, \phi_{j-1} \;+\; \int_0^1 \phi_{m-1}(\cdot - s)\, f^{(m)}(s)\, \mathrm{d}s \tag{4.3}$$

in the orthogonal sum decomposition $(\ker D^m) \oplus (\ker D^m)^\perp$ of X, with

$$\phi_j \;:=\; [\,\cdot\,)_+^j / j! \;:=\; \max(0, \cdot)^j / j!, \tag{4.4}$$

the **normalized truncated power** function. Note that ϕ_0 is left continuous since $\phi_0(0) = [\,0\,)_+^0 = 1$, and that $D^k \phi_j = \phi_{j-k}$ when $j \geq k$. Moreover $D^m Rf = f^{(m)}$, and so

$$\|f\|_X^2 \;=\; \|Pf\|_{\ker D^m}^2 + \|D^m Rf\|_Y^2 \;=\; \|Pf\|_{\ker D^m}^2 + \|D^m f\|_Y^2.$$

Then, since $J(f^l) = \|D^m f^l\|_Y^2$, it follows that (f^l) is bounded when (Pf^l) and $(J(f^l))$ are bounded.

Lemma 4.5. *Suppose that $\Lambda^{-1} K \neq \emptyset$. Let $((f^l, t^l))$ be a sequence in $\Lambda^{-1} K$ on which J is bounded. Suppose that $t^l \longrightarrow t \in \Delta_n^-$ and that $\ker D^m \cap \ker w\Lambda_t {=} \{0\}$ with $w_{ij} := 1$ if $\varepsilon_{ij} := \operatorname{diam}(K_{ij}) < \infty$ and $w_{ij} := 0$ otherwise. Then, (Pf^l) is bounded.*

Proof: We first show that the sequence $(w\Lambda_{t^l} Pf^l)$ is bounded, whether or not $\ker D^m \cap \ker w\Lambda_t {=} \{0\}$. For this, we have by (4.3) that since $Pf^l = f^l - Rf^l$, and since $\lambda_{ij}^l f^l \in K_{ij}$ for each i and j, it follows that

$$|\lambda_{ij}^l Pf^l| \;=\; |\lambda_{ij}^l f^l - \lambda_{ij}^l Rf^l| \;\leq\; \operatorname{dist}(\{0\}, K_{ij}) + \varepsilon_{ij} + |\lambda_{ij}^l Rf^l|.$$

By Holder's inequality, and since $|\phi_{m-j}| \le 1$ on $[0\,.\,.1]$, it follows that

$$\begin{aligned}|\lambda^l_{ij} R f^l|^2 &= |\int_0^1 \phi_{m-j}(t^l_i - s)\, D^m f^l(s)\,\mathrm{d}s|^2 \\ &\le \int_0^1 |\phi_{m-j}(t^l_i - s)|^2\,\mathrm{d}s \int_0^1 |D^m f^l(s)|^2\,\mathrm{d}s \\ &\le \int_0^1 |D^m f^l(s)|^2\,\mathrm{d}s \;=\; J(f^l).\end{aligned}$$

Therefore,

$$|\lambda^l_{ij} P f^l| \;\le\; \mathrm{dist}(\{0\}, K_{ij}) + \varepsilon_{ij} + J(f^l)^{1/2},$$

and so $(\lambda^l_{ij} P f^l)$ is bounded when $(J(f^l))$ is bounded and $\varepsilon_{ij} < \infty$, i.e., when $w_{ij} \ne 0$. Hence, $(w\Lambda_{t^l} P f^l)$ is bounded in this case.

Since $\ker D^m \cap \ker\, w\Lambda_t{=}\{0\}$, then $\|w\Lambda_t \cdot\|_Z$ is a norm on $\ker D^m$. Further, since $w\Lambda_{t^l}$ converges to $w\Lambda_t$ as $t^l \longrightarrow t$, then it does so in norm on the finite-dimensional linear space $\ker D^m$. Hence, there is some neighborhood O of $t \in \Delta^-_n$ on which $\sup_{t^l \in O} \|w\Lambda_{t^l} - w\Lambda_t\| < .5$ (with the norm the map-norm on $\ker D^m$). Since t^l converges to t, it follows that, for sufficiently large l,

$$\|w\Lambda_t P f^l\|_Z \le \|(w\Lambda_t - w\Lambda_{t^l}) P f^l\|_Z + \|w\Lambda_{t^l} P f^l\|_Z \le .5\|w\Lambda_t P f^l\|_Z + \|w\Lambda_{t^l} P f^l\|_Z,$$

hence

$$\|w\Lambda_t P f^l\|_Z \le 2\|w\Lambda_{t^l} P f^l\|_Z$$

for all large l. Therefore $(w\Lambda_t P f^l)$ is bounded in Z, and so $(P f^l)$ is bounded in $\ker D^m$. □

The conditions for the boundedness of $(P f^l)$ in Lemma 4.5 depend on certain limiting data sites t. Since such data sites are generally not known until a solution to problem (B) has been determined, we would like a condition that does not involve t, similar to what was done in [10] for the problem of best interpolation.

Definition 4.6. *We say that K is* ***m*-order** *for fixed $t \in \Delta^-_n$ if*

$$\mathrm{dist}(\Lambda_t \ker D^m, K) \;=\; 0,$$

and **near *m*-order** *if*

$$\inf_{t \in \Delta^-_n} \mathrm{dist}(\Lambda_t \ker D^m, K) \;=\; 0.$$

The property of near m-order reduces to what is termed "asymptotically polynomial of order m" in [10] when $\varepsilon = 0$ and $\Lambda_t : f \longmapsto (f(t_i))$. To handle the derivative maps in near-interpolation, we pay attention only to those constraints ij for which $\varepsilon_{ik} := \mathrm{diam}(K_{ik}) < \infty$ for $k{=}1{:}j$, a "Hermite"-subset of the constraints. For this, let

$$\varepsilon^H_{ij} := \begin{cases} \varepsilon_{ij}, & \text{if } \varepsilon_{i1}, \ldots, \varepsilon_{ij} < \infty; \\ \infty, & \text{otherwise.}\end{cases} \tag{4.7}$$

Then, with $K^H_{ij} := K_{ij}$ if $\varepsilon^H_{ij} < \infty$ and $\mathrm{I\!R}^d$ otherwise, let $K^H := \times_{ij} K^H_{ij}$.

Lemma 4.8. *Suppose that $\Lambda^{-1}K \neq \emptyset$. Let $((f^l, t^l))$ be a sequence in $\Lambda^{-1}K$ on which J is bounded. If K^H is not near m-order, then (Pf^l) is bounded.*

Proof: Since (t^l) is bounded in $\mathbb{R}^n$, it suffices to establish the result on an arbitrary subsequence of $((f^l, t^l))$ on which (t^l) converges. Hence, assume from here on that $t^l \longrightarrow t \in \Delta_n^-$. Let $w_{ij} := 1$ whenever $\varepsilon_{ij} < \infty$ and 0 otherwise, and let $w_{ij}^H := 1$ whenever $\varepsilon_{ij}^H < \infty$ and 0 otherwise. Since $w_{ij}^H \neq 0$ implies $w_{ij} \neq 0$, then $\ker D^m \cap \ker\, w^H\Lambda_t = \{0\}$ implies $\ker D^m \cap \ker\, w\Lambda_t = \{0\}$. Hence, by Lemma 4.5, (Pf^l) is bounded when $\ker D^m \cap \ker\, w^H\Lambda_t = \{0\}$, and so our goal is to show that $\ker D^m \cap \ker\, w^H\Lambda_t = \{0\}$ when K^H is not near m-order.

By way of contradiction, we will show that K^H is near m-order when $\ker D^m \cap \ker\, w^H\Lambda_t \neq \{0\}$. To do so, we proceed like in [10] for the problem of best interpolation, and construct a sequence (p^l) in $\ker D^m$ such that $\operatorname{dist}((\Lambda_{t^l} p^l), K^H) \longrightarrow 0$. Recalling by (4.3) that $f^l = Pf^l + Rf^l$, we expect that a sequence (p^l) with $p^l := Pf^l + p$ for some $p \in \ker D^m$ such that $\|w^H\, \Lambda_{t^l}(Rf^l - p)\|_Z \longrightarrow 0$ will do the job, while passing to a subsequence if necessary.

To find such a p, we will first obtain a convergent subsequence and corresponding limit point of $(w^H\, \Lambda_{t^l} Rf^l)$. For this, note that since $J(f^l) = \|D^m f^l\|_Y^2$ is bounded and since Y is a reflexive Banach space, we may assume on passing to a subsequence that $(D^m f^l)$ converges weakly to some point y in Y (see [11: Theorem 10.6]), i.e., $\langle D^m f^l, \cdot\rangle_Y \longrightarrow \langle y, \cdot\rangle_Y$. Let

$$g := \int_0^1 \phi_{m-1}(\cdot - s)\, y(s)\, \mathrm{d}s.$$

Since Rf^l and g are in $(\ker D^m)^\perp$, and since $D^m Rf^l = D^m f^l$ and $D^m g = y$, then

$$\langle Rf^l, \cdot\rangle_X = \langle D^m f^l, D^m\cdot\rangle_Y \longrightarrow \langle y, D^m\cdot\rangle_Y = \langle D^m g, D^m\cdot\rangle_Y = \langle g, \cdot\rangle_X$$

on X. Hence,

$$Rf^l \xrightarrow{w} g$$

in X, which by Lemma 4.1 implies that

$$\Lambda_{t^l} Rf^l \longrightarrow \Lambda_t g.$$

Next, we will first show that $\operatorname{ran} w^H\Lambda_t = w^H\Lambda_t \ker D^m$, which will then allow us to select $p \in \ker D^m$ such that $w^H\Lambda_t p = w^H\Lambda_t g$. To account for possible repetitions of the data sites, let

$$t' := (t_i : t_i \neq t_{i-1}) \tag{4.9}$$

the maximal strictly increasing subsequence of t, and with $n' := \#t'$, let

$$w'^H_{ij} := \sum_{t_k = t'_i} w^H_{kj}$$

for i=1:n'. Then, $w'^H_{ij} = 0$ iff $w^H_{kj} = 0$ for all k such that $t_k = t'_i$, and so $\ker D^m \cap \ker\, w^H\Lambda_t \neq \{0\}$ implies $\ker D^m \cap \ker\, w'^H\Lambda_{t'} \neq \{0\}$. Moreover, since $\ker D^k \cap \ker\, w'^H\Lambda_{t'} = \{0\}$ with $k := \#(w'^H_{ij} > 0)$ (a "Hermite system" in [3]), it follows that $k < m$. Consequently, for this k,

$$k \geq \dim(\operatorname{ran} w'^H\, \Lambda_{t'}) \geq \dim(w'^H\, \Lambda_{t'} \ker D^m) \geq \dim(w'^H\, \Lambda_{t'} \ker D^k) = k,$$

and, since $\ker D^m \subset X$, it follows that $\operatorname{ran} w'^H \Lambda_{t'} = w'^H \Lambda_{t'} \ker D^m$. Moreover, since (λ_{ij}) is the sequence (λ'_{ij}) except for possible repetitions, it follows that $\operatorname{ran} w^H \Lambda_t = w^H \Lambda_t \ker D^m$. In particular, there exists $p \in \ker D^m$ such that $w^H \Lambda_t p = w^H \Lambda_t g$.

So far, we have a subsequence of $((f^l, t^l))$ on which $t^l \longrightarrow t$ in Δ_n^- and $\Lambda_{t^l} R f^l \longrightarrow \Lambda_t g$ in Y, and we have $p \in \ker D^m$ such that $w^H \Lambda_t p = w^H \Lambda_t g$. Then, for $p^l := P f^l + p = f^l - R f^l + p$,

$$\begin{aligned}
\operatorname{dist}(\{\Lambda_{t^l} p^l\}, K^H) &= \operatorname{dist}(\{w^H \Lambda_{t^l} p^l\}, K^H) \\
&= \operatorname{dist}(\{w^H \Lambda_{t^l} (f^l - R f^l + p)\}, K^H) \\
&\le \operatorname{dist}(\{w^H \Lambda_{t^l} f^l\}, K^H) + \|w^H \Lambda_{t^l} (R f^l - p)\|_Z \\
&\le \operatorname{dist}(\{w^H \Lambda_{t^l} f^l\}, K^H) + \|w^H (\Lambda_{t^l} R f^l - \Lambda_t g)\|_Z + \|w^H (\Lambda_t g - \Lambda_{t^l} p)\|_Z,
\end{aligned}$$

which goes to 0 as $l \longrightarrow \infty$ since $\Lambda_{t^l} f^l$ is in $K \subset K^H$ for each l and since $\Lambda_{t^l} R f^l \longrightarrow \Lambda_t g$ and $w^H \Lambda_{t^l} p \longrightarrow w^H \Lambda_t p = w^H \Lambda_t g$ as $l \longrightarrow \infty$. Therefore, K^H is near m-order.

To conclude, we have shown that $\Lambda^{-1} K$ is near m-order when $\ker D^m \cap \ker w^H \Lambda_t \neq \{0\}$. Conversely, if $\Lambda^{-1} K$ is not near m-order, then $\ker D^m \cap \ker w^H \Lambda_t = \{0\}$, from which it follows by Lemma 4.5 that $(P f^l)$ is bounded. □

We are now in the position to show that sequences in $\Lambda^{-1} K$ on which J is bounded have subsequences that converge weakly in $\Lambda^{-1} K$. Before establishing existence, we need to show that the value of the objective functional is "reduced" at these weak limits. For this, we need that J is "weakly lower semi-continuous", as verified in the next lemma.

Lemma 4.10. *Suppose that $f^l \xrightarrow{w} f$ in X. Then, $J(f) \le \varliminf J(f^l)$. That is, J is weakly sequentially lower semi-continuous on X.*

Proof: Let $y \in Y$ and $g := \int_0^1 \phi_{m-1}(\cdot - s)\, y(s)\, \mathrm{d}s$. Then, $g \in (\ker D^m)^\perp \subset X$ and $D^m g = y$, and so

$$\langle D^m f^l, y \rangle_Y = \langle D^m f^l, D^m g \rangle_Y = \langle f^l, g \rangle_X \longrightarrow \langle f, g \rangle_X = \langle D^m f, D^m g \rangle_Y = \langle D^m f, y \rangle_Y.$$

Since y was arbitrary, it follows that $D^m f^l \xrightarrow{w} D^m f$ in Y.

By [4: Theorem 4.10.7], $\|\cdot\|_Y$ is weakly sequentially lower semicontinuous on Y when $d = 1$, which generalizes by linearity to the case $d > 1$. Therefore,

$$J(f) = \|D^m f\|_Y^2 \le \varliminf \|D^m f^l\|_Y^2 = \varliminf J(f^l).$$

□

The main existence result is stated next. For this, we say that $((f^l, t^l))$ is a "minimizing sequence" for problem (B) if $(f^l, t^l) \in \Lambda^{-1} K$ for each l and $J(f^l) \longrightarrow \inf J(\Lambda^{-1} K)$.

Theorem 4.11. *Suppose that K is closed in Z and $\Lambda^{-1} K \neq \emptyset$.*

(i) Let ε^H as in (4.7). There exists a solution to problem (B) if K is m-order or if K^H is not near m-order, and only if K is m-order or not near m-order.

(ii) Suppose that $((f^l, t^l))$ is a minimizing sequence for (B) such that $t^l \longrightarrow t \in \Delta_n^-$, and suppose that $\ker D^m \cap \ker w\Lambda_t = \{0\}$ with $w \in \mathbb{R}^{m\times n}$ such that $w_{ij} := 1$ when $\varepsilon_{ij} < \infty$ and 0 otherwise. Then, solutions to problem (B) exist.

Proof: We first establish part (i). Suppose that K is m-order. Then, $J(f) = 0$ for some $(f,t) \in \Lambda^{-1}K$, and so minimizers trivially exist.

Suppose that K^H is not near m-order. Let $((f^l, t^l))$ be a minimizing sequence for (B). In particular, $J(f^l) = \|D^m f^l\|_Y^2$ is bounded, and by Lemma 4.8, (Pf^l) is also bounded. Then, (f^l) is bounded in X since

$$\|f^l\|_X^2 = \|Pf^l\|_{\ker D^m}^2 + \|D^m f^l\|_Y^2,$$

and (t^l) is bounded in $\mathbb{R}^n$ since $t^l \in \Delta_n^-$ for each l, and so $((f^l, t^l))$ is bounded in $X \times \mathbb{R}^n$. Since $X \times \mathbb{R}^n$ is a reflexive Banach space, we may assume on passing to a convergent subsequence that $(f^l, t^l) \xrightarrow{w} (\eta, t)$ in $X \times \mathbb{R}^n$ (see [11: Theorem 10.6]), and by Lemma 4.2, $(\eta, t) \in \Lambda^{-1}K$. Since $f^l \xrightarrow{w} \eta$, it follows by Lemma 4.10 that $J(\eta) \leq \underline{\lim}\, J(f^l)$, and since (f^l, t^l) is a minimizing sequence for (B) then $J(f^l) \longrightarrow \inf J(\Lambda^{-1}K)$. Therefore, (η, t) solves problem (B).

For necessity, $\inf J(\Lambda^{-1}K) = 0$ when K is near m-order, in which case minimizers are in $\ker D^m = \ker J$ when they exist. Hence, K is either m-order or not near m-order when minimizers exist.

For part (ii), boundedness follows from Lemma 4.5, and the rest of the proof proceeds as in (i). □

As a special case, Theorem 4.11 applies to the problem of best interpolation by curves studied in [10] when $K_{i1} = \{z_{i1}\}$ for some points $z_{i1} \in \mathbb{R}^d$, and $K_{ij} = \mathbb{R}^d$ for $j>1$, $i=1{:}n$. Moreover, this theorem allows for the possibility of repeated data sites (where $t_i \neq t_{i+1}$ for some i), a situation that was excluded in [10] by the assumption $z_{i1} \cap z_{i+1,1} = \emptyset$ for $i=1{:}n-1$. In the next result, a similar condition is stated that excludes the possibility of repeated data sites in near-interpolation.

Corollary 4.12. *Suppose that $(f,t) \in \Lambda^{-1}K \neq \emptyset$. Assume that, for some i and j, $K_{ij} \cap K_{i+1,j} = \emptyset$. Then, $t_i < t_{i+1}$ for this i.*

Proof: Let i and j as in the hypothesis. Since K_{ij} and $K_{i+1,j}$ are closed and disjoint, then $\mathrm{dist}(K_{ij}, K_{i+1,j}) > \epsilon$ for some $\epsilon > 0$. Since $\lambda_{ij} f \in K_{ij}$ and $\lambda_{i+1,j} f \in K_{i+1,j}$, it follows that $|\lambda_{ij} f - \lambda_{i+1,j} f| > \epsilon$, and so $t_i \neq t_{i+1}$. □

5. Convergence. To verify existence in the previous section, certain sequences $((f^l, t^l))$ in $\Lambda^{-1}K$ were shown to be weakly, sequentially convergent. In this section, sufficient conditions are given for which such sequences converge in norm (in X), even as data sites coalesce (i.e., $|t_{i+1}^l - t_i^l| \longrightarrow 0$ for some i). It is expected that the results can be applied or extended to other problems in nonlinear best approximation.

As discussed in Section 3, and as in [6], the minimizers to problems (A) and (B) are polynomial splines. More precisely, if η solves (A) at fixed t or if (η, t) solves (B), then, with

$$t' := (t_i : t_i \neq t_{i-1}) \quad \text{(as in (4.9))},$$

η is in the linear space

$$\$_{2m,t'} := \$_{2m,t'}([0\,..1] \longrightarrow \mathrm{IR}^d)$$

of piecewise polynomial *curves* on $[0\,..1]$, into IR^d, of order $2m$ (degree $2m-1$), and with $m-1$ continuous derivatives at the breakpoints t'_i in t'. With respect to t and t', let

$$\begin{aligned} n' &:= \#t', \\ Z' &:= \mathrm{IR}^{m\times n'}, \\ S &:= \{\alpha \in Z : \alpha_{ij} = \alpha_{i-1,j} \text{ if } t_i = t_{i-1}\}, \\ \pi &: S \longrightarrow Z' : \alpha \longmapsto (\alpha_{ij} : t_i \neq t_{i-1}). \end{aligned} \tag{5.1}$$

In particular, π is a bijection on S, and so π^{-1} is well-defined. We do a similar construction for each l. That is, $t^{l'}$, n'_l, Z'_l, S_l, π_l and $\$_{2m,t^{l'}}$ are defined as above, but with respect to t^l.

Lemma 5.2. *Suppose that $t^l \longrightarrow t$ in Δ_n^-, and that, whenever $\#(t_k : t_k = t_i) > 1$ for some i, the intersections $\cap_{t_k=t_i} K_{kj}$ have a nonempty interior for j=1:m. Let $f \in \$_{2m,t'} \cap \Lambda_{t'}^{-1}K$. Then, there exists a sequence (f^l) in X such that $(f^l, t^l) \in \Lambda^{-1}K$ for each l, $J(f^l) \longrightarrow J(f)$, and $f^l \longrightarrow f$ (in norm) in X.*

Proof: Let $V_{t'}$ be the basis-map for $\$_{2m,t'}$ that is dual to $\Lambda_{t'}$, and let $V_{t^{l'}}$ be the basis-map for $\$_{2m,t^{l'}}$ that is dual to $\Lambda_{t^{l'}}$. Since $f \in \$_{2m,t'}$, we let $\alpha \in S$ such that $f = V_{t'}\pi(\alpha)$, with π as defined in (5.1). Our goal is to choose coefficients α^l in S_l such that $f^l = V_{t^{l'}}\pi_l(\alpha^l)$ is in $\Lambda^{-1}K$ for each l and converges to f in norm. In particular, the coefficients are chosen such that $\alpha^l \longrightarrow \alpha$.

Consider two cases. In the first case, suppose that $t_{i-1} \neq t_i \neq t_{i+1}$ for some i. Then $t^l_{i-1} \neq t^l_i \neq t^l_{i+1}$ for large enough l since $t^l \longrightarrow t$, and we set $\alpha^l_{ij} := \alpha_{ij}$ for j=1:m for large l.

In the second case, suppose that $t_{i-1} \neq t_i = \cdots = t_{i+k} \neq t_{i+k+1}$ for some i and k. By hypothesis, the interior of $\Omega := K_{ij} \cap \cdots \cap K_{i+k,j}$ is not empty. Then $t^l_{i-1} \neq t^l_i$ and $t^l_{i+k} \neq t^l_{i+k+1}$ for large enough l. We may assume, without loss of generality, that t^l is ordered such that $|t^{l+1}_{i+k} - t^{l+1}_i| \leq |t^l_{i+k} - t^l_i|$. Let

$$p_i^{\hat{l}l} := \sum_{j=1}^{m} \alpha^{\hat{l}}_{ij} (\cdot - t^l_i)^{j-1} \in \ker D^m$$

with $\alpha^{\hat{l}}_{ij}$ chosen in the interior of Ω, and such that $\alpha^{\hat{l}}_{ij} \longrightarrow \alpha_{ij}$ as $\hat{l} \longrightarrow \infty$. Let $l = l(\hat{l})$ be the first index l such that $l(\hat{l}) > l(\hat{l}-1)$ when $\hat{l} > 1$ and $D^{j-1} p_i^{\hat{l}l}(s)$ is in Ω for all $s \in [t^l_i \,..\, t^l_{i+k}]$ and j=1:m. Such an $l = l(\hat{l})$ exists for each $\hat{l}$ since $|t^l_{i+k} - t^l_i| \longrightarrow 0$ and since the Taylor polynomial coefficients $\alpha^{\hat{l}}_{ij}$ are in the *interior* of Ω. Moreover, $l(\hat{l}) \longrightarrow \infty$ as $\hat{l} \longrightarrow \infty$, strictly monotonically. Also, let

$$\hat{l}(l) := \{\hat{l} : l(\hat{l}) \leq l < l(\hat{l}+1)\}.$$

Then, $\hat{l}(l) \longrightarrow \infty$ as $l \longrightarrow \infty$, not necessarily strictly. Finally, let $\alpha^l_{oj} := D^{j-1}p_i^{\hat{l}(l),l}(t^l_o)$ for $o{=}i{:}i{+}k$ and $j{=}1{:}m$. In particular, $\alpha^l_{oj} = \alpha^l_{o+1,j}$ when $t^l_o = t^l_{o+1}$.

We have constructed the tail of a sequence $((\alpha^l, t^l))$ of coefficients and data sites such that $\alpha^l \in S_l \cap K$ for large l and $\alpha^l \longrightarrow \alpha$. On this tail let $f^l := V_{t^{l\prime}}\pi_l(\alpha^l)$, and on the head of the sequence (a finite set) let (α^l) be the coefficients corresponding to any feasible function f^l from $\$_{2m,t^{\prime l}}$. By construction, if $t_i = t_{i+1}$, then the restriction of f^l to the interval $(t^l_i \,..\, t^l_{i+1})$ is in $\ker D^m$ for large l, in which case

$$\int_{t^l_i}^{t^l_{i+1}} |D^m f^l|^2 = 0.$$

On the other hand, if $t_i \neq t_{i+1}$, then the restriction of $D^m f^l$ to the intervals $(t^l_i..t^l_{i+1})$ are polynomials curves in $\ker D^{m-1}$ that converge uniformly (component-wise) to $D^m f$ on $(t_i \,..\, t_{i+1})$, and so

$$\int_{t^l_i}^{t^l_{i+1}} |D^m f^l|^2 \longrightarrow \int_{t_i}^{t_{i+1}} |D^m f|^2.$$

Therefore,

$$J(f^l) = \sum_{i=1}^{n-1} \int_{t^l_i}^{t^l_{i+1}} \|D^m f^l\|^2 = \sum_{t_i \neq t_{i+1}} \int_{t^l_i}^{t^l_{i+1}} \|D^m f^l\|^2$$
$$\longrightarrow \sum_{t_i \neq t_{i+1}} \int_{t_i}^{t_{i+1}} \|D^m f\|^2 = \sum_{i=1}^{n-1} \int_{t_i}^{t_{i+1}} \|D^m f\|^2 = J(f),$$

and so $J(f^l) \longrightarrow J(f)$. Moreover, since $\Lambda_{t^l} f^l = \alpha^l \longrightarrow \alpha = \Lambda_t f$, then $D^{j-1}f^l(0) \longrightarrow D^{j-1}f(0)$ for $j{=}1{:}m$, and so $\|f^l - f\|_{\ker D^m} \longrightarrow 0$. Hence,

$$\|f^l - f\|^2_X = \|f^l - f\|^2_{\ker D^m} + \|D^m f^l - D^m f\|^2_Y \longrightarrow 0.$$

□

The next proposition is based mainly on results from Section 4. The sequence (f^l) constructed in Lemma 5.2 will be used in conjunction with this proposition to prove Corollary 5.4.

Proposition 5.3. *Suppose that K is closed in Z and not near m-order, $\Lambda^{-1}K \neq \emptyset$, $t^l \longrightarrow t$ in Δ^-_n, η^l solves problem (A) with respect to t^l for each l, and η solves (A) with respect to t. Assume, moreover, that η is the unique solution to (A) for these data sites t. Suppose that there exists a sequence $((f^l, t^l))$ in $\Lambda^{-1}K$ such that $J(f^l) \longrightarrow J(\eta)$. Then $J(\eta^l) \longrightarrow J(\eta)$ and $\eta^l \longrightarrow \eta$ (in norm) in X.*

Proof: Since η^l solves problem (A) with respect to t^l, then $J(\eta^l) \leq J(f^l)$. In particular, $(J(\eta^l))$ is bounded, and so, by Lemma 4.8, $(P\eta^l)$ is also bounded. Therefore, (η^l) is bounded in X. Since X is a reflexive Banach space, (η^l) has a weakly convergent subsequence. After passing to such a subsequence, we may assume that $(\eta^l, t^l) \xrightarrow{w} (f, t) \in X$ for some $f \in X$. For this f, $(f, t) \in \Lambda^{-1}K$ by Lemma 4.2, and by Lemma 4.10 $J(f) \leq \underline{\lim}\, J(\eta^l)$. Now, since η solves (A) for the data sites t, then $J(\eta) \leq J(f)$, and similarly, $J(\eta^l) \leq J(f^l)$. Therefore,

$$J(\eta) = \lim J(f^l) \geq \underline{\lim}\, J(\eta^l) \geq J(f) \geq J(\eta),$$

and so $J(\eta^l) \longrightarrow J(\eta)$. Moreover, since η *uniquely* solves (A) for fixed t, then $f = \eta$ for all such weak limits, and so the original sequence (η^l) converges weakly to η.

Since $\eta^l \xrightarrow{w} \eta$ in X, then $D^m\eta^l \xrightarrow{w} D^m\eta$ in Y, and since $J(\eta^l) \longrightarrow J(\eta)$, it follows that

$$\begin{aligned}\|D^m\eta^l - D^m\eta\|_Y^2 &= \langle D^m\eta^l - D^m\eta, D^m\eta^l - D^m\eta\rangle_Y \\ &= \langle D^m\eta^l, D^m\eta^l\rangle_Y - 2\langle D^m\eta^l, D^m\eta\rangle_Y + \langle D^m\eta, D^m\eta\rangle_Y \\ &= J(\eta^l) - 2\langle D^m\eta^l, D^m\eta\rangle_Y + J(\eta) \\ &\longrightarrow J(\eta) - 2\langle D^m\eta, D^m\eta\rangle_Y + J(\eta) \\ &= 0.\end{aligned}$$

Moreover, by Lemma 4.1 $\Lambda_{t^l}\eta^l \longrightarrow \Lambda_t\eta$, which implies that $D^{j-1}\eta^l(0) \longrightarrow D^{j-1}\eta(0)$ for j=1:m, and so $\|\eta^l - \eta\|_{\ker D^m} \longrightarrow 0$. Therefore,

$$\|\eta^l - \eta\|_X^2 = \|\eta^l - \eta\|_{\ker D^m}^2 + \|D^m\eta^l - D^m\eta\|_Y^2 \longrightarrow 0.$$

That is, η^l converges (in norm) to η in X. □

Next, we state the main convergence result, as a corollary to Lemma 5.2 and Proposition 5.3.

Corollary 5.4. *Suppose that $t^l \longrightarrow t$ in Δ_n^-, and that, whenever $\#(t_k : t_k = t_i) > 1$ for some i, the intersections $\cap_{t_k=t_i} K_{kj}$ have a nonempty interior for j=1:m. Suppose that η^l solves problem (A) with respect to t^l for each l, and η uniquely solves (A) with respect to t. Then $J(\eta^l) \longrightarrow J(\eta)$ and $\eta^l \longrightarrow \eta$ (in norm) in X.*

Proof: Since η solves problem (A) for fixed t, then $\eta \in \$_{2m,t'}$. By Lemma 5.2, there exists a sequence (f^l) in X such that $(f^l, t^l) \in \Lambda^{-1}K$ for each l and $J(f^l) \longrightarrow J(\eta)$. By Proposition 5.3, $J(\eta^l) \longrightarrow J(\eta)$ and $\eta^l \longrightarrow \eta$ (in norm) in X. □

Acknowledgement. I wish to thank Carl de Boor for his input on the writing and on some of the results in this paper.

REFERENCES

[1] de Boor, C. (1969), "On the approximation by γ-polynomials", in *Approximation with Special Emphasis on Spline Functions* (I. J. Schoenberg, ed), Academic Press (New York), 157–183.
[2] Braess, D. (1986), *Nonlinear Approximation Theory*, Springer Verlag (Berlin).
[3] Ferguson, D. (1969), "The question of uniqueness for G. D. Birkhoff interpolation problems", *J. Approx. Theory* **2**, 1–28.
[4] Friedman, A. (1982), *Foundations of modern analysis*, Dover publications.
[5] Hobby, C. R., and J. R. Rice (1967), "Approximation from a curve of functions", *Arch. Rational Mech. Anal.* **24**, 91–106.
[6] Kersey, S. (1999), "Best near-interpolation by curves: optimality conditions", *submitted* **xx**, xx–xx.
[7] Marin, S. P. (1984), "An approach to data parametrization in parametric cubic spline interpolation problems", *J. Approx. Theory* **41**, 64–86.

[8] Rice, J. R. (1969), *The Approximation of Functions Vol. II*, Addison-Wesley (Reading, Mass.).
[9] Rudin, W. (1976), *Principles of mathematical analysis*, 3nd ed., McGraw-Hill.
[10] Scherer, K., and P. W. Smith (1989), "Existence of best parametric interpolation by curves", *SIAM J. Math. Anal.* **20 (1)**, 160–168.
[11] Taylor, A. E., and D. C. Lay (1980), *Introduction to functional analysis (second edition)*, John Wiley and Sons, Inc..
[12] Zeidler, E. (1985), *Nonlinear functional analysis and its applications. III*, Springer Verlag.