MATHEMATICS OF COMPUTATION
Volume 65, Number 214
April 1996, Pages 647–660

UNICITY IN PIECEWISE POLYNOMIAL L^1-APPROXIMATION VIA AN ALGORITHM

R. C. GAYLE AND J. M. WOLFE

ABSTRACT. Our main result shows that certain generalized convex functions on a real interval possess a unique best L^1 approximation from the family of piecewise polynomial functions of fixed degree with varying knots. This result was anticipated by Kioustelidis in [11]; however the proof given there is nonconstructive and uses topological degree as the primary tool, in a fashion similar to the proof the comparable result for the L^2 case in [5]. By contrast, the proof given here proceeds by demonstrating the global convergence of an algorithm to calculate a best approximation over the domain of all possible knot vectors. The proof uses the contraction mapping theorem to simultaneously establish convergence and uniqueness. This algorithm was suggested by Kioustelidis [10]. In addition, an asymptotic uniqueness result and a nonuniqueness result are indicated, which analogize known results in the L^2 case.

1. INTRODUCTION

In this paper we examine the question of the unicity of best L^1 approximations of $n+1$ times continuously differentiable functions. The nonlinear approximating family consists of piecewise polynomial functions of degree at most n with k varying points of discontinuity. In 1978, Barrow et al. [2] showed:

Theorem 1. *Let $f \in C^2[0,1]$ with $f'' > 0$ on $[0,1]$ and suppose that $\log f''$ is concave on $(0,1)$. Then f has unique best L^1 and L^2 approximations from S_k^2, the nonlinear family of all second-order (piecewise linear) spline functions with at most k variable knots in $[0,1]$.*

In fact, if we denote by $\mathcal{P}_{k,1}$ the likewise nonlinear family of piecewise linear functions with at most k points of discontinuity in $[0,1]$, the same hypotheses on f suffice to show that f has a unique best approximation in the L^1 sense or L^2 sense from $\mathcal{P}_{k,1}$. This observation follows at once from the formula for the derivative of the error functional $F(\mathbf{x}) = \|f - p(\mathbf{x})\|_1$ with respect to the components of $\mathbf{x} = (x_1, x_2, \dots, x_k) \in R^k$. Here, $\mathbf{x}$ is the k-tuple of possible discontinuity points in $[0,1]$ of the approximant $p(\mathbf{x})$. One sees that at critical points of $F(\mathbf{x})$, the function $p(\mathbf{x})$ is in fact continuous and therefore in S_k^2. From this observation one is naturally led to ask whether, under conditions on f analogous to these in Theorem 1, an f might have a unique best L^p approximation from $P_{k,n}$, the family of all piecewise polynomials of degree at most n with at most k variable points of discontinuity in

Received by the editor April 13, 1994 and, in revised form, January 10, 1995.
1991 *Mathematics Subject Classification.* Primary 41A15, 41A52; Secondary 41A05.
Key words and phrases. Polynomial approximation, Lagrange interpolation, L^1 approximation.

$[0,1]$. Chow [5] answered the question affirmatively for $p = 2$, using the methods of [2]. The question has an affirmative answer in the case $p = 1$ as well, a fact which was established, unbeknownst originally to the authors of the present paper, by Kioustelidis in [11]. The method of proof was similar to that used in [5] for the L^2 case and is nonconstructive. In the present paper we have established this result, already known to Kioustelidis, using methods which are constructive and which utilize an algorithm proposed by Kioustelidis in [10]. We would like to thank the referees for bringing the results contained in [11] to our attention. Our main achievement in this paper is in the manner of establishing the following theorem.

Theorem 2. *Suppose $f \in C^{n+2}[0,1]$ and either*

i. *$f^{(n+1)} > 0$ on $[0,1]$ and $f^{(n+2)}/f^{(n+1)}$ is nonincreasing on $(0,1)$ (i.e., $\log f^{(n+1)}$ is concave), or*

ii. *$f^{(n+1)} < 0$ on $[0,1]$ and $f^{(n+2)}/f^{(n+1)}$ is nondecreasing on $(0,1)$ (i.e., $\log(-f^{(n+1)})$ is concave); then f has a unique best L^1 approximation from $\mathcal{P}_{k,n}$.*

In addition, we have the following theorems which analogize those of [2] and [5] to the L^1 case for piecewise polynomials of arbitrary degree with arbitrarily many knots.

Theorem 3. *Suppose $f \in C^{n+3}[0,1]$ and $f^{(n+1)} > 0$ on $[0,1]$. Then, for sufficiently large k, f has a unique best L^1 approximation from $\mathcal{P}_{k,n}$.*

Theorem 4. *There exists a C^∞ function f such that $f^{(n+1)} > 0$ throughout $[0,1]$ so that f has more than one $L^q[-1,1]$ approximation from the family $\mathcal{P}_{k,n}$, $1 \le q < \infty$.*

Before proceeding to the proof of Theorem 2, which is the main result of this paper, we shall give some formal definitions as well as some known results. In §3 we prove Theorem 2 and in §4 Theorem 3. The proof of Theorem 4 is just as the one already given in [2] and [5] for the L^2 case and is omitted.

2. Preliminaries

We begin by defining the approximating family $\mathcal{P}_{k,n}$, which will be the main object of study. Accordingly, let

$$\Sigma^k = \{\mathbf{x} \in \mathbb{R}^k \mid 0 < x_1 < x_2 < \cdots < x_k < 1\},$$

where the x_i are the components of $\mathbf{x}$. Thus, Σ^k is an open simplex in $\mathbb{R}^k$ and is, therefore, convex. For a fixed $\mathbf{x}$ in the closure of Σ^k, define

$$\mathcal{P}_{k,n,\mathbf{x}} = \{f \in L^\infty[0,1] : f|_{[x_{i-1},x_i]} \in \Pi_n \text{ for } i = 1, 2, \ldots, k+1\},$$

where $x_0 = 0$, $x_{k+1} = 1$ and Π_n denotes the real $(n+1)$-dimensional space of all polynomials with real coefficients of degree at most n. Finally, let

$$\mathcal{P}_{k,n} = \bigcup_{\mathbf{x} \in \overline{\Sigma^k}} \mathcal{P}_{k,n,\mathbf{x}}.$$

We shall refer to this set of functions as the collection of piecewise polynomials of degree at most n with k variable knots, although the term "knot" implies a continuity which the elements of $\mathcal{P}_{k,n}$ need not possess. We observe that the family $\mathcal{P}_{k,n}$ is not a linear space but is approximatively compact.

The following lemma will be useful. Its simple proof is omitted.

Lemma 5. *If $f^{(n+1)} > 0$ throughout $[a,b]$ (or $f^{(n+1)} < 0$ throughout $[a,b]$), then the unique best L^1 approximation from Π_n in $L^1[a,b]$ is given by the Lagrange interpolating polynomial to f at the points $t_j = \dfrac{b-a}{2}\zeta_j + \dfrac{b+a}{2}$, where $\zeta_j = \cos\theta_j$ and $\theta_j = \dfrac{j+1}{n+2}\pi$. The points t_j are known as the L^1 canonical points for the interval $[a,b]$.*

The error between this Lagrange interpolating polynomial and the function f has a well-known form in terms of divided differences:

$$f(t) - p(t) = f[t_0, t_1, \ldots, t_n, t] \prod_{j=0}^{n} (t - t_j),$$

where $f[t_0, t_1, \ldots, t_n, t]$ denotes the $(n+1)$-st order divided difference of f based on the points $\{t_0, t_1, \ldots, t_n, t\} \subset [a,b]$ ([8]). We shall abbreviate this as

$$f(t) - p(t) = f[\mathbf{t}, t] \prod_{j=0}^{n} (t - t_j).$$

Lemma 5 together with the following result allows the succinct reformulation of the piecewise polynomial approximation problem as a nonlinear optimization problem for those functions satisfying the hypothesis of Lemma 5, which we may view as possessing a generalized convexity (or concavity).

Lemma 6. *Suppose $f \in C^1[a,b]$ and, for each i, p_i is the unique best L^q approximation to f from Π_n over the interval $[x_{i-1}, x_i]$. Then, if $1 \le q < \infty$ and*

$$F_q(f, \mathbf{x}) = \sum_{i=1}^{k+1} \int_{x_{i-1}}^{x_i} |f(t) - p_i(t)|^q \, dt,$$

we have

$$\frac{\partial}{\partial x_i} F_q(f, \mathbf{x}) = |f(x_i) - p_i(x_i)|^q - |f(x_i) - p_{i+1}(x_i)|^q.$$

Proof. First note from the form of $F(f, \mathbf{x})$ that its partial derivative with respect to x_i will only involve

$$\frac{\partial}{\partial x_i}\Big(\int_{x_{i-1}}^{x_i} |f(t) - p_i(t)|^q \, dt - \int_{x_i}^{x_{i+1}} |f(t) - p_{i+1}(t)|^q \, dt \Big).$$

We shall use the easily verified fact that if $G(u,t)\colon [a,b] \times [a,b] \to \mathbb{R}$ is continuous and continuously differentiable with respect to u on $[a,b]$, then

$$\frac{\partial}{\partial u}\Big(\int_a^u G(u,t)\, dt \Big)\Big|_{u=x} = G(x,x) + \int_a^u \Big(\frac{\partial}{\partial u} G(u,t)\big|_{u=x} \Big)\, dt.$$

Note that, if p_b is the best L^q approximation to a given continuous f over the interval $[a,b]$, then the quantity

$$|f(t) - p_b(t)|^q$$

is a continuously differentiable function of b and in fact has derivative

$$q\,\mathrm{sgn}(f(t)-p_b(t))|f(t)-p_b(t)|^{q-1}\frac{\partial}{\partial b}(p_b(t)).$$

We apply this observation and the aforementioned fact to obtain

$$\begin{aligned}\frac{\partial}{\partial x_i}\int_{x_{i-1}}^{x_i}|f(t)-p_i(t)|^q\,dt &= |f(x_i)-p_i(x_i)|^q \\ &\quad + \int_{x_{i-1}}^{x_i}\mathrm{sgn}(f(t)-p_i(t))|f(t)-p_i(t)|^{q-1}\frac{\partial}{\partial x_i}(p_i(t))\,dt.\end{aligned}$$

Observe that $\dfrac{\partial}{\partial x_i}(p_i(t))$ is a polynomial of degree at most n and therefore is orthogonal to $\mathrm{sgn}(f(t)-p_i(t))|f(t)-p_i(t)|^{q-1}$ over the interval $[x_{i-1},x_i]$(cf., e.g., [14, p. 168]). Hence, the integral portion of the partial derivative above is zero. Similar remarks apply to the second term involved in the differentiation; thus the lemma is proved. □

Of course, best approximations to f satisfying the hypothesis in Lemma 5 will be found amongst critical points of the functional

$$\begin{aligned}F(f,\mathbf{x}) &= \sum_{i=1}^{k+1}\| f-p_i \|_{1,[x_{i-1},x_i]} \\ &= \sum_{i=1}^{k+1}\int_{x_{i-1}}^{x_i}|f(t)-p_i(t)|\,dt,\end{aligned}$$

that is, amongst those $\mathbf{x}\in\Sigma^k$ satisfying

$$|f[\mathbf{t}_i,x_i]\prod_{j=0}^{n}(x_i-t_{j,i})|-|f[\mathbf{t}_{i+1},t]\prod_{j=0}^{n}(x_i-t_{j,i+1})|=0,$$

where $\mathbf{t}_i=\{t_{0,i},t_{1,i},\dots,t_{n,i}\}$, the L^1 canonical points for the interval $[x_{i-1},x_i]$.

We close this section with a lemma concerning divided differences.

Lemma 7. *Suppose $f\in C^1[a,b]$ and $\{t_0,t_1,\dots,t_n\}\subset[a,b]$ consists of $n+1$ distinct points. Then*

$$\sum_{j=0}^{n} f[t_0,\dots,t_j,t_j,\dots,t_n]=f'[t_0,\dots,t_n]. \tag{2.1}$$

Proof. We proceed by induction on n. If $n=0$, we seek to verify

$$f'[t_0]=f[t_0,t_0],$$

which, for a continuously differentiable function, is true by the definition of the right-hand side. Suppose now that for any collection of n distinct elements of $[a,b]$

the conclusion of the lemma is valid. We wish to show (2.1). We begin with the right-hand side. By the recursive definition of higher-order divided differences

$$\begin{aligned} f'[t_0,\dots,t_n] &= \frac{f'[t_1,\dots,t_n]-f'[t_0,\dots,t_{n-1}]}{t_n-t_0} \\ &= \frac{\sum_{j=1}^{n} f[t_1,\dots,t_j,t_j,\dots,t_n]-\sum_{j=0}^{n-1} f[t_0,\dots,t_j,t_j,\dots,t_{n-1}]}{t_n-t_0}, \end{aligned}$$

where one obtains the second equality by applying the induction hypothesis. We now combine the two sums, producing

$$\begin{aligned} f'[t_0,\dots,t_n] &= \sum_{j=1}^{n-1} \frac{f[t_1,\dots,t_j,t_j,\dots,t_n]-f[t_0,\dots,t_j,t_j,\dots,t_{n-1}]}{t_n-t_0} \\ &\quad + \frac{f[t_1,\dots,t_{n-1},t_n,t_n]-f[t_0,t_0,t_1,\dots,t_{n-1}]}{t_n-t_0} \\ &= \sum_{j=1}^{n-1} f[t_0,\dots,t_j,t_j,\dots,t_{n-1}] \\ &\quad + \frac{f[t_1,\dots,t_{n-1},t_n,t_n]-f[t_0,,t_1,\dots,t_n]}{t_n-t_0} \\ &\qquad + \frac{f[t_0,t_1,\dots,t_n]-f[t_0,t_0,t_1,\dots,t_{n-1}]}{t_n-t_0} \\ &= \sum_{j=0}^{n} f[t_0,\dots,t_j,t_j,\dots,t_{n-1}], \end{aligned}$$

precisely as required. □

3. Proof of Theorem 2

The proof of Theorem 2 involves an algorithm. We begin with a description of the algorithm, which was proposed some time ago by Kioustelidis [10] without any proof of convergence, followed by a demonstration that it is well defined in our setting. Thereafter, we examine its convergence properties under the special assumptions of Theorem 2.

Algorithm. *Fix f such that $f^{(n+1)} > 0$ on the interval* $[0,1]$.

Step I. Select $\mathbf{x} \in \Sigma^k \backslash \partial\Sigma^k$.

Step II. On each subinterval $[x_{i-1}, x_i]$ of $[0,1]$ determined by $\mathbf{x}$, compute $p_{i,q}(t)$, the unique best approximant to f from Π_n in the L^q norm, for $i = 1,2,\dots,k+1$.

Step III. For each $i = 1,2,\dots,k$, locate the zero, z_i say, of

$$g_i(t) = |f(t)-p_{i,q}(t)|^q - |f(t)-p_{i+1,q}(t)|^q$$

which lies nearest x_i.

Step IV. Set $x_i = z_i$, $i = 1,2,\dots,k$, and repeat.

Lemma 8. *For f which satisfies $f^{(n+1)} > 0$, the algorithm just described is well defined at each iteration.*

Proof. We must verify that the map

$$\mathbf{x} \longmapsto \mathbf{z} \text{ where } \mathbf{z} = (z_i, z_2, \ldots, z_k)$$

is well defined. Since $p_{i,q}$ is the best $L^q[x_{i-1}, x_i]$ approximation to f and $f^{(n+1)} > 0$ on that interval, $f - p_{i,q}$ must have precisely $n+1$ simple zeros. Thus, by using Rolle's Theorem and the fact that $f^{(n+1)} > 0$ once again, we conclude that there are no zeros of $f' - p'_{i,q}$ before the first or after the last zero of $f - p_{i,q}$. Denoting these zeros $\tau_{i,0}$ and $\tau_{i,n}$, respectively, for each $i = 1, 2, \ldots, k$, we thus have that $|f - p_{i,q}|^q$ is monotone increasing on the interval $[\tau_{i,n}, \infty]$ whereas $|f - p_{i+1,q}|^q$ is monotone decreasing on $[-\infty, \tau_{i+1,0}]$. Since $g_i(\tau_{i,n}) < 0$ and $g_i(\tau_{i+1,0}) > 0$, it follows that g_i has a unique zero, z_i, which in fact lies on the interval $[\tau_{i,n}, \tau_{i+1,0}]$. $\square$

We now prove Theorem 2. Our proof assumes that the hypotheses of case (i) hold; a virtually identical proof works in case (ii).

Proof of Theorem 2. We shall show that the map

$$\mathbf{x} \longmapsto \mathbf{z}$$

is a contraction on $\bar{\Sigma}^k$ and hence has a unique fixed point on that domain. In fact, since best approximations cannot come from $\partial\Sigma^k$, the unique fixed point must be in Σ^k itself. Clearly, if $\mathbf{x}$ is a critical point, then the algorithm maps $\mathbf{x}$ to itself and hence it is the unique critical point for the minimization problem, and so the element of $\mathcal{P}_{k,n}$ determined by $\mathbf{x}$ will be the unique best approximation to f from $\mathcal{P}_{k,n}$.

Note first that all of the following derivatives exist and are positive:

$$\frac{d}{ds} g_i(s)\Big|_{s=z_i},$$

$$\frac{\partial z_i}{\partial x_{i-1}}, \ \frac{\partial z_i}{\partial x_i}, \text{ and } \frac{\partial z_i}{\partial x_{i+1}}.$$

The first is positive by our earlier discussion, and thus the implicit function theorem guarantees the existence of the final three. These last three are positive because as any one of the parameters x_{i-1}, x_i, or x_{i+1} increases, so too do the canonical points on the intervals $[x_{i-1}, x_i]$ and $[x_i, x_{i+1}]$, and so also the crossings z_i of the error functions.

We now differentiate the expression

$$g_i(z_i) := |f(z_i) - p_i(z_i)| - |f(z_i) - p_{i+1}(z_i)| = 0, \tag{3.1}$$

which defines the value z_i with respect to the parameters x_{i-1}, x_i, or x_{i+1} to obtain the derivative of the map defined in the algorithm. Note that all other possible partial derivatives are zero. Before doing so, we rewrite (3.1), using divided differences, as

$$g_i(z_i) := f[\mathbf{t}_i, z_i]Q_i(z_i) + (-1)^n f[\mathbf{t}_{i+1}, z_i]Q_{i+1}(z_i) = 0, \tag{3.2}$$

where $\mathbf{t}_i$ is the vector $(t_{i,0}, t_{i.1}, \ldots, t_{i,n})$ of L^1 canonical points for the interval $[x_{i-1}, x_i]$ for the family Π_n,

$$Q_i(t) = \prod_{j=0}^{n}(t - t_{i,j}),$$

and we have used the fact that $z_i \in [t_{i,n}, t_{i+1,0}]$ to infer the sign of the terms in question. Proceeding with the differentiation, we obtain

$$\frac{\partial z_i}{\partial x_{i-1}} = -\frac{1}{2g_i'(z_i)}\big[Q_i(z_i)\sum_{j=0}^{n} f[\mathbf{t}_i, t_{i,j}, z_i](1-\zeta_j) - f[\mathbf{t}_i, z_i]\sum_{j=0}^{n} Q_{i,j}(z_i)(1-\zeta_j)\big],$$

$$\begin{aligned}\frac{\partial z_i}{\partial x_{i+1}} = -\frac{(-1)^n}{2g_i'(z_i)}\big[&Q_i(z_i)\sum_{j=0}^{n} f[\mathbf{t}_{i+1}, t_{i+1,j}, z_i](1+\zeta_j)\\ &- f[\mathbf{t}_{i+1}, z_i]\sum_{j=0}^{n} Q_{i+1,j}(z_i)(1+\zeta_j)\big],\end{aligned}$$

and

$$\begin{aligned}\frac{\partial z_i}{\partial x_i} = &-\frac{1}{2g_i'(z_i)}\big[Q_i(z_i)\sum_{j=0}^{n} f[\mathbf{t}_i, t_{i,j}, z_i](1+\zeta_j) - f[\mathbf{t}_i, z_i]\sum_{j=0}^{n} Q_{i,j}(z_i)(1+\zeta_j)\big]\\ &-\frac{(-1)^n}{2g_i'(z_i)}\big[Q_{i+1}(z_i)\sum_{j=0}^{n} f[\mathbf{t}_{i+1}, t_{i+1,j}, z_i](1-\zeta_j)\\ &\qquad - f[\mathbf{t}_{i+1}, z_i]\sum_{j=0}^{n} Q_{i+1,j}(z_i)(1-\zeta_j)\big],\end{aligned}$$

where

$$Q_{i,j} = \prod_{l=0, l\neq j}^{n}(t - t_{i,j}),$$

and

$$\begin{aligned}g_i'(z_i) = &Q_i(z_i) f[\mathbf{t}_i, z_i, z_i] + f[\mathbf{t}_i, z_i]\sum_{j=0}^{n} Q_{i,j}(z_i)\\ &+ (-1)^n\big[Q_{i+1}(z_i) f[\mathbf{t}_{i+1}, z_i, z_i] + f[\mathbf{t}_{i+1}, z_i]\sum_{j=0}^{n} Q_{i+1,j}(z_i)\big].\end{aligned}$$

Our interest is in the row norm of the matrix $\mathcal{Z} = \{\frac{\partial z_i}{\partial x_l}\}_{1\le i\le k, 1\le l\le k}$ and hence in the sum of the absolute values of the partial derivatives above. By an earlier remark, however, this is merely their sum. Thus,

$$\begin{aligned}
&\Big|\frac{\partial z_i}{\partial x_{i-1}}\Big| + \Big|\frac{\partial z_i}{\partial x_{i+1}}\Big| + \Big|\frac{\partial z_i}{\partial x_i}\Big| \\
&\quad = \frac{1}{g_i'(z_i)}\Big[- Q_i(z_i)\sum_{j=0}^{n} f[\mathbf{t}_i, t_{i,j}, z_i] + f[\mathbf{t}_i, z_i]\sum_{j=0}^{n} Q_{i,j}(z_i) \\
&\qquad\qquad - (-1)^n Q_{i+1}(z_i)\sum_{j=0}^{n} f[\mathbf{t}_{i+1}, t_{i+1,j}, z_i] \\
&\qquad\qquad + (-1)^n f[\mathbf{t}_{i+1}, z_i]\sum_{j=0}^{n} Q_{i+1,j}(z_i)\Big],
\end{aligned}$$

where products involving ζ_j have been cancelled in the addition. We shall show that under the conditions imposed on f, the sum above is less than or equal to 1 in rows two through k and is strictly less in the first and last rows. Now

$$\begin{aligned}
&\Big|\frac{\partial z_i}{\partial x_{i-1}}\Big| + \Big|\frac{\partial z_i}{\partial x_{i+1}}\Big| + \Big|\frac{\partial z_i}{\partial x_i}\Big| \le 1 \\
&\iff -Q_i(z_i)\sum_{j=0}^{n} f[\mathbf{t}_i, t_{i,j}, z_i] + f[\mathbf{t}_i, z_i]\sum_{j=0}^{n} Q_{i,j}(z_i) \\
&\quad - (-1)^n Q_{i+1}(z_i)\sum_{j=0}^{n} f[\mathbf{t}_{i+1}, t_{i+1,j}, z_i] + (-1)^n f[\mathbf{t}_{i+1}, z_i]\sum_{j=0}^{n} Q_{i+1,j}(z_i) \\
&\qquad \le g_i'(z_i) = Q_i(z_i) f[\mathbf{t}_i, z_i, z_i] + f[\mathbf{t}_i, z_i]\sum_{j=0}^{n} Q_{i,j}(z_i) \\
&\qquad\quad + (-1)^n Q_{i+1}(z_i) f[\mathbf{t}_{i+1}, z_i, z_i] + (-1)^n f[\mathbf{t}_{i+1}, z_i]\sum_{j=0}^{n} Q_{i+1,j}(z_i) \\
&\iff -Q_i(z_i)\sum_{j=0}^{n} f[\mathbf{t}_i, t_{i,j}, z_i] - (-1)^n Q_{i+1}(z_i)\sum_{j=0}^{n} f[\mathbf{t}_{i+1}, t_{i+1,j}, z_i] \\
&\qquad \le Q_i(z_i) f[\mathbf{t}_i, z_i, z_i] + (-1)^n Q_{i+1}(z_i) f[\mathbf{t}_{i+1}, z_i, z_i].
\end{aligned}$$

By Lemma 7 we have

$$\sum_{j=0}^{n} f[\mathbf{t}_i, t_{i,j}, z_i] = f'[\mathbf{t}_i, z_i] - f[\mathbf{t}_i, z_i, z_i]$$

and

$$\sum_{j=0}^{n} f[\mathbf{t}_{i+1}, t_{i+1,j}, z_i] = f'[\mathbf{t}_{i+1}, z_i] - f[\mathbf{t}_{i+1}, z_i, z_i].$$

Applying this observation allows us to infer that the above inequality is equivalent to

$$0 \le Q_i(z_i) f'[\mathbf{t}_i, z_i] + (-1)^n Q_{i+1}(z_i) f'[\mathbf{t}_{i+1}, z_i]$$

or

$$Q_i(z_i)f'[\mathbf{t}_i, z_i] \geq (-1)^{(n+1)}Q_{i+1}(z_i)f'[\mathbf{t}_{i+1}, z_i]$$
$$\iff f'[\mathbf{t}_i, z_i] \geq \frac{(-1)^{(n+1)}Q_{i+1}(z_i)}{Q_i(z_i)}f'[\mathbf{t}_{i+1}, z_i].$$

Recalling (3.2), we find this is equivalent to

$$f'[\mathbf{t}_i, z_i] \geq \frac{f[\mathbf{t}_i, z_i]}{f[\mathbf{t}_{i+1}, z_i]}f'[\mathbf{t}_{i+1}, z_i]$$
$$\iff \frac{f'[\mathbf{t}_i, z_i]}{f[\mathbf{t}_i, z_i]} \geq \frac{f'[\mathbf{t}_{i+1}, z_i]}{f[\mathbf{t}_{i+1}, z_i]}.$$

An argument using the Cauchy Mean Value theorem shows that this condition is implied by the log concavity condition imposed on $f^{(n+1)}$. Hence, the row norm of $\mathcal{Z}$ is less than or equal to one. We note, further, that the sums in rows one and k are strictly less than one. We claim that this, in conjunction with the special form of the matrix $\mathcal{Z}$, implies that its spectral norm, the product of its eigenvalues, is strictly less than one. Indeed all the eigenvalues of $\mathcal{Z}$ are real since $\mathcal{Z}$ is symmetric and, since the row norm of $\mathcal{Z}$ dominates the spectral, all are of absolute value less than one. If $\lambda = \pm 1$ is an eigenvalue of $\mathcal{Z}$, then $\lambda I - \mathcal{Z}$ is singular. But $\lambda I - \mathcal{Z}$ is an irreducible diagonally dominant matrix which is strictly so in either row one or row k, hence nonsingular, a contradiction. Thus, the spectral norm of $\mathcal{Z}$ is less than one as claimed, and the map

$$\mathbf{x} \longmapsto \mathbf{z}$$

is a contraction as we indicated. □

Note that the algorithm is well defined in the general L^q case, though no convergence results are shown here in cases other than $q = 1$. When $q \neq 1$ or 2, the calculation of the polynomial of best L^q approximation to f across each subinterval is a difficult nonlinear problem in general. Iterative methods for arriving at approximate solutions to this problem are, however, available [15]. One notes as well that the algorithm just described decreases the L^p norm of the error curve at each step provided the sign of $f^{(n+1)}$ is constant. Finally, preliminary calculations indicate that standard methods may successfully be applied to accelerate the rate of convergence.

4. Proof of Theorem 3

The proof of Theorem 3 is, in outline, as the proof of the analogous theorems in [2] and [5]. One shows that the matrix of second partials of the functional $F(f, \mathbf{x})$ has positive determinant at any critical point $\mathbf{x} \in \Sigma^k$ for sufficiently large k, whenever $f^{(n+1)} > 0$. This, together with the unicity of the best approximation to t^{n+1} from $\mathcal{P}_{k,n}$ in all the L^q norms $1 \leq q \leq \infty$, allows one to infer through a topological degree of mapping argument that f has only one critical point over Σ^k and thus a unique best approximation. Aside from establishing a particular and convenient form for the entries of the matrix of second partials of the functional $F(f, \mathbf{x})$, the proof is the same in any L^q norm and is given in [2] and [5]. Hence, we shall give here only those elements peculiar to the L^1 case which are new here. The first step is to establish a convenient form for the entries of $J(f, \mathbf{x})$, the matrix of the second partials of $F(f, \mathbf{x})$, in terms of integrals against a kernel function which depends on n but on neither f nor k. The following lemma furthers this aim.

Proposition 9. *Suppose $f \in C^{(n+1)}[a,b]$ and p is the unique best approximation to p from Π_n in the L^1 sense. Suppose*

$$L(f,[a,b]) = f(a) - p(a)$$

and

$$R(f,[a,b]) = f(b) - p(b).$$

Then

$$L(f,[a,b]) = \frac{(-1)^{(n+1)}(b-a)^n}{n!}\int_0^1 f^{(n+1)}((b-a)t+a)K_n^+(t)\,dt$$

and

$$R(f,[a,b]) = \frac{(b-a)^n}{n!}\int_0^1 f^{(n+1)}((b-a)t+a)K_n^-(t)\,dt,$$

where

$$K_n^+(t) = \sum_{j=0}^n A_j(\xi_j - t)_+^n,$$

$$K_n^-(t) = \sum_{j=0}^n A_j(\xi_j - t)_-^n,$$

$$A_j = \prod_{k=0,k\neq j}^n \frac{1-\xi_k}{\xi_j - \xi_k},$$

and $(x-t)_+^n$ denotes the truncated power of $(x-t)^n$ given by

$$(x-t)_+^n = \begin{cases} (x-t)^n & \text{if } x \geq t, \\ 0 & \text{otherwise,} \end{cases}$$

and the values $\xi_k{}_{k=0}^n$ are the L^1-canonical points for Π_n for the interval $[0,1]$ indexed so that $\xi_n < \xi_{n-1} < \cdots < \xi_0$.

Proof. We will use the Peano Kernel Theorem, a statement of which may be found in Davis [6, Theorem 9.7.1]. As was discussed earlier, for a given f over $[a,b]$, p is the uniquely determined $p \in \Pi_n$ which interpolates f at the points given by

$$t_k = (b-a)\xi_k + a.$$

Accordingly,

$$L((\cdot - \tau)_+^n, [a,b]) = (a-\tau)_+^n - \sum_{j=0}^n (t_j - \tau)_+^n \prod_{k\neq j} \frac{a - t_k}{t_j - t_k}$$

and

$$R((\cdot - \tau)_+^n, [a,b]) = (b-\tau)_+^n - \sum_{j=0}^n (t_j - \tau)_+^n \prod_{k\neq j} \frac{b - t_k}{t_j - t_k}.$$

Noting that $(a-\tau)_+^n = 0$ and that $(b-\tau)_+^n = (b-\tau)^n$ whenever $\tau \in [a,b]$, the above forms become

$$L((\cdot-\tau)_+^n,[a,b]) = -\sum_{j=0}^{n}(t_j-\tau)_+^n \prod_{k\neq j}\frac{a-t_k}{t_j-t_k}$$

and

$$R((\cdot-\tau)_+^n,[a,b]) = (b-\tau)^n - \sum_{j=0}^{n}(t_j-\tau)_+^n \prod_{k\neq j}\frac{b-t_k}{t_j-t_k}.$$

Thus by the Peano Kernel Theorem,

$$L(f,[a,b]) = \frac{1}{n!}\int_a^b f^{(n+1)}(\tau)\big(-\sum_{j=0}^{n}(t_j-\tau)_+^n \prod_{k\neq j}\frac{a-t_k}{t_j-t_k}\big)\,d\tau$$

and

$$R(f,[a,b]) = \frac{1}{n!}\int_a^b f^{(n+1)}(\tau)\big((b-\tau)^n - \sum_{j=0}^{n}(t_j-\tau)_+^n \prod_{k\neq j}\frac{a-t_k}{t_j-t_k}\big)\,d\tau.$$

Now we convert the above integrals to integrals over the interval $[0,1]$ by making the change of variables $\tau = (b-a)t+a$. Noting that $t_j = (b-a)\xi_j + a$, we obtain

$$L(f,[a,b]) = \frac{-(b-a)^{(n+1)}}{n!}\int_0^1 f^{(n+1)}((b-a)t+a)\big(\sum_{j=0}^{n}(\xi_j-t)_+^n \prod_{k\neq j}\frac{-\xi_k}{\xi_j-\xi_k}\big)\,dt$$

and

$$R(f,[a,b]) = \frac{(b-a)^{(n+1)}}{n!}\int_0^1 f^{(n+1)}((b-a)t+a)$$
$$\cdot\big((1-t)_+^n - \sum_{j=0}^{n}(\xi_j-t)_+^n \prod_{k\neq j}\frac{1-\xi_k}{\xi_j-\xi_k}\big)\,dt.$$

Now the symmetry of the ξ_j yields

$$-\xi_k = -(1-\xi_k) \quad \text{and} \quad \xi_j - \xi_k = -(\xi_{n-j}-\xi_{n-k}),$$

which in turn implies

$$\prod_{k\neq j}\frac{-\xi_k}{\xi_j-\xi_k} = (-1)^n \prod_{k\neq j}\frac{1-\xi_k}{\xi_j-\xi_k},$$

where the symmetry of the canonical points has been again employed. Substituting this into the form for $L(f,[a,b])$ produces

$$L(f,[a,b]) = \frac{(-1)^{(n+1)}(b-a)^{(n+1)}}{n!}\int_0^1 f^{(n+1)}((b-a)t+a)\big(\sum_{j=0}^{n}A_j(\xi_j-t)_+^n\big)\,dt$$

and

$$R(f,[a,b]) = \frac{(b-a)^{(n+1)}}{n!} \int_0^1 f^{(n+1)}((b-a)t+a)\big((1-t)^n - \sum_{j=0}^{n} A_j(\xi_j - t)_+^n\big)\,dt.$$

Now we expand $(1-t)^n$ in terms of the polynomials $q_j(t) = (\xi_j - t)^n$, using the fact that the latter form a Haar system relative to any real interval, so as to obtain a comparable kernel function for both $L(f,[a,b])$ and $R(f,[a,b])$. The useful expansion is, remarkably,

$$(1-t)^n = \sum_{j=0}^{n} A_j(\xi_j - t)^n.$$

To verify this, for each fixed t consider the polynomial in u of degree n, $(u-t)^n$. The Lagrange interpolating form for this polynomial in u using the points $\xi_0,\dots,\xi_n$ is

$$(u-t)^n = \sum_{j=0}^{n} (\xi_j - t)^n \prod_{k\neq j} \frac{u-\xi_k}{\xi_j - \xi_k},$$

an identity in u. If $u=1$, we obtain the form claimed. Using this, we find

$$\begin{aligned} R(f,[a,b]) &= \frac{(b-a)^{(n+1)}}{n!} \int_0^1 f^{(n+1)}((b-a)t+a) \textstyle\sum_{j=0}^{n} A_j\big((\xi_j-t)^n - (\xi_j - t)_+^n\big)\,dt \\ &= \frac{(b-a)^{n+1}}{n!} \int_0^1 f^{(n+1)}((b-a)t+a) \sum_{j=0}^{n} A_j(\xi_j - t)_-^n\,dt, \end{aligned}$$

which completes the proof of the proposition. □

Henceforth we shall denote $\dfrac{\partial}{\partial x_i} F(f,\mathbf{x})$ by $F_i(f,\mathbf{x})$. With this notation we have the following corollary to Proposition 9.

Corollary 10. *There holds*

$$\begin{aligned} F_i(f,\mathbf{x}) = &\frac{h_i^{(n+1)}}{n!} \int_0^1 f^{(n+1)}(x_i - th_i)K_n(t)\,dt \\ &- \frac{h_{i+1}^{(n+1)}}{n!} \int_0^1 f^{(n+1)}(x_i + th_{i+1})K_n(t)\,dt, \end{aligned}$$

where $K_n(t) = K_n^+(t)$ as given in the preceding proposition, $\mathbf{x} = (x_1, x_2, \dots, x_k)$, and $h_i = x_i - x_{i-1}$.

Proof. Using the hypothesis that $f^{(n+1)} > 0$ and Rolle's Theorem, one shows that

$$F_i(f,\mathbf{x}) = (f(x_i) - p_i(x_i)) - (f(x_i) - p_{i+1}(x_i))$$

if n is odd, and

$$F_i(f,\mathbf{x}) = (f(x_i) - p_i(x_i)) - (p_{i+1}(x_i) - f(x_i))$$

if n is even. Thus, in the notation of Proposition 9,

$$F_i(f,\mathbf{x}) = \begin{cases} R(f,[x_{i-1},x_i]) - L(f,[x_i,x_{i+1}]) & \text{if } n \text{ is odd,} \\ R(f,[x_{i-1},x_i]) + L(f,[x_i,x_{i+1}]) & \text{if } n \text{ is even.} \end{cases}$$

Hence, applying Proposition 9, we get

$$\begin{aligned} F_i(f,\mathbf{x}) &= \frac{h_i^{(n+1)}}{n!}\int_0^1 f^{(n+1)}(x_i+th_i)K_n^-(t)\,dt \\ &\quad - (-1)^{n+1}\frac{h_{i+1}^{(n+1)}}{n!}\int_0^1 f^{(n+1)}(x_i+th_{i+1})K_n^+(t)\,dt \\ &= \frac{h_i^{(n+1)}}{n!}\int_0^1 f^{(n+1)}(x_i+th_i)K_n^-(t)\,dt \\ &\quad - \frac{h_{i+1}^{(n+1)}}{n!}\int_0^1 f^{(n+1)}(x_i+th_{i+1})K_n^+(t)\,dt \end{aligned}$$

if n is odd, and

$$\begin{aligned} F_i(f,\mathbf{x}) &= \frac{h_i^{(n+1)}}{n!}\int_0^1 f^{(n+1)}(x_i+th_i)K_n^-(t)\,dt \\ &\quad + (-1)^{n+1}\frac{h_{i+1}^{(n+1)}}{n!}\int_0^1 f^{(n+1)}(x_i+th_{i+1})K_n^+(t)\,dt \\ &= \frac{h_i^{(n+1)}}{n!}\int_0^1 f^{(n+1)}(x_i+th_i)K_n^-(t)\,dt \\ &\quad - \frac{h_{i+1}^{(n+1)}}{n!}\int_0^1 f^{(n+1)}(x_i+th_{i+1})K_n^+(t)\,dt \end{aligned}$$

if n is even. Thus, $F_i(f,\mathbf{x})$ has the same form for even and odd n. Finally, note that

$$\frac{h_i^{(n+1)}}{n!}\int_0^1 f^{(n+1)}(x_i+th_i)K_n^-(t)\,dt = \frac{h_i^{(n+1)}}{n!}\int_0^1 f^{(n+1)}(x_i-th_i)K_n^+(t)\,dt$$

to obtain the corollary. □

Using these forms, one can show, as in [2] and [5], that the matrix of second partials is a tridiagonal matrix whose entries $\{a_{i,j}\}_{1\le i,j\le k}$ satisfy the inequality

$$a_{ii}a_{i-1,i-1}\big(1+O(\Delta^3)\big) \ge 4a_{i,i-1}a_{i-1,i}$$

at any critical point $\mathbf{x}$, where $\Delta = \max\{h_i | i = 1,\dots,k\}$. This, in turn, implies that $\det J(f,\mathbf{x}) > 0$ at such a point. The proof of Theorem 3 then follows by the degree of mapping argument as given in [2] and [5].

References

1. J. Al-Siwan, *Best segmented L_2-approximation with free knots*, Dissertation, University of Oregon, Eugene, March 1986.
2. D. L. Barrow, C. K. Chui, P. W. Smith, and J. D. Ward, *Unicity of best mean approximation by second order splines with variable knots*, Math. Comp. **32** (1978), 1132–1148. MR **58:**1853
3. D. Braess, *"Nonlinear Approximation Theory"*, Springer-Verlag, Berlin-New York, 1986. MR **88e:**41002
4. H. G. Burchard, *Piecewise polynomial approximation on optimal meshes*, Journal of Approximation Theory **14** (1975), 128–147. MR **51:**10957
5. J. Chow, *On the uniqueness of best $L_2[0,1]$ approximation by piecewise polynomials with variable breakpoints*, Math. Comp. **39** (1982), 571–585. MR **83j:**41009
6. P. J. Davis, *"Interpolation and Approximation"*, Blaisdell, New York, 1963. MR **28:**393
7. R. L. Eubank, *On the relationship between functions with the same optimal knots in spline and piecewise polynomial approximation*, Journal of Approximation Theory **40** (1984), 327–332. MR **85j:**41016
8. K. Jetter, *L_1 Approximation verallgemeinerter konvexer Funktionen durch Splines mit freien Knoten*, Math. Z. **164** (1978), 55–66. MR **80f:**41007
9. J. B. Kioustelidis, *Optimal segmented approximations*, Computing **24** (1980), 1-8. MR **82i:** 41032
10. J. B. Kioustelidis, *Optimal segmented polynomial L^p-approximation*, Computing **26** (1981), 239–246. MR **82h:**41006
11. J. B. Kioustelidis, *Uniqueness of optimal piecewise polynomial L_1 approximations for generalized convex functions*, from "Functional Analysis and Approximation", Internat. Ser. Numer. Math., vol. 60 (1981), Birkhäuser Verlag, Basel, 421–432. MR **83d:**41036
12. J. M. Ortega, *"Numerical Analysis, a Second Course"*, Academic Press, New York, 1972. MR **53:**6967
13. A. M. Pinkus, *"On L^1-Approximation"*, Cambridge Univ. Press, Cambridge and New York, 1989. MR **90j:**41046
14. M. J. D. Powell, *"Approximation Theory and Methods"*, Cambridge University Press, Cambridge, 1981. MR **82f:**41001
15. J. M. Wolfe, *On the convergence of an algorithm for discrete L_p approximation*, Numer. Math. **32** (1979), 439–459. MR **80i:**65021

Department of Science and Mathematics, Montana State University-Northern, P. O. Box 7751, Havre, Montana 59501

E-mail address: `gayle@nmc1.nmclites.edu`

Department of Mathematics, University of Oregon, Eugene, Oregon 97403-1222

E-mail address: `wolfe@bright.uoregon.edu`