# On monotone and convex approximation by splines with free knots

D. Leviatan*

*School of Mathematics, Raymond and Beverley Sackler Faculty of Exact Sciences*
*Tel Aviv University, 69978 Tel Aviv, Israel*

A. Shadrin*†

*Computing Center, Siberian Branch, Russian Academy of Sciences,*
*pr. Lavrentieva 6, Novosibirsk 630090, Russia*

To appear in Annals of Numer. Math. (1996) (Ch. Micchelli, Editor)

**Abstract.** We prove that the degree of shape preserving free knot spline approximation in $L_p[a,b]$, $0 < p \leq \infty$ is essentially the same as that of the non-constrained case. This is in sharp contrast to the well known phenomenon we have in shape preserving approximation by splines with equidistant knots and by polynomials. The results obtained are valid both for piecewise polynomials and for smooth splines with the highest smoothness.

*Supported by the Office of Naval Research Contract N0014-91-J1343 and the National Foundation Grant EHR 9108772.

†Supported by the Russian Foundation of Fundamental Researches under Contract 95-01-00949a.

*AMS classification:* 41A15, 41A25, 41A29.

*Keywords and phrases:* Shape preserving approximation, Free knot splines, Degree of approximation.

## 0. Introduction

Recent years have seen a growing interest in questions of shape preserving approximation. In particular there has been extensive activity in questions of estimating the degree of approximation of monotone and convex functions $f \in L_p[a,b]$ by means of polynomials $p_n \in \Pi_n$, the space of algebraic polynomials of degree not exceeding $n$, and splines of degree $k$ and $n$ equidistant knots, which preserve the monotonicity and the convexity, respectively. Here and in the sequel, we mean by $L_\infty[a,b]$, the space $C[a,b]$. Also, in the sequel we will use for $[a,b]$, the generic interval $[0,1]$. These approximation processes are linear in the sense that the approximants for each fixed $n$ are taken from a fixed linear set of elements. The degree of such constrained approximation is in general worse than the degree of nonconstrained approximation of the same functions. This phenomenon is well illustrated in that while for any $f \in C^1[0,1]$ we have

$$E_n(f)_\infty \le \frac{C}{n} E_{n-1}(f')_\infty \ ,$$

where $E_n(f)_\infty$ denotes the degree of approximation by polynomials of degree not exceeding $n$, a recent result of Shevchuk [5] shows that there exists an absolute constant $C > 0$ such that for any $n \ge 1$, a nondecreasing $f = f_n \in C^1[0,1]$ exists such that

$$\inf_{\substack{p_n \in \Pi_n \\ p_n \nearrow}} \|f - p_n\|_\infty \ge C E_{n-1}(f')_\infty \ .$$

Nevertheless the degree of constrained approximation can be estimated by Jackson-type estimates, namely, involving the moduli of smoothness of the function, but unlike the nonconstrained approximation by polynomials and such splines, the estimates of monotone and convex approximation involve only moduli of smoothness of very restricted orders, $\omega_2$ and $\omega_3$, respectively (for details see [2] and [1]).

It is well-known that for nonconstrained approximation, nonlinear methods usually achieve a given degree of approximation for a much wider class of functions than linear methods do (see discussion in the book by Petrushev and Popov [4] and especially Chapter 7 therein). Naturally, one would like to know whether the same phenomenon occurs for constrained approximation, and indeed we prove that this is so for the approximation by splines with free knots. Maybe this is to be expected since splines with free knots seem to be most amenable to shape preservation requirements. It would be interesting to see whether constrained rational approximation can provide the same degree of approximation as the nonconstrained does. This seems to be a very difficult question and the only result we are aware of in this direction is a (non-optimal) estimate of convex rational approximation to a convex continuously differentiable function (see [3]). The striking conclusion of this work is that the constrained approximation by splines with free knots is in a sense as good as the nonconstrained approximation, unlike the situation described above for the approximation by polynomials and splines with fixed knots. (We refer the reader to [4] where the various classes of functions with a given degree of approximation by splines with free knots are characterized as Besov spaces.) We will show that if we allow the number of knots to be some constant multiple of the original one, then we preserve monotonicity and convexity while guaranteeing the same degree of approximation as in the nonconstrained case. We first prove these assertions in §1 and §2 for splines with free knots without any continuity assumptions at the knots. Then in §3 we show that the same holds for such splines with maximal smoothness at the knots. At the same time we show in §4 that if one wishes to preserve the monotonicity or convexity with some smaller constant multiple of the original number of knots, even without any continuity assumptions at the knots, then one cannot obtain the same degree of approximation.

We denote by $S(N,k)$, the collection of splines of degree $k$ with at most $N-1$ knots, i.e., with at most $N$ polynomial pieces, where we do not assume any smoothness at the knots. We further denote by $\tilde{S}(N,k)$, the sub-collection of those splines with maximum continuity, i.e., which are in $C^{k-1}[0,1]$. Now we let $M^r = M^r[0,1]$ be the space of all functions $f$ defined on $[0,1]$ for which the $r$th forward difference is nonnegative in $[0,1]$. In particular $M^1$ is the space of monotone functions in $[0,1]$, $M^2$ is the space of convex functions in $[0,1]$ and in general we call $f \in M^r$ an $r$-monotone function. Finally, we set $M_p^r = M_p^r[0,1] := L_p[0,1] \cap M^r[0,1]$, $0 < p \leq \infty$. (Recall that by $L_\infty[0,1]$ we mean the space $C[0,1]$.)

For $f \in M_p^r$, we denote the degree of approximation by splines in $S(N,k)$, by $S_{N,k}(f)_p$, and the degree of approximation by $r$-monotone splines in $S(N,k)$, by $S_{N,k}^{(r)}(f)_p$. Similarly, we denote by $\tilde{S}_{N,k}^{(r)}(f)_p$, the degree of approximation by $r$-monotone splines in $\tilde{S}(N,k)$.

## 1. Monotone spline approximation

Our result on monotone approximation is the following.

**Theorem 1.** *For any $k \in \mathbf{N}$ there exists a constant $A = A(k)$ such that for every $f \in M^1_p[0,1]$, $0 < p \le \infty$, and each $N \in \mathbf{N}$ we have*

$$S^{(1)}_{AN,k}(f)_p \le S_{N,k}(f)_p \ .$$

Theorem 1 is an immediate consequence of the following.

**Proposition 1.** *For any $k \in \mathbf{N}$ there exists a constant $A = A(k)$ such that for every $f \in M^1_p$, $0 < p \le \infty$, $N \in \mathbf{N}$, and $s \in S(N,k)$, there exists a spline $\sigma$ such that*

$$\sigma \in S(AN,k) \cap M^1[0,1],$$

*and*

$$|f(x) - \sigma(x)| \le |f(x) - s(x)|, \quad \forall x \in [0,1].$$

The proof of Proposition 1 follows from

**Lemma 1.** *For any $k \in \mathbf{N}$ there exists a constant $A = A(k)$ such that for every $f \in M^1$, and $p \in \Pi_k$, there exists a spline $\pi$ such that*

$$\pi \in S(A,k) \cap M^1[0,1],$$

$$|f(x) - \pi(x)| \le |f(x) - p(x)|, \quad \forall x \in [0,1],$$

*and*

$$f(0) \le \pi(0), \quad \pi(1) \le f(1).$$

**Proof of Proposition 1.** We apply Lemma 1 to each polynomial piece of the spline $s \in S(N,k)$.

Finally, in order to prove Lemma 1 we divide it into two parts, Lemmas 1a and 1b, from which Lemma 1 readily follows.

**Lemma 1a.** *Let $f \in M^1[0,1]$ and $p \in \Pi_k$ be given. Then there exists $g$, such that*

$$g \in M^1[0,1], \tag{A}$$

*the difference $g(x) - p(x)$ changes its sign at most $k$ times in $[0,1]$,* $(B)$

*$g(x)$ is equal to either $f(x)$ or $p(x)$ $\forall x \in [0,1]$ so that*

$$|g(x) - p(x)| \leq |f(x) - p(x)|, \quad \forall x \in [0,1], \tag{C}$$

*and*

$$f(0) = g(0), \quad g(1) = f(1). \tag{D}$$

**Lemma 1b.** *Let $g \in M^1$ and $p \in \Pi_k$ be such that the difference $g(x) - p(x)$ changes its sign at most $k$ times. Then there exists a spline $\pi$ such that*

$$\pi \in M^1[0,1], \tag{a}$$

$$\pi \in S(A,k), \quad A = A(k), \tag{b}$$

$$|g(x) - \pi(x)| \leq |g(x) - p(x)|, \quad \forall x \in [0,1], \tag{c}$$

*and*

$$g(0) \leq \pi(0), \quad \pi(1) \leq g(1). \tag{d}$$

**Proof of Lemma 1a.** Divide the interval $[0,1]$ into $l$ subintervals of monotonicity of the given polynomial $p$, i.e.,

$$[0,1] = \cup_{i=l}^{l} I_i, \quad I_i =: [x_i, x_{i+1}],$$

where $x_1 = 0$, $x_{l+1} = 1$, and the other $x_i$ are exactly the points at which the derivative $p'$ changes its sign. We denote $I_i$ by $I_i^+$ and $I_i^-$ if on $I_i$, the polynomial $p$ is increasing and decreasing, respectively.

First, on $I_i^-$, the graph of the decreasing polynomial $p$ intersects the graph of the nondecreasing function $f$ at most once, i.e., the difference $f(x) - p(x)$ changes sign at most once. We set

$$g(x) = f(x), \quad x \in I_i^-.$$

On $I_i^+$, the graphs of the two non-decreasing functions $p$ and $f$ may intersect at any number of points, even an infinite numbers of points $\{t_\alpha\}$. If the number of intersections is $\le 1$, then we still set $g(x) = f(x)$. Otherwise, we put

$$\gamma_i := \inf_\alpha t_\alpha, \quad \delta_i := \sup_\alpha t_\alpha,$$

and set

$$g(x) = \begin{cases} p(x), & x \in (\gamma_i, \delta_i); \\ f(x), & \text{otherwise.} \end{cases}$$

By construction, the difference $g(x) - p(x)$ changes its sign on $I_i^+$ at most once.

We are ready to verify the properties of $g$.

$(A)$. The function $g$ is equal to $f \in M^1$ on $[\delta_i, \gamma_{i+1}]$, and to a non-decreasing part of $p$ on $(\gamma_i, \delta_i)$. At the points of $\gamma_i$ we have

$$g(\gamma_i-) = f(\gamma_i-) \le f(\gamma_i+) = p(\gamma_i+) = g(\gamma_i+),$$

and similarly at the points $\delta_i$. Hence $g \in M^1$.

$(B)$. The difference $g(x) - p(x)$ has at most one change of sign on each of the intervals $I_i$ of monotonicity of $p \in \Pi_k$, and the total number $l$ of such intervals does not exceed $k$.

$(C)$. The inequality holds since $g(x)$ is equal either to $f(x)$ or to $p(x)$.

$(D)$. The function $g$ coincides with $f$ at the endpoints of all intervals $I_i$, in particular 0 and 1.

This completes the proof of Lemma 1a.

**Proof of Lemma 1b.** Let $\{y_i\}_1^l$ be the points where the difference $g(x) - p(x)$ changes its sign. We set

$$y_0 = 0, \quad y_{l+1} = 1, \quad J_i = [y_{i-1}, y_i],$$

and denote $J_i$ by $J_i^+$ and $J_i^-$, if the sign of $p(x) - g(x)$ on $J_i$ is non-negative and non-positive, respectively.

We obtain the spline $\pi$ on $[0,1]$ by constructing its parts $\pi_i$ on each of the intervals $J_i$ with the following properties

$$\pi_i \in M^1[J_i], \tag{a1}$$

$$\pi_i \in S(A_i, k), \tag{$b1$}$$

where $A_i = A_i(k)$,

$$|g(x) - \pi_i(x)| \leq |g(x) - p(x)|, \quad \forall x \in J_i, \tag{$c1$}$$

and

$$g(y_{i-1}) \leq \pi_i(y_{i-1}), \quad \pi_i(y_i) \leq g(y_i). \tag{$d1$}$$

Then, the monotonicity of $\pi$, i.e., $(a)$, follows by $(a1)$ and $(d1)$.
Further, $(b)$ follows by $(b1)$ with

$$A = \sum A_i.$$

Now $(c)$ follows trivially from $(c1)$, and $(d)$ is trivially implied by $(d1)$ since $0 = y_0$ and $1 = y_{l+1}$.

We first define the spline $\pi_i$ on $J_i^+$. Recall that $x \in J_i^+$ implies $p(x) \geq g(x)$.

Denote by $\{\xi_j\}_{j=1}^{k_i}$ the points of local minima of the polynomial $p$ inside $J_i^+$, in increasing order, and set $\xi_0 = y_{i-1}$, $\xi_{k_i+1} = y_i$.

Then for $x = y_i = \xi_{k_i+1}$ let

$$\pi_i(y_i) := g(y_i) \leq p(y_i).$$

We continue by induction defining for $x \in [\xi_{j-1}, \xi_j)$,

$$\pi_i(x) := \min\left(p(x), \pi_i(\xi_j)\right), \quad j = k_i + 1, \ldots, 1.$$

Then for $J_i^+$, we readily have,

$(a1)$. The monotonicity follows from the fact that $\pi_i$ is equal either to a non-decreasing part of $p$ or to a constant which are so defined that monotonicity is preserved.

$(b1)$. Each subinterval $[\xi_{j-1}, \xi_j]$ of $J_i$, consists of two intervals (one of which may be empty) with $\pi_i$ being $p \in \Pi_k$ in the left one and a constant in the right one. The number of these subintervals is equal to $k_i + 1 \leq k + 1$. Hence,

$$A_i \leq 2(k_i + 1).$$

$(c1)$. Our construction yields,

$$g(x) \leq \pi_i(x) \leq p(x), \quad \forall x \in J_i^+.$$

Hence the required inequality.

$(d1)$. We have

$$g(y_{i-1}) \leq \pi_i(y_{i-1}), \quad \pi_i(y_i) = g(y_i).$$

We proceed in a similar way to define the spline $\pi_i$ on $J_i^-$, where we recall that $x \in J_i^-$ implies $p(x) \leq g(x)$. This time $\pi_i$ is defined moving from the left to the right, and taking maxima instead of minima. Namely, we denote by $\{\eta_j\}_{j=1}^{k_i}$ the points of local maxima of the polynomial $p$ which are inside $J_i^-$ and we set $\eta_0 = y_{j-1}$, $\eta_{k_i+1} = y_i$.

Then for $x = \eta_0 = y_{i-1}$ we let

$$\pi_i(x) := g(y_{i-1}) \geq p(y_{i-1}),$$

and for $x \in (\xi_j, \xi_{j+1}]$ we let

$$\pi_i(x) := \max(\pi_i(\xi_{j-1}), p(x)), \quad j = 1, \dots, k_i.$$

This completes the proof of Lemma 1b but we would like to get an estimate of $A(k)$. To this end we see that

$$A = \sum_1^{k+1} A_i = 2 \sum_1^{k+1} (k_i + 1) = 2(k+1) + 2 \sum k_i.$$

But by definition $k_i$ is the number of either local minima or local maxima of the polynomial $p \in \Pi_k$ in the interval $J_i$. Thus $\sum k_i$ does not exceed the total number of the local extrema of $p \in \Pi_k$, which is $\leq k-1$. Hence,

$$A = A_k \leq 4k.$$

In fact closer analysis shows that actually $A \leq 2k$.

*Remark.* The definition of $\pi_i$ on $J_i^+$ is equivalent to the following:

$$\pi_i(x) := \sup\ q(x),$$

where supremum is taken over all functions $q$ such that

$$q \in M^1[J_i^+], \quad q(t) \le p(t), \forall t \in J_i^+, \quad q(y_i) = g(y_i).$$

Similarly for $J_i^-$ (of course taking infimum on $q \ge p$, instead of supremum). We have chosen to give the more constructive proof for the monotone case and we use the other approach later for the convex case where it significantly simplifies part of the proof of Lemma 2b.

## 2. Convex spline approximation

In this section we obtain estimates for convex approximation. We prove

**Theorem 2.** *For any $k \in \mathbf{N}$ there exists a constant $A = A(k)$ such that for every $f \in M_p^2[0,1]$, $0 < p \le \infty$, and any $N \in \mathbf{N}$ we have*

$$S^{(2)}_{AN,k}(f)_p \le S_{N,k}(f)_p \ .$$

Theorem 2 is an immediate consequence of the following.

**Proposition 2.** *For any $k \in \mathbf{N}$ there exists a constant $A = A(k)$ such that for every $f \in M_p^2$, $0 < p \le \infty$, $N \in \mathbf{N}$, and $s \in S(N,k)$, there exists a spline $\sigma$ such that*

$$\sigma \in S(AN,k) \cap M^2[0,1],$$

$$|f(x) - \sigma(x)| \le |f(x) - s(x)|, \quad \forall x \in [0,1].$$

The proof of Proposition 2 follows from

**Lemma 2.** *For any $k \in \mathbf{N}$ there exists a constant $A = A(k)$ such that for every $f \in M^2$, and $p \in \Pi_k$, there exists a spline $\pi$ such that*

$$\pi \in S(A,k) \cap M^2[0,1],$$

$$|f(x) - \pi(x)| \le |f(x) - p(x)|, \quad \forall x \in [0,1],$$

$$f(0) = \pi(0), \quad \pi(1) = f(1),$$

*and*

$$f'(0+) \leq \pi'(0+), \quad \pi'(1-) \leq f'(1-).$$

**Proof of Proposition 2.** We apply Lemma 2 to each polynomial part of the spline $s \in S(N,k)$.

Finally as before, in order to prove Lemma 2 we divide it into two parts, Lemmas 2a and 2b from which Lemma 2 readily follows.

**Lemma 2a.** *Let $f \in M^2[0,1]$ and $p \in \Pi_k$ be given. Then there exists $g$, such that*

$$g \in M^2, \tag{E}$$

*the difference $g(x) - p(x)$ changes its sign at most $2k$ times in $[0,1]$,* $(F)$

*$g(x)$ is equal to either $f(x)$ or $p(x)$ $\forall x \in [0,1]$ so that*

$$|g(x) - p(x)| \leq |f(x) - p(x)|, \quad \forall x \in [0,1], \tag{G}$$

$$f(0) = g(0), \quad g(1) = f(1), \tag{H}$$

*and*

$$f'(0+) \leq g'(0+), \quad g'(1-) \leq f'(1-). \tag{K}$$

**Lemma 2b.** *Let $g \in M^2[0,1]$ and $p \in \Pi_k$ be such that the difference $g(x) - p(x)$ changes its sign at most $2k$ times. Then there exists a spline $\pi$ such that*

$$\pi \in M^2[0,1], \tag{e}$$

$$\pi \in S(A,k), \quad A = A(k), \tag{f}$$

$$|g(x) - \pi(x)| \leq |g(x) - p(x)|, \quad \forall x \in [0,1], \tag{g}$$

$$g(0) = \pi(0), \quad \pi(1) = g(1), \tag{h}$$

*and*

$$g'(0+) \leq \pi'(0+), \quad \pi'(1-) \leq g'(1-). \tag{k}$$

**Proof of Lemma 2a.** Divide the interval $[0,1]$ into $l$ subintervals of convexity and concavity of the given polynomial $p$, i.e.,

$$[0,1] = \cup_{i=l}^{l} I_i, \quad I_i = [x_i, x_{i+1}],$$

where $x_1 = 0$, $x_{l+1} = 1$, and the other $x_i$ are exactly the points at which the second derivative $p''$ changes sign. Denote $I_i$ by $I_i^+$ and $I_i^-$, if on $I_i$, $p$ is convex and concave, respectively.

On $I_i^-$, the graph of the concave polynomial $p$ intersects the graph of the convex function $f$ at most twice, i.e., the difference $f(x) - p(x)$ changes sign at most twice. So we set

$$g(x) = f(x), \quad x \in I_i^-.$$

On the other hand on $I_i^+$, the graphs of the two convex functions $p$ and $f$ may intersect at any number of points, even at an infinite number of points $\{t_\alpha\}$. If the number of intersections is $\leq 2$, then we still set $g(x) = f(x)$. Otherwise, we put

$$\gamma_i := \inf \{ t_\alpha : \ f'(t_\alpha +) \leq p'(t_\alpha +) \},$$

$$\delta_i := \sup \{ t_\alpha : \ p'(t_\alpha -) \leq f'(t_\alpha -) \},$$

and we set

$$g(x) = \begin{cases} p(x), & x \in (\gamma_i, \delta_i); \\ f(x), & \text{otherwise.} \end{cases}$$

Note that at two consecutive $t_\alpha$'s the above inequalities are reversed. Therefore there is at most one $t_\alpha < \gamma_i$ for which $p'(t_\alpha +) \leq f'(t_\alpha + ax)$, and there is at most one $t_\alpha > \delta_i$ for which $f'(t_\alpha -) \leq p'(t_\alpha -)$. Thus there is at most one intersection between $f$ and $p$ on either side of $(\gamma_i, \delta_i)$. Hence, the difference $g(x) - p(x)$ changes sign at most twice in $I_i^+$.

We now verify that $g$ has the required properties.

$(E)$. The function $g$ is evidently continuous. It is equal to $f \in M^2$ on $[\delta_i, \gamma_{i+1}]$, and to a convex part of $p$ on $(\gamma_i, \delta_i)$. At the points $\gamma_i$, we have

$$g'(\gamma_i -) \leq f'(\gamma_i -) \leq f'(\gamma_i +) \leq p'(\gamma_i +) = g'(\gamma_i +), \tag{2.1}$$

i.e.,

$$g'(\gamma_i -) \leq g'(\gamma_i +).$$

Similarly for the points $\delta_i$. Hence $g \in M^2$.

$(F)$. The difference $g(x) - p(x)$ has at most two changes of sign in each of the intervals $I_i$, of convexity and concavity of $p \in \Pi_k$, and the number of such intervals is less than $k$.

$(G)$. The inequality holds since $g(x)$ is equal either to $f(x)$ or to $p(x)$.

$(H)$. The function $g$ coincides with $f$ at the endpoints of each of the intervals $I_i$, in particular at 0 and 1.

$(K)$. If $0 < \gamma_1$, then $g$ coincides with $f$ in the neighborhood of 0, hence the first derivatives coincide too. The same with 1, if $\delta_{l+1} < 1$. Otherwise, the inequalities follow by (2.1) and its counterpart for $\delta_i$.

This completes the proof of Lemma 2a.

**Proof of Lemma 2b.** Let $\{y_i\}_1^l$ be the points where the difference $g(x) - p(x)$ changes its sign. We set

$$y_0 = 0, \quad y_{l+1} = 1, \quad J_i = [y_{i-1}, y_i],$$

and denote $J_i$ by $J_i^+$ and $J_i^-$, if on $J_i$, the sign of $p(x) - g(x)$ is non-negative and non-positive, respectively.

We obtain the spline $\pi$ on $[0,1]$ by constructing its parts $\pi_i$ on each of the intervals $J_i$ with the following properties

$$\pi_i \in M^2[J_i], \tag{e1}$$

$$\pi_i \in S(A_i, k), \tag{f1}$$

where $A_i = A_i(k)$,

$$|g(x) - \pi_i(x)| \le |g(x) - p(x)|, \quad \forall x \in J_i, \tag{g1}$$

$$g(y_{i-1}) = \pi_i(y_{i-1}), \quad \pi_i(y_i) = g(y_i), \tag{h1}$$

and

$$g'(y_{i-1}+) \le \pi_i'(y_{i-1}+), \quad \pi_i'(y_i-) \le g'(y_i-). \tag{k1}$$

Then, the convexity of $\pi$, that is $(e)$ follows from $(e1)$, $(h1)$ and $(k1)$.

Further, $(f)$ follows by $(f1)$ with

$$A = \sum A_i,$$

Now, $(g)$ trivially follows from $(g1)$ and finally $(h)$ and $(k)$ follow from $(h1)$ and $(k1)$, respectively since $0 = y_0$ and $1 = y_{l+1}$.

On $J_i^+$ we have $\quad g(x) \le p(x)$, and we set

$$\pi_i(x) := \sup\ q(x)$$

where supremum is taken over all functions $q$ such that

$$q \in M^2[J_i^+]; \quad q(t) \le p(t), \quad \forall t \in J_i^+; \quad q(u) = g(u), \quad u = y_{i-1},\, y_i.$$

Then the convexity of $\pi_i(x)$ readily follows as the maximum of convex functions is convex, hence $(e1)$.

As for the other requirements.

$(f1)$. We have to show that $\pi_i$ is a piecewise polynomial, and evaluate $A_i$.

Denote by $\{\xi_j\}_{j=1}^{k_i}$ the points where the second derivative $p''$ changes its sign, and set $\xi_0 = y_{i-1}$, $\xi_{k_i+1} = y_i$. Then we have

$$J_i^+ = \cup_{j=0}^{k_i} T_j, \quad \text{where} \quad T_j := [\xi_j, \xi_{j+1}],$$

and on each of the intervals $T_j$ the polynomial $p$ is alternatingly convex and concave. Denote $T_j$ by $T_j^+$ and $T_j^-$ if, respectively, it is an interval of convexity or concavity of $p$.

By definition, $\pi_i(x) \le p(x)$, so that $J_i^+ = G \cup E$, where

$$G := \{\, x \in J_i^+ :\ \pi_i(x) < p(x)\,\}, \quad \text{and} \quad E := \{\, x \in J_i^+ :\ \pi_i(x) = p(x)\,\}.$$

Obviously $G$ is open, hence it is a union of open intervals. On each of these intervals the convex function $\pi_i(x)$ is linear, for otherwise we can replace part of it by a linear function in a way that preserves both the convexity and the inequality $\pi_i(x) < p(x)$. A contradiction to the definition of $\pi_i$.

It is also evident that the intervals $T_j^-$ are contained in $G$. Hence the set $E$ consists of some closed subintervals where $p$ is convex. We claim that there can be at most one such subinterval in $T_j^+$. For if we assume the contrary and we take two neighboring disjoint subintervals in $T_j^+$, then between them $\pi_i(x)$ is a linear function which coincides with $p$ at the endpoints. But $p$ is convex, thus $\pi_i(x) \ge p(x)$, a contradiction.

Thus, for $x \in J_i^+$ the function $\pi_i(x)$ defined above coincides with $p(x)$ on at most one subinterval of each $T_j^+$, and is equal to a linear function between two such neighboring subintervals, in particular on $T_j^-$. Hence

$$\pi_i \in S(A_i, k), \quad \text{where} \quad A_i \le \sum_{j:\ T_j \in J_i^+} 1 = k_i + 1,$$

where $k_i$ is the number of sign changes of $p''$ on $J_i^+$.

Now, $(g1)$ and $(k1)$ follow since by construction

$$g(x) \le \pi_i(x) \le p(x), \quad \forall x \in J_i^+,$$

and $(h1)$ holds for all functions $q$ involved in the definition of $\pi_i$, hence holds for $\pi_i$. The definition of $\pi_i$ and the proofs are more complicated on $J_i^-$, where $p(x) \le g(x)$.

Again divide the interval $J_i^-$ into subintervals $T_j^+$ and $T_j^-$ of convexity and concavity of the polynomial $p$, respectively, and set

$$T_j = [\xi_j, \xi_{j+1}], \quad j = 0, \dots, k_i.$$

On each interval $T_j^-$ we prescribe one point $\eta_j$ at which the distance between $p$ and $g$ is minimal, that is,

$$g(\eta_j) - p(\eta_j) \le g(x) - p(x), \quad \forall x \in T_j^-.$$

If there are more than one such a point, then we arbitrarily take one.

We define the linear functions

$$\begin{aligned} \ell_{-1}(x): \quad & \ell_{-1}(y_{i-1}) = g(y_{i-1}), \quad \ell'_{-1}(y_{i-1}) = g'(y_{i-1}+), \\ \ell_{k_i+1}(x): \quad & \ell_{k_i+1}(y_i) = g(y_i), \quad \ell'_{k_i+1}(y_i) = g'(y_i-), \end{aligned}$$

and for $j$ such that $T_j = T_j^-$,

$$\ell_j(x): \begin{cases} \ell_j^{(r)}(\eta_j) = p^{(r)}(\eta_j), \quad r = 0, 1, & \text{if } \eta_j \text{ is interior point}, \\ \ell_j(\eta_j) = p(\eta_j),\ \ell'_j(\eta_j) = g'(\eta_j+) & \text{if } \eta_j = \xi_j, \\ \ell_j(\eta_j) = p(\eta_j),\ \ell'_j(\eta_j) = g'(\eta_j-) & \text{if } \eta_j = \xi_{j+1}. \end{cases}$$

We observe that it follows from the minimality property of $\eta_j$, that $g'(\eta_j-) \le p'(\eta_j) \le g'(\eta_j+)$, whenever $\eta_j$ is interior to $T_j^-$, and when $\eta_j$ is one of the endpoints, then the right-hand inequality holds or the left-hand inequality holds when $\eta_j$ is the left-hand endpoint or the right-hand endpoint, respectively. This together with the convexity of $g$, implies that the slope of $\ell_{j-1}$ is smaller than the slope of $\ell_{j+1}$, and that

$$\ell_j(x) \leq g(x), \quad \forall x \in J_i^-.$$

We now set

$$\ell(x) := \max_j \ell_j(x),$$

and finally,

$$\pi_i(x) := \max\{\, \ell(x),\, p(x)\,\}.$$

$(e1)$. First, since $\ell_j$ is the tangent to $p$ at a point of concavity, then

$$\ell(x) \geq \ell_j(x) \geq p(x), \quad \forall x \in T_j^-, \quad j \in \{0, \ldots, k_i\}.$$

This means that the inequality

$$\ell(z_0) < p(z_0),$$

implies that $z_0 \in T_j^+$ for some $j$. Therefore if we let

$$\phi_j(x) = \begin{cases} \max\{\, \ell_{j-1}(x),\, p(x),\, \ell_{j+1}(x)\,\}, & x \in T_j^+; \\ \max\{\, \ell_{j-1},\, \ell_{j+1}(x)\,\}, & x \notin T_j^+, \end{cases}$$

then evidently

$$\pi_i(x) = \max_j \phi_j(x).$$

In order to complete the proof of $(e1)$, we first prove that for each $j$ we have $\phi_j \in M^2[J_i^-]$. Then $\pi_i$ is convex as the maximum of convex functions and the proof of $(e1)$ is complete.

To this end we note that since

$$\ell_{j-1}(\xi_j) \geq p(\xi_j), \quad \ell_{j+1}(\xi_{j+1}) \geq p(\xi_{j+1}),$$

each of the equations

$$\ell_{j-1}(x) - p(x) = 0, \quad \ell_{j+1}(x) - p(x) = 0, \tag{2.2}$$

has at most one solution in $T_j^+ = [\xi_j, \xi_{j+1}]$. Moreover, if these solutions exist and are $\beta_{j-1}$, $\beta_{j+1}$ respectively, then we have

$$\ell'_{j-1}(\beta_{j-1}) \le p'(\beta_{j-1}), \quad p'(\beta_{j+1}) \le l'_{j+1}(\beta_{j+1}). \tag{2.3}$$

If there is no solution to one or any of the equations (2.2), or if their solutions satisfy the inequality

$$\beta_{j-1} \ge \beta_{j+1},$$

then $\phi_j(x) = \max\{\, \ell_{j-1}(x),\, \ell_{j+1}(x)\}$ and it is convex since the slope of $\ell_{j-1}$ is smaller than the slope of $\ell_{j+1}$.

If on the other hand, each of the equations (2.2) has a solution in $T_j^+$, and these solutions satisfy the inequality

$$\beta_{j-1} < \beta_{j+1},$$

then

$$\phi_j(x) = \begin{cases} \ell_{j-1}(x), & x \in [y_{i-1}, \beta_{j-1}]; \\ p(x), & x \in [\beta_{j-1}, \beta_{j+1}]; \\ \ell_{j+1}(x), & x \in [\beta_{j+1}, y_i], \end{cases}$$

and by virtue of (2.3), $\phi_j$ is convex.

$(f1)$. Recall that

$$\ell(x) := \max_j \ell_j(x), \quad \ell_j \in \Pi_1,$$

where $\ell_{-1}$ and $\ell_{k_i+1}$ correspond to the endpoints of $J_i^-$ and the others to the intervals $T_j^-$ of concavity of the polynomial $p$. Hence

$$\ell \in S(A'_i, 1), \quad A'_i \le 2 + \sum_{j:\ T_j = T_j^-} 1.$$

Since

$$\pi_i(x) = \max\{\, \ell(x),\, p(x)\,\},$$

we change $\ell$ into $p$ at most once in each $T_j^+$. Thus,

$$\pi_i \in S(A_i, k),$$

where
$$A_i \le A_i' + \sum_{j:\ T_j=T_j^+} 1 \le 2 + \sum_{j:\ T_j\subset J_i^-} 1 = k_i + 3.$$
Now for $(g1)$ we have
$$p(x) \le \phi_j(x), \quad x \in T_{j-1}^- \cup T_j^+ \cup T_{j+1}^-,$$
$$\phi_j(x) \le g(x), \quad x \in J_i^-,$$
and
$$\pi_i(x) = \max_j \ \phi_j(x).$$
Hence
$$p(x) \le \pi_i(x) \le g(x), \quad x \in J_i^-.$$
Finally $(h1)$ and $(k1)$ follow by the construction.

It remains to evaluate $A$. We have at most $2k$ sign changes of $g(x)-p(x)$, hence at most $2k+1$ intervals $J_i$. Therefore
$$A \le \sum_1^{2k+1} A_i \le \sum_1^{2k+1} (k_i+3) = 3(2k+1) + \sum k_i.$$
By definition $k_i$ is the number of sign changes of $p''$ on the interval $J_i$, so that $\sum k_i$ does not exceed the total number of zeros of the second derivative of $p \in \Pi_k$, that is, $\le k-2$. Hence,
$$A = A_k \le 7k+1.$$
This concludes the proof of Lemma 2b.

## 3. Approximation by smooth splines

**Theorem 3.** *Let $r = 1,2$. For any $k \in \mathbf{N}$ there exists a constant $B = B(k)$ such that for every $f \in M_p^r[0,1]$, $0 < p \le \infty$, and each $N \in \mathbf{N}$ we have*
$$\tilde{S}_{BN,k}^{(r)}(f)_p \le S_{N,k}(f)_p\ .$$

Theorem 3 readily follows from Theorems 1 and 2, and the following lemma.

**Lemma 3.** *Let $r = 1, 2$. For each $k \in \mathbf{N}$, there exists a constant $B = B(k)$ with the following property. For every $f \in M_p^r[0,1]$, $0 < p \leq \infty$, and all $N \in \mathbf{N}$ and $\epsilon_0 > 0$, if a (non-smooth) spline $s \in S(N,k) \cap M^r[0,1]$, is such that*

$$\|f - s\|_p < \epsilon_0, \tag{3.1}$$

*then for every $\epsilon > 0$, a (smooth) spline $s_\epsilon \in \tilde{S}(BN,k) \cap M^r[0,1]$ exists, satisfying*

$$\|f - s_\epsilon\|_p < \epsilon_0 + \epsilon.$$

**Proof of Theorem 3.** The proof follows from the above by letting $\epsilon_0 \searrow S_N(f)_p$, which we can do by Theorems 1 and 2, and $\epsilon \to 0$.

**Proof of Lemma 3.** We first prove the lemma for monotone approximation, which we later apply to the convex case.

**The monotone case.** We begin by replacing the spline $s \in S(N,k) \cap M^1[0,1]$ with a continuous spline $\tilde{s} \in S(2N,k) \cap M^1[0,1]$ such that

$$\|f - \tilde{s}\|_p < \epsilon_0 + \epsilon/3. \tag{3.2}$$

Let $x_1 < \cdots < x_{N-1}$ be the knots of $s$ and set $x_N := 1$. If $p = \infty$, let $\delta(\epsilon) > 0$ be such that

$$|f(x) - f(y)| < \epsilon/3, \qquad |x - y| \leq \delta, \tag{3.3}$$

and sufficiently small so that $x_i + \delta < x_{i+1}$, $i = 1, \ldots, N-1$. Otherwise, we just assume the latter on $\delta$.

If $x_i$ is any of the points of discontinuity of $s$, then we replace $s$ on $[x_i, x_i + \delta]$ by the straight line connecting $s(x_i-)$ and $s(x_i + \delta)$. The new spline $\tilde{s}$ is evidently in $S(2N,k) \cap M^1[0,1]$, and it differs from $s$ only on the intervals $[x_i, x_i + \delta]$. Now, the boundedness of $s$ implies that for $0 < p < \infty$, we may choose $\delta$ so small that

$$\|s - \tilde{s}\|_p < \epsilon/3. \tag{3.4}$$

This together with (3.1), implies (3.2). If $p = \infty$, then it follows by (3.1) and (3.3) that for $x_i \le x \le x_i + \delta$,

$$|s(x_i-) - f(x)| < \epsilon_0 + \epsilon/3$$

and

$$|s(x_i + \delta) - f(x)| < \epsilon_0 + \epsilon/3.$$

Hence (3.2) follows.

Thus, replacing $\epsilon_0$ by $\epsilon_0 + \epsilon/3$, which we rename $\epsilon_0$, we may assume that $s$ satisfies (3.1) and is continuous. By adding $\epsilon x/3$ to $s$ we obtain $\bar{s}$ such that $|s(x) - \bar{s}(x)| \le \epsilon/3$ and $\bar{s}' \ge \eta$, where $\eta := \epsilon/3$. Therefore, we may assume that $s' \ge \eta$ for some fixed small $\eta > 0$.

We now smooth $s'$ at each of the knots $x_i$, $i = 1, \ldots, N-1$, retaining the nonnegativity and keeping close to it in the $L_1$ norm.

Let $x_i$ be any of the knots of $s$, and we wish to smooth it there. Then at $x_i$ there meet two polynomial pieces of $s$, the polynomial $p_1$ on the left and $p_2$ on the right. Recall that $p_1' \ge \eta$ in some left neighborhood of $x_i$ and $p_2' \ge \eta$ in some right neighborhood of $x_i$. By [4, Lemma 7.12] it follows that for the above $\eta$, there exists a $\delta_0 > 0$, such that for each $0 < \delta \le \delta_0$, a spline $s_i' \in \tilde{S}(k+1, k-1)$ exists, with knots at $x_i = y_1 < y_2 < \cdots < y_k = x_i + \delta$, such that

$$s_i'(x) = \begin{cases} p_1'(x), & x \le x_i, \\ p_2'(x), & x \ge x_i + \delta, \end{cases}$$

and

$$0 \le s_i'(x) \le \max\{p_1'(x_i), p_2'(x_i + \delta)\} + \eta, \quad x_i < x < x_i + \delta.$$

We denote the smooth spline obtained in this way $s_\epsilon'$ and observe that in view of the boundedness of $s'$, for sufficiently small $\delta \le \delta_0$, we have

$$\|s' - s_\epsilon'\|_1 < \epsilon/3.$$

Hence by integration we get a spline $s_\epsilon \in \tilde{S}(k(N-1)+1, k) \cap M^1[0,1]$, such that

$$\|s - s_\epsilon\|_\infty < \epsilon/3,$$

hence for all $0 < p \le \infty$

$$\|s - s_\epsilon\|_p < \epsilon/3. \tag{3.5}$$

We conclude the proof of the monotone case by combining this with (3.1).

**The convex case.** We observe that the convex spline $s$ is continuous at the knots, but that $s'$ which is monotone, may have jumps there. We first smooth the monotone $s'$ by the previous scheme so that for the given $\epsilon$, we obtain $s'_\epsilon \in \tilde{S}(BN, k-1)$, such that together, (3.4) and (3.5) in the $L_1$-norm for $s'$ become

$$\|s' - s'_\epsilon\|_1 < \epsilon.$$

Note that $\epsilon_0$ of (3.1) is not involved in any way since we do not use the sup-norm. Now we integrate and obtain a convex $s_\epsilon \in \tilde{S}(BN, k)$, such that for all $0 < p \leq \infty$,

$$\|s - s_\epsilon\|_p < \epsilon.$$

Now, by virtue of (3.1)

$$\begin{aligned} \|f - s_\epsilon\|_p &\leq \|f - s\|_p + \|s - s_\epsilon\|_p \\ &< \epsilon_0 + \epsilon, \end{aligned}$$

and the proof is concluded.

## 4. A negative result

In this section we show that the upper bounds for the constants $A^{(r)} := A(k)$, $r = 1, 2$, in Theorems 1 and 2, respectively, namely,

$$A^{(1)} \leq 2k, \quad A^{(2)} \leq 7k + 1,$$

are exact up to the order with respect to $k$, the degree of the splines. Specifically, we obtain the following lower estimate

$$A^{(r)} \geq \left[\frac{k-r}{2}\right], \quad k > r + 1.$$

Note that the result is valid for $r$-monotone approximation, $r \geq 1$ so we state it accordingly. This means that for a given $r \in \mathbf{N}$, we can guarantee the same error bounds for shape-preserving approximation by spline with free $N_1$ knots, as for the best approximation by spline with free $N$ knots only if

$$N_1 \geq c\,k\,N,$$

for some $c > 0$.

The proof of the following theorem is a slight modification of the arguments by Shvedov [6] for the case of shape preserving approximation by polynomials.

**Theorem 4.** *For any $k, r \in \mathbf{N}$, $k > r+1$, a constant $a = a(k,r)$ exists, such that for any $N \in \mathbf{N}$ and $0 < p \leq \infty$, and each $C > 0$, there exists $F \in M^r[0,1]$, for which*

$$S^{(r)}_{aN,k}(F)_p > C\, S_{N,k}(F)_p.$$

**Proof.** Given $k, r, N$ as above, we set

$$m := [(k-r)/2], \quad \text{and} \quad N_1 := Nm.$$

Let $q_1 \in \Pi_{2m}$ be the polynomial

$$q_1(x) := \prod_{i=1}^{m} (x - i/m)^2, \quad x \in [0,1],$$

and

$$q_N(x) := q_1(Nx - j), \quad x \in (j/N,\, (j+1)/N], \quad j = 0, \ldots, N-1.$$

Finally for $0 < \epsilon < \frac{1}{(2m)^{2m}}$ to be prescribed, let

$$q(x) := q(x;\, N, \epsilon) := q_N(x) - \epsilon,$$

and

$$f(x) := f(x;\, N, \epsilon) := \max\{\, q(x), 0\,\}.$$

The functions $q(x)$ and $f(x)$ have the following properties.

$$q \in S(N, 2m), \tag{4.1}$$

$$f(x) \geq 0, \quad x \in [0,1], \tag{4.2}$$

$$q(\nu/N_1) = -\epsilon < 0, \quad \nu = 1, \ldots, N_1, \tag{4.3}$$

$$|q(x) - f(x)| \leq \epsilon, \quad x \in [0,1], \tag{4.4}$$

and

$$\text{mes}\,\{x:\; q(x) \neq f(x)\} \leq c_1(m)\sqrt{\epsilon}. \tag{4.5}$$

Only (4.5) needs explanation. First observe that

$$\begin{aligned} \text{mes}\{x : q(x) \neq f(x)\} &= \text{mes}\{x : q(x) < 0\} \\ &= \text{mes}\{x : q_N(x) < \epsilon\} \\ &= N\text{mes}\{x \in [0, 1/N] : q_N(x) < \epsilon\} \\ &= \text{mes}\{x : q_1(x) < \epsilon\}. \end{aligned}$$

For $x \in [\frac{1}{2m}, 1]$, there is exactly one $\nu$ such that $|x - \frac{\nu}{m}| \leq \frac{1}{2m}$. For all $j \neq \nu$ we have $|x - \frac{j}{m}| > \frac{1}{2m}$. Thus $q_1(x) < \epsilon$ implies

$$\left(\frac{1}{2m}\right)^{m-1} |x - \frac{\nu}{m}| < \sqrt{\epsilon}.$$

Hence

$$|x - \frac{\nu}{m}| < (2m)^{m-1}\sqrt{\epsilon} < \frac{1}{2m},$$

by the choice of $\epsilon$. Since

$$q_1(x) > \left(\frac{1}{2m}\right)^{2m} > \epsilon, \quad 0 \leq x < \frac{1}{2m},$$

we conclude that

$$\text{mes}\{x : q_1(x) < \epsilon\} \leq \frac{1}{2}(2m)^m\sqrt{\epsilon},$$

so (4.5) is proved.

Next, define the functions $Q$ and $F$ as the $r$th integrals of $q$ and $f$ respectively, i.e.,

$$Q(x) = \int_0^x \int_0^{t_r} \ldots \int_0^{t_2} q(t_1)\, dt_1 \,\ldots\, dt_r\,,$$

and

$$F(x) = \int_0^x \int_0^{t_r} \ldots \int_0^{t_2} f(t_1)\, dt_1 \,\ldots\, dt_r\,.$$

It immediately follows by (4.1) through (4.5) that

$$Q \in S(N, k), \tag{4.6}$$

$$F \in M^r[0,1], \tag{4.7}$$

$$Q^{(r)}(\nu/N_1) = -\epsilon < 0, \quad \nu = 1, \ldots, N_1, \tag{4.8}$$

and

$$|Q(x) - F(x)| \le c_1(m)\epsilon\sqrt{\epsilon}. \tag{4.9}$$

Finally, we show that for any $k$, $r$, $k > r+1$, $N$, $0 < p \le \infty$ and $C > 0$ we can prescribe $\epsilon$ so that for any $R \in S(N_1, k) \cap M^r$ the inequality

$$\|F - R\|_p > C\,\|F - Q\|_p$$

is valid. This proves Theorem 4 with

$$a(k,r) \ge m = \left[\frac{k-r}{2}\right].$$

To this end let

$$R \in S(N_1, k) \cap M^r[0,1].$$

Then, since the total number of its polynomial pieces is $\le N_1$, there exists an interval $I_0$, such that

$$R\,|_{I_0} \in \Pi_k, \quad |I_0| := \operatorname{mes} I_0 = 1/N_1,$$

and thus for some $\mu \in \{1, \ldots, N_1 - 1\}$

$$x_\mu := \mu/N_1 \in \bar{I}_0.$$

We are now going to apply the relation

$$R^{(r)}(x_\mu) \ge 0,$$

which is due to the assumption that $R \in M^r[0,1]$; Markov's inequality for $p \in \Pi_k(I)$, namely,

$$\|p^{(r)}\|_{L_\infty(I)} \le c_2(p, r, k, |I|)\,\|p\|_{L_p(I)};$$

and (4.6) through (4.9) in order to obtain

$$
\begin{aligned}
\|F - Q\|_{L_p[0,1]} &\leq \|F - Q\|_{L_\infty[0,1]} \\
&\leq c_1\, \epsilon\sqrt{\epsilon} \\
&= c_1 \sqrt{\epsilon}\, |Q^{(r)}(x_\mu)| \\
&\leq c_1 \sqrt{\epsilon}\, |Q^{(r)}(x_\mu) - R^{(r)}(x_\mu)| \\
&\leq c_1 \sqrt{\epsilon}\, \|Q^{(r)} - R^{(r)}\|_{L_\infty(I_0)} \\
&\leq c_1\, c_2 \sqrt{\epsilon}\, \|Q - R\|_{L_p(I_0)} \\
&\leq c_1\, c_2 \sqrt{\epsilon}\, \|Q - R\|_{L_p[0,1]} \\
&\leq c_1\, c_2\, c_3(p) \sqrt{\epsilon}\, (\, \|Q - F\|_{L_p[0,1]} + \|F - R\|_{L_p[0,1]}\, ).
\end{aligned}
$$

Therefore we conclude that

$$
\|F - Q\|_{L_p[0,1]} \leq c_4 \sqrt{\epsilon}\, (\, \|Q - F\|_{L_p[0,1]} + \|F - R\|_{L_p[0,1]}\, ),
$$

where

$$
c_4 := c_1\, c_2\, c_3 = c(k, p, r, N).
$$

Hence,

$$
\|F - R\|_{L_p[0,1]}) \geq c_5\, \|F - Q\|_{L_p[0,1]},
$$

with

$$
c_5 = \frac{1}{c_4\sqrt{\epsilon}} - 1,
$$

Thus we can readily prescribe $\epsilon$ small enough so that

$$
c_5 > C.
$$

This completes the proof.

*Acknowledgement.* The results of this paper were obtained in February 1995 while both authors were visiting Prof. R. A. DeVore at the University of South Carolina. The authors are indebted to Ron DeVore for making their visit possible and to him and to Pencho Petrushev for valuable discussions and helpful comments on the subject.

## References

1. R. A. DeVore, Y. K. Hu and D. Leviatan, *Convex polynomial and spline approximation in $L_p$, $0 < p < \infty$*, Constr. Approx. (to appear).

2. R. A. DeVore, D. Leviatan and X. M. Yu, *Polynomial Approximation in $L_p$ $(0 < p < 1)$*, Constr. Approx. 8(1992), 187–201.

3. Bo Gao, D. J. Newman and V. A. Popov, *Convex approximation by rational functions*, SIAM J. Math. Anal. 26(1995), 488–499.

4. P. P. Petrushev and V. A. Popov, *Rational approximation of real functions*, Cambridge University Press, Cambridge 1987.

5. I. A. Shevchuk, *One example in monotone approximation*, J. Approx. Theory (to appear).

6. A. S. Shvedov, *Orders of coapproximation of functions by algebraic polynomials*, Mat. Zametki 29(1981), 117-130; English transl. in Math. Notes 29(1981), 63-70.