UNIVERSITY OF AUCKLAND
DEPARTMENT OF MATHEMATICS

On Bernstein's comparison theorem, Peano kernels of constant sign and near–minimax approximation

Shayne Waldron* (*waldron@math.auckland.ac.nz*)

Technical Report
May 1997

ABSTRACT

Some basic properties of what are called 'B(ernstein)–monotone' seminorms are investigated. These lie between the classes of monotone and sign–monotone seminorms. It is seen that these seminorms arise naturally in Bernstein's comparison theorem, the description of Peano kernels of constant sign, and in near–minimax approximations. A number of new results are obtained including some sufficient conditions for a projection to be near–minimax which are easily seen to be satisfied by all the known examples, and a characterisation of the Peano kernels of constant sign where derivatives are replaced by divided differences.

Key Words: monotone seminorm, B–monotone seminorm, sign–monotone seminorm, property B of order k, Bernstein's comparison theorem, Peano kernel, divided difference, Hermite interpolation, Steffensen's differentiation formula, positive linear operator, Bernstein operator, variation diminishing splines, near–minimax approximation, Chebyshev polynomial

AMS (MOS) Subject Classifications: primary 41A10, 41A65, 41A80, secondary 41A05, 41A55, 46E99, 65J05

* This work was supported by the Israel Council for Higher Education.

1. Introduction

In this section we give some basic properties of what are called 'B(ernstein)–monotone' seminorms. These lie between the classes of monotone and sign–monotone seminorms. In the rest of this paper it is shown how these seminorms arise naturally in Bernstein's comparison theorem, the description of Peano kernels of constant sign, and in near–minimax approximations. In Section 2 we give a characterisation of the Peano kernels of constant sign. The analoguous result where divided differences at distinct points replace derivatives is obtained in Section 4. In Section 3 sufficient conditions for a projection to be near minimax which are easily seen to be satisfied by all known examples are given.

B–monotone seminorms

Here we consider seminorms defined on $C(T)$, where T is some topological space. This covers the cases of interest in this paper, namely $T = [a, b]$, and to deal with divided differences (at distinct points)

$$T = \{(t_0, \ldots, t_{k+1}) : a \leq t_0 < t_1 < \cdots < t_{k+1} \leq b\}. \tag{1.1}$$

Definition 1.2. *A seminorm $E : C(T) \to \mathbb{R}^+$ is said to be monotone if*

$$|f| \leq |g| \quad \Longrightarrow \quad E(f) \leq E(g),$$

more generally B–monotone if it satisfies (the weaker condition)

$$|f| \leq g \quad \Longrightarrow \quad E(f) \leq E(g),$$

and even more generally sign–monotone if

$$0 \leq f \leq g \quad \Longrightarrow \quad E(f) \leq E(g).$$

Further, E is said to be strictly B–monotone if

$$|f| < g \quad \Longrightarrow \quad E(f) < E(g),$$

and strictly sign–monotone if

$$0 \leq f < g \quad \Longrightarrow \quad E(f) < E(g).$$

A strictly B–monotone seminorm is B–monotone, since

$$\begin{aligned} |f| \leq g \quad &\Longrightarrow \quad |f| < g + \varepsilon, \quad \forall \varepsilon > 0 \quad \Longrightarrow \quad E(f) < E(g) + \varepsilon E(1), \quad \forall \varepsilon > 0 \\ &\Longrightarrow \quad E(f) \leq E(g), \end{aligned} \tag{1.3}$$

and similarly a strictly sign–monotone seminorm is sign–monotone. If T is compact, which (1.1) isn't, then the classes of strictly B–monotone and B–monotone seminorms coincide, since if E is B–monotone

$$|f| < g \quad \Longrightarrow \quad |f| \leq \theta g, \quad \theta := \sup \frac{|f|}{g} < 1 \quad \Longrightarrow \quad E(f) \leq \theta E(g) < E(g). \tag{1.4}$$

Similarly, the classes of strictly sign–monotone and sign–monotone seminorms are the same when T is compact.

It immediately follows that if E is B–monotone, then

$$E(f) \leq E(|f|), \tag{1.5}$$

and if E is monotone, then

$$E(f) = E(|f|).$$

If E is sign–monotone, then $f \mapsto E(|f|)$ is a monotone seminorm, and so a sign–monotone seminorm is monotone if and only if $E(f) = E(|f|)$. From this observation it is easily seen that there are B–monotone seminorms which are not monotone, e.g.,

$$E : C[a,b] \to \mathbb{R}^+ : f \mapsto \left| \int_a^b f(t)\, dt \right|,$$

or, more generally, see Lemma 1.9. Further, from (1.5) it follows that there are sign–monotone seminorms which are not B–monotone. For example (as in Kimchi and Dyn [KD79] where sign–monotone norms were introduced), let $N : C[a,b] \to \mathbb{R}^+$ be the sign–monotone norm given by

$$N(f) := \|f^+\|_\infty + \|f^-\|_\infty, \tag{1.6}$$

where f^+, f^- are the positive and negative parts of f. If $\|f^+\|_\infty, \|f^-\|_\infty \neq 0$, then

$$N(f) = \|f^+\|_\infty + \|f^-\|_\infty > \|f\|_\infty = N(|f|),$$

and so N is not B–monotone.

B–monotone seminorms arise in connection with Peano kernels of constant sign and near–minimax approximations because of the following consequence of the intermediate value theorem.

Lemma 1.7. *Suppose that E is a strictly B–monotone seminorm defined on a subspace of $C(T)$, where T is a connected topological space. Then it satisfies the equivalent conditions*

(a) *For $\forall f$, $\exists t \in T$ with*

$$E(f) = |f(t)|E(1).$$

(b) *There $\exists A$ such that $\forall f$, $\exists t \in T$ with*

$$E(f) = |f(t)|A.$$

Proof: Choosing $f = 1$ in the second equation gives $E(1) = A$, which proves the equivalence of (a) and (b). Let

$$m := \inf_{\xi \in [a,b]} |f(\xi)|, \qquad M := \sup_{\xi \in [a,b]} |f(\xi)|.$$

Case 1. Suppose that f has a zero. If M is attained, then

$$|f| \leq M \quad \Longrightarrow \quad mE(1) = 0 \leq E(f) \leq ME(1),$$

and by the intermediate value theorem there is $t \in T$ with $E(f) = |f(t)|E(1)$. If M is not attained, then

$$|f| < M \quad \Longrightarrow \quad mE(1) = 0 \le E(f) < ME(1),$$

and since $f(T) = [m, M)$ there is $t \in T$ with $E(f) = |f(t)|E(1)$ (the intermediate value theorem).
Case 2. Suppose that f has no zeroes. Since $E(f) = E(-f)$ we may assume wlog that $f > 0$. Depending on whether m, M are attained or not, we have either

$$\begin{array}{rcl} 0 < m \le f \le M & \Longrightarrow & mE(1) \le E(f) \le ME(1), \\ 0 \le m < f \le M & \Longrightarrow & mE(1) < E(f) \le ME(1), \\ 0 < m \le f < M & \Longrightarrow & mE(1) \le E(f) < ME(1), \\ 0 \le m < f < M & \Longrightarrow & mE(1) < E(f) < ME(1), \end{array}$$

and the intermediate value theorem can be applied in each case. □

A similar result for $T = [a, b]$ is given in Holland, Phillips and Taylor [HPT80:Th.1]. The sign–monotone norm given by (1.6) does not satisfy the conditions (a),(b) of Lemma 1.7. The author does not know whether the conditions (a),(b) are sufficient to characterise the class of B–monotone seminorms. We now show that for certain seminorms at least, this is the case.

Definition 1.8. *A linear functional $L : C(T) \to \mathbb{R}$ is said to be positive if*

$$f \ge 0 \quad \Longrightarrow \quad L(f) \ge 0,$$

strictly positive if

$$f > 0 \quad \Longrightarrow \quad L(f) > 0,$$

and is negative, strictly negative if $-L$ is positive, strictly positive, respectively.

It can easily be shown that strictly positive linear functionals are positive, and that if T is compact then the classes of strictly positive and positive linear functionals coincide (with the same results holding for negative linear functionals).

Lemma 1.9. *Suppose that L is a nonzero linear functional defined on a subspace of $C(T)$, where T is a connected topological space. Let E be the seminorm given by*

$$E : f \mapsto |Lf|.$$

Then the following are equivalent
(a) *E is strictly B–monotone.*
(b) *E is strictly sign–monotone.*
(c) *L is strictly positive or strictly negative.*
(d) *For $\forall f$, $\exists t \in T$ with*

$$E(f) = |f(t)|E(1).$$

(e) *There* $\exists A$ *such that* $\forall f$, $\exists t \in T$ *with*

$$E(f) = |f(t)|A.$$

(f) *For* $\forall f$, $\exists t \in T$ *with*

$$L(f) = f(t)L(1).$$

(g) *There* $\exists B$ *such that* $\forall f$, $\exists t \in T$ *with*

$$L(f) = f(t)B.$$

Furthermore, if T *is a subset of* $\mathbb{R}^n$, *then* E *is monotone if and only if the support of* L *consists of exactly one point, i.e.,* E *is of the form*

$$E : f \mapsto c|f(t)|, \qquad c > 0,\ t \in T.$$

Proof: (a)$\Longrightarrow$(b) Immediate.

(b)$\Longrightarrow$(c) By contraposition. Suppose L is not strictly positive or strictly negative, and (by considering $-L$ if necessary) choose $f, g > 0$ with $L(f) \geq 0$ and $L(g) < 0$. If $L(f) = 0$, then E is not strictly sign–monotone, so suppose $L(f) > 0$ and choose $\beta > 0$ such that $L(f + \beta g) = L(f) + \beta L(g) = 0$. Since

$$0 \leq f < f + \beta g, \qquad E(f) > E(f + \beta g) = 0,$$

the seminorm E is again not strictly sign–monotone.

(a)$\Longleftarrow$(c) Since $|Lf| = |-Lf|$, assume wlog that L is strictly positive. Then

$$|f| < g \quad \Longrightarrow \quad L(|f|) < L(g),$$
$$\pm f \leq |f| \quad \Longrightarrow \quad \pm Lf \leq L(|f|),$$

giving

$$E(f) = |Lf| \leq L(|f|) < L(g) = E(g),$$

and so E is strictly B–monotone.

(a)$\Longrightarrow$(d)$\Longleftrightarrow$(e) Follows from Lemma 1.7.

(c)$\Longleftarrow$(d) By contraposition. Suppose L is not strictly positive or strictly negative, and (by considering $-L$ if necessary) choose $f, g > 0$ with $L(f) \geq 0$ and $L(g) < 0$. If $L(f) = 0$, then (d) implies $E(1) = 0 \Longrightarrow E = 0$, a contradiction. If on the otherhand $L(f) > 0$, choose $\beta > 0$ such that $L(f + \beta g) = 0$, then using the function $f + \beta g > 0$ obtain the same contradiction as before.

(c)$\Longrightarrow$(f) Suppose wlog that L is strictly positive, then L is a strictly B–monotone seminorm and Lemma 1.7 can be applied.

(c)$\Longleftarrow$(f) Since L is nonzero, $L(f) = f(t)L(1)$ implies that $L(1) \neq 0$, and so if $\pm L(1) > 0$, then $\pm L$ is strictly positive.

(f)$\Longleftrightarrow$(g) Choose $f = 1$ in $L(f) = f(t)B$ to obtain $L(1) = B$.

Finally, suppose E is monotone, and the support of L contains two points. Then it is possible to choose f, g with disjoint supports and $L(f), L(g) \neq 0$. Let $\beta \neq 0$ be such that $L(f + \beta g) = Lf + \beta Lg = 0$. Since

$$|f| \leq |f + \beta g|, \qquad E(f) > E(f + \beta g) = 0,$$

this contradicts E being monotone, and so E is monotone if and only if the support of L is empty ($L = 0$), or consists of one point (L is a scalar multiple of point evaluation at that point, which clearly gives a monotone seminorm). $\square$

This result forms the basis for the characterisation of Peano kernels of constant sign given next.

2. Peano kernels of constant sign

In this section quotient theorems for seminorms and linear functionals (Peano's kernel theorem) are used to translate the properties of B–monotone seminorms into results about seminorms that satisfy 'property B of order k' and Peano kernels of constant sign.

Definition 2.1. *A seminorm $E : C^k[a,b] \to \mathbb{R}^+$ is said to satisfy property B of order k (cf Holland, Phillips and Taylor [HPT80]) if*

$$|D^k f| \le D^k g \quad \Longrightarrow \quad E(f) \le E(g). \tag{2.2}$$

The term 'property B' was used because of the connection with *Bernstein's comparison theorem* [B26:p8] which states that the seminorm

$$E(f) := \inf_{p \in \Pi_k} \|f - p\|_\infty \qquad (\Pi_k \text{ the polynomials of degree} \le k) \tag{2.3}$$

satisfies property B of order $k+1$. Our use of the term property B differs slightly from the definition given in [HPT80], where for $k \ge 1$ the E of (2.2) would be said to satisfy property B of order $k-1$. The seminorms that satisfy property B of order 0 (in our sense) are precisely the B–monotone seminorms, hence our choice of the term B(ernstein)–monotone.

Quotient theorem (for seminorms). *There is a 1–1 correspondence between the seminorms $E : C^{k+1}[a,b] \to \mathbb{R}^+$ that vanish on Π_k and the seminorms $Q : C[a,b] \to \mathbb{R}^+$ which is given by*

$$E(f) = Q(D^{k+1} f), \qquad \forall f.$$

In particular, the seminorms that satisfy property B of order $k+1$ vanish on Π_k, and correspond to the B–monotone seminorms.

Proof: Since $D^{k+1} : C^{k+1}[a,b] \to C[a,b]$ is onto, with Π_k in its kernel, the function

$$Q : C[a,b] \to \mathbb{R}^+ : D^{k+1} f \mapsto E(f), \qquad \forall f \in C^{k+1}[a,b]$$

is well defined and is a seminorm. Conversely for a given seminorm $Q : C[a,b] \to \mathbb{R}^+$ the map $E := Q \circ D^{k+1} : C^{k+1}[a,b] \to \mathbb{R}^+$ is a seminorm which vanishes on Π_k. This establishes the 1–1 correspondence.

If E satisfies property B of order $k+1$, then

$$|D^{k+1} p| \le D^{k+1} 0, \ \forall p \in \Pi_k \quad \Longrightarrow \quad E(p) \le E(0) = 0,$$

so E vanishes on Π_k, and the corresponding Q is B–monotone. Conversely, if Q is B–monotone then $E := Q \circ D^{k+1}$ satisfies property B of order $k+1$. $\square$

Thus, the seminorms satifying property B of order $k+1$ are precisely those of the form $E(f) = |D^{k+1}f|$, $\forall f$, where $|\cdot|$ is a B–monotone seminorm.

Peano kernel theorem (cf Sard [S63:p.314]). *There is a 1–1 (linear) correspondence between the linear functionals $\lambda : C^{k+1}[a,b] \to \mathrm{I\!R}$ that vanish on Π_k and the linear functionals $L : C[a,b] \to \mathrm{I\!R}$ which is given by*

$$\lambda(f) = L(D^{k+1}f), \qquad \forall f.$$

In particular, the continuous linear functionals $\lambda \in (C^{k+1}[a,b])^*$ correspond to the continuous linear functionals $L \in (C[a,b])^*$.

Let $(\cdot)^k$ denote the monomial $x \mapsto x^k$.

Theorem 2.4 (Peano kernels of constant sign). *Suppose that $\lambda : C^{k+1}[a,b] \to \mathrm{I\!R}$ is a linear functional. Let E be the seminorm given by*

$$E : C^{k+1}[a,b] \to \mathrm{I\!R}^+ : f \mapsto |\lambda f|.$$

Then the following are equivalent

(a) *λ vanishes on Π_k, and the linear functional $L : C[a,b] \to \mathrm{I\!R}$ defined by $\lambda = L \circ D^{k+1}$ is positive or negative.*

(b) *E satisfies property B of order $k+1$.*

(c) *For $\forall f$, $\exists \xi \in [a,b]$ with*

$$\lambda(f) = \frac{D^{k+1}f(\xi)}{(k+1)!}\lambda\left((\cdot)^{k+1}\right).$$

(d) *For $\forall f$, $\exists \xi \in [a,b]$ with*

$$E(f) = \frac{|D^{k+1}f(\xi)|}{(k+1)!}E\left((\cdot)^{k+1}\right).$$

(e) *There $\exists A$ such that $\forall f$, $\exists \xi \in [a,b]$ with*

$$\lambda(f) = \frac{D^{k+1}f(\xi)}{(k+1)!}A.$$

(f) *There $\exists B$ such that $\forall f$, $\exists \xi \in [a,b]$ with*

$$E(f) = \frac{|D^{k+1}f(\xi)|}{(k+1)!}B.$$

(g) *There is a positive or negative linear map $L : C[a,b] \to \mathrm{I\!R}$ for which $\lambda = L \circ D^{k+1}$.*

(h) *There is a B–monotone seminorm $|\cdot|$ for which $E(f) = |D^{k+1}f|$, $\forall f$.*

(i) *There is a sign–monotone seminorm $|\cdot|$ for which $E(f) = |D^{k+1}f|$, $\forall f$.*

(j) *The map $f \mapsto \lambda(f)$ does not change sign on the set*

$$\{f \in C^{k+1}[a,b] : D^{k+1}f \geq 0\}.$$

Further, if $E : C^{k+1}[a,b] \to \mathbb{R}^+$ is any seminorm (not necessarily of the form $f \mapsto |\lambda f|$), then the implications

$$\text{(i)} \Longleftarrow \text{(h)} \Longleftrightarrow \text{(b)} \Longrightarrow \text{(d)} \Longleftrightarrow \text{(f)} \tag{2.5}$$

remain valid.

Proof: Recall that since $T = [a,b]$ is compact the strict and nonstrict versions of B–montone, positive, etc are equivalent. Since $f \in \Pi_k \Longrightarrow D^{k+1}f = 0$, each of (c),(d),(e),(f),(g),(h),(i) implies that λ and E vanish on Π_k. Thus, in view of the quotient and Peano kernel theorems, Lemma 1.9 gives the equivalence of all conditions except (j), and Lemma 1.7 gives the implications (2.5).
The equivalence of (j). The implication (a) $\Longrightarrow$ (j) is immediate. Suppose (j) holds, then

$$f \in \Pi_k \quad \Longrightarrow \quad \pm D^{k+1}f \geq 0 \quad \Longrightarrow \quad \pm\lambda(f) \geq 0, \text{ or } \pm\lambda(f) \leq 0 \quad \Longrightarrow \quad \lambda(f) = 0,$$

so λ vanishes on Π_k. Since λ does not change sign on $\{f : D^{k+1}f \geq 0\}$, the linear functional L defined by $\lambda = L \circ D^{k+1}$ is either positive or negative, which gives (a). $\square$

In [HPT84] the implications (a) $\Longleftrightarrow$ (b) $\Longleftrightarrow$ (c) were proved for linear functionals λ from a subspace of $(C^{k+1}[a,b])^*$ called the 'Peano functionals on $C^{k+1}[a,b]$'.

Suppose that any one of (a), (b), $\ldots$, (j) in Theorem 2.4 holds. If $\lambda \in (C^{k+1}[a,b])^*$, then the L for which $\lambda = L \circ D^{k+1}$ belongs to $(C[a,b])^*$, and hence can be represented as integration against a Riemann–Stieltjes measure dw given by a $w \in \mathrm{NBV}[a,b]$. For such an L the condition that it be positive or negative is equivalent to dw being a positive or negative measure, which is equivalent to w being a monotone function. Further, if $\lambda \in (C^k[a,b])^* \subset (C^{k+1}[a,b])^*$, then L can be represented as integration against a Peano kernel $K \in L_\infty[a,b]$, i.e.,

$$\lambda(f) = L(D^{k+1}f) = \int_a^b D^{k+1}f(t)K(t)\,dt, \qquad \forall f, \tag{2.6}$$

and the condition that L be positive or negative is equivalent to K being positive or negative. In this situation K is usually referred to as a Peano kernel of constant sign (or as one that does not change sign), and by some authors as definite of order $k+1$.

Example 1: The *error in quadrature rules exact for* Π_k. The error in many quadrature rules that are exact for Π_k (including the Gauss and Newton–Cotes rules) have Peano kernels of constant sign. This can be seen from the error estimates of the forms (c) and (d) that are given in many numerical analysis texts.

Example 2: The $(k+1)$*–th order divided differences* of f at a multiset of $k+1$ points in $[a,b]$

$$\lambda : f \mapsto [\theta_0, \ldots, \theta_{k+1}]f$$

satisfies the conditions (a),(c),(e),(g),(j). This is easily seen from the Hermite–Genocchi formula

$$\lambda(f) = \int_0^1 \int_0^{s_1} \cdots \int_0^{s_k} D^{k+1}f(\theta_0 + s_1(\theta_1 - \theta_0) + \cdots + s_{k+1}(\theta_{k+1} - \theta_k))\,ds_{k+1} \cdots ds_2\,ds_1,$$

which shows that L is positive. The Peano kernel K representing L (cf (2.6)) is the (*normalised*) *B–spline* of order $k+1$ with knots $\{\theta_0, \ldots, \theta_{k+1}\}$ (provided not all of the knots are the same).

Example 3: The *error in Hermite interpolation from* Π_k. Let $H_\Theta : C^{k+1}[a,b] \to \Pi_k$ be the operator of Hermite interpolation to the multiset $\Theta := \{\theta_0, \ldots, \theta_k\}$ of points in $[a,b]$, and $\omega_\Theta(x) := \prod_{j=0}^{k}(x-\theta_j)$. Since the pointwise error at $x \in [a,b]$

$$\lambda : f \mapsto f(x) - H_\Theta f(x) = \omega_\Theta(x)[\theta_0, \ldots, \theta_k, x]f$$

is a scalar multiple of a $(k+1)$–th order divided difference, it satisfies (a),(c),(e),(g),(j), and so

$$|D^{k+1}f| \le D^{k+1}g \quad \Longrightarrow \quad |f - H_\Theta f| \le |g - H_\Theta g|$$

(see, e.g., Kimichi and Dyn [KD78:Lem.2.2]). Thus, if $\|\cdot\|$ is any monotone seminorm, then

$$|D^{k+1}f| \le D^{k+1}g \quad \Longrightarrow \quad \|f - H_\Theta f\| \le \|g - H_\Theta g\|. \tag{2.7}$$

Example 4: The *Steffensen differentiation formula.* Let H_Θ be the Hermite interpolation operator as above. In Steffenson [St27:p64] it is shown that if $x \in [a, \min\Theta] \cup [\max\Theta, b]$, and $j = 0, \ldots, k+1$, then $\forall f$, $\exists \xi \in [a,b]$ with

$$D^j f(x) = D^j(H_\Theta f)(x) + \frac{D^{k+1}f(\xi)}{(k+1)!} D^j \omega_\Theta(x).$$

Thus, the linear functional

$$\lambda : f \mapsto D^j(f - H_\Theta f)(x)$$

satisfies (e), and so the seminorm

$$E : f \mapsto |D^j(f - H_\Theta f)(x)| \tag{2.8}$$

satisfies property B of order $k+1$, etc. A complete description of when (2.8) satisfies property B of order $k+1$ (in terms of the sign of the Peano kernel) was recently given in Shadrin [Sh95].

Example 5: The *error in best approximation from* Π_k *in a monotone norm.* Let $\|\cdot\|$ be a monotone norm, e.g., any of the $L_p[a,b]$–norms, and Pg be a best approximation from Π_k to g in this norm, i.e.,

$$E(g) := E_{k,\|\cdot\|}(g) := \inf_{p \in \Pi_k} \|g - p\| = \|g - Pg\|.$$

Since the error $g - Pg$ has at least $k+1$ zeroes counting multiplicities (this was proved in Kimichi and Dyn [KD79:Th.2.2] for any sign–monotone norm), i.e., $Pg = H_\Theta g$ for some multiset $\Theta := \{\theta_0, \ldots, \theta_k\}$ (depending on g), we have by (2.7) the following generalisation of Bernstein's comparison theorem

$$|D^{k+1} f| \le D^{k+1} g \quad \Longrightarrow \quad E(f) \le \|f - H_\Theta f\| \le \|g - H_\Theta g\| = \|g - Pg\| = E(g). \quad (2.9)$$

In other words, as proved in [KD78:Th.2.2], $E_{k,\|\cdot\|}$ satisfies property B of order $k+1$.

Example 6: The *error in approximation by a positive linear operator which reproduces the linear polynomials*, e.g., the Bernstein operators, or more generally Schoenberg's variation diminishing spline operators. An operator L is positive (monotone) if

$$f \ge 0 \quad \Longrightarrow \quad Lf \ge 0.$$

If $L : C[a,b] \to C[a,b]$ is a positive linear operator which reproduces the linear polynomials, then

$$f - Lf = -Q(D^2 f), \qquad \forall f \in C^2[a,b], \tag{2.10}$$

where $Q : C[a,b] \to C[a,b]$ is a *positive* linear operator.

The existence of a Q satisfying (2.10) follows (pointwise) by applying the Peano kernel theorem to the error functionals $f \mapsto f(x) - Lf(x)$, $x \in [a,b]$. That Q is positive, i.e., the Peano kernel of each error functional has constant (negative) sign, can be proved as follows. Suppose that $D^2 f \ge 0$. Then f is convex, and so p the tangent line to f at x lies below f, i.e., $p \le f$. Since L is positive and reproduces the linear polynomials (such as p)

$$p \le f \quad \Longrightarrow \quad p = Lp \le f.$$

Evaluating this at x gives $f(x) = p(x) \le Lf(x)$, so that $Q(D^2 f) = Lf - f \ge 0$, i.e., Q is positive.

Though the representation (2.10) (with Q positive) is not mentioned in the literature, the corresponding sharp error bounds

$$E(f,x) := |f(x) - Lf(x)| \le \alpha^2(x)\, \|D^2 f\|_\infty, \qquad \forall f \in C^2[a,b], \tag{2.11}$$

where

$$\alpha(x) := \frac{1}{2} L\big((\cdot - x)^2\big)(x) = \frac{1}{2} E((\cdot)^2, x)$$

are well known (see, e.g., DeVore [De72:Theorem 2.5, p. 39]). Other positive linear operators reproducing the linear polynomials for which (2.10), (2.11) and

$$E(f,x) := |f(x) - Lf(x)| = \alpha^2(x)\, |D^2 f(\xi)|, \qquad \exists \xi \in [a,b],$$

etc hold include (see, e.g., Păltănea [Pa88]) the *Meyer–König and Zeller operators* on $C[0,1]$, and the *Szasz–Mirakjan* and *Baskakov operators* on $C[0,\infty)$ (with $[0,\infty)$ replacing $[a,b]$ in the natural way).

3. Near–minimax approximations

In this section we give examples which are related to 'near–minimax' approximations.

Definition 3.1. *A linear projector P from $C[-1,1]$ (or $C^{k+1}[-1,1]$) onto Π_k is said to be near–minimax if for $\forall f \in C^{k+1}[-1,1]$, $\exists \xi \in [-1,1]$ with*

$$\|f - Pf\|_\infty := \|f - Pf\|_{L_\infty[-1,1]} = \frac{|D^{k+1} f(\xi)|}{2^k (k+1)!}. \tag{3.2}$$

This definition is due to Phillips [P92], whose motivation was that the 'minimax' error satisfies: $\forall f$, $\exists \xi$ with

$$\inf_{g \in \Pi_k} \|f - g\|_\infty = \frac{|D^{k+1} f(\xi)|}{2^k (k+1)!}$$

(which follows from Bernstein's comparison theorem and (b) $\Longrightarrow$ (f) of Theorem 2.4). There are a number of other notions of 'near–minimax' approximation in the literature, see, e.g., Mason [M73].

Example 7: In Phillips [P92] it is shown that Lagrange interpolation at the zeroes of T_{k+1} the *Chebyshev polynomial* of degree $k+1$ is a near–minimax approximation. Recall that the infimum

$$\inf_{p \in \Pi_k} \|(\cdot)^{k+1} - p\|_\infty = \frac{1}{2^k}$$

is achieved if and only if

$$(\cdot)^{k+1} - p = 2^{-k} T_{k+1}, \tag{3.3}$$

and that

$$\|T_{k+1}\|_\infty = 1. \tag{3.4}$$

In view of (2.7), the seminorm E given by

$$E(f) := \|f - H_\Theta f\|_\infty$$

satifies property B of order $k+1$, and so, by (b) $\Longrightarrow$ (d) of Theorem 2.4, the Hermite interpolation operator H_Θ (see Example 3) is near–minimax if and only if

$$\|(\cdot)^{k+1} - H_\Theta\left((\cdot)^{k+1}\right)\|_\infty = \|\omega_\Theta\|_\infty = \frac{1}{2^k},$$

which, by (3.3), holds if and only if $\omega_\Theta = 2^{-k} T_{k+1}$. Thus, Lagrange interpolation at the zeroes of the Chebyshev polynomial of degree $k+1$ is the *only* Hermite interpolation operator which is near–minimax. Even more generally, this shows the following:

Theorem 3.5 (sufficient conditions for near–minimax approximation). *Let P be any projector of $C^{k+1}[-1,1]$ onto Π_k (possibly nonlinear like the $\|\cdot\|_\infty$–best approximation operator) for which $E : f \mapsto \|f - Pf\|_\infty$ is a seminorm. If*

$$(\cdot)^{k+1} - P\left((\cdot)^{k+1}\right) = 2^{-k} T_{k+1}, \tag{3.6}$$

and $f - Pf$ has (at least) $k+1$ zeroes counting multiplicities for all f, then P is near–minimax (in the sense of Phillips [P92]). Notice that if P is linear, then

$$(3.6) \quad \Longrightarrow \quad P(T_{k+1}) = 0.$$

Proof: Suppose $f - Pf$ has $k+1$ zeroes Θ, then $Pf = H_\Theta f$, and by (2.7) the seminorm E satisfies property B of order $k+1$. Thus, by (b) $\Longleftrightarrow$ (d), (e) of Theorem 2.4, P is near–minimax if and only if

$$E\left((\cdot)^{k+1}\right) = \|(\cdot)^{k+1} - P\left((\cdot)^{k+1}\right)\|_\infty = \frac{1}{2^k},$$

which, by (3.3), holds only if $(\cdot)^{k+1} - P\left((\cdot)^{k+1}\right) = 2^{-k} T_{k+1}$.

If P is a linear projector, then applying P to both sides of (3.6) gives $0 = 2^{-k} P(T_{k+1})$. □

Example 8: The paper of Brass [Br84] shows the *truncated Chebyshev series* is near–minimax. An alternative proof of this fact was given independently by Elliot, Paget, Phillips and Taylor [EPPT87]. Since the error in the best approximation from Π_k in a weighted L_2–norm vanishes $k+1$ times, it is easily seen that the truncated Chebyshev series satisfies the conditions of Theorem 3.5.

Example 9: In Phillips and Taylor [PT82] it is shown that the projector $f \mapsto p \in \Pi_k$ obtained by choosing the unique $p \in \Pi_k$ for which the error $f - p$ equioscillates on the point set consisting of the $k+2$ extrema of T_{k+1} is near–minimax. Again it is easily seen that this (nonlinear) projector satisfies the conditions of Theorem 3.5.

Example 10: In the paper of Chalmers, Phillips and Taylor [CPT88] two necessary conditions for a continuous *linear* projector $P : C[-1,1] \to \Pi_k$ to be near–minimax are given. The first of these is that

$$P(T_{k+1}) = 0. \tag{3.7}$$

In our next result we show that infact the stronger condition (3.6) of Theorem 3.5 is necessary for *all* projectors $P : C^{k+1}[-1,1] \to \Pi_k$ (these need not be linear or continuous). It is also shown that a stronger form of the second condition given is necessary.

Theorem 3.8 (necessary conditions for near–minimax approximation). *For a (possibly nonlinear) projector $P : C^{k+1}[-1,1] \to \Pi_k$ to be near–minimax it is necessary that*

$$(\cdot)^{k+1} - P\left((\cdot)^{k+1}\right) = 2^{-k} T_{k+1}. \tag{3.9}$$

Further, if P is a continuous linear projector, then for P to be near–minimax each of the (continuous) linear functionals

$$\lambda : C^{k+1}[-1,1] \to \mathrm{I\!R} : f \mapsto f(z_i) - Pf(z_i), \qquad z_0 < z_1 < \cdots < z_{k+1} \text{ the extrema of } T_{k+1}$$

must satisfy the (equivalent) conditions of Theorem 2.4 (with the corresponding L changing between being positive/negative for consecutive extrema z_i and z_{i+1}). This implies, for example, that $f - Pf$ must have $k+1$ zeroes (counting multiplicities) for all f from the set $\{f : D^{k+1}f \geq 0\}$.

Proof: Suppose that P is near–minimax. Substituting $f = (\cdot)^{k+1}$ into (3.2) gives

$$\|(\cdot)^{k+1} - P\left((\cdot)^{k+1}\right)\|_\infty = \frac{1}{2^k},$$

which, by (3.3), implies (3.9).

Now suppose P is also a continuous linear map, and so each of

$$\lambda_x : C^{k+1}[-1,1] \to \mathrm{I\!R} : f \mapsto f(x) - Pf(x), \qquad x \in [-1,1] \tag{3.10}$$

is a continuous linear functional (which vanishes on Π_k). By the Peano kernel theorem of Section 2, each λ_x has a (unique) representation of the form

$$\lambda_x(f) = L_x(D^{k+1}f), \qquad \forall f,$$

where $L_x : C[-1,1] \to \mathrm{I\!R}$ is a continuous linear functional, which can therefore be represented by a Riemann–Stieltjes integral

$$L_x(f) = \int_{-1}^{1} f(t)\,dw_x(t), \qquad \forall f, \qquad w_x \in \mathrm{BV}[-1,1]$$

with

$$\|L_x\| = \mathrm{Var}(w_x) \qquad \text{(the total variation of } w_x).$$

Since $|\lambda_x(f)| \leq \|f - Pf\|_\infty$, the near–minimax condition (3.2) implies that

$$\mathrm{Var}(w_x) = \|L_x\| \leq \frac{1}{2^k(k+1)!}. \tag{3.11}$$

Also, by (3.9),

$$\|L_x\| \geq \frac{|\lambda_x\left((\cdot)^{k+1}\right)|}{\|D^{k+1}(\cdot)^{k+1}\|_\infty} = \frac{2^{-k}|T_{k+1}(x)|}{(k+1)!}, \tag{3.12}$$

and

$$\frac{|\lambda_x\left((\cdot)^{k+1}\right)|}{\|D^{k+1}(\cdot)^{k+1}\|_\infty} = \left|\int_{-1}^{1} dw_x(t)\right| = |w_x(b) - w_x(a)| = |\mathrm{Var}^+(w_x) - \mathrm{Var}^-(w_x)|. \tag{3.13}$$

Here $\mathrm{Var}^+(w_x)$, $\mathrm{Var}^-(w_x)$ denote the positive and negative variations of w_x. Since (3.4) implies that $|T_{k+1}(z_i)| = 1$, it follows from (3.11),(3.12),(3.13) that

$$\mathrm{Var}(w_x) = \mathrm{Var}^+(w_x) + \mathrm{Var}^-(w_x) = |\mathrm{Var}^+(w_x) - \mathrm{Var}^-(w_x)| = \frac{1}{2^k(k+1)!}, \qquad \forall x \in \{z_i\}.$$

Thus, when $x \in \{z_i\}_{i=0}^{k+1}$ either the positive or negative variation of w_x must be zero, i.e., w_x is monotone, and so L_x is either a positive or negative linear functional. This shows that $\lambda : f \mapsto f(z_i) - Pf(z_i)$ satisfies condition (a) of Theorem 2.4, with L_{z_i} being positive when $T_{k+1}(z_i) = 1$, and negative otherwise.

Finally, suppose that $D^{k+1}f \geq 0$. By condition (d) of Theorem 2.4 and (3.9), for each z_i

$$f(z_i) - Pf(z_i) = \frac{D^{k+1}f(\xi)}{(k+1)!} 2^{-k} T_{k+1}(z_i), \qquad \exists \xi \in [-1,1],$$

which (since $T_{k+1}(z_i) = (-1)^{k+1-i}$) implies

$$(-1)^{k+1-i}\left(f(z_i) - Pf(z_i)\right) \geq 0, \qquad i = 0, \ldots, k+1, \tag{3.14}$$

and so $f - Pf$ must have at least $k+1$ zeroes counting multiplicities (including at least one zero in each of the intervals $[z_i, z_{i+1}]$, $i = 0, \ldots, k$, with any zero at $z_1, \ldots, z_k$ being of double multiplicity). $\square$

In Chalmers, Phillips and Taylor [CPT88] the second necessary condition for a continuous linear projector $P : C[-1,1] \to \Pi_k$ to be near–minimax is given as:

(ii) *For each $t \in [-1,1]$, with $f := g_t := (\cdot - t)_+^k/k!$, the error $f - Pf$ must have $k+1$ zeroes $x_i \in [z_i, z_{i+1}]$ (counting multiplicities).*

By Theorem 3.8, the Peano kernel for λ_x of (3.10) which is given by

$$K_x : t \mapsto g_t(x) - Pg_t(x)$$

is of constant sign for $x \in \{z_i\}$ (and this sign changes for consecutive extrema z_i and z_{i+1}). Thus, with $f := g_t$, the error $f - Pf$ satisfies (3.14), and (as in the proof of Theorem 3.8) it must therefore have $k+1$ zeroes of the form given in (ii) above.

4. Divided difference formulation

In this section we give a variation of Theorem 2.4 which uses divided differences of order $k+1$ instead of derivatives of order $k+1$. This generalises a number of results in the literature including a classical theorem of Popoviciu [Po59] and recent work of Elliott and Taylor [ET91].

Let T be the collection of (ordered) sequences of $k+2$ distinct points in $[a,b]$,

$$T := T(k+1, [a,b]) := \{(t_0, \ldots, t_{k+1}) : a \leq t_0 < t_1 < \cdots < t_{k+1} \leq b\},$$

and let $\mathcal{D}^{k+1} : C[a,b] \to \mathbb{R}^T$ be the linear operator which associates with each $f \in C[a,b]$ the collection of its divided differences of order $k+1$

$$\mathcal{D}^{k+1} f : T \to \mathbb{R} : (t_0, \ldots, t_{k+1}) \mapsto [t_0, \ldots, t_{k+1}] f.$$

This will replace the linear operator $D^{k+1} : C^{k+1}[a,b] \to C[a,b]$ in Theorem 2.4 which associates with each $f \in C^{k+1}[a,b]$ its $k+1$ order derivative $D^{k+1} f \in C[a,b]$. Notice that the range of $\mathcal{D}^{k+1}$ consists of continuous functions on T (with the Euclidean metric), i.e.,

$$\operatorname{ran}(\mathcal{D}^{k+1}) \subset C(T)$$

(by the well known property that a divided difference is a continuous function of its distinct arguments). The kernel of $\mathcal{D}^{k+1}$ is Π_k, since if all the $k+1$ order divided differences of the function $f \in C[a,b]$ vanish, then $p \in \Pi_k$ its Lagrange interpolant at the $k+1$ points $a \le t_0 < \cdots < t_k \le b$ satisfies

$$f(t) - p(t) = (t - t_0)(t - t_1) \cdots (t - t_k) [t_0, t_1, \ldots, t_k, t] f = 0, \qquad \forall t \notin \{t_0, \ldots, t_k\},$$

and so $f = p \in \Pi_k$. Since the kernel of $\mathcal{D}^{k+1}$ is Π_k, similarly to the case of D^{k+1}, we obtain the following factorisation theorems.

Peano kernel theorem (for $\mathcal{D}^{k+1}$). *There is a 1–1 (linear) correspondence between the linear functionals $\lambda : C[a,b] \to \mathbb{R}$ that vanish on Π_k and the linear functionals $\mathcal{L}$: $\operatorname{ran}(\mathcal{D}^{k+1}) \to \mathbb{R}$ which is given by $\lambda = \mathcal{L} \circ \mathcal{D}^{k+1}$. (cf Sard [S63:p.311])*

Seminorm quotient theorem (for $\mathcal{D}^{k+1}$). *There is a 1–1 correspondence between the seminorms $E : C[a,b] \to \mathbb{R}^+$ that vanish on Π_k and the seminorms $\mathcal{Q} : \operatorname{ran}(\mathcal{D}^{k+1}) \to \mathbb{R}^+$ which is given by $E = \mathcal{Q} \circ \mathcal{D}^{k+1}$. In particular, the seminorms that satisfy the strict property B of order $k+1$, i.e.,*

$$\big|[t_0, \ldots, t_{k+1}] f\big| < [t_0, \ldots, t_{k+1}] g, \quad \forall a \le t_0 < \cdots < t_{k+1} \le b \quad \Longrightarrow \quad E(f) < E(g),$$

vanish on Π_k, and correspond to the strictly B–monotone seminorms.

Though T is bounded and connnected, it is not closed (e.g., the points $(t_0, \ldots, t_0)$, $t_0 \in [a,b]$ lie on its boundary). It is for this reason the "strict" forms of Lemmas 1.7 and 1.9 were proved, which allows the relevant functionals

$$L : \operatorname{ran}(\mathcal{D}^{k+1}) \to \mathbb{R}$$

to be handled.

Theorem 4.1. *Suppose that $\lambda : C[a,b] \to \mathbb{R}$ is a linear functional. Let E be the seminorm given by*

$$E : C[a,b] \to \mathbb{R}^+ : f \mapsto |\lambda f|.$$

Then the following are equivalent

(a) λ *vanishes on* Π_k, *and the linear functional* $\mathcal{L} : \mathrm{ran}(\mathcal{D}^{k+1}) \to \mathbb{R}$ *defined by* $\lambda = \mathcal{L} \circ \mathcal{D}^{k+1}$ *is strictly positive or strictly negative.*

(b) E *satisfies*

$$\big|[t_0, \dots, t_{k+1}]f\big| < [t_0, \dots, t_{k+1}]g, \quad \forall a \le t_0 < \cdots < t_{k+1} \le b \quad \Longrightarrow \quad E(f) < E(g).$$

(c) *For* $\forall f \in C[a,b]$, $\exists a \le t_0 < t_1 < \cdots < t_{k+1} \le b$ *with*

$$\lambda(f) = [t_0, \dots, t_{k+1}]f\, \lambda\left((\cdot)^{k+1}\right).$$

(d) *For* $\forall f \in C[a,b]$, $\exists a \le t_0 < t_1 < \cdots < t_{k+1} \le b$ *with*

$$E(f) = \big|[t_0, \dots, t_{k+1}]f\big|\, E\left((\cdot)^{k+1}\right).$$

(e) *There* $\exists A$ *such that* $\forall f \in C[a,b]$, $\exists a \le t_0 < t_1 < \cdots < t_{k+1} \le b$ *with*

$$\lambda(f) = [t_0, \dots, t_{k+1}]f\, A.$$

(f) *There* $\exists B$ *such that* $\forall f \in C[a,b]$, $\exists a \le t_0 < t_1 < \cdots < t_{k+1} \le b$ *with*

$$E(f) = \big|[t_0, \dots, t_{k+1}]f\big|\, B.$$

(g) *There is a strictly positive or strictly negative linear map* $\mathcal{L} : \mathrm{ran}(\mathcal{D}^{k+1}) \to \mathbb{R}$ *for which* $\lambda = \mathcal{L} \circ \mathcal{D}^{k+1}$.

(h) *There is a strictly B–monotone seminorm* $|\!\cdot\!| : \mathrm{ran}(\mathcal{D}^{k+1}) \to \mathbb{R}^+$ *for which* $E(f) = |\mathcal{D}^{k+1} f|$, $\forall f \in C[a,b]$.

(i) *There is a strictly sign–monotone seminorm* $|\!\cdot\!| : \mathrm{ran}(\mathcal{D}^{k+1}) \to \mathbb{R}^+$ *for which* $E(f) = |D^{k+1} f|$, $\forall f \in C[a,b]$.

(j) *The map* $f \mapsto \lambda(f)$ *is of one sign on the set*

$$\{f \in C[a,b] : [t_0, \dots, t_{k+1}]f > 0, \forall a \le t_0 < t_1 < \cdots < t_{k+1} \le b\}.$$

The conditions above imply the (equivalent) conditions (a), ..., (j) *of Theorem 2.4. Further, if* $E : C[a,b] \to \mathbb{R}^+$ *is any seminorm (not necessarily of the form* $f \mapsto |\lambda f|$*), then the implications*

$$\text{(i)} \Longleftarrow \text{(h)} \Longleftrightarrow \text{(b)} \Longrightarrow \text{(d)} \Longleftrightarrow \text{(f)} \tag{4.2}$$

remain valid.

Proof: Since $f \in \Pi_k \Longrightarrow \mathcal{D}^{k+1} f = 0$, each of (c),(d),(e),(f),(g),(h),(i) implies that λ and E vanish on Π_k. Thus, in view of the quotient and Peano kernel theorems for $\mathcal{D}^{k+1}$, Lemma 1.9 gives the equivalence of all conditions except (j), and Lemma 1.7 gives the implications (4.2).

The equivalence of (j). The implication (a) $\Longrightarrow$ (j) is immediate. Suppose (j) holds, then

$$f \in \Pi_k \quad \Longrightarrow \quad \pm \mathcal{D}^{k+1} f \ge 0 \quad \Longrightarrow \quad \pm\lambda(f) \ge 0, \text{ or } \pm\lambda(f) \le 0 \quad \Longrightarrow \quad \lambda(f) = 0,$$

so λ vanishes on Π_k. Since λ is of one sign on $\{f : \mathcal{D}^{k+1} f \ge 0\}$, the linear functional $\mathcal{L}$ defined by $\lambda = \mathcal{L} \circ \mathcal{D}^{k+1}$ is either positive or negative, which gives (a).

For $f \in C^{k+1} f[a,b]$, it is well known (cf Example 2) that

$$[t_0, \dots, t_{k+1}]f = \frac{D^{k+1} f(\xi)}{(k+1)!}, \qquad \exists \xi \in [t_0, t_{k+1}] \subset [a,b],$$

and so, for example, the equivalence (c) above implies (c) of Theorem 2.4. $\square$

Example 11: In Popoviciu [Po59] a function $f \in C[a,b]$ is said to be *convex of order k on* $[a,b]$ if $\mathcal{D}^{k+1}f > 0$, i.e.,

$$[t_0, \ldots, t_{k+1}]f > 0, \quad \forall a \le t_0 < t_1 < \cdots < t_{k+1} \le b,$$

and a linear functional $\lambda : C[a,b] \to \mathbb{R}$ is *of the simple form* if condition (e) holds, i.e., $\exists A \neq 0$ (independent of f) for which $\forall f$, $\exists a \le t_0 < t_1 < \cdots < t_{k+1} \le b$ with

$$\lambda(f) = [t_0, \ldots, t_{k+1}]f\, A.$$

There it is proved [Po59:Theorem 5] that a linear functional is of the simple form if and only if λ does not vanish on the convex functions of order k. This is the implications (e) $\iff$ (j) of Theorem 4.1.

Example 12: The *error in best approximation from Π_k in a monotone norm* $\|\cdot\|$ (see Example 5). As in Example 3,

$$\begin{aligned} &|[t_0, \ldots, t_{k+1}]f| < [t_0, \ldots, t_{k+1}]g, \quad \forall a \le t_0 < t_1 < \cdots < t_{k+1} \le b \\ &\quad \Longrightarrow \quad |f - H_\Theta f| < |g - H_\Theta g| \\ &\quad \Longrightarrow \quad \|f - H_\Theta f\| < \|g - H_\Theta g\|, \end{aligned}$$

and so, with the notation of (2.9)

$$\begin{aligned} &|[t_0, \ldots, t_{k+1}]f| < [t_0, \ldots, t_{k+1}]g, \quad \forall a \le t_0 < t_1 < \cdots < t_{k+1} \le b \\ &\quad \Longrightarrow \quad E(f) \le \|f - H_\Theta f\| < \|g - H_\Theta g\| = \|g - Pg\| = E(g), \end{aligned} \tag{4.3}$$

in other words $E_{k,\|\cdot\|}$ satisfies the condition (b) of Theorem 4.1. By (b) $\Longrightarrow$ (d) we conclude that $\forall f \in C[a,b]$, $\exists a \le t_0 < t_1 < \cdots < t_{k+1} \le b$ with

$$E_{k,\|\cdot\|}(f) = |[t_0, \ldots, t_{k+1}]f|\, E_{k,\|\cdot\|}\big((\cdot)^{k+1}\big).$$

This result was proved in Elliott and Taylor [ET91:Theorem 2.1] for the case when $\|\cdot\|$ is the uniform norm.

Example 13: The *error in near–minimax approximation.* If P is a linear projector of $C[-1,1]$ onto Π_k which for all $f \in C[-1,1]$ satisfies the sufficient conditions for near–minimax approximation given by Theorem 3.5, then, using the notation of its proof, it follows using (4.3) that $E : f \mapsto \|f - Pf\|_\infty$ satisfies condition (b). Hence by (b) $\Longrightarrow$ (d) it follows that $\forall f \in C[-1,1]$, $\exists -1 \le t_0 < t_1 < \cdots < t_{k+1} \le 1$ with

$$\|f - Pf\|_\infty = 2^{-k}|[t_0, \ldots, t_{k+1}]f|. \tag{4.4}$$

For P equal to the truncated Chebyshev series operator (see Example 8) it is shown in Elliott and Taylor [ET91:Theorem 4.1] that (4.4) holds for every $f \in C^2[-1,1]$ with points $-1 \le t_0 \le t_1 \le \cdots \le t_{k+1} \le 1$ of which at least k are distinct.

References

[B26] S. BERNSTEIN, Leçons sur les propriétés extrémales et la meilleure approximation des fonctions analytiques d'une variable réele, Gauthier–Villars, Paris, 1926.

[Br84] H. BRASS, Error estimates for least squares approximation by polynomials, *J. Approx. Theory* **41**(1984), 345–349.

[CPT88] B. L. CHALMERS, G. M. PHILLIPS AND P. J. TAYLOR, Polynomial approximation using projections whose kernels contain the Chebyshev polynomials, *J. Approx. Theory* **53**(1988), 321–334.

[De72] R. A. DEVORE, The Approximation of Continuous Functions by Positive Linear Operators, Springer–Verlag, Berlin, 1972.

[D79] N. RICHTER–DYN, On best nonlinear approximation in sign–monotone norms and in norms induced by inner products, *SIAM J. Numer. Anal.* **16(4)**(1979), 612–622.

[EPPT87] D. ELLIOT, D. F. PAGET, G. M. PHILLIPS AND P. J. TAYLOR, Error of truncated Chebshev series and other near minimax polynomial approximations, *J. Approx. Theory* **50**(1987), 49–57.

[ET91] D. ELLIOTT AND P. J. TAYLOR, Polynomial approximation errors for functions of low-order continuity, *Constr. Approx.* **7**(1991), 381–387.

[HPT80] A. S. B. HOLLAND, G. M. PHILLIPS AND P. J. TAYLOR, Generalisation of a theorem of Bernstein, In: Approximation Theory III (E. W. Cheney, Ed), pp. 513–516, Academic Press, New York, 1980.

[HPT84] A. S. B. HOLLAND, G. M. PHILLIPS AND P. J. TAYLOR, A relation between Peano kernels and a theorem of Berstein, *J. Approx. Theory* **41**(1984), 386–389.

[KD78] E. KIMCHI AND N. RICHTER-DYN, Restricted range approximation of k-convex functions in monotone norms, *SIAM J. Numer. Anal.* **15**(1978), 1030–1038.

[KD79] E. KIMCHI AND N. RICHTER-DYN, A necessary condition for best approximation in monotone and sign–monotone norms, *J. Approx. Theory* **25**(1979), 169–175.

[M73] J.C. MASON, Some methods of near–minimax approximation using Laguerre polynomials, *SIAM J. Numer. Anal.* **10**(1973), 470–477.

[Pa88] R. PĂLTĂNEA, Improved estimates with the second order modulus of continuity in approximation by linear positive operators, *Anal. Numér. Théor. Approx.* **17(2)**(1988), 171–179.

[P92] G. M. PHILLIPS, Error estimates for near-minimax approximations, In: Approximation Theory, Spline Functions and Applications (S. P. Singh, Ed), pp. 223–241, Kluwer, Dordrecht, Netherlands, 1992.

[PT82] G. M. PHILLIPS AND P. J. TAYLOR, Polynomial approximation using equioscillation on the extreme points of Chebyshev polynomials, *J. Approx. Theory* **36**(1982), 257–264.

[Po59] T. POPOVICIU, Sur le reste dans certaines formules linéaires d'approximation de l'analyse, *Mathematica (Cluj)* **1**(1959), 95–142.

[S63] A. SARD, Linear approximation, Math. Survey 9, AMS, , 1963.

[Sh95] A. YU. SHADRIN, Error bounds for Lagrange interpolation, *J. Approx. Theory* **80**(1995), 25–49.

[St27] J. F. STEFFENSEN, Interpolation, Williams and Wilkins, Baltimore, 1927.