Journal of Approximation Theory **101**, 359–403 (1999)
Article ID jath.1999.3372, available online at http://www.idealibrary.com on IDEAL®

Weak Type Inequalities for Best Simultaneous Approximation in Banach Spaces

Stefan Jansche

SAP AG, P.O. Box 1461, D-69190 Walldorf, Germany
E-mail: stefan.jansche@sap.com

Communicated by Vilmos Totik

Received April 20, 1998; accepted in revised form March 25, 1999

For arbitrary Banach spaces Butzer and Scherer in 1968 showed that the approximation order of best approximation can characterized by the order of certain K-functionals. This general theorem has many applications such as the characterization of the best approximation of algebraic polynomials by moduli of smoothness involving the Legendre, Chebyshev, or more general the Jacobi transform. In this paper we introduce a family of seminorms on the underlying approximation space which leads to a generalization of the Butzer–Scherer theorems. Now the characterization of the weighted best algebraic approximation in terms of the so-called main part modulus of Ditzian and Totik is included in our frame as another particular application. The goal of the paper is to show that for the characterization of the orders of best approximation, simultaneous approximation (in different spaces), reduction theorems, and K-functionals one has (essentially) only to verify three types of inequalities, namely inequalities of Jackson-, Bernstein-type and an equivalence condition which guarantees the equivalence of the seminorm and the underlying norm on certain subspaces. All the results are given in weak-type estimates for almost arbitrary approximation orders, the proofs use only functional analytic methods. © 1999 Academic Press

Key Words: best approximation; simultaneous approximation; K-functionals; rate of convergence; best weighted algebraic approximation.

1. INTRODUCTION

Denoting by $E_n(f; C_{2\pi}) := \inf_{t_n \in \Pi_n} \|f - t_n\|_{C_{2\pi}}$ the (error of the) best approximation of a 2π-periodic continuous function $f \in C_{2\pi}$ by the space Π_n of trigonometric polynomials of degree not exceeding $n \in \mathbb{N}$, the classical theorem of K. Weierstrass (1885) states that $\lim_{n\to\infty} E_n(f; C_{2\pi}) = 0$. In 1908 C. de la Vallée Poussin proposed the problem of characterizing of the rate of convergence of the best approximation. This was solved by the famous direct theorem of D. Jackson (1911) and the inverse theorem of S. N. Bernstein (1911/1922). The behaviour of $E_n(f; C_{2\pi})$ depends heavily

0021-9045/99 \$30.00

on the smoothness properties of the given function f, which can be measured by the so called rth modulus of smoothness $\omega_r(f, t; C_{2\pi}) := \sup_{0 \leqslant |h| \leqslant t} \|\Delta_h^r f\|_{C_{2\pi}}$, where Δ_h^r denotes the rth usual difference operator with increment $h \in \mathbb{R}$. Then the following theorem holds:

THEOREM 1.1. *For $f \in C_{2\pi}$ and $s < \sigma < r$, $s, r \in \mathbb{N}_0$, the following five assertions are equivalent*:

(i) $$E_n(f; C_{2\pi}) = \mathcal{O}(n^{-\sigma}), \qquad n \to \infty;$$

(ii) $$\omega_r(f, t; C_{2\pi}) = \mathcal{O}(t^{\sigma}), \qquad t \to 0^+;$$

(iii) $$f \in C_{2\pi}^s \quad \textit{and} \quad E_n(f^{(s)}; C_{2\pi}) = \mathcal{O}(n^{s-\sigma}), \qquad n \to \infty;$$

(iv) $$f \in C_{2\pi}^s \quad \text{and} \quad \begin{cases} \omega_1(f^{(s)}, t; C_{2\pi}) = \mathcal{O}(t^{\sigma - s}), & t \to 0^+,\ 0 < \sigma - s < 1 \\ \omega_2(f^{(s)}, t; C_{2\pi}) = \mathcal{O}(t^{\sigma - s}), & t \to 0^+,\ 0 < \sigma - s < 2; \end{cases}$$

(v) $$\|(t_n^{\circ})^{(r)}\|_{C_{2\pi}} = \mathcal{O}(n^{r-\sigma}), \qquad n \to \infty.$$

Above, $C_{2\pi}^s$ denotes the space of all s times continuously differentiable 2π-periodic functions and $t_n^{\circ} \in \Pi_n$ the polynomial of best approximation of f. The implication (i) to (iv) is due to D. Jackson [15, Abschnitt I.4]; (for $r = 1$), its converse was proved by S. N. Bernstein [2, 56. Théorème]. Concerning the second order modulus, this part was established by A. Zygmund [33], where one can find a partial proof of the the equivalence of (ii) and (iv). M. Zamansky [32, p. 26] showed that the assumption $\|f - t_n\|_{C_{2\pi}} = \mathcal{O}(n^{-\sigma})$ implies $\|t_n^{(r)}\|_{C_{2\pi}} = \mathcal{O}(n^{r-\sigma})$, which readily yields (v) from (i). In the fundamental paper [28] of S. B. Stečkin it was shown that the same assumption already ensures the existence of $f^{(s)}$, and $\|f^{(s)} - t_n^{(s)}\|_{C_{2\pi}} = \mathcal{O}(n^{s-\sigma})$. An application of the Bernstein theorem to $f^{(s)}$ delivers (iv) from (iii). Finally, the implication from (v) to (ii) goes back to G. Sunouchi [29, 30]. All these five equivalences in connection are first appeared in P. L. Butzer and K. Scherer [7, 8] (see also P. L. Butzer and S. Pawelke [6] for $L_{2\pi}^2$), where it was pointed out that the proof of the theorem above follows essentially from the validity of two fundamental inequalities, namely the so-called Jackson inequality [15, Satz II],

$$E_n(f; C_{2\pi}) \leqslant M n^{-r} \|f^{(r)}\|_{C_{2\pi}}, \qquad f \in C_{2\pi}^r, \tag{J}$$

and the Bernstein inequality [2, Sect. 12],

$$\|t_n^{(r)}\|_{C_{2\pi}} \leqslant n^r \|t_n\|_{C_{2\pi}}, \qquad t_n \in \Pi_n. \tag{B}$$

Theorem 1.1 does not only hold for continuous functions; the assertions remain valid if one considers 2π-periodic Lebesgue integrable functions $f \in L_{2\pi}^p$ as well. Furthermore, in their papers [7, 8] Butzer and Scherer showed that the equivalence theorem even holds in arbitrary Banach spaces X provided that, apart from some natural assumptions, suitable generalizations of the Jackson and Bernstein inequalities are given. To this end the set Π_n of trigonometric polynomials has to be replaced by a family $\mathscr{S} = \{S_t\}_{t \in (0,1]}$ of nested linear subspaces of X, i.e., $S_h \subset S_t$, $0 < t \leqslant h \leqslant 1$, which play the role of approximands. Thus, in this frame the best approximation for $f \in X$ is defined by $E_t(f; X) := \inf_{g_t \in S_t} \| f - g_t \|_X$. The modulus of smoothness makes no longer sense in arbitrary spaces, a substitute is given by the (Peetre) K-functional $K(f, t^r; X, Y) := \inf_{g \in Y} \{ \| f - g \|_X + t^r |g|_Y \}$, $Y \subset X$ denoting a linear subspace with semi-norm $|\cdot|_Y$. Assuming now that there hold (essentially) as modifications of (J) and (B) the Jackson inequality

$$E_t(f; X) \leqslant M t^r |f|_Y, \qquad f \in Y, \tag{J_X}$$

as well as the Bernstein inequality

$$|g_t|_Y \leqslant M t^{-r} \|g_t\|_X, \qquad g_t \in S_t, \tag{B_{X}}$$

then there follows for $f \in X$ and $\sigma < r$ the equivalence of

$$E_t(f; X) = \mathcal{O}(t^\sigma), \qquad t \to 0^+;$$

$$K(f, t^r; X, Y) = \mathcal{O}(t^\sigma), \qquad t \to 0^+.$$

This is the counterpart of (i) and (ii) in Theorem 1.1, and has many applications in various specific Banach spaces X. In particular, choosing $X := C_{2\pi}$, $Y := C^r_{2\pi}$, and $S_t := \Pi_{[1/t]-1}$, then on noting that $K(f, t^r; C_{2\pi}, C^r_{2\pi}) \sim \omega_r(f, t; C_{2\pi})$, the equivalence of (i) and (ii) follows from the equivalence above.

Unfortunately, the general theorems of Butzer–Scherer do not apply to the important case of weighted algebraic approximation by algebraic polynomials in connection with the powerful Ditzian–Totik moduli of smoothness, the usefulness of which is pointed out in their monograph [13]. The main difficulty concerning algebraic approximation is the fact that the accuracy of the best approximation by algebraic polynomials behaves not uniformly over the interval, say $[0, 1]$, it becomes better towards the endpoints 0, 1 of $[0, 1]$, instead. This observation already due to S. M. Nikolskiĭ [25] and has to be taken in consideration to define a suitable modulus of smoothness. To establish a characterization of the best algebraic approximation $E_n(f, L^p_\mu) := \inf_{p_n \in \mathscr{P}_n} \| f - p_n \|_{p,\mu}$ of an element f belonging to the weighted Lebesgue spaces L^p_μ (for definitions and ranges of the parameters see Section 7 below), Ditzian and Totik replaced the usual difference $\Delta_h f(x)$ by $\Delta_{h\varphi(x)} f(x) := f(x + h\varphi(x)/2) - f(x - h\varphi(x)/2)$,

$\varphi(x) := \sqrt{x(1-x)}$. The increment $h\varphi(x)$ decreases towards the endpoints of $[0, 1]$, the function may be of lesser smoothness there. Furthermore the norm is taken only over a subinterval $[2r^2h^2, 1-2r^2h^2]$ of $[0, 1]$, which forms a exhaustion of $[0, 1]$ for the limit $h \to 0^+$. This leads to the definition of the rth main part modulus of Ditzian and Totik, namely

$$\Omega_{r,\mu}(f, t) := \sup_{0<h\leqslant t} |\Delta^r_{h\varphi(\cdot)} f(\cdot)|_{p,\mu,h},$$

where $|f|_{p,\mu,h} := \|f\varphi^\mu\|_{L^p(2r^2h^2, 1-2r^2h^2)}$. Then the following theorem holds (cf. Z. Ditzian and V. Totik [13, Theorem 8.2.1.]):

Theorem 1.2. *For $f \in L^p_\mu$ and $r \in \mathbb{N}$, $t \in (0, 1]$ there holds*

$$E_n(f; L^p_\mu) \leqslant M \int_0^{1/n} \Omega_{r,\mu}(f, u) \frac{du}{u},$$

and

$$\Omega_{r,\mu}(f, t) \leqslant Mt^r \sum_{0\leqslant k\leqslant 1/t} (k+1)^{r-1} E_k(f; L^p_\mu).$$

Both estimates are called weak type inequalities, since the right hand terms are means of integrals and harmonic sums of the best approximation and main part modulus, respectively. The constants $M > 0$ above are independent of f, therefore weak type inequalities are stronger than the corresponding $\mathcal{O}$-assertions in Theorem 1.1 where the constants involved by the Landau symbol $\mathcal{O}$ may depend of f.

The main part modulus is no longer equivalent to a K-functional in the classical sense, in order to find a functional analytic approach of approximation theorems such that the algebraic approximation applies as well, we have to modify the general setup in Banach spaces. The first observation in this direction is that $|\cdot|_{p,\mu,t}$ defines a family of seminorms on L^p_μ which converges for $t \to 0^+$ towards the norm $\|\cdot\|_{p,\mu}$. In particular $|f|_{p,\mu,t}$ decreases in t, and $|f|_{p,\mu,t} = 0$ for all $0 < t < 1/(2r)$ implies $f = 0$ a.e. Now, in Banach spaces X we do not only consider a norm $\|\cdot\|_X$ on X, additionally we introduce a family $\{|\cdot|_{X,t}\,;\, t \in (0, 1]\}$ of *seminorms*, which has the *exhaustion property*, i.e., we have monotonicity

$$|f|_{X,h} \leqslant |f|_{X,t} \leqslant M\,\|f\|_X, \qquad f \in X, \quad 0 < t \leqslant h \leqslant 1,$$

and

$$\text{if } f \in X \text{ such that } |f|_{X,t} = 0 \text{ for all } 0 < t \leqslant 1, \text{ then } f = 0.$$

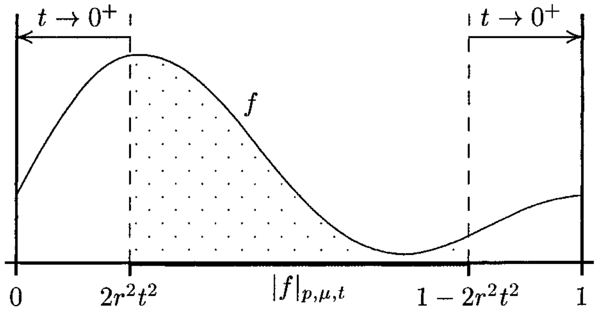

FIG. 1. Seminorms converging to the norm.

This exhaustion property immediately implies

$$\text{if } \lim_{t \to 0^+} |f|_{X,t} = 0 \text{ then } \|f\|_X = 0.$$

In this sense the one can say that the seminorms converge towards the norm (Fig. 1), see also [4]. Another motivation for introducing the seminorms is the fact that the exhaustion property can be used to weaken the Jackson inequality, whose verification is in fact usually the most difficult part in applications. To this end the parameters t of the seminorms and of the approximands S_t will be linked in the definition of the best modified approximation, namely

$$E_t^*(f; X) := \inf_{g_t \in S_t} |f - g_t|_{X,t}.$$

In order to characterize the best approximation $E_t(f; X)$ in terms of its modified counterpart $E_t^*(f; X)$, we need to control the norm by the seminorms. Thus we have to claim the equivalence condition

$$\|g_t\|_X \leqslant M\, |g_t|_{X, t/C^*}, \qquad g_t \in S_t, \tag{Eq}$$

for a constant $C^* \in (0, 1)$. This assumption implies

$$\|g_t\|_X \sim |g_t|_{X,t}, \qquad g_t \in S_t,$$

i.e., the seminorms and the norm are equivalent on the subspaces S_t. It will be shown that under the claim above it is sufficient to verify the weak Jackson inequality

$$E_t^*(f; X) \leqslant M\psi(t)\, |f|_Y, \qquad f \in Y, \tag{J_X^*}$$

instead of (J_X). ψ denotes an order function like $\psi(t) = t^r$. In the particular application of weighted algebraic approximation by verification of

(J_X^*) we can avoid estimates at the endpoints of the underlying interval, which would cause difficulties, since the function φ vanishes in 0, 1. The main goal in introducing the seminorms is the fact that we are now in position to define a modified K-functional, which is equivalent to the main part modulus $\Omega_{r,\mu}$ in that specific example. This K^*-functional is given by

$$K^*(f, \psi(t); X, Y) := \sup_{0<h\leqslant t} \inf_{g\in Y} \{|f-g|_{X,h} + \psi(h)\, |g|_Y\},$$

where the parameter of the order function ψ is also linked with the seminorms.

A Banach space X endowed with a family of seminorms $\{|\cdot|_{X,t};\, t\in(0,1]\}$, which satisfies the exhaustion property and the equivalence condition (Eq), is said to be an approximation space (A-space), provided for all $f\in X$ that the best approximation $E_t(f; X)$ vanishes for $t\to 0^+$, cf. Definition 3.1. Furthermore, a linear subspace $Y\subset X$ is called a subspace of order ψ, iff there hold the weak Jackson inequality (J_X^*) and the and the Bernstein inequality

$$|g_t|_Y \leqslant M\, \frac{1}{\psi(t)}\, \|g_t\|_X, \qquad g_t\in S_t. \tag{B_X^*}$$

In A-space we can prove among others the following direct and inverse estimates

$$E_t(f; X) \leqslant MK(f, \psi(t); X, Y),$$

$$E_t(f; X) \leqslant M \int_0^t K^*(f, \psi(u); X, Y)\, \frac{du}{u},$$

$$K(f, \psi(t); X, Y) \leqslant M\psi(t) \left\{ \|f\|_X + \sum_{1\leqslant k\leqslant 1/t} \frac{1}{k\psi(k^{-1})}\, E_{1/k}(f) \right\}.$$

Concerning simultaneous approximation and reduction theorems, i.e., theorems involving assertions of types (iii) and (iv) in Theorem 1.1, in the case of weighted algebraic approximation (in opposite of the trigonometric case) the problem arises that the derivative $f^{(r)}$ of a function $f\in L^p_\mu$ usually belongs to the different space $L^p_{\mu+r}$, instead of L^p_μ. Since the differentiation operator maps into a different space, in our setting we have to consider another A-space $\bar{X}$, such that in the application mentioned, the spaces X and $\bar{X}$ can be identified with L^p_μ and $L^p_{\mu+r}$, respectively. Note that $\bar{X}$ denotes another A-space and not the closure of X. The generalization of the derivative will be realized by a closed operator

$$D\colon X_D \to \bar{X},$$

satisfying $S_t \subset X_D \subset X$, $D(S_t) = \bar{S}_t$, and

$$\frac{1}{M}\, |f|_{X_D} \leqslant \|Df\|_{\bar{X}} \leqslant M\, |f|_{X_D}, \qquad f \in X_D,$$

$\bar{S}_t$ denoting the approximands of $\bar{X}$. If X_D forms a subspace of order ψ_D, then D is called an abstract derivative of the order ψ_D. On using this terminology, we can prove weak type inequalities between the best approximation and K-functionals (of derivatives). This leads to new estimates in the Ditzian–Totik frame as well. For convenience and in contrast to Theorem 1.1, we summarize the results of the following sections (all established as weak type estimates) in terms of $\mathcal{O}$ assertions.

COROLLARY 1.3. *Let X, $\bar{X}$ A-spaces, $Y \subset X$, $\bar{Y} \subset \bar{X}$ be subspaces of order ψ, $\bar{\psi} \prec 1$, respectively, and $D\colon X_D \to \bar{X}$ an abstract derivative of order $\psi_D \prec 1$. Then for $f \in X$ and $\phi \in \Phi$, satisfying $\psi \prec \phi \prec \psi_D$, $\phi \succ \bar{\psi}\psi_D$, the following six assertions are equivalent*:

$$\text{(i)} \qquad E_t(f; X) = \mathcal{O}(\phi(t)), \qquad t \to 0^+;$$

$$\text{(ii)} \qquad K(f, \psi(t); X, Y) = \mathcal{O}(\phi(t)), \qquad t \to 0^+;$$

$$\text{(ii)*} \qquad K^*(f, \psi(t); X, Y) = \mathcal{O}(\phi(t)), \qquad t \to 0^+;$$

$$\text{(iii)} \qquad Df \in \bar{X} \quad \textit{and} \quad E_t(Df; \bar{X}) = \mathcal{O}\left(\frac{\phi(t)}{\psi_D(t)}\right), \qquad t \to 0^+;$$

$$\text{(iv)} \qquad Df \in \bar{X} \quad \textit{and} \quad K(Df, \bar{\psi}(t); \bar{X}, \bar{Y}) = \mathcal{O}\left(\frac{\phi(t)}{\psi_D(t)}\right), \qquad t \to 0^+;$$

$$\text{(iv)*} \qquad Df \in \bar{X} \quad \textit{and} \quad K^*(Df, \bar{\psi}(t); \bar{X}, \bar{Y}) = \mathcal{O}\left(\frac{\phi(t)}{\psi_D(t)}\right), \qquad t \to 0^+.$$

Above, $\prec$ denotes an order relation (cf. Definition 2.3) for $\mathcal{O}$-regularly varying functions, introduced in [16, 17]. This relation provides an easy tool in comparison the growth of functions like $\phi(t) = t^\sigma\, |\log t|^\nu$. The main facts of this relation and the underlying function classes are collected in the next section.

In Sections 3–5 we want to establish—in extension of the Butzer–Scherer approach—the direct and inverse theorems, theorems concerning the simultaneous approximation, as well as reduction theorems in arbitrary A-spaces. In the last two sections, we wish to apply this theorems for the particular cases of trigonometric and algebraic approximation. The aim of this paper is to show on using functional analytic methods, i.e., on using soft analysis, that for the characterization of the best approximation

$E_t(f; X)$ one has (essentially) only to verify three types of inequalities, namely the equivalence condition (Eq), the Jackson, and the Bernstein inequality (J_X^*), (B_X^*). These estimates have to shown under the use of the underlying (hard) analysis or the particular applications.

Beside the classical applications of the Butzer–Scherer theorems, our approach by including the exhaustion method by seminorms, covers also many results in the weighted and non-weighted Ditzian–Totik setting (cf. [13, Chaps. 6–8]), and delivers new results for simultaneous approximation and reduction theorems. Further applications will be considered in a forthcoming paper.

2. ORDER FUNCTIONS

The approximation behaviour of the best approximation or the order of Jackson or Bernstein inequalities are described by order functions such as power functions $\phi(t) = t^\sigma$, for instance. In this section we generalize the admissible order functions used in approximation theorems to an Abelian group of functions Φ, and we introduce a relation $\prec$ in Φ which gives an easy tool to compare the different orders of elements of Φ. Several attempts to extend power functions can be found in papers on approximation theory, see e.g., P. L. Butzer and K. Scherer [9], P. L. Butzer *et al.* [4], Z. Ditzian and X. Zhou [14] as well as in S. Jansche and R. L. Stens [18] and E. van Wickeren [31]. A rigorous application of the theory of regularly varying functions seems to have been first carried out in [17], and in [16], where one finds a detailed elaboration and proofs of the material of this section. The textbooks of E. Seneta [27] and N. H. Bingham, *et al.* [3] provide a systematic treatment of the theory of regularly varying functions.

We make use of the Landau symbol in the common way, and write $f \sim g$ if $f = \mathcal{O}(g)$ and $g = \mathcal{O}(f)$. Furthermore, we will not distinguish the constants in various estimates; the value of a constant $M > 0$ may differ in each occurrence of a chain of inequalities, but always independent of the varying parameters. If necessary, we indicate dependencies by $M(C)$ etc. The class of order functions is given by the following definition:

DEFINITION 2.1. A positive (Lebesgue-) measurable function ϕ: $(0, 1] \to \mathbb{R}$ is said to be a *$\mathcal{O}$-regularly varying function* (*$\mathcal{O}$-RV function*), or *order function*, for short, if for each $t_0 \in (0, 1)$

$$\phi \sim 1 \qquad \text{on } [t_0, 1], \tag{2.1}$$

and if

$$0 < \liminf_{t \to 0^+} \frac{\phi(Ct)}{\phi(t)}, \qquad \limsup_{t \to 0^+} \frac{\phi(Ct)}{\phi(t)} < \infty \tag{2.2}$$

for all $C \in (0, 1]$. The class of all order functions is denoted by Φ.

The above definition can be found in S. Aljančić and D. Arandelović [1], except that the limit $t \to \infty$ is considered instead of $t \to 0^+$ here. A typical example of an order function is given by

$$\phi(t) := t^\sigma \, |\log t|^\rho, \quad t \in (0, 1/2]; \qquad \phi(t) := 1, \quad t \in (1/2, 1],$$

σ, ρ being arbitrary real numbers. Note that functions having exponential growth like $\phi(t) := e^{-1/t}$ do not belong to Φ. It is convenient to define $\phi^*, \phi_*: (0, 1] \to \mathbb{R}$

$$\phi^*(C) := \limsup_{t \to 0^+} \frac{\phi(Ct)}{\phi(t)}, \qquad \phi_*(C) := \liminf_{t \to 0^+} \frac{\phi(Ct)}{\phi(t)}.$$

Now we collect some properties of order functions, which can be found in [17].

Proposition 2.1. *If $\phi \in \Phi$, then for all $C \in (0, 1)$ we have*

$$\phi^* \sim 1, \qquad \phi_* \sim 1 \qquad \text{on} \quad [C, 1].$$

In particular, there exists a constant $M = M(C) > 0$ such that

$$\frac{1}{M} \phi(t) \leqslant \phi(h) \leqslant M \phi(t) \tag{2.3}$$

for all $C \leqslant h/t \leqslant 1$, $h, t \in (0, 1]$.

Concerning weak type inequalities, functions are often estimated by certain means of harmonic or geometric sums or by integrals. If the underlying functions belong to Φ, one can switch between these types of estimates, as the following lemma shows.

Lemma 2.2. *Let $\phi \in \Phi$ and $C \in (0, 1)$; then for $t \in (0, 1)$ we have*

$$\int_0^t \phi(u) \, \frac{du}{u} \sim \sum_{k \geqslant 1/t} \frac{1}{k} \, \phi(k^{-1}) \sim \sum_{j=0}^{\infty} \phi(C^j t), \tag{2.4}$$

$$\int_t^1 \phi(u) \, \frac{du}{u} \sim \sum_{1 \leqslant k \leqslant 1/t} \frac{1}{k} \, \phi(k^{-1}) \sim \sum_{j;\, t \leqslant C^j \leqslant 1} \phi(C^j), \tag{2.5}$$

provided that the integral exists or one of the series converges. The constants induced by $\sim$ depend on C.

Remark 2.2. In particular, if a family $\{\phi_\lambda\}_{\lambda \in \Lambda} \subset \Phi$, Λ an arbitrary index set, satisfies

$$\frac{1}{M}\,\phi_\lambda(t) \leqslant \phi_\lambda(h) \leqslant M\phi_\lambda(t)$$

for all $\lambda \in \Lambda$ and $h, t \in (0, 1]$, $C \leqslant t/h \leqslant 1$, then the estimates (2.4) and (2.5) hold uniformly in $\lambda \in \Lambda$.

If χ: $(0, 1] \to \mathbb{R}$ is just a non-negative almost increasing function, and $\phi \in \Phi$, then for $C \in (0, 1)$ there exists a $M = M(C) > 0$ such that

$$\sum_{j=0}^{\infty} \frac{\chi(C^j t)}{\phi(C^j t)} \leqslant M \sum_{k \geqslant [1/t]} \frac{1}{k}\,\frac{\chi(k^{-1})}{\phi(k^{-1})}, \qquad t \in (0, 1], \tag{2.6}$$

$$\sum_{j;\; t \leqslant C^j \leqslant 1} \frac{\chi(C^j)}{\phi(C^j)} \leqslant M \sum_{1 \leqslant k \leqslant 1/t} \frac{1}{k}\,\frac{\chi(k^{-1})}{\phi(k^{-1})}, \qquad t \in (0, 1]. \tag{2.7}$$

Concerning the behaviour of order functions in a neighbourhood of the origin, for $\phi \in \Phi$ the numbers

$$\alpha(\phi) := \sup_{C \in (0,1)} \frac{\log \phi^*(C)}{\log C}, \qquad \beta(\phi) := \inf_{C \in (0,1)} \frac{\log \phi_*(C)}{\log C}$$

are called the *upper* and *lower Matuzewska index*, respectively; see W. Matuzewska and W. Orlicz [22] and the literature of the authors cited there. These indices extract the rate of growth of ϕ at the origin. For the power function $\phi(t) = t^\sigma$, as an example, we note that $\phi^*(C) = \phi_*(C) = C^\sigma$ imply $\alpha(\phi) = \beta(\phi) = \sigma$. For arbitrary functions, which may oscillate, the Matuzewska indices cover the rate of growth above and below in some sense. Some basic properties of the indices are collected in the following

LEMMA 2.3. *Let $\alpha = \alpha(\phi)$ and $\beta = \beta(\phi)$ be the Matuzewska indices of $\phi \in \Phi$.*

(a) *The indices are real numbers $\alpha, \beta \in \mathbb{R}$, $\alpha \leqslant \beta$, and there hold*

$$\alpha(\phi) = \lim_{C \to 0^+} \frac{\log \phi^*(C)}{\log C}, \qquad \beta(\phi) = \lim_{C \to 0^+} \frac{\log \phi_*(C)}{\log C}. \tag{2.8}$$

(b) *For a given $\varepsilon > 0$ there exists some $C_0 \in (0, 1)$ such that*

$$C^\alpha \leqslant \phi^*(C) < C^{\alpha - \varepsilon}, \qquad C^{\beta + \varepsilon} < \phi_*(C) \leqslant C^\beta, \qquad C \in (0, C_0]. \tag{2.9}$$

In particular,

$$\phi_*(C) \leqslant C^\beta \leqslant C^\alpha \leqslant \phi^*(C) \qquad \forall C \in (0, 1]. \tag{2.10}$$

(c) *The Matuzewska indices are related by*

$$\alpha\left(\frac{1}{\phi}\right) = -\beta(\phi), \qquad \beta\left(\frac{1}{\phi}\right) = -\alpha(\phi), \tag{2.11}$$

and for the product of order functions $\phi_1, \phi_2 \in \Phi$ *we have* $\phi_1\phi_2 \in \Phi$, *satisfying*

$$-\infty < \alpha(\phi_1) + \alpha(\phi_2) \leqslant \alpha(\phi_1\phi_2) \leqslant \beta(\phi_1\phi_2) \leqslant \beta(\phi_1) + \beta(\phi_2) < \infty. \tag{2.12}$$

For approximation theorems we need a tool for estimating different order functions. To this end we define a growth relation $\prec$ in Φ, which allows a comparison of the growth of order functions at the origin.

Definition 2.3. Let $\phi_1, \phi_2 \colon (0, 1] \to \mathbb{R}$ be positive. The *growth relation* $\phi_1 \prec \phi_2$ holds iff there exists a constant $C \in (0, 1)$ such that

$$\limsup_{t \to 0^+} \frac{\phi_1(Ct)\ \phi_2(t)}{\phi_1(t)\ \phi_2(Ct)} < 1.$$

We write $\phi_1 \preccurlyeq \phi_2$, iff $t^\varepsilon \phi_1(t) \prec \phi_2(t)$ for all $\varepsilon > 0$.

For instance, let $\sigma_i, \rho_i \in \mathbb{R}$, $i = 1, 2$, and

$$\phi_i(t) := t^{\sigma_i} |\log t|^{\rho_i}, \quad t \in (0, 1/2], \qquad \phi_i(t) := 1, \quad t \in (1/2, 1].$$

Then it is obvious that $\phi_1 \prec \phi_2$ iff $\sigma_1 > \sigma_2$, and $\phi_1 \preccurlyeq \phi_2$ iff $\sigma_1 \geqslant \sigma_2$. In particular, for $\phi \in \Phi$ we have $\phi \prec 1$ iff $\phi^*(C) < 1$. Roughly speaking, $\phi_1 \prec \phi_2$ means that the growth of ϕ_1 and ϕ_2 differ by a power t^ε at the origin. Changes of growth in terms of powers of the logarithm are not effected by the growth relation $\prec$. The connections between the Matuzewska indices and our growth relation are given by the following

Proposition 2.4. *Let* $\phi_1, \phi_2 \in \Phi$; *then the following four assertions are equivalent*:

(i) $\phi_1 \prec \phi_2$;

(ii) *There exist constants* $\alpha_0 > 0$ *and* $C_0 \in (0, 1]$ *such that*

$$\frac{\phi_1(Ct)}{\phi_1(t)}\ \frac{\phi_2(t)}{\phi_2(Ct)} < C^{\alpha_0}, \qquad 0 < t < t_0(C),$$

for all $C \in (0, C_0]$ *and some* $t(C) \in (0, 1]$;

(iii) $\alpha(\phi_1/\phi_2) > 0$;

(iv) $\beta(\phi_1/\phi_2) < 0$.

Remark 2.4. The class Φ equipped with pointwise multiplication forms a multiplicative Abelian group with identity element $1(t) := 1$.

The relation $\prec$ in Φ is transitive, and Φ forms with $\prec$ a (partial) ordered group. In particular, for $\phi_1, \phi_2, \phi_3 \in \Phi$ satisfying $\phi_1 \prec \phi_2$ and $\phi_2 \prec \phi_3$, we have $\phi_1 \prec \phi_3$, and $\phi_1 \prec \phi_2$ implies $\phi_1 \phi_3 \prec \phi_2 \phi_3$.

The identity element 1 separates the convergent and divergent elements of Φ; we have

$$\lim_{t \to 0^+} \phi(t) = 0, \qquad \lim_{t \to 0^+} \phi(t) = \infty,$$

for $\phi \prec 1$ and $\phi \succ 1$, respectively. Since we are interested in convergent elements of Φ, we define $\Psi := \{\psi \in \Phi;\ \psi \prec 1\}$.

Now we want to give some characterizations for $\mathcal{O}$-RV functions, in combination with our growth relation. Owing to the group property it is sufficient to compare only one member of Φ with the identity element. A function on $(0, 1]$ is said to be *almost increasing* or *almost decreasing*, if there is some constant $M > 0$ such that $f(t_1) \leqslant M f(t_2)$ or $f(t_2) \leqslant M f(t_1)$ for all $t_1, t_2 \in I$, $t_1 \leqslant t_2$.

Proposition 2.5. *Let* $\phi \in \Phi$.

(a) *If ϕ is almost increasing, or if*

$$\phi(t) \sim \int_0^t \phi(u) \frac{du}{u}, \qquad t \to 0^+, \tag{2.13}$$

then there holds $\phi \preccurlyeq 1$.

(b) *Conversely, for each $\phi \prec 1$, the function ϕ is bounded, almost increasing, and we have* (2.13).

(c) *If ϕ is almost decreasing, or if ϕ satisfies*

$$\phi(t) \sim \int_t^1 \phi(u) \frac{du}{u}, \qquad t \to 0^+, \tag{2.14}$$

then $\phi \succcurlyeq 1$.

(d) $\phi \succ 1$ *implies that ϕ is almost decreasing, and* (2.14).

In both cases the integrals in (2.13) *and* (2.14) *can be replaced by the corresponding sums in* (2.4) *and* (2.5), *respectively.*

In the next sections we will frequently use that

$$\sum_{1 \leqslant k \leqslant 1/t} \frac{1}{k\psi(k^{-1})} \leqslant M \frac{1}{\psi(t)} \tag{2.15}$$

hold for all $\psi \in \Psi$ and $t \in (0, 1]$.

3. APPROXIMATION SPACES

In this section we provide the basic definitions and properties which are needed for approximation theorems in Banach spaces. In particular, we wish to introduce a family of seminorms to establish a seminorm exhaustion method required in the important application of weighted algebraic approximation.

Let X be a Banach space endowed with a norm $\|\cdot\|_X$ and let $\{|\cdot|_{X,t};\ t \in (0, 1]\}$ be a family of seminorms defined on X, satisfying the two conditions

$$|f|_{X,h} \leqslant |f|_{X,t} \leqslant M \|f\|_X, \qquad f \in X, \quad 0 < t \leqslant h \leqslant 1, \tag{3.1}$$

for some constant $M > 0$ independent of f and the parameters t, h; and

$$\text{if } f \in X \text{ with } |f|_{X,t} = 0 \qquad \text{for all} \quad 0 < t \leqslant 1, \text{ then } f = 0. \tag{3.2}$$

For simplicity we write $|f|_{X,t} := |f|_{X,1}$ for $t \geqslant 1$. A family $\{|\cdot|_{X,t};\ t \in (0, 1]\}$ of seminorms possesses the *exhaustion property* if the two conditions 3.1 and 3.2 are satisfied. Thus, by the first condition the seminorms $|\cdot|_{X,t}$ are bounded by $M \|f\|_X$ and decreasing in t. The property 3.2 is a separation property, also used in the theory of topological vector spaces to give a characterization of Hausdorff spaces defined by seminorms. From the exhaustion property we immediately obtain

$$\text{if } \lim_{t \to 0^+} |f|_{X,t} = 0, \text{ then } \|f\|_X = 0.$$

In this sense one can say that the seminorms converge to the norm for $t \to 0^+$, or the norm is exhausted by the seminorms.

The elements of approximation are given by a nested family of linear subspaces $\mathscr{S} = \{S_t\}_{t \in (0,1]}$ in X satisfying

$$S_h \subset S_t \subset X, \qquad 0 < t \leqslant h \leqslant 1. \tag{3.3}$$

The (*error of*) *best approximation* to $f \in X$ by elements of S_t is defined by

$$E_t(f) \equiv E_t(f; X) := \inf_{g_t \in S_t} \|f - g_t\|_X, \qquad t \in (0, 1]. \tag{3.4}$$

We also define the *modified best approximation* with respect to the seminorms by

$$E_t^*(f) \equiv E_t^*(f; X) := \inf_{g_t \in S_t} |f - g_t|_{X, t}, \qquad t \in (0, 1]. \tag{3.5}$$

Note that the parameter t of the seminorms and S_t coincide, and the (modified) best approximations are sublinear functionals bounded by $\|f\|_X$ and $M \|f\|_X$, respectively.

To guarantee that each element of X can be approximated by elements of $S_0 := \bigcup_{0 < t \leqslant 1} S_t$, we require the *Weierstrass property*, namely that

$$\lim_{t \to 0^+} E_t(f) = 0, \qquad f \in X. \tag{3.6}$$

By (3.1) it is obvious that

$$E_t^*(f) \leqslant M E_t(f), \qquad t \in (0, 1]. \tag{3.7}$$

To establish an estimate of $E_t(f)$ by the modified best approximation $E_t^*(f)$, we need a sort of converse of (3.1). The goal is to postulate the required counterpart only on subspaces of S_t. $\mathscr{S} = \{S_t\}_{t \in (0, 1]}$ is said to satisfy the *equivalence condition*, if there exists some constants $M \geqslant 0$, $C^* \in (0, 1)$ and $t_0 \in (0, 1]$ such that

$$\|g_t\|_X \leqslant M \, |g_t|_{X, t/C^*}, \qquad g_t \in S_t, \; t \in (0, t_0]. \tag{3.8}$$

Throughout, these constants C^*, t_0 are assumed to be fixed. Indeed, on using the exhaustion property, we have

$$\|g_t\|_X \sim |g_t|_{X, t/C^*}, \qquad g_t \in S_t,$$

justifying the name equivalence condition. Note that in applications the S_t are usually finite dimensional, therefore (3.8) is much weaker than an equivalence of $\|\cdot\|_X$ and $|\cdot|_{X, t}$ on the whole space X. We summarize these properties in the following

Definition 3.1. Let X be a Banach space satisfying (3.1), (3.2), and $\mathscr{S} = \{S_t\}_{t \in (0, 1]}$ be a family of linear subspaces in X such that (3.3), (3.6) as well (3.8) hold for fixed constants $C^* \in (0, 1)$, $t_0 \in (0, 1]$. Then the tuple $(X, \mathscr{S})$ or simply X is called an *approximation space* or *A-space*.

For the discrete case, i.e., $t = 1/n$, $n \in \mathbb{N}$, the properties and assumptions of an A-space are essentially introduced in [4]. First we prove the weak equivalence of the best and modified best approximation.

PROPOSITION 3.1. *Let X be an A-space; then for each $C \in [C^*, 1)$ there exists constants $M = M(C) \geqslant 0$ and $\tilde{M} \geqslant 0$ such that*

$$\tilde{M} E_t^*(f) \leqslant E_t(f) \leqslant M \sum_{j=0}^{\infty} E_{C^j t}^*(f) \tag{3.9}$$

for all $f \in X$ and $t \in (0, t_0]$.

Proof. The left hand estimate is given by (3.7). To estimate the best approximation we can assume that the sum is finite. For an $\varepsilon > 0$, $t \in (0, t_0]$ we choose $g_{C^j t} \in S_{C^j t}$ satisfying

$$|f - g_{C^j t}|_{X, C^j t} \leqslant E_{C^j t}^*(f) + C^j \varepsilon.$$

Using (3.8) and (3.1) it follows

$$\begin{aligned} \|g_{C^{j+1}t} - g_{C^j t}\|_X &\leqslant M |g_{C^{j+1}t} - g_{C^j t}|_{X, C^{j+1}t/C^*} \\ &\leqslant M\{|g_{C^{j+1}t} - f|_{X, C^{j+1}t} + |g_{C^j t} - f|_{X, C^j t}\} \\ &\leqslant M\{E_{C^{j+1}t}^*(f) + C^{j+1}\varepsilon + E_{C^j t}^*(f) + C^j \varepsilon\}, \end{aligned}$$

hence

$$\sum_{j=0}^{\infty} \|g_{C^{j+1}t} - g_{C^j t}\|_X \leqslant M \left\{\varepsilon + \sum_{j=0}^{\infty} E_{C^j t}^*(f)\right\}. \tag{3.10}$$

By the finiteness of the right hand side $\{g_{C^j t}\}_{j \in \mathbb{N}_0}$ forms a Cauchy sequence in the complete space X, converging to a $g \in X$. Then (3.1) yields for $h \in (0, 1]$

$$\begin{aligned} |f - g|_{X, h} &\leqslant \lim_{j \to \infty} \{|f - g_{C^j t}|_{X, h} + |g_{C^j t} - g|_{X, h}\} \\ &\leqslant M \lim_{j \to \infty} \{|f - g_{C^j t}|_{X, C^j t} + \|g_{C^j t} - g\|_X\} \\ &\leqslant M \lim_{j \to \infty} \{E_{C^j t}^*(f) + C^j \varepsilon + \|g_{C^j t} - g\|_X\} = 0. \end{aligned}$$

The separation property (3.2) then implies $f = g$. Therefore we have

$$f - g_t = \sum_{j=0}^{\infty} (g_{C^{j+1}t} - g_{C^j t}),$$

which implies (3.9) after using (3.10). ∎

The rate of convergence of the best approximation depends on the smoothness of a given $f \in X$, which has to be interpreted in a suitable way in arbitrary Banach spaces. Instead of moduli of smoothness, which are only meaningful in specified function spaces, we will use modifications of the Peetre K-functional. These K-functionals measure the *distance* of f to a given subspace $Y \subset X$. If this subspace represents a class of smooth elements, e.g., differentiable functions, then the K-functionals become measures of the smoothness of f.

To be concrete, throughout the paper let X be a A-space and $Y \subset X$ a linear subspace with seminorm $|\cdot|_Y$ and associate norm $\|\cdot\|_Y := |\cdot|_Y + \|\cdot\|_X$, such that Y is complete with respect to $\|\cdot\|_Y$, satisfying

$$S_0 \subset Y \subset X. \tag{3.11}$$

The *K-functional* $K(f, \psi(t))$ is defined for given $f \in X$, $t \in (0, 1]$, and $\psi \in \Psi = \{\phi \in \Phi;\ \phi \prec 1\}$, by

$$K(f, \psi(t)) \equiv K(f, \psi(t); X, Y) := \inf_{g \in Y} \{\|f - g\|_X + \psi(t)\, |g|_Y\}.$$

In some applications it is more advantageous to use a *modified K-functional*, given by

$$K^*(f, \psi(t)) \equiv K^*(f, \psi(t); X, Y) := \sup_{0 < h \leqslant t}\ \inf_{g \in Y} \{|f - g|_{X,h} + \psi(h)\, |g|_Y\}.$$

Since the seminorm $|f - g|_{X,h}$ is decreasing and $\psi(h)$ is almost increasing in h, the supremum is taken to yield monotonicity. K-functionals involving families of seminorms in special function spaces are already used in approximation theory; see, e.g., Ditzian and Totik [13], and H. Mevissen and R. J. Nessel [20, 21] The definition above first appeared in [4].

In the following lemma we collect some elementary properties of the K-functionals.

LEMMA 3.2. *For $\psi \in \Psi$ the K-functionals $K(f, \psi(t))$ and $K^*(f, \psi(t))$ are sublinear bounded functional in $f \in X$, and we have either $K(f, \psi(t)) \in \Phi$ or $K(f, \psi(t)) \equiv 0$, satisfying for all fixed $C \in (0, 1]$*

$$\frac{1}{M} K(f, \psi(t)) \leqslant K(f, \psi(h)) \leqslant M K(f, \psi(t)) \tag{3.12}$$

for all $t, h \in (0, 1]$, $C \leqslant t/h \leqslant 1$. Furthermore, for $\phi \in \Phi$ we have

$$\begin{aligned}\int_0^t \phi(u)\, K(f, \psi(u)) \frac{du}{u} &\sim \sum_{k \geqslant 1/t} \frac{1}{k}\, \phi(k^{-1})\, K(f, \psi(k^{-1})) \\ &\sim \sum_{j=0}^{\infty} \phi(C^j t)\, K(f, \psi(C^j t))\end{aligned}$$

as well as

$$\int_t^1 \phi(u)\, K(f, \psi(u))\, \frac{du}{u} \sim \sum_{1 \leqslant k \leqslant 1/t} \frac{1}{k}\, \phi(k^{-1})\, K(f, \psi(k^{-1}))$$

$$\sim \sum_{j;\, t \leqslant C^j \leqslant 1} \phi(C^j)\, K(f, \psi(C^j)),$$

where the constants, involved by $\sim$, *are independent of* $f \in X$ *and* $t \in (0, 1]$. *All these assertions remain valid, if* $K(f, \psi(t))$ *is replaced by the modified* K^**-functional* $K^*(f, \psi(t))$. *Finally it holds*

$$K^*(f, \psi(t)) \leqslant MK(f, \psi(t)), \qquad f \in X, \quad t \in (0, 1]. \tag{3.13}$$

Proof. On using Proposition 2.1, we obtain for arbitrary $g \in Y$

$$\|f - g\|_X + \frac{1}{M}\, \psi(t)\, |g|_Y \leqslant \|f - g\|_X + \psi(h)\, |g|_Y \leqslant \|f - g\|_X + M\psi(t)\, |g|_Y,$$

for all $t, h \in (0, 1]$, $C \leqslant t/h \leqslant 1$. Taking the infimum over all $g \in Y$, there follows (3.12). This also implies that $K(f, \psi(\cdot)) \in \Phi$. Concerning the modified K-functional, because of the monotonicity, we have only to establish the right hand side of (3.12) for $K^*(f, \psi(t))$. Let $t \leqslant u \leqslant h$; then by Proposition 2.1 for arbitrary $g \in Y$,

$$|f - g|_{X, u} + \psi(u)\, |g|_Y \leqslant M\{|f - g|_{X, t} + \psi(t)\, |g|_Y\},$$

and therefore

$$\inf_{g \in Y} \{|f - g|_{X, u} + \psi(u)\, |g|_Y\} \leqslant MK^*(f, \psi(t)).$$

Thus (3.12) follows from

$$K^*(f, \psi(h)) = \max\{K^*(f, \psi(t)), \sup_{t \leqslant u \leqslant h} \inf_{g \in Y} \{|f - g|_{X, u} + \psi(u)\, |g|_Y\}\}$$

$$\leqslant MK^*(f, \psi(t)).$$

Recalling Remark 2.2, the equivalences between the integral and sums above are an immediate consequence of (3.12). As to (3.13), for $0 < u < t$ and $\tilde{g} \in Y$, we have on noting that ψ is almost increasing according to Proposition 2.5,

$$|f - \tilde{g}|_{X, u} + \psi(u)\, |\tilde{g}|_Y \leqslant M\{\|f - \tilde{g}\|_X + \psi(t)\, |\tilde{g}|_Y\},$$

and,

$$\sup_{0<u\leqslant t} \inf_{g\in Y} \{|f-g|_{X,u}+\psi(u)\,|g|_Y\} \leqslant M\{\|f-\tilde{g}\|_X+\psi(t)\,|\tilde{g}|_Y\}.$$

Taking the infimum over $\tilde{g}\in Y$ gives (3.13). ∎

4. DIRECT AND INVERSE THEOREMS

In the following, let X be an A-space in the sense of Definition 3.1. As mentioned above, the K-functionals become a measure of smoothness if the subspace Y can be interpreted as a subspace of smooth elements of X. This is naturally satisfied, if the best approximation of elements of Y vanishes sufficiently rapidly.

The subspace Y is said to satisfy a *Jackson*(*-type*) *inequality* of order $\psi\in\Psi$, if

$$E_t(f)\leqslant M\psi(t)\,|f|_Y, \qquad f\in Y, \quad t\in(0, t_0], \tag{4.1}$$

and Y satisfies a *weak Jackson*(*-type*) *inequality* of order ψ, if

$$E_t^*(f)\leqslant M\psi(t)\,|f|_Y, \qquad f\in Y, \quad t\in(0, t_0]. \tag{4.2}$$

In the next fundamental lemma we show, that these two types of Jackson inequalities are equivalent. Thus in applications it suffices to verify the weaker form (4.2), and in the following we can require the stronger inequality (4.1).

LEMMA 4.1. *If $\psi\in\Psi$, then the Jackson inequality* (4.1) *is satisfied iff the weak Jackson inequality* (4.2) *holds.*

Proof. According to Proposition 3.1 we just have to assume (4.2). Then again by Propositions 3.1 and 2.5 we obtain the estimate

$$E_t(f)\leqslant M\sum_{j=0}^{\infty} E^*_{C^jt}(f)\leqslant M\,|f|_Y\sum_{j=0}^{\infty}\psi(C^jt)\leqslant M\psi(t)\,|f|_Y$$

for $t\in(0, t_0]$ and $C\in[C^*, 1)$. ∎

The following estimates of the best approximation by K-functionals are referred to as direct or Jackson(-type) theorems.

THEOREM 4.2. *Let $Y \subset X$ satisfy a Jackson inequality of order $\psi \in \Psi$. Then for all $f \in X$, $t \in (0, t_0]$ we have*

$$E_t(f) \leqslant MK(f, \psi(t)), \tag{4.3}$$

$$E_t(f) \leqslant M \int_0^t K^*(f, \psi(u)) \frac{du}{u}. \tag{4.4}$$

Proof. The sublinearity of $E_t(f)$ together with the Jackson inequality (4.1) imply for arbitrary $g \in Y$

$$E_t(f) \leqslant E_t(f-g) + E_t(g) \leqslant M\{\|f-g\|_X + \psi(t)\,|g|_Y\}.$$

Taking the infimum over all $g \in Y$ we obtain (4.3). By the same arguments we also obtain

$$E_t^*(f) \leqslant MK^*(f, \psi(t)).$$

Inserting this estimate into (3.9) we derive (4.4) after using Lemma 3.2. ∎

The main ingredient to establish inverse theorems is a Bernstein inequality which controls the behaviour of the Y seminorms applied to the smooth approximands $g_t \in S_t$ by their X norms.

A subspace $Y \subset X$ satisfies a *Bernstein*(*-type*) *inequality* of order $\psi \in \Psi$, if

$$|g_t|_Y \leqslant M \frac{1}{\psi(t)} \|g_t\|_X, \qquad g_t \in S_t,\ t \in (0, 1]. \tag{4.5}$$

A corresponding weak Bernstein inequality is obsolete, because the norm $\|\cdot\|_X$ and the seminorm $|\cdot|_{X,t}$ are equivalent on S_t. Typical methods establishing estimates of measures of smoothness in terms of the best approximation are telescoping arguments.

LEMMA 4.3. *Let $Y \subset X$ satisfy the Bernstein inequality* (4.5) *of order $\psi \in \Psi$. Then there exists a constant $M \geqslant 0$ such that for all $g_t \in S_t$, $t \in (0, 1]$,*

$$|g_t|_Y \leqslant M \left\{ \|f\|_X + \frac{1}{\psi(t)} \|f - g_t\|_X + \sum_{1 \leqslant k \leqslant 1/t} \frac{1}{k\psi(k^{-1})} E_{1/k}(f) \right\}, \qquad f \in X.$$

Proof. Let $\varepsilon > 0$ and choose $N = N(t) \in \mathbb{N}_0$ satisfying $2^{-N-1} < t \leqslant 2^{-N}$. Then there exist $\tilde{g}_{2^{-j}} \in S_{2^{-j}}$ such that the inequality

$$\|f - \tilde{g}_{2^{-j}}\|_X \leqslant E_{2^{-j}}(f) + 2^{-j}\psi(2^{-j})\,\varepsilon$$

holds for all $j \in \mathbb{N}$. Observing that $g_t - \tilde{g}_{2^{-N}} \in S_t$ we have, on using the Bernstein inequality (4.5) and Proposition 2.1,

$$\begin{aligned}
|g_t - \tilde{g}_{2^{-N}}|_Y &\leqslant M \frac{1}{\psi(t)} \|g_t - \tilde{g}_{2^{-N}}\|_X \\
&\leqslant M \frac{1}{\psi(t)} \|f - g_t\|_X + M \frac{1}{\psi(2^{-N})} \|f - \tilde{g}_{2^{-N}}\|_X \\
&\leqslant M \left\{ \frac{1}{\psi(t)} \|f - g_t\|_X + \frac{1}{\psi(2^{-N})} E_{2^{-N}}(f) + 2^{-N}\varepsilon \right\}.
\end{aligned}$$

Correspondingly we find

$$\begin{aligned}
|\tilde{g}_{2^{-j-1}} - \tilde{g}_{2^{-j}}|_Y &\leqslant M \frac{1}{\psi(2^{-j-1})} \|f - \tilde{g}_{2^{-j-1}}\|_X + M \frac{1}{\psi(2^{-j})} \|f - \tilde{g}_{2^{-j}}\|_X \\
&\leqslant M \left\{ \frac{1}{\psi(2^{-j-1})} E_{2^{-j-1}}(f) + 2^{-j-1}\varepsilon + \frac{1}{\psi(2^{-j})} E_{2^{-j}}(f) + 2^{-j}\varepsilon \right\},
\end{aligned}$$

as well as

$$\begin{aligned}
|\tilde{g}_1|_Y &\leqslant M \frac{1}{\psi(1)} \|\tilde{g}_1\|_X \leqslant M \|f\|_X + M \frac{1}{\psi(1)} \|f - \tilde{g}_1\|_X \\
&\leqslant M \left\{ \|f\|_X + \frac{1}{\psi(1)} E_1(f) + \varepsilon \right\},
\end{aligned}$$

After summation of these estimates, the monotonicity of the best approximation allows us to apply (2.7) to deduce

$$\begin{aligned}
|g_t|_Y &\leqslant |\tilde{g}_1|_Y + \sum_{j=0}^{N-1} |\tilde{g}_{2^{-j-1}} - \tilde{g}_{2^{-j}}|_Y + |g_t - \tilde{g}_{2^{-N}}|_Y \\
&\leqslant M \left\{ \varepsilon + \|f\|_X + \frac{1}{\psi(t)} \|f - g_t\|_X + \sum_{j=0}^{N} \frac{1}{\psi(2^{-j})} E_{2^{-j}}(f) \right\} \\
&\leqslant M \left\{ \varepsilon + \|f\|_X + \frac{1}{\psi(t)} \|f - g_t\|_X + \sum_{1 \leqslant k \leqslant 1/t} \frac{1}{k\psi(k^{-1})} E_{1/k}(f) \right\}.
\end{aligned}$$

Thus the assertion is proved, because $\varepsilon > 0$ is independent of the last term. ∎

The inverse theorem now reads

THEOREM 4.4. *If a subspace $Y \subset X$ satisfies a Bernstein inequality of order $\psi \in \Psi$, then for $t \in (0, 1]$ we have the weak type estimate*

$$K(f, \psi(t)) \leqslant M\psi(t) \left\{ \|f\|_X + \sum_{1 \leqslant k \leqslant 1/t} \frac{1}{k\psi(k^{-1})} E_{1/k}(f) \right\}, \qquad f \in X. \tag{4.6}$$

Furthermore, for each fixed $t_1 \in (0, 1]$ the summation over $1 \leqslant k \leqslant 1/t$ can be replaced by summation over $1/t_1 \leqslant k \leqslant 1/t$.

Proof. For a given $\varepsilon > 0$ we choose elements $g_t \in S_t$ such that

$$\|f - g_t\|_X \leqslant E_t(f) + \varepsilon.$$

An application Lemma 4.3 yields immediately

$$\begin{aligned} K(f, \psi(t)) &\leqslant \|f - g_t\|_X + \psi(t) \, |g_t|_Y \\ &\leqslant E_t(f) + \varepsilon + M\psi(t) \\ &\quad \times \left\{ \|f\|_X + \frac{1}{\psi(t)} (E_t(f) + \varepsilon) + \sum_{1 \leqslant k \leqslant 1/t} \frac{1}{k\psi(k^{-1})} E_{1/k}(f) \right\}, \end{aligned}$$

hence (4.6) after using (2.3) and the monotonicity of $E_t(f)$. The boundedness of the best approximation and ψ on $[t_1, 1]$ imply for a constant $M = M(t_1) > 0$

$$\sum_{1 \leqslant k \leqslant 1/t_1} \frac{1}{k\psi(k^{-1})} E_{1/k}(f) \leqslant M \, \|f\|_X,$$

completing the proof. ∎

Remark 4.1. Recalling that $K^*(f, \psi(t)) \leqslant MK(f, \psi(t))$, the K-functional may be replaced by the K^*-functional. If in particular $|g|_Y = 0$ for all $g \in S_1$, then we get $K(f - g, \psi(t)) = K(f, \psi(t))$, $g \in S_1$, by the definition of the K-functional. In this case we obtain on applying Theorem 4.4 to $f - g$ and taking the infimum over all $g \in S_1$,

$$K^*(f, \psi(t)) \leqslant MK(f, \psi(t)) \leqslant M\psi(t) \sum_{1 \leqslant k \leqslant 1/t} \frac{1}{k\psi(k^{-1})} E_{1/k}(f) \qquad \forall f \in X. \tag{4.7}$$

Thus if the approximands of S_1 are contained in the null-space $\mathcal{N}(Y) := \{g \in Y;\ |g|_Y = 0\}$ of Y, the term $\psi(t) \, \|f\|_X$ (of order ψ) can be dropped.

Obviously the behaviour of the best approximation $E_t(f)$ depends only on the underlying space X and the family $\mathscr{S}$ of approximands. In particular, $E_t(f)$ is independent of any choice of the subspace Y or of the order ψ of the required Jackson and Bernstein inequalities. On the other hand, the latter inequalities give a characterization of the smoothness of the elements of Y. This gives a justification for the following

DEFINITION 4.2. Let Y be a linear subspace of an A-space $(X, \mathscr{S})$ such that $S_0 \subset Y \subset X$, and Y is complete with respect to $\|\cdot\|_Y$. Then Y is called a subspace of order $\psi \in \Psi$, if Y satisfies a weak Jackson inequality (4.1) as well as a Bernstein inequality (4.5) of order ψ. If the corresponding Matuzewska indices of ψ coincide, we identify the order with the values $\alpha(\psi) = \beta(\psi)$.

In applications the order usually corresponds to the order of differentiability, e.g., the space $C^r_{2\pi}$ is a subspace of $C_{2\pi}$ of order $\alpha(t^r) = \beta(t^r) = r$.

By combining the direct and inverse theorem we are now able to prove estimates between the classical and modified K-functionals.

PROPOSITION 4.5. *Let $Y \subset X$ be a subspace of order $\psi \in \Psi$. Then there exists a constant $M > 0$ such that*

$$\begin{aligned} \frac{1}{M} K^*(f, \psi(t)) &\leqslant K(f, \psi(t)) \\ &\leqslant M \left\{ \psi(t) \|f\|_X + \int_0^t K^*(f, \psi(u)) \frac{du}{u} \right. \\ &\quad \left. + \psi(t) \int_t^1 \frac{1}{\psi(u)} K^*(f, \psi(u)) \frac{du}{u} \right\}, \end{aligned} \tag{4.8}$$

for all $f \in X$ and $t \in (0, 1]$.

Proof. The left hand inequality has already been established in Lemma 3.2. On $[t_0, 1]$ we obtain on noting $\psi \sim 1$

$$K(f, \psi(t)) \leqslant M \|f\|_X \leqslant M\psi(t) \|f\|_X, \qquad t \in [t_0, 1].$$

Therefore it remains to verify the right hand estimate for $t \in (0, t_0]$. To see this we use Lemma 3.2 to insert the direct theorem (4.4) into Theorem 4.4 yielding

$$K(f, \psi(t)) \leqslant M\psi(t) \left\{ \|f\|_X + \sum_{1/t_0 \leqslant j \leqslant 1/t} \frac{1}{j\psi(j^{-1})} \sum_{k \geqslant j} \frac{1}{k} K^*(f, \psi(k^{-1})) \right\} \tag{4.9}$$

for each $j \in \mathbb{N}$ with $1/t_0 \leqslant j \leqslant 1/t$. Now we apply (2.15) for $h = t$ and $h = k^{-1}$ to obtain after splitting the inner sum and changing the order of summation

$$\begin{aligned}
\sum_{1 \leqslant j \leqslant 1/t} \frac{1}{j\psi(j^{-1})} \sum_{k \geqslant j} \frac{1}{k} K^*(f, \psi(k^{-1}))
&= \sum_{k > 1/t} \Bigg(\sum_{1 \leqslant j \leqslant k} \frac{1}{j\psi(j^{-1})} \Bigg) \frac{1}{k} K^*(f, \psi^{-1})) \\
&\quad + \sum_{1 \leqslant k \leqslant 1/t} \Bigg(\sum_{1 \leqslant j \leqslant 1/t} \frac{1}{j\psi(j^{-1})} \Bigg) \frac{1}{k} K^*(f, \psi^{-1})) \\
&\leqslant M \frac{1}{\psi(t)} \sum_{k \geqslant 1/t} \frac{1}{k} K^*(f, \psi(k^{-1})) \\
&\quad + M \sum_{1 \leqslant k \leqslant 1/t} \frac{1}{k\psi(k^{-1})} K^*(f, \psi(k^{-1})).
\end{aligned}$$

Passing over to the corresponding integral representation the assertion follows by inserting this inequality into (4.9). ▮

In preparation for the next sections we state the following

COROLLARY 4.6. *Let $Y \subset X$ be a subspace of order $\psi \in \Psi$ and $\psi \prec \tilde{\psi} \prec 1$. Then for $f \in X$ we have*

$$\begin{aligned}
\frac{1}{M} \int_0^1 \frac{1}{\tilde{\psi}(t)} K^*(f, \psi(t)) \frac{dt}{t} &\leqslant \int_0^1 \frac{1}{\tilde{\psi}(t)} K(f, \psi(t)) \frac{dt}{t} \\
&\leqslant M \left\{ \|f\|_X + \int_0^1 \frac{1}{\tilde{\psi}(t)} K^*(f, \psi(t)) \frac{dt}{t} \right\}.
\end{aligned}$$

In particular, if one of these integrals is finite, so are the others.

Proof. The left hand inequality is a trivial consequence of Proposition 4.5. Integrating the right estimate of 4.8, we obtain after changing the order of integration

$$\begin{aligned}
&\int_0^1 \frac{1}{\tilde{\psi}(t)} K(f, \psi(t)) \frac{dt}{t} \\
&\quad \leqslant M \left\{ \|f\|_X \int_0^1 \frac{1}{\tilde{\psi}(t)} \frac{dt}{t} + \int_0^1 K^*(f, \psi(u)) \int_u^1 \frac{1}{\tilde{\psi}(t)} \frac{dt}{t} \frac{du}{u} \right. \\
&\qquad \left. + \int_0^1 \frac{1}{\psi(u)} K^*(f, \psi(u)) \int_0^u \frac{1}{\tilde{\psi}(t)} \frac{dt}{t} \frac{du}{u} \right\}.
\end{aligned}$$

Applying Proposition 2.5 on $1/\tilde{\psi} \succ 1$ and $\psi/\tilde{\psi} \prec 1$, respectively, it follows

$$\int_0^u \frac{\psi(t)}{\tilde{\psi}(t)} \frac{dt}{t} \sim \frac{\psi(u)}{\tilde{\psi}(u)}, \qquad \int_u^1 \frac{1}{\tilde{\psi}(t)} \frac{dt}{t} \sim \frac{1}{\tilde{\psi}(u)},$$

proving the right inequality. ∎

The interplay between the direct and the inverse theorem becomes clearer if we use Proposition 2.5 to formulate the foregoing theorems as $\mathcal{O}$-statements.

COROLLARY 4.7. *Let X be an A-space, $\phi \in \Phi$ and $Y \subset X$ a subspace of order $\psi \in \Psi$.*

(a) *If $\psi \prec \phi$, then the following statements are equivalent for $f \in X$:*

$$\text{(i)} \qquad E_t(f; X) = \mathcal{O}(\phi(t)), \qquad t \to 0^+;$$

$$\text{(ii)} \quad K(f, \psi(t); X, Y) = \mathcal{O}(\phi(t)), \qquad t \to 0^+.$$

(b) *If additionally $\phi \prec 1$, then* (i) *and* (ii) *of part* (a) *are equivalent to*

$$\text{(iii)*} \quad K^*(f, \psi(t); X, Y) = \mathcal{O}(\phi(t)), \qquad t \to 0^+.$$

5. SIMULTANEOUS APPROXIMATION AND REDUCTION THEOREMS

In function spaces, the problem of characterizing the best approximation of derivatives of a given function f in terms of moduli of smoothness applied to f, is referred to as *simultaneous approximation*. In arbitrary Banach spaces X there is no way to define the derivative of an element $f \in X$ in a classical manner by limits of differences. In order to make the study of simultaneous approximation in Banach spaces possible, we define instead a closed operator D on a subspace $X_D \subset X$ such that in applications D can be identified with suitable differential operators. Using this operator we have to estimate the best approximation of Df by K-functionals applied to f. We will see below that in the case of weighted algebraic approximation the rth derivative of a function $f \in L^p$ has to be understood as an element of a different weighted space $L^p_{\varphi^r}$ instead of L^p itself. To include weighted algebraic approximation as an application of our general approach, we have to allow that our operator D maps X_D into a different Banach space $\bar{X}$. Again, throughout this section we use X and $\bar{X}$ as abbreviations for A-spaces $(X, \mathcal{S})$ and $(\bar{X}, \bar{\mathcal{S}})$, respectively.

DEFINITION 5.1. Let $(X, \mathscr{S})$, $(\bar{X}, \bar{\mathscr{S}})$ be two A-spaces and $X_D \subset X$ a subspace of order $\psi_D \in \Psi$. If there exists a closed linear operator

$$D: X_D \to \bar{X},$$

such that

$$\frac{1}{M} |f|_{X_D} \leqslant \|Df\|_{\bar{X}} \leqslant M |f|_{X_D}, \qquad f \in X_D \tag{5.1}$$

for a constant $M > 0$, and if for $t \in (0, 1]$

$$S_t \subset X_D, \qquad D(S_t) = \bar{S}_t, \tag{5.2}$$

then D is said to be an *abstract differential operator* (from X_D in $\bar{X}$) or an *abstract derivative* of order ψ.

It should be mentioned that $\bar{X}$ denotes not denote the closure of X. For $f \in X_D$, i.e., f possesses a derivative of order ψ_D, we use also the notation $Df \in \bar{X}$. Throughout this section let $(X, \mathscr{S})$ and $(\bar{X}, \bar{\mathscr{S}})$ be two A-spaces with approximands $\mathscr{S} = \{S_t\}$, $\bar{\mathscr{S}} = \{\bar{S}_t\}$, respectively. First of all we examine the relationship between the best approximation of f in X and the best approximation of Df in $\bar{X}$. It turns out that the approximation order of Df reduces exactly by the order of the subspace X_D.

PROPOSITION 5.1. *Let D be an abstract derivative of order $\psi_D \in \Psi$ from $X_D \subset X$ into $\bar{X}$. If $f \in X_D$, then*

$$E_t(f; X) \leqslant M\psi_D(t)\, E_t(Df; \bar{X}), \qquad t \in (0, t_0]. \tag{5.3}$$

Conversely, if for $f \in X$ the series

$$\sum_{k=1}^{\infty} \frac{1}{k\psi_D(k^{-1})} E_{1/k}(f; X) \tag{5.4}$$

is convergent, then there hold $Df \in \bar{X}$ and for $t \in (0, 1]$ also

$$E_t(Df; \bar{X}) \leqslant M \sum_{k \geqslant [1/t]} \frac{1}{k\psi_D(k^{-1})} E_{1/k}(f; X) \tag{5.5}$$

as well as

$$E_t(Df; \bar{X}) \leqslant M \sum_{j=0}^{\infty} \frac{1}{\psi_D(2^{-j}t)} E_{2^{-j}t}(f; X). \tag{5.6}$$

Proof. First we obtain on using the estimate (5.1) and (5.2),

$$\frac{1}{M} \inf_{g_t \in S_t} |f-g_t|_{X_D} \leqslant E_t(Df; \bar{X}) \leqslant M \inf_{g_t \in S_t} |f-g_t|_{X_D} \tag{5.7}$$

for all $f \in X_D$ and $t \in (0, 1]$. To verify (5.3), we get for arbitrary $g_t \in S_t$ by using the Jackson inequality (4.1) with respect to X_D,

$$E_t(f; X) = E_t(f-g_t; X) \leqslant M\psi_D(t)\, |f-g_t|_{X_D}, \qquad t \in (0, t_0].$$

Taking the infimum over all $g_t \in S_t$, (5.3) is established. Now, let $f \in X$. We have to show that $f \in X_D$, if the series (5.4) is finite. For a given $\varepsilon > 0$ we choose approximands $g_h \in S_h$, $h \in (0, 1]$, satisfying

$$\|f-g_h\|_X \leqslant E_h(f; X) + h\psi_D(h)\varepsilon. \tag{5.8}$$

Similarly as in the proof of Lemma 4.3 we obtain on using the Bernstein inequality with respect to X_D,

$$\begin{aligned} |g_{2^{-j-1}} - g_{2^{-j}}|_{X_D} &\leqslant M \left\{ \frac{1}{\psi_D(2^{-j-1}t)} \|f-g_{2^{-j-1}t}\|_X + \frac{1}{\psi_D(2^{-j}t)} \|f-g_{2^{-j}t}\|_X \right\} \\ &\leqslant M \left\{ \frac{1}{\psi_D(2^{-j-1}t)} E_{2^{-j-1}t}(f; X) + 2^{-j-1}\varepsilon \right. \\ &\quad \left. + \frac{1}{\psi_D(2^{-j})} E_{2^{-j}t}(f; X) + 2^{-j}\varepsilon \right\}. \end{aligned}$$

Taking the sum over $j \in \mathbb{N}$ yields

$$\begin{aligned} \sum_{j=0}^{\infty} |g_{2^{-j-1}t} - g_{2^{-j}t}|_Z &\leqslant M \left\{ \varepsilon + \sum_{j=0}^{\infty} \frac{1}{\psi_D(2^{-j}t)} E_{2^{-j}t}(f; X) \right\} \\ &\leqslant M \left\{ \varepsilon + \sum_{k \geqslant [1/t]} \frac{1}{k\psi_D(k^{-1})} E_{1/k}(f; X) \right\}. \end{aligned} \tag{5.9}$$

Thus, by assumption (5.4), $\{g_{2^{-j}t}\}_{j \in \mathbb{N}_0}$ forms a Cauchy sequence with respect to $|\cdot|_{X_D}$, and, according to the Weierstraß property (3.6), with respect to $\|\cdot\|_{X_D}$ as well. Denoting its limit by $g \in X_D$ and using the Weierstraß property again, we find

$$\|f-g\|_X \leqslant \lim_{j \to \infty} \{\|f-g_{2^{-j}t}\|_X + \|g_{2^{-j}t} - g\|_X\} = 0,$$

concluding $f=g\in X_D$, and

$$|f-g_t|_{X_D}\leqslant\sum_{j=0}^{\infty}|g_{2^{-j-1}t}-g_{2^{-j}t}|_{X_D}.$$

Since ε is arbitrary, (5.9) implies

$$\begin{aligned}\inf_{h_t\in S_t}|f-h_t|_{X_D}&\leqslant M\sum_{j=0}^{\infty}\frac{1}{\psi_D(2^{-j}t)}E_{2^{-j}t}(f;X)\\&\leqslant M\sum_{k\geqslant[1/t]}\frac{1}{k\psi_D(k^{-1})}E_{1/k}(f;X),\end{aligned}$$

proving (5.5) and (5.6). ∎

Now we can state the theorem on simultaneous approximation, i.e., direct and inverse theorems involving the best approximation $E_t(Df;\bar{X})$ of Df.

THEOREM 5.2. *Let $Y\subset X$ be a subspace of order $\psi\in\Psi$ and D an abstract derivative of order $\psi_D\in\Psi$ from $X_D\subset X$ into $\bar{X}$ such that $\psi\prec\psi_D$. If the integral*

$$\int_0^1\frac{1}{\psi_D(u)}K(f,\psi(u);X,Y)\frac{du}{u}\tag{5.10}$$

is finite for $f\in X$, then $Df\in\bar{X}$, and there holds the direct estimate

$$E_t(Df;\bar{X})\leqslant M\int_0^t\frac{1}{\psi_D(u)}K(f,\psi(u);X,Y)\frac{du}{u},\qquad t\in(0,t_0].\tag{5.11}$$

For each $f\in X_D$ and $t\in(0,1]$ we have the inverse inequality

$$K(f,\psi(t);X,Y)\leqslant M\psi(t)\left\{\|f\|_X+\sum_{1\leqslant k\leqslant 1/t}\frac{\psi_D(k^{-1})}{k\psi(k^{-1})}E_{1/k}(Df;\bar{X})\right\}.\tag{5.12}$$

The statements of the theorem remain valid, if in (5.10), (5.11), or (5.12) the K-functional is replaced by the K^*-functional. Furthermore, the summation over $1\leqslant k\leqslant 1/t$ in estimate (5.12) can be replaced by summation over $1/t_1\leqslant k\leqslant 1/t$ for each fixed $t_1\in(0,1]$.

Proof. First, by Corollary 4.6 the expression (5.10) is finite iff the corresponding integral over the K^*-functional is finite. Inserting the direct estimate (4.3) into (5.5), the first estimate (5.11) follows immediately. In particular, the convergence of the above integral (5.3) implies $f\in X_D$.

Concerning the modified K^*-functional, estimating the term $E_{2^{-j}t}$ in (5.6) by the associated sum of the direct estimate (4.4), we obtain

$$\begin{aligned} E_t(Df;\bar{X}) &\leqslant M \sum_{j=0}^{\infty} \frac{1}{\psi_D(2^{-j}t)} \left(\sum_{k=0}^{\infty} K^*(f, \psi(2^{-k-j}t); X, Y) \right) \\ &= M \sum_{k=0}^{\infty} \left(\sum_{j=0}^{k} \frac{1}{\psi_D(2^{-j}t)} \right) K^*(f, \psi(2^{-k}t); X, Y) \\ &\leqslant M \sum_{k=0}^{\infty} \frac{1}{\psi_D(2^{-k}t)} K^*(f, \psi(2^{-k}t); X, Y). \end{aligned}$$

In the last step above, we used in view of Propositions 2.1 and 2.5 that

$$\sum_{j=0}^{k} \frac{1}{\psi_D(2^{-j}t)} \leqslant M \sum_{j;\ 2^{-k}t \leqslant 2^{-j} \leqslant 1} \frac{1}{\psi_D(2^{-j})} \leqslant M \frac{1}{\psi_D(2^{-k}t)}.$$

Thus, (5.11) holds for the K^*-functional also. It remains to show the inverse estimate. For $t \in [t_0, 1]$ the K-functionals are bounded by $M\psi(t)\,\|f\|_X$, and if $t \in (0, t_0]$ we can insert the inequality (5.3) into the inverse Theorem 4.4 to obtain the assertion (5.12) immediately. ∎

Theorems relating the behaviour of moduli of smoothness of functions and their derivatives are called reduction theorems, because by increasing the order of the derivative one can reduce the order of its modulus of smoothness. If for instance the modulus of smoothness $\omega_r(f, t)$ of a given function $f \in C_{2\pi}$ vanishes sufficiently rapidly, then $f^{(s)} \in C_{2\pi}$, $s < r$, and the

Again, the K-functional in (5.13) *or* (5.14) *may be replaced by the modified K^*-functional.*

Conversely, for $f \in X_D$ and $t \in (0, 1]$ there hold the estimates

$$K(f, \psi(t); X, Y) \leqslant M\psi(t) \left\{ \|f\|_X + \int_t^1 \frac{\psi_D(u)}{\psi(u)} K(Df, \bar{\psi}(u); \bar{X}, \bar{Y}) \frac{du}{u} \right\}, \quad (5.15)$$

$$K^*(f, \psi(t); X, Y) \leqslant M \left\{ \psi(t) \|f\|_X + \psi_D(t) \int_0^t K^*(Df, \bar{\psi}(u); \bar{X}, \bar{Y}) \frac{du}{u} \right.$$
$$\left. + \psi(t) \int_t^1 \frac{\psi_D(u)}{\psi(u)} K^*(Df, \bar{\psi}(u); \bar{X}, \bar{Y}) \frac{du}{u} \right\}. \quad (5.16)$$

Proof. Assuming (5.13) for the K- or K^*-functional the existence of the derivative $Df \in \bar{X}$ was already shown in Theorem 5.2. And again, on using Proposition 2.1 it suffices to verify the estimates for $t \in (0, t^*]$, $t^* := \min\{t_0, \bar{t}_0\}$. The direct estimate of Theorem 5.2 yields for $j \in \mathbb{N}$, $j \geqslant 1/t^*$

$$E_{1/j}(Df; \bar{X}) \leqslant M \sum_{k \geqslant j} \frac{1}{k\psi_D(k^{-1})} K^*(f, \psi(k^{-1}); X, Y),$$

and by Theorem 4.4 applied to $\bar{X}$, $\bar{Y}$ instead of X, Y it follows that

$$K(Df, \bar{\psi}(t); \bar{X}, \bar{Y}) \leqslant M\bar{\psi}(t) \left\{ \|Df\|_{\bar{X}} + \sum_{1/t_0 \leqslant j \leqslant 1/t} \frac{1}{j\bar{\psi}(j^{-1})} \right.$$
$$\left. \times \left(\sum_{k \geqslant j} \frac{1}{k\psi_D(k^{-1})} K^*(f, \psi(k^{-1}); X, Y) \right) \right\}.$$

We change the order of summation and apply (2.15) for $h = 1/k$ and $h = t$ to deduce

$$K(Df, \bar{\psi}(t); \bar{X}, \bar{Y})$$
$$\leqslant M\bar{\psi}(t) \left\{ \|Df\|_{\bar{X}} + \left(\sum_{1 \leqslant k \leqslant 1/t} \sum_{1 \leqslant j \leqslant k} + \sum_{k \geqslant 1/t} \sum_{1 \leqslant j \leqslant 1/t} \right) \right.$$
$$\left. \times \frac{1}{j\bar{\psi}(j^{-1})} \frac{1}{k\psi_D(k^{-1})} K^*(f, \psi(k^{-1}); X, Y) \right) \right\}$$
$$\leqslant M\bar{\psi}(t) \left\{ \|Df\|_{\bar{X}} + \sum_{1 \leqslant k \leqslant 1/t} \frac{1}{k\bar{\psi}(k^{-1})\, \psi_D(k^{-1})} \frac{1}{k\psi_D(k^{-1})} \right.$$
$$\left. \times K^*(f, \psi(k^{-1}); X, Y) + \frac{1}{\bar{\psi}(t)} \sum_{k \geqslant 1/t} \frac{1}{k\psi_D(k^{-1})} K^*(f, \psi(k^{-1}); X, Y) \right) \right\}$$

The corresponding integral representation together with (3.13) imply (5.14) for the K- and K^*-functionals.

The inverse estimate (5.15) follows from inequality (5.3) in Proposition 5.1 and Theorem 4.4 applied to the A-space $(\bar{X}, \bar{\mathscr{S}})$. Finally, to show the last assertion (5.16), we insert the estimate

$$E_{1/j}(Df; \bar{X}) \leqslant M \sum_{k \geqslant j} \frac{1}{k} K^*(Df, \bar{\psi}(k^{-1}); \bar{X}, \bar{Y}), \qquad j \geqslant 1/t^*,$$

which follows from Theorem 4.2, into the inverse inequality of Theorem 5.2 to obtain in the same way as above

$$\begin{aligned} K^*(f, \psi(t); X, Y) &\leqslant M\psi(t) \left\{ \|f\|_X + \sum_{1/t^* \leqslant j \leqslant 1/t} \frac{\psi_D(j^{-1})}{j\psi(j^{-1})} \right. \\ &\quad \left. \times \sum_{k \geqslant j} \frac{1}{k} K^*(Df, \bar{\psi}(k^{-1}); \bar{X}, \bar{Y}) \right\} \\ &\leqslant M\psi(t) \left\{ \|f\|_X + \left(\sum_{k \geqslant 1/t} \sum_{1 \leqslant j \leqslant 1/t} + \sum_{1 \leqslant k \leqslant 1/t} \sum_{1 \leqslant j \leqslant k} \right) \right. \\ &\quad \left. \times \frac{\psi_D(j^{-1})}{j\psi(j^{-1})} \frac{1}{k} K^*(Df, \bar{\psi}(k^{-1}); \bar{X}, \bar{Y}) \right\} \\ &\leqslant M \left\{ \psi(t) \|f\|_X + \psi_D(t) \sum_{k \geqslant 1/t} K^*(Df, \bar{\psi}(k^{-1}); \bar{X}, \bar{Y}) \right. \\ &\quad \left. + \psi(t) \sum_{1 \leqslant k \leqslant 1/t} \frac{\psi_D(k^{-1})}{k\psi(k^{-1})} K^*(Df, \bar{\psi}(k^{-1}); \bar{X}, \bar{Y}) \right\} \end{aligned}$$

Thus the theorem has been verified. ∎

Finally we establish abstract versions of the so called de la Vallée Poussin-Stečkin-type estimates.

THEOREM 5.4. *Let $\bar{Y} \subset \bar{X}$ be a subspace of order $\bar{\psi} \in \Psi$, and let D be a derivative of order $\psi_D \in \Psi$ on X_D into $\bar{X}$. For $f \in X_D$ and $t \in (0, t^*]$, $t^* := \min\{t_0, \bar{t}_0\}$, we have*

$$E_t(f; X) \leqslant M\psi_D(t)\, K(Df, \bar{\psi}(t); \bar{X}, \bar{Y}), \tag{5.17}$$

$$E_t(f; X) \leqslant M\psi_D(t) \int_0^t K^*(Df, \bar{\psi}(u); \bar{X}, \bar{Y}) \frac{du}{u}. \tag{5.18}$$

Conversely, if the series

$$\sum_{k=1}^{\infty} \frac{1}{k\psi_D(k^{-1})} E_{1/k}(f; X) \tag{5.19}$$

converges for a given $f \in X$, *then the derivative* $Df \in \bar{X}$ *exists, and the following estimates hold for* $t \in (0, 1]$

$$\begin{aligned} &K(Df, \bar{\psi}(t); \bar{X}, \bar{Y}) \\ &\quad \leqslant M \left\{ \bar{\psi}(t) \, \|Df\|_{\bar{X}} + \bar{\psi}(t) \sum_{1 \leqslant k \leqslant 1/t} \frac{1}{k\bar{\psi}(k^{-1}) \, \psi_D(k^{-1})} E_{1/k}(f; X) \right. \\ &\qquad \left. + \sum_{k \geqslant 1/t} \frac{1}{k\psi_D(k^{-1})} E_{1/k}(f; X) \right\}, \end{aligned} \tag{5.20}$$

as well as,

$$\begin{aligned} &K^*(Df, \bar{\psi}(t); \bar{X}, \bar{Y}) \\ &\quad \leqslant M \left\{ \bar{\psi}(t) \, \|Df\|_{\bar{X}} + \bar{\psi}(t) \sum_{1 \leqslant k \leqslant 1/t} \frac{1}{k\bar{\psi}(k^{-1})\psi_D(k^{-1})} E_{1/k}(f; X) \right. \\ &\qquad \left. + \sum_{k \geqslant 1/t} \frac{1}{k\psi_D(k^{-1})} E_{1/k}(f; X) \right\}. \end{aligned} \tag{5.21}$$

Proof. An application of Theorem 4.2 for $\bar{X}$ to (5.3) and (5.5) implies the assertions (5.17) and (5.18) immediately. Assuming (5.19), on using Proposition 5.1 we obtain $f \in X_D$, and analogously as in the proof of the last theorem we deduce the remaining estimates by inserting (5.5) for $t = 1/j$ into the inverse Theorem 4.4. ▌

Estimates between K-functionals with respect to subspaces of different orders are referred to as Marchaud-type inequalities.

THEOREM 5.5. *Let* X *be an* A*-space and let* $Y, Z \subset X$ *be subspaces of orders* $\psi, \psi_Z \in \Psi$. *Then for* $f \in X$ *and* $t \in (0, 1]$ *we have*

$$K(f, \psi(t); X, Y) \leqslant M\psi(t) \left\{ \|f\|_X + \int_t^1 \frac{1}{\psi(u)} K(f, \psi_Z(u); X, Z) \frac{du}{u} \right\}, \tag{5.22}$$

and

$$K^*(f, \psi(t); X, Y) \leqslant M \left\{ \psi(t) \, \|f\|_X + \int_0^t K^*(f, \psi_Z(u); X, Z) \, \frac{du}{u} \right.$$
$$\left. + \psi(t) \int_t^1 \frac{1}{\psi(u)} K^*(f, \psi_Z(u); X, Z) \, \frac{du}{u} \right\}. \quad (5.23)$$

If $\psi \prec \psi_Z$, *then* $Y \subset Z$, *and*

$$K(f, \psi(t); X, Y) \leqslant M\{\psi(t) \, \|f\|_X + K(f, \psi_Z(t); X, Z)\}, \qquad t \in (0, 1]. \quad (5.24)$$

Proof. As before, we have only to consider $t \in (0, t_0]$. The estimate (5.22) can be obtained readily by combining the direct and inverse theorems (4.3) and (4.6). Similarly, on using (4.4) instead of (4.3) with the inverse Theorem 4.4, we obtain

$$\begin{aligned} &K^*(f, \psi(t); X, Y) \\ &\quad \leqslant M\psi(t) \left\{ \|f\|_X + \sum_{1/t_0 \leqslant j \leqslant 1/t} \frac{1}{j\psi(j^{-1})} E_{1/j}(f) \right\} \\ &\quad \leqslant M\psi(t) \left\{ \|f\|_X + \sum_{1 \leqslant j \leqslant 1/t} \frac{1}{j\psi(j^{-1})} \sum_{k \geqslant j} \frac{1}{k} K^*(f, \psi_Z(k^{-1}); X, Z) \right\}. \end{aligned}$$

As before, on applying (2.15) and changing the order of summation, we get

$$\begin{aligned} K^*(f, \psi(t); X, Y) &\leqslant M\psi(t) \left\{ \|f\|_X + \left(\sum_{k \geqslant 1/t} \sum_{1 \leqslant j \leqslant 1/t} + \sum_{1 \leqslant k \leqslant 1/t} \sum_{1 \leqslant j \leqslant k} \right) \right. \\ &\qquad \left. \times \frac{1}{j\psi(j^{-1})} \frac{1}{k} K^*(f, \psi_Z(k^{-1}); X, Z) \right\} \\ &\leqslant M \left\{ \psi(t) \, \|f\|_X + \sum_{k \geqslant 1/t} \frac{1}{k} K^*(f, \psi_Z(k^{-1}); X, Z) \right. \\ &\qquad \left. + \psi(t) \sum_{1 \leqslant k \leqslant 1/t} \frac{1}{k\psi(k^{-1})} K^*(f, \psi_Z(k^{-1}); X, Z) \right\}, \end{aligned}$$

establishing (5.23).

Now let $\psi \prec \psi_Z$; then $1/\psi \succ 1$, $\psi_Z/\psi \succ 1$. Following the proof of the Jackson theorem, the Jackson inequality (4.1) implies for arbitrary $g \in Z$,

$$E_t(f) \leqslant E_t(f - g) + E_t(g) \leqslant \|f - g\|_X + \psi_Z(t) \, |g|_Z.$$

An application of the direct Theorem 4.4 and Proposition 2.5 yield

$$\begin{aligned} &K(f, \psi(t); X, Y) \\ &\quad \leqslant M\psi(t) \left\{ \|f\|_X + \sum_{1/t_0 \leqslant k \leqslant 1/t} \frac{1}{k\psi(k^{-1})} (\|f-g\|_X + \psi_Z(k^{-1}) \, |g|_Z) \right\} \\ &\quad \leqslant M\{\psi(t) \, \|f\|_X + \|f-g\|_X + \psi_Z(t) \, |g|_Z\}. \end{aligned}$$

We now take the infimum over $g \in Z$ to deduce (5.24). Finally, using the Jackson inequality again, for $f \in Y$ we have

$$E_t(f) \leqslant M\psi(t) \, |f|_Y = \mathcal{O}(\psi(t)), \qquad t \in (0, t_0];$$

thus Proposition 5.1 implies $f \in Z$, since $\psi \prec \psi_Z$. ∎

The interplay of the approximation orders between the theorems on simultaneous approximation and the reduction theorems become clearer, if we use Proposition 2.5 to formulate the foregoing theorems as $\mathcal{O}$-statements.

COROLLARY 5.6. *Let X, $\bar{X}$ be A-spaces, $Y \subset X$, $\bar{Y} \subset \bar{X}$ subspaces of orders $\psi, \bar{\psi} \in \Psi$, respectively, and D: $X_D \to \bar{X}$ an abstract derivative of order $\psi_D \in \Psi$. Then for $f \in X$ and $\phi \in \Phi$ satisfying $\psi \prec \phi \prec \psi_D$ and $\phi \succ \bar{\psi}\psi_D$ the following statements are equivalent*:

(i) $\qquad E_t(f; X) = \mathcal{O}(\phi(t)), \qquad t \to 0^+;$

(ii) $\qquad K(f, \psi(t); X, Y) = \mathcal{O}(\phi(t)), \qquad t \to 0^+;$

(ii)* $\qquad K^*(f, \psi(t); X, Y) = \mathcal{O}(\phi(t)), \qquad t \to 0^+;$

(iii) $\qquad Df \in \bar{X}$ and $E_t(Df; \bar{X}) = \mathcal{O}\left(\dfrac{\phi(t)}{\psi_D(t)}\right), \qquad t \to 0^+;$

(iv) $\qquad Df \in \bar{X}$ and $K(Df, \bar{\psi}(t); \bar{X}, \bar{Y}) = \mathcal{O}\left(\dfrac{\phi(t)}{\psi_D(t)}\right), \qquad t \to 0^+;$

(iv)* $\qquad Df \in \bar{X}$ and $K^*(Df, \bar{\psi}(t); \bar{X}, \bar{Y}) = \mathcal{O}\left(\dfrac{\phi(t)}{\psi_D(t)}\right), \qquad t \to 0^+.$

The equivalences between (i) and (ii), (ii)* are the direct and inverse theorems. The simultaneous approximation is given by the equivalences of (iii) and (ii), (ii)*, and the implications concerning (ii), (ii)* and (iv), (iv)* are called reduction theorems. Finally the theorems involving the equivalences of (i) and (iv), (iv)* are referred to as theorems of de la Vallée Poussin-Stečkin-type.

6. BEST APPROXIMATION BY TRIGONOMETRIC POLYNOMIALS

In this section we wish to apply the abstract theory to the problem of best approximation by trigonometric polynomials. To this end let $L^p_{2\pi}$, $1 \leqslant p < \infty$, be the Banach space of the 2π-periodic Lebesgue measurable functions $f\colon \mathbb{R} \to \mathbb{C}$ endowed with the norm

$$\|f\|_p := \left\{ \int_{-\pi}^{\pi} |f(x)|^p \, dx \right\}^{1/p}.$$

For simplicity we identify $L^\infty_{2\pi}$ with the space of all 2π-periodic continuous functions on $\mathbb{R}$ equipped with the usual supremum norm $\|f\|_\infty := \sup_{x \in \mathbb{R}} |f(x)|$. In the following let $1 \leqslant p \leqslant \infty$. Because of the periodicity there is no need to separate the endpoints; thus we can identify the norms by $\|\cdot\|_X = \|\cdot\|_{X,t} := \|\cdot\|_p$ for all $t \in (0, 1]$. If we denote by Π_n the set of trigonometric polynomials $t_n(x) = \sum_{k=-n}^{n} a_k e^{ikx}$, $a_k \in \mathbb{C}$, of degree not exceeding $n \in \mathbb{N}_0$, then the *best approximation* of $f \in L^p_{2\pi}$ is given by

$$E_n(f, L^p_{2\pi}) := \inf_{t_n \in \Pi_n} \|f - t_n\|_p.$$

Since the union of the subspaces Π_n is dense in $L^p_{2\pi}$, the couple $(L^p_{2\pi}, \{\Pi_n\}_{n \in \mathbb{N}_0})$ forms a approximation space in the sense of Definition 3.1 if we use the discretization $n = [1/t]$. We denote by $W^{p,r}_{2\pi}$ the Sobolev space of all functions $f \in L^p_{2\pi}$ which coincide almost everywhere with an $(r-1)$-times continuously differentiable function g, $g^{(r-1)}$ being absolutely continuous with $g^{(r)} \in L^p_{2\pi}$, i.e., $f \in W^{p,r}_{2\pi}$ iff $f^{(r)}$ exists almost everywhere and belongs to $L^p_{2\pi}$. Of course, if $p = \infty$, $W^{\infty,r}_{2\pi}$ is the space of all r-times continuously differentiable 2π-periodic functions on $\mathbb{R}$. In $L^p_{2\pi}$ we can identify the K-functionals with moduli of smoothness. Defining for $r \in \mathbb{N}$ the rth (centered) difference of f with increment $h > 0$ by $\Delta^1_h f(x) := f(x + h/2) - f(x - h/2)$, $\Delta^{r+1}_h := \Delta^1_h \Delta^r_h$, then the *r-th modulus of smoothness* is given by

$$\omega_r(f, t) := \sup_{0 < h \leqslant t} \|\Delta^r_h f\|_p, \qquad t > 0.$$

The following well known equivalence can be found in, e.g., R. A. DeVore and G. G. Lorentz [10, Chap. 7, Sect. 2].

PROPOSITION 6.1. *For $r \in \mathbb{N}$ there holds*

$$\omega_r(f, t) \sim K(f, t^r; L^p_{2\pi}, W^{p,r}_{2\pi}), \qquad t \in (0, 1],$$

where the constants induced are independent of f and t.

The Sobolev spaces $W^{p,r}_{2\pi}$ form subspaces of order t^r or r for short; this is an immediate consequence of the validity of the Jackson and Bernstein inequalities below (cf. P. L. Butzer and R. J. Nessel [5, p. 99] or [10, pp. 97, 202]).

PROPOSITION 6.2. *For $r \in \mathbb{N}$ there hold*

$$E_n(f) \leqslant Mn^{-r} \|f^{(r)}\|_p, \qquad f \in W^{p,r}_{2\pi},\ n \in \mathbb{N}_0, \tag{6.1}$$

$$\|t_n^{(r)}\|_p \leqslant Mn^r \|t_n\|_p, \qquad t_n \in \Pi_n,\ n \in \mathbb{N}_0. \tag{6.2}$$

Now, an application of Theorems 4.2, 4.4 implies the direct and inverse theorem on best trigonometric approximation.

COROLLARY 6.3. *If $f \in L^p_{2\pi}$ and $r \in \mathbb{N}$, then*

$$E_n(f; L^p_{2\pi}) \leqslant M\omega_r(f, n^{-1}), \qquad n \in \mathbb{N}, \tag{6.3}$$

as well as

$$\omega_r(f, t) \leqslant Mt^r \sum_{0 \leqslant k \leqslant 1/t} (k+1)^{r-1} E_k(f; L^p_{2\pi}), \qquad t \in (0, 1]. \tag{6.4}$$

Turning over to simultaneous approximation we have to identify abstract differentiation with the ordinary derivative. Noting that differentiation of trigonometric polynomials does not reduce the degree of the polynomial, the null-th Fourier coefficient vanishes instead. Therefore we denote by $\bar{L}^p_{2\pi}$, $\bar{W}^{p,r}_{2\pi}$, and $\bar{\Pi}_n$ the spaces of all elements f of $L^p_{2\pi}$, $W^{p,r}_{2\pi}$, and Π_n, respectively, satisfying $f^\wedge(0) = 1/(2\pi) \int_{-\pi}^{\pi} f(u)\, du = 0$. These spaces have the same approximation properties as do their non-bared counterparts. To see this, we use for $f \in \bar{L}^p_{2\pi}$ and $t_n \in \Pi_n$ the estimate

$$\|f - (t_n - t_n{}^\wedge(0))\|_p = \|f - t_n - (f^\wedge(0) - t_n{}^\wedge(0))\|_p \leqslant M \|f - t_n\|_p,$$

to obtain

$$E_n(f; L^p_{2\pi}) \leqslant E_n(f; \bar{L}^p_{2\pi}) \leqslant ME_n(f; L^p_{2\pi}), \qquad f \in \bar{L}^p_{2\pi},\ n \in \mathbb{N}_0. \tag{6.5}$$

Similarly, for $f \in \bar{L}^p_{2\pi}$ and $r \in \mathbb{N}$ there hold

$$K(f, t^r; L^p_{2\pi}, W^{p,r}_{2\pi}) \leqslant K(f, t^r; \bar{L}^p_{2\pi}, \bar{W}^{p,r}_{2\pi}) \leqslant MK(f, t^r; L^p_{2\pi}, W^{p,r}_{2\pi}),$$

yielding

$$\frac{1}{M}\, \omega_r(f, t) \leqslant K(f, t^r; \bar{L}^p_{2\pi}, \bar{W}^{p,r}_{2\pi}) \leqslant M\omega_r(f, t), \qquad f \in \bar{L}^p_{2\pi}, \quad t > 0. \tag{6.6}$$

Hence, it is sufficient to consider only the best approximation $E_n(f; L^p_{2\pi})$. Furthermore, it follows that the spaces $(\bar{L}^p_{2\pi}, \{\bar{\Pi}_n\}_{n\in\mathbb{N}_0})$ form A-spaces, and the Sobolev spaces $\bar{W}^{p,r}_{2\pi}$ are subspaces of order $r \in \mathbb{N}$. In particular, the operator

$$D^r\colon W^{p,r}_{2\pi} \to \bar{L}^p_{2\pi}, \qquad f \mapsto f^{(r)}$$

is an abstract derivative of order r (or t^r).

On using Theorem 5.2 and the fact that $t^r \prec t^l$ iff $l < r$, the direct and inverse theorems on simultaneous approximation now read

Corollary 6.4. *Let $r, l \in \mathbb{N}$ such that $l < r$. If for $f \in L^p_{2\pi}$ the integral*

$$\int_0^1 u^{-l-1}\omega_r(f, u)\, du$$

is finite, then f belongs to $f \in W^{p,l}_{2\pi}$, and

$$E_n(f^{(l)}; L^p_{2\pi}) \leqslant M \int_0^{1/n} u^{-l-1}\omega_r(f, u)\, du.$$

Conversely, if $f \in W^{p,l}_{2\pi}$, then we have

$$\omega_r(f, t) \leqslant Mt^r \sum_{0 \leqslant k \leqslant 1/t} (k+1)^{r-l-1} E_k(f^{(l)}; L^p_{2\pi}), \qquad t \in (0, 1].$$

It is well known that sufficiently fast convergence of the best approximation implies the differentiability of the underlying function. The corresponding weak type inequalities of de la Vallée Poussin-Stečkin type follow from Theorem 5.4

Corollary 6.5. *Let $l, r \in \mathbb{N}$. If for $f \in L^p_{2\pi}$ the series*

$$\sum_{k=1}^{\infty} k^{l-1} E_k(f; L^p_{2\pi})$$

converges, then $f \in W^{p,l}_{2\pi}$. Additionally, for all $f \in W^{p,l}_{2\pi}$ there hold the estimates

$$E_n(f; L^p_{2\pi}) \leqslant Mn^{-l}\omega_r(f^{(l)}, 1/n)$$

and

$$\omega_r(f^{(l)}, t) \leqslant Mt^r \sum_{0 \leqslant k \leqslant 1/t} (k+1)^{r+l-1} E_k(f; L^p_{2\pi}) + M \sum_{k \geqslant 1/t} k^{l-1} E_k(f; L^p_{2\pi}).$$

7. BEST WEIGHTED ALGEBRAIC APPROXIMATION

We now wish to apply the abstract approximation theorems to the much more delicate case of best approximation by algebraic polynomials in weighted L_p spaces. This requires the full use of the exhaustion method by a family of seminorms. Let $\varphi(x) := \sqrt{x(1-x)}$, $x \in [0, 1]$, (in order to keep the common notations, we distinguish between φ and order functions ϕ); then for $1 \leqslant p < \infty$ and $\mu > -2/p$ the Banach spaces L^p_μ consist of all measurable functions $f\colon [0, 1] \to \mathbb{C}$ with finite norm

$$\|f\|_{p,\mu} := \left\{ \int_0^1 |f(x)\, \varphi^\mu(x)|^p \, dx \right\}^{1/p}.$$

As in the preceding section, for $p = \infty$ we denote by L^∞_μ, $\mu \geqslant 0$, the space of all $f\colon [0, 1] \to \mathbb{C}$ for which $\varphi^\mu f$ is continuous on $[0, 1]$ and the associated supremum norm $\|f\|_{\infty,\mu} := \sup_{x \in [0,1]} |f(x)\, \varphi^\mu(x)|$ is finite. The spaces L^p_μ are exactly the Gegenbauer or ultrasperical weighted spaces with weight $\varphi^{\mu p}$, to which we restrict ourselves; however, the following will work in Jacobi weighted spaces as well.

In order to separate the endpoints of the interval $[0, 1]$, we define for a fixed chosen constant $c > 0$ the seminorms $|\cdot|_{p,\mu,t}$, $t \in (0, 1]$, on L^p_μ by

$$|f|_{p,\mu,t} := \begin{cases} \left\{ \displaystyle\int_{ct^2}^{1-ct^2} |f(x)\, \varphi^\mu(x)|^p \, dx \right\}^{1/p}, & 1 \leqslant p < \infty \\ \displaystyle\sup_{x \in [ct^2, 1-ct^2]} |f(x)\, \varphi^\mu(x)|, & p = \infty. \end{cases} \tag{7.1}$$

For convenience we let $|\cdot|_{p,\mu,t} = 0$ if $c^{-1/2}/2 \leqslant t \leqslant 1$. Then the seminorms obviously satisfy the exhaustion property (3.1) and (3.2).

The set of the algebraic polynomials $p_n(x) := \sum_{k=0}^n a_k x^k$, $a_k \in \mathbb{C}$, of degree not exceeding $n \in \mathbb{N}_0$ is denoted by $\mathscr{P}_n$; thus the best and best modified approximation of $f \in L^p_\mu$ are given by

$$E_n(f; L^p_\mu) := \inf_{p_n \in \mathscr{P}_n} \|f - p_n\|_{p,\mu}, \qquad E^*_n(f; L^p_\mu) := \inf_{p_n \in \mathscr{P}_n} |f - p_n|_{p,\mu,1/n},$$

respectively. It is well known that the Gegenbauer polynomials span the whole space L^p_μ, so that the Weierstraß property is satisfied. The following equivalence condition for the seminorms has been proved by P. G. Nevai [23] and by Z. Ditzian and V. Totik [13, Theorem 8.4.7], based on estimates of M. K. Potapov [26].

PROPOSITION 7.1. *For all $c>0$ exists some constant $M=M(c)>0$ such that*

$$\|p_n\|_{p,\mu} \leqslant M\, |p_n|_{p,\mu,2/n}, \qquad p_n \in \mathscr{P}_n, \tag{7.2}$$

for all $n\in\mathbb{N}$, $n>2\sqrt{c}$.

In contrast to the trigonometric case we have to operate with weighted derivatives. Therefore we define for $r\in\mathbb{N}$ the weighted Sobolev spaces by the set of all $f\in L^p_\mu$ which coincide with an $(r-1)$-times continuously locally differentiable function g such that $\varphi^r g^{(r)}$ belongs to L^p_μ. Thus for $f\in W^{p,r}_\mu$ we have $f^{(r)}\in L^p_{\mu+r}$, and the associated seminorm is given by $\|f\|_{W^{p,r}_\mu} := \|f^{(r)}\|_{p,\mu+r}$. We need the following Jackson and Bernstein-type inequalities to ensure that the spaces $W^{p,r}_\mu$ are subspaces of order r (or t^r).

PROPOSITION 7.2. *Let $r\in\mathbb{N}$. For all $f\in W^{p,r}_\mu$ we have*

$$E^*_n(f;L^p_\mu) \leqslant Mn^{-r}\,\|f^{(r)}\|_{p,\mu+r}, \qquad n>r, \tag{7.3}$$

and for all $p_n\in\mathscr{P}_n$

$$\|p_n^{(r)}\|_{p,\mu+r} \leqslant Mn^r\,\|p_n\|_{p,\mu}, \qquad n\in\mathbb{N}_0. \tag{7.4}$$

The Jackson inequality (7.3) due to R. A. DeVore and L. R. Scott [11] for $p=1$, the general case can be found in P. L. Butzer, S. Jansche and R. L. Stens [4]. The Bernstein inequality (7.4) was established by B. A. Khalilova [19].

We now have to identify the spaces above with A-spaces $(X_\mu, \mathscr{S}_\mu)$ in the sense of Definition 3.1. For the underlying A-spaces we set $X_\mu := L^p_\mu$ and the subspaces are given by $Y_\mu := W^{p,r}_\mu$. It turns out to be useful for simultaneous approximation to use a discretization for the approximands S_t depending on the parameter μ. In fact, we set $S_{t,\mu} := \mathscr{P}_{[1/t-\mu-1]_+}$, where

$$x_+ := \begin{cases} x, & x\geqslant 0\\ 0, & x<0, \end{cases}$$

$[x]$ denoting the integer part of $x\in\mathbb{R}$. We use the notation $\{\mathscr{P}_{[n-\mu]_+}\}_{n\in\mathbb{N}_0}$ instead of $\{S_{t,\mu}\}_{t\in(0,1]}$ to prove

THEOREM 7.3. *The spaces $(L^p_\mu, \{\mathscr{P}_{[n-\mu]_+}\}_{n\in\mathbb{N}_0})$ endowed with the norm $\|\cdot\|_{p,\mu}$ and the family of seminorms $\{|\cdot|_{p,\mu,t};\, 0<t\leqslant 1\}$ are A-spaces, and the*

Sobolev spaces $W_\mu^{p,r}$ *are complete subspaces of order* $r \in \mathbb{N}$. *The associated (discrete) constants* $n_0 = n_0(r)$ *and* $C^* = C^*(\mu)$ *are given by*

$$n_0 := \max\{r+1, [2\sqrt{c}\,]+1\}, \qquad C^* := \begin{cases} \frac{1}{2}, & \mu \geqslant -1 \\ -\mu/2, & \mu < -1. \end{cases} \tag{7.5}$$

Furthermore, the operators

$$D^r \colon W_\mu^{p,r} \to L^p_{\mu+r}, \qquad f \mapsto f^{(r)},$$

are derivatives of order $r \in \mathbb{N}$.

Proof. We have to show that the conditions of Definition 3.1 are satisfied for the spaces $(X_\mu, \{S_{t,\mu}\})$. As mentioned above, the exhaustion property (3.1), (3.2), and (3.3) as well as the Weierstraß property (3.6) are given. Now let $0 < t \leqslant 1/([2\sqrt{c}\,]+\mu+2)$; then we have $t/C^* \leqslant 2/[1/t-\mu-1]_+$, and on using Proposition 7.1 there follows

$$\|g\|_{p,\mu} \leqslant M\,|g|_{p,\mu,2/[1/t-\mu-1]_+} \leqslant M\,|g|_{p,\mu,t/C^*},$$

which proves (3.8).

Noting that $[1/t-\mu-1]_+ \sim 1/t-\mu-1$, $t<1/(\mu+1)$, the weak Jackson inequality (4.2) (for $E_t^*(f, X_\mu) = \inf_{g_t \in S_t} |f-g_t|_{X_\mu, t}$) holds for $t \in (0, 1/(r+\mu+2)]$, while the Bernstein inequality (4.5) is valid for $t \in (0, 1]$. It can be easily shown that the operator $f \mapsto \varphi^r f^{(r)}$ is closed, concluding that the spaces $W_\mu^{p,r}$ are subspaces of order r. In particular, the constant t_0 of Section 3 can be chosen as

$$t_0 = t_0(r,\mu) := \min\left\{\frac{1}{[2\sqrt{c}\,]+\mu+2}, \frac{1}{r+\mu+2}\right\}.$$

This gives $n_0 = n_0(r) := \max\{r+1, [2\sqrt{c}\,]+1\}$ for our discretization. Finally, we have

$$D^r(S_{t,\mu}) = D^r(\mathscr{P}_{[1/t-\mu-1]_+}) = \mathscr{P}_{[1/t-\mu-1-r]_+} = S_{t,\mu+r}, \qquad t \in (0, 1].$$

This together with the definition of the seminorm imply that D^r is an abstract derivative of order r. ∎

S. M. Niloskiĭ [24] has already shown that the accuracy of approximation by algebraic polynomials increases towards the endpoints of the interval. This has to be taken into account by the definition of a suitable modulus of smoothness. For $f \in L^p_\mu$ and $r \in \mathbb{N}$ the main part modulus of Ditzian and Totik is given by

$$\Omega_{r,\mu}(f,t) \equiv \Omega_{r,\mu}(f,t; L^p_\mu) := \sup_{0<h\leqslant t} |\Delta^r_{h\varphi(x)} f(x)|_{p,\mu,h}, \quad 0<t\leqslant 1, \tag{7.6}$$

where we set $c=2r^2$ in (7.1). If $\mu=0$, we simply write $\Omega_r(f, t)=\Omega_{r,0}(f, t)$; and we define the ordinary Ditzian–Totik modulus by

$$\omega_r(f, t)\equiv\omega_r(f, t; L^p) := \sup_{0<h\leqslant t} \|\Delta^r_{h\varphi(x)} f(x)\|_p, \qquad 0<t\leqslant 1. \tag{7.7}$$

with the convention that $\Delta^r_{h\varphi(x)} f(x) := 0$, if $x\pm r/2h\varphi(x)\notin[0, 1]$. The link to the abstract theory is the following equivalence between the K-functionals and the moduli of smoothness.

PROPOSITION 7.4. *For $r\in\mathbb{N}$ we have*

$$\Omega_{r,\mu}(f, t)\sim K^*(f, t^r; L^p_\mu, W^{p,r}_\mu), \qquad 0\leqslant t\leqslant 1, \tag{7.8}$$

independently of $f\in L^p_\mu$. Similarly in the case $\mu=0$ there holds for $f\in L^p$

$$\omega_r(f, t)\sim K(f, t^r; L^p, W^{p,r}), \qquad 0\leqslant t\leqslant 1. \tag{7.9}$$

The first equivalence (7.8) was proven in P. L. Butzer, S. Jansche and R. L. Stens [4] by using Theorem 6.2.1. in Z. Ditzian and V. Totik [13]. The second equivalence is given in [13, Theorem 2.1.1]. In the definition of the main part modulus, the constant c of the seminorms is fixed by $c=2r^2$. But, on using Lemma 3.2, the constant $c>0$ in the definition of the seminorm of the K^*-functional can be chosen arbitrarily provided that the parameter $t>0$ is sufficiently small. Therefore we can set $c=r^2/4$ in (7.1), which gives $n_0=n_0(r):=r+1$ for its discrete analogue.

Now we are in position to apply the abstract theorems to algebraic approximation. In the non-weighted case $\mu=0$ Proposition 4.5 gives estimates between the ordinary and the main part modulus of Ditzian and Totik.

COROLLARY 7.5. *If $f\in L^p$, then*

$$\begin{aligned}\Omega_r(f, t)&\leqslant\omega_r(f, t)\\ &\leqslant M\left\{t^r\,\|f\|_{p,\mu}+\int_0^t\Omega_r(f, u)\,\frac{du}{u}+t^r\int_t^1 u^{-r-1}\Omega_r(f, u)\,du\right\}.\end{aligned}$$

Now we wish to prove Theorem 1.2, namely the direct and inverse inequalities of best weighted algebraic approximation.

Proof of Theorem 1.2. Let $n\geqslant r+1$ and choose $t:=1/(n+\mu+1)$; then, noting that $nt=1-t(\mu+1)\sim 1$, we deduce from Theorem 4.2, Propositions 7.4, and 2.1

$$E_n(f; L^p_\mu)\leqslant M\int_0^t K^*(f, u^r; L^p_\mu, W^{p,r}_\mu)\,\frac{du}{u}\leqslant M\int_0^{1/n}\Omega_{r,\mu}(f, u)\,\frac{du}{u}.$$

This proves the direct estimate. Recalling the definition of our discretization,

$$\inf_{g \in S_{1/k,\mu}} \|f-g\|_{p,\mu} = E_{[k-\mu-1]_+}(f; L^p_\mu).$$

Inserting this into the inverse Theorem 4.4, we obtain

$$\Omega_{r,\mu}(f,t) \leqslant M t^r \sum_{1 \leqslant k \leqslant 1/t} k^{r-1} E_{[k-\mu-1]_+}(f; L^p_\mu), \qquad t \in (0,1].$$

Assuming that $1/t \geqslant [\mu]+2 \geqslant 0$, for $k \geqslant [\mu]+2$ the monotonicity of the best approximation implies $E_{[k-\mu-1]_+}(f; L^p_\mu) \leqslant E_{k-[\mu]-2}(f; L^p_\mu)$, yielding with $j = k-[\mu]-2$,

$$\begin{aligned}\Omega_{r,\mu}(f,t) &\leqslant M t^r \sum_{0 \leqslant j \leqslant 1/t-[\mu]-2} (j+1)^{r-1} E_j(f; L^p_\mu)\\ &\leqslant M t^r \sum_{0 \leqslant j \leqslant 1/t} (j+1)^{r-1} E_j(f; L^p_\mu).\end{aligned}$$

Thus the inverse estimate of Theorem 1.2 is proved also. Concerning the case $1/t < [\mu]+2$, we use the fact that $\Omega_{r,\mu}(f,t) = \Omega_{r,\mu}(f, t+p_0)$ for all $p_0 \in \mathscr{P}_0$ to deduce $\Omega_{r,\mu}(f,t) \leqslant M E_0(f; L^p_\mu)$, $1/t < [\mu]+2$, completing the proof. ▮

On using Theorem 5.2 instead of Theorems 4.2 and 4.4, the direct and inverse estimates concerning simultaneous approximation can be proved along the same lines.

COROLLARY 7.6. *Let $r, l \in \mathbb{N}$ satisfy $l < r$. If for $f \in L^p_\mu$ the integral*

$$\int_0^1 u^{-l-1} \Omega_{r,\mu}(f,u)\, du$$

is finite, then we have $f^{(l)} \in L^p_{\mu+l}$, and

$$E_n(f^{(l)}; L^p_{\mu+l}) \leqslant M \int_0^{1/n} u^{-l-1} \Omega_{r,\mu}(f,u) \frac{du}{u}, \qquad n \geqslant r+1.$$

Conversely, for all $f \in W^{p,l}_\mu$, $t \in (0,1]$,

$$\Omega_{r,\mu}(f,t) \leqslant M t^r \left\{ \|f\|_{p,\mu} + \sum_{0 \leqslant k \leqslant 1/t} (k+1)^{r-l-1} E_k(f^{(l)}; L^p_{\mu+l}) \right\}, \quad t \in (0,1].$$

In the case $\mu = 0$ the estimates remain valid, if Ω_r is replaced by ω_r.

An application of Theorem 5.4 yields the algebraic counterparts of the de la Vallée Poussin-Stečkin theorems.

COROLLARY 7.7. *Let* $l, r \in \mathbb{N}$. *If for* $f \in L^p_\mu$ *the series*

$$\sum_{k=0}^{\infty} k^{l-1} E_{1/k}(f; L^p_\mu)$$

is convergent, then we have $f \in W^{p,l}_\mu$, *satisfying for* $t \in (0, 1]$

$$\Omega_{r,\mu+l}(f^{(l)}, t) \leqslant M t^r \sum_{0 \leqslant k \leqslant 1/t} k^{r+l-1} E_k(f; L^p_\mu) + M \sum_{k \geqslant 1/t} k^{l-1} E_k(f; L^p_\mu).$$

For all $f \in W^{p,l}_\mu$ *and* $n \geqslant \max\{r, l\} + 1$ *we have*

$$E_n(f; L^p_\mu) \leqslant M n^{-l} \int_0^{1/n} \Omega_{r,\mu+l}(f^{(l)}, u) \frac{du}{u}.$$

Inequalities of Marchaud type and reduction theorems are collected in the following

COROLLARY 7.8. *Let* $f \in L^p_\mu$, $l, r, s \in \mathbb{N}$, *and* $t \in (0, 1]$.

(a) *We have*

$$\Omega_{r,\mu}(f, t) \leqslant M \left\{ t^r \|f\|_{p,\mu} + \int_0^t \Omega_{s,\mu}(f, u) \frac{du}{u} + t^r \int_t^1 u^{-r-1} \Omega_{s,\mu}(f, u)\, du \right\},$$

and in the case $\mu = 0$,

$$\omega_r(f, t) \leqslant M t^r \left\{ \|f\|_p + \int_t^1 u^{-r-1} \omega_s(f, u)\, du \right\}.$$

In particular, if $s < r$, *then*

$$\omega_r(f, t) \leqslant M\{ t^r \|f\|_p + \omega_s(f, t)\}.$$

(b) *If for* $l < r$ *the integral*

$$\int_0^1 u^{-l-1} \Omega_{r,\mu}(f, u)\, du$$

is convergent, then $f^{(l)} \in L^p_{\mu+l}$, *and*

$$\Omega_{s,\mu+l}(f^{(l)}, t) \leqslant M \left\{ t^s \|f\|_{p,\mu} + \int_0^t u^{-l-1} \Omega_{r,\mu}(f, u)\, du + t^s \int_t^1 u^{-s-l-1} \Omega_{r,\mu}(f, u)\, du \right\}.$$

(c) *If* $l < r$, *then for all* $f \in W^{p,l}_\mu$,

$$\Omega_{r,\mu}(f, t) \leqslant M \left\{ t^r \|f\|_{p,\mu} + t^l \int_0^t \Omega_{s,\mu+l}(f^{(l)}, u)\, \frac{du}{u} + t^r \int_t^1 u^{l-r-1} \Omega_{s,\mu+l}(f^{(l)}, u)\, du \right\}.$$

Proof. Using the equivalences between the K-functionals and the moduli of smoothness, (a) follows from Theorem 5.5. The estimates in (b) and (c) are immediate consequences of Theorem 5.3, if the spaces X, $\bar{X}$ are identified with L^p_μ and $L^p_{\mu+r}$, respectively, noting that $D^l\colon W^{p,l}_\mu \to L^p_{\mu+l}$ forms an abstract derivative of order l in the sense of Definition 5.1. ∎

Finally, the assertions of the preceding corollaries can be collected in terms of $\mathcal{O}$-equivalences.

COROLLARY 7.9. *Let* $l, r, s \in \mathbb{N}$, $l < r, s$ *and* $\phi \in \Phi$ *such that* $t^r \prec \phi(t) \prec t^l$ *and* $t^{s+l} \prec \phi(t)$. *Then the following four assertions are equivalent for* $f \in L^p_\mu$:

(i) $E_n(f, L^p_\mu) = \mathcal{O}(\phi(n^{-1})), \quad n \to \infty;$

(ii) $\Omega_{r,\mu}(f, t) = \mathcal{O}(\phi(t)), \quad t \to 0^+;$

(iii) $f^{(l)} \in L^p_{\mu+l}$ and $E_n(f^{(l)}, L^p_{\mu+l}) = \mathcal{O}(n^l \phi(n^{-1})), \quad n \to \infty;$

(iv) $f^{(l)} \in L^p_{\mu+l}$ and $\Omega_{s,\mu+l}(f^{(l)}, t) = \mathcal{O}(t^{-l}\phi(t)), \quad t \to 0^+.$

In particular we can choose

$$\phi(t) = t^\alpha |\log t|^\beta, \qquad t \in (0, 1),$$

for $l < \alpha < r$, $\alpha < s + l$, and $\beta \in \mathbb{R}$.

ACKNOWLEDGMENT

The author thanks P. L. Butzer (Aachen) for his careful examination of the manuscript and suggestions for improving the presentation of this paper.

REFERENCES

1. S. Aljančić and D. Aran$\bar{\text{d}}$elović, O-regularly varying functions, *Publ. Inst. Math. (Beograd) (N.S.)* **22**, No. 36 (1977), 5–22.
2. S. N. Bernstein, Sur l'ordre de la meilleure approximation des fonctions continues par des polynômes de degré donné, *Mém. Cl. Sci. Acad. Roy. Belg. (2)* **4** (1922), 1–103.
3. N. H. Bingham, C. M. Goldie, and J. L. Teugels, "Regular Variation," Encyclopedia Math. Appl., Vol. 27, Cambridge Univ. Press, Cambridge, 1987.
4. P. L. Butzer, S. Jansche, and R. L. Stens, Functional analytic methods in the solution of the fundamental theorems on best algebraic approximation, *in* "Approximation Theory" (G. A. Anastassiou, Ed.), Proc. 6th Southeastern Approximation Theorists Annual Conf., Memphis, 1991, Lecture Notes in Pure and Appl. Math., Vol. 138, pp. 151–205, Dekker, New York, 1992.
5. P. L. Butzer and R. J. Nessel, "Fourier Analysis and Approximation," Lehrbücher und Monographien aus dem Gebiete der exakten Wissenschaften, Mathematische Reihe, Vol. 40, Birkhäuser, Basel, 1971.
6. P. L. Butzer and S. Pawelke, Ableitungen von trigonometrischen Approximationsprozessen, *Acta Sci. Math. (Szeged)* **28** (1967), 173–183.
7. P. L. Butzer and K. Scherer, On the fundamental approximation theorems of D. Jackson, S. N. Bernstein and theorems of M. Zamansky and S. B. Stečkin, *Aequationes Math.* **3** (1969), 170–185.
8. P. L. Butzer and K. Scherer, Über die Fundamentalsätze der klassischen Approximationstheorie in abstrakten Räumen, *in* "Abstract Spaces and Approximation" (P. L. Butzer and B. Sz. Nagy, Eds.), Proc. Internat. Conf., Math. Res. Inst. Oberwolfach, 1968, Internat. Ser. Numer. Math., Vol. 10, pp. 113–125, Birkhäuser, Basel, 1969.
9. P. L. Butzer and K. Scherer, Jackson and Bernstein-type inequalities for families of commutative operators in Banach spaces, *J. Approx. Theory* **5** (1972), 308–342.
10. R. A. DeVore and G. G. Lorentz, "Constructive Approximation," Grundlehren Math. Wiss., Vol. 303, Springer-Verlag, Berlin, 1991.
11. R. A. DeVore and L. R. Scott, Error bounds for Gaussian quadrature and weighted L^1-polynomial approximation, *SIAM J. Numer. Anal.* **21** (1984), 400–412.
12. Z. Ditzian, Equivalence of rate of approximation and smoothness, *J. Approx. Theory* **62** (1990), 316–339.
13. Z. Ditzian and V. Totik, "Moduli of Smoothness," Springer Ser. Comput. Math., Vol. 9, Springer-Verlag, New York, 1987.
14. Z. Ditzian and X. Zhou, Kantorovich–Bernstein polynomials, *Constr. Approx.* **6** (1990), 421–435.
15. D. Jackson, "Über die Genauigkeit der Annäherung stetiger Funktionen durch ganze rationale Funktionen gegebenen Grades und trigonometrische Summen gegebener Ordnung," Preisschrift und Dissertation, Univ. Göttingen, 1911.
16. S. Jansche, "Weak-Type-Ungleichungen in Approximationsräumen," Shaker Verlag, Aachen, 1996.
17. S. Jansche, $\mathcal{O}$-regularly varying functions in approximation theory, *J. Inequalities Appl.* **1** (1997), 253–274.

18. S. Jansche and R. L. Stens, Best weighted polynomial approximation on the real line; a functional-analytic approach, *J. Comput. Appl. Math.* **40** (1992), 199–213.
19. B. A. Khalilova, On some estimates for polynomials, *Izv. Akad. Nauk Azerbaĭdzhan. SSR Ser. Fiz.-Tekhn. Mat. Nauk* **2** (1974), 46–55. [In Russian]
20. H. Mevissen and R. J. Nessel, On the best approximation of Riemann integrable functions by trigonometric polynomials, *Math. Balkanica (N.S.)* **4** (1990), 289–299.
21. H. Mevissen and R. J. Nessel, Approximation of Riemann integrable functions by trigonometric convolution processes, *Funct. Approx. Comment. Math.* **20** (1992), 79–97.
22. W. Matuzewska and W. Orlicz, On some classes of functions with regard to their orders of growth, *Studia Math.* **26** (1965), 11–24.
23. P. G. Nevai, Bernstein's inequality in L^p for $0<p<1$, *J. Approx. Theory* **27** (1979), 239–243.
24. S. M. Nikolskiĭ, Approximation of functions in the mean by trigonometrical polynomials, *Izv. Akad. Nauk SSSR Ser. Mat.* **10** (1946), 207–256. [In Russian, English summary]
25. S. M. Nikolskiĭ, On the best approximation of functions satisfying Lipschitz's conditions by polynomials, *Izv. Akad. Nauk SSSR Ser. Mat.* **10** (1946), 295–322. [In Russian, English summary]
26. M. K. Potapov, Some inequalities for polynomials and their derivatives, *Vestnik Moskov. Univ. Ser. I Mat. Mekh.* **2** (1960), 10–20. [In Russian, English summary]
27. E. Seneta, "Functions of Regular Variation," Lecture Notes in Math., Vol. 506, Springer-Verlag, New York, 1976.
28. S. B. Stečkin, On the order of best approximations of continuous functions, *Izv. Akad. Nauk SSSR Ser. Mat.* **15** (1951), 219–242. [In Russian]
29. G. Sunouchi, Derivatives of a polynomial of best approximation, *Jahresber. Deutsch. Math.-Verein.* **70** (1968), 165–166.
30. G. Sunouchi, Derivatives of a trigonometric polynomial of best approximation, *in* "Abstract Spaces and Approximation" (P. L. Butzer and B. Sz. Nagy, Eds.), Proc. Internat. Conf., Math. Res. Inst. Oberwolfach, 1968, Internat. Ser. Numer. Math., Vol. 10, pp. 233–241, Birkhäuser, Basel, 1969.
31. E. van Wickeren, Weak-type inequalities for Kantorovitch polynomials and related operators, *Nederl. Akad. Wetensch. Indag. Math.* **49** (1987), 111–120.
32. M. Zamansky, Classes de saturation de certains procédés d'approximation des séries de Fourier des fonctions continues et applications à quelques problèmes d'approximation, *Ann. Sci. École Norm. Sup. (3)* **66** (1949), 19–93.
33. A. Zygmund, Smooth functions, *Duke Math. J.* **12** (1945), 47–76.