JOURNAL OF APPROXIMATION THEORY 62, 316–339 (1990)

# Equivalence of Rate of Approximation and Smoothness

Z. DITZIAN*

*Department of Mathematics, University of Alberta, Edmonton T6G 2G1, Canada*

*Communicated by V. Totik*

Received July 26, 1988; revised December 1, 1988

## 1. INTRODUCTION

To introduce the problem dealt with in this paper, a description of the situation in the special but very important approximation process, namely, the best $n$th degree trigonometric approximation in $C(T)$, will be given. We denote

$$E_n^*(f)_X \equiv \inf_{T_n \in \mathscr{T}_n} \|f - T_n\|_X, \tag{1.1}$$

where $X$ is a space of functions (or distributions) on $T$ and $\mathscr{T}_n$ is the space of trigonometric polynomials of degree $n$. The Jackson inequality

$$E_n^*(f)_X \leqslant C\omega^r(f, n^{-1})_X, \tag{1.2}$$

where

$$\begin{aligned} \omega^r(f, t)_X &\equiv \sup_{0<h\leqslant t} \|\Delta_h^r f\|_X, \\ \Delta_h^r f(x) &= \Delta_h^{r-1}\left(f\left(x+\frac{h}{2}\right) - f\left(x-\frac{h}{2}\right)\right) \end{aligned} \tag{1.3}$$

($\Delta_h^0 f(x) = f(x)$) and the Bernstein inequality

$$\|T_n^{(r)}\|_X \leqslant n^r \|T\|_X \qquad \text{for} \quad T_n \in \mathscr{T}_n \tag{1.4}$$

are valid for many spaces of functions on $T$. (In particular (1.2) and (1.4) are satisfied for spaces of functions on $T$ for which the translation $T(t)$ is continuous in $t$ and $\|T(t)\| = 1$.) These inequalities imply for $0 < \alpha < r$,

$$\omega^r(f, t)_X = O(t^\alpha) \Leftrightarrow E_n^*(f)_X = O(n^{-\alpha}). \tag{1.5}$$

* Supported by NSERC of Canada Grant A4816.



0021-9045/90 $3.00

For the sake of brevity, we will write $\omega^r(f, t)$, $E_n^*(f)$ and $\|f\|$ instead of $\omega^r(f, t)_X$, $E_n^*(f)_X$, and $\|f\|_X$ when $X$ will be understood from the context.

At first, one might think that in the situation above, i.e., when $\omega^r(f, t) = O(t^\alpha)$, we will have $\omega^r(f, 1/n) \sim E_n^*(f)$, which implies the exact converse of (1.2). Unfortunately, this turns out to be false, as can be shown by modifying an example given by Boman [3] or using the classical proof (by constructive contradiction) of the Banach–Steinhaus theorem as done by Dickmeis, Nessel, and Wickeren in many articles (see, for example, [4, 9]). A question of this type was posed by Stechkin in a conference on constructive function theory, 1977, in Blagoevgrad, Bulgaria. Stechkin asked if for $f \in C(T)$ which does not belong $C^\alpha(T)$ there exists an $r$ such that $\omega^r(f, 1/n) \leqslant CE_n^*(f)$. That question was answered negatively by Boman [3]. It seems natural to look at the same question in a positive manner and ask if there is a simple condition on the behaviour of $\omega^r(f, t)$ or of $E_n^*(f)$ that will imply for some $r$,

$$\omega^r(f, t) \sim E_n^*(f). \tag{1.6}$$

In fact this type of question was answered for trigonometric polynomials on $C(T)$ and probably other spaces by several Russian mathematicians, notably N. K. Bari and S. B. Stechkin [2]. In [2] (see Lemma 7, p. 513) it was shown that for a nondecreasing continuous function in $[0, \pi]$, $\psi_r$ satisfying $\psi_r(0) = 0$ and $\psi_r(t) \neq 0$ for $t \neq 0$ satisfies the condition

$$\delta^r \int_\delta^c \frac{\psi_r(u)}{u^{r+1}}\, du \sim \psi_r(\delta), \tag{1.7}$$

one has

$$E_n^*(f) \sim \psi(1/n) \Leftrightarrow \omega^r(f, t) \sim \psi(t), \tag{1.8}$$

and therefore, (1.6) is satisfied.

A result of this type for general non-linear processes of best approximation that possess Jackson- and Bernstein-type inequalities will be shown. Moreover, analogous results will be achieved relating the best weighted algebraic polynomial approximation on $[-1, 1]$ or $R$ to moduli that do not satisfy the Jackson inequality but a weaker inequality that should perhaps be called the weak Jackson inequality (see also [8]).

For linear approximation processes that satisfy Jackson- and Bernstein-type inequalities, analogous results are also achieved in Section 3.

It will be the applications that will guide the investigation and conditions in this paper, and many applications will be given.

## 2. Best Approximation and Smoothness

Results of the type described in Section 1, i.e., (1.8) for $\psi$ satisfying (1.7), are not particular to best $n$th degree trigonometric approximation. A general framework under which they are valid will be described in this section. Let $X$ and $Y$ be Banach spaces with $Y \subset X$. The $K$-functional of the pair $(X, Y)$ is given by

$$K(f, t) = \inf_{g \in Y} (\|f - g\|_X + t\Phi(g)), \tag{2.1}$$

where $\Phi$ is a seminorm for which

$$Y = \{f \in X; \Phi(f) < \infty\}.$$

For $\{X_n\}$ a sequence of subspace of $X$ satisfying $X_n \subset X_{n+1}$, we write

$$E_n(f) = E_n(f)_X = \inf_{\varphi \in X_n} \|f - \varphi\|_X. \tag{2.2}$$

A Jackson-type inequality is given by

$$E_n(f) \leqslant CK(f, \sigma_n), \qquad f \in X. \tag{2.3}$$

A Bernstein-type inequality is given by

$$\Phi(\varphi) \leqslant C_1 \sigma_n^{-1} \|\varphi\|_X, \qquad \varphi \in X_n \subset Y. \tag{2.4}$$

In fact, while in the statement of the Jackson and the Bernstein inequality, there is no reason to use the same $X$, $Y$, or $\sigma_n$, we (and almost everybody else) are interested in a matching pair, i.e., Jackson- and Bernstein-type inequalities with the same $X$, $Y$, $X_n$, and $\sigma_n$.

We should also remark that if we denote the best approximant of $X_n$ to $f$ (or one of them) by $A_n f$, that is,

$$\|A_n f - f\|_X \equiv \inf_{\varphi \in X_n} \|\varphi - f\|_X, \tag{2.5}$$

we have

$$\|A_n f\| \leqslant \|f\|$$

and

$$\|A_n(f + g) - (f + g)\| \leqslant \|A_n f - f\| + \|A_n g - g\|.$$

Therefore, the inequality

$$E_n(g) \leqslant C_2 \sigma_n \Phi(g) \qquad \text{for} \quad g \in Y \tag{2.6}$$

can replace the Jackson inequality (2.3). This inequality is sometimes called the Favard inequality and sometimes the Jackson inequality, and therefore, we may call it the Jackson–Favard inequality.

To prove the theorems, it is necessary that the sequence $\sigma_n$ not tend to zero at a faster than geometric rate, i.e.,

$$\sigma_{n+1} \geqslant c\sigma_n \qquad \text{for some} \quad c>0 \qquad \text{and} \qquad n=1,2,... \tag{2.7}$$

(As $\sigma_n = o(1)$, $0<c<1$.) The sequence $\sigma_n = n^{-\gamma}$ is prevalent in applications. (Most times $\gamma$ is an integer but in Section 8 we have a case for which $\gamma$ is not an integer.) In fact in most works the space $X_n$ first worked with is of about $2^n$ dimensions and the sequence is $\sigma_n = 2^{-n\lambda}$, and only later is an adjustment made to $n$ (or $2n+1$) dimensional $X_n$. (Note that here we did not specify the dimension of $X_n$.)

We follow N. Bari and S. B. Stechkin [2] and define a class of functions $\psi$ which satisfy the steadiness condition which we call $S$ ($S$ for steady).

Definition 2.1. A continuous nondecreasing function $\psi(t)$ is of class $S$ if $0=\psi(0)<\psi(t)$ for $t>0$ and the condition

$$\delta \int_\delta^1 \frac{\psi(t)}{t^2}\,dt \sim \psi(\delta) \tag{2.8}$$

is satisfied.

In fact the steadiness of $\psi$ is seen much more clearly from the following equivalent condition (see [2]).

Theorem A. *For a nondecreasing continuous function $\psi(t)$, $0=\psi(0)<\psi(t)$ for $t>0$, the assumption that there exist $A_0$, $A_0>0$ such that*

$$\frac{\psi(A\delta)}{\psi(\delta)} \leqslant \frac{1}{2}A \qquad \textit{for} \quad A \geqslant A_0 \qquad \textit{and} \qquad A\delta \leqslant \frac{1}{2} \tag{2.9}$$

*is equivalent to* (2.8).

While in [2] the statement is somewhat different than (2.9), it actually implies the same facts and the equivalence between (2.8) and (2.9) is given in the proof there before a less convenient (in my opinion) equivalent condition is stated.

It can be noted that the assumption

$$\delta \int_\delta^1 \frac{\psi(t)}{t^2}\,dt = O(\psi(\delta))$$

already implies (2.8). Furthermore, one can note that the continuous nondecreasing function $\psi_r(t)$ for which $0=\psi_r(0)<\psi_r(t)$ for $t>0$ satisfies

$$\delta^r \int_\delta^1 \frac{\psi_r(t)}{t^{r+1}}\,dt \sim \psi_r(\delta)$$

if and only if $\psi_r(t^{1/r})\equiv\psi(t)$ with $\psi(t)$ satisfying $\psi\in S$.

THEOREM 2.2. *Suppose the Jackson inequality* (2.3) *and the Bernstein inequality* (2.4) *are satisfied for the Banach spaces* $X$, $Y$, *and* $X_n$ *satisfying* $X_n\subset X_{n+1}\subset Y\subset X$, *and suppose the sequence* $\sigma_n$ *satisfies* (2.7) (*i.e.*, $\sigma_{n+1}\geqslant c\sigma_n$). *Then for* $\psi\in S$, *the condition*

$$K(f,t)\sim\psi(t) \qquad \textit{for} \quad t\in[0,1]$$

*and the condition*

$$E_n(f)\sim\psi(\sigma_n) \qquad \textit{for} \quad n\geqslant n_0$$

*are equivalent, and either one of them implies*

$$E_n(f)\sim K(f,\sigma_n) \qquad \textit{for} \quad n\geqslant n_0, \tag{2.10}$$

*where* $n_0$ *is independent of* $f$.

As $K(f,t)$ is an increasing function and $E_n(f)$ a decreasing sequence, we also have the following immediate corollary of Theorem 2.2.

COROLLARY 2.3. (A) *Under the conditions of Theorem* 2.2,

$$K(f,2^j t)\leqslant CK(f,t), \qquad C<2^j$$

*for some* $j$ *implies* (2.10).

(B) *Under the conditions of Theorem* 2.2 *with* $\sigma_n=n^{-\lambda}$,

$$E_n(f)\leqslant CE_{2^j n}(f), \qquad C<2^{j\lambda}$$

*for some* $j$ *implies* (2.10).

*Proof of Corollary* 2.3. To prove (A), we set $K(f,t)=\psi(t)$, and the assumption on $K(f,t)$ implies $\psi(t)\in S$. To prove (B), we set $E_n(f)=\psi(\sigma_n)$ and define $\psi$ linearly elsewhere. The assumption on $E_n(f)$ now implies $\psi(t)\in S$. ∎

*Proof of Theorem* 2.2. Using (2.7), we choose a subsequence of $n$, $n_\nu$, such that $n_0=1$ and

$$c^{\nu-1}\sigma_1<\sigma_{n_\nu}\leqslant c^\nu\sigma_1 \qquad \text{for} \quad \nu=1,2,\ldots.$$

(For $\sigma_n = n^{-\lambda}$, $n_\nu = 2^\nu$, and for $\sigma_n = 2^{-\lambda n}$, $n_\nu = \nu$.) With no loss of generality we may assume $\sigma_1 = 1$, and therefore, $c^{\nu-1} < \sigma_{n_\nu} \leqslant c^\nu$ for $\nu = 1, 2, \ldots$. We now write

$$A_{n_k} f = \sum_{\nu=1}^{k} (A_{n_\nu} f - A_{n_{\nu-1}} f) + A_1 f,$$

where

$$\|A_m f - f\| = \inf_{\varphi \in X_m} \|\varphi - f\|.$$

We apply the Bernstein inequality to the above and obtain

$$\begin{aligned} K(f, t) &= E_{n_k}(f) + t\Phi(A_{n_k} f) \\ &\leqslant E_{n_k}(f) + Ct \sum_{\nu=1}^{k} c^{-\nu} E_{n_{\nu-1}}(f) + Ct\|f\|, \end{aligned} \tag{2.11}$$

where $C = C_1 2c^{-1}$. Frequently the last term in (2.11) does not appear in applications as often we have $\Phi(\varphi) = 0$ for $\varphi \in X_1$.

We now prove that $K(f, t) \sim \psi(t)$ implies $E_n(f) \sim \psi(\sigma_n)$. Using Jackson's inequality and

$$C_3^{-1}\psi(t) \leqslant K(f, t) \leqslant C_3\psi(t),$$

we have only to show $E_n(f) \geqslant A\psi(\sigma_n)$. We use (2.11) to write

$$\begin{aligned} C_3^{-1}\psi(t) \leqslant K(f, t) &\leqslant E_{n_k}(f) + Ct \sum_{\nu=1}^{k} c^{-\nu} K(f, \sigma_{n_{\nu-1}}) + Ct\|f\| \\ &\leqslant E_{n_k}(f) + C(1)\, t \sum_{\nu=2}^{k} c^{-\nu}\psi(c^{\nu-2}) + C(1)\, t\|f\|. \end{aligned}$$

We now use $\psi \in S$ (Definition 2.1), and therefore,

$$\sum_{\nu=0}^{k} c^{-\nu}\psi(c^\nu) \leqslant C_4 \psi(c^k)\, c^{-k}.$$

This now implies (for $k \geqslant k_0$)

$$C_3^{-1}\psi(t) \leqslant E_{n_k}(f) + C(2)\, tc^{-k+2}\psi(c^{k-2}) + C(1)\, t\|f\|.$$

We will choose $t = c^m$ with $m = m(k) > k \geqslant k_0$ for which

(I) $$C(2)\, c^m c^{-k+2}\psi(c^{k-2}) < \tfrac{1}{4} C_3^{-1}\psi(c^m),$$

(II) $$C(1)\, c^m \|f\| \leqslant \tfrac{1}{4} C_3^{-1}\psi(c^m)$$

and at the same time

$$\text{(III)} \qquad |m(k)-k|<N,$$

where $N$ does not depend on $k$.

If a choice $m=m(k)$ satisfying (I), (II), and (III) is possible, we have, using the inequality

$$K(f, at)\leqslant aK(f, t)$$

which is valid for $0<a\leqslant 1$,

$$\begin{aligned} E_{n_k}(f) &\geqslant C_3^{-1}\psi(c^m)\geqslant C_3^{-2}K(f, c^m) \\ &\geqslant C_3^{-2}c^{m-k+1}K(f, c^{k-1})\geqslant C_3^{-3}c^{m-k+1}\psi(c^{k-1}) \\ &\geqslant C_3^{-3}c^{m-k+1}\psi(\sigma_{n_k}) \end{aligned}$$

for $k\geqslant k_0$. We now use for $n_k<n<n_{k+1}$, $k\geqslant k_0$,

$$E_n(f)\geqslant E_{n_{k+1}}(f)\geqslant C\psi(\sigma_{n_k}),$$

and using the same consideration as above,

$$\psi(\sigma_{n_k})\geqslant C\psi(\sigma_n) \qquad \text{for} \quad n_k<n<n_{k+1},$$

which yields

$$E_n(f)\geqslant M\psi(\sigma_n), \qquad n>n_{k_0}.$$

Note that $n_0$ is independent of $f$.

Therefore, to complete the proof that $E_n(f)\sim\psi(\sigma_n)$ follows from $K(f, t)\sim\psi(t)$, we need to show that we can choose $m\equiv m(k)$ that satisfy (I), (II), and (III).

To choose $m-k$ we first choose an integer $l$ such that

$$2^{-l}\leqslant\tfrac{1}{4}C_3^{-1}C(2)^{-1},$$

where $C_3$ and $C(2)$ were given in (I). We now choose $m-k$ so that for $A_0$ given in (2.9)

$$A_0^l\leqslant c^{-m+k-2}<A_0^{l+1},$$

and therefore, (III) is satisfied, and for $k\geqslant k_0$, (I) follows from (2.9) with $A^l=c^{-m+k-2}$ as

$$\frac{\psi(c^{k-2})}{\psi(c^m)}\leqslant\left(\frac{1}{2}\right)^l c^{-m+k-2}\leqslant(c^{m+2-k}\cdot C_3\cdot C(2)\,4)^{-1}.$$

The inequality (2.9) which is equivalent to $\psi \in S$ now implies $c^k = o(\psi(c^k))$, $k \to \infty$, and therefore, for $m \geqslant k \geqslant k_0$, (II) is satisfied. We have now proved that a choice of $m(k)$ is possible for $k \geqslant k_0$.

We now prove that (2.10) follows from $E_n(f) \sim \psi(\sigma_n)$ $(n \geqslant n_0)$. We write

$$M^{-1}\psi(\sigma_n) \leqslant E_n(f) \leqslant M\psi(\sigma_n),$$

and following (2.11),

$$\begin{aligned} K(f, t) &\leqslant E_{n_k}(f) + Ct \sum_{\nu=1}^{k} c^{-\nu} E_{n_{\nu-1}}(f) + Ct\|f\| \\ &\leqslant M\psi(\sigma_{n_k}) + C(1)\, t \sum_{\nu=1}^{k} c^{-\nu+1}\psi(\sigma_{n_{\nu-1}}) + Ct\|f\|. \end{aligned}$$

We set $t = c^{k-2}$ and recall $\psi \in S$ to obtain

$$\begin{aligned} K(f, c^{k-2}) &\leqslant M\psi(c^{k-2}) + C(2)\, \psi(c^{k-2}) + Cc^{k-2}\|f\| \\ &= C(3)(\psi(c^{k-2}) + c^{k-2}\|f\|). \end{aligned}$$

Using the equivalent form to $\psi \in S$ in Theorem A,

$$c^{k-2} = o(\psi(c^{k-2})), \qquad k \to \infty,$$

and therefore,

$$c^{k-2} \leqslant M_1 \psi(c^{k-2}).$$

This now implies

$$K(f, c^k) \leqslant C(4)\, \psi(c^k),$$

and therefore,

$$K(f, t) \leqslant C(5)\, \psi(t),$$

which completes the proof of the theorem. ∎

*Remark* 2.4. It is the existence of an inequality of weak type like (2.11) that together with (2.3) is sufficient for the proof of Theorem 2.1 or Theorem 2.2. A Bernstein-type inequality is the crucial tool to achieve (2.11) but sometimes an inequality such as (2.11) can be proved without explicit use of the corresponding Bernstein inequality. This type of situation is given in [7, Chap. 12] (see also Section 9 below).

## 3. Rate of Approximation of Linear Processes and Smoothness

In this section an analogous result to that of the last section will be achieved for linear processes of approximation. While there is a similarity in the problem and some of the ideas, we lose monotonicity and the Bernstein-type inequality is of somewhat different character. We first note that the rate of convergence for a sequence of linear operators (even standard ones) is not necessarily monotone. For a $K$-functional of the pair of spaces $(X, Y)$ given in (2.1), the Jackson inequality is given by

$$\|A_n f - f\|_X \leqslant CK(f, \sigma_n). \tag{3.1}$$

The Jackson inequality for linear processes usually follows from

$$\|A_n f\|_X \leqslant M\|f\|_X, \tag{3.2}$$

which means that the sequence of operators $A_n$ is uniformly bounded, and

$$\|A_n g - g\|_X \leqslant C\sigma_n \Phi(g) \qquad \text{for} \quad g \in Y. \tag{3.3}$$

(Recall that $\Phi$ is a seminorm which defined $Y$ and was used in (2.1).) The inequality (3.3) is sometimes called the Jackson inequality as well, and sometimes the Favard inequality. The Bernstein-type inequality is given by

$$\Phi(A_n f) \leqslant C_1 \sigma_n^{-1} \|f\|_X \tag{3.4}$$

which assumes implicitly that $A_n f \in Y$. For linear approximation processes

$$\Phi(A_n g) \leqslant C_2 \Phi(g) \qquad \text{for} \quad g \in Y \tag{3.5}$$

is also necessary. The inequality (3.5) means that $A_n$ is a uniformly bounded sequence of operators in $Y$ as well. The Bernstein-type inequality here looks somewhat different from (2.4) but the similarity of the results and applications (see later sections) will, it is hoped, convince the reader (who is not already convinced) that the identical name for these somewhat different inequalities is justified. In fact, the inequalities (3.4) and (2.4) are both commonly referred to in the literature as Bernstein-type inequalities.

THEOREM 3.1. *Suppose $A_n$ is a sequence of linear operators on $X$ satisfying $A_n f \in Y$ and the inequalities* (3.1), (3.2), (3.4), *and* (3.5) *with respect to the pair of spaces $X$ and $Y$ and a sequence $\sigma_n \searrow 0$. Then the condition*

$$K(f, t) \sim \psi(t) \qquad \textit{for} \quad \psi \in S$$

*for $t \leqslant t_0$ implies*

$$\|A_n f - f\| \sim K(f, \sigma_n). \tag{3.6}$$

*Remark.* For $K(f, t)$ satisfying

$$K(f, 2^j t) \leqslant CK(f, t), \qquad \text{for some} \quad C < 2^j \tag{3.7}$$

we have $K(f, t) \in S$, and therefore (3.6).

*Proof of Theorem* 3.1. Using (3.1), we have only to show that

$$\|A_n f - f\| \geqslant MK(f, \sigma_n).$$

Combining

$$K(f, t) \leqslant \|A_n f - f\| + t\Phi(A_n f) \tag{3.8}$$

with (3.4) and (3.5), we obtain the common form

$$K(f, t) \leqslant \|A_n f - f\| + Lt\sigma_n^{-1} K(f, \sigma_n). \tag{3.9}$$

We choose $t = \delta\sigma_n$ with some $0 < \delta < 1$ for which

$$Lt\sigma_n^{-1} K(f, \sigma_n) < \tfrac{1}{2} K(f, t)$$

which is possible for $n \geqslant n_0$ as $K(f, t) \sim \psi(t)$ and $\psi(t) \in S$, and therefore, $\psi(t)$ and $K(f, t)$ satisfy (2.9).

We now have

$$\|A_n f - f\| \geqslant \frac{1}{2} K(f, t) \geqslant \frac{\delta}{2} K(f, \sigma_n)$$

and this completes the proof when we observe that $\delta$ does not depend on $n \geqslant n_0$. ∎

The condition $K(f, t) \sim \psi(t)$ for $\psi \in S$ can be replaced by a condition on $\|A_n f - f\|$ as is shown in the following theorem.

THEOREM 3.2. *Suppose $A_n$ satisfy* (3.2), (3.3), (3.4), *and* (3.5), $\psi \in S$ (*see Definition* 2.1) *and $\sigma_{n+1} > c\sigma_n$ for some $0 < c < 1$. Then*

$$\|A_n f - f\| \sim \psi(\sigma_n) \tag{3.10}$$

*implies* (3.6).

*Proof.* We have to show only that

$$K(f, \sigma_n) \leqslant M\psi(\sigma_n)$$

and in fact it is sufficient to prove it for $n \geqslant N$. Using (3.4) and (3.5), we write

$$\begin{aligned} K(f, \sigma_n) &\leqslant \|f - A_k f\| + \sigma_n \Phi(\sigma_k f) \\ &\leqslant \|f - A_k f\| + L\sigma_n(\sigma_k^{-1} \|f - g\| + \Phi(g)) \\ &\leqslant \|f - A_k f\| + L\sigma_n \sigma_k^{-1} K(f, \sigma_k). \end{aligned}$$

We now choose a subsequence of $n$, $n-1, ..., 1$, $n_1 > n_2 > \cdots > n_{i_0}$ such that

$$c^{-im}\sigma_n \leqslant \sigma_{n_i} < \sigma_n c^{-(i+1)m},$$

where $c^{-m} \geqslant A_0^l$ and $2^l > L$ with $c$ of (2.7) (or the statement of the present Theorem) $A_0$ of (2.9) and $L$ given above. Obviously $\{n_i\}$ is a finite sequence, and moreover, we choose $i_0$ so that $\sigma_n c^{-(i_0+1)m} \leqslant \frac{1}{2} < \sigma_n c^{-(i_0+2)m}$ unless we exhausted the sequence $n, ..., 1$ before that.

Using

$$\|f - A_k f\| \leqslant M_1 \psi(\sigma_k)$$

and

$$\Phi(A_{n_{i_0}} f) \leqslant C(1)\|f\| \leqslant M\psi(\sigma_{i_0}),$$

we write

$$\begin{aligned} K(f, \sigma_n) &\leqslant \|f - A_{n_1} f\| + \sum_{i=1}^{i_0} L^i \sigma_n \sigma_{n_i}^{-1} \|f - A_{n_i} f\| + L^{i_0} \sigma_n \sigma_{n_0}^{-1} \Phi(A_{n_0} f) \\ &\leqslant M_1 \psi(\sigma_{n_1}) + M_1 \sum_{i=1}^{i_0} L^i \sigma_n \sigma_{n_i}^{-1} \psi(\sigma_{n_i}) + M_1 L^{i_0} \sigma_n \sigma_{n_{i_0}}^{-1} \psi(\sigma_{n_{i_0}}) \\ &\leqslant M_2 \psi(\sigma_n) + M_2 \sum_{i=0}^{i_0} (L/2^l)^i \psi(\sigma_n) + M_2 \psi(\sigma_n) \\ &\leqslant M_3 \psi(\sigma_n). \quad \blacksquare \end{aligned}$$

## 4. Weak Jackson Inequality

In Section 2, we assumed a Jackson-type inequality with respect to the $K$-functional. In many cases of theorems about best approximation, such an inequality is satisfied and the $K$-functional is equivalent to a satisfactory measure of smoothness. However, recently while investigating weighted algebraic polynomial approximation on $[-1, 1]$ and on $R$ (with different weights of course), two natural (and indispensable for some weights) measures of smoothness emerged (see [7]) which were not equivalent to the corresponding $K$-functional and did not satisfy the Jackson inequality. In [7], such measures of smoothness were named main-part moduli of smoothness and many of their properties were proved. We denote main-part moduli of smoothness by $\Omega^r(f, t)$. The Jackson inequality was replaced by

$$E_n(f) \leqslant C_1 \int_0^{\sigma_n} (\Omega^r(f, t)/t)\, dt \tag{4.1}$$

which we call the weak Jackson inequality. We further use a form of weak-type estimate given by

$$\Omega^r(f, t) \leqslant C_2 \left( E_{n_k}(f) + t^r \sum_{\nu=0}^{k} \sigma_{n_\nu}^{-r} E_{n_\nu}(f) \right) \tag{4.2}$$

for all $t$ and a subsequence of $n\{n_\nu\}$ for which $\sigma_{n_\nu} \sim \beta^\nu$ for some $0 < \beta < 1$. Recall that (2.11), which corresponds to (4.2), was instrumental in the proof of Theorem 2.2. In the applications, both $E_n(f)$ and $\Omega^r(f, t)$ will have other subscripts to indicate the space $X$ in which $f$ lives, the weight in question, and perhaps the sequence of subspaces $X_n$. Here, we derive from (4.1) and (4.2) a somewhat weaker result than that derived from (2.3) and (2.11) in Theorem 2.2 using the same methods and steps of the proof of that theorem.

DEFINITION 4.1. The increasing function $\psi(t)$ belongs to class $S^*$ if $\psi \in S$, that is,

$$\delta \int_\delta^1 \frac{\psi(u)}{u^2}\, du \sim \psi(\delta) \tag{4.3}$$

and in addition

$$\int_0^t (\psi(\tau)/\tau)\, d\tau \sim \psi(t) \tag{4.4}$$

for $0 < t \leqslant t_0$.

THEOREM 4.2. *Suppose for the space $X$, a sequence of subspaces $X_n$, $X_n \subset X_{n+1}$, and a sequence $\sigma_{n+1} > c\sigma_n$, a relation between $\Omega^r(f, t)$ and $E_n(f)$ is given by* (4.1) *and* (4.2), *the latter for some subsequence of $n$, $n_k$ for which $\sigma_{n_k} \sim \beta^k$. Then for $\psi(t) \in S^*$*

$$\Omega^r(f, t) \sim \psi(t^r) \tag{4.5}$$

*if and only if*

$$E_n(f) \sim \psi(\sigma_n^r). \tag{4.6}$$

*Remark* 4.3. We may restate Theorem 4.2 as follows: Under the assumptions of Theorem 4.2, either (4.5) or (4.6) implies

$$E_n(f) \sim \Omega^r(f, \sigma_n). \tag{4.7}$$

*Proof.* To prove that (4.5) implies (4.6) we first observe that (4.4) implies

$$\int_0^t (\psi(\tau^\alpha)/\tau)\, d\tau \sim \psi(t^\alpha) \tag{4.8}$$

for all positive $\alpha$. We then use (4.1), (4.2), (4.5), and (4.8) with $\alpha = r$ to obtain

$$\psi(t^r) \leqslant A\left(E_{n_k}(f) + t^r \sum_{\nu=0}^{k} \sigma_{n_\nu}{}^{-r}\, \psi(\sigma_{n_\nu}^r)\right).$$

We can now follow the technique of the corresponding part of Theorem 2.2 to obtain

$$E_{n_k}(f) \geqslant M\psi(\sigma_{n_k}^r).$$

As we already have $E_{n_k}(f) \leqslant M_1 \psi(\sigma_{n_k}^r)$, this implies

$$E_{n_k}(f) \sim \psi(\sigma_{n_k}^r).$$

Hence, using the conditions (4.3) and monotonicity of $E_n(f)$ and $\psi(t)$, we have

$$E_n(f) \sim \psi(\sigma_n^r).$$

The proof of the implication

$$E_n(f) \sim \psi(\sigma_n^r) \Rightarrow \Omega^r(f, \sigma_n) \leqslant ME_n(f) \leqslant M^*\psi(\sigma_n^r)$$

is similar to that used in Theorem 2.2 and will be omitted. Using (4.1) and monotonicity of $\Omega^r(f, t)$, we have

$$E_n(f) \leqslant M_3 \int_0^{\sigma_n} (\Omega^r(f, t)/t)\, dt \leqslant M_4 \Omega^r(f, \sigma_n)$$

and therefore,

$$\Omega^r(f, \sigma_n) \sim E_n(f). \quad \blacksquare$$

## 5. Trigonometric Polynomial Approximation

In this section, we deal with trigonometric polynomial approximation in a Banach space $X$ of functions or distributions on $T$ (the "circle" $[-\pi, \pi]$). The translation or shift operator $S_h$ on $X$ is given by

$$S_h f(x) = f(x+h) \qquad \text{if } f \text{ is a function}$$

or by

$$\langle S_h f, g\rangle = \langle f, S_{-h} g\rangle,$$

where $g \in \mathscr{S}$ if $f \in \mathscr{S}'$. (Recall that $\mathscr{S}'$ is the space of tempered distribution dual to the space of test functions $\mathscr{S}$.) We further assume that $S_h$ is an isometry and that either $S_h$ is weakly continuous, that is,

$$\langle S_h f - f, g\rangle \to 0, \qquad \text{as} \quad h \to 0, \text{ for all } g \in X^* \text{ (dual to } X), \tag{5.1}$$

or that $S_h$ is weakly* continuous which means that $X = B^*$ and

$$\langle g, S_h f - f\rangle \to 0 \qquad \text{as} \quad h \to 0, \text{ for all } g \in B. \tag{5.2}$$

(Of course, $S_h$ is strongly continuous implies that $S_h$ is weakly continuous.) Under these assumptions, one has the Jackson inequality

$$E_n^*(f)_X \leqslant C\omega^r(f, n^{-1})_X, \tag{5.3}$$

where $E_n^*(f)_X$ and $\omega^r(f, t)$ are given by (1.2) and (1.3) and $C$ is independent of $n, f$ and $X$ (see [5, 6], for example).

Moreover, for the above situation $T_n \in X$ implies $T_n' \in X$ and hence, $T_n^{(r)} \in X$ and the Bernstein inequality,

$$\|T_n^{(r)}\|_X \leqslant n^r \|T_n\|_X \qquad \text{for all} \quad T_n \in \mathscr{T}_n, \tag{5.4}$$

is satisfied (see [1, pp. 140–144; 5; 6]).

The $K$-functional is given by

$$K_r(f, t^r)_X = \inf_{g \in Y} (\|f-g\|_X + t^r \|g^{(r)}\|_X), \tag{5.5}$$

where $Y$ is the collection of $g$ such that $g^{(r)}$, taken in the distributional sense, satisfies $g^{(r)} \in X$. Obviously, we can choose $g \in Y$ such that

$$\omega^r(f, t)_X \leqslant 2^r \|f-g\|_X + t^r \|g^{(r)}\| \leqslant 2^{r+1} K_r(f, t^r)_X \tag{5.6}$$

and therefore, the conditions of Theorem 3.1 are valid. Following the standard proof for $C(T)$ or $L_p(T)$, one also has

$$K_r(f, t^r)_X \leqslant C\omega^r(f, t)_X. \tag{5.7}$$

THEOREM 5.1. *Suppose $X$ is a space of functions or distributions on $T$ for which a shift is an isometry satisfying* (5.1) *or* (5.2) *and $\psi \in S$. Then the conditions*

$$\omega^r(f, t)_X \sim \psi(t^r) \tag{5.8}$$

*and*

$$E_n(f)_X \sim \psi(1/n^r) \tag{5.9}$$

*are equivalent, and either* (5.8) *or* (5.9) *implies*

$$E_n(f)_X \sim \omega^r(f, 1/n)_X.$$

*Proof.* This theorem is an immediate corollary of Theorem 2.1. To satisfy the conditions of Theorem 2.1, we recall (5.3) and (5.4), replace $t$ by $t^r$, and observe that

$$K_r(f, t^r)_X \sim \omega^r(f, t)_X$$

which follows from (5.6) and (5.7). ▮

## 6. ALGEBRAIC POLYNOMIAL APPROXIMATION ON $[-1, 1]$

In [7, Chap. 7], the rate of approximation of algebraic polynomials was discussed for $L_p[-1, 1]$, $1 \leqslant p \leqslant \infty$. The $K$-functional

$$K_{r,\varphi}(f, t^r)_p = \inf_{g^{(r-1)} \in A \cdot C_{loc}} (\|f - g\|_p + t^r \|\varphi^r g^{(r)}\|_p), \qquad \varphi(x) = \sqrt{1 - x^2} \tag{6.1}$$

was shown to be equivalent to a modulus of smoothness

$$\omega_\varphi^r(f, t)_p = \sup_{0 < h \leqslant t} \|\Delta_{h\varphi}^r f\|_p, \qquad \varphi(x) = \sqrt{1 - x^2} \tag{6.2}$$

(with the understanding that $\Delta_h^r f = 0$ if $(x - \eta r/2, x + \eta r/2) \not\subset [-1, 1]$). As the Jackson inequality

$$E_n(f)_p = \inf_{g \in \mathscr{P}_n} \|g - f\|_p \leqslant C\omega_\varphi^r(f, 1/n)_p \leqslant C_1 K_{r,\varphi}(f, n^{-r})_p \tag{6.3}$$

[3, p. 79) and the Bernstein inequality

$$\|\varphi^r P^{(r)}\|_p \leqslant Cn^r \|P\|_p, \qquad P \in \mathscr{P}_n, \ 1 \leqslant p \leqslant \infty \tag{6.4}$$

(see [7, p. 107] with $w = 1$) were already proved, we have the following theorem as a corollary.

THEOREM 6.1. *For $\psi \in S$, $\omega^r_\varphi(f, t)_p$ given by (6.2), and $E_n(f)_p$ given by (6.3), the conditions*

$$\omega^r_\varphi(f, t)_p \sim \psi(t^r)$$

*and*

$$E_n(f)_p \sim \psi(n^{-r})$$

*are equivalent.*

*Remark.* For some particular functions $\psi(t)$ such as $t^\alpha$ or $t^\alpha(\log t)^\beta$ it was proved in [7] that $\omega^r_\varphi(f, t)_p \sim \psi(t^r)$ implies $E_n(f)_p \sim \psi(n^{-r})$.

## 7. BEST WEIGHTED ALGEBRAIC POLYNOMIAL APPROXIMATION IN $L_p[-1, 1]$

The rate of best weighted algebraic polynomial approximation in $L_p$ given by

$$E_n(f)_{w,p} \equiv \inf_{P \in \mathscr{P}_n} \|w(f - P)\|_{L_p[-1,1]} \tag{7.1}$$

was investigated in [7, Chap. 8] for weights $w \in J^*_p$, where $J^*_p$ includes the Jacobi weights $w(x) = (1-x)^{\gamma_1}(1+x)^{\gamma_2}$ if $\gamma_i > -1/p$. For estimating $E_n(f)_{w,p}$, the main-part moduli $\Omega^r_\varphi(f, t)_{w,p}$ given by (for $\varphi(x) = \sqrt{1-x^2}$)

$$\Omega^r_\varphi(f, t)_{w,p} = \sup_{0 < h \leqslant t} \|w\Delta^r_{h\varphi} f\|_{L_p[-1+2r^2h^2,\, 1-2r^2h^2]} \tag{7.2}$$

were indispensable. The weak Jackson inequality

$$E_n(f)_{w,p} \leqslant C \int_0^{1/n} (\Omega^r_\varphi(f, t)/t)\, dt \tag{7.3}$$

was proved in [7, p. 94]. Therefore, we have the following result.

THEOREM 7.1. *Suppose $E_n(f)_{w,p}$ and $\Omega^r_\varphi(f, t)_{w,p}$ are given by (7.1) and (7.2), respectively, $w(x) = (1+x)^{\gamma_1}(1-x)^{\gamma_2}$ with $\gamma_i > -1/p$, $\varphi(x) = \sqrt{1-x^2}$, and $\psi \in S^*$ (see Definition 4.1). Then*

$$\Omega^r_\varphi(f, t)_{w,p} \sim \psi(t^r)$$

*and*

$$E_n(f)_{w,p} \sim \psi(n^{-r})$$

*are equivalent.*

*Remark* 7.2. Theorem 7.1 is valid for a somewhat more general class of weights called $J_p^*$ in [7].

*Proof.* We recall that for $w(x) = (1+x)^{\gamma_1}(1-x)^{\gamma_2}$, $\gamma_i > -1/p$, $1 \leqslant p \leqslant \infty$, (and actually for the wider class $J_p^*$ which includes for instance $w(x) = (1+x)^{\gamma_1}(1-x)^{\gamma_2}(|\log(1+x)|^{\beta_1}|\log(1-x)|^{\beta_2}, \gamma_i > -1/p)$ one has

$$\|w\varphi^r P^{(r)}\|_{L_p[-1,1]} \leqslant Cn^r \|wP\|_{L_p[-1,1]}, \qquad P \in \mathscr{P}_n \tag{7.4}$$

(see [7, p. 107]). We can now follow the proof of Theorem 8.2.1 of [7, p. 96] to obtain

$$\begin{aligned} &\|w\Delta_{h\varphi}^r f\|_{L_p[-1+2h^2r^2,\,1-2h^2r^2]} \\ &\qquad \leqslant M(\|w(f-P_{2^l})\|_{L_p[-1,1]} + h^r\|w\varphi^r P_{2^l}^{(r)}\|_{L_p[-1,1]}) \end{aligned} \tag{7.5}$$

from which one can easily deduce

$$\Omega_\varphi^r(f,t)_{w,p} \leqslant M_1 \left[ E_{2^k}(f)_{w,p} + t^r \sum_{v=0}^{k} 2^{vr} E_{2^v}(f)_{w,p} \right]. \tag{7.6}$$

The formula (7.6) is the weak-type result (4.2) ($\beta = 1/2$) for our case and hence, we can follow the result of Section 4 to complete the proof. ∎

## 8. Weighted Polynomial Approximation in $L_p(R)$

We will apply here the result of Section 4 to weighted polynomial approximation with Freud's weight

$$W_\lambda(x) \equiv \exp(-|x|^\lambda), \qquad \lambda > 1.$$

The rate of best polynomial approximation is given by

$$E_n(f)_{W_\lambda,p} = \inf_{P \in \mathscr{P}_n} \|W_\lambda(f-P)\|_{L_p(R)}. \tag{8.1}$$

We note that more general situations than $W_\lambda$ were investigated for which analogous results would be achieved with minor but messy modifications of the results in Section 4. The related main-part moduli are given by

$$\Omega^r(f,t)_{W_\lambda,p} \equiv \sup_{0<h\leqslant t} \|W_\lambda \Delta_h^r f\|_{L_p[-h^{1/(1-\lambda)},\,h^{1/(1-\lambda)}]}. \tag{8.2}$$

In [7, Chap. 11], the weak Jackson inequality

$$E_n(f)_{W_\lambda,p} \leqslant C \int_0^{n^{(1-\lambda)/\lambda}} (\Omega^r(f,t)_{W_\lambda,p}/t)\,dt + Ce^{-cn^{(\lambda-1)/\lambda}} \|W_\lambda f\|_p \quad (8.3)$$

was proved.

The Bernstein inequality (see for instance [7, p. 185])

$$\|W_\lambda P'\|_{L_p(R)} \leqslant Cn^{(\lambda-1)/\lambda} \|W_\lambda P\|_{L_p(R)}, \qquad P \in \mathscr{P}_n \quad (8.4)$$

will complete the necessary prerequisites for a result of the type given in Section 4 and hence, we can derive the following theorem.

THEOREM 8.1. *Suppose* $\Omega^r(f,t)_{W_\lambda,p}$ *and* $E_n(f)_{W_\lambda,p}$ *are given by* (8.2) *and* (8.1) *where* $W_\lambda(x) = \exp(-|x|^\lambda)$, $\lambda > 1$, *and* $1 \leqslant p \leqslant \infty$. *Then for* $\psi \in S^*$, *the conditions*

$$\Omega^r(f,t)_{W_\lambda,p} \sim \psi(t^r)$$

*and*

$$E_n(f)_{W_\lambda,p} \sim \psi(n^{-r(1-(1/\lambda))})$$

*are equivalent.*

*Remark* 8.2. In [7, Chap. 11], the concepts $K_r(f,t^r)_{W_\lambda,p}$ and $\omega_r^*(f,t)_{W_\lambda,p}$ were also discussed. If these concepts rather than $\Omega^r(f,t)_{W_\lambda,p}$ were used in Theorem 8.1, we could relax the condition on $\psi$ and assume only $\psi \in S$ as we could apply Theorem 2.2 rather than 4.2.

We also should remark that for some particular functions $\psi(t)$, the implication $\Omega^r(f,t)_{W_\lambda,p} \sim \psi(t^r)$ implies $E_n(f)_{W_\lambda,p} \sim \psi(n^{-r(1-(1/\lambda))})$ was shown in [7].

## 9. MULTIVARIATE BEST APPROXIMATION

On the domain $S \subset R^d$, we can define best $n$th degree algebraic polynomial approximation by

$$E_n(f)_{L_p(S)} = \inf\{\|f-P\|;\, P \text{ a polynomial of total degree } n\}. \quad (9.1)$$

We recall that a polytope is the convex hull of finitely many points and a simple polytope $S \subset R^d$ is a polytope with an interior point for which any one of its vertices is connected to other vertices by exactly $d$ edges. For a simple polytope, the moduli of smoothness $\omega_S^r(f,t)_p$ and $\tilde{\omega}_S^r(f,t)_p$ were defined in [7, p. 202]. As a corollary of results in [7] and in Section 2, we obtain the following theorem.

THEOREM 9.1. *Suppose $S$ is a simple polytope, $S \subset R^d$, $1 \leqslant p \leqslant \infty$, $\omega_S^r(f, t)_p$, and $\tilde{\omega}_S^r(f, t)_p$ are as defined in [7, p. 202] and $\psi \in S$. Then*

$$E_n(f)_{L_p(S)} \sim \psi(1/n^r),$$

$$\omega_S^r(f, t)_p \sim \psi(t^r),$$

*and*

$$\tilde{\omega}_S^r(f, t)_p \sim \psi(t^r)$$

*are equivalent.*

*Remark* 9.2. When discussing $\tilde{\omega}_S^r(f, t)_p$ and $\omega_S^r(f, t)_p$, we should recall the definitions of these measures of smoothness. For $1 \leqslant p < \infty$

$$\tilde{\omega}_S^r(f, t)_p = \sum_{\substack{e \in V^{d-1} \\ 0 < h \leqslant t}} \Lambda \int_{x \in e^{\perp}} \int_{-\infty}^{\infty} \left| \Delta^r_{eh\tilde{d}_S(e, x+\lambda e)^{1/2}} f(x + \lambda e) \right|^p d\lambda \, dm_e(x) \Big\}^{1/p}, \tag{9.2}$$

where $V^{d-1}$ is the set of unit vectors in $R^d$, $m_e$ is the $(m-1)$ dimensional Lebesgue measure on $e^{\perp}$, and

$$\tilde{d}_S(e, x) = \big(\min_{x + \lambda e \notin S} d(x, x + \lambda e)\big)\big(\max_{x + \lambda_i \in S} d(x + \lambda_1 e, x + \lambda_2 e)\big). \tag{9.3}$$

Similarly, $\tilde{\omega}_S^r(f, t)_\infty$ is defined. The modulus $\omega_S^r(f, t)_p$ is defined using (2.9) where instead of taking the supremum for all $e \in V^{d-1}$, we take it only in directions parallel to the edges of the simplex.

*Proof.* The Jackson inequality

$$\begin{aligned} E_n(f)_{L_p(S)} &\leqslant M[\omega_S^r(f, 1/n)_p + n^{-r}\|f\|_p] \\ &\leqslant M[\tilde{\omega}_S^r(f, 1/n)_p + n^{-r}\|f\|_p] \end{aligned} \tag{9.4}$$

is given in [7, (12.2.3)] and the weak-type result

$$\begin{aligned} \omega_S^r(f, t)_p &\leqslant \tilde{\omega}_S^r(f, t)_p \\ &\leqslant C\left[ E_{2^l}(f)_{L_p(S)} + t^r \sum_{v=0}^{k} 2^{vr} E_{2^v}(f)_{L_p(S)} + t^r \|f\|_{L_p(S)} \right] \end{aligned} \tag{9.5}$$

is achieved as a step in the proof of (12.2.3) in [7] (see [7, p. 206] used without the restriction $2^l < 1/t < 2^{l+1}$).

We use (9.5) and (9.4) to obtain our result. We recall that we may deal with the moduli of smoothness rather than the $K$-functional as was done in Section 4. ∎

*Remark* 9.3. The Bernstein inequality of the approximation process by best algebraic polynomials on a simple polytope can be given explicitly but the inequality (9.5) is much simpler and is what we need here.

For a Banach function space $X$ on $T^d$ satisfying conditions (5.1) or (5.2), we define

$$E_n(f)_X = \inf_{\tau \in \mathcal{T}_n} \|f - \tau\|_X, \tag{9.6}$$

where $\mathcal{T}_n$ is the set of trigonometric polynomials of degree $n$ in each direction. We define

$$\omega^r(f, t)_X = \sup_{0 < |v| \leqslant t} \|\Delta_v^r f\|_X, \tag{9.7}$$

where

$$\Delta_v^r f(u) = \Delta_v^{r-1}(f(u + (v/2)) - f(u - (v/2))), \qquad u, v \in R^d. \tag{9.8}$$

We now have the following generalization of Theorem 5.1.

THEOREM 9.4. *Suppose $X$ is a Banach space of functions or distributions on $T^d$ satisfying* (5.1) *or* (5.2) *and $\psi \in S$. Then*

$$E_n(f)_X \sim \psi(n^{-r})$$

*and*

$$\omega^r(f, t)_X \sim \psi(t^r)$$

*are equivalent for $E_n(f)_X$ and $\omega^r(f, t)_X$ given by* (9.6) *and* (9.7), *respectively.*

*Proof.* For $X = L_\infty(T^d)$, we have

$$E_n(f)_X \leqslant C\omega^r(f, 1/n)_X \tag{9.9}$$

and

$$\omega^r(f, t)_X \leqslant 2^r E_{2^k}(f)_X + Ct^r \sum_{v=0}^{k} 2^{vr} E_{2^v}(f)_X. \tag{9.10}$$

While we could not find the statements (9.9) and (9.10), they can be shown following the more complicated case proved in [7, Chap. 12], and probably are known (at least for $f \in L_\infty(T^d)$). In [10, p. 273, 350], much more complicated formulas appear which cannot be used here. From (9.9) and (9.10) for $X = L_\infty(T^d)$, we get these formulas for $X$ satisfying (5.1) or (5.2) using the same steps used in [5, 6]. We now follow earlier sections to obtain our theorem from (9.9) and (9.10). ∎

## 10. Some Linear Approximation

The theorem in Section 3 was tailor-made to fit many linear approximation processes.

For example, for convolution approximation processes we have:

THEOREM 10.1. *Suppose the sequence $G_n(x)$ of functions on $R$ or $T$ satisfy*

(a) $$\|G_n\|_{L_1} \leqslant M,$$

(b) $$\int G_n(t)\, t^i\, dt = o(\sigma_n^r) \qquad \text{for} \quad 0 < i < r,$$

(c) $$\int G_n(t)\, dt = 1,$$

(d) $$\int |t|^r\, |G_n(t)|\, dt = O(\sigma_n^r), \text{ and}$$

(e) $$\int |G_n'(t)|\, dt = O(\sigma_n^{-1})$$

*for $\sigma_n = o(1)$ satisfying $\sigma_{n+1} \geqslant c\sigma_n$ for some $c > 0$.*

*Then for $\psi \in S$ and $B$ a Banach space of functions on $R$ or $T$ for which translation is a continuous isometry*

$$\|G_n * f - f\|_B \sim \psi(\sigma_n^r)$$

*and*

$$\omega^r(f, t)_B \sim \psi(t^r)$$

*are equivalent, and each imply*

$$\omega^r(f, \sigma_n)_B \sim \|G_n * f - f\|_B.$$

*Proof.* We define

$$G_{n,r}(t) \equiv G_{n,r-1} * G_n(t), \qquad G_{n,1}(t) \equiv G_n(t), \qquad G_{n,0}(t) \equiv I. \quad (10.1)$$

Assumption (a) now implies

$$\|G_{n,k} * f\|_B \leqslant M\|G_{n,k-1} * f\|_B \leqslant M^k\|f\|_B, \tag{10.2}$$

and

$$\begin{aligned}\|G_{n,k} * f - f\|_B &\leqslant \sum_{\nu=0}^{k} \|G_{n,\nu} * (G_n * f - f)\|_B \\ &\leqslant (1 + M + \dots M^{m-1})\|G_n * f - \ \|_B.\end{aligned} \tag{10.3}$$

The assumptions (b), (c), and (d) imply for $g \in B$ where $g^{(r)}$ the $r$th strong derivative of $g$ in $B$ satisfies $g^{(r)} \in B$,

$$\|G_n * g - g\|_B \leqslant C\sigma_n^r\|g^{(r)}\|_B,$$

and therefore,

$$\|G_{n,k} * g - g\|_B \leqslant C(1 + \cdots + M^{k-1})\,\sigma_n^r\|g^{(r)}\|_B \leqslant C_1\sigma_n^r\|g^{(r)}\|_B. \tag{10.4}$$

The assumption (e) implies

$$\left\|\frac{d}{dx} G_n * f\right\|_B \leqslant C_2\sigma_n^{-1}\|f\|_B,$$

and therefore, for $k \geqslant r$,

$$\left\|\left(\frac{d}{dx}\right)^r G_{n,k} * f\right\|_B \leqslant C_2\sigma_n^{-r}\|f\|_B. \tag{10.5}$$

For $g \in B^r$ the subspace for which the strong $r$ derivative of $g$ in $B$ exists and $\Phi(g) = \|g^{(r)}\|_B < \infty$, we have

$$\left\|\left(\frac{d}{dx}\right)^r G_{n,k} * g\right\|_B \leqslant M^r\|g^{(r)}\|_B. \tag{10.6}$$

Using Theorem 3.1, we have for $k \geqslant r$

$$\omega^r(f, t)_B \sim \psi(t^r) \Rightarrow \|G_{n,k} * f - f\|_B \sim \psi(\sigma_n^r).$$

Using (10.3), we have

$$\|G_{n,k} * f - f\|_B \geqslant A\psi(\sigma_n^r).$$

We also have

$$\begin{aligned}\|G_n * f - f\|_B &\leqslant \|G_n * (f - G_{n,r} * f)\|_B + \|G_{n,r+1} * f - f\|_B \\ &\leqslant M\|G_{n,r} * f - f\|_B + \|G_{n,r+1} * f - f\|_B,\end{aligned} \tag{10.7}$$

and therefore,

$$\|G_n * f - f\|_B \leqslant A_1 \psi(\sigma_n^r).$$

We now deduce from $\|G_n * f - f\|_B \sim \psi(\sigma_n^r)$ using (10.3) and (10.7) that

$$A_2 \psi(\sigma_n^r) \leqslant \max(\|G_{n,r} * f - f\|_B, \|G_{n,r+1} * f - f\|_B) \leqslant A_3 \psi(\sigma_n^r).$$

Now $G_n^*$ which is equal to $G_{n,r}$ or $G_{n,r+1}$, whichever achieves the maximum above, satisfies the conditions of Theorem 3.2 and implies

$$\omega^r(f, t)_B \sim \psi(t^r). \quad \blacksquare$$

In [7, Chap. 9], combinations of Bernstein-type operators $L_{n,r}(f)$ were discussed and related to the moduli $\omega_\varphi^{2r}(f, t)_p$ where $\varphi$ depended on the particular approximation process. (For instance for combinations of Bernstein polynomials, $\varphi(x)^2 = x(1-x)$.) We use the inequality (9.3.1) of [7] for the Jackson inequality with $\omega_\varphi^{2r}(f, t)_p$ taking the place of $K(f, t^{2r})_p$ in (3.1). We use (9.3.2) of [7] for the crucial inequality (3.9) (again with $\omega_\varphi^{2r}(f, t)_p$ standing for $K(f, t^{2r})_p$). With the above, we now have the following theorem.

THEOREM 10.2. *For $L_{n,r}(f)$ given in [7, Theorem 9.3.2] and $\psi(t) \in S$*

$$\omega_\varphi^{2r}(f, t)_p \sim \psi(t^{2r}) \tag{10.8}$$

*and*

$$\|L_{n,r} f - f\|_p \sim \psi(n^{-r}) \tag{10.9}$$

*are equivalent.*

In [7, Corollary 9.3.8], it was shown that a somewhat stronger condition than (10.8) implies (10.9).

For the reader who is not familiar with combinations of exponential type operators as given in [7, Chap. 9], we give the following special case of Theorem 10.2.

COROLLARY 10.3. *For $\psi \in S$ the conditions*

$$K_\varphi(f, t^4)_C \equiv \inf_g (\|f - g\|_{C[0,1]} + t^4 \|x^2(1-x)^2 g^{(4)}\|_{C[0,1]}) \sim \psi(t^4) \tag{10.10}$$

*and*

$$\|2B_{2n} f - B_n f - f\|_{C[0,1]} \sim \psi(n^{-2}), \tag{10.11}$$

*where*

$$B_n f(x) \equiv \sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k} f\left(\frac{k}{n}\right)$$

*are equivalent.*

Exponential type operators include other operators like Baskakov, Szasz–Mirakjan, Gauss–Weierstrass, and Post–Widder operators. $L_{n,r}(f)$ of Theorem 10.2 include also modification following Kantorovich to accommodate $L_p$ spaces as well as combinations that accommodate any fixed degree of smoothness.

## References


1. N. I. Achieser, "Theory of Approximation," Ungar, New York, 1956 (translated from Russian).
2. N. K. Bari and S. B. Stechkin, Best approximation and differential properties of two conjugate functions, *Trudy Moskov. Mat. Obshch.* **5** (1956), 483–522.
3. J. Boman, A problem of Stečkin on trigonometric approximation, *in* "Constructive Function Theory," pp. 269–272, Proceedings of Conference held in 1977, Blagoevgrad, Bulgaria, 1980.
4. W. Dickmeis and R. J. Nessel, Condensation principles with rates, *Studia Math.* **75** (1982), 55–68.
5. Z. Ditzian, Some remarks on approximation theorems on various Banach spaces, *J. Math. Anal. Appl.* **77** (1980), 567–576.
6. Z. Ditzian, On Lipschitz classes and derivative inequality in various Banach spaces, *in* "Functional Analysis, Holomorphy and Approximation" (G. I. Zapata, Ed.), pp. 57–67, Proceedings of Conference held in Rio de Janeiro, Brazil, 1981, North-Holland, Amsterdam, 1984.
7. Z. Ditzian and V. Totik, "Moduli of Smoothness," Springer-Verlag, New York/Berlin, 1987.
8. Z. Ditzian and V. Totik, Remarks on Besov spaces and best polynomial approximation, *Proc. Amer. Math. Soc.* **104** (1988), 1059–1066.
9. R. J. Nessel and E. Wickeren, Negative results in connection with Favard's problem on the comparison of approximation processes, *in* "Anniversary Volume on Approximation Theory and Functional Analysis" (P. L. Butzer, R. I. Stens, and B. Sz. Nagy, Eds.), pp. 189–200, Birkhäuser-Verlag, Basel, 1984.
10. A. F. Timan, "Theory of Approximation of Functions of a Real Variable," Pergamon, Elmsford, NY, 1963 (translated from Russian).