A TAUBERIAN THEOREM RELATED TO APPROXIMATION THEORY

BY

HAROLD S. SHAPIRO

The University of Michigan, Ann Arbor, Mich., U.S.A.

As pointed out in [4], a number of known results in the theory of approximation may be interpreted in terms of the behaviour, for small values of the parameter, of an expression (called by us the σ-deviation) based on convolving a function with a measure σ. More specifically, so called direct, inverse, and saturation theorems of approximation theory are all interpretable in terms of comparison (for fixed f) of the σ_1-deviation with the σ_2-deviation, where σ_1 and σ_2 are suitably chosen measures. This is basically a Tauberian problem; results of considerable generality are proved in the present paper, and some applications to approximation theory are given (a more systematic account of applications will appear elsewhere).

1.1. *Notation and preliminaries.* By $R^m (m \geqslant 1)$ we denote real Euclidean m-space. If $x=(x_1, \ldots, x_m)$ and $y=(y_1, \ldots, y_m)$ are elements of R^m, $x \cdot y$ denotes $\sum_{i=1}^m x_i y_i$ and $\|x\| = (x \cdot x)^{\frac{1}{2}}$. The letters $t, u, v, \ldots, y$ always denote elements of R^m. By C we denote the class of complex-valued bounded continuous functions on R^m. M denotes the class of complex-valued finite countably additive measures on the Borel sets of R^m, made into a ring under the convolution $\ast$ characterized by: $\varrho = \sigma \ast \tau$ if and only if, for all(1) $g \in G$, $\int g d\varrho = \int\int g(t+u)\, d\sigma_t d\tau_u$. (Greek letters shall always denote measures.) W denotes the isomorphic (Wiener) ring (with respect to ordinary multiplication) of Fourier transforms $\hat{\sigma} : \hat{\sigma}(x) = \int e^{-ix \cdot t} d\sigma_t$ of elements of M. $V(\sigma)$ denotes the total variation of σ, and W is normed by taking $\|\hat{\sigma}\|_W = V(\sigma)$. Observe that for any positive scalar a, $\hat{\sigma}(ax)$ has the same norm as $\hat{\sigma}$.

For $p \in L^1(R^m)$ we write $\hat{p}$ to denote the Fourier transform of the measure $p d\lambda$

(1) A subscript on a measure indicates the variable with respect to which to integrate.

($\lambda=\lambda^m$: Lebesgue measure on R^m), i.e., $\hat{p}(x)=\int e^{-ix\cdot t}p(t)\,d\lambda_t$. δ always denotes a unit mass concentrated at the origin.

For $\sigma \in M$, and a positive scalar a we define the *σ-deviation* of a function $f \in C$ by

$$D_\sigma(f;a)=\sup_{t\in R^m}\left|\int f(t-au)\,d\sigma_u\right|. \tag{1}$$

An evident consequence of the definition is

$$D_{\sigma+\tau}(f;a)\leqslant D_\sigma(f;a)+D_\tau(f;a). \tag{2}$$

LEMMA 1. *If σ divides ϱ (in the ring M), say $\sigma * \tau=\varrho$,*

$$D_\varrho(f;a)\leqslant V(\tau)\,D_\sigma(f;a). \tag{3}$$

Proof.

$$\left|\int f(t-au)\,d\varrho_u\right|=\left|\int\int f(t-au-av)\,d\sigma_u\,d\tau_v\right|\leqslant V(\tau)\sup_{v\in R^m}\left|\int f(t-au-av)\,d\sigma_u\right|$$
$$=V(\tau)\,D_\sigma(f;a)$$

and the result follows.

Combining this with (2) above gives the more general result[1]

LEMMA 2. *If ϱ belongs to the ideal generated by $\sigma_1,\dots,\sigma_r$, say $\varrho=\sum_1^r\sigma_i*\tau_i$,*

$$D_\varrho(f;a)\leqslant\sum_{i=1}^{r}V(\tau_i)\,D_{\sigma_i}(f;a). \tag{4}$$

Clearly the hypothesis of Lemma 2 can be phrased: "If $\hat{\varrho}$ belongs to the ideal (in W) generated by $\hat{\sigma}_1,\dots,\hat{\sigma}_r$." This formulation is more convenient for us. Our main result is that an inequality slightly weaker than (4) is true if we assume only that $\hat{\varrho}$ belongs *locally* to the ideal generated by $\hat{\sigma}_1,\dots,\hat{\sigma}_r$, i.e., that for some $\tau_i\in M$, $\hat{\varrho}-\sum_{i=1}^r\hat{\tau}_i\hat{\sigma}_i$ vanishes in a neighborhood of the origin (for a precise statement see Theorem 2).

1.2. We leave to the reader the proof of

LEMMA 3. *If a pair of measures satisfy the relation $\hat{\varrho}(x)=\hat{\sigma}(cx)$, where c is a positive scalar, then for any f and $a>0$, $D_\varrho(f;a)=D_\sigma(f;ca)$.*

Remark. Since σ and $\hat{\sigma}$ uniquely determine each other, we may, by an abuse of language, speak of $D_\sigma(f;a)$ also as the *$\hat{\sigma}$-deviation* of f. This is convenient since we

(1) Conversely, an inequality of the form (4) implies p belongs to the ideal generated by $\sigma_1,\dots,\sigma_n$ (this was shown and kindly communicated to us by D. L. Ragozin).

shall speak mostly not of measures but of their Fourier transforms. Thus, Lemma 3 says simply that $\hat{\sigma}$ and the $\hat{\sigma}$-deviation are *covariant* with respect to scale change in the independent variable.

We shall say a sequence $\sigma_n \in M$ is *s-convergent* to σ, written $\sigma_n \xrightarrow{s} \sigma$ to mean

$$\lim_{n\to\infty} \int f\,d\sigma_n = \int f\,d\sigma, \quad \text{all } f \in C. \tag{5}$$

The notion of s-convergence when n is a continuous parameter is defined similarly. By abuse of language we shall also speak of s-convergence of elements of W, that is $\hat{\sigma}_n \xrightarrow{s} \hat{\sigma}$, in case (5) holds. An example, which we shall need later, is given by

LEMMA 4. *For $k \in L^1(R^m)$ we have as $a \to +\infty$*

$$a^m k(at)\,d\lambda \xrightarrow{s} \hat{k}(0)\,d\delta$$

$$\hat{k}\left(\frac{x}{a}\right) \xrightarrow{s} \hat{k}(0).$$

Proof. The two relations are equivalent, and we prove the first. For $f \in C$ we have (all integrations being over R^m)

$$\int f(t)\,a^m k(at)\,d\lambda = \int f\left(\frac{u}{a}\right) k(u)\,d\lambda \to f(0)\int k(u)\,d\lambda$$

as $a \to +\infty$ by dominated convergence.

Some useful properties of s-convergence are summed up in

LEMMA 5. *Let R_n, S_n denote elements of W. Then*

(i) *If $R_n \xrightarrow{s} R$, $S_n \xrightarrow{s} S$, then $aR_n + bS_n \xrightarrow{s} aR + bS$ for any real scalars a, b.*

(ii) *If $R_n \xrightarrow{s} R$, then $SR_n \xrightarrow{s} SR$ for any $S \in W$.*

(iii) *If $R_n \xrightarrow{s} R$, then $R_n(ax) \xrightarrow{s} R(ax)$ for $a > 0$.*

(iv) *If $\sum_{n=1}^{\infty} \|R_n\| < \infty$, and $T(x) = \sum_{n=1}^{\infty} R_n(x)$, $T_n = \sum_{j=1}^{n} R_j$ is s-convergent to T.*

Proof. (i) and (iii) are immediate and left to the reader. As for (ii), let R_n, R, S be the Fourier transforms of ϱ_n, ϱ, σ respectively. For any $f \in C$ we have

$$\int f d(\sigma * \varrho_n) = \iint f(t+u)\,d\sigma(t)\,d\varrho_n(u) = \int g\,d\varrho_n \to \int g\,d\varrho = \iint f(t+u)\,d\sigma(t)\,d\varrho(u) = \int f d(\sigma * \varrho)$$

(where we have written $g(u)$ for $\int f(t+u)\,d\sigma(t)$), showing that $\sigma * \varrho_n \xrightarrow{s} \sigma * \varrho$.

As for (iv), write $T = T_n + Z_n$, where $\|Z_n\| \to 0$. Letting T, T_n, Z_n be the Fourier transforms of τ, τ_n, ζ_n respectively, we have for $f \in C$

$$\left| \int f d\tau - \int f d\tau_n \right| = \left| \int f d\zeta_n \right| \leqslant \sup |f| \, V(\zeta_n)$$

which tends to zero as $n \to \infty$.

LEMMA 6. *If* $\sigma_n \xrightarrow{s} \sigma$, *then for any* $f \in C$, $a > 0$

$$D_\sigma(f; a) \leqslant \liminf_{n\to\infty} D_{\sigma_n}(f; a).$$

Proof. Let $\varepsilon > 0$, and choose t_0 so that

$$\left| \int f(t_0 - au)\, d\sigma(u) \right| > D_\sigma(f; a) - \varepsilon.$$

By hypothesis,

$$\left| \int f(t_0 - au)\, d\sigma(u) \right| = \lim_{n\to\infty} \left| \int f(t_0 - au)\, d\sigma_n(u) \right| \leqslant \liminf_{n\to\infty} D_{\sigma_n}(f; a),$$

hence
$$D_\sigma(f; a) - \varepsilon \leqslant \liminf_{n\to\infty} D_{\sigma_n}(f; a)$$

and the result follows.

We also require the following well-known result on Fourier transforms:

LEMMA 7. *Let* $R, S_1, S_2, \ldots, S_r$ *be elements of* W, *and* K *a compact set in* R^m *such that*

(a) $R(x) = 0$ *outside* K,

(b) $\sum_{i=1}^{r} |S_i(x)| > 0$ *for* $x \in K$.

Then R *belongs to the ideal in* W *generated by* $S_1, \ldots, S_r$.

For the proof we refer the reader to the literature (e.g., Rudin [3]).

1.3. *The Tauberian theorem*

LEMMA 8. *If* $\hat{\sigma}$ *is different from zero for* $\|x\| = 1$, *and* $\hat{\sigma}$ *divides* $\hat{\varrho}$ *at the origin (that is,* $\hat{\varrho} - \hat{\sigma}\hat{\tau}$ *vanishes in a neighborhood of the origin, for some* $\tau \in M$), *we have for any* $f \in C$ *and* $a > 0$

$$D_\varrho(f; a) \leqslant A D_\sigma(f; a) + B \sum_{i=0}^{\infty} D_\sigma(f; Cb^i a). \tag{6}$$

Here A, B, C *are positive constants depending only on* σ *and* ϱ, *and* b $(0 < b < 1)$ *depends only on* σ.

Proof. By continuity, there is a number $c>1$ such that $\hat{\sigma}(x)\neq 0$ for any $x\in E$, where $E=\{x \mid 1\leqslant \|x\|\leqslant c\}$. Let P be a function of class C^∞ equal to one for $\|x\|\leqslant 1$ and to zero for $\|x\|\geqslant c^{\frac{1}{2}}$. Then, setting $b=c^{-\frac{1}{2}}$, $P(bx)-P(x)$ vanishes for all x in the complement of E and therefore by the case $r=1$ of Lemma 7,

$$P(bx)-P(x)=\hat{\sigma}(x)\,Q(x),\quad Q\in W. \tag{7}$$

Now, $$\sum_{i=0}^{n}(P(b^{i+1}x)-P(b^i x))=P(b^{n+1}x)-P(x)\xrightarrow{s}1-P(x)$$

as $n\to\infty$, by Lemma 4. Therefore, in view of (7)

$$P(x)+\sum_{i=0}^{n}\hat{\sigma}(b^i x)\,Q(b^i x)\xrightarrow{s}1.$$

Therefore, by Lemma 5, (iii), for any $C>0$

$$P(Cx)+\sum_{i=0}^{n}\hat{\sigma}(Cb^i x)\,Q(Cb^i x)\xrightarrow{s}1$$

and applying Lemma 5, part (ii)

$$P(Cx)\,\hat{\varrho}(x)+\sum_{i=0}^{n}\hat{\sigma}(Cb^i x)\,Q(Cb^i x)\,\hat{\varrho}(x)\xrightarrow{s}\hat{\varrho}(x). \tag{8}$$

Finally, recall that by hypothesis, for some $T\in W$,

$$\hat{\sigma}(x)\,T(x)=\hat{\varrho}(x)$$

holds in some neighborhood N of the origin, and therefore we can find a positive constant C such that

$$P(Cx)\,\hat{\varrho}(x)=P(Cx)\,\hat{\sigma}(x)\,T(x) \tag{9}$$

holds for *all* x, since when x is outside N, $P(Cx)$ vanishes, if C has been chosen large enough (P being of compact support). From (8) and (9) we have, with such a choice of C

$$P(Cx)\,T(x)\,\hat{\sigma}(x)+\sum_{i=0}^{n}Q(Cb^i x)\,\hat{\varrho}(x)\,\hat{\sigma}(Cb^i x)\xrightarrow{s}\hat{\varrho}(x). \tag{10}$$

We now apply Lemmas 2, 3 and 6 to (10), and we get, for any $f\in C$ and $a>0$

$$D_\varrho(f;a)\leqslant\|P\|_W\,\|T\|_W\,D_\sigma(f;a)+\|Q\|_W\,V(\varrho)\sum_{i=0}^{\infty}D_\sigma(f;Cb^i a)$$

and the Lemma is proved.

Remarks. It should be observed (for purposes of application) that, for a concrete choice of σ, ϱ the various constants occurring can usually be estimated quite explicitly.

Also, if $\hat{\sigma}$ is fairly smooth (this is generally the case in applications) the depth of Lemma 7 is not required (since then the quotient of $P(x)-P(bx)$ by $\hat{\sigma}(x)$ is smooth and of compact support, which guarantees, by elementary theorems, that it is an L^1 transform). We are indebted to C. O. Kiselman for calling this point to our attention.

LEMMA 9. *If $\sigma_1, \ldots, \sigma_r$ are measures such that $\sum_{i=1}^r |\hat{\sigma}_i(x)|$ is different from zero for $\|x\|=1$, and $\hat{\varrho}$ belongs locally (at $x=0$) to the ideal generated by $\hat{\sigma}_1, \ldots, \hat{\sigma}_r$, then*

$$D_\varrho(f; a) \leqslant \sum_{i=1}^r A_i D_{\sigma_i}(f; a) + B \sum_{k=0}^\infty \sum_{i=1}^r D_{\sigma_i}(f; Cb^k a),$$

where A_i, B, C are positive constants depending only on ϱ and σ, and b $(0<b<1)$ depends only on σ.

Proof. We construct P as in the previous proof, and now get in place of (7)

$$P(bx)-P(x) = \sum_{i=1}^r Q_i(x)\, \hat{\sigma}_i(x).$$

The rest of the proof is similar to the proof of Lemma 8 and we omit it.

LEMMA 10. *Consider any family of continuous complex-valued functions on a compact topological space. If every finite subcollection has a common zero, there is a point where all the functions of the family vanish.*

Proof. Assume the contrary. Then to each point x in the underlying space X we can associate a function f_x in the given family which is different from zero in a neighborhood N_x of x. By compactness, $X = \bigcup_{i=1}^n N_{x_i}$ for suitably chosen x_i. The associated f_{x_i} $(i=1, \ldots, n)$ have, by hypothesis, a common zero y. Since $y \in N_{x_j}$ for some j $(1 \leqslant j \leqslant n)$ we have $f_{x_j}(y) \neq 0$. Contradiction.

Definition. A function F on R^m satisfies the *Tauberian condition* if, on each half-ray through the origin, there is a point where F does not vanish. (In other words, if for every x with $\|x\|=1$ there exists $c \geqslant 0$ such that $F(cx) \neq 0$.)

A measure σ will be said to satisfy the Tauberian condition if $\hat{\sigma}$ does.

Examples. (a) Any *real* non-null measure σ on R^1 satisfies the Tauberian condition, since $\hat{\sigma}(-x) = \overline{\hat{\sigma}(x)}$, therefore there exist at least one positive and one negative value of x for which $\hat{\sigma}(x) \neq 0$.

(b) On R^1, the measure $dt/(t+i)^2$ doesn't satisfy the Tauberian condition, because its Fourier transform vanishes for $x \leqslant 0$.

(c) If δ denotes a unit mass concentrated at the origin, and $K \in L^1(R^m)$, $d\delta - K d\lambda$ satisfies the Tauberian condition, since its Fourier transform is $1-\hat{K}$, which is non-zero for all sufficiently large $\|x\|$, by the Riemann–Lebesgue theorem.

LEMMA 11. *A continuous complex-valued function F on R^m satisfies the Tauberian condition if and only if there exist positive numbers $c_1 < \ldots < c_r$ such that*

$$\sum_{i=1}^{r} |F(c_i x)| > 0, \quad \textit{all} \quad \|x\| = 1. \tag{11}$$

Proof: If F fails to satisfy the Tauberian condition there exists x_0 with $\|x_0\| = 1$ and $F(cx_0) = 0$ for all $c \geqslant 0$, so (11) cannot hold. Suppose on the other hand F satisfies the Tauberian condition. Then the functions $F_c(x) = F(cx)$, where c ranges over all positive numbers, do not all vanish at any point of the unit sphere. By Lemma 10, there is a finite subcollection of the F_c with no common zero on the unit sphere, and this gives (11).

We come now to our main result.

THEOREM 1. *If $\hat{\sigma}$ satisfies the Tauberian condition, and $\hat{\sigma}$ divides $\hat{\varrho}$ at the origin, we have for any $f \in C$ and $a > 0$*

$$D_\varrho(f; a) \leqslant A \sum_{k=0}^{\infty} \sum_{j=1}^{n} D_\sigma(f; B_j b^k a). \tag{12}$$

Here A, n, B_j are constants depending only on ϱ and σ, and b $(0 < b < 1)$ depends only on σ.

COROLLARY. *If $\hat{\sigma}$ satisfies the Tauberian condition, and $\hat{\varrho}$ vanishes in a neighborhood of* 0, *then* (12) *holds.*

Remark. The hypothesis that σ satisfies the Tauberian condition cannot be dropped. For a counterexample, take $m = 1$ and consider any measure σ such that $\hat{\sigma}(x) = 0$ for $x \geqslant 0$. Suppose the conclusion of Theorem 1 were to hold for this σ. Choosing $f(t) = e^{it}$, we have for any $a > 0$,

$$\int f(t-au)\, d\sigma(u) = \int e^{i(t-au)}\, d\sigma(u) = e^{it}\, \hat{\sigma}(a) = 0$$

so that $D_\sigma(f; a)$ vanishes for all $a > 0$. If Theorem 1 were applicable to this σ, we should have $D_\varrho(e^{it}; a) = 0$, and hence $\hat{\varrho}(a) = 0$, for all $a > 0$ and any ϱ such that $\hat{\varrho}$ vanishes in a neighborhood of zero, which is obviously untrue. A similar example can be constructed in R^m.

Theorem 1 is the case $r = 1$ of the following theorem (we have stated it separately because it is important for applications).

THEOREM 2. *If $\sigma_1, \ldots, \sigma_r$ are measures such that $\sum_{i=1}^r |\hat{\sigma}_i|$ satisfies the Tauberian condition, and $\hat{\varrho}$ belongs locally (at $x=0$) to the ideal in W generated by $\hat{\sigma}_1, \ldots, \hat{\sigma}_r$ then for any $f \in C$ and $a > 0$*

$$D_\varrho(f; a) \leqslant A \sum_{k=0}^{\infty} \sum_{j=1}^{n} \sum_{i=1}^{r} D_{\sigma_i}(f; B_j b^k a). \tag{13}$$

Here A, n, B_j are constants depending only on ϱ and the σ_i, and b $(0 < b < 1)$ depends only on the σ_i.

Proof. By Lemma 11 there are positive constants $c_1 < \ldots < c_n$ such that

$$\sum_{i=1}^{r} \sum_{j=1}^{n} |\hat{\sigma}_i(c_j x)| > 0 \quad \text{for } \|x\| = 1.$$

Also, $\hat{\varrho}(c_1 x)$ belongs locally (at 0) to the ideal generated by $\{\sigma_i(c_1 x)\}_{i=1,\ldots,r}$ and *a fortiori* to the ideal generated by $\{\sigma_i(c_j x)\}_{i=1,\ldots,r; j=1,\ldots,n}$. Therefore, by Lemmas 3 and 9

$$D_\varrho(f; c_1 a) \leqslant \sum_{j=1}^{n} \left(\sum_{i=1}^{r} A_i D_{\sigma_i}(f; c_j a) + B \sum_{k=0}^{\infty} \sum_{i=1}^{r} D_{\sigma_i}(f; C b^k c_j a) \right)$$

and replacing a by a/c_1 gives the result, with a slight change of notation.

Remarks. In the most important applications we will have $D_{\sigma_i}(f; a) = O(a^p)$, $i = 1, \ldots, r$ for some positive p. (13) then gives $D_\varrho(f; a) = O(a^p)$.

An example of a situation where Theorem 2 is applicable but not Theorem 1 is obtained by taking $m = 2$, and $\hat{\sigma}_n(x) = (1 - e^{-iy_n \cdot x})^2$, $n = 1, 2, 3$ where the y_n are pairwise linearly independent vectors in R^2 and $\hat{\varrho}(x) = (1 - e^{-iy \cdot x})^2$. Then $\sum_1^3 |\hat{\sigma}_n|$ satisfies the Tauberian condition and it is not hard to verify that $\hat{\varrho}$ belongs locally (at 0) to the ideal generated by the $\hat{\sigma}_n$. From this one can deduce that if the second-order moduli of smoothness (see [2]) of the restrictions of an $f \in C(R^2)$ to each of three lines through the origin (1) are $O(a^q)$, the same is true for the restriction of f to every line through the origin (for more general results of this type see Boman [1]).

2. We shall now deduce some useful results from Theorem 1.

THEOREM 3. *Let σ, ϱ be measures on R^m. Suppose σ satisfies the Tauberian condition, and for some $f \in C$, $D_\sigma(f; a) = O(a^q)$ with $q > 0$, as $a \to 0$. Suppose there exists a function P defined and positive-homogeneous of degree $r > 0$ in a neighborhood N of the origin (i.e., $P(bx) = b^r P(x)$ for small positive b) such that both $P(x)$ and $\hat{\varrho}(x)/P(x)$ coincide in N with elements of W.*

(1) Actually *two* directions suffice, as further analysis shows (remarked by D. L. Ragozin).

Then $D_\varrho(f; a)$ *is large O of* a^q, $a^q|\log a|$, *or* a^r *according as* q *is less than, equal to, or greater than* r.

Proof. We remark first that there is no loss of generality in assuming that $\hat{\varrho}$ is itself positive-homogeneous in N. Indeed, suppose the theorem were known in this case. Writing $P(x) = \hat{\pi}(x)$ for $x \in N$ we obtain the desired conclusion first for $D_\pi(f; a)$, then for $D_\varrho(f; a)$ by an application(1) of Theorem 1. Consider now the function

$$T(x) = \hat{\varrho}(x) - 2^{-r}\hat{\varrho}(2x).$$

Because of the homogeneity of $\hat{\varrho}$, T vanishes in a neighborhood of 0, hence by the Corollary to Theorem 1,

$$D_\tau(f; a) \leqslant Aa^q \quad (\hat{\tau} = T),$$

where A does not depend on f or a. Now,

$$\hat{\varrho}(x) = \sum_{n=0}^{\infty} 2^{-nr} T(2^n x),$$

where the series on the right is s-convergent by Lemma 5, part (iv). Therefore, by Lemmas 3 and 6

$$D_\varrho(f; a) \leqslant \sum_{n=0}^{\infty} 2^{-nr} D_\tau(f; 2^n a) \leqslant A_1 \sum_{n=0}^{\infty} 2^{-nr} \min\left((2^n a)^q, 1\right)$$

where A_1 is a constant (as also $A_2, \ldots$ below). If $q < r$ the sum is less than

$$A_1 \sum_{n=0}^{\infty} (2^{q-r})^n a^q = A_2 a^q.$$

If $q \geqslant r$ observe that for each m

$$D_\varrho(f; a) \leqslant A_1 \sum_{n=0}^{m} 2^{(q-r)n} a^q + A_2 2^{-mr}. \tag{1}$$

For $q = r$ this is $O(ma^q + 2^{-mr})$ which is $O(a^q|\log a|)$ if we take $m = \left[\log_2 \frac{1}{a}\right]$. Finally, if $q > r$, the right side of (1) is $O(2^{(q-r)m} a^q + 2^{-mr})$ which is $O(a^r)$ if we choose $m = \left[\log_2 \frac{1}{a}\right]$. This completes the proof of Theorem 3.

Remark. It is clear that similar results could be obtained with a^q replaced by more general functions of a.

COROLLARY. *If* $f \in C$, *and* σ *is a non-null real measure on* R^1 *such that* $D_\sigma(f; a) = O(a^q)$, *then the modulus of smoothness of* f, *of order* r, *is large O of* a^q, $a^q \log(1/a)$, *or* a^r *as* $a \to 0$ *according as* q *is less than, equal to, or greater than* r.

(1) We must arrange that $\hat{\pi}$ satisfies the Tauberian condition, which we can certainly do without disturbing its values in N.

Proof. The modulus of smoothness of order r is defined by

$$\omega_r(f;s) = \sup_{0\leqslant a\leqslant s} D_{\beta_r}(f;a),$$

where $\beta=\beta_r$ is the discrete measure with "mass" $(-1)^j\binom{r}{j}$ at the point j $(j=0,1,\ldots,r)$, $\hat{\beta}(x)=(1-e^{-ix})^r$. The function $P(x)=(ix)^r$ is positive-homogeneous of degree r, and both $P(x)$ and $\hat{\beta}(x)/P(x)$ coincide with elements of W in a neighborhood of 0. Since σ satisfies the Tauberian condition (as remarked earlier, any real non-null measure on R^1 does), we have only to take β for ϱ in Theorem 3, and the Corollary follows.

We give some applications to approximation theory of the Corollary.

THEOREM 4. *If* $f\in C(R^1)$, $K\in L^1(R^1)$ *and for some* $q>0$

$$\left|f(t)-\int_{-\infty}^{\infty} f(t-u)\,sK(su)\,du\right| = O(s^{-q}), \quad s\to+\infty \tag{2}$$

uniformly in t, *the modulus of smoothness of* f, *of order* r, *is large* O *of* a^q, $a^q\log(1/a)$, *or* a^r *as* $a\to 0$, *according as* q *is less than, equal to, or greater than* r.

Remark. We could not find this result in the literature (nor the following theorem) even for special kernels K, e.g., $K(t)=(1/\pi)(\sin t/t)^2$ in which case the integral in (2), for f of period 2π, reduces to the Fejér sum of order s of the Fourier series of f. It is rather striking that in this theorem no hypothesis beyond integrability need be imposed on K.

Proof of Theorem 4. Making the change of variable $v=su$, and writing $a=1/s$, we may reformulate (2) as saying that the σ-deviation of f is $O(a^q)$ as $a\to 0$, where $d\sigma=d\delta-Kd\lambda$ (δ: unit mass at 0, λ: Lebesgue measure on R^1), and the result follows from the Corollary.

THEOREM 5. *If* $f\in C(R^1)$, $K\in L^1(R^1)$ *and* K *is not identically zero and has an integrable derivative of order* n, *and for some* q $(0<q<n)$

$$\frac{d^n}{dt^n}\int_{-\infty}^{\infty} f(t-u)\,sK(su)\,du = O(s^q), \quad s\to+\infty \tag{3}$$

uniformly in t, *the modulus of smoothness of* f, *of order* r, *is large* O *of* a^{n-q}, $a^r\log(1/a)$ *or* a^r *as* $a\to 0$, *according as* r *is greater than, equal to, or less than* $n-q$.

Proof. The left side of (3) equals

$$\frac{d^n}{dt^n}\int_{-\infty}^{\infty} f(u)\,sK(st-su)\,du = s^n\int_{-\infty}^{\infty} f(u)\,sK^{(n)}(st-su)\,du = a^{-n}\int_{-\infty}^{\infty} f(t-au)K^{(n)}(u)\,du \quad (a=s^{-1}).$$

Hence, taking $d\sigma = K^{(n)}du$, the σ-deviation of f is $O(a^{n-q})$, and the conclusion follows from the Corollary.

It is easy also to deduce from the Corollary "inverse theorems" of Bernstein and Zygmund, i.e., to infer from the existence of trigonometric polynomials (or entire functions of exponential type) which approximate f corresponding upper bounds for the moduli of smoothness (see [4]). Theorem 3 provides a convenient tool for proving higher-dimensional versions of these theorems. This method has perhaps some methodological interest in that the "Bernstein inequality" for the derivative of a trigonometric polynomial is not used.

3. L^p theory

It is easy to extend the provious analysis to L^p, rather than uniform, norms. We outline the main steps. In this section $\tilde{\sigma}$ always denotes the total variation of the complex measure σ. If α is any non-negative countably additive measure on the Borel sets of R^m, we define $L^p(\alpha)$ to be the set of complex-valued Borel measurable functions on R^m such that $\int |f|^p d\alpha < \infty$. By $\|f\|_p$ we always denote the usual L^p norm: $\|f\|_p = (\int |f|^p d\lambda^m)^{1/p}$, λ^m: Lebesgue measure on R^m, and L^p always denotes $L^p(\lambda^m)$. When there is no ambiguity we write λ for λ^m.

We shall always suppose $1 \leqslant p < \infty$. It will be convenient to introduce the following notations.

$$F_\sigma(t; a) = \int f(t - au)\, d\sigma_u \quad \text{for} \quad f \in L^p,\ \sigma \in M, \tag{1}$$

$$D_{\sigma, p}(f; a) = \|F_\sigma(\ ; a)\|_p. \tag{2}$$

$D_{\sigma, p}$ (as a function of a) we call *the σ, p deviation* of f. Let us first verify that for $f \in L^p$, $F_\sigma(\ ; a) \in L^p$. We have from (1),

$$|F_\sigma(t; a)| \leqslant \int |f(t - au)|\, d\tilde{\sigma}_u,$$

$$|F_\sigma(t; a)|^p \leqslant V(\tilde{\sigma})^{p-1} \int |f(t - au)|^p\, d\tilde{\sigma}_u,$$

$$\int |F_\sigma(t; a)|^p\, d\lambda \leqslant V(\tilde{\sigma})^{p-1} \int\int |f(t - au)|^p\, d\tilde{\sigma}_u\, d\lambda_t$$

and, carrying out the t integration first we see that the double integral on the right equals $V(\tilde{\sigma}) \int |f|^p d\lambda$, and we have proved that $F_\sigma(\ ; a) \in L^p$ and

$$D_{\sigma, p}(f; a) \leqslant V(\sigma)\, \|f\|_p. \tag{3}$$

Thus the σ, p modulus of an L^p function is finite and bounded. The reader may easily verify that it is uniformly continuous in a and vanishes at $a=0$.

We omit the simple verifications of

$$D_{\sigma,p}(f+g;a) \leqslant D_{\sigma,p}(f;a) + D_{\sigma,p}(g;a), \tag{4}$$

$$D_{\sigma+\tau,p}(f;a) \leqslant D_{\sigma,p}(f;a) + D_{\tau,p}(f;a). \tag{5}$$

LEMMA 12. *If σ divides ϱ (in the ring M), $\sigma \ast \tau = \varrho$, then*

$$D_{\varrho,p}(f;a) \leqslant V(\tau)\, D_{\sigma,p}(f;a) \tag{6}$$

for all $f \in L^p$.

Proof. For any $g \in L^p$ we have $\int g(t+u)\, d\sigma_t\, d\tau_u = \int g\, d\varrho$. Hence

$$\left|\int f(t-au)\, d\varrho_u\right| = \left|\int\int f(t-au-av)\, d\sigma_u\, d\tau_v\right| = \left|\int F_\sigma(t-av;a)\, d\tau_v\right|,$$

hence
$$|F_\varrho(t;a)|^p \leqslant V(\tau)^{p-1}\int |F_\sigma(t-av;a)|^p\, d\tilde{\tau}_v.$$

Multiplying by $d\lambda_t$ and integrating, and on the right side performing the t-integration first gives

$$\int |F_\varrho(t;a)|^p\, d\lambda \leqslant V(\tau)\int |F_\sigma(t;a)|^p\, d\lambda$$

and this implies (6).

From (5) and Lemma 12 we obtain the analog of Lemma 2. We can now easily prove

LEMMA 13. *Let ϱ, σ be as in Lemma 8. Then for any $f \in L^p$ and $a>0$ we have*

$$D_{\varrho,p}(f;a) \leqslant A D_{\sigma,p}(f;a) + B\sum_{i=0}^{\infty} D_{\sigma,p}(f; Cb^i a),$$

where A, B, C are positive constants depending only on σ and ϱ, and b $(0<b<1)$ depends only on σ.

Proof. With the same notations as in the proof of Lemma 8 we have

$$P(Cx)\,\hat{\varrho}(x) + \sum_{i=0}^{n} \hat{\sigma}(Cb^i x)\, Q(Cb^i x)\,\hat{\varrho}(x) = \hat{\varrho}(x)\, P(b^{n+1}x).$$

The only place where the proof of Lemma 8 must be modified is in showing

$$D_{\varrho,p}(f;a) \leqslant \liminf_{n\to\infty} D_{\varrho_n,p}(f;a), \tag{7}$$

where $\hat{\varrho}_n(x) = \hat{\varrho}(x)\, P(b^{n+1}x)$. Now, if $f \in L^p$,

$$\int f(t-au)\, d\varrho_n(u) = \iint f(t-au-av)\, s^m k(sv)\, d\varrho_u\, d\lambda_v,$$

where $s = b^{-n-1}$, since $P(x/s)$ is the Fourier transform of $s^m k(st)\, d\lambda^m$ (here we write $P = \hat{k}$, rather than $\hat{p}$, to avoid notational confusion, and $d\varrho_n(u)$ to indicate that integration is with respect to u).

$$F_{\varrho_n}(t; a) = \int F_\varrho(t-av; a)\, s^m k(sv)\, d\lambda_v$$

$$= \int F_\varrho(t-b^{n+1}au; a)\, k(u)\, d\lambda_u.$$

Let us for brevity write $F(t)$ in place of $F_\varrho(t; a)$, and $c_n = b^{n+1}a$. Then

$$F(t) = \int [F(t) - F(t-c_n u)]\, k(u)\, d\lambda_u + F_{\varrho_n}(t; a). \tag{8}$$

Let G_n denote the first term on the right; Hölder's inequality gives

$$|G_n(t)|^p \leqslant A \int |F(t) - F(t-c_n u)|^p\, |k(u)|\, d\lambda_u,$$

where A depends only on k. Therefore

$$\int |G_n(t)|^p\, d\lambda \leqslant A \iint |F(t) - F(t-c_n u)|^p\, |k(u)|\, d\lambda_t\, d\lambda_u.$$

Performing first the t-integration on the right, and noting that as $n \to \infty$, $c_n \to 0$, and consequently $\int |F(t) - F(t-c_n u)|^p\, |k(u)|\, d\lambda_t \to 0$ for each fixed u, we see by dominated convergence that $\|G_n\|_p \to 0$. Now, from (8), $\|F\|_p \leqslant \|G_n\|_p + \|F_{\varrho_n}(\ ; a)\|_p$, i.e.,

$$D_{\varrho, p}(f; a) \leqslant \|G_n\|_p + D_{\varrho_n, p}(f; a)$$

which implies (7). Thus Lemma 13 is proved.

We can now extend all the theorems of the preceding sections from C to L^p. One has only to replace C by L^p and D_σ by $D_{\sigma, p}$ in each theorem.

4. Concluding remarks

4.1. It is natural to ask whether one can prove similar results for other spaces of functions. For example, suppose X is a linear topological space of functions on R^m, which contains the exponentials $\{e^{-ix \cdot t}\}_{x \in R^m}$, and the linear span of these is dense in X.

Suppose also that $f(t-au) \in X$ (as a function of u) whenever $f \in X$. If σ is a continuous linear functional on X, and the function $\sigma_u f(t-au) = F(t; a)$ belongs to L^p, we may define the σ, p deviation of f to be the L^p norm of F. Since $\hat{\sigma}(x) = \sigma_t(e^{-ix\cdot t})$ can be defined in this situation we may ask whether Theorem 1 holds or not. Preliminary investigations show that for certain choices of X (e.g., $X = C^2$ with a quite natural topology) Theorem 1 is no longer true. On the other hand, the much more trivial Lemma 2 (i.e., the case where $\hat{\varrho}$ is *globally*, and not merely locally, equal to a linear form in the $\hat{\sigma}_i$ with elements of W as coefficients) is widely extendible. This point of view leads to simple proofs of Jackson's Theorem in higher dimensions, Sobolyev's inequality, certain saturation theorems, etc. These results shall be presented elsewhere.

4.2. *Acknowledgements.* The author wishes to express his deep thanks to Jan Boman for many valuable criticisms of an earlier draft of the present paper. Theorems 1 and 2 are due to Boman, and replace earlier, weaker theorems of the author (which, however, suffice for the known applications). We also wish to thank D. L. Ragozin for several valuable comments.

References

[1]. Boman, J., Partial regularity of mappings between Euclidean spaces. *Acta Math.* 119 (1967), 1–25.
[2]. Lorentz, G. G., *Approximation of functions.* Holt, Rinehart and Winston, 1966.
[3]. Rudin, W., *Fourier analysis on groups.* Interscience, New York, 1962.
[4]. Shapiro, H. S., Some Tauberian theorems with applications to approximation theory. To appear in *Bull. Amer. Math. Soc.*

Received December 1, 1967