SOME TAUBERIAN THEOREMS WITH APPLICATIONS TO APPROXIMATION THEORY

BY HAROLD S. SHAPIRO

Communicated by L. Cesari, November 29, 1967

In this note we state some results, mostly without proof, concerning the comparison of integral inequalities of a certain type. As applications of these results we can deduce from a unified source theorems of approximation theory due to Jackson, Bernstein, and Zygmund, comparison theorems for moduli of smoothness of different orders, and theorems of Hardy and Littlewood and of Zygmund concerning harmonic functions. Moreover, our results yield "inverse theorems" for arbitrary integral kernels which even for many classical kernels (Fejér's, etc.) seem not to be known; thus we complete in some essential respects a program outlined by P. L. Butzer in a series of papers [1], [2], [3] (see esp. [1, p. 95]).

A preliminary version of this work giving further details may be found in the author's mimeographed lecture notes [5].

1.1. We denote by $\mathbb{R}$ the real line, by $C$ the class of functions real, bounded and uniformly continuous on $\mathbb{R}$, and by $\sigma$ a real finite signed measure on $\mathbb{R}$ satisfying $\sigma(\mathbb{R}) = 0$. Define for $f \in C$ and $u \geq 0$

$$D_\sigma(f; u) = \sup_{t \in \mathbb{R}} \left| \int f(t - u)v \, d\sigma(v) \right|$$

and for $t \geq 0$,

$$\omega_\sigma(f; t) = \sup_{0 \leq s \leq t} D_\sigma(f; u).$$

It is easily seen that $D_\sigma(f; u)$ (and hence also $\omega_\sigma(f; u)$) tends to zero as $u \to 0$. The primary purpose of this note is to compare (for fixed $f$) the rate of decrease of these functions for different choices of $\sigma$. We call $D_\sigma, \omega_\sigma$ the $\sigma$-deviation and $\sigma$-modulus of $f$, respectively.

1.2. Examples. (i) Take for $\sigma$ the "binomial measure" $\beta_r$ where $r$ is a positive integer, i.e. $\beta_r$ is the discrete measure with mass $(-1)^n \binom{r}{n}$ at the point $n$ ($n = 0, 1, \ldots, r$). The $\sigma$-modulus is then the modulus of smoothness of order $r$ of the function $f$ ([4, p. 47]). We write $\omega_r$ in place of $\omega_{\beta_r}; \omega_1$ is also called the modulus of continuity of $f$.

---

1 For full details see the author's forthcoming paper A Tauberian theorem related to approximation theory in Acta Math. There a discussion of the $L^p$- and many-variable cases is also given.

1 $C_{r,n}$ here denotes $r!/(n!(r-n)!)$.
(ii) Let \( K \in L^1(R) \), \( \int Kdv = 1 \) (\( dv \) = Lebesgue measure on \( R \)) and
\[
\frac{dH}{dv} = Kdv
\]
where \( dH \) denotes a unit mass at \( t = 0 \) ("Dirac measure"). In this case \( D_{e_\lambda}(f; u) = \|f - f^\lambda\| \), where \( \lambda = 1/\mu \), the norm denotes supremum over \( R \), and \( f^\lambda \) is the "singular" convolution integral
\[
f^\lambda(t) = \int f\left(t - \frac{v}{\lambda}\right)K(v)dv = \int f(t - v)\lambda K(\lambda v)dv.
\]
Thus, \( D_{e_\lambda} \) denotes the "error of approximation" (measured in the sup norm) to \( f \) by a convolution integral depending in a familiar way upon a ("large") parameter \( \lambda \).

1.3. Further conventions concerning terminology and notation. In what follows, we denote by \( W \) the normed ring of Fourier transforms \( S = \mathcal{F} \) of finite measures \( \sigma \) on \( R \). By "\( S_1 \) divides \( S_2 \)" for \( S_1 \in W \) we mean as usual that there is some \( S \in W \) such that \( S_2 = S_1S \), and by "\( S_1 \) divides \( S_2 \) at \( x_0 \)" we mean that there is an \( S \in W \) such that \( S_2(x) = S_1(x)S(x) \) for all \( x \) in some neighborhood of \( x_0 \). For \( S = \mathcal{F} \in W \) we write \( \|S\|_W \) to denote the total variation of \( \sigma \). A kernel means an integrable function \( K \) on \( R \) such that \( \int Kdv = 1 \). By \( K_\lambda \) we always denote the function defined by \( K_\lambda(t) = \lambda K(\lambda t) \). By \( * \) we denote the usual convolution product in \( L^1(R) \), and by a slight abuse of notation we write \( \hat{K} \) to denote the Fourier transform of the measure \( Kdv \), i.e. the ordinary Fourier transform of \( K \). A norm symbol not subscripted by \( W \) refers always to the \( C \)-norm, i.e. the supremum over \( R \).

2.1. We begin with some rather trivial but none the less useful results, which also serve to motivate the later theorems.

**Theorem 1.** If \( \mathcal{F}_1 \) divides \( \mathcal{F}_2 \), say \( \mathcal{F}_2 = \mathcal{F}_1S \), and \( A = \|S\|_W \), we have for all \( f \in C \)
\[
D_{e_\lambda}(f; u) \leq AD_{e_\lambda}(f; u).
\]
From this theorem we get (in view of the discussion in 2.1):

**Theorem 2.** Let \( J, K \) be kernels such that \( 1 - J \) divides \( 1 - \hat{K} \) in \( W \). Then, for every \( f \in C \) and every \( \lambda > 0 \)
\[
\|f - (f * K_\lambda)\| \leq A\|f - (f * J_\lambda)\|
\]
where \( A \) is a constant depending only on the kernels \( J, K \).

**Theorem 3.** If \( J, K \) denote the Cauchy and Fejér kernels (see [1])
\[
A_1\|f - (f * K_\lambda)\| \leq \|f - (f * J_\lambda)\| \leq A_2\|f - (f * K_\lambda)\|
\]
where $A_1$ and $A_2$ are numerical constants independent of $\lambda$ and $f$.

(3) may be said to express the equivalence of the Cauchy and Fejer kernels from the standpoint of sup-norm approximation theory.

2.2. The main theorems. If $\delta_1$ does not divide $\delta_2$ we cannot in general expect an inequality of the type (1). We can, however, establish a weaker inequality under the weaker hypothesis of divisibility at $x = 0$.

**Theorem 4.** If $\delta_1$ divides $\delta_2$ at $x = 0$, then

$$D_{\nu}(f; u) \leq A \sum_{t=0}^{\infty} D_{\nu}(f; B\theta u)$$

where $A$, $B$ are constants depending only on $\sigma_1$, $\sigma_2$ and $\theta < 1$ is a positive constant depending only on $\sigma_1$.

**Corollary.** If $\delta_1$ divides $\delta_2$ at $x = 0$, and $f$ has a $\sigma_1$-modulus which is $O(u^\alpha)$ for some $\alpha > 0$, then the $\sigma_2$-modulus is also $O(u^\alpha)$.

**Example.** Choose $\sigma_1 = \beta_2$ (see §1.2 above), $d\sigma_2 = dH - Ldv$, where

$$L(t) = \pi^{-1/2}e^{-t^2}$$

(Weierstrass kernel).

Here $\delta_1(x) = (1 - e^{-ix})^2$, $\delta_2(x) = 1 - e^{-x^2/4}$. It is readily checked that $\delta_1$ divides $\delta_2$ at $x = 0$ (although not globally) hence by the Corollary we obtain the familiar result: $\omega_2(f; u) = O(u^\alpha)$ implies $\|f - (f * L)\| = O(\lambda^{-\alpha})$. Since in this case $\delta_2$ divides $\delta_1$ (globally) the implication in the reverse direction is also true by Theorem 1; this gives the "saturation class" for the Weierstrass kernel (see [1]).

**Theorem 5.** Let $\sigma$ be any nonnull measure, $\alpha > 0$ and suppose for some $f \in C, D_\nu(f; u) \leq A u^\alpha$ holds for $u > 0$, where $A$ is independent of $u$. Then for every positive integer $r$, if $\rho$ is a measure such that, $\beta(x)/x^r$ coincides with an element of $W$ in some neighborhood of 0,

$$D_\nu(f; t) \leq \left( \frac{B_1}{\alpha - r} \right) A t^r \quad \text{if} \quad r < \alpha,$$

$$\leq B_2 \left( \frac{r + \alpha + 1}{\alpha} \right) t^\alpha (A | \log t | + 1) \quad \text{if} \quad r = \alpha,$$

$$\leq \left( B_3^a + \frac{B_4}{r - \alpha} \right) A t^\alpha \quad \text{if} \quad r > \alpha,$$

where the $B_i$ are constants depending only on $\sigma$ and $\rho$.  

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use
EXAMPLE. Specializing $\sigma$ and $\rho$ to be the measures $\beta_m$ and $\beta_n$ of §1.2 gives a comparison theorem for moduli of smoothness of different orders, e.g. $\omega_m(t) = O(t)$ implies $\omega_1(t) = O(t \log (1/t))$ (case $r = \alpha = 1$).

Choosing $d\sigma = dH - K \mu$ (see 1.2) and $\rho = \beta$, gives

**Theorem 6.** Let $K$ be any $L^1$ kernel, $f \in C$ and $||f - (f \ast K)\| = O(\lambda^{-\alpha})$, then $\omega_r(f; t)$ is large $O$ of $t^r$, $t^r \log(1/t)$ or $t^r$ according as $r < \alpha$, $r = \alpha$, or $r > \alpha$.

**Remark.** We emphasize that in Theorem 6 there is no hypothesis whatever about the kernel. Even for various special kernels (e.g. Fejér’s) there seem to be almost no “inverse” theorems in the literature of the type of Theorem 6, except when the assumed degree of approximation is the “saturation” order (see [1]).

It is easy to deduce from Theorem 5 also

**Theorem 7.** Suppose $\beta > 0$, $m$ is a positive integer larger than $\beta$, and $K$ an $m$ times differentiable kernel with $K^{(m)} \in L^1 (\mathbb{R})$. If $f \in C$ and the sup norm of the $m$th derivative of $f \ast K$ is $O(\lambda^m)$, then $\omega_r(f; t)$ is large $O$ of $t^r$, $t^r \log(1/t)$, or $t^{m-\beta}$ according as $r$ is less than, equal to, or greater than $m-\beta$.

**Examples.** Choosing $K$ to be the Cauchy kernel, $m = 1$, $r = 1$ gives (re-interpreting $f \ast K$ as a harmonic function in the half-plane $y > 0$ where $y = 1/\lambda$): If $U(t, y)$ is continuous for $y \geq 0$ and harmonic for $y > 0$, and $\partial U/\partial t$ is $O(y^{-\beta})$ uniformly for $t \in \mathbb{R}$, then $U(t, 0)$ satisfies a Hölder condition of order $1 - \beta$ (Hardy and Littlewood, cf. [6, p. 263]). The choice $m = 2$, $\beta = 1$, $r = 2$ gives: If $\partial^2 U/\partial t^2$ is $O(y^{-1})$ uniformly for $t \in \mathbb{R}$, then $f(t) = U(t, 0)$ satisfies $\omega_2(f; t) = O(t)$ (this theorem, or rather the analogous version for functions harmonic in a circle, is due to Zygmund [6, p. 263]. Previous proofs used special properties of harmonic functions.}

3. It is easy to deduce Jackson’s Theorem (in the form given in [4, p. 58]) from the corollary to Theorem 4. The deduction is based on

**Lemma.** Let $K$ be a kernel such that $K(x) = 0$ for $|x| \geq 1$. If $f$ has period $2\pi$, then $f \ast K$ is a trigonometric polynomial of degree less than $\lambda$.

Now, given $r$, it is easy to construct a kernel $K$ such that (i) $K$ vanishes for $|x| \geq 1$, and (ii) $(1 - e^{-ix})^r$ divides $1 - K$ at $x = 0$. With such choice of $K$, $f \ast K_n = T_n f$ is a trigonometric polynomial of degree less than $n$. If now for some $\alpha > 0$, $\omega_r(f; t) = O(t^\alpha)$, then by the corollary to Theorem 4, $||f - T_n||$ is $O(n^{-\alpha})$; from this result the Jackson Theorem in full generality is easily deduced.
Many classical "inverse theorems" of approximation theory are deducible from Theorem 5. Suppose, for instance, there exists for each \( n \) a trigonometric polynomial \( T_n \) of degree at most \( n - 1 \) with \( \| f - T_n \| \leq A/n \). Let \( \sigma \) be any nonnull measure such that \( \sigma(x) \) vanishes for \( |x| \leq 1 \). Writing \( T_n(t) = \sum_{|k| < n} c_k e^{ikt} \), we get
\[
\int f(t - uv) d\sigma(v) = \int [f(t - uv) - T_n(t - uv)] d\sigma(v) + \sum_{|k| < n} c_k e^{ikt} \sigma(ku).
\]
Now, given \( u > 0 \) fix \( n \) so that \( 1/u^n < 1/u + 1 \). Then the second term on the right vanishes, giving
\[
\left| \int f(t - uv) d\sigma(v) \right| \leq \| f - T_n \| \cdot \int |d\sigma| \leq \frac{B}{n} \leq Bu
\]
where \( B = A \int |d\sigma| \). Applying Theorem 5 with \( \alpha = 1, r = 2, \rho = \beta_2 \), gives \( \omega_2(f; t) = O(t) \). This is the celebrated inverse theorem of Zygmund [4, p. 61, Theorem 7].

4. Generalizations. The above theorems may be generalized to Euclidean \( m \)-space and complex measures (then, however, a certain "Tauberian condition" must be imposed on \( \sigma \), also to \( L^p \) norms.\(^8\)

REFERENCES


\(^8\) The more trivial Theorem 1 generalizes also to distributions. This leads to a particularly illuminating proof of Jackson's theorem.