# REGULARITY PROPERTIES OF THE ELEMENT OF CLOSEST APPROXIMATION

BY

#### HAROLD S. SHAPIRO

ABSTRACT. Given an element  $f \in L^p(T)$ , 1 , and a closed translation invariant subspace <math>S of  $L^p(T)$ , we investigate the regularity (smoothness) properties of the element of S which is closest to f. The regularity of this element is in general less than that of f. The problem reveals a surprising connection with a hitherto unstudied class of extremal Fourier multipliers.

1. Introduction. Although there is a vast literature on the subject of best approximation, comprising both general results and detailed studies for concrete spaces of functions, we could not find any results of the following kind: suppose S is some set of functions and  $f_0$  the best approximation from S (in some specified sense) to f; can we infer, if f has some "regularity" properties (such as differentiablity, etc.), that these are inherited by  $f_0$ ? The motivation for studying such a problem is the empirical "principle" that the solution of an extremal problem (not necessarily one of closest approximation) tends to be "regular" if the data determining the problem are. Of course, in any attempt to embody this idea in definite theorems, very strong restrictions will have to be imposed on the nature of the extremal problem, as immediate counterexamples show. Nonetheless, it seems not without interest to study some special situations where the "principle" can be rigorously formulated and its validity, as well as its limitations, established.

In the present paper I have mainly considered the problem of best approximation, in the  $L^p$  metric, of a given function by elements of a closed convex subset S of the  $2\pi$ -periodic functions. In order to guarantee existence and uniqueness of the best approximation I suppose  $1 . Actually, I work entirely in the range <math>2 \le p < \infty$ , but this is only to avoid duplication of proofs, since the main tool (Clarkson's inequality) takes a different form for p < 2 than for p > 2. The main positive result is Corollary 2 to Theorem 1 below, showing some inheritance of regularity, under the drastic assumption that S contains, with each function,

Received by the editors February 8, 1972.

AMS (MOS) subject classifications (1970). Primary 41A50, 42A08, 42A18.

Key words and phrases. Best approximation,  $L^p$  modulus of continuity, regularity, metric projection, translation invariant space, Fourier multiplier.

all its translates. This restriction is maintained throughout the paper, except for §6.3. While perhaps ad hoc, it is satisfied in many cases of interest; moreover it guarantees that S is large enough so that, on the one hand, it contains some "regular" functions (without which our problem would be senseless) and (except when S is finite dimensional) some very "irregular" ones (without which our problem would be trivial). In Theorem 2 we show that some degree of regularity can actually be lost in passing to the best approximation. In Theorem 3, where now S is assumed to be a subspace, it is pointed out that a very strong positive result holds when p = 2. Here the mapping onto the closest element is linear, and we make contact with well studied problems. § 5 treats a problem of pure Fourier analysis that arose in the present work, whereby Theorems 1 and 2 find an unexpected application. §6 contains variants (Theorem 6 deals with regularity of the minimal solution to a moment problem), generalizations, and open questions. Of the latter there is certainly no shortage in this circle of ideas, and it is hoped that the present study, by virture of its very incompleteness, may inspire others to provide more answers.

Notations. By **R** we denote the real line; by **Z** the integers, and  $T = R/2\pi Z$  (circle group). By  $L^p(T)$  we denote, as usual, the space of complex-valued  $2\pi$ -periodic functions f such that

$$||f||_{p} = \left(\left(\frac{1}{2\pi}\right)\int_{0}^{2\pi}|f(t)|^{p} dt\right)^{1/p}$$

is finite. For  $f \in L^p(\mathbb{R})$  we use the same symbol  $\|f\|_p$  to denote the usual norm in  $L^p(\mathbb{R})$ , which should cause no confusion. We also use the now standard "roof" symbol  $\hat{}$  to denote Fourier transformation.

Since we are concerned with questions of regularity (smoothness) a word about the relevant function classes is in order. Denoting by  $T_u$  the operator of translation by u,  $(T_u f)(t) = f(t-u)$ , we denote by  $\widetilde{\Lambda}^p_\alpha$ , for  $0 < \alpha \le 2$ , the set of  $f \in L^p$  such that

where  $\Delta_a$  denotes  $T_{-a}$  minus the identity (difference operator).

In  $\S 4$ , other closely related smoothness classes  $\Lambda^p_\alpha$  will be defined; we may remark that  $\Lambda^p_\alpha$  and  $\widetilde{\Lambda}^p_\alpha$  coincide for  $0<\alpha<2$ . For  $0<\alpha<1$ ,  $\widetilde{\Lambda}^p_\alpha$  coincides with the space denoted by  $\Lambda^p_\alpha$  in [18, p. 45] (i.e. functions whose  $L^p$  modulus of continuity is  $O(a^\alpha)$ ); whereas our  $\widetilde{\Lambda}^p_1$  is Zygmund's  $\Lambda^p_*$ , and  $\widetilde{\Lambda}^p_2$  is the class of twice integrated  $L^p$  functions.

Finally, we have allowed ourselves the literary license to use the terms, "best approximation", and "closest element", to designate the same thing. The

metric projection of f on S denotes the mapping from f to the set of best approximations in S (which in our case will always be a set consisting of one element); by abuse of language, the element of best approximation is itself sometimes also called the metric projection of f on S.

Acknowledgements. The deduction given of Corollary 2 from Theorem 1 is somewhat simpler than in our original presentation, thanks to a useful comment by Hermann Rost. Mlle. Martine Joublin kindly pointed out an error in the original form of Corollary 3. I also wish to thank B. O. Björnestål for valuable comments.

## 2. Metric projection on certain convex sets in $L^p$ .

**Theorem 1.** Let S denote a closed convex subset of  $L^p(T)$ ,  $2 \le p < \infty$ . Let  $f \in L^p(T)$ , and denote by  $f_0$  the closest element to f in S. Then, if U is any linear operator on  $L^p(T)$  having norm 1 and mapping S into S,

(2.1) 
$$||f_0 - Uf_0||^p \le 2^{p-1} [(d + ||f - Uf||)^p - d^p]$$

where d denotes  $||f - f_0||$ .

Before proving Theorem 1, we turn to some corollaries thereof.

Corollary 1. Under the hypotheses of Theorem 1,

$$||f_0 - Uf_0|| \le 6 d^{1 - (1/p)} ||f - Uf||^{1/p}$$

*provided* ||f - Uf|| < d/(p-1).

To deduce (2.2) from (2.1), we apply the elementary inequality  $(1+x)^p \le 1 + 3px$ , for  $0 \le x \le 1/(p-1)$ . Writing x = ||f - Uf||/d, and assuming  $x \le 1/(p-1)$ , we deduce from (2.1),

$$||f_0 - Uf_0|| \le 2^{1-(1/p)} 3^{1/p} p^{1/p} d^{1-(1/p)} ||f - Uf||^{1/p}$$

which implies (2.2).

We now state our main result:

Corollary 2. Let S denote a closed convex subset of  $L^p(T)$ ,  $2 \le p < \infty$ , which is translation invariant (that is,  $g \in S$  implies  $T_u g \in S$  for all real u). Suppose  $0 < \alpha \le 2$ , and f is an element of  $L^p(T)$  which belongs to  $\widetilde{\Lambda}^p$ . Then, the element of S which lies closest to f belongs to  $\widetilde{\Lambda}^p_{\alpha/p}$ .

Remarks. I do not know whether the number  $\alpha/p$  appearing in the conclusion is the largest possible (when p=2 it is not, see §4). However, examples given below will show that, even when S is a subspace, the metric projection may actually have less regularity (as measured by the  $L^p$  Lipschitz condition) than the original element.

To prove Corollary 2, choose  $U = (T_a + T_{-a})/2$ . The hypotheses on S guarantee that  $US \subseteq S$ , so that (2.2) is applicable, and yields, for small a > 0,

$$\|\Delta_a^2 f_0\| \le C_1 \|\Delta_a^2 f\|^{1/p} \le C_2 a^{\alpha/p}.$$

Corollary 3. If, in the hypotheses of the preceding corollary,  $p \le 4$  and  $\alpha > p/2$ , the closest element to f is (after correction on a set of measure zero) continuous and even has an absolutely convergent Fourier series; it satisfies a Lipschitz condition (in the classical, sup norm sense) of order  $\alpha/p - 1/2$ .

Indeed,  $f_0$  belongs to  $\widetilde{\Lambda}^p_{\alpha/p}$  and hence, since p>2, to  $\widetilde{\Lambda}^2_{\alpha/p}$  which implies (by Bernstein's theorem) that  $f_0$  has an absolutely convergent Fourier expansion. Also, by a theorem of Hardy and Littlewood [4],  $f_0 \in \widetilde{\Lambda}^\infty_\beta$  with  $\beta = \alpha/p - 1/2$ .

**Proof of Theorem 1.** We shall require the 'parallelogram inequality' of Clarkson [2]:

valid for any x, y in  $L^p$ . Since U has norm 1,  $||Uf - Uf_0|| \le d$ , so that

$$||f - Uf_0|| \le d + ||f - Uf||.$$

Now, applying (2.3) with  $x = f - f_0$  and  $y = f - Uf_0$ ,

$$\|f - (1/2)(f_0 + Uf_0)\|^p + \|(f_0 - Uf_0)/2\|^p \le (d^p + \|f - Uf_0\|^p)/2$$

and since the hypotheses imply that  $(1/2)(f_0 + Uf_0)$  is an element of S, the first term on the left is not less than  $d^p$ , since d is the distance from f to S. Hence

$$\|(f_0 - Uf_0)/2\|^p \le (1/2)(\|f - Uf_0\|^p - d^p) \le (1/2)[(d + \|f - Uf\|)^p - d^p]$$

by virtue of (2.4), and so (2.1) is established.

Remark. An almost identical argument establishes that the metric projection operator  $P_S$  onto S is Hölder continuous, with Hölder exponent 1/p, at each point of  $L^p$ . (This in turn implies Corollary 2 when  $\alpha < 1$ .) This is valid in all  $L^p$ -spaces, being based only on the Clarkson inequalities. For 1 one obtains (as B. O. Björnestål pointed out to me) the Hölder exponent <math>1/2. It is not known to me whether those exponents are improvable (except that for p = 2,  $P_S$  is a contraction, cf. §4). It is a consequence of the result in §3 that when p > 2 the Hölder exponent (for a suitable  $S \subset L^p$  (T) of codimension 2) cannot exceed (1/p) + 1/(p-1).

## 3. An example exhibiting loss of regularity.

**Theorem 2.** Let  $2 , and denote by S the closed subspace of <math>L^p(T)$  spanned by the exponentials  $\{e^{int}\}, n \in \mathbb{Z}\setminus\{0, 1\}$ . There exists a trigonometric

polynomial f (in fact, of the form  $f(t) = 1 + ce^{it}$ ) whose metric projection on S is in  $\widetilde{\Lambda}^p_{\beta}$ , where  $\beta = p^{-1} + (p-1)^{-1}$ , but not in  $\widetilde{\Lambda}^p_{\alpha}$  for any  $\alpha > \beta$ .

Remark. In particular,  $f \in \widetilde{\Lambda}_2^p$ ; since  $\beta < 2$ , we have thus an example to show that the number  $\alpha/p$  in Corollary 2 above certainly cannot be replaced by  $\alpha$ . Observe that  $p^{-1} + (p-1)^{-1} > 2p^{-1}$ , so the present example leaves unsettled the question of the sharpness of Corollary 2.

**Proof of Theorem 2.** Let us recall that an element g of a Banach space B is said to be *orthogonal* to the subspace X if  $\|g - b\| \ge \|g\|$  whenever  $b \in X$ . It is easy to show that the function

$$g(t) = (|1 - e^{it}|^{p'})/(1 - e^{it})$$

where p' denotes p/(p-1), is orthogonal to S in  $L^p(T)$ . (As an almost identical proposition is established in [14], we refer to that paper for details.) Consider now the function  $f(t) = \hat{g}(0) + \hat{g}(1)e^{it}$ ; observe that  $f_0 = f - g$  lies in S, and since  $f - f_0$  is orthogonal to S,  $f_0$  is the closest element to f in S. To complete the proof, it is enough to establish, introducing the notation

$$D(F, a) = \left(\left(\frac{1}{2\pi}\right) \int_{-\pi}^{\pi} |F(t+a) - 2F(t) + F(t-a)|^{p} dt\right)^{1/p},$$

that

(3.1) 
$$D(f_0, a) = Ca^{\beta} + O(a^c), \quad a \to 0+,$$

holds for suitable constants C, c with  $c > \beta$ .

Before going into the details of the proof, we record two simple remarks:

- (i)  $D(F + G, a) \leq D(F, a) + D(G, a)$ , by virtue of Minkowski's inequality.
- (ii)  $D(F, a) = O(a^2)$  whenever F is twice continuously differentiable on  $[-3\pi/2, 3\pi/2]$  (note that we may suppose always  $0 \le a \le \pi/2$ ).

Since  $f_0$  differs from g be a trigonometric polynomial, it is enough, in view of the foregoing remarks, to prove (3.1) with  $f_0$  replaced by g. Now, we may write  $g = g_1 + g_2 + g_3$  where

$$g_1(t) = -|1 - e^{it}|^{p'}/it$$
,  $g_2(t) = (1/2)|1 - e^{it}|^{p'}$ ,   
 $g_3(t) = |1 - e^{it}|^{p'}[(1 - e^{it})^{-1} + (it)^{-1} - (1/2)]$ .

Because p' > 1, and the bracketed term in the last equation is twice differentiable and vanishes at t = 0,  $g_3 \in C^2$ . Moreover,

$$g_1(t) = -(|t|^{p'}/it) + g_4(t), \qquad g_2(t) = (1/2)|t|^{p'} + g_5(t),$$

where

$$g_{\delta}(t) = (|t|^{p'} - |1 - e^{it}|^{p'})/it$$
 and  $g_{5}(t) = (1/2)(|1 - e^{it}|^{p'} - |t|^{p'})$ 

are twice continuously differentiable (we leave to the reader the verification of these facts). Therefore, the proof will be completed if we establish the estimates

(3.2) 
$$D(G, a) = Ca^{\beta}(1 + O(a^{\lambda_1})),$$

$$D(H, a) = O(a^{\beta + \lambda_2})$$

with  $\lambda_1$  and  $\lambda_2$  positive, where  $G(t) = |t|^{p'}/t$ ,  $H(t) = |t|^{p'}$ . Now,

$$\int_{-\pi}^{\pi} |G(t+a) - 2G(t) + G(t-a)|^{p} dt = a \int_{-\pi/a}^{\pi/a} |G(as+a) - 2G(as) + G(as-a)|^{p} ds$$

$$= a^{1+(p'-1)p} \int_{-\pi/a}^{\pi/a} |G(s+1)-2G(s)+G(s-1)|^p ds,$$

since G is positive-homogeneous of degree p'-1. Now,

$$|G(s+1)-2G(s)+G(s-1)|=O(|s|^{p'-3})$$

for large |s|; therefore the last integral equals

$$\int_{-\infty}^{\infty} |G(s+1) - 2G(s) + G(s-1)|^p ds + O(a^{(3-p')p-1}).$$

Thus we have established

$$\int_{-\pi}^{\pi} |G(t+a) - 2G(t) + G(t-a)|^{p} dt = C_{1}a^{1+(p'-1)p}(1 + O(a^{\lambda_{1}}))$$

where  $C_1$  is a positive absolute constant, and  $\lambda_1 = (3 - p')p - 1 > 0$ , which implies (3.2).

Finally, to prove (3.3) we proceed in a similar fashion:

$$\int_{-\pi}^{\pi} |H(t+a) - 2H(t) + H(t-a)|^p dt = a^{1+pp'} \int_{-\pi/a}^{\pi/a} |H(s+1) - 2H(s) + H(s-1)|^p ds$$
and since

$$|H(s+1)-2H(s)+H(s-1)| < C_2(1+|s|)^{p'-2}$$

the integral on the right is

$$O(a^{(2-p')p-1}), \quad \text{if } (2-p')p < 1,$$

$$O(\log(1/a)), \quad \text{if } (2-p')p \ge 1.$$

In the first case, we obtain the estimate  $D(H, a) = O(a^2)$ , in the second  $D(H, a) = O(a^{p'+(1/p)-\epsilon}) = O(a^{\beta+1-\epsilon})$  for every  $\epsilon > 0$ . This establishes (3.3), and completes the proof of Theorem 2.

4. An extension in the case p=2. Since the hypotheses of our main result, Corollary 2, require  $\alpha \le 2$ , the conclusion can never yield information about  $f_0$ ,

the metric projection of f on S, stronger than its inclusion in  $\widetilde{\Lambda}_{2/p}^{p}$ , not even if f is assumed to possess very high (say,  $C^{\infty}$ ) regularity. In view of this, it is natural to look for stronger hypotheses on S that will enable stronger conclusions about  $f_0$  to be drawn when f has high regularity; one would, for example, like a theorem saying that  $f \in C^{\infty}$  implies  $f_0 \in C^{\infty}$ .

Results of this kind are valid when p is 1, 2, or  $\infty$  and  $S = H^p = \{\hat{f} \in L^p(T): \}$  $\hat{f}(n) = 0$  for n < 0. For example, consider the case  $S = H^{\infty}$  ( $H^{\infty}$  is not uniformly convex, nevertheless, the existence of the metric projection, and its uniqueness in case f is continuous, can be established by ad hoc arguments, see 13, Chapter 8]). Recently Adamyan, Arov and Krein [1], and independently and by other methods Carleson and Jacobs (see [6]) established that  $f \in C^{\infty}$  implies its metric projection  $f_0$  on  $H^{\infty}$  is in  $C^{\infty}$ ; moreover, a number of similar results were established (for instance, if f satisfies a Lipschitz condition of order  $\alpha$ ,  $0 < \alpha < 1$ , so does  $f_0$ , see [6]). Already in my thesis (cf. [11]) I had proved that if  $f(t) = \phi(e^{it})$ , where  $\phi(z)$  is holomorphic in a neighborhood of the point  $e^{it0}$ , then also  $f_0(t) =$  $\psi(e^{it})$  where  $\psi(z)$  is holomorphic in a neighborhood of  $e^{it0}$ . To a certain extent, these results can be extended to  $H^1$ . However, the proofs rely essentially on the theory of holomorphic functions, and shall not be further discussed here, insofar as they lie outside the general framework envisaged in the present paper. case  $S = H^p$ , when p is not equal to 1, 2, or  $\infty$  presents peculiar features, in particular a " $C^{\infty}$  implies  $C^{\infty}$ " theorem does not hold; none the less some refinement of Corollary 2 is possible for this choice of S, see  $\S 6.4$ .

In the case p=2, one can establish easily, by Fourier methods, a very general result. (1) It is, in principle, well known (see the remarks at the end of this section) but its inclusion here is important for purposes of orientation.

To formulate it as compactly as possible, we wish to define in a convenient manner the analogues of the classes  $\widetilde{\Lambda}^p_\alpha$  when  $\alpha > 2$ . Let E denote the set of complex-valued functions  $\phi$  in  $C^\infty(\mathbf{R})$ , rapidly decreasing in the sense of L. Schwartz, and such that

- (i)  $\hat{\phi}(x)$  vanishes for all x in some neighborhood of 0,
- (ii)  $\hat{\phi}$  has compact support,
- (iii) there exist  $x_1 < 0$  and  $x_2 > 0$  such that  $\hat{\phi}(x_1) \neq 0$ ,  $\hat{\phi}(x_2) \neq 0$ . Denoting by  $\phi_{(a)}$  the function whose value at t is  $(1/a)\phi(t/a)$ , where a is a positive parameter, we define for every  $\alpha > 0$ , and  $1 \leq p \leq \infty$ ,

<sup>(1)</sup> Let us also recall the well-known fact that if S is a closed convex subset of a Hilbert space, the metric projection operator  $P = P_S$  onto S satisfies  $\|Px - Py\| \le \|x - y\|$  (cf. R. B. Holmes, A course on optimization and best approximation, Lecture Notes in Math., vol. 257, Springer-Verlag, Berlin and New York, 1972, p. 157. An immediate consequence is that in the conclusion of Corollary 2 (in the case p = 2) we may replace  $\alpha/2$  by  $\alpha$  provided  $\alpha < 1$ .

$$\Lambda_{\alpha}^{p} = \{ f \in L^{p}(\mathbb{T}) : \| f * \phi_{(a)} \|_{p} = O(a^{\alpha}), \ a \longrightarrow 0 + \}.$$

Here the \* denotes convolution over R, which is meaningful since f has a periodic extension to R. Or, more simply,  $f * \phi_{(a)}$  is simply the trigonometric polynomial  $\sum_{n=-\infty}^{\infty} \hat{\phi}(na) \hat{f}(n) e^{int}$ . (Note: this definition need not be restricted to  $L^p(T)$ ; the analogous classes  $\Lambda_a$  can also be defined for any "homogeneous Banach space" of functions on  $R^n$ , cf. [13, Chapter 9].) The following propositions are consequences of results in [13, Chapter 9]:

- (a) The class  $\Lambda^p_\alpha$  is independent of the particular choice of  $\phi \in E$  used in its definition.
- (b)  $\Lambda_{\alpha}^{p}$  coincides with the class of functions which we denoted by  $\widetilde{\Lambda}_{\alpha}^{p}$  in the Introduction, if  $0 < \alpha < 2$ .
- (c) If  $\alpha = n + \beta$  where n is a positive integer and  $0 < \beta \le 1$ ,  $\Lambda^p_\alpha$  coincides with the set of functions  $f \in L^p$  such that  $f^{(n)} \in \Lambda^p_\beta$ .
  - (d)  $C^{\infty} = \bigcap_{\alpha > 0} \Lambda_{\alpha}^{p}$ .
  - (e)  $\Lambda_a^{\infty}$  is the classical space Lip  $\alpha$ , for  $0 < \alpha < 1$ .

As is clearly necessary with regularity classes defined in terms of integrals, (d) and (e) (and several statements made elsewhere in this paper) must be understood modulo the usual convention in distribution theory, e.g. in (d) it is meant that an element in  $\bigcap_{\alpha>0} \Lambda^p_\alpha$ , after eventual correction on a set of measure zero, belongs to  $C^\infty$ .

**Theorem 3.** Let S be any closed, translation-invariant subspace of  $L^2(T)$ . If  $f \in L^2(T)$  belongs, for some  $\alpha > 0$ , to  $\Lambda^2_{\alpha}$ , then so does  $f_0$ , the metric projection (which coincides with the orthogonal projection) of  $f_0$  on S.

Corollary 4. With the hypotheses of the preceding theorem, if  $f \in C^{\infty}(T)$  then  $f_0 \in C^{\infty}(T)$ .

That the corollary follows from the theorem is clear, by virtue of (d) above. The corollary is in any case rather trivial (nor would we wish to claim that the theorem is much deeper!). For, as is well known, every closed translation-invariant subspace of  $L^p(T)$ , when  $1 \le p < \infty$ , has the form  $S(M) = \{f \in L^p(T): \hat{f}(n) = 0 \text{ whenever } n \in M\}$  where M is some subset of the integers. Now,  $f \in C^{\infty}(T)$  is equivalent to

(3.1) 
$$\hat{f}(n) = O(|n|^{-\alpha}), \quad |n| \to \infty, \quad \text{all } \alpha > 0.$$

Since  $f_0$ , the orthogonal projection of f on S(M), satisfies

$$\hat{f_0}(n) = \begin{cases} f(n), & n \in M, \\ 0, & n \notin M, \end{cases}$$

it is clear that  $f_0$  satisfies (3.1) whenever f does,

**Proof of Theorem 3.** We have, recalling that S = S(M) for a suitable  $M \subseteq \mathbb{Z}$ ,

$$(f*\phi_{(a)})(t) = \sum \hat{\phi}(an)\hat{f}(n) e^{int}, \qquad (f_0*\phi_{(a)})(t) = \sum_{n \in M} \hat{\phi}(an)\hat{f}(n) e^{int}$$

and it is now immediate, by Parseval's formula, that  $\|f_0 * \phi_{(a)}\|_2 \le \|f * \phi_{(a)}\|_2$ , which implies the conclusion of the theorem.

We also obtain a result if we measure regularity with the aid of sup norms:

Corollary 5. With S as in Theorem 3, if  $f \in \Lambda_{\alpha}^{\infty}$  and  $\alpha > 1/2$ , the orthogonal projection f on S in the Hilbert space  $L^2(T)$  is in  $\Lambda_{\alpha-(1/2)}^{\infty}$ .

Corollary 5 is an immediate consequence of the inclusions  $\Lambda_{\alpha}^{\infty} \subset \Lambda_{\alpha}^{2} \subset \Lambda_{\alpha-(1/2)}^{\infty}$ , the first of which is trivial, and the second of which follows from the theorem of Hardy and Littlewood referred to earlier. Indeed, if  $f \in \Lambda_{\alpha}^{\infty}$ , then  $f \in \Lambda_{\alpha}^{2}$ , hence  $f_0 \in \Lambda_{\alpha}^{2}$  (by Theorem 3), hence  $f_0 \in \Lambda_{\alpha-(1/2)}^{\infty}$ .

In the conclusion of Corollary 5, the 1/2 cannot be replaced by a smaller number. This may be seen from the following example, kindly pointed out to me by Y. Katznelson, and which I include here with his permission. As is well known, there exists a sequence  $\{a_n\}_{n=0}^{\infty}$ , each element of which is 1 or -1, and such that  $|\sum_{n=0}^{m} a_n e^{int}|$  has sup norm  $\leq C(m+1)^{1/2}$  for every m, where C is an absolute constant [7, p. 34]. Now, using this, it is possible to deduce that  $f(t) = \sum_{n=0}^{\infty} (n+1)^{-1} a_n e^{int}$  belongs to  $\Lambda_{1/2}^{\infty}$ . Let

$$M_{+} = \{n : a_{n} = 1\}, \qquad M_{-} = \{n : a_{n} = -1\},$$

and denote by  $f_+$ ,  $f_-$  respectively the orthogonal projections, in  $L^2(T)$ , of f on  $S(M_+)$ ,  $S(M_-)$ . Now, at least one of the functions  $f_+$ ,  $f_-$  is unbounded, since  $f_+ - f_-$  has the Fourier series  $\sum_{n=0}^{\infty} (n+1)^{-1} e^{int}$ .

In concluding this section, I wish to point out that Theorem 3 is in principle well known; indeed, with suitable identifications it is a very special case of Lemma 1 on p. 825 of Taibleson [16]. It is also closely connected in spirit with the circle of problems treated in Zygmund [17], that is, with sequences (in our case, arising from orthogonal projection, these sequences consisted only of the numbers 0 and 1) which multiply the Fourier coefficients of every function in  $\Lambda^2_{\alpha}$  into the Fourier coefficients of some function in  $\Lambda^r_{\beta}$ . This problem is in essence completely solved (another treatment of it, in terms of the definition of  $\Lambda^p_{\alpha}$  given in this section, will appear in a forthcoming paper by Nestor Rivière). However, there seems little reason to hope that these results can in any way carry over to the (nonlinear) metric projections on translation-invariant subspaces of  $L^p$  when  $p \neq 2$ .

5. Relation to Fourier multipliers and saturation. Throughout this section S shall denote a closed translation-invariant subspace of  $L^p(T)$ . By combining Theorem 1 with the counterexample constructed in  $\S 3$ , we shall deduce a result

concerning convolution operators in  $L^p$ , that can also be interpreted as a saturation theorem.

Let  $k \in L^1(\mathbb{R})$ , and denote by  $k_{(a)}$  the function

$$k_{(a)}(t) = (1/a)k(t/a)$$
.

It is well known (see for instance, [13, Chapter 9] that for  $1 \le p \le \infty$  the map (5.1)  $f \mapsto k * f$ 

where

$$(k * f)(t) = \int_{-\infty}^{\infty} k(t - u)f(u) du$$

maps  $L^p(\mathbf{R})$  into  $L^p(\mathbf{R})$ , and also  $L^p(\mathbf{T})$  into  $L^p(\mathbf{T})$  (in the latter case, we interpret the elements of  $L^p(\mathbf{T})$  as functions on  $\mathbf{R}$  of period  $2\pi$ ); moreover, in each case the bound of the operator does not exceed  $\|k\|_1$ .

**Definition.** For  $1 \le p \le \infty$ ,  $V_p$  is the set of  $k \in L^1(\mathbb{R})$  such that

$$\int k(t) dt = 1$$

and the norm of the operator (5.1), from  $L^p(R)$  to  $L^p(R)$ , is 1.

We have  $V_p = V_{p'}$  for p' = p/(p-1); moreover,  $V_2$  is just the set of k satisfying (5.2) with  $\|\hat{k}\|_{\infty} = 1$ . As is easily seen,  $V_1 = V_{\infty}$  is the set of nonnegative integrable functions satisfying (5.2). As for  $V_p$ ,  $2 , this class is known to contain functions which are not of constant sign (this is a consequence of results in [14], which paper is closely related to the material in the present section). Nonetheless, it is a remarkable fact that the elements of <math>V_p$ , 2 , present features reminiscent of positive kernels. The relevant result can be stated as a pure theorem on Fourier transforms, or re-interpreted as a saturation theorem. First the "Fourier" version:

Theorem 4. Let  $k \in V_p$ , 2 . If

(5.3) 
$$1 - \hat{k}(x) = O(x^{\alpha}), \quad x \to 0 +,$$

then  $\alpha \le 2 + (p-1)^{-1}$  (and so  $\alpha \le 3$ ).

**Remark.** For  $p=\infty$  the result is also true (and known) even in the stronger form that an element k of  $V_{\infty}$  cannot satisf  $1-\hat{k}(x)=o(x^2), x \to 0+$  (compare the proof of Theorem 4.1.1 in [9, p. 68]).

**Proof of Theorem 4.** Let us consider the particular subspace S, and function f, defined in Theorem 2, and let  $U_a$  denote the transformation

$$g \mapsto k_{(a)} * g, \quad g \in L^p(\mathbf{T}).$$

Now,  $U_a$  has norm one from  $L^p(\mathbf{R})$  to  $L^p(\mathbf{R})$ , and hence also from  $L^p(\mathbf{T})$  to

 $L^p(\mathbf{T})$ , see [8]. Since also  $U_a$  maps S into itself, we may apply Corollary 1 with  $U=U_a$ , and get for small a,

$$\|f_0 - U_a f_0\|_p \le C_p |1 - \hat{k}(a)|^{1/p}$$

where  $C_p$  depends only on p. Hence, by (5.3),

By an "inverse theorem of approximation theory" (see [13, Chapter 9]), this implies  $f_0 \in \Lambda^p_{\alpha/p}$ . Hence, by Theorem 2,

(5.5) 
$$\alpha/p \le (1/p) + (1/(p-1))$$

and the theorem is proved.

Thus a kernel  $k \in V_p$ , with p > 2, cannot satisfy  $1 - \hat{k}(x) = O(|x|^{\alpha})$  for any  $\alpha > 3$ . From this we can readily deduce the following "saturation theorem", cf. [12, p. 36, Theorem 15]:

If  $k \in V_p$ , where  $2 and <math>g \in L^p(\mathbf{T})$  and  $\|g - (k_{(a)} * g)\|_p = O(a^{\alpha})$  for some  $\alpha > 3$ , then g is constant.

It is natural to seek a generalization of Theorem 4, whereby k is permitted to be a distribution in place of an integrable function; let us first discuss the case of a *measure*, which avoids certain technical difficulties of the distributional problem (discussed further below).

Theorem 5. Let  $\mu$  denote a bounded complex measure on R such that the map  $f \mapsto f * \mu$  has norm one from  $L^p(R)$  to  $L^p(R)$ , where  $2 . If <math>1 - \hat{\mu}(x) = O(x^{\alpha})$ ,  $x \to 0+$ , where  $\alpha > 2 + (p-1)^{-1}$ , then  $\hat{\mu}(x) = 1$  for all  $x \ge 0$ .

**Proof.** We proceed as in the preceding proof, and obtain in place of (5.4),

(5.5) 
$$||f_0 - (\mu_{(a)} * f_0)||_p = O(a^{\alpha/p}), \quad a \to 0_+,$$

where  $\mu_{(a)}$  denotes the measure whose Fourier transform is  $\hat{\mu}(ax)$ . It is deducible from Theorem 9.4.4.5 of [13] that either  $f \in \Lambda^p_{\alpha/p}$ , or else  $1 - \hat{\mu}(x)$  fails to satisfy the "Tauberian condition", that is, it vanishes on a half-line through the origin. If the former case holds, then, reasoning as in the proof of Theorem 4 we deduce a contradiction, hence we may assume  $\hat{\mu}(x) = 1$  on a half-line through the origin. If this is the positive half-axis, we are done. In the contrary case, we may assume  $\hat{\mu}(x) = 1$  for all  $x \leq 0$ , whereas  $\hat{\mu}(x_1) \neq 1$  for some  $x_1 > 0$ . Since the  $L^p$   $\rightarrow L^p$  operator norm of  $\mu$  is 1, the measure  $\nu$  defined by  $\hat{\nu}(x) = (1/2)(\hat{\mu}(x) + \hat{\mu}(-x))$  has operator norm not exceeding 1. Also,  $\hat{\nu}$  satisfies the Tauberian condition, and  $1 - \hat{\nu}(x) = (1/2)(1 - \hat{\mu}(x)) = O(x^a)$ ,  $x \to 0+$ . But now we are reduced to a previously treated case, and arrive at a contradiction. The proof is finished.

In seeking to extend this theorem from measures to the most general distri-

butions which convolve  $L^p$  into  $L^p$ , a difficulty arises in connection with the "inverse theorem". Namely, if  $\mu$  is only assumed to be a distribution convolving  $L^p$  into  $L^p$ , and  $\hat{\mu}$  satisfies the Tauberian condition (in some reasonable sense; recall that  $\hat{\mu}$  is defined only a.e.), there seems to be no extant theorem which would enable us to conclude from (5.5) that  $f_0$  is in  $\Lambda^p_{\alpha/p}$ ; some extra regularity assumption on  $\hat{\mu}(x)$  near x=0, in any case, suffices. Thus, Theorem 5 remains valid for any distribution  $\mu$  such that, for some  $\epsilon>0$ , the restriction of  $\hat{\mu}$  to  $(-\epsilon,\epsilon)$  coincides with the Fourier transform of a bounded measure; we omit the simple proof. (2)

Of much greater interest is the question whether  $2 + (p-1)^{-1}$  in the conclusion of Theorem 5 can be replaced by a smaller number (the present theorem is an improvement over an earlier, unpublished result of N. Rivière and myself; we obtained using other means, "4" in the conclusion). It seems reasonable that  $2 + (p-1)^{-1}$  can be replaced by "2", which would then be the definitive value. It is clear from examining the method here employed that an improvement in Theorem 2 would yield a corresponding improvement in Theorem 5. Indeed, it is not hard to see that if we could find a trigonometric polynomial, or even some sufficiently regular function, whose metric projection on some closed translation-invariant subspace of  $L^p(T)$  fails to belong to  $\Lambda^p_a$ , for every  $\alpha > 2/p$ , we would obtain Theorem 5 in a stronger version, with "2" in the conclusion.

In concluding this section, I cannot refrain from remarking that the proof of Theorem 5 seems very indirect, almost bizarre. Any more direct approach would be of interest.

### 6. Concluding remarks and open questions.

- 6.1. Open questions. We have already indicated some open questions, especially the gap between Corollary 2 and Theorem 2. Here are a few others.
- 6.1.1. Localization. For special choices of S (notably  $S=H^p$ ) the property of f to belong in the neighborhood of some point to a regularity class entails a corresponding property of its metric projection on S. Results of great precision are known for  $H^\infty$  (cf. [6]) and also for  $H^2$ , when the metric projection is closely related to the Hilbert transform (I do not know any reference, but it is easy to derive such results). Even for p=2 the local regularity problem for orthogonal projection on S(M) does not seem quite trivial; it is related to the location of the singularities of the distribution  $\Sigma_{t\in M} e^{int}$ , which in the case of  $H^2$  (that is,  $M=\mathbf{Z}^+$ ) has singularity only at t=0.

<sup>(2)</sup> Added in proof. Y. Katznelson has pointed out to me how the "inverse theorem" can be proved in the desired degree of generality. Thus, Theorem 5 is valid if in place of a measure  $\mu$  we have any tempered distribution, provided  $\hat{\mu}$  is a continuous (or even "regulated", in the sense of [8]) function on  $\hat{\mathbf{R}}$ .

- 6.1.2. Saturation. Can one find nontrivial S such that  $f \in C^{\infty}$  implies  $f_0 \in \Lambda^p_{\alpha}$  with  $\alpha > 2/p$ ? In particular, does  $H^p$  have this property (cf. 6.4)?
- 6.1.3. Translation-invariant sets. We have assumed, in Theorem 1, that S is a closed convex translation-invariant subset of  $L^p$ . The structure of these sets seems to be of some independent interest: can one describe these sets, say, in terms of Fourier transforms? (As we remarked, this problem has a simple and well known solution when S is a subspace.) Even the case p=2 seems non-trivial.
- 6.2. Generalizations. Most results of this paper hold, mutatis mutandis in  $L^p(\mathbf{R})$ , also in  $L^p(\mathbf{R}^n)$  and  $L^p(\mathbf{T}^n)$ ; the proofs are the same, indeed we could work more abstractly, in a homogeneous Banach space on  $\mathbf{R}^n$  (cf. [13, Chapter 9]) for which the Clarkson inequality is assumed to hold. Here a variant that is convenient is to choose for U, in the proof of Corollary 2, the operator  $f \to k_{(a)} * f$  rather than  $(T_a + T_{-a})/2$ ;  $k(t) = \pi^{-n/2} \exp(-|t|^2)$  would be a suitable kernel, and application of "direct" and "inverse" theorems of approximation yields the desired conclusion (now the definition of the classes  $\Lambda^p_a$  given in §4 is especially convenient); the details should be easy to fill in by anyone who has read thus far, and are thus omitted.
- 6.3. Regularity of extremal functions, a related result. There exist subspaces X of  $L^p(T)$  of which one can assert that every element orthogonal (we emphasize that we use this term throughout in the  $L^p$  sense) to X enjoys a certain degree of regularity. For instance, if X is the span of  $\{e^{int}\}_{n\neq 0}$  any orthogonal element is easily seen be be constant; somewhat less trivially, if X is the span of  $\{e^{int}\}_{n\neq 0,1}$ , the orthogonal elements have the form  $|a+be^{it}|^{p'}/(a+be^{it})$  and so (as follows from the proof of Theorem 2) belong to  $\Lambda^p_\beta$ ,  $\beta=(1/p)+1/(p-1)$ . I can prove a slight generalization of this, from which one can infer some regularity of the minimal solution of a moment problem with regular data. Observe that now X is not assumed translation-invariant.

First a definition: we say that a  $2\pi$ -period function f is *highly regular* if it is the restriction to R of a function analytic in a strip about R.

Theorem 6. Let 
$$\phi_1, \dots, \phi_m$$
 be highly regular functions,  $2 , and 
$$X = \left\{ f \in L^p(T) : \int f(t)\phi_j(t) dt = 0, \ j = 1, \dots, m \right\}.$$$ 

Any function in  $L^p(T)$  orthogonal, in the  $L^p$  sense, to X belongs to  $\Lambda^p_{\beta}$ , where  $\beta = (1/p) + 1/(p-1)$ .

Corollary 6. Let  $\phi_1, \dots, \phi_m$  be highly regular functions, linearly independent, and  $\lambda_1, \dots, \lambda_m$  complex numbers not all zero. Let 2 , and

$$Y = \left\{ f \in L^p(\mathbf{T}) : \int f(t)\phi_j(t) dt = \lambda_j, \ j = 1, \cdots, m \right\}.$$

Then, Y contains a unique element  $f^*$  of minimal norm, and  $f^* \in \Lambda^p_\beta$ , where  $\beta = (1/p) + 1/(p-1)$ .

Let us first deduce the corollary from the theorem. The existence and uniqueness of  $f^*$  are apparent, since Y is a closed convex set in a uniformly convex Banach space. A routine variational argument shows that  $f^*$  is orthogonal to X, and the desired conclusion follows.

For the proof of the theorem we need some lemmas.

Lemma 1. Let f and g belong to  $L^{\infty}(T) \cap \Lambda_a^b$ . Then  $fg \in \Lambda_a^b$ .

**Proof.** There exists, for each n, a trigonometric polynomial  $u_n$  such that  $\|f-u_n\|_p = O(n^{-\alpha})$  (and corresponding trigonometric polynomials  $v_n$  associated with g). Moreover, we may assume  $\|u_n\|_{\infty}$  are bounded (for, we can choose  $u_n = k_{(1/n)} * f$ , where  $k \in L^1(\mathbb{R})$ , and  $\hat{k}(x)$  vanishes outside [-1,1] and equals 1 in a neighborhood of x=0; cf. [12, p. 57, Theorem 21]). Hence

$$\|fg - u_n v_n\|_p \le \|g(g - u_n)\|_p + \|u_n(g - v_n)\|_p = O(n^{-\alpha}), \quad n \to \infty.$$

Since  $u_n v_n$  is a trigonometric polynomial of degree at most 2n, the desired conclusion follows from the converse to Jackson's theorem.

**Lemma 2.** For fixed p, 2 , let <math>J denote the map which takes the function f into the function F, where

$$F(t) = \begin{cases} |F(t)|^{p'}/F(t), & F(t) \neq 0, \\ 0, & F(t) = 0. \end{cases}$$

If f is highly regular, If  $\in \Lambda^p_\beta$  where  $\beta = (1/p) + 1/(p-1)$ .

**Proof.** We can write  $f=g_1\cdots g_r h$ , where  $h\in C^2$  is nonvanishing, and  $g_j(t)=e^{i(t-t_j)}-1$  with  $t_j$  real. Now,  $Jf=(Jg_1)\cdots(Jg_r)(Jh)$  and, in view of Lemma 1, it is enough to show that each factor on the right belongs to  $\Lambda^p_\alpha$ . Clearly,  $Jh\in C^2\subset \Lambda^p_\beta\subset \Lambda^p_\beta$ . As for  $g_j$ , since  $\Lambda^p_\beta$  is translation-invariant it is enough to consider  $t_j=0$ , i.e. to show that  $|e^{it}-1|^{p'}/(e^{it}-1)$  is in  $\Lambda^p_\beta$ . But, this was shown in the proof of Theorem 2.

**Proof of Theorem 6.** Every  $f \in L^p(T)$  orthogonal to X has the form  $|g|^{p'}/g$  where  $g \in L^{p'}(T)$  and

$$\int u(t)g(t) dt = 0, \quad \text{all } u \in X$$

(cf.  $\S 2$  of [14]). By simple linear algebra, g must be a linear combination of the  $\phi_j$ , hence highly regular. Therefore,  $f = Jg \in \Lambda^p_\beta$ , by Lemma 2.

Remarks. (1) Probably these results could be extended. The difficulty is in finding broader additive classes of functions than the "highly regular" ones whose

image under the *J*-operator lies in some specified  $\Lambda^p_\alpha$ .

- (2) Here, as elsewhere, analogous results are valid for 1 . For <math>p = 2, stronger results are (trivially) true, e.g. in Theorem 6 elements orthogonal to X are merely linear combinations of the  $\phi_j$ , hence highly regular
- (3) The value given for  $\beta$  in Theorem 6 and its corollary is the largest possible, as the subspace S in Theorem 2 shows.
  - 6.4. A refinement of Corollary 2 when  $S = H^p$ .

When  $S = H^p$ , and f is holomorphic in a strip about the real axis (that is,  $F(z) = \sum_{n=-\infty}^{\infty} \hat{f}(n)z^n$  is holomorphic in some annulus  $\{\rho < |z| < \rho^{-1}\}$ ), the theory of "dual extremal problems" [3, Chapter 8] can be employed to give something a little better than  $f_0 \in \Lambda_{2/p}^p$ , which is all we can get out of Corollary 2. Although the analysis in [3, §8.4] is carried out for rational f, a very similar reasoning (which we omit) establishes, in case f is holomorphic on |z| = 1, that the best approximation to f from f is of the form f is a strip about the real axis. By computations similar to those carried out in the preceding section and in the proof of Theorem 2, we deduce f f f is and for f distinct from 1, 2, and f the number f is in general the largest possible. It would be of interest here to establish the same conclusion about f with a weaker (for example, f by hypothesis about f.

6.5. Addendum (April 1973). A detailed account of the Carleson-Jacobs results referred to in §4 has now appeared in Ark. Mat. 10 (1972), 219-229.

A detailed study of metric projections, principally in translation-invariant subspaces of  $L^1(T)$  (in particular, inheritance of regularity properties), has been made by Jean-Pierre Kahane in the following papers:

Projection métrique de  $L^1(T)$  sur des sous-espaces sermés invariants par translation, to appear in the same volume as [14].

Projections métriques dans  $L^1(T)$ , C. R. Acad. Sci. Paris Sér. A 276 (1973), A621-A623.

Heredity in metric projections (preprint).

#### BIBLIOGRAPHY

- 1. V. M. Adamjan, D. Z. Arov and M. G. Krein, Infinite Hankel matrices and generalized problems of Carathéodory-Fejér and F. Riesz, Funkcional. Anal. i Priložen. 2 (1968), no. 1-19. (Russian) MR 38 #2591.
- 2. J. Clarkson, Uniformly convex spaces, Trans. Amer. Math. Soc. 40 (1936), 396-414.
- 3. P. L. Duren, Theory of H<sup>b</sup> spaces, Pure and Appl. Math., vol. 38, Academic Press, New York, 1970. MR 42 #3552.
- 4. G. H. Hardy and J. E. Littlewood, A convergence criterion for Fourier series, Math. Z. 28 (1928), 612-634.

- 5. L. Hörmander, Estimates for translation invariant operators in L<sup>p</sup> spaces, Acta Math. 104 (1960), 93-140. MR 22 #12389.
- 6. S. Jacobs, An extremal problem for analytic functions in multiply connected regions, Uppsala, Feb. 1969, preprint. (Swedish)
- 7. J.-P. Kahane, Séries de Fourier absolument convergentes, Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 50, Springer-Verlag, Berlin and New York, 1970.
- 8. K. deLeeuw, On  $L^p$  multipliers, Ann. of Math. (2) 81 (1965), 364-379. MR 30 #5127.
  - 9. E. Lukacs, Characteristic functions, 2nd ed., Hafner, New York, 1970.
  - 10. J. Peetre, Reflections on Besov spaces, Lund, 1966 (mimeographed lecture notes).
- 11. H. S. Shapiro, Applications of normed linear spaces to function-theoretic extremal problems, Lectures on Functions of a Complex Variable, Univ. of Michigan Press, Ann Arbor, Mich., 1955, pp. 399-404. MR 17, 25.
- 12. ——, Smoothing and approximation of functions, Van Nostrand Reinhold Math. Studies, no. 24, Van Nostrand Reinhold, New York, 1969.
- 13. ——, Topics in approximation theory, Lecture Notes in Math., vol. 187, Springer-Verlag, Berlin and New York, 1971.
- 14. ——, Fourier multipliers whose multiplier norm is an attained value, Proc. Sympos. Linear Operators and Approximation (Oberwolfach, Germany, Aug. 1971) (to appear).
- 15. O. Szasz, Über die Fourierschen Reihen gewisser Funktionenklassen, Math. Ann. 100 (1928), 530-536.
- 16. M. H. Taibleson, On the theory of Lipschitz spaces of distributions on Euclidean n-space. II. Translation invariant operators, duality and interpolation, J. Math. Mech. 14 (1965), 821-839. MR 31 #5087.
- 17. A. Zygmund, On the preservation of classes of functions, J. Math. Mech. 8 (1959), 889-895; erratum, ibid. 9 (1960), 663. MR 22 #8277.
- 18. ———, Trigonometrical series, 2nd rev. ed., Cambridge Univ. Press, New York, 1968. MR 38 #4882.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF MICHIGAN, ANN ARBOR, MICHIGAN 48104

Current address: Mathematics Institute, Royal Institute of Technology, (KTH), S-100 44 Stockholm 70, Sweden