SOME PROBLEMS ARISING FROM PREDICTION THEORY AND A THEOREM OF KOLMOGOROV

by

JOSE J. GUADALUPE HERNANDEZ and JOSE L. RUBIO DE FRANCIA

1. INTRODUCTION.

Let $T = \{ e^{2\pi i x} \mid 0 \leqslant x < 1 \}$ denote the unit circle identified in the natural way with the interval [0, 1] and provided with normalized Lebesgue measure m, so that m (T) = 1. Denote by P the vector space of all trigonometric polynomials, and consider the subspaces

$$P_+ = \{ g \in P \mid \hat{g}(k) = 0 \quad \text{for all} \quad k \leqslant 0 \}$$

$$P_0 = \{ g \in P \mid \hat{g}(0) = 0 \}$$

If μ is a positive finite Borel measure in T, the distance in $L^p(\mu)$ between 1 and P_+ is given by the beautiful formula

$$\inf_{g \in P_+} \int |1 + g(x)|^p d\mu(x) = \exp \int \log w(x)\, dx \qquad (0 < p < \infty) \tag{1}$$

where $d\mu(x) = w(x)\, dm(x) + d\mu_s(x)$ is the Lebesgue decomposition of μ. The identity (1) has important consequences in the Theory of Functions, Fourier Analysis, Orthogonal Polynomials and Prediction Theory. It was first proved by Szegö in the case: $p = 2$, $\mu << m$, and subsequently generalized by several authors (see Helson [4]). The same question can be asked for P_0 insted of P_+, i.e., which is the value of

$$d_p(\mu) = \inf_{g \in P_0} \int |1 + g(x)|^p d\mu(x) \tag{2}$$

Observe that $d_p(\mu) > 0$ means that no character in T can be approximated in the metric of $L^p(\mu)$ by linear combinations of characters different from it. In terms of Prediction Theory, $d_2(\mu) > 0$ means that, for a certain stochastic process represented by μ, the strict past and the strict future together do not determine the present (see [10]). When $p = 2$, an answer to this question was given by Kolmogorov:

Theorem 0: *Let* $d\mu(x) = w(x)\,dm(x)$. *Then*

$$d_2(\mu) = \left(\int_T w(x)^{-1}\,dm(x)\right)^{-1}$$

In particular, $d_2(\mu) > 0$ *if and only if* $w(x) > 0$ *a.e. and* $w^{-1} \in L^1(T)$.

The proof of theorem 0 is simple and based on orthogonality arguments in the Hilbert space $L^2(\mu)$. Our aim here is to present a different approach which is suitable to deal with $L^p(\mu)$, $0 < p < \infty$, and even with Orlicz spaces. Some other related problems will be considered along the way.

2. THE MAIN RESULT FOR ORLICZ SPACES.

The following notation will be used throughout. Φ and Ψ will be two conjugate Young functions in $[0, \infty)$, and $L_\Phi(\mu)$, $L_\Psi(\mu)$ will denote the corresponding Orlicz spaces defined on the measure space (T, μ). Here μ is a positive finite Borel measure in T with Lebesgue decomposition:

$$d\mu(x) = w(x)\,dm(x) + d\mu_s(x).$$

When μ is absolutely continuous, we identify μ with the weight $w(x)$, and write $L_\Phi(w)$ instead of $L_\Phi(\mu)$. Finally, for every trigonometric polynomial $f \in P$, we denote by $Hf(x) = \tilde{f}(x)$ the conjugate of f, and by

$$S_n f(x) = \sum_{|k| \leq n} \hat{f}(k) \exp(2\pi i k x)$$

the n-th partial sum of f.

Theorem 1: *The following statements are equivalent:*

a) $\inf\{\|1 + g\|_{L_\Phi(\mu)} : g \in P_0\} = c > 0$

b) $w(x) > 0$ a.e. and $w^{-1} \in L_\Psi(w)$

c) $m(\{x: |Hf(x)| > t\}) \leqslant Ct^{-1} \|f\|_{L_\Phi(\mu)}$ $(f \in P)$

d) $m(\{x: |S_n f(x)| > t\}) \leqslant Ct^{-1} \|f\|_{L_\Phi(\mu)}$ $(f \in P;\ n \geqslant 0)$

where $C > 0$ *denotes an absolute constant which depends only on* μ.

Proof: (a) implies (b). For every f e P we have the inequality

$$|\hat{f}(0)| \leqslant c^{-1} \|f\|_{L_\Phi(\mu)} \tag{3}$$

In fact, this is obvious when $\hat{f}(0) = 0$, and when $\hat{f}(0) \neq 0$ we apply (a) to the trigonometric polynomial $[\hat{f}(0)]^{-1} f \in 1 + P_0$. Let $E_\Phi(\mu)$ denote the closure of P in $L_\Phi(\mu)$, which coincides with the closure of $L^\infty(\mu)$ in $L_\Phi(\mu)$ (by Lusin's theorem and Weierstrass' approximation theorem). By (3), the mapping $f \to \hat{f}(0)$ extends to a continuous linear functional on $E_\Phi(\mu)$, and it is known ([6]) that all such functionals are represented by functions in $L_\Psi(\mu)$, i.e., there exists $h \in L_\Psi(\mu)$ such that

$$\int f(x)\, dm(x) = \hat{f}(0) = \int f(x)\, h(x)\, d\mu(x) \qquad (f \in P) \tag{4}$$

It follows that $h d\mu = dm$, i.e.

$$h = 0 \quad \mu_s\text{- a.e.,} \quad h(x) = w(x)^{-1} \ m \quad \text{a.e.}$$

Therefore, $h \in L_\Psi(\mu)$ is equivalent to $w^{-1} \in L_\Psi(w)$.

(b) implies (c) and (d). We shall need the extension of Hölder's inequality

$$\|fg\|_{L^1} \leqslant \|f\|_{L_\Phi} \|g\|_{L_\Psi}$$

where $\|\cdot\|_{L_\Phi}$ and $\|\cdot\|_{L_\Psi}$ are the Orlicz norms. By Kolmogorov's inequality (see [11])

$$m(\{x: |Hf(x)| > t\}) \leqslant At^{-1} \int |f(x)|\, dm(x) =$$

$$= At^{-1} \int |f|\, w^{-1}\, w dm \leqslant A \|w^{-1}\|_{L_\Psi(w)}\, t^{-1} \|f\|_{L_\Phi}(w)$$

Since $\|w^{-1}\|_{L_\Phi(w)} < \infty$ and $\|\cdot\|_{L_\Phi(w)} \leqslant \|\cdot\|_{L_\Phi(\mu)}$, we have proved (c). The same argument can be applied to $\{S_n\}_{n \in N}$ instead of H to prove (d).

(c) implies (a). For every f e P we have

$$Hf(x) \cdot e^{2\pi ix} H(fe^{-2\pi i\cdot})(x) = -i[\hat{f}(0) + \hat{f}(1)e^{2\pi ix}]$$

and from (c) we obtain

$$m(\{x: |\hat{f}(0) + \hat{f}(1)e^{2\pi ix}| \geq t\}) \leq 4Ct^{-1}\|f\|_{L_\Phi(\mu)} \qquad (f \in P)$$

Now, if $\hat{f}(0) = 1$, there is an interval of length 1/2 such that, for all x in that interval: $\mathrm{Re}(\hat{f}(1)e^{2\pi ix}) \leq 0$, and therefore $|1 + \hat{f}(1)e^{2\pi ix}| \geq 1$. Then

$$1/2 \leq 4C\|f\|_{L_\Phi(\mu)} \qquad (f \in 1 + P_0)$$

which is (a).

(d) implies (a). It suffices to use the inequality for $S_0 f = \hat{f}(0)$, which gives (3) for every $f \in P$, and (3) is equivalent to (a).

Remarks: 1.– A retrospective look at the proof of the theorem shows that the best C in (c) and (d) is equivalent to 1/c and to $\|w^{-1}\|_{L_\Psi(w)}$, i.e.

$$1/c \leq k_1\|w^{-1}\|_{L_\Psi(w)} \leq k_2 C \leq k_3/c$$

for some constants $k_1, k_2, k_3 > 0$ independent of μ.

2.- If Ψ satisfies the Δ_2 condition:

$$\Psi(2t) \leq (\mathrm{Const.})\,\Psi(t),\ t > 0,$$

then (b) can be written as

$$\int \Psi(w(x)^{-1})\,w(x)\,dm(x) < \infty$$

3. THE CASE OF L^p SPACES.

The proof of theorem 1 shows that, in general

$$\inf\{\|1 + g\|_{L_\Phi(\mu)}: g \in P_0\}^{-1}$$

equals the norm of w^{-1} as an element of $[E_\Phi(w)]^*$. When $\Phi(t) = t^p$, $1 < p < \infty$, this is exactly $\|w^{-1}\|_{L^{p'}(w)}$. Thus we have

Theorem 2: *Let* $d_p(\mu)$ *be defined by* (2). *If* $1 < p < \infty$, *we have* $d_p(\mu) > 0$ *if and only if* $w^{-p'/p} \in L^1(T)$, *and, more precisely*

$$d_p(\mu) = \{\int w(x)^{-p'/p}\, dm(x)\}^{-p/p'}$$

Observe that Kolmogorov's theorem 0 is a particular case of theorem 2. The case $p = 1$ can be treated exactly as in theorem 1 (and it is even simpler), with L_Φ and L_Ψ replaced by L^1 and L^∞ respectively. We limit ourselves to state the results:

Theorem 3: *The following statements are equivalent:*

a) $d_1(\mu) = \inf\{\int|1 + g|\, d\mu : g \in P_0\} > 0$

b) *there exists* $k > 0$ *such that* $k \leqslant w(x)$ a.e.

c) H *extends to a bounded operator from* $L^1(\mu)$ *into* $L^1_*(T) = \text{weak-}L^1(T)$.

Moreover, we have the identity

$$d_1(\mu) = \|w^{-1}\|_\infty^{-1} = \operatorname*{ess\,inf}_{x \in T} w(x)$$

Theorem 4: $d_p(\mu) = 0$ *for all* $0 < p < 1$.

Proof: If $d_p(\mu) > 0$ for some $p < 1$, then $f \to \hat{f}(0)$ extends to a continuous linear functional on $L^p(\mu)$, and (4) will be verified for some function h which is a countable linear combination of μ-atoms ([2]). In particular, supp (h) $\subset$ supp (μ_s). (Since there are no μ-atoms in the absolutely continuous part of μ) and this makes (4) impossible.

The result of theorem 2, and in particular, the formula for $d_p(\mu)$, merits further discussion. First, if we recall that, in a probaility space, the norms $\|g\|_r$ converge to the geometric mean of $|g|$ as $r \to 0$ (see [5], 13.32) we have

Corollary 1: $d_p(\mu)$ *is an increasing function of* p, *and*

$$\lim_{p \to \infty} d_p(\mu) = \exp \int \log w(x)\, dm(x)$$

If one combines this with Szegö's formula, it turns out that, although the left hand side of (1) is greater than $d_p(\mu)$ for each $p > 0$, both terms converge to the same limit when $p \to \infty$.

Corollary 2: *The following statements are equivalent:*

a) $\log w \in L^1(T)$

b) $\inf\{\int|1+g|^p\,d\mu : g \in P_+\} > 0$ for all $p > 0$

c) $\inf\{\int|1+g|^p\,d\mu : g \in P_+\} > 0$ for some $p > 0$

d) $\inf\{\int|1+g|^p\,d\mu : g \in P_o\} > 0$ for some $p > 0$

e) $\inf\{\|1+g\|_{\exp L(\mu)} : g \in P_o\} > 0$

In fact, it is obvious from (1) that (a) $\Leftrightarrow$ (b) $\Leftrightarrow$ (c), the equivalence (a) $\Leftrightarrow$ (d) follows from Corollary 1, and the remark 2 of the previous section applied to

$\Psi(t) = t\log^+ t$, yields (a) $\Leftrightarrow$ (e).

There is a final remark to make if the formula for $d_p(\mu)$ is written as

$$\inf_{g \in P_o} \int|1+g(x)/p|^p\, w(x)\,dm(x) = \{\int w(x)^{-p'/p}\,dm(x)\}^{-p/p} \qquad (1 < p < \infty)$$

Letting $p \to \infty$ and proceeding formally (i.e., replacing lim (inf...) by inf (lim...)) one gets

$$\inf_{g \in P_o} \int e^{\mathrm{Re}\, g(x)}\, w(x)\,dm(x) = \exp\int\log w(x)\,dm(x) \tag{5}$$

It turns out that (5) actually holds, and it is a well known identity ([9], 8.3).

4. A RELATED PROBLEM.

The condition $w^{-p'/p} \in L^1(T)$ appearing in theorem 2, which is equivalent to $L^p(w) \hookrightarrow L^1(T)$, seems to be also the answer to several other questions. For instance, it is necessary and sufficient for the existence of some $u(x) > 0$ such that

$$\int|Hf(x)|^p u(x)\,dm(x) \leqslant \int|f(x)|^p\, w(x)\,dm(x) \qquad (f \in P)$$

(see [7], [8], [1]). The problem that we shall consider here is the following. For a given $w \in L^1_+(T)$, $H^p(w)$ and $\widetilde{H}^p(w)$ denote, respectively, the closures in $L^p(w)$ of the subspaces

$$\mathbb{C} \oplus P_+ = \{ g \in P \mid \hat{g}(k) = 0 \quad \text{for all} \quad k < 0 \}$$

$$\mathbb{C} \oplus P_- = \{ g \in P \mid \hat{g}(k) = 0 \quad \text{for all} \quad k > 0 \}$$

and we ask ourselves, for which w is $H^p(w) \cap \widetilde{H}^p(w) = \mathbb{C} = \{$ constant functions $\}$. In terms of Prediction Theory, this condition means, in the case $p = 2$, that the intersection of the past and the future is the present (where past and future refer to the stochastic process spectrally represented by wdm). For some similar questions, see [10].

Theorem 5 : If $w^{-p'/p} \in L^1(T)$, *then* $H^p(w) \cap \widetilde{H}^p(w) = \mathbb{C}$. *However, if* $0 < q < p'/p$, *there exists* $w(x) > 0$ *such that* $w^{-q} \in L^1(T)$ *and*

$$H^p(w) \cap \widetilde{H}^p(w) \underset{\neq}{\supset} \mathbb{C}.$$

Proof:

Assume that

$$k = \{ \int w(x)^{-p'/p} \, dm(x) \}^{1/p'} < \infty$$

If $f \in H^p(w) \cap \widetilde{H}^p(w)$, then $f, \bar{f} \in H^p(w)$, and there exist trigonometric polynomials $\{P_n\}$ and $\{Q_n\}$ of analytic type such that

$$\lim_n \|P_n - f\|_{L^p(w)} = \lim_n \|Q_n - \bar{f}\|_{L^p(w)} = 0$$

On the other hand, Holder's inequality gives us

$$\int |P_n(x) - f(x)| \, dm(x) \leqslant k \{ \int |P_n(x) - f(x)|^p w(x) \, dm(x) \}^{1/p} \to 0$$

$$(n \to \infty)$$

i.e., $f \in H^1(T)$. By the same reason, $\bar{f} \in H^1(T)$, and both togehter imply $f(x) =$ constant, a.e.

On the other hand, given $q < p'/p$, we choose s such that $p - 1 < s < 1/q$, and define $w(x) = |1 - e^{2\pi i x}|^s$. The condition $w^{-q} \in L^1(T)$ is then verified. We shall need the following description of $H^p(w)$ (see [3]):

$$H^p(w) = K \cdot H^p(T) = \{ K(x) g(x) \mid g \in H^p(T) \}$$

where K is the boundary value of the outer function $K(z)$ such that

$$|K(x)| = w(x)^{-1/p} \qquad (\text{a.e. } x \in T)$$

Then $g(z) = K(z)^{-1}(1+z)/(1-z)$ is an analytic function in the unit disc with radial limit $g \in L^p(T)$, since

$$\int |g(x)|^p \, dm(x) = \int |\text{sen } x/(1 - \cos x)|^p w(x) \, dm(x) \leqslant C \int_0^1 x^{s-p} dx < \infty$$

Therefore, $g \in H^p(T)$, and $f(x) = K(x)\, g(x) = \text{sen } x/(1 - \cos x)$ belongs to $H^p(w)$. Since $f(x)$ is real valued, it follows that $f \in \widetilde{H}^p(w)$. Thus, $H^p(w) \cap \widetilde{H}^p(w)$ contains non constant functions.

This teorem leads naturally to formulate the following

Conjecture: $H^p(w) \cap \widetilde{H}^p(w) = \mathbb{C}$ *if and only if* $w^{-p'/p} \in L^1(T)$.

Let us finally observe that all the results proved in this paper make sense and are also true if T is replaced by an arbitrary compact Abelian group with an ordered dual (see [4] and [9] for the extension of some classical facts to this context). Moreover, if one drops the statements (c) and (d) of theorem 1, (c) of theorem 3 and (b), (c) in Corollary 2, the whole sections 2 and 3 can be generalized to arbitrary compact groups.

REFERENCES

1. L. Carleson, P. Jones: *Weighted norm inequalities and a theorem of Koosis.* Inst. Mittag-Leffler, Report No. 2 (1981).

2. M. M. Day: *The spaces L^p with $0 < p < 1$.* Bull. Amer. Math. Soc. 46 (1940), 816-823.

3. J. J. Guadalupe: *Invariant súbspaces and H^p spaces with respect to arbitrary measures.* Boll. Unione Mat. Italiana (6) 1 - B (1982). 1067-1077.

4. H. Helson: *Analyticity on compact abelian groups.* In "Algebras in Analysis", Ed. Williamson, Academic Press 1975.

5. E. Hewitt, K. Stromberg: *Real and Abstract Analysis.* Springer-Verlag, Berlín 1969.

6. M. A. Krasnosel'skii, J. B. Rutickii: *Convex functions and Orlicz spaces.* Noordhoff, Groningen 1961.

7. P. Koosis: *Moyennes quadratiques pondérées de fonctions périodiques et de leurs conjuguées harmoniques.* C. R. Acad. Sc. París 291 (1980), sér A, 255-257.

8. J. L. Rubio de Francia: *Boundedness of maximal functions and singular integrals in weighted L^p spaces.* Proc. Amer. Math. Soc. 83 (1981). 673-679.

9. W. Rudin: *Fourier Analysis on Groups.* J. Wiley, Interscience, N.Y. 1962.

10. D. Sarason: *Function Theory on the unit cricle.* Notes of Lectures, Virginia Polytechnic Inst., 1978.

11. A. Zygmgund: *Trigonometric Series, I.* Cambridge Univ. Press, 1968.

J. J. Guadalupe Hernández
Colegio Universitario
Logroño

J. L. Rubio de Francia
División de Matemáticas
Universidad Autónoma de Madrid