JOURNAL OF APPROXIMATION THEORY 63, 338–367 (1990)

Best Approximants in Modular Function Spaces

SHELBY J. KILMER AND WOJCIECH M. KOZLOWSKI*

Department of Mathematics,
Southwest Missouri State University, Springfield, Missouri 65804

AND

GRZEGORZ LEWICKI

Department of Mathematics,
The Jagiellonian University, Reymonta 4, 30-059 Kraków, Poland

Communicated by E. W. Cheney

Received August 14, 1989

In this paper we consider the existence of best approximants in modular function spaces by elements of sublattices. Modular function spaces are the natural generalization of L_p, $p>0$, Orlicz, Lorentz, and Köthe spaces. Let ρ be a pseudomodular, L_ρ the corresponding modular function space, and C a sublattice of L_ρ. Given a function $f \in L_\rho$ we consider the minimization problem of finding $h \in C$ such that $\rho(f-h) = \inf\{f-g: g \in C\}$. Such an h is called a *best approximant.* Problems of finding best approximants are important in approximation theory and probability theory. In the case where C is $L_\rho(\mathscr{B})$ for some σ-subalgebra $\mathscr{B}$ of the original σ-algebra, finding best approximants is closely related to the problem of nonlinear prediction. Throughout most of the paper we assume only that ρ is a pseudomodular and except in one section, we do not assume ρ to be orthogonally additive. This allows, for instance, application to Lorentz type L_p spaces. If ρ is a semimodular or a modular, then L_ρ can be equipped with an F-norm $\|\cdot\|_\rho$ and one considers the corresponding F-norm minimization problem. This paper gives several existence theorems relating to this problem, a theorem comparing the set of all best ρ-approximants with the set of all best $\|\cdot\|_\rho$-approximants and a uniqueness theorem.

INTRODUCTION

In this paper we consider the existence of best approximants in modular function spaces by elements of sublattices. Modular function spaces are the natural generalization of L_p, $p>0$, Orlicz, Lorentz, Marcinkiewicz, and Köthe spaces. In Preliminaries, we give some basic concepts and facts of the theory. For further information the reader is referred to [10–12]. In

* Fulbright Scholar, on leave from the Jagiellonian University at Kraków, Poland.

0021-9045/90 $3.00

[23], Musielak gives a thorough exposition on the general theory of both modular spaces and generalized Orlicz spaces. For information about classical Orlicz spaces see [16, 18, 30], and for some generalizations see [8, 9, 29]. The theory of modular spaces has proven to be useful in approximation theory [14, 15, 20–22, 24, 25], as well as in interpolation theory [5, 12, 17], and in operator theory [6, 13].

Let ρ be a function pseudomodular satisfying the Fatou property (see the remark after Definition 1.5), L_ρ the corresponding modular function space, and C a sublattice of L_ρ. Given a function $f \in L_\rho$, we consider the minimization problem of finding $h \in C$ such that $\rho(f-h) = \inf\{f-g: g \in C\}$. Such an h is called a *best ρ-approximant*. For example, if L_ρ is an Musielak–Orlicz space (see Example 0.11), this problem is the problem of finding $h \in C$ such that

$$\int \varphi(x, f(x)-h(x))\, d\mu(x) = \inf_{g \in C} \int \varphi(x, f(x)-g(x))\, d\mu(x).$$

Problems of finding best approximants are important in approximation theory and in probability theory. In the case where C is $L_\rho(\mathscr{B})$ for some σ-subalgebra $\mathscr{B}$ of the original σ-algebra, finding best approximants is closely related to the problem of nonlinear prediction (see, e.g., [3]). For instance, if $\mathscr{B}$ is the σ-subalgebra generated by $\{B_k\}$, this problem of finding best approximants can be stated as follows: Given a random variable $f \in L_\rho$, find a function h, constant on each B_k, such that $\rho(f-h)$ is minimal. In many cases $\rho(f-h)$ represents the loss of information or the average error suffered when f is replaced by h.

Best ρ-approximants are known by many different names in specific situations. When C is $L_\rho(\mathscr{B})$, for a σ-subalgebra $\mathscr{B}$, best approximants in L_2 are known as conditional expectations; in L_p, for $p > 1$, as p-predictors [1]; and in L_1 as conditional medians [27]. When C is an order closed sublattice of L_p, they are known as p-means [2], and in Orlicz spaces as φ-approximants [19]. In this paper ρ is assumed to be a pseudomodular; hence our results are applicable in all the above spaces as well as in many others. For example, ρ need not be of symmetric type, so our results are applicable in Musielak–Orlicz spaces. Moreover, except in parts of Section 4, we do not assume ρ to be orthogonally additive. This allows, for instance, application to Lorentz spaces.

If ρ is a semimodular or a modular, then L_ρ can be equipped with an F-norm $\|\cdot\|_\rho$ (Definition 0.6), and one considers the corresponding F-norm minimization problem. This paper gives several existence theorems relating to this problem and Theorem 5.4 compares the set of all best ρ-approximants with the set of all best $\|\cdot\|_\rho$-approximants. Let us emphasize that best approximants are usually not unique; however, Section 5 gives some exceptions.

The existence theorems presented in this paper can be used for proving some convergence results that are closely related to the theory of martingales. Moreover, by using the existence of best approximants we can describe some properties of modular function spaces. Sets of best approximants in Musielak–Orlicz spaces are described in [7].

PRELIMINARIES: MODULAR FUNCTION SPACES

Let us begin with basic definitions and well-known properties of modular function spaces. Before giving the definition of a function modular let us first recall the following.

DEFINITION 0.1 [23]. Let V be a vector space over $\mathbb{R}$.

(a) If $\rho: V \to [0, \infty]$ satisfies

(1) $\rho(0) = 0$,

(2) $\rho(-v) = \rho(v)$ for every $v \in V$, and

(3) $\rho(\alpha u + \beta v) \leqslant \rho(u) + \rho(v)$, for every $u, v \in V$ whenever $\alpha, \beta \geqslant 0$ and $\alpha + \beta = 1$,

then ρ is a *pseudomodular.*

(b) If a pseudomodular ρ satisfies

(4) $v = 0$, whenever $\rho(\lambda v) = 0$ for all $\lambda > 0$,

then ρ is a *semimodular.*

(c) If a semidomodular ρ satisfies

(5) $v = 0$, whenever $\rho(v) = 0$,

then ρ is a *modular.*

Let X be a nonempty set, Σ a σ-algebra of subsets of X and $\mathscr{P} \subset \Sigma$ a δ-ring such that

(i) $\mathscr{P}$ is an ideal in Σ, that is $E \cap A \in \mathscr{P}$, whenever $E \in \mathscr{P}$ and $A \in \Sigma$, and

(ii) there exists a nondecreasing sequence of sets $\{X_k\}_1^\infty \subset \mathscr{P}$ such that $X = \bigcup_{k=1}^\infty X_k$.

By $\mathscr{E}$ we denote the linear space of all simple real valued functions of the form

$$s = \sum_{k=1}^{n} r_k 1_{A_k},$$

where each $r_k \in \mathbb{R}$, $\{A_k\}_1^n \subset \mathscr{P}$ is a disjoint family and 1_A denotes the characteristic function of a set A. By

$$M_\infty(X, \Sigma, \mathscr{P})$$

we mean the set of all functions $f: X \to [-\infty, \infty]$ such that there exists a sequence of simple functions from $\mathscr{E}$ converging to f pointwise. Similarly

$$M(X, \Sigma, \mathscr{P}) = \{f \in M_\infty(X, \Sigma, \mathscr{P}): |f(x)| < \infty \text{ for each } x \in X\}.$$

DEFINITION 0.2. (Cf. [10, 11, 12]).

(a) A mapping $\rho: \mathscr{E} \times X \to [0, \infty]$ is called a *function pseudomodular* if it satisfies the following properties:

(1) $\rho(0, A) = 0$ for each $A \in \Sigma$.

(2) $\rho(f, A) \leqslant \rho(g, A)$, if $|f(x)| \leqslant |g(x)|$ for every $x \in A$, and $A \in \Sigma$.

(3) $\rho(f, \cdot): \Sigma \to [0, \infty]$ is a σ-subadditive measure for each $f \in \mathscr{E}$.

(4) $\rho(\alpha, A) \to 0$ whenever $\alpha \to 0$ for every $A \in \mathscr{P}$. (Here α denotes the constant function with value α.)

(5) $\rho(\alpha, A_n) \to 0$ for every $\alpha \in \mathbb{R}$, whenever $A_n \downarrow \phi$ and $\{A_n\}_1^\infty \subset \mathscr{P}$.

(b) A function pseudomodular ρ is called a *function semimodular* if it satisfies the following property:

(6) There exists $\alpha_0 \geqslant 0$ such that $\rho(\beta, A) = 0$ for every $\beta \in \mathbb{R}$ whenever $A \in \mathscr{P}$ and $\rho(\alpha, A) = 0$ for some $\alpha > \alpha_0$.

(c) A function semimodular ρ satisfying (6) above with $\alpha_0 = 0$ is called a *function modular.*

(d) The definition of ρ is then extended to all functions $f \in M_\infty(X, \Sigma, \mathscr{P})$ and $E \in \Sigma$ by defining that

$$\rho(f, E) = \sup\{\rho(g, E): g \in \mathscr{E} \text{ and } |g| \leqslant |f| \text{ on } E\}.$$

For the sake of simplicity, $\rho(f)$ is written in place of $\rho(f, X)$.

Some examples are given at the end of this section.

THEOREM 0.3 [10]. *Each function pseudomodular (respectively function semimodular and function modular) is a pseudomodular (respectively semimodular and modular).*

Two important basic notions are those of ρ-null sets and the relation of equality ρ-a.e. They play the same role as sets of measure zero and equality a.e. in L^p and Orlicz spaces.

DEFINITION 0.4. Let ρ be a function pseudomodular.

(a) A set $A \in \Sigma$ is said to be *ρ-null* if $\rho(g, A) = 0$ for every $g \in \mathscr{E}$.

(b) A property $P(x)$ is said to hold *ρ-almost everywhere*, (ρ-a.e.), if the set

$$\{x \in X : P(x) \text{ does not hold}\}$$

is ρ-null.

(c) The set of all ρ-null sets from Σ is denoted by $\mathscr{N}_\rho$.

As usual we identify any pair of measurable sets whose symmetric difference is ρ-null as well as any pair of measurable functions differing only on a ρ-null set. With this in mind we make the following definition.

DEFINITION 0.5. We define

$$M(X, \Sigma, \mathscr{P}, \rho) = \{f \in M_\infty(X, \Sigma, \mathscr{P}) : f \text{ is finite } \rho\text{-a.e.}\},$$

where each $f \in M(X, \Sigma, \mathscr{P}, \rho)$ is actually an equivalence class of functions equal ρ-a.e.

Where no confusion exists M or $M(X, \Sigma)$ is written in place of $M(X, \Sigma, \mathscr{P}, \rho)$.

DEFINITION 0.6 [10]. Let ρ be a pseudomodular.

(a) A *modular function space* is the vector space $L_\rho(X, \Sigma)$, or briefly L_ρ, defined by

$$L_\rho = \{f \in M : \rho(\lambda f) \to 0 \text{ as } \lambda \to 0\}.$$

(b) If ρ is a function semimodular, then the formula

$$\|f\|_\rho = \inf\{\alpha > 0 : \rho(f/\alpha) \leqslant \alpha\}$$

defines the *ρ-norm* in L_ρ.

THEOREM 0.7 [10, 11, 12]. *Let ρ be a function semimodular.*

(a) *$(L_\rho, \|\cdot\|_\rho)$ is an F-space, (i.e., $\|\cdot\|_\rho$ is an F-norm and the metric space L_ρ with $d(f, g) = \|f - g\|_\rho$ is complete).*

(b) *$\|f_n\|_\rho \to 0$ if and only if $\rho(\alpha f_n) \to 0$ for every $\alpha > 0$.*

We also use another type of convergence in L_ρ.

DEFINITION 0.8. Let ρ be a pseudomodular.

(a) We say $\{f_n\}_1^\infty$ *ρ-converges to f* and write $f_n \xrightarrow{\rho} f$, if there exists $\lambda > 0$ such that $\rho(\lambda(f_n - f)) \to 0$ as $n \to \infty$.

(b) A set $D \subset L_\rho$ is *ρ-closed* if $f \in D$, whenever $f \in L_\rho$ and $f_n \xrightarrow{\rho} f$ for some $\{f_n\}_1^\infty \subset D$. Note that by Theorem 0.7(b), $\|\cdot\|_\rho$-convergence implies ρ-convergence and consequently a set is $\|\cdot\|_\rho$-closed if it is ρ-closed.

DEFINITION 0.9. Let ρ be a function pseudomodular.

(a) ρ is *left continuous* if for each $f \in M$,

$$\rho(\lambda f) \uparrow \rho(f) \qquad \text{as} \quad \lambda \uparrow 1.$$

(b) ρ is *right continuous* if for each $f \in M$,

$$\rho(\lambda f) \downarrow \rho(f) \qquad \text{as} \quad \lambda \downarrow 1.$$

(c) ρ has the *Fatou property* if $\rho(f_n) \uparrow \rho(f)$, whenever $|f_n| \uparrow |f|$ ρ-a.e. for $f, f_n \in M$.

PROPOSITION 0.10 [10].

(i) *A function pseudomodular ρ is left continuous if and only if ρ satisfies the Fatou property.*

(ii) *If ρ is a left continuous function semimodular, then*

(a) $\|f_n\|_\rho \uparrow \|f\|_\rho$ *whenever* $|f_n| \uparrow |f|$ ρ-a.e.

(b) $\|\cdot\|_\rho$ *is a function modular, that is* $\tilde{\rho}: \mathscr{E} \times \Sigma \to [0, \infty]$ *defined by*

$$\tilde{\rho}(f, A) = \|f 1_A\|_\rho$$

is a function modular, $\sup\{\tilde{\rho}(g, A): g \in \mathscr{E}$ *and* $|g| \leqslant |f|$ *on* $A\} = \|f 1_A\|_\rho$ *for each* $f \in M$ *and* $L_\rho = L_{\tilde{\rho}}$.

(c) $\rho(f) \leqslant 1$ *if and only if* $\|f\|_\rho \leqslant 1$, *and*

(d) $\rho(f/\|f\|_\rho) \leqslant \|f\|_\rho$.

This section is concluded by some examples.

EXAMPLE 0.11 (The Musielak–Orlicz Modular). See, e.g., [23]. Let

$$\rho(f, A) = \int_A \varphi(x, f(x))\, d\mu(x),$$

where μ, a σ-finite measure on X, and $\varphi: X \times \mathbb{R} \to [0, \infty)$ satisfy the following.

(a) $\varphi(x, u)$ is a continuous even function of u, nondecreasing on $\mathbb{R}^+$, such that $\varphi(x, 0)=0$, $\varphi(x, u)>0$ for $u \neq 0$, and $\varphi(x, u) \to \infty$, as $u \to \infty$.

(b) $\varphi(x, u)$ is a measurable function of x for each $u \in \mathbb{R}$.

The corresponding modular function space is called a Musielak–Orlicz space (or a generalized Orlicz space), and is denoted by L^φ. If φ does not depend on the first variable, then L^φ is called an Orlicz space. If $\varphi(u)=|u|^p$, for $p>0$, then L^φ is isomorphic to L^p.

EXAMPLE 0.12 [4, 10]. Let

$$\rho(f, A)=\sup_{\mu \in \Omega} \int_A \varphi(x, f(x))\, d\mu(x),$$

where φ is as in Example 0.11 and Ω is a family of positive measures such that $\sup_{\mu \in \Omega} \mu(X)<\infty$. Then ρ is a function modular.

EXAMPLE 0.13 (Lorentz type L^p-spaces [4, 10]). Let

$$\rho(f, A)=\sup_{\tau \in \mathscr{T}} \int_A |f(x)|^p\, d\mu_\tau(x),$$

where μ is a fixed σ-finite measure on X, $\mathscr{T}$ is a family of measurable transformations $\tau: X \to X$, and

$$\mu_\tau(E)=\mu(\tau^{-1}(E)).$$

Then ρ is a function modular.

EXAMPLE 0.14. Let $X=\mathbb{N}$, let Σ be the σ-algebra of all subsets of $\mathbb{N}$, and let $\mathscr{P}$ be the δ-ring of all finite subsets of $\mathbb{N}$. Let $I_n=\{1, 2, ..., n\}$ and define

$$\rho(f, A)=\sup_n \frac{1}{n} \sum_{k \in A \cap I_n} (e^{|f(k)|}-|f(k)|-1).$$

Then ρ is a function modular.

EXAMPLE 0.15. Given a sequence of function semimodulars $\{\rho_k\}_1^\infty$, such that $\mathscr{N}_{\rho_k}=\mathscr{N}_{\rho_m}$ for each m and k, define

$$\rho(f, A)=\sum_{k=1}^{\infty} \frac{1}{2^k} \frac{\rho_k(f, A)}{1+\rho_k(f, A)}.$$

If we follow the convention that $\infty/\infty=1$, then we obtain a function semimodular again.

EXAMPLE 0.16. In the same situation as (0.15), define

$$\rho(f, A) = \sup_k \rho_k(f, A).$$

ρ is a function semimodular if and only if

(a) $\sup_k \rho_k(\alpha, A_m) \to 0$ as $A_m \downarrow \phi$ for each $\alpha \in \mathbb{R}$, where each $A_m \in \mathscr{P}$,

(b) $\sup_k \rho_k(\alpha, A) \to 0$ as $\alpha \to 0$ for each $A \in \mathscr{P}$.

This construction is used in Section 4.

SECTION 1

The following notions are new and are of frequent use.

DEFINITION 1.1. If ρ is a nonzero function pseudomodular on X, we define

$$m_\rho(E) = \sup\{\rho(g, E) \colon g \in M\} \in [0, \infty].$$

Remarks 1.2. (a) By convention $m_\rho = m_\rho(X)$.

(b) If $\mathscr{E} \subset B \subset M$, then for each $E \in \Sigma$,

$$m_\rho(E) = \sup_{g \in B} \rho(g, E).$$

For example B could be L_ρ.

(c) $E \in \Sigma$ is ρ-null if and only if $m_\rho(E) = 0$.

(d) If ∞ denotes the constant function with value ∞, then

$$\rho(\infty, E) = m_\rho(E) \qquad \text{for all} \quad E \subset \Sigma.$$

DEFINITION 1.3. For any function $f \in L_\rho$, define

$$\beta_f = \sup\{\beta \geqslant 0 \colon \rho(\beta f) < m_\rho\} \in [0, \infty].$$

DEFINITION 1.4. Given $f \in M$, $f \colon X \to \mathbb{R}$, we define the function $r_f \colon [0, \beta_f] \to [0, m_\rho]$ by

$$r_f(t) = \rho(tf),$$

where by convention $r_f(\infty) = m_\rho$.

DEFINITION 1.5. Let $\mathscr{R}_m$ (respectively $\mathscr{R}_s$ and $\mathscr{R}_p$), be the set of all nonzero function modulars ρ (respectively semimodulars and pseudomodulars), such that for every $f \in M$, $r_f \colon [0, \beta_f] \to [0, m_\rho]$ is continuous.

Remarks. (a) In this paper we usually assume that $\rho \in \mathscr{R}_p$, $\mathscr{R}_s$, or $\mathscr{R}_m$. From this assumption it follows immediately that ρ is a left continuous pseudomodular and therefore by Proposition 0.10 has the Fatou property.

(b) If $\beta_f > 1$, then $\rho(\lambda_n f) \downarrow \rho(f)$ whenever $\lambda_n \downarrow 1$; i.e., ρ is right continuous at f. This does not mean that ρ is right continuous at all functions (not even those from L_ρ). For instance, consider an Orlicz space L^φ such that φ does not satisfy Δ_2. Then $m_\rho = \infty$ and $\beta_f > 1$ for some functions (for example, bounded functions), while $\beta_f \leqslant 1$ for some other functions [16].

PROPOSITION 1.6. *Let $\rho \in \mathscr{R}_p$ and $f \in M$. Then $f \in L_\rho$ if and only if $\beta_f > 0$.*

Proof. Let $f \in L_\rho$; then $\rho(\lambda f) \to 0$ as $\lambda \to 0$ and consequently $\rho(\lambda f) < m_\rho$ for λ close to zero. (Recall: $m_\rho > 0$.) Conversely, take $f \in M$ and assume that $\beta_f > 0$. Let $0 \leqslant \lambda_n \to 0$; then for n sufficiently large, $0 \leqslant \lambda_n \leqslant \beta_f$ and by the preceeding remark we have $\rho(\lambda_n f) = r_f(\lambda_n) \to 0$, i.e., $f \in L_\rho$.

The above proposition immediately implies our next result, which is used frequently throughout the paper.

PROPOSITION 1.7. *If $\rho \in \mathscr{R}_p$, then $L_\rho = \{f \in M : \exists \lambda > 0$ with $\rho(\lambda f) < m_\rho\}$.*

EXAMPLE 1.8. In Musielak–Orlicz spaces, as defined in Example 0.11, $\rho \in \mathscr{R}_m$ and $m_\rho = \infty$.

EXAMPLE 1.9. If $\Psi : [0, \infty) \to [0, a]$ is strictly increasing and continuous, with $\Psi(0) = 0$ and if ρ_1 is a Musielak–Orlicz modular, then $\rho \in \mathscr{R}_m$, where $\rho(f) = \Psi(\rho_1(f))$, and satisfies that $m_\rho = a$. For instance, if $\rho(f) = \rho_1(f)/(1 + \rho_1(f))$ (with the convention $\infty/\infty = 1$), then $m_\rho = 1$.

EXAMPLE 1.10. In Lorentz type L^p-spaces (see Example 0.13),

$$\rho(f, A) = \sup_{\tau \in \mathscr{T}} \int_A |f(x)|^p \, d\mu_\tau(x)$$

is in $\mathscr{R}_m$ and $m_\rho = \infty$.

The following result gives some versions of the Fatou property that we use frequently. The proofs are standard and are omitted.

LEMMA 1.11. *Let $\rho \in \mathscr{R}_p$.*

(a) *Let $g_n \to g$, as $n \to \infty$, with $g_n, g \in M$ and* $\liminf_n \rho(g_n) = a$. *Then $\rho(g) \leqslant a$. Furthermore if $a < m_\rho$, then $g \in L_\rho$.*

(b) *If $\{g_n\}_1^\infty \subset M$ and each $g_n \geqslant 0$ then* $\rho(\liminf_n g_n) \leqslant \liminf_n \rho(g_n)$.

DEFINITION 1.12. A subset C of L_ρ is called a *lattice* if $f_1 \wedge f_2 \in C$ and $f_1 \vee f_2 \in C$, whenever $f_1, f_2 \in C$. Recall that $(f_1 \wedge f_2)(t) = \min\{f_1(t), f_2(t)\}$ and $(f_1 \vee f_2)(t) = \max\{f_1(t), f_2(t)\}$.

DEFINITION 1.13. A set $C \subset L_\rho$ is called *order closed* in L_ρ, if whenever $f \in L_\rho$ with $\{f_n\}_1^\infty \subset C$ such that $f_n \uparrow f$ or $f_n \downarrow f$, then $f \in C$.

DEFINITION 1.14. We define $\mathscr{L}_i$ to be the family of all lattices C in L_ρ such that for any sequence $\{g_k\}_1^\infty \subset C$, $\bigwedge_{k=1}^\infty g_k \in L_\rho$; and $\mathscr{L}_s$ to be the family of all lattices C in L_ρ such that for any sequence $\{g_k\}_1^\infty \subset C$, $\bigvee_{k=1}^\infty g_k \in L_\rho$.

DEFINITION 1.15. (a) We understand the *ρ-distance*, respectively *$\|\cdot\|_\rho$-distance*, from an $f \in L_\rho$ to a set $D \subset L_\rho$ to be the quantities

$$\operatorname{dist}_\rho(f, D) = \inf\{\rho(f-h) \colon h \in D\}$$

and

$$\operatorname{dist}_{\|\cdot\|_\rho}(f, D) = \inf\{\|f-g\|_\rho \colon g \in D\}.$$

(b) The set of all *best ρ-approximants*, respectively *best $\|\cdot\|_\rho$-approximants*, of f with respect to D are denoted by

$$P_\rho(f, D) = \{g \in D \colon \rho(f-g) = \operatorname{dist}_\rho(f, D)\}$$

and

$$P_{\|\cdot\|_\rho}(f, D) = \{g \in D \colon \|f-g\|_\rho = \operatorname{dist}_{\|\cdot\|_\rho}(f, D)\}.$$

(c) If D satisfies $P_\rho(f, D) \neq \phi$, respectively $P_{\|\cdot\|_\rho}(f, D) \neq \phi$, for every $f \in L_\rho$, we say D is *ρ-proximinal*, respectively *$\|\cdot\|_\rho$-proximinal* (see, e.g., [28]).

PROPOSITION 1.16. *Let $\rho \in \mathscr{R}_s$ and let V be any vector subspace of L_ρ. Then*

$$\sup_{h \in V} \|h\|_\rho = \sup_{g \in V} \rho(g).$$

Proof. In order to prove that $\sup_{h \in V} \|h\|_\rho \leqslant \sup_{g \in V} \rho(g)$ we may without loss of generality assume that

$$\sup_{g \in V} \rho(g) = \alpha < \infty.$$

Let $g \in V$. Since V is a vector space, $g/\alpha \in V$ and hence

$$\rho(g/\alpha) \leqslant \sup_{h \in V} \rho(h) = \alpha.$$

By the definition of $\|\cdot\|_\rho$ it follows that $\|g\|_\rho \leqslant \alpha = \sup_{h \in V} \rho(h)$ and because we chose g arbitrarily from V,

$$\sup_{g \in V} \|g\|_\rho \leqslant \sup_{h \in V} \rho(h).$$

In order to prove that $\sup_{h \in V} \|h\|_\rho \geqslant \sup_{g \in V} \rho(g)$, let $\alpha = \sup_{h \in V} \|h\|_\rho$. Define a left continuous function semimodular ρ_1 by $\rho_1(f) = (1/\alpha)\, \rho(f)$. It is well known that $\|g\|_{\rho_1} = (1/\alpha)\, \|\alpha g\|_\rho$ for every $g \in V$. From this it follows that

$$\sup_{h \in V} \|h\|_{\rho_1} = \sup_{h \in V} \frac{1}{\alpha}\, \|\alpha h\|_\rho = \frac{1}{\alpha} \sup_{h \in V} \|h\|_\rho = 1$$

and that $1/\|h\|_{\rho_1} \geqslant 1$ for every $h \in V$. Therefore $\rho_1(h) \leqslant \rho_1(h/\|h\|_{\rho_1})$ for every $h \in V$. By the left continuity of ρ_1, $\rho_1(h/\|h\|_{\rho_1}) \leqslant \|h\|_{\rho_1}$ for every $h \in V$. Combining these we obtain that

$$\frac{1}{\alpha}\, \rho(h) = \rho_1(h) \leqslant \rho_1\left(\frac{h}{\|h\|_{\rho_1}}\right) \leqslant \|h\|_{\rho_1} = \frac{1}{\alpha}\, \|\alpha h\|_\rho$$

for every $h \in V$. Thus

$$\sup_{h \in V} \rho(h) \leqslant \sup_{h \in V} \|\alpha h\|_\rho = \sup_{h \in V} \|h\|_\rho,$$

which completes the proof.

PROPOSITION 1.17. *Let $\rho \in \mathscr{R}_s$, let $f \in M$, and let V be any vector subspace of L_ρ.*

(a) *If $\|f\|_\rho < m_\rho$, then $f \in L_\rho$.*

(b) *If $\|f\|_\rho \geqslant m_\rho$, then $\|f\|_\rho \geqslant \sup_{g \in V} \|g\|_\rho$.*

Proof of (a). Since ρ is left continuous, it follows that

$$\rho\left(\frac{f}{\|f\|_\rho}\right) \leqslant \|f\|_\rho < m_\rho.$$

Hence it follows from Proposition 1.7, that $f \in L_\rho$.

Proof of (b). Immediate from Proposition 1.16.

PROPOSITION 1.18. *If $\rho \in \mathscr{R}_p$, the following assertions are equivalent:*

(a) *If $E \in \Sigma \backslash \mathscr{N}_\rho$, then $m_\rho(E) = m_\rho$.*

(b) *If $\mathscr{E} \subset B \subset M$, then $\sup_{g \in B} \rho(g, E) = \sup_{h \in B} \rho(h)$ for every $E \in \Sigma \backslash \mathscr{N}_\rho$.*

(c) *If $f \in M_\infty$ and $\rho(\lambda f) < m_\rho$ for some $\lambda > 0$, then $f \in L_\rho$.*

Proof. Proof that (a) and (b) are equivalent is easy and is omitted. See Remark 1.2.

To prove that (c) follows from (a), consider $f \in M_\infty$ such that $\rho(\lambda f) < m_\rho$ for some $\lambda > 0$. Denote

$$E = \{x : |f(x)| = \infty\}.$$

By Proposition 1.7 it suffices to show that $E \in \mathscr{N}_\rho$. By the definition of $m_\rho(E)$

$$m_\rho(E) = \rho(\lambda f, E) \leqslant \rho(\lambda f) < m_\rho;$$

hence by (a), $E \in \mathscr{N}_\rho$.

In order to prove that (a) follows from (c), suppose that $E \in \Sigma$ and that $m_\rho(E) < m_\rho$. Define

$$f(x) = \begin{cases} \infty & \text{if } \ x \in E \\ 0 & \text{if } \ x \in X \backslash E. \end{cases}$$

Then $\rho(f) = \rho(f, E) = m_\rho(E) < m_\rho$ and by (c) $f \in L_\rho$, which forces f to be finite ρ-a.e. This shows that $E \in \mathscr{N}_\rho$, completing the proof.

DEFINITION 1.19. If ρ satisfies any of the equivalent statements in Proposition 1.18, ρ is said to have property (K).

Most interesting function pseudomodulars have property (K); for instance, Examples 0.11–0.14. See also Lemma 4.4. In (3.8) we present an example that does not have this property.

SECTION 2

In this section we consider the lattice $C \subset L_\rho$ to be the subspace of all $\mathscr{B}$-measurable functions in L_ρ, where $\mathscr{B}$ is a σ-subalgebra of Σ. If L_ρ is L^p or L^φ and best approximants are unique, then the operator P that assigns the unique best approximant to each $f \in L_\rho$ is known as the prediction operator. See [1, 26].

DEFINITION 2.1. Let $Y \subset X$, let $Y \in \Sigma$, and let $\mathscr{B} \subset \Sigma$ be a σ-algebra of subsets of Y such that there exists a sequence of sets $Y_k \in \mathscr{B} \cap \mathscr{P}$ with $Y = \bigcup_{k=1}^{\infty} Y_i$. For each $h \in M(Y, \mathscr{B})$, we define

$$\tilde{h}(x) = \begin{cases} h(x) & \text{if } \ x \in Y \\ 0 & \text{if } \ x \in X \setminus Y. \end{cases}$$

We consider $M(Y, \mathscr{B}) \subset M(X, \Sigma)$ by the embedding $h \to \tilde{h}$.

Define $\rho_{\mathscr{B}} \colon \mathscr{B} \times \mathscr{E}(Y, \mathscr{B}) \to [0, \infty]$ by

$$\rho_{\mathscr{B}}(h, E) = \rho(\tilde{h}, E) \qquad \text{for} \quad h \in \mathscr{E}(Y, \mathscr{B}) \qquad \text{and} \qquad E \in \mathscr{B}.$$

We call $L_{\rho_{\mathscr{B}}}(Y, \mathscr{B})$ a *modular function subspace* of $L_\rho(X, \Sigma)$.

LEMMA 2.2. *Let $\rho \in \mathscr{R}_p$. If $L_{\rho_{\mathscr{B}}}(Y, \mathscr{B})$ is a modular function subspace of $L_\rho(X, \Sigma)$, then*

(a) $\rho_{\mathscr{B}} \in \mathscr{R}_p$,

(b) $\rho_{\mathscr{B}}(h) = \rho(\tilde{h})$ *for each* $h \in M(Y, \mathscr{B})$, *and*

(c) $L_{\rho_{\mathscr{B}}}(Y, \mathscr{B}) \subset L_\rho(X, \Sigma)$.

(d) *If* $h \in M(Y, \mathscr{B})$ *and* $\tilde{h} \in L_\rho(X, \Sigma)$, *then* $h \in L_{\rho_{\mathscr{B}}}(Y, \mathscr{B})$.

The proof of this lemma is standard and is omitted.

In the next result we characterize $\mathscr{B}$-measurable functions in term of *$\mathscr{B}$-atoms*. By a $\mathscr{B}$-atom, we mean a nonempty set $A \subset \mathscr{B}$ such that whenever a nonempty set $D \subset A$ is $\mathscr{B}$-measurable, $D = A$. The proof is also standard and is omitted as well.

LEMMA 2.3. *Let $Y = \bigcup_{k=1}^{\infty} A_k$, where $\mathscr{B}$ is a σ-algebra on Y and each A_k is a $\mathscr{B}$-atom. Then given $h \colon Y \to \mathbb{R}$, $h \in M(Y, \mathscr{B})$ if and only if h is constant on the $\mathscr{B}$-atoms of Y.*

We are now ready to present our first existence result.

THEOREM 2.4. *Let $\rho \in \mathscr{R}_p$ and let $L_{\rho_{\mathscr{B}}}(Y, \mathscr{B}) \neq \varnothing$ be a modular function subspace of $L_\rho(X, \Sigma)$ such that $Y = \bigcup_{k=1}^{\infty} A_k$, where each A_k is a $\mathscr{B}$-atom. Then whenever $f \in L_\rho(X, \Sigma)$ is bounded on each A_k,*

$$P_\rho(f, L_{\rho_{\mathscr{B}}}(Y, \mathscr{B})) \neq \phi.$$

Proof. Suppose that $f \in L_\rho(X, \Sigma)$ is bounded on each A_k. If

$$\operatorname{dist}_\rho(f, L_{\rho_{\mathscr{B}}}(Y, \mathscr{B})) = \sup_{w \in L_\rho} \rho(w),$$

then

$$\rho(f-g)=\sup_{w\in L_\rho}\rho(w) \qquad \text{for each} \quad g\in L_{\rho_{\mathscr{B}}}(Y,\mathscr{B})$$

and hence

$$P_\rho(f, L_{\rho_{\mathscr{B}}}(Y,\mathscr{B}))=L_{\rho_{\mathscr{B}}}(Y,\mathscr{B})\neq\phi.$$

We may therefore assume that

$$\text{(a)} \qquad \operatorname{dist}_\rho(f, L_{\rho_{\mathscr{B}}}(Y,\mathscr{B}))<\sup_{w\in L_\rho}\rho(w).$$

For each n, choose $h_n\in L_{\rho_{\mathscr{B}}}(Y,\mathscr{B})$ so that

$$\rho(\tilde{h}_n-f)<\operatorname{dist}_\rho(f, L_{\rho_{\mathscr{B}}}(Y,\mathscr{B}))+\frac{1}{n}.$$

Since h_n is $\mathscr{B}$-measurable, h_n is constant on A_k by Lemma 2.3.

We can choose, for every k, a subsequence $\{h_n^k\}_{n=1}^\infty\subset\{h_n\}_{n=1}^\infty$ such that

$$\{h_n^{k+1}\}_{n=1}^\infty\subset\{h_n^k\}_{n=1}^\infty \qquad \text{for every } k$$

and so that whenever $y\in A_k$ either

$$h_n^k(y)\to h(A_k) \qquad \text{as} \quad n\to\infty \qquad \text{for some} \quad h(A_k)\in\mathbb{R}$$

or

$$|h_n^k(y)-f(y)|\uparrow\infty \qquad \text{as} \quad n\to\infty.$$

These choices are made inductively. Given $y\in A_k$, the former choice is made when $\{h_n^{k-1}(y)\}_{n=1}^\infty$ is bounded and the latter choice when it is not. In this second case $\{h_n^k(y)\}_{n=1}^\infty$ is chosen as follows. By Lemma 2.3, if $\{h_n^{k-1}(y)\}_{n=1}^\infty$ is unbounded for some $y\in A_k$, it is unbounded for every $y\in A_k$. Therefore, since f is bounded on each A_k, $\{h_n^{k-1}(y)-f(y)\}_{n=1}^\infty$ is unbounded for every $y\in A_k$ as well. In fact

$$\text{(b)} \qquad \{\inf_{y\in A_k}|h_n^{k-1}(y)-f(y)|\}_{n=1}^\infty \qquad \text{is unbounded.}$$

Suppose $\{h_j^k\}_{j=1}^{n-1}$ has been chosen from $\{h_j^{k-1}\}_{j=1}^\infty$. Since h_{n-1}^k is constant on A_k and f is bounded on A_k, $\{|h_{n-1}^k(y)-f(y)|: y\in A_k\}$ has a supremum, say β. Now by (b) we can choose $h_n^k\in\{h_j^{k-1}\}_{j=1}^\infty$, so that

$$\inf_{y\in A_k}|h_n^k(y)-f(y)|>\beta.$$

Hence we have that

$$|h_n^k(y) - f(y)| \geqslant |h_{n-1}^k(y) - f(y)| \qquad \text{for every} \quad y \in A_k,$$

completing the induction.

Denote

$$G = \bigcup \{A_k : h_n^k(y) \to h(A_k) \text{ for each } y \in A_k\}$$

and $H = Y \backslash G$. Note that G and $H \in \mathscr{B}$.

Define g_n by the diagonal sequence, that is $g_n = h_n^n$. Define $f_n \colon Y \to \mathbb{R}$ by

$$f_n(y) = \begin{cases} g_n(y) & \text{if} \quad y \in G \\ g_k(y) & \text{if} \quad y \in H \cap A_k \end{cases}$$

and finally $h \colon Y \to \mathbb{R}$ by

$$h(y) = \lim_n f_n(y) = \begin{cases} h(A_k) & \text{if} \quad y \in A_k \quad \text{and} \quad A_k \subset G \\ g_k(y) & \text{if} \quad y \in A_k \quad \text{and} \quad A_k \subset H. \end{cases}$$

Each $f_n \in M(Y, \mathscr{B})$ by Lemma 2.3 and hence

$$\text{(c)} \qquad h \in M(Y, \mathscr{B}),$$

because $h = \lim_n f_n$.

For each n, $g_n = h_{m_n}$ for some $m_n \geqslant n$; hence for $y \in A_k$ and each $n \geqslant k$ we have that

$$\begin{aligned} |f_n(y) - f(y)| &= \begin{cases} |g_n(y) - f(y)| & \text{if} \quad y \in G \\ |g_k(y) - f(y)| & \text{if} \quad y \in H \end{cases} \\ &\leqslant |g_n(y) - f(y)| \\ &= |h_{m_n}(y) - f(y)|. \end{aligned}$$

From this it follows that for every $y \in Y$,

$$\text{(d)} \qquad \liminf_n |f_n(y) - f(y)| \leqslant \liminf_n |h_{m_n}(y) - f(y)|.$$

Now by (d) and Lemma 1.11(b),

$$\begin{aligned} \text{(e)} \qquad \rho(\tilde{h} - f) &= \rho(|\tilde{h} - f|) = \rho(\liminf_n |\tilde{f}_n - f|) \\ &\leqslant \rho(\liminf_n |\tilde{h}_{m_n} - f|) \leqslant \liminf_n \rho(|\tilde{h}_{m_n} - f|) \\ &= \operatorname{dist}_\rho(f, L_\rho(Y, \mathscr{B})). \end{aligned}$$

We claim that $h \in P_\rho(f, L_{\rho_{\mathscr{B}}}(Y, \mathscr{B}))$. Recall (a); we are assuming $\operatorname{dist}_\rho(f, L_{\rho_{\mathscr{B}}}(Y, \mathscr{B})) < m_\rho$. From this, Proposition 1.7, and (e) it follows that $\tilde{h} - f \in L_\rho(X, \Sigma)$ and hence that

$$\text{(f)} \qquad \tilde{h} \in L_\rho(X, \Sigma).$$

By (c), (f), and Lemma 2.2(d) we have that $h \in L_{\rho_{\mathscr{B}}}(Y, \mathscr{B})$. This and (e) yield $h \in P_\rho(f, L_{\rho_{\mathscr{B}}}(Y, \mathscr{B}))$, completing the proof.

If ρ is any left continuous function semimodular then it is easy to check that the F-norm $\|\cdot\|_\rho \in \mathscr{R}_m$. (See Prop. 0.10(b).) Hence the following theorem is an immediate corollary of Theorem 2.4.

THEOREM 2.5. *Let ρ be any left continuous function semimodular and let $L_{\rho_{\mathscr{B}}}(Y, \mathscr{B}) \neq \phi$ be a modular function subspace such that $Y = \bigcup_{k=1}^{\infty} A_k$, where each A_k is a $\mathscr{B}$-atom. Then whenever $f \in L_\rho(X, \Sigma)$ is bounded on each A_k, $P_{\|\cdot\|_\rho}(f, L_{\rho_{\mathscr{B}}}(Y, \mathscr{B})) \neq \phi$.*

The following are simple examples of spaces where Theorem 2.4 can be applied to show the existence of a best approximant. This existence is often not evident a priori.

EXAMPLE 2.6. Let $X = Y = \mathbb{R}$, let $\mathscr{B}$ the σ-algebra generated by $\{[n, n+1) : n \in \mathbb{Z}\}$, let ρ be any function pseudomodular, and let f be any continuous function on $\mathbb{R}$. Then by Theorem 2.4 there exists a best ρ-approximant of f by a function constant on each interval $[n, n+1)$.

EXAMPLE 2.7. Let $X = \mathbb{N}$; let $\rho(h) = \sum_{k=1}^{\infty} \varphi(h(k))$, where φ is an Orlicz function (for example $\rho(h) = \sum_{k=1}^{\infty} |h(k)|^p$); and let $Y = \{1, 2, 3, ..., 300\}$. If we define Σ to be all subsets of X and $\mathscr{B}$ to be all subsets of Y, then for any $f \in L_\rho(X, \Sigma)$ we can apply Theorem 2.4 to obtain that $g \in L_\varphi(Y, \mathscr{B})$ (i.e., $g(k) = 0$, for $k > 300$), such that

$$\rho(f - g) = \inf\{\rho(f - h) : h \in L_\varphi(Y, \mathscr{B})\}.$$

(Clearly $g(k) = f(k)$ for $k = 1, 2, ..., 300$ and zero otherwise.)

EXAMPLE 2.8. Modify example (2.6) above by taking $\mathscr{B}$ to be the σ-algebra generated by $\{\{1, 2, 3\}, \{4, 5, 6\}, ..., \{298, 299, 300\}\}$. Given $f \in L_\rho(X, \Sigma)$, there is by Theorem 2.4, a best ρ-approximant of f by a function constant on each set $\{1, 2, 3\}, \{4, 5, 6\}, ..., \{298, 299, 300\}$ and zero otherwise.

EXAMPLE 2.9. Let $Y = X = \mathbb{N}$, let Σ be the σ-algebra of all subsets of X, and let $\mathscr{B}$ be the σ-algebra generated by $\{\{1\}, \{2, 3\}, \{4, 5, 6\}, \{7, 8, 9, 10\}, ...\}$. For each n define the probability measure μ_n by

$$\mu_n(\{k\}) = \begin{cases} \dfrac{1}{n} & \text{for} \quad k = 1, 2, ..., n \\ 0 & \text{for} \quad k = n+1, n+2, ... \end{cases}$$

and then ρ by

$$\rho(h) = \sup_n \int_X |h| \, d\mu_n = \sup_n \sum_{k=1}^{\infty} \mu_n(\{k\}) \, |h(k)| = \sup_n \frac{1}{n} \sum_{k=1}^{n} |h(k)|.$$

Then given $f \in L_\rho(X, \Sigma)$, there is by Theorem 2.4, a best ρ-approximant of f by a function constant on the sets $\{\{1\}, \{2, 3\}, \{4, 5, 6\}, \{7, 8, 9, 10\}, ...\}$.

Section 3

In most of this section we assume X to be a countable set. Under this additional hypothesis we prove several theorems on the existence of best ρ-approximants. Among these is a theorem about modular function subspaces in which we eliminate the hypothesis of f being bounded on $\mathscr{B}$-atoms. We also obtain existence for order closed lattices and then conclude with an interesting consequence of having existence.

Lemma 3.1. *Let $\rho \in \mathscr{R}_p$ have property* (K). *If $X = \{x_k\}_1^\infty$, Σ is the σ-algebra of all subsets of X and $C \subset L_\rho$ is a nonempty sublattice of L_ρ, then for each $f \in L_\rho$, there exists a sequence $\{h_n\}_1^\infty \subset C$ and an $h \in L_\rho$ such that*

(a) *$h_n \to h$ pointwise on X and*

(b) *$\rho(h - f) \leqslant \lim_n \rho(h_n - f) = \operatorname{dist}_\rho(f, C)$.*

Proof. Fix $f \in L_\rho$. If

$$\operatorname{dist}_\rho(f, C) = \sup_{g \in L_\rho} \rho(g),$$

then

$$\rho(f - w) = \sup_{g \in L_\rho} \rho(g) \qquad \text{for each} \quad w \in L_\rho$$

and we may take $h = h_n$ to be any element of C. We may therefore assume that

$$\text{(c)} \qquad \operatorname{dist}_\rho(f, C) < \sup_{g \in L_\rho} \rho(g).$$

Define $D = C - f$ and note that

$$\operatorname{dist}_\rho(f, C) = \operatorname{dist}_\rho(0, D) = \inf_{g \in D} \rho(g).$$

Choose $\{w_n\}_1^\infty \subset D$ such that

$$\rho(w_n) < \operatorname{dist}_\rho(0, D) + 1/n \qquad \text{for each } n.$$

We claim that for each k, $\{|w_n(x_k)|\}_{n=1}^\infty$ is a bounded set. If we suppose otherwise, by passing to an appropriate subsequence, we may assume $|w_n(x_k)| \uparrow \infty$ for some fixed k.

Let $g \in L_\rho$. For sufficiently large n, $|w_n(x_k)| > |g(x_k)|$, which shows that

$$\text{(d)} \qquad \sup_{g \in L_\rho} \rho(g 1_{\{x_k\}}) = \sup_n \rho(|w_n(x_k)|\, 1_{\{x_k\}}).$$

We can now obtain a contradiction as follows:

$$\begin{aligned}
\sup_{g \in L_\rho} \rho(g) &= \sup_{g \in L_\rho} \rho(g 1_{\{x_k\}}) && \text{property (K)} \\
&= \sup_n \rho(|w_n(x_k)|\, 1_{\{x_k\}}) && \text{by (d)} \\
&= \lim_n \rho(|w_n(x_k)|\, 1_{\{x_k\}}) && \text{since} \quad |w_n(x_k)| \to \infty \\
&\leqslant \liminf_n \rho(w_n) && \text{monotonicity of } \rho \\
&\leqslant \operatorname{dist}_\rho(0, D) && \text{since} \quad w_n \in D \\
&< \sup_{g \in L_\rho} \rho(g) && \text{by (c).}
\end{aligned}$$

This contradiction establishes our claim and makes it possible, via a diagonalization argument, to choose a subsequence $\{v_n\}_1^\infty \subset \{w_n\}_1^\infty$ and to define a function w on X, so that for each k, $v_n(x_k) \to w(x_k)$.

We claim $w \in L_\rho$. Since each $v_n = w_m$ for some $m \geqslant n$,

$$\operatorname{dist}_\rho(0, D) \leqslant \rho(v_n) < \operatorname{dist}_\rho(0, D) + 1/n,$$

which shows via Lemma 1.11(a) that

$$\text{(e)} \qquad \rho(w) \leqslant \lim_n \rho(v_n) = \operatorname{dist}_\rho(0, D) < \sup_{g \in L_\rho} \rho(g) \leqslant m_\rho.$$

Thus $w \in L_\rho$ by property (K). (See 1.18(c).)

By defining $h_n = v_n + f$ for all n and $h = w + f$ we immediately satisfy (a) and (e) yields that

$$\rho(h - f) \leqslant \lim_n \rho(h_n - f) = \operatorname{dist}_\rho(0, D) = \operatorname{dist}_\rho(f, C),$$

satisfying (b).

THEOREM 3.2. *Let $X=\{x_k\}_1^\infty$, and let Σ be the σ-algebra of all subsets of X. If $\rho\in\mathscr{R}_p$ has property* (K), *then $L_{\rho_{\mathscr{B}}}(Y,\mathscr{B})$ is ρ-proximinal.*

Proof. By Lemma 3.1, there exists $\{h_n\}\in L_{\rho_{\mathscr{B}}}(Y,\mathscr{B})$ and $h\in L_\rho(Y,\Sigma)$ such that

(a) $h_n\to h$ pointwise on Y and

(b) $\rho(h-f)\leqslant\lim_n\rho(h_n-f)=\operatorname{dist}_\rho(f,C)$.

From (a) we see that $h\in M(Y,\mathscr{B})$ and by Lemma 2.2(b),

$$\rho_{\mathscr{B}}(\lambda h)=\rho(\lambda h)\to 0 \qquad \text{as} \quad \lambda\downarrow 0.$$

Thus $h\in L_{\rho_{\mathscr{B}}}(Y,\mathscr{B})$ and by (b) $h\in P_\rho(f,L_{\rho_{\mathscr{B}}}(Y,\mathscr{B}))$.

THEOREM 3.3. *Let ρ be a left continuous function semimodular, $X=\{x_k\}_1^\infty$, and Σ the σ-algebra of all subsets of X. If ρ has property* (K), *then $L_{\rho_{\mathscr{B}}}(Y,\mathscr{B})$ is $\|\cdot\|_\rho$-proximinal.*

Proof. It is easy to see that $\|\cdot\|_\rho$ is a continuous function modular with property (K). Hence the proof is immediate by Theorem 3.2.

Theorem 3.2 is valid in the following example, although Theorem 2.4 is not.

EXAMPLE 3.4. Let $Y=X=\mathbb{N}$, let Σ be the σ-algebra of all subsets of X, and let $\mathscr{B}$ be the σ-algebra $\{\{\text{evens}\},\{\text{odds}\},\mathbb{N},\phi\}$. Given a probability measure (μ,X,Σ) define $\rho(h)=\sum_{k=1}^\infty\varphi(h(k))\,\mu(\{k\})$, where φ is an Orlicz function. Then given $f\in L_\rho(X,\Sigma)$, there is by Theorem 3.2, a best ρ-approximant of f by a function constant on the evens and on the odds.

THEOREM 3.5. *Let $\rho\in\mathscr{R}_p$, let $X=\{x_k\}_1^\infty$, and let Σ be the σ-algebra of all subsets of X. In addition assume that ρ has property* (K), *that C is a non-empty order closed sublattice of L_ρ, and that*

(a) *$C\in\mathscr{L}_i$ or*

(b) *$C\in\mathscr{L}_s$ (see Definition* 1.14).

Then C is ρ-proximal.

Proof. By Lemma 3.1 there exists a sequence $\{h_n\}\subset C$ and $h\in L_\rho$ such that

(c) $h_n\to h$ pointwise on X and

(d) $\rho(h-f)\leqslant\lim_n\rho(h_n-f)=\operatorname{dist}_\rho(f,C)$.

In order to prove that $h\in P_\rho(f,C)$ it suffices to show that $h\in C$. By hypothesis we have that (a) $\bigwedge_{k=1}^\infty h_k\in L_\rho$ or (b) $\bigvee_{k=1}^\infty h_k\in L_\rho$. Since the other case is similar, we assume (a). Since $\bigwedge_{k=1}^{m} h_k \downarrow \bigwedge_{k=1}^{\infty} h_k$ as $m \to \infty$, and C is order closed, we have that $\bigwedge_{k=1}^{\infty} h_k \in C$. By a well-known inequality $|\bigwedge_{k=n}^{\infty} h_k| \leqslant |h_n| + |\bigwedge_{k=1}^{\infty} h_k|$, so for each n, $\bigwedge_{k=n}^{\infty} h_k \in L_\rho$. Because C is order closed and $\bigwedge_{k=n}^{m} h_k \downarrow \bigwedge_{k=n}^{\infty} h_k$ as $m \to \infty$, we have that $\bigwedge_{k=n}^{\infty} h_k \in C$. The proof is now complete, since C is order closed and $\bigwedge_{k=n}^{\infty} h_k \uparrow h$.

THEOREM 3.6. *Let ρ be a left continuous function semimodular with property* (K). *In addition let $X = \{x_k\}_1^\infty$, and let Σ be the σ-algebra of all subsets of X. If C is a nonempty order closed sublattice of L_ρ and*

(a) $C \in \mathscr{L}_i$ *or*

(b) $C \in \mathscr{L}_s$,

then C is $\|\cdot\|_\rho$-proximinal.

Proof. It is easy to see that $\|\cdot\|_\rho$ is a continuous function modular with property (K). Hence the proof is immediate by Theorem 3.5.

EXAMPLE 3.7. Let L_ρ be any modular function space such that $\rho \in \mathscr{R}_p$ has property (K) and let

$$C = \{h \in L_\rho : h \geqslant 0, \text{ is nondecreasing}\}.$$

Then $C \in \mathscr{L}_i$.

The next example shows the necessity of property (K) in Theorem 3.5.

EXAMPLE 3.8. Let $X = \mathbb{N}$, let Σ be the σ-algebra of all subsets of X, and let

$$\rho(h) = \arctan |h(1)| + 100 \sum_{k=2}^{\infty} \frac{|h(k)|}{2^k}.$$

Furthermore let $C = \{g_n\}_1^\infty$ be the lattice of functions on X satisfying $g_n(1) = n$, and $g_n(k) = 1 - 1/n$ for $k > 1$. Then define $f(k) = 1$ for every $k \in X$. All the hypotheses of Theorem 3.5 are satisfied except ρ does not have property (K). For each g_n in C we have that

$$\rho(f - g_n) = \arctan(n-1) + \frac{100}{n} \sum_{k=2}^{\infty} \frac{1}{2^k} \to \frac{\pi}{2} \qquad \text{as} \quad n \to \infty.$$

Since

$$0 \leqslant \operatorname{dist}_\rho(f, C) \leqslant \rho(f - g_n) \downarrow \frac{\pi}{2} \qquad \text{as} \quad n \to \infty,$$

we have that $\operatorname{dist}_\rho(f, C) \leqslant \pi/2$. However for each $g_n \in C$, $\rho(f - g_n) > \pi/2$. Hence $P_\rho(f, C) = \varnothing$.

It seems reasonable that in order for best approximants of a function f to exist in a lattice C, the lattice must in some sense be closed. Surprisingly the following theorem shows that C must not only be $\|\cdot\|_\rho$-closed, but ρ-closed as well.

THEOREM 3.9. *Let $\rho \in \mathscr{R}_m$ and let $C \subset L_\rho$ be an order closed lattice. If for every order closed lattice DCL_ρ and for every $f \in L_\rho$, $P_\rho(f, D) \neq \varnothing$, then C is ρ-closed.*

Proof. First note that λC is an order closed lattice for each $\lambda > 0$.

Let $\{g_n\}_1^\infty \subset C$, $g \in L_\rho \backslash C$, and $\lambda > 0$ be such that $\rho(\lambda(g_n - g) \to 0$. Then $\operatorname{dist}_\rho(\lambda g, \lambda C) = 0$. However ρ is a function modular, and since $g \neq h$ for every $h \in C$, $\rho(\lambda g - \lambda h) > 0$ for every $h \in C$, which implies that $P_\rho(\lambda_g, \lambda C) = \varnothing$. This contradiction shows that C is ρ-closed, completing the proof.

This result has application in Section 4. The proof of the following theorem is similar.

THEOREM 3.10. *Let $\rho \in \mathscr{R}_s$ and let $C \subset L_\rho$. If $P_{\|\cdot\|_\rho}(f, C) \neq \varnothing$ for each $f \in L_\rho$, then C is $\|\cdot\|_\rho$-closed.*

SECTION 4

Recall that a function pseudomodular ρ is orthogonally additive if $\rho(f, A \cup B) = \rho(f, A) + \rho(f, B)$ for disjoint A and B. The class of all orthogonally additive function pseudomodulars include such important pseudomodulars as Lebesgue, Orlicz, and Musielak–Orlicz pseudomodulars. For such pseudomodulars, we can obtain more general existence theorems, but first we need a somewhat technical lemma.

LEMMA 4.1. *If $\rho \in \mathscr{R}_p$ is orthogonally additive and if f, f_1, and $f_2 \in L_\rho$, then*

$$\rho(f - f_1 \wedge f_2) + \rho(f - f_1 \vee f_2) = \rho(f - f_1) + \rho(f - f_2).$$

Proof. Define $X_1 = \{x \in X : f_1(x) \geqslant f_2(x)\}$ and $X_2 = X \backslash X_1$. The following computation completes the proof of the lemma:

$$\begin{aligned}\rho(f - f_1 \wedge f_2) + \rho(f - f_1 \vee f_2) &= \rho(f - f_1 \wedge f_2, X_1) + \rho(f - f_1 \wedge f_2, X_2) \\ &\quad + \rho(f - f_1 \vee f_2, X_1) + \rho(f - f_1 \vee f_2, X_2) \\ &= \rho(f - f_2, X_1) + \rho(f - f_1, X_2) \\ &\quad + \rho(f - f_1, X_1) + \rho(f - f_2, X_2) \\ &= \rho(f - f_1) + \rho(f - f_2).\end{aligned}$$

THEOREM 4.2. *If $\rho \in \mathscr{R}_p$ is orthogonally additive and has property* (K) *and if $C \subset L_\rho$ is a nonempty order closed lattice, then C is ρ-proximinal.*

Proof. Fix arbitrary $f \in L_\rho$. If $\operatorname{dist}_\rho(f, C) = m_\rho$, then $P_\rho(f, C) = C$; hence without loss of generality we may assume that there exists $a \geqslant 0$ such that

(a) $\operatorname{dist}_\rho(f, C) = a < m_\rho$.

Let $\{h_n\}_1^\infty \subset C$ satisfy

(b) $\rho(f - h_n) < a + (1/2^n)$ for each n.

Whenever $k \geqslant n$, define $g_n^k = \bigwedge_{j=n}^k h_j$, and then define $g_n = \bigwedge_{j=n}^\infty h_j$. For $x \in X$, let $h(x) = \liminf_n h_n(x)$.

Fix two natural numbers $k \geqslant n$. Then for each $j \geqslant n$, by Lemma 4.1, we have that

(c) $\rho(f - g_n^j \wedge h_{j+1}) + \rho(f - g_n^j \vee h_{j+1}) = \rho(f - g_n^j) + \rho(f - h_{j+1})$.

By (b) and since $g_n^j \vee h_{j+1} \in C$, we have that

$$\rho(f - h_{j+1}) \leqslant a + \frac{1}{2^{j+1}} \leqslant \rho(f - g_n^j \vee h_{j+1}) + \frac{1}{2^{j+1}}.$$

After combining this with (c) we obtain

$$\begin{aligned}\rho(f - g_n^j \wedge h_{j+1}) &+ \rho(f - g_n^j \vee h_{j+1}) \\ &\leqslant \rho(f - g_n^j) + \rho(f - g_n^j \vee h_{j+1}) + \frac{1}{2^{j+1}}.\end{aligned}$$

Consequently we have for $n \leqslant j \leqslant k$ that

$$\rho(f - g_n^{j+1}) = \rho(f - g_n^j \wedge h_{j+1}) \leqslant \rho(f - g_n^j) + \frac{1}{2^{j+1}}.$$

Applying this repeatedly for $j = n, ..., k$, while noting that $g_n^n = h_n$, we obtain

$$\rho(f - g_n^{k+1}) \leqslant \rho(f - g_n^n) + \sum_{n+1}^{k+1} \frac{1}{2^j} < \rho(f - h_n) + \frac{1}{2^n} \leqslant a + \frac{1}{2^{n-1}}.$$

This proves that

$$\liminf_k \rho(f - g_n^{k+1}) \leqslant a + \frac{1}{2^{n-1}},$$

but

$$f - g_n^{k+1} \to f - g_n \qquad \text{as} \quad k \to \infty;$$

hence by Lemma 1.11(a),

$$\rho(f-g_n) \leqslant a + \frac{1}{2^{n-1}} < m_\rho \qquad \text{for sufficiently large } n.$$

By Proposition 1.18(c), $f-g_n$, and hence $g_n \in L_\rho$ for sufficiently large n.

On the other hand C is an order closed lattice in L_ρ and $g_n^k \downarrow g_n$ as $k \to \infty$; hence $g_n \in C$ for sufficiently large n. By Lemma 1.11(a),

$$\text{(d)} \quad \rho(f-h) = \rho(\lim_n (f-g_n)) \leqslant \liminf_n \rho(f-g_n) \leqslant a < m_\rho.$$

Furthermore $g_n = \bigwedge_{j=n}^{\infty} h_j \uparrow \liminf_n h_n = h$. By Proposition 1.18(c) $f-h$, hence $h \in L_\rho$. Therefore, since C is order closed, we have that $h \in C$. This along with (d) completes the proof of the theorem.

We wish to extend the previous result in the following sense. Consider a sequence of orthogonally additive function pseudomodulars ρ_n and $\rho = \sup \rho_n$. Examples show that ρ need not be orthogonally additive; however, Theorem 4.5 shows the existence of best approximants for such ρ.

LEMMA 4.3. *Let $\{\rho_n\}_1^\infty \subset \mathcal{R}_p$ be a family of orthogonally additive function pseudomodulars, such that $\mathcal{N}_{\rho_k} = \mathcal{N}_{\rho_m}$ for all k and m and suppose that $\rho = \sup_n \rho_n \in \mathcal{R}_s$. (See 0.16.) Let $f \in L_\rho \backslash C$ and let C be a nonempty lattice in L_ρ, such that $\{\rho_n\}_1^\infty$ is increasing on $f-C$. That is*

$$\text{(a)} \quad \rho_n(f-g) \leqslant \rho_{n+1}(f-g) \qquad \textit{for each } n \textit{ and every } g \in C.$$

If $h_n \in P_{\rho_n}(f, C)$ for each n, $h = \liminf_n h_n$ and

$$\text{(b)} \qquad \rho_n(f-h_n) = \operatorname{dist}_{\rho_n}(f, C) \leqslant \delta_n \qquad \textit{for all } n,$$

then $\rho(f-h) \leqslant \liminf_n \delta_n$.

Proof. By Lemma 4.1 and then by (b) we have that for every $g \in C$ and every n,

$$\begin{aligned} \rho_n(f-h_n \wedge g) + \rho_n(f-h_n \vee g) &= \rho_n(f-h_n) + \rho_n(f-g) \\ &\leqslant \rho_n(f-h_n \vee g) + \rho_n(f-g), \end{aligned}$$

and hence that $\rho_n(f-h_n \wedge g) \leqslant \rho_n(f-g)$ for every $g \in C$.

Using this and the monotonicity of $\{\rho_n\}_1^\infty$ repeatedly yields that for $k > n$,

$$\begin{aligned} \text{(c)} \quad \rho_n\left(f - \bigwedge_{j=n}^{k} h_j\right) &\leqslant \rho_{n+1}\left(f - \bigwedge_{j=n+1}^{k} h_j\right) \\ &\leqslant \rho_{n+2}\left(f - \bigwedge_{j=n+2}^{k} h_j\right) \leqslant \rho_k(f-h_k). \end{aligned}$$

We now apply Lemma 1.11(a) and (c) to obtain

$$\text{(d)} \qquad \rho_n\left(f - \bigwedge_{j=n}^{\infty} h_j\right) = \rho_n\left(\lim_k \left(f - \bigwedge_{j=n}^{k} h_j\right)\right)$$
$$\leqslant \liminf_k \rho_n\left(f - \bigwedge_{j=n}^{k} h_j\right)$$
$$\leqslant \liminf_k \rho_k(f - h_k) \leqslant \liminf_k \delta_k.$$

We now obtain the desired result by Lemma 1.11(a), the monotonicity of $\{\rho_n\}_1^\infty$ and (d) as follows:

$$\rho(f-h) = \sup_k \rho_k\left(f - \lim_n \bigwedge_{j=n}^{\infty} h_j\right)$$
$$\leqslant \sup_k \liminf_n \rho_k\left(f - \bigwedge_{j=n}^{\infty} h_j\right)$$
$$\leqslant \sup_k \liminf_n \rho_n\left(f - \bigwedge_{j=n}^{\infty} h_j\right) \leqslant \liminf_k \delta_k.$$

LEMMA 4.4. *Let $\{\rho_n\}_1^\infty \subset \mathcal{R}_s$ be a family of function semimodulars with property* (K). *If $\mathcal{N}_{\rho_k} = \mathcal{N}_{\rho_m}$ for all k and m and $\rho = \sup_n \rho_n \in \mathcal{R}_s$, then ρ has property* (K) *as well.*

Proof. By hypothesis, $\mathcal{N}_{\rho_n} = \mathcal{N}_\rho$ for each n; hence $M(X, \Sigma, \rho_n) = M(X, \Sigma, \rho) = M$ for all n. Thus for each $E \in \Sigma \backslash \mathcal{N}_\rho$,

$$\text{(a)} \qquad m_\rho(E) = \sup_{g \in M} \rho(g, E) = \sup_{g \in M} \sup_n \rho_n(g, E)$$
$$= \sup_n \sup_{g \in M} \rho_n(g, E) = \sup_n m_{\rho_n}(E).$$

In particular

$$\text{(b)} \qquad \sup_n m_{\rho_n} = m_\rho.$$

By (a), Proposition 1.18, and (b) we obtain that

$$m_\rho(E) = \sup_n m_{\rho_n}(E) = \sup_n m_{\rho_n} = m_\rho,$$

which shows via Proposition 1.18 that ρ has property (K).

Applying Lemmas 4.3 and 4.4 we can prove the following existence theorem.

THEOREM 4.5. *Let* $\{\rho_n\}_1^\infty \subset \mathscr{R}_p$ *be a family of orthogonally additive function pseudomodulars with property* (K) *and let* $\rho = \sup_n \rho_n \in \mathscr{R}_s$.

(a) *If* C, *a nonempty lattice in* L_ρ, *and* $f \in L_\rho \backslash C$ *are such that* $\{\rho_n\}_1^\infty$ *is increasing on* $f - C$, *then* $P_\rho(f, C) \neq \varnothing$.

(b) *Moreover if* $h_n \in P_{\rho_n}(f, C)$ *for each* n, $h = \liminf_n h_n$ *and* $\operatorname{dist}_\rho(f, C) < m_\rho$ *then* $h \in P_\rho(f, C)$.

Proof. Let us prove (b) first.

Define $\delta_n = \operatorname{dist}_{\rho_n}(f, C)$ for each n. By Lemma 4.3 and the definition of ρ,

(c) $\rho(f-h) \leqslant \liminf_n \delta_n \leqslant \operatorname{dist}_\rho(f, C)$.

Therefore by hypothesis, $\rho(f-h) < m_\rho$. By Lemma 4.4, ρ has property (K), so Proposition 1.18(c) applies, yielding that $f-h$ and hence $h \in L_\rho$. Since C is order closed, $h \in C$. This and (c) complete the proof of part (b).

Now for the proof of (a). If $\operatorname{dist}_\rho(f, C) = m_\rho$, then $P_\rho(f, C) = C$; hence without loss of generality assume that $\operatorname{dist}_\rho(f, C) < m_\rho$. There exists $h_n \in P_{\rho_n}(f, C)$ by Theorem 4.2 for each n. Hence if we define $h = \liminf_n h_n$, the proof of (a) is complete by part (b).

If ρ is an orthogonally additive function modular, then although $\|\cdot\|_\rho$ is a function modular, it is not necessarily orthogonally additive. We cannot therefore apply Theorem 4.2 directly in proving the existence of best $\|\cdot\|_\rho$-approximants; however we can apply Lemma 4.3.

THEOREM 4.6. *Let* $\rho \in \mathscr{R}_s$ *be orthogonally additive and have property* (K). *If* $C \subset L_\rho$ *is a nonempty order closed lattice, then* C *is* $\|\cdot\|_\rho$*-proximinal.*

Proof. Let $f \in L_\rho$. If $f \in C$, the assertion obviously holds; hence without loss of generality we assume $f \notin C$.

Denote $\delta = \operatorname{dist}_{\|\cdot\|_\rho}(f, C)$. Because C is order closed, Theorems 4.2 and 3.9 apply to yield that C is ρ-closed. Therefore $\delta \neq 0$.

First suppose that $\delta \geqslant m_\rho$. By Proposition 1.16, for every $g \in C$,

$$m_\rho \leqslant \delta \leqslant \inf_{w \in C} \|f-w\|_\rho \leqslant \|f-g\|_\rho \leqslant \sup_{w \in L_\rho} \|w\|_\rho \leqslant \sup_{w \in L_\rho} \rho(w) \leqslant m_\rho.$$

Thus $\|f-g\|_\rho = m_\rho$ for every $g \in C$ and $C = P_{\|\cdot\|_\rho}(f, C)$.

Now we assume that $\delta \in (0, m_\rho)$. Take $\{g_n\}_1^\infty \subset C$ such that $\|f-g_n\| \downarrow \delta$. For each n define $\delta_n = \|f-g_n\|_\rho + (1/n)$ and ρ_n by $\rho_n(g) = \rho(g/\delta_n)$. Clearly each ρ_n is a function semimodular. For each $E \in \Sigma$, $m_{\rho_n}(E) = m_\rho(E)$; hence by 1.18(a), each ρ_n has property (K). Note that by the definition of L_ρ, $L_\rho = L_{\rho_n}$. Since $1/\delta_n \uparrow 1/\delta$ as $n \to \infty$, by the left continuity of ρ,

$$\sup_n \rho_n(g) = \sup_n \rho\left(\frac{g}{\delta_n}\right) = \rho\left(\frac{g}{\delta}\right) \qquad \text{for each} \quad g \in L_\rho.$$

Moreover, for every n and for each $g \in C$,

$$\left|\frac{f-g}{\delta_n}\right| \leqslant \left|\frac{f-g}{\delta_{n+1}}\right|, \qquad \text{thus} \quad \rho_n(f-g) \leqslant \rho_{n+1}(f-g);$$

that is, $\{\rho_n\}_1^\infty$ is nondecreasing on $f-C$.

By Theorem 4.2 for every n, there exists $h_n \in P_{\rho_n}(f, C)$. We define $h = \liminf_n h_n$. For each n,

$$\inf\left\{\alpha : \rho\left(\frac{f-g_n}{\alpha}\right) \leqslant \alpha\right\} = \|f-g_n\|_\rho \leqslant \delta_n,$$

hence $\rho_n(f-g_n) = \rho((f-g_n)/\delta_n) \leqslant \delta_n$. From this it now follows that

$$\rho_n(f-h_n) = \operatorname{dist}_{\rho_n}(f, C) \leqslant \rho_n(f-g_n) \leqslant \delta_n.$$

Hence by Lemma 4.3

$$\sup_n \rho_n(f-h) = \rho(f-h) \leqslant \liminf_n \delta_n = \delta.$$

This means that for every n

$$\text{(a)} \qquad \rho\left(\frac{f-h}{\delta_n}\right) = \rho_n(f-h) \leqslant \delta.$$

Recall that as $n \to \infty$, $1/\delta_n \uparrow 1/\delta$; hence by the left continuity of ρ,

$$\rho\left(\frac{f-h}{\delta_n}\right) \uparrow \rho\left(\frac{f-h}{\delta}\right).$$

This combined with (a) yields that

$$\text{(b)} \qquad \rho\left(\frac{f-h}{\delta}\right) \leqslant \delta < m_\rho,$$

and hence $\|f-h\|_\rho \leqslant \delta < m_\rho$. From (b) and Prop. 1.18(c) we obtain that $(f-h)/\delta$, and hence $h \in L_\rho$. Since C is order closed, $h \in C$. This and (b) imply that $h \in P_{\|\cdot\|_\rho}(f, C)$, completing the proof.

SECTION 5

In order to state a uniqueness theorem, let us first recall the following definition [23].

DEFINITION 5.1. A function modular ρ is called *strictly convex* if ρ is convex and if whenever $h, g \in L_\rho$ satisfy

$$\rho(h) = \rho(g) \qquad \text{and} \qquad \rho\left(\frac{h+g}{2}\right) = \frac{\rho(h)+\rho(g)}{2},$$

then $h = g$ ρ-a.e.

EXAMPLE 5.2. An Orlicz modular ρ_φ is strictly convex if and only if φ is strictly convex [23].

THEOREM 5.3. *Let $\rho \in \mathscr{R}_m$ be strictly convex and let $C \subset L_\rho$ be a nonempty convex lattice that is order closed in L_ρ. Let $f \in L_\rho$ be such that* $\operatorname{dist}_\rho(f, C) < \infty$. *Then the set $P_\rho(f, C)$ consists of at most one element.*

Proof. Assume that $P_\rho(f, C)$ contains g and h. That is

$$\rho(f-g) = \rho(f-h) = \operatorname{dist}_\rho(f, C).$$

By the convexity of ρ,

$$\rho\left(\frac{f-g+f-h}{2}\right) \leqslant \frac{\rho(f-g)}{2} + \frac{\rho(f-h)}{2} = \operatorname{dist}_\rho(f, C).$$

However,

$$\frac{g+h}{2} \in C,$$

so

$$\operatorname{dist}_\rho(f, C) \leqslant \rho\left(\frac{f-g+f-h}{2}\right),$$

as well. Thus

$$\rho\left(\frac{f-g+f-h}{2}\right) = \frac{\rho(f-g)}{2} + \frac{\rho(f-h)}{2};$$

hence by the strict convexity of ρ, $g = h$ ρ-a.e., completing the proof.

In the next result we compare the sets of best ρ-approximants and best $\|\cdot\|_\rho$-approximants.

THEOREM 5.4. *Let $\rho \in \mathscr{R}_m$, let $f \in L_\rho$ and let $D \subset L_\rho$ be a nonempty set such that*

(i) $\delta = \operatorname{dist}_{\|\cdot\|_\rho}(f, D) > 0$, *and*

(ii) $P_{\|\cdot\|_\rho}(f, D) \neq \varnothing$.

Then

(a) $\delta P_\rho(f/\delta, D/\delta) \subset P_{\|\cdot\|_\rho}(f, D)$ *and*

(b) *if* $\beta_{f-g} > 1/\delta$ *for each* $g \in D$, *then* $\delta P_\rho(f/\delta, D/\delta) = P_{\|\cdot\|_\rho}(f, D)$.

Proof of (a). Suppose $g/\delta \in P_\rho(f/\delta, D/\delta)$. Then

$$\text{(c)} \qquad \rho\left(\frac{f-g}{\delta}\right) \leqslant \rho\left(\frac{f-h}{\delta}\right) \qquad \text{for each} \quad h \in P_{\|\cdot\|_\rho}(f, D).$$

Furthermore $\|f-h\|_\rho = \delta$ for each $h \in P_{\|\cdot\|_\rho}(f, D)$, hence by the left continuity of ρ,

$$\text{(d)} \qquad \rho\left(\frac{f-h}{\delta}\right) \leqslant \delta \qquad \text{for each} \quad h \in P_{\|\cdot\|_\rho}(f, D).$$

Together (c) and (d) show that $\rho((f-g)/\delta) \leqslant \delta$ and hence that $\|f-g\|_\rho \leqslant \delta$, which completes the proof.

Proof of (b). Let $h \in P_{\|\cdot\|_\rho}(f, D)$. By (a) it suffices to show that $h \in \delta P_\rho(f/\delta, D/\delta)$. To this end let g be any function from D. Hence $\|f-g\|_\rho \geqslant \delta$. Take $0 < \gamma \uparrow \delta$, then $1/\gamma \downarrow 1/\delta$. By hypothesis $1/\gamma \in [0, \beta_{f-g})$, so by (1.5),

$$\rho\left(\frac{f-g}{\gamma}\right) \to \rho\left(\frac{f-g}{\delta}\right) \qquad \text{as} \quad \gamma \uparrow \delta.$$

Now, since $\gamma < \delta \leqslant \|f-g\|_\rho$, we have that $\rho((f-g)/\gamma) > \gamma$. Therefore

$$\delta = \lim \gamma \leqslant \lim \rho\left(\frac{f-g}{\gamma}\right) = \rho\left(\frac{f-g}{\delta}\right).$$

By the left continuity of ρ we have that

$$\rho\left(\frac{f-h}{\|f-h\|_\rho}\right) \leqslant \|f-h\|_\rho,$$

and finally for each $g \in D$,

$$\rho\left(\frac{f-h}{\delta}\right) \leqslant \rho\left(\frac{f-h}{\|f-h\|_\rho}\right) \leqslant \|f-h\|_\rho = \delta \leqslant \rho\left(\frac{f-g}{\delta}\right),$$

that is $h/\delta \in P_\rho(f/\delta, D/\delta)$, as desired.

Remarks. (a) Observe that whenever $\beta_{f-g} = \infty$, $\delta P_\rho(f/\delta, D/\delta) = P_{\|\cdot\|_\rho}(f, D)$. For example, if ρ is a Musielak–Orlicz modular satisfying Δ_2 (see [23]), or if $f - D \subset E_\rho$.

(b) If $\beta_{f-g} \leqslant 1/\delta$, then the inclusion in (a) can be strict.

(c) If $\beta_{f-g} > 1/\delta$, then we conclude from Theorem 5.4 that

$$P_{\|\cdot\|_\rho}(f, C) \neq \varnothing \qquad \text{implies} \quad P_\rho\left(\frac{f}{\delta}, \frac{C}{\delta}\right) \neq \varnothing.$$

ACKNOWLEDGMENT

The authors express their gratitude to Anna Kaminska who suggested the topic of this paper.

REFERENCES

1. T. ANDO AND I. AMEMIYA, Almost everywhere convergence of prediction sequences in L_p $(1 < p < \infty)$, *Z. Wahrscheinlichkeitstheor. verw. Gebiete* **4** (1965), 113–120.
2. H. D. BRUNK, Uniform inequalities for conditional p-means given σ-lattices, *Ann. Probab.* **3** (1975), 1025–1030.
3. M. H. DEGROOT AND M. M. RAO, Multidimensional information inequalities, *in* "Proceedings of an International Symposium on Multivariate Analysis held in Dayton, Ohio, June 14–19, 1965" (P. P. Krishnaiah, Ed.), pp. 287–313, Academic Press, New York/London, 1967.
4. L. DREWNOWSKI AND A. KAMINSKA, Orlicz spaces of vector functions generated by a family of measures. *Comment. Math.* **22** (1981), 175–186.
5. H. HUDZIK, J. MUSIELAK, AND R. URBANSKI, Interpolation of compact sublinear operators in generalized Orlicz spaces of nonsymmetric type, *in* "Topology," Vol. I, Proceedings of the 4th Colloquium, Budapest, 1978; *Coll. Math. Ja. Bolyai* **23** (1980), 625–638.
6. M. A. KHAMSI, W. M. KOZLOWSKI, AND S. REICH, Fixed point theory in modular function spaces. Nonlinear Analysis, *Theory, Methods, and Applications* **14**, No. 11 (1990), 935–953.
7. S. KILMER AND W. M. KOZLOWSKI, On some properties of sets of best approximants in Musielak–Orlicz spaces, *Comment. Math.*, to appear.
8. A. KOZEK, Orlicz spaces of functions with values in Banach spaces, *Comment. Math.* **19** (1976), 259–288.
9. A. KOZEK, Convex integral functionals on Orlicz spaces, *Comment. Math.* **21** (1979), 109–135.
10. W. M. KOZLOWSKI, "Modular Function Spaces," Series on Monographs and Textbooks in Pure and Applied Mathematics, Vol. 122, Dekker, New York/Basel, 1988.
11. W. M. KOZLOWSKI, Notes on modular function spaces, I, *Comment. Math.* **28**, No. 1 (1988), 87–100.
12. W. M. KOZLOWSKI, Notes on modular function spaces, II, *Comment. Math.* **28**, No. 1 (1988), 101–116.
13. W. M. KOZLOWSKI, On domains of some nonlinear operators, *J. Math. Anal. Appl.* **138** (1989), 243–267.
14. W. M. KOZLOWSKI AND G. LEWICKI, On polynomial approximation in modular function spaces, *in* "Proceedings of the International Conference "Function Spaces" held in Poznan, August 1986, Texte zür Mathematik" (J. Musielak, Ed.), Band 103, pp. 63–68, Teubner, Stuttgart, 1988.

15. W. M. Kozlowski and G. Lewicki, Analyticity and polynomial approximation in modular function spaces, *J. Approx. Theory*, **58** (1989), 15–35.
16. M. A. Krasnosel'skii and Y. B. Rutickii, "Convex Function and Orlicz Spaces," Noordhoff, Groningen, 1961.
17. M. Krbec, Modular interpolation spaces, *Z. Anal. Anwendungen* **1** (1982), 25–40.
18. A. Kufner, O. John and S. Fucik, "Function Spaces," Academia, Prague, and Nordhoff Leyden, 1977.
19. D. Landers and L. Rogge, Best approximants in L_φ-spaces, *Z. Wahrscheinlichkeitstheor. verw. Gebiete* **51** (1980), 215–237.
20. G. Lewicki, Bernsteins "lethargy" theorem in metrizable topological linear spaces, submitted for publication.
21. J. Musielak, An application of modular spaces to approximation, *Comment. Math.*, Tomus Specialis in honorem Ladislai Orlicz, I (1979), 251–259.
22. J. Musielak, Modular approximation by a filtered family of linear operators, *in* "Proceed. Confer. Oberwolfach, August 9–16, 1980, Birkhäuser Verlag, 99–110, 1981."
23. J. Musielak, Orlicz spaces and modular spaces, *in* "Lecture Notes in Mathematics," Vol. 1034, Springer-Verlag, Berlin/Heidelberg/New York/Tokyo, 1983.
24. W. Plesniak, Quasianalyticity in F-spaces of Integrable Functions *in* "Approximation and Function Spaces: Proceedings of the International Conference in Gdansk, August 1979" (P. Ciesielski, Ed.), pp. 558–571, PWN Warszawa, North-Holland, Amsterdam/New York/Oxford, 1981.
25. W. Plesniak, Leja's type polynomial condition and polynomial approximation in Orlicz spaces, *Ann. Polon. Math.* **46** (1985), 268–278.
26. M. M. Rao, Abstract nonlinear prediction and operator martingales, *J. Multivariate Anal.* **1** (1971), 129–157.
27. T. Shintani and T. Ando, Best approximants in L_1 space, *Z. Wahrscheinlichkeitstheor. verw. Gebiete* **33** (1975), 33–39.
28. I. Singer, "Best approximation in normed linear spaces by elements of linear subspaces," Springer, Berlin/Heidelberg/New York, 1970.
29. B. Turett, Fenchel–Orlicz spaces, *Dissertationes Math.* **181** (1980), 1–60.
30. A. C. Zaanen, "Integration," North-Holland, Amsterdam/London, 1967.