Nieuw Archief voor Wiskunde (3), XI, 77–98 (1964)

LINEAR FUNCTIONALS ON ORLICZ SPACES [1)]

BY

M. M. RAO

0. *Introduction.* Let $\Phi(\cdot)$ and $\Psi(\cdot)$ be nonnegative convex functions vanishing at the origin and complementary to each other in the sense of Young. (Precise definitions will be given later.) If L^Φ and L^Ψ are Orlicz (function) spaces, which include the usual Lebesgue spaces L^p, $1 \leq p \leq \infty$, defined on some measure space (Ω, Σ, μ), then the representation of continuous linear functionals on L^Φ (and L^Ψ) is an important problem in their study. It has been considered in the past and the solution was given if $\Phi(\cdot)$ satisfies some growth condition (e.g., $\Phi(2x) \leq C\Phi(x)$) *and* the measure μ is at most σ-finite (cf., [3], [10], [11]). Some improvement on the growth of $\Phi(\cdot)$ has been found if Ω is a subset of the line and μ is the Lebesgue measure in [6]. However, the general representation problem for the functionals has not been considered if $\Phi(\cdot)$ (or $\Psi(\cdot)$) is allowed to grow exponentially fast, and the measure is not σ-finite. These considerations are of interest for at least two reasons. First, the corresponding results for L^p spaces are known for arbitrary measures and the analogous results for L^Φ are of interest for comparison and extension purposes. Second, the growth restrictions of $\Phi(\cdot)$ are not natural and seem dictated only by the limitations of certain calculations, the general $\Phi(\cdot)$ being of use in applications (cf., last section). The study of the general problem is moreover of interest in itself.

The purpose of the present paper is to present a global representation for bounded linear functionals on L^Φ without any restrictions on $\Phi(\cdot)$, $\Psi(\cdot)$ or the measure μ (Theorem 3). From this various

---

1) This work was supported under contract No. DA–36–061–ORD–477 and NSF grant GP–1349. Reproduction in whole or in part is permitted for any purpose of the U.S. Government.

specialized results are then obtained when (i) $\Phi(\cdot)$ is restricted but μ is not, and (ii) both $\Phi(\cdot)$ and μ are restricted. These include all the previously known cases. A brief summary of the results is as follows.

The next section contains some preliminaries and needed lemmas. In Section 2 the "conjugate" space of L^{Φ} is introduced, and then the general representation of bounded linear functionals on L^{Φ} is given in Section 3. Section 4 is devoted to various specializations together with a discussion of reflexivity. The final section contains an application to a problem in statistical theory extending an earlier result in [7].

In the present work, the concept of "quasi-functions" (cf., [5], also called "cross-sections" in [9]) and some consequences, as well as the integration relative to finitely additive set functions, were found useful in several places. It should be noted, however, that the demonstrations sometimes take longer and are even more difficult than in the usual special situations, because of the generality here. The generalization of the paper is suggested, in part, by a study of [6], [11] and ([2], Chapters III and IV).

1. *Some lemmas.* In this section a few results, known for finite or σ-finite measures, will be extended for the non σ-finite case as they are needed for the later work. Let $\Phi(\cdot)$ and $\Psi(\cdot)$ be two non-negative symmetric convex functions such that $\Phi(0) = 0 = \Psi(0)$, $\lim_{x\to+\infty} \Phi(x) = \lim_{x\to+\infty} \Psi(x) = +\infty$, and satisfying the (Young's) inequality

$$xy \leq \Phi(x) + \Psi(y), \tag{1}$$

for all x, y. Then $\Phi(\cdot)$ and $\Psi(\cdot)$ are termed complementary Young's functions. If (Ω, Σ, μ) is a measure space, then (following [9]) μ is said to have the *finite subset property* (FSP) if every measurable set of positive μ measure contains a measurable subset of finite positive measure. Let L^{Φ} be the linear space of equivalence classes of measurable functions on (Ω, Σ, μ) such that $f \in L^{\Phi}$ implies $\|f\|_{\Phi} < \infty$, where (cf., [11])

$$\|f\|_{\Phi} = \sup_g \textstyle\int_{\Omega} |fg|\, d\mu, \text{ with } \rho(g) = \textstyle\int_{\Omega} \Psi(g)\, d\mu \leq 1. \tag{2}$$

[Here and in what follows, $f \in L^{\Phi}$ means that f is any member of the equivalence class to which it belongs.]

It is readily verified that $\|\cdot\|_{\Phi}$ is a semi-norm in general and a norm if μ has the FSP and in the latter case L^{Φ} is a Banach (or B-)

space. (See also, [11], p. 79.) Similarly L^{Ψ} is defined and L^{Φ}, L^{Ψ} are called *Orlicz spaces*. In general, for $f \in L^{\Phi}$, if

$$N_{\Phi}(f) = \inf\left\{K > 0, \int_{\Omega} \Phi\left(\frac{f}{K}\right) d\mu \leq 1\right\}, \tag{3}$$

then $N_{\Phi}(\cdot)$ is always a norm and with it L^{Φ} becomes a B-space. If μ has the FSP, then (2) and (3) define the same topology for L^{Φ} in the sense that (cf., [8])

$$N_{\Phi}(f) \leq \|f\|_{\Phi} \leq 2N_{\Phi}(f), \quad f \in L^{\Phi}. \tag{4}$$

In the following for any set $B \subset \Omega$, the symbol B' stands for $\Omega - B$.

Throughout this paper, the finite subset property of μ will be *assumed* without further comment. This assumption rather eliminates the nuisance of sets of infinite measure without any subsets of finite positive measure than really restricting the generality. (See [5] for a discussion on this point.) In what follows, the norms (2) and (3) and the relation (4) will be used according to convenience. Also, hereafter all the functions are taken to be real-valued, but all the results hold for complex functions as well, with simple (mostly notational) modifications.

The first lemma is proved for the σ-finite μ in ([11], p. 80), and for localizable μ in ([8], p. 44) where $\lim_{x\to\infty} \Phi'(x) = \infty$ is also assumed. [Φ' and Ψ' stand in this paper for the derivatives of Φ and Ψ, which exist a.e., (Lebesgue).]

Lemma 1. Let $\Phi(\cdot)$ be a Young's function. Then for every $f(\neq 0)$ in L^{Φ},

$$\int_{\Omega} \Phi\left(\frac{f}{\|f\|_{\Phi}}\right) d\mu \leq 1. \tag{5}$$

Proof. If $\rho(g)$ is as in (2) and $\rho'(g) = \max(1, \rho(g))$, one has

$$\textstyle\int_{\Omega} |fg|\, d\mu \leq \rho'(g)\|f\|_{\Phi}, \tag{6}$$

which follows from the definition of norm in (2) if $\rho(g) \leq 1$ and replacing g by $g/\rho(g)$ if $\rho(g) > 1$. For any A in Σ, $\mu(A) < \infty$, if $f_A = f\chi_A$, then using (6) with f_A, the first part of the classical proof ([11], p. 80) gives (5). This means (5) holds with $\Omega = A$ there.

To prove the general case, let $\nu(E) = \int_E \Phi(f/\|f\|_{\Phi})\, d\mu$ for any E in Σ. Then ν is a measure and has the FSP. Let Σ_1 be the class of all sets A in Σ of finite μ-measure. Then by the preceding paragraph $\sup_{A\in\Sigma_1} \nu(A) = \alpha \leq 1$. Since $\nu(\cdot)$ is a measure there exists a sequence $\{A_n\} \subset \Sigma_1$, $A_n \subset A_{n+1}$, such that $\lim_n \nu(A_n) = \alpha$. If $B = \bigcup_{n=1}^{\infty} A_n$, then $\alpha = \nu(B) \leq 1$. To complete the proof it remains to show that $\nu(B') = 0$. If E_0 is the set on which $\Phi(f/\|f\|_\Phi) = 0$, then $\nu(E_0) = 0$, so consider $E^* = B' \cap E_0'$. If $F \in \Sigma_1$ and $F \subset E^*$, then $\mu(F) = 0$ (and hence $\nu(F) = 0$) as otherwise

$$\alpha < \nu(B) + \nu(F) = \nu(B \cup F) = \lim_n \nu(A_n \cup F) \leq \sup_{D \in \Sigma_1} \nu(D) = \alpha,$$

gives a contradiction. Thus $\mu(E^*) > 0$ is impossible and so $\mu(E^*) = 0$, implying $\nu(E^*) = 0$ or since $B' = E^* \cup (B' \cap E_0)$, it follows that

$$\nu(\Omega) = \nu(B) + \nu(B') = \alpha \leq 1,$$

completing the proof.

The following consequence of the lemma [with (3) and (4)] will be needed: "Hölder inequality": If $f \in L^\Phi$, $g \in L^\Psi$ then

$$\int_\Omega |fg|\, d\mu \leq \|f\|_\Phi N_\Psi(g) \leq \|f\|_\Phi \|g\|_\Psi.$$

Remark. By the above lemma, L^Φ can also be characterized as the set of measurable functions on (Ω, Σ, μ) satisfying

$$\int_\Omega \Phi(Kf)\, d\mu < \infty$$

for *some* $K > 0$. Then the norm is given by (3) with the relation (4). Note also that by the monotonicity of $\Phi(\cdot)$ if $\|f\|_\Phi \leq 1$,

$$\int_\Omega \Phi(f)\, d\mu \leq \|f\|_\Phi \int_\Omega \Phi\left(\frac{f}{\|f\|_\Phi}\right) d\mu, \qquad (\leq \|f\|_\Phi), \tag{7}$$

and in any case (1) and (2) imply $\|f\|_\Phi \leq \int_\Omega \Phi(f)\, d\mu + 1$.

Definition. The set $\tilde{L}^\Phi$ denotes the class of measurable functions on (Ω, Σ, μ) such that $\int_\Omega \Phi(f)\, d\mu < \infty$. (Hence $\tilde{L}^\Phi \subset L^\Phi$.)

Lemma 2. Let M^Φ denote the closed subspace of L^Φ determined by the set of all μ-simple functions in L^Φ. If Φ is continuous then $M^\Phi \subset \tilde{L}^\Phi$ $(\subset L^\Phi)$, and if further $\tilde{L}^\Phi$ is linear, then $M^\Phi = L^\Phi$, i.e., the μ-simple functions in L^Φ are everywhere dense.

Proof. Let $f \in M^\Phi$ be arbitrary. Given $\varepsilon > 0$, there exists a μ-simple function f_ε in M^Φ, by definition, such that $\|f - f_\varepsilon\|_\Phi < \varepsilon$. Since f_ε vanishes outside a set of finite μ-measure and (being in M^Φ) is bounded it follows that $a.f_\varepsilon \in \tilde{L}^\Phi$ for any constant a where $\Phi(K) < \infty$ for $K < \infty$ (a consequence of continuity of Φ on the line) is used. If $0 < \varepsilon \leq \frac{1}{2}$, then (7) implies $2(f - f_\varepsilon) \in \tilde{L}^\Phi$. But $\tilde{L}^\Phi$ is a convex set (since $\Phi(\cdot)$ is a convex function) so taking $a = 2$ above, it follows that $f = \frac{1}{2}[2f_\varepsilon] + \frac{1}{2}2(f - f_\varepsilon)$ is in $\tilde{L}^\Phi$ proving the first part of the lemma.

For the second part, let $\tilde{L}^\Phi$ be linear. Since $\tilde{L}^\Phi \subset L^\Phi$, the reverse inequality follows from (5) and the fact that $\int_\Omega \Phi(Kf)\, d\mu < \infty$ for every $K > 0$ whenever $\int_\Omega \Phi(f)\, d\mu < \infty$, due to the linearity of $\tilde{L}^\Phi$ (which fact in turn characterizes the linearity of $\tilde{L}^\Phi$). So $\tilde{L}^\Phi = L^\Phi$. Also since $M^\Phi \subset L^\Phi$, the opposite inequality, and with it the lemma, will be proved if the following steps (i), (ii) and (iii) are established.

Let $f \in L^\Phi$ be arbitrary. By the structure of measurable functions there exists an increasing sequence $\{f_n\}$ of measurable functions, each taking a finite number of values such that $0 \leq f_n \to |f|$ everywhere. Then the following statements hold:

(i) $f_n \in L^\Phi$, for,

$$\|f_n\|_\Phi = \sup_{\rho(g) \leq 1} \int_\Omega |f_n g|\, d\mu \leq \sup_{\rho(g) \leq 1} \int_\Omega |f|\,|g|\, d\mu = \|f\|_\Phi < \infty.$$

(ii) $f_n \to |f|$ in L^Φ. For, given $\varepsilon > 0$, set $K = 2/\varepsilon$. Since $f_n \to |f|$ pointwise and $\Phi(\cdot)$ is continuous, $\Phi(K(|f| - f_n))$ tends monotonely to zero everywhere. But $|f| - f_n \in L^\Phi = \tilde{L}^\Phi$, and so

$$\int_\Omega \Phi(K(|f| - f_n))\, d\mu$$

exists for every n, and by monotone convergence theorem tends to zero. So choosing n_0 such that $n \geq n_0$ implies $\int_\Omega \Phi(K(|f| - f_n))\, d\mu \leq 1$, one has on using (1) (and the fact that $\|\cdot\|_\Phi$ is a norm),

$$\|K(|f| - f_n)\|_\Phi \leq \sup_{\rho(g) \leq 1} \int_\Omega |(|f| - f_n) g|\, d\mu \leq$$
$$\leq \int_\Omega \Phi(K(|f| - f_n))\, d\mu + 1 \leq 2.$$

Consequently, $\||f| - f_n\|_\Phi \leq 2/K = \varepsilon$. Thus (ii) follows since $\varepsilon > 0$ is arbitrary.

(iii) $f \in L^\Phi$ implies $f \in M^\Phi$. Since $L^\Phi = \tilde{L}^\Phi$, $\int_\Omega \Phi(Kf)\, d\mu < \infty$ for every $K > 0$. So, if $K = 4/\varepsilon$ since $\Phi(Kf) \in L^1$, by ([2]; III.2.20(c)) there exists a set $A \in \Sigma_1$, $\mu(A) < \infty$ and $\int_{A'} \Phi(Kf)\, d\mu \leq 1$. Consequently,

$$\|Kf\chi_{A'}\|_\Phi \leq \int_\Omega \Phi(Kf\chi_{A'})\, d\mu + 1 = \int_{A'} \Phi(Kf)\, d\mu + 1 \leq 2,$$

and hence $\|f\chi_{A'}\|_\Phi \leq 2/K = \varepsilon/2$. Since $f\chi_A \in L^\Phi$, applying the results of (i) and (ii) to the function $f\chi_A$ (if necessary to the positive and negative parts separately), the resulting sequence f_n is μ-simple (since $\mu(A) < \infty$), $f_n \in M^\Phi$, and there is n_ε such that $n \geq n_\varepsilon$ implies $\|f\chi_A - f_n\|_\Phi < \varepsilon/2$. Thus for this n one has

$$\|f - f_n\|_\Phi \leq \|f\chi_A - f_n\|_\Phi + \|f - f\chi_A\|_\Phi \leq$$
$$\leq \|f\chi_{A'}\|_\Phi + \varepsilon/2 \leq \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

From the arbitrariness of ε, it results that $f \in M^\Phi$, as was to be proved.

Remark. If $\tilde{L}^{\Phi}$ is not linear, the inclusions between M^{Φ}, $\tilde{L}^{\Phi}$ and L^{Φ} can be proper as simple examples show. This is true even if $\mu(\Omega) < \infty$.

The following result is proved in ([11]; p. 136) for the σ-finite μ.

Lemma 3. Let f be a measurable function such that fg is integrable for every $g \in L^{\Psi}$. Then $f \in L^{\Phi}$, where Φ, Ψ are Young's complementary functions.

Proof. Let Σ_1 be the class of sets in Σ of finite μ measure, and let $f_A = f\chi_A$ for $A \in \Sigma_1$. Since $|f_A|$ is measurable, there is an increasing sequence of measurable functions $f_{n,A}$, $0 \leq f_{n,A} \to |f_A|$, pointwise. Define a family of linear functionals $l_{n,A}$ on L^{Ψ} as

$$l_{n,A}(g) = \int_{\Omega} f_{n,A} g \, d\mu, \quad g \in L^{\Psi}.$$

This class of functionals is well defined on L^{Ψ} and by monotone convergence theorem,

$$|l_{n,A}(g)| \leq \int_{\Omega} |f_{n,A} g| \, d\mu \leq \int_{\Omega} |f_A| \, |g| \, d\mu \leq \int_{\Omega} |fg| \, d\mu < \infty.$$

Thus $\{l_{n,A}(g), n \geq 1, A \in \Sigma_1\}$ is bounded for each $g \in L^{\Psi}$. Hence by the uniform boundedness principle ([2], p. 59), there is a positive M $(< \infty)$, such that $|l_{n,A}(g)| \leq M \|g\|_{\Psi}$, for all n, and A (in Σ_1). Then using $\|g\|_{\Psi} \leq 2$, for $\rho(g) \leq 1$,

$$\|f_A\|_{\Phi} = \sup_{\rho(g) \leq 1} \int_{\Omega} |f_A g| \, d\mu = \sup_{\rho(g) \leq 1} \left| \int_{\Omega} f_A g \, d\mu \right| \leq \sup_{\rho(g) \leq 1} [\lim_n |l_{n,A}(g)|] \leq 2M < \infty.$$

Hence $f_A \in L^{\Phi}$, for all $A \in \Sigma_1$.

In the general case, consider a fixed but arbitrary $g \in L^{\Psi}$, with $\rho(g) \leq 1$. Since by definition $f_A = f_B = f_{A \cap B}$ on $A \cap B$, $A, B \in \Sigma_1$, if

$$\nu_g(E) = \int_E |f_E g| \, d\mu = \int_E |fg| \, d\mu, \quad E \in \Sigma_1,$$

then $\nu_g(\cdot)$ is a measure on Σ_1. The preceding proof implies

$$\sup_{E \in \Sigma_1} \nu_g(E) \leq \sup_{E \in \Sigma_1} [\|f_E\|_{\Phi}] \leq 2M < \infty.$$

Hence, as in the proof of Lemma 1, there is a sequence $\{A_n\} \subset \Sigma_1$, $A_n \subset A_{n+1}$, such that if $F = \bigcup_{n=1}^{\infty} A_n$, $\nu_g(F) = \lim_n \nu_g(A_n) = \sup_{E \in \Sigma_1} \nu_g(E) \leq 2M$, and $\nu_g(F') = 0$. [The omitted details parallel those of Lemma 1.] Thus

$$\|f\|_{\Phi} = \sup_{\rho(g) \leq 1} \int_{\Omega} |fg| \, d\mu = \sup_{\rho(g) \leq 1} [\nu_g(\Omega)] = \sup_{\rho(g) \leq 1} [\nu_g(F)] \leq 2M < \infty.$$

and so $f \in L^{\Phi}$. This completes the proof.

Remark. The assumption that f be measurable may be dropped in the above result if $M^{\Phi} = \tilde{L}^{\Phi}$ or μ is σ-finite. This becomes evident after seeing the proof of Lemma 5 below. In any case, f will be a "quasi-function" as defined there. (Compare with ([2], IV.13.7).)

Definition. Let (Ω, Σ, μ) be an arbitrary measure space and $\Phi(\cdot)$ be a Young's function. Then a real (or complex) set function G on Σ, vanishing on μ-null sets, is said to be of *Φ-bounded variation on E* (relative to μ), if $E \in \Sigma$ and $\{A_i\}$ are finite disjoint measurable collection of sets, of finite positive μ-measure, contained in E, then

$$I_{\Phi}(G; E) = \sup \sum_{i=1}^{n} \Phi\left(\frac{G(A_i)}{\mu(A_i)}\right) \mu(A_i) \tag{8}$$

exists, where the supremum is taken relative to all such sets $\{A_i\}$ of E. If $I_{\Phi}(G) = I_{\Phi}(G; \Omega) < \infty$, G is said to be of *Φ-bounded variation.* (If $\Phi(x) = |x|^p$, $p \geq 1$, this reduces to the well-known concept of p-bounded variation. If $G(\cdot)$ is a real-valued point function with Ω as the line, this is called a generalized Hellinger integral, in [6].)

Lemma 4. If $\Phi(\cdot)$ is a continuous Young's function, $g(\cdot) \in \tilde{L}^{\Phi}$, and $G(E) = \int_E g(x) \, d\mu$, $E \in \Sigma$ where (Ω, Σ, μ) is a measure space, then $I_{\Phi}(G)$ exists and moreover the following conclusions hold:

(a) $$I_{\Phi}(G; E) = \int_E \Phi(g) \, d\mu, \quad E \in \Sigma.$$

(b) $\lim_{\mu(A) \to 0} I_{\Phi}(G; A) = 0$; (b′) $I_{\Phi}(G) = 0$, if and only if $g = 0$, a.e., $[\mu]$.

(c) For each $\varepsilon > 0$, there is an $A \in \Sigma$ such that $\mu(A) < \infty$, $I_{\Phi}(G; A') < \varepsilon$.

The proof of this result is parallel to that of III.2.20 of [2], and will be omitted. The following difference may, however, be noted. First, the proof that $I_{\Phi}(G; E) \leq \int_E \Phi(g) \, d\mu$ is the usual one and for the reverse inequality, first the proof is obtained for the simple functions g_m (where $g_m \to g$, everywhere), as in ([2], p. 110). However, if g_m determines g it is not necessarily true that $\Phi(g_m)$ determines $\Phi(g)$ in the sense of [2]. But here μ is a measure, and hence by Fatou's lemma the opposite inequality can be proven from the result already obtained for g_m. The second part of (b) utilizes the FSP of μ and the proof of the other parts is similar to that given in [2]. It is remarked that the lemma may not be true if μ is only finitely additive, but remains valid provided a growth condition on Φ is imposed. For instance, $M^{\Phi} = \tilde{L}^{\Phi}$ may be stipulated.

The next result provides a sort of converse to the preceding. For this, the concept of "quasi-function" is needed and is given in

Definition. A real (or complex) function f on (Ω, Σ, μ) is said to be a (measurable) *quasi-function* relative to μ if for every $E \in \Sigma$ of finite μ-measure it is equivalent to a measurable function f_E vanishing outside of E, and if E and F are in Σ and of finite μ-measure, then the corresponding functions f_E and f_F, equivalent to f on E and F, are themselves equivalent on $E \cap F$. A quasi-function f will be denoted by f^*. [As usual, two functions are equivalent if they differ on a set of measure zero. Thus quasi-functions are measurable if μ is σ-finite or "localizable", but need not be so in general.]

The term "quasi-function" is used in [5] and "cross-section" in [9] for the same concept. The fact that the integrals are defined for the quasi-functions with the usual properties will be used below. (See [5] and [9], for further discussion.) The same symbols L^{Φ}, L^{Ψ} will be used even if they contain quasi-functions. As the context shows, this causes no difficulty.

Lemma 5. If $G(\cdot)$ is a countably additive set function, on (Ω, Σ, μ), of Φ-bounded variation, then there exists a quasi-function g^* such that

$$I_{\Phi}(G) = \int_{\Omega} \Phi(g^*)\, d\mu.$$

If, in addition, either (i) μ is σ-finite (more generally "localizable"), or (ii) $\tilde{L}^{\Phi} = M^{\Phi}$, in the notation of Lemma 2, g^* can be chosen to be measurable. [A measure μ is *localizable* means: for each non-empty class $\mathscr{K} \subset \Sigma$, there is a measurable set $B \in \Sigma$ – supremum – such that (a) $A \in \mathscr{K}$ implies $\mu(A - B) = 0$, and (b) if $C \in \Sigma$, $\mu(A - C) = 0$ for every $A \in \mathscr{K}$, then $\mu(B - C) = 0$.]

Proof. By hypothesis, G is countably additive and vanishes on μ null sets. Then by the general Radon-Nikodým theorem ([3], p. 336; [9], p. 173) there is a unique quasi-function g^* such that

$$G(E) = \int_E g^*(x)\, d\mu, \quad E \in \Sigma.$$

The representation for $I_{\Phi}(G)$ then follows from this and the Φ-boundedness as in Lemma 4 (a), first proving it for sets E of finite μ-measure (since on such E, g^* is equivalent to a measurable function), and then extending it using the properties of quasi-functions with a similar procedure used in the proof of Lemma 1. If μ is localizable, then (by [5] and [9]) g^* can be taken to be measurable. If, alternately, $\tilde{L}^{\Phi} = M^{\Phi}$, since $I_{\Phi}(G) < \infty$, it follows that $\int_{\Omega} \Phi(Kg^*)\, d\mu < \infty$ for all $K > 0$. So (cf., Lemma 2) there

exists a sequence of μ-simple functions g_n such that $\|g_n - g^*\|_{\Phi} \to 0$, and hence g_n tends in measure to g, a measurable function. This implies $g = g^*$, a.e., $[\mu]$, as was to be shown.

Finally, even if $G(\cdot)$ is only additive there is another property of interest:

Lemma 6. If $G(\cdot)$ is an additive set function of Φ-bounded variation, and the complementary function Ψ of Φ satisfies the condition $\Psi(x) < \infty$ for $|x| < \infty$, then G is μ-continuous, i.e., $\lim_{\mu(A) \to 0} G(A) = 0$.

Proof. The proof is similar to the absolute continuity of point functions, as, for instance, given in [6] (see also [2], p. 115). Given that $I_{\Phi}(G) < \infty$. For any $\varepsilon > 0$, choose an $M \geq 1$, such that $(I_{\Phi}(G)/M) \leq \varepsilon/2$. Now let $A \in \Sigma$ be any set such that $\mu(A)\Psi(M) < \varepsilon/2$. This is possible since $\Psi(M) < \infty$ for all $M < \infty$. If $\{A_i\}$ is any measurable dissection of A, $\mu(A_i) > 0$, denote by $x_i = G(A_i)/\mu(A_i)$ and $x(s) = x_i$, $s \in A_i$, $i = 1, \ldots, n$, and $= 0$ otherwise. Then,

$$|G(A)| \leq \sum_{i=1}^{n} |G(A_i)| = \sum_{i=1}^{n} |x_i|\mu(A_i) = \int_A M \left|\frac{x(s)}{M}\right| d\mu$$

$$\leq \int_A \Phi\left(\frac{x(s)}{M}\right) d\mu + \int_A \Psi(M)\, d\mu, \text{ by (1)},$$

$$\leq \frac{1}{M} \sum_{i=1}^{n} \Phi(x_i)\mu(A_i) + \Psi(M)\mu(A), \text{ by convexity of } \Phi,$$

$$\leq \frac{1}{M} I_{\Phi}(G) + \Psi(M)\mu(A) < \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, the lemma is proved.

Remark. If $G(\cdot)$ is not countably additive, then even with the hypothesis of this lemma, the existence of g^* of Lemma 5 cannot be asserted. From the convexity of Φ it follows that $I_{\Phi}(\cdot)$ [cf. (8) for definition] is also convex. If further $\Phi(\cdot)$ is continuous, $G_n(E) \to G(E)$ for all $E \in \Sigma_1$, as $n \to \infty$, and G is an additive set function on Σ_1, then it can be shown (as in [6], p. 599) that

$$\underline{\lim}_n\, I_{\Phi}(G_n) \geq I_{\Phi}(G).$$

2. *A space of set functions.* This section is devoted to the introduction of a linear space of set functions that will be useful in the study of the conjugate space of L^{Φ}.

Definition. If (Ω, Σ, μ) is a measure space, then $A_{\Phi}(\mu)$ stands for the class of finitely additive scalar set functions G on Σ that vanish on μ null sets and such that $I_\Phi(G/K) \leq 1$ for some $K > 0$. The class $A_\Psi(\mu)$ is defined similarly. Here Φ, Ψ are as usual Young's complementary functions and $I_\Phi(\cdot)$ is the same symbol of (8).

It is clear that $A_\Phi(\mu)$ is linear. It is normed according to (3) specialized as

Definition. For any $G \in A_\Phi(\mu)$, define $\|\cdot\|_\Phi$ (which is a norm) by

$$\|G\|_\Phi = \inf\left\{K > 0, I_\Phi\left(\frac{G}{K}\right) \leq 1\right\}. \tag{3'}$$

The space $A_\Phi(\mu)$ contains the Orlicz space L^Φ in the following sense. If the subclass of G's in $A_\Phi(\mu)$ that are countably additive are considered, then as in Lemma 5, $G(E) = \int_E g^*\, d\mu$ and if μ is localizable or $M^\Phi = \tilde{L}^\Phi$, then g is measurable and then (g^* is g)

$$I_\Phi\left(\frac{G}{K}\right) = \int_\Omega \Phi\left(\frac{g}{K}\right) d\mu \leq 1,$$

gives by (3) the equation $\|G\|_\Phi = N_\Phi(g)$. Conversely the $G(\cdot)$ defined by a g in L^Φ as above defines an element in the subclass of $A_\Phi(\mu)$. Thus the subclass of A_Φ consisting of countably additive set functions is isometrically isomorphic to L^Φ, and the A_Φ itself is a much larger set.

The normed linear space $A_\Phi(\mu)$ is actually a B-space. This follows from the representation theorem of the next section where it is shown that $A_\Phi(\mu)$ is isomorphic and (topologically) equivalent to the conjugate space of L^Ψ. From the discussion of the preceding paragraph it follows that $\rho(g) \leq 1$ if and only if $\|\tilde{G}\|_\Psi \leq 1$ where $\tilde{G}(E) = \int_E g\, d\mu$, $E \in \Sigma$, $g \in L^\Psi$. (Cf. also [3], p. 79 and [6].) This observation gives a new definition of (2) as

$$\|f\|_\Phi = \sup_{G \in S} \left|\int_\Omega f\, dG\right|, \quad f \in L^\Phi, \tag{2'}$$

where $S \subset A_\Psi(\mu)$ is a set of countably additive G's with $\|G\|_\Psi \leq 1$. The corresponding L^Ψ contains all quasi-functions on (Ω, Σ, μ) which are the usual functions under the hypothesis of Lemma 5.

3. *Global representation theorem for functionals on L^Φ.* Let L^Φ be an Orlicz space on (Ω, Σ, μ). In what follows integrals relative to finitely additive set functions appear and their basic properties to be utilized are those given in [1] and ([2], Chapters III and IV).

Theorem 1. Let Φ and Ψ be complementary Young's functions and L^Φ, $A_\Psi(\mu)$ be the spaces on (Ω, Σ, μ) introduced before. Then the integral

$$x^*(f) = \int_\Omega f\, dG, \quad f \in L^\Phi, \tag{9}$$

exists and defines a continuous linear functional on L^Φ for each G in $A_\Psi(\mu)$. Moreover, $\|G\|_\Psi \leq \|x^*\| \leq 4\|G\|_\Psi$, and $x^*(f) = 0$ for all f in L^Φ, if and only if $G = 0$.

Proof. It is first noted that there is one special case in which (9) follows immediately, and that is when G is countably additive. Since then, as in Lemma 5, $G(E) = \int_E g^*\, d\mu$, $E \in \Sigma$ and g^* is a quasi-function in L^Ψ. The integral in (9) can be rewritten and the Hölder inequality (stated after Lemma 1) applied to obtain

$$|x^*(f)| = \left|\int_\Omega f\, dG\right| = \left|\int_\Omega f g^*\, d\mu\right| \leq \|f\|_\Phi \|g^*\|_\Psi. \tag{10}$$

Consequently x^* is a bounded linear functional on L^Φ. There is no other case for which the conclusion follows so easily. Even for the case $\Phi(x) < \infty$ for $x \leq x_0 (< \infty)$ and $\Phi(x) = \infty$ for $x > x_0$, so that $L^\Phi \subset L^\infty$, the known result on the representation of functionals on L^∞ (cf., [2], p. 296) is not directly applicable here. The trouble is that $\Psi(x) \leq x_0|x|$ and the Ψ-bounded variation of G does not give the usual bounded variation (due to the wrong inequality) unless $\mu(\Omega) < \infty$. The result will now be established as follows.

It suffices to consider $f \geq 0$, and $f \in L^\Phi$, since any f in L^Φ can be decomposed into positive and negative parts and this result is then applicable, due to the linearity of the integral and that of the space L^Φ. Let $\Sigma_1 \subset \Sigma$, be the class of all μ-finite sets. If $E \in \Sigma_1$, let f_E be $f\chi_E$. Then $f_E \in L^\Phi$, and since it is measurable, there is an increasing sequence $\{f_n\}$ of μ-simple functions such that $0 \leq f_n \to f_E$. Consequently this sequence converges to f_E (by Egorov's theorem) μ-almost uniformly as well, and this fact is used below. Since G vanishes on μ-null sets, it follows easily that f_n are measurable relative to G also. The proof of (9) will be completed in two stages.

Let $f_n = \sum_{i=1}^m a_{ni}\chi_{E_{ni}}$ be a typical representation of f_n on E. Then one has, since, $0 \leq f_n \uparrow f_E$,

$$\left|\int_E f_n\, dG\right| = \left|\sum_{i=1}^m a_{ni} G(E_{ni})\right| \leq \sum_{i=1}^m a_{ni} |G(E_{ni})| =$$

$$= K\|f_E\|_\Phi \int_E \Phi\left(\frac{f_n}{\|f_E\|_\Phi}\right) \frac{h}{K}\, d\mu, \text{ where } h(x) = \frac{|G(E_{ni})|}{\mu(E_{ni})},$$

$$x \in E_{ni},\ i = 1, \ldots, m, \qquad = 0, \text{ otherwise},$$

$$\le K\|f_E\|_\Phi\left[\int_E \Phi\left(\frac{f_n}{\|f_E\|_\Phi}\right)d\mu + \int_E \Psi\left(\frac{h}{K}\right)d\mu\right], \text{ by (1)},$$

$$\le K\|f_E\|_\Phi\left[\int_E \Phi\left(\frac{f_E}{\|f_E\|_\Phi}\right)d\mu + I_\Psi\left(\frac{G}{K}, E\right)\right]$$

$$\le 2\|f_E\|_\Phi K_E \le 2\|f\|_\Phi\|G\|_\Psi, \text{ with } K_E = \|G_E\|_\Psi,$$

and Lemma 1. [G_E is G restricted to E.] (11)

This shows that for any $E \in \Sigma_1$, $\int_E f_n\, dG$ exists and is bounded by the right side constant of (11). Now two cases arise.

(a) If $\Phi(\cdot)$ is continuous then, by Lemma 6, G is μ-continuous, and hence also is $v(G)$, the variation of G. Thus the f_n-sequence which is μ-almost uniformly convergent (shown in the preceding paragraph) is also $v(G)$, and so G, -almost uniformly convergent to f_E. Therefore by ([1], Theorem 8), $\int_E f\, dG = \lim_n \int_E f_n\, dG$ exists and by (11) it follows that for all $E \in \Sigma_1$,

$$|\textstyle\int_E f\, dG| \le 2\|f_E\|_\Phi\|G_E\|_\Psi \le 2\|f\|_\Phi\|G\|_\Psi. \qquad (12)$$

(b) If $\Phi(\cdot)$ is discontinuous, then the only possible form is that $\Phi(x) < \infty$ for $x < x_0$ but $\Phi(x) = \infty$ for $x > x_0$ $(0 < x_0 < \infty)$. But then $L^\Phi \subset L^\infty$ and f is essentially bounded. Considering the subspace of L^Φ of functions restricted to E, say $L^\Phi(E)$, one has $L^\Phi(E) = L^\infty(E)$, since $\mu(E) < \infty$. But then however by definition, the Ψ-bounded variation of a set function reduces to the ordinary bounded variation on E, and that $\|f_E\|_\Phi = \|f_E\|_\infty$, $\|G_E\|_\Psi = \|G_E\|_1 = v(G, E)$. In this case the result of ([2], p. 296) is applicable and yields,

$$|\textstyle\int_E f\, dG| \le \int_E |f|\, dv(G) \le \|f_E\|_\Phi\|G_E\|_\Psi \le 2\|f\|_\Phi\|G\|_\Psi. \qquad (13)$$

Thus in all cases (12) holds for all $E \in \Sigma_1$. Note that the final bound does not involve E.

Now it is to be shown that in (12) E may be replaced by Ω. Let $\lambda(E) = \int_E f\, dG$. Then λ is a bounded additive set function on Σ_1 and vanishes on μ- (and G-) null sets, and so is $v(\lambda)$ on Σ_1. Let $\alpha = \sup_{E\in\Sigma_1}|\lambda(E)|$, then $0 < \alpha \le 2\|f\|_\Phi\|G\|_\Psi$. But by ([2], IV.9.11) there is a measurable space $(S, \mathscr{A})$, where S is a (totally disconnected) compact Hausdorff space and $\mathscr{A}$ is a σ-field on S, and a regular countably additive measure $\tilde\lambda$ on $(S, \mathscr{A})$ and an isometric isomorphism $T: v(\lambda) \to \tilde\lambda$. If $\mathscr{A}_1$ corresponds to Σ_1, then this implies $\alpha = \sup_{A\in\mathscr{A}_1}\tilde\lambda(A)$. Since $\tilde\lambda$ is a measure there is a monotone increasing sequence $\{A_n\} \subset \mathscr{A}_1$ such that $\alpha = \lim_n \tilde\lambda(A_n) = \tilde\lambda(B)$, where $B = \bigcup_{n=1}^\infty A_n$. Now applying the procedure used in the proof of Lemma 1, it can be concluded that $\tilde\lambda(B') = 0$, so that $\tilde\lambda(S) = \alpha$. Using the isometry again, one has $v(\lambda; \Omega) = \alpha$. It therefore follows that $|\lambda(\Omega)| \le \alpha \le 2\|f\|_\Phi\|G\|_\Psi$. In other words (12) holds with E replaced by Ω. This completes the proof that (9) is a bounded linear functional on L^Φ, as well as $\|x^*\| \le 4\|G\|_\Psi$, taking the nonpositive case into account.

The opposite inequality is obtained as follows. Writing $K = \|G\|_\Psi$, one has $I_\Psi(G/K) = 1$. (The modifications required in the following if there is inequality here will be obvious.) By definition of $I_\Psi(\cdot)$, for any $\delta > 0$, there exists a partition $E_1, \ldots, E_n$, such that (the trivial case $K = 0$ being excluded)

$$\sum_{i=1}^n \Psi\left(\frac{G(E_i)}{K\mu(E_i)}\right)\mu(E_i) > 1 - \delta/K. \qquad (14)$$

Define

$$a_i = \Psi'\left(\frac{G(E_i)}{K\mu(E_i)}\right), \text{ and } f_0 = \sum_{i=1}^n a_i\chi_{E_i},$$

where Ψ' is the derivative of Ψ. Since f_0 is a μ-simple function it is in L^Φ, and

$$\|f_0\|_\Phi\|x^*\| \ge x^*(f_0) = \int_\Omega f_0\, dG = K\sum_{i=1}^n a_i\left(\frac{G(E_i)}{K\mu(E_i)}\right)\mu(E_i)$$

$$= K\sum_{i=1}^n\left[\Phi(a_i) + \Psi\left(\frac{G(E_i)}{K\mu(E_i)}\right)\right]\mu(E_i), \text{ by equality in (1)},$$

$$> K[\textstyle\int_\Omega \Phi(f_0)\, d\mu + 1 - \delta/K], \text{ by (14)},$$

$$\ge K\|f_0\|_\Phi - \delta, \text{ since } \|f\|_\Phi \le \textstyle\int_\Omega \Phi(f)\, d\mu + 1.$$

Since $\delta > 0$ is arbitrary, it follows that $\|x^*\| \ge K = \|G\|_\Psi$. Combining this with the previous inequality one has $\|G\|_\Psi \le \|x^*\| \le 4\|G\|_\Psi$. From this it also follows immediately that $x^*(f) = 0$ for all f in L^Φ if and only if $G = 0$. This completes the proof of the theorem.

Remark. If G is countably additive, using (2′) in Section 2, one has $|x^*(f)| = K|\int_\Omega f\, d(G/K)| \le K\|f\|_\Phi$, so that $\|x^*\| \le K$. Thus in this case $\|x^*\| = \|G\|_\Psi (= K)$ and the proof can also be simplified. (The same is true if Φ is discontinuous, by (13), even though G may not be countably additive.) In general, however, only the inequalities given in the theorem can be proved. (If the functions are allowed to take complex values also, then the right side inequality will have 8 instead of 4.)

Theorem 2. If $x^*(\cdot)$ is a continuous linear functional on L^Φ, then there exists a (additive) set function G in $A_\Psi(\mu)$ such that

$$x^*(f) = \int_\Omega f \, dG, \quad f \in L^\Phi. \tag{15}$$

Proof. Let $\Sigma_1 \subset \Sigma$ be the sets of finite μ-measure as before. So $\chi_E \in L^\Phi$ for $E \in \Sigma_1$, and if $x^*(\chi_E) = G(E)$, then G is a finitely additive set function, since x^* is linear. Also since $\chi_A = 0$, a.e., for μ-null sets A, it follows that G vanishes on μ-null sets. To show that $G \in A_\Psi(\mu)$, the boundedness of x^* will be used. For any $E \in \Sigma$, if $E_1, \ldots, E_n$ is a measurable disjoint collection of sets, $E_i \subset E$, and $0 < \mu(E_i) < \infty$; define f^0 in $\tilde{L}^\Phi$ as

$$f^0 = \sum_{i=1}^n b_i \chi_{E_i}, \quad b_i = \Psi'\left(\frac{G(E_i)}{\|x^*\| \mu(E_i)}\right). \tag{16}$$

Since $\|f\|_\Phi \leq \int_\Omega \Phi(f) \, d\mu + 1$, the following inequalities hold.

$$1 + \int_\Omega \Phi(f^0) \, d\mu \geq \|f^0\|_\Phi \geq \frac{x^*(f^0)}{\|x^*\|} = \sum_{i=1}^n b_i \left(\frac{G(E_i)}{\|x^*\| \mu(E_i)}\right) \mu(E_i)$$

$$= \sum_{i=1}^n \left[\Phi(b_i) + \Psi\left(\frac{G(E_i)}{\|x^*\| \mu(E_i)}\right)\right] \mu(E_i), \text{ by equality of (1)},$$

$$= \int_\Omega \Phi(f^0) \, d\mu + \sum_{i=1}^n \Psi\left(\frac{G(E_i)}{\|x^*\| \mu(E_i)}\right) \mu(E_i).$$

This implies, since $E \in \Sigma$ is arbitrary, $I_\Psi(G/\|x^*\|) \leq 1$, so that $\|G\|_\Psi \leq \|x^*\| < \infty$. It follows that G is in $A_\Psi(\mu)$.

It remains to show that (15) holds for this G. If $f = \chi_E$, $E \in \Sigma_1$, then with this G, (15) is true. If $f = \sum_{i=1}^k a_i \chi_{E_i}$, $E_i \in \Sigma_1$, then also (15) holds for such f, by the linearity of the integral. In the general case, let $f_E = f\chi_E$, $E \in \Sigma_1$. Again it suffices to consider $f \geq 0$. Then there is an increasing sequence $\{f_n\}$ of μ-simple functions on E such that $0 \leq f_n \to f_E$, a.e., $[\mu]$, and since $\mu(E) < \infty$, the convergence is also μ-almost uniform. Here again two cases arise.

(a) $\Phi(\cdot)$ is continuous. Then G is μ-continuous by Lemma 6, and so $f_n \to f_E$, G-almost uniformly and, as in the proof of (12), $\lim_n \int_E f_n \, dG = \int_E f \, dG$ is true. But $L^\Phi(E) \subset L^1(E)$ for $E \in \Sigma_1$. [This follows from the supporting line property of the convex Φ, $c_0 + c_1 x \leq \Phi(x)$ for some constants c_0, c_1. Hence there is a $c > 0$, with $\|f\|_1 \leq c\|f\|_\Phi$.] However the linear space $L^\Phi(E)$ is not in general closed in L^1-norm. So, let $(L^\Phi(E))_1$ be the L^1-closure of $L^\Phi(E)$ and considering x^* as bounded operator in L^1 with domain $L^\Phi(E)$, let y^* be its (norm-preserving) extension to $(L^\Phi(E))_1$. Thus $y^*(f) = x^*(f)$ for all f in $L^\Phi(E)$, and y^* is a continuous linear functional on the closed subspace $(L^\Phi(E))_1$ of L^1. Hence

$$(f_n, f_E \in L^\Phi(E) \subset (L^\Phi(E))_1 \subset L^1),$$

$$|y^*(f_E) - y^*(f_n)| \leq \|y^*\| \|f_E - f_n\|_1 = \\ = \|y^*\| \int_E (f_E - f_n) \, d\mu \to 0, \tag{17}$$

as $n \to \infty$, by the Lebesgue monotone convergence theorem. But $y^*(f_E) = x^*(f_E)$ and $y^*(f_n) = x^*(f_n)$ since the f_E, f_n are in $L^\Phi(E)$. Hence from (17),

$$x^*(f_E) = y^*(f_E) = \lim_n y^*(f_n) = \lim_n x^*(f_n) = \\ = \lim_n \int_E f_n \, dG = \int_E f \, dG, \quad E \in \Sigma_1. \tag{18}$$

It follows that (15) holds for f_E, $E \in \Sigma_1$, when $\Phi(\cdot)$ is continuous.

(b) If $\Phi(\cdot)$ is discontinuous, then, as noted before, $L^\Phi(E) = L^\infty(E)$, for $E \in \Sigma_1$. In this case however, a result of ([2], p. 296) is again applicable to yield $x^*(f_E) = \int_E f \, dG$. This means, if $x_E{}^*$ is the restriction of x^* to E, then (15) takes the form, $x_E{}^*(f) = \int_E f \, dG$, for all $E \in \Sigma_1$. Thus (15) is true in general if $\mu(\Omega) < \infty$.

It remains to show that the above result holds for general Ω (i.e., $\mu(\Omega) = \infty$). For this, again the procedure of the paragraph following (13) of Theorem 1 is applicable *mutatis mutandis*. So $\int_\Omega f \, dG$ is well-defined ,and if $y^*(f) = \int_\Omega f \, dG$ then y^* is a continuous linear functional on L^Φ and $x_E{}^* = y_E{}^*$ for all E in Σ_1. By the FSP of μ, this implies that x^* and y^* agree everywhere and (15) holds as stated. The proof is therefore complete.

The preceding results enable the presentation of the global representation theorem for continuous linear functions on L^Φ as follows.

Theorem 3. Let (Ω, Σ, μ) be a measure space. If Φ and Ψ are complementary Young's functions and L^Φ, $A_\Psi(\mu)$ are the Orlicz space and the space of Ψ-bounded additive set functions on (Ω, Σ, μ), then for every continuous linear functional x^* on L^Φ, there is a unique G in $A_\Psi(\mu)$ such that

$$x^*(f) = \int_\Omega f \, dG, \quad f \in L^\Phi, \tag{19}$$

and moreover,

$$\|G\|_\Psi \leq \|x^*\| \leq 4\|G\|_\Psi. \tag{20}$$

Proof. By Theorem 2, every x^* of the theorem has the form (19), and by Theorem 1, the norm inequalities (20) hold. Only uniqueness needs to be proven. If G_1, G_2 are two elements in A_Ψ for which (19) holds, then

$$x^*(f) = \int_\Omega f\, dG_1 = \int_\Omega f\, dG_2, \quad f \in L^\Phi.$$

So,

$$\int_\Omega f\, d(G_1 - G_2) = 0, \quad f \in L^\Phi. \tag{21}$$

Since $G_1 - G_2 \in A_\Psi(\mu)$, (21) represents a zero functional on L^Φ, by Theorem 1. It follows by (20) that $\|G_1 - G_2\|_\Psi = 0$, so that $G_1 = G_2$, as was to be proved.

Remarks. 1. The result of this theorem asserts that there is an isomorphism between $(L^\Phi)^*$ – the conjugate space of L^Φ – and A_Ψ, and that the spaces $(L^\Phi)^*$, A_Ψ are (topologically) equivalent. As a corollary, it follows that A_Ψ is a B-space.

2. Without further hypothesis G may not be countably additive and hence $x^*(f) = \int_\Omega fg\, d\mu$, $f \in L^\Phi$, $g \in L^\Psi$, may not hold. The conditions under which this special result holds, its falsity in the general case, and the possibility $\|G\|_\Psi = \|x^*\|$ in (20), are discussed in the next section.

4. *Special cases.* Linear functionals on certain subspaces of L^Φ can be given a more specific form by specializing Theorem 3. Thus the following result holds and it includes the corresponding known results, (cf., [3], [6], [11]).

Theorem 4. Let L^Φ and $A_\Psi(\mu)$ be the B-spaces on (Ω, Σ, μ), Φ be continuous, and Φ, Ψ be complementary Young's functions. If x^* is a continuous linear functional on M^Φ, the closed subspace determined by the μ-integrable simple functions in L^Φ, then there is a subspace $A_\Psi{}^0$ of $A_\Psi(\mu)$ containing countably additive set functions, which is isometrically equivalent to $(M^\Phi)^*$, the conjugate space of M^Φ. Moreover,

$$x^*(f) = \int_\Omega f\, dG, \quad (= \int_\Omega fg^*\, d\mu), \quad f \in M^\Phi, \quad G \in A_\Psi{}^0, \tag{22}$$

and

$$\|x^*\| = \|G\|_\Psi. \tag{23}$$

In (22), g^* is a quasi-function, $g^* \in L^\Psi$, and it can be taken as a measurable function in L^Ψ if and only if, either (i) μ is localizable, or (ii) $M^\Psi = \tilde{L}^\Psi$. [(ii) implies that Ψ is continuous and $M^\Psi = L^\Psi$.]

Proof. Consider the set $M^\Phi = \{f \in L^\Phi | \int_\Omega \Phi(Kf)\, d\mu < \infty$ for all $K > 0\}$. Following the proof of Lemma 2, it is readily verified that this set is a closed subspace of L^Φ and is determined by the μ-integrable simple functions in L^Φ, that is, the same set as in the statement of the theorem. Since, by Theorem 3, the representation (22) holds, one only has to show that $G \in A_\Psi{}^0$ if $x^* \in (M^\Phi)^*$.

Let E, E_i, $i = 1, 2, \ldots$, $E = \bigcup_{n=1}^\infty E_n$, be sets of finite μ-measure and E_i be disjoint. Then χ_E, χ_{E_i} are in M^Φ. Let $F_m = \bigcup_{n=m+1}^\infty E_n$. Also

$$x^*(\chi_E) = \int_E dG = \textstyle\sum_{i=1}^m G(E_i) + G(F_m),$$

since G is finitely additive. To show that $G \in A_\Psi{}^0$, it is enough to establish $|G(F_m)| = |x^*(\chi_{F_m})| \to 0$ as $m \to \infty$. However, E_i are disjoint and this implies that $\chi_{F_m} \to 0$ a.e. $[\mu]$, and for any K, there is an m_0, such that for $m \geq m_0$ one has

$$\|K\chi_{F_m}\|_\Phi \leq \int_\Omega \Phi(K\chi_{F_m})\, d\mu + 1 \leq 2, \qquad (\chi_{F_m} \in M^\Phi).$$

So $\|\chi_{F_m}\| \leq 2/K$, and since K is arbitrary, $\|\chi_{F_m}\|_\Phi \to 0$, and this in turn gives $|G(F_m)| \leq \|x^*\| \|\chi_{F_m}\|_\Phi \to 0$, as $m \to \infty$. Hence G is countably additive. Therefore, by the remark following the proof of Theorem 1, (23) holds. The remaining conclusions are immediate consequences of Lemma 5. This completes the proof.

Corollary. If $\Phi(2x) \leq C\Phi(x)$, $x \geq 0$, and L^Φ, L^Ψ are Orlicz spaces on (Ω, Σ, μ), a localizable measure space, (Φ, Ψ being complementary Young's functions) then for every $x^* \in (L^\Phi)^*$, there is a unique $g \in L^\Psi$ such that

$$x^*(f) = \int_\Omega fg\, d\mu, \quad \text{and} \quad \|x^*\| = N_\Psi(g). \tag{24}$$

The condition on Φ implies that $M^\Phi = \tilde{L}^\Phi$, and the result follows from the theorem, and (3) and (3′) defining $N_\Psi(g)$ $(= \|G\|_\Psi)$.

Remarks. 1. It is to be noted that in these results, the norm for L^Φ is that of (2) while the one for L^Ψ is (3). Using (4) these relations can be expressed in terms of the norm (2) itself. Thus one gets for (24),

$$\tfrac{1}{2}\|g\|_\Psi \leq \|x^*\| \leq \|g\|_\Psi, \tag{25}$$

a form given in [10], and all earlier studies on the subject. The above corollary was originally proven in this form in ([10], and [11]).

2. The above formula (24) may not hold if M^Φ is a proper subset of L^Φ even if $\mu(\Omega) < \infty$ and $\Phi(\cdot)$ is continuous where x^* is in $(L^\Phi)^*$. The following two sentences are well-known and are added to make this point clearer: If $0 \neq f_0 \in L^\Phi - M^\Phi$, and (24) holds for all $f \in L^\Phi$, and all $x^* \in (L^\Phi)^*$, then by a corollary to Hahn-Banach

theorem ([2], p. 64) there exists a functional $x_0^* \in (L^\Phi)^*$, such that $x_0^*(M^\Phi) = 0$, $x_0^*(f_0) = 1$, and by (24) $x_0^*(f) = \int_\Omega f g_0 \, d\mu$, $g_0 \in L^\Psi$. Then, if $f_n = g_0$ whenever $|g_0| \leq n$, $= 0$ otherwise, one has the bounded functions $f_n \in M^\Phi$ (since $\mu(\Omega) < \infty$), and $0 = x_0^*(f_n) = \int_\Omega f_n g_0 \, d\mu = \int_{[-n \leq g_0 \leq n]} g_0^2 \, d\mu$, so that $g_0 = 0$, for all n, implying $g_0 = 0$ a.e. But this makes $x_0^* = 0$ which is impossible since $x_0^*(f_0) = 1$. This shows Theorems 3 and 4 cannot be improved.

Theorem 5. Let L^Φ, L^Ψ be Orlicz spaces on a measure space (Ω, Σ, μ). If the complementary Young's functions Φ, Ψ are such that $M^\Phi = \tilde{L}^\Phi$ and $M^\Psi = \tilde{L}^\Psi$ hold simultaneously, then $L^\Phi(L^\Psi)$ is reflexive with $(L^\Phi)^*$ $((L^\Psi)^*)$ being topologically equivalent to $L^\Psi(L^\Phi)$.

Proof. If $x^* \in (L^\Phi)^*$, by Theorem 4, there is a $g \in L^\Psi$ such that

$$x^*(f) = \int_\Omega fg \, d\mu = \int_\Omega gf \, d\mu = x^{**}(g) = x^{**}(x^*),$$

where $x^{**} \in (L^\Phi)^{**}$. In the notation of [2], writing $x^*(f) = \hat{f}(x^*)$ where $\hat{f}$ is the image of f in $(L^\Phi)^{**}$ under a natural embedding, the above equation shows that (since $x^* \leftrightarrow g$, $x^{**} \leftrightarrow f$), L^Φ and $(L^\Phi)^{**}$ are isomorphic. Since by (25), $\frac{1}{2}\|f\|_\Phi \leq \|x^{**}\| \leq \|f\|_\Phi$, the topological equivalence of L^Φ and $(L^\Phi)^{**}$ follows, and completes the proof, as the remaining conclusions are immediate from the preceding result.

This result is proved, in ([11], p. 154), under the condition that $\Phi(2x) \leq C_1\Phi(x)$, $\Psi(2y) \leq C_2\Psi(y)$, $x, y \geq 0$, (which imply that $M^\Phi = \tilde{L}^\Phi$, $M^\Psi = \tilde{L}^\Psi$) and the *further* assumption that μ be σ-finite. All the previously known results on the reflexivity impose this additional σ-finiteness restriction on μ. No such restriction was needed in the above theorem, and it thus generalizes these results.

If $\Phi(x) = |x|^p$, $1 < p < \infty$, then Theorems 4 and 5 reduce to the usual results for Lebesgue spaces (cf., [2]). For $p = 1$, Theorem 4 takes an interesting form, and in that form it was recently proved in ([5], p. 337). This may be stated as,

Theorem 6. If $\Phi(x) = |x|$, and $\Psi(\cdot)$ is its complementary function, then for every $x^* \in (L^\Phi)^*$, one has

$$x^*(f) = \int_\Omega f \, dG, \quad (= \int_\Omega f g^* \, d\mu), \quad f \in L^\Phi, \tag{26}$$

where G is a bounded countably additive, μ-continuous unique set function, and (in the second form) g^* is an essentially bounded unique quasi-function. Moreover, g^* is measurable if and only if μ is localizable.

This is a restatement of Theorem 4, where $M^\Phi = \tilde{L}^\Phi$ holds, but $\tilde{L}^\Psi$ is in general a proper subset of L^Ψ. The μ-continuity of G is a consequence of Lemma 6.

Finally it may be noted that Theorem 3 reduces to the corresponding result, if $\Psi(x) = |x|$, of ([2], p. 296). In this case the result coincides with [2], and of course is not really a generalization as it is already in its general form.

5. *Application to statistical theory*. The preceding work enables a clarification and extension of certain results in statistical theory. More particularly, a result in ([7], Theorem 9) will be considered here and a more general form of it can now be given.

Let $\{P_\theta, \theta \in \Theta\}$ be a family of probability measures on (Ω, Σ), with Θ as a subset of the line. If $X(\cdot)$ on Ω is a random variable with the distribution determined by P_θ, an important statistical problem is to estimate θ by a suitable measurable function T of X (called an estimator of θ based on an observation on X) in an "optimal" way. This last term means that if θ is the true value, the discrepancy $|T - \theta|$ $(= |\overline{T}|$, say) or a nonnegative function $W(\cdot)$ of it, called a loss function, has the minimum average value relative to P_θ for $\theta \in \Theta$. Much used examples of $W(\cdot)$ are $W(x) = |x|^p$, $p = 1, 2$, or $p \geq 1$. Thus a general function to consider is that $W(\cdot)$ is a nonnegative symmetric convex function with $W(0) = 0$. Then the quantity to be minimized is $R(T, \theta)$, called risk function, where

$$R(T, \theta) = \int_\Omega W(\overline{T}) \, dP_\theta. \tag{27}$$

It is clear that this problem is closely related to the Orlicz space theory. By Lemma 5, $R(T, \theta)$ of (27) is the same as $I_W(G_\theta)$ where $G_\theta(E) = \int_E \overline{T} \, dP_\theta$. To avoid notational confusion, hereafter $W(\cdot)$ will be denoted by $\Phi(\cdot)$. The problem of statistical interest is whether the subset of L^Φ, containing the optimal $\overline{T}$ is nonempty and then to characterize it. Since the measure μ of the preceding work is to be identified with P_θ, $(P_\theta(\Omega) = 1)$ it follows that $L^\Phi \subset L^1$, $L^\Psi \subset L^1$ (by the support line property and finiteness of measures, as used before). Let $\{D_i\} \subset L^\Psi$ be a sequence of elements such that $(D_0 = 1, T \in L^\Phi)$,

$$\text{(i)} \int_\Omega D_i \, dP_\theta = 0, \quad i = 1, 2, \ldots, \qquad \text{(ii)} \int_\Omega T D_i \, dP_\theta = \alpha_i(\theta), \tag{28}$$

where α_i are known constants. Such elements D_i exist. For instance, if the family $\{P_\theta, \theta \in \Theta\}$ is dominated by a fixed σ-finite measure λ_0, and $p(x, \theta)$ are the (Radon-Nikodým) densities of P_θ relative to λ_0, then let $p(x, \theta)$ be differentiable functions of θ. If $D_i = \partial p / \partial \theta^i$, $D_i \in L^\Psi$ and $p(x, \theta)$ satisfies the standard conditions for the interchange of the integral and the derivative, then D_i satisfies the conditions given in (28). Several other functions D_i can be constructed with the properties of (28), even with θ multidimensional. (Cf., [7] for such constructions and related discussion on these points.)

Instead of $R(T, \theta)$, it is technically more convenient (and corresponds naturally to the case $\Phi(x) = |x|^p$) to consider $K(T, \theta)$ ($= K$, say) as the criterion for the optimality of T, where $K = N_\Phi(\overline{T})$, i.e., $K = \inf \{k > 0, I_\Phi(G_\theta/k) \leq 1\}$, G_θ being the set function of the preceding paragraph. Let $\mathfrak{m}(\overline{T})$ be the set of elements of L^Φ, such that $K \leq K_0$, where $K = N_\Phi(\overline{T})$, i.e., the "risk" associated with the choice of T. This set is characterized below. The result generalizes ([7], Theorem 9) and the proof is similar to the special case. It will be sketched here for completeness.

Theorem 7. Let Φ, Ψ be Young's complementary functions, Ψ be continuous, and L^Φ, L^Ψ be the Orlicz spaces on $(\Omega, \Sigma, P_\theta)$. If $\mathfrak{m}(\overline{T}) \subset L^\Phi$, $\{D_i\} \subset L^\Psi$ are as defined above, ($\overline{T} = T - \theta$), then the following conclusions hold:

(a) $\mathfrak{m}(\overline{T})$ is nonempty if and only if (i) for the given K_0, every finite set $\{D_{i_1}, \ldots, D_{i_m}\}$ of $\{D_i\}$, and any scalars $a_1, \ldots, a_n$,

$$|\Sigma_{j=1}^n a_j \alpha_{i_j}(\theta)| \leq K_0 \|\Sigma_{j=1}^n a_j D_{i_j}\|_\Psi, \tag{29}$$

and (ii) $\{D_i\} \subset M^\Psi$. [M^Ψ is the subspace of L^Ψ as in Theorem 4.]

(b) For any $\overline{T} \in \mathfrak{m}(T)$, $K_0 \geq N_\Phi(\overline{T}) \geq C_0$ ($= \inf K > 0$ satisfying (29)).

(c) If $\overline{T}_0 \in \mathfrak{m}(\overline{T})$, satisfies $N_\Phi(\overline{T}_0) = C_0$, then it is essentially unique.

Note. In the special result of [7] the condition $\Psi(2x) \leq C\Psi(x)$, was imposed (so that $M^\Psi = \tilde{L}^\Psi$ and (a) (ii) above is implied). But this is more restrictive than even the condition $M^\Psi = \tilde{L}^\Psi$ (which is not assumed here) since there are functions satisfying this but not $\Psi(2x) \leq C\Psi(x)$. (As pointed out in [6], $\Phi(x) = e^{|x|} - 1$, is such an example.) From the preceding theory, it follows that the present conditions are the most general ones for this problem. (In [7], for this condition, Ψ and Φ were interchanged due to a typographical error.)

Proof. (a) If $\overline{T} \in \mathfrak{m}(\overline{T})$, then it is immediate that (29) holds for $D_i \in L^\Psi$ and (ii) is not necessary, since (by Hölder inequality),

$$|\Sigma_{j=1}^n a_j \alpha_{i_j}(\theta)| = |\textstyle\int_\Omega \overline{T}(\Sigma_{j=1}^n a_j D_{i_j})\, dP_\theta| \leq N_\Phi(\overline{T}) \|\Sigma_{j=1}^n a_j D_{i_j}\|_\Psi.$$

Thus if $\mathfrak{m}(\overline{T})$ is nonempty, (29) holds, no matter what Φ and Ψ are.

Conversely, if (29) holds for any finite set of D_i's in M^Ψ and the scalars, a_i's, then by a result of Hahn (a consequence of Hahn-Banach theorem, cf., [2], p. 86), there exists a bounded linear functional x^* on M^Ψ such that $x^*(D_i) = \alpha_i$ for all i, considering the closed subspace M^Ψ of L^Ψ as a B-space in its own right. Hence by Theorem 4, (cf., (22)) there exists a unique G_θ in $A_\Phi{}^0$, such that

$$x^*(D_i) = \textstyle\int_\Omega D_i\, dG_\theta = \int_\Omega D_i \overline{T}\, dP_\theta. \tag{30}$$

(Here $\overline{T}$ is actually measurable, since P_θ is a finite measure. Note that this statement would be impossible if D_i are not in M^Ψ, by Theorem 4.) By (23) applied to (30), it results that $\|x^*\| = N_\Phi(\overline{T})$ ($= \|G\|_\Phi$) $\leq K_0$. Since $D_0 = 1$, $x^*(D_0) = \alpha_0(\theta)$, it follows that $\overline{T} \in \mathfrak{m}(\overline{T})$, and $K_\theta(\overline{T}) \leq K_0$. This proves (a).

Part (b) is clear, and (c) follows from the fact that if $\overline{T}_1 \in \mathfrak{m}(\overline{T})$ such that $N_\Phi(\overline{T}_0) = N_\Phi(\overline{T}_1) = C_0$, then $T_2 = \alpha\overline{T}_0 + \beta\overline{T}_2$, $\alpha + \beta = 1$, $\alpha \geq 0$, is also in $\mathfrak{m}(\overline{T})$, and $C_0 \leq N_\Phi(\alpha\overline{T}_0 + \beta\overline{T}_1) \leq \alpha N_\Phi(\overline{T}_0) + \beta N_\Phi(\overline{T}_1) = C_0$ gives $\overline{T}_0$ and $\overline{T}_1$ to be proportional. From this it follows readily that $\overline{T}_1 = \overline{T}_0$, and completes the proof.

Final remarks. 1. When the existence of estimators is thus settled in the above theorem, the actual construction of the best estimators T (i.e., elements of $\mathfrak{m}(\overline{T})$) is an interesting, and in general nontrivial, problem in statistical theory. [For a discussion on statistical implications, with $\Phi(x) = |x|^p$, see E. W. Barankin, *Ann. Math. Statist.*, **20** (1949), 477–501.]

2. As seen from the work of this paper, and motivated by the study of [4], it would be interesting to work out the representation theory, and other problems, of the B-spaces $A_\Phi(\mu)$ and $A_\Psi(\mu)$ which may be called the "Orlicz spaces of set functions". This will both generalize [4] and give a better understanding of the structure of these spaces.

After the preparation of this paper, the author learned of a paper by T. Andô (Linear functionals on Orlicz spaces, *Nieuw Arch. Wisk.*, **8** (1960), 1–16) in which another special case of Theorem 3 was proved. The relation between these results is as follows: Let $\mu(\Omega) < \infty$. Then (and only then) Φ (or Ψ)-bounded variation of a set function implies bounded variation. So the function G of Theorem 3 will be a bounded additive set function and in this case, by Yosida-Hewitt theorem ([2], III.7.8), G can be uniquely decomposed as $G = G_1 + G_2$ where G_1 is (bounded) countably additive and G_2 is purely finitely additive (bounded) set function on Σ. Since G_1 and G_2 vanish on μ-null sets, Theorem 3 reduces to

$$x^*(f) = \int_\Omega fg\, d\mu + \int_\Omega f\, dG_2, \quad f \in L^\Phi, \quad g \in L^\Psi, \tag{*}$$

where the first integral is a consequence of the case (i) of Theorem 4. The result (*) includes Andô's theorem where, in addition to $\mu(\Omega) < \infty$, it was also assumed that Φ, Ψ satisfy the conditions

$$\frac{\Phi(x)}{x} \to \infty, \quad \frac{\Psi(x)}{x} \to \infty \quad \text{as} \quad x \to \infty.$$

It may be seen that the proofs of Theorems 1 and 2 simplify considerably with these additional assumptions. It appears that all the work of the present paper is needed to generalize Andô's result to that of Theorem 3, since the above mentioned (simplifying) assumptions were used in a crucial manner in his paper.

REFERENCES

[1] R. G. Bartle, "A general bilinear vector integral", *Studia Math.*, **15** (1956), 337–352.

[2] N. Dunford and J. T. Schwartz, *Linear Operations, Part I: General Theory*, Interscience, New York, 1958.

[3] M. A. Krasnosel'skii and Ya. B. Rutickii, *Convex Functions and Orlicz Spaces*, (Translation) P. Noordhoff Ltd., Groningen, 1961.

[4] S. Leader, "The Theory of L^p spaces for finitely additive set functions", *Ann. of Math.*, **58** (1953), 528–543.

[5] E. J. McShane, "Families of measures and representations of algebras of operators", *Trans. Amer. Math. Soc.*, **102** (1962), 328–345.

[6] M. Morse and W. Transue, "Functionals F bilinear over the product $A \times B$ of two pseudo-normed vector spaces, II. Admissible spaces A", *Ann. of Math.*, **51** (1950), 576–614.

[7] M. M. Rao, "Theory of lower bounds for risk functions in estimation", *Math. Annalen*, **143** (1961), 379–398.

[8] G. Weiss, "A note on Orlicz spaces". *Portugaliae Math.*, **15** (1956), 35–47.

[9] A. C. Zaanen, "The Radon-Nikodým theorem, I, II", *Indag. Math.*, **23** (1961), 157–187.

[10] A. C. Zaanen, "On a class of Banach spaces", *Ann. of Math.*, **47** (1946), 645–666.

[11] A. C. Zaanen, *Linear Analysis*, North-Holland Publishing Co., Amsterdam, 1953.

(Received, March 10, 1964)

Carnegie Institute of Technology
Pittsburgh, Pennsylvania.