A brief introduction to N–functions and Orlicz function spaces

John Alexopoulos
Kent State University, Stark Campus

March 12, 2004

Contents

Chapter 1

Introduction: Uniform integrability

1.1 Definitions

Throughout these notes, unless stated otherwise, all measures are of bounded variation and countably additive. In particular (Ω, Σ, μ) will denote a probability space.

Recall that *a subset $\mathcal{K}$ of $L^1(\mu)$ is called uniformly integrable if*

$$\lim_{c\to\infty} \sup \left\{ \int_{[|f| \geq c]} |f| \, d\mu : f \in \mathcal{K} \right\} = 0.$$

That is given $\varepsilon > 0$ there is a $c_\varepsilon > 0$ so that for each $f \in \mathcal{K}$ and each $c \geq c_\varepsilon$ we have

$$\int_{[|f| \geq c]} |f| \, d\mu < \varepsilon.$$

Alternatively *A subset $\mathcal{K}$ of $L^1(\mu)$ is uniformly integrable if and only if it is L^1–bounded and for each $\varepsilon > 0$ there is a $\delta > 0$ so that $\sup \left\{ \int_A |f| \, d\mu : f \in \mathcal{K} \right\} < \varepsilon$ for all $A \in \Sigma$ with $\mu(A) < \delta$.*

In order to establish the equivalence of the two notions above, first note that for all measurable A, $f \in \mathcal{K}, \ c > 0$ we have

$$\int_A |f| \, d\mu = \int_{A \cap [|f| < c]} |f| \ d\mu + \int_{A \cap [|f| \geq c]} |f| \, d\mu \leq c\mu(A) + \int_{[|f| \geq c]} |f| \ d\mu.$$

Fix $\varepsilon > 0$ and choose $c_0 > 0$ so that $\sup \left\{ \int_{[|f| \geq c]} |f| \ d\mu : f \in \mathcal{K} \right\} < \frac{\varepsilon}{2}$ whenever $c \geq c_0$. Then for all $f \in \mathcal{K}$ we have

$$\int_\Omega |f| \; d\mu \leq c_0\mu(\Omega) + \int_{[|f|\geq c_0]} |f| \; d\mu \leq c_0 + \frac{\varepsilon}{2}.$$

and thus $\mathcal{K}$ is L^1 bounded. Now let $0 < \delta < \frac{\varepsilon}{2c_0}$. Then for all measurable A with $\mu(A) < \delta$ and all $f \in \mathcal{K}$ we have

$$\int_A |f| \; d\mu \leq c_0\mu(A) + \int_{[|f|\geq c_0]} |f| \; d\mu < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

To establish the converse, fix $\varepsilon > 0$ and choose $\delta > 0$ so that $\sup\left\{\int_A |f| \; d\mu : f \in \mathcal{K}\right\} < \varepsilon$ whenever A is measurable with $\mu(A) < \delta$. Let $M = \sup\left\{\int_\Omega |f| \; d\mu : f \in \mathcal{K}\right\}$ and choose $c_0 > 0$ so that $\frac{M}{c_0} < \delta$. Then for all $f \in \mathcal{K}$ and all $c \geq c_0$ we have

$$\mu([|f| \geq c]) \leq \frac{1}{c} \int_{[|f|\geq c]} |f| \; d\mu \leq \frac{M}{c_0} < \delta.$$

So $\int_{[|f|\geq c]} |f| \; \mathrm{d}\mu < \varepsilon$ and so we are done.

1.2 The theorem of De La Vallée Poussin

One characterization of uniformly integrable sets is an old theorem that finds its roots in Harmonic Analysis and Potential theory and it is due to De La Vallée Poussin.

Theorem 1.1 (De La Vallée Poussin) *A subset $\mathcal{K}$ of $L^1(\mu)$ is uniformly integrable if and only if there is a non-negative and convex function Q with $\lim_{t\to\infty} \frac{Q(t)}{t} = \infty$ so that*

$$\sup\left\{\int_\Omega Q(|f|)\, d\mu : f \in \mathcal{K}\right\} < \infty\ .$$

Proof. Suppose that $\mathcal{K}$ is a uniformly integrable subset of $L^1(\mu)$. We will construct a non-negative and non-decreasing function q that is constant on $[n, n+1)$ for $n = 0, 1, \ldots$ with $\lim_{t\to\infty} q(t) = \infty$ and we will set $Q(x) = \int_0^x q(t)\, dt$ for $x > 0$. Use the hypothesis to choose a subsequence (c_n) of the positive integers so that

$$\sup\left\{\int_{[|f|\geq c_n]} |f| \; d\mu : f \in \mathcal{K}\right\} < \frac{1}{2^n} \;\forall n = 1, 2, \ldots$$

Then for each $f \in \mathcal{K}$ and all $n = 1, 2, \ldots$ we have

$$\begin{aligned}\int_{[|f| \geq c_n]} |f| \, d\mu &= \sum_{m=c_n}^{\infty} \int_{[m \leq |f| < m+1]} |f| \, d\mu \\ &\geq \sum_{m=c_n}^{\infty} m\mu([m \leq |f| < m+1]) \\ &\geq \sum_{m=c_n}^{\infty} \mu([|f| \geq m]) \, .\end{aligned}$$

So for all $f \in \mathcal{K}$ we have

$$\sum_{n=1}^{\infty} \sum_{m=c_n}^{\infty} \mu([| \, f \, | \geq m]) \leq 1 \, .$$

Now for $m = 1, 2, \ldots$ let q_m be the number of the positive integers n, for which $c_n \leq m$. Then $q_m \nearrow \infty$. Furthermore observe that

$$\sum_{n=1}^{\infty} \sum_{m=c_n}^{\infty} \mu([|f| \geq m]) = \sum_{k=1}^{\infty} q_k \mu([|f| \geq k]) \, .$$

Let $q_0 = 0$ and define $q(t) = q_n$ if $t \in [n, n+1)$ for $n = 0, 1, 2, \ldots$. Then if $Q(x) = \int_0^x q(t) \, dt$ we have

$$\begin{aligned}\int_{\Omega} Q(|f|) \, d\mu &= \sum_{n=0}^{\infty} \int_{[n \leq |f| < n+1]} Q(|f|) \, d\mu \\ &\leq \sum_{n=0}^{\infty} \left(\sum_{m=0}^{n} q_m \right) \cdot \mu \left([n \leq |f| < n+1] \right) \\ &= q_0 \cdot \mu([0 \leq |f| < 1]) + (q_0 + q_1) \cdot \mu([1 \leq |f| < 2]) + \cdots \\ &= \sum_{n=0}^{\infty} q_n \mu([|f| \geq n]) \\ &\leq 1.\end{aligned}$$

So $\sup \left\{ \int_{\Omega} Q(| \, f \, |) \, d\mu : f \in \mathcal{K} \right\} < \infty$.

To see that Q is convex, fix $0 \leq x_1 < x_2$. We then have

$$\begin{aligned}
Q\left(\frac{1}{2}(x_1+x_2)\right) &= \int_0^{\frac{1}{2}(x_1+x_2)} q(t)\,dt \\
&= \int_0^{x_1} q(t)\,dt + \int_{x_1}^{\frac{1}{2}(x_1+x_2)} q(t)\,dt \\
&\leq \int_0^{x_1} q(t)\,dt + \frac{1}{2}\int_{x_1}^{\frac{1}{2}(x_1+x_2)} q(t)\,dt + \frac{1}{2}\int_{\frac{1}{2}(x_1+x_2)}^{x_2} q(t)\,dt \\
&= \frac{1}{2}\int_0^{x_1} q(t)\,dt + \frac{1}{2}\int_0^{x_2} q(t)\,dt \\
&= \frac{1}{2}(Q(x_1)+Q(x_2))\ .
\end{aligned}$$

Finally observe that

$$Q(x) = \int_0^x q(t)\,dt \geq \int_{\frac{x}{2}}^x q(t)\,dt \geq \frac{x}{2}q(\frac{x}{2})$$

and thus $\frac{Q(x)}{x} \geq \frac{1}{2}q(\frac{x}{2}) \to \infty$ as $x \to \infty$.

We now prove the converse. Let $M = \sup\left\{\int_\Omega Q(|f|)\,d\mu : f \in \mathcal{K}\right\}$. Let $\varepsilon > 0$ and choose $c_0 > 0$ so that $\frac{Q(t)}{t} > \frac{M}{\varepsilon}$ whenever $t \geq c_0$. Then for $f \in \mathcal{K}$ and $c \geq c_0$ we have that $|f| < \frac{\varepsilon}{M}Q(|f|)$ on the set $[|f| \geq c]$. Thus

$$\int_{[|f|\geq c]} |f|\ d\mu \leq \frac{\varepsilon}{M}\int_{[|f|\geq c]} Q(|f|)\,d\mu \leq \frac{\varepsilon}{M}M = \varepsilon$$

and so we are done. ■

1.3 The Dunford-Pettis theorem

The well known theorem of Dunford and Pettis which states that, *a subset $\mathcal{K}$ of $L^1(\mu)$ is uniformly integrable if and only if it is relatively weakly compact*, gives some deep insight to the notion of uniform integrability. Our proof of this theorem will more or less be a combination of those in [3] and [2].

1.3.1 The space $\Sigma(\mu)$

Define a (pseudo)metric on the σ-algebra Σ by $d(A,B) = \mu(A \triangle B)$ for all $A, B \in \Sigma$. It is clear that $d(A,B) = \mu(A \triangle B) = \mu(B \triangle A) = d(B,A)$ and

$$\begin{aligned} d(A,B) + d(B,C) &= \mu(A \triangle B) + \mu(B \triangle C) \\ &\geq \mu((A \triangle B) \cup (B \triangle C)) \\ &= \mu(AB^c \cup A^c B \cup BC^c \cup B^c C) \\ &= \mu(B^c (A \cup C) \cup B (AC)^c) \\ &\geq \mu(B^c (A \triangle C) \cup B (A \triangle C)) \\ &= \mu(A \triangle C) \\ &= d(A,C). \end{aligned}$$

Further, the (pseudo)metric space $(\Sigma(\mu), d)$ is a complete one, because the map $A \longmapsto \chi_A$ is an isometry $\Sigma(\mu) \to L^1(\mu)$ onto the closed set of all the characteristic functions (or more directly, if (E_n) is a Cauchy sequence in $\Sigma(\mu)$ then $E_n \to \limsup E_n = \liminf E_n$, μ-almost surely).

Also notice that if λ is an absolutely continuous measure with respect to μ, then λ is a continuous real valued function on the (pseudo)metric space $\Sigma(\mu)$, for if $E_n \to E$ in $\Sigma(\mu)$ then both $\mu(E \setminus EE_n) \to 0$ and $\mu(E_n \setminus EE_n) \to 0$ and so

$$\begin{aligned} |\lambda(E_n) - \lambda(E)| &= |\lambda(E_n) - \lambda(EE_n) + \lambda(EE_n) - \lambda(E)| \\ &\leq \lambda(E_n \setminus EE_n) + \lambda(E \setminus EE_n) \to 0. \end{aligned}$$

It is also noteworthy that the set-theoretic operations of union, intersection, symmetric difference and complementation are continuous.

Theorem 1.2 (Vitali-Hahn-Saks) *Let (Ω, Σ, μ) be a probability space and (λ_n) a sequence of μ-continuous measures. If $\lim_n \lambda_n(E)$ exists for each $E \in \Sigma$ then*

$$\lim_{\mu(E)\to 0} \sup_n |\lambda_n(E)| = 0$$

Proof. Fix $\varepsilon > 0$. As each λ_n is continuous on $\Sigma(\mu)$, the sets $\Sigma_{n,m} = \left\{E \in \Sigma : |\lambda_n(E) - \lambda_m(E)| \leq \frac{\varepsilon}{3}\right\}$ are closed for all positive integers n, m. Hence the sets $\Sigma_p = \bigcap_{n,m \geq p} \Sigma_{n,m}$ are also closed for each positive integer p. Since $\lim_n \lambda_n(E)$ exists for each $E \in \Sigma$ then the complete metric space $\Sigma(\mu) = \bigcup_{p=1}^{\infty} \Sigma_p$ and thus, thanks to the Baire Category theorem, there is a positive integer q such that the closed set Σ_q has non-empty interior. That is, there are $r > 0,\ A \in \Sigma_q$ so that the ball $\mathcal{B}(A,r) \subset \Sigma_q$ i.e. $|\lambda_n(E) - \lambda_m(E)| \leq \frac{\varepsilon}{3}$ whenever $\mu(A \triangle E) < r$ and $n, m \geq q$.

Now choose $0 < \delta < r$ so that $\max_{n=1,\ldots,q} |\lambda_n(B)| < \frac{\varepsilon}{3}$ for all $B \in \Sigma$ with $\mu(B) < \delta$. Notice that if $\mu(B) < \delta$ then $A \cup B,\ A \setminus B \in \mathcal{B}(A,r)$. Furthermore note that $B = (A \cup B) \setminus (A \setminus B)$ and so for all positive integers n and all $B \in \Sigma$ with $\mu(B) < \delta$ we have

$$\begin{aligned}
|\lambda_n(B)| &\leq |\lambda_q(B)| + |\lambda_n(B) - \lambda_q(B)| \\
&= |\lambda_q(B)| + |\lambda_n(A \cup B) - \lambda_n(A \setminus B) - \lambda_q(A \cup B) + \lambda_q(A \setminus B)| \\
&\leq |\lambda_q(B)| + |\lambda_n(A \cup B) - \lambda_q(A \cup B)| + |\lambda_n(A \setminus B) - \lambda_q(A \setminus B)| \\
&< \varepsilon
\end{aligned}$$

■

Now we are ready for the Dunford-Pettis theorem:

Theorem 1.3 (Dunford-Pettis) *A subset $\mathcal{K}$ of $L^1(\mu)$ is uniformly integrable if and only if it is relatively weakly compact.*

Proof. **($\Longrightarrow$):** Suppose that $\mathcal{K} \subset L^1(\mu)$ is uniformly integrable. Let $\lambda \in \overline{\mathcal{K}}^{\text{weak}^*} \subset (L^\infty(\mu))^* = (L^1(\mu))^{**}$. For simplicity set $\lambda(\chi_E) \equiv \lambda(E)$ for all $E \in \Sigma$. Notice that λ can be viewed as a finitely additive set function on Σ. As $\lambda \in \overline{\mathcal{K}}^{\text{weak}^*}$ we have that

$$|\lambda(E)| \leq \sup_{f \in \mathcal{K}} \left| \int_E f d\mu \right| \leq \sup_{f \in \mathcal{K}} \int_E |f| \, d\mu$$

and so λ is μ-continuous (hence countably additive as well), thanks to $\mathcal{K}$'s uniform integrability. Thus by the Radon-Nikodym Theorem, $\lambda(E) = \int_E \frac{d\lambda}{d\mu} d\mu$ for all $E \in \Sigma$. Passing to simple functions as well as an elementary density argument, convinces as that $\lambda(g) = \int_\Omega g \frac{d\lambda}{d\mu} d\mu$ for all $g \in L^\infty(\mu)$. Consequently $\lambda \in L^1(\mu)$ and so, by a simple comparison of topologies, $\overline{\mathcal{K}}^{\text{weak}^*}$and $\overline{\mathcal{K}}^{\text{weak}}$ are equal and topologically identical. Hence $\overline{\mathcal{K}}^{\text{weak}}$ is (weakly) compact thanks to Alaoglou's theorem.

($\Longleftarrow$): Now assume that $\mathcal{K}$ is relatively weakly compact and suppose that $\mathcal{K}$ is not uniformly integrable. Then there is an exceptional $\varepsilon_0 > 0$, a sequence $(f_n) \subset \mathcal{K}$ and a sequence (E_n) of measurable sets with $\mu(E_n) \to 0$ such that $\left|\int_{E_n} f_n d\mu\right| \geq \varepsilon_0$ for all positive integers n. As $\mathcal{K}$ is relatively weakly compact, the Eberlein-Smulian theorem guarantees the existence of a subsequence (n_k) of the positive integers and that of a
function $f \in L^1(\mu)$ so that $f_{n_k} \overset{\text{weakly}}{\to} f$. For each k and each $E \in \Sigma$, set $\lambda_k(E) = \int_E f_{n_k} d\mu$ and notice that $\lim_k \lambda_k(E) = \int_E f d\mu$. Hence by the Vitali-Hahn-Saks theorem $\lim\limits_j \sup\limits_k \left|\lambda_k\left(E_{n_j}\right)\right| = 0$. In particular $\lim\limits_j \left|\lambda_j\left(E_{n_j}\right)\right| = \left|\int_{E_{n_j}} f_{n_j} d\mu\right| = 0$ contradicting the fact that $\left|\int_{E_n} f_n d\mu\right| \geq \varepsilon_0$ for all positive integers n. ■

Chapter 2

N–Functions

2.1 Definitions and elementary results

In this section we will summarize the necessary facts about a special class of convex functions called N–functions. For a detailed account of these facts, the reader could consult [4] or [5].

Definition 2.1 *Let $f : [0, \infty) \to [0, \infty)$ be a right continuous, monotone increasing function with*

1. $f(0) = 0$;
2. $\lim_{t\to\infty} f(t) = \infty$;
3. $f(t) > 0$ *whenever* $t > 0$;

then the function defined by

$$F(x) = \int_0^{|x|} f(t)\,dt$$

is called an N–function. Alternatively, the function F is an N–function if and only if F is continuous, even and convex with

1. $\lim_{x\to 0} \frac{F(x)}{x} = 0$
2. $\lim_{x\to\infty} \frac{F(x)}{x} = \infty$
3. $F(x) > 0$ *if* $x > 0$.

In that case if $f = F'_+$*, the right derivative of* F *then* f *satisfies* $f(0) = 0$*;* $\lim_{t\to\infty} f(t) = \infty$*;* $f(t) > 0$ *whenever* $t > 0$*; and* $F(x) = \int_0^{|x|} f(t)\,dt$*.*

It is not hard to see that the composition of two N–functions is an N–function. A little more thought convinces us about the truth of the converse. i.e. *every* N*–function* F *is the composition of two other* N*–functions. That is, there are* N*–functions* F_1 *and* F_2 *so that* $F = F_2 \circ F_1$*.* Here is why:
Given the N–function F_1 then F_2 is uniquely determined by $F_2 = F \circ F_1^{-1}$. Since for $x > 0$ we have that $f_2(x) = \frac{f(F_1^{-1}(x))}{f_1(F_1^{-1}(x))}$ and F_1^{-1} is increasing, tends to zero as $x \to 0$ and to infinity as $x \to \infty$ it is necessary and sufficient for F_2 to be an N–function if $\frac{f}{f_1}$ satisfies all the conditions that right derivatives of N–functions satisfy. Take $f_1 = f^p$ for any $0 < p < 1$ and the rest follows.
N–functions come in *mutually complementary pairs.* In fact we have the following

Definition 2.2 *For an* N*–function* F *define*

$$G(x) = \int_0^x g(t)\,dt$$

where g *is the right inverse of the right derivative* f *of* F *(see figure 2.1).* G *is an* N*–function called the complement of* F*. Furthermore it is plain that the complement of* G *is* F*.*

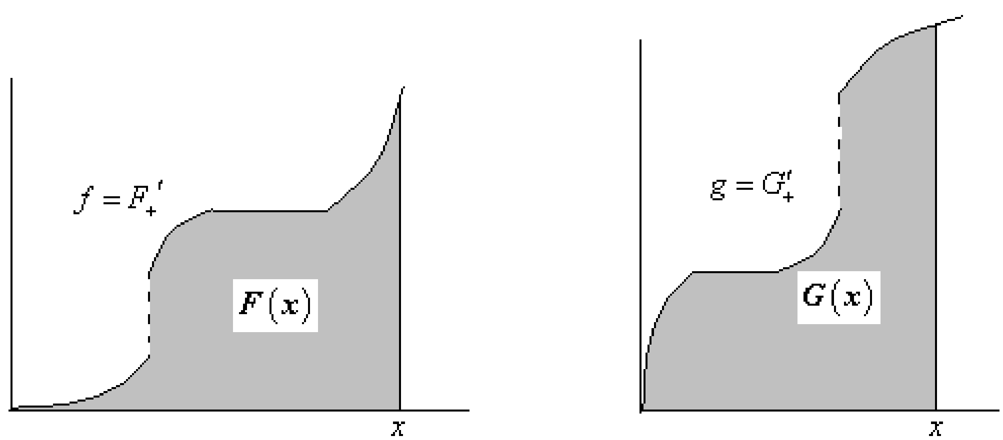

Figure 2.1: A pair of complementary N-functions

Complementary pairs of N-functions satisfy,

Theorem 2.1 (Young's Inequality) *If F and G are two mutually complementary N-functions then*

$$uv \leq F(u) + G(v)\ \forall u, v \in \mathbf{R}$$

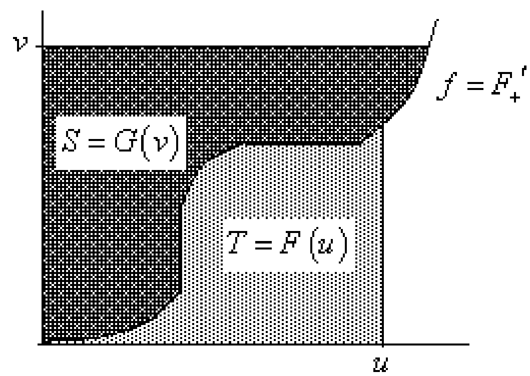

Figure 2.2: A geometric interpretation of Young's Inequality.

Figure 2.2 above, makes Young's Inequality geometrically clear. It is also clear from the figure that equality is attained when $v = f(|u|)\text{sgn }u$ or $u = g(|v|)\text{sgn }v$. In particular we have

$$|u|f(|u|) = F(u) + G(f(|u|))$$

and

$$|v|g(|v|) = F(g(|v|)) + G(v).$$

Consequently we have an alternative definition for the complementary function G:

$$G(x) = \max\{t|x| - F(t) : t \geq 0\}$$

Young's Inequality gives rise to the following

Theorem 2.2 *Suppose that F_1, F_2 are N–functions with complements G_1 and G_2 respectively. Suppose that $F_1(x) \leq F_2(x)$ for $x \geq x_0$. Then $G_2(y) \leq G_1(y)$ for $y \geq y_0 = f_2(x_0) = F'_{2+}(x_0)$.*

Proof. Let f_1, f_2, g_1 and g_2 be the right derivatives of F_1, F_2, G_1 and G_2 respectively. Then $g_2(y) \geq x_0$ whenever $y \geq y_0 = f_2(x_0)$. Note that $yg_2(y) = G_2(y) + F_2(g_2(y))$ (equality case of Young's Inequality). Furthermore by Young's Inequality $yg_2(y) \leq G_1(y) + F_1(g_2(y))$ and so $G_2(y) + F_2(g_2(y)) \leq G_1(y) + F_1(g_2(y)$. But since $F_1(g_2(y) \leq F_2(g_2(y))$ we have that $G_2(y) \leq G_1(y)$. ■

N-functions grow at different rates. The following definition makes their comparison possible.

Definition 2.3 *For N-functions F_1, F_2 we write $F_1 \prec F_2$ if there is a $K > 0$ so that $F_1(x) \leq F_2(Kx)$ for large values of x. If $F_1 \prec F_2$ and $F_2 \prec F_1$ then we say that F_1 and F_2 are equivalent and we write $F_1 \sim F_2$.*

If two N-functions are comparable then so are their complements in the reverse. Indeed *if $F_1 \prec F_2$ then $G_2 \prec G_1$, where G_i is the complement of F_i. In particular if $F_1(x) \leq F_2(x)$ for large values of x then $G_2(x) \leq G_1(x)$ for large values of x.*

It is worth noting at this stage that every N–function F is equivalent to the N–function Φ defined by $\Phi(x) = \int_0^x \frac{F(t)}{t}\,dt$. After all

$$\Phi(x) = \int_0^x \frac{F(t)}{t}\,dt \leq \int_0^x \frac{tf(t)}{t}\,dt = F(x).$$

Furthermore

$$\frac{F(t)}{t} = \frac{1}{t}\int_0^t f(s)\,ds \geq \frac{1}{t}\int_{\frac{t}{2}}^t f(s)\,ds \geq \frac{1}{2} f\left(\frac{t}{2}\right)$$

and so

$$\Phi(2x) = \int_0^{2x} \frac{F(t)}{t}\,dt \geq \frac{1}{2}\int_0^{2x} f\left(\frac{t}{2}\right)\,dt = \int_0^x f(s)\,ds = F(x)$$

Now the convexity of F ensures that $F(\alpha x) \leq \alpha F(x)$ for $0 < \alpha < 1$ and so $\frac{F(t)}{t}$ is increasing. A convex function Q is called the principal part of an N-function F, if $F(x) = Q(x)$ for large x. All convex functions of the "De La Vallée Poussin" type are principal parts of N-functions. Specifically we have

Theorem 2.3 *If Q is convex with $\lim_{x\to\infty} \frac{Q(x)}{x} = \infty$ then Q is the principal part of some N-function.*

Proof. Since $\lim_{x\to\infty} \frac{Q(x)}{x} = \infty$ then $\lim_{x\to\infty} Q(x) = \infty$ and so there is x_0 so that $Q(x) > 0$ for $x \geq x_0$. Thus $Q(x) - Q(x_0) = \int_{x_0}^x q(t)\,dt$ where q is the right derivative of Q. Of course

q is non–decreasing and right–continuous. Furthermore $\lim_{t\to\infty} q(t) = \infty$ since otherwise $q(t) \le b$ would imply $Q(x) \le b(x - x_0) + q(x_0)$ contradicting the fact that $\lim_{x\to\infty} \frac{Q(x)}{x} = \infty$. Without loss of generality assume that $q(t) > 0$ for $t \ge x_0$. Now since $\lim_{t\to\infty} q(t) = \infty$ there is $x_1 \ge x_0 + 1$ so that $q(x_1) > q(x_0 + 1) + Q(x_0)$. Then

$$\begin{aligned} Q(x_1) &= \int_{x_0}^{x_0+1} q(t)\,dt + \int_{x_0+1}^{x_1} q(t)\,dt + Q(x_0) \\ &\le q(x_0+1) + Q(x_0) + q(x_1)(x_1 - x_0 - 1) \\ &< q(x_1)(x_1 - x_0) \end{aligned}$$

and thus $\alpha = \frac{x_1 q(x_1)}{Q(x_1)} > 1$. Now define F by

$$F(x) = \begin{cases} \frac{Q(x_1)}{x_1^\alpha}|x|^\alpha & \text{for} \quad |x| \le x_1 \\ Q(x) & \text{for} \quad |x| \ge x_1 \end{cases}$$

Now F is an N–function since its right derivative

$$f(x) = F'_+(x) = \begin{cases} \frac{\alpha Q(x_1)}{x_1^\alpha}|x|^{\alpha-1} & \text{for} \quad 0 \le |x| \le x_1 \\ q(x) & \text{for} \quad |x| \ge x_1 \end{cases}$$

is right continuous for $x \ge 0$ satisfying $f(0) = 0$, $\lim_{t\to\infty} f(t) = \infty$ and $f(t) > 0$ whenever $t > 0$. ■

2.2 Conditions on N–functions

There are several important classes of N-functions. Among other things, these conditions relate to the growth of N-functions. Here are some of the most important definitions:

Definition 2.4 *Let F be an N-function and let G denote its complement. Then*

1. *F is said to satisfy the Δ_2 condition ($F \in \Delta_2$) if $\limsup_{x\to\infty} \frac{F(2x)}{F(x)} < \infty$. That is, there is a $K > 0$ so that $F(2x) \le KF(x)$ for large values of x. If $G \in \Delta_2$ we say that $F \in \nabla_2$.*

2. *F is said to satisfy the Δ' condition ($F \in \Delta'$) if there is a $K > 0$ so that $F(xy) \le KF(x)F(y)$ for large values of x and y. If $G \in \Delta'$ we say that $F \in \nabla'$.*

3. *F is said to satisfy the Δ_3 condition ($F \in \Delta_3$) if there is a $K > 0$ so that $xF(x) \leq F(Kx)$ for large values of x. If $G \in \Delta_3$ we say that $F \in \nabla_3$.*

4. *F is said to satisfy the Δ^2 condition ($F \in \Delta^2$) if there is a $K > 0$ so that $(F(x))^2 \leq F(Kx)$ for large values of x. If $G \in \Delta^2$ we say that $F \in \nabla^2$.*

It is plain that all the classes defined above are closed under the equivalence of N–functions.

2.2.1 The Δ_2 condition

Among all these conditions, the Δ_2 condition is perhaps the most important. It is worth noting that $F \in \Delta_2$ is equivalent to $F(cx) \leq k_c F(x)$ for large values of x, where c can be any number greater than 1. Indeed for $2^n \geq c$ and large enough x we have $F(cx) \leq F(2^n x) \leq K^n F(x) = k_c F(x)$. Conversely, if $2 \leq c^n$ we have $F(2x) \leq F(c^n x) \leq k_c^n F(x)$ for large values of x.

N-functions that satisfy the Δ_2 condition have growth rates less than that of power functions as we can see by the following theorem:

Theorem 2.4 *If $F \in \Delta_2$ then there are constants $\alpha > 1$ and $c > 0$ so that $F(x) \leq c|x|^\alpha$ for large values of x.*

The proof of this theorem follows easily from the following independently useful lemma:

Lemma 2.5 *$F \in \Delta_2$ iff there are constants $\alpha > 1$ and x_0 so that*

$$\frac{xf(x)}{F(x)} < \alpha \text{ for all } x \geq x_0$$

where f is the right derivative of F.

Proof. First note that

$$kF(x) \geq F(2x) = \int_0^{2x} f(t)\,dt > \int_x^{2x} f(t)\,dt > xf(x) \text{ for large enough } x$$

and so necessity follows.

To see the converse, observe that since $xf(x) > F(x)$ for all x, $\alpha > 1$. Now for $x \geq x_0$ we have that

$$\int_x^{2x} \frac{f(t)}{F(t)}\,dt < \alpha \int_x^{2x} \frac{1}{t}\,dt = \alpha \log 2.$$

Hence

$$\log \frac{F(2x)}{F(x)} = \int_{F(x)}^{F(2x)} \frac{1}{t}\, dt < \alpha \log 2$$

and so $F(2x) < 2^\alpha F(x)$. ■

Now the proof of the theorem follows from the fact that

$$\int_{x_0}^{x} \frac{f(t)}{F(t)}\, dt < \alpha \int_{x_0}^{x} \frac{1}{t}\, dt \ \text{ for all } x > x_0$$

i.e. that $F(x) < \frac{F(x_0)}{x_0^\alpha} x^\alpha$.

Lemma 2.5 offers a test for the Δ_2 condition. It is often useful to have a direct test to determine when *the complement* of an N–function satisfies the Δ_2 condition (i.e. when does an N–function belong to ∇_2). The following theorem offers such a test:

Theorem 2.6 *$G \in \nabla_2$ iff there exist constants $\beta > 1$ and $x_0 \geq 0$ such that*

$$G(x) \leq \frac{1}{2\beta} G(\beta x) \ \text{ for all } x \geq x_0\,.$$

Lets first isolate the following useful lemma:

Lemma 2.7 *If $F_1(x) = aF(bx)$ where a and b are positive then the complement G_1 of F_1 is given by $G_1(x) = aG\left(\frac{x}{ab}\right)$, where G is the complement of F.*

Proof. Notice that the right derivative f_1 of F_1 is given by $f_1(t) = abf(bt)$ where f is the right derivative of F. Then the right derivative g_1 of G_1 is given by $g_1(s) = \frac{1}{b} g\left(\frac{s}{ab}\right)$, where g is the right derivative of G. So

$$G_1(x) = \int_0^{|x|} g_1(s)\, ds = \frac{1}{b} \int_0^{|x|} g\left(\frac{s}{ab}\right)\, ds = a \int_0^{\frac{|x|}{ab}} g(r) dr = aG\left(\frac{x}{ab}\right)$$

which is what we wanted. ■

Proof of theorem *2.6*. First suppose that $G \in \nabla_2$. Then F, the complement of G, satisfies the Δ_2 condition and thus there is a constant $k > 2$ so that $F(2x) \leq kF(x)$ for large values of x. So in virtue of the previous lemma and theorem 2.2, $kG(x/k) \leq G(x/2)$ or equivalently $kG(x) \leq G\left(\frac{kx}{2}\right)$ for large values of x. So the forward implication follows by setting $\beta = \frac{k}{2}$. The reverse implication is also a direct consequence of the lemma and theorem 2.2 and its proof is left as an exercise. ■

2.2.2 The Δ_3 condition

First note that if $F \in \Delta_3$ then F increases more rapidly than any power function. Indeed for any positive integer n and $x \geq k^n x_0$ we have:

$$F(x) > \frac{x}{k} F\left(\frac{x}{k}\right) > \frac{x^2}{k^3} F\left(\frac{x}{k^2}\right) > \cdots > \frac{x^n}{k^{\frac{n(n+1)}{2}}} F\left(\frac{x}{k^n}\right) > \frac{F(x_0)x^n}{k^{\frac{n(n+1)}{2}}}$$

Functions satisfying the Δ_3 condition are equivalent to their integrals. In particular we have that *if* $\Phi(x) = \int_0^x F(t)\,dt$ *and* $F \in \Delta_3$ *then* $\Phi \sim F$. In order to see this first observe that $\Phi(x) = \int_0^x F(t)\,dt < xF(x) \leq F(kx)$, for sufficiently large x. Furthermore for $x > 1$ we have

$$\Phi(2x) = \int_0^{2x} F(t)\,dt \geq \int_x^{2x} F(t)\,dt > xF(x) > F(x).$$

From this we obtain the following theorem:

Theorem 2.8 *If $F \in \Delta_3$ and G denotes the complement of F then there are constants $k_1 < k_2$ so that*

$$k_1 x F^{-1}(k_1 x) \leq G(x) \leq k_2 x F^{-1}(k_2 x)$$

for large values of x.

Proof. Let $\Phi(x) = \int_0^x F(t)\,dt$ and $\Psi(x) = \int_0^x F^{-1}(t)\,dt$. Then Φ and Ψ are complementary and as $\Phi \sim F$ we conclude $\Psi \sim G$. Now note that

$$\Psi(x) = \int_0^x F^{-1}(t)\,dt < xF^{-1}(x)$$

while

$$\Psi(x) = \int_0^x F^{-1}(t)\,dt > \int_{\frac{x}{2}}^x F^{-1}(t)\,dt > \frac{x}{2} F^{-1}\left(\frac{x}{2}\right).$$

Since $\Psi \sim G$ the result follows. ■

2.2.3 Some implications

There is a plethora of results pertaining to the different conditions on N–functions. Again the reader should consult [4] and [5] for a detailed account of the subject. Next, we summarize some of the most important relations between the different classes of N–functions:

Theorem 2.9 *Let F be an N-function and let G be its complement; then the following hold.*

1. *If $F \in \Delta'$ then $F \in \Delta_2$.*
2. *IF $F \in \Delta^2$ then $F \in \Delta_3$.*
3. *If $F \in \Delta_3$ then its complement $G \in \Delta_2$ (i.e. $F \in \nabla_2$).*
4. *If $F \in \Delta^2$ then its complement $G \in \Delta'$ (i.e. $F \in \nabla'$).*

Proof. (1) and (2) are obvious. For (3) let k_1 and k_2 as in theorem 2.8. Then since F^{-1} is concave and $\frac{2k_2}{k_1} > 1$ we have

$$F^{-1}\left(\frac{2k_2}{k_1}x\right) < \frac{2k_2}{k_1}F^{-1}(x).$$

Thus by theorem 2.8 we obtain

$$G(2x) \leq 2k_2xF^{-1}(2k_2x) < 2k_2x\frac{2k_2}{k_1}F^{-1}(k_1x) \leq \left(\frac{2k_2}{k_1}\right)^2 G(x)$$

for large values of x.

We continue into showing (4): So assume $F(kx) \geq F^2(x)$ for large values of x. Then for sufficiently large x and y with $x \geq y$ we have $F(kxy) > F(kx) \geq F^2(x) \geq F(x)F(y)$. By setting $x = F^{-1}(u)$ and $y = F^{-1}(v)$ we have $F(kF^{-1}(u)F^{-1}(v) > uv$ and thus

$$F^{-1}(uv) \leq kF^{-1}(u)F^{-1}(v) \text{ for large } u,\ v$$

Now since $F \in \Delta^2$ then $F \in \Delta_3$ and so by theorem 2.8 we have that $k_1xF^{-1}(k_1x) \leq G(x) \leq k_2xF^{-1}(k_2x)$ for large values of x. So for sufficiently large u and v we get

$$\begin{aligned}
G(uv) &\leq k_2uvF^{-1}(k_2uv)\\
&= (\sqrt{k_2}u)(\sqrt{k_2}v)F^{-1}(\sqrt{k_2}u\sqrt{k_2}v)\\
&\leq k\sqrt{k_2}uF^{-1}(\sqrt{k_2}u)\sqrt{k_2}vF^{-1}(\sqrt{k_2}v)\\
&\leq kG\left(\frac{\sqrt{k_2}}{k_1}u\right)G\left(\frac{\sqrt{k_2}}{k_1}v\right).
\end{aligned}$$

Since $G \in \Delta_2$ the result follows. ■

Last and not least we prove the following theorem:

Theorem 2.10 *Given an N-function F, there is an N-function $H \in \nabla^2$, and thus $H \in \Delta'$, so that $H(H(x)) \leq F(x)$ for large values of x.*

Proof. Write $F = F_1 \circ F_2$, where F_1, F_2 are N-functions and let G_i be the complement of F_i. Let $Q(x) = e^{G_1(x)+G_2(x)}$. The function Q is convex, with $\lim_{x\to\infty} \frac{Q(x)}{x} = \infty$. Hence there is an N-function K whose principal part is Q. Clearly $K \in \Delta^2$ and $G_i(x) \leq K(x)$ for large x. So if H is complementary to K, we must have $H \in \Delta'$ and $H(x) \leq F_i(x)$ for large x. Thus $H(H(x)) \leq F_1(F_2(x)) = F(x)$ for large values of x. ∎

2.2.4 Some examples

In the next example we construct a pair of complementary N–functions neither of which satisfies the Δ_2 condition, yet both grow slower than a power of x.

Example 2.1 *Let f be defined by*

$$f(t) = \begin{cases} t & \text{if} \quad 0 \leq t < 1 \\ k! & \text{if} \quad (k-1)! \leq t < k! \quad k = 2, 3, \ldots \end{cases}$$

Clearly $F(x) = \int_0^x f(t)\,dt$ is an N–function. Furthermore for each n let $x_n = n!$. Then

$$F(2x_n) = \int_0^{2n!} f(t)\,dt > \int_{n!}^{2n!} f(t)\,dt = (n+1)! \cdot n!$$

while

$$F(x_n) = \int_0^{n!} f(t)\,dt < n! \cdot n!$$

So $\frac{F(2x_n)}{F(x_n)} > \frac{(n+1)!\cdot n!}{n!\cdot n!} = n+1$ and thus $\limsup_{x\to\infty} \frac{F(2x)}{F(x)} = \infty$. Hence $F \notin \Delta_2$.
Now observe that if g is the right inverse of f then

$$g(t) = \begin{cases} t & \text{if} \quad 0 \leq t < 1 \\ (k-1)! & \text{if} \quad (k-1)! \leq t < k! \quad k = 2, 3, \ldots \end{cases}$$

So the complement G of F is given by $G(x) = \int_0^x g(t)\,dt$. Again let $x_n = n!$ and note that

$$G(2x_n) = \int_0^{2n!} g(t)\,dt > \int_{n!}^{2n!} g(t)\,dt = n! \cdot n!$$

while

$$G(x_n) = \int_0^{n!} g(t)\,dt < (n-1)! \cdot n!$$

So $\frac{G(2x_n)}{G(x_n)} > \frac{n!\cdot n!}{(n-1)!\cdot n!} = n$ *and thus* $\limsup_{x\to\infty} \frac{G(2x)}{G(x)} = \infty$. *Hence* $G \notin \Delta_2$.
Now it is plain to see that $f(t) < t^2$ *and* $g(t) < t$ *for large values of* t. *Hence* F *and* G *grow slower than* x^3 *and* x^2 *respectively.*

We have seen that every N–function F satisfying the Δ_3 condition grows faster than all power functions. The next example shows that the converse is not true even if F satisfies the ∇_2 condition.

Example 2.2 *Let* F *be an* N*–function whose principal part is given by* $x^{\sqrt{\log x}}$. *It is plain that* F *grows faster than any power function. Furthermore for large values of* x *we have that*

$$\frac{1}{4}F(2x) = \frac{1}{4}(2x)^{\sqrt{\log 2x}} \geq x^{\sqrt{\log 2x}} > x^{\sqrt{\log x}} = F(x)$$

and so $F \in \nabla_2$ *by theorem 2.6.*
Now notice that for any positive k *and* p *we have that* $k^{\log kx} = \left(k^{\frac{\log kx}{\log k}}\right)^{\log k} = (kx)^{\log k}$ *and so*

$$k^{\sqrt{\log kx}} = (kx)^{\frac{\log k}{\sqrt{\log kx}}} < x^p$$

for large values of x. *Similarly* $x^{\log kx} = x^{\log k} x^{\log x}$ *and thus*

$$x^{\sqrt{\log kx}} < x^p x^{\frac{\log x}{\sqrt{\log kx}}} < x^p x^{\sqrt{\log x}}$$

for large values of x.

$$\frac{F(kx)}{xF(x)} = \frac{(kx)^{\sqrt{\log kx}}}{x x^{\sqrt{\log x}}} < \frac{x^{2p}}{x}$$

for large values of x.
Thus $\lim_{x\to\infty} \frac{F(kx)}{xF(x)} = 0$ *for all positive constants* k. *Hence* $F \notin \Delta_3$.

Next we give an example of an N–function F in $\Delta_2 \setminus \Delta'$.

Example 2.3 *Let* $F(x) = \frac{x^2}{\log(|x|+e)}$. *It is a matter of calculus to show that* F *is an* N*–function. The reader can also verify that* $\lim_{x\to\infty} \frac{F(2x)}{F(x)} = 4$ *and* $\lim_{x\to\infty} \frac{F(x^2)}{F^2(x)} = \infty$. *Hence* $F \in \Delta_2$ *but* $F \notin \Delta'$.

Finally we give an example of an N–function F in $\Delta_3 \setminus \Delta^2$.

Example 2.4 *An N–function F whose principal part is $x^{\log x}$ is in Δ_3 but not in Δ^2. The details are left to the reader.*

Chapter 3

Orlicz Spaces

3.1 Orlicz classes

In this section we summarize the necessary definitions and results about Orlicz classes. Again, for a detailed account the reader should consult [4] or [5]. Throughout the remaining material we are going to assume that we are working with a non–atomic probability space (Ω, Σ, μ).

Definition 3.1 *For an N-function F and a measurable u define*

$$\mathbf{F}(u) = \int_\Omega F(u) d\mu.$$

Let $\tilde{L}^F = \{u \text{ measurable} : \mathbf{F}(u) < \infty\}$. The set $\tilde{L}^F$ is called an Orlicz class.

The theorem of De La Vallée Poussin establishes that every relatively weakly compact subset of L^1 is a bounded subset of some Orlicz class. It also establishes the fact that L^1 is the union of all Orlicz classes. But it does not specify just how well the function F can be chosen. We begin by noting the following improvement to De La Vallée Poussin's theorem (see [1]):

Theorem 3.1 *A subset $\mathcal{K}$ of $L^1(\mu)$ is uniformly integrable if and only if there exists an N–function $H \in \nabla^2$ so that $H(\mathcal{K}) = \{H(\mid u \mid) : \ u \in \mathcal{K}\}$ is uniformly integrable.*

Proof. By the theorem of De La Vallée Poussin there is a non-negative and convex function Q with $\lim_{t\to\infty} \frac{Q(t)}{t} = \infty$ so that $\sup \left\{ \int_\Omega Q(\mid u \mid)\, d\mu : u \in \mathcal{K} \right\} < \infty$. Since Q is the principal part of some N–function there is no loss in assuming that Q is actually an N–function. By theorem 2.10 there is an N–function $H \in \nabla^2$ so that $H(H(x)) \leq Q(x)$ for large values of x. Hence $\sup\left\{\int_\Omega H(H(|\,u\,|))\,d\mu : u \in \mathcal{K}\right\} < \infty$ and so $H(\mathcal{K}) = \{H(|\,u\,|) : u \in \mathcal{K}\}$ is uniformly integrable by De La Vallée Poussin's theorem. The converse is just De La Vallée Poussin's theorem again. ■

The famous *Jensen's inequality*, proved in almost every text in analysis or probability theory, is often a tool of great value:

Theorem 3.2 (Jensen's inequality) *For Φ convex and u measurable we have*

$$\Phi\left(\frac{1}{\mu(A)}\int_A u\,d\mu\right) \leq \frac{1}{\mu(A)}\int_A \Phi(u)\,d\mu$$

for all $A \in \Sigma$ with $\mu(A) > 0$.

Theorem 3.3 *Let F_1 and F_2 be N–functions. Then $\tilde{L}^{F_1} \subset \tilde{L}^{F_2}$ if and only if there positive constants c and x_0 so that $F_2(x) \leq cF_1(x)$ for all $x \geq x_0$.*

Proof. In order to see the sufficiency let $u \in \tilde{L}^{F_1}$. Then

$$\mathbf{F_2}(u) = \int_\Omega F_2(u)\,d\mu = \int_{[|u|<x_0]} F_2(u)d\mu + \int_{[|u|\geq x_0]} F_2(u)\,d\mu \leq F_2(x_0) + c\int_\Omega F_1(u)\,d\mu < \infty$$

and so $u \in \tilde{L}^{F_2}$.

In order to establish the converse, suppose that there is a sequence (x_n) with $x_n \nearrow \infty$ so that $F_2(x_n) > 2^n F_1(x_n)$. Since $(\Omega,\, \Sigma,\, \mu)$ is non–atomic, choose a sequence (E_n) of pairwise disjoint sets with $\mu(E_n) = \frac{F_1(x_1)}{2^n F_1(x_n)}$ and let $u = \sum_{n=1}^\infty x_n \chi_{E_n}$. Then $u \in \tilde{L}^{F_1}$ since

$$\int_\Omega F_1(u)\,d\mu = \sum_{n=1}^\infty F_1(x_n)\mu(E_n) = \sum_{n=1}^\infty F_1(x_n)\frac{F_1(x_1)}{2^n F_1(x_n)} = \sum_{n=1}^\infty \frac{F_1(x_1)}{2^n} < F_1(x_1) < \infty$$

But $u \notin \tilde{L}^{F_2}$ since

$$\int_\Omega F_2(u)\,d\mu = \sum_{n=1}^\infty F_2(x_n)\mu(E_n) \geq \sum_{n=1}^\infty 2^n F_1(x_n)\frac{F_1(x_1)}{2^n F_1(x_n)} = \sum_{n=1}^\infty F_1(x_1) = \infty$$

Hence the result is established. ■

It is plain that each Orlicz class $\tilde{L}^F$ is an absolutely convex set. In general, $\tilde{L}^F$ *is not* a linear space. Our next result establishes exactly when this is so.

Theorem 3.4 *$\tilde{L}^F$ is a linear space if and only if $F \in \Delta_2$.*

Proof. If $F \in \Delta_2$ then for each scalar c there is a constant k_c and $x_c > 0$ so that $F(cx) \leq k_c F(x)$ for all $x \geq x_c$. So for any $u \in \tilde{L}^F$ we have

$$\mathbf{F}(cu) = \int_\Omega F(cu)\, d\mu = \int_{[|u|<x_c]} F(cu)\, d\mu + \int_{[|u|\geq x_c]} F(cu)\, d\mu \leq F(cx_c) + k_c \int_\Omega F(u)\, d\mu < \infty$$

and so $cu \in \tilde{L}^F$. For closure under addition, let $u_1,\, u_2 \in \tilde{L}^F$, set $u = \frac{1}{2}(u_1 + u_2)$ and notice that

$$\begin{aligned}
\mathbf{F}(u_1 + u_2) &= \int_\Omega F(u_1 + u_2)\, d\mu \\
&= \int_\Omega F(2u)\, d\mu \\
&= \int_{[|u|<x_2]} F(2u)\, d\mu + \int_{[|u|\geq x_2]} F(2u)\, d\mu \\
&\leq F(2x_2) + k_2 \int_\Omega F(u)\, d\mu \\
&= F(2x_2) + k_2 \int_\Omega F(\frac{1}{2}(u_1 + u_2))\, d\mu \\
&\leq F(2x_2) + \frac{k_2}{2}\mathbf{F}(u_1) + \frac{k_2}{2}\mathbf{F}(u_2) < \infty
\end{aligned}$$

and so $u_1 + u_2 \in \tilde{L}^F$.

Now assume that $\tilde{L}^F$ is a linear space. Let $\Phi(x) = F(2x)$. Note that for $u \in \tilde{L}^F$ we have that $2u \in \tilde{L}^F$ and thus $u \in \tilde{L}^\Phi$. Hence, by theorem 3.3 we have that there is a constant $k > 0$ so that $\Phi(x) = F(2x) \leq kF(x)$ for large values of x. That is $F \in \Delta_2$. ∎

3.2 The Orlicz space L^F

3.2.1 The Orlicz norm

Given an N–function F and its complement G, let

$$L^F = \left\{ u \text{ measurable} : \int_\Omega uv\, d\mu < \infty \text{ for all } v \in \tilde{L}^G \right\}$$

It is plain that L^F is a vector space. Furthermore, Young's inequality ensures that $\tilde{L}^F \subset L^F$.

Theorem 3.5 *For each $u \in L^F$ we have*

$$\sup_{\mathbf{G}(v)\leq 1}\left|\int_\Omega uv\,d\mu\right| < \infty$$

Proof. Assume not. Then without loss of generality there is a non–negative $u \in L^F$ and a sequence of non–negative (v_n) in $\tilde{L}^G$ with $\mathbf{G}(v_n) \leq 1$ so that

$$\int_\Omega uv_n\,d\mu > 2^n$$

Let $v = \sum_{n=1}^{\infty} \frac{1}{2^n} v_n$. Then by the convexity of G we have that $\mathbf{G}(v) \leq 1$, yet $\int_\Omega uv\,d\mu = \infty$. ∎

Now define *the Orlicz norm* of $u \in L^F$ by

$$\|u\|_F = \sup_{\mathbf{G}(v)\leq 1}\left|\int_\Omega uv\,d\mu\right|$$

It is very easy to see that all the axioms for a norm are satisfied. Equipped with this norm L^F is called an *Orlicz space*. An obvious but important property of the norm is that is that $\|u_1\|_F \leq \|u_2\|_F$ whenever $u_1 \leq u_2$ a.s. We leave it as an exercise to the reader to verify that the Orlicz norm is complete, thus making an Orlicz space a Banach space.

Example 3.1 (The Orlicz norm of a characteristic function) *Notice that if $E \in \Sigma$ and $v \in \tilde{L}^G$ with $\mathbf{G}(v) \leq 1$ then by Jensen's inequality we have*

$$G\left(\frac{1}{\mu(E)}\int_\Omega \chi_E v\,d\mu\right) \leq \frac{1}{\mu(E)}\int_E G(v)\,d\mu \leq \frac{1}{\mu(E)}$$

and so

$$\|\chi_E\|_F = \sup_{\mathbf{G}(v)\leq 1}\left|\int_\Omega \chi_E v\,d\mu\right| \leq \mu(E)G^{-1}\left(\frac{1}{\mu(E)}\right).$$

On the other hand if $v_0 = G^{-1}\left(\frac{1}{\mu(E)}\right)\chi_E$ then $\mathbf{G}(v_0) = 1$ and $\int_\Omega \chi_E v_0\,d\mu = \mu(E)G^{-1}\left(\frac{1}{\mu(E)}\right)$. So

$$\|\chi_E\|_F = \mu(E)G^{-1}\left(\frac{1}{\mu(E)}\right).$$

Example 3.2 (The classical L^p spaces) *Let p, q be numbers greater than one and such that $\frac{1}{p} + \frac{1}{q} = 1$. If the N–function F is given by $F(x) = \frac{|x|^p}{p}$ then the complement G of F is* *given by* $G(x) = \frac{|x|^q}{q}$. *For* $u \in L^F$ *with* $\|f\|_p = 1$ *and* $v \in \tilde{L}^G$ *with* $\mathbf{G}(v) \leq 1$, *the classical Hölder's inequality yields*

$$\left| \int_\Omega uv \, d\mu \right| \leq \|u\|_p \cdot \|v\|_q \leq q^{\frac{1}{q}}.$$

On the other hand if $v_0 = q^{\frac{1}{q}} |u|^{p-1} \text{sgn}\, u$ *then* $\mathbf{G}(v_0) = 1$, $\int_\Omega uv_0 \, d\mu = q^{\frac{1}{q}}$ *and so*

$$\|u\|_F = q^{\frac{1}{q}}$$

So for any $u \in L^F$ *we have*

$$\|u\|_F = q^{\frac{1}{q}} \|u\|_p.$$

Our next goal is to derive a generalized Hölder's inequality. We will need several lemmas which will be useful in other occasions.

Lemma 3.6 *For every* $u \in L^F$ *and* $v \in \tilde{L}^G$ *with* $\mathbf{G}(v) > 1$ *we have*

$$\left| \int_\Omega uv \, d\mu \right| \leq \|u\|_F \cdot \mathbf{G}(v)$$

Proof. Notice that $G\left(\frac{v}{\mathbf{G}(v)}\right) \leq \frac{1}{\mathbf{G}(v)} G(v)$. Hence

$$\int_\Omega G\left(\frac{v}{\mathbf{G}(v)}\right) \, d\mu \leq \frac{1}{\mathbf{G}(v)} \int_\Omega G(v) \, d\mu = 1$$

Thus

$$\int_\Omega u \frac{v}{\mathbf{G}(v)} \, d\mu \leq \|f\|_F$$

and so the result follows. ∎

Lemma 3.7 *Let* F *be an* N*–function and let* f *be its right derivative. Let* $u \in L^F$ *with* $\|u\|_F \leq 1$. *Then* $v = f(|u|) \in \tilde{L}^G$ *and* $\mathbf{G}(v) \leq 1$.

Proof. Suppose that $\mathbf{G}(v) > 1$. Then there is a positive integer n so that if $u_n = u\chi_{[|u| \leq n]}$ we have that

$$\int_\Omega G(f(|u_n|)) \, d\mu > 1$$

Hence

$$G(f(|u_n|)) < F(u_n) + G(f(|u_n|)) = |u_n| \cdot f(|u_n|)$$

thanks to the equality case of Young's inequality. So by integrating the last inequality and using the previous lemma we obtain

$$\int_\Omega G(f(|u_n|))\,d\mu < \int_\Omega |u_n| \cdot f(|u_n|)\,d\mu \leq \|u_n\|_F \cdot \int_\Omega G(f(|u_n|))\,d\mu$$

and so $\|u\|_F \geq \|u_n\|_F > 1$ which is obviously a contradiction. ∎

Lemma 3.8 *Let $u \in L^F$ with $\|u\|_F \leq 1$. Then $u \in \tilde{L}^F$ and*

$$\mathbf{F}(u) \leq \|u\|_F$$

Consequently for every $u \in L^F$ we have

$$\mathbf{F}\left(\frac{u}{\|u\|_F}\right) \leq 1$$

Proof. Set $v = f(|u|)\text{sgn}\, u$. Then by the previous lemma we have $\mathbf{G}(v) \leq 1$. By the equality case of Young's inequality once again we have $uv = F(u) + G(v)$ and so

$$\int_\Omega F(u)\,d\mu \leq \int_\Omega F(u)\,d\mu + \int_\Omega G(v)\,d\mu = \int_\Omega uv\,d\mu \leq \|u\|_F$$

Now the second conclusion of the lemma is obvious. ∎

Theorem 3.9 (Hölder's inequality) *For every $u \in L^F$ and $v \in L^G$ we have*

$$\left|\int_\Omega uv\,d\mu\right| \leq \|u\|_F \cdot \|v\|_G$$

Proof. From the previous lemma we have that

$$\mathbf{F}\left(\frac{u}{\|u\|_F}\right) \leq 1$$

Hence

$$\left|\int_\Omega \frac{u}{\|u\|_F} v\,d\mu\right| \leq \|v\|_G.$$

∎

3.2.2 The Luxemburg norm

Definition 3.2 *In light of lemma 3.8, given an N–function F, the set $B_{(F)} = \{u \text{ measurable} : \mathbf{F}(u) \leq 1\}$ is an absolutely convex, balanced and absorbing set within its corresponding Orlicz space*

L^F. *The well known* Minkowski functional *defines a different norm* $\|\cdot\|_{(F)}$ *on* L^F *which is called the* Luxemburg norm. *Specifically for* $u \in L^F$ *set*

$$\|u\|_{(F)} = \inf\left\{k > 0 \,:\, \mathbf{F}\left(\frac{u}{k}\right) \leq 1\right\}.$$

It is plain from lemma 3.8 that $\|u\|_{(F)} \leq \|u\|_F$. Furthermore the infimum is attained whenever $u \in L^F$ is not 0 a.s. and we have

$$\mathbf{F}\left(\frac{u}{\|u\|_{(F)}}\right) \leq 1.$$

The reader can verify these facts through the use of Fatou's lemma or the monotone convergence theorem.

Example 3.3 (The Luxemburg norm of a characteristic function) *Let* $E \in \Sigma$ *with* $\mu(E) > 0$. *Then* $\mathbf{F}\left(F^{-1}\left(\frac{1}{\mu(E)}\right)\chi_E\right) = 1$ *and so*

$$\|\chi_E\|_{(F)} = \frac{1}{F^{-1}\left(\frac{1}{\mu(E)}\right)}$$

Theorem 3.10 *The unit–ball of the Orlicz space* L^F *endowed with the Luxemburg norm is* $B_{(F)} = \{u \text{ measurable} : \mathbf{F}(u) \leq 1\}$. *Furthermore for any* $u \in \tilde{L}^F$ *we have that*

$$\begin{cases} \mathbf{F}(u) \leq \|u\|_{(F)} & \text{whenever} \quad \|u\|_{(F)} \leq 1 \\ \mathbf{F}(u) \geq \|u\|_{(F)} & \text{whenever} \quad \|u\|_{(F)} > 1 \end{cases}$$

Proof. If $\|u\|_{(F)} \leq 1$ then

$$\frac{1}{\|u\|_{(F)}}\int_\Omega F(u)\,d\mu \leq \int_\Omega F\left(\frac{u}{\|u\|_{(F)}}\right)\,d\mu \leq 1$$

and so $\mathbf{F}(u) \leq \|u\|_{(F)}$. On the other hand, if $\|u\|_{(F)} > 1$ then for ε small enough we have

$$\frac{1}{\|u\|_{(F)} - \varepsilon}\int_\Omega F(u)\,d\mu \geq \int_\Omega F\left(\frac{u}{\|u\|_{(F)} - \varepsilon}\right)\,d\mu > 1$$

hence $\mathbf{F}(u) \geq \|u\|_{(F)}$. ■

We have noted that the Orlicz norm is larger than the Luxemburg norm. In fact these norms are equivalent as established by the following theorem.

Theorem 3.11 *For any* $u \in L^F$ *we have* $\|u\|_{(F)} \leq \|u\|_F \leq 2\|u\|_{(F)}$.

Proof. The first inequality is already established. In order to see the second inequality, note that from Young's inequality we have that

$$\|u\|_F = \sup_{\mathbf{G}(v)\le 1} \int_\Omega uv\, d\mu \le \mathbf{F}(u) + 1$$

whenever $u \in \tilde{L}^F$. Thus

$$\left\| \frac{u}{\|u\|_{(F)}} \right\|_F \le \int_\Omega F\left(\frac{u}{\|u\|_{(F)}}\right) d\mu + 1 \le 2$$

and so the second inequality is done. ■

At this stage notice that theorem 3.10 gives an alternative formula for the Orlicz norm:

$$\sup_{\|v\|_{(G)}\le 1} \left| \int_\Omega uv\, d\mu \right|$$

from which we can easily obtain the following stronger versions of Hölder's inequality:

Theorem 3.12 (Hölder's inequality) *For every $u \in L^F$ and $v \in L^G$ we have*

$$\left| \int_\Omega uv\, d\mu \right| \le \|u\|_{(F)} \cdot \|v\|_G$$

and For every $u \in L^F$ and $v \in L^G$ we have

$$\left| \int_\Omega uv\, d\mu \right| \le \|u\|_F \cdot \|v\|_{(G)}$$

3.2.3 Mean and norm convergence

Definition 3.3 *We say that a sequence of functions (u_n) in L^F converges in mean to a function $u \in L^F$ iff $\mathbf{F}(u_n - u) \to 0$ as $n \to \infty$.*

In view of lemma 3.8, it is plain that norm convergence implies mean convergence. If $F \in \Delta_2$ then we have the following theorem:

Theorem 3.13 *If $\mathbf{F}(u_n - u) \to 0$ as $n \to \infty$ for $F \in \Delta_2$ then $\|u_n - u\|_F \to 0$ as $n \to \infty$.*

Proof. Fix $\varepsilon > 0$ and choose a positive integer k so that $\frac{1}{2^{k-1}} < \varepsilon$. Since $F \in \Delta_2$ and $\lim_{n\to\infty} \mathbf{F}(u_n - u) = 0$, we have that $\lim_{n\to\infty} \mathbf{F}(2^k (u_n - u)) = 0$ and thus there is a positive integer N so that $\mathbf{F}(2^k(u_n-u))<1$ whenever $n\geq N$. Thus for $n\geq N$

$$\begin{aligned}\left\|2^k(u_n-u)\right\|_F &= \sup_{\mathbf{G}(v)\leq 1}\left|\int_\Omega 2^k(u_n-u)\,v\,d\mu\right| \\ &\leq \sup_{\mathbf{G}(v)\leq 1}\int_\Omega \left|2^k(u_n-u)\right||v|\;d\mu \\ &\leq \sup_{\mathbf{G}(v)\leq 1}\left[\mathbf{F}(2^k(u_n-u))+\mathbf{G}(v)\right] \text{ (by Young's inequality)} \\ &< 2\end{aligned}$$

Hence $\|u_n-u\|_F<\frac{1}{2^{k-1}}<\varepsilon$ and the result is established. ■

Remark 3.1 *If $F\notin\Delta_2$, choose an increasing sequence of numbers (a_n) with $F(2a_n)>2^nF(a_n)$ and $F(a_1)>1$.Since $(\Omega,\ \Sigma,\ \mu)$ is non-atomic, for each n choose pairwise disjoint sets $A_1^n,\ldots,A_n^n\in\Sigma$ with $\mu(A_k^n)=\frac{1}{n2^kF(a_k)}$ and let $u_n=\sum_{k=1}^n a_k\chi_{A_k^n}$. Then*

$$\mathbf{F}(u_n)=\sum_{k=1}^n F(a_k)\,\mu(A_k^n)<\frac{1}{n}$$

Hence (u_n) converges to zero in mean. On the other hand, (u_n) does not converge to zero in norm: For otherwise $\|u_n\|_F\to 0$ implies $\|2u_n\|_F\to 0$ and so $\mathbf{F}(2u_n)\to 0$. But

$$\mathbf{F}(2u_n)=\sum_{k=1}^n F(2a_k)\,\mu(A_k^n)>\sum_{k=1}^n 2^kF(a_k)\,\frac{1}{n2^kF(a_k)}=1$$

which is obviously a contradiction.

3.3 The closure of $L^{\infty}(\mu)$ in $L^{F}(\mu)$

Notation 3.1 *Let E^F denote the closure of $L^{\infty}(\mu)$ in the Orlicz space $L^F(\mu)$.*

Note that if $u\in\tilde{L}^F$ then the sequence of bounded functions $\left(u\chi_{[|u|\leq n]}\right)$ converges to u in mean. Indeed

$$\begin{aligned}\mathbf{F}(u\chi_{[|u|\leq n]}-u) &= \int_\Omega F\left(u\chi_{[|u|\leq n]}-u\right)d\mu \\ &= \int_{[|u|\leq n]} F\left(u\chi_{[|u|\leq n]}-u\right)d\mu+\int_{[|u|>n]} F(u)\,d\mu \\ &= \int_{[|u|>n]} F(u)\,d\mu\to 0 \text{ as } n\to\infty\end{aligned}$$

In light of theorems 3.4 and 3.13 we have the following

Proposition 3.14 *If $F \in \Delta_2$ then $E^F = L^F$.*

In general we have that $E^F \subseteq \tilde{L}^F$ for if $u \in E^F$ choose a bounded function r with $\|u - r\|_F < \frac{1}{2}$. Then $\|2u - 2r\|_F < 1$ and so by lemma 3.8 we have $2u - 2r \in \tilde{L}^F$ and $\mathbf{F}(2u - 2r) \leq 2\|u - r\|_F < 1$. As $2r \in \tilde{L}^F$, convexity ensures that $u = \frac{1}{2}[(2u - 2r) + 2r] \in \tilde{L}^F$.

Theorem 3.15 *The space $E^F[0,1]$ is separable.*

Proof. First note that if u is a bounded function with $\|u\|_\infty = M$ then by Lousin's theorem, there is a sequence of continuous functions (u_n) with $\|u_n\|_\infty \leq M$ and a sequence of sets (A_n) with $\mu(A_n) < \frac{1}{n}$ so that $u - u_n = 0$ on $[0,1] \setminus A_n$. Thus

$$\begin{aligned} \|u_n - u\|_F &= \sup_{\mathbf{G}(v) \leq 1} \int_0^1 |u_n - u|\,|v|\; d\mu = \sup_{\mathbf{G}(v) \leq 1} \int_{A_n} |u_n - u|\,|v|\; d\mu \\ &\leq 2M \sup_{\mathbf{G}(v) \leq 1} \int_{A_n} |v|\; d\mu = 2M\,\|\chi_{A_n}\|_F \\ &= 2M\mu(A_n)\,G^{-1}\left(\frac{1}{\mu(A_n)}\right) \to 0 \text{ as } n \to \infty\ . \end{aligned}$$

Now separability follows from the fact that every continuous function is the uniform limit of polynomials with rational coefficients. ∎

Definition 3.4 *We say that a function $u \in L^F$ has absolutely continuous norm if and only if $\forall \varepsilon > 0$ there is $\delta > 0$ such that $\|u\chi_A\|_F < \varepsilon$ for any measurable set A with $\mu(A) < \delta$.*

Theorem 3.16 *A function $u \in L^F$ has absolutely continuous norm if and only if $u \in E^F$.*

Proof. Suppose that $u \in L^F$ has absolutely continuous norm. Since $u \in L^1$ then $\mu([|u| > n]) \to 0$ as $n \to \infty$ and thus $\left\|u - u\chi_{[|u| \leq n]}\right\|_F = \left\|u\chi_{[|u| > n]}\right\|_F \to 0$ as $n \to \infty$. As each of the functions $u\chi_{[|u| \leq n]}$ is bounded, the forward direction is established.
Now suppose that $u \in E^F$. Fix $\varepsilon > 0$ and choose a bounded function r with $\|u - r\|_F < \frac{\varepsilon}{2}$. Let $M = \|r\|_\infty$ and choose $\delta > 0$ such that $xG^{-1}\left(\frac{1}{x}\right) < \frac{\varepsilon}{2M}$, whenever $0 < |x| < \delta$ (after all $\lim_{x \to 0} xG^{-1}\left(\frac{1}{x}\right) = 0$). Now for any measurable set A with $\mu(A) < \delta$ we have $\|\chi_A\|_F = \mu(A)\,G^{-1}\left(\frac{1}{\mu(A)}\right) < \frac{\varepsilon}{2M}$ thanks to example 3.1. Thus

$$\|u\chi_A\|_F \leq \|(u - r)\,\chi_A\|_F + \|r\chi_A\|_F \leq \|u - r\|_F + M\,\|\chi_A\|_F < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

and so we are done. ■

We now turn our attention to N-functions that do not satisfy the Δ_2 condition. In the following important example we follow the lead of B. Turett in [6]:

Example 3.4 *Suppose that $F \notin \Delta_2$. Then, choose an increasing sequence of numbers (a_k) satisfying*

- $\lim_{k\to\infty} a_k = \infty$
- $F(2a_k) > 2^{k+1} F(a_k)$

Since $(\Omega,\ \Sigma,\ \mu)$ is non-atomic, choose a sequence of pairwise disjoint measurable sets (A_k) with $\mu(A_k) = \frac{1}{2^k F(a_k)}$. Let $u_k = a_k \chi_{A_k}$ and observe that

1. $\mathbf{F}(u_k) = \int_\Omega F(u_k)\, d\mu = \frac{1}{2^k} \leq 1$ *while* $\mathbf{F}(2u_k) = \int_{A_k} F(2a_k)\, d\mu \geq 2^{k+1} F(a_k)\, \mu(A_k) = 2$. *Thus* $\frac{1}{2} < \|u_k\|_{(F)} \leq 1$.
2. $\mathbf{F}\left(\sum_{k=1}^{\infty} u_k\right) = 1$ *and thus* $\left\|\sum_{k=1}^{\infty} u_k\right\|_{(F)} = 1$. *Thus it follows from (1), that* $\sum_{k=1}^{\infty} u_k$ *does not have absolutely continuous norm and so* $\sum_{k=1}^{\infty} u_k \notin E^F$.
3. *The function* $2\sum_{k=1}^{\infty} u_k \notin \tilde{L}^F$
4. *The map* $T : \ell^\infty \to L^F$ *defined by* $T(x_n) = \sum_{k=1}^{\infty} x_k u_k$ *is an isomorphism:*

$$\|T(x_n)\|_{(F)} = \left\|\sum_{k=1}^{\infty} x_k u_k\right\|_{(F)} \leq \left\|\sum_{k=1}^{\infty} \|(x_n)\|_\infty u_k\right\|_{(F)} = \|(x_n)\|_\infty \cdot \left\|\sum_{k=1}^{\infty} u_k\right\|_{(F)} = \|(x_n)\|_\infty$$

On the other hand if $|x_{k_0}| > \frac{1}{2}\|(x_n)\|_\infty$ we have

$$\|T(x_n)\|_{(F)} = \left\|\sum_{k=1}^{\infty} x_k u_k\right\|_{(F)} \geq \|x_{k_0} u_k\|_{(F)} \geq \frac{1}{2}\|(x_n)\|_\infty \|u_k\|_{(F)} \geq \frac{1}{4}\|(x_n)\|_\infty$$

In fact with a bit more work we can get an isometric copy of ℓ^∞ inside L^F. Here is how:

Split Ω into pairwise disjoint sets (A_n) with $\mu(A_n) = \frac{1}{2^n}$. Again since $F \notin \Delta_2$, choose an increasing sequence of numbers (a_k) satisfying $\lim_{k\to\infty} a_k = \infty$ and $F\left(\frac{k+1}{k} a_k\right) > 2^{k+1} F(a_k)$. Now inside each A_n choose a sequence of pairwise disjoint sets $\left(A_n^k\right)$ with $\mu\left(A_n^k\right) = \frac{1}{2^n} \cdot \frac{1}{2^k F(a_k)}$ and let $u_n = \sum_{k=1}^{\infty} a_k \chi_{A_k^n}$. Notice that

(a) *For each* n, $\mathbf{F}(u_n) = \int_\Omega F(u_n)\,d\mu = \frac{1}{2^n}$. *Yet if* $c > 1$ *we have that* $F(ca_k) \geq F\left(\frac{k+1}{k}a_k\right) > 2^{k+1}F(a_k)$ *for large enough* k *and so* $\mathbf{F}(cu_n) = \infty$. *Thus* $\|u_n\|_{(F)} = 1$

(b) $\mathbf{F}\left(\sum_{n=1}^\infty u_n\right) = 1$ *and thus* $\left\|\sum_{n=1}^\infty u_n\right\|_{(F)} = 1$.

(c) *The map* $T : \ell^\infty \to L^F$ *defined by* $T(x_n) = \sum_{n=1}^\infty x_n u_n$ *is an isometry:*

$$\|T(x_n)\|_{(F)} = \left\|\sum_{k=1}^\infty x_k u_k\right\|_{(F)} \leq \left\|\sum_{k=1}^\infty \|(x_n)\|_\infty u_k\right\|_{(F)} = \|(x_n)\|_\infty \cdot \left\|\sum_{k=1}^\infty u_k\right\|_{(F)} = \|(x_n)\|_\infty$$

on the other hand

$$\|T(x_n)\|_{(F)} = \left\|\sum_{k=1}^\infty x_k u_k\right\|_{(F)} \geq \sup_k \|x_k u_k\|_{(F)} = \sup_k |x_k| \cdot \|u_k\|_{(F)} = \|(x_n)\|_\infty$$

Thus we have

Proposition 3.17 *The following are equivalent for an* N*-function* F*:*

1. $L^F[0,1]$ *is separable*
2. $L^F = E^F$
3. $\tilde{L}^F = L^F$
4. *Every function* $u \in L^F$ *has absolutely continuous norm*
5. $F \in \Delta_2$

3.4 Orlicz spaces are dual spaces

Theorem 3.18 *If* F *and* G *are complementary* N*-functions then* $\left(E^F, \|\cdot\|_{(F)}\right)^* = \left(L^G, \|\cdot\|_G\right)$ *and* $\left(E^F, \|\cdot\|_F\right)^* = \left(L^G, \|\cdot\|_{(G)}\right)$.

Proof. First notice that for each $v \in L^G$ the map $\rho : u \longmapsto \int_\Omega uv d\mu$ defines a bounded linear functional in $\left(E^F, \|\cdot\|_{(F)}\right)^*$. Furthermore by Hölder's inequality, $\|\rho\| \leq \|v\|_G$. Now since $\|v\|_G = \sup_{\|u\|_{(F)} \leq 1} \left|\int_\Omega uv\,d\mu\right|$, for each $\varepsilon > 0$ choose u_0 with $\|u_0\|_{(F)} \leq 1$ and $\|v\|_G < \left|\int_\Omega u_0 v\,d\mu\right| + \frac{\varepsilon}{2}$. Then for sufficiently large n we have that $\int_\Omega |u_0 v|\ d\mu \leq \int_{[|u_0| \leq n]} |u_0 v|\ d\mu + \frac{\varepsilon}{2} =$

$\int_\Omega \left|\chi_{[|u_0|\leq n]} u_0 v\right| \, d\mu + \frac{\varepsilon}{2}$. Hence by setting $u_1 = \operatorname{sign}(v)\, \chi_{[|u_0|\leq n]} |u_0|$ we have that $u_1 \in E^F$ and $\|u_1\|_{(F)} \leq \|u_0\|_{(F)} \leq 1$ and so

$$\|v\|_G < \left|\int_\Omega u_0 v \, d\mu\right| + \frac{\varepsilon}{2} \leq \int_\Omega |u_0 v| \, d\mu + \frac{\varepsilon}{2} \leq \int_\Omega \left|\chi_{[|u_0|\leq n]} u_0 v\right| \, d\mu + \varepsilon = \rho(u_1) + \epsilon \leq \|\rho\| + \varepsilon$$

Therefore $\|v\|_G = \|\rho\|$.

Now if $\rho \in \left(E^F, \|\cdot\|_{(F)}\right)^*$ define $\lambda(E) = \rho(\chi_E)$ for all $E \in \Sigma$. Then

$$|\lambda(E)| = |\rho(\chi_E)| \leq \|\rho\| \, \|\chi_E\|_{(F)} = \frac{\|\rho\|}{F^{-1}\left(\frac{1}{\mu(E)}\right)} \to 0 \text{ uniformly as } \mu(E) \to 0$$

Thus λ is absolutely continuous with respect to μ and so by the Radon-Nicodym theorem, there is $v \in L^1(\mu)$ so that $\rho(\chi_E) = \lambda(E) = \int_\Omega \chi_E v d\mu$ for all $E \in \Sigma$. Since the simple functions are dense in E^F we have that $\rho(u) = \int_\Omega uv d\mu$ for all $u \in E^F$. Furthermore, for all $u \in \tilde{L}^F$ we have that

$$\begin{aligned}\left|\int_\Omega uv d\mu\right| &\leq \int_\Omega |uv| \, d\mu \leq \lim\inf_{n\to\infty} \int_{[|u|\leq n]} |uv| \, d\mu \leq \|\rho\| \sup_n \left\|\chi_{[|u|\leq n]} u\right\|_{(F)} \\ &\leq \|\rho\| \, \|u\|_{(F)} < \infty \text{ thanks to Fatou's lemma}\end{aligned}$$

and so $v \in L^G$.

A similar argument establishes the fact that $\left(E^F, \|\cdot\|_F\right)^* = \left(L^G, \|\cdot\|_{(G)}\right)$. ■

Bibliography

[1] J. Alexopoulos, *De La Vallée Poussin's theorem and weakly compact sets in Orlicz spaces*, QM (1994), no. 17, 231–248.

[2] J. Diestel, *Sequences and series in Banach spaces*, Springer Verlag, 1984.

[3] N. Dunford and J. Schwartz, *Linear operators part I: General theory*, Wiley Classics library edition, Wiley Interscience, 1988.

[4] M. A. Krasnoselskii and Ya. B. Rutickii, *Convex functions and Orlicz spaces*, Noorhoff Ltd., Groningen, 1961.

[5] M. M. Rao and Z. D. Ren, *Theory of Orlicz spaces*, Marcel Dekker, 1991.

[6] B. Turett, *Rotundity in Orlicz spaces*, Indag. Math., 38 (1976), no. 38, 462–469.