Chapter 2
A Framework for Function Spaces

In this chapter we study *modular spaces* and *Musielak–Orlicz spaces* which provide the framework for a variety of different function spaces, including classical (weighted) Lebesgue and Orlicz spaces and variable exponent Lebesgue spaces. Although our aim mainly is to study the latter, it is important to see the connections between all of these spaces. Many of the results in this chapter can be found in a similar form in [307], but we include them to make this exposition self-contained. Research in the field of Musielak–Orlicz functions is still active and we refer to [69] for newer results and references.

Our first two sections deal with the more general case of semimodular spaces. Then we move to basic properties of Musielak–Orlicz spaces in Sect. 2.3. Sections 2.4 and 2.5 deal with the uniform convexity and the separability of the Musielak–Orlicz spaces. In Sects. 2.6 and 2.7 we study dual spaces, and a related concept, associate spaces. Finally, we consider embeddings in Sect. 2.8.

2.1 Basic Properties of Semimodular Spaces

For the investigation of weighted Lebesgue spaces it is enough to stay in the framework of Banach spaces. In particular, the space and its topology is described in terms of a norm. However, in the context of Orlicz spaces this is not the best way. Instead, it is better to start with the so-called modular which then induces a norm. In the case of classical Lebesgue spaces the modular is $\int |f(x)|^p\,dx$ compared to the norm $\left(\int |f(x)|^p\,dx\right)^{\frac{1}{p}}$. In some cases the modular has certain advantages compared to the norm, since it inherits all the good properties of the integral. The *modular spaces* defined below capture this advantage.

We are mainly interested in vector spaces defined over $\mathbb{R}$. However, there is no big difference in the definition of real valued and complex valued modular spaces. To avoid a double definition we let $\mathbb{K}$ be either $\mathbb{R}$ or $\mathbb{C}$.

The function ϱ is said to be *left-continuous* if the mapping $\lambda \mapsto \varrho(\lambda x)$ is left-continuous on $[0,\infty)$ for every $x \in X$, i.e. $\lim_{\lambda\to 1^-} \varrho(\lambda x) = \varrho(x)$. Here

L. Diening et al., *Lebesgue and Sobolev Spaces with Variable Exponents*,
Lecture Notes in Mathematics 2017, DOI 10.1007/978-3-642-18363-8_2,

$a \to b^-$ means that a tends to b from below, i.e. $a < b$ and $a \to b$; $a \to b^+$ is defined analogously.

Definition 2.1.1. Let X be a $\mathbb{K}$-vector space. A function $\varrho\colon X \to [0, \infty]$ is called a *semimodular* on X if the following properties hold.

(a) $\varrho(0) = 0$.
(b) $\varrho(\lambda x) = \varrho(x)$ for all $x \in X, \lambda \in \mathbb{K}$ with $|\lambda| = 1$.
(c) ϱ is convex.
(d) ϱ is left-continuous.
(e) $\varrho(\lambda x) = 0$ for all $\lambda > 0$ implies $x = 0$.

A semimodular ϱ is called a *modular* if

(f) $\varrho(x) = 0$ implies $x = 0$.

A semimodular ϱ is called *continuous* if

(g) the mapping $\lambda \mapsto \varrho(\lambda x)$ is continuous on $[0, \infty)$ for every $x \in X$.

Remark 2.1.2. Note that our semimodulars are always convex, in contrast to some other sources.

Before we proceed let us provide a few examples.

Definition 2.1.3. Let (A, Σ, μ) be a σ-finite, complete measure space. Then by $L^0(A, \mu)$ we denote the space of all $\mathbb{K}$-valued, μ-measurable functions on A. Two functions are identical, if they agree almost everywhere.

In the special case that μ is the n-dimensional Lebesgue measure, Ω is a μ-measurable subset of $\mathbb{R}^n$, and Σ is the σ-algebra of μ-measurable subsets of Ω we abbreviate $L^0(\Omega) := L^0(\Omega, \mu)$.

Example 2.1.4.

(a) If $1 \leqslant p < \infty$, then

$$\varrho_p(f) := \int_\Omega |f(x)|^p \, dx$$

defines a continuous modular on $L^0(\Omega)$.

(b) Let $\varphi_\infty(t) := \infty \cdot \chi_{(1,\infty)}(t)$ for $t \geqslant 0$, i.e. $\varphi_\infty(t) = 0$ for $t \in [0, 1]$ and $\varphi_\infty(t) = \infty$ for $t \in (1, \infty)$. Then

$$\varrho_\infty(f) := \int_\Omega \varphi_\infty(|f(x)|) \, dx$$

defines a semimodular on $L^0(\Omega)$ which is not continuous.

(c) Let $\omega \in L^1_{\mathrm{loc}}(\Omega)$ with $\omega > 0$ almost everywhere and $1 \leqslant p < \infty$. Then

$$\varrho(f) := \int_\Omega |f(x)|^p \omega(x)\, dx$$

defines a continuous modular on $L^0(\Omega)$.

(d) The integral expression

$$\varrho(f) := \int_\Omega \exp(|f(x)|) - 1\, dx$$

defines a modular on $L^0(\Omega)$. It is not continuous: if $f \in L^2(\Omega)$ is such that $|f| > 2$ and $|f| \notin L^p(\Omega)$ for any $p > 2$, then $\varrho(\lambda \log |f|) = \infty$ for $\lambda > 2$ but $\varrho(2 \log |f|) < \infty$.

(e) If $1 \leqslant p < \infty$, then

$$\varrho_p\big((x_j)\big) := \sum_{j=0}^{\infty} |x_j|^p\, dx$$

defines a continuous modular on $\mathbb{R}^{\mathbb{N}}$.

(f) For $f \in L^0(\Omega)$ we define the decreasing rearrangement, $f^*\colon [0,\infty) \to [0,\infty)$ by the formula $f^*(s) := \sup\{t\colon |\{|f| > t\}| > s\}$. For $1 \leqslant q \leqslant p < \infty$ the expression

$$\varrho(f) := \int_0^\infty \left|f^*(s^{p/q})\right|^q ds$$

defines a continuous modular on $L^0(\Omega)$.

Let ϱ be a semimodular on X. Then by convexity and non-negative of ϱ and $\varrho(0) = 0$ it follows that $\lambda \mapsto \varrho(\lambda x)$ is non-decreasing on $[0,\infty)$ for every $x \in X$. Moreover,

$$\begin{aligned} \varrho(\lambda x) &= \varrho(|\lambda|\, x) \leqslant |\lambda|\, \varrho(x) &\quad& \text{for all } |\lambda| \leqslant 1, \\ \varrho(\lambda x) &= \varrho(|\lambda|\, x) \geqslant |\lambda|\, \varrho(x) && \text{for all } |\lambda| \geqslant 1. \end{aligned} \tag{2.1.5}$$

In the definition of a semimodular or modular the set X is usually chosen to be larger than necessary. The idea behind this is to choose the same large set X for different modulars like in our Examples 2.1.4(a), (b), (c), (d) and (f). Then depending on the modular we pick interesting subsets from this set X.

Definition 2.1.6. If ϱ be a semimodular or modular on X, then

$$X_\varrho := \Big\{x \in X\colon \lim_{\lambda \to 0} \varrho(\lambda x) = 0\Big\}$$

is called a *semimodular space* or *modular space*, respectively. The limit $\lambda \to 0$ takes place in $\mathbb{K}$.

Since $\varrho(\lambda x) = \varrho(|\lambda|\, x)$ it is enough to require $\lim_{\lambda\to 0} \varrho(\lambda x)$ with $\lambda \in (0,\infty)$. Due to (2.1.5) we can alternatively define X_ϱ by

$$X_\varrho := \big\{x \in X \colon \varrho(\lambda x) < \infty \text{ for some } \lambda > 0\big\},$$

since for $\lambda' < \lambda$ we have by (2.1.5) that

$$\varrho(\lambda' x) = \varrho\Big(\frac{\lambda'}{\lambda}\lambda x\Big) \leqslant \frac{\lambda'}{\lambda}\varrho(\lambda x) \to 0$$

as $\lambda' \to 0$.

In the next theorem, like elsewhere, the infimum of the empty set is by definition infinity.

Theorem 2.1.7. *Let ϱ be a semimodular on X. Then X_ϱ is a normed $\mathbb{K}$-vector space. The norm, called the* Luxemburg norm, *is defined by*

$$\|x\|_\varrho := \inf\Big\{\lambda > 0 \colon \varrho\Big(\frac{1}{\lambda}x\Big) \leqslant 1\Big\}.$$

Proof. We begin with the vector space property of X_ϱ. Let $x, y \in X_\varrho$ and $\alpha \in \mathbb{K} \setminus \{0\}$. From the definition of X_ϱ and $\varrho(\alpha x) = \varrho(|\alpha|\, x)$ it is clear that $\alpha x \in X_\varrho$. By the convexity of ϱ we estimate

$$0 \leqslant \varrho\big(\lambda(x+y)\big) \leqslant \tfrac12\varrho(2\lambda x) + \tfrac12\varrho(2\lambda y) \xrightarrow{\lambda\to 0} 0.$$

Hence, $x + y \in X_\varrho$. It is clear that $0 \in X_\varrho$. This proves that X_ϱ is a $\mathbb{K}$-vector space.

It is clear that $\|x\|_\varrho < \infty$ for all $x \in X_\varrho$ and $\|0\|_\varrho = 0$. For $\alpha \in \mathbb{K}$ we have

$$\begin{aligned}\|\alpha x\|_\varrho &= \inf\Big\{\lambda > 0 \colon \varrho\Big(\frac{\alpha x}{\lambda}\Big) \leqslant 1\Big\} = |\alpha| \inf\Big\{\lambda > 0 \colon \varrho\Big(\frac{1}{\lambda}x\Big) \leqslant 1\Big\}\\ &= |\alpha|\, \|x\|_\varrho.\end{aligned}$$

Let $x, y \in X$ and $u > \|x\|_\varrho$ and $v > \|y\|_\varrho$. Then $\varrho(x/u) \leqslant 1$ and $\varrho(y/v) \leqslant 1$, hence, by the convexity of ϱ,

$$\varrho\Big(\frac{x+y}{u+v}\Big) = \varrho\Big(\frac{u}{u+v}\frac{x}{u} + \frac{v}{u+v}\frac{y}{v}\Big) \leqslant \frac{u}{u+v}\varrho\Big(\frac{x}{u}\Big) + \frac{v}{u+v}\varrho\Big(\frac{y}{v}\Big) \leqslant 1.$$

Thus $\|x+y\|_\varrho \leqslant u + v$, and we obtain $\|x+y\|_\varrho \leqslant \|x\|_\varrho + \|y\|_\varrho$.

If $\|x\|_\varrho = 0$, then $\varrho(\alpha x) \leqslant 1$ for all $\alpha > 0$. Therefore,

$$\varrho(\lambda x) \leqslant \beta\varrho\Big(\frac{\lambda x}{\beta}\Big) \leqslant \beta$$

for all $\lambda > 0$ and $\beta \in (0,1]$, where we have used (2.1.5). This implies $\varrho(\lambda x) = 0$ for all $\lambda > 0$. Thus $x = 0$. $\square$

The norm in the previous theorem is more generally known as the *Minkowski functional* of the set $\{x \in X : \varrho(x) \leqslant 1\}$, see Remark 2.1.16. The Minkowski functional was first introduced by Kolmogorov in [253] long before the appearance of the Luxemburg norm. Nevertheless, we use the name "Luxemburg norm" as it is customary in the theory of Orlicz spaces.

In the following example we use the notation of Example 2.1.4.

Example 2.1.8 (Classical Lebesgue spaces). Let $1 \leqslant p < \infty$. Then the corresponding modular space $(L^0(\Omega))_{\varrho_p}$ coincides with the classical Lebesgue space L^p, i.e.

$$\|f\|_p := \|f\|_{\varrho_p} = \Big(\int\limits_{\Omega} |f(x)|^p \, dx\Big)^{\frac{1}{p}}.$$

Similarly, the corresponding semimodular space $(L^0(\Omega))_{\varrho_\infty}$ coincides with the classical Lebesgue space L^∞, i.e.

$$\|f\|_\infty := \|f\|_{\varrho_\infty} = \operatorname*{ess\,sup}_{x\in\Omega} |f(x)|.$$

The norm $\|\cdot\|_\varrho$ defines our standard topology on X_ϱ. So for $x_k, x \in X_\varrho$ we say that x_k *converges strongly or in norm to* x if $\|x_k - x\|_\varrho \to 0$. In this case we write $x_k \to x$. The next lemma characterizes this topology in terms of the semimodular. Here it suffices to study null-sequences.

Lemma 2.1.9. *Let ϱ be a semimodular on X and $x_k \in X_\varrho$. Then $x_k \to 0$ for $k \to \infty$ if and only if $\lim_{k\to\infty} \varrho(\lambda x_k) = 0$ for all $\lambda > 0$.*

Proof. Assume that $\|x_k\|_\varrho \to 0$ and $\lambda > 0$. Then $\|K\,\lambda x_k\|_\varrho < 1$ for all $K > 1$ and large k. Thus $\varrho(K\,\lambda x_k) \leqslant 1$ for large k, hence

$$\varrho(\lambda x_k) \leqslant \frac{1}{K}\varrho(K\,\lambda x_k) \leqslant \frac{1}{K}$$

for large k, by (2.1.5). This implies $\varrho(\lambda x_k) \to 0$.

Assume now that $\varrho(\lambda x_k) \to 0$ for all $\lambda > 0$. Then $\varrho(\lambda x_k) \leqslant 1$ for large k. In particular, $\|x_k\|_\varrho \leqslant 1/\lambda$ for large k. Since $\lambda > 0$ was arbitrary, we get $\|x_k\|_\varrho \to 0$. In other words $x_k \to 0$. $\square$

Apart from our standard topology on X_ϱ, which was induced by the norm, it is possible to define another type of convergence by means of the semimodular.

Definition 2.1.10. Let ϱ be a semimodular on X and $x_k, x \in X_\varrho$. Then we say that x_k is modular convergent (ϱ-convergent) to x if there exists $\lambda > 0$ such that $\varrho(\lambda(x_k - x)) \to 0$. We denote this by $x_k \xrightarrow{\varrho} x$.

It is clear from Lemma 2.1.9 that modular convergence is weaker than norm convergence. Indeed, for norm convergence we have $\lim_{k\to\infty} \varrho(\lambda(x_k - y)) = 0$ for all $\lambda > 0$, while for modular convergence this only has to hold for some $\lambda > 0$.

For some semimodular spaces modular convergence and norm convergence coincide and for others they differ:

Lemma 2.1.11. *Let X_ϱ be a semimodular space. Then modular convergence and norm convergence are equivalent if and only if $\varrho(x_k) \to 0$ implies $\varrho(2x_k) \to 0$.*

Proof. "$\Rightarrow$": Let modular convergence and norm convergence be equivalent and let $\varrho(x_k) \to 0$ with $x_k \in X_\varrho$. Then $x_k \to 0$ and by Lemma 2.1.9 it follows that $\varrho(2x_k) \to 0$.

"$\Leftarrow$": Let $x_k \in X_\varrho$ with $\varrho(x_k) \to 0$. We have to show that $\varrho(\lambda x_k) \to 0$ for all $\lambda > 0$. For fixed $\lambda > 0$ choose $m \in \mathbb{N}$ such that $2^m \geqslant \lambda$. Then by repeated application of the assumption we get $\lim_{k\to\infty} \varrho(2^m x_k) = 0$. Then $0 \leqslant \lim_{k\to\infty} \varrho(\lambda x_k) \leqslant \lambda 2^{-m} \lim_{k\to\infty} \varrho(2^m x_k) = 0$ by (2.1.5). This proves that $x_k \to 0$. □

If either of the equivalent conditions in the previous lemma hold, then we say that the semimodular satisfies the *weak Δ_2-condition*.

If ϱ is a semimodular that satisfies the weak Δ_2-condition, then ϱ is already a modular. Indeed, if $\varrho(x) = 0$, then the constant sequence x is modular convergent to 0 and therefore convergent to 0 with respect to the norm, but this implies $x = 0$.

Lemma 2.1.12. *Let be a semimodular on X that satisfies the weak Δ_2-condition. Then for every $\varepsilon > 0$ there exists $\delta > 0$ such that $\varrho(f) \leqslant \delta$ implies $\|f\|_\varrho \leqslant \varepsilon$.*

Proof. This is an immediate consequence of the equivalence of modular and norm convergence. □

Example 2.1.13. The weak Δ_2-condition of modulars is satisfied in Examples 2.1.4 (a) and (c). Examples 2.1.4 (b) and (d) do not satisfy this condition.

Let us study the closed and open unit ball of X_ϱ. The left-continuity of ϱ is of special significance. The following lemma is of great technical importance. We will invoke it by mentioning the *unit ball property*, or, when more clarity is needed, the *norm-modular unit ball property*.

Lemma 2.1.14 (Norm-modular unit ball property). *Let ϱ be a semimodular on X. Then $\|x\|_\varrho \leqslant 1$ and $\varrho(x) \leqslant 1$ are equivalent. If ϱ is continuous, then also $\|x\|_\varrho < 1$ and $\varrho(x) < 1$ are equivalent, as are $\|x\|_\varrho = 1$ and $\varrho(x) = 1$.*

Proof. If $\varrho(x) \leqslant 1$, then $\|x\|_\varrho \leqslant 1$ by definition of $\|\cdot\|_\varrho$. If on the other hand $\|x\|_\varrho \leqslant 1$, then $\varrho(x/\lambda) \leqslant 1$ for all $\lambda > 1$. Since ϱ is left-continuous it follows that $\varrho(x) \leqslant 1$.

Let ϱ be continuous. If $\|x\|_\varrho < 1$, then there exists $\lambda < 1$ with $\varrho(x/\lambda) \leqslant 1$. Hence by (2.1.5) it follows that $\varrho(x) \leqslant \lambda\varrho(x/\lambda) \leqslant \lambda < 1$. If on the other hand $\varrho(x) < 1$, then by the continuity of ϱ there exists $\gamma > 1$ with $\varrho(\gamma x) < 1$. Hence $\|\gamma x\|_\varrho \leqslant 1$ and $\|x\|_\varrho \leqslant 1/\gamma < 1$. The equivalence of $\|x\|_\varrho = 1$ and $\varrho(x) = 1$ now follows immediately from the cases "$\leqslant 1$" and "< 1". □

A simple example of a semimodular which is left-continuous but not continuous is given by $\varrho_\infty(t) = \infty \cdot \chi_{(1,\infty)}(t)$ on $X = \mathbb{R}$. This is a semimodular on $\mathbb{R}$ and $\|x\|_{\varrho_\infty} = |x|$.

Corollary 2.1.15. *Let ϱ be a semimodular on X and $x \in X_\varrho$.*

(a) *If $\|x\|_\varrho \leqslant 1$, then $\varrho(x) \leqslant \|x\|_\varrho$.*
(b) *If $1 < \|x\|_\varrho$, then $\|x\|_\varrho \leqslant \varrho(x)$.*
(c) *$\|x\|_\varrho \leqslant \varrho(x) + 1$.*

Proof. (a) The claim is obvious for $x = 0$, so let us assume that $0 < \|x\|_\varrho \leqslant 1$. By the unit ball property (Lemma 2.1.14) and $\|x/\|x\|_\varrho\|_\varrho = 1$ it follows that $\varrho(x/\|x\|_\varrho) \leqslant 1$. Since $\|x\|_\varrho \leqslant 1$, it follows from (2.1.5) that $\varrho(x)/\|x\|_\varrho \leqslant 1$.

(b) Assume that $\|x\|_\varrho > 1$. Then $\varrho(x/\lambda) > 1$ for $1 < \lambda < \|x\|_\varrho$ and by (2.1.5) it follows that $1 < \varrho(x)/\lambda$. Since λ was arbitrary, $\varrho(x) \geqslant \|x\|_\varrho$.

(c) This follows immediately from (b). □

Remark 2.1.16. Let $K := \{x \in X_\varrho \colon \varrho(x) \leqslant 1\}$. Then the unit ball property states that $K = \overline{B(0,1)}$, the closed unit ball with respect to the norm. This provides an alternative proof of the fact that $\|\cdot\|_\varrho$ is a norm. Indeed, K is a balanced, i.e. $\lambda K := \{\lambda x \colon x \in K\} \subset K$ for all $|\lambda| \leqslant 1$, convex set. Moreover, by definition of X_ϱ the set K is absorbing for X_ϱ, i.e. $\bigcup_{\lambda>0}(\lambda K) = X_\varrho$. Therefore, the *Minkowski functional* of K, namely $x \mapsto \inf\{\lambda > 0 \colon \frac{1}{\lambda}x \in K\}$, defines a norm on X_ϱ. But this functional is exactly $\|\cdot\|_\varrho$ which is therefore a norm on X_ϱ.

We have seen in Remark 2.1.16 that $\{x \in X_\varrho \colon \varrho(x) \leqslant 1\}$ is closed. This raises the question whether $\{x \in X \colon \varrho(x) \leqslant \alpha\}$ is closed for every $\alpha \in [0, \infty)$. This is equivalent to the lower semicontinuity of ϱ on X_ϱ, hence the next theorem gives a positive answer.

Theorem 2.1.17. *Let ϱ be a semimodular on X. Then ϱ is lower semicontinuous on X_ϱ, i.e.*

$$\varrho(x) \leqslant \liminf_{k\to\infty} \varrho(x_k)$$

for all $x_k, x \in X_\varrho$ with $x_k \to x$ (in norm) for $k \to \infty$.

Proof. Let $x_k, x \in X_\varrho$ with $x_k \to x$ for $k \to \infty$. We begin with the case $\varrho(x) < \infty$. By Lemma 2.1.9, $\lim_{k\to\infty} \varrho(\gamma(x - x_k)) = 0$ for all $\gamma > 0$. Let $\varepsilon \in (0, \frac{1}{2})$. Then, by convexity of ϱ,

$$\begin{aligned}
\varrho\big((1-\varepsilon)x\big) &= \varrho\bigg(\frac{1}{2}x + \frac{1-2\varepsilon}{2}(x - x_k) + \frac{1-2\varepsilon}{2}x_k\bigg) \\
&\leqslant \frac{1}{2}\varrho(x) + \frac{1}{2}\varrho\Big((1-2\varepsilon)(x-x_k) + (1-2\varepsilon)x_k\Big) \\
&\leqslant \frac{1}{2}\varrho(x) + \frac{2\varepsilon}{2}\varrho\bigg(\frac{1-2\varepsilon}{2\varepsilon}(x-x_k)\bigg) + \frac{1-2\varepsilon}{2}\varrho(x_k).
\end{aligned}$$

We pass to the limit $k \to \infty$:

$$\varrho\big((1-\varepsilon)x\big) \leqslant \frac{1}{2}\varrho(x) + \frac{1-2\varepsilon}{2}\liminf_{k\to\infty} \varrho(x_k).$$

Now letting $\varepsilon \to 0^+$ and using the left-continuity of ϱ, we get

$$\varrho\big(x\big) \leqslant \frac{1}{2}\varrho(x) + \frac{1}{2}\liminf_{k\to\infty} \varrho(x_k).$$

Since $\varrho(x) < \infty$, we get $\varrho(x) \leqslant \liminf_{k\to\infty} \varrho(x_k)$. This completes the proof in the case $\varrho(x) < \infty$.

Assume now that $x \in X_\varrho$ with $\varrho(x) = \infty$. If $\liminf_{k\to\infty} \varrho(x_k) = \infty$, then there is nothing to show. So we can assume $\liminf_{k\to\infty} \varrho(x_k) < \infty$. Let $\lambda_0 := \sup\{\lambda > 0 \colon \varrho(\lambda x) < \infty\}$. Since $x \in X_\varrho$, we have $\lambda_0 > 0$. Moreover, $\varrho(x) = \infty$ implies $\lambda_0 \leqslant 1$. For all $\lambda \in (0, \lambda_0)$ the inequality $\varrho(\lambda x) < \infty$ holds, so

$$\varrho(\lambda x) \leqslant \liminf_{k\to\infty} \varrho(\lambda x_k) \leqslant \liminf_{k\to\infty} \varrho(x_k)$$

for all $\lambda \in (0, \lambda_0)$ by the first part of the proof. The left-continuity of ϱ implies that

$$\varrho(\lambda_0 x) \leqslant \liminf_{k\to\infty} \varrho(x_k).$$

If $\lambda_0 = 1$, then the proof is finished. Finally we show, by contradiction, that $\lambda_0 \notin (0, 1)$. So let $\lambda_0 \in (0, 1)$. Choose $\lambda_1 \in (\lambda_0, 1)$ and $\alpha \in (0, 1)$ such that

$$\frac{\lambda_1 - \lambda_0}{\lambda_0} + \alpha + \lambda_0 = 1.$$

The convexity of ϱ implies

$$\begin{aligned}\varrho(\lambda_1 x) &= \varrho\Big((\lambda_1-\lambda_0)x + \lambda_0(x-x_k) + \lambda_0 x_k\Big)\\ &\leqslant \frac{\lambda_1-\lambda_0}{\lambda_0}\varrho(\lambda_0 x) + \alpha\varrho\Big(\frac{\lambda_0}{\alpha}(x-x_k)\Big) + \lambda_0\varrho(x_k).\end{aligned}$$

We pass to the limit $k \to \infty$:

$$\varrho(\lambda_1 x) \leqslant \frac{\lambda_1-\lambda_0}{\lambda_0}\varrho(\lambda_0 x) + \lambda_0 \liminf_{k\to\infty}\varrho(x_k) \leqslant (1-\alpha)\liminf_{k\to\infty}\varrho(x_k).$$

Since $\liminf_k \varrho(x_k) < \infty$, we get $\varrho(\lambda_1 x) < \infty$. But this and $\lambda_1 > \lambda_0$ contradict the definition of λ_0. $\square$

Remark 2.1.18. It follows from Theorem 2.1.17 that the sets $\{x \in X : \varrho(x) \leqslant \alpha\}$ are closed for every $\alpha \in [0,\infty)$. Since these sets are convex, it follows that they are also closed with respect to the weak topology of X_ϱ (cf. Sect. 1.4, Functional analysis).

Remark 2.1.19. Let ϱ be a semimodular on X. Then

$$|||x|||_\varrho := \inf_{\lambda>0} \lambda\Big(1+\varrho\Big(\frac{1}{\lambda}x\Big)\Big)$$

defines a norm on X_ϱ and

$$\|x\|_\varrho \leqslant |||x|||_\varrho \leqslant 2\|x\|_\varrho.$$

This norm is called the *Amemiya norm*. For a proof see [307].

2.2 Conjugate Modulars and Dual Semimodular Spaces

The dual space of a normed space X is the set of all linear, bounded functionals from X to $\mathbb{K}$. It is denoted by X^*. It is well known that X^* equipped with the norm

$$\|x^*\|_{X^*} := \sup_{\|x\|_X \leqslant 1} |\langle x^*, x\rangle|$$

is a Banach space. Here we use the notation $\langle x^*, x\rangle := x^*(x)$. The study of the dual of X is a standard tool to get a better understanding of the space X itself. In this section we examine the dual space of X_ϱ.

Lemma 2.2.1. *Let ϱ be a semimodular on X. A linear functional x^* on X_ϱ is bounded with respect to $\|\cdot\|_\varrho$ if and only if there exists $\gamma > 0$ such that for every $x \in X_\varrho$*

$$|\langle x^*, x\rangle| \leqslant \gamma\big(\varrho(x) + 1\big).$$

Proof. If $x^* \in X_\varrho^*$ and $x \in X_\varrho$, then $\langle x^*, x\rangle \leqslant \|x^*\|_{X_\varrho^*}\|x\|_{X_\varrho} \leqslant \|x^*\|_{X_\varrho^*}(1+\varrho(x))$ by Corollary 2.1.15. Assume conversely that the inequality holds. Then

$$\Big|\Big\langle x^*, \frac{x}{\|x\|_\varrho + \varepsilon}\Big\rangle\Big| \leqslant \gamma\bigg(\varrho\Big(\frac{x}{\|x\|_\varrho + \varepsilon}\Big) + 1\bigg) \leqslant 2\gamma$$

for every $\varepsilon > 0$, hence $\|x^*\|_{X_\varrho^*} \leqslant 2\gamma$. □

Definition 2.2.2. Let ϱ be a semimodular on X. Then by X_ϱ^* we denote the dual space of $(X_\varrho, \|\cdot\|_\varrho)$. Furthermore, we define $\varrho^* \colon X_\varrho^* \to [0,\infty]$ by

$$\varrho^*(x^*) := \sup_{x \in X_\varrho} \big(|\langle x^*, x\rangle| - \varrho(x)\big).$$

We call ϱ^* the *conjugate semimodular* of ϱ.

Note the difference between the spaces X_ϱ^* and X_{ϱ^*}: the former is the dual space of X_ϱ, whereas the latter is the semimodular space defined by ϱ^*.

By definition of the functional ϱ^* we have

$$|\langle x^*, x\rangle| \leqslant \varrho(x) + \varrho^*(x^*) \tag{2.2.3}$$

for all $x \in X_\varrho$ and $x^* \in X_\varrho^*$. This inequality is a generalized version of the classical Young inequality.

Theorem 2.2.4. *Let ϱ be a semimodular on X. Then ϱ^* is a semimodular on X_ϱ^*.*

Proof. It is easily seen that $\varrho^*(0) = 0$, $\varrho^*(\lambda x^*) = \varrho^*(x^*)$ for $|\lambda| = 1$, and $\varrho^*(x^*) \geqslant 0$ for every $x^* \in X_\varrho^*$. Let $x_0^*, x_1^* \in X_\varrho^*$ and $\theta \in (0,1)$. Then

$$\begin{aligned}
\varrho^*\big((1-\theta)x_0^* + \theta x_1^*\big) &= \sup_{x\in X}\big(\big|\big\langle (1-\theta)x_0^* + \theta x_1^*, x\big\rangle\big| - \varrho(x)\big) \\
&\leqslant (1-\theta)\sup_{x\in X}\big(\big|\big\langle x_0^*, x\big\rangle\big| - \varrho(x)\big) \\
&\quad + \theta \sup_{x\in X}\big(\big|\big\langle x_1^*, x\big\rangle\big| - \varrho(x)\big) \\
&= (1-\theta)\varrho^*(x_0^*) + \theta\varrho^*(x_1^*).
\end{aligned}$$

Finally, let $\varrho^*(\lambda x^*) = 0$ for every $\lambda > 0$. For $x \in X_\varrho$ choose $\eta > 0$ such that $\varrho(\eta x) < \infty$. Then by (2.2.3)

$$\lambda\eta\,|\langle x^*, x\rangle| \leqslant \varrho(\eta x) + \varrho^*(\lambda x^*) = \varrho(\eta x).$$

Taking $\lambda \to \infty$ we obtain $|\langle x^*, x\rangle| = 0$. Hence $x^* = 0$. It remains to show that ϱ^* is left-continuous. For $\lambda \to 1^-$ and $x^* \in X_\varrho^*$ we have

$$\begin{aligned}
\lim_{\lambda\to 1^-} \varrho^*(\lambda x^*) &= \lim_{\lambda\to 1^-} \sup_{x\in X} \big(|\langle \lambda x^*, x\rangle| - \varrho(x)\big)\\
&= \sup_{0<\lambda<1} \sup_{x\in X} \big(|\lambda|\,|\langle x^*, x\rangle| - \varrho(x)\big)\\
&= \sup_{x\in X} \big(|\langle x^*, x\rangle| - \varrho(x)\big) = \varrho^*(x).
\end{aligned}$$

Thus ϱ^* is left-continuous. □

For a semimodular ϱ on X we have defined the conjugate semimodular ϱ^* on X_ϱ^*. By duality we can proceed further and define ϱ^{**} the conjugate semimodular of ϱ^* on the bidual $X_\varrho^{**} := (X_\varrho^*)^*$. The functional ϱ^{**} is called the *biconjugate semimodular of ϱ on X_ϱ^{**}*. Using the natural injection ι of X_ϱ into its bidual X_ϱ^{**}, the mapping $x \mapsto \varrho^{**}(\iota x)$ defines a semimodular on X_ϱ, which we call the *biconjugate semimodular of ϱ on X_ϱ*. For simplicity of notation it is also denoted by ϱ^{**} neglecting the extra injection ι. In particular, we have

$$\varrho^{**}(x) = \sup_{x^*\in X_\varrho^*} \big(|\langle x^*, x\rangle| - \varrho^*(x^*)\big) \tag{2.2.5}$$

for all $x \in X_\varrho$. Certainly the formula is also valid for all $x \in X_\varrho^{**}$, by the definition of ϱ^{**} on X_ϱ^{**}, if we interpret $\langle x^*, x\rangle$ as $\langle x, x^*\rangle_{X_\varrho^{**}\times X_\varrho^*}$.

Analogously to the fact that $\iota : X_\varrho \to X_\varrho^{**}$ is an isometry, it turns out that the biconjugate ϱ^{**} and ϱ coincide on X_ϱ.

Theorem 2.2.6. *Let ϱ be a semimodular on X. Then $\varrho^{**} = \varrho$ on X_ϱ.*

Proof. Exactly as in the proof of Theorem 2.2.4 we can prove that ϱ^{**} is a semimodular on X_ϱ. By definition of ϱ^{**} and (2.2.3) we get for $x \in X_\varrho$

$$\begin{aligned}
\varrho^{**}(x) &= \sup_{x^*\in X_\varrho^*} \big(|\langle x^*, x\rangle| - \varrho^*(x^*)\big)\\
&= \sup_{x^*\in X_\varrho^*, \varrho^*(x^*)<\infty} \big(|\langle x^*, x\rangle| - \varrho^*(x^*)\big)\\
&\leqslant \sup_{x^*\in X_\varrho^*, \varrho^*(x^*)<\infty} \big(\varrho(x) + \varrho^*(x^*) - \varrho^*(x^*)\big)\\
&= \varrho(x).
\end{aligned}$$

It remains to show $\varrho^{**}(x) \geqslant \varrho(x)$. We prove this by contradiction. Assume to the contrary that there exists $x_0 \in X_\varrho$ with $\varrho^{**}(x_0) < \varrho(x_0)$. In particular, $\varrho^{**}(x_0) < \infty$. We define the epigraph of ϱ by

$$\operatorname{epi}(\varrho) := \bigcup_{\lambda \in \mathbb{R}} \big\{(x,\gamma) \in X_\varrho \times \mathbb{R} \colon \gamma \geqslant \varrho(x)\big\}.$$

Since ϱ is convex and lower semicontinuous (Theorem 2.1.17), the set $\operatorname{epi}(\varrho)$ is convex and closed (cf. [58, Sect. I.3]). Moreover, due to $\varrho^{**}(x_0) < \varrho(x_0)$ the point $(x_0, \varrho^{**}(x_0))$ is not contained in $\operatorname{epi}(\varrho)$. So by the Hahn–Banach Theorem 1.4.2 there exists a functional on $X_\varrho \times \mathbb{R}$ which strictly separates $\operatorname{epi}(\varrho)$ from $(x_0, \varrho^{**}(x_0))$. So there exist $\alpha, \beta \in \mathbb{R}$ and $x^* \in X_\varrho^*$ with

$$\langle x^*, x\rangle - \beta\varrho(x) < \alpha < \langle x^*, x_0\rangle - \beta\varrho^{**}(x_0)$$

for all $x \in X_\varrho$. The choice $x = x_0$ and the estimate $\varrho^{**}(x_0) < \varrho(x_0)$ imply $\beta > 0$. We multiply by $\frac{1}{\beta}$ and get

$$\Big\langle \frac{x^*}{\beta}, x\Big\rangle - \varrho(x) < \frac{\alpha}{\beta} < \Big\langle \frac{x^*}{\beta}, x_0\Big\rangle - \varrho^{**}(x_0)$$

for all $x \in X_\varrho$. Due to (2.2.5) the right-hand side is bounded by $\varrho^*(\frac{x^*}{\beta})$. Now, taking the supremum on the left-hand side over $x \in X_\varrho$ implies

$$\varrho^*\Big(\frac{x^*}{\beta}\Big) \leqslant \frac{\alpha}{\beta} < \varrho^*\Big(\frac{x^*}{\beta}\Big).$$

This is the desired contradiction. □

For two semimodulars ϱ, κ on X we write $\varrho \leqslant \kappa$ if $\varrho(f) \leqslant \kappa(f)$ for every $f \in X$.

Corollary 2.2.7. *Let ϱ, κ be semimodulars on X. Then $\varrho \leqslant \kappa$ if and only if $\kappa^* \leqslant \varrho^*$.*

Proof. If $\varrho \leqslant \kappa$, then by definition of the conjugate semimodular follows easily $\kappa^* \leqslant \varrho^*$. If however $\kappa^* \leqslant \varrho^*$, then $\varrho^{**} \leqslant \kappa^{**}$ and by Theorem 2.2.6 follows $\varrho \leqslant \kappa$. □

From Theorem 2.1.17 we already know that the modular ϱ is lower semicontinuous on X_ϱ with respect to convergence in norm. This raises the question of whether ϱ is also lower semicontinuous on X_ϱ with respect to weak convergence. Let $f_k, f \in X_\varrho$. As usual we say that f_k *converges weakly* to f if $\langle g^*, f_k\rangle \to \langle g^*, f\rangle$ for all $g^* \in X_\varrho^*$. In this case we write $f_k \rightharpoonup f$.

Theorem 2.2.8. *Let ϱ be a semimodular on X, then the semimodular ϱ is weakly (sequentially) lower semicontinuous, i.e. if $f_k \rightharpoonup f$ weakly in X_ϱ, then $\varrho(f) \leqslant \liminf_{k\to\infty} \varrho(f_k)$.*

Proof. Let $f_k, f \in X_\varrho$ with $f_k \rightharpoonup f$. Then, by Theorem 2.2.6, $\varrho = \varrho^{**}$, which implies

$$\begin{aligned}
\varrho(f) = \varrho^{**}(f) &= \sup_{g^* \in X_\varrho^*} \big(|\langle g^*, f\rangle| - \varrho^*(g^*)\big) \\
&= \sup_{g^* \in X_\varrho^*} \big(\lim_{k\to\infty} |\langle g^*, f_k\rangle| - \varrho^*(g^*)\big) \\
&\leqslant \liminf_{k\to\infty} \Big(\sup_{g^* \in X_\varrho^*} \big(|\langle g^*, f_k\rangle| - \varrho^*(g^*)\big)\Big) \\
&= \liminf_{k\to\infty} \varrho^{**}(f_k) \\
&= \liminf_{k\to\infty} \varrho(f_k).
\end{aligned}$$

□

In the definition of ϱ^* the supremum is taken over all $x \in X_\varrho$. However, it is possible to restrict this to the closed unit ball of X_ϱ.

Lemma 2.2.9. *If ϱ is a semimodular on X, then*

$$\varrho^*(x^*) = \sup_{x\in X_\varrho, \|x\|_\varrho \leqslant 1} \big(|\langle x^*, x\rangle| - \varrho(x)\big) = \sup_{x\in X_\varrho, \varrho(x)\leqslant 1} \big(|\langle x^*, x\rangle| - \varrho(x)\big)$$

for $x^* \in X_\varrho^*$ with $\|x^*\|_{X_\varrho^*} \leqslant 1$.

Proof. The equivalence of the suprema follows from the unit ball property (Lemma 2.1.14). Let $\|x^*\|_{X_\varrho^*} \leqslant 1$. By the definition of the dual norm we have

$$\begin{aligned}
\sup_{\|x\|_\varrho > 1} \big(|\langle x^*, x\rangle| - \varrho(x)\big) &\leqslant \sup_{\|x\|_\varrho > 1} \big(\|x^*\|_{X_\varrho^*}\|x\|_\varrho - \varrho(x)\big) \\
&\leqslant \sup_{\|x\|_\varrho > 1} \big(\|x\|_\varrho - \varrho(x)\big).
\end{aligned}$$

If $\|x\|_\varrho > 1$, then $\varrho(x) \geqslant \|x\|_\varrho$ by Corollary 2.1.15, and so the right-hand side of the previous inequality is non-positive. Since ϱ^* is defined as a supremum, and is always non-negative, we see that the points with $\|x\|_\varrho > 1$ do not affect the supremum, and so the claim follows. □

Since ϱ^* is a semimodular on X_ϱ^*, it defines another norm $\|\cdot\|_{\varrho^*}$ on X_ϱ^*. We next want to compare it with the norm $\|\cdot\|_{X_\varrho^*}$.

Theorem 2.2.10. *If ϱ be a semimodular on X, then for every $x^* \in X_\varrho^*$*

$$\|x^*\|_{\varrho^*} \leqslant \|x^*\|_{X_\varrho^*} \leqslant 2\|x^*\|_{\varrho^*}.$$

Proof. We first prove the second inequality. By the unit ball property (Lemma 2.1.14) the inequalities $\|x\|_\varrho \leqslant 1$ and $\varrho(x) \leqslant 1$ are equivalent. Hence,

$$\|x^*\|_{X_\varrho^*} = \sup_{\|x\|_\varrho \leqslant 1} |\langle x^*, x\rangle| \leqslant \sup_{\varrho(x)\leqslant 1} \big(\varrho^*(x^*) + \varrho(x)\big) \leqslant \varrho^*(x^*) + 1.$$

If $\|x^*\|_{\varrho^*} \leqslant 1$, then $\varrho^*(x^*) \leqslant 1$ by the unit ball property and we conclude that $\|x^*\|_{X_\varrho^*} \leqslant 2$. The conclusion follows from this by a *scaling argument*: if $\|x^*\|_{\varrho^*} > 0$, then set $y^* := x^*/\|x^*\|_{\varrho^*}$. Since $\|y^*\|_{\varrho^*} = 1$, we conclude that $\|y^*\|_{X_\varrho^*} \leqslant 2\|y^*\|_{\varrho^*}$. Multiplying by $\|x^*\|_{\varrho^*}$ gives the result.

Assume now that $\|x^*\|_{X_\varrho^*} \leqslant 1$. Then by Lemma 2.2.9 and Corollary 2.1.15 (c)

$$\varrho^*(x^*) = \sup_{x \in X_\varrho, \varrho(x) \leqslant 1} \big(|\langle x^*, x\rangle| - \varrho(x)\big) \leqslant \sup_{x \in X_\varrho, \varrho(x) \leqslant 1} \big(\|x\|_\varrho - \varrho(x)\big) \leqslant 1.$$

Hence, $\|x^*\|_{\varrho^*} \leqslant 1$. The scaling argument gives $\|x^*\|_{\varrho^*} \leqslant \|x^*\|_{X_\varrho^*}$ □

Note the scaling argument technique used in the previous proof. It is one of the central methods for dealing with these kind of spaces, and it will be used often in what follows.

With the help of the conjugate semimodular ϱ^* it is also possible to define yet another norm on X_ϱ by means of duality. Luckily this norm is equivalent to the norm $\|\cdot\|_\varrho$.

Theorem 2.2.11. *Let ϱ be a semimodular on X. Then*

$$\begin{aligned} \|x\|'_\varrho &:= \sup\big\{|\langle x^*, x\rangle| \colon x^* \in X_\varrho^*, \|x^*\|_{\varrho^*} \leqslant 1\big\} \\ &= \sup\big\{|\langle x^*, x\rangle| \colon x^* \in X_\varrho^*, \varrho^*(x^*) \leqslant 1\big\} \end{aligned}$$

defines a norm on X_ϱ. This norm is called the Orlicz norm. *For all $x \in X_\varrho$ we have $\|x\|_\varrho \leqslant \|x\|'_\varrho \leqslant 2\|x\|_\varrho$.*

Proof. By the unit ball property (Lemma 2.1.14) the two suprema are equal. If $\|x\|_\varrho \leqslant 1$ and $\|x^*\|_{\varrho^*} \leqslant 1$, then $\varrho(x) \leqslant 1$ and $\varrho^*(x^*) \leqslant 1$. Hence, $|\langle x, x^*\rangle| \leqslant \varrho(x) + \varrho^*(x^*) \leqslant 2$. Therefore $\|x\|'_\varrho \leqslant 2$. A scaling argument proves $\|x\|'_\varrho \leqslant 2\|x\|_\varrho$.

If $\|x\|'_\varrho \leqslant 1$, then $|\langle x^*, x\rangle| \leqslant 1$ for all $x^* \in X_\varrho^*$ with $\|x^*\|_{\varrho^*} \leqslant 1$. In particular, by Theorem 2.2.10 we have $|\langle x^*, x\rangle| \leqslant 1$ for all $x^* \in X_\varrho^*$ with $\|x^*\|_{X_\varrho^*} \leqslant 1$. Hence, Corollary 1.4.3 implies $\|x\|_\varrho \leqslant 1$. We have thus shown that $\|x\|_\varrho \leqslant \|x\|'_\varrho$. □

2.3 Musielak–Orlicz Spaces: Basic Properties

In this section we start our journey towards more concrete spaces. Instead of general semimodular spaces, we will consider spaces where the modular is given by the integral of a real-valued function.

Definition 2.3.1. A convex, left-continuous function $\varphi\colon [0,\infty) \to [0,\infty]$ with $\varphi(0) = 0$, $\lim_{t\to 0^+} \varphi(t) = 0$, and $\lim_{t\to\infty} \varphi(t) = \infty$ is called a Φ-*function.* It is called *positive* if $\varphi(t) > 0$ for all $t > 0$.

In fact, there is a very close relationship between Φ-functions and semimodulars on $\mathbb{R}$.

Lemma 2.3.2. *Let $\varphi\colon [0,\infty) \to [0,\infty]$ and let ϱ denote its even extension to $\mathbb{R}$, i.e. $\varrho(t) := \varphi(|t|)$ for all $t \in \mathbb{R}$. Then φ is a Φ-function if and only if ϱ is a semimodular on $\mathbb{R}$ with $X_\varrho = \mathbb{R}$. Moreover, φ is a positive Φ-function if and only if ϱ is a modular on $\mathbb{R}$ with $X_\varrho = \mathbb{R}$.*

Proof. "⇒": Let φ be a Φ-function. Since $\lim_{t\to 0^+} \varphi(t) = 0$, we have $X_\varrho = \mathbb{R}$. To prove that ϱ is a semimodular on $\mathbb{R}$ it remains to prove that $\varrho(\lambda t_0) = 0$ for all $\lambda > 0$ implies $t_0 = 0$. So assume that $\varrho(\lambda t_0) = 0$ for all $\lambda > 0$. Since $\lim_{t\to\infty} \varphi(t) = \infty$, there exists $t_1 > 0$ with $\varphi(t_1) > 0$. Thus there exists no $\lambda > 0$ such that $t_1 = \lambda t_0$, which implies that $t_0 = 0$. Hence ϱ is a semimodular. Assume that φ is additionally positive. If $\varrho(s) = 0$, then $\varphi(|s|) = 0$ and therefore $s = 0$. This proves that ϱ is a modular.

"⇐": Let ϱ be a semimodular on $\mathbb{R}$ with $X_\varrho = \mathbb{R}$. Since $X_\varrho = \mathbb{R}$, there exists $t_2 > 0$ such that $\varrho(t_2) < \infty$. From (2.1.5) follows that $0 \leqslant \varphi(t) \leqslant t/t_2\varphi(t_2)$ for all $t \in [0, t_2]$, which implies that $\lim_{t\to 0^+} \varphi(t) = 0$. Since $1 \neq 0$, there exists $\lambda > 0$ such that $\varrho(\lambda \cdot 1) \neq 0$. In particular there exists $t_3 > 0$ with $\varphi(t_3) > 0$ and $\varphi(kt_3) \geqslant k\varphi(t_3) > 0$ by (2.1.5) for all $k \in \mathbb{N}$. Since k is arbitrary, we get $\lim_{t\to\infty} \varphi(t) = \infty$. We have proved that φ is a Φ-function. Assume additionally that ϱ is a modular. In particular $\varrho(t) = \varphi(|t|) = 0$ implies $t = 0$. Hence by negation we get that $t > 0$ implies $\varphi(t) > 0$, so φ is positive. □

Let us remark that if φ is a Φ-function then on the set $\{t \geqslant 0\colon \varphi(t) < \infty\}$ it has the form

$$\varphi(t) = \int_0^t a(\tau)\, d\tau, \tag{2.3.3}$$

where $a(t)$ is the right-derivative of $\varphi(t)$ (see [330], Theorem 1.3.1). Moreover, the function $a(t)$ is non-increasing and right-continuous.

The following lemma is an easy consequence of the left-continuity, convexity, and monotonicity of φ. However, it is also possible to use Lemma 2.3.2 and Theorem 2.1.17 to prove this.

Lemma 2.3.4. *Every Φ-function is lower semicontinuous.*

Example 2.3.5. Let $1 \leqslant p < \infty$. Define

$$\varphi_p(t) := \frac{1}{p} t^p,$$
$$\varphi_\infty(t) := \infty \cdot \chi_{(1,\infty)}(t)$$

for all $t \geqslant 0$. Then φ_p and φ_∞ are Φ-functions. Moreover, φ_p is continuous and positive, while φ_∞ is only left-continuous and lower semicontinuous but not positive.

Remark 2.3.6. Let φ be a Φ-function. As a lower semicontinuous function φ satisfies

$$\varphi\big(\inf A\big) \leqslant \inf \varphi(A)$$

for every non-empty set $A \subset [0,\infty)$. The reverse estimate might fail as the example φ_∞ with $A := (1,\infty)$ shows. However, for every $\lambda > 1$ we have

$$\inf \varphi(A) \leqslant \varphi\big(\lambda \inf A\big).$$

Indeed, if $\inf A = 0$, then the claim follows by $\lim_{t\to 0^+} \varphi(t) = 0$. If $\inf A > 0$, then we can find $t \in A$ such that $\inf A \leqslant t \leqslant \lambda \inf A$. Now, the monotonicity of φ implies $\inf \varphi(A) \leqslant \varphi(t) \leqslant \varphi\big(\lambda \inf A\big)$.

Remark 2.3.7. Let φ be a Φ-function. Then, as a convex function, φ is continuous if and only if φ is finite on $[0,\infty)$.

The following properties of Φ-functions are very useful:

$$\begin{aligned} \varphi(rt) &\leqslant r\varphi(t), \\ \varphi(st) &\geqslant s\varphi(t), \end{aligned} \tag{2.3.8}$$

for any $r \in [0,1]$, $s \in [1,\infty)$ and $t \geqslant 0$ (compare with (2.1.5)). This is a simple consequence of the convexity of φ and $\varphi(0) = 0$. Inequality (2.3.8) further implies that $\varphi(a) + \varphi(b) \leqslant \frac{a}{a+b}\varphi(a+b) + \frac{b}{a+b}\varphi(a+b) = \varphi(a+b)$ for $a+b > 0$ for all $a, b \geqslant 0$ which combined with convexity yields

$$\varphi(a) + \varphi(b) \leqslant \varphi(a+b) \leqslant \tfrac{1}{2}(\varphi(2a) + \varphi(2b)).$$

Although it is possible to define function spaces using Φ-functions, these are not sufficiently general for our needs. In the case of variable exponent Lebesgue spaces (see Chap. 3) we need our function φ to depend also on the location in the space. So we need to generalize Φ-functions in such a way that they may depend on the space variable.

Definition 2.3.9. Let (A, Σ, μ) be a σ-finite, complete measure space. A real function $\varphi\colon A \times [0,\infty) \to [0,\infty]$ is said to be a *generalized Φ-function* on (A, Σ, μ) if:

(a) $\varphi(y,\cdot)$ is a Φ-function for every $y \in A$.
(b) $y \mapsto \varphi(y,t)$ is measurable for every $t \geqslant 0$.

If φ is a generalized Φ-function on (A, Σ, μ), we write $\varphi \in \Phi(A,\mu)$. If Ω is an open subset of $\mathbb{R}^n$ and μ is the n-dimensional Lebesgue measure we abbreviate this as $\varphi \in \Phi(\Omega)$ or say that φ is a generalized Φ-function on Ω.

In what follows we always make the natural assumption that our measure μ is not identically zero.

Certainly every Φ-function is a generalized Φ-function if we set $\varphi(y,t) := \varphi(t)$ for $y \in A$ and $t \in [0,\infty)$. Also, from (2.3.8) and Lemma 2.3.4 we see that $\varphi(y,\cdot)$ is non-decreasing and lower semicontinuous on $[0,\infty)$ for every $y \in A$.

We say that a function is *simple* if it is the linear combination of characteristic functions of measurable sets with finite measure, $\sum_{i=1}^k s_i \chi_{A_i}(x)$ with $\mu(A_1),\dots,\mu(A_k) < \infty$, $s_1,\dots,s_k \in \mathbb{K}$. We denote the set of simple functions by $S(A,\mu)$. If Ω is an open subset of $\mathbb{R}^n$ and μ is the n-dimensional Lebesgue measure we abbreviate this by $S(\Omega)$.

We next show that every generalized Φ-function generates a semimodular on $L^0(A,\mu)$.

Lemma 2.3.10. *If $\varphi \in \Phi(A,\mu)$ and $f \in L^0(A,\mu)$, then $y \mapsto \varphi(y,|f(y)|)$ is μ-measurable and*

$$\varrho_\varphi(f) := \int_A \varphi(y,|f(y)|)\,d\mu(y)$$

is a semimodular on $L^0(A,\mu)$. If φ is positive, then ϱ_φ is a modular. We call ϱ_φ the semimodular induced by *φ.*

Proof. By splitting the function into its positive and negative (real and imaginary) part it suffices to consider the case $f \geqslant 0$. Let $f_k \nearrow f$ point-wise where f_k are non-negative simple functions. Then

$$\varphi(y,|f_k(y)|) = \sum_j \varphi(y,\alpha_j^k)\cdot \chi_{A_j^k}(y),$$

which is measurable and $\varphi(y,f_k(y)) \nearrow \varphi(y,f(y))$. Thus $\varphi(\cdot,f(\cdot))$ is measurable.

Obviously, $\varrho_\varphi(0) = 0$ and $\varrho_\varphi(\lambda x) = \varrho_\varphi(x)$ for $|\lambda| = 1$. The convexity of ϱ_φ is a direct consequence of the convexity of φ. Let us show the left-continuity of ϱ_φ: if $\lambda_k \to 1^-$ and $y \in A$, then $0 \leqslant \varphi(y,\lambda_k f(y)) \to \varphi(y,f(y))$ by the left-continuity and monotonicity of $\varphi(y,\cdot)$. Hence $\varrho_\varphi(\lambda_k f) \to \varrho_\varphi(f)$, by the theorem of monotone convergence. So ϱ_φ is left-continuous in the sense of Definition 2.1.1 (d).

Assume now that $f \in L^0(A,\mu)$ is such that $\varrho_\varphi(\lambda f) = 0$ for all $\lambda > 0$. So for any $k \in \mathbb{N}$ we have $\varphi(y,kf(y)) = 0$ for almost all $y \in A$. Since $\mathbb{N}$ is countable we deduce that $\varphi(y,kf(y)) = 0$ for almost all $y \in A$ and all $k \in \mathbb{N}$. The convexity of φ and $\varphi(y,0) = 0$ imply that $\varphi(y,\lambda f(y)) = 0$ for almost all $y \in A$ and all $\lambda > 0$. Since $\lim_{t\to\infty} \varphi(y,t) = \infty$ for all $y \in A$, this implies that $|f(y)| = 0$ for almost all $y \in A$, hence $f = 0$. So ϱ_φ is a semimodular on $L^0(A,\mu)$.

Assume now that φ is positive and that $\varrho_\varphi(f) = 0$. Then $\varphi(y, f(y)) = 0$ for almost all $y \in A$. Since φ is positive, $f(y) = 0$ for almost all $y \in A$, thus $f = 0$. This proves that ϱ_φ is a modular on $L^0(A, \mu)$. □

Since every $\varphi \in \Phi(A, \mu)$ generates a semimodular it is natural to study the corresponding semimodular space.

Definition 2.3.11. Let $\varphi \in \Phi(A, \mu)$ and let ϱ_φ be given by

$$\varrho_\varphi(f) := \int\limits_A \varphi(y, |f(y)|)\, d\mu(y)$$

for all $f \in L^0(A, \mu)$. Then the semimodular space

$$\begin{aligned}(L^0(A,\mu))_{\varrho_\varphi} &= \{f \in L^0(A,\mu)\colon \lim_{\lambda\to 0} \varrho_\varphi(\lambda f) = 0\}\\ &= \big\{f \in L^0(A,\mu) : \varrho_\varphi(\lambda f) < \infty \text{ for some } \lambda > 0\big\}\end{aligned}$$

will be called *Musielak–Orlicz space* and denoted by $L^\varphi(A, \mu)$ or L^φ, for short. The norm $\|\cdot\|_{\varrho_\varphi}$ is denoted by $\|\cdot\|_\varphi$, thus

$$\|f\|_\varphi = \inf\Big\{\lambda > 0\colon \varrho_\varphi\Big(\frac{x}{\lambda}\Big) \leqslant 1\Big\}.$$

The Musielak–Orlicz spaces are also called *generalized Orlicz spaces*. They provide a good framework for many function spaces. Here are some examples.

Example 2.3.12. Let (A, Σ, μ) be a σ-finite, complete measure space.

(a) The (semi)modulars given in Example 2.1.4 (a)–(c) give rise to (weighted) Lebesgue spaces.

(b) Let φ be a Φ-function. Then

$$\varrho_\varphi(f) = \int\limits_A \varphi(|f(y)|)\, d\mu(y)$$

is a semimodular on $L^0(A, \mu)$. If φ is positive, then ϱ is a modular on $L^0(A, \mu)$ and the space $L^\varphi(A, \mu)$ is called an *Orlicz space*.
With suitable choices of φ, A and μ, this includes all modulars in Example 2.1.4 except (f).

(c) Example 2.1.4 (f) is not a Musielak–Orlicz space.

As a semimodular space, $L^\varphi = (L^\varphi, \|\cdot\|_\varphi)$ is a normed space, which, in fact, is complete.

Theorem 2.3.13. *Let $\varphi \in \Phi(A, \mu)$. Then $L^\varphi(A, \mu)$ is a Banach space.*

Before we get to the proof of Theorem 2.3.13 we need to prove two useful lemmas.

Lemma 2.3.14. *Let $\varphi \in \Phi(A,\mu)$ and $\mu(A) < \infty$. Then every $\|\cdot\|_\varphi$-Cauchy sequence is also a Cauchy sequence with respect to convergence in measure.*

Proof. Fix $\varepsilon > 0$ and let $V_t := \{y \in A\colon \varphi(y,t) = 0\}$ for $t > 0$. Then V_t is measurable. For all $y \in A$ the function $t \mapsto \varphi(y,t)$ is non-decreasing and $\lim_{t\to\infty} \varphi(y,t) = \infty$, so $V_t \searrow \emptyset$ as $t \to \infty$. Therefore, $\lim_{k\to\infty} \mu(V_k) = \mu(\emptyset) = 0$, where we have used that $\mu(A) < \infty$. Thus, there exists $K \in \mathbb{N}$ such that $\mu(V_K) < \varepsilon$. Note that if φ is positive then $V_t = \emptyset$ for all $t > 0$ and we do not need this step in the proof.

For a μ-measurable set $E \subset A$ define

$$\nu_K(E) := \varrho_\varphi(K\,\chi_E) = \int\limits_E \varphi(y,K)\,d\mu(y).$$

If E is μ-measurable with $\nu_K(E) = 0$, then $\varphi(y,K) = 0$ for μ-almost every $y \in E$. Thus $\mu(E \setminus V_K) = 0$ by the definition of V_K. Hence, E is a $\mu|_{A\setminus V_K}$-null set, which means that the measure $\mu|_{A\setminus V_K}$ is absolutely continuous with respect to ν_K.

Since $\mu(A \setminus V_K) \leqslant \mu(A) < \infty$ and $\mu|_{A\setminus V_K}$ is absolutely continuous with respect to ν_K, there exists $\delta \in (0,1)$ such that $\nu_K(E) \leqslant \delta$ implies $\mu(E \setminus V_K) \leqslant \varepsilon$ (cf. [184, Theorem 30.B]). Since f_k is a $\|\cdot\|_\varphi$-Cauchy sequence, there exists $k_0 \in \mathbb{N}$ such that $\|K\,\varepsilon^{-1}\delta^{-1}(f_m - f_k)\|_\varphi \leqslant 1$ for all $m,k \geqslant k_0$. Assume in the following $m,k \geqslant k_0$, then by (2.1.5) and the norm-modular unit ball property (Lemma 2.1.14)

$$\varrho_\varphi\big(K\,\varepsilon^{-1}(f_m - f_k)\big) \leqslant \delta\varrho_\varphi\big(K\,\varepsilon^{-1}\delta^{-1}(f_m - f_k)\big) \leqslant \delta.$$

Let us write $E_{m,k,\varepsilon} := \{y \in A\colon |f_m(y) - f_k(y)| \geqslant \varepsilon\}$. Then

$$\nu_K(E_{m,k,\varepsilon}) = \int\limits_{E_{m,k,\varepsilon}} \varphi(y,K)\,d\mu(y) \leqslant \varrho_\varphi\big(K\,\varepsilon^{-1}(f_m - f_k)\big) \leqslant \delta.$$

By the choice of δ, this implies that $\mu(E_{m,k,\varepsilon} \setminus V_K) \leqslant \varepsilon$. With $\mu(V_K) < \varepsilon$ we have $\mu(E_{m,k,\varepsilon}) \leqslant 2\varepsilon$. Since $\varepsilon > 0$ was arbitrary, this proves that f_k is a Cauchy sequence with respect to convergence in measure.

If $\|f_k\|_\varphi \to 0$, then as above there exists $K \in \mathbb{N}$ such that $\mu(\{|f_k| \geqslant \varepsilon\}) \leqslant 2\varepsilon$ for all $k \geqslant K$. This proves $f_k \to 0$ in measure. $\square$

Lemma 2.3.15. *Let $\varphi \in \Phi(A,\mu)$. Then every $\|\cdot\|_\varphi$-Cauchy sequence $(f_k) \subset L^\varphi$ has a subsequence which converges μ-almost everywhere to a measurable function f.*

Proof. Recall that μ is σ-finite. Let $A = \bigcup_{i=1}^{\infty} A_i$ with A_i pairwise disjoint and $\mu(A_i) < \infty$ for all $i \in \mathbb{N}$. Then, by Lemma 2.3.14, (f_k) is a Cauchy sequence with respect to convergence in measure on A_1. Therefore there exists a measurable function $f\colon A_1 \to \mathbb{K}$ and a subsequence of f_k which converges to f μ-almost everywhere. Repeating this argument for every A_i and passing to the diagonal sequence we get a subsequence (f_{k_j}) and a μ-measurable function $f\colon A \to \mathbb{K}$ such that $f_{k_j} \to f$ μ-almost everywhere. □

Let us now get to the proof of the completeness of L^φ.

Proof of Theorem 2.3.13. Let (f_k) be a Cauchy sequence. By Lemma 2.3.15 there exists a subsequence f_{k_j} and a μ-measurable function $f\colon A \to \mathbb{K}$ such that $f_{k_j} \to f$ for μ-almost every $y \in A$. This implies $\varphi(y, |f_{k_j}(y) - f(y)|) \to 0$ μ-almost everywhere. Let $\lambda > 0$ and $0 < \varepsilon < 1$. Since (f_k) is a Cauchy sequence, there exists $K = K(\lambda, \varepsilon) \in \mathbb{N}$ such that $\|\lambda(f_m - f_k)\|_\varphi < \varepsilon$ for all $m, k \geqslant N$, which implies $\varrho_\varphi(\lambda(f_m - f_k)) \leqslant \varepsilon$ by Corollary 2.1.15. Therefore by Fatou's lemma

$$\begin{aligned}
\varrho_\varphi\big(\lambda(f_m - f)\big) &= \int_A \lim_{j\to\infty} \varphi\big(y, \lambda|f_m(y) - f_{k_j}(y)|\big)\, d\mu(y) \\
&\leqslant \liminf_{j\to\infty} \int_A \varphi\big(y, \lambda|f_m(y) - f_{k_j}(y)|\big)\, d\mu(y) \\
&= \liminf_{j\to\infty} \varrho_\varphi\big(\lambda(f_m - f_{k_j})\big) \\
&\leqslant \varepsilon.
\end{aligned}$$

So $\varrho_\varphi(\lambda(f_m - f)) \to 0$ for $m \to \infty$ and all $\lambda > 0$ and $\|f_k - f\|_\varphi \to 0$ by Lemma 2.1.9. Thus every Cauchy sequence converges in L^φ, as was to be shown. □

The next lemma collects analogues of the classical Lebesgue integral convergence results.

Lemma 2.3.16. *Let $\varphi \in \Phi(A, \mu)$ and $f_k, f, g \in L^0(A, \mu)$.*

(a) *If $f_k \to f$ μ-almost everywhere, then $\varrho_\varphi(f) \leqslant \liminf_{k\to\infty} \varrho_\varphi(f_k)$.*
(b) *If $|f_k| \nearrow |f|$ μ-almost everywhere, then $\varrho_\varphi(f) = \lim_{k\to\infty} \varrho_\varphi(f_k)$.*
(c) *If $f_k \to f$ μ-almost everywhere and $|f_k| \leqslant |g|$ μ-almost everywhere, and $\varrho_\varphi(\lambda g) < \infty$ for every $\lambda > 0$, then $f_k \to f$ in L^φ.*

These properties are called Fatou's lemma (for the modular), monotone convergence *and* dominated convergence*, respectively.*

Proof. By Lemma 2.3.4 the mappings $\varphi(y, \cdot)$ are lower semicontinuous. Thus Fatou's lemma implies

$$\begin{aligned}\varrho_\varphi(f) &= \int_A \varphi(y, \lim_{k\to\infty} |f_k(y)|)\, d\mu(y) \\ &\leqslant \int_A \liminf_{k\to\infty} \varphi(y, |f_k(y)|)\, d\mu(y) \\ &\leqslant \liminf_{k\to\infty} \int_A \varphi(y, |f_k(y)|)\, d\mu(y) \\ &= \liminf_{k\to\infty} \varrho_\varphi(f_k).\end{aligned}$$

This proves (a).

To prove (b) let $|f_k| \nearrow |f|$. Then by the left-continuity and monotonicity of $\varphi(y,\cdot)$, we have $0 \leqslant \varphi(\cdot, |f_k(\cdot)|) \nearrow \varphi(\cdot, |f(\cdot)|)$ almost everywhere. So, the theorem of monotone convergence gives

$$\begin{aligned}\varrho_\varphi(f) &= \int_A \varphi(y, \lim_{k\to\infty} |f_k(y)|)\, d\mu(y) \\ &= \int_A \lim_{k\to\infty} \varphi(y, |f_k(y)|)\, d\mu(y) \\ &= \lim_{k\to\infty} \int_A \varphi(y, |f_k(y)|)\, d\mu(y) \\ &= \lim_{k\to\infty} \varrho_\varphi(f_k).\end{aligned}$$

To prove (c) assume that $f_k \to f$ almost everywhere, $|f_k| \leqslant |g|$, and $\varrho(\lambda g) < \infty$ for every $\lambda > 0$. Then $|f_k - f| \to 0$ almost everywhere, $|f| \leqslant |g|$ and $|f_k - f| \leqslant 2|g|$. Since $\varrho_\varphi(2\lambda g) < \infty$, we can use the theorem of dominated convergence to conclude that

$$\lim_{k\to\infty} \varrho_\varphi(\lambda|f - f_k|) = \int_A \varphi\Big(y, \lim_{k\to\infty} \lambda|f(y) - f_k(y)|\Big)\, d\mu(y) = 0.$$

Since $\lambda > 0$ was arbitrary, Lemma 2.1.9 implies that $f_k \to f$ in L^φ. $\square$

Let us summarize a few additional properties of L^φ. Properties (a), (b), (c) and (d) of the next theorem are known as *circularity*, *solidity*, *Fatou's lemma (for the norm)*, and the *Fatou property*, respectively.

Theorem 2.3.17. *Let $\varphi \in \Phi(A,\mu)$. Then the following hold.*

(a) $\|f\|_\varphi = \big\| |f| \big\|_\varphi$ *for all $f \in L^\varphi$.*

(b) *If $f \in L^\varphi$, $g \in L^0(A,\mu)$, and $0 \leqslant |g| \leqslant |f|$ μ-almost everywhere, then $g \in L^\varphi$ and $\|g\|_\varphi \leqslant \|f\|_\varphi$.*

(c) *If* $f_k \to f$ *almost everywhere, then* $\|f\|_\varphi \leqslant \liminf_{k\to\infty} \|f_k\|_\varphi$.
(d) *If* $|f_k| \nearrow |f|$ μ*-almost everywhere with* $f_k \in L^\varphi(A,\mu)$ *and* $\sup_k \|f_k\|_\varphi < \infty$, *then* $f \in L^\varphi(A,\mu)$ *and* $\|f_k\|_\varphi \nearrow \|f\|_\varphi$.

Proof. The properties (a) and (b) are obvious. Let us now prove (c). So let $f_k \to f$ μ-almost everywhere. There is nothing to prove for $\liminf_{k\to\infty} \|f_k\|_\varphi = \infty$. Let $\lambda > \liminf_{k\to\infty} \|f_k\|_\varphi$. Then $\|f_k\|_\varphi < \lambda$ for large k. Thus by the unit ball property $\varrho_\varphi(f_k/\lambda) \leqslant 1$ for large k. Now Fatou's lemma for the modular (Lemma 2.3.16) implies $\varrho_\varphi(f/\lambda) \leqslant 1$. So $\|f\|_\varphi \leqslant \lambda$ again by the unit ball property, which implies $\|f\|_\varphi \leqslant \liminf_{k\to\infty} \|f_k\|_\varphi$.

It remains to prove (d). So let $|f_k| \nearrow |f|$ μ-almost everywhere with $\sup_k \|f_k\|_\varphi < \infty$. From (a) and (c) follows $\|f\|_\varphi \leqslant \liminf_{k\to\infty} \|f_k\|_\varphi \leqslant \sup_k \|f_k\|_\varphi < \infty$, which proves $f \in L^\varphi$. On the other hand $|f_k| \nearrow |f|$ and (b) implies that $\|f_k\|_\varphi \nearrow \limsup_{k\to\infty} \|f_k\|_\varphi \leqslant \|f\|_\varphi$. Thus $\lim_{k\to\infty} \|f_k\|_\varphi = \|f\|_\varphi$ and $\|f_k\|_\varphi \nearrow \|f\|_\varphi$. □

2.4 Uniform Convexity

In this section we study sufficient conditions for the uniform convexity of a modular space X_ϱ and the Musielak–Orlicz space L^φ. We first show that the uniform convexity of the Φ-function implies that of the modular; and that the uniform convexity of the semimodular combined with the Δ_2-condition implies the uniform convexity of the norm. The section is concluded by some further properties of uniformly convex modulars. Let us start with the Δ_2-condition of the Φ-function and some implications.

Definition 2.4.1. We say that $\varphi \in \Phi(A,\mu)$ satisfies the Δ_2-*condition* if there exists $K \geqslant 2$ such that

$$\varphi(y, 2t) \leqslant K\varphi(y,t)$$

for all $y \in A$ and all $t \geqslant 0$. The smallest such K is called the Δ_2-*constant of* φ.

Analogously, we say that a semimodular ϱ on X satisfies the Δ_2-*condition* if there exists $K \geqslant 2$ such that $\varrho(2f) \leqslant K\,\varrho(f)$ for all $f \in X_\varrho$. Again, the smallest such K is called the Δ_2-*constant of* ϱ.

If $\varphi \in \Phi(A,\mu)$ satisfy the Δ_2-condition, then ϱ_φ satisfies the Δ_2-condition with the same constant. Moreover, ϱ_φ satisfies the weak Δ_2-condition for modulars, so by Lemma 2.1.11 modular convergence and norm convergence are equivalent; and $E \subset L^\varphi(\Omega,\mu)$ is bounded with respect to the norm if and only if it is bounded with respect to the modular, i.e. $\sup_{f\in E} \|f\| < \infty$ if and only if $\sup_{f\in E} \varrho_\varphi(f) < \infty$.

Corollary 2.1.15 shows that a small norm implies a small modular. The following result shows the reverse implication.

Lemma 2.4.2. *Let ϱ be a semimodular on X that satisfies the Δ_2-condition. Let K be the Δ_2-constant of ϱ. Then for every $\varepsilon > 0$ there exists $\delta = \delta(\varepsilon, K) > 0$ such that $\varrho(f) \leqslant \delta$ implies $\|f\|_\varrho \leqslant \varepsilon$.*

Proof. For $\varepsilon > 0$ choose $j \in \mathbb{N}$ with $2^{-j} \leqslant \varepsilon$. Let $\delta := K^j$ and $\varrho(f) \leqslant \delta$. Then $\varrho(2^j f) \leqslant K^j \varrho(f) \leqslant 1$ and the unit ball property yields $\|f\|_\varrho \leqslant 2^{-j} \leqslant \varepsilon$. □

Lemma 2.4.3. *Let ϱ be a semimodular on X that satisfies the Δ_2-condition with constant K. Then ϱ is a continuous modular and for every $\varepsilon > 0$ there exists $\delta = \delta(\varepsilon, K) > 0$ such that $\varrho(f) \leqslant 1-\varepsilon$ implies $\|f\|_\varrho \leqslant 1-\delta$ for $f \in X_\varrho$.*

Proof. If $\varrho(f) = 0$, then $\varrho(2^m f) \leqslant K^m \varrho(f) = 0$, where K is the Δ_2-constant of φ. This proves $f = 0$, so ϱ is a modular. We already know that ϱ is left-continuous, so it suffices to show $\varrho(x) = \lim_{\lambda \to 1^+} \varrho(\lambda x)$. By monotonicity we have $\varrho(x) \leqslant \liminf_{\lambda \to 1^+} \varrho(\lambda x)$. It follows by convexity of ϱ that

$$\begin{aligned} \varrho(af) &\leqslant (2-a)\varrho(f) + (a-1)\varrho(2f) \leqslant \big((2-a) + K(a-1)\big)\varrho(f) \\ &\leqslant \big(1 + (K-1)(a-1)\big)\varrho(f) \end{aligned}$$

for every $a \in [1, 2]$. Hence $\varrho(x) \geqslant \liminf_{\lambda \to 1^+} \varrho(\lambda x)$, which completes the proof of continuity.

Let $\varepsilon > 0$ and $f \in X_\varrho$ with $\varrho(f) \leqslant 1 - \varepsilon$. Fix $a = a(K, \varepsilon) \in (1, 2)$ such that the right-hand side of the previous inequality is bounded by one. Then $\varrho(af) \leqslant 1$ and the unit ball property implies $\|af\|_\varrho \leqslant 1$. The claim follows with $1 - \delta := \frac{1}{a}$. □

In the previous sections we worked with general $\varphi \in \Phi(A, \mu)$. The corresponding Musielak–Orlicz spaces include the classical spaces L^p with $1 \leqslant p \leqslant \infty$, see Example 2.1.8. Sometimes, however, it is better to work with a subclass of $\Phi(A, \mu)$, called N-functions. These functions will have better properties (N stands for *nice*) but the special cases $p = 1$ and $p = \infty$ are excluded. This corresponds to the experience that also in the classical case the "borderline" cases $p = 1$ and $p = \infty$ are often treated differently.

Definition 2.4.4. A Φ-function φ is said to be an *N-function* if it is continuous and positive and satisfies $\lim_{t\to 0} \frac{\varphi(t)}{t} = 0$ and $\lim_{t\to\infty} \frac{\varphi(t)}{t} = \infty$.

A function $\varphi \in \Phi(A, \mu)$ is said to be a *generalized N-function* if $\varphi(y, \cdot)$ is for every $y \in \Omega$ an N-function.

If φ is a generalized N-function on (A, μ), we write $\varphi \in N(A, \mu)$ for short. If Ω is an open subset of $\mathbb{R}^n$ and μ is the n-dimensional Lebesgue measure we abbreviate $\varphi \in N(\Omega)$.

Definition 2.4.5. A function $\varphi \in N(A, \mu)$ is called *uniformly convex* if for every $\varepsilon > 0$ there exists $\delta > 0$ such that

$$|u - v| \leqslant \varepsilon \max\{u, v\} \quad \text{or} \quad \varphi\Big(y, \frac{u+v}{2}\Big) \leqslant (1-\delta)\frac{\varphi(y,u) + \varphi(y,v)}{2}$$

for all $u, v \geqslant 0$ and every $y \in A$.

Remark 2.4.6. If $\varphi(x,t) = t^q$ with $q \in (1,\infty)$, then φ is uniformly convex. To prove this, we have to show that for $u,v \geqslant 0$ the estimate $|u-v| > \varepsilon \max\{v,u\}$ implies $(\frac{u+v}{2})^q \leqslant (1-\delta(\varepsilon))\frac{1}{2}(u^q+v^q)$ with $\delta(\varepsilon) > 0$ for every $\varepsilon > 0$. Without loss of generality we can assume $\varepsilon \in (0,\frac{1}{2})$. By homogeneity it suffices to consider the case $v = 1$ and $0 \leqslant u \leqslant 1$. So we have to show that $u \in [0,1-\varepsilon)$ implies $(\frac{1+u}{2})^q \leqslant (1-\delta(\varepsilon))\frac{1}{2}(1+u^q)$. Define $f(\tau) := 2^{1-q}\frac{(1+u)^q}{(1+u^q)}$. Then f is continuous on $[0,1]$ and has its maximum at 1. This proves as desired $f(u) \leqslant \delta(\varepsilon)$ for all $u \in [0,1-\varepsilon)$.

It follows by division with q that $\varphi(x,t) = \frac{1}{q}t^q$ with $1 < q < \infty$ is also uniformly convex.

Definition 2.4.5 is formulated for $u,v \geqslant 0$. However, the following lemma shows that this can be relaxed to values in $\mathbb{K}$.

Lemma 2.4.7. *Let $\varphi \in N(A,\mu)$ be uniformly convex. Then for every $\varepsilon_2 > 0$ there exists $\delta_2 > 0$ such that*

$$|a-b| \leqslant \varepsilon_2 \max\{|a|,|b|\} \quad \textit{or} \quad \varphi\Big(y,\Big|\frac{a+b}{2}\Big|\Big) \leqslant (1-\delta_2)\frac{\varphi(y,|a|)+\varphi(y,|b|)}{2}.$$

for all $a,b \in \mathbb{K}$ and every $y \in A$.

Proof. Fix $\varepsilon_2 > 0$. For $\varepsilon := \varepsilon_2/2$ let $\delta > 0$ be as in Definition 2.4.5. Let $|a-b| > \varepsilon_2 \max\{|a|,|b|\}$. If $\big||a|-|b|\big| > \varepsilon \max\{|a|,|b|\}$, then the claim follows by $|a+b| \leqslant |a|+|b|$ and choice of δ with $\delta_2 = \delta$. So assume in the following $||a|-|b|| \leqslant \varepsilon \max\{|a|,|b|\}$. Then

$$|a-b| > \varepsilon_2 \max\{|a|,|b|\} = 2\varepsilon \max\{|a|,|b|\} \geqslant 2||a|-|b||.$$

Therefore,

$$\begin{aligned}\Big|\frac{a+b}{2}\Big|^2 &= \frac{|a|^2}{2} + \frac{|b|^2}{2} - \Big|\frac{a-b}{2}\Big|^2 \\ &\leqslant \frac{|a|^2}{2} + \frac{|b|^2}{2} - \frac{3}{4}\Big|\frac{a-b}{2}\Big|^2 - \Big(\frac{|a|-|b|}{2}\Big)^2 \\ &= \Big(\frac{|a|+|b|}{2}\Big)^2 - \frac{3}{4}\Big|\frac{a-b}{2}\Big|^2.\end{aligned}$$

Since $|a-b| > \varepsilon_2 \max\{|a|,|b|\} \geqslant \varepsilon_2(|a|+|b|)/2$, it follows that

$$\Big|\frac{a+b}{2}\Big|^2 \leqslant \Big(1-\frac{3\varepsilon_2^2}{16}\Big)\Big(\frac{|a|+|b|}{2}\Big)^2.$$

Let $\delta_2 := 1-\sqrt{1-\frac{3\varepsilon_2^2}{16}} > 0$, then $|\frac{a+b}{2}| \leqslant (1-\delta_2)\frac{|a|+|b|}{2}$. This, (2.1.5) and the convexity of φ imply

$$\varphi\Big(y, \Big|\frac{a+b}{2}\Big|\Big) \leqslant (1-\delta_2)\varphi\Big(y, \frac{|a|+|b|}{2}\Big) \leqslant (1-\delta_2)\frac{\varphi(y,|a|)+\varphi(y,|b|)}{2}. \qquad \square$$

Remark 2.4.8. If $u, v \in \mathbb{K}$ satisfies $|a-b| \leqslant \varepsilon_2 \max\{|a|,|b|\}$ with $\varepsilon_2 \in (0,1)$, then $\frac{|a-b|}{2} \leqslant \varepsilon_2 \frac{|a|+|b|}{2}$ and by the convexity of φ follows

$$\varphi\Big(y, \frac{|a-b|}{2}\Big) \leqslant \varepsilon_2 \frac{\varphi(y,|a|)+\varphi(y,|b|)}{2}. \tag{2.4.9}$$

Therefore, we can replace the first alternative in Lemma 2.4.7 by the weaker version (2.4.9).

We need the following concept of uniform convexity for the semimodular.

Definition 2.4.10. A semimodular ϱ on X is called *uniformly convex* if for every $\varepsilon > 0$ there exists $\delta > 0$ such that

$$\varrho\bigg(\frac{f-g}{2}\bigg) \leqslant \varepsilon \frac{\varrho(f)+\varrho(g)}{2} \quad \text{or} \quad \varrho\bigg(\frac{f+g}{2}\bigg) \leqslant (1-\delta)\frac{\varrho(f)+\varrho(g)}{2}$$

for all $f, g \in X_\varrho$.

Theorem 2.4.11. *Let $\varphi \in N(A,\mu)$ be uniformly convex. Then ϱ_φ is uniformly convex.*

Proof. Let $\varepsilon_2, \delta_2 > 0$ be as in Lemma 2.4.7 and let $\varepsilon := 2\,\varepsilon_2$. There is nothing to show if $\varrho_\varphi(f) = \infty$ or $\varrho_\varphi(g) = \infty$. So in the following let $\varrho_\varphi(f), \varrho_\varphi(g) < \infty$, which implies by convexity $\varrho(\frac{f+g}{2}), \varrho(\frac{f-g}{2}) < \infty$.

Assume that $\varrho_\varphi(\frac{f-g}{2}) > \varepsilon \frac{\varrho_\varphi(f)+\varrho_\varphi(g)}{2}$. We show that

$$\varrho_\varphi\Big(\frac{f+g}{2}\Big) \leqslant \Big(1 - \frac{\delta_2\varepsilon}{2}\Big)\frac{\varrho_\varphi(f)+\varrho_\varphi(g)}{2},$$

which proves that ϱ_φ is uniformly convex. Define

$$E := \Big\{y \in A \colon |f(y)-g(y)| > \frac{\varepsilon}{2}\max\big\{|f(y)|,|g(y)|\big\}\Big\}.$$

It follows from Remark 2.4.8 that (2.4.9) holds for almost all $y \in A \setminus E$. In particular,

$$\varrho_\varphi\bigg(\chi_{A\setminus E}\frac{f-g}{2}\bigg) \leqslant \frac{\varepsilon}{2}\frac{\varrho_\varphi(\chi_{A\setminus E}f)+\varrho_\varphi(\chi_{A\setminus E}g)}{2} \leqslant \frac{\varepsilon}{2}\frac{\varrho_\varphi(f)+\varrho_\varphi(g)}{2}.$$

This and $\varrho_\varphi(\frac{f-g}{2}) > \varepsilon\frac{\varrho_\varphi(f)+\varrho_\varphi(g)}{2}$ imply

$$\varrho_\varphi\Big(\chi_E\frac{f-g}{2}\Big) = \varrho_\varphi\Big(\frac{f-g}{2}\Big) - \varrho_\varphi\Big(\chi_{A\setminus E}\frac{f-g}{2}\Big) > \frac{\varepsilon}{2}\frac{\varrho_\varphi(f)+\varrho_\varphi(g)}{2}. \tag{2.4.12}$$

On the other hand it follows by the definition of E and the choice of δ_2 in Lemma 2.4.7 that

$$\varrho_\varphi\bigg(\chi_E\frac{f+g}{2}\bigg) \leqslant (1-\delta_2)\frac{\varrho_\varphi(\chi_E f)+\varrho_\varphi(\chi_E g)}{2}. \tag{2.4.13}$$

We estimate

$$\frac{\varrho_\varphi(f)+\varrho_\varphi(g)}{2} - \varrho_\varphi\Big(\frac{f+g}{2}\Big) \geqslant \frac{\varrho_\varphi(\chi_E f)+\varrho_\varphi(\chi_E g)}{2} - \varrho_\varphi\Big(\chi_E\frac{f+g}{2}\Big),$$

where we have split the domain of the involved integrals into the sets E and $A\setminus E$ and have used $\frac{1}{2}(\varphi(f)+\varphi(g)) - \varphi(\frac{f+g}{2}) \geqslant 0$ on $A\setminus E$. This, (2.4.13), the convexity and (2.4.12) imply

$$\begin{aligned}\frac{\varrho_\varphi(f)+\varrho_\varphi(g)}{2} - \varrho_\varphi\Big(\frac{f+g}{2}\Big) &\geqslant \delta_2\frac{\varrho_\varphi(\chi_E f)+\varrho_\varphi(\chi_E g)}{2}\\ &\geqslant \delta_2\varrho_\varphi\Big(\chi_E\frac{f-g}{2}\Big)\\ &\geqslant \frac{\delta_2\varepsilon}{2}\frac{\varrho_\varphi(f)+\varrho_\varphi(g)}{2}.\end{aligned}$$

□

The question arises if uniform convexity of the semimodular ϱ implies the uniform convexity of X_ϱ. This turns out to be true under the Δ_2-condition.

Theorem 2.4.14. *Let ϱ be a uniformly convex semimodular on X that satisfies the Δ_2-condition. Then the norm $\|\cdot\|_\varrho$ on X_ϱ is uniformly convex. Hence, X_ϱ is uniformly convex.*

Proof. Fix $\varepsilon > 0$. Let $x, y \in X$ with $\|x\|_\varrho, \|y\|_\varrho \leqslant 1$ and $\|x-y\|_\varrho > \varepsilon$. Then $\|\frac{x-y}{2}\| > \frac{\varepsilon}{2}$ and by Lemma 2.4.2 there exists $\alpha = \alpha(\varepsilon) > 0$ such that $\varrho(\frac{x-y}{2}) > \alpha$. By the unit ball property we have $\varrho(x), \varrho(y) \leqslant 1$, so $\varrho(\frac{x-y}{2}) > \alpha\frac{\varrho(x)+\varrho(y)}{2}$. Since ϱ is uniformly convex, there exists $\beta = \beta(\alpha) > 0$ such that $\varrho(\frac{x+y}{2}) \leqslant (1-\beta)\frac{\varrho(x)+\varrho(y)}{2} \leqslant 1-\beta$. Now Lemma 2.4.3 implies the existence of $\delta = \delta(K,\beta) > 0$ with $\|\frac{x+y}{2}\|_\varrho \leqslant 1-\delta$. This proves the uniform convexity of $\|\cdot\|_\varrho$. □

Remark 2.4.15. If $\varphi \in N(A,\mu)$ is uniformly convex and satisfies the Δ_2-condition, then it follows by the combination of Theorems 2.4.11 and 2.4.14 that the norm $\|\cdot\|_\varphi$ of $L^\varphi(A,\mu)$ is uniformly convex. Hence, $L^\varphi(A,\mu)$ is also uniformly convex.

We will later need that the sum of uniformly convex semimodulars is again uniformly convex.

Lemma 2.4.16. *If ϱ_1, ϱ_2 are uniformly convex semimodulars on X, then $\varrho := \varrho_1 + \varrho_2$ is uniformly convex.*

Proof. If $\varepsilon > 0$, then there exists $\delta > 0$ such that

$$\varrho_j\Big(\frac{f-g}{2}\Big) \leqslant \varepsilon\,\frac{\varrho_j(f)+\varrho_j(g)}{2} \quad \text{or} \quad \varrho_j\Big(\frac{f+g}{2}\Big) \leqslant (1-\delta)\frac{\varrho_j(f)+\varrho_j(g)}{2}$$

for $j = 1, 2$. We show that

$$\varrho\Big(\frac{f-g}{2}\Big) \leqslant 2\,\varepsilon\,\frac{\varrho(f)+\varrho(g)}{2} \quad \text{or} \quad \varrho\Big(\frac{f+g}{2}\Big) \leqslant (1-\delta\varepsilon)\frac{\varrho(f)+\varrho(g)}{2},$$

since this proves the uniform convexity of ϱ. Fix f and g and assume that $\varrho(\frac{f-g}{2}) > 2\,\varepsilon\,\frac{\varrho(f)+\varrho(g)}{2}$. Without loss of generality, we can assume that $\varrho_1(\frac{f-g}{2}) \geqslant \varrho_2(\frac{f-g}{2})$ for this specific choice of f and g. Therefore, $\varrho_1(\frac{f-g}{2}) > \varepsilon\,\frac{\varrho(f)+\varrho(g)}{2} \geqslant \varepsilon\,\frac{\varrho_1(f)+\varrho_1(g)}{2}$. So the choice of δ implies

$$\varrho_1\Big(\frac{f+g}{2}\Big) \leqslant (1-\delta)\frac{\varrho_1(f)+\varrho_1(g)}{2}.$$

Taking into account the convexity of ϱ_2, we obtain

$$\varrho\Big(\frac{f+g}{2}\Big) \leqslant \frac{\varrho(f)+\varrho(g)}{2} - \delta\,\frac{\varrho_1(f)+\varrho_1(g)}{2}.$$

Since $\frac{\varrho_1(f)+\varrho_1(g)}{2} \geqslant \varrho_1(\frac{f-g}{2}) > \varepsilon\,\frac{\varrho(f)+\varrho(g)}{2}$, this implies

$$\varrho\Big(\frac{f+g}{2}\Big) \leqslant (1-\delta\varepsilon)\frac{\varrho(f)+\varrho(g)}{2}. \qquad \square$$

It is well known that on uniformly convex spaces weak convergence $x_k \rightharpoonup x$ combined with convergence of the norms $\|x_k\| \to \|x\|$ implies strong convergence $x_k \to x$. The following lemma is in this spirit.

Lemma 2.4.17. *Let ϱ be a uniformly convex semimodular on X. Let $x_k, x \in X_\varrho$ such that $x_k \rightharpoonup x$, $\varrho(x_k) \to \varrho(x)$ and $\varrho(x) < \infty$. Then*

$$\varrho\Big(\frac{x_k - x}{2}\Big) \to 0.$$

Proof. We proceed by contradiction. Assume that the claim is wrong and there exists $\varepsilon > 0$ and a subsequence x_{k_j} such that

$$\varrho\Big(\frac{x_{k_j} - x}{2}\Big) > \varepsilon \tag{2.4.18}$$

for all $j \in \mathbb{N}$. Since ϱ is uniformly continuous, there exists $\delta > 0$ such that

$$\varrho\Big(\frac{x_k - x}{2}\Big) \leqslant \varepsilon \quad \text{or} \quad \varrho\Big(\frac{x_k + x}{2}\Big) \leqslant (1-\delta)\,\frac{\varrho(x_k) + \varrho(x)}{2}.$$

In particular, our subsequence always satisfies the second alternative. Together with $\frac{1}{2}(x_k + x) \rightharpoonup x$, the weak lower semicontinuity of ϱ (Theorem 2.2.8) and $\varrho(x_k) \to \varrho(x)$ implies that

$$\varrho(x) \leqslant \liminf_{j\to\infty} \varrho\Big(\frac{x_{k_j} + x}{2}\Big) \leqslant (1-\delta)\liminf_{j\to\infty} \frac{\varrho(x_{k_j}) + \varrho(x)}{2} = (1-\delta)\varrho(x).$$

Using $\varrho(x) < \infty$ we get $\varrho(x) = 0$. It follows by convexity and $\varrho(x_k) \to \varrho(x)$ that

$$\varrho\Big(\frac{x_k - x}{2}\Big) \leqslant \frac{\varrho(x_k) + \varrho(x)}{2} \to \varrho(x) = 0$$

for $n \to \infty$. This contradicts (2.4.18). □

Remark 2.4.19. If ϱ satisfies the (weak) Δ_2-condition, then under the conditions of the previous lemma, $\varrho(\lambda(x_k - x)) \to 0$ for all $\lambda > 0$ and $x_k \to x$ in X_ϱ by Lemma 2.1.11.

2.5 Separability

We next prove basic properties of Musielak–Orlicz spaces that require some additional structure. Since these properties do not even hold for the full range $p \in [1, \infty]$ of classical Lebesgue spaces, it is clear that some restrictions are necessary. In this section we consider separability.

We first define some function classes related to L^φ. The set E^φ of finite elements will be later important in the approximability by simple functions, see Theorem 2.5.9.

Definition 2.5.1. Let $\varphi \in \Phi(A, \mu)$. The set

$$L^\varphi_{OC} := L^\varphi_{OC}(A, \mu) := \{f \in L^\varphi \colon \varrho_\varphi(f) < \infty\} \tag{2.5.2}$$

is called the *Musielak–Orlicz class.* Let

$$E^\varphi := E^\varphi(A,\mu) := \{f \in L^\varphi\colon \varrho_\varphi(\lambda f) < \infty \text{ for all } \lambda > 0\}. \tag{2.5.3}$$

The elements of $E^\varphi(A,\mu)$ are called *finite.*

Let us start with a few examples:

(a) Let $\varphi(y,t) = t^p$ with $1 \leqslant p < \infty$. Then $E^\varphi = L^\varphi_{OC} = L^\varphi = L^p$.

(b) Let $\varphi(y,t) = \infty \cdot \chi_{(1,\infty)}(t)$. Then

$$\begin{aligned} E^\varphi &= \{0\}, \\ L^\varphi_{OC} &= \{f\colon |f| \leqslant 1 \text{ almost everywhere}\}, \\ L^\varphi &= L^\infty. \end{aligned}$$

(c) Let $\varphi(y,t) = \exp(t) - 1$ and $\Omega = (0,1)$. Then $\varphi \in \Phi(\Omega)$ is positive and continuous but $E^\varphi \neq L^\varphi_{OC} \neq L^\varphi$. Indeed, if $f := \sum_{k=1}^\infty \frac{k}{2}\chi_{(2^{-k},2^{-k+1})}$, then $f \in L^\varphi_{OC} \setminus E^\varphi$ and $2f \in L^\varphi \setminus L^\varphi_{OC}$.

By definition of E^φ, L^φ_{OC}, and L^φ it is clear that $E^\varphi \subset L^\varphi_{OC} \subset L^\varphi$. Moreover, by convexity of φ the set L^φ_{OC} is convex and the sets E^φ and L^φ are linear subspaces of L^0. There is a special relation of E^φ and L^φ to L^φ_{OC}: E^φ is the biggest vector space in L^φ_{OC} and L^φ is the smallest vector space in L^0 containing L^φ_{OC}.

In some cases the inclusions $E^\varphi \subset L^\varphi_{OC} \subset L^\varphi$ are strict and in other cases equality holds. In fact, it is easily seen that $E^\varphi = L^\varphi_{OC} = L^\varphi$ is equivalent to the implication $f \in L^\varphi_{OC} \Rightarrow 2f \in L^\varphi_{OC}$. The Δ_2-condition (see Definition 2.4.1) implies that $\varrho_\varphi(2^m f) \leqslant K^m \varrho_\varphi(f)$, where K is the Δ_2-constant, from which we conclude that

$$E^\varphi(A,\mu) = L^\varphi_{OC}(A,\mu) = L^\varphi(A,\mu).$$

Remark 2.5.4. The set E^φ is a closed subset of L^φ. Indeed, let $f_k \to f$ in L^φ with $f_k \in E^\varphi$. For $\lambda > 0$ we have $\varrho_\varphi(2\lambda(f_k - f)) \to 0$ as $k \to \infty$. In particular, $\varrho_\varphi(2\lambda(f_{k_\lambda} - f)) \leqslant 1$ for some k_λ. By convexity $\varrho_\varphi(\lambda f) \leqslant \frac{1}{2}\varrho_\varphi\big(2\lambda(f_{k_\lambda} - f)\big) + \frac{1}{2}\varrho_\varphi(2\lambda f_{k_\lambda}) \leqslant \frac{1}{2} + \frac{1}{2}\varrho_\varphi(2\lambda f_{k_\lambda}) < \infty$, which shows that $f \in E^\varphi$.

In the approximation of measurable functions it is very useful to work with simple functions. To be able to approximate a function f by simple functions we have to assume an additional property of φ:

Definition 2.5.5. A function $\varphi \in \Phi(A,\mu)$ is called *locally integrable* on A if $\varrho_\varphi(t\chi_E) < \infty$ for all $t \geqslant 0$ and all μ-measurable $E \subset A$ with $\mu(E) < \infty$.

Note that local integrability in the previous definition differs from the one used in L^1_{loc}, where we assume integrability over compact subsets.

If $\varphi \in \Phi(A,\mu)$ is locally integrable, then the set of simple functions $S(A,\mu)$ is contained in E^φ. Actually, the property $S(A,\mu) \subset E^\varphi$ is equivalent to the local integrability of φ.

Example 2.5.6. Let $\varphi \in \Phi(A,\mu)$ with $\varphi(y,t) = \psi(t)$ where ψ is a continuous Φ-function. Then φ is locally integrable. Indeed, due to the continuity we know that $t \mapsto \psi(t)$ is everywhere finite on $[0,\infty)$. Therefore, $\varrho_\varphi(t\chi_E) = \mu(E)\psi(t) < \infty$ for all $t \geqslant 0$ and $\mu(E) < \infty$.

Proposition 2.5.7. *Let $\varphi \in \Phi(A,\mu)$ be locally integrable. Then for every $\lambda > 0$ and $\varepsilon > 0$ there exists $\delta > 0$ such that $\mu(E) \leqslant \delta$ implies $\varrho_\varphi(\lambda\chi_E) \leqslant \varepsilon$ and $\|\chi_E\|_\varphi \leqslant \frac{1}{\lambda}$.*

Proof. We begin with the proof of $\varrho_\varphi(\lambda\chi_E) \leqslant \varepsilon$ by contradiction. Assume to the contrary that there exist $\lambda > 0$ and $\varepsilon > 0$ and a sequence (E_k) such that $\mu(E_k) \leqslant 2^{-k}$ and $\varrho_\varphi(\lambda\chi_{E_k}) > \varepsilon$. Let $G_k := \bigcup_{m=k}^\infty E_m$, and note that $\mu(G_k) \leqslant \sum_{m=k}^\infty 2^{-m} = 2^{1-k} \to 0$ as $k \to \infty$. Since φ is locally integrable and $\mu(G_1) \leqslant 1$, we have $\varrho_\varphi(\lambda\chi_{G_1}) < \infty$. Moreover, $\lambda\chi_{G_k} \leqslant \lambda\chi_{G_1}$ and $\lambda\chi_{G_k} \to 0$ almost everywhere. Thus, we conclude by dominated convergence that $\varrho_\varphi(\lambda\chi_{G_k}) \to 0$. This contradicts $\varrho_\varphi(\lambda\chi_{G_k}) \geqslant \varrho_\varphi(\lambda\chi_{E_k}) > \varepsilon$ for every k.

The claim $\|\chi_E\|_\varphi \leqslant \frac{1}{\lambda}$ follows from $\varrho_\varphi(\lambda\chi_E) \leqslant \varepsilon$ by the choice $\varepsilon = 1$ and the unit ball property. □

Remark 2.5.8. If $f \in L^\varphi$ has the property that $\|\chi_{E_k} f\|_\varphi \to 0$ if $E_k \searrow \emptyset$, then we say that f *has absolutely continuous norm*. If follows easily by dominated convergence (Lemma 2.3.16) that every $f \in E^\varphi$ has absolutely continuous norm.

Theorem 2.5.9. *Let $\varphi \in \Phi(A,\mu)$ be locally integrable and let $S := S(A,\mu)$ be the set of simple functions. Then $\overline{S}^{\|\cdot\|_\varphi} = E^\varphi(A,\mu)$.*

Proof. The local integrability implies that $S \subset E^\varphi$. Since E^φ is closed by Remark 2.5.4, it suffices to show that every $f \in E^\varphi$ is in the closure of S. Let $f \in E^\varphi$ with $f \geqslant 0$. Since $f \in L^0(A)$, there exist $f_k \in S$ with $0 \leqslant f_k \nearrow f$ almost everywhere. So $f_k \to f$ in L^φ by the theorem of dominated convergence (Lemma 2.3.16). Thus, f is in the closure of S. If we drop the assumption $f \geqslant 0$, then we split x into positive and negative parts (real and imaginary parts) which belong again to E^φ. □

We now investigate the problem of separability of E^φ. Let (A,Σ,μ) be a σ-finite, complete measure space. Here, we need the notion of separable measures: recall that a measure μ is called *separable* if there exists a sequence $(E_k) \subset \Sigma$ with the following properties:

(a) $\mu(E_k) < \infty$ for all $k \in \mathbb{N}$.
(b) For every $E \in \Sigma$ with $\mu(E) < \infty$ and every $\varepsilon > 0$ there exists an index k such that $\mu(E \triangle E_k) < \varepsilon$, where $\triangle$ denotes the symmetric difference defined through $E \triangle E_k := (E \setminus E_k) \cup (E_k \setminus E)$.

For instance the Lebesgue measure on $\mathbb{R}^n$ and the counting measure on $\mathbb{Z}^n$ are separable. Recall that a Banach space is separable if it contains a dense, countable subset.

Theorem 2.5.10. *Let $\varphi \in \Phi(A,\mu)$ be locally integrable and let μ be separable. Then $E^{\varphi}(A,\mu)$ is separable.*

Proof. Let S_0 be the set of all simple functions g of the form $g = \sum_{i=1}^{k} a_i \chi_{E_i}$ with $a_i \in \mathbb{Q}$ and E_i is as in the definition of a separable measure. By Theorem 2.5.9 it suffices to prove that S_0 is dense in S. Let $f \in S$. Then we can write f in the form $f = \sum_{i=1}^{k} b_i \chi_{B_i}$ with $b_i \in \mathbb{R}$, $B_i \in \Sigma$ pairwise disjoint and $\mu(B_i) < \infty$. Let $\lambda > 0$ be arbitrary and define $b := \max_{1\leqslant i\leqslant k} |b_i|$. Since φ is locally integrable, we know by Proposition 2.5.7 that the integral of $y \mapsto \varphi(y, 4k\lambda b)$ is small over small sets. Hence, by the separability of μ we find measurable sets $E_{j_1}, \dots, E_{j_k}$ of finite measure such that

$$\int\limits_{E_{j_i} \triangle B_i} \varphi(y, 4k\lambda b)\, d\mu(y) \leqslant 1.$$

Let $B := \bigcup_{i=1}^{k} B_i$. Then $\int_B \varphi(y, 2\lambda\eta)\, d\mu(y) \to 0$ for $\eta \to 0$, since $\mu(B) < \infty$ and φ is locally integrable. Let $\delta > 0$ be such that $\int_B \varphi(y, 2\lambda\delta)\, d\mu(y) \leqslant 1$. Choose rational numbers $a_1, \dots, a_k$ such that $|b_i - a_i| < \delta$ and $|a_i| \leqslant 2b$ for $i = 1, \dots, k$. Let $g := \sum_{i=1}^{k} a_i\, \chi_{E_{j_i}}$. Then

$$\begin{aligned} |f-g| &= \left|\sum_{i=1}^{k} (b_i - a_i)\chi_{B_i}\right| + \left|\sum_{i=1}^{k} a_i \big(\chi_{B_i} - \chi_{E_{j_i}}\big)\right| \\ &\leqslant \sum_{i=1}^{k} |b_i - a_i| \chi_{B_i} + \sum_{i=1}^{k} |a_i|\, \chi_{E_{j_i} \triangle B_i} \\ &\leqslant \delta\, \chi_B + \sum_{i=1}^{k} 2b\, \chi_{E_i \triangle B_i}. \end{aligned}$$

Hence, by the previous estimate and convexity,

$$\begin{aligned} \varrho_{\varphi}\big(\lambda(f-g)\big) &\leqslant \frac{1}{2}\varrho_{\varphi}\big(2\lambda\delta\, \chi_B\big) + \frac{1}{2k}\sum_{i=1}^{k} \varrho_{\varphi}\left(4k\lambda b\, \chi_{E_i \triangle B_i}\right) \\ &= \frac{1}{2}\int\limits_{B} \varphi(y, 2\lambda\delta)\, d\mu(y) + \frac{1}{2k}\sum_{i=1}^{k} \int\limits_{E_i \triangle B_i} \varphi(y, 4k\lambda b)\, d\mu(y). \end{aligned}$$

The right-hand side of the previous estimate is at most 1 and so $\|f-g\|_{\varphi} \leqslant \frac{1}{\lambda}$ by the unit ball property. Since $\lambda > 0$ was arbitrary, this completes the proof. $\square$

2.6 Conjugate Φ-Functions

In this section we specialize the results from Sect. 2.2 Conjugate modulars and dual semimodular spaces to Φ-functions and generalized Φ-functions. Apart from the general results, we are also able to prove stronger results in this special case.

By Lemma 2.3.2 we know that every Φ-function defines (by even extension) a semimodular on $\mathbb{R}$. This motivates to transfer the definition of a conjugate semimodular in a point-wise sense to generalized Φ-functions:

Definition 2.6.1. Let $\varphi \in \Phi(A, \mu)$. Then for any $y \in A$ we denote by $\varphi^*(y, \cdot)$ the *conjugate function* of $\varphi(y, \cdot)$ which is defined by

$$\varphi^*(y, u) = \sup_{t \geqslant 0} \big(tu - \varphi(y, t)\big)$$

for all $u \geqslant 0$ and $y \in \Omega$.

This definition applies in particular in the case when φ is a (non-generalized) Φ-function, in which case

$$\varphi^*(u) = \sup_{t \geqslant 0} \big(tu - \varphi(t)\big)$$

concurs with the *Legendre transformation of* φ. By definition of φ^*,

$$tu \leqslant \varphi(t) + \varphi^*(u) \tag{2.6.2}$$

for every $t, u \geqslant 0$. This inequality is called *Young's inequality.* If φ is a Φ-function and $\varrho(t) := \varphi(|t|)$ is its even extension to $\mathbb{R}$, then $\varrho^*(t) = \varphi^*(|t|)$ for all $t \in \mathbb{R}$.

As a special case of Theorem 2.2.6 we have

Corollary 2.6.3. *Let $\varphi \in \Phi(A, \mu)$. Then $(\varphi^*)^* = \varphi$. In particular,*

$$\varphi(y, t) = \sup_{u \geqslant 0} \big(tu - \varphi^*(y, u)\big)$$

for all $y \in \Omega$ and all $t \geqslant 0$.

Lemma 2.6.4. *Let φ, ψ be Φ-functions.*

(a) *The estimate $\varphi(t) \leqslant \psi(t)$ holds for all $t \geqslant 0$ if and only if $\psi^*(u) \leqslant \varphi^*(u)$ for all $u \geqslant 0$.*
(b) *Let $a, b > 0$. If $\psi(t) = a\varphi(bt)$ for all $t \geqslant 0$, then $\psi^*(u) = a\varphi^*(\frac{u}{ab})$ for all $u \geqslant 0$.*

Proof. We begin with the proof of (a). Let $\varphi(t) \leqslant \psi(t)$ for all $t \geqslant 0$. Then

$$\psi^*(u) = \sup_{t\geqslant 0} \big(tu - \psi(t)\big) \leqslant \sup_{t\geqslant 0} \big(tu - \varphi(t)\big) = \varphi^*(u)$$

for all $u \geqslant 0$. The reverse claim follows using $\varphi^{**} = \varphi$ and $\psi^{**} = \psi$ by Corollary 2.6.3. Let us now prove (b). Let $a, b > 0$ and $\psi(t) = a\varphi(bt)$ for all $t \geqslant 0$. Then

$$\begin{aligned}\psi^*(u) &= \sup_{t\geqslant 0} \big(tu - \psi(t)\big) = \sup_{t\geqslant 0} \big(tu - a\varphi(bt)\big) = \sup_{t\geqslant 0} a\Big(t\frac{u}{ab} - \varphi(t)\Big)\\ &= a\psi^*\Big(\frac{u}{ab}\Big)\end{aligned}$$

for all $u \geqslant 0$. □

The following result is the generalization of the classical Hölder inequality $\int |f||g|\,d\mu \leqslant \|f\|_q \|g\|_{q'}$ to the Musielak–Orlicz spaces. The extra constant 2 cannot be omitted.

Lemma 2.6.5 (Hölder's inequality). *Let $\varphi \in \Phi(A,\mu)$. Then*

$$\int_A |f|\,|g|\,d\mu(y) \leqslant 2\|f\|_\varphi\,\|g\|_{\varphi^*}$$

for all $f \in L^\varphi(A,\mu)$ and $g \in L^{\varphi^*}(A,\mu)$.

Proof. Let $f \in L^\varphi$ and $g \in L^{\varphi^*}$. The claim is obvious for $f = 0$ or $g = 0$, so we can assume $f \neq 0$ and $g \neq 0$. Due to the unit ball property, $\varrho_\varphi(f/\|f\|_\varphi) \leqslant 1$ and $\varrho_{\varphi^*}(g/\|g\|_{\varphi^*}) \leqslant 1$. Thus, using Young's inequality (2.6.2), we obtain

$$\begin{aligned}\int_A \frac{|f(y)|}{\|f\|_\varphi}\,\frac{|g(y)|}{\|g\|_{\varphi^*}}\,d\mu(y) &\leqslant \int_A \varphi\Big(y, \frac{|f(y)|}{\|f\|_\varphi}\Big) + \varphi^*\Big(y, \frac{|g(y)|}{\|g\|_{\varphi^*}}\Big)\,d\mu(y)\\ &= \varrho_\varphi(f/\|f\|_\varphi) + \varrho_{\varphi^*}(g/\|g\|_{\varphi^*})\\ &\leqslant 2.\end{aligned}$$

Multiplying through by $\|f\|_\varphi \|g\|_{\varphi^*}$ yields the claim. □

Let us recall the definitions of N-function and generalized N-function from Definition 2.4.4. A Φ-function φ is said to be an N-function if it is continuous and positive and satisfies $\lim_{t\to 0} \frac{\varphi(t)}{t} = 0$ and $\lim_{t\to\infty} \frac{\varphi(t)}{t} = \infty$. A function $\varphi \in \Phi(A,\mu)$ is said to be a generalized N-function if $\varphi(y,\cdot)$ is for every $y \in \Omega$ an N-function.

Note that by continuity N-functions only take values in $[0,\infty)$. Let $\varphi \in N(A,\mu)$ be an N-function. As was noted in (2.3.3), the function has a right-derivative, denoted by $\varphi'(y,\cdot)$, and

$$\varphi(y,t) = \int_0^t \varphi'(y,\tau)\,d\tau$$

for all $y \in A$ and all $t \geqslant 0$. The right-derivative $\varphi'(y,\cdot)$ is non-decreasing and right-continuous.

Lemma 2.6.6. *Let φ be an N-function. Then*

$$\frac{t}{2}\varphi'\Big(\frac{t}{2}\Big) \leqslant \varphi(t) \leqslant t\varphi'(t)$$

for all $t \geqslant 0$

Proof. Using the monotonicity of φ' we get

$$\varphi(t) = \int_0^t \varphi'(\tau)\,d\tau \leqslant \int_0^t \varphi'(t)\,d\tau = t\varphi'(t),$$
$$\varphi(t) = \int_0^t \varphi'(\tau)\,d\tau \geqslant \int_{t/2}^t \varphi'(t/2)\,d\tau = \frac{t}{2}\varphi'\Big(\frac{t}{2}\Big)$$

for all $t \geqslant 0$. □

Remark 2.6.7. If φ is a generalized N-function, which satisfies the Δ_2-condition (Definition 2.4.1), then Lemma 2.6.6 implies $\varphi(y,t) \approx \varphi'(y,t)t$ uniformly in $y \in A$ and $t \geqslant 0$.

Let $\varphi \in N(A,\mu)$. Then we already know that $\varphi'(y,\cdot)$ is for any $y \in A$ non-decreasing right-continuous, $\varphi'(y,0) = 0$, and $\lim_{t\to\infty} \varphi'(y,t) = \infty$. Define

$$b(y,u) := \inf\,\{t \geqslant 0 \colon \varphi'(y,t) > u\}.$$

Then $b(y,\cdot)$ has the same properties, i.e. $b(y,\cdot)$ is for any $y \in A$ non-decreasing, right-continuous, $b(y,0) = 0$, and $\lim_{t\to\infty} b(y,t) = \infty$. The function $b(y,\cdot)$ is the *right-continuous inverse* of $\varphi'(y,\cdot)$ and we therefore denote it by $(\varphi')^{-1}(y,u)$. It is easy to see that the right-continuous inverse of $(\varphi')^{-1}$ is again φ', i.e. $((\varphi')^{-1})^{-1} = \varphi'$. The function $(\varphi')^{-1}$ is important, since we can use it to represent the conjugate function φ^*.

Theorem 2.6.8. *If $\varphi \in N(A,\mu)$, then $\varphi^* \in N(A,\mu)$ and $(\varphi^*)' = (\varphi')^{-1}$. In particular,*

$$\varphi^*(y,t) = \int_0^t (\varphi')^{-1}(y,\tau)\,d\tau$$

for all $y \in A$ and $t \geqslant 0$.

Proof. It suffices to prove the claim point-wise, and thus we may assume without loss of generality that $\varphi(y,t)$ is independent of y, i.e. an N-function.

It is easy to see that φ' is non-decreasing, right-continuous and satisfies $(\varphi')^{-1}(0)=0$, $(\varphi')^{-1}(t)>0$ for $t>0$, and $\lim_{t\to\infty}(\varphi')^{-1}(t)=\infty$. Thus,

$$\psi(t) := \int\limits_0^t (\varphi')^{-1}(\tau)\,d\tau$$

for $t \geqslant 0$ defines an N-function. In particular, φ and ψ are finite.

Note that $\sigma < \varphi'(\tau)$ is equivalent to $(\varphi')^{-1}(\sigma) < \tau$. Hence, the sets

$$\begin{gathered}\big\{(\tau,\sigma)\in[0,\infty)\times[0,\infty)\colon \sigma<\varphi'(\tau)\big\}\\ \big\{(\tau,\sigma)\in[0,\infty)\times[0,\infty)\colon (\varphi')^{-1}(\sigma)\geqslant\tau\big\}\end{gathered}$$

are complementary with respect to $[0,\infty)\times[0,\infty)$. Therefore, we can estimate with the help of the theorem of Fubini

$$\begin{aligned}0 \leqslant tu &= \int\limits_0^t\int\limits_0^u d\sigma\,d\tau\\ &= \iint\limits_{\{0\leqslant\tau\leqslant t,\sigma\leqslant u\colon 0\leqslant\sigma<\varphi'(\tau)\}} d\sigma\,d\tau + \iint\limits_{\{0\leqslant\tau\leqslant t,0\leqslant\sigma\leqslant u\colon (\varphi')^{-1}(\sigma)\geqslant\tau\}} d\sigma\,d\tau\\ &= \int\limits_0^t \int\limits_0^{\min\{u,\varphi'(\tau)\}} d\sigma\,d\tau + \int\limits_0^u \int\limits_0^{\min\{t,(\varphi')^{-1}(\sigma)\}} d\tau\,d\sigma\\ &\leqslant \int\limits_0^t \varphi'(\tau)\,d\tau + \int\limits_0^u (\varphi')^{-1}(\sigma)\,d\sigma\\ &= \varphi(t)+\psi(u).\end{aligned}$$

If $u=\varphi'(t)$ or $t=(\varphi')^{-1}(u)$, then $\min\{u,\varphi'(\tau)\}=\varphi'(\tau)$ and $\min\{t,(\varphi')^{-1}(\sigma)\}=(\varphi')^{-1}(\sigma)$ in the integrals of the third line. So in this case we have equality in the penultimate step. Since $\varphi^*(u)=\sup_t(ut-\varphi(t))$ it follows that $\varphi^*=\psi$. □

Remark 2.6.9. Let φ be an N-function. Then it follows from the previous proof that the right-derivative $(\varphi^*)'$ of φ^* satisfies $(\varphi^*)'=(\varphi')^{-1}$ for all $t\geqslant 0$. Young's inequality $tu\leqslant\varphi(t)+\varphi^*(u)$ holds with equality if $u=\varphi'(t)$ or $t=(\varphi')^{-1}(u)$.

Theorem 2.6.8 enables us to calculate the conjugate function of N-functions. Let us present three examples:

(a) Let $\varphi(t) = e^t - t - 1$. Then $\varphi'(t) = e^t - 1$ and $(\varphi^*)'(u) = (\varphi')^{-1}(u) = \log(1+u)$. Integration over u implies $\varphi^*(u) = (1+t)\log(1+t) - t$.
(b) Let $\varphi(t) = \frac{1}{p}t^p$ for $1 < p < \infty$. Then $\varphi'(t) = t^{p-1}$ and $(\varphi^*)'(u) = (\varphi')^{-1}(u) = u^{\frac{1}{p-1}} = u^{p'-1}$ with $\frac{1}{p} + \frac{1}{p'} = 1$. Integration over u implies $\varphi^*(u) = \frac{1}{p'}u^{p'}$.
(c) Let $\varphi(t) = t^p$ for $1 < p < \infty$. Then $\varphi'(t) = pt^{p-1}$ and $(\varphi^*)'(u) = (\varphi')^{-1}(u) = (u/p)^{\frac{1}{p-1}} = p^{\frac{1}{1-p}}u^{p'-1}$ with $\frac{1}{p} + \frac{1}{p'} = 1$. Integration over u implies $\varphi^*(u) = p^{\frac{1}{1-p}}\frac{1}{p'}u^{p'} = p^{-p'}(p-1)u^{p'}$.

Remark 2.6.10. We have seen that the supremum in Remark 2.6.9 is attained for any N-function φ. However, this is not the case if φ is only a Φ-function. Indeed, if $\varphi(t) = t$, then $\varphi^*(u) = \infty \cdot \chi_{\{u>1\}}(u)$. However, $tu = \varphi_1(t) + (\varphi_1)^*(u)$ only holds if $u = 1$ and $t \geqslant 0$ or if $u \in [0,1]$ and $t = 0$.

There are a lot of nice estimates for N-functions. Let us collect a few.

Lemma 2.6.11. *Let φ be an N-function. Then for all $t \geqslant 0$ and all $\varepsilon > 0$*

$$t \leqslant \varphi^{-1}(t)(\varphi^*)^{-1}(t) \leqslant 2t, \tag{2.6.12}$$
$$(\varphi^*)'\big(\varphi'(t) - \varepsilon\big) \leqslant t \leqslant (\varphi^*)'\big(\varphi'(t)\big), \tag{2.6.13}$$
$$\varphi'\big((\varphi^*)'(t) - \varepsilon\big) \leqslant t \leqslant \varphi'\big((\varphi^*)'(t)\big), \tag{2.6.14}$$
$$\varphi^*\big(\varphi'(t)\big) \leqslant t\varphi'(t), \tag{2.6.15}$$
$$\varphi^*\Big(\frac{\varphi(t)}{t}\Big) \leqslant \varphi(t), \tag{2.6.16}$$

where we assumed $t > 0$ in (2.6.16).

Proof. We first note that $(\varphi^*)' = (\varphi')^{-1}$ by Remark 2.6.9. Let $t \geqslant 0$ and $\varepsilon > 0$. The first part of (2.6.13) follows from

$$(\varphi^*)'\big(\varphi'(t) - \varepsilon\big) = \inf\{a \geqslant 0 \colon \varphi'(a) > \varphi'(t) - \varepsilon\} \leqslant t.$$

The second part of (2.6.13) follows from

$$\begin{aligned}\varphi'\big((\varphi^*)'(t)\big) &= \varphi'\Big(\inf\{a \geqslant 0 \colon \varphi'(a) > t\}\Big)\\ &= \inf\{\varphi'(a) \geqslant 0 \colon \varphi'(a) > t\} \geqslant t,\end{aligned}$$

where we have used that φ' is right-continuous and non-decreasing.

Now, (2.6.14) is a consequence of (2.6.13) using $(\varphi^*)^* = \varphi$. By Young's inequality (2.6.2) we estimate

$$\varphi^{-1}(t)(\varphi^*)^{-1}(t) \leqslant t + t = 2t.$$

With Lemma 2.6.6 for φ and φ^* and (2.6.13) we deduce

$$\varphi^*\big(\varphi'(t)\big) \leqslant (\varphi^*)'(\varphi'(t))\varphi'(t) \leqslant t\varphi'(t),$$
$$\varphi^*\Big(\frac{\varphi(t)}{t} - \varepsilon\Big) \leqslant \Big(\frac{\varphi(t)}{t} - \varepsilon\Big)(\varphi^*)'\Big(\frac{\varphi(t)}{t} - \varepsilon\Big) \leqslant \frac{\varphi(t)}{t}(\varphi^*)'\big(\varphi'(t) - \varepsilon\big) \leqslant \varphi(t).$$

Letting $\varepsilon \to 0$ in the latter inequality yields (2.6.16). Setting $t = \varphi^{-1}(u)$ in (2.6.16) gives

$$\varphi^*\Big(\frac{u}{\varphi^{-1}(u)}\Big) \leqslant u.$$

From this it follows that $u \leqslant \varphi^{-1}(u)(\varphi^*)^{-1}(u)$. □

Note that if φ and φ^* satisfy the Δ_2-condition (Definition 2.4.1), than all the "$\leqslant$"-signs in Lemma 2.6.11 can be replaced by "$\approx$"-signs.

2.7 Associate Spaces and Dual Spaces

In the case of classical Lebesgue spaces it is well known that there is a natural embedding of $L^{q'}$ into $(L^q)^*$ for $1 \leqslant q \leqslant \infty$ and $\frac{1}{q} + \frac{1}{q'} = 1$. In particular, for every $g \in L^{q'}$ the mapping $J_g \colon f \mapsto \int fg\,d\mu$ is an element of $(L^q)^*$. Even more, if $1 \leqslant q < \infty$, then the mapping $g \mapsto J_g$ is an isometry from $L^{q'}$ to $(L^q)^*$. Besides the nice characterization of the dual space, this has the consequence that

$$\|f\|_q = \sup_{\|g\|_{q'} \leqslant 1} \int f\,g\,d\mu$$

for every $1 \leqslant q \leqslant \infty$. This formula is often called the *norm conjugate formula.* The cases $q = 1$ and $q = \infty$ need special attention, since $(L^1)^* = L^\infty$ but $(L^\infty)^* \neq L^1$. However, the isometry $(L^1)^* = L^\infty$ suffices for the proof of the formula when $q = 1$ and $q = \infty$.

In the case of Musielak–Orlicz spaces we have a similar situation. We will see that L^{φ^*} can be naturally embedded into $(L^\varphi)^*$. Moreover, the mapping $g \mapsto J_g$ is an isomorphism under certain assumptions on φ, which exclude for example the case $L^\varphi = L^\infty$. The mapping is not an isometry but its operator norm lies in the interval $[1, 2]$.

The norm conjugate formula above requires more attention in the case of Musielak–Orlicz spaces. Certainly, we cannot expect equality but only equivalence up to a factor of 2. Since the space L^φ can partly behave like L^1 and partly like L^∞, there are cases, where $(L^\varphi)^* \neq L^{\varphi^*}$ and $L^\varphi \neq (L^{\varphi^*})^*$. This is in particular the case for our generalized Lebesgue spaces $L^{p(\cdot)}$ (see Chap. 3) when p take the values 1 and ∞ on some subsets. To derive an equivalent of the norm conjugate formula for L^{φ}, we need to study the associate space, which is a closed subspaces of $(L^{\varphi})^*$ generated by measurable functions.

Definition 2.7.1. Let $\varphi \in \Phi(A, \mu)$. Then

$$\big(L^{\varphi}(A,\mu)\big)' := \{g \in L^0(A,\mu) \,:\, \|g\|_{(L^{\varphi}(A,\mu))'} < \infty\}$$

with norm

$$\|g\|_{(L^{\varphi}(A,\mu))'} := \sup_{f \in L^{\varphi}\,:\, \|f\|_{\varphi} \leqslant 1} \int_A |f|\,|g|\,d\mu,$$

will be called the *associate space of* $L^{\varphi}(A,\mu)$ or $(L^{\varphi})'$ for short.

In the definition of the norm of the associate space $(L^{\varphi})'$ it suffices to take the supremum over simple function from L^{φ}:

Lemma 2.7.2. *Let* $\varphi \in \Phi(A, \mu)$. *Then*

$$\|g\|_{(L^{\varphi})'} = \sup_{f \in S \cap L^{\varphi}\,:\, \|f\|_{\varphi} \leqslant 1} \int_A |f|\,|g|\,d\mu$$

for all $g \in (L^{\varphi}(A,\mu))'$.

Proof. For $g \in (L^{\varphi})'$ let $|||g|||$ in this proof denote the right-hand side of the expression in the lemma. It is obvious that $|||g|||_{\varphi} \leqslant \|g\|_{(L^{\varphi})'}$. To prove the reverse let $f \in L^{\varphi}$ with $\|f\|_{\varphi} \leqslant 1$. We have to prove that $\int |f|\,|g|\,d\mu \leqslant |||g|||$. Let (f_k) be a sequence of simple functions such that $0 \leqslant f_k \nearrow |f|$ almost everywhere. In particular, $f_k \in S(A,\mu) \cap L^{\varphi}$ and $\|f_k\|_{\varphi} \leqslant \|f\|_{\varphi} \leqslant 1$, since L^{φ} is solid (Theorem 2.3.17 (b)). Since $0 \leqslant f_k|g| \nearrow |f||g|$, we can conclude by the theorem of monotone convergence and the definition of $|||g|||$ that

$$\int |f||g|\,d\mu = \lim_{k\to\infty} \int f_k|g|\,d\mu \leqslant |||g|||.$$

The claim follows by taking the supremum over all possible f. □

As an immediate consequence of Hölder's inequality (Lemma 2.6.5) we have
$L^{\varphi^*}(A,\mu) \hookrightarrow (L^{\varphi}(A,\mu))'$ and

$$\|g\|_{(L^{\varphi})'} \leqslant 2\,\|g\|_{\varphi^*}$$

for every $g \in L^{\varphi^*}(A,\mu)$.

If $g \in (L^\varphi)'$ and $f \in L^\varphi$, then $fg \in L^1$ by definition of the associate space. In particular, the integral $\int fg\,d\mu$ is well defined and

$$\left| \int fg\,d\mu \right| \leqslant \|g\|_{(L^\varphi)'} \|f\|_\varphi.$$

Thus $f \mapsto \int fg\,d\mu$ defines an element of the dual space $(L^\varphi)^*$ with $\|g\|_{(L^\varphi)^*} \leqslant \|g\|_{(L^\varphi)'}$. Therefore, for every $g \in (L^\varphi)'$ we can define an element $J_g \in (L^\varphi)^*$ by

$$J_g : f \mapsto \int fg\,d\mu \tag{2.7.3}$$

and we have $\|J_g\|_{(L^\varphi)^*} \leqslant \|g\|_{(L^\varphi)'}$. Since L^φ is circular (Theorem 2.3.17 (a)), we even have

$$\begin{aligned} \|J_g\|_{(L^\varphi)^*} &= \sup_{f \in L^\varphi : \|f\|_\varphi \leqslant 1} \left| \int fg\,d\mu \right| \\ &= \sup_{f \in L^\varphi : \|f\|_\varphi \leqslant 1} \int |f|\,|g|\,d\mu = \|g\|_{(L^\varphi)'} \end{aligned}$$

for every $g \in (L^\varphi)'$. Obviously, $g \mapsto J_g$ is linear. Hence, $g \mapsto J_g$ defines an isometric, *natural embedding* of $(L^\varphi)' \hookrightarrow (L^\varphi)^*$. So the associate space $(L^\varphi)'$ is isometrically isomorphic to a closed subset of the dual space $(L^\varphi)^*$ and therefore itself a Banach space. It is easy to see that $(L^\varphi)'$ is circular and solid. We have the following inclusions of Banach spaces

$$L^{\varphi^*} \hookrightarrow (L^\varphi)' \hookrightarrow (L^\varphi)^*.$$

Under rather few assumptions on φ, we will see that the first inclusion is surjective and therefore an isomorphism even if L^{φ^*} is not isomorphic to the dual space $(L^\varphi)^*$. Therefore, the notion of the associate space is more flexible than that of the dual space.

The mapping $g \mapsto J_g$ can also be used to define natural embeddings

$$L^\varphi = L^{\varphi^{**}} \hookrightarrow (L^{\varphi^*})' \hookrightarrow (L^{\varphi^*})^*.$$

if we replace above φ by φ^* and use $\varphi = \varphi^{**}$ (Corollary 2.6.3).

Since $L^{\varphi^*} \hookrightarrow (L^\varphi)' \hookrightarrow (L^\varphi)^*$ via the embedding $g \mapsto J_g$, we can evaluate the conjugate semimodular $(\varrho_\varphi)^*$ at J_g for every $g \in L^{\varphi^*}$. As a direct consequence of Young's inequality (2.6.2) we have

$$(\varrho_\varphi)^*(J_g) = \sup_{f \in L^\varphi} \big(J_g(f) - \varrho_\varphi(f) \big) \leqslant \varrho_{\varphi^*}(g).$$

Theorem 2.7.4. *Let $\varphi \in \Phi(A,\mu)$ be such that $S(A,\mu) \subset L^{\varphi}(A,\mu)$. Then $L^{\varphi^*}(A,\mu) = (L^{\varphi}(A,\mu))'$, $\varrho_{\varphi^*}(g) = (\varrho_\varphi)^*(J_g)$ and*

$$\|g\|_{\varphi^*} \leqslant \|g\|_{(L^\varphi)'} = \|J_g\|_{(L^\varphi)^*} \leqslant 2\,\|g\|_{\varphi^*}$$

for every $g \in L^{\varphi^*}(A,\mu)$, where $J_g : f \mapsto \int_A fg\,d\mu$. (or complex-valued functions, the constant 2 should be replaced by 4.)

Proof. For the sake of simplicity we assume $\mathbb{K} = \mathbb{R}$. In the case $\mathbb{K} = \mathbb{C}$ we can proceed analogously by splitting g into its real and imaginary part.

We already know that $L^{\varphi^*} \subset (L^\varphi)'$, $\|g\|_{(L^\varphi)'} = \|J_g\|_{(L^\varphi)^*} \leqslant 2\,\|g\|_{\varphi^*}$, and $(\varrho_\varphi)^*(J_g) \leqslant \varrho_{\varphi^*}(g)$ for every $g \in L^{\varphi^*}$. Fix $g \in (L^\varphi)'$. We claim that $g \in L^{\varphi^*}$ and $\varrho_{\varphi^*}(g) \leqslant (\varrho_\varphi)^*(J_g)$.

Since μ is σ-finite, we find measurable sets $A_k \subset A$ with $\mu(A_k) < \infty$ and $A_1 \subset A_2 \subset \dots$ such that $A = \bigcup_{k=1}^\infty A_k$. Let $\{q_1, q_2, \dots\}$ be a countable, dense subset of $[0,\infty)$ with $q_j \neq q_k$ for $j \neq k$ and $q_1 = 0$. For $k \in \mathbb{N}$ and $y \in A$ define

$$r_k(y) := \chi_{A_k}(y) \max_{j=1,\dots,k} \big(q_j\,|g(y)| - \varphi(y,q_j)\big).$$

The special choice $q_1 = 0$ implies $r_k(y) \geqslant 0$ for all $y \geqslant 0$. Since $\{q_1, q_2, \dots\}$ is dense in $[0,\infty)$ and $\varphi(y,\cdot)$ is left-continuous, $r_k(y) \nearrow \varphi^*(y, |g(y)|)$ for any $y \in A$ as $k \to \infty$. For every $k \in \mathbb{N}$ there exists a simple function f_k with $f_k(A) \subset \{q_1, \dots, q_k\}$ and $f_k(y) = 0$ for all $y \in A \setminus A_k$ such that

$$r_k(y) = f_k(y)\,|g(y)| - \varphi(y, f_k(y))$$

for all $y \in A$. As a simple function, f_k belongs by assumption to $L^\varphi(A,\mu)$. Define $h_k(y) := f_k(y)\,\operatorname{sgn}(g(y))$ for $y \in A$, where $\operatorname{sgn}(a)$ denotes the sign of a. Then also h_k is a simple function (here we use $\mathbb{K} = \mathbb{R}$) and therefore

$$(\varrho_\varphi)^*(J_g) \geqslant J_g(h_k) - \varrho_\varphi(h_k) = \int_A g(y)h_k(y) - \varphi(y, |h_k(y)|)\,d\mu(y).$$

By the definition of h_k it follows that

$$(\varrho_\varphi)^*(J_g) \geqslant \int_A |g(y)|\,f_k(y) - \varphi(y, |f_k(y)|)\,d\mu(y) = \int_A r_k(y)\,d\mu(y).$$

Since $r_k \geqslant 0$ and $r_k(y) \nearrow \varphi^*(y, |g(y)|)$, we get by the theorem of monotone convergence that

$$(\varrho_\varphi)^*(J_g) \geqslant \limsup_{k\to\infty} \int_A r_k(y)\,d\mu(y) = \int_A \varphi^*(y, |g(y)|)\,d\mu(y) = \varrho_{\varphi^*}(g).$$

Together with $(\varrho_\varphi)^*(J_g) \leqslant \varrho_{\varphi^*}(g)$ we get $(\varrho_\varphi)^*(J_g) = \varrho_{\varphi^*}(g)$.

Since $g \mapsto J_g$ is linear, it follows that $(\varrho_\varphi)^*(\lambda J_g) = \varrho_{\varphi^*}(\lambda g)$ for every $\lambda > 0$ and therefore $\|g\|_{\varphi^*} = \|J_g\|_{(\varrho_\varphi)^*} \leqslant \|J_g\|_{(L^\varphi)^*} = \|g\|_{(L^\varphi)'}$ using in the second step Theorem 2.2.10. □

Theorem 2.7.4 allows us to generalize the norm conjugate formula to L^φ.

Corollary 2.7.5 (Norm conjugate formula). *Let $\varphi \in \Phi(A,\mu)$. If $S(A,\mu) \subset L^{\varphi^*}(A,\mu)$, then*

$$\|f\|_\varphi \leqslant \sup_{g \in L^{\varphi^*} : \|g\|_{\varphi^*} \leqslant 1} \int |f|\,|g|\,d\mu \leqslant 2\,\|f\|_\varphi$$

for every $f \in L^0(A,\mu)$. The supremum is unchanged if we replace the condition $g \in L^{\varphi^*}$ by $g \in S(A,\mu)$.

Proof. Applying Theorem 2.7.4 to φ^* and taking into account that $\varphi^{**} = \varphi$, we have

$$\|f\|_\varphi \leqslant \|f\|_{(L^{\varphi^*})'} \leqslant 2\,\|f\|_\varphi$$

for $f \in L^\varphi$. That the supremum does not change for $g \in S(A,\mu)$ follows by Lemma 2.7.2. The claim also follows in the case $f \in L^0 \setminus L^{\varphi^*} = L^0 \setminus (L^\varphi)'$, since both sides of the formula are infinite. □

Remark 2.7.6. Since μ is σ-finite it suffices in Theorem 2.7.4 and Corollary 2.7.5 to assume $S(A_k,\mu) \subset L^\varphi(A,\mu)$, where (A_k) is a sequence with $A_k \nearrow A$ and $\mu(A_k) < \infty$ for all k. This is important for example in weighted Lebesgue spaces $L^q_\omega(\mathbb{R}^n)$ with Muckenhoupt weights.

Definition 2.7.7. A normed space $(Y, \|\cdot\|_Y)$ with $Y \subset L^0(A,\mu)$ is called a *Banach function space*, if

(a) $(Y, \|\cdot\|_Y)$ is circular, solid and satisfies the Fatou property.
(b) If $\mu(E) < \infty$, then $\chi_E \in Y$.
(c) If $\mu(E) < \infty$, then $\chi_E \in Y'$, i.e. $\int_E |f|\,d\mu \leqslant c(E)\|f\|_Y$ for all $f \in Y$.

From Theorem 2.3.17 we know that L^φ satisfies (a) for every $\varphi \in \Phi(A,\mu)$ so one need only check (b) and (c). These properties are equivalent to $S \subset L^\varphi$ and $S \subset (L^\varphi)'$, where S is the set of simple functions. These inclusions may or may not hold, depending on the function φ.

Definition 2.7.8. A generalized Φ-function $\varphi \in \Phi(A,\mu)$ is called *proper* if the set of simple functions $S(A,\mu)$ satisfies $S(A,\mu) \subset L^\varphi(A,\mu) \cap (L^\varphi(A,\mu))'$.

So φ is proper if and only if L^φ is a Banach function space. Moreover, if φ is proper then the norm conjugate formula for L^φ and L^{φ^*} holds (Corollary 2.7.5) and $L^{\varphi^*} = (L^\varphi)'$.

Corollary 2.7.9. *Let $\varphi \in \Phi(A,\mu)$. Then the following are equivalent:*

(a) *φ is proper.*
(b) *φ^* is proper.*
(c) *$S(A,\mu) \subset L^{\varphi}(A,\mu) \cap L^{\varphi^*}(A,\mu)$.*

Proof. If (a) or (c) holds, then $S \subset L^{\varphi}$. Hence $(L^{\varphi})' = L^{\varphi^*}$ by Theorem 2.7.4, which obviously implies the equivalence of (a) and (c).

Applying this equivalence for the function φ^*, and taking into account that $\varphi^{**} = \varphi$, yields the equivalence of (b) and (c). □

Remark 2.7.10. The conditions $\chi_E \in L^{\varphi}$ and $\chi_E \in (L^{\varphi})'$ for $\mu(E) < \infty$ in Definition 2.7.7 can be interpreted in terms of embeddings. Indeed, $\chi_E \in L^{\varphi}$ implies $L^{\varphi^*} \hookrightarrow L^1(E)$. The condition $\chi_E \in (L^{\varphi})'$ is equivalent to $L^{\varphi}(E) \hookrightarrow L^1(E)$. In particular, if φ is proper, then $L^{\varphi}(\Omega) \hookrightarrow L^1_{\mathrm{loc}}(\Omega)$ and $L^{\varphi^*}(\Omega) \hookrightarrow L^1_{\mathrm{loc}}(\Omega)$.

Remark 2.7.11. Let $\varphi \in \Phi$ be proper; so L^{φ} is a Banach function space. It has been shown in [43, Proposition 3.6] that $f \in L^{\varphi}$ has absolutely continuous norm (see Remark 2.5.8) if and only if f has the following property: If g_k, $g \in L^0$ with $|g_k| \leqslant |f|$ and $g_k \to g$ almost everywhere, then $g_k \to g$ in L^{φ}. Thus, f acts as a majorant in the theorem of dominated convergence.

It has been shown by Lorentz and Luxemburg that the second associate space X'' of a Banach function space coincides with X with equality of norms, see [43, Theorem 2.7]. In particular, $(L^{\varphi})'' = L^{\varphi}$ with equality of norms if φ is proper. For the sake of completeness we include a proof of this result in our setting.

Theorem 2.7.12. *Let $\varphi \in \Phi(A,\mu)$ be proper. Then $L^{\varphi^*}(A,\mu) = (L^{\varphi}(A,\mu))'$ and $(L^{\varphi^*}(A,\mu))' = L^{\varphi}(A,\mu)$. Moreover, $(L^{\varphi}(A,\mu))'' = L^{\varphi}(A,\mu)$ with equality of norms, i.e. $\|f\|_{\varphi} = \|f\|_{(L^{\varphi})''}$ for all $f \in L^{\varphi}(A,\mu)$.*

Proof. The equalities $L^{\varphi^*} = (L^{\varphi})'$ and $(L^{\varphi^*})' = L^{\varphi}$ follow by Theorem 2.7.4 and as a consequence $(L^{\varphi})'' = (L^{\varphi^*})' = L^{\varphi^{**}} = L^{\varphi}$ using $\varphi^{**} = \varphi$. It only remains to prove the equality of norms. Let $f \in L^{\varphi}$, then

$$\|f\|_{(L^{\varphi})''} = \sup_{g \in (L^{\varphi})' : \|g\|_{(L^{\varphi})'} \leqslant 1} \int |f|\,|g|\,d\mu \leqslant \|f\|_{\varphi}.$$

We now prove $\|f\|_{\varphi} \leqslant \|f\|_{(L^{\varphi})''}$. We begin with the case $\mu(A) < \infty$. If $f = 0$, there is nothing to show, so assume $f \neq 0$. Let B denote the unit ball of L^{φ}. Due to Remark 2.7.10 and $\mu(A) < \infty$, we have $L^{\varphi}(A) \hookrightarrow L^1(A)$, so $B \subset L^1(A)$. Moreover, B is a closed, convex subset of $L^1(A)$. Indeed, if $u_k \in B$ with $u_k \to u$ in $L^1(A)$, then $u_k \to u$ μ-almost everywhere for a subsequence, so Fatou's lemma for the norm (Theorem 2.3.17) implies $u \in B$.

Let $h := \lambda f/\|f\|_\varphi$ with $\lambda > 1$, then $h \notin B$, so by the Hahn–Banach Theorem 1.4.2 there exists a functional on $(L^1(A))^*$ separating B and f. In other words, there exists a function $g \in L^\infty(A)$ and $\gamma \in \mathbb{R}$ such that

$$\operatorname{Re}\bigg(\int vg\,d\mu\bigg) \leqslant \gamma < \operatorname{Re}\bigg(\int hg\,d\mu\bigg)$$

for all $v \in B$, where we have used the representation of $(L^1(A))^*$ by $L^\infty(A)$. From $g \in L^\infty(A)$ and $\chi_A \in (L^\varphi)'$ it follows by solidity of $(L^\varphi)'$ that $g \in (L^\varphi)'$. Moreover, the circularity of L^φ implies that

$$\int |v||g|\,d\mu \leqslant \gamma < \int |h||g|\,d\mu = \frac{\lambda}{\|f\|_\varphi}\int |f||g|\,d\mu \leqslant \frac{\lambda\|f\|_{(L^\varphi)''}\|g\|_{(L^\varphi)'}}{\|f\|_\varphi}$$

for all $v \in B$. In other words,

$$\|g\|_{(L^\varphi)'} \leqslant \frac{\lambda\|f\|_{(L^\varphi)''}\|g\|_{(L^\varphi)'}}{\|f\|_\varphi}.$$

Using $\|g\|_{(L^\varphi)'} < \infty$ we get $\|f\|_\varphi \leqslant \lambda\|f\|_{(L^\varphi)''}$. This proves $\|f\|_\varphi \leqslant \|f\|_{(L^\varphi)''}$ and therefore $\|f\|_\varphi = \|f\|_{(L^\varphi)''}$.

It remains to consider the case $\mu(A)=\infty$. Choose $A_k \subset A$ with $\mu(A_k)<\infty$, $A_1 \subset A_2 \subset \dots$, and $A = \bigcup_{k=1}^\infty A_k$. Then $\|f\chi_{A_k}\|_\varphi = \|f\|_{L^\varphi(A_k)} = \|f\|_{(L^\varphi(A_k))''} = \|f\chi_{A_k}\|_{(L^\varphi(A))''}$ by the first part. Now, with the Fatou property of L^φ and $(L^\varphi)''$ we conclude $\|f\|_\varphi = \|f\|_{(L^\varphi)''}$. □

Remark 2.7.13. Let $\varphi \in \Phi(A,\mu)$ be proper. Then we can use Theorem 2.7.12 Hölder's inequality to derive the formula

$$\frac{1}{2}\|f\|_\varphi \leqslant \sup_{h\in L^{\varphi^*}\,:\,\|h\|_{\varphi^*}\leqslant 1}\int |f|\,|h|\,d\mu \leqslant 2\,\|f\|_\varphi.$$

for all $f \in L^0(A,\mu)$. This is a weaker version of the norm conjugate formula in Corollary 2.7.5, with an extra factor $\frac{1}{2}$ on the left-hand side.

We are now able to characterize the dual space of L^φ.

Theorem 2.7.14. *Let $\varphi \in \Phi(A,\mu)$ be proper and locally integrable, and suppose that $E^\varphi = L^\varphi$. Then $V : g \mapsto J_g$ is an isomorphism from $L^{\varphi^*}(A,\mu)$ to $(L^\varphi(A,\mu))^*$.*

Proof. By Theorem 2.7.4 V is an isomorphism from L^{φ^*} onto its image $\operatorname{Im}(V) \subset (L^\varphi)^*$. In particular, $\operatorname{Im}(V)$ is a closed subspace of $(L^\varphi)^*$. Since φ is locally integrable $\overline{S} = E^\varphi$ by Theorem 2.5.9, so that $\overline{S} = E^\varphi = L^\varphi$.

We have to show that V is surjective. We begin with the case $\mu(A) < \infty$. Let $J \in (L^\varphi)^*$. For any measurable set $E \subset A$ we define $\tau(E) := J(\chi_E)$, which is well defined since $S \subset L^\varphi$. We claim that τ is a signed, finite measure on A. Obviously, τ is a set function with $\tau(E_1 \cup E_2) = \tau(E_1) + \tau(E_2)$ for E_1, E_2 disjoint measurable sets. Let (E_j) be sequence of pairwise disjoint, measurable sets. Let $E := \bigcup_{j=1}^\infty E_j$. Then $\sum_{j=1}^k \chi_{E_j} \to \chi_E$ almost everywhere and by dominated convergence (Lemma 2.3.16) using $\chi_E \in L^\varphi = E^\varphi$ we find that $\sum_{j=1}^k \chi_{E_j} \to \chi_E$ in L^φ. This and the continuity of J imply

$$\sum_{j=1}^\infty \tau(E_j) = \sum_{j=1}^\infty J(\chi_{E_j}) = J(\chi_E) = \tau(E),$$

which proves that τ is σ-additive. The estimate

$$|\tau(E)| = |J(\chi_E)| \leqslant \|J\|_{(L^\varphi)^*} \|\chi_E\|_\varphi \leqslant \|J\|_{(L^\varphi)^*} \|\chi_A\|_\varphi$$

for all measurable E, proves that τ is a signed, finite measure. If $\mu(E) = 0$, then $\tau(E) = J(\chi_E) = 0$, so τ is absolutely continuous with respect to μ. Thus by the Radon–Nikodym Theorem 1.4.13 there exists a function $g \in L^1(A)$ such that

$$J(f) = \int_A f g \, d\mu \tag{2.7.15}$$

for all $f = \chi_E$ with E measurable and therefore by linearity for all $f \in S$. We claim that $\|g\|_{(L^\varphi)'} \leqslant \|J\|_{(L^\varphi)^*}$. Due to Lemma 2.7.2 it suffices to show that $\int |f|\,|g| \, d\mu \leqslant \|J\|_{(L^\varphi)^*}$ for every $f \in S = S \cap L^\varphi$ with $\|f\|_\varphi \leqslant 1$. Fix such an f. If $\mathbb{K} = \mathbb{R}$, then $\operatorname{sgn} g$ is a simple function. However, to include the case $\mathbb{K} = \mathbb{C}$, we need to approximate $\operatorname{sgn} g$ by simple function as follows. Since $\operatorname{sgn} g \in L^\infty$, we find a sequence (h_k) of simple functions with $h_k \to \operatorname{sgn} g$ almost everywhere and $|h_k| \leqslant 1$. Since $|f| h_k \in S$ and $\| \, |f| \, h_k\|_\varphi \leqslant \|f\|_\varphi \leqslant 1$, we estimate $\int |f| h_k g \, dx = J(|f| h_k) \leqslant \|J\|_{(L^\varphi)^*}$ using (2.7.15). We have $|f| h_k g \to |f| \, |g|$ almost everywhere and $|f h_k g| \leqslant |f||g| \in L^1$, since $g \in L^1$ and $f \in L^\infty$ as a simple function. Therefore, by the theorem of dominated convergence we conclude $\int |f||g| \, dx = \lim_{k \to \infty} \int |f| h_k g \, dx \leqslant \|J\|_{(L^\varphi)^*}$. This yields $\|g\|_{(L^\varphi)'} \leqslant \|J\|_{(L^\varphi)^*}$. Then $g \in L^{\varphi^*}$ follows from $(L^\varphi)' = L^{\varphi^*}$ by Theorem 2.7.4. By (2.7.3) and (2.7.15) the functionals J_g and J agree on the set S. So the continuity of J and J_g and $\overline{S} = L^\varphi$ imply $J = J_g$ proving the surjectivity of $g \mapsto J_g$ in the case $\mu(A) < \infty$.

It remains to prove the surjectivity for μ σ-finite. Choose $A_k \subset A$ with $\mu(A_k) < \infty$, $A_1 \subset A_2 \subset \dots$, and $A = \bigcup_{k=1}^\infty A_k$. By restriction we see that $J \in (L^\varphi(A_k))^*$ for each $J \in (L^\varphi(A))^*$. Since $\mu(A_k) < \infty$, there exists $g_k \in L^{\varphi^*}(A_k)$ such that $J(f) = J_{g_k}(f)$ for any $f \in L^\varphi(A_k)$ and $\|g_k\|_{\varphi^*} \leqslant \|J\|_{(L^\varphi)^*}$. The injectivity of $g \mapsto J_g$ implies $g_j = g_k$ on A_j for all $k \geqslant j$. So $g := g_k$ on A_k is well defined and $J(f) = J_g(f)$ for all $f \in L^\varphi(A_k)$ and every

k. Since $|g_k| \nearrow |g|$ almost everywhere and $\sup_k \|g_k\|_{\varphi^*} \leqslant \|J\|_{(L^\varphi)^*}$, it follows by the Fatou property of L^{φ^*} that $\|g\|_{\varphi^*} \leqslant \|J\|_{(L^\varphi)^*}$.

It remains to prove $J = J_g$. Let $f \in L^\varphi$. Then by Fatou's lemma (Lemma 2.3.16), $f\,\chi_{A_k} \to f$ in L^φ. Hence, the continuity of J and J_g and $J(f\,\chi_{A_k}) = J_g(f\,\chi_{A_k})$ yields $J(f) = J_g(f)$ as desired. □

Remark 2.7.16. (a) If φ is proper and locally integrable, then the condition $L^\varphi = E^\varphi$ is equivalent to the density of the set S of simple functions in L^φ, see Theorem 2.5.9.

(b) If μ is atom-free, then the assumptions "locally integrable" and "$E^\varphi = L^\varphi$" are also necessary for $V : g \mapsto J_g$ from $L^{\varphi^*} = (L^\varphi)'$ to $(L^\varphi)^*$ to be an isomorphism. Indeed, if V is an isomorphism, then it has been shown in [43, Theorem 4.1] that every function $f \in L^\varphi$ has absolutely continuous norm (see Remark 2.5.8). In particular, every χ_E with $\mu(E) < \infty$ has absolutely continuous norm. We prove that φ is locally integrable by contradiction, so assume that there exists a measurable set E and $\lambda > 0$ such that $\mu(E) < \infty$ and $\varrho_\varphi(\lambda\chi_E) = \infty$. Since μ is atom-free there exists a sequence (E_k) of pairwise disjoint, measurable sets such that $E_k \searrow \emptyset$ and $\varrho_\varphi(\lambda\chi_{E_k}) = \infty$. In particular, $\|\chi_{E_k}\|_\varphi \geqslant \frac{1}{\lambda}$. However, since χ_E has absolutely continuous norm, we should have $\|\chi_{E_k}\|_\varphi = \|\chi_E\chi_{E_k}\|_\varphi \to 0$, which gives the desired contradiction. Thus, φ is locally integrable. If follows from Theorem 2.5.9 that $E^\varphi = \overline{S}$, where S are the simple functions. Moreover, since V is an isomorphism, by the norm conjugate formula in Lemma 2.7.2 it follows that $S^\circ = \{0\}$, where S° is the annihilator of S. This implies $E^\varphi = \overline{S} = S^{\circ\circ} = L^{\varphi^*}$.

The reflexivity of L^φ can be reduced to the characterization of $(L^\varphi)^*$ and $(L^{\varphi^*})^*$.

Lemma 2.7.17. *Let $\varphi \in \Phi(A,\mu)$ be proper. Then L^φ is reflexive, if and only if the natural embeddings $V \colon L^{\varphi^*} \to (L^\varphi)^*$ and $U \colon L^\varphi \to (L^{\varphi^*})^*$ are isomorphisms.*

Proof. Let ι denote the natural injection of L^φ into its bidual $(L^\varphi)^{**}$. It is easy to see that $V^* \circ \iota = U$. Indeed,

$$\langle V^*\iota f, g\rangle = \langle \iota f, Vg\rangle = \langle Vg, f\rangle = \int f(x)g(x)\,d\mu = \langle Uf, g\rangle$$

for $f \in L^\varphi$ and $g \in L^{\varphi^*}$. If V and U are isomorphisms, then $\iota = (V^*)^{-1} \circ U$ must be an isomorphism and L^φ is reflexive.

Assume now that L^φ is reflexive. We have to show that U and V are isomorphisms. We already know from Theorem 2.7.4 (since φ is proper) that U and V are isomorphisms from L^φ and L^{φ^*} to their images $\mathrm{Im}(U)$ and $\mathrm{Im}(V)$, respectively. In particular, V is a closed operator and as a consequence $\mathrm{Im}(V^*) = (\ker(V))^\circ$. The injectivity of V implies that V^* is surjective. So

$U = V^* \circ \iota$ is surjective as well. This proves that U is an isomorphism. The formula $U = V^* \circ \iota$ implies that V^* is also an isomorphism. Since V is a closed operator, we have $\mathrm{Im}(V) = (\ker(V^*))^\circ$. The injectivity of V^* proves that V is surjective and therefore an isomorphism. □

By Theorem 2.7.14 and Lemma 2.7.17 we immediately get the reflexivity of L^φ.

Corollary 2.7.18. *Let $\varphi \in \Phi(A,\mu)$ be proper. If φ and φ^* are locally integrable, $E^\varphi = L^\varphi$ and $E^{\varphi^*} = L^{\varphi^*}$, then L^φ is reflexive.*

2.8 Embeddings and Operators

In this section we characterize bounded, linear operators from one Musielak–Orlicz space to another. Recall that the operator S is said to be bounded from L^φ to L^ψ if $\|Sf\|_\varphi \leqslant C\,\|f\|_\psi$. We want to characterize this in terms of the modular. The study of embeddings is especially important to us, i.e. we want to know when the identity is a bounded operator. Such embeddings, which are denoted by $L^\varphi \hookrightarrow L^\psi$, can be characterized by comparing φ pointwise with ψ.

Let us begin with a characterization of bounded, sub-linear operators. Let $\varphi, \psi \in \Phi(A,\mu)$ and let $S\colon L^\varphi(A,\mu) \to L^\psi(A,\mu)$ be sub-linear. By the norm-modular unit ball property, S is bounded if and only if there exist $c > 0$ such that

$$\varrho_\varphi(f) \leqslant 1 \implies \varrho_\psi(Sf/c) \leqslant 1.$$

If φ and ψ satisfy the Δ_2-condition, then this is equivalent to the existence of $c_1, c_2 > 0$ such that

$$\varrho_\varphi(f) \leqslant c_1 \implies \varrho_\psi(Sf) \leqslant c_2$$

(since the Δ_2-condition allows us to move constants out of the modular).

Theorem 2.8.1. *Let $\varphi, \psi \in \Phi(A,\mu)$ and let the measure μ be atom-less. Then $L^\varphi(A,\mu) \hookrightarrow L^\psi(A,\mu)$ if and only if there exists $c' > 0$ and $h \in L^1(A,\mu)$ with $\|h\|_1 \leqslant 1$ such that*

$$\psi\Big(y, \frac{t}{c'}\Big) \leqslant \varphi(y,t) + h(y)$$

for almost all $y \in A$ and all $t \geqslant 0$.

Moreover, c' is bounded by the embedding constant, whereas the embedding constant is bounded by $2c'$.

Proof. Let us start by showing that the inequality implies the embedding. Let $\|f\|_\varphi \leqslant 1$, which yields by the unit ball property that $\varrho_\varphi(f) \leqslant 1$. Then

$$\varrho_\psi\Big(\frac{f}{2c'}\Big) \leqslant \frac{1}{2}\varrho_\psi\Big(\frac{f}{c'}\Big) \leqslant \frac{1}{2}\varrho_\varphi(f) + \frac{1}{2}\int\limits_A h(y)\,dy \leqslant 1.$$

This and the unit ball property yield $\|f/(2c')\|_\psi \leqslant 1$. Then the embedding follows by the scaling argument.

Assume next that the embedding holds with embedding constant c_1. For $y \in A$ and $t \geqslant 0$ define

$$\alpha(y,t) := \begin{cases} \psi(y, \frac{t}{c_1}) - \varphi(y,t) & \text{if } \varphi(y,t) < \infty, \\ 0 & \text{if } \varphi(y,t) = \infty. \end{cases}$$

Since $\varphi(y,\cdot)$ and $\psi(y,\cdot)$ are left-continuous for all $y \in A$, also $\alpha(y,\cdot)$ is left-continuous for all $y \in A$. Let (r_k) be a sequence of distinct numbers with $\{r_k : k \in \mathbb{N}\} = \mathbb{Q} \cap [0,\infty)$ and $r_1 = 0$. Then

$$\psi(y, \tfrac{r_k}{c_1}) \leqslant \varphi(y, r_k) + \alpha(y, r_k)$$

for all $k \in \mathbb{N}$ and $y \in A$. Define

$$b_k(y) := \max_{1 \leqslant j \leqslant k} \alpha(y, r_j).$$

Since $r_1 = 0$ and $\alpha(y,0) = 0$, we have $b_k \geqslant 0$. Moreover, the functions b_k are measurable and nondecreasing in k. The function $b := \sup_k b_k$ is measurable, non-negative, and satisfies

$$b(y) = \sup_{t \geqslant 0} \alpha(y,t),$$
$$\psi(y, \tfrac{t}{c_1}) \leqslant \varphi(y,t) + b(y)$$

for all $y \in A$ and all $t \geqslant 0$, where we have used that $\alpha(y,\cdot)$ is left-continuous and the density of $\{r_k : k \in \mathbb{N}\}$ in $[0,\infty)$.

We now show that $b \in L^1(A,\mu)$ with $\|b\|_1 \leqslant 1$. We consider first the case $|b| < \infty$ a.e., and assume to the contrary that there exists $\varepsilon > 0$ such that

$$\int\limits_A b(y)\,d\mu(y) \geqslant 1 + 2\varepsilon.$$

Define

$$V_k := \big\{y \in A : \ \alpha(y, r_k) > \tfrac{1}{1+\varepsilon} b(y)\big\},$$
$$W_{k+1} := V_{k+1} \setminus \big(V_1 \cup \cdots \cup V_k\big)$$

for all $k \in \mathbb{N}$. Note that $V_1 = \emptyset$ due to the special choice $r_1 = 0$. Since $\{r_k : k \in \mathbb{N}\}$ is dense in $[0, \infty)$ and $\alpha(y, \cdot)$ is left-continuous for every $y \in A$, we have $\bigcup_{k=1}^{\infty} V_k = \bigcup_{k=2}^{\infty} W_k = \{y \in A : b(y) > 0\}$.

Let $f := \sum_{k=2}^{\infty} r_k \, \chi_{W_k}$. For every $y \in W_k$ we have $\alpha(y, r_k) > 0$ and therefore $\varphi(y, r_k) < \infty$. If y is outside of $\bigcup_{k=2}^{\infty} W_k$, then $\varphi(y, |f(y)|) = 0$. This implies that $\varphi(y, |f(y)|)$ is everywhere finite. Moreover, by the definition of W_k and α we get

$$\psi\Big(y, \frac{|f(y)|}{c_1}\Big) \geqslant \varphi\big(y, |f(y)|\big) + \frac{1}{1+\varepsilon} b(y) \tag{2.8.2}$$

for all $y \in A$.

If $\varrho_\varphi(f) \leqslant 1$, then $\varrho_\psi(\frac{f}{c_1}) \leqslant 1$ by the unit ball property since c_1 is the embedding constant. However, this contradicts

$$\varrho_\psi(\tfrac{f}{c_1}) \geqslant \varrho_\varphi(f) + \frac{1}{1+\varepsilon} \int\limits_A b(y) \, d\mu(y) \geqslant \frac{1+2\varepsilon}{1+\varepsilon} > 1,$$

where we have used (2.8.2) and $\bigcup_{k=2}^{\infty} W_k = \{y \in A : b(y) > 0\}$. So we can assume that $\varrho_\varphi(f) > 1$. Since μ is atom-less and $\varphi(y, |f(y)|)$ is almost everywhere finite, there exists $U \subset A$ with $\varrho_\varphi(f \chi_U) = 1$. Thus

$$\begin{aligned} \varrho_\psi(\tfrac{f}{c_1} \, \chi_U) &\geqslant \varrho_\varphi(f \, \chi_U) + \tfrac{1}{1+\varepsilon} \int\limits_U b(y) \, d\mu(y) \\ &= 1 + \tfrac{1}{1+\varepsilon} \int\limits_U b(y) \, d\mu(y). \end{aligned} \tag{2.8.3}$$

Now, $\varrho_\varphi(f \chi_U) = 1$ implies that $\mu(U \cap \{f \neq 0\}) > 0$. Since $\{f \neq 0\} = \bigcup_{k=2}^{\infty} W_k = \{y \in A : b(y) > 0\}$ we get $\mu(U \cap \{y \in A : b(y) > 0\}) > 0$ and

$$\int\limits_U b(y) \, d\mu(y) > 0.$$

This and (2.8.3) imply that

$$\varrho_\psi(f/c_1 \, \chi_U) > 1.$$

which contradicts $\varrho_\psi(f/c_1) \leqslant 1$. Thus the case where $|b| < \infty$ a.e. is complete.

If we assume that there exists $E \subset A$ with $b|_E = \infty$ and $\mu(E) > 0$, then a similar argument with $V_k := \{y \in E : \alpha(y, r_k) \geqslant \frac{2}{\mu(E)}\}$ yields a contradiction. Hence this case cannot occur, and the proof is complete by what was shown previously. $\square$

http://www.springer.com/978-3-642-18362-1

Lebesgue and Sobolev Spaces with Variable Exponents
Diening, L.; Harjulehto, P.; Hästö, P.; Ruzicka, M.
2011, IX, 509 p. 10 illus., Softcover
ISBN: 978-3-642-18362-1