Известия НАН Армении. Математика, том 44, н. 5, 2009, стр. 27-70.

STRONG SOLVABILITY IN ORLICZ SPACES

PETER KOSMOL AND DIETER MÜLLER-WICHARDS

University of Kiel, Germany
Hamburg University of Applied Sciences, Germany
E-mail: *kosmol@math.uni-kiel.de, muewi@etech.haw-hamburg.de*

Аннотация. The main objective of the authors is to characterize strong solvability of optimization problems where convergence of the values to the optimum already implies norm-convergence of the approximations to the minimal solution. It turns out that strong solvability can be geometrically characterized by the local uniform convexity of the corresponding convex functional (local uniform convexity being appropriately defined). For bounded functionals we establish that in reflexive Banach spaces strong solvability is characterized by the Fréchet-differentiability of the convex conjugate. These results are based in part on a paper of Asplund and Rockafellar on the duality of A-differentiability and B-convexity of conjugate pairs of convex functions, where B is the polar of A. Before we apply these results to Orlicz spaces, we turn to E-spaces introduced by Fan and Glicksberg. Using the properties of E-spaces we can show that for finite not purely atomic measures Fréchet differentiability of an Orlicz space already implies its reflexivity. The main theorem gives - in 17 equivalent statements - a characterization of strong solvability, local uniform convexity, and Fréchet differentiability of the dual space, in case L^{Φ} is reflexive. It is remarkable that all these properties can also be equivalently expressed by the differentiability of Φ or the strict convexity of Ψ. In particular, L^{Φ} is an E-space, if L^{Φ} is reflexive and Φ is strictly convex.
We discuss applications that refer to

- Tychonov-regularization: local uniformly convex regularisations are sufficient to ensure convergence. As we have given a complete description of local uniform convexity in Orlicz spaces we can state such regularizing functionals explicitly.
- Ritz method: it is well known that the Ritz procedure generates a minimizing sequence. Actual convergence of the minimal solutions on each subspace is achieved if the original problem is strongly solvable.
- Greedy algorithms: the convergence proof makes use of the Kadec-Klee property of E-spaces.

1. INTRODUCTION

In this introductory section we recall some properties of Orlicz spaces necessary for what follows.

Definition 1. *An even, lower semi-continuous, nontrivial convex function* $\Phi : \mathbb{R} \to \overline{\mathbb{R}}_{\geq 0}$ *with* $\Phi(0) = 0$*, where* 0 *is an interior point of* $\mathrm{Dom}(\Phi)$ *is called a Young function. If* $\Phi(s) > 0$ *for* $s > 0$ *then* Φ *is said to be definite.*

Let (T, Σ, μ) be an arbitrary measure space and let E be the set of all μ-measurable real-valued functions on T. Then for a given Young function Φ we can define a modular $f^{\Phi} : E \to \overline{\mathbb{R}}$ by

$$f^{\Phi}(x) = \int_T \Phi(x) d\mu.$$

Minimization w.r.t. this functional on subsets of μ-measurable functions can be viewed as generalizations of L^p-minimizations, where $\Phi(s) = |s|^p/p$.

The Minkowski functional of the level set $S_{f^{\Phi}}(1)$ is defined by $p_{\Phi} : E \to \bar{\mathbb{R}}$ where

$$p_{\Phi}(x) := \inf \left\{ c > 0 \,\middle|\, \int_{t\epsilon T} \Phi\left(\frac{x(t)}{c}\right) d\mu \leq 1 \right\}.$$

The Orlicz-space $L^{\Phi}(\mu)$ is given by the subspace of E, where p_{Φ} is finite:

$$L^{\Phi}(\mu) := \left\{ x\epsilon E \,\middle|\, \exists \alpha > 0 : \int_{t\epsilon T} \Phi(\alpha x(t)) d\mu < \infty \right\}.$$

The above p_{Φ} defines a norm on $L^{\Phi}(\mu)$, called Luxemburg-norm, to be denoted by $\| \cdot \|_{(\Phi)}$. It is well known (s. e.g. [31] that $(L^{\Phi}(\mu), \| \cdot \|_{(\Phi)})$ is a Banach space.

Special cases:

- $\Phi_p(s) = |s|^p/p$ then $L^{\Phi} = L^p$ $(1 \leq p < \infty)$.
- $\Phi_{\infty}(s) := \begin{cases} \infty \text{ for } |s| > 1 \\ 0 \text{ for } |s| < 1 \end{cases}$ then $L^{\Phi} = L^{\infty}$.

By $\mathcal{M}^{\Phi}(\mu)$ we denote the closure of the subspace spanned by the step functions with finite support in $L^{\Phi}(\mu)$.

If Φ is finite, then $f^{\Phi} : \mathcal{M}^{\Phi}(\mu) \to \mathbb{R}$ is continuous, because f^{Φ} is bounded on the unit sphere of $\mathcal{M}^{\Phi}(\mu)$ (s. [31], p. 219).

The proof of the theorem that follows can be obtained by use of the two-norm theorem (s. [39]).

Theorem 1. *Let* Ψ *be not finite and* $\mu(T) < \infty$*, then the following statements hold:*

(1) $L^{\Phi}(\mu) = L^1(\mu)$,
(2) $\| \cdot \|_{(\Phi)}$ *is equivalent to* $\| \cdot \|_1$,
(3) $L^{\Psi}(\mu) = L^{\infty}(\mu)$,
(4) $\| \cdot \|_{(\Psi)}$ *is equivalent to* $\| \cdot \|_{\infty}$.

The following well known growth conditions for Young functions are of central importance for our discussion of the properties of Orlicz spaces:

Definition 2. *We say that a Young function satisfies the conditions*

(1) Δ_2*-condition if there is* $\lambda\epsilon\mathbb{R}$*, such that* $\forall s\epsilon\mathbb{R}$*:* $\Phi(2s) \leq \lambda\Phi(s)$*,*
(2) Δ_2^∞*-condition if there is* $\lambda\epsilon\mathbb{R}$ *and* $k > 0$ *such that* $\forall s \geq k$*:* $\Phi(2s) \leq \lambda\Phi(s)$*,*
(3) Δ_2^0*-condition if there is* $\lambda\epsilon\mathbb{R}$ *and* $k > 0$ *such that* $\forall 0 \leq s \leq k$*:* $\Phi(2s) \leq \lambda\Phi(s)$*.*

Let ℓ^Φ denote the Orlicz-sequence space. The following theorem can be found in [52]:

Theorem 2 (Lindenstrauss-Tsafriri)**.** *For a finite Young function* Φ *the following statements are equivalent*

(a) Φ *satisfies the* Δ_2^0*-condition,*
(b) $\ell^\Phi = m^\Phi$*,*
(c) ℓ^Φ *is separabel,*
(d) ℓ^Φ *contains no subspace isomorphic to* ℓ^∞*.*

For not purely atomic measure spaces the above isomorphy is in fact isometrical (see [80])

Theorem 3 (Turett)**.** *Let* Φ *be a finite Young function and let* (T, Σ, μ) *be a not purely atomic measure space. Then the following statement holds: if* Φ *does not satisfy the* Δ_2^∞*-condition, then* $L^\Phi(\mu)$ *contains a subspace isometrically isomorphic to* ℓ^∞*.*

The following theorem can be found e.g. in [39] (comp. also [47]):

Theorem 4. *Let* (T, Σ, μ) *be a not purely atomic measure space with* $\mu(T) < \infty$ *and* Φ *be finite. Then the following statements are equivalent:*

(1) Φ *satisfies the* Δ_2^∞*- condition,*
(2) $\mathcal{M}^\Phi = L^\Phi$*.*

Remark 1. *The above theorem also holds if* Φ *satisfies the* Δ_2 *- condition and* $\mu(T\backslash A) = \infty$*, where* A *denotes the set of atoms in* T*.*

If Φ satisfies the Δ_2-condition, then convergence in the norm is equivalent to convergence w.r.t. the modular (s. [39]):

Theorem 5. *If Φ satisfies the Δ_2-condition and $(x_n)_n$ be a sequence in $L^\Phi(\mu)$, then*

$$f^\Phi(x_n) \underset{n\to\infty}{\longrightarrow} 0 \iff \|x_n\|_{(\Phi)} \underset{n\to\infty}{\longrightarrow} 0.$$

This is also true for finite measure, if Φ is definite and satisfies the Δ_2^∞-condition.

1.1. Duality.

Definition 3. *Let Φ be a Young function, Ψ be its convex conjugate. E be the space of the equivalence classes of μ-measurable functions. Then we define the functional $N_\Phi : E \to \overline{\mathbb{R}}$ by*

$$N_\Phi(u) := \sup\left\{\left|\int_T v \cdot u d\mu\right| \;\middle|\; v \epsilon S_{f^\Psi}(1)\right\}.$$

It turns out that N_Φ is a finite norm on L^Φ; it is called Orlicz-norm:

$$N_\Phi(x) = \|x\|_\Phi = \sup\left\{\left|\int_T x \cdot y d\mu\right| \;\middle|\; \|y\|_{(\Psi)} \leq 1\right\}.$$

Thus the Orlicz-norm is the canonical norm of the dual space and $(L^\Phi(\mu), \|\cdot\|_\Phi)$ is isometrically embedded into the dual space of $L^\Psi(\mu)$.

Luxemburg- and Orlicz-Norm are equivalent. Moreover, Hölder's inequality holds (s. e.g. [47]):

Theorem 6. *Let Φ be the conjugate of Ψ, then the following statements hold*

(1) $\|x\|_{(\Phi)} \leq \|x\|_\Phi \leq 2\|x\|_{(\Phi)}$ *for all* $x \epsilon L^\Phi(\mu)$,
(2) *Hölder's inequality:* $\left|\int_T x \cdot y d\mu\right| \leq \|x\|_\Phi \cdot \|y\|_{(\Psi)}$.

As the conjugate of Ψ is again Φ, one can exchange the roles of Φ and Ψ and obtains $L^\Psi(\mu)$ as a subspace of $L^\Phi(\mu)^*$.

A precise description of the dual space of $\mathcal{M}^\Phi(\mu)$ is given by the following well known theorem, the proof of which for finite Lebesgue measure can be found in [47], s. [39]:

Theorem 7 (Duality). *Let (T, Σ, μ) be a σ-finite measure space and let Φ be finite, then every continuous linear functional f on $\mathcal{M}^\Phi(\mu)$ is represented by a function $y \in L^\Psi(\mu)$ via the formula*

$$\langle f, x\rangle = \int_T y(t)x(t)d\mu, \quad x \in L^\Phi.$$

If $\mathcal{M}^\Phi(\mu)$ is equipped with the Luxemburg-norm then $\|f\| = \|y\|_\Psi$, i.e.:

$$\left(\mathcal{M}^\Phi(\mu), \|\cdot\|_{(\Phi)}\right)^* = \left(L^\Psi(\mu), \|\cdot\|_\Psi\right).$$

Examples

1. For $1 < p < \infty$ and $\Phi(s) = |s|^p/p$ we have $\Psi(s) = |s|^q/q$, where $1/p+1/q = 1$, while Φ satisfies the Δ_2-condition and hence
$$\big(L^p(\mu), \|\cdot\|_p\big)^* = \big(L^q(\mu), \|\cdot\|_q\big).$$
2. For $\Phi(s) = |s|$ we have $\Psi(s) = \Phi_\infty(s)$ and
$$\big(L^1(\mu), \|\cdot\|_1\big)^* = \big(L^\infty(\mu), \|\cdot\|_\infty\big).$$

To discuss Frechét-differentiability and reflexivity we need a duality theorem, which states the conditions rendering Luxemburg- and Orlicz-norm changeable:

Theorem 8. *If Φ and Ψ are finite and $\mathcal{M}^\Phi(\mu) = L^\Phi(\mu)$, then*

(1) $\big(\mathcal{M}^\Psi(\mu), \|\cdot\|_\Psi\big)^* = \big(L^\Phi(\mu), \|\cdot\|_{(\Phi)}\big)$,
(2) $\big(\mathcal{M}^\Psi(\mu), \|\cdot\|_\Psi\big)^{**} = \big(L^\Psi(\mu), \|\cdot\|_\Psi\big)$.

Proof: Let $X := \big(\mathcal{M}^\Phi(\mu), \|\cdot\|_{(\Phi)}\big)$, then according to the previous theorem $X^* = (L^\Psi(\mu), \|\cdot\|_\Psi)$. Let now $U := (\mathcal{M}^\Psi(\mu), \|\cdot\|_\Psi)$ and $f \in U^*$. Due to the equivalence of Luxemburg and Orlicz norm there is - according to the duality theorem - a function $y \in L^\Phi(\mu) = \mathcal{M}^\Phi(\mu)$ with $\langle f, u\rangle = \int_T u \cdot y d\mu$ for all $u \in U$. Hence we obtain

$$\begin{aligned}\|f\| &= \sup\left\{\langle f, u\rangle \middle| u \in \mathcal{M}^\Psi, \|u\|_\Psi \le 1\right\}\\ &= \sup\left\{\int_T uy\, d\mu \;\middle|\; u \in \mathcal{M}^\Psi,\ \|u\|_\Psi \le 1\right\}.\end{aligned} \tag{1.1}$$

For $z \in S_{f^\Psi}(1)$, we construct a sequence

$$z_n(t) := \begin{cases} z(t) & \text{for } |z(t)| \le n \text{ and } t \in B^n, \\ 0 & \text{otherwise.}\end{cases}$$

Since z_n is bounded and has finite support, it can be approximated by step functions as in the proof of the previous theorem, hence $z_n \in \mathcal{M}^\Psi(\mu)$. Due to the monotonicity of the Orlicz norm by Fatou lemma:

$$\begin{aligned}\left|\int_T zyd\mu\right| \le \int_T |zy|d\mu &\le \sup_n \int_T |z_n y|d\mu = \sup_n \langle f, |z_n| \text{ sign } (y)\rangle\\ &\le \sup_n \|f\|\|z_n\|_\Psi \le \|f\|\|z\|_\Psi\end{aligned}$$

and hence by equation (1.1),

$$\|f\| = \sup\left\{\left|\int_T zyd\mu\right| \;\middle|\; z \in L^\Psi,\ \|z\|_\Psi \le 1\right\}.$$

We define $\langle f, z\rangle := \int_T zyd\mu$ for $z \in L^\Psi$. Because of Hölder's inequality, f is a continuous functional on L^Ψ. According to the duality theorem we have $L^\Psi = X^*$, i.e. $f \in X^{**}$. Due to [23] p.181, theorem 41.1 the canonical mapping $I : X \to X^{**}$ with $y \longmapsto \int_T (\cdot) y d\mu$ is norm-preserving, i.e. $\|f\| = \|y\|_{(\Phi)}$. Hence 1. holds. Due to

$$\big(L^\Phi(\mu), \|\cdot\|_{(\Phi)}\big)^* = \big(\mathcal{M}^\Phi(\mu), \|\cdot\|_{(\Phi)}\big)^* = \big(L^\Psi(\mu), \|\cdot\|_\Psi\big)$$

we obtain 2.

1.2. **Reflexivity.** Reflexivity is closely related to the the Δ_2-condition.

Theorem 9. *Let Φ and Ψ be conjugate Young functions and let the measure space be not purely atomic.*

(1) *If $\mu(T) < \infty$, then $L^\Phi(\mu)$ is reflexive if and only if Φ and Ψ satisfy the Δ_2^∞-condition,*
(2) *if $\mu(T) = \infty$ with $\mu(T \backslash A) = \infty$ (A set of atoms in T), then $L^\Phi(\mu)$ is reflexive if and only if Φ and Ψ satisfy the Δ_2-condition.*

For sequence spaces a similar, but weaker theorem is available:

Theorem 10. *If Φ and Ψ are finite then ℓ^Φ is reflexive if and only if Φ and Ψ satisfy the Δ_2^0-condition.*

1.3. **Separability.**

Theorem 11 (Lusin). *Let T be a compact Hausdorff space, $(T, \sum, \mu)$ the corresponding Bair measure space and let x be a measurable function. Then for every $\varepsilon > 0$ there is a continuous function y on T such that*

$$\mu\big(\{t \in T | x(t) - y(t) \neq 0\}\big) < \varepsilon.$$

Furthermore, if $\|x\|_\infty < \infty$ then y can be chosen to satisfy $\|y\|_\infty \leq \|x\|_\infty$.

The next theorem is a consequence of Lusin's theorem (compare [47]):

Theorem 12. *Let T be a compact Hausdorff space, (T, Σ, μ) be the corresponding Bair measure space and Φ and Ψ be finite. Then the continuous on T functions are dense in $\mathcal{M}^\Phi(\mu)$.*

Using the Stone-Weierstraß-theorem we obtain

Theorem 13. *If Φ and Ψ are finite, T be a compact subset of $\mathbb{R}^m$ and μ be the Lebesgue-measure, then $\mathcal{M}^\Phi(\mu)$ is separabel.*

2. FLAT CONVEXITY AND WEAK DIFFERENTIABILITY

Let X be a real normed space and $f : X \to \mathbb{R}$ be a continuous convex function, then the subdifferential

$$\partial f(x_0) := \{\phi \in X^* \mid \phi(x - x_0) \leq f(x) - f(x_0)\}$$

is a non-empty convex set (s. e.g. [31]). For $\phi \in \partial f(x_0)$ the graph of $[f(x_0) + \phi(\cdot - x_0)]$ is a supporting hyperplain of the epigraph $\{(x, s) \in X \times R \mid f(x) \leq s\}$ in $(x_0, f(x_0))$, and each supporting hyperplain can be represented in such a way.

The right-handed derivative

$$f'_+(x_0, x) := \lim_{t \downarrow 0} \frac{f(x_0 + tx) - f(x_0)}{t}$$

always exists and is finite (s.e.g. [31]) and the equation $f'_-(x_0, x) = -f'_+(x_0, -x)$ holds. By the theorem of Moreau-Pschenichnii [31]

$$f'_+(x_0, x) = \max_{\phi \in \partial f(x_0)} \phi(x), \quad f'_-(x_0, x) = \min_{\phi \in \partial f(x_0)} \phi(x).$$

Definition 4. *A convex set (with non-empty interior) is called flat, if every boundary point has a unique supporting hyperplain.*

The next theorem (s. ([35]) gives a characterization of flat level sets of a continuous convex function:

Theorem 14. *Let X be a real normed space and $f : X \to \mathbb{R}$ be a continuous convex function. Then for $r > \inf\{f(x) \mid x \in X\}$ the following statements (a) and (b) are equivalent:*

(a) *The convex set $S_f(r) := \{x \in X \mid f(x) \leq r\}$ is flat.*

(b) *For all boundary points x_0 of $S_f(r)$*

  (i) $[f'_+(x_0, \cdot) + f'_-(x_0, \cdot)] \in X^*$,

  (ii) *there exists a $c > 0$ with*

$$f'_-(x_0, x) = cf'_+(x_0, x)$$

  for all x with $f'_+(x_0, x) \geq 0$.

In particular, $S_f(r)$ is flat, if f Gâteaux-differentiable in $\{x \in X \mid f(x) = r\}$.

Proof: Let f be not constant. Due to the continuity of f the set $S_f(r)$ has a non-empty interior, and for every boundary point x_0 of $S_f(r)$ one has $f(x_0) = r > \inf f(x)$, hence $0 \notin \partial f(x_0)$.

(a) $\to$ (b). Let H be the unique supporting hyperplain of $S_f(r)$ in x_0. For every $\phi \in \partial f(x_0)$, $x_0 + \operatorname{Ker}\phi$ is a supporting hyperplain, hence

$$x_0 + \operatorname{Ker}\phi = H.$$

If $\phi_0 \in X^*$ represents the hyperplain H, then $\phi = \lambda\phi_0$ for a $\lambda \in \mathbb{R}$. The theorem of Moreau-Pschenichnii yields

$$\partial f(x_0) = \{\lambda\phi_0 \mid \lambda_1 \le \lambda \le \lambda_2\},$$

hence

$$\begin{aligned} &\left.\begin{aligned} f'_+(x_0,x) &= \lambda_2\phi_0(x) \\ f'_-(x_0,x) &= \lambda_1\phi_0(x) \end{aligned}\right\} \text{ for } \phi_0(x) \ge 0, \\ &\left.\begin{aligned} f'_+(x_0,x) &= \lambda_1\phi_0(x) \\ f'_-(x_0,x) &= \lambda_2\phi_0(x) \end{aligned}\right\} \text{ for } \phi_0(x) < 0. \end{aligned} \tag{2.1}$$

We conclude (i):

$$\left[f'_+(x_0,\cdot) + f'_-(x_0,\cdot)\right] = (\lambda_1+\lambda_2)\cdot\phi_0(\cdot) \in X^*$$

It remains to verify (ii): as $0 \notin \partial f(x_0)$, the relation sign λ_1 = sign $\lambda_2 \ne 0$ holds, consequently

$$f'_-(x_0,x) = \left(\lambda_1\lambda_2^{-1}\right)^{\operatorname{sign}\lambda_1} f'_+(x_0,x)$$

for x with $f'_+(x_0,x) \ge 0$.

(b) $\to$ (a). For $\phi \in \partial f(x_0)$ we have

$$f'_+(x_0,x) \ge \phi(x) \ge f'_-(x_0,x) \tag{2.2}$$

for all $x \in X$ and because of (ii)

$$f'_+(x_0,x) \ge \phi(x) \ge cf'_+(x_0,x) = f'_-(x_0,x) \tag{2.3}$$

for those x with $f'_+(x_0,x) \ge 0$. From $\phi(x) = 0$, (2.2), $f'_+(x_0,x) \ge 0$ and (2.3) it follows

$$f'_+(x_0,x) = f'_-(x_0,x) = 0,$$

i.e.

$$\operatorname{Ker}\phi \subset \operatorname{Ker}(f'_+ + f'_-) := H_0.$$

As $f'_+(x_0,\cdot) \ge f'_-(x_0,\cdot)$ and $0 \notin \partial f(x_0)$ it follows from (2.3) $H_0 \ne X$ and hence

$$\operatorname{Ker}\phi = H_0.$$

Let now H be a support hyperplain of $S_f(r)$ in x_0. Due to the Theorem of Mazur [31] the affine supporting set $H \times \{f(x_0)\}$ of the epigraph of f in $(x_0, f(x_0))$ can be extended to a support hyperplain in $X \times \mathbb{R}$, thus there exists $\phi \in \partial f(x_0)$ with

$$H \times \{f(x_0)\} \subset \big\{(x, f(x_0) + \phi(x - x_0)) \mid x \in X\big\}.$$

Hence

$$H \subset \big\{x \mid \phi(x - x_0) = 0\big\} = x_0 + \operatorname{Ker}\phi = x_0 + H_0,$$

consequently

$$H = x_0 + H_0.$$

Of particular interest is the positive homogeneous case (s. [35]):

Theorem 15. *If f is nonnegativ and positively homogeneous, then $S_f(r)$ is flat convex if and only if f is Gâteaux-differentiable in $\{x \in X \mid f(x) > 0\}$.*

Proof: If $S_f(r)$ is flat convex for an $r > 0$, then, because f is positive homogeneous, all level sets $S_f(r)$ are flat convex. Now we have

$$\begin{aligned} f'_+(x_0, x_0) &= \lim_{t \downarrow 0} \frac{f(x_0 + tx_0) - f(x_0)}{t} = f(x_0) \\ &= \lim_{t \downarrow 0} \frac{f(x_0 - tx_0) - f(x_0)}{-t} = f'_-(x_0, x_0) \end{aligned}$$

Hence $c = 1$ in (ii) and thus $f'_+(x_0, x) = f'_-(x_0, x)$ for $f'_+(x_0, x) \geq 0$.

If $f'_+(x_0, x) < 0$, then

$$f'_+(x_0, -x) \geq f'_-(x_0, -x) = -f'_+(x_0, x) > 0,$$

and with $c = 1$ in (ii) it follows that

$$f'_+(x_0, -x) = f'_-(x_0, -x).$$

Therefore we obtain using (i):

$$\begin{aligned} 2f'_+(x_0, x) &= -2f'_-(x_0, -x) = -(f'_+(x_0, -x) + f'_-(x_0, -x)) \\ &= f'_+(x_0, x) + f'_-(x_0, x). \end{aligned}$$

Hence $f'_+(x_0, x) = f'_-(x_0, x)$.

Definition 5. *A normed space is called flat convex, if the closed unit ball is flat.*

If f is a norm, the above yields the following theorem.

Theorem 16 (Mazur)**.** *A normed space X is flat convex if and only if the norm is Gâteaux-differentiable in $X \setminus \{0\}$.*

3. FLAT CONVEXITY AND GÂTEAUX -DIFFERNTIABILITY OF ORLICZ SPACES

Let $\Phi : \mathbb{R} \to \mathbb{R}$ a finite Young function. Let (T, Σ, μ) be a σ-finite measure space, and $L^\Phi(\mu)$ be the Orlicz space determined via Φ and μ, equipped with the Luxemburg norm $\|\cdot\|_{(\Phi)}$.

We consider the unit ball in $\mathcal{M}^\Phi(\mu)$. If $x \in \mathcal{M}^\Phi(\mu)$, then

$$\|x\|_{(\Phi)} = 1 \text{ if and only if } \int_T \Phi(x)d\mu = 1. \tag{3.1}$$

According to Theorem 14 the level set $S_f(r)$ is flat, if f is Gâteaux-differentiable in $\{x \in X \mid f(x) = r\}$.

Lemma 1 ([41]). *The right and left-sided derivatives of the modular $f^\Phi : \mathcal{M}^\Phi(\mu) \to \mathbb{R}$ defined by*

$$f^\Phi(x) = \int_T \Phi(x)d\mu$$

for $x_0 \in \mathcal{M}^\Phi(\mu)$ can be represented as follows:

$$\begin{aligned} (f^\Phi)'_+(x_0, x) &= \int_{\{x>0\}} x\Phi'_+(x_0)d\mu + \int_{\{x<0\}} x\Phi'_-(x_0)d\mu \\ (f^\Phi)'_-(x_0, x) &= \int_{\{x>0\}} x\Phi'_-(x_0)d\mu + \int_{\{x<0\}} x\Phi'_+(x_0)d\mu. \end{aligned} \tag{3.2}$$

If Φ is differentiable, then f^Φ is Gâteaux-differentiable and

$$(f^\Phi)'(x_0, x) = \int_T x\Phi'(x_0)d\mu. \tag{3.3}$$

Proof: For the difference quotient we obtain:

$$\begin{aligned} \frac{f^\Phi(x_0 + \tau x) - f^\Phi(x_0)}{\tau} &= \int_T \frac{\Phi(x_0(t) + \tau x(t)) - \Phi(x_0(t))}{\tau} d\mu \\ &= \int_{x(t)>0} \frac{\Phi(x_0(t) + \tau x(t)) - \Phi(x_0(t))}{\tau x(t)} x(t)d\mu \\ &\quad + \int_{x(t)<0} \frac{\Phi(x_0(t) + \tau x(t)) - \Phi(x_0(t))}{\tau x(t)} x(t)d\mu. \end{aligned}$$

By the monotonicity w.r.t. τ of

$$\frac{\Phi(s_0 + \tau s) - \Phi(s_0)}{\tau}$$

for all $s_0, s \in \mathbb{R}$, the above representation (3.2) of $(f^\Phi)'_+$ and $(f^\Phi)'_-$ follows. Let Φ be differentiable. According to (3.2) $f'(x_0, x)$ exists and we have (3.3).

Lemma 2 ([41]). *Let T_0, T_1 and $T_2 \in \Sigma$ be disjoint sets with $0 < \mu(T_i) < \infty$ for $i = 0, 1, 2$. If there exists an $s_0 \geq 0$ for which $\Phi'_+(s_0) \neq \Phi'_-(s_0)$ and $\Phi(s_0) \cdot \mu(T_0) < 1$, then $\mathcal{M}^\Phi(\mu)$ is not flat convex.*

Proof: The set of discontinuities of a function, monotonically increasing on $[a, b]$ is at most countable (s. [64], S. 229). Hence there exists $s_1 > 0$ with $\Phi'_+(s_1) = \Phi'_-(s_1)$ and

$$1 - \Phi(s_0)\mu(T_0) - \Phi(s_1)\mu(T_1) > 0.$$

As Φ is continuous, one can choose $s_2 \in \mathbb{R}$ such that

$$\Phi(s_0) \cdot \mu(T_0) + \Phi(s_1) \cdot \mu(T_1) + \Phi(s_2) \cdot \mu(T_2) = 1$$

For the functions

$$\begin{aligned} x_0 &= s_0\chi_{T_0} + s_1\chi_{T_1} + s_2\chi_{T_2}, \\ x_1 &= \chi_{T_0}, \\ x_2 &= \chi_{T_1} \end{aligned}$$

we have

$$\begin{aligned} 0 < (f^\Phi)'_+(x_0, x_1) = \mu(T_0) \cdot \Phi'_+(s_0) \neq (f^\Phi)'_-(x_0, x_1) = \mu(T_0) \cdot \Phi'_-(s_0), \\ 0 < (f^\Phi)'_+(x_0, x_2) = \mu(T_1) \cdot \Phi'_+(s_1) = \mu(T_1) \cdot \Phi'_-(s_1) = (f^\Phi)'_-(x_0, x_2). \end{aligned}$$

Hence condition (ii) of Theorem 14 is not satisfied, and thus $\mathcal{M}^\Phi(\mu)$ not flat convex.

Theorem 17 ([41]). *Let μ be not purely atomic. Then $\mathcal{M}^\Phi(\mu)$ is flat convex if and only if Φ is differentiable.*

Proof: follows from Lemma 2 and Theorem 1.

The derivative of the norm $\|\cdot\|_{(\Phi)}$ is defined as follows. Let $x_0 \in \mathcal{M}^\Phi(\mu)$ and $\|x_0\|_{(\Phi)} = 1$. The graph of the function $x \to (f^\Phi)'(x_0, x - x_0) + f^\Phi(x_0)$ is a supporting hyperplain of the epigraph of f^Φ in $(x_0, f^\Phi(x_0)) = (x_0, 1) \in \mathcal{M}^\Phi \times \mathbb{R}$. This means that the hyperplain $\{x \in \mathcal{M}^\Phi \mid (f^\Phi)'(x_0, x - x_0) = 0\}$ supports the unit ball of $\mathcal{M}^\Phi$ in x_0. If we denote $\|\cdot\|_{(\Phi)}$ by p_Φ then $p'_\Phi(x_0, x_0) = 1$ (compare proof of Theorem 15) and $p'_\Phi(x_0, \cdot)$ is a multiple of $(f^\Phi)'(x_0, \cdot)$. Taking into account that

$$\int_T \Phi'(x_0)x_0 d\mu \geq \int_T \Phi(x_0) d\mu = 1,$$

we obtain that

$$x \longmapsto \frac{\displaystyle\int x\Phi'(x_0)d\mu}{\displaystyle\int x_0\Phi'(x_0)d\mu}, \quad \|x_0\|_{(\Phi)} = 1, \tag{3.4}$$

is the derivative of $\|\cdot\|_{(\Phi)}$ (in $\mathcal{M}^\Phi$).

If the measure μ is purely atomic, then the differentiability of Φ is not a necessary condition for flat convexity of $\mathcal{M}^\Phi(\mu)$.

In the sequel let $S = \{s \in \mathbb{R} \mid \Phi \text{ in } s \text{ not differentiable}\}$.

Theorem 18 ([41]). *Let (T, Σ, μ) be purely atomic and consist of more than 2 atoms. Then $\mathcal{M}^\Phi(\mu)$ is flat convex if and only if for all $s \in S$ and all atoms $A \in \Sigma$*

$$\Phi(s) \cdot \mu(A) \geq 1.$$

Proof: Necessity follows immediately from Lemma 2.

Let $x_0 \in \mathcal{M}^\Phi(\mu)$ and $\|x_0\|_{(\Phi)} = 1$. According to Lemma 1 and the condition $\Phi(s) \cdot \mu(A) \geq 1$, it is sufficient to consider $x_0 = r\chi_A$ for an atom $A \in \Sigma$ and $r \in S$.

As $\Phi(r) \cdot \mu(A) = 1$ we have $\Phi'_+(r) \neq 0$.

As $0 \notin S$ and $\Phi'(0) = 0$ it follows from Lemma 1 that

$$\begin{aligned}(f^\Phi)'_+(x_0, x) &= \int_{\{x>0\}} x\Phi'_+(x_0)d\mu + \int_{\{x<0\}} x\Phi'_-(x_0)d\mu \\ &= \Phi'_+(r)\int_{\{x>0\}} x\cdot\chi_A d\mu + \Phi'_-(r)\int_{\{x<0\}} x\cdot\chi_A d\mu,\end{aligned}$$

and in in the same way

$$(f^\Phi)'_-(x_0, x) = \Phi'_-(r)\int_{\{x>0\}} x\chi_A d\mu + \Phi'_+(r)\int_{\{x<0\}} x\chi_A d\mu.$$

Hence we obtain (i) with $f'_+(x_0, x) + f'_-(x_0, x) = \mu(A)[\Phi'_+(r) + \Phi'_-(r)]x\mid_A$, where A is an atom, i.e. x is a constant on A. This implies

$$x \longmapsto f'_+(x_0, x) + f'_-(x_0, x) = \mu(A)\big[\Phi'_+(r) + \Phi'_-(r)\big]x\mid_A \in (M^\Phi)^*.$$

(ii) If $x > 0$ on A, then $f'_-(x_0, x) = \frac{\Phi'_-(r)}{\Phi'_+(r)} f'_+(x_0, x)$, because in that case

$$\int_{\{x<0\}} x\cdot\chi_A d\mu = 0 \quad \text{and} \quad \Phi'_+(r) > 0.$$

If, however, $x < 0$ on A, a suitable multiplyer is $\frac{\Phi'_+(r)}{\Phi'_-(r)}$ with $\Phi'_+(r) < 0$. From Theorem 14 it then follows that $\mathcal{M}^\Phi(\mu)$ is flat convex, and the proof is complete.

4. LOCAL UNIFORM CONVEXITY, STRONG SOLVABILITY AND FRECHET-DIFFERENTIABILITY OF THE CONJUGATE

There are several different ways to generalize the notion of uniform convexity of functions [50] by allowing for different convexity modules at different points of the space.

Definition 6. *A monotonically increasing function $\tau : \mathbb{R}_+ \to \mathbb{R}_+$ with $\tau(0) = 0$, $\tau(s) > 0$ for $s > 0$ and $\tau(s)/s \underset{s\to\infty}{\longrightarrow} \infty$ is called a convexity module. A continuous function $f : X \to \mathbb{R}$, where X is a normed space, is called locally uniformly convex if*

(a) *for all $x \in X$ a convexity module τ_x exists, such that for all $y \in X$*
$$\frac{1}{2}\big(f(y) + f(x)\big) \geq f\left(\frac{x+y}{2}\right) + \tau_x(\|x-y\|);$$

(b) *for all $x \in X$ and all $x^* \in \partial f(x)$ a convexity module τ_{x,x^*} exists, such that for all $y \in X$*
$$f(y) - f(x) \geq \langle y-x, x^*\rangle + \tau_{x,x^*}(\|x-y\|);$$

(c) *for all $x \in X$ a convexity module τ_x exists, such that for all $y \in X$*
$$\frac{1}{2}(f(x+y) + f(x-y)) \geq f(x) + \tau_x(\|y\|).$$

It is easily seen that: a) $\Rightarrow$ b) and b) $\Rightarrow$ c). If the function f satisfies a), then by the subgradient inequality
$$\begin{aligned}\frac{1}{2}\big(f(x) + f(y)\big) &\geq f\left(\frac{x+y}{2}\right) + \tau_x(\|x-y\|)\\ &\geq f(x) + \left\langle \frac{x+y}{2} - x, x^*\right\rangle + \tau_x(\|x-y\|)\end{aligned}$$
defining $\tau_{x,x^*}(s) := 2\tau_x(s)$, and the property b) follows. If a function f satisfies b), then by the subgradient inequality:
$$\begin{aligned}\frac{1}{2}f(x+y) &\geq \frac{1}{2}\big(\langle y, x^*\rangle + f(x) + \tau_{x,x^*}(\|y\|)\big)\\ &\geq \frac{1}{2}\big(f(x) - f(x-y) + f(x) + \tau_{x,x^*}(\|y\|)\big),\end{aligned}$$
hence the property (c) for $\tau_x(s) := \tau_{x,x^*}(s)/2$.

The converse, however is not true, because the function $f(x) = e^x$ satisfies (c) using the convexity module $\tau_x(s) := e^x(\cosh(s) - 1)$, yet does not have bounded level sets, hence cannot satisfy b).

The strictly convex function $f(x) := (x+1)\log(x+1) - x$ satisfies property (b), because for $h > 0$, due to the strict convexity of f

$$\begin{aligned}0 &< \frac{1}{h}\big(f(x+h) - f(x) - f'(x)h\big)\\ &= \frac{1}{h}\left((x+1)\log\frac{x+h+1}{x+1} + h\left(\log\frac{x+h+1}{x+1} - 1\right)\right) \underset{h\to\infty}{\longrightarrow} \infty.\end{aligned}$$

However, it violates (a) at the origin, because

$$\lim_{y\to\infty}\frac{1}{y}\left(\frac{1}{2}f(y) - f\left(\frac{y}{2}\right)\right) = \frac{1}{2}\log 2 < \infty.$$

In the sequel we will mean by locally uniformly convex functions always those with property (b).

Remark 2. *Lovaglia in* [53] *investigates locally another lass of uniformly convex norms. The squares of these norms are, however, locally uniformly convex in the sense of (b),* [45].

Definition 7. *A function f has a strong minimum k_0 on a subset K of a Banach space X, if the set of minimal solutions $M(f,K)$ of f on K consists of $\{k_0\}$ and if for every sequence $\{k_n\}_{n=1}^{\infty} \subset K$ with*

$$\lim_{n\to\infty} f(k_n) = f(k_0) \quad \text{it follows that:} \quad \lim_{n\to\infty} k_n = k_0.$$

The problem of minimizing f on K is then called strongly solvable.

Lemma 3 ([45])**.** *Let X be a reflexive Banach space and let $f : X \to \mathbb{R}$ be convex and continuous then the following statements are equivalent:*

(1) *f has a strong minimum on every closed hyperplain and on X;*
(2) *f has a strong minimum on every closed half-space and on X;*
(3) *f has a strong minimum on every closed convex subset of X.*

Proof: Without loosing generality we can assume that $f(0) = 0$ and $f(x) > 0$ for $x \neq 0$. Otherwise we can consider $g(x) := f(x_0 - x) - f(x_0)$, where $f(x_0) = \min\{f(x)|x \in X\}$.

1. $\Rightarrow$ 2. Let $G_\alpha := \{x|\langle x_0^*, x\rangle \geq \alpha\}$. If $0 \in G_\alpha$, then $g_0 = 0$ is the strong minimum of f on G_α. If $0 \notin G_\alpha$, then it follows that $\alpha \neq 0$. Let x be an interior point of G_α, then $f(x) > 0$ and with $x_\lambda := (1-\lambda)\cdot 0 + \lambda x$ for $\lambda \in (0,1)$ it follows that $f(x_\lambda) \leq \lambda f(x) < f(x)$. Hence x cannot be a minimum of f on G_α. Let g_0 be the strong minimum of f on $H_\alpha := \{x|\langle x_0^*, x\rangle = \alpha\}$. Then by reasoning as above $f(g_0) = \inf\{f(g)|g \in G_\alpha\}$.

Let now $f(g_n) \to f(g_0)$ with $g_n \in G_\alpha$. Then $\langle x_0^*, g_n\rangle \to \alpha$. To prove it assume there is $\varepsilon > 0$ and a subsequence $\langle x_0^*, g_{n_k}\rangle \geq \alpha + \varepsilon$. Then $g_{n_k} \in G_{\alpha+\varepsilon}$, but we have for every $\varepsilon > 0$

$$\min\left\{f(x) | \langle x_0^*, x\rangle \geq \alpha + \varepsilon\right\} > f(g_0).$$

Let $f(g_\varepsilon) = \inf\{f(g) | g \in G_{\alpha+\varepsilon}$, then we obtain in the same way as above $\langle x_0^*, g_\varepsilon\rangle = \alpha + \varepsilon$. As 0 is the strong minimum of f on X, the mapping $\lambda \mapsto f(\lambda g_\varepsilon)$ is strictly monotonically increasing on $[0, 1]$. The mapping $\lambda \mapsto \phi(\lambda) = \langle x_0^*, \lambda g_\varepsilon\rangle$ is continuous on $[0, 1]$ and we have: $\phi(0) = 0$ and $\phi(1) = \alpha + \varepsilon$. Hence there is a $\lambda_\alpha \in (0, 1)$ with $\phi(\lambda_\alpha) = \alpha$, i.e. $g_\alpha := \lambda_\alpha g_\varepsilon \in H_\alpha$. Therefore we obtain

$$f(g_0) \leq f(g_\alpha) < f(g_\varepsilon) = \min\left\{f(x) | \langle x_0^*, x\rangle \geq \alpha + \varepsilon\right\}$$

hence $f(g_{n_k}) \geq f(g_\varepsilon)$, a contradiction.
Since $\alpha / \langle x_0^*, g_n\rangle \leq 1$ we have

$$f\left(\frac{\alpha}{\langle x_0^*, g_n\rangle} g_n\right) \leq \frac{\alpha}{\langle x_0^*, g_n\rangle} f(g_n) \to f(g_0).$$

Moreover

$$\frac{\alpha}{\langle x_0^*, g_n\rangle} g_n \in \left\{x | \langle x_0^*, x\rangle = \alpha\right\},$$

and hence $f(g_0) \leq f\left(\frac{\alpha}{\langle x_0^*, g_n\rangle} g_n\right)$. We conclude that $\left\{\frac{\alpha}{\langle x_0^*, g_n\rangle} g_n\right\}$ is a minimizing sequence of f on H_α, and hence

$$\frac{\alpha}{\langle x_0^*, g_n\rangle} g_n \to g_0,$$

thus $g_n \to g_0$.

2. $\Rightarrow$ 3. If K is convex and closed, $0 \notin K$ and $r := \inf f(K)$, then $r > 0$ and the interior of $S_f(r)$ is non empty. According to the separation theorem of Eidelheit [31], there is a half space G_α with $K \subset G_\alpha$ and $G_\alpha \cap \operatorname{Int}(S_f(r)) = \emptyset$. If g_0 is the strong minimum of f on G_α, then $f(g_0) \leq r$, but $f(g_0) < r$ is impossible, because in that case g_0 would belong to the interior of $S_f(r)$. Thus g_0 is the minimal solution of f on K and because $K \subset G_\alpha$ it is also the strong minimum.

If $0 \in K$ and (k_n) is a sequence in K with $f(k_n) \to 0$, we can assume that K is a subset of a half-space G. The point 0 is a minimum and hence the strong minimum of f on G and therefore also the strong minimum of f on K.

Definition 8. *A convex function is called bounded, if the image of every bounded set is bounded.*

In order to establish a relation between strong solvability and local uniform convexity, we need

Lemma 4. *Let X be a reflexive Banach space and $f : X \to \mathbb{R}$ be convex and continuous.*

Then $\dfrac{f(x)}{\|x\|} \underset{\|x\|\to\infty}{\longrightarrow} \infty$ *holds, if and only if the convex conjugate f^* is bounded.*

Proof: Let f^* be bounded. Suppose there is a sequence $\{x_n\}_{n=1}^{\infty}$ with $\|x_n\| \underset{n\to\infty}{\longrightarrow} \infty$ for which always $f(x_n)/\|x_n\| < M$ for some $M \in \mathbb{R}$. Then there is a sequence

$$\{x_n^*\}_{n=1}^{\infty} \subset X^* \quad \text{with} \quad \|x_n^*\| = 1 \quad \text{and} \quad \langle x_n^*, x_n\rangle = \|x_n\|.$$

However

$$\begin{aligned} f^*\big(2Mx_n^*\big) &= \sup_{x\in X} \big\{2M\langle x_n^*, x\rangle - f(x)\big\} \\ &\geq \|x_n\| \left(2M\left\langle x_n^*, \frac{x_n}{\|x_n\|}\right\rangle - \frac{f(x_n)}{\|x_n\|}\right) \geq M\|x_n\|, \end{aligned}$$

and because of the boundedness of f^* we get a contradiction.

Conversely let $\|x^*\| \leq r$, then there is $\varrho \in \mathbb{R}$, such that $f(x)/\|x\| \geq r$ for $\|x\| \geq \varrho$. Therefore

$$\begin{aligned} f^*(x^*) &= \sup_{x\in X} \{\langle x, x^*\rangle - f(x)\} \leq \sup_{x\in X} \left\{\left(r - \frac{f(x)}{\|x\|}\right)\|x\|\right\} \\ &\leq \sup_{\|x\|\leq\varrho} \left\{\left(r - \frac{f(x)}{\|x\|}\right)\|x\|\right\} + \sup_{\|x\|\geq\varrho} \left\{\left(r - \frac{f(x)}{\|x\|}\right)\|x\|\right\} \\ &\leq \sup_{\|x\|\leq\varrho} \big\{r\|x\| - f(x)\big\} \leq r\varrho - \inf f\big(\{x \mid \|x\| \leq \varrho\}\big). \end{aligned}$$

In order to estimate $-\inf f\big(\{x \mid \|x\| \leq \varrho\}\big)$ from above, let finally $x_0^* \in \partial f(0)$, then $f(x) \geq f(0) - \|x\|\|x_0^*\| \geq f(0) - \rho\|x_0^*\|$ and thus

$$-\inf f\big(\{x \mid \|x\| \leq \varrho\}\big) \leq \rho\|x_0^*\| - f(0).$$

Lemma 5. *If X is a reflexive Banach space, $f : X \to \mathbb{R}$ is continuous and convex, and f^* is bounded, then for every closed convex set K the set of minimal solutions $M(f, K) \neq \emptyset$.*

Proof: As f^* is bounded, then according to Lemma 4 all level sets of f are bounded. So due to the theorem of Mazur-Schauder (s. e.g. [31]) $M(f, K) \neq \emptyset$ for an arbitrary closed convex set K.

The following lemma can be found in [74]:

Lemma 6. *Let $f : X \to \mathbb{R}$ be convex and bounded, then f is Lipschitz-continuous on bounded sets.*

Proof: Let B be the unit ball in X, $\varepsilon > 0$, and S be a bounded subset of X. Then $S + \varepsilon B$ is bounded. Let α_1 and α_2 be lower and upper bounds of f on $S + \varepsilon B$. Let $x, y \in S$ with $x \neq y$ and $\lambda := \frac{\|y-x\|}{\varepsilon+\|y-x\|}$. Writting further $z := y + \frac{\varepsilon}{\|y-x\|}(y-x)$, then $z \in S + \varepsilon B$ and $y = (1-\lambda)x + \lambda z$. From the convexity of f we obtain:

$$f(y) \leq (1-\lambda)f(x) + \lambda f(z) = f(x) + \lambda\big(f(z) - f(x)\big),$$

and hence

$$f(y) - f(x) \leq \lambda(\alpha_2 - \alpha_1) = \frac{\alpha_2 - \alpha_1}{\varepsilon}\frac{\varepsilon}{\varepsilon + \|y-x\|}\|y-x\| < \frac{\alpha_2 - \alpha_1}{\varepsilon}\|y-x\|.$$

Exchanges the roles of x and y, one obtains the Lipschitz-continuity f on S.

Theorem 19. *Let X be a reflexive Banach space and $f : X \to \mathbb{R}$ a strictly convex and bounded function, whose conjugate f^* is also bounded. Then f has a strong minimum on every closed convex set, if and only if f^* is Fréchet-differentiable.*

Proof: Let f^* be Fréchet-differentiable. According to Lemma 3 it suffices to show that f has a strong minimum on every closed hyperplain and on X. By [3] f^* is Fréchet-differentiable at x^*, if and only if the function $f - \langle x^*, \cdot\rangle$ has a strong minimum on X. As f^* is differentiable at 0, f has a strong minimum on X.

Let $H := \{x | \langle x_0^*, x\rangle = r\}$ for $x_0^* \neq 0$. By Lemma 5 we have $M(f, H) \neq \emptyset$. Let $h_0 \in M(f, H)$. According to [31] there is a $x_1^* \in \partial f(h_0)$ with $\langle x_1^*, h - h_0\rangle = 0$ for all $h \in H$. Due to the subgradient inequality $\langle x_1^*, x - h_0\rangle \leq f(x) - f(h_0)$. Setting $f_1 := f - \langle x_1^*, \cdot\rangle$, it follows that $f_1(x) \geq f_1(h_0)$ for all $x \in X$. As f^* is differentiable at x_1^*, h_0 is the the strong minimum of f_1 on X. In particular h_0 is also the strong minimum of f on H, due to

$$f|H = (f_1 + \langle x_1^*, h_0\rangle)|H$$

because for all $h \in H$ we have:

$$f_1(h) + \langle x_1^*, h_0\rangle = f(h) - \langle x_1^*, h\rangle + \langle x_1^*, h_0\rangle = f(h) - \langle x_1^*, h - h_0\rangle = f(h).$$

In order to prove the converse let $x_0^* \in X^*$. According to [3] we have to show that $f_1 := f - \langle x_0^*, \cdot\rangle$ has a strong minimum on X. As $f_1^*(x^*) = f^*(x_0^* + x^*)$ for all $x^* \in X^*$, apparently f_1^* is bounded and hence f_1 has, according to Lemma 5, a minimum x_0 on X. If $x_0^* = 0$, then $f_1 = f$ and hence x_0 is the strong minimum of f_1 on X.

Let $x_0^* \neq 0$ and $f_1(x_n) \to f_1(x_0)$. For $\varepsilon > 0$,

$$K_1 := \big\{x \in X | \langle x_0^*, x\rangle \geq \langle x_0^*, x_0\rangle + \varepsilon\big\}, \quad K_2 := \big\{x \in X | \langle x_0^*, x\rangle \leq \langle x_0^*, x_0\rangle - \varepsilon\big\},$$

because of the strict convexity of f_1:

$$\min\big\{\min(f_1, K_1), \min(f_1, K_2)\big\} > f_1(x_0).$$

It follows $\langle x_0^*, x_n\rangle \to \langle x_0^*, x_0\rangle$ (otherwise there is a subsequence $\{x_{n_k}\}$ in K_1 or K_2, contradicting $f_1(x_n) \to f_1(x_0)$).

For $H := \{h | \langle x_0^*, x_0\rangle = \langle x_0^*, h\rangle\}$ by Ascoli formula [31]

$$\min\big\{\|x_n - h\| | h \in H\big\} = \frac{|\langle x_0^*, x_n\rangle - \langle x_0^*, x_0\rangle|}{\|x_0^*\|} \to 0.$$

For $h_n \in M(\|x_n - \cdot\|, H)$ we conclude that $x_n - h_n \to 0$. The level sets of f_1 are bounded due to Lemma 4, hence so are the sequences (x_n) and (h_n).

According to Lemma 6, f is Lipschitz-continuous on bounded sets, i.e. there is a constant L such that

$$|f(x_n) - f(h_n)| \leq L\|x_n - h_n\| \to 0.$$

We obtain:

$$\begin{aligned} f_1(h_n) &= f(h_n) - \langle x_0^*, h_n\rangle \\ &= \big(f(h_n) - f(x_n)\big) + \big(f(x_n) - \langle x_0^*, x_n\rangle\big) - \langle x_0^*, h_n - x_n\rangle \to f_1(x_0). \end{aligned}$$

On H the functions f and f_1 differ only by a constant. As f has a strong minimum on H, this also holds for f_1, hence $h_n \to x_0$ and thus $x_n \to x_0$. This completes the proof.

Theorem 20. *Let X be a reflexive Banach space and f be a continuous convex function on X. If f is locally uniformly convex and f^* is bounded, then f has a strong minimum on every closed convex set.*

Proof: Let f be locally uniformly convex. As f^* is bounded, then due to Theorem 4 all level sets of f are bounded and hence according to Lemma 5 $M(f, K) \neq \emptyset$ for an arbitrary closed convex set K. Let now $k_0 \in M(f, K)$, then by the characterization theorem of convex optimization [31] and the theorem of Moreau–Pschenitschnii there exists

$$x_0^* \in \partial f(k_0) \quad \text{with} \quad \langle k - k_0, x_0^*\rangle \geq 0 \quad \text{for all} \quad k \in K.$$

If τ is the convexity module of f belonging to k_0 and x_0^*, then for an arbitrary minimizing sequence $\{k_n\}$ we have

$$f(k_n) - f(k_0) \geq \langle k_n - k_0, x_0^* \rangle + \tau(\|k_n - k_0\|) \geq \tau(\|k_n - k_0\|).$$

Thus $\lim_{n\to\infty} k_n = k_0$, and the proof is complete.

Corollary 1. *Let X be a reflexive Banach space and f and g be continuous convex functions on X. If f is locally uniformly convex and f^* is bounded, then $f + g$ is locally uniformly convex and has a strong minimum on every closed convex set.*

Proof: First we show that $f + g$ is locally uniformly convex. Let $x \in X$ and $x_f^* \in \partial f(x)$, then there is a convexity module τ_{x,x_f^*}, such that for all $y \in X$:

$$f(y) - f(x) \geq \langle y - x, x_f^* \rangle + \tau_{x,x_f^*}(\|x - y\|).$$

If $x_g^* \in \partial g(x)$, then by the subgradient inequality:

$$f(y) + g(y) - (f(x) + g(x)) \geq \langle y - x, x_f^* + x_g^* \rangle + \tau_{x,x_f^*}(\|x - y\|).$$

Having established the local uniform convexity of $f + g$ we now show that $(f + g)^*$ is bounded. According to Theorem 4 this happens if and only if

$$\frac{f(x) + g(x))}{\|x\|} \underset{\|x\|\to\infty}{\longrightarrow} \infty.$$

Now for $x_g^* \in \partial g(0)$ we have $g(x) - g(0) \geq \langle x - 0, x_g^* \rangle \geq -\|x\|\|x_g^*\|$, hence $g(x)/\|x\| \geq g(0)/\|x\| - \|x_g^*\| \geq c \in \mathbb{R}$ for $\|x\| \geq r > 0$. Therefore

$$\frac{f(x) + g(x))}{\|x\|} \geq \frac{f(x)}{\|x\|} + c \underset{\|x\|\to\infty}{\longrightarrow} \infty,$$

and the proof is complete.

In [45] the main statement of the theorem that follows is proved for f^2, without requiring the boundedness of f^*. For the equivalence of strong solvability and local uniform convexity we need in addition the boundedness of f.

Theorem 21. *Let X be a reflexive Banach space, and $f : X \to \mathbb{R}$ be a bounded, strictly convex function, whose conjugate f^* is also bounded. Then f has a strong minimum on every closed convex set, if and only if f is locally uniformly convex.*

Proof: If f is locally uniformly convex, then strong solvability follows from Theorem 20.

Conversely, let f have the strong minimum property. Then by Theorem 19 f^* is Fréchet-differentiable. If x_0^* is the Fréchet-gradient of f in x_0, then $f(\cdot) - \langle \cdot, x_0^* \rangle$ has [3] the global strong minimum at x_0. We write

$$\tau(s) := \inf_{\|y\|=s} \left\{ f(x_0 + y) - f(x_0) - \langle y, x_0^* \rangle \right\}.$$

In fat $\tau(0) = 0$ and $\tau(s) > 0$ for $s > 0$, and τ is monotonically increasing. Let $s_2 > s_1 > 0$ and $\|z\| = s_1$, and $y := \frac{s_2}{s_1} z$ due to the monotonicity of the difference quotient

$$\frac{f(x_0 + \frac{s_1}{s_2} y) - f(x_0)}{s_1/s_2} \leq f(x_0 + y) - f(x_0),$$

hence

$$\begin{aligned} f(x_0 + z) - f(x_0) - \langle z, x_0^* \rangle &\leq \frac{s_1}{s_2} (f(x_0 + y) - f(x_0) - \langle y, x_0^* \rangle) \\ &\leq f(x_0 + y) - f(x_0) - \langle y, x_0^* \rangle. \end{aligned}$$

Finally, using Theorem 4

$$\begin{aligned} \frac{\tau(s)}{s} &= -\frac{f(x_0)}{s} + \inf_{\|y\|=s} \left\{ \frac{f(x_0 + y)}{\|y\|} - \langle \frac{y}{\|y\|}, x_0^* \rangle \right\} \\ &\geq -\frac{f(x_0)}{s} - \|x_0^*\| + \inf_{\|y\|=s} \left\{ \frac{f(x_0 + y)}{\|y\|} \right\} \underset{s \to \infty}{\longrightarrow} \infty. \end{aligned}$$

On a reflexive Banach space the conjugate of a bounded convex function does not have to be bounded.

Example 1. *Let $f : l^2 \to \mathbb{R}$ be defined by $f(x) = \sum_{i=1}^{\infty} \varphi_i(x^{(i)})$, where $\varphi_i : \mathbb{R} \to \mathbb{R}$ is given by*

$$\varphi_i(s) := \begin{cases} \dfrac{s^2}{2} & \text{for} \quad |s| \leq 1. \\ \dfrac{i}{i+1} |s|^{\frac{i+1}{i}} + \dfrac{1-i}{2(i+1)} & \text{for} \quad |s| > 1. \end{cases}$$

f is bounded, because $f(x) \leq \|x\|_2^2$ for all $x \in l^2$. The conjugate function of f is

$$f^*(x) = \sum_{i=1}^{\infty} \Psi_i(x^{(i)}), \quad \textit{where} \quad \Psi_i(s) := \begin{cases} s^2/2 & \textit{for} \quad |s| \leq 1 \\ \dfrac{|s|^{i+1}}{i+1} + \dfrac{i-1}{2(i+1)} & \textit{for} \quad |s| > 1. \end{cases}$$

Writing e_i for the i-th unit vector in l^2, we obtain

$$f^*(2e_i) = \frac{2^{i+1}}{i+1} + \frac{i-1}{2(i+1)} \underset{i \to \infty}{\longrightarrow} \infty,$$

so f^* is continuous on l^2 and Gâteaux-differentiable.

4.1. **E-Spaces.** Following Holmes [26], we summarize a number of characterizations of a particular class of Banach spaces, the so-called E-spaces, where all convex norm-minimization problems are *strongly solvable.*

Definition 9. *Let (Ω, d) be a metric space and Ω_0 be a subset of Ω. Ω_0 is called approximatively compact, if for each $x \in \Omega$ every minimizing sequence in Ω_0 (i.e. every sequence $\{x_n\} \subset \Omega_0$, for which $d(x, x_n) \to d(x, \Omega_0)$ holds) has a point of accumulation in Ω_0.*

Definition 10. *A Banach space X is an E-space, if X is strictly convex and if every weakly closed subset is approximatively compact.*

Such spaces were originally introduced by Fan and Glicksberg [18].
The next theorems can be found in [26]:

Theorem 22. *A Banach space X is an E-space if and only if X is reflexive, strictly convex and from $x_n, x \in S(X)$ with $x_n \rightharpoonup x$ it follows that $\|x_n - x\| \to 0$.*

The E-space property is closely related to the Kadec-Klee property:

Definition 11. *The Banach space X has the Kadec-Klee property, if from $x_n \rightharpoonup x$ and $\|x_n\| \to \|x\|$ it follows that $\|x_n - x\| \to 0$.*

A strictly convex and reflexive Banach space with Kadec-Klee property is thus an E-space. The E-space property can be characterized by the strong differentiability of the dual space.

Theorem 23 (Anderson)**.** *A Banach space X is an E-space if and only if X^* is Fréchet-differentiable.*

Theorem 24. *A Banach space X is an E-space if and only if for every closed convex set K the problem $\min(\|\cdot\|, K)$ is strongly solvable.*

5. FRECHET–DIFFERENTIABILITY AND LOCAL UNIFORM CONVEXITY IN ORLICZ SPACES

In this section we prove that in a reflexive Orlicz space strong and weak differentiability of Luxemburg and Orlicz norm, as well as strict and local uniform convexity coincide.

5.1. **Fréchet-Differentiability of Modular and Luxemburg norm.** If Φ is differentiable, then f^Φ and $\|\cdot\|_{(\Phi)}$ are Gâteaux-differentiable and for the Gâteaux-derivatives we have (s. Lemma 1):

$$\big\langle (f^\Phi)'(x_0), x\big\rangle = \int_T x\Phi'(x_0)d\mu,$$

and (see (3.4))

$$\Big\langle \|x_0\|'_{(\Phi)}, x\Big\rangle = \frac{\langle f'_\Phi(y_0), x\rangle}{\langle f'_\Phi(y_0), y_0\rangle},$$

where $y_0 := x_0/\|x_0\|_{(\Phi)}$ for $x_0 \neq 0$. If the conjugate function of Φ satisfies the Δ_2-condition we can prove the continuity of the above Gâteaux-derivatives.

The following theorem was proved in a different way by Krasnosielskij [47] for the Lebesgue measure on a compact subset of the $\mathbb{R}^n$.

Theorem 25. *Let (T, Σ, μ) be a σ-finite measure space, Φ a differentiable Young function. If its conjugate function Ψ satisfies the Δ_2-condition, then the Gâteaux-derivatives of f^Φ and $\|\cdot\|_{(\Phi)}$ are continuous mappings from*

$$\mathcal{M}^\Phi(\mu) \quad \textit{resp.} \quad \mathcal{M}^\Phi(\mu)\setminus\{0\} \quad \textit{to} \quad L^\Psi(\mu).$$

Proof: Let $(x_n)_{n=1}^\infty$ be a sequence in $\mathcal{M}^\Phi(\mu)$ with $\lim x_n = x_0$. First we show that in $L^\Psi(\mu)$ the relation $\lim_{n\to\infty} \Phi'(x_n) = \Phi'(x_0)$ holds. As Ψ satisfies the Δ_2-condition, convergence in the norm is equivalent to convergence w.r.t. the modular f^Ψ (s. Theorem 5), i.e. we have to show that

$$\lim_{n\to\infty} f^\Psi\big(\Phi'(x_n) - \Phi'(x_0)\big) = 0.$$

Let now T be represented as a countable union of pairwise disjoint sets T_i, $i = 1, 2, \ldots$, of finite positive measure.

We define

$$S_k := \bigcup_{i=1}^k T_i, \quad S'_k := \big\{t \in T \mid |x_0(t)| \le k\big\}, \quad D_k := S_k \cap S'_k,$$

and finally

$$R_k := T \setminus D_k.$$

We show that $\int_{R_k} x_0\Phi'(x_0)d\mu \to_{k\to\infty} 0$. As $|x_0(t)| > k$ on R_k we obtain:

$$\infty > \int_T x_0\Phi'(x_0)d\mu \ge \int_{R_k} x_0\Phi'(x_0)d\mu \ge \mu(R_k)k\Phi'(k).$$

Since $k\Phi'(k) \ge \Phi(k)\underset{k\to\infty}{\longrightarrow}$, it follows that $\mu(R_k)\underset{k\to\infty}{\longrightarrow} 0$. As $D_k \subset D_{k+1}$, the sequence $\big\{x_0\Phi'(x_0)\chi_{D_k}\big\}$ converges monotonically increasing pointwise almost everywhere to

$x_0\Phi'(x_0)$. With $\int_{D_k} x_0\Phi'(x_0)d\mu \underset{k\to\infty}{\to} \int_T x_0\Phi'(x_0)d\mu$ it follows that $\int_{R_k} x_0\Phi'(x_0)d\mu \underset{k\to\infty}{\to} 0$. For given $\varepsilon > 0$ we now choose k large enough to have

$$\int_{R_k} x_0\Phi'(x_0)d\mu \leq \varepsilon \quad \text{and} \quad \mu(D_k) > 0. \tag{5.1}$$

As Φ' is uniformly continuous on $I := [-k-1, k+1]$, there exists a $\delta \in (0,1)$, such that

$$\left|\Phi'(s) - \Phi'(r)\right| \leq \Psi^{-1}\left(\frac{\varepsilon}{\mu(D_k)}\right)$$

for $|s-r| \leq \delta$ and $s, r \in I$.

According to [47], p. 71, the sequence $\{x_n\}_{n=1}^{\infty}$ converges to x_0 in measure, i.e. there is a sequence of sets $(Q_n)_{n=1}^{\infty}$ with $\lim_{n\to\infty} \mu(Q_n) = 0$, such that

$$\lim_{n\to\infty} \sup_{t\in T\setminus Q_n} |x_0(t) - x_n(t)| = 0.$$

In particular there is a natural number N, such that for $n \geq N$

$$\left|x_n(t) - x_0(t)\right| \leq \delta \quad \text{for} \quad t \in T \setminus Q_n$$

and

$$\mu(Q_n) \leq \frac{\varepsilon}{k\Phi'(k)}. \tag{5.2}$$

Thus we obtain

$$\begin{aligned} \int_{T\setminus(Q_n\cup R_k)} \Psi\big(\Phi'(x_n) - \Phi'(x_0)\big)d\mu &\leq \int_{T\setminus(Q_n\cup R_k)} \Psi\left(\Psi^{-1}\left(\frac{\varepsilon}{\mu(D_k)}\right)\right)d\mu \\ &= \frac{\mu(T\setminus(Q_n\cup R_k))}{\mu(D_k)}\varepsilon \leq \varepsilon. \end{aligned} \tag{5.3}$$

According to [82] the derivative of a continuous convex function is semi-continuous if the sequence $\{w_n\}_{n=1}^{\infty}$ converges to w_0 in $\mathcal{M}^{\Phi}(\mu)$, then $\left\{\Phi'(w_n)\right\}_{n=1}^{\infty}$ converges $*$-weak to $\Phi'(w_0)$. According to the uniform boundedness theorem of Banach (s. e.g. [31]) the sequence $\left\{\Phi'(w_n)\right\}_{n=1}^{\infty}$ is bounded in $L^{\Psi}(\mu)$, and we obtain

$$\begin{aligned} &\int_T w_n\Phi'(w_n)d\mu - \int_T w_0\Phi'(w_0)d\mu \\ &\qquad = \int_T (w_n - w_0)\Phi'(w_n)d\mu + \int_T w_0\big(\Phi'(w_n) - \Phi'(w_0)\big)d\mu. \end{aligned}$$

By Hölder's inequality

$$\begin{aligned} \left|\int_T w_n\Phi'(w_n)d\mu - \int_T w_0\Phi'(w_0)d\mu\right| &\leq \|w_n - w_0\|_{(\Phi)}\|\Phi'(w_n)\|_{\Psi} \\ &\quad + \left|\int_T w_0\big(\Phi'(w_n) - \Phi'(w_0)\big)d\mu\right|, \end{aligned}$$

where the last expression on the right hand side converges to zero because of the weak convergence of $\Phi'(w_n)$ to $\Phi'(w_0)$. In this way we obtain the relation

$$\lim_{n\to\infty}\int_T w_n\Phi'(w_n)d\mu = \int_T w_0\Phi'(w_0)d\mu\ . \tag{5.4}$$

If $w_n := \chi_{Q_n\cup R_k}\cdot x_n$, $w_0 := \chi_{R_k}\cdot x_0$, and $v_n := \chi_{Q_n\cup R_k}\cdot x_0$. Then we obtain:

$$w_n - w_0 = w_n - v_n + v_n - w_0 = (x_n - x_0)\chi_{Q_n\cup R_k} + x_0(\chi_{Q_n\cup R_k} - \chi_{R_k}).$$

We have $|x_n - x_0| \geq |x_n - x_0|\chi_{Q_n\cup R_k}$, so using the monotonicity of the Luxemburg norm, we obtain $|x_n - x_0|\chi_{Q_n\cup R_k} \underset{n\to\infty}{\longrightarrow} 0$. Using (5.2)

$$|v_n - w_0| = \chi_{Q_n\setminus R_k}|x_0| \leq \chi_{Q_n}\cdot k \leq \frac{\varepsilon}{\Phi'(k)} \leq \frac{\varepsilon}{\Phi(k)/k} \underset{n\to\infty}{\longrightarrow} 0,$$

as Ψ is in particular finite. We conclude that $w_n \underset{n\to\infty}{\longrightarrow} w_0$. Taking into account

$$\begin{aligned}&\int_{Q_n\cup R_k} x_n\Phi'(x_n)d\mu - \int_{R_k} x_0\Phi'(x_0)d\mu + \int_{R_k} x_0\Phi'(x_0)d\mu - \int_{Q_n\cup R_k} x_0\Phi'(x_0)d\mu\\ &= \left(\int_T w_n\Phi'(w_n)d\mu - \int_T w_0\Phi'(w_0)d\mu\right) + \left(\int_T w_0\Phi'(w_0)d\mu - \int_T v_n\Phi'(v_n)d\mu\right)\end{aligned}$$

it follows using (5.4)

$$\lim_{n\to\infty}\left(\int_{Q_n\cup R_k} x_n\Phi'(x_n)d\mu - \int_{Q_n\cup R_k} x_0\Phi'(x_0)d\mu\right) = 0.$$

The Δ_2-condition for Ψ together with Young's equality almost everywhere yields

$$\begin{aligned}\Psi\big(\Phi'(x_n) - \Phi'(x_0)\big) &\leq \frac{1}{2}\big(\Psi(2\Phi'(x_n)) + \Psi(2\Phi'(x_0))\big)\\ &\leq \frac{\lambda}{2}\big(\Psi(\Phi'(x_n)) + \Psi(\Phi'(x_0))\big)\\ &\leq \frac{\lambda}{2}\big(x_n\Phi'(x_n) + x_0\Phi'(x_0)\big).\end{aligned}$$

Therefore for n sufficiently large

$$\begin{aligned}&\int_{Q_n\cup R_k} \Psi\big(\Phi'(x_n) - \Phi'(x_0)\big)d\mu\\ &\qquad\leq \frac{\lambda}{2}\left(\int_{Q_n\cup R_k} x_n\Phi'(x_n)d\mu + \int_{Q_n\cup R_k} x_0\Phi'(x_0)d\mu\right)\\ &\qquad\leq \lambda\left(\int_{Q_n\cup R_k} x_0\Phi'(x_0)d\mu + \varepsilon\right).\end{aligned}$$

Using (5.1) and (5.2), we obtain

$$\int_{Q_n\cup R_k} x_0\Phi'(x_0)d\mu = \int_{R_k} x_0\Phi'(x_0)d\mu + \int_{Q_n\setminus R_k} x_0\Phi'(x_0)d\mu \leq 2\varepsilon.$$

Together with (5.3) the first part of the statement follows.

Let now $x_0 \in \mathcal{M}^\Phi(\mu) \setminus \{0\}$ and $(x_n)_{n=1}^\infty$ be a sequence that converges in $\mathcal{M}^\Phi(\mu)$ to x_0.

Writing $y_n := x_n/\|x_n\|_{(\Phi)}$ for $n \in \mathbb{N} \cup \{0\}$, then $(y_n)_{n=1}^\infty$ converges to y_0, i.e. $\lim_{n\to\infty} \Phi'(y_n) = \Phi'(y_0)$ and, because of (5.4) we get

$$\lim_{n\to\infty} \int_T y_n \Phi'(y_n) d\mu = \int_T y_0 \Phi'(y_0) d\mu.$$

Remark 3. *To prove in a similar way the above Theorem 25 for the sequence space l^Φ, only the Δ_2^0-condition for Ψ is required.*

Remark 4. *If T in Theorem 25 has finite measure, only the Δ_2^∞-condition for Ψ is needed.*

Proof: Φ is differentiable, hence Ψ is strict convex, and in particular definite.

Let now $\Psi(2s) \le \lambda \Psi(s)$ for all $s \ge s_0 \ge 0$. For $D_k := S'_k$ and k large enough

$$\mu(R_k) \le \frac{\varepsilon}{2\,(\Psi(2s_0)+1)}$$

furthermore $\int_{R_k} x_0 \Phi'(x_0) d\mu \le \varepsilon$ and $\Psi(2s_0)/k\Phi'(k) \le 1$.

If we set

$$\begin{aligned} P_n &:= \big\{ t \in Q_n \cup R_k \mid |\Phi'(x_n(t))| \le s_0 \big\}, \\ P'_n &:= \big\{ t \in Q_n \cup R_k \mid |\Phi'(x_0(t))| \le s_0 \big\}, \end{aligned}$$

then as in Theorem 25 for n sufficiently large

$$\begin{aligned} &\int_{Q_n \cup R_k} \Psi\big(\Phi'(x_n) - \Phi'(x_0)\big) d\mu \\ &\quad \le \frac{1}{2} \left(\int_{Q_n \cup R_k} \Psi\big(2\Phi'(x_n)\big) d\mu + \int_{Q_n \cup R_k} \Psi\big(2\Phi'(x_0)\big) d\mu \right) \\ &\quad = \frac{1}{2} \Bigg(\int_{P_n} \Psi\big(2\Phi'(x_n)\big) d\mu + \int_{(Q_n \cup R_k)\setminus P_n} \Psi\big(2\Phi'(x_n)\big) d\mu \\ &\qquad\quad + \int_{P'_n} \Psi(2\Phi'(x_0)) d\mu + \int_{(Q_n \cup R_k)\setminus P'_n} \Psi\big(2\Phi'(x_0)\big) d\mu \Bigg) \\ &\quad \le \frac{1}{2} \left(\Psi(2s_0)(\mu(P_n) + \mu(P'_n)\right) \\ &\qquad + \lambda \left(\int_{(Q_n \cup R_k)\setminus P_n} x_n \Phi'(x_n) d\mu + \int_{(Q_n \cup R_k)\setminus P'_n} x_0 \Phi'(x_0) d\mu \right) \\ &\quad \le \frac{1}{2} \big(\mu(R_k) + \mu(Q_n)\big) \Psi(2s_0) + \lambda \left(\int_{Q_n \cup R_k} x_0 \Phi'(x_0) d\mu + \varepsilon \right) \le \varepsilon(1 + 3\lambda). \end{aligned}$$

This completes the proof.

Now our purpose is to demonstrate that for a finite, not purely atomic measure space the following statement holds: if $(L^\Phi(\mu), \|\cdot\|_{(\Phi)})$ is Fréchet-differentiable, then $(L^\Phi(\mu), \|\cdot\|_{(\Phi)})$ is already reflexive.

Lemma 7. *Let (T, Σ, μ) be a σ-finite, not purely atomic measure space. If $\big(L^\Phi(\mu), \|\cdot\|_{(\Phi)}\big)$ is Fréchet-differentiable, then Ψ is finite.*

Proof: Let Ψ not be finite, then Φ' is bounded on $\mathbb{R}$, i.e. there are positive numbers a, b, c such that $a(s-c) \leq \Phi(s) \leq bs$ for $s \geq c$. If T has finite measure, then Theorem 1 implies

(1) $L^\Phi(\mu) = L^1(\mu)$, and the norms $\|\cdot\|_1$ and $\|\cdot\|_{(\Phi)}$ are equivalent;
(2) $L^\Psi(\mu = L^\infty(\mu)$, and the norms $\|\cdot\|_\infty$ and $\|\cdot\|_\Psi$ are equivalent.

Clearly Φ satisfies the Δ_2^∞-condition, hence $\mathcal{M}^\Phi(\mu) = L^\Phi(\mu)$ (see Theorem 4). Therefore

$$\big(L^\Psi(\mu), \|\cdot\|_\Psi\big) = \big(L^\Phi(\mu), \|\cdot\|_{(\Phi)}\big)^*.$$

Let now A be the set of atoms in T and let $\lambda := \mu(T \setminus A)$, then we choose disjoint sets G_k in $T \backslash A$ with $\mu(G_k) = 2^{-k-1}\lambda$ for $k = 1, 2, \ldots$ [85]. We take $s_0 \in \mathbb{R}_+$ large enough

$$\Phi'\left(\frac{s_0}{e^2\Phi^{-1}\left(\left(\frac{\lambda}{2} + \mu(A)\right)^{-1}\right)}\right) > 0,$$

and define the functions x_n on T by

$$x_n(t) := \begin{cases} (1-\frac{1}{k})^n s_0 & \text{for} \quad t \in G_k,\ k = 1, 2, \ldots, \\ 1 & \text{otherwise} \end{cases}$$

for $n = 1, 2, \ldots$ and put

$$x_0(t) := \begin{cases} 0 & \text{for} \quad t \in G_k,\ k = 1, 2, \ldots, \\ 1 & \text{otherwise.} \end{cases}$$

The sequence $\{x_n\}$ converges to x_0 in the L^1-norm. To see this we observe

$$\int_T |x_0 - x_n| d\mu = \sum_{k=1}^\infty \int_{G_k} |x_0 - x_n| d\mu = \lambda s_0 \sum_{k=1}^\infty 2^{-k-1}\left(1 - \frac{1}{k}\right)^n.$$

Given $\varepsilon > 0$, we choose a natural number r such that $\sum_{k=r+1}^\infty 2^{-k-1} < \varepsilon/2$. Then

$$\int_T |x_0 - x_n| d\mu \leq \lambda s_0 \left(\left(1 - \frac{1}{r}\right)^n \sum_{k=1}^r 2^{-k-1} + \frac{\varepsilon}{2}\right) \leq s_0 \lambda \varepsilon$$

for sufficiently large n, hence $x_n \to x_0$ in the L^1-norm.

Thus the number sequence $\{\|x_n\|_{(\Phi)}\}$ converges to $\|x_0\|_{(\Phi)}$. Since x_0 is the characteristic function of $T\backslash\cup_{k=1}^{\infty}$ we obtain, because of $\sum_k \mu(G_k)=\lambda/2$,

$$\|x_0\|_{(\Phi)}=\frac{1}{\Phi^{-1}\left(\left(\frac{\lambda}{2}+\mu(A)\right)^{-1}\right)}.$$

We now set $y_n:=x_n/\|x_n\|_{(\Phi)}$ for $n=0,1,2,\dots$. The Gâteaux-gradient of $\|\cdot\|_{(\Phi)}$ at x_n is

$$\frac{y_n}{\int_T y_n\Phi'(y_n)\,d\mu}.$$

We first consider the sequence $\{\Phi'(y_n)\}_{n=1}^{\infty}$ in $L^\infty(\mu)$. Because of the monotonicity of Φ' and the choice of s_0, for n sufficiently large we obtain

$$\begin{aligned}\big\|\Phi'(y_n)-\Phi'(y_0)\big\|_\infty &\geq \operatorname*{ess\,sup}_{t\in G_n}\big|\Phi'\big(y_n(t)\big)-\Phi'\big(y_0(t)\big)\big|\\ &=\Phi'\left(\frac{(1-1/n)^n s_0}{\|x_n\|_{(\Phi)}}\right)\geq\Phi'\left(\frac{e^{-2}s_0}{\|x_0\|_{(\Phi)}}\right)>0.\end{aligned}$$

But

$$\left\{\int_T y_n\Phi'(y_n)d\mu\right\}_{n=1}^{\infty}\quad\text{converges to}\quad\int_T y_0\Phi'(y_0)d\mu,$$

because apparently $\{\Phi'(y_n)|n\in\mathbb{N}\}$ is uniformly bounded in L^∞ and we have

$$\begin{aligned}&\int_T y_0\big(\Phi'(y_n)-\Phi'(y_0)\big)d\mu=\int_{T\backslash\cup_{k=1}^{\infty}G_k}y_0\big(\Phi'(y_n)-\Phi'(y_0)\big)d\mu\\ &=\frac{1}{\|x_0\|_{(\Phi)}}\left(\Phi'\left(\frac{1}{\|x_n\|_{(\Phi)}}\right)-\Phi'\left(\frac{1}{\|x_0\|_{(\Phi)}}\right)\right)\left(\frac{\lambda}{2}+\mu(A)\right).\end{aligned}$$

Thus the Gâteaux-derivative of $\|\cdot\|_{(\Phi)}$ is not continuous in y_0 and therefore $\|\cdot\|_{(\Phi)}$ according to [66] is not Fréchet-differentiable. If T has infinite measure, then we choose a not purely-atomic subset T_0 of T with finite measure.

If Σ_0 is the subalgebra of Σ, that consists of the elements of Σ contained in T_0, then we denote the restriction of μ to Σ_0 by μ_0. In the same way as above we can construct a sequence (y_n) of elements of the unit sphere of $L^\Phi(\mu_0)$ converging to y_0 in $L^\Phi(\mu_0)$, for which the sequence $\{\Phi'(y_n)\}$ does not converge to $\Phi'(y_0)$ in $L^\Psi(\mu_0)$. In fact $L^\Phi(\mu_0)$ and $L^\Psi(\mu_0)$ are subspaces of $L^\Phi(\mu)$ and $L^\Psi(\mu)$ resp., so the restriction of the Luxemburg norm of $L^\Phi(\mu)$ to $L^\Phi(\mu_0)$ and of the Orlicz norm of $L^\Psi(\mu)$ to $L^\Psi(\mu_0)$ is equal to the Luxemburg norm on $L^\Phi(\mu_0)$ and equal to the Orlicz norm on $L^\Psi(\mu_0)$ resp. Thus for Ψ not finite and $\mu(T)=\infty$ the Luxemburg norm $\|\cdot\|_{(\Phi)}$ is also not Fréchet-differentiable on $L^\Phi(\mu)\setminus\{0\}$, and the proof is complete.

Theorem 26. *Let (T,Σ,μ) be a σ-finite, not purely-atomic measure space. If $(L^\Phi(\mu),\|\cdot\|_{(\Phi)})$ is Fréchet-differentiable, then Φ satisfies Δ_2^∞-condition.*

Proof: Assume that Φ does not satisfy the Δ_2^∞-condition, then L^Φ contains according to Theorem 3 a subspace, which is isometrically isomorphic to ℓ^∞. But ℓ^∞ is not Fréchet-differentiable, because otherwise ℓ^1 (according Anderson's Theorem 23) is an E-space, hence in particular is reflexive. We came to a contradiction.

Theorem 27. *Let (T, Σ, μ) be a finite, not purely-atomic measure space. If $(L^\Phi(\mu), \|\cdot\|_{(\Phi)})$ is Fréchet-differentiable, then $(L^\Phi(\mu), \|\cdot\|_{(\Phi)})$ is reflexive.*

Proof: According to Theorem 26 Φ satisfies the Δ_2^∞-condition. Then, due to Theorem 4 we have $L^\Phi(\mu) = \mathcal{M}^\Phi(\mu)$. Therefore, if $\|\cdot\|_{(\Phi)}$ is Fréchet-differentiable on $L^\Phi(\mu) \setminus \{0\}$, then Ψ is finite, according to Lemma 7, hence by Theorem 8 $(\mathcal{M}^\Psi(\mu), \|\cdot\|_\Psi)^* = (L^\Phi(\mu), \|\cdot\|_{(\Phi)})$ and thus, according to the theorem of Anderson 23, $(\mathcal{M}^\Psi(\mu), \|\cdot\|_\Psi)$ is an E-space, hence in particular reflexive. Using Hölder's inequality we obtain

$$\big(L^\Psi(\mu), \|\cdot\|_\Psi\big) \subset \big(L^\Phi(\mu), \|\cdot\|_{(\Phi)}\big)^* = \big(\mathcal{M}^\Psi(\mu), \|\cdot\|_\Psi\big)$$

and so we conclude $L^\Psi(\mu) = \mathcal{M}^\Psi(\mu)$.

We are now in the position to characterize the Fréchet-differentiable Luxemburg norms.

Definition 12. *Let A be the set of atoms. We call the measure space (T, Σ, μ) essentially not purely-atomic, if $\mu(T \backslash A) > 0$, and for $\mu(T) = \infty$ holds $\mu(T \backslash A) = \infty$.*

Theorem 28. *Let (T, Σ, μ) be a σ-finite, essentially not purely-atomic measure space, and $(L^\Phi(\mu), \|\cdot\|_{(\Phi)})$ be reflexive. Then the following statement are equivalent*

(a) *$\big(L^\Phi(\mu), \|\cdot\|_{(\Phi)}\big)$ is flat convex;*
(b) *f_Φ is continuously Fréchet-differentiable on $L^\Phi(\mu)$;*
(c) *$\|\cdot\|_{(\Phi)}$ is continuously Fréchet-differentiable on $L^\Phi(\mu) \setminus \{0\}$;*
(d) *Φ is differentiable.*

Proof: a)$\Rightarrow$ b) and c). According to Theorem 17 flat convexity for a not purely-atomic measure space implies differentiability of Φ and by Mazur's Theorem 16 the Gâteaux-differentiability of $\|\cdot\|_{(\Phi)}$.

By Theorem 9, reflexivity of $\big(L^\Phi(\mu), \|\cdot\|_{(\Phi)}\big)$ implies Δ_2 or Δ_2^∞ condition for Φ and Ψ, if T has infinite or finite measure respectively, because otherwise L^Φ has a (closed) subspace, isomorphic to ℓ^∞, which according to [25] is also reflexive, meaning a contradiction. We conclude, using Theorem 4 $\mathcal{M}^\Phi(\mu) = L^\Phi(\mu)$.

By Δ_2 or Δ_2^∞ for Ψ and Theorem 25 and Remark 4 it follows that the Gâteaux-derivative of f^Φ is continuous on $\mathcal{M}^\Phi(\mu)$, and as in Theorem 25, the corresponding property for $\|\cdot\|_{(\Phi)}$ holds. By [66] continuous Gâteaux-differentiability and continuous Fréchet-differentiability are equivalent.

(a) $\iff$ (d). According to Theorem 17 flat convexity for a not purely atomic measure is equivalent to differentiability of Φ.

(b) $\implies$ (a). Due to Theorem 14 the level set $S_{f^\Phi}(1)$ is flat convex. On the other hand $S_{f^\Phi}(1)$ is identical to the unit sphere of $\mathcal{M}^\Phi = L^\Phi$. Hence $(L^\Phi, \|\cdot\|_{(\Phi)})$ is flat convex.

(c) $\implies$ (a). If $\|\cdot\|_{(\Phi)}$ is Fréchet-differentiable on $L^\Phi(\mu)\setminus\{0\}$, then $(L^\Phi(\mu),\ \|\cdot\|_{(\Phi)})$ is flat convex due to Mazur's Theorem 16.

For the sequence space l^Φ the above theorem can be proved in somewhat weaker form.

Theorem 29. *Let the Young function Φ be finite and l^Φ be reflexive. Then the following statements are equivalent*

(a) $(l^\Phi,\ \|\cdot\|_{(\Phi)})$ *is flat convex;*

(b) $\|\cdot\|_{(\Phi)}$ *is Fréchet-differentiable on* $l^\Phi\setminus\{0\}$*;*

(c) Φ *is differentiable for all s with* $|s|<\Phi^{-1}(1)$.

Proof. The equivalence of (a) and (c) follows from Theorems 18 and 10. Let now $(l^\Phi, \|\cdot\|_{(\Phi)})$ be flat convex. The left-sided derivative Φ'_- of Φ in $\Phi^{-1}(1)$ is finite. If we continue Φ'_- in a continuous, linear, and strictly increasing way beyond $\Phi^{-1}(1)$, and denote that primitive function which is zero at the origin by $\tilde{\Phi}$, then $\tilde{\Phi}$ is a differentiable Young function. Apparently $l^\Phi = l^{\tilde{\Phi}}$ and $\|\cdot\|_{(\tilde{\Phi})} = \|\cdot\|_{(\Phi)}$. If $\tilde{\Psi}$ is the convex conjugate of $\tilde{\Phi}$ then by construction $\tilde{\Psi}$ is finite. Φ and hence $\tilde{\Phi}$ satisfies the Δ_2^0-condition, because otherwise $(l^\Phi, \|\cdot\|_{(\Phi)})$ contains, according to the theorem of Lindenstrauss-Tsafriri a subspace isomorphic to ℓ^∞, in contradiction to the reflexivity of $(l^\Phi, \|\cdot\|_{(\Phi)})$. It follows that $m^{\tilde{\Phi}} = \ell^{\tilde{\Phi}}$. Due to Theorem 8 we have $(l^{\tilde{\Phi}}, \|\cdot\|_{\tilde{\Phi}})^* = (l^{\tilde{\Psi}}, \|\cdot\|_{(\tilde{\Psi})})$. The equivalence of the norms implies that $(l^{\tilde{\Phi}}, \|\cdot\|_{\tilde{\Phi}})$ is also reflexive. Hence $(l^{\tilde{\Psi}}, \|\cdot\|_{(\tilde{\Psi})})$, the dual space of a reflexive space, is also reflexive. As demonstrated above the Δ_2^0-condition for $\tilde{\Psi}$ follows. Using Remark 3 we obtain (b). By Theorem 15 a) follows from (b).

5.2. **Frechet-differentiability and local uniform convexity of Orlicz spaces.** We need a characterization of strict convexity of the Luxemburg norm on $\mathcal{M}^\Phi(\mu)$.

Theorem 30. *Let (T, Σ, μ) be a not purely atomic measure space and let Φ a finite Young function. Then Φ is strictly convex, if and only if $\left(\mathcal{M}^\Phi(\mu), \|\cdot\|_{(\Phi)}\right)$ is strictly convex.*

Proof: Let A be the set of atoms of T. If Φ is not strictly convex, then there are different positive numbers s and t, such that

$$\Phi\left(\frac{s+t}{2}\right) = \frac{1}{2}\big(\Phi(s) + \Phi(t)\big).$$

We choose pairwise disjoint sets $T_1, T_2, T_3 \in \Sigma$ of finite measure with

$$0 < \mu(T_1) = \mu(T_2) \le \min\left(\frac{\mu(T \setminus A)}{3}, \frac{1}{2(\Phi(s) + \Phi(t))}\right)$$

and

$$0 < \mu(T_3) \le \mu(T \setminus A) - 2\mu(T_1).$$

We put

$$u := \Phi^{-1}\left(\frac{1 - \mu(T_1)(\Phi(s) + \Phi(t))}{\mu(T_3)}\right).$$

The functions $x := s\chi_{T_1} + t\chi_{T_2} + u\chi_{T_3}$ and $y := t\chi_{T_1} + s\chi_{T_2} + u\chi_{T_3}$ are then elements of the unit sphere, because we have

$$\int_T \Phi(x) d\mu = \mu(T_1)\Phi(s) + \mu(T_2)\Phi(t) + \mu(T_3)\Phi(u)$$
$$= \mu(T_2)\Phi(s) + \mu(T_1)\Phi(t) + \mu(T_3)\Phi(u) = \int_T \Phi(y) d\mu = 1.$$

Now, the properties of s, t, T_1 and T_2 imply

$$\int_T \Phi\left(\frac{x+y}{2}\right) d\mu = \mu(T_1)\Phi\left(\frac{s+t}{2}\right) + \mu(T_2)\Phi\left(\frac{s+t}{2}\right) + \mu(T_3)\Phi(u) = 1.$$

Conversely, let Φ be strictly convex, and x_1, x_2 be elements of the unit sphere $\|(x_1 + x_2)/2\|_{(\Phi)} = 1$. Then

$$1 = \int_T \Phi(x_1) d\mu = \int_T \Phi(x_2) d\mu = \int_T \Phi\left(\frac{x_1 + x_2}{2}\right) d\mu.$$

Because of convexity of Φ, we obtain

$$\Phi((x_1 + x_2)/2) - \Phi(x_1) - \Phi(x_2) = 0$$

almost everywhere. Due to the strict convexity $x_1 = x_2$ almost everywhere, i.e. the Luxemburg norm is strictly convex on $\mathcal{M}^\Phi(\mu)$.

Corollary 2. Let (T, Σ, μ) a not purely atomic measure space and Φ be a finite Young function. If f^Φ strictly convex, then Φ is strictly convex (because then also $\mathcal{M}^\Phi$ is strictly convex).

The next theorem (compare with [30] p. 350) describes the duality between strict and flat convexity:

Theorem 31. *If the dual space X^* of a normed space X is strict or flat convex, then X itself is flat or strict convex respectively.*

The norm of the dual space (or more precisely its square) of a normed space X can be interpreted as the convex conjugate of the square of the norm in X:

Remark 5. *([31]) Let $\Phi : \mathbb{R} \to \mathbb{R}$ be convex with $\Phi(0) = 0$ and $(X, \|\cdot\|)$ be a normed space. We consider the convex function $f : X \to \mathbb{R}$, given by $f(x) := \Phi(\|x\|)$. Its conjugate $f^* : (X^*, \|\cdot\|_d) \to \overline{\mathbb{R}}$ is given by*

$$f^*(y) = \Phi^*(\|y\|_d),$$

in particular

$$\left(\frac{\|\cdot\|^2}{2}\right)^* = \frac{\|\cdot\|_d^2}{2}.$$

Conjugate Young functions lead to conjugate modulars as is stated in the next theorem (see [39]):

Theorem 32. *Let Φ be a Young function and Ψ its conjugate. Then for arbitrary $x \in L^\Phi(\mu)$*

$$f^\Phi(x) = \sup_{y \in L^\Psi(\mu)} \left\{ \int_T xy d\mu - f^\Psi(y) \right\} = (f^\Psi)^*(x).$$

The modulars f^Φ is bounded if Φ satisfies a Δ_2-conditions (see [39]):

Theorem 33. *If Φ satisfies Δ_2 condition (or for $\mu(T) < \infty$ the Δ_2^∞ condition), then $f^\Phi : L^\Phi(\mu) \to \mathbb{R}$ is a bounded function.*

Proof: Let Φ satisfy the Δ_2 condition, $M > 0$ and $x \in K_{(\Phi)}(0, M)$. If $2^n \geq M$, then

$$f^\Phi(x) \leq f^\Phi\left(\frac{2^n}{M}x\right) \leq \lambda^n f^\Phi\left(\frac{x}{M}\right) \leq \lambda^n.$$

Let now T be of finite measure and Φ satisfy merely the Δ_2^∞ condition. We set

$$T_0 := \left\{ t \in T | \frac{|x(t)|}{M} \geq s_0 \right\}$$

to obtain

$$f^\Phi(x) = \int_{T_0} \Phi(x) d\mu + \int_{T \setminus T_0} \Phi(x) d\mu \leq \int_{T_0} \Phi\left(\frac{2^n}{M}x\right) d\mu + \mu(T)\Phi(M s_0)$$
$$\leq \lambda^n \int_{T_0} \Phi\left(\frac{x}{M}\right) d\mu + \mu(T)\Phi(M s_0) \leq \lambda^n + \mu(T)\Phi(M s_0).$$

A similar statement holds for sequence spaces:

Theorem 34. *If Φ is finite and Φ satisfies the Δ_2^0 condition, then $f^\Phi : \ell^\Phi \to \mathbb{R}$ is a bounded function.*

Proof: Let Φ satisfy the Δ_2^0 condition on every interval $[0, a]$ let $M > 0$ and $x \in K_{(\Phi)}(0, M)$. Then $\sum_{i=1}^\infty \Phi(x_i/M) \leq 1$ and hence $|x_i|/M \leq \Phi^{-1}(1)$. Let further $2^n \geq M$, then for $a := 2^n\Phi^{-1}(1)$

$$f^\Phi(x) \leq f^\Phi\left(\frac{2^n}{M}x\right) \leq \lambda^n f^\Phi\left(\frac{x}{M}\right) \leq \lambda^n,$$

and the proof is complete.

We are going to characterize reflexive and locally uniformly convex Orlicz spaces w.r.t. the Orlicz norm.

Theorem 35. *Let (T, Σ, μ) a σ-finite, essentially not purely atomic measure space and $L^\Phi(\mu)$ be reflexive. Then the following statements are equivalent:*

(a) *Φ is strictly convex;*
(b) *$(L^\Phi(\mu), \|\cdot\|_\Phi)$ is strictly convex;*
(c) *$\|\cdot\|_\Phi^2$ is locally uniformly convex;*
(d) *f^Φ is locally uniformly convex.*

Proof: The reflexivity implies in particular that Φ and Ψ are finite.

(a) $\Longrightarrow$ b): If Φ is strictly convex, then its conjugate Ψ is differentiable. Then due to Theorem 17 $(L^\Psi(\mu), \|\cdot\|_{(\Psi)})$ is flat convex, hence $(L^\Phi(\mu), \|\cdot\|_\Phi)$ is strictly convex (see Theorem 31).

(b) $\Longleftrightarrow$ (c): If $(L^\Phi(\mu), \|\cdot\|_\Phi)$ is strictly convex, then $(L^\Psi(\mu), \|\cdot\|_{(\Psi)})$ is flat convex (see Theorem 31). Hence due to Theorem 28 $\|\cdot\|_{(\Psi)}^2/2$ is Fréchet-differentiable. As $\|\cdot\|_\Phi^2/2$ is the conjugate function of $\|\cdot\|_{(\Psi^2)}^2/2$ (see Remark 5), then, according to Theorems 19 and 21 $\|\cdot\|_\Phi^2$ is locally uniformly convex. Apparently (c) follows immediately from (b).

(b) $\Longrightarrow$ (d): From flat convexity of $(L^\Psi(\mu), \|\cdot\|_{(\Psi)})$ and reflexivity it follows by Theorem 28 that f^Ψ is Fréchet-differentiable. Φ and Ψ satisfy the Δ_2 or the Δ_2^∞ condition respectively, depending on $\mu(T)$ being infinite or finite (see Theorem 9). Thus f^Ψ and its conjugate f^Φ are bounded (see Theorems 32 and 33). Due to Theorems 19 and 21 this implies the local uniform convexity of f^Φ.

(d) $\Longrightarrow$ (a): This follows immediately from Corollary 2.

Remark 6. *Milnes in* [60]*, page 1482) gives an example of a reflexive and strictly convex Orlicz space w.r.t. the Orlicz norm, which is not uniformly convex (compare.*

For the sequence space l^Φ the above theorem can be proved in a somewhat weaker version.

Theorem 36. *Let Φ and Ψ be finite and let l^Φ be reflexive. Then the following statements are equivalent:*

(a) $(l^\Phi, \|\cdot\|_\Phi)$ *is strictly convex;*

(b) $\|\cdot\|_\Phi^2$ *is locally uniformly convex.*

Proof: As in Theorem 35 by use of Theorem 29.

5.3. **Fréchet-Differentiability of the Orlicz norm and local uniform convexity of Luxemburg norm.** Using the relationships between Fréchet-differentiability, local uniform convexity and strong solvability presented in the previous section, we describe the Fréchet-differentiability of the Orlicz norm.

Theorem 37. *Let (T, Σ, μ) be an essentially not purely atomic, σ-finite measure space, and let $L^\Phi(\mu)$ be reflexive. Then the following statements are equivalent:*

(a) $(L^\Phi(\mu), \|\cdot\|_\Phi)$ *is flat convex;*

(b) Φ *is differentiable,*

c) $\|\cdot\|_\Phi$ *is Fréchet-differentiable on* $L^\Phi(\mu) \setminus \{0\}$*;*

(c) $\|\cdot\|_\Phi$ *is Fréchet-differentiable on* $L^\Phi(\mu) \setminus \{0\}$.

Proof: (a) $\Longrightarrow$ (b): Let Ψ be the conjugate of Φ. If $\big(L^\Phi(\mu), \|\cdot\|_\Phi\big)$ is flat convex, then $(L^\Psi(\mu), \|\cdot\|_{(\Psi)})$ is strictly convex. Due to Theorem 30 Ψ is strictly convex and hence Φ is differentiable.

(b) $\Longrightarrow$ (c): From the differentiability of Φ it follows by Theorem 28 that f^Φ is Fréchet-differentiable. As in the proof of Theorem 35, the local uniform convexity of f^Ψ follows. Strong and weak sequential convergence agree on the set $S := \{x \mid f^\Psi(x) = 1\}$ because from $x_n \rightharpoonup x$ for $x_n,\, x \in S$ it follows for $x^* \in \partial f^\Psi(x)$:

$$0 = f^\Psi(x_n) - f^\Psi(x) \geq \langle x_n - x, x^*\rangle + \tau\big(\|x_n - x\|_{(\Psi)}\big),$$

where τ is the convexity module of f^Ψ belonging to x and x^*, and thus $x_n \to x$. As S is the unit sphere of $L^\Psi(\mu)$ w.r.t. the Luxemburg norm, $(L^\Psi(\mu), \|\cdot\|_{(\Psi)})$ is an E-space according to Theorem 22, hence $\|\cdot\|_{(\Psi)}$ has a strong minimum on every closed convex set due to Theorem 24. Apparently, this also holds for $\|\cdot\|_{(\Psi)}^2/2$. Theorems 19 and 21 then imply the Fréchet-differentiability of $\|\cdot\|_\Phi^2/2$ and hence of $\|\cdot\|_\Phi$ in $L^\Phi(\mu)\setminus\{0\}$.

(c) $\Longrightarrow$ (a): Follows from the theorem of Mazur, and the proof is complete.

It is now a simple task to characterize the locally uniformly convex, reflexive Orlicz spaces w.r.t. the Luxemburg norm.

Theorem 38. *Let (T,Σ,μ) be σ-finite, essentially not purely atomic measure space and $L^\Phi(\mu)$ be reflexive. Then the following statements are equivalent:*

(a) *Φ is strict convex;*
(b) *$\big(L^\Phi(\mu),\|\cdot\|_{(\Phi)}\big)$ is strictly convex;*
(c) *$\|\cdot\|_{(\Phi)}^2$ is locally uniformly convex.*

Proof: Because of $\mathcal{M}^\Phi(\mu) = L^\Phi(\mu)$ the equivalence of a) and b) follows from Theorem 30. If $(L^\Phi(\mu),\|\cdot\|_{(\Phi)})$ is strictly convex, then $(L^\Psi(\mu),\|\cdot\|_\Psi)$ is flat convex and therefore due to Theorem 37 $\|\cdot\|_\Psi$ is Fréchet-differentiable. Theorems 19 and 21 now imply (c).

(c) $\Longrightarrow$ (b) is obvious. This completes the proof.

The theorems corresponding to Theorems 37 and 38 for l^Φ can be stated in the weaker form.

Theorem 39. *Let l^Φ be reflexive, Φ differentiable and Ψ finite, then $\|\cdot\|_\Phi$ is Fréchet-differentiable on $l^\Phi\setminus\{0\}$.*

Proof: Because of Theorem 10 Ψ satisfies the Δ_2^0 condition. Hence f^Φ is, due to Remark 3, Fréchet-differentiable. The remaining reasoning follows the lines of Theorem 37, (b) $\Longrightarrow$ (c).

Remark 7. *If the conditions of Theorem 39 are satisfied, then strong and weak differentiability of the Orlicz norm on l^Φ agree.*

Theorem 40. *Let l^Φ be reflexive, Φ be strictly convex and Ψ be finite. Then $\|\cdot\|_{(\Phi)}^2$ is locally uniformly convex.*

Proof: Ψ is differentiable and hence $\|\cdot\|_\Psi$ according to Theorem 39 Fréchet-differentiable. By Theorems 19 and 21 the statement follows.

5.4. **Summary.** We now describe Fréchet-differentiability and local uniform convexity by a list of equivalent statements.

Theorem 41. *Let (T, Σ, μ) be a σ-finite essentially not purely atomic measure space, Φ be a Young function and Ψ its conjugate, and $L^\Phi(\mu)$ be reflexive. If K is a closed convex subset of $L^\Psi(\mu)$, then the following statements are equivalent:*

(a) *Φ is differentiable;*
(b) *$(L^\Phi(\mu), \|\cdot\|_\Phi)$ is flat convex;*
(c) *$(L^\Phi(\mu), \|\cdot\|_{(\Phi)})$ is flat convex;*
(d) *$\|\cdot\|_\Phi$ is continuously Fréchet-differentiable on $L^\Phi(\mu) \setminus \{0\}$;*
(e) *$\|\cdot\|_{(\Phi)}$ is continuously Fréchet-differentiable on $L^\Phi(\mu) \setminus \{0\}$;*
(f) *f_Φ is continuously Fréchet-differentiable on $L^\Phi(\mu)$;*
(g) *Ψ is strictly convex;*
(h) *$L^\Psi(\mu), \|\cdot\|_\Psi)$ is strictly convex;*
(i) *$L^\Psi(\mu), \|\cdot\|_{(\Psi)})$ is strictly convex;*
(j) *$\|\cdot\|_\Psi^2$ is locally uniformly convex;*
(k) *$\|\cdot\|_{(\Psi)}^2$ is locally uniformly convex;*
(l) *f^Ψ is locally uniformly convex;*
(m) *$\|\cdot\|_\Psi$ has a strong minimum on K;*
(n) *$\|\cdot\|_{(\Psi)}$ has a strong minimum on K;*
(o) *f^Ψ has a strong minimum on K;*
(p) *$L^\Psi(\mu), \|\cdot\|_{(\Psi)})$ is an E-space;*
(q) *$L^\Psi(\mu), \|\cdot\|_\Psi)$ is an E-space.*

Proof: (a), (c), (e) (f) are equivalent according to Theorem 28; (a), (b), (d) according to Theorem 37; (g), (h), (j), (l) according to Theorem 35; (g), (i), (k) according to Theorem 38. The equivalence of (a) and (g) is well known, the equivalence of (j) and (m), (k) and (n) as well as of (l) and (o) follow from Theorem 21. Equivalence of (m) and (q) and of (n) and (p) follows from Theorem 24, and the proof is complete.

Theorem 42. *Let Φ be differentiable, Ψ finite and let l^Φ be reflexive. Then the following statements hold:*

(a) *$\|\cdot\|_\Phi$ is continuously Fréchet-differentiable on $l^\Phi \setminus \{0\}$;*
(b) *$\|\cdot\|_{(\Phi)}$ is continuously Fréchet-differentiable on $l^\Phi \setminus \{0\}$;*
(c) *f^Φ is continuously Fréchet-differentiable;*
(d) *$\|\cdot\|_\Psi^2$ is locally uniformly convex;*

(e) $\|\cdot\|^2_{(\Psi)}$ *is locally uniformly convex;*

(f) f^{Ψ} *is locally uniformly convex;*

(g) $\|\cdot\|_{\Psi}$ *has a strong minimum on* K*;*

(h) $\|\cdot\|_{(\Psi)}$ *has a strong minimum on* K*;*

(i) f^{Ψ} *has a strong minimum on* K

for every closed convex subset K *of* l^{Ψ}.

Proof: (b) follows from Theorem 29, (d) by Theorem 40, (a) by Theorem 39. From reflexivity we obtain using Remark 3 (namely, its statement (c)) and using Theorems 19 and 21 thereby (f). Finally (e) follows from (a). The statements (g), (h) and (i) follow from Theorems 19 and 21.

6. APPLICATIONS

We discuss

- Tychonov-regularization: this method was introduced for the treatment of ill-posed problems. The convergence of the method was proved by Levitin and Polyak for uniformly convex regularizing functionals. We show that locally uniformly convex regularizations are sufficient for that purpose. As we have given a complete description of local uniform convexity in Orlicz spaces we propose such regularizing functionals explicitly.
- Ritz method: the Ritz method plays an important role in many applications (e.g. in FEM-methods). It is well known that the Ritz procedure generates a minimizing sequence. Actual convergence of the minimal solutions on each subspace is only achieved if the original problem is strongly solvable.
- Greedy algorithms have drawn a growing attention and experienced a rapid development in recent years (see e.g. Temlyakov). The aim is to arrive at a 'compressed' representation of a function in terms of its dominating "frequencies".

In the convergence proof of the Tychonov regularization method we make explicit use of local uniform convexity. The convergence of the Ritz method follows from strong solvability, whereas the convergence proof of the greedy algorithm follows from the Kadec-Klee property. So three different aspects of E-spaces come into play.

6.1. **Regularisation of Tychonov Type.**

Lemma 8. *Let X be a Banach space, $f : X \to \mathbb{R}$ a continuous locally uniformly convex function, then for all $x, y \in X$, all $x^* \in \partial f(x)$ and all $y^* \in \partial f(y)$:*

$$\tau_{x,x^*}(\|x-y\|) \leq \langle x-y, x^*-y^* \rangle$$

where τ_{x,x^*} denotes the convexity module belonging to f at x, x^*.

Proof: As f is locally uniformly convex we have:

$$\tau_{x,x^*}(\|x-y\|) + \langle y-x, x^* \rangle \leq f(y) - f(x).$$

On the other hand the subgradient inequality yields: $\langle x-y, y^* \rangle \leq f(x) - f(y)$, i.e.

$$\tau_{x,x^*}(\|x-y\|) + \langle y-x, x^* \rangle \leq \langle y-x, y^* \rangle$$

as claimed.

Theorem 43. *Let X be a reflexive Banach space and f and g be continuous, Gateaux-differentiable convex functions on X. Let further f be locally uniformly convex with bounded conjugate f^* and $S := M(g, X) \neq \emptyset$.*

Let now α_n be a positive sequence tending to zero and $f_n := \alpha_n f + g$. Let finally x_n be the (uniquely determined) minimal solution of f_n on X. Then the sequence $\{x_n\}$ converges to the (uniquely determined) minimal solutions of f on S.

Proof: By Theorem 1 f_n is locally uniformly convex and f_n^* is bounded, hence $M(f_n, X)$ consists of the unique element x_n.

For $x \in S$, because of monotonicity of the derivative of g

$$\begin{aligned} &\underbrace{\langle x_n - x, g'(x_n) - g'(x) \rangle}_{\geq 0} + \alpha_n \langle x_n - x, f'(x_n) \rangle \\ &\qquad\qquad = \langle x_n - x, \underbrace{\alpha_n f'(x_n) + g'(x_n)}_{=0} - \underbrace{g'(x)}_{=0} \rangle = 0. \end{aligned}$$

It follows

$$\langle x_n - x, f'(x_n) \rangle \leq 0. \tag{6.1}$$

For $\bar{x} \in S$ arbitrary

$$0 \geq f_n(x_n) - f_n(\bar{x}) = g(x_n) - g(\bar{x}) + \alpha_n(f(x_n) - f(\bar{x})) \geq \alpha_n(f(x_n) - f(\bar{x})).$$

This implies $f(x_n) \leq f(\bar{x})$, hence $x_n \in S_f(f(\bar{x}))$ for all $n \in \mathbb{N}$. As f^* is bounded, it follows according to Theorem 4 that the sequence $\{x_n\}$ is bounded. Let now (x_k) be a subsequence converging weakly to x_0.

First we show that $g'(x_0) = 0$, i.e. $x_0 \in S$. For $y \in X$

$$\underbrace{\langle y - x_k, g'(y) - g'(x_k)\rangle}_{\geq 0} + \alpha_k \langle x_k - y, f'(x_k)\rangle$$
$$= \langle y - x_k, g'(y) - \underbrace{f'_k(x_k)}_{=0}\rangle = \langle y - x_k, g'(y)\rangle.$$

For fixed y the expression $\langle x_k - y, f'(x_k)\rangle$ is bounded from below

$$\langle x_k - y, f'(x_k)\rangle = \underbrace{\langle x_k - y, f'(x_k) - f'(y)\rangle}_{\geq 0} + \langle x_k - y, f'(y)\rangle$$
$$\geq -\|f'(y)\|\|x_k - y\| \geq C.$$

Hence we obtain

$$-C\alpha_k \leq \alpha_k \big\langle x_k - y, f'(x_k)\big\rangle \leq \big\langle y - x_k, g'(y)\big\rangle.$$

On the other hand the weak convergence of $x_k \rightharpoonup x_0$ implies:

$$\big\langle y - x_k, g'(y)\big\rangle \underset{k\to\infty}{\longrightarrow} \big\langle y - x_0, g'(y)\big\rangle$$

and thus for all $y \in X$

$$\big\langle y - x_0, g'(y)\big\rangle \geq 0.$$

Let now $t > 0$,$z \in X$ be arbitrary and $y = x_0 + tz$. Then the continuity of $t \longmapsto \langle z, g'(x_0{+}tz)\rangle = \frac{d}{dt} g(x_0{+}tz)$ implies $0 \leq \langle z, g'(x_0{+}tz)\rangle \underset{t\to 0}{\longrightarrow} \langle z, g'(x_0)\rangle$, hence $g'(x_0) = 0$, i.e. $x_0 \in S$.

Now we show the strong convergence of (x_k) to x_0. Due to Lemma 8 the weak convergence and inequality (6.1) yield

$$\begin{aligned}
&\tau_{x_0, f'(x_0)}(\|x_0 - x_k\|) \leq \big\langle x_0 - x_k, f'(x_0) - f'(x_k)\big\rangle \\
&= \langle x_0 - x_k, f'(x_0)\rangle + \underbrace{\langle x_k - x_0, f'(x_k)\rangle}_{\leq 0} \\
&\leq \langle x_0 - x_k, f'(x_0)\rangle \to 0.
\end{aligned}$$

Hence $x_k \to x_0$ and thus also $x_n \to x_0$.

It remains to be shown that x_0 is the minimal solution of f on S. Because of the semi-continuity [82] of f' it follows with inequality (6.1) for $x \in S$

$$0 \leq \big\langle x - x_n, f'(x_n)\big\rangle \to \big\langle x - x_0, f'(x_0)\big\rangle.$$

By the characterization theorem of convex optimization [31] the assertion of the theorem follows.

Remark 8. *the proof of above theorem easily carries over to semi-continuous monotone operators (compare* [43]*).*

Theorem 44. *Let (T, Σ, μ) be a σ-finite, essentially not purely-atomic measure space, $L^\Phi(\mu)$ be reflexive, and Φ be strictly convex and differentiable. Let either $f := f^\Phi$ or $f := \frac{1}{2}\| \cdot \|^2_{(\Phi)}$ or $f := \frac{1}{2}\| \cdot \|^2_\Phi$ and $g : L^\Phi(\mu) \to \mathbb{R}$ be convex, continuous, and Gâteaux-differentiable and $S := M(g, L^\Phi(\mu)) \neq \emptyset$.*

Let now α_n be a positive sequence tending to zero and $f_n := \alpha_n f + g$. Let finally x_n be the (uniquely determined) minimal solution of f_n on $L^\Phi(\mu)$. Then the sequence $\{x_n\}$ converges to the(uniquely determined) minimal solution of f on S.

Proof: By Theorem 43 in conjunction with Theorem 41.

6.2. **Ritz's Method.** The following method of minimizing a functional on an increasing sequence of subspaces of a separable space is due to Ritz (compare e.g. [82]).

Theorem 45. *Let X be a separable normed space, $X = \overline{\text{span } \{\varphi_i, i \in \mathbb{N}\}}$, and $X_n :=$ span $\{\varphi_1, \ldots, \varphi_n\}$. Let further $f : X \to \mathbb{R}$ be upper semi-continuous and bounded from below. If $d := \inf f(X)$ and $d_n := \inf f(X_n)$ for $n \in \mathbb{N}$, then $\lim_{n\to\infty} d_n = d$.*

Proof: $d_n \geq d_{n+1}$ for all $n \in \mathbb{N}$, hence $d_n \to a \in \mathbb{R}$. Suppose $a > d$. Let $(d-a)/2 > \epsilon > 0$ and $x \in X$ with $f(x) \leq d + \varepsilon$. As f is upper semi-continuous, there is a neighbourhood $U(x)$ with $f(y) \leq f(x) + \varepsilon$ for all $y \in U(x)$, in particular there exists $y_m \in U(x)$ with $y_m \in X_m$. It follows that

$$a \leq d_m \leq f(y_m) \leq f(x) + \epsilon \leq d + 2\epsilon.$$

We came to a contradiction.

Corollary 3. Let $d_n := \inf f(X_n)$, and (δ_n) be a sequence of positive numbers tending to zero. If $x_n \in X_n$ is chosen in such a way that $f(x_n) \leq d_n + \delta_n$, then (x_n) is a minimizing sequence for the minimization of f on X.

Theorem 46. *Let X be a separable reflexive Banach space with*

$$X = \overline{\text{span } \{\varphi_i, i \in \mathbb{N}\}}, \quad X_n := \text{span } \{\varphi_1, \ldots, \varphi_n\}.$$

Let f and g be continuous convex function on X. Let further f be locally uniformly convex and f^* be bounded. Let $d := \inf(f + g)(X)$ and $d_n := \inf(f + g)(X_n)$ for $n \in \mathbb{N}$, (δ_n) be a sequence of positive numbers and $x_n \in X_n$ be chosen to have $f(x_n)+g(x_n) \leq d_n+\delta_n$. Then the minimizing sequence (x_n) converges to the (uniquely determined) minimal solution of $f + g$ on X.

Proof: Corollary of 3, with Theorems 1 and 20.

Remark 9. *The above theorem enables to regularize the minimization problem* min (g, X) *by adding a positive multiple* αf *of a locally uniformly convex function, i.e. one replaces the above problem by* $\min(\alpha f + g, X)$ *(compare with Theorem 43).*

Theorem 47. *Let* (T, Σ, μ) *be a* σ*-finite, essentially not purely-atomic measure,* $L^\Phi(\mu)$ *be separable and reflexive,* Φ *be strictly convex. Let* $L^\Phi(\mu) = \overline{\text{span }\{\varphi_i, i \in \mathbb{N}\}}$ *and* $X_n := \text{span }\{\varphi_1, \ldots, \varphi_n\}$. *Let either* $f := f^\Phi$ *or* $f := \|\cdot\|^2_{(\Phi)}$ *or* $f := \|\cdot\|^2_\Phi$ *and* $g : L^\Phi(\mu) \to \mathbb{R}$ *be convex and continuous.*

If $d := \inf(f + g)(L^\Phi(\mu))$ *and* $d_n := \inf(f + g)(X_n)$ *for* $n \in \mathbb{N}$, (δ_n) *is a sequence of positive numbers tending to zero and* $x_n \in X_n$ *is chosen to have* $f(x_n) + g(x_n) \leq d_n + \delta_n$, *then the minimizing sequence* $\{x_n\}$ *converges to the (uniquely determined) minimal solution of* $f + g$ *on* $L^\Phi(\mu)$.

Proof: Theorems 46 and 41.

6.3. **A Greedy Algorithm in Orlicz Space.** For a compressed approximate representation of a given function in $L^2[a, b]$ by harmonic oscillations it is reasonable to take into account only the frequencies with dominating Fourier-coefficients. This leads to nonlinear algorithms, whose generalization to Banach spaces was considered by V. N. Temlyakov. We will discuss the situation in Orlicz spaces.

Definition 13. *Let* X *be a Banach space, then* $\mathrm{D} \subset X$ *is called a dictionary if*

(1) $\|\varphi\| = 1$ *for all* $\varphi \in \mathrm{D}$;
(2) *from* $\varphi \in \mathrm{D}$ *it follows that* $-\varphi \in \mathrm{D}$;
(3) $X = \overline{\text{span (D)}}$.

We consider an algorithm from the class of nonlinear m-term algorithms [78]. In [76] it is called Weak Chebyshev Greedy Algorithm (WCGA).

Algorithm *(WGA) Let* X *be strictly convex and* $\tau = (t_k)_{k=1}^\infty$ *with* $0 < t_k \leq 1$ *for all* $k \in \mathbb{N}$. *Let* $x \in X$ *be arbitrary, and* $F_x \in S(X^*)$ *be a functional with* $F_x(x) = \|x\|$.

Given $x \in X \setminus \{0\}$, *set* $r_0 := x$, *and for* $m \geq 1$

(1) *choose* $\varphi_m \in \mathrm{D}$ *with*

$$F_{r_{m-1}}(\varphi_m) \geq t_m \sup\{F_{r_{m-1}}(\varphi) | \varphi \in \mathrm{D}\},$$

(2) *for* $U_m := \text{span }\{\varphi_j, j = 1, \ldots, m\}$ *let* x_m *be the best approximation of* x *w.r.t.* U_m,

(3) *Set* $r_m := x - x_m$ *and* $m \leftarrow m + 1$, *goto 1.*

If X is flat convex, then F_x is given by the gradient $\nabla\|x\|$ of the norm at x (s. Mazur's Theorem 16). Apparently: $\|F_x\| = 1$.

For the case $\tau = (t)$ with $0 < t \leq 1$ we denote the corresponding algorithm by GA.

The following theorem is proved by V. N. Temlyakov in [76]:

Theorem 48. *Let X be a strict convex and reflexive Banach space with Kadec-Klee property (s. Theorem 22), whose norm is Frêchet-differentiable. Then GA converges for every dictionary* D *and every $x \in X$.*

The above theorem can be restated:

Corollary 4. *Let X be an E-space, whose norm is Fréchet-differentiable. Then GA converges for every dictionary* D *and every $x \in X$.*

Proof: According to Theorem 22 X has the Kadec-Klee property and the proof is complete.

In Orlicz-spaces the above theorem assumes the following formulation:

Theorem 49 ((Convergence of GA in Orlicz spaces)). *Let (T, Σ, μ) be a σ-finite, essentially not purely-atomic measure space, Φ be a differentiable, strictly convex Young function, and $L^{\Phi}(\mu)$ be separable and reflexive. Then GA converges for every dictionary* D *and every x in $L^{\Phi}(\mu)$.*

Proof: If Ψ is the conjugate of Φ, then Ψ is differentiable. Hence by Theorem 41 $(L^{\Psi}(\mu), \|\cdot\|_{(\Psi)})$ and $(L^{\Psi}(\mu), \|\cdot\|_{\Psi})$ resp. are Fréchet-differentiable. Due to Anderson's Theorem 23 $(L^{\Phi}(\mu), \|\cdot\|_{(\Phi)})$ and $(L^{\Phi}(\mu), \|\cdot\|_{\Phi})$ resp. are E-spaces, whose norms are by Theorem 41 Fréchet-differentiable.

Remark 10. *Depending on the measure, reflexivity of $L^{\Phi}(\mu)$ can be characterized by appropriate Δ_2 conditions (see Theorem 9).*

Список литературы

[1] N. I. Achieser, *Vorlesungen über Approximationstheorie*,(Akademie Verlag, Berlin, 1953).

[2] B. A. Akimovich, "On the Uniform Convexity and Uniform Smoothness of Orlicz Spaces", Teoriya Funkci Funk. Anal. i Priloz. **15**, 114-120 (1972).

[3] E. Asplund, R. T. Rockafellar, "Gradients of Convex Functions", Trans. Amer. Math. Soc. **139**, 443-467 (1969).

[4] C. Bennet, R. Sharpley, *Interpolation of Operators*, Pure and Appl. Math., **129** (Ac. Press, 1988).

[5] Z. W. Birnbaum, W. Orlicz, "Über die Verallgemeinerung des Begriffes der zueinander konjugierten Potenzen", Studia Math. **3**, 1-64 (1931).
[6] E. Blum, W. Oettli, *Mathematische Optimierung* (Springer-Verlag, Berlin, Heidelberg, New York, 1975).
[7] D.W. Boyd, "The Hilbert Transform on Rearrangement-Invariant Spaces", Canad. J. of Math. **19**, 599-616 (1967).
[8] F. Browder, *Problèmes Nonlinéaires* (Univ. of Montreal Press, 1966).
[9] F. E. Browder, "Nonlinear Operators and Nonlinear Equations of Evolution in Banach Spaces", in: Proc. Symp. Nonlinear Funct. Analysis, Chicago, 890-892, Amer. Math. Soc. **81** (1975).
[10] S. Chen, *Geometry of Orlicz Spaces*, Dissertationes Math. **356** (1996).
[11] S. Chen, S. Huiying, "Reflexive Orlicz Spaces Have Uniformly Normal Structure", Studia Math. **109** (2), 197-208 (1994).
[12] E. E. Cheney, *Introduction to Approximation Theory* (Mc Graw Hill, 1966).
[13] G. Choquet, *Lectures on Analysis* I, II (Amsderdam - New York, 1969).
[14] M. M. Day, *Normed Linear Spaces*, third edition, Ergebnisse der Mathematik u. ihrer Grenzgebiete **21** (Springer, 1973).
[15] I. Descloux, "Approximation in L^p and Tschebyscheff Approximation", SIAM J. Appl. Math. **11**, 1017-1026 (1963).
[16] J. B. Diaz, F. T. Metcalf, "On the Structure of the Set of Subsequential Limit Points of Successive Approximations", Bull. Amer. Math. Soc. **73**, 516-519 (1967).
[17] T. Dominguez, H. Hudzik, G. Lopez, M. Mastylo and B. Sims, "Complete Characterization of Kadec-Klee Properties in Orlicz Spaces", Houston Journal of Mathematics **29** (4), 1027-1044 (2003).
[18] K. Fan and I. Glicksberg, "Some Geometric Properties of the Spheres in a Normed Linear Space", Duke Math. J. **25**, 553-568 (1958).
[19] T. Figiel, "On Moduli of Convexity and Smoothness", Studia Math. **56**, 121-155 (1976).
[20] G. Garrigos, E. Hernandez, J. Martell, "Wavelets, Orlicz Spaces, and Greedy Bases", submitted Jan. 2007.
[21] S. A. Gustavson, "Nonlinear Systems in Semi-Infinite Programming", In: Numerical Solution of Systems of of Nonlinear Algebraic Equations, Ed. G.D. Byren and C.A. Hall (Academic Press, N.Y. and London, 1973).
[22] M. D. Hebden, "A Bound on the Difference Between the Chebyshev Norm and the Hölder Norms of a Function", SIAM J. of Numerical Analysis **8**, 271-277 (1971).
[23] H. Heuser, *Funktionalanalysis* (B.G. Teubner, Stuttgart, 1975).
[24] E. Hewitt, K. Stromberg, *Real and Abstract Analysis* (Springer, 1965).
[25] F. Hirzebruch and W. Scharlau, *Einführung in die Funktionalanalysis*, BI Bd. **296** (Mannheim 1971).
[26] R. B. Holmes, *A Course on Optimization and Best Approximation*, Lecture Notes in Math. **257** (Springer 1972).
[27] H. Hudzik, L. Maligandra, "Amemiya Norm Equals Orlicz Norm in General", Indag. Math. **11**, 573-585 (2000).
[28] A. Kaminska, "On Uniform Rotundity of Orlicz Spaces", Indag. Math. **A85**, 27-36 (1982).
[29] L. Karlovitz, "Construction of Nearest Points in the L_p, p Even, and L^∞ Norms", J. Approx. Theory **3**, 125-127 (1970).
[30] G. Köthe, *Topologische Lineare Räume*, I (Springer Verlag, Berlin-Göttingen-Heidelberg, 1966).
[31] P. Kosmol, *Optimierung und Approximation* (de Gruyter Lehrbuch, Berlin, New York, 1991).
[32] P. Kosmol, *Methoden zur numerischen Behandlung nichtlinearer Gleichungen und Optimierungsaufgaben* (B.G. Teubner Studienbücher, Stuttgart, zweite Auflage, 1993).
[33] P. Kosmol, "Projection Methods for Linear Optimization", Izv. NAN Armenii, Matematika [Journal of Contemporary Mathematical Analysis (Armenian Academy of Sciences)] **36** (6), 49-56 (2001).
[34] P. Kosmol, "Nichtexpansive Abbildungen bei Optimierungsverfahren", First Symposium on Operations Research, University of Heidelberg, Sept. 1-3, 1976, in: Operations Research Verfahren/

Methods of Operations Research XXV, 88-92, R. Henn et al. (Ed.) (Verlag Anton Hain, Meisenheim, 1976).
[35] P. Kosmol, "Flache Konvexität und schwache Differenzierbarkeit", Manuscripta Mathematica **8**, 267-270 (1973).
[36] P. Kosmol, D. Müller-Wichards, "Homotopy Methods for Optimization in Orlicz Space", in: Proceedings of the 12-th Baikal Int. Conf. **1**, 224-230, June 24-July 1 (2001).
[37] P. Kosmol, D. Müller-Wichards, "Homotopic Methods for Semi-infinite Optimization", Izv. NAN Armenii, Matematika [Journal of Contemporary Mathematical Analysis (Armenian Academy of Sciences)] **36** (6), 35-51 (2001).
[38] P. Kosmol, D. Müller-Wichards, "Stability for Families of Nonlinear Equations", Izv. NAN Armenii, Matematika [Journal of Contemporary Mathematical Analysis (Armenian Academy of Sciences)] **41** (1), 49-58 (2006).
[39] P. Kosmol, D. Müller-Wichards, "Optimierung in Orlicz-Räumen", Manuscript, University of Kiel (2008).
[40] P. Kosmol, "Ein Algorithmus für Variationsungleichungen und lineare Optimierungsaufgaben", in: Deutsch - französisches Treffen zur Optimierungstheorie, Hamburg (1986).
[41] P. Kosmol, "Flache Konvexität von Orliczräumen", Colloq. Math. **31**, 249-252 (1974).
[42] P. Kosmol, "Optimierung konvexer Funktionen mit Stabilitätsbetrachtungen", Dissertationes Mathematicae CXL, 1-38 (1976).
[43] P. Kosmol, "Regularization of Optimization Problems and Operator Equations", in: Lecture Notes in Economics and Mathematical Systems **117**, Optimization and Operations Research, Oberwolfach, 1975, 161-170 (Springer, 1976).
[44] P. Kosmol, "Über Approximation stetiger Funktionen in Orlicz-Räumen", J. of Approx. Theory **8**, 67-83 (1973).
[45] P. Kosmol, M. Wriedt, "Starke Lösbarkeit von Optimierungsaufgaben", Math. Nachrichten **83**, 191-195 (1978).
[46] W. Krabs, *Optimierung und Approximation* (Teubner, Stuttgart, 1972).
[47] M. A. Krasnosielskii, Ya. B. Ruticki, *Convex Functions and Orlicz Spaces* (P. Noordhoff, Groningen, 1961).
[48] B. Kripke, "Best Approximation with Respect to Nearby Norms", Numer. Math., **6**, 103-105 (1964).
[49] W. Kurc, "Strictly and Uniformly Monotone Musielak-Orlicz Spaces and Applications to Best Approximation", J. Approx. Theory **69**, 173-187 (1992).
[50] E. S. Levitin, B. T. Polyak, "Convergence of Minimizing Sequences in Conditional Extremum Problems", Soviet Math. Dokl. **7**, 764-767 (1966).
[51] J. Lindenstrauss, "On the Modulus of Smoothness and Divergent Series in Banach Spaces", Mich. Math. J. **10**, 241-252 (1963).
[52] J. Lindenstrauss, L. Tsafriri, *Classical Banach Spaces*, I (Springer Verlag, 1977).
[53] A. R. Lovaglia, "Locally Uniform Convex Banach Spaces", Trans. Amer. Math. **XVIII**, 225-238 (1955).
[54] D. G. Luenberger, *Optimization by Vector Space Methods* (Wiley and Sons, NY, 1969).
[55] W. A. J. Luxemburg, "Banach Function Spaces", Doctoral Thesis, Delft (1955).
[56] W. A. J. Luxemburg and A. C. Zaanen, *Riesz Spaces*, I (Amsterdam- London 1971).
[57] W. A. Luxemburg, A. Zaanen, "Conjugate Spaces of Orlicz Spaces", Indag. Math. **XVIII**, 217-228 (1956).
[58] R. M. Maleev, S. L. Troyanski, "On the Moduli of Convexity and Smoothness in Orlicz Spaces", Studia Math **54**, 131-141 (1975).
[59] Y. Meyer, *Wavelets and Operators* (Cambridge Studies in Advanced Mathematics **37**, 1992).
[60] H. W. Milnes, "Convexity of Orlicz Spaces", Pacific J. Math. **VII**, 1451-1483 (1957).
[61] D. Müller-Wichards, "Über die Konvergenz von Optimierungsmethoden in Orlicz-Räumen", Diss. Uni. Kiel (1976).
[62] D. Müller-Wichards, "Regularisierung von verallgemeinerten Polya Algorithmen", in: Operations Research Verfahren, Verlag Anton Hain, Meisenheim **XXIV**, 119-136 (1978).

[63] J. Musielak, *Orlicz Spaces and Modular Spaces*, Lecture Notes in Mathematics **1034** (Springer-Verlag, 1983).

[64] L. P. Natanson, *Theorie der Funktionen einer reellen Veränderlichen* (Akademie Verlag Berlin, 1961).

[65] I. Peetre, "Approximation of Norms", J. Approx. Theory, **3**, 243-260 (1970).

[66] R. R. Phelps, *Convex Functions, Monotone Operators and Differentiability*, Lecture Notes in Mathematics **1364** (Springer-Verlag, 1989).

[67] M. M. Rao, "Smoothness of Orlicz Spaces", Indagationes Mathematicae, **27**, 671-689 (1965).

[68] M. M. Rao, "Linear Functionals on Orlicz Spaces", Nieuw. Arch. Wisk. **12**, 77-98 (1964).

[69] M. M. Rao, "Linear Functionals on Orlicz Spaces, General Theory", Pacific J. Math. **25**, 553-585 (1968).

[70] M. M. Rao, Z. D. Ren, *Theory of Orlicz Spaces*, Pure and Applied Mathematics (Marcel Dekker, New York, 1991).

[71] M. M. Rao, Z. D. Ren, *Applications of Orlicz Spaces*, Pure and Applied Mathematics (Marcel Dekker, New York, 2002).

[72] J. R. Rice, *Approximation of Functions*, I and II (Addison Wesley, 1964 and 1969).

[73] R. T. Rockafellar, "Characterization of the Subdifferentials of Convex Functions", Pac. J. Math. **17**, 497-510 (1966).

[74] R. T. Rockafellar, *Convex Analysis* (Princeton, 1970).

[75] P. Soardi, "Wavelet Bases in Rearrangement Invariant Function Spaces", Proc. Amer. Math. Soc. **125** (12), 3669-3673 (1997).

[76] M. M. Temlyakov, "Convergence of Some Greedy Algorithms in Banach Spaces", J. of Fourier Analysis and Appl. **8,5**, 489-505 (2002).

[77] M. M. Temlyakov, "Greedy Algorithms in Banach Spaces", Advances in Computational Mathematics **14**, 277-292 (2001).

[78] M. M. Temlyakov, "Nonlinear Methods of Approximation", Foundations of Computational Mathematics **3**, 33-107 (2003).

[79] I. V. Tsenov, "Some Questions in the Theory of Functions", (Russian) Mat. Sbornik **28**, 473-478 (1951).

[80] B. Turett, "Fenchel-Orlicz Spaces", Diss. Math. **181** (1980).

[81] A. N. Tychonoff, "Solution of Incorrectly Formulated Problems and the Regularization Method", Soviet Math. Dokl. **4**, 1035-1038 (1963).

[82] M. M. Vainberg, *Variational Method and Method of Monotone Operators in the Theory of Nonlinear Equations* (Wiley and Sons, 1973).

[83] D. Werner, *Funktionalanalysis* (Springer, 6. Auflage, 2007).

[84] H. Werner, *Vorlesungen über Approximationstheorie*, Lect. Notes in Math. **14** (Springer Verlag, 1966).

[85] A. C. Zaanen, *Integration* (North Holland, 1967).

Поступила 8 декабря 2008