

A Note on the Approximate Symmetry of Bregman Distances

Stefan Kindermann

*Industrial Mathematics Institute, Johannes Kepler University Linz,
Altenbergergasse 69, 4040 Linz, Austria
kindermann@indmath.uni-linz.ac.at*

Received: August 21, 2018

Accepted: November 28, 2018

The Bregman distance $B_{\xi_x}(y, x)$, $\xi_x \in \partial J(y)$, associated to a convex sub-differentiable functional J is known to be in general non-symmetric in its arguments x, y . In this note we address the question when Bregman distances can be bounded against each other when the arguments are switched, i.e., if some constant $C > 0$ exists such that for all x, y on a convex set M it holds that $\frac{1}{C}B_{\xi_x}(y, x) \leq B_{\xi_y}(x, y) \leq CB_{\xi_x}(y, x)$. We state sufficient conditions for such an inequality and prove in particular that it holds for the p -powers of the ℓ_p and L^p -norms when $1 < p < \infty$.

Keywords: Bregman distance, symmetry, strong monotonicity, convexity.

2010 Mathematics Subject Classification: 52A41, 47H05, 26B25.

1. Introduction

For a convex sub-differentiable functional J on a convex subset M of a Banach space X , the Bregman distance between $x, y \in M$ with a chosen element $\xi_x \in \partial J(x)$ is defined by

$$B_{\xi_x}(y, x) = J(y) - J(x) - \langle \xi_x, y - x \rangle \quad \xi_x \in \partial J(x).$$

It is a useful tool in the analysis of optimization problems, and in particular, in the Banach space theory of variational regularization, it has become a useful tool to measure errors; see, e.g., [5, 3]. Note that for Hilbert spaces X with $J = \frac{1}{2}\|\cdot\|^2$, the Bregman distance equals $B_{\xi_x}(y, x) = \frac{1}{2}\|x - y\|^2$, hence convergence in Bregman distance is often used to generalize results on norm convergence in Hilbert spaces.

However, in general, in contrast to the Hilbert space case, the Bregman distance is not symmetric in its arguments, i.e., $B_{\xi_x}(y, x) \neq B_{\xi_y}(x, y)$. It has been observed that the conditions for convergence rates for Tikhonov regularization are not the same when convergence is measured in the Bregman distance and in its switched version [2, 4]. It is hence of interest, to study the question when a switching of the arguments does not change the topology of Bregman convergence, i.e., when the switched Bregman distance can be bounded by a constant times the original one.

More precisely, we investigate conditions, when there exists a constant C_M such that for all $x, y \in M$ it holds that

$$\frac{1}{C_M} B_{\xi_x}(y, x) \leq B_{\xi_y}(x, y) \leq C_M B_{\xi_x}(y, x), \quad \forall \xi_x \in \partial J(x), \xi_y \in \partial J(y), \quad (1)$$

where M is some convex set in X . If (1) holds, then, when investigating the Bregman convergence of an approximating sequence $y = x_n$ to x , it does not matter which of the two variants, i.e., $B_{\xi_x}(x_n, x)$ or $B_{\xi_{x_n}}(x, x_n)$ one considers.

It is trivial that (1) holds with the constant $C = 1$ for $J = \frac{1}{2} \|\cdot\|_H^2$ with $\|\cdot\|_H$ a Hilbert space norm. However, the inequality does not hold in general, as some simple counterexamples show.

A practically useful result that we show in this paper is that for p -powers of the ℓ^p - or L^p -norms we can establish the inequality (1). The corresponding result reads as follows:

Theorem 1.1. *For $1 < p < \infty$, let X be the space of p -summable sequences $X = \ell_p$ or p -integrable functions $X = L^p(\Omega)$, $\Omega \subset \mathbb{R}^n$. Let J be the p -power of the corresponding norms $J = \|\cdot\|^p$. Then there exist a constant C_p such that*

$$\frac{1}{C_p} B_{\xi_x}(y, x) \leq B_{\xi_y}(x, y) \leq C_p B_{\xi_x}(y, x) \quad (2)$$

for all $x, y \in X$ and for all $\xi_x \in \partial J(x)$, $\xi_y \in \partial J(y)$. A constant C_p is given by $C_p = 2 \max\{\frac{1}{p-1}, p-1\}$.

2. Sufficient conditions for approximate symmetry

Before we investigate some sufficient conditions for (1), we illustrate the problem by some (simple) examples.

Consider the abs-functional on $\mathbb{R}$: $J(x) = |x|$. Its subgradient is $\partial J(x) = \text{sign}(x)$ for $x \neq 0$ and multi-valued $\partial J(x) \in [-1, 1]$ at $x = 0$. Hence, for $x \neq 0$,

$$B_{\xi_0}(x, 0) = |x| - [-1, 1]x = |x|[0, 2] \quad B_{\xi_x}(0, x) = -|x| + \text{sign}(x)x = 0.$$

Thus, (1) cannot hold in this case. The fact that the subgradient is here multi-valued at one of the arguments is not responsible for the violation of inequality (1) as the example $x > 0$, $y = -\epsilon$, $\epsilon > 0$ shows:

$$B_{\xi_\epsilon}(x, \epsilon) = 2x \quad B_{\xi_x}(\epsilon, x) = 2\epsilon.$$

Since ϵ can be chosen arbitrary small, no constant for (1) exists.

Furthermore, let us illustrate that the lack of differentiability is not a sole reason for a violation of (1). Indeed, we might introduce a Huber-type smoothing of the previous example:

$$J(x) = \begin{cases} |x| & |x| \geq 1, \\ x^2 & |x| < 1, \end{cases}$$

such that $J(x) \in C^1(\mathbb{R})$. Then, for $x > 1$ and $y = 1 - \epsilon$ with $1 > \epsilon > 0$, we have

$$B_{\xi_y}(x, y) = x - 2xy + y^2 = (1 - \epsilon)^2 - x(1 - 2\epsilon) = 1 - x + O(\epsilon)$$

while

$$B_{\xi_x}(y, x) = (1 - \epsilon)\epsilon = O(\epsilon).$$

Thus, (1) cannot hold uniformly in ϵ . Note that similar counter-examples to (1) can be constructed where J is in C^∞ .

The main inequality (1) can certainly be established if appropriate upper and lower bounds for the Bregman distances can be verified. For functionals involving powers of norms, Xu and Roach [6] established useful estimates that relate upper and lower bounds for Bregman distances to smoothness and strict convexity. It is thus not surprising that C_M will be related to the ratio of quantities representing smoothness and strict convexity. Before we elaborate on that, let us state that Theorem 1.1 cannot be obtained by a simple application of the Xu-Roach inequalities but requires a more detailed analysis. Indeed, for the l^p - or L^p -case, the Xu-Roach inequalities imply the following estimates on bounded sets:

$$B_{\xi_x}(y, x) \leq C\|x - y\|^{\min\{p, 2\}} \quad B_{\xi_x}(y, x) \geq C\|x - y\|^{\max\{p, 2\}}.$$

Thus, except for the trivial Hilbert space case $p = 2$, the exponents do not match to establish (1) in a simple manner.

Before we state a general result, we give some reformulations of the main inequality: Obviously, for (1) to hold, it is enough that

$$B_{\xi_y}(x, y) \leq C_M B_{\xi_x}(y, x) \quad (3)$$

for all $x, y \in M$ and all $\xi_y \in \partial J(y)$, $\xi_x \in \partial J(x)$ as the other inequality follows easily from that one by switching arguments. We may also introduce the *symmetric Bregman distance* with $\xi_x \in \partial J(x)$ and $\xi_y \in \partial J(y)$

$$B_S(x, y) = B_{\xi_y}(x, y) + B_{\xi_x}(y, x) = \langle \xi_x - \xi_y, x - y \rangle.$$

Then the main inequality can be expressed in terms of the symmetric version. Indeed, by adding $C_M B_{\xi_y}(x, y)$ to (3), we get the following lemma:

Lemma 2.1. *Inequality (3) holds with some constant C_M if and only if*

$$B_{\xi_y}(x, y) \leq \eta_M B_S(x, y) \quad (4)$$

holds with some constant $\eta_M < 1$. The constants are related by

$$C_M = \frac{\eta_M}{1 - \eta_M} \quad \text{and} \quad \eta_M = \frac{C_M}{C_M + 1}.$$

Note that (4) always trivially holds with $\eta = 1$. We also point out that the approximate symmetry holds for a functional J if it holds for its Legendre-Fenchel conjugate J^* :

$$J^* : X^* \rightarrow \mathbb{R}, \quad J^*(x^*) := \sup_{x \in \text{dom}(J)} \langle x, x^* \rangle - J(x).$$

Indeed, from the well-known relations [1]

$$\langle \xi_y, y \rangle = J(y) + J^*(\xi_y), \quad x^* \in \partial J(x) \Leftrightarrow x \in \partial J^*(x^*),$$

we can conclude that $B_{\xi_y}(x, y) = B_x^*(\xi_y, \xi_x)$, (5)

where B_x^* is the Bregman distance associated to the conjugate functional J^* and ξ_y, ξ_x are elements of the subgradients at y and x , respectively. Thus, we have:

Lemma 2.2. *The inequality (1) holds for the Bregman distance of a functional J on M with a constant C_M if and only if it holds with the same constant for the Bregman distance of the conjugate functional J^* on $\partial J(M)$.*

Let us now state a general result on approximate symmetry:

Theorem 2.3. *Suppose that ∂J is strongly monotone on M*

$$\langle \partial J(x) - \partial J(y), x - y \rangle \geq c_0 \|x - y\|^2 \quad \forall x, y \in M,$$

and Lipschitz continuous on M with Lipschitz constant L . Then (3) holds true for all $x, y \in M$ with a constant $C_M = \frac{L}{c_0}$.

Proof. The result follows from the estimates

$$\begin{aligned} B_{\xi_x}(y, x) &= \int_0^1 \langle \partial J(x + t(y - x)) - \partial J(x), y - x \rangle dt \\ &\leq L \int_0^1 t dt \|x - y\|^2 = \frac{L}{2} \|x - y\|^2 \quad \text{and} \\ B_{\xi_y}(x, y) &= \int_0^1 \langle \partial J(y + t(x - y)) - \partial J(y), x - y \rangle dt \\ &= \int_0^1 \frac{1}{t} \langle \partial J(y + t(x - y)) - \partial J(y), (y + t(x - y)) - y \rangle dt \\ &\geq c_0 \int_0^1 \frac{1}{t} \|y + t(x - y) - y\|^2 dt = \frac{c_0}{2} \|x - y\|^2. \quad \square \end{aligned}$$

Note that for a $C^2(M)$ -functional with a strongly monotone gradient, we can estimate the constants in the previous theorem by upper and lower bounds for the second derivative:

$$C_M \leq \frac{\sup_{x \in M, \|h\| \leq 1} |\langle \nabla^2 J(x) h, h \rangle|}{\inf_{x \in M, \|h\| \leq 1} |\langle \nabla^2 J(x) h, h \rangle|}. \quad (6)$$

As an example, we consider the square-root regularization of the abs-functional, which is often employed when a differentiable approximation to the ℓ^1 -norm is needed:

$$J : \mathbb{R} \rightarrow \mathbb{R}, \quad x \rightarrow \sqrt{x^2 + \epsilon}, \quad \epsilon > 0.$$

We show that the associated Bregman distance satisfies (3) on bounded sets: We have for the derivatives

$$J'(x) = \frac{x}{\sqrt{x^2 + \epsilon}} \quad J''(x) = \frac{\epsilon}{(x^2 + \epsilon)^{\frac{3}{2}}}.$$

A direct calculation reveals a constant $C_M = \sup_{x,y \in M} \frac{\sqrt{y^2 + \epsilon}}{\sqrt{x^2 + \epsilon}}$.

In particular, if we restrict the functional to bounded sets, $|x|, |y| \leq R$ then

$$C_M \leq \sqrt{1 + \frac{R^2}{\epsilon}} \leq 1 + \frac{R}{\sqrt{\epsilon}}.$$

Using (6) gives a slightly worse bound for $\frac{R}{\sqrt{\epsilon}} \gg 1$:

$$C_M \leq \frac{\sup_{|x| \leq R} J''(x)}{\inf_{|x| \leq R} J''(x)} = \frac{\sup_{|x| \leq R} (|x|^2 + \epsilon)^{\frac{3}{2}}}{\inf_{|x| \leq R} (|x|^2 + \epsilon)^{\frac{3}{2}}} \leq \left(\frac{R^2}{\epsilon} + 1 \right)^{\frac{3}{2}} \leq 1 + O\left(\frac{R}{\sqrt{\epsilon}}\right)^{\frac{3}{2}}.$$

The same theorem can be used to prove (1) for the p -power norm-functional in Theorem 1.1 if we restrict x, y to bounded sets. Below, however, we obtain the result without the restriction to bounded sets by a direct calculation.

For the sake of generalizations, we slightly improve Theorem 2.3 by showing that the conditions in (2.3) can be localized in the following sense:

Proposition 2.4. *Suppose that (1) holds for all x, y in $M_0 = \{\|x - y\| \leq \epsilon\}$. Then there exists a constant C_M such that (1) holds in $M = \{\|x - y\| \leq R\}$, $R > \epsilon$, with*

$$C_M = \frac{R - \epsilon}{\epsilon} + \frac{RC_{M_0}}{\epsilon}.$$

Proof. Denote by $B_R(x) := \{x + z \mid \|z\| \leq R\}$ a ball around x with radius R , and fix $x \in X$. According to Lemma 2.1 and the hypothesis, we have a constant $\eta_{M_0} = \frac{C_{M_0}}{1+C_{M_0}}$ such that (4) holds for all z in $B_\epsilon(x)$. Take $y \in B_R(x)$, and set $z = x + \lambda(y - x)$ and $\lambda = \frac{\epsilon}{R} < 1$. Note that

$$z - x = \lambda(y - x), \quad y - z = (1 - \lambda)(y - x).$$

Then $z \in B_\epsilon(x)$. We have

$$\begin{aligned} B_{\xi_x}(y, x) &= J(y) - J(x) - \langle \xi_x, y - x \rangle \\ &= J(z) - J(x) - \langle \xi_x, z - x \rangle + J(y) - J(z) - \langle \xi_x, y - z \rangle \\ &\leq \eta_{M_0} \langle \xi_z - \xi_x, z - x \rangle \\ &\quad + (J(y) - J(z) - \langle \xi_z, y - z \rangle) + \langle \xi_z - \xi_x, y - z \rangle \\ &\leq \eta_{M_0} \langle \xi_z - \xi_x, z - x \rangle + \langle \xi_y - \xi_z, y - z \rangle + \langle \xi_z - \xi_x, y - z \rangle \end{aligned}$$

$$\begin{aligned}
&= (\eta_{M_0}\lambda + (1 - \lambda)) \langle \xi_z - \xi_x, y - x \rangle + (1 - \lambda) \langle \xi_y - \xi_z, y - x \rangle \\
&\leq \max \{ \eta_{M_0}\lambda + 1 - \lambda, 1 - \lambda \} (\langle \xi_z - \xi_x, y - x \rangle + \langle \xi_y - \xi_z, y - x \rangle) \\
&= \left(\frac{\eta_{M_0}}{\lambda} + 1 - \lambda \right) \langle \xi_y - \xi_x, y - x \rangle = (\eta_{M_0}\lambda + 1 - \lambda) B_S(y, x),
\end{aligned}$$

which proves (4) with $\eta_M = (\eta_{M_0}\lambda + 1 - \lambda)$ for all $y \in B_R(x)$, from which the constant C_M can be calculated. Note that the estimate in the penultimate line is valid because both expression with brackets $\langle \cdot, \cdot \rangle$ in the previous line are nonnegative. Switching the role of x, y proves the assertion. $\square$

3. The proof of Theorem 1.1

We prove Theorem 1.1 by a direct calculation of the constant C_M . Slightly more general as stated in the theorem, we consider the p -power of a Hilbert space norm. For $p \in (1, \infty)$, define

$$J : H \rightarrow \mathbb{R} \quad x \rightarrow \frac{1}{p} \|x\|^p$$

with $\|\cdot\|$ the norm on H . Note that J is continuously differentiable with

$$\partial J(x) = e_x \|x\|^{p-1} \quad e_x := \begin{cases} \frac{x}{\|x\|} & x \neq 0, \\ 0 & x = 0. \end{cases}$$

To prove the theorem, we establish the following technical lemma:

Lemma 3.1. *Let $1 < p < 2$, and define the functions*

$$\begin{aligned}
f_\theta(r) &= \frac{1}{p} r^p + \left(1 - \frac{1}{p}\right) - r\theta, & r \geq 0, \theta \in [-1, 1], \\
g_\theta(r) &= \frac{1}{p} + \left(1 - \frac{1}{p}\right) r^p - r^{p-1}\theta & r \geq 0, \theta \in [-1, 1].
\end{aligned}$$

$$\text{Then} \quad \max_{r \geq 0, \theta \in [-1, 1]} \frac{f_\theta(r)}{g_\theta(r)} \leq \frac{1}{p-1} \max_{r \geq 1} \frac{r^{p-1} + 1}{r^{p-1} + r^{p-2}} \leq \frac{2}{p-1}. \quad (7)$$

Proof. Both functions are positive with a zero only at $r = 1$ and $\theta = 1$. The ratio function $\frac{f_\theta(r)}{g_\theta(r)}$ is thus a positive smooth function except for $r = 1, \theta = 1$ and has the following limit values for any $\theta \in [-1, 1]$

$$\lim_{r \rightarrow 0} \frac{f_\theta(r)}{g_\theta(r)} = p-1, \quad \lim_{r \rightarrow \infty} \frac{f_\theta(r)}{g_\theta(r)} = \frac{1}{p-1}, \quad \lim_{r \rightarrow 1} \frac{f_\theta(r)}{g_\theta(r)} = 1.$$

In case $\theta = 1$, the last limit can be evaluated by twice applying de l'Hospital's rule. Thus, for fixed θ the ratio function can be continuously extended to all $r \geq 0$.

For a fixed value $r \neq 1$, we calculate the derivative

$$\begin{aligned}
\frac{\partial}{\partial \theta} \frac{f_\theta(r)}{g_\theta(r)} &= \frac{1}{g_\theta(r)^2} (f_\theta(r)r^{p-1} - rg_\theta(r)) \\
&= \frac{1}{g_\theta(r)^2} \left(\frac{1}{p} r^{2p-1} + \left(1 - \frac{1}{p}\right) r^{p-1} - \frac{1}{p} r + \left(1 - \frac{1}{p}\right) r^{p+1} \right)
\end{aligned}$$

and observe that this value is always of the same sign for a fixed r . Thus the maximal value with respect to θ is attained at the boundary of $[-1, 1]$:

$$\max_{\theta \in [-1, 1]} \frac{f_\theta(r)}{g_\theta(r)} \leq \max_{\theta \in \{-1, 1\}} \frac{f_\theta(r)}{g_\theta(r)}.$$

We now investigate those interior maxima of the ratio function with respect to r in the interval $(0, 1)$ and $(1, \infty)$ that have a value larger than the largest currently found $\frac{f_\theta(\infty)}{g_\theta(\infty)} = \frac{1}{p-1}$. As the ratio is a smooth function, we may conclude from the optimality condition of first order that at such a maximum r_* , we have

$$f'_\theta(r_*) = \frac{f_\theta(r_*)}{g_\theta(r_*)} g'_\theta(r_*).$$

Setting $\lambda = \frac{f_\theta(r_*)}{g_\theta(r_*)} > \frac{1}{p-1}$ (recall that only maxima larger than $\frac{1}{p-1}$ are of interest) yields the condition

$$r_*^{p-1} - \theta = \lambda(p-1)(r_*^{p-1} - \theta r_*^{p-2}), \quad r_* \geq 0, \quad \lambda(p-1) \geq 1, \quad \theta \in \{-1, 1\}.$$

We investigate the cases $\theta = 1$ and $\theta = -1$ separately.

In case that $\theta = 1$, the optimality condition reads

$$1 - \frac{1}{(r_*)^{p-1}} = \lambda(p-1)\left(1 - \frac{1}{r_*}\right).$$

Thus, for $\lambda(p-1) > 1$ to hold, the following inequalities must be satisfied:

$$\begin{cases} 1 - \frac{1}{(r_*)^{p-1}} > 1 - \frac{1}{r_*} \Leftrightarrow r_* < (r_*)^{p-1} & \text{if } r_* > 1, \\ \frac{1}{(r_*)^{p-1}} - 1 > \frac{1}{r_*} - 1 \Leftrightarrow r_* > (r_*)^{p-1} & \text{if } r_* < 1. \end{cases}$$

But in both cases the corresponding inequality cannot be true since $p-1 \in (0, 1)$. Thus in this case no interior maximum with a value larger than $\frac{1}{p-1}$ can exist.

In case that $\theta = -1$ the optimality condition reads

$$r_*^{p-1} + 1 = \lambda(p-1)(r_*^{p-1} + r_*^{p-2}) \geq (r_*^{p-1} + r_*^{p-2}),$$

thus, $r_* \geq 1$. Hence we have that

$$\frac{f_\theta(r_*)}{g_\theta(r_*)} = \lambda \leq \frac{1}{p-1} \max_{r \geq 1} \frac{r^{p-1} + 1}{r^{p-1} + r^{p-2}} = \frac{1}{p-1} \left(1 + \max_{r \geq 1} \frac{1 - r^{p-2}}{r^{p-1} + r^{p-2}}\right). \quad (8)$$

Since for $r \geq 1$

$$\frac{r^{p-1} + 1}{r^{p-1} + r^{p-2}} = \frac{r^{p-1} + 1}{r^{p-1}} \frac{r}{r+1} \leq 2,$$

we have established the upper bound $\frac{2}{p-1}$. □

We now continue with the proof of Theorem 1.1. We observe that for the l^p - or L^p -case, we can express the Bregman distances componentwise:

$$B_{\xi_y, \frac{1}{p} \|\cdot\|_{l^p}^p}(x, y) = \sum_{i=1}^{\infty} B_{\xi_{y_i}}(x_i, y_i),$$

where $B_{\xi_{y_i}}(x_i, y_i)$ is the Bregman distance for the functional $J: \mathbb{R} \rightarrow \mathbb{R}$, where $J(x) = \frac{1}{p}|x|^p$. Thus, for (3) to hold it is enough to prove the corresponding inequality for the Bregman distance of this functional, $B_{\xi_y}(x, y)$, $x, y \in \mathbb{R}$.

We have for $y \neq 0$ $B_{\xi_y}(x, y) = \|y\|^p B_{\xi_{e_y}}(\frac{x}{\|y\|}, e_y)$

and for $x \neq 0$ $B_{\xi_y}(x, y) = \|x\|^p B_{\xi_z}(e_x, \frac{y}{\|x\|})$, $\xi_z \in \partial J(\frac{y}{\|x\|})$.

Thus, for $y \neq 0$, (3) is equivalent to

$$B_{\xi_{e_y}}(z, e_y) \leq C_M B_{\xi_z}(e_y, z)$$

with $z = \frac{x}{\|y\|}$. We may calculate for $y \neq 0$ and $\theta = \langle e_y, e_z \rangle = \text{sign}(y)\text{sign}(z)$ and $z = \|z\|e_z$ that

$$B_{\xi_{e_y}}(z, e_y) = f(\|z\|, \theta) \quad B_{\xi_z}(e_y, z) = g(\|z\|, \theta).$$

Thus, for the case $1 < p < 2$, and $y \neq 0$, the theorem follows from Lemma 3.1. The case of $y = 0$ can be estimated directly by $B_{\xi_0}(x, 0) \leq \frac{1}{1-p} B_{\xi_x}(0, x)$.

Consider now the case $2 < p < \infty$. From the duality formula (5) we obtain that (3) is equivalent to

$$B_x^*(\xi_y, \xi_x) \leq C_M B_y^*(\xi_x, \xi_y).$$

Since $\partial J(\mathbb{R}) = \mathbb{R}$, and $J^* = \frac{1}{q} \|\cdot\|^q$, $\frac{1}{q} + \frac{1}{p} = 1$ we obtain that (3) holds with the constant

$$\frac{2}{q-1} = 2(p-1),$$

which finishes the proof. □

The analysis furthermore gives the more precise estimate

$$\frac{1}{t-1} \leq C_p \leq \frac{1}{t-1} \max_{r \geq 1} \frac{r^{t-1} + 1}{r^{t-1} + r^{t-2}}, \quad \text{where } t = \begin{cases} p & 1 < p < 2 \\ \frac{p}{p-1} & p \geq 2 \end{cases}.$$

The proof can be extended verbatim with the same constant to the case of X being the space of sequences with values in a Hilbert space and

$$J(x) = \frac{1}{p} \|x\|_{\ell^p(N, H)}^p = \sum_{i=1}^n \frac{1}{p} \|x_i\|_H^p$$

or Hilbert space-valued functions on an interval $L^p(I, H)$.

Acknowledgement. The work was partly supported by the Austrian Science Fund (FWF) Project P 30157-N31.

References

- [1] I. Ekeland, R. Témam: *Convex Analysis and Variational Problems*, SIAM, Philadelphia, (1999).
- [2] J. Flemming: *A converse result for Banach space convergence rates in Tikhonov-type convex regularization of ill-posed linear equations*, J. Inverse Ill-Posed Problems 26(5) (2018) 639–646.
- [3] B. Hofmann, B. Kaltenbacher, C. Pöschl, O. Scherzer: *A convergence rates result for Tikhonov regularization in Banach spaces with non-smooth operators*, Inverse Problems 23(3) (2007) 987–1010.
- [4] S. Kindermann: *Convex Tikhonov regularization in Banach spaces: new results on convergence rates*, J. Inverse Ill-Posed Problems 24(3) (2016) 341–350.
- [5] E. Resmerita, O. Scherzer: *Error estimates for non-quadratic regularization and the relation to enhancement*, Inverse Problems 22(3) (2006) 801–814.
- [6] Z. B. Xu, G. F. Roach: *Characteristic inequalities of uniformly convex and uniformly smooth Banach spaces*, J. Math. Anal. Appl. 157(1) (1991) 189–210.